Amended Draft new Test Guideline for the Rapid Estrogen ACTivity *In Vivo* assay (REACTIV)

For second WNT-review by 11 March 2024

Draft New Test Guideline For The Rapid Estrogen ACTivity In Vivo (REACTIV) Assay

January 2024

INTRODUCTION

1. The Rapid Estrogen ACTivity In Vivo (REACTIV) Assay test guideline describes an aquatic assay that utilises transgenic *Oryzias latipes* (Japanese medaka) eleutheroembryos at day post hatch zero (DPH0; see Annex 1 for abbreviations and definitions), in a multi-well plate format to identify chemicals active on the estrogen axis. The REACTIV assay was designed as a screening assay to provide a medium throughput and short-term assay to measure the response of eleutheroembryos to chemicals potentially active on the estrogen axis (Spirhanzlova *et al.*, 2016). A description of the modes of action known to be covered by the assay can be found below (see §10). The REACTIV assay is intended to classify chemicals into potentially active on the estrogen axis or inactive but the REACTIV assay was not designed to establish NOAEC or ECx values. The REACTIV assay is intended to be placed at level 3 of the OECD conceptual framework for the testing of endocrine disrupters (OECD, 2018).
2. The Japanese medaka fish, *O. latipes*, is the test species selected for the REACTIV assay. This species is utilized in a number of validated OECD Test Guidelines including: OECD TG 203 (Fish Acute Toxicity Test; OECD, 2019), OECD TG 210 (Fish Early Life Stage Toxicity Test; OECD, 2013), OECD TG 212 (Fish Short Term Toxicity Test on Embryo and Sac-fry Stages; OECD, 1998), OECD TG 229 (Fish Short Term Reproduction Assay; OECD, 2012), OECD TG 230 (21-day Fish Assay; OECD, 2009), OECD TG 234 (Fish Sexual Development Test; OECD, 2011); OECD TG 240 (Medaka Extended One Generation Reproduction Test; OECD, 2023) and OECD TG 251 (Rapid Androgen Disruption Activity Reporter; OECD, 2022).
3. The REACTIV assay is transcription-based and uses a transgenic medaka line harbouring the *chgh-gfp* genetic construct. The *chgh-gfp* transgenic line used in the REACTIV assay harbours 2.047 kb of the medaka choriogenin H gene promoter immediately upstream of the start codon driving expression of Green Fluorescent Protein (GFP) coding sequence. The *chgh-gfp* transgene is expressed in the liver of the medaka in response to activation of estrogen axis signalling. There is also a non-inducible ectopic expression of GFP in some cells of the heart and head at eleutheroembryonic life stages. This allows visual confirmation that the developing fry are transgenic.
4. The promoter region present in the transgene has been shown to contain putative estrogen response elements (ERE) and the expression of the transgene has been demonstrated to be significantly modulated in the presence of estrogen receptor (ER) agonists, antagonists and compounds inducing or inhibiting steroidogenic enzymes (Kurauchi *et al.*, 2005, 2008; Spirhanzlova *et al.*, 2016).
5. Choriogenin genes, much like vitellogenin, are required for egg production in fish. Their expression is upregulated in response to estrogen axis signalling. As a terminal step, their expression and the expression of GFP in the *chgh-gfp* medaka line represents the overall or net effects of both endogenous and exogenous factors altering estrogen axis signalling (alterations in production, transport, metabolism and excretion of hormones as well as activation and inhibition of ER).
6. Before performing the REACTIV assay, the laboratory should verify that it has the certifications that may be required by local regulations on the use of transgenic organisms. The REACTIV assay should be performed using the *chgh-gfp* transgenic line used for the test guideline development, which is commercially available (OECD, REACTIV assay validation report). The use of another transgenic line based on the Choriogenin H promoter driving the expression of GFP or another reporter gene requires a complete OECD

validation to adapt the validation criteria, the statistical analysis and the fluorescence thresholds as well as the decision logic. Therefore, other transgenic lines could not be considered as appropriate for the implementation of the REACTIV assay.

7. This guideline proposal is based on an international interlaboratory validation study conducted from 2020 to 2022 (OECD, REACTIV assay validation report). The test has been validated in six laboratories with 18 mono-constituent test substances. Of these: four were tested in six laboratories; another six in five laboratories; another two in four laboratories; another one in three laboratories; another four in two laboratories and another one in one laboratory.

8. The endpoint measured is fluorescence in the liver of eleutheroembryos. A very low level of fluorescence is observed in unexposed eleutheroembryos. When transcription of the genetic construct is activated or inhibited following chemical exposure, eleutheroembryos express more or less GFP and, therefore, emit more or less fluorescence. The level of fluorescence of eleutheroembryos exposed to the test chemical is compared to that of eleutheroembryos not exposed to the test chemical.

9. The test chemical is tested in the presence and absence of 30 µg/L of testosterone (T). As circulating estrogen levels remain very low at this eleutheroembryonic life stage, adding T to the test medium allows the detection of substances affecting T availability or antagonising ERs as it is metabolised *in vivo* into estradiol by the cytochrome P450 enzyme aromatase (CYP19). The concentration of T used for the co-treatment was determined empirically. The chosen concentration (30 µg/L) is the lowest concentration of T inducing a statistically significant increase in fluorescence following a 24 h exposure. The differential gene expression induced by the combination of T and the tested chemical is, therefore, a laboratory induced phenomenon, not observed in the absence of exogenous T at this developmental stage, and thus is only indicative of the capacity of the test item to induce an (anti-)estrogenic activity and is currently not considered predictive of a physiological outcome. It does, however, allow mechanisms of action to be detected that would not be revealed in the absence of an aromatisable androgen such as alterations in aromatase activity or ER antagonism.

INITIAL CONSIDERATIONS AND LIMITATIONS

10. The assay measures the ability of a chemical to activate or inhibit transcription of the *chgh-gfp* genetic construct, whether directly through binding to ER or modifying the binding of estrogens to the ER, or indirectly by modifying the amount of estrogen available to activate the ER and thereby transcription of the *chgh-gfp* construct. To date the REACTIV assay has been shown to detect chemicals acting through various mechanisms of action including: ER agonists (e.g. estradiol, estrone); selective estrogen response modulators (e.g. tamoxifen); modulators of steroidogenesis including aromatase enzyme inhibitors (e.g. anastrozole and fadrozole), aromatase transcriptional inhibitors (e.g. prochloraz) and aromatase transcriptional inducers (e.g. estrogens) and chemicals requiring metabolic activation (e.g. T) (OECD, REACTIV assay validation report; Spirhanzlova *et al.*, 2016). In addition, the REACTIV assay potentially detects modulators of estrogen transport via interaction with plasma binding proteins. The REACTIV assay does not distinguish between the different mechanisms of action but provides information on whether a chemical acts as a global activator or inhibitor of the estrogen axis in the *O. latipes* eleutheroembryos.

11. As the transcription of the *chgh-gfp* construct requires the direct action of ER on the Choriogenin H promotor, chemicals affecting ER signalling through alternative signalling pathways that do not lead to an alteration in the interaction between ER and DNA (i.e., "non-genomic actions") are not expected to be detected by the REACTIV assay. This includes rapid estrogen signalling through membrane-localised ER. The relative prominence of non-genomic ER signalling is poorly understood at present.

12. A number of publications have supported the idea that early life stages of medaka are metabolically competent, although current data are insufficient to conclude on the full breadth of metabolic competency. The liver forms between DPF2 and 4, approximately 7 days before hatching and initiation of a REACTIV assay (Iwamatsu, 2004). Prior to liver formation at day post fertilisation 1 (DPF) it has been demonstrated that embryonic medaka could transform benzo(a)pyrene (BaP) into metabolites including BaP-3-glucuronide demonstrating UDP-glucuronosyltransferase (Hornung *et al.*, 2007). Strong cytochrome P450 (CYP) 1A activity has also been identified in the liver, gills and other organs in DPH1 medaka (Kashiwada *et al.*, 2007). In addition, CYP3A40 is expressed throughout medaka development, with CYP3A38 (the post-embryonic form) being expressed from DPH1 (Kullman and Hinton, 2001). Exposure of pre-hatch medaka to imidacloprid resulted in detection of hydroxyl and olefin metabolites at hatch, indicating the presence of CYP3A4 activity (Vignet *et al.*, 2019; Schulz-Jander *et al.*, 2002a; Schulz-Jander *et al.*, 2002b). In addition, urea-imidacloprid was also detected suggesting activity of CYP1A2, CYP2B6, CYP2D6 and/or CYP2E1. Expression of the steroidogenic enzymes P450 aromatase, 11β-hydroxylase and 3β-hydroxysteroid-dehydrogenase has been detected prior to hatch (Schiller et al., 2014). Indeed, pre-hatch medaka have been proposed as a model for studying the metabolism of anabolic steroids and have been demonstrated to produce a number of metabolites when exposed to metandienone including three mono-hydroxylated and one reduced metabolite that are produced by humans (Liu *et al.*, 2022).

13. This test guideline relies on the quantification of fluorescence in the whole eleutheroembryo. A limitation of this test guideline is that it should not be used for test chemicals emitting fluorescence between 500 and 550 nm (λ_{EM} = 500–550 nm) when excited at wavelengths between 450 and 500 nm (λ_{EX} = 450–500 nm) and is fluorescent and fluoresces within the eleutheroembryos. Test chemicals sharing these two properties may induce a fluorescence which could be interpreted as GFP signal, leading to the test chemical being incorrectly identified as active on the estrogen axis. A simple protocol to determine if the test chemical emits fluorescence is proposed in §31. This protocol requires the use of wild-type *O. latipes* eleutheroembryos.

14. The REACTIV assay should not be used to test chemicals falling outside of its applicability domain. The REACTIV assay is suitable for testing non-volatile substances. When considering testing mixtures or difficult test chemicals, upfront consideration should be given to whether such testing will yield results that are scientifically reliable. If the test guideline is used for the testing of a mixture, a UVCB (substances of unknown or variable composition, complex reaction products or biological materials) or a multi-constituent substance, its composition should, as far as possible, be characterized, e.g., by the chemical identity of its constituents, their quantitative occurrence and their substance-specific properties. Recommendations about the testing of difficult test chemicals (e.g., mixtures, UVCB or multi-constituent substances) are given in Guidance Document No. 23 (OECD, 2019a).

PRINCIPLE OF THE TEST

General experimental design

15. The general experimental design entails exposing DPH0 transgenic *chgh-gfp* medaka eleutheroembryos in six-well plates to a test chemical in the presence ("spiked mode") and absence ("unspiked mode") of a co-treatment with 30 µg/L of T. Three independent runs should be performed for each assay. It is recommended to use a minimum of five concentrations plus non-optional controls (a test medium control and/or solvent control, a 488 ng/L 17α-ethinylestradiol [EE2] control, a T control, an induction control for spiked groups and an inhibition control for spiked groups) per run. The test uses eight eleutheroembryos distributed in a single well per test condition (test concentrations and controls except the T control which comprises of two wells of eight eleutheroembryos), under a static regime. All six wells can be used on each six-well plate. It is not problematic to have two different test or control groups occupying the same plate as volatile chemicals are excluded, however, care should be taken to avoid cross-contamination. With five test concentrations and the non-optional controls, performed in three runs, the REACTIV assay uses 128 eleutheroembryos per run (136 if test medium and solvent control groups are both required), therefore, 384 eleutheroembryos are required for all three runs constituting an experiment (see Figure 1 and §17) or 408 if test medium and solvent control groups are both required. The exposure duration is 24 h with a 14:10 light: dark cycle. The assay measures GFP fluorescence in transgenic *chgh-gfp* eleutheroembryos by fluorescence imaging that transforms the fluorescence signal to a numerical format. A detailed overview of test conditions can be found in Annex 2.

Controls

16. The REACTIV assay requires the following non-optional control groups, all of which, except the test medium control, should have the same concentration of organic solvent (if one is used). Likewise, all groups exposed to test chemical should be exposed to the same concentration of solvent as the control groups.

a. Test medium and/or solvent control: 1 well with 8 organisms/well is exposed to test medium. This control defines the basal fluorescence level in the test medium. If a solvent is used, then this group is exposed to test medium plus the solvent used at the same concentration as all other groups. In some cases, such as a solvent being used with no historical data available, both a test medium and a solvent control group may be required, this is recommended to ensure mutual acceptance of data.
b. EE2 488 ng/L: 1 well with 8 organisms/well is exposed to 488 ng/L of EE2. This control establishes a close to maximal fluorescence observable for most mechanisms of action. It is also equivalent to the lowest concentration of EE2 inducing a statistically significant reduction in fecundity in a published 21-day medaka assay (Seki *et al.*, 2002).

c. T 30 µg/L: Two wells with 8 organisms/well are exposed to 30 µg/L of T. This control serves to induce estrogen axis signalling via endogenous conversion of T to estradiol. Induction of estrogen signalling in "T spiked mode" allows inhibition of estrogen axis signalling through ER antagonism, aromatase inhibition or repression of aromatase expression to be detected. It also allows induction of estrogen axis signalling through mechanisms such as increased aromatase expression to be detected. Data from two wells are pooled for this control to increase confidence in the mean fluorescence value.
d. Induction control for spiked groups: 1 well with 8 organisms/well is exposed to 64 ng/L of EE2 plus 30 µg/L of T. This control group confirms that an induction of fluorescence can be observed above that of the T 30 µg/L control group. Under 21-day flow-through conditions (OECD, 2009) in medaka, 64 ng/L of EE2 is the lowest concentration shown to have a physiological effect, consisting of testis-ova in one third of male fish (Seki *et al.*, 2002).
e. Inhibition control for spiked groups: 1 well with 8 organisms/well is exposed to 10 µg/L of fadrozole plus 30 µg/L of T. At 10 µg/L, fadrozole induces a modification in the gonadosomatic ratio of male fish within an OECD testing protocol (OECD TG 229; OECD, 2012)) (Ankley *et al.*, 2002).

The following additional control groups are optional, but are recommended for calibration of reading parameters in naïve laboratories as well as for quality control purposes. They constitute an EE2 standard curve and can also be used to derive a concentration-response relationship for EE2 allowing the results to be expressed in EE2 equivalents. The calculation of equivalence values is not required and is for informative purposes only as the result of the assay is that the test chemical is active or inactive only. If equivalence values are to be calculated, the optional controls below should be included in each run.

f. EE2 34 ng/L: 1 well with 8 organisms/well is exposed to 34 ng/L of EE2.
g. EE2 51 ng/L: 1 well with 8 organisms/well is exposed to 51 ng/L of EE2.
h. EE2 76 ng/L: 1 well with 8 organisms/well is exposed to 76 ng/L of EE2.
i. EE2 114 ng/L: 1 well with 8 organisms/well is exposed to 114 ng/L of EE2.
j. EE2 171 ng/L: 1 well with 8 organisms/well is exposed to 171 ng/L of EE2.

If the assay is to be performed with a solvent, it should be determined whether the results for the control groups pass validity criteria with the imaging system used for the readout, if not the experiment is considered invalid (see also §37).

Experimental runs

17. One test is composed of three independent and valid runs using 1 well x 8 organisms/treatment group/run (see Figure 1). At least five concentrations of the test chemical should be evaluated in the presence and absence of T. The same concentrations of the test item must be evaluated in each run. Each run should be performed using independent solutions (see §42). The runs should be conducted using eleutheroembryos from different spawnings. They can be performed sequentially or concurrently. The raw

data for a given test chemical are obtained by pooling the data from the three runs to ideally obtain n=24 fluorescence values in each treatment group, except the T control which ideally will provide n=48 values. Pooling of the data is obligatory for this test and is performed irrespective of whether the individual runs show positive or negative responses. It is performed to provide an improved estimate of the mean fluorescence value for each experimental group.

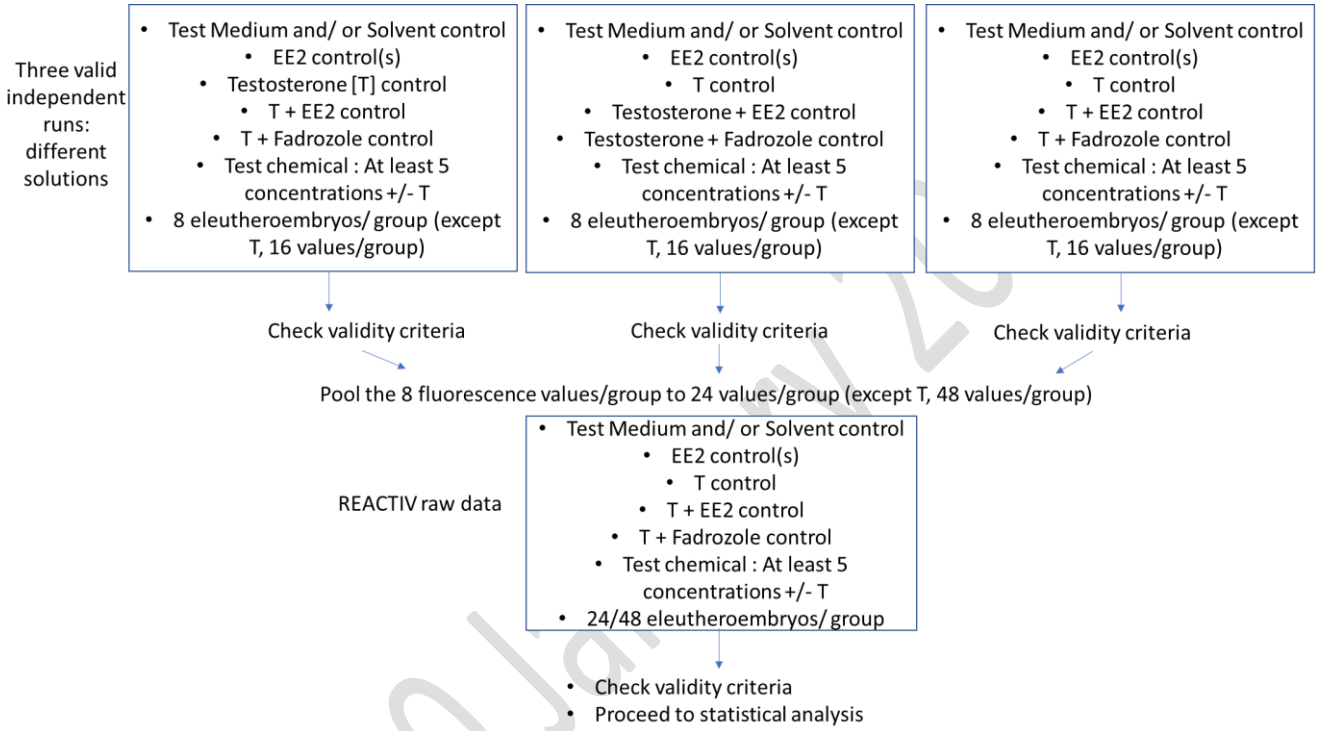

Figure 1: Overview of the REACTIV assay. ("+/- T" refers to spiked and unspiked groups).

INFORMATION ON THE TEST CHEMICAL

18. Available information on the test chemical should be reported (see §59).

19. Whenever possible, the solubility of the test chemical in the test medium should be known and a validated analytical method, of known accuracy, precision, and sensitivity, should be available for the quantification of the test chemical in the test solutions with reported efficiency and limit of quantification. Guidance for the validation of quantitative analytical methods can be found in the GD 204 (OECD, 2014a). Analytical determination of the test chemical concentration should be performed as described in §43.

DEMONSTRATION OF PROFICIENCY

Fluorescence quantification

20. The REACTIV assay relies on the quantification of the fluorescence emitted by each organism. To ensure that a proper and accurate quantification can be achieved, preliminary experiments should be conducted. These experiments are performed to calibrate the fluorescence imaging system and to ensure that a suitable dynamic range of fluorescence measurements can be read by the equipment. These experiments are detailed in Annex 3 and should be performed when a change in equipment or equipment settings has occurred. If an alternative system for fluorescence measurement is used, it should be calibrated and validated in the same way as detailed for a fluorescence imaging system (Annex 3). However, use of a fluorescence microscope equipped with an appropriate camera is the preferred method as this allows a quality control step to be performed on the pictures to identify misplaced eleutheroembryos or fluorescence signal not related to estrogen axis activation (fluorescent dust or fibres, fluorescent test chemical accumulated in the eleutheroembryo, abnormal fluorescent pattern).

Proficiency chemicals

21. Prior to routine use of this test guideline, laboratories should demonstrate technical proficiency by correctly categorising the four proficiency chemicals listed in Table 1. The expected statistical significance limits in Table 1 refer to the fluorescence of the group exposed to the indicated concentration of reference chemicals when compared to the relevant control. These limits were determined from the OECD REACTIV assay validation exercise (OECD, REACTIV assay validation report).

Table 1: Proficiency chemicals, anastrozole, tamoxifen, atenolol and saccharin.

Chemical	CAS No.	Category	Concentrations to test	Expected statistical significance limit
Anastrozole	120511-73-1	Active	20, 4, 0.8, 0.16, 0.032 µg/L	4 µg/L
Tamoxifen	10540-29-1	Active	483, 242, 121, 60.4, 30.2 µg/L	483 µg/L
Atenolol	29122-68-7	Inert	100, 10, 1, 0.1, 0.01 mg/L	Inert
Saccharin	82385-42-0	Inert	100, 10, 1, 0.1, 0.01 mg/L	Inert

Validity of the test

22. For the test to be valid, the following criteria should be met for each run, and if they are not, the run is considered invalid:

- The mortality or overt sublethal toxicity such as immobilisation should not exceed two eleutheroembryos for the T control group.
- In all other non-optional control groups and in at least the four lowest test concentration groups in the presence and absence of T, mortality or overt sublethal toxicity such as immobilisation should not exceed one eleutheroembryo. Any groups other than the non-optional control groups and the four lowest test concentration groups not meeting these criteria are considered compromised and data from these groups should be excluded from the final analysis.
- Invalid data due to poorly positioned eleutheroembryos (see Annex 7) should not exceed two eleutheroembryos in the T control group.
- In all other non-optional control groups and in at least the four lowest test concentration groups in the presence and absence of T, invalid data due to poorly positioned eleutheroembryos should not exceed one eleutheroembryo. Any groups other than the non-optional control groups and the four lowest test concentration groups not meeting these criteria are considered compromised and data from these groups should be excluded from the final analysis.
- A statistically significant fluorescence induction for the EE2 488 ng/L and T 30 µg/L controls compared to the solvent control if one is present or the water control in the absence of a solvent control. The fluorescence value for the EE2 488 ng/L control should be at least 5-fold that of the relevant negative control. The fluorescence value for the T 30 µg/L control should be at least 2-fold the relevant negative control.

If one or more runs are invalidated, one or more additional runs can be performed in order to obtain three valid runs.

For the test to be valid, the following criteria should be met for the pool of the three runs, and if they are not, all three runs are considered invalid:

- A statistically significant fluorescence induction for the T plus EE2 control compared to the T control.

- A statistically significant fluorescence inhibition for the T and fadrozole control compared to the T control.
- If a minor deviation from the validity criteria is observed, the consequences should be considered in relation to the reliability of the test data and these considerations should be included in the report.

DESCRIPTION OF THE METHOD

Apparatus

23. Standard laboratory equipment and in particular the following:
    - laboratory incubator or any adequate apparatus for temperature and light control;
    - transparent cell culture grade 6-well plates made of a chemically inert material;
    - clear bottomed black 96-well plates certified for fluorescence quantification if eleutheroembryos are imaged from below or a black plastic surface suitable for fluorescence quantification if eleutheroembryos are imaged from above;
    - pH meter;
    - stereomicroscope equipped with a light source (for embryo and eleutheroembryo sorting);
    - fluorescent microscope equipped for fluorescence quantification with GFP long-pass filters and a colour camera (OECD, REACTIV assay validation report);
    - Image analysis software;
    - analytical instrumentation appropriate for the test chemical or contracted analytical services.

If plastic well plates are not appropriate for a given test chemical, alternative glass vessels (e.g., small diameter Petri dishes) should be used.

Test organism

24. The test organisms for the REACTIV assay are homozygous O. *latipes*, Japanese medaka eleutheroembryos of the *chgh-gfp* transgenic line. These organisms should be produced by mating two homozygous *chgh-gfp* Japanese medaka. The *chgh-gfp* transgenic line is maintained in several laboratories (Annex 10) and can be obtained upon subscribing to a license agreement. When a test chemical is shown to be fluorescent, wild type Japanese medaka eleutheroembryos could also be required to verify if the test chemical fluoresces within the eleutheroembryos (see §31).
25. The exposure phase of the test is initiated with DPH0 eleutheroembryos (approximately 10 days post fertilisation at 26°C or 7 days post fertilisation at 30°C). Although the eleutheroembryos must be DPH0, they can have a different number of DPF. The difference should not be more than one DPF in a single run. All eleutheroembryos should be randomly selected for the different test groups. Eleutheroembryos should either be: bred within the laboratory from stock animals; or eggs can be shipped from another laboratory (see Annex 10) and received as early as possible in development to allow for the longest possible

recovery period before beginning the test. Acclimation and batch acceptance criteria are outlined in Annex 4.

26. Housing, breeding and care of *O. latipes* are described in a number of sources, for example, Medaka: Biology, Management, and Experimental Protocols volumes 1 and 2 (Kinoshita *et al.*, 2009; Murata *et al.*, 2019) or the United States Environmental Protection Agency Guidelines for Culturing the Japanese Medaka, *Oryzias latipes* (Denny *et al.*, 1991).

27. The integrity of the *chgh-gfp* transgenic line should be verified every generation by running a full set of controls including the optional controls (§16) and ensuring that all validity criteria are met and that an expected response profile is obtained for the EE2 controls (§16). The transgene transmission and GFP response have been stable over more than 20 generations.

28. A quality control check on the developmental stage of randomly selected eleutheroembryos should be performed once a year to ensure that developmental stage of the eleutheroembryos at the end of the assay is not higher than stage 41.

Test medium

29. The test medium could be medaka medium (Annex 5), glass bottled still mineral water, spring water, well water and charcoal-filtered tap water. Because local water quality can differ substantially from one area to another, analysis of water quality should be undertaken to screen for potential contaminants (including heavy metals) and chemicals likely to interfere with the assay, particularly if historical data on the appropriateness of the water for raising *O. latipes* are not available. Special attention should be given to copper, chlorine and chloramine, all of which are toxic to *O. latipes* eleutheroembryos. Chelating agents should not be used. Results from analysis of water quality should be reported. Some chemical characteristics of an acceptable water suitable for *O. latipes* can be found in Annex 5. However, any medium that supports the normal growth and development of *O. latipes* and allows the test validity criteria to be met is suitable as a test medium.

Feeding

30. Eleutheroembryos between developmental stages DPH0 (beginning of the test) and DPH1 (end of the test) are used for this test. They are not fed before or during the test as the test is terminated at stage 40 (Iwamatsu, 2004). Yolk is still present until stage 41/42 and is used as the source of energy for the development of the eleutheroembryo.

Determining potential fluorescence of the test chemical

31. This test guideline should not be used for test chemicals emitting fluorescence between 500 and 550 nm (λ_{EM} = 500–550 nm) when excited at wavelengths between 450 and 500 nm (λ_{EX} = 450–500 nm) and able to fluoresce within the eleutheroembryos. Test chemicals sharing these two properties may induce a fluorescence which could be interpreted as GFP

signal, leading to the test chemical being incorrectly identified as active on the estrogen axis. A simple protocol to determine if the test chemical emits fluorescence at these wavelengths is to place 200 µL/well of a solution of the test chemical at the highest concentration intended to be tested in the REACTIV assay into ten wells of a 96-well plate. An additional ten wells of a 96-well plate should then be filled with 200 µL/well of test medium. The fluorescence should then be quantified using the same apparatus and settings as for the quantification of eleutheroembryo fluorescence. Potential differences in fluorescence between the test medium and the test chemical should be evaluated by statistical analysis. First, a D'Agostino-Pearson normality test should be performed. If the fluorescence data for both the test medium and test chemical follow a normal distribution, a two-tailed T-test should be performed to determine whether there is a statistically significant difference in fluorescence. If one or both sets of data do not follow a normal distribution, a Mann-Whitney test should be performed. If a fluorescent chemical is identified, 20 wild type *O. latipes* eleutheroembryos should be exposed at 26 ± 1°C for 24 ± 1 h with the highest concentration of the test chemical intended to be tested in the REACTIV assay. The fluorescence should then be quantified and compared to the fluorescence of a group of 20 wild type eleutheroembryos exposed to test medium only in the same conditions. Statistical analysis should be performed as detailed previously in this paragraph for comparing the test medium to the test chemical. If a statistically significant difference in fluorescence is present, the chemical is fluorescent and fluoresces within the eleutheroembryos and should not be tested using the REACTIV assay. In cases where the test chemical induces fluorescence in both unspiked and spiked modes in a REACTIV assay, then it cannot be excluded that it is metabolised into a fluorescent metabolite. In these cases, the images should be examined to identify whether the fluorescence is limited to the liver. If this is not the case, then the procedure described above for exposing wild-type eleutheroembryos should be performed to identify whether the chemical is metabolised into a fluorescent metabolite.

Selection of test concentrations

Establishing the maximum test concentration

32. The maximum tolerated concentration (MTC) is theoretically defined as the highest test concentration of the chemical which results in less than two mortalities or overt sublethal toxicity such as immobilisation in each of the three individual runs (less than two mortalities per group per run). The laboratory should perform a range-finding test with wild-type or preferably *chgh-gfp O. latipes* eleutheroembryos to evaluate possible toxicity.
33. The range-finding should consist of at least three test concentrations. They should be arranged in a geometric series with a separation factor not exceeding 10. Only one run is required with the chosen test concentrations and control. The range-finding test is performed with eight eleutheroembryos and 8 mL of exposure solution per well, with one well per test concentration and one well for the control. The percentage of eleutheroembryos exhibiting mortality or overt sublethal toxicity such as immobilisation is calculated from all eight eleutheroembryos exposed to the same test concentration or control. The highest concentration tested in the range-finding test must result in more than one case of mortality or overt sublethal toxicity such as immobilisation, unless the

highest tested concentration is 100 mg/L or the solubility limit of the test chemical. In order to be valid, no more than one mortality or overt sublethal toxicity such as immobilisation should occur in the control group of the range-finding test. One valid run is generally sufficient to determine the MTC.

34. The maximum test concentration should be set by the solubility limit of the test chemical in the test medium, the MTC, or a maximum concentration of 100 mg/L, whichever is lowest.

Test concentration range

35. There is a required minimum of five test concentrations. Generally, a concentration separation (spacing factor) of 3- to 10-fold between two adjacent test concentrations is recommended.

Test solutions

36. Test solutions of the chosen concentrations are usually prepared by dilution of a stock solution. The pH of each test solution should be adjusted to a pH comprised between 6.5 and 8.0. Stock solutions should be prepared by dissolving the test chemical using mechanical means if needed such as agitation, stirring or ultrasonication, or other appropriate methods. For difficult to test chemicals, the OECD Guidance Document No. 23 on aqueous-phase aquatic toxicity testing of difficult test chemicals should be consulted (OECD, 2019a).

37. It is possible to prepare the exposure solutions with no solvent or a maximum solvent concentration of 100 µL/L (0.01%) in line with OECD Guidance Document 23 (OECD, 2019a) if it is confirmed that the solvent and concentration of solvent used to dissolve the test item allow all validity criteria to be met. These validity criteria include eleutheroembryo survival but also the performance of the control groups (see §22). The test guideline was validated using dimethyl sulfoxide (DMSO) exclusively, at a final concentration of 0.2%, without the generation of false positive results. Therefore, the test can be conducted with 0.01% DMSO in line with OECD Guidance Document 23 (OECD, 2019a) as long as the validity criteria are fulfilled.

38. If a solvent is used, the concentration of solvent should be equal in all test concentrations and in all controls. The selection of an appropriate solvent depends on the physico-chemical properties of the test chemical and on the sensitivity of *O. latipes*, which should preferably be determined in a previous study to determine the maximum concentration of solvent showing an absence of mortality or overt sublethal toxicity such as immobilisation and an absence of endocrine activity. Possible actions of the solvent on the reproductive axis should also be considered (Hutchinson *et al.*, 2006).

39. Control solutions should be prepared on the first day of a run. The same preparation of control or test solutions should not be used across independent runs. Solutions that have been stored at 4°C should be allowed to reach 26 ± 1°C before being placed in contact with the eleutheroembryos to prevent thermal shock.

PROCEDURE

Exposure conditions

40. The organisms are exposed in chemically inert plastic cell culture grade 6-well plates (typically wells of 34 mm internal diameter and 20 mm height). Each well should contain eight organisms in 8 mL of exposure or control solution (see §16 for the list of control groups).
41. Eleutheroembryos are maintained in an incubator for 24 ± 1 h at 26 ± 1°C with a 14:10 light: dark cycle.
42. A new set of exposure solutions should be prepared for each of the three runs of the REACTIV assay.

Analytical measurements

43. As a static 24 h exposure method is used, the stability of the test chemical concentration should be documented. The stability of the test chemical should ideally allow the exposure concentration to remain within ± 20% of the nominal concentration in a 24 h time frame. The minimum requirement for analytical measures is the minimum scientifically justifiable set of samples as determined by the needs of the regulatory authority. OECD Guidance Document No. 23 provides guidance on issue (OECD, 2019a). If concentrations cannot be maintained within ± 20% in the test system, renewal of exposure solutions could be considered. The use of the geometric mean of measured concentrations is allowed for chemicals that do not remain within 80-120% of the nominal concentration; see Chapter 5 in the OECD Guidance Document No. 23 for more details (OECD, 2019a).

Test initiation and conduct

Day 0 Test initiation

44. The exposure should be initiated on the day that the eleutheroembryos hatch (DPH0).
45. For selection of test organisms, eleutheroembryos should be observed and those exhibiting grossly visible malformations or physical injury (e.g., damage of the tail, oedema, scoliosis) should be excluded from the assay (Annex 6). Healthy and normal looking eleutheroembryos of the stock population should be pooled in a single vessel containing an appropriate volume of test medium. The selected organisms should be homogenous in size, eleutheroembryos presenting a visually obvious difference in size should be removed. Batches of eleutheroembryos that contain less than 80% of normal and healthy eleutheroembryos at DPH0 (not including any dead or unfertilised eggs that were removed after egg collection) should not be used for the test. This should be determined whilst removing dead and malformed eleutheroembryos from the batch prior to performing the assay.
46. To start the experiment, eight eleutheroembryos should be randomly selected and placed into each well of a 6-well plate or glass vessel in drops of test medium (see §29) using a

transfer pipet. Excess test medium should be removed and the test chemical solutions added for the first time. One should pay attention to work with one plate at a time to avoid drying out the eleutheroembryos.

Day 1 Fluorescence quantification

47. The fluorescence of each organism is quantified after 24 ± 1 h of exposure. Immediately prior to this, dead organisms should be removed and the exposure medium should be replaced with test medium (see §29). This is to prevent the person reading the fluorescence from being exposed to the test chemical. All observations should be recorded. If more than one mortality or overt sublethal toxicity such as immobilisation is encountered in one of the non-optional control groups or in one or more of the four lowest concentration treatment groups, then the on-going independent run is considered compromised and should be terminated. The data of compromised groups should not be considered for analysis. If the eleutheroembryos are required to be anaesthetised for imaging, they should be anesthetised by adding 2 mL of 1 g/L buffered MS222 (tricaine methylsulfonate) into the wells of the six-well plates. Anaesthesia is recommended in all cases where the eleutheroembryos are placed in a drop of liquid for imaging. It is only not recommended if they are imaged whilst swimming freely, such as in a well of a 96-well clear-bottomed plate. To avoid excessive anaesthesia, only the number of organisms that can be read in one series should be anaesthetised. After the onset of anaesthesia (1 to 5 min) if required, the eleutheroembryos are transferred to the support to be used for imaging such as a black plastic surface for imaging from above or clear-bottomed 96-well plates for imaging from below. They are then imaged with a colour camera and GFP long pass filters. An image of the ventral region including the liver of each organism should be captured using the parameters identified during the calibration (see Annex 7 for examples of the expected positioning of the eleutheroembryos for imaging).

Terminating the experiment

48. After reading the fluorescence, each eleutheroembryo is euthanised by exposing it to 1 g/L of buffered MS222 for at least 20 min.

Analysis of data / Evaluation of test results

Data analysis considerations

49. Fluorescence measurements from images of poorly positioned eleutheroembryos (see Annex 7) should be removed from the data before analysis.
50. Treatment of the colour images of the eleutheroembryos to extract a numerical value for GFP fluorescence should be performed using appropriate software. An open-source option is ImageJ or the more recent version Fiji (Schindelin *et al.*, 2012). In order to exclude autofluorescence (non-GFP endogenous fluorescence of the eleutheroembryos) from the images it is recommended to separate the red, green and blue colour layers of the images.

The red layer can then be subtracted from the green layer or the values of the red layer can be doubled and subtracted from the green layer. An intensity threshold can then be applied to the resulting image to reduce background caused by endogenous pigmentation. The sum of the fluorescence of all pixels in the resulting image should then be quantified. This technique is an efficient way to restrict the measurement to GFP and not endogenous (auto-) fluorescence. As GFP-related fluorescence will only appear in the green layer, but yellow fluorescence will appear in both the green and red layer. Doubling the red layer is useful depending on the imaging system if some endogenous fluorescence remains after subtracting the undoubled red layer. Other techniques to reduce the impact of endogenous pigmentation on the quantification of GFP signal can be applied depending on the imaging system and fluorescence filters used. Once an image analysis workflow has been demonstrated to allow validation criteria to be met for a given fluorescence imagery system, it should be applied for all future experiments (see §22 and Annex 3).

51. Data from the three independent runs are pooled to obtain 18 to 24 fluorescence values for each valid test concentration and control (36 to 48 for the T control). The maximum number of values is 24 (48 for the T control) as each test condition or control is made up of eight eleutheroembryos per run (16 for the T control) and the REACTIV assay consists of three runs. The lower threshold of 18 values (36 for the T control) represents the limit of one mortality or overt sublethal toxicity such as immobilisation in each run and one poorly positioned eleutheroembryo per run, therefore, six values per run (12 for the T control).

52. Three independent runs are performed to increase robustness of the assay. Only the pooled data are considered when evaluating the test chemical as active or inert.

53. If a solvent is used in the experiment, an evaluation of the potential effects of the solvent should be performed. This is done through a statistical comparison of the solvent control group and the test medium control group. If a statistically significant difference is identified between the test medium control and the solvent control for the pool of the three runs, then consideration should be made as to whether the solvent interfered with the integrity of the test and whether the results meet the purposes for which the data are intended. It is important to verify that all validity criteria are met with the chosen solvent (§22, §37). If historical data exist indicating that the chosen solvent, at the chosen concentration, does not elicit a statistically significant difference when compared to the test medium control, then the test medium control may not be required.

Statistical analysis

54. Appropriate statistical methods should be used according to OECD Document 54 on the Current Approaches in the Statistical Analysis of Ecotoxicity Data: A Guidance to Application (OECD, 2006). In general, effects on the fluorescence of the test chemical compared to the control are investigated using two-tailed hypothesis testing at $p < 0.01$.

55. The recommended statistical approach, which was evaluated during the interlaboratory validation exercise, is to determine whether the data for each exposure group is normally distributed by performing a D'Agostino-Pearson normality test, then performing either an ANOVA test followed by a Dunnett's test if the data are normally distributed with equal variances or a Kruskal-Wallis test followed by a Dunn's test if the data do not follow a normal distribution or if the homogeneous variance assumption is violated (see Annex 8 for a more detailed description). Alternatively, a mixed ANOVA (also referred to as nested ANOVA) approach can be carried out. This approach is described in detail in Annex 8. In

contrast to the approach mentioned earlier, the mixed ANOVA does account for the variability between the runs and the interaction of run and treatment. This is an advantage because it leads to a more accurate testing by regarding the dependency structure of the data.

Decision logic

56. In unspiked mode, an active concentration is defined as a concentration giving a statistically significant increase in fluorescence compared to the test medium control/ solvent control (see §53).
57. In T-spiked mode, an active concentration is defined as a concentration giving a statistically significant increase or decrease in fluorescence compared to the 30 µg/L T control.
58. A decision logic flowchart was developed for the REACTIV assay to provide assistance in the conduct and interpretation of the results of the assay (Figure 2). This decision logic is based on three valid runs pooled for statistical analysis (see Figure 1 and §15). A test chemical is considered to give a positive result in the REACTIV assay if at least one tested concentration is active in either unspiked or T-spiked mode and a monotonous concentration-response relationship is observed (i.e., this is the highest tested concentration). A test chemical is also considered to be active if at least two tested concentrations are active in either unspiked or T-spiked mode if a non-monotonic concentration-response relationship is observed, provided that at least two adjacent concentrations are active. In unspiked mode, at least two adjacent concentrations must show a statistically significant increase in fluorescence. In T-spiked mode, at least two adjacent active concentrations must both show a statistically significant increase in fluorescence or they must both show a statistically significant decrease in fluorescence.

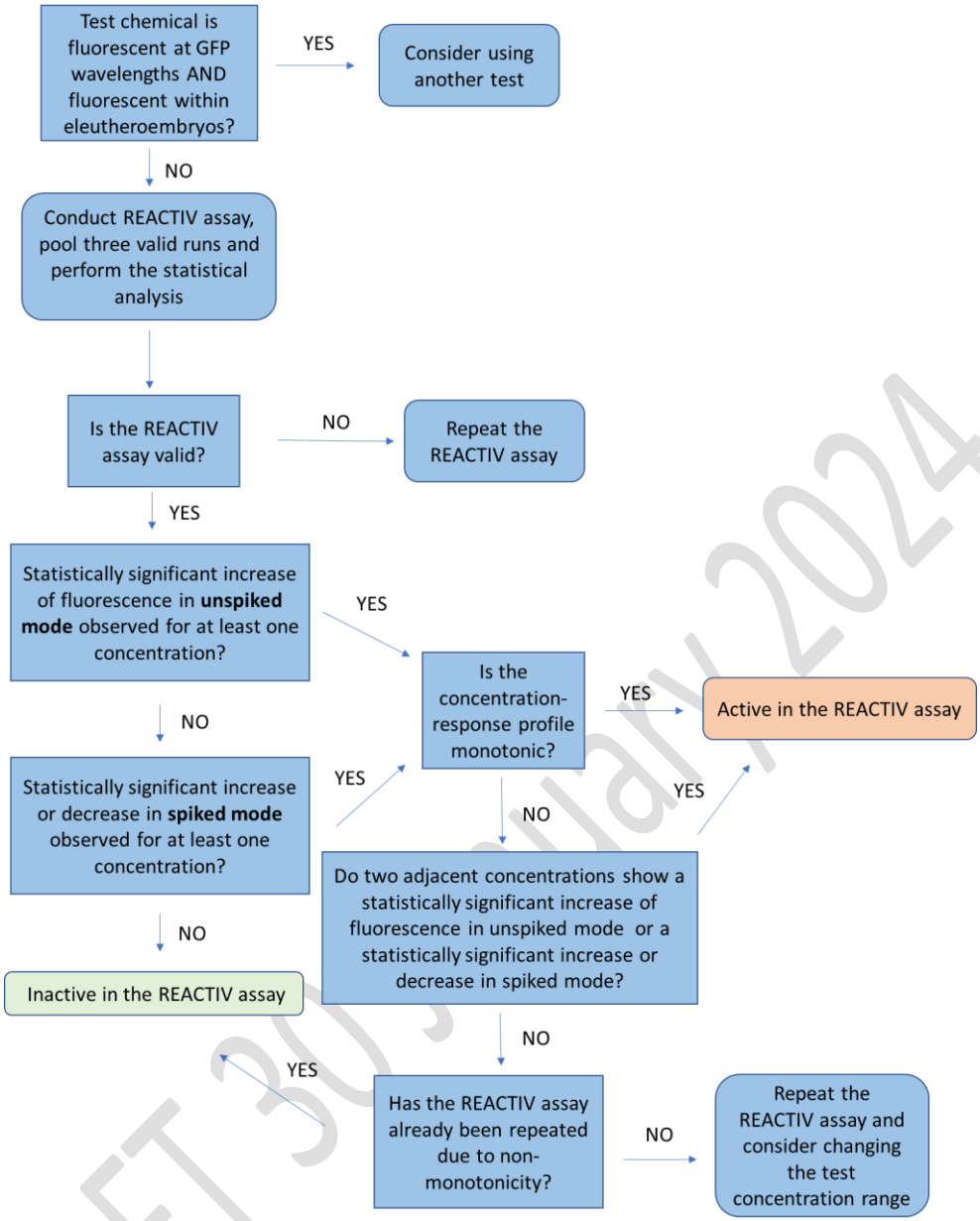

Figure 2: Decision logic for the interpretation of the result of the REACTIV assay.

Test report

59. The test report should include the following information:

Test chemical

- Mono-constituent substance: physical appearance, water solubility, and additional relevant physico-chemical properties; chemical identification, such as IUPAC or CAS name, CAS number, SMILES or InChI code, structural formula, purity, chemical identity of impurities as appropriate and practically feasible, etc. (including the organic carbon content, if appropriate). Also, if available, stability in light, stability under the test conditions, pKa, Kow, information on the fate of the test chemical and its potential for being rapidly degraded in the test system e.g., results of a biodegradability test, see OECD TG 301 (OECD, 1992) and TG 310 (OECD, 2014b).
- Multi-constituent substance, UVCBs and mixtures: characterised as far as possible by chemical identity (see above), quantitative occurrence and relevant physico-chemical properties of the constituents.
- Analytical method for quantification of the test chemical, including quantification limit.
- Available data or results from any preliminary studies on the stability or solubility of the test chemical.
- Results of any tests performed to determine potential fluorescence of the test chemical.

Test species

- Scientific name, transgenic line, supplier or source, and culture conditions.
- The percentage of dead and malformed eleutheroembryos removed from the batch immediately prior to performing the assay.

Test conditions

- Test procedure used (e.g., concentrations tested, temperature, duration, static exposure, volume, number of organisms per mL).
- Details of test medium characteristics (reference of mineral water or spring water, description of tap water treatment (e.g., charcoal filtration…) or artificial test medium used and any measurements made.
- Method of preparation of stock solutions and frequency of renewal if performed (the solvent and its concentration should be given, when used).
- Brand and references of 6-well plates used for exposure and any plates used for fluorescence quantification.
- References and settings of the fluorescence microscope used for quantification. The method used for image analysis should also be provided.

Results

- Results of the range-finding test(s) that allow the determination of the MTC and/or the selection of the test concentrations for the definitive test.
- The nominal test concentrations and, where possible, results of all chemical analyses to determine the concentration of the test chemical in the test vessels;

the measured exposure concentration as an appropriate statistical average (e.g., arithmetic mean, time-weighted mean etc.) where appropriate; the recovery efficiency of the analytical method and the limit of quantification should also be reported.

- The numbers of dead organisms in each run and the group(s) and days on which they occurred.
- Fluorescence quantification raw data (e.g., individual fluorescence raw data). Ideally, data should be collected in tab or comma separated format with the following metadata present in the file: date; chemical name; concentration used; solvent; machine name; signal collection parameters for the machine, laboratory name, eleutheroembryo batch number and fluorescence values.
- Approach for the statistical analysis and treatment of data including statistical test used and whether and why any data censuring was conducted.
- Demonstration that all validity criteria of the guideline were met.
- The means of fluorescence of each experimental group including all control and test chemical concentrations and their SEM (standard error of the mean) should be presented both by a graphical representation and also in a table together with the sample size.
- The percentage increase or decrease of fluorescence for each concentration compared to its respective control in spiked and unspiked modes.
- Optionally and where appropriate, results of the evaluation of the potential effects of the solvent: a statistical comparison of the solvent control group and the test medium control group if included in the present study or a result from a previous study.
- Other observed biological effects or measurements: report any other biological effects which were observed or measured (e.g., abnormal behaviour, malformations or abnormal pigmentation).
- An explanation for any deviation from the test guideline or deviation from the validity criteria, and considerations of potential consequences on the outcome of the test.
- Where appropriate, a discussion presenting the concentrations found active in spiked and/or unspiked mode.
- A conclusion presenting whether the test chemical is found to be active or inactive on the estrogen axis in the REACTIV assay.

REFERENCES

Ankley, G.T. *et al.* (2002) Evaluation of the aromatase inhibitor fadrozole in a short-term reproduction assay with the fathead minnow (*Pimephales promelas*). *Toxicol. Sci.*, **67**, 121–130.

Bates, D. *et al.* (2015) Fitting Linear Mixed-Effects Models Using lme4. J. Stat. Softw., **67**, 1–48.

Denny, J.R. *et al.* (1991) Guidelines for the culture of the medaka, *Oryzias latipes*. U.S. Environmental Protection Agency, Washington, D.C., EPA/600/3-91/064 (NTIS PB92137496).

Green, J. W. *et al.* (2018). Statistical Analysis of Ecotoxicity Studies. John Wiley & Sons, Ltd.

Hornung, M.W. *et al.* (2007) Tissue distribution and metabolism of benzo[a]pyrene in embryonic and larval medaka (*Oryzias latipes*). *Toxicol. Sci.*, **100**, 393–405.

Hutchinson, T.H. *et al.* (2006) Acute and chronic effects of carrier solvents in aquatic organisms: a critical review. *Aquat. Toxicol.*, **76**, 69–92.

Iwamatsu, T. (2004) Stages of normal development in the medaka *Oryzias latipes*. *Mech. Dev.*, **121**, 605–618.

Kashiwada, S. *et al.* (2007) Age-dependent *in situ* hepatic and gill CYP1A activity in the see-through medaka (*Oryzias latipes*). *Comp. Biochem. Physiol. C Toxicol. Pharmacol.*, **145**, 96–102.

Kinoshita, M. *et al.* (2009) Medaka: Biology, Management, and Experimental Protocols John Wiley & Sons.

Kullman, S.W. and Hinton, D.E. (2001) Identification, characterization, and ontogeny of a second cytochrome P450 3A gene from the fresh water teleost medaka (*Oryzias latipes*). *Mol. Reprod. Dev.*, **58**, 149–158.

Kurauchi, K. *et al.* (2008) Characteristics of ChgH-GFP transgenic medaka lines, an in vivo estrogenic compound detection system. *Mar. Pollut. Bull.*, **57**, 441–444.

Kurauchi, K. *et al.* (2005) *In vivo* visual reporter system for detection of estrogen-like substances by transgenic medaka. *Environ. Sci. Technol.*, **39**, 2762–2768.

Liu et al. (2022) Medaka embryos as a model for metabolism of anabolic steroids. *Arch Toxicol.*, **96**(7), 1963-1974.

Murata, K. *et al.* (2019) Medaka: Biology, Management, and Experimental Protocols John Wiley & Sons.

OECD (1992). Test No. 301: Ready Biodegradability, OECD Guidelines for the Testing of Chemicals, Section 3, OECD Publishing, Paris, https://doi.org/10.1787/9789264070349-en.

OECD (1998) Test No. 212: Fish, Short-term Toxicity Test on Embryo and Sac-Fry Stages, OECD Guidelines for the Testing of Chemicals, Section 2, OECD Publishing, Paris, https://doi.org/10.1787/9789264070141-en.

OECD (2006) Current Approaches in the Statistical Analysis of Ecotoxicity Data: A guidance to application (annexes to this publication exist as a separate document), OECD Series on Testing and Assessment, No. 54, OECD Publishing, Paris, https://doi.org/10.1787/9789264085275-en.

OECD (2009) Test No. 230: 21-day Fish Assay: A Short-Term Screening for Oestrogenic and Androgenic Activity, and Aromatase Inhibition, OECD Guidelines for the Testing of Chemicals, Section 2, OECD Publishing, Paris, https://doi.org/10.1787/9789264076228-en.

OECD (2011) Test No. 234: Fish Sexual Development Test, OECD Guidelines for the Testing of Chemicals, Section 2, OECD Publishing, Paris, https://doi.org/10.1787/9789264122369-en.

OECD (2012) Test No. 229: Fish Short Term Reproduction Assay, OECD Guidelines for the Testing of Chemicals, Section 2, OECD Publishing, Paris, https://doi.org/10.1787/9789264185265-en.

OECD (2013) Test No. 210: Fish, Early-life Stage Toxicity Test, OECD Guidelines for the Testing of Chemicals, Section 2, OECD Publishing, Paris, https://doi.org/10.1787/9789264203785-en.

OECD (2014a) Guidance document No. 204 for single laboratory validation of quantitative analytical methods – Guidance used in support of pre-and-post-registration data requirements for plant protection and biocidal products, OECD Series on Testing and Assessment No. 204, series on Biocides No. 9.

OECD (2014b) Test No. 310: Ready Biodegradability - CO2 in sealed vessels (Headspace Test), OECD Guidelines for the Testing of Chemicals, Section 3, OECD Publishing, Paris, https://doi.org/10.1787/9789264224506-en.

OECD (2014c) Fish Toxicity Testing Framework, OECD Series on Testing and Assessment, No. 177, OECD Publishing, Paris, https://doi.org/10.1787/9789264221437-en.

OECD (2018) Revised Guidance Document 150 on Standardised Test Guidelines for Evaluating Chemicals for Endocrine Disruption, OECD Series on Testing and Assessment, No. 150, OECD Publishing, Paris, https://doi.org/10.1787/9789264304741-en.

OECD (2019a) Guidance Document on Aquatic Toxicity Testing of Difficult Substances and Mixtures, OECD Series on Testing and Assessment, OECD Publishing, Paris, https://doi.org/10.1787/0ed2f88e-en.

OECD (2019b) Test No. 203: Fish, Acute Toxicity Test, OECD Guidelines for the Testing of Chemicals, Section 2, OECD Publishing, Paris, https://doi.org/10.1787/9789264069961-en.

OECD (2019c) Validation Report of the Xenopus Eleutheroembryonic Thyroid Signaling Assay (XETA) for the Detection of Thyroid Active Substances, Series on Testing and Assessment No. 302.

OECD (2019d) Test No. 248: Xenopus Eleutheroembryonic Thyroid Assay (XETA), OECD Guidelines for the Testing of Chemicals, Section 2, OECD Publishing, Paris, https://doi.org/10.1787/a13f80ee-en.

OECD (2022) Test No. 251: Rapid Androgen Disruption Activity Reporter (RADAR) assay, OECD Guidelines for the Testing of Chemicals, Section 2, OECD Publishing, Paris, https://doi.org/10.1787/da264d82-en.OECD (2023) Test No. 240: Medaka Extended One Generation Reproduction Test (MEOGRT), OECD Guidelines for the Testing of Chemicals, Section 2, OECD Publishing, Paris, https://doi.org/10.1787/9789264242258-en.

OECD REACTIV assay validation report: To be determined.

Schiller et al. (2014) Species-specific considerations in using the fish embryo test as an alternative to identify endocrine disruption. *Aquat Toxicol.*, **155**, 62-72.

Schindelin, J. *et al.* (2012) Fiji: An open-source platform for biological-image analysis. *Nat. Methods*, **9**, 676–682.

Schulz-Jander *et al.* (2002a) Neonicotinoid insecticides: reduction and cleavage of imidacloprid nitroimine substituent by liver microsomal and cytosolic enzymes. *Chem Res Toxicol.,* **15**(9), 1158-65.

Schulz-Jander *et al.* (2002b) Imidacloprid insecticide metabolism: human cytochrome P450 isozymes differ in selectivity for imidazolidine oxidation versus nitroimine reduction. *Toxicol Lett.*, **132**(1), 65-70.

Sébillot, A. *et al.* (2014) Rapid fluorescent detection of (anti)androgens with *spiggin-gfp* medaka. *Environ. Sci. Technol.*, **48**, 10919–10928.

Seki, M. *et al.* (2002) Effect of ethinylestradiol on the reproduction and induction of vitellogenin and testis-ova in medaka (*Oryzias latipes*). *Environ. Toxicol. Chem.*, **21**, 1692–1698.

Spirhanzlova, P. *et al.* (2016) Oestrogen reporter transgenic medaka for non-invasive evaluation of aromatase activity. *Comp. Biochem. Physiol. C Toxicol. Pharmacol.*, **179**, 64–71.

Vignet *et al.* (2019) Imidacloprid induces adverse effects on fish early life stages that are more severe in Japanese medaka (Oryzias latipes) than in zebrafish (Danio rerio). *Toxicol Lett.* **7**;132(1):65-70.

ANNEX 1: ABREVIATIONS AND DEFINITIONS

***chgh-gfp*:** Transgenic medaka line harbouring a genetic construction consisting of 2.047 kb of the medaka choriogenin H gene promoter upstream of GFP coding sequence.

DMSO: Dimethyl Sulfoxide.

DPF: Day Post Fertilisation.

DPH: Day Post Hatch.

EE2: 17α-Ethinylestradiol, a synthetic estrogen receptor agonist.

Eleutheroembryo: The eleutheroembryonic life stage is post-hatch, but before the embryo is capable of independently feeding on exogenous food supplies and is a stage of on-going embryonic development. In some regulatory jurisdictions, the eleutheroembryonic period is regarded as a non-protected life stage in this context (OECD, 2014c). Applying this definition to *O. latipes* positions this period of development from stage 39 (hatching stage) to stage 42 (formation of structures required for prey capture including the teeth of the upper jaw, the otolith, and the shape of all fins) (Iwamatsu, 2004).

Estrogen axis: In this context, refers to downstream steroidogenesis and estrogen receptor activation/antagonism. No data is currently available on the responsiveness of the REACTIV assay to modulators of upstream steroidogenesis.

Fad: Fadrozole, a pharmaceutical aromatase inhibitor.

GFP: Green Fluorescent Protein.

LCX: Median Lethal Concentration is the concentration of a test chemical that is estimated to be lethal to X% of the test organisms within the test duration.

LOEC: The Lowest Observed Effect Concentration is the lowest tested concentration at which the test chemical is observed to have a statistically significant effect (at $p < 0.05$).

MS-222: Tricaine methanesulfonate; CAS: 886-86-2.

MTC: Maximum tolerated concentration. MTC is defined as the highest test concentration of the chemical which results in less than two eleutheroembryos displaying mortality or overt sublethal toxicity such as immobilisation.

NOEC: The No Observed Effect Concentration is the tested concentration immediately below the LOEC.

Run: A run is defined here as an experiment performed using independent solutions.

SEM: Standard Error of the Mean.

SMILES: Simplified Molecular Input Line Entry Specification.

Spiked mode: Part of a REACTIV assay run in the presence of 30 µg/l of T.

T: Testosterone.

Transgenic organism: Organism that contains novel genetic material, e.g. originally derived from different species or synthetic, that has been inserted into the genome using recombinant DNA techniques.

Unspiked mode: Part of a REACTIV assay run in the absence of T.

UVCB: Substances of unknown or variable composition, complex reaction products or biological materials.

ANNEX 2: OVERVIEW OF TEST CONDITIONS OF THE REACTIV ASSAY

Table 2: Overview of the test conditions for the REACTIV assay.

Test animal	*chgh-gfp O. latipes* eleutheroembryo
Endpoint	Fluorescence of individual eleutheroembryos
Exposure period	DPH0 (beginning of the test) to DPH1 (end of the test)
Exposure duration	24 ± 1 h
Exposure regime	Static renewal. No feeding
pH	6.5 to 8
Incubation conditions during exposure	26 ± 1°C, 14:10 light:dark cycle
Eleutheroembryos per test condition and control group	8 organisms per well (6-well plate) x 1 wells (total of 8 organisms per concentration and run) except the testosterone control which comprises of two wells with 8 organisms per well.
Volume of test medium	8 mL per well
Test medium	Water permitting normal growth and development of *O. latipes* (refer to §29)
Number of experiments	Experiments are run 3 times for each chemical with freshly prepared solutions.
Criteria for selecting test individuals	Developmental stage (DPH0), health of organisms (alive and no malformations).
Validity criteria	For each run: Mortality or overt sublethal toxicity such as immobilisation of ≤ 1 eleutheroembryo in all non-optional control groups and at least the four lowest test concentration groups in the presence and absence of T (≤ 2 for the T control). Invalid data due to poorly positioned eleutheroembryos ≤ 1 eleutheroembryo in all non-optional control groups and at least the lowest four test concentration groups in the presence and absence of T (≤ 2 for the T control). For the pool of three runs: a statistically significant fluorescence induction for the EE2 488 ng/L and T controls compared to the relevant solvent or water control. The mean fluorescence value should be at least 5-fold higher than that of the relevant negative

		control for the EE2 488 ng/L control and at least 2-fold for the T control; a statistically significant fluorescence induction for the T plus EE2 control compared to the testosterone control and a statistically significant fluorescence inhibition for the T and fadrozole control compared to the T control. At least four uncompromised test concentrations. These should include the lowest four test concentrations. A test concentration is considered uncompromised for the purpose of the test when this test concentration is considered uncompromised in each of the three runs of the test. A test concentration (8 individuals) is considered uncompromised in a run when mortality or overt sublethal toxicity such as immobilisation in the group is ≤ 1 eleutheroembryo (≤ 2 for the T control) and invalid data due to poorly positioned eleutheroembryos ≤ 1 eleutheroembryo (≤ 2 for the T control).
Test chemical concentration standard		If the test chemical concentration remains within 20% of nominal at all time points, the nominal concentration is used. Otherwise, the result should be considered using the determined concentrations. For instance, geometric means of each set of new/old concentrations could be calculated. The arithmetic mean of these geometric means should then be used for data interpretation.
Controls	Test medium and/ or solvent control	Test medium and/or test medium plus solvent
	17α-Ethinylestradiol (EE2)	EE2 (488 ng/L)
	Testosterone (T)	T 30 µg/L (2 wells of 8 eleutheroembryos)
	T + EE2	T (30 µg/L) + EE2 (64 ng/L)
	T + Fadrozole	T (30 µg/L) + Fadrozole (10 µg/L)

886

ANNEX 3: CALIBRATION: DETERMINATION OF THE OPTIMAL IMAGING SETTINGS

The goal of the calibration step is to ensure that the imaging equipment is working to the correct parameters for the REACTIV assay. The calibration requires two steps:

1) Determining the optimal imaging settings to allow a satisfactory amplitude of GFP induction to be obtained with a concentration of 488 ng/L of EE2.

2) Applying these settings for the quantitation of three runs of a concentration-response experiment with six concentrations of EE2 as well as the other assay controls (T, T + EE2 and T + fadrozole) to check the amplitude of induction and sensitivity with increasing concentrations of T and to ensure that the other assay controls elicit a detectable GFP response.

The example protocol, described in two steps below, involves the use of 0.2% DMSO in all exposure solutions. This is an example; the same procedure can be performed with an alternative solvent or alternative concentration of solvent. The calibration procedure does not need to be repeated if the solvent is changed when performing a REACTIV assay or if the assay is performed for the first time without a solvent.

1- Selecting image capture settings

The first step is to determine the correct image capture settings for the calibration experiment. In order to select the image capture settings, expose 40 eleutheroembryos to EE2 at 488 µg/L and adjust the settings as indicated in the following protocol. A single replicate experiment is required for this step.

- Setting up the exposure media
    - The test group consists of 5 wells, with each well containing 8 eleutheroembryos of the *chgh-gfp* line.
    - The final concentration of DMSO is 0.2% in all wells.
    - Prepare a solution of 488 µg/L EE2 in DMSO.
        - Aliquot the solution of 488 µg/L EE2 with 200 µL per aliquot.
        - Conserve the aliquots at -20°C for a maximum of 6 months.
    - Prepare the following exposure solution of 488 ng/L EE2 containing 0.2% DMSO.

Test Medium	49.9 mL
EE2 488 µg/L in DMSO	50 µL
DMSO	50 µL

- Starting the exposure
    - Add 8 *chgh-gfp* transgenic eleutheroembryos to each well.
    - Remove the maximum amount of liquid without drying the eleutheroembryos (maximum remaining volume 800 µL).
    - Fill each well with 8 mL of the exposure solution.
    - Incubate the plates at 26 °C in a 14:10 light:dark cycle. Do not feed the eleutheroembryos during the experiment.

- Rinsing eleutheroembryos at 24 h
    - Prepare 6-well rinsing plates containing 8 mL of water permitting normal growth and development of *O. latipes* (refer to §29) in each well.
    - Transfer all eleutheroembryos from an exposure group from their treatment plate to the rinsing plate.

- Reading eleutheroembryos at 24 h
    - If necessary, anesthetise the eleutheroembryos exposed to 488 ng/L of EE2 by placing 2 mL of MS222 at 1 g/L in each well of the 6-well plates. Be careful to anesthetise only 1 plate at a time.
    - Place the eleutheroembryos so that the ventral side can be imaged by the imaging system.
    - Adjust the zoom and focus on the fluorescence microscope to determine the maximal zoom that allows imaging of the entire liver.
    - Check the other eleutheroembryos on the plate to ensure that the selected zoom allows the entire liver to be visualised in a single image. If this is not the case readjust the zoom and begin the process again.
    - If possible, reset the white balance of the camera.
    - Set the gain on the camera settings to zero and adjust the exposure time to the point where the liver is as bright as possible without appearing white.
    - If the exposure needs to be set above 100 ms to result in saturation of the GFP signal (white areas in the GFP signal), increase the gain and restart.
    - Check the other eleutheroembryos on the plate to ensure that the selected exposure time does not result in a significant portion of the liver to be white. If this is not the case adjust the exposure time and begin the process again.
    - Save and note the selected settings for the camera and conserve the settings file to be recalled at each future imaging session.
    - Capture an image of each eleutheroembryo.

- After all images are taken, euthanise the eleutheroembryos.
- Analyse the images by following the instructions in §49 to §55.
- Example images of eleutheroembryos after exposure to an estrogen (ventral view) are shown below (Annex 7).

2- Determining linearity and sensitivity to EE2

The second step is to determine the linearity and sensitivity to EE2. In order to perform this step, groups of 8 eleutheroembryos are exposed to a concentration range of EE2. Three independent runs are required for this step.

- Setting up the exposure media
  - Each test group consists of 1 well, with each well containing 8 eleutheroembryos of the *chgh-gfp* line.
  - The final concentration of DMSO is 0.2% in all wells.
  - Prepare a solution of 488 µg/L EE2 in DMSO.
    - Aliquot the solution of 488 µg/L EE2 with 200 µL per aliquot.
    - Conserve the aliquots at -20°C for a maximum of 6 months.
  - Prepare a stock solution of 30 mg/L T in DMSO.
    - Aliquot the solution of 30 mg/L T with 300 µL per aliquot.
    - Conserve the aliquots at -20°C for a maximum of 3 months.
  - Prepare a stock solution of 10 mg/L fadrozole in DMSO.
    - Aliquot the solution of 30 mg/L T with 300 µL per aliquot.
    - Conserve the aliquots at -20°C for a maximum of 12 months.
  - Prepare the test solutions according to Table 3.

The test groups are:

Solvent control: test medium + 0.2% DMSO

34 ng/L EE2 0.2% DMSO

51 ng/L EE2 0.2% DMSO

76 ng/L EE2 0.2% DMSO

114 ng/L EE2 0.2% DMSO

171 ng/L EE2 0.2% DMSO

488 ng/L EE2 0.2% DMSO

30 µg/L T 0.2% DMSO

30 µg/L T + 64 ng/L EE2 0.2% DMSO

30 µg/L T + 10 µg/L fadrozole 0.2% DMSO

Table 3: Preparation of test solutions and intermediate solutions (grey background).

Solution Name	Intermediary volume to prepare (mL)	Solutions to mix	Final Volume (mL)
Test medium 0.1% DMSO	120	119.88 mL of test medium + 120 µL of DMSO	20
Solvent control	70	69.93 mL of test medium 0.1% DMSO + 70 µL of DMSO	12
T 30 µg/L 0.1% DMSO	50	49.95 mL of test medium + 50 µL of T 30 mg/L	10
T 30 µg/L 0.2% DMSO	20	19.98 mL of T 30 µg/L 0.1% DMSO + 20 µL of DMSO	20
EE2 488 ng/L	30	29.07 mL of test medium 0.1% DMSO + 30 µL of EE2 488 µg/L	16.1
EE2 171 ng/L	17	5.96 mL of EE2 488 ng/L + 11.04 mL of solvent control	17
EE2 114 ng/L	17	3.97 mL of EE2 488 ng/L + 13.03 mL of solvent control	17
EE2 76 ng/L	13	1.87 mL of EE2 488 ng/L + 11.13 mL of solvent control	13
EE2 51 ng/L	10	1.05 mL of EE2 488 ng/L + 8.95 mL of solvent control	10
EE2 34ng/L	15	1.05 mL of EE2 488 ng/L + 13.95 mL of solvent control	15
T 30 µg/L + EE2 64 ng/L	10	9.99 mL of T 30 µg/L 0.1% DMSO + 10 µL of EE2 64 µg/L	10
T 30 µg/L + fadrozole 10 µg/L	10	9.99 mL of T 30 µg/L 0.1% DMSO + 10 µL of fadrozole 10 mg/L	10

- Starting the exposure
  - Add 8 *chgh-gfp* transgenic eleutheroembryos to each well.
  - Remove the maximum amount of liquid without drying the eleutheroembryos (maximum remaining volume 800 µL).
  - Proceed with the treatment of the solvent control, then the EE2 groups and then the following controls in order: T, T + EE2 and T + fadrozole.
  - Fill each well with 8 mL of each preparation.
  - Incubate the plates at 26 °C in a 14: 10 light: dark cycle. Do not feed the eleutheroembryos during the experiment.
  -
- Rinsing eleutheroembryos at 24 h
  - Prepare 6-well rinsing plates containing 8 ml of water permitting normal growth and development of *O. latipes* (refer to §29) in each well.
  - Transfer all eleutheroembryos from an exposure group from their treatment plate to the rinsing plate.
  -
- Reading eleutheroembryos at 24 h
  - Load the image capture parameters that were saved at the end of the first step of the calibration experiment.
  - If necessary, anesthetise the eleutheroembryos exposed to the solvent control solution by placing 2 ml of MS222 at 1 g/L in each well of the 6-well plates. Be careful to anesthetise only 1 plate at a time.
  - After the onset of anaesthesia (1 to 5 min) if required, the eleutheroembryos are transferred to the support to be used for imaging such as a black plastic surface or black 96-well plates.
  - Place the eleutheroembryos so that the ventral side can be imaged by the imaging system.
  - Capture an image of each eleutheroembryo.
  - After all images are taken for an exposure group, euthanise the eleutheroembryos.
  - Continue until all groups are read.
  - Analyse the images by following the instructions in sections §49 to §55.
- Interpreting the results
  - Once the pooled data has been statistically analysed and graphed, the lowest observed effect concentration (LOEC) should be noted for EE2.
  - The LOEC should be at least 114 ng/L for EE2 and the T, T + EE2 and T + fadrozole controls should be statistically significantly different to the relevant controls.
  - The EE2 controls should exhibit a concentration-response relationship over the range of concentrations tested.
  - If a concentration-response relationship is not apparent due to either poor sensitivity at lower concentrations or signal saturation at higher concentrations, then efforts should be made to adjust the image capture parameters to improve the concentration-response relationship.

- If values of zero are present in the raw data for the fluorescence measurements for the solvent or water control, then efforts should be made to adjust the image capture parameters to ensure that all eleutheroembryos in the negative control group give values >0.

ANNEX 4: RECEIVING EMBRYOS: ACCLIMATION AND BATCH ACCEPTANCE

- Embryos should be received no later than 3 days before the test begins to allow a proper recovery and acclimation.
- Batches should be accepted only if dead or abnormal embryos represent less than 20% of the total number between the reception of the batch and the start of the exposure.

Guidance for embryos received three days before the start of the REACTIV assay:

- Do not mix embryos fertilised on different days.
- Sort embryos to remove dead and abnormal embryos, these embryos should represent less than 20% otherwise the batch should not be used to perform the REACTIV assay.
- Transfer only the living and normal embryos to a 1.4 L crystalliser or 15 cm Petri dish containing water suitable for raising medaka embryos (see Annex 5).
- The maximum density per crystalliser is 500 embryos, the maximum density per Petri dish is 200 embryos.
- Incubate embryos with illumination at approximately 26°C with a 14:10 h light:dark cycle. The temperature should be adjusted as required in order for the embryos to hatch around DPF7-10 (tolerance DPF 7-12). Although the eleutheroembryos must be DPH0, they can have a different number of DPF. The difference should not be more than one DPF in a single run. All eleutheroembryos should be randomly selected for the different exposure groups.
- The medium that the embryos are raised in should be changed at least once during the period of embryonic development leading to hatching.

ANNEX 5: SOME CHEMICAL CHARACTERISTICS OF AN ACCEPTABLE WATER FOR RAISING MEDAKA EMBRYOS

Table 4: Characteristics of water suitable for raising medaka embryos to hatch.

Characteristic	Recommended range	Tolerance
Dechlorinated	-	Essential
Particle filtered	25 µm	Recommended
Activated charcoal filtered	-	Recommended
Conductivity	230-290 micro Siemens	
Temperature	26°C	26-30°C
Methylene blue	1 ml of 1 g/L stock per L	Recommended
pH	7.2-8.2	Essential

Alternatively, if an artificial medium is to be used, one option which has been extensively tested including within the OECD interlaboratory validation exercise is detailed here:

A stock solution of 10x Medaka Medium has the following composition:

- NaCl 5 g/L
- $CaCl_2$ 0.151 g/L
- $MgSO_4$ 0.098 g/L
- KCl 0.15 g/L
- NaOH 1N 1.25 mL/L

This solution should then be diluted ten-fold with reverse osmosis water to obtain the 1x working solution. The pH should be adjusted to between 7.2-8.0 with a solution of 1N NaOH.

In addition to artificial media, medaka embryos can also be raised in glass bottled still mineral water, spring water, well water or charcoal-filtered tap water or any medium that supports the normal growth and development of *O. latipes*.

ANNEX 6: PHOTOGRAPHIC GUIDANCE FOR IDENTIFICATION OF NORMAL VERSUS ABNORMAL ELEUTHEROEMBRYOS

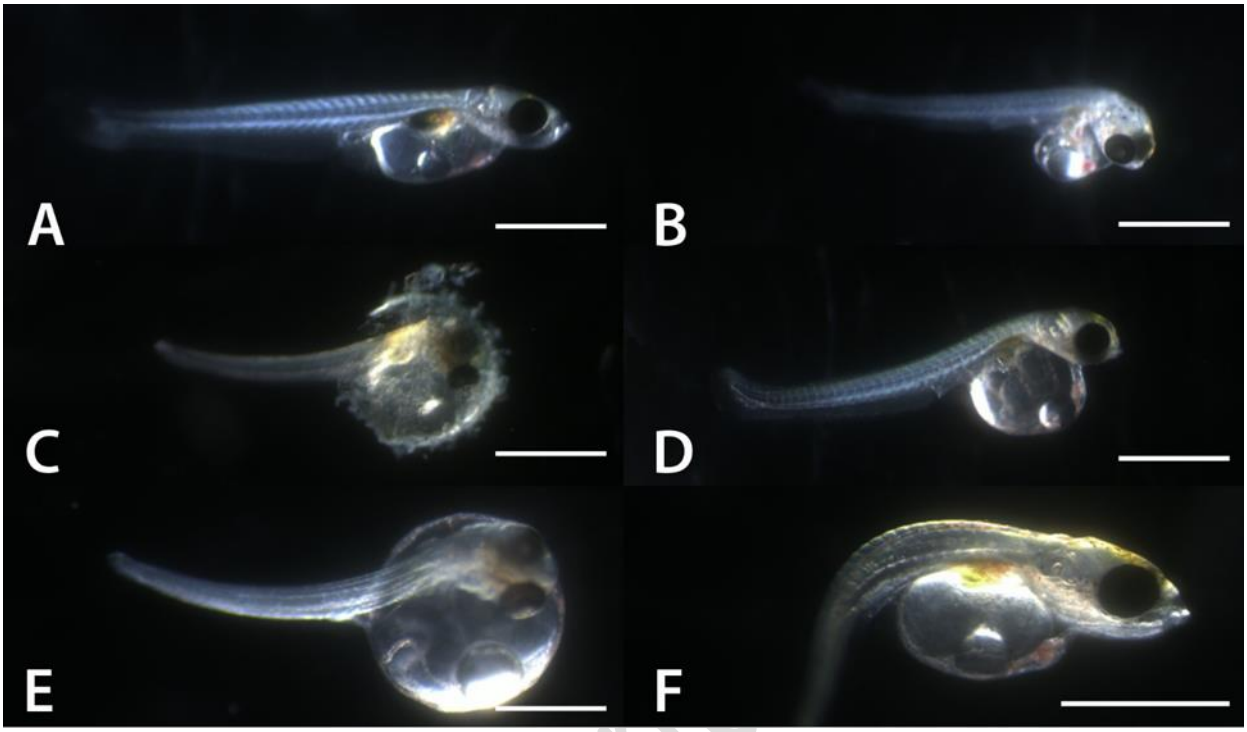

Figure 3: Photographic guidance for identification of normal versus abnormal eleutheroembryos. (A) Normal eleutheroembryo. Abnormal eleutheroembryos: (B) small, the eleutheroembryo clearly has a shorter length than other eleutheroembryos from the same batch; (C) partially hatched, the eleutheroembryo has not yet completely emerged from its egg; under developed, (D and E) both exhibit extremely large yolk sacs for a hatched medaka which still have a spherical shape; (F) malformed, the tail is curved downwards. Scale bars indicate 1 mm.

ANNEX 7: ELEUTHEROEMBRYO POSITIONING

Figure 4 below shows the expected positioning of the eleutheroembryos for imaging. Eleutheroembryos are considered as correctly positioned if they are in a position that allows imaging of the ventral region including the area where the liver is positioned.

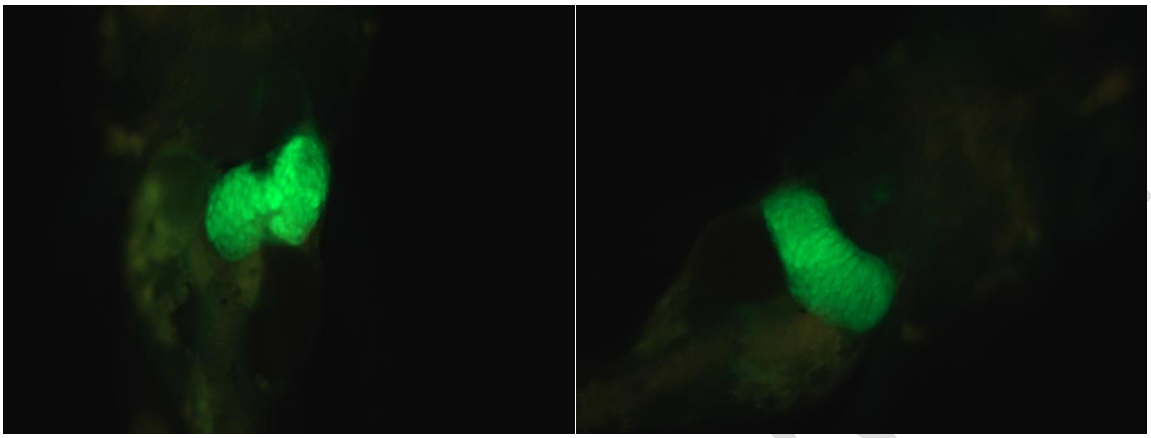

Figure 4: A and B) Ventral views of two *chgh-gfp* medaka eleutheroembryos displaying green fluorescent protein (GFP) signal in the livers. A) The head of the eleutheroembryo is partly out of view at the top of the image. B) The head of the eleutheroembryo is partly out of view at the top-right of the image.

ANNEX 8: METHODS FOR THE STATISTICAL ANALYSIS OF REACTIV ASSAY DATA

Method 1

The recommended statistical approach (Figure 5), which was evaluated during the interlaboratory validation exercise, is to first determine whether the data for each exposure group is normally distributed by performing a D'Agostino-Pearson normality test. To determine whether variance is homogenous, a homoscedasticity test (e.g., Levene's test) should be performed.

If the data are normally distributed and homogeneous variance assumption is not violated, then an ANOVA test should be performed on the unspiked test chemical groups and the negative control (solvent control or test medium control if no solvent it used), followed by a Dunnett's post-hoc-test. Likewise, an ANOVA test should be performed on the spiked test chemical groups and the 30 µg/L testosterone control, followed by a Dunnett's post-hoc-test.

If the data follow a normal distribution but the equal variance assumption is violated, a Kruskal-Wallis test should be performed on the unspiked test chemical groups and the negative control (solvent control or test medium control if no solvent it used), followed by a post-hoc Dunn's test or Welchs's many-to-one comparison test. Likewise, a Kruskal-Wallis test should be performed on the spiked test chemical groups and the 30 µg/L T control, followed by a Dunn's post-hoc-test test or Welchs's many-to-one comparison test.

If the data do not follow a normal distribution, a Kruskal-Wallis test should be performed on the unspiked test chemical groups and the negative control, solvent control or test medium control if no solvent it used, followed by a Dunn's post-hoc-test. Likewise, a Kruskal-Wallis test should be performed on the spiked test chemical groups and the 30 µg/L T control, followed by a Dunn's post-hoc-test.

Differences in mean fluorescence values are considered statistically significant if $P<0.01$ (normally denoted by **).

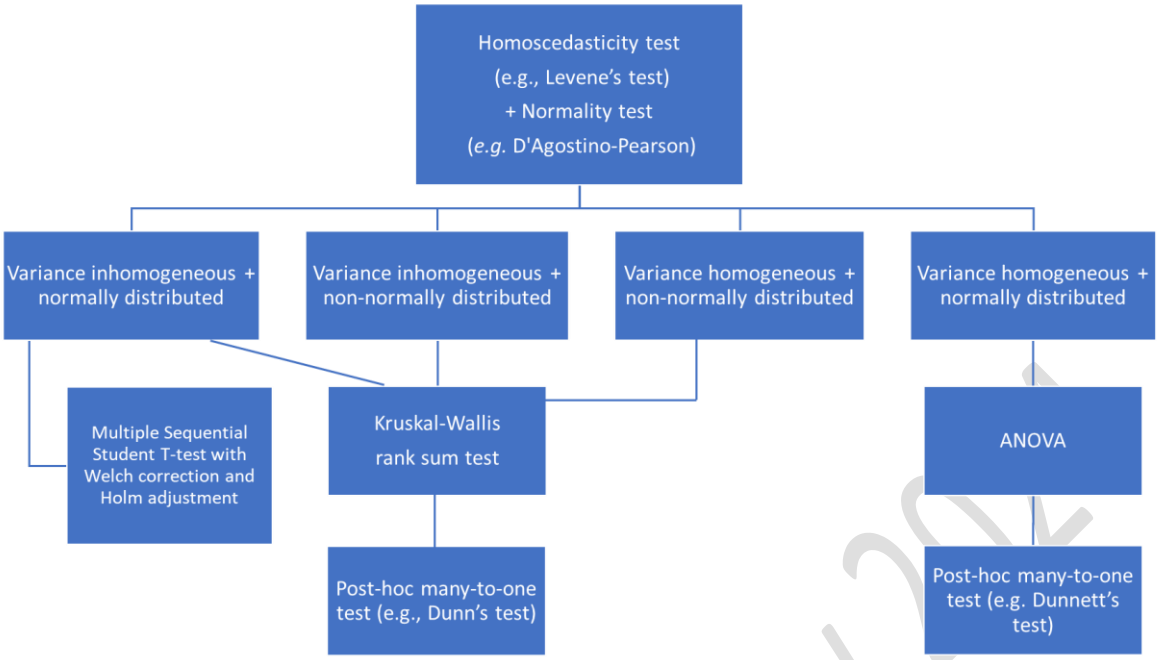

Figure 5: The recommended statistical workflow for comparing more than two groups when analysing the REACTIV assay.

Method 2

Alternatively, a mixed/nested ANOVA approach can be used for the statistical analysis of the data. In this case, a visual inspection of the data per run is strongly advised.

This statistical approach is based on a mixed/nested ANOVA model with the following structure:

$$y_{ijk} = \mu + \alpha_i + \beta_i + \beta_{ij} + \varepsilon_{ijk},$$

where **y_{ijk}** is the measured fluorescence of sample **k** in the run **i** treated with concentration **j**. The model contains a single fixed factor (treatment, **α_i**) and two random factors (the run **β_i** and the run-treatment interaction **β_{ij}**). **ε_{ijk}** describes the error term of the model.

This approach is comparable to the recommendation in Annex 13 of the OECD Test No. 248 for the XETA assay, where a similar experimental set-up is carried out with treatments being nested in runs, leading to variance components for run and run-by-treatment. This is different from the ecotoxicity experimental designs used in most OECD guidelines where replicates are nested within each treatment dose. Analysis of REACTIV data treating replicates/runs incorrectly as nested within treatment has significant effects on the power properties of the tests (OECD 2019c Annex 3).

If R is used to analyze the study, the mixed ANOVA model could be constructed using the lme4 R package (Bates *et al.*, 2015):

```
lme4::lmer(Fluorescence ~ Treatment + (1|Run) + (1| Run:Treatment), REML=TRUE)
```

It is also possible, to treat Run as a fixed effect instead of random effect, which allows the analysis per Run. However, the core analysis should be focused on the population effect of the treatment. Investigations per Run do not provide unbiased information about the effect on the population level.

As default, the treatment group estimates from the mixed ANOVA model should be compared to the control response, using pairwise Dunnett's test at alpha level 0.05. A two-sided test should be carried out unless there is scientific justification to expect only a change in one direction.

Only if a clearly monotonically increasing or decreasing treatment response relationship is detected, can a Williams test be conducted. Therefore, the standard error of each mean difference between each Treatment and the Control is taken from the Dunnett's test result table. Those standard errors and pooled degrees of freedoms (e.g. Kenward-Rogers dfs) are used in an otherwise standard Williams test (Green *et al.*, 2018; OECD 2006; OECD 2019 Annex 13d). In cases where a clearly increasing or decreasing dose-response relationship is detected, it is already known in which direction the effect should be tested for (at alpha level 0.05). This recommendation is deviating from the statement that a trend test should be conducted one-sided in each direction at the 0.025 alpha level, when the direction of testing is not clear (OECD 2006). However, the here provided recommendation is based on practicality since the alpha value of the standard Williams test can often times not be adjusted to a value other than 0.05 (Green *et al.*, 2018).

Deviation from monotonicity can be identified by visual inspection, by issues with the PAVA algorithm of the Williams test (e.g. when the majority of treatment means are amalgamated) and/or by a monotonicity test (Green *et al.*, 2018; OECD 2006). When applying a monotonicity test, it is recommended to assume monotonicity solely when the linear contrast is significant.

When pre-tests are used to test for normality and variance homogeneity among treatment groups, this should be done with the residuals of the mixed ANOVA model. Normality can be assessed using e.g. a Shapiro-Wilk test and variance homogeneity with e.g. a Levene`s test. The alpha value should be 0.01. Visual investigation of residual- and quantile-quantile plots is recommended. In case of deviations from normality and variance homogeneity, outlier removal (e.g. by applying the Tukey rule (Green *et al.*, 2018) and data transformation (for example log- or square-root) can be conducted.

An advantage of the mixed ANOVA approach compared to method 1 is that method 1 does not account for the variability of the interaction between run and treatment. By properly accounting for this source of variability, the mixed ANOVA model can help to make more accurate inferences about the treatment effects on the measured fluorescence.

ANNEX 9: TYPICAL CONCENTRATION-RESPONSE CURVES AND THEIR INTERPRETATION

To aid with interpretation of the REACTIV assay, example histograms are shown below of results obtained during the OECD validation study for the four proficiency chemicals. The interpretation of each result is discussed briefly. It should be noted that during the validation study, all controls including optional controls, were performed by all laboratories.

Anastrozole

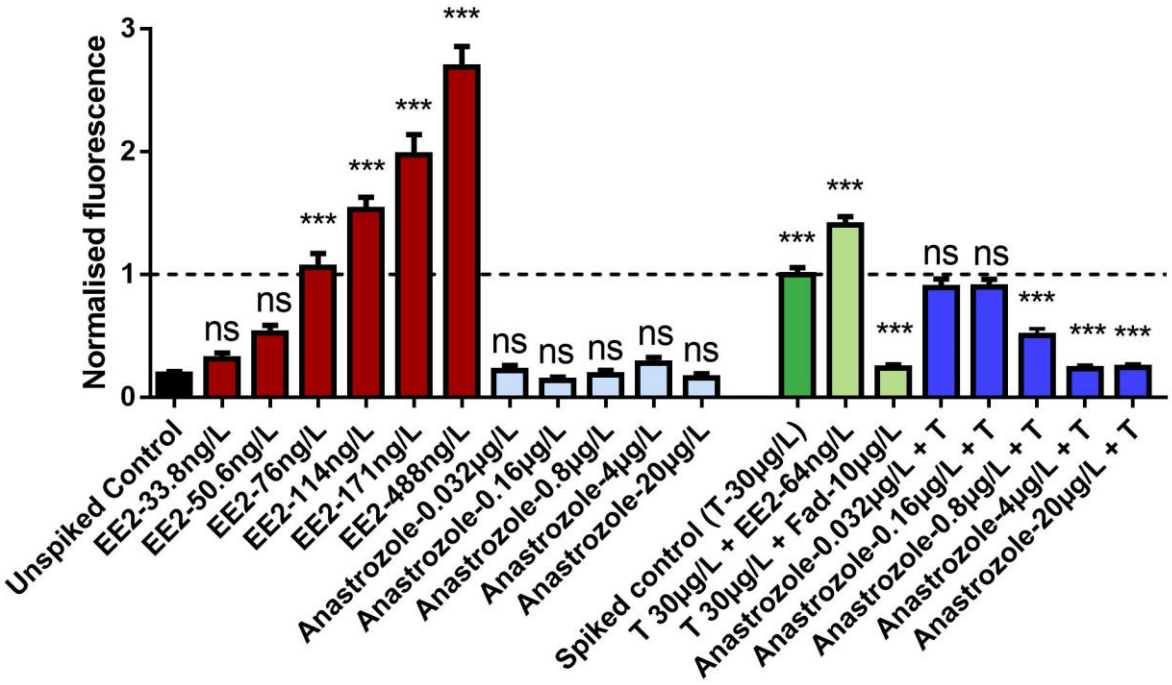

Figure 6: An example result obtained with the proficiency chemical anastrozole during the OECD validation study. Fluorescence was normalised to the mean fluorescence value of the 30 µg/L T control. Statistical significance is shown as: * : p < 0.05 ; ** : p < 0.01 ; *** : p < 0.001 ; ns : not significant p > 0.05. Changes in fluorescence are considered as significant at p < 0.01.

Validity criteria had already been met for the individual runs. Figure 6 shows that all validity criteria related to the performance of the controls in the pooled data set were met by the laboratory performing this REACTIV assay. The mean normalised fluorescence of the 488 ng/L EE2 and spiked control groups were statistically significantly different to the unspiked control group by at least P<0.01. Likewise, the T + EE2 and T + fadrozole control groups were statistically significantly different to the spiked control group by at least P<0.01.

The normalised mean fluorescence of at least one concentration of anastrozole in spiked mode (dark blue bars) was statistically significantly different to the spiked control group (dark green bar) and a monotonic concentration-response profile was observed. Therefore, it was concluded that anastrozole is active in the REACTIV assay.

Tamoxifen

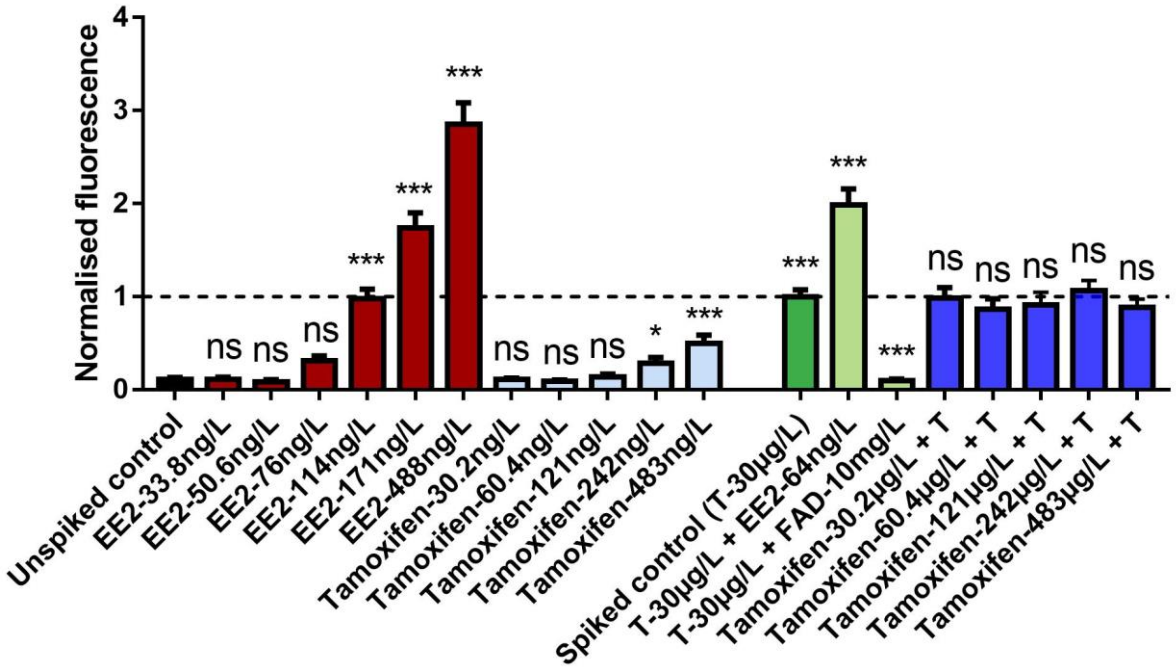

Figure 7: An example result obtained with the proficiency chemical tamoxifen during the OECD validation study. Fluorescence was normalised to the mean fluorescence value of the 30 µg/L T control. Statistical significance is shown as: * : p < 0.05 ; ** : p < 0.01 ; *** : p < 0.001 ; ns : not significant p > 0.05. Changes in fluorescence are considered as significant at p < 0.01.

Validity criteria had already been met for the individual runs. Figure 7 shows that all validity criteria related to the performance of the controls in the pooled data set were met by the laboratory performing this REACTIV assay. The mean normalised fluorescence of the 488 ng/L EE2 and spiked control groups were statistically significantly different to the unspiked control group by at least P<0.01. Likewise, the T + EE2 and T + fadrozole control groups were statistically significantly different to the spiked control group by at least P<0.01.

The normalised mean fluorescence of at least one concentration of tamoxifen in unspiked mode (light blue bars) was statistically significantly different to the unspiked control group (black bar) and a monotonic concentration-response profile was observed. Therefore, it was concluded that tamoxifen is active in the REACTIV assay.

Atenolol

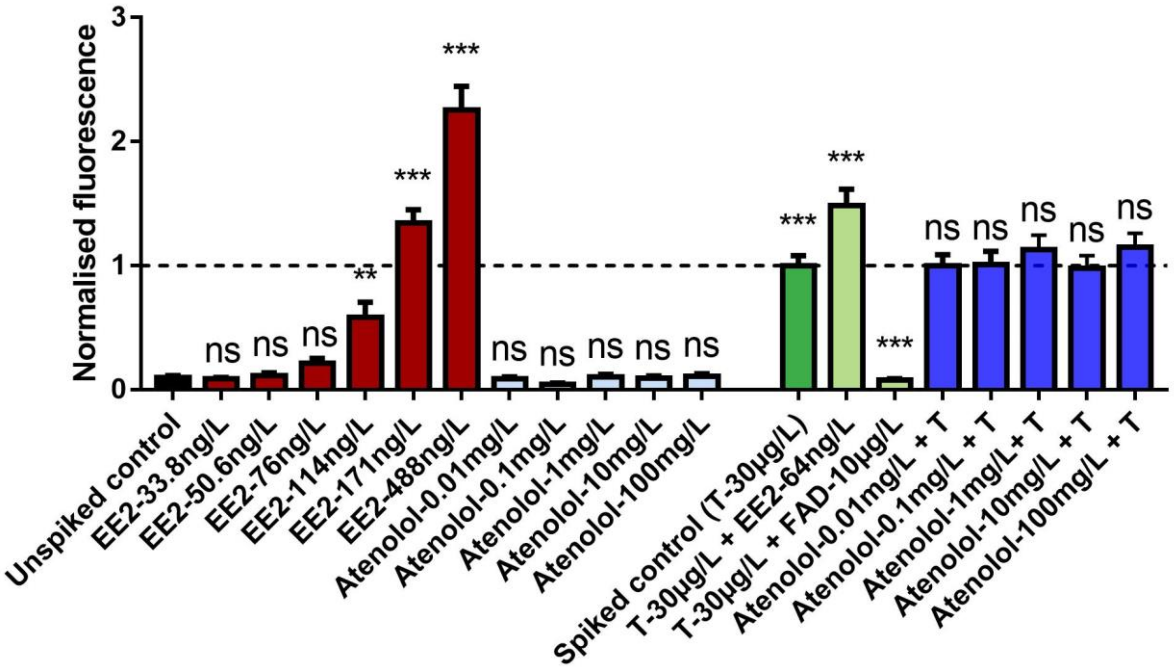

Figure 8: An example result obtained with the proficiency chemical atenolol during the OECD validation study. Fluorescence was normalised to the mean fluorescence value of the 30 µg/L T control. Statistical significance is shown as: * : p < 0.05 ; ** : p < 0.01 ; *** : p < 0.001 ; ns : not significant p > 0.05. Changes in fluorescence are considered as significant at p < 0.01.

Validity criteria had already been met for the individual runs. Figure 8 shows that all validity criteria related to the performance of the controls in the pooled data set were met by the laboratory performing this REACTIV assay. The mean normalised fluorescence of the 488 ng/L EE2 and spiked control groups were statistically significantly different to the unspiked control group by at least P<0.01. Likewise, the T + EE2 and T + fadrozole control groups were statistically significantly different to the spiked control group by at least P<0.01.

None of the tested concentrations of atenolol elicited a statistically significant difference in normalised mean fluorescence in unspiked mode (light blue bars) when compared to the unspiked control group (black bar).

None of the tested concentrations of atenolol elicited a statistically significant difference in normalised mean fluorescence in spiked mode (dark blue bars) when compared to the spiked control group (dark green bar).

Therefore, it was concluded that atenolol is inactive in the REACTIV assay.

Saccharin

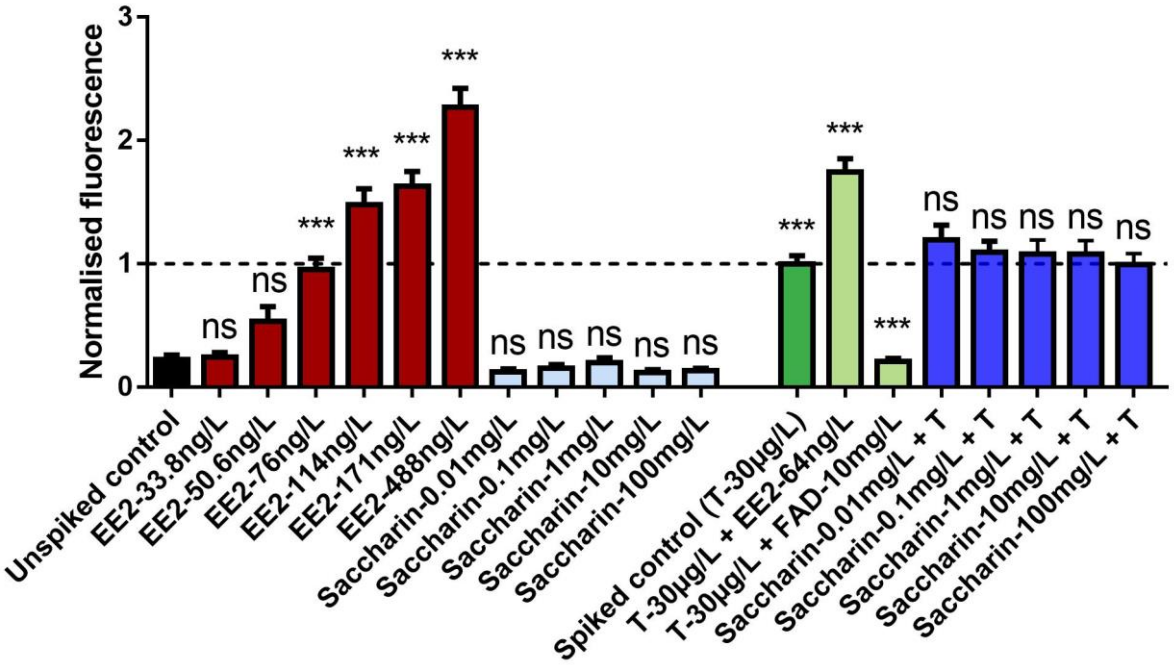

Figure 9: An example result obtained with the proficiency chemical saccharin during the OECD validation study. Fluorescence was normalised to the mean fluorescence value of the 30 µg/L T control. Statistical significance is shown as: * : p < 0.05 ; ** : p < 0.01 ; *** : p < 0.001 ; ns : not significant p > 0.05. Changes in fluorescence are considered as significant at p < 0.01.

Validity criteria had already been met for the individual runs. Figure 9 shows that all validity criteria related to the performance of the controls in the pooled data set were met by the laboratory performing this REACTIV assay. The mean normalised fluorescence of the 488 ng/L EE2 and spiked control groups were statistically significantly different to the unspiked control group by at least P<0.01. Likewise, the T + EE2 and T + fadrozole control groups were statistically significantly different to the spiked control group by at least P<0.01.

None of the tested concentrations of saccharin elicited a statistically significant difference in normalised mean fluorescence in unspiked mode (light blue bars) when compared to the unspiked control group (black bar).

None of the tested concentrations of saccharin elicited a statistically significant difference in normalised mean fluorescence in spiked mode (dark blue bars) when compared to the spiked control group (dark green bar).

Therefore, it was concluded that saccharin is inactive in the REACTIV assay.

ANNEX 10: AVAILABILITY OF THE CHGH-GFP LINE

Concerning access to the *chgh-gfp* Japanese medaka transgenic line, it will be accessible to laboratories from OECD member countries through WatchFrog as well as through partner laboratories. It is envisaged that these partner laboratories will form a network of distributors, possibly including the participants of the ring test as well as stock centres (TEFOR, France; The National BioResource Project, Japan; The National Museum of Natural History, France) as with the XETA assay (TG 248) and RADAR assay (TG 251). A similar network of contract research organisations to the XETA and RADAR assays will also be offered the opportunity to distribute the test independently of the method developer.

Access to this line requires a licensing agreement. The method developer has already signed a legal document committing to applying a FRAND policy established by the OECD to the use of this method. A similar approach has already been successfully applied to the XETA assay (TG 248) and a number of *in vitro* assays.

Establishing this licensing agreement will ensure that the line is the validated line by allowing a legitimate supplier to be identified.