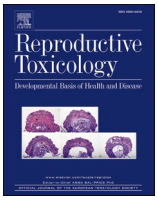
Reproductive Toxicology 127 (2024) 108615

Contents lists available at ScienceDirect

Reproductive Toxicology

journal homepage: www.elsevier.com/locate/reprotox



Biological variability hampers the use of skeletal staining methods in zebrafish embryo developmental toxicity assays

Jente Hoyberghs [a,1], Jonathan Ball [b,1], Maciej Trznadel [b], Manon Beekhuijzen [c], Matthew Burbank [d], Pia Wilhelmi [e], Arantza Muriana [f], Nicola Powles-Glover [g], Ainhoa Letamendia [h], Steven Van Cruchten [a,*]

[a] University of Antwerp, Wilrijk, Belgium
[b] University of Exeter, Exeter, United Kingdom
[c] Charles River Laboratories, 's Hertogenbosch, the Netherlands
[d] L'Oréal Research & Innovation, Paris, France
[e] BASF SE, Ludwigshafen am Rhein, Germany
[f] BBD BioPhenix S.L. - Biobide, San Sebastian, Spain
[g] Astrazeneca UK Ltd, Cambridge, UK
[h] Bayer S.A.S., Lyon, France

ARTICLE INFO

Handling Editor: Dr Piersma Aldert

Keywords:
NAM
Developmental toxicity
Zebrafish larvae
Bone staining
Cartilage staining

ABSTRACT

Zebrafish embryo assays are used by pharmaceutical and chemical companies as new approach methodologies (NAMs) in developmental toxicity screening. Despite an overall high concordance of zebrafish embryo assays with *in vivo* mammalian studies, false negative and false positive results have been reported. False negative results in risk assessment models are of particular concern for human safety, as developmental anomalies may be missed. Interestingly, for several chemicals and drugs that were reported to be false negative in zebrafish, skeletal findings were noted in the *in vivo* studies. As the number of skeletal endpoints assessed in zebrafish is very limited compared to the *in vivo* mammalian studies, the aim of this study was to investigate whether the sensitivity could be increased by including a skeletal staining method. Three staining methods were tested on zebrafish embryos that were exposed to four teratogens that caused skeletal anomalies in rats and/or rabbits and were false negative in zebrafish embryo assays. These methods included a fixed alizarin red-alcian blue staining, a calcein staining, and a live alizarin red staining. The results showed a high variability in staining intensity of larvae exposed to mammalian skeletal teratogens, as well as variability between control larvae originating from the same clutch of zebrafish. Hence, biological variability in (onset of) bone development in zebrafish hampers the detection of (subtle) treatment-related bone effects that are not picked-up by gross morphology. In conclusion, the used skeletal staining methods did not increase the sensitivity of zebrafish embryo developmental toxicity assays.

1. Introduction

In recent years, new approach methodologies (NAMs) for hazard and risk assessment of xenobiotics have received a lot of attention [1]. At this moment, three NAMs for the assessment of developmental toxicity have been validated by the European Centre for the Validation of Alternative Methods (ECVAM): the rat Whole Embryo Culture Test (rWEC), the mouse Embryonic Stem Cell Test (mEST) and the Micromass Teratogen Test (MM) [2,3]. Although not validated yet, developmental toxicity assays using zebrafish embryos are currently also used for screening purposes by several pharmaceutical, (agro)chemical and cosmetic companies [4–15]. Its greatest advantage compared to the other three NAMs is that developmental effects can be assessed in a whole vertebrate organism during the main organogenesis period [16,17]. After all, the mEST and MM do not allow assessment in a whole organism, while the WEC only allows assessment during a short period of the organogenesis period (i.e., only for 24–48 h). Moreover, in the European Union, studies on zebrafish embryos are not legally considered as animal

* Correspondence to: Universiteitsplein 1, Wilrijk, Antwerp 2610, Belgium.
E-mail address: steven.vancruchten@uantwerpen.be (S. Van Cruchten).
[1] Both authors contributed equally.

https://doi.org/10.1016/j.reprotox.2024.108615
Received 29 January 2024; Received in revised form 18 April 2024; Accepted 23 May 2024
Available online 28 May 2024

Table 1
Overview of the different experiments. For each staining method, the compound concentrations, control media tested, and the developmental stage at which the staining was performed are indicated. Abbreviations: alcian blue (AB), alizarin red (AR), days post-fertilization (dpf), not applicable (NA).

Staining method	Chemical	Concentration	Controls	Age
	Compound 5	10, 100 and 1,000 μM	Medium, DMSO	5 dpf
	Compound 9	10, 100 and 1,000 μM	Medium, DMSO	5 dpf
	Compound A	1, 10 and 100 μM	Medium, DMSO	5 dpf
AR-AB (fixed)	Levetiracetam	10, 100 and 1,000 μM	Medium, DMSO	5 dpf
Calcein (live)	NA	NA	Medium, DMSO	5, 9 dpf
AR (live)	Compound 5	100 and 1,000 μM	DMSO	5, 9 dpf

Table 2
List of bone structures that can be scored in 5 dpf and 9 dpf zebrafish larvae when using an AR-AB fixed or an AR live staining. Bones that are likely to be present according to literature [32–35] and/or are clearly visible when stained are included in this list.

Bone structure	Fixed – 5 dpf	AR live – 5 dpf	AR live – 9 dpf
Notochord (n)	X	x	x
Parasphenoid (ps)	X	x	x
Vertebrae (vb)	/	/	x
Ceratobranchial 5 (cb5)	X	x	x
Pharyngeal teeth (t)	X	-[1]	-[1]
Utricular otolith (uot)	X	x/-	x/-
Circle saccular otolith (cot)	X	x/-	x/-
Entopterygoid (en)	X	x	x
Opercle (op)	X	x	x
Cleithrum (c)	X	x	x
Branchiostegal rays (brs)	X	x	x
Branchiostegal rays II (brs2)	/[2]	/[2]	x
Hyomandibular bone (hmb)	/[2]	/	x
Ceratohyal bone (chb)	/[2]	/	x
Dentary (den)	/[2]	/	x
Maxilla (max)	/[2]	/	x
Anguloarticular (aa)	/[2]	/	x

Often present/scored (x), less likely to be present/scored if present (/), not present or visible/not scored (-), only one side could be scored (x/-). [1]Structure not distinguishable from cb5. [2] Structure was never visible at this age using this staining. Abbreviations: alcian blue (AB), alizarin red (AR), days post-fertilization (dpf).

Table 3
List of cartilage structures that can be scored in 5 dpf zebrafish larvae when using an alizarin red-alcian blue fixed staining. Cartilages that are present according to literature [34–37] and/or are clearly visible when stained are included in this list.

Cartilage structure	5 dpf
Ethmoid plate (eth)	X
Basihyal (bh)	X
Basibranchial (bb)	X
Meckel's cartilage (mk)	X
Palatoquadrate (qu)	X
Hyosymplectic (hys)	X
Ceratohyal (ch)	X
Ceratobranchials (cb)	X
Auditory capsule (aud)	X
Pectoral fin (pec fin)	X
Pterygoid process of the quadrate (pty)	X

Should be present/scored (x). Abbreviations: days post-fertilization (dpf).

experiments up to the free feeding stage (i.e., 5 days post-fertilization or dpf) (EU Directive 2010/63). Thus, the zebrafish embryo represents a holistic model that aligns well with the 3R principle for the development of NAMs [5,18].

To date, the use of NAMs in regulatory submissions for developmental and reproductive toxicology (DART) testing of pharmaceuticals has been, and is still very limited, although the third revision of the ICH S5 guideline on detection of toxicity to reproduction for human pharmaceuticals provides opportunities to do so. The guideline does not list any specific NAMs to be used, but says they should be properly qualified [19]: *"If properly qualified, alternative assays have the potential to defer or replace (in certain circumstances) conventional in vivo studies. Approaches that incorporate alternative assays should provide a level of confidence for human safety assurance at least equivalent to that provided by the current testing paradigms."* From the above, it is clear that for zebrafish embryo assays to be considered for regulatory submissions, its potential to detect human teratogens should be well qualified.

Despite an overall high concordance (80–85% [9,20]) of zebrafish embryo assays with the *in vivo* mammalian studies, false negative and false positive results were reported [4,6–12,14,15,17,21]. In particular these false negative results hinder the use for regulatory purposes, as potential teratogens may be missed. For many chemicals and drugs that were reported to be false negative in zebrafish embryo assays, skeletal findings were noted in the *in vivo* studies. In contrast to the exhaustive list of skeletal endpoints assessed in rat and rabbit embryofetal development (EFD) studies [22], in which skeletal staining is a standard procedure during examination, no skeletal endpoints are routinely assessed in zebrafish embryo developmental toxicity assays. As skeletal staining of zebrafish larvae is feasible and different methods have been described before (see Supplementary table 1) [23,24], the aim of this study was to investigate whether the sensitivity (i.e., the ability to detect true teratogens) of zebrafish embryo developmental toxicity assays can be increased by including an extended skeletal assessment. For this purpose, we selected four pharmaceutical compounds (i.e., levetiracetam and proprietary compounds 5, 9, and A) that showed skeletal malformations in rat and/or rabbit fetuses, but were false negative in zebrafish embryo assays [8,10]. As several staining methods for zebrafish embryos are reported in literature, each with different (dis)advantages (see Supplementary table 1), three (i.e., alizarin red (AR)-alcian blue (AB), calcein and alizarin red live) were tested to identify the most suitable staining method for increasing the sensitivity of zebrafish embryo assays.

2. Materials and methods

2.1. Chemicals and solutions

Unless otherwise stated, all test chemicals were purchased from Sigma-Aldrich (Missouri, USA).

- Embryo medium (EM), a 0.3x Danieau's solution, was prepared from a 10x stock solution containing: 580 mM NaCl, 7.0 mM KCl, 4 mM $MgSO_4$, 6 mM $Ca(NO_3)_2$, and 50 mM HEPES (Invitrogen, Massachusetts, USA). The filtered (0.2 μM filter) 0.3x Danieau's solution was made from this 10x stock by adding ultrapure water, and the pH was adjusted to 7.3 ± 0.2 with 1 M NaOH (Thermo fisher, New Hampshire, USA).
- The MS-222 solutions (4 mg/mL and 0.2 mg/mL) were made by dissolving methyl ethane sulfonate in EM, and the pH was adjusted to 7.4 ± 0.3 with 1 M NaOH.
- Four mammalian skeletal teratogens that showed skeletal malformations in EFD studies in rat and/or rabbit but were false negative in zebrafish assays were selected to expose the zebrafish embryos to. Levetiracetam (L-8668–50MG, Sigma-Aldrich) (10, 100 and 1,000 μM), "proprietary compound 5" (10, 100 and 1,000 μM), "proprietary compound 9" (10, 100 and 1,000 μM) and "proprietary

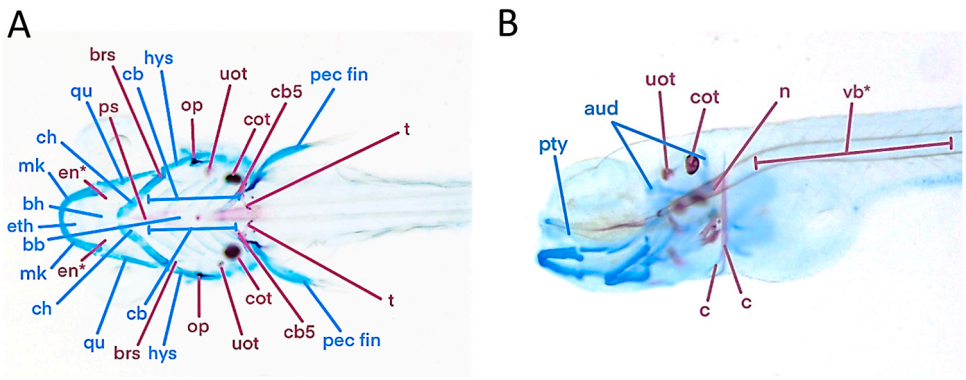


Fig. 1. Zebrafish larvae at 5 dpf with bone (red) and cartilage (blue) structures that were stained with the alizarin red- alcian blue staining. Left panel (A) shows a ventral view. Right panel (B) shows a lateral view. Each structure (except for uot and cot) is indicated in only one of the orientations, namely in the orientation where it was scored. Uot and cot were scored by looking at both positions. The abbreviations are depicted in Table 2 and Table 3. The structures vb* and en* are not present in these images.

Table 4
Intensity and shape scores and their meaning.

Intensity scoring		Shape scoring	
0	structure not stained/not present	**0**	structure is normal
1	structure is weakly stained	**1**	structure is malformed
2	structure is moderately stained		
3	structure is heavily stained		

compound A" (1 μM, 10 μM and 100 μM) were used in the AR-AB staining experiments, and "proprietary compound 5" (100 and 1,000 μM) in the AR live staining experiments [8,10]. The test concentrations were prepared by dissolving the compound in EM containing 0.5% DMSO.

The chemicals and solutions needed for the staining protocol are described in Section 2.6 Staining protocol.

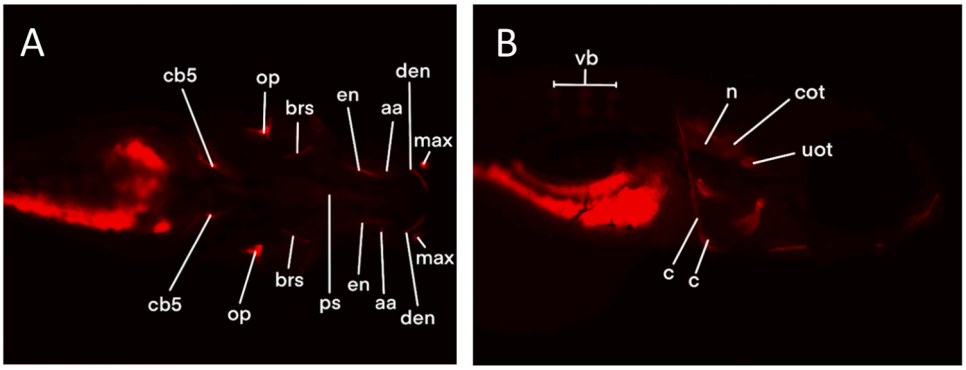


Fig. 2. Zebrafish larvae at 5 dpf with bone structures that were stained with the alizarin red live staining. Left panel (A) shows a ventral view. Right panel (B) shows a lateral view. Each structure is indicated in only one of the orientations, namely in the orientation where it was scored. The abbreviations are depicted in Table 2.

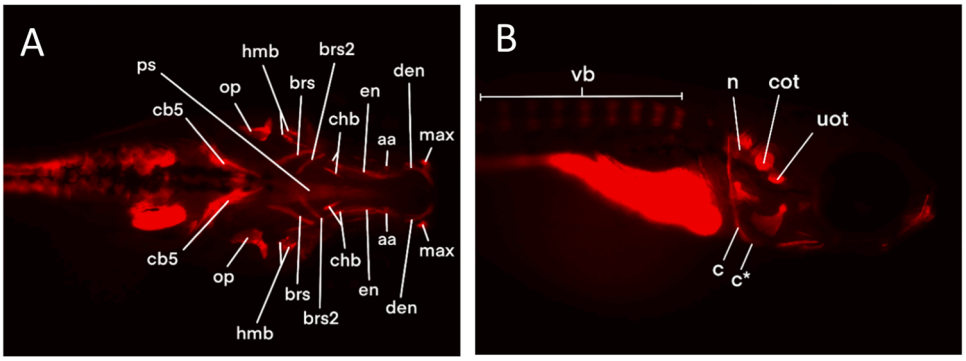


Fig. 3. Zebrafish larvae at 9 dpf with bone structures that were stained with the alizarin red live staining. Left panel (A) shows a ventral view. Right panel (B) shows a lateral view. Each structure is indicated in only one of the orientations, namely in the orientation where it was scored. The abbreviations are depicted in Table 2. The left cleithrum (c*) is hidden behind the right cleithrum in this image.

Table 5
Uptake by zebrafish at 1 dpf (A) and 5 dpf (B).

A) Exposure 0–1 dpf Compound	Conc. in each Larva (μM)	Uptake as %[1]	Conc. in each Larva (μM)	Uptake as %[1]	Conc. in each Larva (μM)	Uptake as %[1]
	100 μM		10 μM		1 μM	
Compound A	0.667 ± 0.00036[a]	1 ± 0.36[a]	0.953 ± 0.00012	10 ± 1.19	<LOQ	
Levetiracetam	18.023 ± 0.00079	18 ± 0.79	1.401 ± 0.00032	14 ± 3.24	<LOQ	
Compound 9	30.651 ± 0.00365	31 ± 3.65	3.114 ± 0.00011	31 ± 1.13	0.458 ± 0.00054	46 ± 5.36
Compound 5	28.967 ± 0.00247	29 ± 2.47	2.463 ± 0.00012	24 ± 1.17	0.326 ± 0.00001	33 ± 1.02

B) Exposure 0–5 dpf Compound	Conc. in each Larva (μM)	Uptake as %[1]	Conc. in each Larva (μM)	Uptake as %[1]	Conc. in each Larva (μM)	Uptake as %[1]	Stability
	100 μM		10 μM		1 μM		Day 5 as % of Day 0
Compound A	1.494 ± 0.00037[a]	1 ± 0.37[a]	1.773 ± 0.00014	18 ± 1.44	<LOQ		104
Levetiracetam	7.498 ± 0.00093	7 ± 0.92	0.480 ± 0.00007	5 ± 0.7	<LOQ		110
Compound 9	1.420 ± 0.00008	1 ± 0.08	<LOQ		<LOQ		106
Compound 5	15.818 ± 0.00148	16 ± 1.48	1.0548 ± 0.00007	11 ± 0.67	<LOQ		105

<LOQ The analysis concentrates below the level of quantification.
[1] Uptake as % is the ratio of measured concentration in each larva (μM)/the nominal concentration *100 (e.g., 0.953 μM/10*100)
The concentration in each larva represents the mean of three replicates, of which each replicate consists out of 4 larvae (i.e., the value of each replicate was divided by 4 to obtain estimations for each larva). The SEM of the replicates is indicated.
Uptake assessment at 1,000 μM was not performed due to excessive precipitation of the compounds in the solution, which would impact the background level of compound carryover in solution when assessing uptake.
Abbreviation: days post-fertilization (dpf).
[a] Precipitation observed at point of sampling of tissue for analysis.

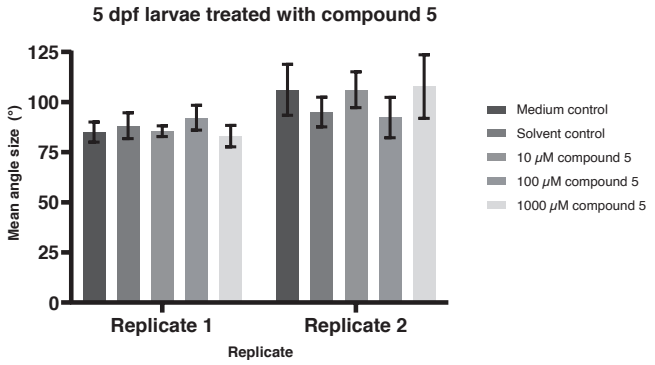


Fig. 4. Mean angle sizes between ceratohyal cartilages in 5 dpf larvae treated with compound 5. All groups were compared to the solvent control group. No significant differences were observed between the 3 test groups or the medium control group and the solvent control group.

2.2. Animal care and egg collection

Breeding stocks of healthy, unexposed adult zebrafish from the wild-type WIK strain (sourced from a historical line at Brixham AstraZeneca and outcrossed with WIK ZERC) were used to produce fertilized eggs. Water was maintained at pH 7.35 ± 0.65 and 28 ± 1°C. Fish were cultured in the aquarium facility with a 14 h light:10 h dark light cycle. Adult fish at a ratio of 2:1 females to males were placed into spawning tanks on the evening prior to the day of culture, and egg traps were positioned within each tank. Eggs were collected the next morning soon after spawning and incubated in system water as detailed in Paull et al., 2008 [25] at 28 (±1)°C for approximately 1–2 h, and then treated against fungal infection using a diluted Chloramine T bleaching solution (10 g/L) for 60 s with gentle periodic agitation. Following bleaching, the embryos were washed twice in rig water with constant agitation and then transferred into a Petri dish containing 0.3x Danieau's solution. Between 100 and 200 fertilized eggs of the same developmental stage (i.e., a stage before 4 hpf) were transferred into a separate Petri dish containing 0.3x Danieau's solution and maintained at 28 (±1)°C. Embryos were staged for development according to Kimmel et al. (1995) [26].

2.3. Uptake assessment

Uptake of the test solutions was assessed as detailed previously in Gustafson et al. (2012) [9]. Briefly, liquid chromatography–tandem mass spectrometry (LC–MS/MS) was used to determine the uptake of the compound into embryos of 1 dpf and larvae of 5 dpf. Uptake assessment in embryos and larvae was originally planned for 4 concentrations (i.e., 1, 10, 100 and 1,000 μM) of each test compound, and for a corresponding 0.5% DMSO solvent control (see 2.4 Exposure of zebrafish embryos). However, uptake at 1,000 μM was not assessed due to excessive precipitation of the test compounds in the solution (see results section for more details). After 24 h (30 hpf) or 5 days (120 hpf) of exposure the embryos/larvae were assessed for viability. For this, the following criteria for lethality were used: coagulation at 24 hpf, and coagulation, the lack of somite formation and the lack of heat beat at 120 hpf. After viability assessment, the embryos/larvae are transferred to a filter plate (3–4 embryos/larvae per well) in a final volume of 300 μL of EM, in triplicates. Embryos were washed using system water (containing 0.5 g/L tricaine at pH 7 (± 0.5)) under vacuum, before being transferred to a deep-well plate containing 300 μL of rig water and homogenized (Geno/grinder, SPEX Certiprep L.L.C., USA). 300 μL acetonitrile (HPLC grade Thermo Fisher, New Hampshire, USA) containing 50 nM of internal standard was added to each well followed by a second homogenization, and then addition of 900 μL (HPLC grade) water (Thermo Fisher, New Hampshire, USA). Samples were then mixed using a plate shaker and centrifuged for 30 min at 3,220 x g. An aliquot of each supernatant (700 μL) was then transferred to a separate deep-well plate for LC–MS/MS analysis. Compound calibration standards were prepared in 80:20 water:acetonitrile (HPLC grade) and covered a concentration range over three orders of magnitude in semi-log steps. Standards and extracts were analyzed using reverse phase liquid chromatography-tandem mass spectrometry (TSQ Quantum Access, Thermo Fisher, USA), operated in positive electrospray ionization mode (ESI +).

Quantification was achieved by reference to calibration standards using an internal standard method. The measured per-embryo/larvae concentration was then expressed as a percentage of the nominal exposure (well plate) concentration. Stability of the exposure solution was determined by comparison of the day 0 solution concentration to the exposure solution concentration at 5 dpf. For this, samples were taken from the 10 μM concentrations by combining medium from

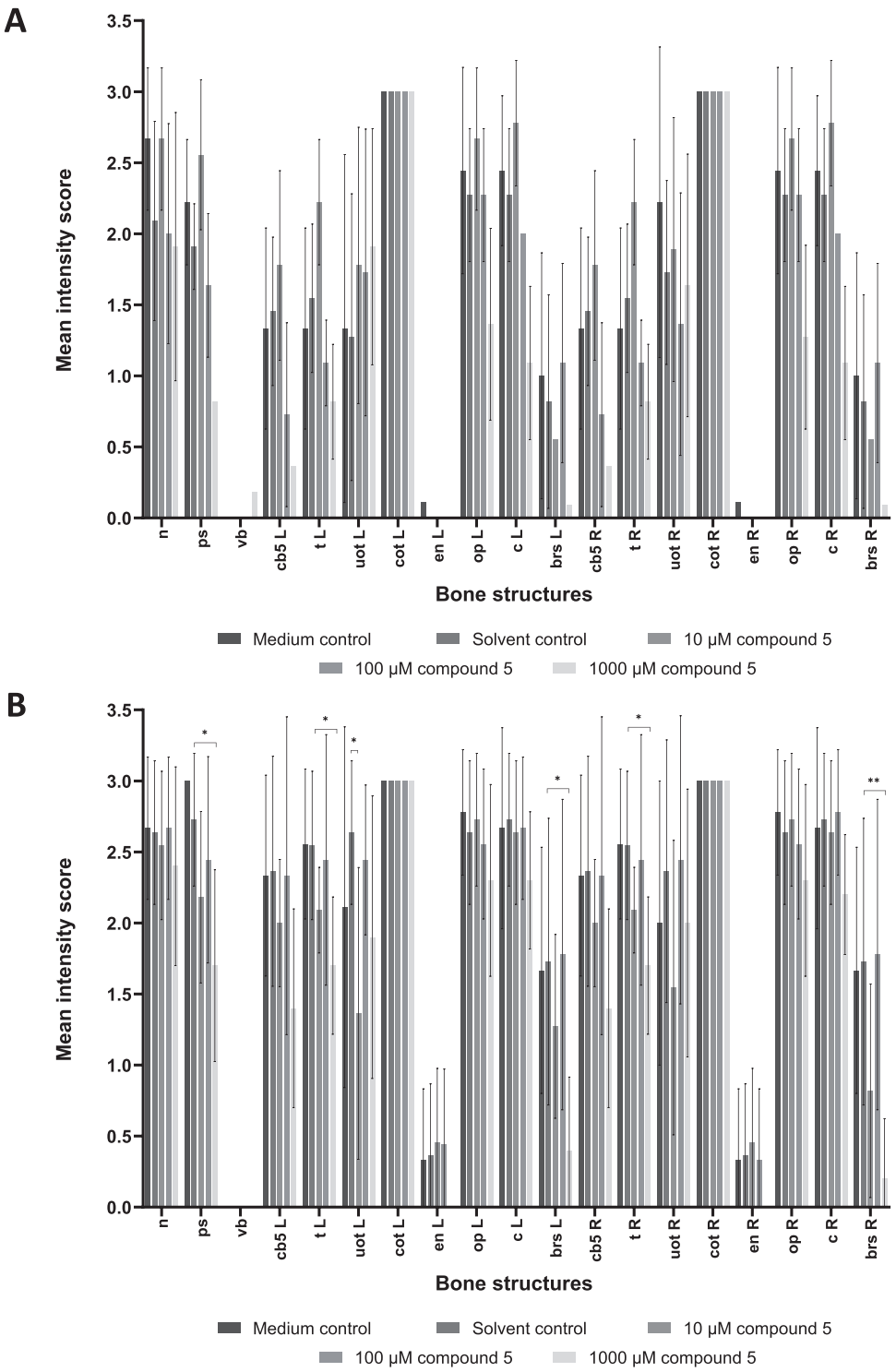


Fig. 5. Mean intensity score of fixed AR-stained bones of 5 dpf larvae treated with proprietary compound 5. A) shows the results of replicate 1 and B) shows the results of replicate 2. All groups were compared to the solvent control group. Significant differences between the solvent control and any of the other groups are indicated. $P \leq 0.05$ (*) and $p \leq 0.01$ (**).

multiple wells.

2.4. Exposure of zebrafish embryos

2.4.1. Compound administration to zebrafish embryos

At 4–6 hpf, zebrafish embryos were transferred individually into the wells of a 24-well plate (BD Falcon, NJ, USA) containing test compound solution, medium and/or vehicle (DMSO) controls in a final volume of 1 mL/embryo. In total, two replicates from a different clutch of fish were used per staining experiment, and in each replicate 24 embryos/group were exposed. The solvent (DMSO) concentration was 0.5% (i.e., lower than the maximum final concentration of DMSO that is considered to be safe to be used in zebrafish assays [27]). If precipitation was observed, stock solutions were pH adjusted with 1 M NaOH or 1 M HCl within a range of pH 4–10 to facilitate dissolution with final adjustment to pH 6.4–8.4. Any precipitation in compound solutions and/or the well was

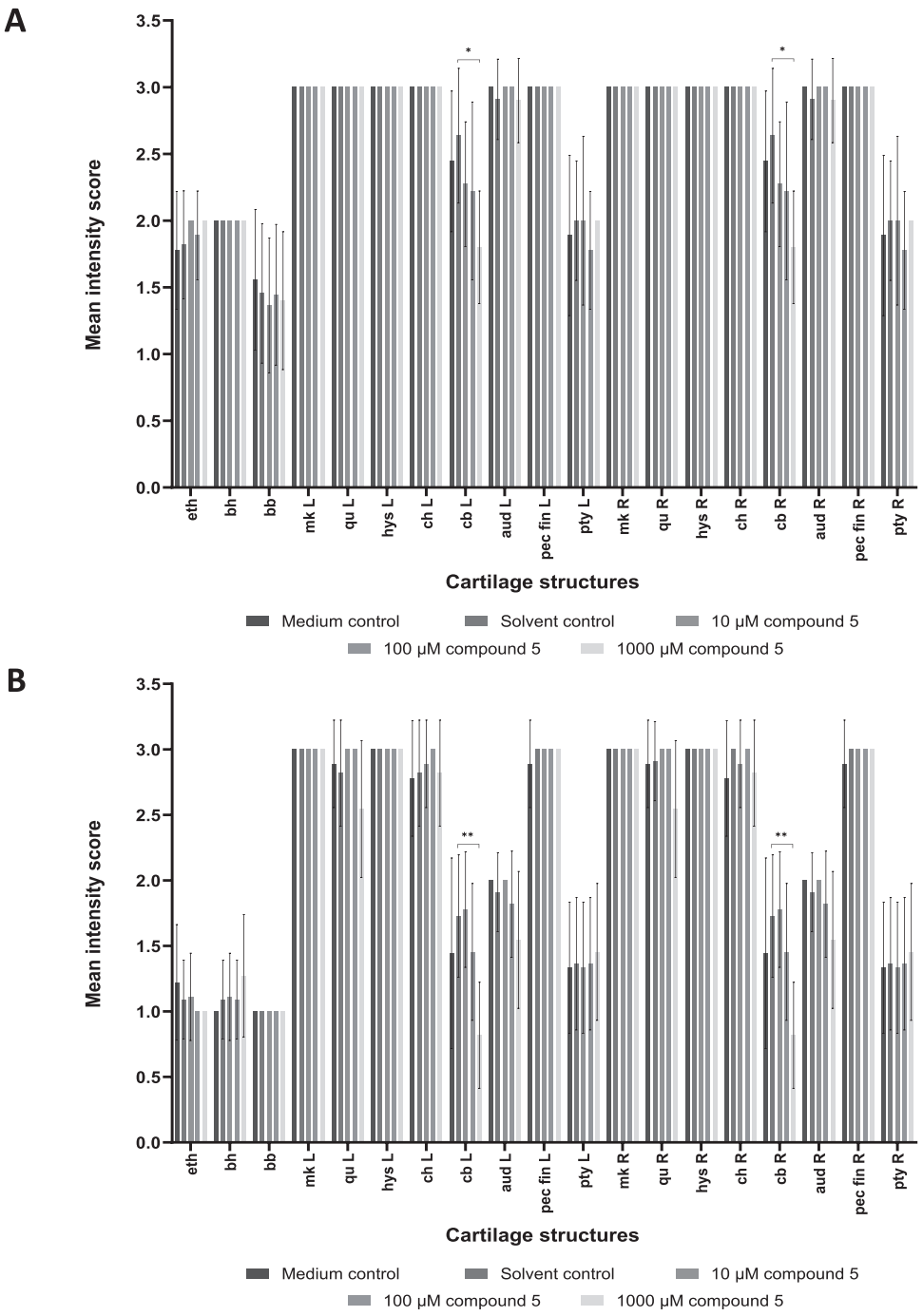


Fig. 6. Mean intensity score of fixed AB-stained cartilages of 5 dpf larvae treated with proprietary compound 5. A) shows the results of replicate 1 and B) shows the results of replicate 2. All groups were compared to the solvent control group. Significant differences between the solvent control and any of the other groups are indicated. $P \leq 0.05$ (*) and $p \leq 0.01$ (**).

recorded at the time of plating and at 5 dpf. All embryos were incubated at 28 (±1) °C for 5 days. More details about the exposure and chemicals that were used in the different experiments can be found in Table 1.

For the live staining experiments, the exposure was stopped at 5 dpf and the larvae were transferred to small crystalline dishes and were further reared in rig water until 9 dpf. Viability was checked daily. The crystalline dishes had a 50% water change daily with fresh rig water after the first feed of the day. The larvae were fed three times a day with 4 mg ZM (Zebrafish Management Ltd, Hampshire, UK) dry particle larval food (5–8 dpf ZM-000, 9 dpf 50:50 ZM-000:ZM-100).

2.5. Viability, morphological evaluation, and length assessment

The viability of all larvae was assessed at 5 dpf, and also at 9 dpf for the live staining experiments. After the viability assessment on 5 dpf (i.e. using the same criteria for lethality as mentioned in 2.3), larvae were anesthetized using tricaine (1 mM) and a morphological assessment was conducted for the same endpoints as used in Gustafson et al., 2012 [28] and Ball et al. (2014) [8]. For this assessment, a numerical system that has been previously described by Panzica-Kelly et al. was used [29]. The standard length (SL; in mm) at 5 and 9 dpf of all hatched larval zebrafish was determined using a Leica M205C stereomicroscope (Leica, UK). Images were captured using a Leica DMC4500 digital camera. Image

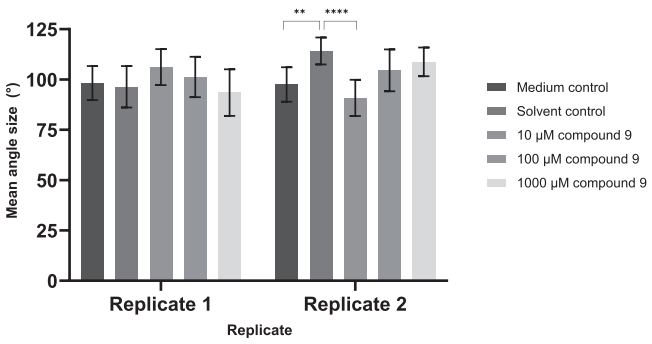


Fig. 7. Mean angle sizes between ceratohyal cartilages in 5 dpf larvae treated with compound 9. All groups were compared to the solvent control group. Significant differences between the solvent control and any of the other groups are indicated. $P \leq 0.01$ (**) and $p \leq 0.0001$ (****).

analysis was conducted by applying Leica LAS X core and LAS X measurements® (see 2.8).

2.6. Staining protocol

2.6.1. Fixed tissue staining (alizarin red and alcian blue)

Two-color acid free staining of the zebrafish larvae was adapted from Walker and Kimmel (2007) [23]. In brief, after euthanasia in 4 mg/mL MS-222, half of the hatched larvae of each replicate (i.e., a maximum of 12/replicate) were fixed in 4% formalin (containing 0.1 M phosphate buffer) for 1 day at room temperature and then stored in 70% ethanol (Thermo Fisher, New Hampshire, USA) overnight at 4°C. Samples were then washed for 5 min in 50% methanol followed by 80% methanol. The larvae were then stained in alcian blue solution (0.02% w/v alcian blue (Sigma-Aldrich, Missouri, USA), 80 mM $MgCl_2$ (Sigma-Aldrich) in 73.5% ethanol) for 1 h. This was followed by a 5-min wash in 50% ethanol and two washes in water containing 0.2% v/v Triton ™ X-100 (Thermo Fisher). The larvae were bleached (0.8% KOH (Thermo Fisher) and 0.9% H_2O_2 (Sigma-Aldrich) in 0.2% Triton ™ X-100 water) for 30 min, while monitoring for pigmentation loss. This was followed by two 5-min washes in water containing 0.2% Triton ™ X-100. Stained larvae were immersed in 100% Borax (saturated sodium tetraborate (Thermo Fisher)) solution for 10 min followed by 1 h in digestion solution (1% trypsin (Sigma-Aldrich) in 60% borax solution with 0.08% Triton ™ X-100). Post digestion, larvae were stained for 3-hours in alizarin red solution (0.003% alizarin red (w/v) (Sigma-Aldrich) in 1% KOH solution) and then 20 min in wash and clear solution (20% glycerol (Thermo Fisher), 0.8% KOH in 0.2% Triton ™ X-100). Samples were then stored in the dark at 4°C in 70% glycerol until imaging. Imaging was conducted for each batch within 72 h after staining. Transfer of larvae between different staining solutions was conducted using Netwell ™ permeable supports (15 mm insert with 74 µm polyester mesh (Corning, USA)).

2.6.2. Live calcein staining

Larvae (24/group) were stained as detailed in Du et al. (2001) [24]. Calcein powder (C0875–5 G, Sigma-Aldrich) was solubilized in deionized water at 2 mg/mL and adjusted to pH 7.4 ± 0.3 with 1 M NaOH (Thermo Fisher). Larvae were transferred to the calcein stain solution for 10 min, then washed in 3 volumes of rig water. The solution was replaced with 0.2 mg/mL MS-222 (pH 7.5) for 5 min. The larvae were embedded in 1% low melting point agarose and imaged immediately (see 2.8).

2.6.3. Live alizarin red staining

Alizarin red powder (A5533–25 G, Sigma-Aldrich) was solubilized in 0.3x Danieau's solution at 0.5% w/v stock solution and stored in the dark. The alizarin red stain stock was diluted 1:100 in 0.3x Danieau's solution and pH adjusted with fresh 1 M KOH (Thermo Fisher) to pH 7.4 ± 0.2 (i.e., the AR solution). Larvae (24/group) were transferred to the AR solution for 1 h and then washed in a 3 times volume of rig water. The solution was then replaced with 0.2 mg/mL MS-222 (pH 7.5) for 5 min until loss of dorsoventral balance. The larvae were then embedded in 1% low melting point agarose and captured immediately (see below).

2.7. Embedding protocol

The fluorescent bone staining was performed in larvae at 5 and 9 dpf using the method of Parker et al., 2014 [30]. Each larva was anesthetized in MS-222 (0.2 mg/mL, pH 7.5) until the loss of dorsoventral balance. The larvae were then transferred into low melting point agarose (1 g/100 mL in the same MS-222) before being deposited in a total volume of 80 µL into a well created by a press-to-seal silicon isolator (Sigma-Aldrich) on a clear microscope slide. Each larva was then gently orientated onto its side with the head to the left, the agarose solidified by very brief (1–2 s) exposure to a cooling plate (5°C), and two drops of MS-222 placed on top to minimize agarose shrinkage during image. Post imaging the larva was released from the agarose into 0.02 g/L MS-222 and then re-embedded in agarose (1 g/100 mL) with the dorsal side down and the head to the left. At 5 dpf, the larvae were then released into clean rig water and maintained till 9 dpf. The imaging of both the lateral and dorsoventral views was repeated at 9 dpf. At the end, the larvae were terminated in an overdose of anesthetic and secondary confirmation of termination was done by the destruction of the brain tissue.

2.8. Image capture

Live larval zebrafish were imaged using a Leica M205C stereomicroscope (Leica, UK) combined with a Leica DMC4500 digital camera. Image analysis was conducted by Leica LAS X core and LAS X measurements®. Larvae were anesthetized in MS-222 (0.2 mg/mL, pH 7.5) until the loss of dorsoventral balance. After imaging, the larvae were returned to the exposure medium or euthanized with an overdose of anesthetic.

Live fluorescent bone staining was imaged using an Olympus SZX16 scope (Olympus, UK) at 6.3x magnification with Prior 200 Lumen illumination (100%). To do so, anesthetized larvae were embedded in 1 % low gelling temperature agar containing 0.2 g/L MS-222.

Alizarin red fluorescence was captured with a red fluorescent protein (RFP) light cube filter (620 nm with a bandwidth of 0 nm (excitation HQ545/30x, emission: HQ620/60x)) (Chroma Technology Corporation. VT USA) and by using 100 ms exposure in Micromanager (v1.4). In total 8 images per stack were captured at a 3 s interval between the images using a Zyla 4.2 sCMOS camera (Andor, Oxford Instruments, UK).

Calcein (GFP) fluorescence was captured with a green fluorescent protein (GFP) light cube filter (460 nm with a bandwidth of 35 nm (excitation BP460 T2, emission: BP495 T2)) (Chroma Technology Corporation. VT USA) and by using 20 or 9.84 ms exposure in Micromanager (v1.4). In total 8 images per stack were captured at a 3 s interval between the images using a Zyla 4.2 sCMOS camera (Andor, Oxford Instruments, UK).

Fixed, stained larval zebrafish larvae were imaged using a Leica M205C stereomicroscope (Leica, UK). Still images were captured using a Leica DMC4500 digital camera. Image analysis was performed using Leica LAS X core and LAS X measurements®.

2.9. Skeletal evaluation

2.9.1. Alizarin red and alcian blue stained larvae

Each bone and cartilage structure (see Table 2, Table 3 and Fig. 1) of each larva was scored for staining intensity and shape by allocating a representative score (see Table 4). Each structure was scored in the position (i.e., lateral or dorsoventral) where it was most visible/stained.

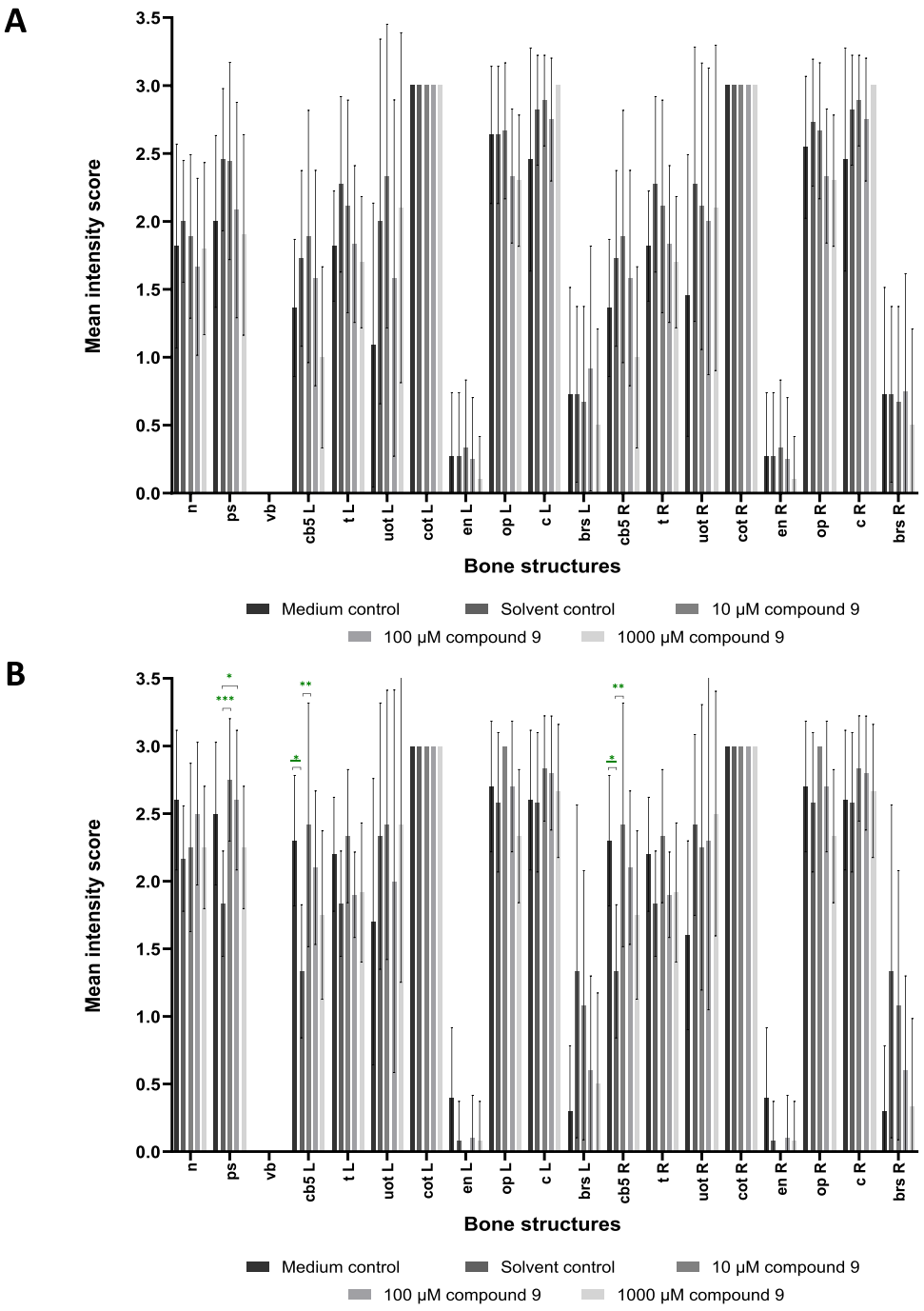


Fig. 8. Mean intensity score of fixed AR-stained bones of 5 dpf larvae treated with proprietary compound 9. A) shows the results of replicate 1 and B) shows the results of replicate 2. All groups were compared to the solvent control group. Significant differences between the solvent control and any of the other groups are indicated. $P \leq 0.05$ (*), $p \leq 0.01$ (**) and $p \leq 0.001$ (***). The color of the asterisks indicate that the solvent control was significantly more (black) or less (green) intense stained than the other group. The green underlined asterisks indicate that the solvent control was significantly less stained than the medium control.

Moreover, the angle between the ceratohyal cartilages was measured. Previous studies showed that this angle may increase after exposure to xenobiotics, and therefore, can be an indication for xenobiotic toxicity [31].

2.9.2. Calcein stained larvae

Six stacks of images (i.e., one stack of 8 images per position) for each orientation were taken for each larva. Two different aperture durations, i.e., 20 ms and 9.84 ms, with fluorescence intensity power at 100, 50 and 10% power were used to prevent over exposure of the bone tissue. However, the larvae could not be scored for individual bone structures, as will be further explained in Section 3.4 due to a high degree of variability in staining.

2.9.3. Alizarin red live stained larvae

For each of the larvae, two stacks of images (i.e., one stack of 8 images per position) were processed in ImageJ and the intensity and shape of the bone structures (see Table 2, Fig. 2 and Fig. 3) of each of the larva were evaluated and a representative score was allocated (see Table 4) [16]. For each bone structure of each larva in each group, the image with the highest intensity score was selected to determine the final score of the bone.

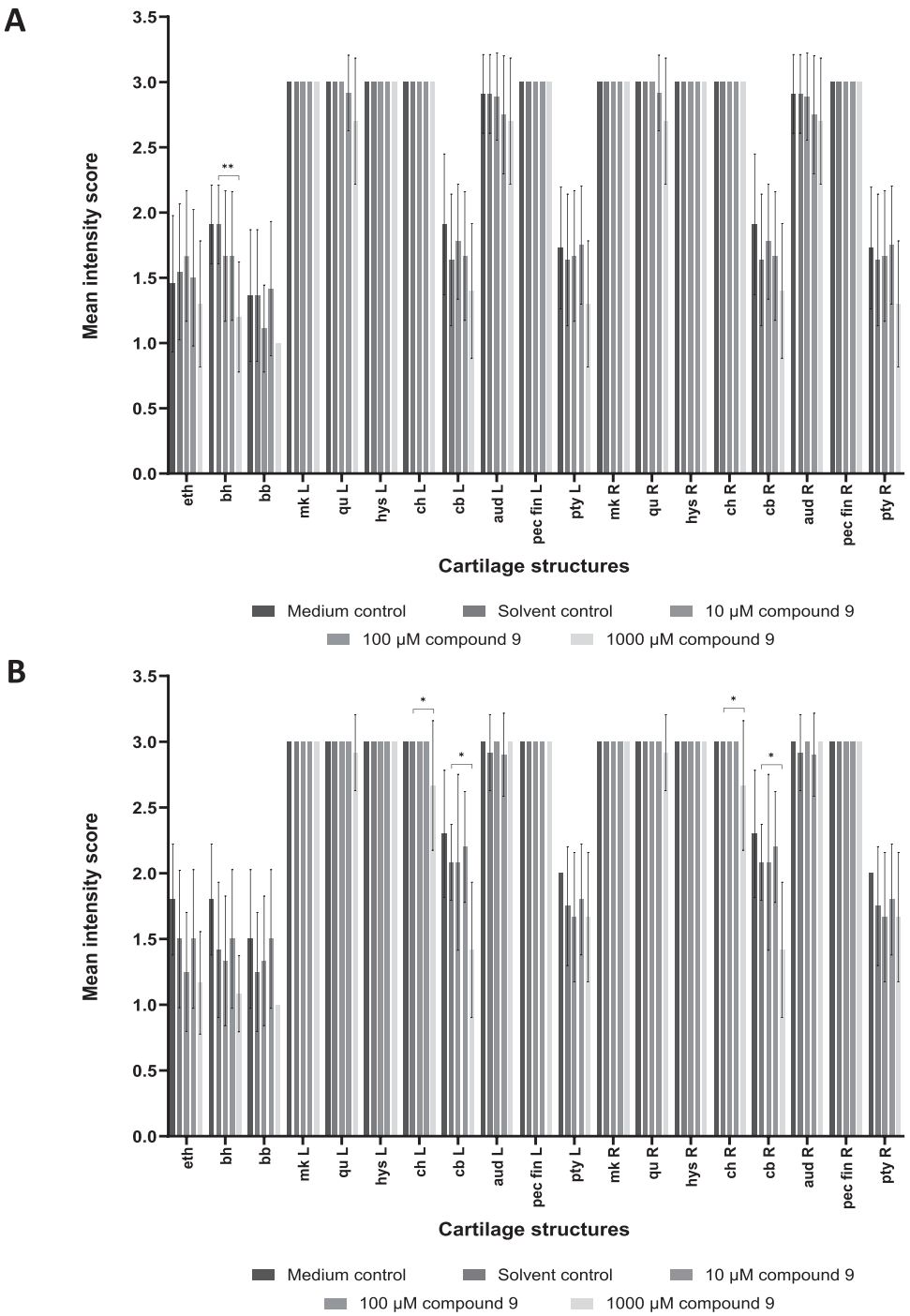


Fig. 9. Mean intensity score of fixed AB-stained cartilages of 5 dpf larvae treated with proprietary compound 9. A) shows the results of replicate 1 and B) shows the results of replicate 2. All groups were compared to the solvent control group. Significant differences between the solvent control and any of the other groups are indicated. $P \leq 0.05$ (*) and $p \leq 0.01$ (**).

2.10. Using solvent controls to investigate the cause of variability between replicates

Due to inconsistent results between replicates of compound-treated zebrafish larvae, additional experiments and/or analyses were performed to investigate the cause of this variability.

For the AR live staining experiments, the variability in the intensity of solvent controls was first checked by comparing the solvent controls of both replicates of compound 5 experiments at 5 dpf and 9 dpf. However, to make sure the time gap between the two replicates was not causing additional variability, new solvent control experiments were carried out. First, 0.5% DMSO treated (solvent control) larvae of two clutches of zebrafish (same spawning day, n=24) were compared. Second, the larvae of each of these clutches were divided into 3 groups of 8 fish (i.e., larva 1–8 in subgroup 1, larva 9–16 in subgroup 2, and larva 17–24 in subgroup 3) and these subgroups were compared to each other. In addition, the solvent control group of the second replicate of the compound 5 experiment (5 dpf) was also divided into 3 subgroups of 8 larvae each, so they could also be compared to each other.

For the fixed staining experiments, the variability in intensity and bone shape of the different medium control groups was checked by comparing the medium control groups of all previously conducted

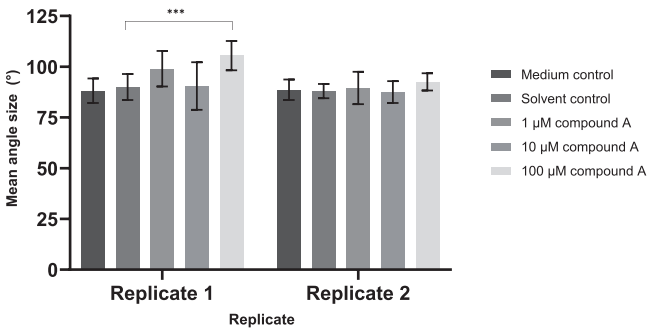


Fig. 10. Mean angle sizes between ceratohyal cartilages in 5 dpf larvae treated with compound A. All groups were compared to the solvent control group. Significant differences between the solvent control and any of the other groups are indicated. P ≤ 0.001 (***).

experiments (compound 5, 9, A, and levetiracetam; 2 replicates each). Thus, 8 control groups of maximum 12 zebrafish larvae were compared to each other.

2.11. Statistical analysis

Length measurements between different treatments were analyzed for statistical differences induced by compound treatment from the pooled experimental replicates. Minitab 21 Statistical software (computer software; www.minitab.com) was used. Normal distribution of the data was determined by Levene's and Bartlett's tests. If data were normally distributed, assessment of significant difference was performed using by Tukey ANOVA test. If no normal distribution was found, a Kruskal-Wallis test was applied and then an individual Mann-Whitney test was applied to each treatment group. For the binary data of the viability and the shape scoring a Fisher Exact test was performed and the Relative Risk was calculated. For the ordinal data (i.e., 0, 1, 2, or 3) of the intensity scoring a Kruskal-Wallis test with correction for multiple comparisons (Dunn's multiple comparisons test) was used to check if there was a significant difference between any of the test groups and the control group (exposure experiments) or between any of the control groups (variability experiments). Also, for the angle size data a Kruskal-Wallis test with correction for multiple comparisons (Dunn's test) was used. If there were only two groups, a Mann Whitney U test was used instead of the Kruskal-Wallis test. Except for length measurement analysis, all statistical analyses were performed using GraphPad 8.4.0 or newer versions.

3. Results

3.1. Uptake assessment

Uptake of all four test compounds was detected in zebrafish at the highest nominal exposure concentration of each compound between 1% and 31% of the nominal concentration at both 1 and 5 dpf (see Table 5). Precipitation was observed in the analysis solution of compound A at 100 µM nominal concentration at both 1 and 5 dpf. Therefore, to avoid excessive precipitation, the highest tested concentration used for exposure was 100 µM, and not higher (i.e., no 1,000 µM). The other compounds did not display any precipitation at 100 µM, and therefore, also 1,000 µM was included for evaluating the stainings. All compounds were stable in the medium over the 5 days (see Table 5). There was a reduction in the recorded uptake of all compounds between the 1 and 5 dpf sampling, except for 10 µM of compound A.

3.2. Viability, morphological evaluation, and length assessment

3.2.1. 5 dpf larvae exposed to compound 5, 9, A and levetiracetam (fixed experiments)

No significant difference in viability was observed between the test concentrations and the control groups (see Supplementary table 2).

Exposure to levetiracetam (LTC) up to 1,000 µM did not induce any malformations at 5 dpf (see Supplementary table 3). Levetiracetam also showed no effect on larval length.

At 5 dpf, no treatment related malformations were observed after exposure to compound A up to 100 µM (see Supplementary table 4). A significant reduction in length of the 1, 10 and 100 µM test groups was observed in one of the replicates for compound A compared to the DMSO control ($p = 0.000$ for 1 µM and 10 µM and $p = 0.014$ for 100 µM) (see Supplementary table 5), and the 1 µM and 10 µM test groups compared to the medium control ($p = 0.037$ for 1 µM and $p = 0.023$ for 10 µM).

Compound 5 induced treatment-related malformations in the jaw and neural tube, and pericardial edema and slow heart rate at 1,000 µM at 5 dpf (see Supplementary table 6). Also a significant reduction in length of the 1,000 µM exposed group compared to the medium and DMSO control groups ($p = 0.048$ to medium and $p = 0.031$ to DMSO in replicate 1; and $p = 0.000$ to DMSO in replicate 2) was observed (see Supplementary table 5).

Compound 9 did not induce any treatment-related malformations at 5 dpf compared to the medium and DMSO control larvae (see Supplementary table 7). The larval length at 1,000 µM of compound 9 was significantly reduced compared to the DMSO control in both replicates ($p = 0.000$ for replicate 1 and $p = 0.004$ for replicate 2), and the medium control in replicate 1 ($p = 0.002$) (see Supplementary table 5).

3.2.2. 5 and 9 dpf larvae exposed to compound 5 (AR live experiments)

No significant difference in viability was observed between the test concentrations of compound 5 and the DMSO control group at 5 dpf (see Supplementary table 8). The larval length at 1,000 µM of compound 5, however, was significantly reduced compared to the DMSO control ($p = 0.001$), but only in the first replicate (see Supplementary table 9). At 9 dpf, the length was significantly reduced after exposure to 100 µM in the first replicate ($p = 0.030$) and after exposure to 1,000 µM in both replicates ($p = 0.011$ for replicate 1, $p = 0.003$ for replicate 2) (see Supplementary table 9). Also, significant differences in viability were found at 9 dpf between the 1,000 µM group and the control group ($p = 0.0001$ for replicate 1 and $p = 0.0496$ for replicate 2) (see Supplementary table 8).

3.3. AR-AB staining results

3.3.1. 0.5% DMSO as a solvent

In each of the different experiments (see Figs. 4–15), the solvent control was compared to the medium control to check if the use of 0.5% DMSO as a solvent did affect the skeletal development or the AR-AB staining.

For all AR-AB staining experiments, no significant deviations in shape of bone and cartilage structures were observed in the medium and solvent controls (data not shown).

The angle between the ceratohyal cartilages was comparable between the medium and solvent control group of the different experiments (see Fig. 4, Fig. 7, Fig. 10 and Fig. 13). Exceptions were found in replicate 2 for compound 9 (see Fig. 7, mean angle size was significantly larger in the solvent control group) and in replicate 1 of the levetiracetam experiment (see Fig. 13; mean angle size was significantly larger in the medium control group). The other replicate of both experiments showed no significant differences in mean angle size between both control groups.

In both replicates of compound 5 (see Fig. 5 and Fig. 6) and levetiracetam (see Fig. 14 and Fig. 15), there were no significant differences in staining intensity between the solvent and medium control. Also, no

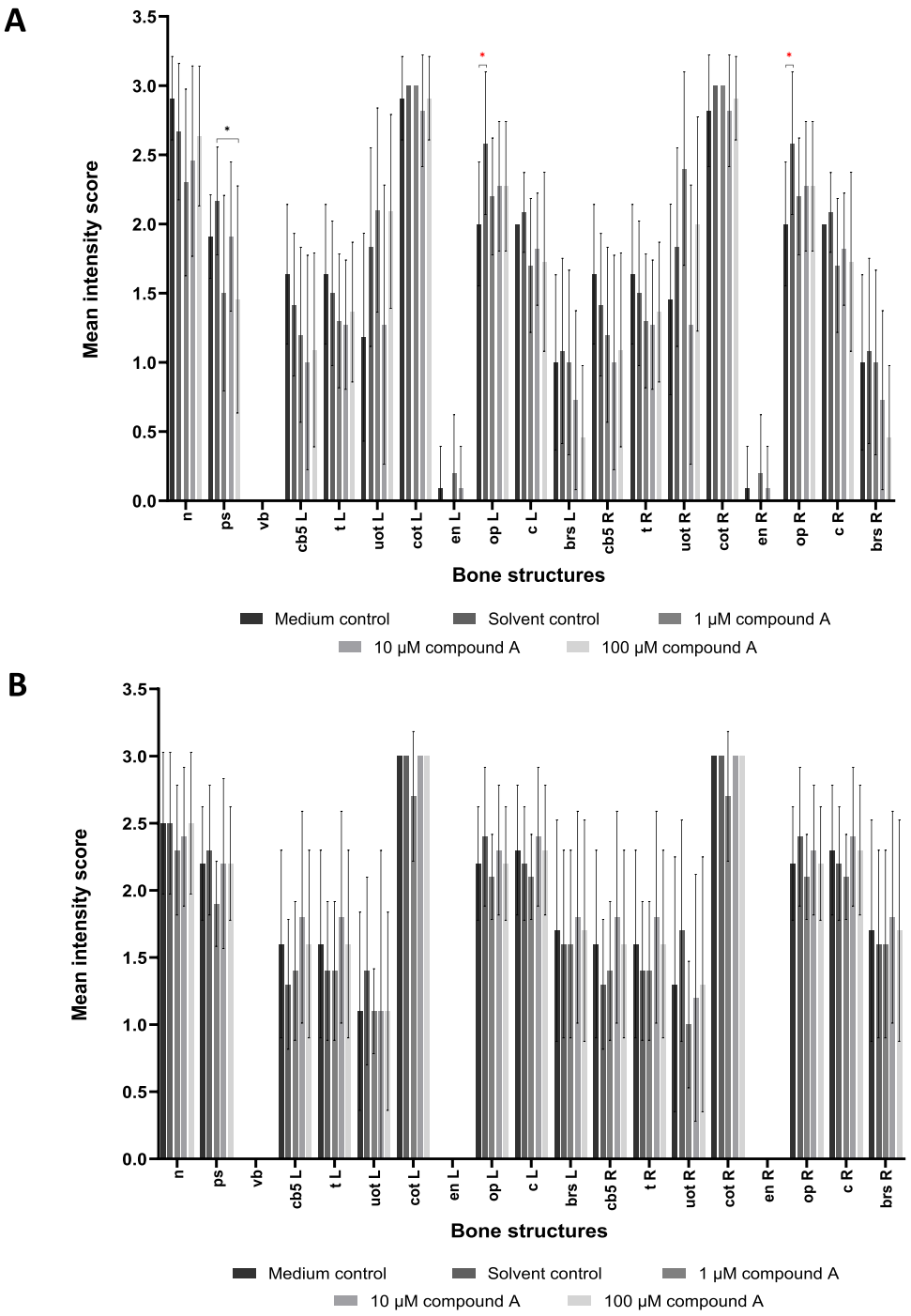


Fig. 11. Mean intensity score of fixed AR-stained bones of 5 dpf larvae treated with proprietary compound A. A) shows the results of replicate 1 and B) shows the results of replicate 2. All groups were compared to the solvent control group. Significant differences between the solvent control and any of the other groups are indicated. $P \leq 0.05$ (*). The black asterisks indicate that the solvent control was significantly more intense stained than the other group. The red asterisks indicate that the solvent control was significantly more stained than the medium control.

differences in intensity of the different bone and cartilage structures were observed when comparing both controls in replicate 1 of the compound 9 experiment (see Fig. 8 and Fig. 9), and replicate 2 of the compound A experiment (see Fig. 11 and Fig. 12). However, in replicate 2 of the compound 9 experiment, the ceratobranchials 5 (left and right) were more intensely stained in the medium control group than in the solvent control group (see Fig. 8B, indicated with a green underlined asterisk). In replicate 1 of the compound A experiment, the opercles (left and right) were more intensely stained in the solvent control group than in the medium control group (see Fig. 11A, indicated by red asterisks). However, in all other replicates and experiments, the intensity of the opercles and ceratobranchials 5 were not significantly different between the solvent and medium controls.

Overall, the results of the intensity, shape and angle size showed that the use of 0.5% DMSO as a solvent did not have an impact on skeletal development nor on AR-AB staining. The results of the different test compound groups could therefore be compared to the solvent control.

3.3.2. Compound 5

No significant differences in shape of the bone and cartilage

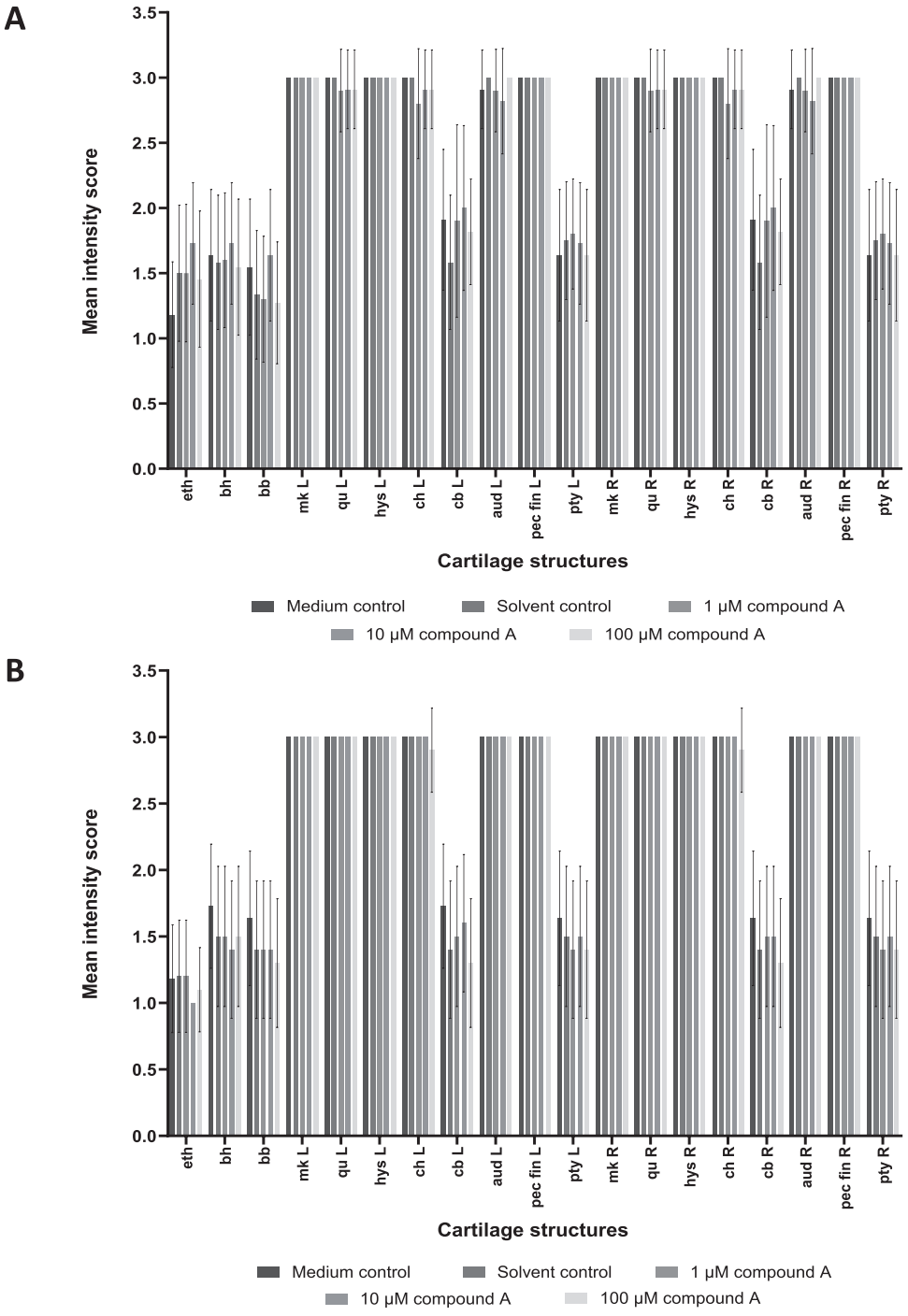


Fig. 12. Mean intensity score of fixed AB-stained cartilages of 5 dpf larvae treated with proprietary compound A. A) shows the results of replicate 1 and B) shows the results of replicate 2. All groups were compared to the solvent control group. No significant differences were observed between the 3 test groups or the medium control group and the solvent control group.

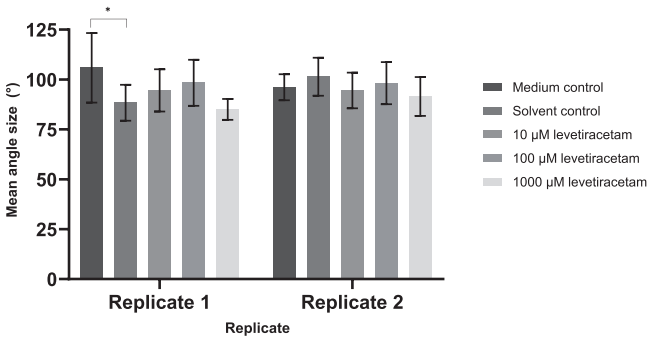


Fig. 13. Mean angle sizes between ceratohyal cartilages in 5 dpf larvae treated with levetiracetam. All groups were compared to the solvent control group. Significant differences between the solvent control and any of the other groups are indicated. $P \leq 0.05$ (*).

structures (data not shown) and ceratohyal angle size (see Fig. 4) were observed when comparing the three test groups with the solvent control.

Compound 5 reduced the staining intensity in some bone structures (ps, t L&R, brs L&R) in replicate 2 at 1,000 µM (see Fig. 5 and Supplementary figure 1). No significant effects were observed at 100 µM. Uot L was less intensely stained at 10 µM in replicate 2 (see Fig. 5). In replicate 1, no significant effects of compound 5 on staining intensity were observed.

For the cartilage structures, cb (left and right) was less intensely stained in both replicates at 1,000 µM (see Fig. 6 and Supplementary figure 1).

3.3.3. Compound 9

No differences in shape of the bone and cartilage structures were observed when comparing the three test groups to the solvent control (data not shown). Differences in mean angle size, however, were observed in replicate 2 (see Fig. 7). The mean angle size of compound 9 was significantly smaller at 10 µM than in the control group. No effects on mean angle size were noted at the higher concentrations.

Compound 9 caused no significant differences in staining intensity at 1,000 µM when compared to the control group (see Fig. 8). In replicate 2 at 100 µM, a significantly more intensely stained parasphenoid was noted than in the control group. An increase in staining intensity was also observed for the parasphenoid and the ceratobranchials 5 at 10 µM in replicate 2. For these structures, no increase in staining was observed at 1,000 µM (ps and cb5) or 100 µM (cb5).

For the cartilage structures, the basihyal (replicate 1), ceratohyals (replicate 2) and ceratobranchials (replicate 2) were less intensely stained in one of the replicates at 1,000 µM. No significant results were observed at 10 µM and 100 µM (see Fig. 9).

3.3.4. Compound A

No abnormalities in the shape of the bones and cartilages were observed after exposure to compound A (data not shown). The mean angle size was significantly increased at 100 µM in replicate 1 (see Fig. 10 and Supplementary figure 2). The parasphenoid was less intensely stained at 100 µM in replicate 1 (see Fig. 11 and Fig. 12).

3.3.5. Levetiracetam

No significant differences in ceratohyal angle size (see Fig. 13) or shape of the skeletal structures (data not shown) were observed when comparing the three test groups with the solvent control.

The staining intensity of skeletal structures after exposure to levetiracetam was very different in both replicates (see Fig. 14). This was especially different at 100 µM and 1,000 µM. A significant decrease in the intensity was observed for the ceratobranchials 5 (L&R) and the pharyngeal teeth (L&R) at 1,000 µM in replicate 1. The same structures were even less intensely stained, and thus more affected, at 100 µM in replicate 1. Moreover, also 3 other structures (i.e., opercle L&R, cleithrum L&R and parasphenoid) were less intensely stained at 100 µM. In contrast, the parasphenoid intensity significantly increased at 100 µM in replicate 2.

For cartilage intensity, only a significant decrease in the basihyal intensity was observed at 100 µM in replicate 1 (Fig. 15).

3.3.6. Conclusion on the AR-AB staining

A lot of variability in staining intensity was observed between the two replicates. Therefore, we decided to assess whether other staining methods may provide more consistent staining results.

3.4. Calcein live staining results

Assessment for the utility of calcein staining for determination of bone formation in the larval zebrafish was undertaken initially on control larvae of 5 dpf and 9 dpf (Fig. 16A). Even before assessing the different bone structures more closely, the control larvae already demonstrated a lot of variability in staining (Fig. 16B). The high degree of variability in fluorescence meant that, when imaging, adjustments in the exposure duration of the image capture had to be made for each larva to avoid overexposure of the image capture. There was also a large degree of autofluorescence (data not shown) within the tissues.

In conclusion, the high degree of variability in staining intensities between control individuals prevented the skeletal assessment. Therefore, calcein was not further used and a third staining method was explored.

3.5. AR live staining results

3.5.1. Compound 5

Based on the results of the AR-AB staining (see 3.3 and discussion), compound 5 was selected to expose zebrafish embryos to AR live, including a longer exposure window (9 dpf) (see discussion).

At 5 dpf, differences in shape of the parasphenoid were observed between both replicates (see Fig. 17). In the first replicate, the parasphenoid was malformed (i.e. shorter) at 1,000 µM ($p = 0.0199$, RR = 1.333). In the other replicate, no malformations were observed.

The staining intensity was significantly reduced in the right utricular otolith at 100 µM in the first replicate at 5 dpf (see Fig. 18). In the second replicate, there was no significant effect on this structure, but an increase in staining intensity was observed in the ceratobranchials 5 L&R. At 1,000 µM, the staining intensity was significantly reduced in almost all the bone structures in the first replicate (see Fig. 18). In the second replicate, however, the staining intensity was only decreased in the vertebrae and the anguloarticulars L&R, and increased in the ceratobranchials 5 L&R.

For both replicates at 9 dpf, no significant differences in shape of the bone structures were found when comparing the two treatment groups with the solvent control (data not shown).

At 9 dpf in the first replicate, several structures were significantly less stained at 100 µM and 1,000 µM when compared to the control group. In the second replicate, there was only a significant decrease in staining intensity of the left hyomandibular bone at 1,000 µM of compound 5 (see Fig. 19).

3.5.2. Conclusion on the AR live staining

In conclusion, the AR live staining also showed variability between replicates. To assess whether biological variability in zebrafish skeletal development could be the cause of the variability between replicates, we performed a post-hoc analysis of all solvent control groups.

3.6. Post-hoc analysis: variability in AR live staining between control groups

When comparing the solvent controls of both replicates of the

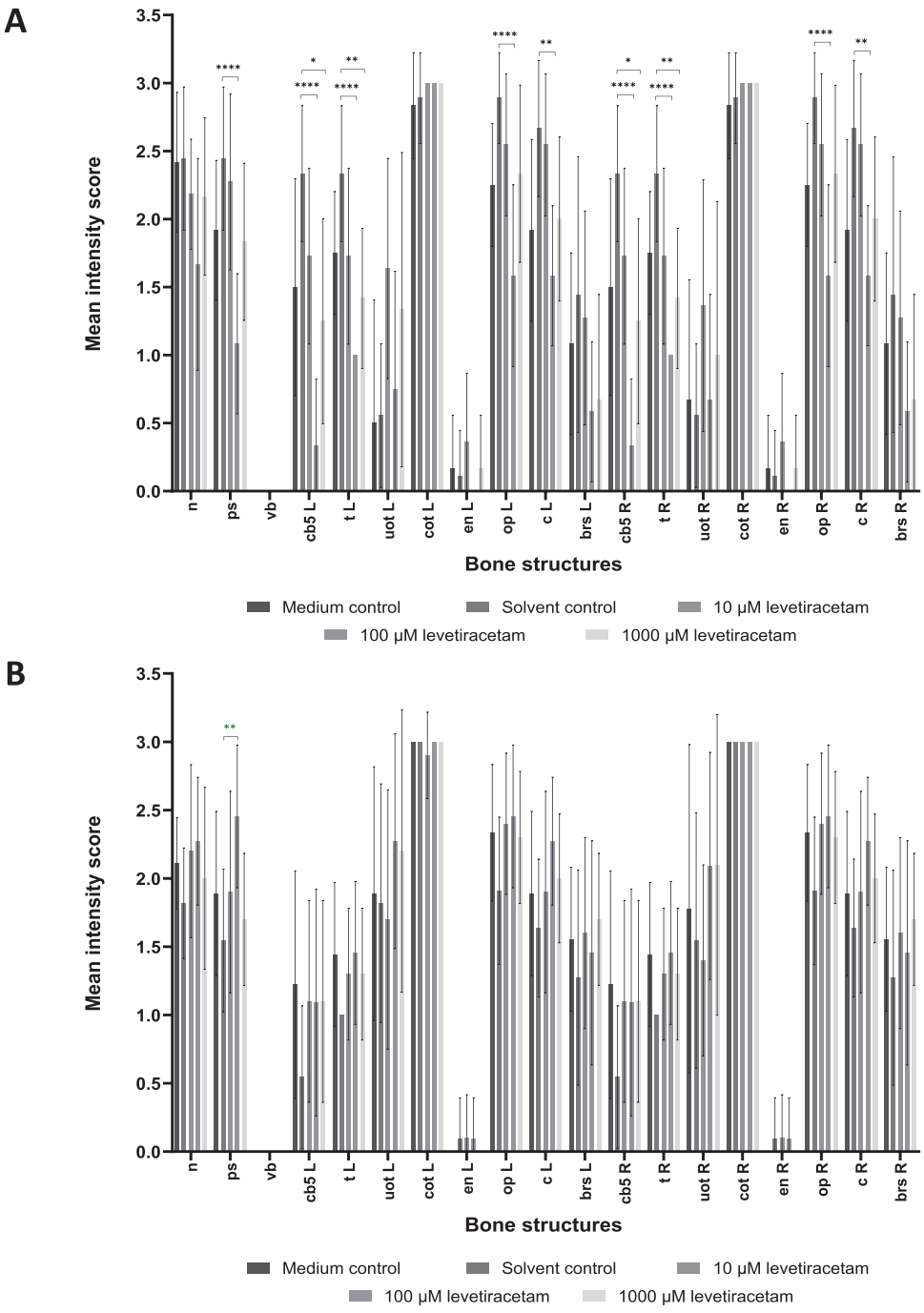


Fig. 14. Mean intensity score of fixed AR-stained bones of 5 dpf larvae treated with levetiracetam. A) shows the results of replicate 1 and B) shows the results of replicate 2. All groups were compared to the solvent control group. Significant differences between the solvent control and any of the other groups are indicated. $P \leq 0.05$ (*), $p \leq 0.01$ (**) and $p \leq 0.0001$ (****). The color of the asterisks indicate that the solvent control was significantly more (black) or less (green) intense stained than the other group.

compound 5 experiment, many differences in bone staining intensity were found (see Fig. 20). At 5 dpf, several bone structures were more intensely stained in the second than in the first replicate. Only the right circle saccular otolith was more intensely stained in the first replicate. In contrast, several structures were significantly more intensely stained in the first replicate at 9 dpf. This was even the case for structures that were observed to be less intensely stained in this first replicate at 5 dpf (i.e., chb R, den L, aa L, and aa R). As both replicates were conducted with 9 months of time in between, the use of different batches of adult fish and the use of freshly prepared staining solutions for both replicates may have contributed to this variability. Therefore, we decided to conduct an additional experiment to assess whether biological variability in (onset of) bone development or technical aspects caused the differences in bone staining intensity between replicates.

3.7. Variability experiments results

3.7.1. Variability between control larvae from 2 clutches of zebrafish (same spawning day)

In the first experiment, differences in bone staining intensity between control larvae from two clutches of fish (same spawning day) were checked (see Fig. 21 and Supplementary figure 3). The results

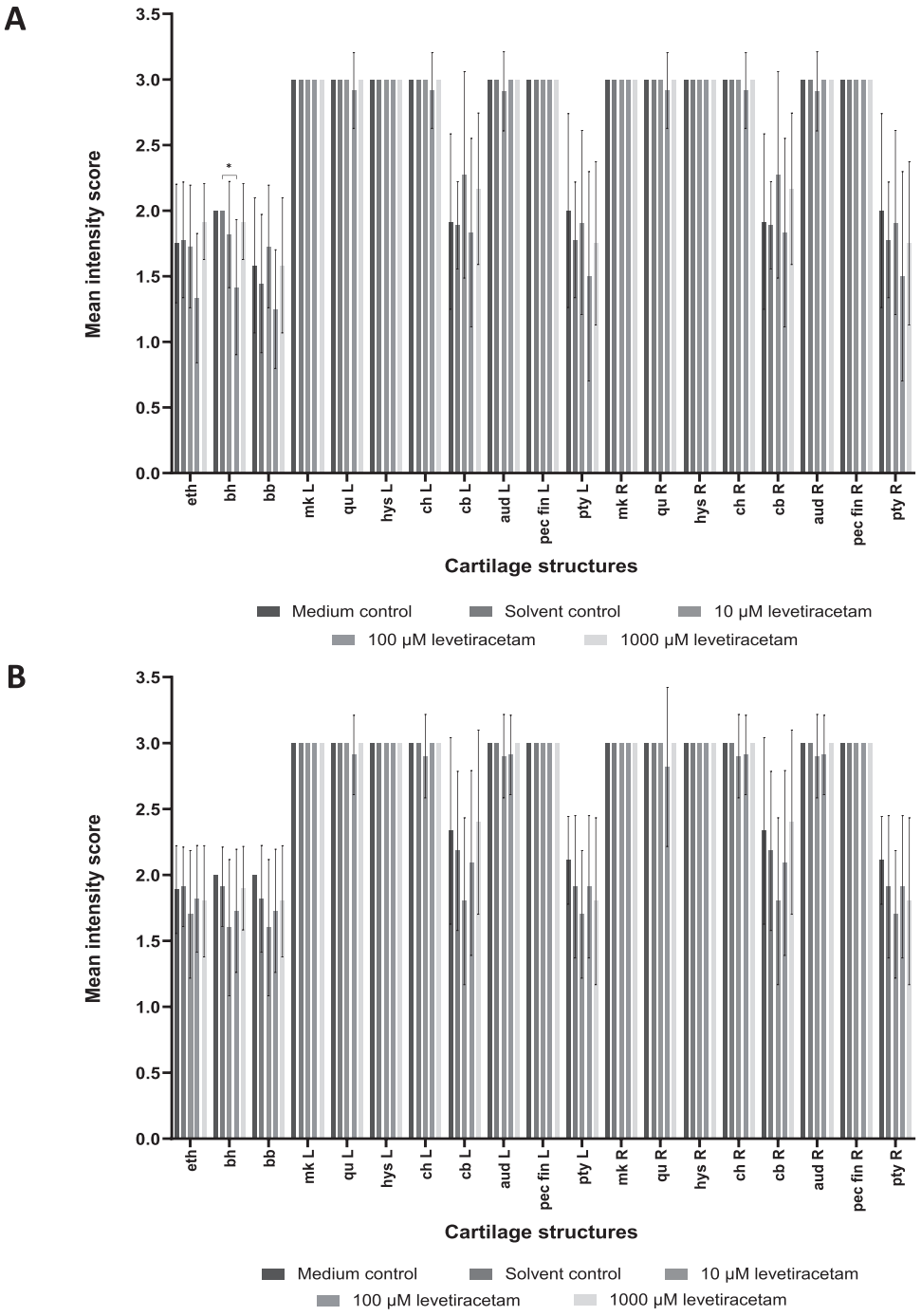


Fig. 15. Mean intensity score of fixed AB-stained cartilages of 5 dpf larvae treated with levetiracetam. A) shows the results of replicate 1 and B) shows the results of replicate 2. All groups were compared to the solvent control group. Significant differences between the solvent control and any of the other groups are indicated. $P \leq 0.05$ (*).

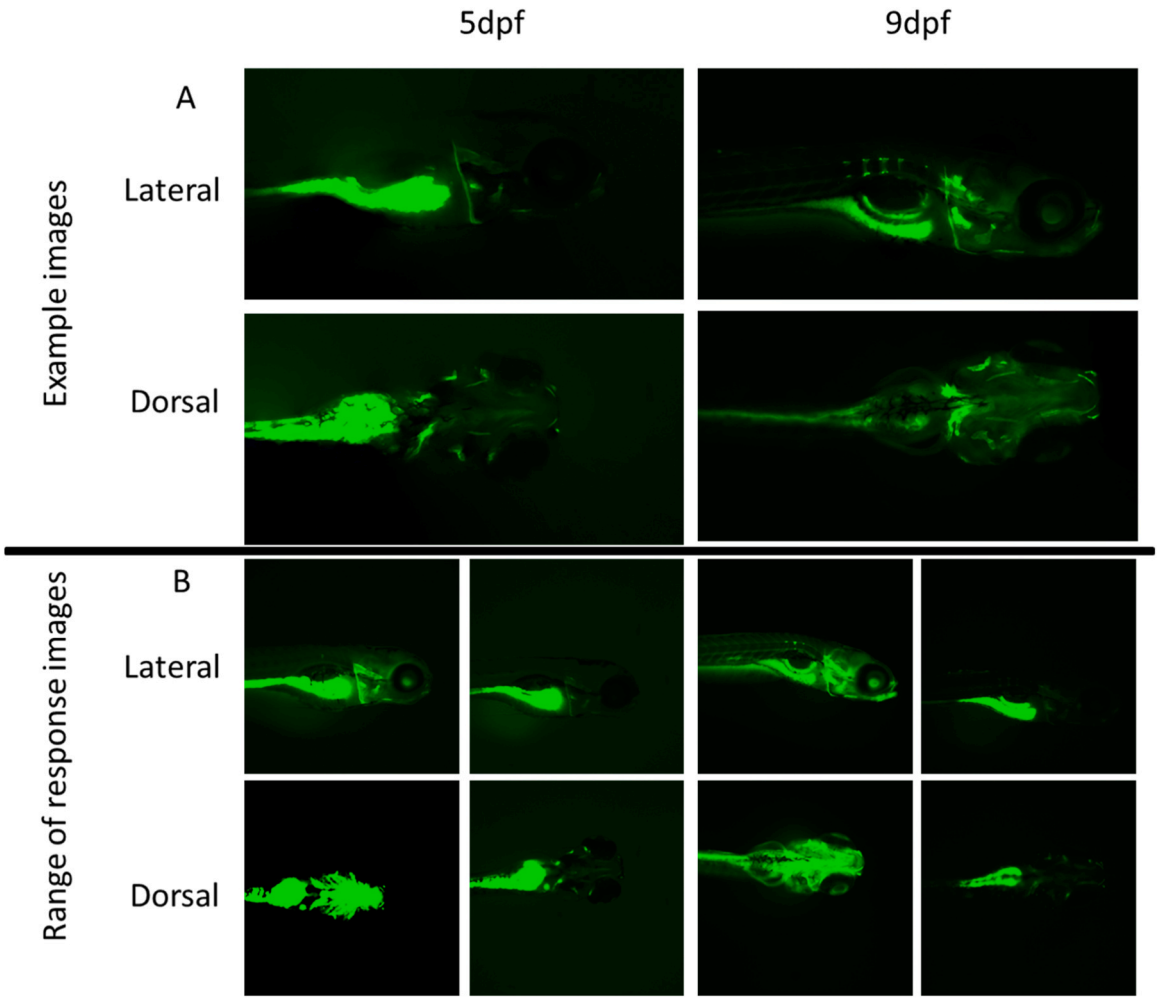


Fig. 16. Calcein stained larval zebrafish at 5 and 9 dpf. Panel A shows the detail of bone formation detectable by calcein staining. Panel B demonstrates the variability between the same spawning group of individuals treated at the same time. Abbreviation: days post-fertilization (dpf).

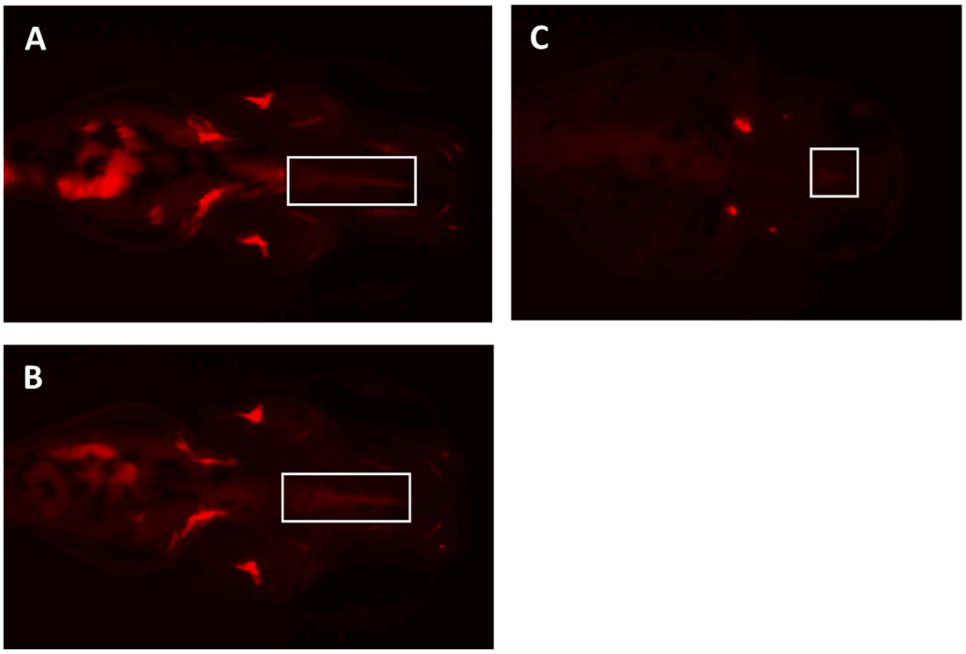


Fig. 17. Images of AR live stained bones of 5 dpf larvae treated with A) the solvent control, B) 100 μM of compound 5 and C) 1,000 μM of compound 5. The parasphenoid, which is normal in A and B, and malformed in C, is indicated with a white box. All pictures are from larvae of the first replicate and show the larvae in dorsoventral position.

revealed that the left and right branchiostegal rays were more intensely stained in larvae that originated from the first clutch.

3.7.2. Variability between control larvae from the same clutch of zebrafish

In a second experiment, differences in staining intensities between larvae from the same clutch of zebrafish were assessed. For this, the larvae of each of the clutches (batch 1 and 2, respectively) of experiment 1 (see 3.7.1) and the solvent control group of the second replicate of the AR live compound 5 experiment (5 dpf) (batch 3) were randomly divided into 3 groups of 8 fish. These 3 subgroups were compared to each other.

For the first batch of experiment 1, no significant differences in staining intensity between the three subgroups were observed (see Fig. 22A). For the other two batches, however, significant differences were observed. In the second batch, the right dentary was significantly less stained in the second subgroup than in the other subgroups (see Fig. 22B). In the third batch, the right utricular otolith was significantly less stained in the third subgroup than in both other subgroups, and the left entopterygoid was significantly less stained than in the first subgroup (see Fig. 22C).

3.7.3. Conclusion on the variability experiments

For the AR live staining, we found variability in staining intensities between larvae of two clutches of the same spawning day, as well as between larvae originating from the same clutch of fish. Thus, biological variability in (onset of) bone development is clearly present in zebrafish larvae.

3.8. Post-hoc analysis: variability in AR-AB staining between control groups

As the results of the AR live staining indicated that biological variability in (onset of) bone development is clearly present in zebrafish larvae, we also performed a post-hoc analysis of the control groups of the AR-AB experiments to check whether the same staining variability can be observed.

The angle size, bone shape and skeletal structure intensity of eight medium control groups were compared to each other. No significant differences in shape were observed (data not shown). Two groups, C5.2 and LTC.1, had significantly increased angle sizes when compared with some other groups (i.e., C5.2 with C5.1, CA.1 and CA.2, and LTC.1 with C5.1) (see Fig. 23). For staining intensity, significant results were observed for several structures (see Table 6). Especially the paraspenoid, branchiostegal rays, and the pharyngeal teeth (bone structures) and the basihyal, ceratohyal, ceratobranchials, and the auditory capsule (cartilage structures) were stained significantly different. So, also for the AR-AB experiments staining variability between different control groups was observed.

4. Discussion

The goal of this study was to investigate whether the sensitivity of zebrafish embryo developmental toxicity assays can be increased by including a skeletal staining. To avoid potential negative results due to the lack of compound uptake by zebrafish larvae, the first aim was to assess the uptake and stability of the four selected compounds. This assessment revealed that each compound was taken up by zebrafish embryos (1 dpf) and larvae (5 dpf) and remained stable in the medium over the tested period of five days.

Once uptake was confirmed, the viability and gross morphology of 5 dpf zebrafish larvae was assessed. Although all four compounds were previously reported to give false negative results [8,10], a re-assessment of the gross morphology in the present study allowed a comparison of the sensitivity with and without the use of skeletal staining methods. Also larval length was determined as an additional endpoint in this study, because it is known that exposure to xenobiotics can cause a reduced body length [38,39].

No significant differences in viability were observed between 5 dpf zebrafish larvae treated with any of the compounds and the control groups, which is in line with what was reported earlier [8]. The morphological assessment of 5 dpf larvae exposed to the highest concentration of compound 5, however, revealed malformations of the jaw, brain and heart, that were not observed earlier by Ball et al. (2014) [8]. The reason why these malformations were not observed before is unknown, as the compound was tested in several labs and the same morphological assessment method was used. Moreover, compounds 5

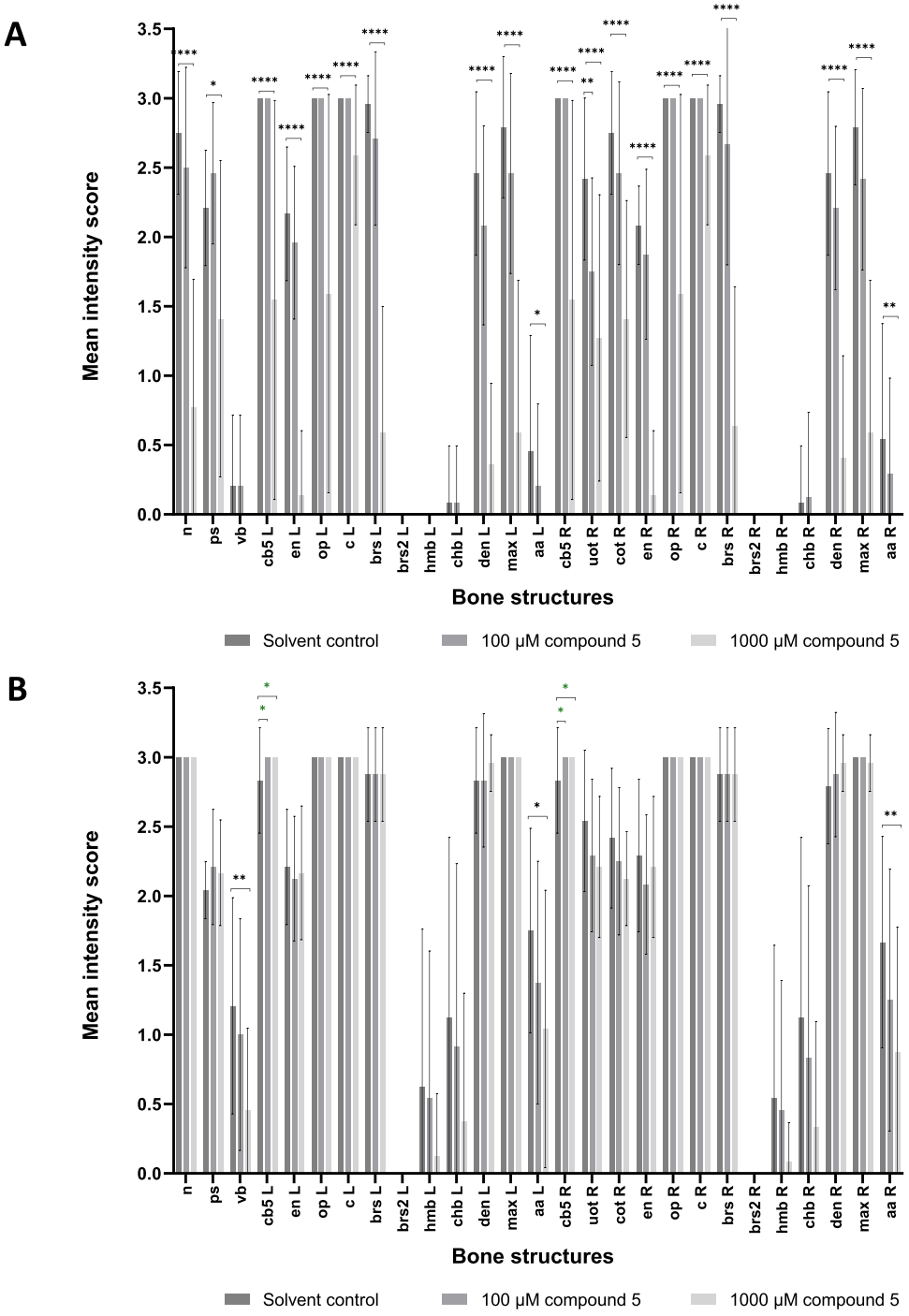


Fig. 18. Mean intensity score of AR live stained bones of 5 dpf larvae treated with proprietary compound 5. A) shows the results of replicate 1 and B) shows the results of replicate 2. Both groups were compared to the solvent control group. Significant differences between the solvent control and the test groups are indicated. P $\leq$ 0.05 (*) and p $\leq$ 0.01 (**). The color of the asterisks indicate that the solvent control was significantly more (black) or less (green) intense stained than the other group.

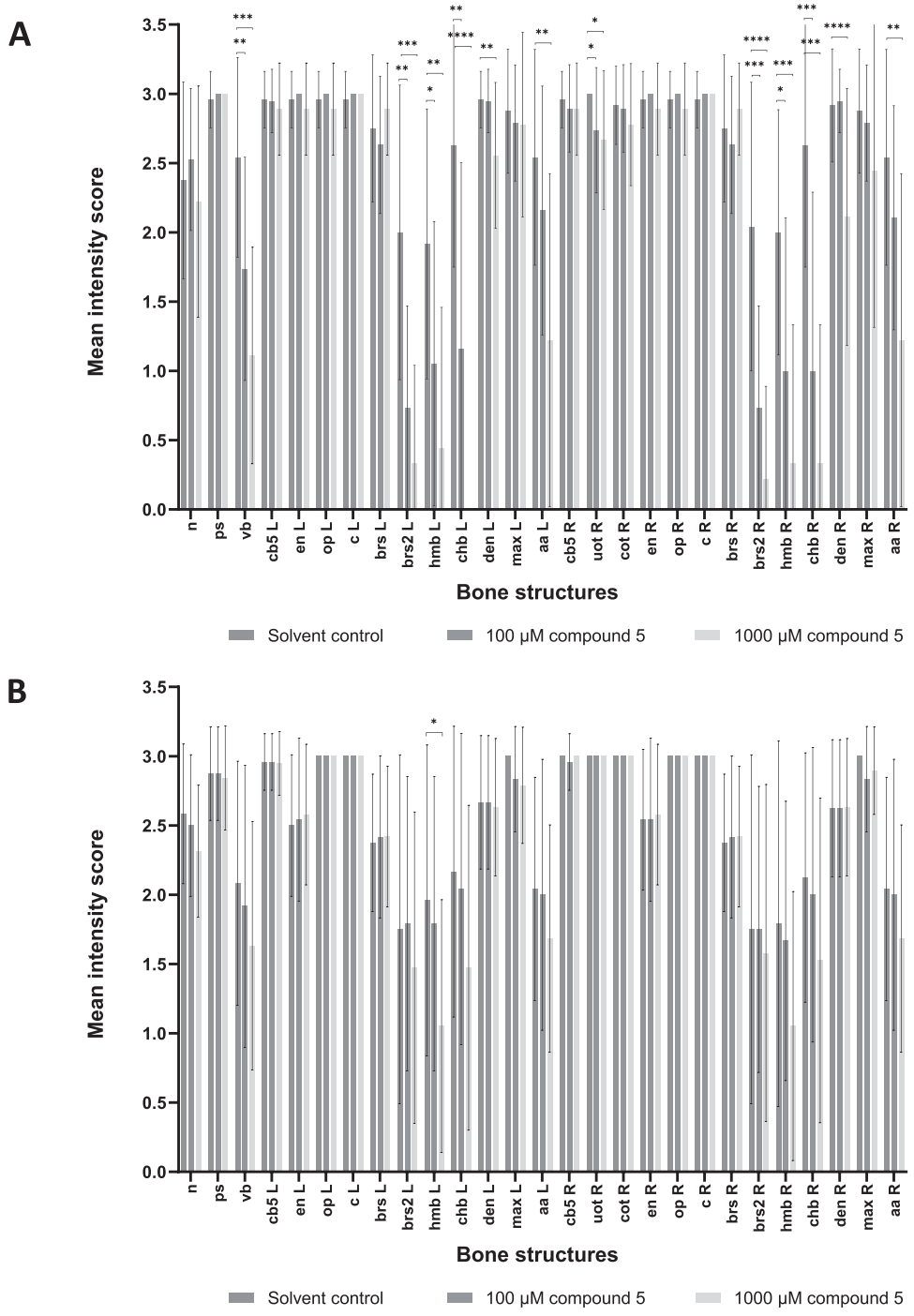


Fig. 19. Mean intensity score of AR live stained bones of 9 dpf larvae treated with proprietary compound 5. A) shows the results of replicate 1 and B) shows the results of replicate 2. Both groups were compared to the solvent control group. Significant differences between the solvent control and the test groups are indicated. P $\leq$ 0.05 (*), p $\leq$ 0.01 (**), p $\leq$ 0.001 (***) and p $\leq$ 0.0001 (****).

and 9 caused a shortened body after exposure to 1,000 µM at 5 dpf. Interestingly, one of the malformations that was detected in EFD studies for compound 9 was shorter long bones in rats [8]. Compound A showed a reduced body length in all concentrations in the first replicate (not dose-dependent), while the larvae in the second replicate were not affected. As larval length was not determined in the studies by Lee, et al. (2013) and Ball, et al. (2014), a comparison with those studies was not possible [8,10]. In our study, the larvae exposed to compound 5 were further reared until 9 dpf to see if skeletal assessment at a later age would make the assay more sensitive than at 5 dpf. At 9 dpf, the body length of larvae exposed to 1,000 µM (2 replicates) and 100 µM (1 replicate) of compound 5 was reduced. Moreover, the viability of the 9 dpf larvae exposed to 1,000 µM of compound 5 was significantly lower. Interestingly, this lethal effect only became visible at 9 dpf, so when the exposure had already been stopped.

Levetiracetam appeared to be non-teratogenic in our study, which is in agreement with the study by Lee, et al. (2013) [10]. The results of compound A are inconclusive due to the differences for larval length between both replicates. In contrast to what was found by Ball, et al. (2014), the results of the viability, length and morphological assessment

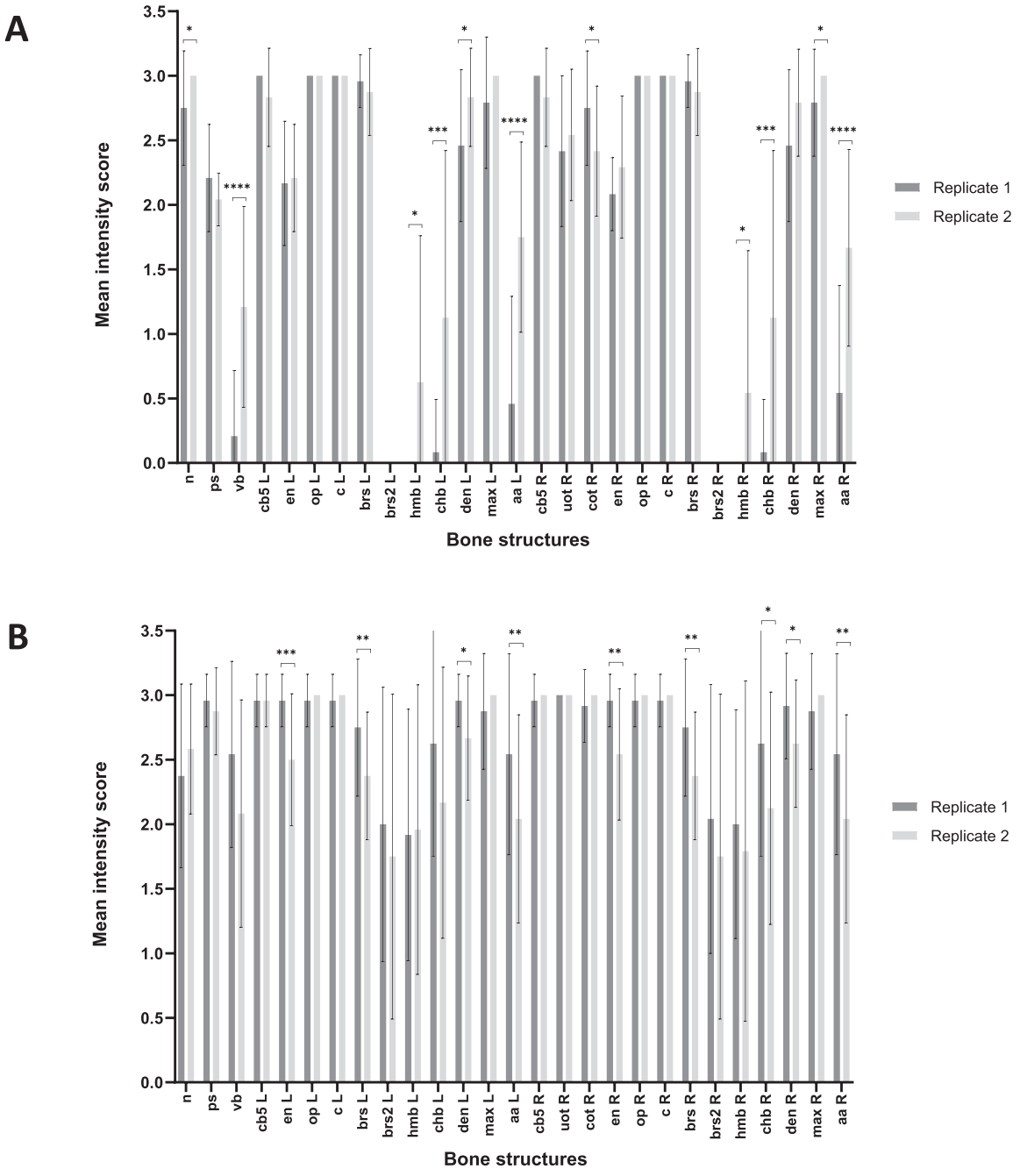


Fig. 20. Variability in staining between control larvae of the two replicates of the AR live experiment. A) shows the results using 5 dpf larvae and B) shows the results using 9 dpf larvae. Significant differences between the two groups of solvent controls are indicated. $P \leq 0.05$ (*), $p \leq 0.01$ (**), $p \leq 0.001$ (***) and $p \leq 0.0001$ (****).

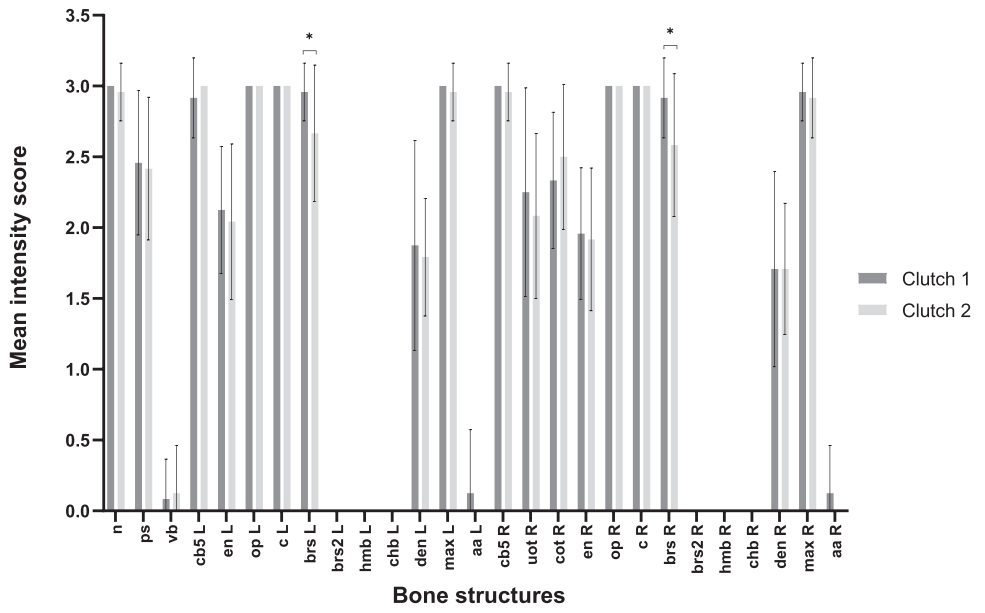


Fig. 21. Variability in staining between two clutches of control larvae (same spawning day). Significant differences between the two clutches are indicated. $P \leq 0.05$ (*).

in our study indicate that compound 5 might be teratogenic in zebrafish larvae. The reduced larval length observed for compound 9, which is a sign of growth retardation, is a first sign that this compound might have a negative impact on zebrafish development. Despite the fact that negative effects were already found for compound 5 and 9 without skeletal staining, we still evaluated whether a skeletal assessment could make the assay more sensitive, i.e., reveal structural effects already at lower concentrations.

The AR-AB fixed staining was the first staining that was tested, since this method allowed visualization of most skeletal structures as both bone and cartilage structures are stained. However, for all four compounds the results were inconsistent due to the variability between replicates. This variability was not found with the gross morphology, length, and viability assessments, except for the length of larvae exposed to compound A. Especially the staining intensity of skeletal structures showed to be very different between the replicates of all four compounds. In contrast, there were hardly any differences in shape and almost no significant differences in angle size. The mean angle size was only significantly increased at 100 μM of compound A in replicate 1. In addition to the differences between replicates, the results of the intensity staining revealed often more pronounced effects after exposure to lower, instead of higher, concentrations of LTC (i.e., not dose-dependent). For compound 9, increased, instead of the expected decreased, staining intensities were observed for some structures (i.e., cb5 and ps). However, this might be due to the fact that the solvent control was significantly less stained for cb5 and ps when compared the medium control group. Therefore, the increase in intensity is likely due to a decreased staining in the control group. Due to the high variability between replicates and inconsistent results, it was difficult to determine whether the AR-AB staining could pick up malformations that were not found by using only a gross morphological, length and viability assessment. In contrast to the use of AR-AB as a skeletal staining, the length assessment was consistent between the replicates of three of the compounds. Therefore, including larval length as a standard endpoint in zebrafish embryo assays could be considered. As a staining method should be able to provide a consistent staining, and especially variability in staining intensity was found to be the most prominent cause of differences between replicates, we decided to explore other staining methods. As most significant differences were found in bones and not in cartilages, the use of skeletal staining methods that only visualize bone might be sufficient to increase the sensitivity. Moreover, we noticed that a manual cartilage assessment is not suitable for zebrafish embryo developmental toxicity assays due to the difficulty in manually assigning scores to structures that were always stained with low intensity (e.g. cb, bh, bb, pty, etc.; see Fig. 1). As assigning a score of e.g. 1 or 2 to these structures was difficult in many cases, this can easily have led to variability between replicates and consequent false differences with controls. On the other hand, the heavily stained structures were easily scored (e.g. mk, qu, hys, etc.), but we noted no difference between the controls and the exposed groups. As such, if the use of a cartilage staining would be further explored in the future, we recommend to use it in combination with an automated scoring method. By developing and using a software program that can analyze the images of the skeletal structures (that were taken by an automated zebrafish handling and imaging system to make sure the structures are imaged in exactly the same position), the observer bias can be reduced, resulting in more correct scoring of the structures that were lightly stained.

The observed malformations of the jaw and the reduced length after exposure to compound 5 may suggest that, of all compounds, an impact on bone development is the most likely for compound 5. Therefore, this compound was used to test the other staining methods.

The calcein live staining caused a high degree of variability in fluorescence between control larvae, which prevented the assessment of the different bone structures. Hence, our findings indicated that calcein cannot be used for bone assessment in developing zebrafish embryos. This is in contrast to what was found in the study by Du, et al. (2001) [24]. They reported that calcein was a more sensitive staining method when compared to AR-AB staining. In their study, calcein revealed most skeletal structures, whereas there was almost no staining from alizarin red and only a subset of the structures could be visualized using alcian blue. However, they did not use an acid-free AR-AB staining method. Walker and Kimmel (2007) reported that the combination of alizarin red and alcian blue is problematic if acidic conditions are used to differentiate the tissue in the alcian blue staining. This acid demineralizes bone, which affects the alizarin red staining as this depends on mineralization of the bone matrix [23]. Hence, this might explain why Du et. al barely found alizarin red stained structures, and therefore reported a higher sensitivity for the calcein staining. Moreover, as alcian blue is a cartilage staining (i.e., it stains the extracellular matrix that is associated with chondrocytes [40]) and not a bone staining, it is reasonable that more

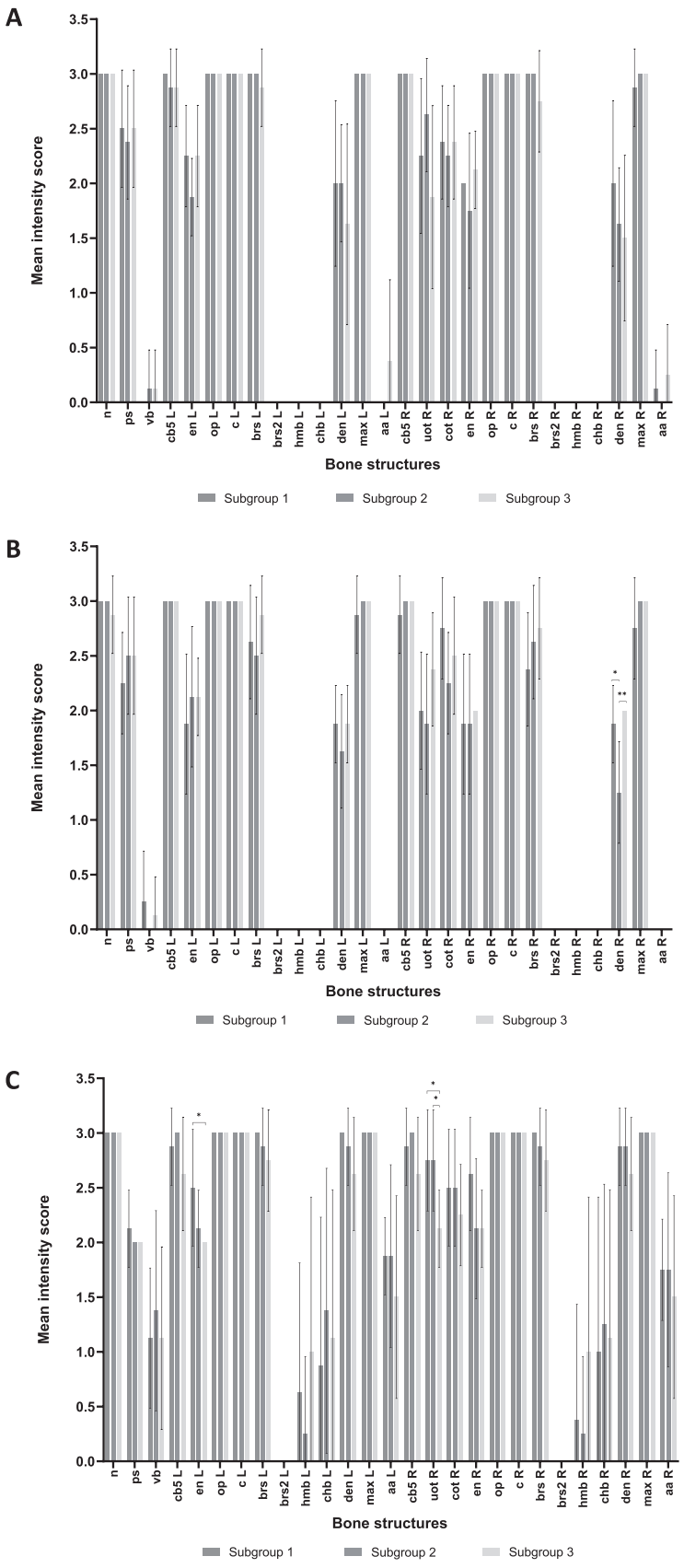


Fig. 22. Variability in staining between subgroups of the same batch of control larvae. In A) the results of batch 1 are shown, in B) the results of batch 2 are shown and in C) the results of another batch (i.e., the solvent control group of the second replicate of the compound 5 experiment at 5 dpf) are shown. For each of these batches, the batch was divided into 3 subgroups of 8 fish. All the subgroups of a batch were compared to each other. Significant differences between the three subgroups are indicated. $P \leq 0.05$ (*) and $p \leq 0.01$ (**).

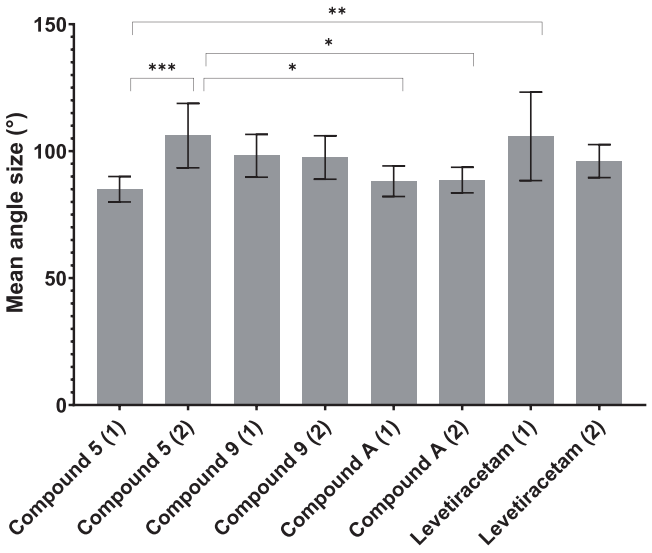


Fig. 23. Differences in mean angle size between the medium controls of all fixed staining experiments. P ≤ 0.05 (*), p ≤ 0.01 (**) and p ≤ 0.001 (***).

calcified skeletal structures were stained with calcein, as the latter specifically binds to calcified skeletal structures. Due to the variability in fluorescence in the controls, calcein was not further used in our study and a third staining, the AR live staining, was explored.

Since an AR live staining allows staining of living larvae, the until 5 dpf exposed larvae could be reared until 9 dpf and checked again for skeletal malformations. The aim was to investigate whether a skeletal assessment at a later age (i.e., 9 dpf) would increase the sensitivity of the assay compared to an assessment at 5 dpf. Although the AR live staining looked promising during an initial test with a test compound using only one replicate in a previous study [16], the variability in staining intensity between replicates of compound 5 appeared to be a problem for this staining as well. This variability was not detected in Hoyberghs, et al. (2020), as only one replicate was used [16]. In our present study, the staining intensity of almost all affected structures was only reduced in one of the replicates at both concentrations at 5 and 9 dpf. In addition, the results indicated that a bone staining is not more sensitive at 9 dpf, as already a reduced body length and increased mortality was observed at this age. In the group treated with the highest concentration of compound 5, only 9 fish survived. At 5 dpf, the AR live staining was not more sensitive in the 1,000 μM group, as gross morphological abnormalities were present. Interestingly, a reduced staining intensity of the dentary and maxilla was observed in replicate 1. This is in line with the observed jaw malformations of the gross morphological assessment. Moreover, in this replicate also a malformation in the shape (i.e., reduced length) of the parasphenoid was found, which is in line with the reduced larval length. Only in the 100 μM group a bone staining at 5 dpf showed to be more sensitive than gross morphology. However, due to the variability in staining between replicates, it is not clear if including an AR live bone staining visualizes malformations at lower concentrations. Nevertheless, the results indicate that a bone staining is not more sensitive at 9 dpf than at 5 dpf.

Due to inconsistent results between replicates of compound-treated zebrafish larvae, additional analyses on AR live and AR-AB-stained control larvae were performed to check whether replicates of control larvae also showed this variability. For both staining methods, many differences in bone staining intensity were found between replicates of control larvae. A reliable staining method should provide consistent results in independent experiments with control larvae. For the AR live staining at 5 and 9 dpf, a lot of differences in bone staining were present between two groups of control larvae. However, as both replicates were conducted with 9 months of time in between, the use of different batches of adult fish and the use of freshly prepared staining solution for both replicates may have caused additional variability. Therefore, additional AR live staining experiments on control larvae were conducted to assess whether biological variability in (onset of) bone development or technical aspects caused the differences in bone staining intensity between replicates. In these experiments, variability in staining between larvae of two clutches of the same spawning day, as well as variability in staining between larvae originating from the same clutch of zebrafish were found. These findings are in line with the study by Cubbage and Mabee (1996) that specifically looked at the development of the zebrafish cranium and paired fins. Using a smaller number of zebrafish larvae per age, inter-individual variability in the onset of ossification was also clearly observed in their study. Moreover, the degree of variability was dependent on the bone structure [32]. The rapid development of zebrafish might be a possible reason for the variability in staining we observed. Variability in skeletal staining also occurs in the *in vivo* mammalian studies. As zebrafish development goes faster than mammalian development (days in zebrafish vs. weeks in mammals), also the staging of early zebrafish larvae is probably more variable due to this rapid development. As such, the pace of zebrafish development could potentially explain the variability in staining we observed.

In conclusion, we found that biological variability in (onset of) bone development is clearly present in zebrafish larvae. This biological variability hampers the detection of (subtle) treatment-related bone effects that are not picked-up by a gross morphological assessment.

Ethics statement

All animal work was carried out in accordance with the EU Directive for the protection of animals used for purposes (2010/63/EU) and UK Animals Scientific Procedures Act (ASPA) 1986. Experimental procedures were carried out under personal and project licenses granted by the UK Home Office under ASPA, and ethically approved by the Animal Welfare and Ethical Review Body at the University of Exeter.

Funding

This work was supported by a special research fund of the University of Antwerp (grant number 43833).

CRediT authorship contribution statement

Pia Wilhelmi: Writing – review & editing, Funding acquisition. **Arantza Muriana:** Writing – review & editing, Funding acquisition. **Manon Beekhuijzen:** Writing – review & editing, Funding acquisition. **Matthew Burbank:** Writing – review & editing, Funding acquisition. **Jonathan Ball:** Writing – review & editing, Writing – original draft, Visualization, Resources, Methodology, Investigation, Funding acquisition, Conceptualization. **Maciej Trznadel:** Methodology, Investigation, Formal analysis. **Jente Hoyberghs:** Writing – review & editing, Writing – original draft, Visualization, Methodology, Investigation, Formal analysis. **Steven Van Cruchten:** Writing – review & editing, Writing – original draft, Supervision, Resources, Project administration, Methodology, Investigation, Funding acquisition, Conceptualization. **Nicola Powles-Glover:** Writing – review & editing, Funding acquisition. **Ainhoa Letamendia:** Writing – review & editing, Funding acquisition.

Declaration of Competing Interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Steven Van Cruchten reports financial support was provided by University of Antwerp. Jonathan Ball reports financial support and equipment, drugs, or supplies were provided by University of Exeter. Steven Van Cruchten is Associate Editor for Reproductive Toxicology and Guest Editor for the Special Issue on NAMs for DART testing If there are other

Table 6
Differences in staining intensity between the medium controls of all fixed staining experiments. All medium control groups were compared to each other and all A) bone and B) cartilage structures that differed significantly are indicated. The asterisks represent the significance. $P \leq 0.05$ (*), $p \leq 0.01$ (**), $p \leq 0.001$ (***) and $p \leq 0.0001$ (****). For vb, all scores were 0 (NA). If no left (L) or right (R) were indicated, the result of both sides was the same. Abbreviations: compound 5 (C5), compound 9 (C9), compound A (CA), levetiracetam (LTC), replicate 1 (.1) and replicate 2 (.2).

A

	N	PS	VB	CB5	T	UOT L	UOT R	COT	EN	OP	C	BRS
C5.1 - C9.1		**	NA									
C5.1 - C9.2			NA									*
C5.1 - C5.2			NA		**							
C5.1 - LTC.1		***	NA			*						
C5.1 - CA.1		***	NA		*					*		
C5.1 - LTC.2		**	NA		**							
C5.1 - CA.2		*	NA									
C9.1 - C9.2			NA									
C9.1 - C5.2			NA									
C9.1 - LTC.1			NA									
C9.1 - CA.1	**		NA									
C9.1 - LTC.2			NA									
C9.1 - CA.2			NA									*
C9.2 - C5.2			NA									
C9.2 - LTC.1			NA									
C9.2 - CA.1			NA									
C9.2 - LTC.2			NA									*
C9.2 - CA.2			NA									***
C5.2 - LTC.1			NA				*					
C5.2 - CA.1			NA									
C5.2 - LTC.2			NA									
C5.2 - CA.2			NA									
LTC.1 - CA.1			NA									
LTC.1 - LTC.2			NA									
LTC.1 - CA.2			NA									
CA.1 - LTC.2	*		NA									
CA.1 - CA.2			NA									
LTC.2 - CA.2			NA									

B

	ETH	BH	BB	MK	QU	HYS	CH	CB	AUD	PEC FIN	PTY
C5.1 - C9.1								**			
C5.1 - C9.2											
C5.1 - C5.2		****						****	****		
C5.1 - LTC.1								**			
C5.1 - CA.1								**			
C5.1 - LTC.2											
C5.1 - CA.2								***			
C9.1 - C9.2											
C9.1 - C5.2		****					*		****		
C9.1 - LTC.1											
C9.1 - CA.1											
C9.1 - LTC.2											
C9.1 - CA.2											
C9.2 - C5.2		**							****		
C9.2 - LTC.1											
C9.2 - CA.1											
C9.2 - LTC.2											
C9.2 - CA.2											
C5.2 - LTC.1		****					*		****		
C5.2 - CA.1		*					*		****		
C5.2 - LTC.2		****	***						****		*
C5.2 - CA.2		**					*		****		
LTC.1 - CA.1											
LTC.1 - LTC.2											
LTC.1 - CA.2											
CA.1 - LTC.2	*										
CA.1 - CA.2											
LTC.2 - CA.2	*										

authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors would like to thank the staff in the Aquatic Resource Centre at the University of Exeter, for the supply and maintenance of the zebrafish.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the

online version at doi:10.1016/j.reprotox.2024.108615.

References

[1] T.B. Knudsen, S.C. Fitzpatrick, K. Nadira De Abrew, L.S. Birnbaum, A. Chappelle, G. P. Daston, D.C. Dolinoy, A. Elder, S. Euling, E.M. Faustman, K.P. Fedinick, J. A. Franzosa, D.E. Haggard, L. Haws, N.C. Kleinstreuer, G.M. Buck Louis, D. L. Mendrick, R. Rudel, K.S. Saili, T.T. Schug, R.L. Tanguay, A.E. Turley, B. A. Wetmore, K.W. White, T.J. Zurlinden, FutureTox IV workshop summary: predictive toxicology for healthy children, Toxicol. Sci. 180 (2021) 198–211, https://doi.org/10.1093/toxsci/kfab013.

[2] A.H. Piersma, E. Genschow, A. Verhoef, M.Q.I. Spanjersberg, N.A. Brown, M. Brady, A. Burns, N. Clemann, A. Seiler, H. Spielmann, Validation of the postimplantation rat whole-embryo culture test in the international ECVAM validation study on three in vitro embryotoxicity tests, Altern. Lab. Anim. 32 (2004) 275–307, https://doi.org/10.1177/026119290403200307.

[3] J. Schumann, Teratogen screening: state of the art, Avicenna J. Med. Biotechnol. 2 (2010) 115–121.

[4] K.C. Brannen, J.M. Panzica-Kelly, T.L. Danberry, K.A. Augustine-Rauch, Development of a zebrafish embryo teratogenicity assay and quantitative prediction model. Birth Defects Res. B. Dev. Reprod. Toxicol. 89 (2010) 66–77, https://doi.org/10.1002/bdrb.20223.

[5] S. Cassar, I. Adatto, J.L. Freeman, J.T. Gamse, I. Iturria, C. Lawrence, A. Muriana, R.T. Peterson, S. Van Cruchten, L.I. Zon, Use of zebrafish in drug discovery toxicology, Chem. Res. Toxicol. 33 (2020) 95–118, https://doi.org/10.1021/acs.chemrestox.9b00335.

[6] A. Yamashita, H. Inada, K. Chihara, T. Yamada, J. Deguchi, H. Funabashi, Improvement of the evaluation method for teratogenicity using zebrafish embryos. J. Toxicol. Sci. 39 (2014) 453–464, https://doi.org/10.2131/jts.39.453.

[7] M. Burbank, F. Gautier, N. Hewitt, A. Detroyer, L. Guillet-Revol, L. Carron, T. Wildemann, T. Bringel, A. Riu, A. Noel-Voisin, N. De Croze, M. Léonard, G. Ouédraogo, Advancing the use of new approach methodologies for assessing teratogenicity: Building a tiered approach, Reprod. Toxicol. 120 (2023) 108454, https://doi.org/10.1016/j.reprotox.2023.108454.

[8] J.S. Ball, D.B. Stedman, J.M. Hillegass, C.X. Zhang, J. Panzica-kelly, A. Coburn, B. P. Enright, B. Tornesi, H.R. Amouzadeh, M. Hetheridge, A.L. Gustafson, K. A. Augustine-rauch, Fishing for teratogens: a consortium effort for a harmonized zebrafish developmental toxicology assay, Toxicol. Sci. 139 (2014) 210–219, https://doi.org/10.1093/toxsci/kfu017.

[9] A.-L. Gustafson, D.B. Stedman, J. Ball, J.M. Hillegass, A. Flood, C.X. Zhang, J. Panzica-Kelly, J. Cao, A. Coburn, B.P. Enright, M.B. Tornesi, M. Hetheridge, K. A. Augustine-Rauch, Inter-laboratory assessment of a harmonized zebrafish developmental toxicology assay - progress report on phase I, Reprod. Toxicol. 33 (2012) 155–164, https://doi.org/10.1016/j.reprotox.2011.12.004.

[10] S.H. Lee, J.W. Kang, T. Lin, J.E. Lee, D. Il Jin, Teratogenic potential of antiepileptic drugs in the zebrafish model, Biomed. Res. Int. 2013 (2013) 726478, https://doi.org/10.1155/2013/726478.

[11] B. Pruvot, Y. Quiroz, A. Voncken, N. Jeanray, A. Piot, J.A. Martial, M. Muller, A panel of biological tests reveals developmental effects of pharmaceutical pollutants on late stage zebrafish embryos, Reprod. Toxicol. 34 (2012) 568–583, https://doi.org/10.1016/j.reprotox.2012.07.010.

[12] I.W.T. Selderslaghs, R. Blust, H.E. Witters, Feasibility study of the zebrafish assay as an alternative method to screen for developmental toxicity and embryotoxicity using a training set of 27 compounds, Reprod. Toxicol. 33 (2012) 142–154, https://doi.org/10.1016/j.reprotox.2011.08.003.

[13] Y.-S. Song, M.-Z. Dai, C.-X. Zhu, Y.-F. Huang, J. Liu, C.-D. Zhang, F. Xie, Y. Peng, Y. Zhang, C.-Q. Li, L.-J. Zhang, Validation, optimization, and application of the zebrafish developmental toxicity assay for pharmaceuticals under the ICH S5(R3) guideline, Front. Cell Dev. Biol. 9 (2021) 2426, https://www.frontiersin.org/article/10.3389/fcell.2021.721130.

[14] E. Teixidó, E. Piqué, J. Gómez-Catalán, J.M. Llobet, Assessment of developmental delay in the zebrafish embryo teratogenicity assay, Toxicol. Vitr. 27 (2013) 469–478, https://doi.org/10.1016/j.tiv.2012.07.010.

[15] K. Van den Bulck, A. Hill, N. Mesens, H. Diekman, L. De Schaepdrijver, L. Lammens, Zebrafish developmental toxicity assay: a fishy solution to reproductive toxicity screening, or just a red herring? Reprod. Toxicol. 32 (2011) 213–219, https://doi.org/10.1016/j.reprotox.2011.06.119.

[16] J. Hoyberghs, C. Bars, C. Pype, K. Foubert, M. Ayuso Hernando, C. Van Ginneken, J. Ball, S. Van Cruchten, Refinement of the zebrafish embryo developmental toxicity assay, MethodsX 7 (2020) 101087, https://doi.org/10.1016/j.mex.2020.101087.

[17] C. Bars, J. Hoyberghs, A. Valenzuela, L. Buyssens, M. Ayuso, C. Van Ginneken, A. J. Labro, K. Foubert, S.J.Van Cruchten, Developmental toxicity and biotransformation of two anti-epileptics in zebrafish embryos and early larvae, Int. J. Mol. Sci. 22 (2021), https://doi.org/10.3390/ijms222312696.

[18] U. Strähle, S. Scholz, R. Geisler, P. Greiner, H. Hollert, S. Rastegar, A. Schumacher, I. Selderslaghs, C. Weiss, H. Witters, T. Braunbeck, Zebrafish embryos as an alternative to animal experiments–a commentary on the definition of the onset of protected life stages in animal welfare regulations, Reprod. Toxicol. 33 (2012) 128–132, https://doi.org/10.1016/j.reprotox.2011.06.121.

[19] EMA, ICH S5 (R3) guideline on reproductive toxicology: Detection of Toxicity to Reproduction for Human Pharmaceuticals Step 5, Eur. Med. Agency. Comm. Med. Prod. Hum. Use. 5 (2020).

[20] K. Augustine-Rauch, C.X. Zhang, J.M. Panzica-Kelly, In vitro developmental toxicology assays: a review of the state of the science of rodent and zebrafish whole embryo culture and embryonic stem cell assays, Birth Defects Res. Part C. - Embryo Today Rev. 90 (2010) 87–98, https://doi.org/10.1002/bdrc.20175.

[21] M. McNerney, D. Potter, K. Augustine-Rauch, P. Barrow, B. Beyer, K. Brannen, S. Engel, B.P. Enright, W.S. Nowland, N. Powles-Glover, S. Powlin, M. J. Schneidkraut, D. Stanislaus, K. Turner, M. Graziano, Concordance of 3 alternative teratogenicity assays with results from corresponding in vivo embryo-fetal development studies: final report from the International Consortium for Innovation and Quality in Pharmaceutical Development (IQ) DruSafe working grou, Regul. Toxicol. Pharmacol. 124 (2021) 104984, https://doi.org/10.1016/j.yrtph.2021.104984.

[22] S.L. Makris, H.M. Solomon, R. Clark, K. Shiota, S. Barbellion, J. Buschmann, M. Ema, M. Fujiwara, K. Grote, K.P. Hazelden, K.W. Hew, M. Horimoto, Y. Ooshima, M. Parkinson, L.D. Wise, Terminology of developmental abnormalities in common laboratory mammals (version 2), Birth Defects Res. Part B - Dev. Reprod. Toxicol. 86 (2009) 227–327, https://doi.org/10.1002/bdrb.20200.

[23] M.B. Walker, C.B. Kimmel, A two-color acid-free cartilage and bone stain for zebrafish larvae, Biotech. Histochem. 82 (2007) 23–28, https://doi.org/10.1080/10520290701333558.

[24] S.J. Du, V. Frenkel, G. Kindschi, Y. Zohar, Visualizing normal and defective bone development in zebrafish embryos using the fluorescent chromophore calcein, Dev. Biol. 238 (2001) 239–246, https://doi.org/10.1006/dbio.2001.0390.

[25] G.C. Paull, K.J.W. Van Look, E.M. Santos, A.L. Filby, D.M. Gray, J.P. Nash, C. R. Tyler, Variability in measures of reproductive success in laboratory-kept colonies of zebrafish and implications for studies addressing population-level effects of environmental chemicals, Aquat. Toxicol. 87 (2008) 115–126, https://doi.org/10.1016/j.aquatox.2008.01.008.

[26] C.B. Kimmel, W.W. Ballard, S.R. Kimmel, B. Ullmann, T.F. Schilling, Stages of embryonic development of the zebrafish, Dev. Dyn. 203 (1995) 253–310, https://doi.org/10.1002/aja.1002030302.

[27] J. Hoyberghs, C. Bars, M. Ayuso, C. Van Ginneken, K. Foubert, S. Van Cruchten, DMSO concentrations up to 1% are safe to be used in the zebrafish embryo developmental toxicity assay, Front. Toxicol. 3 (2021) 804033, https://doi.org/10.3389/ftox.2021.804033.

[28] A.L. Gustafson, D.B. Stedman, J. Ball, J.M. Hillegass, A. Flood, C.X. Zhang, J. Panzica-Kelly, J. Cao, A. Coburn, B.P. Enright, M.B. Tornesi, M. Hetheridge, K. A. Augustine-Rauch, Inter-laboratory assessment of a harmonized zebrafish developmental toxicology assay - progress report on phase I, Reprod. Toxicol. 33 (2012) 155–164, https://doi.org/10.1016/j.reprotox.2011.12.004.

[29] J.M. Panzica-Kelly, C.X. Zhang, T.L. Danberry, A. Flood, J.W. DeLan, K.C. Brannen, K.A. Augustine-Rauch, Morphological score assignment guidelines for the dechorionated zebrafish teratogenicity assay. Birth Defects Res. B. Dev. Reprod. Toxicol. 89 (2010) 382–395, https://doi.org/10.1002/bdrb.20260.

[30] T. Parker, P.-A. Libourel, M.J. Hetheridge, R.I. Cumming, T.P. Sutcliffe, A. C. Goonesinghe, J.S. Ball, S.F. Owen, Y. Chomis, M.J. Winter, A multi-endpoint in vivo larval zebrafish (Danio rerio) model for the assessment of integrated cardiovascular function. J. Pharmacol. Toxicol. Methods 69 (2014) 30–38, https://doi.org/10.1016/j.vascn.2013.10.002.

[31] M. Seda, M. Geerlings, P. Lim, J. Jeyabalan-Srikaran, A.C. Cichon, P.J. Scambler, P. L. Beales, V. Hernandez-Hernandez, A.W. Stoker, D. Jenkins, An FDA-Approved drug screen for compounds influencing craniofacial skeletal development and craniosynostosis, Mol. Syndromol. 10 (2019) 98–114, https://doi.org/10.1159/000491567.

[32] C.C. Cubbage, P.M. Mabee, Development of the cranium and paired fins in the zebrafish Danio rerio (Ostariophysi, Cyprinidae), J. Morphol. 229 (1996) 121–160, https://doi.org/10.1002/(SICI)1097-4687(199608)229:2<121::AID-JMOR1>3.0.CO;2-4.

[33] T. Windhausen, S. Squifflet, J. Renn, M. Muller, BMP signaling regulates bone morphogenesis in zebrafish through promoting osteoblast function as assessed by their nitric oxide production, Molecules 20 (2015) 7586–7601, https://doi.org/10.3390/molecules20057586.

[34] J. Aceto, R. Nourizadeh-Lillabadi, R. Marée, N. Dardenne, N. Jeanray, L. Wehenkel, P. Aleström, J.J.W.A. van Loon, M. Muller, Zebrafish bone and general physiology are differently affected by hormones or changes in gravity, PLoS One 10 (2015) e0126928, https://doi.org/10.1371/journal.pone.0126928.

[35] S.T. Raterman, J.R. Metz, F.A.D.T.G. Wagener, J.W. den Hoff, Zebrafish models of craniofacial malformations: interactions of environmental factors, Front. Cell Dev. Biol. 8 (2020), https://doi.org/10.3389/fcell.2020.600926.

[36] Y. Javidan, T.F. Schilling, Development of cartilage and bone, Methods Cell Biol. 76 (2004) 415–436, https://doi.org/10.1016/s0091-679x(04)76018-5.

[37] S.C. Neuhauss, L. Solnica-Krezel, A.F. Schier, F. Zwartkruis, D.L. Stemple, J. Malicki, S. Abdelilah, D.Y. Stainier, W. Driever, Mutations affecting craniofacial development in zebrafish, Development 123 (1996) 357–367, https://doi.org/10.1242/dev.123.1.357.

[38] Q. Shi, Y. Zhuang, T. Hu, C. Lu, X. Wang, H. Huang, G. Du, Developmental toxicity of triclocarban in zebrafish (Danio rerio) embryos, J. Biochem. Mol. Toxicol. 33 (2019) 1–6, https://doi.org/10.1002/jbt.22289.

[39] Z. Hu, Y. Dang, C. Liu, L. Zhou, H. Liu, Acute exposure to ustiloxin A affects growth and development of early life zebrafish, Danio rerio, Chemosphere 226 (2019) 851–857, https://doi.org/10.1016/j.chemosphere.2019.04.002.

[40] T.F. Schilling, T. Piotrowski, H. Grandel, M. Brand, C. Heisenberg, Y. Jiang, D. Beuchle, M. Hammerschmidt, D.A. Kane, M.C. Mullins, F.J.M. Van Eeden, R.N. Kelsh, M. Furutani-seiki, M. Granato, P. Haffter, J. Odenthal, R.M. Warga, T. Trowe, C. Nüsslein-volhard, Jaw and branchial arch mutants in zebrafish I: branchial arches, (1996) 1–16. papers2://publication/uuid/529A94A6-ECE8-4046-BDC8-84B95B6F9B65.