



Minerva Access is the Institutional Repository of The University of Melbourne

Author/s:

Gagliardi, B;Long, SM;Pettigrove, VJ;Griffin, PC;Hoffmann, AA

Title:

A Re-Evaluation of Chironomid Deformities as an Environmental Stress Response: Avoiding Survivorship Bias and Testing Noncontaminant Biological Factors

Date:

2019-08-01

Citation:

Gagliardi, B., Long, S. M., Pettigrove, V. J., Griffin, P. C. & Hoffmann, A. A. (2019). A Re-Evaluation of Chironomid Deformities as an Environmental Stress Response: Avoiding Survivorship Bias and Testing Noncontaminant Biological Factors. *Environmental Toxicology and Chemistry*, 38 (8), pp.1658-1667. <https://doi.org/10.1002/etc.4446>.

Persistent Link:

<https://hdl.handle.net/11343/285704>

Running head: Chironomid deformities as an environmental stress indicator

A re-evaluation of chironomid deformities as an environmental stress response: avoiding survivorship bias and testing noncontaminant biological factors

Bryant Gagliardi^{a,*}, Sara M. Long^a, Vincent J. Pettigrove^a, Philippa C. Griffin^b, Ary A. Hoffmann^a

^a School of BioSciences, University of Melbourne VIC 3010, Australia

^b EMBL Australia Bioinformatics Resource, University of Melbourne, VIC 3010, Australia

*Address correspondence to bryant.gagliardi@gmail.com

DISCLAIMER

The authors have no conflict of interest to declare.

ACKNOWLEDGEMENTS

This study was funded by the Centre for Aquatic Pollution Identification and Management, and the Holsworth Wildlife Research Endowment. The authors thank Dr W Tyler Mehler for helpful comments on the draft manuscript, and three anonymous reviewers for helpful comments on the submitted manuscript.

This is the author manuscript accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the [Version of Record](#). Please cite this article as [doi: 10.1002/etc.4446](https://doi.org/10.1002/etc.4446).

This article is protected by copyright. All rights reserved.

ABSTRACT

Larval deformities have been observed in chironomids, and are thought to be associated with aquatic contaminant exposure. However, deformities have not been linked with contaminants in the absence of potential confounding variables including mortality in laboratory assays, which introduces a survivorship bias (SB). There is also a paucity of data on noncontaminant causes. Additionally, power analyses are rarely undertaken, meaning effect sizes detectable are usually uncertain. We therefore aimed to clarify factors associated with deformities, by running SB-free (i.e. sublethal) assays, assessing contaminant (copper and imidacloprid) and noncontaminant (malnutrition) stressors, and considering natural biological (metamorphosis) factors in *Chironomus tepperi*. We include *a posteriori* power analyses for all tests. Our assays found no significant association between tested factors and deformity rate. Power analyses indicated that the stressor experiment had moderate power to detect deformity effects. The metamorphosis assay had relatively lower power (due to an unexpectedly high control deformity rate), highlighting the importance of power tests in these types of evaluations. These results, in conjunction with others recently published, raise doubts as to the causal effects of environmental stressors on deformity incidence. By avoiding SB, testing noncontaminant factors, and testing statistical power, we present a more holistic methodology, to resolve ongoing uncertainty in this area. We also discuss possible future directions for chironomid deformity research, and concerns regarding SB in ecotoxicology.

KEYWORDS

Chironomus, deformities, aquatic ecotoxicology, survivorship bias, pupation, malnutrition

INTRODUCTION

Field-collected chironomid (Diptera: Chironomidae) larvae sometimes exhibit deformities of the mouthparts (e.g. mentum, mandibles, and pectin epipharyngis) and antennae (Janssens de Bisthoven and Gerhardt 2005; Di Veroli et al. 2014; Odume et al. 2016). Aquatic chemical contaminant levels and the incidence of chironomid deformities are often spatially associated in field samples, suggesting deformities indicate contaminant exposure and hence are a useful ecotoxicological endpoint [4]. However, over 40 years' worth of laboratory ecotoxicological experiments aiming to assess the causal relationship between contaminants and deformities (using the model genus *Chironomus*) have yielded inconsistent results (Vermeulen 1995; Janssens de Bisthoven and Gerhardt 2005; Arambourou et al. 2014; Salmelin et al. 2015). Meta-analysis shows that overall data across the literature are highly variable, indicating positive, nonsignificant, and even negative associations between contaminant concentrations and deformity incidence; and that overall effect sizes do not significantly differ from zero (no effect) (Gagliardi et al. 2016). The inconsistency of results may suggest that contaminants do not cause deformities, or that experimental conditions vary too widely between studies for an underlying link to be detected. It may also suggest that chironomid deformities are primarily induced in the presence of mixed rather than single stressors (Milošević et al. 2016), although laboratory data do not tend to support this hypothesis (Arambourou et al. 2014; Gagliardi et al. 2016). These concerns have limited the application of this endpoint in ecotoxicology. A related concern in chironomid deformity research is survivorship bias (SB). Deformity incidence is a sublethal response, and accordingly, causal factors should be tested at sublethal levels (Abdel-Moneim et al. 2015). Experiments in which larvae are killed in significant numbers are confounded by mortality (Martinez et al. 2001;

Martinez et al. 2004; Dias et al. 2008). This confounding effect is known as survivorship bias (Qi et al. 2017), which occurs when an analysis focuses on individuals that made it past a selection process (in this case mortality), due to the invisibility of individuals that did not make it past the selection process (in this case, larvae that died).

Mortality does not act randomly: it systematically kills physiologically weaker individuals to cause a selection bias, and hence mortality-inducing assays only assess sublethal effects in the stronger, mortality-resistant proportion of the population. These individuals are not representative of the entire population, so results based upon only surviving individuals cannot legitimately be extrapolated to the whole population. However, this assumption is routinely applied in chironomid deformity (and other ecotoxicological) studies, where deformities are assessed among survivors of lethal doses, with the underlying assumption that these deformity frequencies would have been identical within the killed population fraction.

Different phenotypes between surviving and killed larvae are expected, and some experimental data supports the notion that “deformity likelihood” is one of these phenotypes affected [e.g (Martinez et al. 2001; Dias et al. 2008)]. A high proportion of deformed larvae in a mortality-inducing exposure may simply indicate that deformed larvae can better survive the exposure: it does not unequivocally demonstrate the chemical has physiologically induced the observed deformities.

Experimental data on Dipterans and particularly *Drosophila*, which have assessed fluctuating asymmetry, further support the notion that contaminant-induced mortality can selectively kill individuals bearing certain traits. In one *Drosophila* experiment, an apparent association between sternopleural bristle number and contaminant

concentration appears to represent a false positive result due to selective mortality (Polak et al. 2002).

There is therefore clear evidence that SB can risk false positive or false negative results. False positive results arise in situations where deformed larvae better survive the contaminant exposure. This possibility is suggested by cadmium assays in Figure 2 of Janssens de Bisthoven et al. (2001), which demonstrated a stronger positive deformity-mortality correlation than any contaminant effects observed across multiple experiments. False negatives conversely arise in situations where deformed larvae are physiologically weaker, and hence lost from the analysis due to induced mortality, as suggested by several studies including Martinez et al. (2001) (zinc and lead) and Dias et al. (2008) (uranium), which found unexpected negative correlations between contaminant concentration and deformity rates. Furthermore, variation in the degree of mortality between published studies may contribute to the different patterns seen in different experiments (Gagliardi et al. 2016).

Significant mortality levels across experimental treatments should therefore ideally be avoided. As well as SB concerns, inducing lethality also causes a potential confounding by density, as surviving larvae in lethality-inducing replicates are under lower densities. However, a recent meta-analysis found that >95 % of published laboratory experiments were either definitely or possibly confounded by SB, by their failure to either avoid, statistically test, or in many cases even report mortality levels (Gagliardi et al. 2016). Survivorship bias effects are eliminated by ensuring survival (to the larval instar at which deformities are scored) is high across all control and exposure treatments, with no significant effect of concentration on incidence of mortality. This involves using sublethal chemical concentrations for all treatments,

ensuring laboratory conditions are favourable (so as not to induce mortality), and reporting mortality levels across all experimental treatments (Fukudome et al. 2010).

In addition to issues associated with SB, another difficulty in clearly linking deformities to contaminant exposure is that deformities could be induced by other environmental stressors as well as, or instead of, contaminants. Noncontaminant effects on deformities have rarely been tested, and there is a growing awareness of the importance of understanding the role of noncontaminant stressors in inducing ecotoxicological endpoint responses (Townsend et al. 2012; Dalzochio and Gehlen 2016). Stressors such as malnutrition may induce deformities (de Haas et al. 2005; Janssens de Bisthoven and Gerhardt 2005), however, to our knowledge, deformity assays involving only malnutrition have not been published.

It is also possible that abnormalities observed as larval “deformities” are due to normal developmental processes, as opposed to being environmental “stress” responses. This possibility is supported by the observation of deformities in larvae reared in uncontaminated laboratory conditions (Servia et al. 2000). Insects undergo substantial physiological and morphological changes during moulting and metamorphosis, and changes in the appearance of larval mouthparts and antennae may be induced in the lead-up to these processes. For example, as metamorphosis approaches, insects produce digestive enzymes which break down parts of the cuticle for “recycling” (Gilbert 2000). Several chironomid deformities could conceivably be the result of cuticle dissolution, such as Kohn gaps in mouthparts [e.g. Figure 1C in Vermeulen (1995)], missing mentum teeth [e.g. Figure 2B in Janssens de Bisthoven et al. (1998)], split mentum teeth [e.g. Figure 4H in Warwick and Tisdale (1988)], reduced mentum teeth [e.g. Figure 4B in Warwick and Tisdale (1988)], and the

apparent erosion of terminal antennal segments [e.g. Figure 3G in Warwick (1985)].

Deformities in chironomid larvae may therefore be the result of metamorphosis/moulting-associated biochemical processes occurring as the larva transitions between instars, or towards the pupal stage.

Finally, the uncertainty surrounding deformity causes may also relate to questions of the statistical power of experiments. Sample sizes in published experiments vary widely (Gagliardi et al. 2016), however to our knowledge power analysis has very rarely, if ever, been provided alongside published deformity data.

This study aimed to clarify the causes of chironomid deformities, using the Australian model species *Chironomus tepperi*. An understanding of deformity incidence in laboratory studies has previously been limited by confounding factors associated with SB, a paucity of non-toxicological data on deformities, and a lack of reporting of statistical power. To address this, we considered deformities under sublethal toxicant exposure, and considered non-chemical and developmental factors. We evaluated aqueous metal (copper) and insecticide (imidacloprid) effects under sublethal conditions. These contaminants have previously been suggested as inducing deformities, although results appear inconsistent (Vermeulen 1995; Langer-Jaesrich et al. 2010), potentially due to SB (Gagliardi et al. 2016). We also evaluated a noncontaminant stressor, malnutrition, and a natural biological process, “onset of pupation”, or metamorphosis, as potential deformity-inducing factors under sublethal conditions. To address shortcomings in the literature relating to power analyses, and to determine the loss of power due low-levels of random mortality in our assays, we also conducted *a posteriori* power analyses for all assays.

MATERIALS AND METHODS

Culturing of larvae for experiments

Chironomus tepperi were cultured as described in Jeppe et al. (Jeppe et al. 2014) at 21 °C, using ethanol-rinsed toilet tissue as substrate, Tetramin as food, and “Modified Martins” medium (MM, hardness = 21–24 mg/L as CaCO₃) as culture medium. For the experiment, second instar larvae were reared in 2000 ml beakers, as per Mehler et al. (2017) and Kellar et al. (2014).

Stress experiment

This experiment involved three types of stresses: two aqueous chemical exposure conditions (imidacloprid and copper), a malnutrition condition, as well as well-fed chemical-free controls. We used the metal salt CuSO₄·5H₂O (Univar, analytical reagent) and the granular commercial imidacloprid formulation Confidor (Bayer, 50g/kg active ingredient) in chemical exposures. In experiments, larvae were fed a combination of alfalfa, Spirulina, and α-cellulose powders. This combination has previously been shown to be nutritious for chironomid larvae, conferring high survival, emergence and reproduction in cultures (Gagliardi et al. 2015). A food suspension (FS) was freshly prepared for each feed, comprising 1 g of each of the three powders in 100 ml MM. The food suspension was aliquoted while being mixed with a magnetic stirrer, to ensure an even suspension of food particles.

Deformity results may be confounded by SB: via mortality [if mortality rates are significantly different between treatments (Gagliardi et al. 2016)], instar stage [if larvae in different treatments develop to different instars in significant numbers (Servia et al. 2002)], or life stage [e.g. if individuals in different treatments

pupate/emerge in significantly different numbers (Hudson and Ciborowski 1996; Gagliardi et al. 2016)]. To avoid these confounding factors, we aimed to ensure that survival of larvae to fourth instar did not significantly differ across treatment concentrations within the experiment, and remained $\geq 80\%$ for all experimental treatments. We measured this at the termination of the experiment by counting recovered larvae. If $\geq 80\%$ of introduced larvae in each treatment were recovered (i.e. present and collected), all alive and at the fourth instar stage, we considered SB to be absent. We also ensured that, within each experimental condition, larval recovery rates were not confounded by stress (i.e. food/chemical) level in any treatment (see Methods, *Data analyses*). This was a further measure of SB.

To determine sublethal concentrations for the experiments, we conducted initial pre-experimental rangefinder tests. These involved three chemical concentrations (0.4, 4 and 40 mg/l for copper; 1.3, 13 and 130 $\mu\text{g/l}$ for imidacloprid) and controls (test chemical concentration of 0). The highest rangefinder concentration that produced recovery rates not significantly different from that in controls—and a recovery rate of $\geq 80\%$ —was used as the highest exposure treatment in the full experiment. These sublethal concentrations were 0.4 mg/l (copper) and 1.3 $\mu\text{g/l}$ (imidacloprid). The full experiment involved a series of five exposure treatments (dilution series = 0.6 \times) for each chemical (Table 1). For malnutrition treatments, we conducted a series of initial rangefinder experiments to determine an appropriate range of six FS volumes per feed, that would allow development of $\geq 80\%$ larvae from 2nd to 4th instar, with no significant association between recovery and food level. Larvae were fed twice during rangefinder tests and the experiments. We determined that the range 1600, 1040, 676, 440, 286, 186 μl FS per feed satisfied the criteria. We therefore used 1600 μl FS per feed in controls and chemical treatments to ensure larvae were well-fed, and 1040 –

186 μ l in malnutrition treatments (Table 1). Each stress condition therefore involved a gradient of five (OECD 2004) sublethal food or chemical levels, plus controls.

Exposure vessels were 600 ml glass beakers (three replicate beakers for each food or chemical level), containing 500 ml test solution (spiked MM for chemical exposures, unspiked MM for malnutrition and control conditions), and 10 g acid-rinsed sand (Chem-Supply, grain size = 300-350 μ m) as substrate. Ten (OECD 2004; Kumar and Simpson 2016) second instar larvae were introduced into each beaker with a plastic Pasteur pipette on Day 1 of the experiment. Beakers were incubated at 21 °C under a 16-h:8-h light:dark photoperiod and constant aeration. Experiments ran for five days [in a modification of the 5-d *C. tepperi* growth test procedure in Kumar and Simpson (2016)]. On Day 3, physicochemical water quality parameters [electrical conductivity (EC), dissolved oxygen (DO), pH; Mettler Toledo SevenExcellence multi-probe meter] were measured in each beaker. Each beaker was also given a 50 % test solution change on Day 3, and food added on Days 1 and 3.

On Day 5, experiments were terminated by pouring beaker contents through a sieve (pore size = 250 μ m). Sieve contents were rinsed into a petri dish with MM and survival was recorded. Larvae were scored as dead if they did not move upon gentle probing with a pair of forceps. Larvae not recovered were presumed dead/cannibalised. Larvae were then transferred into 1.5 ml Eppendorf tubes and killed and preserved in 70-100 % ethanol until deformity assessment.

Chemical analyses

The chemical treatments involved preparation of two sets of exposure test solution: one for Day 1, and one for Day 3 (for the solution changes). Chemical concentrations of each set of the test solutions were measured by commercial laboratories. Total copper in water was determined by Australian Laboratory Services (Scoresby, Victoria), using inductively-coupled plasma mass spectrometry (Perkin Elmer Elan 9000, US EPA Method 6020, NATA-accredited). Imidacloprid concentrations were determined by Advanced Analytical Australia (Kensington, Victoria), using liquid chromatography–tandem mass spectrometry (compliant with ISO/IEC 17025). These methods are described fully in Gagliardi et al. (2015). For statistical analysis of experimental results, chemical concentrations were recorded as the average concentration across the two media preparations (Table 1).

Assessment of pupation-associated effects

To assess whether deformity incidence increased as larvae neared pupation, we conducted a 10-d developmental timecourse experiment. Thirty beakers were set up and incubated under the same conditions as the controls in the stress experiment (Table 1). On Days 3, 6, and 9, beakers were given a 50 % solution change and fed 1600 µl FS. On Day 5, four replicate beakers were randomly selected, and larvae were analysed for recovery and deformities as per the stress experiment. These provided the control treatment for this timecourse experiment. Each day thereafter, 3-4 beakers were randomly selected, larvae were collected, and replicates analysed for larval recovery and deformity frequencies. Water quality in each beaker was determined at the time of termination. The experiment was terminated on Day 9, as recovery had reduced to < 80% (due primarily to pupation).

This article is protected by copyright. All rights reserved.

Deformity scoring

Larvae were slide-mounted under a light microscope, in Hoyer's fixative, following Gagliardi and Pettigrove (2013). Head capsules were observed under a compound microscope for deformities of the mentum, antennae and mandibles. Deformities were "blind" scored to avoid biasing results (Kozlov and Zvereva 2015; Salmelin et al. 2015). We assessed larvae for an abnormal number of mentum or mandible teeth, fusion or bifurcation of mentum or mandible teeth, Kohn gaps in the mandibles or mentum, abnormal numbers of antennal segments, and fusion of antennal segments. We did not score as deformed menta or mandibles that appeared worn or broken due to mechanical damage/friction (Salmelin et al. 2015).

If a larva did not have all required features (mentum, both mandibles and both antennae) visible, due to the arrangement of its features on the slide mount, it was scored as "deformed" if any of its visible features were deformed, but excluded from the analysis if none of its visible features were deformed, because it was impossible to determine the deformity status of the non-visible features. We also confirmed the instar of larvae by measuring mentum widths under the compound microscope with a calibrated graticule (McCauley 1974).

Data analyses

At the termination of each experiment, we quantified recovery rate (number of living 4th instar larvae present at the termination of the experiment, of the ten initially introduced larvae) for each beaker. We analysed treatment (food/chemical level, or day, including controls) effects upon recovery with binary logistic generalized linear models (GLMs) (Szöcs and Schäfer 2015) to test whether deformity frequencies were

not confounded by recovery (Gagliardi et al. 2016). A finding of no significant relationship between stress level and recovery rate ($p > 0.05$) would indicate we had successfully avoided SB.

We calculated deformity frequencies (no. recovered larvae with a deformed feature, per total no. recovered larvae) for each beaker. We then tested for differences in deformity frequency between each condition in the stress experiment (imidacloprid, malnutrition, and copper; pooling food/chemical levels within each condition) and controls, by GLM analysis, analysing condition as a categorical factor. This tested whether deformity frequencies were significantly different between control, copper, imidacloprid and malnutrition conditions. We then conducted GLM analysis within each condition, assessing for food/chemical treatment level (as continuous variables) effects among the five treatment levels plus controls, to assess each stressor for dose-response deformity effects.

For the timecourse experiment, we tested the effect of development day on deformity frequency, also using a GLM approach. This analysis included all development days (Days 5 – 9, Day as continuous variable) for which a recovery frequency of $\geq 80\%$ was observed. In all analyses, a significant association between any parameter and deformity frequency was inferred where $p < 0.05$.

Because some larvae were not recovered, and some larval features were not visible on the slide mounts (average nonrecovery plus nonvisibility = 5% larvae per treatment, S1 Supporting Information), sample sizes were in some cases lower than the intended 10 individuals total per replicate beaker (S1, Supporting Information). We therefore performed *a posteriori* power analyses for each of the GLMs investigating chemical, food, or development day effects upon deformity frequency. Our approach was

This article is protected by copyright. All rights reserved.

inspired by the methods described in Floate and Coghlin (2010). In summary, these tests involved simulations to assess what effect size we could detect with 80% power in each experiment. Effect size was defined as the slope of a linear relationship between the x and y axes. For a full description of these power analyses, see S4 (Supporting Information).

RESULTS

Stress experiment

All measured chemical concentrations were within 20 % of nominal concentrations (Table 1). Within each of the three conditions—copper, malnutrition and imidacloprid—no treatment effects upon recovery frequency were detected ($p > 0.05$, Supporting information S1), indicating deformity results were not confounded by mortality, instar, or pupation (Gagliardi et al. 2016). Hence there was no evidence of SB. In all treatments including controls, the number of surviving 4th instar larvae recovered was ≥ 80 % (i.e. ≥ 8 larvae per beaker mean), indicating experimental conditions were nonlethal [39] and further supporting an absence of SB.

Malnutrition and copper clearly induced stress in test larvae, as fourth instar larvae in the highest stress treatment in these experiments (Figures 1B and 1C, respectively) had much smaller bodies than control larvae (Figure 1A). There was no clear size discrepancy in imidacloprid-exposed larvae (Figure 1D), meaning it was less certain whether imidacloprid exposures were at stress-inducing doses. We observed deformities of the mentum, mandibles or antennae in the stress experiment in some individuals, with an overall deformity incidence of 10 % (4 % in controls, Figure 2). The GLM analysis comparing each stress condition (grouped across treatment levels)

indicated no significant deformity frequency differences between copper-exposed larvae and controls (z -score = 1.829, $p = 0.07$), nor between imidacloprid exposures (z -score = 1.146, $p = 0.25$) or malnutrition conditions and controls (z -score = 0.385, $p = 0.70$) (S1, Supporting Information).

In dose-response analyses, no significant effects of copper were observed (z -score = -1.558, $p = 0.12$). No significant dose-response effect of either imidacloprid (z -score = -0.102, $p = 0.92$) or malnutrition (z -score = -1.382, $p = 0.17$) upon deformity frequency levels were observed, despite some variability in deformity frequency across treatments (Figure 2).

Developmental timecourse experiment

Deformities were detected in this experiment, at a higher incidence than in the stress experiment (overall = 27%, control = 26%) (Figure 3). Recovery of larvae was $\geq 80\%$ (i.e. ≥ 8 surviving fourth instar larvae per beaker average) on Days 5 – 9 (Supporting information, S1). We therefore analysed results across Days 5 – 9. Recovery success was not significantly associated with day of development across these five days (z -score = -1.732, $p = 0.08$, S1, Supporting Information), indicating deformity results were not confounded by SB. There was no significant effect of development day (z -score = 0.981, $p = 0.33$, Figure 3) upon deformities, suggesting that onset of pupation is not a driver of chironomid larval deformities.

Power analyses

The chemical exposure experiments had reasonable power to detect underlying linear correlations between contaminant/food level and deformity rate, if they existed. This was assessed by a posteriori simulations (S4, Supplementary Information). For

copper, the experiment had 80% power (S4A, Supplementary Information) to detect a correlation with slope of 0.645 increase in deformity proportion/mgL⁻¹ copper. For imidacloprid, the experiment had 80% power (S4C, Supplementary Information) to detect a correlation with slope of 0.18 increase in deformity proportion/μgL⁻¹ imidacloprid (S4D, Supplementary Information). For both chemical contaminants, this is equivalent to a situation where the lowest treatment concentration tested induces deformities at a rate of approximately 0.08, or 2 × the background rate (as estimated by the deformity rate in the controls) and the highest induces deformities at a rate of approximately 0.28, or 7 × the background rate. The malnutrition experiment had 80% power to detect a relationship with a slope of 0.00017 decrease in deformity proportion/μL food (S4E, Supplementary Information). This corresponds to a situation where the lowest food amount tested induces deformities at approx. 7 × the background (fully-fed control) rate and the second-highest food amount induces deformities at approximately 4 × the background rate (S4F, Supplementary Information).

The developmental timecourse experiment had 80% power (S4G, Supporting Information) to detect a relationship with a slope of 0.0655 (units: increase in deformity proportion per day), or a situation where the larval population showed approx. 2 × the background (Day 5) deformity rate at Day 8 (S3H, Supporting Information).

In all simulations, the observed level of larva loss (through death, or through loss to observation via feature position on the slide) only slightly reduced experimental power compared to a situation where all 10 larvae were recovered for every beaker (S4A, C, E, G, Supporting Information). For example, for the copper experiment, the

full recovery simulation could detect a slope of 0.59 with 80% power, and a slope of 0.645 was detectable with 85% power for this simulation compared to 80% power for the partial recovery simulation (S4A, Supporting Information).

DISCUSSION

Stress experiment

In our experiment neither copper exposure, imidacloprid exposure, nor malnutrition induced chironomid larval deformities. Copper and malnutrition were clearly at stress-inducing levels, given their effect on body size (Figure 1A-C). This is less certain for imidacloprid (Figure 1D), which did not appear to affect body size and could potentially be further tested at higher (but still sublethal) concentrations. Larvae in this experiment exhibited a low [i.e. < 10% (Gagliardi et al. 2016)] control deformity frequency (4%, Figure 2), suggesting an absence of extraneous deformity-inducing factors and therefore a robust experiment. To our knowledge, the only other chemical assays meeting these criteria (i.e. < 10% control deformity level, < 20% mortality levels reported across all treatments, $p > 0.05$ for mortality effects) are the mercury, lead, and β -sitosterol assays of Vermeulen et al. (2000), which also observed no significant association between stressors and deformities. Together with other data recently published (Arambourou et al. 2014; Salmelin et al. 2015), these results cast some doubt upon the suggested links between environmental stress and chironomid deformities.

Although our GLM analysis did not detect dose-response effects, there was some apparent variability in the malnutrition condition. In particular there was an apparent increase in deformity frequency at 1040 μ l FS per feed in the malnutrition

experiment, with the frequency falling to background levels at 1600 μ l FS per feed (Figure 2). Such patterns are sometimes observed in deformity experiments in the literature, and may suggest non-monotonic effects, or simply a “noisy” association between stressors and deformity frequencies (Gagliardi et al. 2016). Further testing may therefore be justified to fully explore any possible nutritional effects.

Developmental timecourse experiment

In this experiment there was no effect of development day on deformity frequency across the five pre-pupation days. This suggests that onset of pupation is not a physiological driver of chironomid larval deformities. Previous studies have suggested a positive association between time and deformity incidence in toxicological assays (Martinez et al. 2004), however these tests may not have been under sublethal conditions, and hence may be confounded by SB.

We observed that larvae approaching pupation underwent whole-body morphological changes, observable to the naked eye. The most obvious of these was a thickening of the “thoracic” body region, noted previously in chironomid developmental descriptions (Vallenduuk and Pillot 2007). Deformities associated with pre-pupation might manifest at this time, however larvae rapidly transition through this phase into pupae (Vallenduuk and Pillot 2007). As such, a 24-h sampling cycle was not sufficient to capture most individuals at this stage of development as, for many individuals, this entire stage had transpired between sampling events. To better test the hypothesis that this whole-body morphological change was associated with mouthpart and antennal deformities, a more time-intensive sampling protocol (several times per day) might better allow for a more comprehensive understanding of changes occurring in larvae during the pre-pupal phase. Additionally, there did appear to be a

slight, though statistically nonsignificant, increase in the deformity rate on Day 9.

Given these factors, we suggest further testing of developmental effects (similar to nutritional effects) may be in order, to comprehensively understand any pupation-deformity links (or lack thereof).

It is noteworthy that there is an apparent difference in control deformity frequencies between the developmental timecourse experiment (26%) and the stress experiment (4%). The relatively high frequency in the timecourse experiment may suggest extraneous deformity-inducing factors compromising this experiment (Gagliardi et al. 2016). It is not clear why there is such a difference between the two experiments, as experimental conditions were identical. It may be a result of the different physicochemical water parameters between the two, although these were only slight (S1, Supporting Information).

Alternatively, a variable control deformity frequency could suggest that deformity frequencies are driven by intrinsic biological factors such as genetic or biochemical differences, independently of environmental conditions. The observation of considerable laboratory control deformity frequency variability across the chironomid deformity literature (0-68.9 %) is in line with this possibility (Gagliardi et al. 2016). If chironomid deformity rates are independent of environmental conditions such as contaminant levels, continued variability in control deformity frequencies will likely be observed across experiments.

Power analyses

In all cases, the modelled 80% power relationship between stressor and deformity rate appeared realistic to achieve within the bounds of the experiment, if such a relationship existed (S4B, D, F, H, Supporting Information). It required observing a mean deformity rate of around 0.3 in the most extreme treatments in the copper, imidacloprid and malnutrition experiments, and a higher mean deformity rate of around 0.8 in the Day 9 cohort of the timecourse experiment. Chironomid deformity frequencies in the literature vary widely. For example, laboratory control levels range from 0% to 68.9% (and nearly one quarter of published studies report a control rate of >10% deformities) (Gagliardi et al. 2016). Power analyses showed our stress tests as capable of detecting effects in experiments in top treatments at the middle and upper ranges of these levels. Hence these tests were of moderate strength.

The relatively higher deformity rate required for detection in the timecourse experiment is largely driven by the relatively high control deformity rate in this experiment, meaning a lesser change would be difficult to detect. This highlights the importance of undertaking power analyses in deformity experiments, given that control deformity frequencies are often high and can weaken the power of analyses [8]. In both experiment sets, weaker effects could be tested with higher levels of replication. Elimination of larvae from the experiment, due to non-recovery or nonvisibility on slide mounts, did not substantially reduce the statistical power of experiments (S4A, C, E, G). Overall, we suggest power analyses such as these can provide an important indicator of robustness in deformity experiments, and should therefore be considered in future studies.

Future work

The present experiment presents an experimental approach for tackling ongoing issues in chironomid deformity assessments, namely SB, understanding of noncontaminant factors, and uncertain statistical power. Although our sublethal stress experiments and previous SB-free results with low control deformity rates (Vermeulen et al. 2000) do not support stressors as causal to deformities, the causal effects of contaminants cannot yet be ruled out given that most other published data (both positive and negative) is apparently confounded by SB. There is therefore value in testing chemical combinations/types/concentrations not tested in these experiments, at sublethal levels (e.g. higher sublethal imidacloprid concentrations). Weaker effects can be tested at higher replication levels than in the present study. If non-significant or variable results continue to be observed in such experiments, it may ultimately suggest that contaminants do not induce chironomid deformities. When evaluating the overall likelihood of a causal association between contaminants and deformities, it is important for researchers to consider previous negative results alongside published positive results (Hanson et al. 2018).

We also suggest the need to understand the involvement of intrinsic biological factors—which are not necessarily linked to environmental conditions—in deformity formation. This may involve examining the genetic and biochemical differences between deformed and nondeformed larvae, within populations under nonstressful laboratory conditions (often referred to as “background deformity” incidence). Inbreeding depression (Vogt et al. 2013) and parental effects (Servia et al. 2002) are two intrinsic factors that have been suggested, and there may be other factors

involved, such as ploidy (Sadler et al. 2001) and differential metabolic profiles (Silva et al. 2011), which are associated with deformities in other taxa.

In addition, other potentially deformity-inducing stress factors are yet to be laboratory-tested. Factors such as larval density (Boglione et al. 2009), developmental rates (Reimer et al. 2017), dietary constituents (Cahu et al. 2003), and pathogens (Johnson et al. 1999) have demonstrated deformity-inducing effects in other taxa, and are worth further consideration. Developmental rates may be of particular interest, given the faint effects suggested in our nutritional and developmental tests. Notably also, parasites have previously been shown to induce morphological abnormalities in adult chironomids (Bhattacharya et al. 2014).

Survivorship bias

Survivorship bias confounds any lethality-inducing experiment that aims to test for a causal link between a chemical contaminant and a sublethal effect. We therefore suggest that any experiments with this aim must use sublethal doses, and that experiments that do induce mortality have questionable results. Concerns relating to SB have been highlighted in chironomid (Martinez et al. 2001; Dias et al. 2008), fish (Janz et al. 2010) and *Drosophila* (Polak et al. 2002) morphological assays, and in other areas of study such as endocrine-disrupting chemical assessments (Barata et al. 2004). However, it appears that the importance of avoiding SB is not universally understood, and should be given greater consideration in ecotoxicological studies. For example, the relatively common practice of running lethality (e.g. LC50) assays, and then assessing sublethal (e.g. EC50) causal effects among the remaining survivors, appears questionable. Although an EC50 is supposed to represent “a concentration that induces effects in 50 % of the population”, in a mortality-inducing experiment it

actually represents “a concentration that induces effects in 50 % of the chemical-resistant proportion of the population”. Avoiding SB instead involves first running an LC50 or rangefinder experiment, and then selecting sublethal concentrations for EC50s experiments based upon the LC50 experimental or rangefinder results. Alternatively, approaches in which lethality-inducing doses are eliminated from statistical analyses of sublethal endpoints also avoid SB.

While “mortality-prone” versus “mortality-resilient” larvae may differ phenotypically, it is not certain if the likelihood of exhibiting deformities is higher among the affected phenotypes (Martinez et al. 2001), although some data point to this possibility (Martinez et al. 2001; Dias et al. 2008). Unfortunately, killed larvae cannot be assessed for deformities (due to their premature death during the exposure). However, differences may be inferred through acute exposures (using short term, semi-lethal concentration exposures) and then determining whether the surviving population has higher or lower deformity frequencies, or by following the *Drosophila* experimental approaches outlined in Polak et al. (2002).

REFERENCES

- Abdel-Moneim A, Moreira-Santos M, Ribeiro R. 2015. A short-term sublethal toxicity assay with zebra fish based on preying rate and its integration with mortality. *Chemosphere* 120:568-574.
- Arambourou H, Beisel J-N, Branchu P, Debat V. 2014. Exposure to sediments from polluted rivers has limited phenotypic effects on larvae and adults of *Chironomus riparius*. *Sci. Total Environ.* 484:92-101.

Barata C, Porte C, Baird DJ. 2004. Experimental designs to assess endocrine disrupting effects in invertebrates a review. *Ecotoxicology* 13:511-517.

Bhattacharya G, Mazumdar A, Hazra N. 2014. Pathological changes in mermithid nematode infected *Stictochironomus polystictus* (Kieffer) (Diptera, Chironomidae, Chironominae). *Proceedings of the Zoological Society* 67:153-157.

Boglione C, Marino G, Giganti M, Longobardi A, De Marzi P, Cataudella S. 2009. Skeletal anomalies in dusky grouper *Epinephelus marginatus* (Lowe 1834) juveniles reared with different methodologies and larval densities. *Aquaculture* 291:48-60.

Cahu C, Infante JZ, Takeuchi T. 2003. Nutritional components affecting skeletal development in fish larvae. *Aquaculture* 227:245-258.

Dalzochio T, Gehlen G. 2016. Confounding factors in biomonitoring using fish. *Ecotoxicol. Environ. Contam.* 11:53-61.

de Haas EM, Van Haaren R, Koelmans AA, Kraak MH, Admiraal W. 2005. Analyzing the causes for the persistence of chironomids in floodplain lake sediments. *Arch. Hydrobiol.* 162:211-228.

Di Veroli A, Santoro F, Pallottini M, Selvaggi R, Scardazza F, Cappelletti D, Goretti E. 2014. Deformities of chironomid larvae and heavy metal pollution: From laboratory to field studies. *Chemosphere* 112:9-17.

Dias V, Vasseur C, Bonzom JM. 2008. Exposure of *Chironomus riparius* larvae to uranium: Effects on survival, development time, growth, and mouthpart deformities. *Chemosphere* 71:574-581.

Floate KD, Coghlin PC. 2010. No support for fluctuating asymmetry as a biomarker of chemical residues in livestock dung. *Can. Entomol.* 142:354-368.

Fukudome EY, Kochanek AR, Li Y, Smith EJ, Liu B, Kheirbek T, Lu J, Kim K, Hamwi K, Velmahos GC. 2010. Pharmacologic resuscitation promotes survival and attenuates hemorrhage-induced activation of extracellular signal-regulated kinase 1/21. *J. Surg. Res.* 163:118-126.

Gagliardi B, Pettigrove V. 2013. Removal of intensive agriculture from the landscape improves aquatic ecosystem health. *Agric. Ecosyst. Environ.* 176:1-8.

Gagliardi BS, Long SM, Pettigrove VJ, Hoffmann AA. 2015. The parthenogenetic cosmopolitan chironomid, *Paratanytarsus grimmii*, as a new standard test species for ecotoxicology: Culturing methodology and sensitivity to aqueous pollutants. *Bull. Environ. Contam. Toxicol.*:1-7.

Gagliardi BS, Pettigrove VJ, Long SM, Hoffmann AA. 2016. A meta-analysis evaluating the relationship between aquatic contaminants and chironomid larval deformities in laboratory studies. *Environ. Sci. Technol.*

Gilbert SF. 2000. Metamorphosis: The hormonal reactivation of development. In Gilbert SF, ed, *Developmental biology*, 6th ed. Sinauer Associates, Sunderland (MA).

Hanson M, Deeth L, Prosser R. 2018. Evidence of citation bias in the pesticide ecotoxicology literature. *Ecotoxicology*:1-7.

Hudson LA, Ciborowski JJH. 1996. Teratogenic and genotoxic responses of larval *Chironomus salinarius* group (Diptera: Chironomidae) to contaminated sediment. *Environ. Toxicol. Chem.* 15:1375-1381.

This article is protected by copyright. All rights reserved.

Janssens de Bisthoven L, Gerhardt A. 2005. A critical appraisal of morphological deformities in midge larvae (Insecta, Nematocera: Chironomidae): Clues for biomonitoring the sediment quality of freshwater ecosystems. *Comp. Biochem. Physiol.*:27-36.

Janssens de Bisthoven L, Nuyts P, Goddeeris B, Ollevier F. 1998. Sublethal parameters in morphologically deformed *Chironomus* larvae: Clues to understanding their bioindicator value. *Freshwater Biol.* 39:179-191.

Janssens de Bisthoven L, Postma J, Vermeulen A, Goemans G, Ollevier F. 2001. Morphological deformities in *Chironomus riparius* Meigen larvae after exposure to cadmium over several generations. *Water, Air, and Soil Pollution* 129:167-179.

Janz DM, DeForest DK, Brooks ML, Chapman PM, Gilron G, Hoff D, Hopkins WA, McIntyre DO, Mebane CA, Palace VP. 2010. Selenium toxicity to aquatic organisms. In Chapman PM, Adams WJ, Brooks ML, Delos CG, Luoma SL, Maher WA, Ohlendorf HM, Presser TS, Shaw P, eds, *Ecological assessment of selenium in the aquatic environment*. pp 141-231.

Jeppe KJ, Carew ME, Long SM, Lee SF, Pettigrove V, Hoffmann AA. 2014. Genes involved in cysteine metabolism of *Chironomus tepperi* are regulated differently by copper and by cadmium. *Comp. Biochem. Physiol. C Pharmacol. Toxicol. Endocrinol.* 162:1-6.

Johnson PT, Lunde KB, Ritchie EG, Launer AE. 1999. The effect of trematode infection on amphibian limb development and survivorship. *Science* 284:802-804.

Kellar CR, Hassell KL, Long SM, Myers JH, Golding L, Rose G, Kumar A, Hoffmann AA, Pettigrove V. 2014. Ecological evidence links adverse biological

This article is protected by copyright. All rights reserved.

effects to pesticide and metal contamination in an urban australian watershed. *J. Appl. Ecol.* 51:426-439.

Kozlov MV, Zvereva EL. 2015. Confirmation bias in studies of fluctuating asymmetry. *Ecol. Indic.* 57:293-297.

Kumar A, Simpson SL. 2016. Appendix h: Protocol for 10-day whole-sediment sub-lethal and acute toxicity tests using the freshwater chironomid *Chironomus tepperi*. In Simpson SL, Batley G, eds, *Sediment quality assessment: A practical guide*. CSIRO Publishing, Clayton South, p. 294.

Langer-Jaesrich M, Koehler H-R, Gerhardt A. 2010. Can mouth part deformities of *Chironomus riparius* serve as indicators for water and sediment pollution? A laboratory approach. *J. Soils Sediments* 10:414-422.

Martinez EA, Moore BC, Schaumloffel J, Dasgupta N. 2004. Effects of exposure to a combination of zinc- and lead-spiked sediments on mouthpart development and growth in *Chironomus tentans*. *Environ. Toxicol. Chem.* 23:662-667.

Martinez EA, Moore BC, Schaumloffel J, Dasgupta N. 2001. Induction of morphological deformities in *Chironomus tentans* exposed to zinc-and lead-spiked sediments. *Environ. Toxicol. Chem.* 20:2475-2481.

McCauley VJ. 1974. Instar differentiation in larval Chironomidae (Diptera). *Can. Entomol.* 106:179-200.

Mehler WT, Keough MJ, Pettigrove V. 2017. Development of whole-sediment toxicity identification and evaluation (TIE) techniques for two australian freshwater

species: *Chironomus tepperi* and *Austrochiltonia subtenuis*. *Environ. Toxicol. Chem.* 36:2476-2484.

Milošević D, Čerba D, Szekeres J, Csányi B, Tubić B, Simić V, Paunović M. 2016. Artificial neural networks as an indicator search engine: The visualization of natural and man-caused taxa variability. *Ecol. Indic.* 61:777-789.

Odume ON, Palmer CG, Arimoro FO, Mensah PK. 2016. Chironomid assemblage structure and morphological response to pollution in an effluent-impacted river, eastern cape, south africa. *Ecol. Indic.* 67:391-402.

OECD. 2004. Test no. 219: Sediment-water chironomid toxicity using spiked water. OECD Publishing.

Polak M, Opoka R, Cartwright IL. 2002. Response of fluctuating asymmetry to arsenic toxicity: Support for the developmental selection hypothesis. *Environ. Pollut.* 118:19-28.

Qi Y, Yan B, Fu G, Guan X, Du L, Li J. 2017. Germination of seeds and seedling growth of *Amaranthus retroflexus* L. Following sublethal exposure of parent plants to herbicides. *Sci. Rep.* 7:157.

Reimer T, Dempster T, Wargelius A, Fjellidal PG, Hansen T, Glover KA, Solberg MF, Swearer SE. 2017. Rapid growth causes abnormal vaterite formation in farmed fish otoliths. *J. Exp. Biol.* 220:2965-2969.

Sadler J, Pankhurst P, King H. 2001. High prevalence of skeletal deformity and reduced gill surface area in triploid atlantic salmon (*Salmo salar* L.). *Aquaculture* 198:369-386.

This article is protected by copyright. All rights reserved.

- Salmelin J, Vuori KM, Hämäläinen H. 2015. Inconsistency in the analysis of morphological deformities in Chironomidae (insecta: Diptera) larvae. *Environ. Toxicol. Chem.* 34:1891-1898.
- Servia MJ, Cobo F, Gonzalez MA. 2000. Incidence and causes of deformities in recently hatched larvae of *Chironomus riparius* meigen, 1804 (diptera, chironomidae). *Arch. Hydrobiol.* 149:387-401.
- Servia MJ, Cobo F, González MA. 2002. Monitoring deformities of *Chironomus riparius* Meigen, 1804 (Diptera, Chironomidae) larvae from first to fourth instar: A potential tool for investigating their origin. *Arch. Hydrobiol.* 154:447-460.
- Silva TS, Cordeiro O, Richard N, Conceição LE, Rodrigues PM. 2011. Changes in the soluble bone proteome of reared white seabream (*Diplodus sargus*) with skeletal deformities. *Comp. Biochem. Physiol. Part D Genomics Proteomics* 6:82-91.
- Szöcs E, Schäfer RB. 2015. Ecotoxicology is not normal. *Environ. Sci. Pollut. Res.* 22:13990-13999.
- Townsend KR, Pettigrove VJ, Hoffmann AA. 2012. Food limitation in *Chironomus tepperi*: Effects on survival, sex ratios and development across two generations. *Ecotoxicol. Environ. Saf.* 84:1-8.
- Vallenduuk HJ, Pillot HKM. 2007. *Chironomidae larvae, vol. 1: Tanypodinae: General ecology and tanypodinae*. Brill.
- Vermeulen AC. 1995. Elaborating chironomid deformities as bioindicators of toxic sediment stress: The potential application of mixture toxicity concepts. *Ann. Zool. Fennici* 32:265-285.

Vermeulen AC, Dall P, Lindegaard C, Olliver F, Goddeeris BR. 2000. Exposure of *Chironomus riparius* larvae (Diptera) to lead, mercury and β -sitosterol: Effect on mouthpart deformation and moulting. *Chemosphere* 41:45-55.

Vogt C, Langer-Jaesrich M, Elsässer O, Schmitt C, Van Dongen S, Köhler HR, Oehlmann J, Nowak C. 2013. Effects of inbreeding on mouthpart deformities of *Chironomus riparius* under sublethal pesticide exposure. *Environ. Toxicol. Chem.* 32:423-425.

Warwick W. 1985. Morphological abnormalities in Chironomidae (Diptera) larvae as measures of toxic stress in freshwater ecosystems: Indexing antennal deformities in *Chironomus* Meigen. *Can. J. Fish Aquat. Sci* 42:1881-1914.

Warwick WF, Tisdale NA. 1988. Morphological deformities in *Chironomus*, *Cryptochironomus*, and *Procladius* larvae (Diptera: Chironomidae) from two differentially stressed sites in Tobin Lake, Saskatchewan. *Can. J. Fish Aquat. Sci* 45:1123-1144.

FIGURES

Figure 1 Fourth instar *Chironomus tepperi* larvae recovered at the termination of the stress experiment. Control larvae (A.) had much larger bodies than larvae exposed to the highest copper concentration (B.), or those exposed to the lowest food amount (C.), but were comparable to larvae exposed to the highest imidacloprid concentration (D.).

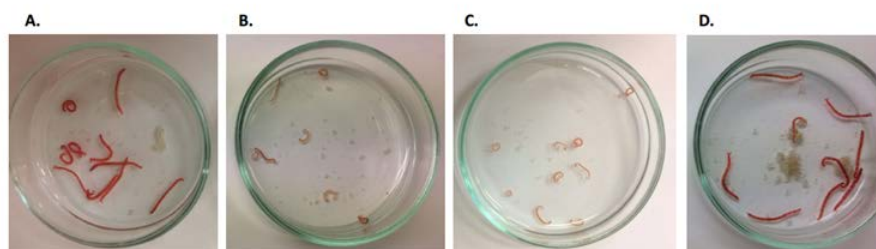


Table 1 Experimental design for stress and developmental timecourse experiments, including chemical concentrations [mean (SD)] for contaminant exposures

Experiment	Condition	Treatment	n (replicate beakers)	Copper concentration (mg/l)		Imidacloprid concentration (µg/l)		Food amount (µl food suspension per feed)	
				Nominal	Measured	Nominal	Measured		
Stress	Control	C	3	0	< 0.001	0	< 0.01	1600	
		Copper exposure	T1	3	0.05	0.05 (0.005)	-	-	1600
	T2		3	0.09	0.08 (0.01)	-	-	1600	
	T3		3	0.14	0.13 (0.01)	-	-	1600	
	T4		3	0.24	0.22 (0.01)	-	-	1600	
	T5		3	0.40	0.36 (0.02)	-	-	1600	
	Imidacloprid exposure	T1	3	-	-	0.17	0.20 (0.2)	1600	
		T2	3	-	-	0.28	0.30 (0.007)	1600	
		T3	3	-	-	0.47	0.49 (0.01)	1600	
		T4	3	-	-	0.78	0.79 (0.007)	1600	
		T5	3	-	-	1.30	1.35 (0.07)	1600	
	Malnutrition	T1	T1	3	-	-	-	-	186
			T2	3	-	-	-	-	286

		T3	3	-	-	-	-	440
		T4	3	-	-	-	-	676
		T5	3	-	-	-	-	1040
Developmental timecourse	Onset of pupation	Day 5 (Control)	4	-	-	-	-	1600
		Day 6	4	-	-	-	-	1600
		Day 7	4	-	-	-	-	1600
		Day 8	4	-	-	-	-	1600
		Day 9	3	-	-	-	-	1600
		Day 10	4	-	-	-	-	1600
