Section 2
Effects on Biotic Systems

Test Guideline No. 248

Xenopus Eleutheroembryo Thyroid Assay (XETA)

18 June 2019

OECD Guidelines for the Testing of Chemicals

OECD
BETTER POLICIES FOR BETTER LIVES

OECD/OCDE 248
Adopted:
18 June 2019

OECD GUIDELINE FOR TESTING OF CHEMICALS

Xenopus Eleutheroembryonic Thyroid Assay (XETA)

INTRODUCTION

1. The Xenopus Eleutheroembryonic Thyroid Assay (XETA) test guideline describes an aquatic assay that utilizes transgenic *Xenopus laevis* (*X. laevis*) eleutheroembryos at stage 45, according to Nieuwkoop and Faber (Nieuwkoop and Faber, 1994), in a multi-well format to detect thyroid active chemicals. The XETA was designed as a screening assay to provide a medium throughput and short-term assay to measure the response of eleutheroembryos to potential thyroid active chemicals (Fini et al, 2007). The XETA is intended to be an amphibian screen classifying the chemicals into potentially thyroid active or inactive but the XETA is not intended to determine toxicity values for risk assessment (e.g., NOEC or ECx). The XETA is placed at level 3 of the OECD conceptual framework for the testing of endocrine disrupters (OECD, 2018a). The OECD GD 150 provides further guidance on the interpretation and extrapolation between taxa of the results of the XETA (OECD, 2018a).

2. The South African clawed frog, *X. laevis*, is the test species selected for the XETA. *X. laevis* is a relevant amphibian model because its early development and thyroid hormones (TH; see Annex 1 for abbreviations) dependent metamorphosis have been extensively studied. In addition, amphibians and higher vertebrates, including mammals, share high genetic homology (Hellsten et al, 2010) as well as similar biotransformation systems and homologous endocrine pathways (Fini et al, 2012), allowing the XETA assay to provide information that may be extrapolated to other taxa. This species is also utilized in the two OECD Test Guidelines using amphibians: the AMA (amphibian metamorphosis assay; OECD TG 231) (OECD, 2009) and the LAGDA (larval amphibian growth and development assay; OECD TG 241) (OECD, 2015).

3. The assay is transcription-based and uses a transgenic *X. laevis* line harbouring the THb/ZIP-GFP genetic construct. This genetic construct comprises of the promoter of the TH/bZIP gene coupled to a reporter gene for Green Fluorescent Protein (GFP). The TH/bZIP gene codes for a transcription factor associated with amphibian metamorphosis, a process controlled by TH. The expression of the TH/bZIP gene is regulated directly by TH at the moment of metamorphosis (Furlow and Brown, 1999). TH/bZIP expression is a trigger for metamorphosis and in part controls its timing. The use of this gene as a biomarker allows the detection of potential modulations of thyroid activity induced by the

test chemical. Before performing the XETA, the laboratory should verify that it has the certifications that may be required by local regulations on the use of transgenic organisms. The XETA should be performed using the THb/ZIP-GFP transgenic line used for the test guideline development, which is commercially available (OECD, forthcoming). The use of another transgenic line based on the THb/ZIP 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 used in the decision logic. Therefore, other transgenic lines could not be considered as "me-too" methods.

4. This guideline proposal is based on a two-phased international interlaboratory validation study conducted between 2012 and 2017 (OECD, forthcoming). The test has been validated in six laboratories with nine mono-constituent substances when considering the two successive validation phases.

5. The endpoint measured is induction of fluorescence in 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 compared to unexposed individuals where fluorescence remains at the basal level.

6. The test chemical is tested in the presence and absence of 3.25 µg/L of the TH triiodothyronine (T3). As TH concentration remains very low at this larval stage, adding T3 to the test medium allows the detection of test chemicals affecting T3 availability or antagonising the thyroid hormone receptor (TR). The differential gene expression induced by the combination of T3 and the tested chemical is therefore a laboratory induced phenomenon, not observed in the absence of exogenous T3 at this developmental stage, and thus not relevant to natural amphibian individuals and populations in the field.

INITIAL CONSIDERATIONS AND LIMITATIONS

7. The assay measures the ability of a chemical to activate or inhibit transcription of the genetic construct, whether directly through binding to the thyroid hormone receptor (TR) or modifying the binding of TH to the TR, or indirectly by modifying the amount of TH available to activate the TR and thereby transcription of the TH/bZIP-GFP construct. The XETA detects chemical effects in thyroid hormone sensitive tissues (i.e. tissues containing thyroid hormone receptors), and not only effects on the hypothalamic-pituitary-thyroid axis. To date the XETA has been shown to detect chemicals acting through various mechanisms of action including TR agonists (e.g., Thyroxine [T4], 3,5,3'-triiodothyroacetic acid [TRIAC]), pharmacological antagonists of the TRs (e.g., NH3; CAS 447415-26-1), modulators of TH clearance including UDP-glucuronosyltransferase modulators (e.g., phenobarbital) and modulators of TH metabolism, including deiodinase inhibitors (e.g., iopanoic acid) (Fini et al, 2007 and OECD, forthcoming). In addition, the XETA potentially detects modulators of TH transport via interaction with TH plasma binding proteins and inhibitors of TH transmembrane transporters. As *X. laevis* NF45 stage eleutheroembryos do not synthesise their own TH, inhibitors of TH synthesis are not intended to be detected by the XETA. The XETA does not distinguish between the different modes of action but provides information on whether a chemical acts as a global activator or inhibitor of the thyroid signalling pathway in the *X. laevis* eleutheroembryo. As the transcription of the TH/bZIP-GFP construct requires the direct action of TR on the TH/bZIP promotor, chemicals affecting TH signalling through alternative signalling pathways that do not lead to an alteration in the interaction between TR and DNA (i.e "non-genomic actions") are not expected to be detected by the XETA.

8. 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 when excited at wavelengths between 450 and 500 nm and able to accumulate in the eleutheroembryo. 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 thyroid active. A simple protocol to determine if the test chemical emits fluorescence is proposed in paragraph 25. This protocol requires the use of wild-type eleutheroembryos.

9. When considering testing of mixtures, difficult-to-test chemicals (e.g. unstable), or test chemicals not clearly within the applicability domain described in this Guideline, upfront consideration should be given to whether the results of such testing will yield results that are meaningful scientifically. 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, 2019). The test design described in this test guideline is not suitable to test volatile chemicals.

PRINCIPLE OF THE TEST

General experimental design

10. The general experimental design entails exposing stage NF45 transgenic THb/ZIP-GFP *X. laevis* eleutheroembryos in 6-well plates to a test chemical in the presence "spiked mode" and absence "unspiked mode" of a co-treatment with 3.25 µg/L of T3. It is recommended to use a minimum of three concentrations plus controls. The test uses 20 eleutheroembryos distributed in two wells (10 organisms per well) per treatment level (test concentrations and controls), under semi-static regime. With three test concentrations and controls, performed in three runs, the XETA uses 540 eleutheroembryos (see Figure 1). The exposure duration is 72 h with a daily renewal (i.e. after 24 h and 48 h) of the exposure solutions. Three independent runs should be performed for each assay. The assay measures GFP fluorescence in the transgenic THb/ZIP-GFP eleutheroembryo by way of a spectrofluorimeter or fluorescence imaging that transforms the fluorescence signal to a numerical format. A detailed overview of test conditions can be found in Annex 2.

Controls

11. The XETA requires the following control groups:

- Test medium control: 2 wells with 10 organisms/well are exposed to test medium only. This control defines the basal fluorescence level in the test medium.
- T3 control: This control establishes the fluorescence level for a T3 concentration of 3.25 µg/l. This concentration is equivalent to the plasma T3 hormone concentration during *X. laevis* metamorphosis and is known to induce morphological changes and TH target genes modulation in premetamorphic tadpoles (Leloup and Buscaglia, 1977 and Shi, 2000). This control serves as a positive control for the groups without T3 co-treatment and a reference control for the group receiving a T3 co-treatment.

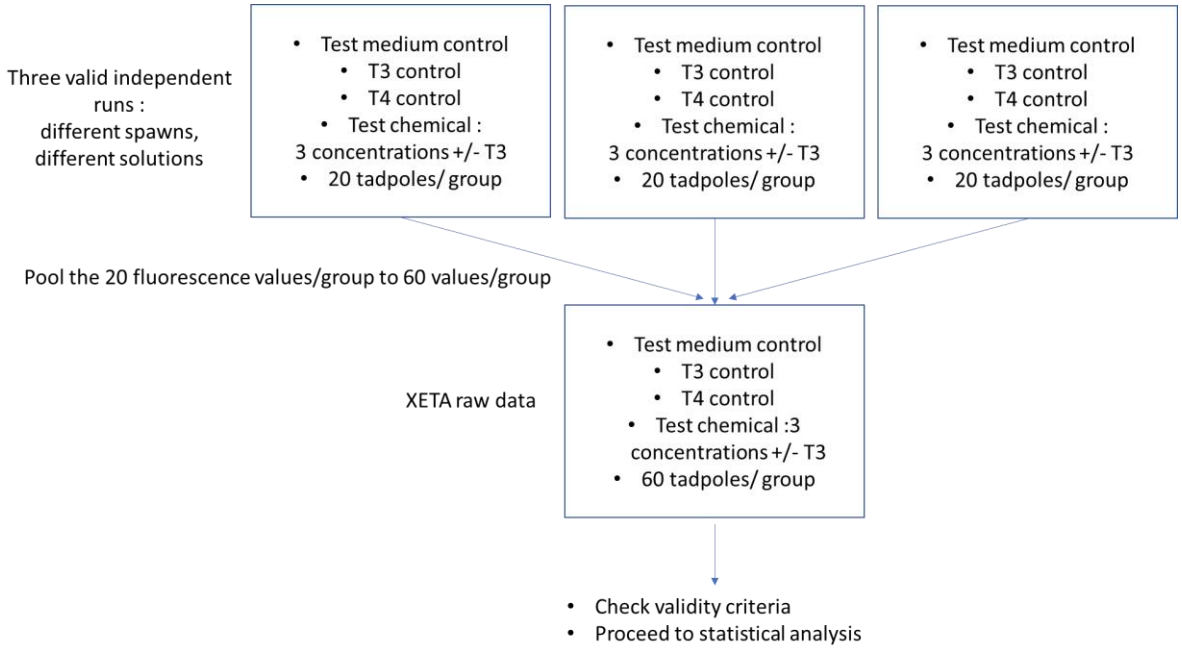
- T4 control (saturation control): This control establishes the maximal fluorescence level that can be quantified in the experiment. This control serves as a positive control for the groups with T3 co-treatment and together with the test medium control defines the dynamic range of fluorescence for a given experiment.

12. If a solvent is used, the test medium control, T3 control and T4 control should receive an equal concentration of solvent (see also §32).

Replication

13. One test is composed of three independent and valid runs using 2 x 10 organisms/treatment group (see figure 1). Each run should be performed using independent solutions and spawn (see paragraph 37). The runs should be ideally conducted sequentially but could be conducted in parallel. The raw data for a given test chemical is obtained by pooling the data from the three runs to obtain n=60 fluorescence values in each treatment group.

Figure 1. Overview of the XETA. ("+/- T3" refers to spiked and unspiked groups).



Note: A XETA assay is composed of three independent runs and utilizes 540 eleutheroembryos in total. A solvent control and a T3 solvent control, should be performed in each run in addition to the test medium control and the T3 control if the solvent has not been tested previously and shown to be negative using the XETA (requires 120 additional eleutheroembryos).

INFORMATION ON THE TEST CHEMICAL

14. Available information on the test chemical should be reported (see paragraph 53) and includes e.g., the structural formula, purity and, if available, stability in light, stability under the test conditions, pKa, Pow, information on the fate of the test chemical and on 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, 2014a).

15. The water 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, 2014b). Analytical determination of the test chemical concentration should be performed before and after test medium renewal (see paragraph 35).

DEMONSTRATION OF PROFICIENCY

Fluorescence quantification

16. The XETA 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 spectrofluorometer and to ensure that a suitable dynamic range of fluorescence measurements can be read by the equipment. These experiments are detailed in Annex 3. Alternatively, the fluorescence emitted by the eleutheroembryo can be quantified by imaging using a fluorescence microscope equipped with an appropriate camera (OECD, forthcoming). In this case the amplitude of fluorescence induction obtained with a range of concentrations of T3 should meet the same criteria as for the spectrofluorometer; optimisation of image acquisition and image analysis parameters is strongly recommended using the same procedure as that detailed for spectrofluorometers in Annex 3. Using a fluorescence microscope equipped with an appropriate camera is the preferable method as this allows a quality control step to be performed on the pictures to identify misplaced eleutheroembryos or fluorescence signal not related to thyroid activation (fluorescent dust or fibers, fluorescent test chemical accumulated in the eleutheroembryo, abnormal fluorescent pattern).

Proficiency chemicals

17. Prior to routine use of this test guideline, laboratories should demonstrate technical proficiency by correctly categorising the four proficiency chemicals listed in Table 1.

Table 1 . Proficiency chemicals. T4 (thyroxine), PTU (Propylthiouracil).

Chemicals	CASNR	Category	Concentrations to test	Expected significant effect
T4	51-48-9	Active	0.001; 0.01; 0.1; 1 and 10 mg/L	0.01 to 0.1 mg/L
PTU	51-52-5	Active	1; 3; 10; 30; 100 mg/L	3 to 100 mg/L
Abamectine	67-64-1	Inert	0.1; 0.5; 1; 5 and 10 mg/L	Inert
Methomyl	16752-77-5	Inert	18; 36; 56; 112; 168 µg/L	Inert

Note: The expected significant effects were determined from the OECD XETA validation (OECD, forthcoming).

VALIDITY OF THE TEST

18. For the test to be valid, the following criteria should be met for each run and for the pool of the three runs:

- A statistically significant induction of fluorescence should be measured between the test medium control group and the T3 control group. The mean fluorescence of the T3 group should be at least 20% higher than mean of fluorescence of the test medium control group.
- A statistically significant induction of fluorescence of at least 70% should be present between the T4 control group and the test medium control.
- The coefficient of variation of the fluorescence intensity measured for the test medium control should not exceed 30%.
- The initial pH of the exposure solutions should be between 6.5 and 8.5 for each renewal.
- The mortality should not exceed 10% in each control group.
- The percentage of malformed organisms should not exceed 10% in each control group.

  These validity criteria are applicable after trimming or picture quality control if a trimming or a picture quality control is performed (see paragraph 45). 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

19. Normal 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;
- conical bottomed black 96-well plates certified for fluorescence quantification;
- pH meter;
- stereomicroscope equipped with a light source (for embryo and eleutheroembryo sorting);
- spectrofluorimeter (96-well plate reader) or fluorescent microscope equipped for GFP fluorescence quantification (OECD, forthcoming);
- analytical instrumentation appropriate for the test chemical or contracted analytical services.

Test organism

20. The test organisms for the XETA are heterozygous *X. laevis* eleutheroembryos of the THb/ZIP-GFP transgenic line. These organisms should be produced by mating a homozygous TH/bZIP-GFP *X. laevis* with a wild-type *X. laevis*. The THb/ZIP-GFP

transgenic line is maintained in several laboratories (OECD, forthcoming) and can be obtained upon subscribing to a license agreement. When a test chemical is shown to be fluorescent, wild type *X. laevis* eleutheroembryos could be also required to check if it could bioaccumulate and produce a fluorescent signal in the eleutheroembryo (see paragraph 25).

21. In a given run, all organisms used as test organisms should be derived from the same spawning. Each spawn should be obtained from a reproduction between one male and one female. The exposure phase of the test is initiated with stage NF45 eleutheroembryos (7 days post fertilisation at 21°C, Annex 4). Eleutheroembryos should ideally be bred within the laboratory from stock animals. Alternatively, eleutheroembryos could be shipped from another laboratory 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 5. At the end of the test, the eleutheroembryos are usually at stage NF47 (Annex 4).

22. Housing and care of *X. laevis* are described by Reed (Reed, 2005). Appropriate care and breeding of *X. laevis* are described by the ASTM (American Society for Testing and Materials) standardized guideline for the FETAX (Frog Embryo Teratogenesis Assay—Xenopus (ASTM, 2012)). A complete breeding protocol is outlined in Annex 6.

Test Medium

23. The test medium could be any water permitting normal growth and development of *X. laevis* including 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 *X. laevis* are not available. Special attention should be given to copper, chlorine and chloramine, all of which are toxic to *X. laevis* eleutheroembryos. Results from analysis of water quality should be reported. Some chemical characteristics of an acceptable test medium suitable for *X. laevis* can be found in Annex 7. However, any medium that supports the normal growth and development of *X. laevis* and meets the test validity criteria is suitable as a test medium.

Feeding

24. Eleutheroembryos between developmental stages NF45 (beginning of the test) and NF47 (end of the test) are used for this test. They are not fed before or during the test as yolk is still present in the intestine from stage NF45 to stage NF47 and is used as the source of energy for the development of the eleutheroembryo (Nieuwkoop and Faber, 1994).

Determining potential fluorescence of the test chemical

25. This test guideline should not be used for test chemicals emitting fluorescence between 500 and 550 nm when excited at wavelengths between 450 and 500 nm and able to accumulate in the eleutheroembryo. 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 thyroid active. A simple protocol to determine if the test chemical emits fluorescence at these wavelengths is to place 200 µL of a solution of the test chemical at the highest concentration intended to be tested in the XETA and 200 µL of test medium into two wells of a 96-well plate and to quantify the fluorescence using the same apparatus and settings as for the quantification of eleutheroembryo fluorescence. If a fluorescent chemical is identified, 20 wild type *X. laevis* eleutheroembryos should be

exposed at 21°C for 72 h with a daily renewal to the highest concentration intended to be tested in the XETA and the fluorescence should be quantified and compared to the fluorescence of a group of 20 wild type eleutheroembryos exposed to test medium only in the same conditions. If a statistically significant difference in fluorescence is present, the chemical is fluorescent and accumulates in the eleutheroembryo body and should not be tested using the XETA.

Selection of test concentrations

Establishing the maximum test concentration

26. For the purposes of this test, the maximum test concentration should be set by the solubility limit of the test chemical in the test medium, the maximum tolerated concentration (MTC), the maximum concentration inducing malformations in less than 10% of eleutheroembryo, or a maximum concentration of 100 mg/L, whichever is lowest. If data from other aquatic toxicity tests are available (e.g. LC50 studies in fish, including fish embryos, QSAR data or possibilities for read-across, or data from other amphibian species), then expert judgment could be used to inform the determination of the maximum test concentration.

27. The MTC is defined as the highest test concentration of the chemical which results in less than 10% acute mortality. A useful approximation of the MTC can be derived from existing acute data as 1/3 of the acute LC50 value. If an *X. laevis* LC50 is not available, than 1/3 of the acute LC50 from other aquatic species (e.g. LC50 studies in fish or other amphibian species), may be used to estimate the MTC. It should be noted that acute toxicity studies are normally performed with adults, therefore, the potential differences in sensitivity between life stages should be accounted for in the estimation of the MTC.

28. If relevant acute toxicity data are not available, before proceeding to a definitive test, the laboratory should perform a range finder test with *X. laevis* eleutheroembryos to evaluate possible toxicity. The range finder should consist of at least 2 test concentrations, including the selected highest test concentration or estimated MTC. The highest concentration tested must result in more than 10% acute mortality.

Test Concentration Range

29. There is a required minimum of three test concentrations and a test medium control (and solvent control if necessary). Generally, a concentration separation (spacing factor) of 3 to 10-fold is recommended.

Test solutions

30. 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 8.5. Stock solutions should be prepared by dissolving the test chemical in the test medium using mechanical means such as agitation, stirring or ultrasonication, or other appropriate methods. If possible, the use of solvents should be avoided. For difficult to test chemicals, including volatile or absorptive chemicals, the OECD Guidance Document No. 23 on aqueous-phase aquatic toxicity testing of difficult test chemicals should be consulted (OECD, 2019).

31. If a solvent is required in order to produce a suitably concentrated and homogeneous stock solution, the maximum solvent level should be at least one order of

magnitude below the appropriate no-observed effect concentration (NOEC) and in any case should not exceed 100 µl/L or 100 mg/L (OECD, 2018b). It is recommended to keep solvent concentration as low as possible (e.g, < 20 µl/L) to avoid potential effects of the solvent on the measured endpoints (Hutchinson et al, 2006).

32. If a solvent is used, the concentration of solvent should be equal in all test concentrations and in all controls. Both the solvent control and test medium control, and both the T3 control and T3 solvent control, should be performed if the solvent was not tested previously and shown to be negative using the XETA. Ideally, a XETA should be performed for each solvent not tested previously with the solvent as a test chemical to determine if the solvent is positive or negative in the XETA. Once it is demonstrated that a solvent does not have an effect, only one control (dependent on the regulatory framework, but normally the water control) is sufficient in the test. The selection of an appropriate solvent depends on the physico-chemical properties of the test chemical and on the sensitivity of *X. laevis*, which should preferably be determined in a previous study. The following solvents have been successfully used in the XETA: dimethylsulfoxyde (DMSO), ethanol, methanol, acetone, acetonitrile. Please note that despite the fact that DMSO has been successfully used in the XETA, it should not be preferentially used as it increases the entrance of the chemical into the organism.

PROCEDURE

Exposure conditions

33. The organisms are exposed in chemically inert cell culture grade 6-well plates (typically wells of 34 mm internal diameter and 20 mm height). Each well should contain ten organisms in 8 mL of solution. In a run, 20 organisms are exposed to each test concentration. Each control group contains 20 organisms (see paragraph 11 for the list of control groups). If plastic well plates are not appropriate for a given test chemical, alternative glass vessels (i.e. small diameter petri dishes) should be used.

34. Eleutheroembryos are maintained in an incubator for 72 ± 2 h at 21 ± 1°C in constant dark throughout the test.

Analytical measurements

35. As a semi-static renewal 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. Analysis of the highest and lowest test concentration at the beginning of the test, at the end of the first renewal cycle (before and after renewal of test solutions), and at the end of the test is considered the minimum requirement. Twenty-four hours ± 2 h renewal periods are the longest periods accepted. If concentrations cannot be maintained within ± 20% in the test system, shortened renewal periods and/or preconditioning of the test vessel with the test solutions for 24 h before initiation to minimize test chemical loss by adsorption could be considered. Use of mean measured concentrations is allowed for test chemicals that do not remain within 80-120% of the nominal concentration; see OECD Guidance Document No. 23 for more details (OECD, 2019).

Day 0

36. The exposure should be initiated when the eleutheroembryos reach developmental stage NF45 (7 days post fertilisation at 21°C; Annex 4).

37. Each run should be performed using a single spawn. Since there are three runs per assay, three spawns should be used per assay. All the eleutheroembryos used for each run should originate from a single reproduction (i.e., spawn from different females should not be mixed). For selection of test animals, eleutheroembryos should be observed under a binocular dissection microscope and those exhibiting grossly visible malformations, abnormal pigmentation, or physical injury (e.g., damage of the tail, oedema, scoliosis) should be excluded from the assay (Annex 8 and 9). 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 an obvious difference in size should be removed. Spawns that contain less than 80% of normal and healthy eleutheroembryos at stage NF45 should not be used for the test.

38. For developmental stage determination, eleutheroembryos should be removed from the pooling tank using a transfer pipette and isolated into a drop of dilution water in a transparent Petri dish. For stage determination, it is preferential not to use anaesthetics. If used, methodology for appropriately using an anaesthetic such as MS-222 (tricaine methanesulfonate) should be obtained from experienced laboratories and reported with the test results. Typically, MS222 is used at 0.1g/L and buffered to pH 7-8. If the eleutheroembryos are anesthetised, the anaesthetic treatment should be applied in the same condition to all eleutheroembryos used for the three runs of the XETA. Eleutheroembryos should be carefully handled during this transfer in order to minimize handling stress and to avoid any injury.

39. The developmental stage is determined using a microscope or by appropriate digital imaging. According to Nieuwkoop and Faber (Nieuwkoop and Faber, 1994), the primary developmental landmark for selecting stage NF45 eleutheroembryos is intestinal morphology. At stage NF45, the intestinal spirals reach one and a half rotations (Annex 4). The morphological characteristics of the intestine should be examined under the microscope. While the complete Nieuwkoop and Faber (Nieuwkoop and Faber, 1994) guide should be consulted for comprehensive information on staging eleutheroembryos, one can reliably determine stage using this prominent morphological landmark.

40. Developmental stage distribution in *X. laevis* is homogenous around stage NF45 (Mengeling et al, 2017). Therefore, it is recommended to stage 10% of the total of organisms required for the study and consider starting the exposure to the test chemical when at least 80% of the staged eleutheroembryos are at stage NF45. Typically, 180 organisms are required for a run. It is recommended that approximately 18 individuals from the test batch are staged. If at least 15 of these organisms have reached stage NF45 the batch is suitable for initiating a test.

41. To start the experiment, 10 eleutheroembryos/well should be placed into 6-well plates in drops of test medium (using a transfer pipet with the tip cut off to avoid damaging the eleutheroembryos). The 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. An example of 6-well plate set up is shown in Annex 10.

Day 1 and day 2

42. The test chemical solutions and the control solutions should be renewed at 24 and 48 h ± 2h. Each well is inspected for organisms with an abnormal appearance (injuries, abnormal swimming behaviour, etc.). Dead organisms or those exhibiting grossly visible malformations (Annex 8) or injuries should be removed and the latter be euthanised as described in §44. All observations should be recorded. If >10% cumulative mortality or >10% cumulative malformation rate is encountered in one of the control groups or one of the treatment groups leaving less than three uncompromised test concentrations, then the on-going independent run is stopped and the source of the mortality or abnormality should be identified.

Day 3 Fluorescence quantification

43. The fluorescence of each organism is quantified after 72 ± 2 h of exposure. Eleutheroembryos should be transferred individually (i.e. one organism per well) into wells of black 96-well plates with conical bottoms suitable for fluorescence quantification. For this purpose, the test solutions should be first renewed with 8 mL of test medium and dead organisms or those exhibiting grossly visible malformations should be removed. All observations should be recorded. If >10% cumulative mortality or >10% cumulative malformation rate is encountered in one of the control groups or one of the treatment group leaving less than three uncompromised test concentrations, then the on-going independent run is stopped and the source of the mortality or abnormality should be identified. Eleutheroembryos should then be anesthetized by adding 2 mL of 1 g/L buffered MS222 into the wells. To avoid excessive anaesthesia, only the number of organisms required to fill one 96-well plate at a time are anesthetised. After complete anaesthesia (2 to 5 min), the eleutheroembryos are placed individually in each well of the 96-well plate. It is recommended to place the organisms from the same well of the 6-well plate in the same row of the 96-well plate (i.e., each row of the 96-well plate is equal to one well of the 6-well plate). An example of 96-well plate is shown in Annex 11. Each organism is placed on its back using a thin transfer pipet (aspirating the medium back and forth will help to turn the eleutheroembryo) and most of the medium is removed while maintaining enough moisture under the organism. The eleutheroembryos are placed separately into the wells with the head in the centre of the well and the dorsal surface in contact with the base of the well. The tail is placed around the eleutheroembryo's head. An image of the correct positioning is shown in Annex 12. The 96-well plate containing the organism is then placed into the spectrofluorometer or under the fluorescent microscope to quantify the fluorescence using the parameters identified during the calibration. Please note that once the eleutheroembryos have been placed on their back and the MS222 removed it is important to proceed quickly to the fluorescence quantification in order to prevent the organisms from drying out. If a spectrofluorometer is used to quantify the signal instead of a fluorescent microscope, there is a risk that unusual patterns of fluorescence, possibly indicative of false positive results, cannot be assessed.

Terminating the experiment

44. After reading the fluorescence, each well of the 96-well plate is filled with a solution of buffered MS222 (1 g/L) using a squeeze bottle to euthanize the eleutheroembryos. The 96-well plates are disposed as required by local laboratory safety protocols.

Analysis of data / Evaluation of test results

Data analysis considerations

45. Fluorescence measurements from wells known to be empty should be removed from the data before any analysis of the data. A trimming could be performed prior to the statistical analysis by omitting in each run the highest and lowest 10% of the fluorescence values for each control group and each group of T3-spiked and unspiked test concentration (i.e. omitting the two highest and two lowest values of each group of 20 values). This trim is intended to remove values that arise in the XETA from several events including missorted eleutheroembryos (abnormal size or pigmentation), misplaced eleutheroembryos in the 96-well plate (upside down or on their side), eleutheroembryos dying after anaesthesia and wells containing no organism. Alternatively, if a fluorescent microscope is used for fluorescence quantification, an image quality control should replace the 10% trim to remove the values corresponding to these events before the statistical analysis. If applicable, the number of images removed from the analysis following the image quality control should be indicated.

46. Data from the three independent runs are pooled to obtain 42 (worst case) to 60 fluorescence values for each concentration and control.

47. If a solvent was used in the experiment for the first time, 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 there is no statistically significant difference between the test medium control and solvent control, the pooled test medium and solvent controls should be used. If a statistically significant difference greater than 12% is detected between the test medium control and solvent control group for the pool of the three runs, the study is compromised. The runs should be then considered individually to determine if the solvent has a reproducible effect among the three runs. If yes, the study is compromised and a new XETA should be performed considering using a new batch of solvent or another solvent. If only one or two runs are compromised, new runs should be performed considering changing the batch of solvent.

Statistical analysis

48. 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 one-tailed hypothesis testing at $p < 0.05$.

49. Step-down trend tests are recommended to determine whether there are significant differences ($p < 0.05$) between the control(s) and the various test chemical concentrations (OECD, 2006). Comparisons should be done either with Williams' test or Dunnett's test to determine any statistical difference (a possible way to perform statistical analyses of the data obtained in the XETA is described in Annex 13).

Decision logic

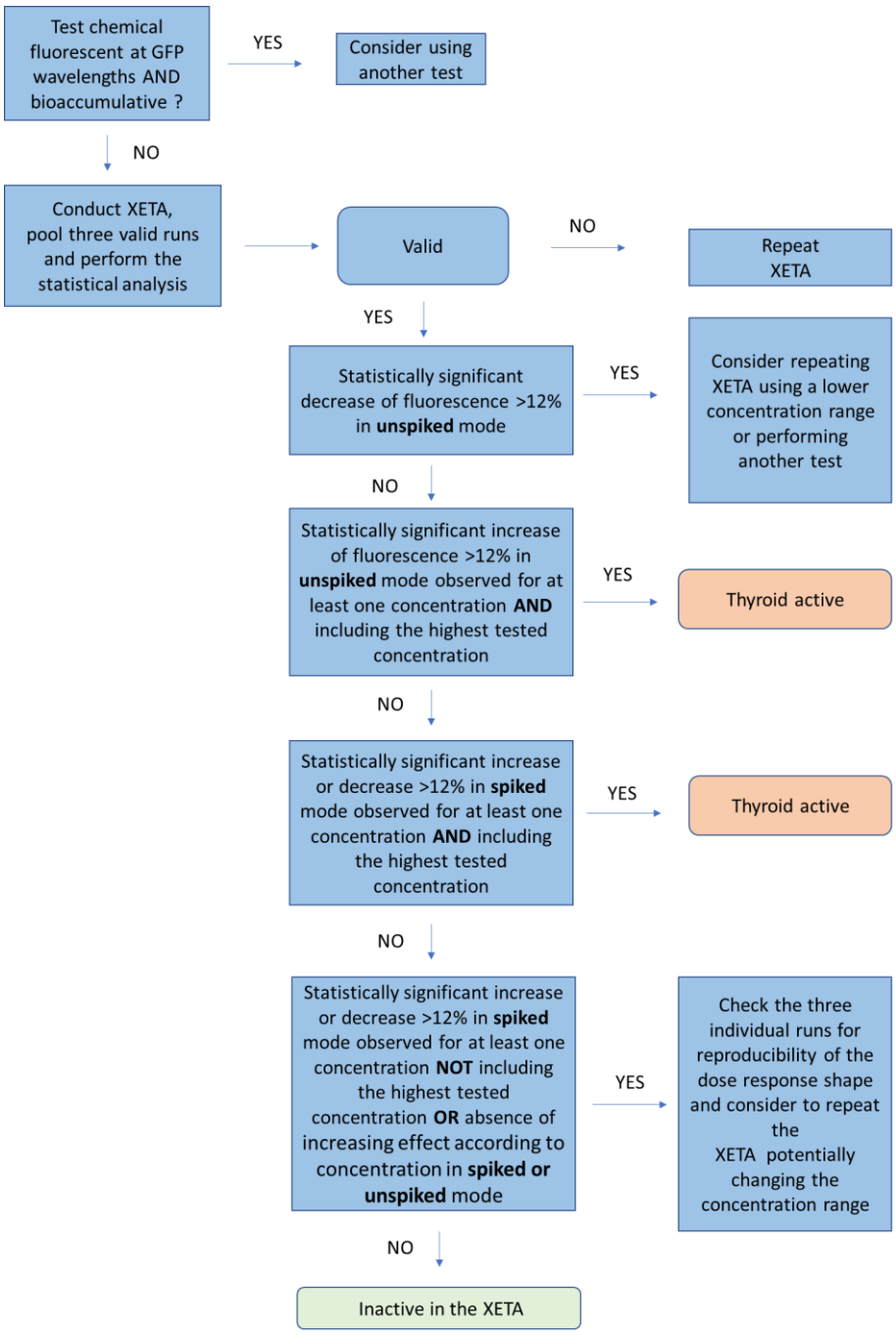
50. A decision logic was developed for the XETA to provide logical assistance in the conduct and interpretation of the result of the bioassay (see flow chart in figure 2). This decision logic is based on three valid runs pooled for statistical analysis (see figure 1 in §13). A test chemical is considered to give a positive result in the XETA if at least one concentration tested including the highest is active in T3-spiked and/or unspiked mode.

- In unspiked mode an active concentration is defined as a concentration giving a statistically significant fluorescence increase of 12% or greater compared to the test medium control.
- In T3-spiked mode an active concentration is defined as a concentration giving a statistically significant fluorescence increase or decrease of 12% or greater compared to the T3 control.
- A 12% threshold was determined from the power analysis which found the power to exceed 80% for all concentration-response shapes simulated. The adequation of this threshold was confirmed by the results for the two validation phases showing no groups of tadpoles treated with a reference inactive chemical giving a statistically significant variation over 12% (OECD, forthcoming).

51. Non-monotonic concentration responses tend to be common in mechanistic assays (Parrott et al, 2017) and could theoretically be obtained using the XETA. If several test concentrations are found to be active, but this activity does not follow a concentration-response relationship, or if one or more concentration(s) tested are found to be active, but the highest tested concentration is inactive, the reproducibility of this result should be confirmed by comparing the results of the three runs of the XETA. If the results are not reproducible, the assay should be repeated, potentially using another dose range, to confirm the result (see flow chart in figure 2). If the result is reproducible the test chemical is considered thyroid active in the XETA.

52. Fluorescence decreases in unspiked mode are not expected as the eleutheroembryos do not synthetize their own TH at this development stage. If a statistically significant fluorescence decrease >12% is observed in the unspiked mode, it could indicate that the test chemical is decreasing the synthesis of GFP by a mode of action that is not thyroid related, in which case the XETA is not appropriate for the test chemical, or a potential problem with the organism or the test conditions which should need further investigations. Individual runs should be considered to determine if the statistically significant fluorescence decrease is present in the three runs and best professional judgement should then be used to decide between repeating: only one run using a new batch of organisms; a complete XETA, possibly using a lower concentration range; or performing a different thyroid activity test.

Figure 2. . Decision logic for the interpretation of the result of the XETA.



Note: The OECD GD 150 provides further guidance on the interpretation and extrapolation between taxa of the results of the XETA (OECD, 2018a).

Test report

53. 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). Information on biodegradation if available.
- 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 regarding lack of fluorescence emission at wavelength of 450 and 500 nm; as well as on lack of accumulation in the eleutheroembryos for substances shown to be fluorescent at these wavelengths.

Test species:

- Scientific name, transgenic line, supplier or source, and culture conditions.

Test conditions:

- Test procedure used (e.g., concentrations tested, temperature, duration, semi-static, volume, number of organism per mL).
- Details of test medium characteristics (reference of mineral water or spring water, description of tap water treatment (e.g., charcoal filtration…)) and any measurements made.
- Method of preparation of stock solutions and frequency of renewal (the solvent and its concentration should be given, when used).
- Brand and references of 6-well and 96-well plates used for exposure and fluorescence quantification.
- References and settings of the spectrofluorometer or fluorescence microscope used for quantification. The method used for image analysis should also be provided for this latter case.

Results:

- Results from any preliminary studies on the LC50 or MTC of the test chemical.
- The nominal test concentrations and 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 and malformed organisms in each run and the day on which they occurred.
- Fluorescence quantification raw data (e.g., individual fluorescence raw data).
- Approach for the statistical analysis and treatment of data including statistical test used and whether and why any data censuring was conducted.
- Evidence that the controls meet the validity criteria.
- Evidence that all groups meet the validity criteria in regards to survival and malformation.
- 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 in a table.
- The percentage increase or decrease of fluorescence for each concentration compared to its respective control in spiked and unspiked modes.
- 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 previous study showing the solvent to be negative in the XETA.
- 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 acceptance criteria, and considerations of potential consequences on the outcome of the test.
- Where appropriate, a discussion presenting the list of concentrations found active in spiked and unspiked mode.
- A conclusion presenting if the test chemical is found to be thyroid active or inactive in the XETA.

REFERENCES

ASTM E1439-12. (2019). Standard Guide for Conducting the Frog Embryo Teratogenesis Assay-Xenopus (FETAX), ASTM International, West Conshohocken, PA. www.astm.org

Fini, J.B., Le Mevel, S., Turque, N., Palmier, K., Zalko, D., Cravedi, J.P., Demeneix, B.A. (2007). An *in vivo* multiwell-based fluorescent screen for monitoring vertebrate thyroid hormone disruption. Environ. Sci. Technol., 41, 5908–5914.

Fini, J.B., Rui, A., Debrauwer, L., Hillenweck, A., Le Mevel, S., Chevolleau, S., Boulahtouf, A., Palmier, K., Balaguer, P., Cravedi, J.P., Demeneix, A., Zalko, D. (2012). Parallel biotransformation of tetrabromobisphenol A in *Xenopus laevis* and mammals: Xenopus as a model for endocrine perturbation studies. Tox. Sci., 125(2): 359-367.

Furlow JD, Brown DD. (1999) *In vitro* and *in vivo* analysis of the regulation of a transcription factor gene by thyroid hormone during *Xenopus laevis* metamorphosis. Mol Endocrinol. December;13(12):2076-89.

Hellsten, U., Harland, R.M, Gilchrist, M.J., Hendrix, D., Jurka, J., Kapitonov, V., Ovcharenko, I., Putnam, N.H., Shu, S., Taher, L., Blitz, I.L., Blumberg, B., Dichmann, D.S., Dubchak, I., Amaya, E., Detter, J.C., Fletcher, R., Gerhard, D.S., Goodstein, D., Graves, T., Grigoriev, I,V., Grimwood, J., Kawashima, T., Lindquist, E., Lucas, S.M., Mead, P.E., Mitros, T., Ogino, H., Ohta, Y., Poliakov, A.V., Pollet, N., Robert J,, Salamov, A., Sater, A.K., Schmutz, J., Terry, A., Vize, P.D., Warren, W.C., Wells, D., Wills, A., Wilson, R.K., Zimmerman, L.B., Zorn, A.M., Grainger, R., Grammer, T., Khokha, M.K., Richardson, P.M., Rokhsar, D.S. (2010). The genome of the Western clawed frog *Xenopus tropicalis*. Science, 328(5978), 633-636.

Hutchinson TH, Shillabeer N, Winter MJ and Pickford DB. (2006). Acute and Chronic Effects of Carrier Solvents in Aquatic Organisms: A Critical Review. Review. Aquat Toxicol 76: 69–92.

Leloup, J., Buscaglia, M. (1977). Triiodothyronine, hormone of amphibian metamorphosis. Comptes Rendus Hebdomadaires des Séances de l'Académie des Sciences Serie D, 284, 2261–2263.

Mengeling BJ, Wei Y, Dobrawa LN, Streekstra M, Louisse J, Singh V, Singh L, Lein PJ, Wulff H, Murk AJ, Furlow JD. (2017) A multi-tiered, in vivo, quantitative assay suite for environmental disruptors of thyroid hormone signaling. Aquat Toxicol. September;190:1-10.

Nieuwkoop, P. D., Faber, J. (1994). Normal Table of *Xenopus laevis* (Daudin). Garland Publishing Inc, New York ISBN 0-8153-1896-0.

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

OECD (2006). Hypothesis Testing, in Current Approaches in the Statistical Analysis of Ecotoxicity Data: A Guidance to Application, Chapter 5. OECD Series on Testing and Assessment no. 54, OECD Publishing, Paris, https://doi.org/10.1787/9789264085275-en.

OECD (2009), Test No. 231: Amphibian Metamorphosis Assay, OECD Guidelines for the Testing of Chemicals, Section 2, OECD Publishing, Paris, https://doi.org/10.1787/9789264076242-en.

OECD (2014a). Test No. 310: Ready Biodegradability - CO_2 in sealed vessels (Headspace Test). OECD Guidelines for the Testing of Chemicals. OECD Publishing, Paris, https://doi.org/10.1787/9789264224506-en.

OECD (2014b). Guidance document 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. Series on testing and assessment no. 204. OECD Publishing, Paris, http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=ENV/JM/MONO(2014)20&doclanguage=en.

OECD (2015), Test No. 241: The Larval Amphibian Growth and Development Assay (LAGDA), OECD Guidelines for the Testing of Chemicals, Section 2, OECD Publishing, Paris, https://doi.org/10.1787/9789264242340-en.

OECD (2018a), Revised Guidance Document 150 on Standardised Test Guidelines for Evaluating Chemicals for Endocrine Disruption, OECD Series on Testing and Assessment, OECD Publishing, Paris. https://doi.org/10.1787/9789264304741-en.

OECD (2019). Guidance document on aqueous-phase aquatic toxicity testing of difficult test chemicals series on testing and assessment no. 23 (second edition). OECD Series on Testing and Assessment, OECD Publishing, Paris, http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=ENV/JM/MONO(2000)6/REV1&doclanguage=en.

OECD (forthcoming). Validation of the xenopus eleutheroembryonic thyroid signaling assay (XETA) for the detection of thyroid active substances. (under development)

Parrott JL, Bjerregaard P, Brugger KE, Gray LE Jr, Iguchi T, Kadlec SM, Weltje L, Wheeler JR. (2017) Uncertainties in biological responses that influence hazard and risk approaches to the regulation of endocrine active substances. Integr Environ Assess Manag. March;13(2):293-301.

Reed, B.T. (2005). Guidance on the housing and care of the African clawed frog *Xenopus laevis*. Research Animals Department. RSCPA.

Shi (2000). Amphibian metamorphosis: from morphology to molecular biology, John Wiley and sons, Inc. ISBN 0-471-24475-9.

ANNEX 1: ABREVIATIONS AND DEFINITIONS

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, 2014). Applying this definition to *X. laevis* positions this period of development from stage NF36 (hatching (Nieuwkoop and Faber, 1954)) to stage NF 48 (yolk exhaustion and beginning of independent feeding (Nieuwkoop and Faber, 1954; Marshall and Dixon, 1978)).

GFP: Green Fluorescent Protein.

LC50: Median Lethal Concentration is the concentration of a test chemical that is estimated to be lethal to 50% 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 than10% acute mortality.

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

PTU: Propylthiouracil (CAS: 51-52-5).

SEM: Standard Error of the Mean.

SMILES: Simplified Molecular Input Line Entry Specification.

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

Spiked mode: Part of the XETA run in the presence of 3.25µg/l of T3.

T3: Triiodothyronine (CAS 6893-02-3).

T4: Thyroxine (CAS : 51-48-9).

TR: Thyroid hormone receptor.

TH: Thyroid hormones, T3 (triiodothyronine) and T4 (thyroxine)**.**

THb/ZIP: Thyroid hormone beta zip transcription factor.

Unspiked mode: Part of the XETA run in the absence of T3.

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

Marshall JA, Dixon KE. (1978) Cell specialization in the epithelium of the small intestine of feeding *Xenopus laevis* tadpoles. J Anat. May;126 (Pt 1):133-44.

Nieuwkoop, P. D., Faber, J. (1994). Normal Table of *Xenopus laevis* (Daudin). Garland Publishing Inc, New York ISBN 0-8153-1896-0.

OECD (2014), Fish Toxicity Testing Framework, OECD Series on Testing and Assessment, n° 177, OECD Publishing, Paris, https://doi.org/10.1787/9789264221437-en.

ANNEX 2: OVERVIEW OF TEST CONDITIONS OF THE XETA

Test Animal		THb/ZIP-GFP *X. laevis* eleutheroembryo
Endpoint		Total fluorescence of individual eleutheroembryo
Exposure period		Stage NF45 (beginning of the test) to stage NF47 (end of the test)
Exposure duration		72 h ± 2 h
Exposure regime		Renewal after 24 h and 48 h. No feeding.
pH		6.5 to 8.5
Incubation conditions during exposure		21 ± 1ºC, dark
Organism per concentration		10 organisms per well (6-well plate) x 2 wells (total of 20 organisms per concentration)
Volume of test medium		8 mL per well
Test medium		Water permitting normal growth and development of *X. laevis* (refer to paragraph 23).
Number of Experiments		Experiments are run 3 times for each test chemical with different spawn and freshly prepared solutions.
Criteria for selecting test individuals		Developmental stage (NF45), health of animal (alive and no malformations), homogenous pigmentation.
Validity criteria		pH 6.5-8.5, mortality ≤10% in all control groups, malformations ≤10% in all control groups Fluorescence induction >20% in the T3 control and >70% in the T4 control. CV<30% in the test medium control. At least three uncompromised test concentrations.
Test chemical concentration standard		Test chemical concentration should remain within 20% of nominal throughout 24 hours otherwise the result should be considered using the determined concentrations.
Controls	Controls	Test medium (basal fluorescence)
	T3 Control	T3 (3.25 µg/L)
	T4 Control	T4 (10 mg/L) (saturation of the fluorescence signal)

ANNEX 3: CALIBRATION: DETERMINATION OF THE OPTIMAL SPECTROFLUORIMETER SETTINGS

The goal of the calibration step is to ensure that an optimal amplitude of response in simple conditions could be achieved using the equipment. It is crucial to ensure that the fluorescence readers allow a suitable dynamic concentration response to be obtained with T3 before testing chemicals.
The calibration will require two steps:

1) Finding the best settings for the spectrofluorimeter to obtain a satisfactory amplitude of GFP induction with a concentration of 3.25 µg/L T3.
2) Apply these settings for the quantitation of three concentration-response experiments with T3 to check the amplitude of induction using increasing concentrations of T3 and the lowest concentration of T3 that elicits a detectable GFP response detectable concentration.

1-Finding the suitable settings for the spectrofluorimeter

- Before starting the experiment:
  - Identify the parameters that could be set on the selected spectrofluorimeter model.
  - Choose a first combination of settings based on the ones used during the interlaboratory validation of the XETA (See XETA validation report (5), annex 2) or/and on any previous experiment with the chosen spectrofluorimeter. Take into account that the settings should lead to a 96-well plate reading time under 30 min.
  - Identify the parameter to optimize first.

- Prepare T3 and T4 stock solutions:
  - T3 STOCK SOLUTION(10-2M) (6.51 G/L)
    - This stock solution can be store 1 year at -20°C.
    - Weigh 100 mg 3, 3', 5-Triiodo-L-thyronine in a weigh boat and gently pour into to a 20 mL volumetric flask.
    - Rinse the weigh boat with 4.47 ml of NaOH 1N and gently add to the 20ml volumetric flask.
    - Mix gently for 1 hour with a stir bar; make sure the solution is limpid.
    - Transfer 4.1 ml of the solution into a new 20 ml volumetric flask; add 9.9 ml of ultrapure water.
    - Agitate 1 hour with a stir bar.
    - Make enough aliquots of 30µl in 0.5ml eppendorf tubes to fill a freezer storage box.
    - Put any remaining solution in several tubes of 1.5 ml.
    - Storage: -20°C for 1 year.

- T4 0.8 G/L STOCK SOLUTION
  - Weigh 40 mg of T4 in a weigh boat and gently pour into to a 50 mL volumetric flask.
  - Rinse the weigh boat with 50 ml of ultrapure water and gently add to the volumetric flask.
  - Add 0.1ml of NaOH 1N.
  - Mix gently for 1 hour with a stir bar; make sure the solution is limpid.
  - Make 1ml aliquots and store at -20°C for 6 months maximum

- Prepare the following solutions

-Exposure medium

- Place 1L of test medium in a 1L glass bottle

-Exposure medium+T3 3.25 µg/L:

- Place 1L of test medium in a 1L glass bottle
- Add 1ml of T3 at 3.25 mg/l

-Exposure medium +T4 10 mg/L
Prepare 300 ml of T4 (10 mg/L):

- Place 296.25 mL of test medium in a 500 ml bottle.
- Add 3.75 ml of T4 at 0.8 g/L.
- Mix by agitation of the bottle.
- Store at 4°C in the dark

- Expose for 72 h exposure 5 groups of 20 controls, 5 groups of T3 controls and 5 groups of T4 controls organisms (in 6-well plate, 10 organisms per well) with a daily renewal as described in the test guideline.
- After 72-hrs, anesthetize the tadpoles with MS-222 and place each study organism in a separate well of a 96-well plate ensuring that none of the eleutheroembryo anesthetized for more than 45 minutes
- Use the spectrofluorimeter with a first set of parameters to quantify the fluorescence. Optimize fluorescence quantification by making appropriate adjustments of settings (one parameter at a time).

Do not keep the same eleutheroembryo anesthetised for more than 45 minutes. Longer times will induce organism death leading to fluorescence artifacts. To continue the calibration after 45 minutes, anesthetise three new groups of organism and load a new 96-well plate.

- After the acquisition of the data, calculate the mean fluorescence and the CV (coefficient of variation) in each group. Calculate the percentage of fluorescence induction in the T3 group and T4 groups.

-Percentage of induction T3 group (RFU = relative fluorescence units)
100*(mean RFU T3 group - mean RFU control group)/ mean RFU control group

-Percentage of induction T4 group =
100*(mean RFU T4 group - mean RFU control group)/ mean RFU control group

Each parameter should to be optimised to:

- Obtain fluorescence induction between the control and the T3 control group as well as between T3 controls and T4 controls close to the ones observed during the validation of the XETA:
- Induction of fluorescence between the control and the T3 control group ranging from 30 to 75% with a mean of 45%.
- Induction of fluorescence between the control and the T4 control group ranging from 70 to 190% with a mean of 105%.
- Ensure that spectrofluorimeter saturation is never reached in the T4 controls.
- Obtain a total 96-well plate reading time under 30 min, ideally 20 min.

When an optimised set of parameters has been determined, repeat the experiment on another spawn to confirm the results.

2- Determination of the amplitude of fluorescence induction and the LOEC using increasing concentrations of T3

This test should be repeated three times on three separate spawns.
Concentrations to test are: 0 mg/L (untreated control), 26 µg/L; 13 µg/L; 6.5 µg/L; 3.25 µg/L; 2 µg/L; 1.5 µg/L; 1 µg/L; 0.65 µg/L.

Preparation of solutions

Preparation of T3 3.25 mg/L solution:

T3-Intermediate Solution (0.65 g/L)

- Remove an aliquot of Stock T3 at 6.51g/L (10^{-2}M) from -20°C
- Make an intermediate solution of T3 at 0.65g/L (250µL):
  - Put 225µL of Test medium in an Eppendorf 1.5ml
  - Add 25 µL of T3 at 6.51g/L
  - Vortex the solution

T3 solution at 3.25 mg/L

- Make a 3.25 mg/L (20ml):
  - Put 19.9 ml of Test medium in a 50 ml falcon tube
  - Add 100 µl of T3 solution at 0.65 g/L
  - Vortex the solution

Preparation of T3 test solutions:

Solution Name	Final Concentration (µg/L)	Intermediary volume to prepare (mL)	Volume stock solution	Volume Test medium (mL)	Final Volume (mL)
T3_26	26	130.0	1.04 mL of T3 3.25mg/L	129.0	65.0
T3_13	13	130.0	65 mL of T3_26	65.0	55.0
T3_6.5	6.5	150.0	75 mL of T3_13	75.0	65.0
T3_3.25	3.25	170.0	85 mL of T3_6.5	85.0	65.4
T3_2	2	170.0	105 mL of T3_3.25	65.4	57.5
T3_1.5	1.5	150.0	113 mL of T3_2	37.5	63.3
T3_1	1	130.0	87 mL of T3_1.5	43.3	65.0
T3_0.65	0.65	100.0	65 mL of T3_1	35.0	100.0

Store the dilutions at 4°C in the dark. Use the dilutions for the exposure and the two renewals.

Preparation of T4 Solutions

- Place 60 mL of test medium in 100 mL bottle
- Remove and discard 0.75 mL
- Add 0.75 ml of T4 0.8 g/L
- Mix by agitation
- Store at 4°C in the dark

Exposure and renewals

Exposure:

- Prepare the test solutions of T3 as described.
- Place sorted eleutheroembryo into the 6-well plates in test medium with 10 organisms per well using a plastic transfer pipet with the tip cut off to avoid damaging the eleutheroembryo. Note: pay attention to include only healthy-looking eleutheroembryo homogenous in size and pigmentation.
- Begin exposure by carefully aspirating the test medium solution using a plastic transfer pipet. Replace test medium solution with 8 ml of the appropriate solution.

Renewal:

Test solutions are renewed at 24 and 48 ± 1 h

- Take the relevant stock solutions out of the fridge and place them at 21 ± 1°C for at least one hour before the renewal.
- Remove any dead or abnormal eleutheroembryos .
- Aspirate carefully approx. 90% of the medium with a 3 mL pipet.
- Fill each well with 8 ml of the appropriate renewal solution.
- Place the plates in an incubator at 21 ± 1°C in the dark.

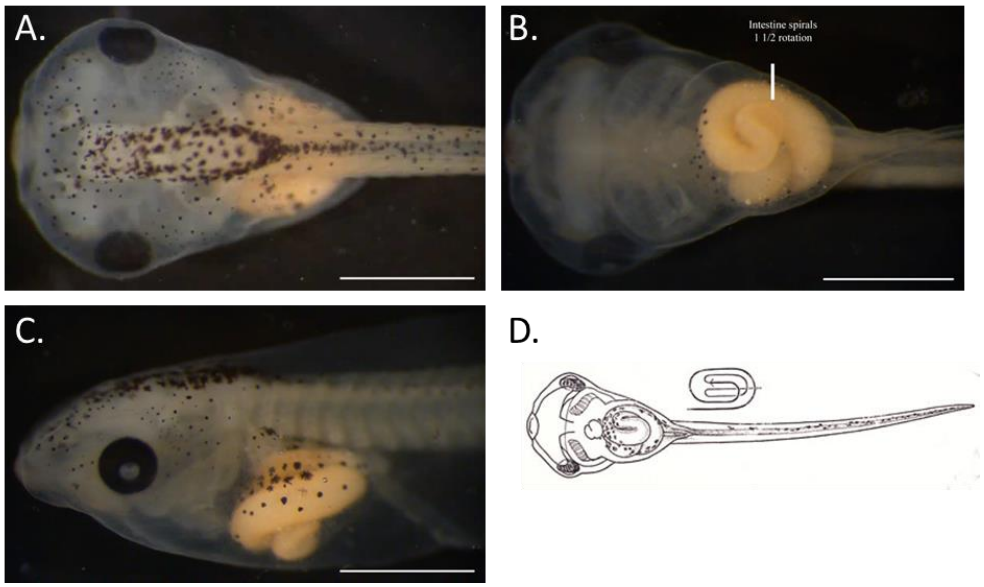
Measuring Fluorescence

After 72 ± 2 h, proceed to the anaesthesia and placement of the organisms in 96 well plates as described in the test guideline and measure the fluorescence. Proceed to data treatment and statistical analysis to obtain a LOEC as described in the test guideline.

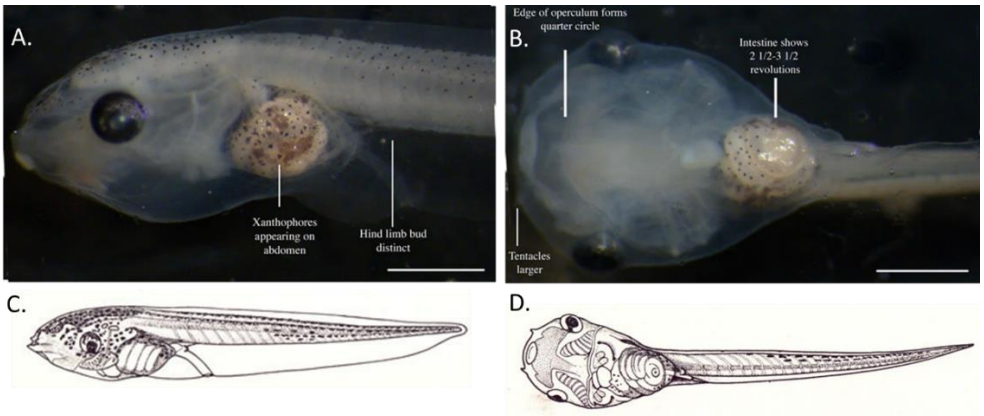
Interpreting the results:

As for the experiment described above the results should show:
- A fluorescence induction between the control and the T3 control group as well as between T3 controls and T4 controls close to the ones observed during the validation (see above).
-That the spectrofluorimeter saturation is never reached in the T4 controls.
The LOEC found for T3 should be 3.25 µg/L or lower.
If the results are not in accordance with these three points, the spectrofluorimeter settings should be refined.

ANNEX 4: STAGING STAGE NF45 AND NF47 ELEUTHEROEMBRYO



Stage NF45 eleutheroembryo : Dorsal view (A), ventral view showing the intestine spirals (B), lateral view (C) and drawing of a ventral view (D). http://wiki.xenbase.org/xenwiki/index.php/Stage_45. Bar =2.5 mm



Stage NF47 eleutheroembryos : At stage NF47, the intestinal spirals reach two and a half to three and a half rotations, xanthophores appears on the abdomen and hind limb buds become distinct. Lateral view showing the xanthophores appearing on abdomen (A), ventral view showing the intestine spirals (B) and drawings of lateral and ventral views (C and D). http://wiki.xenbase.org/xenwiki/index.php/Stage_47

ANNEX 5: 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 the occurrence of dead or abnormal embryos represent less than 20% between the reception of the batch and the start of the exposure.

Guidance for embryos received three days before the start of the XETA:

- **Do not mix embryos from different spawns.**
- 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 XETA.
- Transfer only the living and normal healthy embryos of the same spawn to a 10-liter aquaria containing test medium.
- The maximum density per 10-liter aquarium tank is 800 embryos.
- Incubate embryos without illumination at 21°C ± 1°C (in the dark).
- After the step above, do not renew medium in the tank until the start of the experiment.

ANNEX 6: BREEDING

This section describes how laboratories could breed their own eleutheroembryo.

The materials specifically needed for this procedure include:

- Binocular microscope
- Bleach solution (100%)
- Gentamycin solution
- 2% Cysteine solution, pH 7.5-8
- (des-gly, D-ala, pro-NHEt) LHRH acetate salt, Store at -20 °C
- Human chorionic gonadotrophin (HCG), Store at 4°C
- MMR (Marc's Modified Ringer's) 0.1X solution + Gentamycin 50 µg/L
- Syringes, 5 mL and 1 mL
- Syringe needle (Pink needle); G 18 1 ½;
- Syringe needle (Orange needle); G 25 5/8
- 250 mL beaker to contain the eggs (blastula stage)

To be ready for the following Monday or Tuesday, breeding should begin on the Monday the week prior to the start of the experiment.

Sample Breeding Schedule (beginning on the Monday prior to the start of the experiment)

Day	Activity	Details
Monday prior to start of experiment	Mate male THb/ZIP-GFP and female wild type adult *X. laevis*.	Inject hormones to both sexes (LHRH for male and HCG for female) to induce breeding. Place the *Xenopus* together in the same tank.
Tuesday	De-jelly (dissociate) fertilized eggs with 2% cysteine and sort healthy embryos (at blastula stage) and incubate in the dark at 21± 1 °C.	Keep spawns separate if more than one spawn obtained
Wednesday	Sort healthy embryos and refresh medium with MMR + Gentamycin, Incubate in the dark at 21± 1 °C.	
Thursday	Sort and refresh medium as above, Incubate for the weekend in the dark at 21± 1 °C.	
Following Monday	Start experiment	Eleutheroembryos should be at stage NF45. This should be confirmed by staging the eleutheroembryos.

Preparation of solutions required for the breeding:

MMR (20X) Stock solution
For 2L:

- 234 g NaCl
- 6 g KCl
- 3.8 g $MgCl_2$
- 11.76 g $CaCl_2$
- 48 g HEPES

Place the powders in a 2L glass bottle
Add distilled water QSP 2L
Adjust pH with NaOH 1N to be 7.5-8
Storage at 4°C 2months max

MMR (0.1X) from 20X MMR

For 10 L:

- Pour 50 ml of MMR 20X into a 10 L container
- Fill the container to 10L with distilled water
- Adjust the pH between to 7.5-8 using 1N NaOH
- Storage at room temperature during use(2 weeks max)

MMR (0.1X) + gentamycin 50 µg/L

- Place 2 mL of gentamycin 50 mg/L in a 2 L volumetric flask
- Fill to 2L with MMR 0.1X
- Cover the flask with parafilm
- Mix by inverting several times
- Storage at room temperature during use (2 weeks max)

Cysteine 2%, pH 7.5-8

- For 100 mL
- Weigh 2g of cysteine
- Add 90 mL of MMR 0.1X
- Adjust the pH to 7.5-7.7 using 10N NaOH, then 1N
- Complete to 100 mL with MMR 0.1X
- Use only on the day of the preparation

Preparation of the Gonadotropin-releasing hormone GnRH (Leutenizing hormone-releasing hormone [LHRH] acetate salt) solution:

- Remove the tube of LHRH acetate salt powder (GnRH) provided by the supplier from the -20°C freezer
- Resuspend the powder directly (5mg) in the tube above using 720 µL of sterile water (stock solution: 6.94 µg/µL); Dilute 6 µL of the stock solution in 1 mL of sterile water (final solution 41.64 µg/ml) or for a larger volume, dilute 90 µL of stock solution in 15 mL of sterile water.
- Make 1 mL aliquots in sterile 1.5 mL Eppendorf tubes.
- Store aliquots at -20 ºC.

Preparation of the human chorionic gonadotropin (hCG) solution:

- Obtain a vial of 5000 units of hGC in powder form from the supplier.
- Using a sterile syringe, inject 5 mL of sodium chloride 0.9% into the tube of HCG in order to obtain a concentration of 1 unit per µL.
- Mix by manual agitation.
- Any solution that remains unused can be stored at 4ºC for two weeks maximum.

BREEDING PROCEDURE

Monday*:* Perform hormonal injections to induce the mating of the male and female adult *X. laevis*.
The *X. laevis* male could be injected either with GnRH or with HCG:
Injection of males with GnRH (LHRH acetate):

- For reference, read: p126 &127 of The Laboratory XENOPUS sp. From Sherill L. Green. A subcutaneous injection in dorsal lymph sac can be viewed at http://www.jove.com/index/Details.stp?ID=890 (Cross and Powers 2008)
- Remove an aliquot of GnRH (41.64 µg/mL) from the -20ºC freezer and let it thaw at room temperature.
- To cause minimal distress, immobilize the frog by covering its eyes.
- Using a 25 gauge (orange) needle and a 1ml syringe containing 0.120 mL of the thawed GnRH solution, perform a dorsal subcutaneous injection. In order to assure that the injection volume will be remained completely in the Xenopus the injection needle has to be inserted in parallel to the upper part of the hind leg beneath the skin and then under the skin passing the lateral line organ to reach the dorsal lymph sac. Then the injection can be made and one can easily see that the injection might form a bubble under the skin in the lymph sac. Then after careful removal of the needle the bubble will stay until being absorbed via the lymphatic system without any leaching of the volume injected because the lateral line system is tight to the musculature beneath and thus preventing any loss of the volume injected.

Injection of males with HCG:

- Using a 1 mL syringe and a 25 gauge (orange) needle, aspirate 10-25 units of HCG (depending on the size of the frog).
- Immobilise the frog on a wet paper towel on a bench and mask its eyes as described in the picture below.
- Perform a dorsal subcutaneous injection at the level of the dorsal lymph sac. Follow the same instructions as above for the injection.

Injection of females with HCG:

- Using a 1 mL syringe and a 25 gauge (orange) needle, aspirate 500-700 units of HCG (depending on the size of the frog). *Note: inject 500 units for a frog under 9 cm, and 700 units for a frog 9 cm or above.*
- Immobilize the frog and mask its eyes.
- Perform a dorsal subcutaneous injection at the level of the dorsal lymph sac. Follow the same instructions as above for the injection.

Tuesday*:* (*sort and incubate*)

- Harvest and dejelly the eggs:
    - Harvest the eggs in 250 mL beaker, remove the excess of water
    - Add enough cysteine 2% to cover the eggs, mix the eggs with cysteine by rotating the beaker for 2 or 3 min until eggs are dissociated. It is important to be careful not to leave the cysteine too long, otherwise it will affect the survival. This step is critical and has to be performed with care. Cysteine solution should be acclimated to ~21°C. Temperature impacts the effectiveness of the cysteine treatment and can also affect the health of the embryos if they experience temperature shock.
    - Stop the action of cysteine by adding MMR 0.1X in the beaker (almost to the top).
    - Remove the liquid by pouring it slowly into a beaker of bleach solution and rinse again 3 times with MMR 0.1X.
    - Place the embryos into large 100 mL annotated Petri plates by gently pouring from the beaker.

- Sort the embryos (dead versus alive), most of the embryos should be at the blastula stage:

    - Place the dish under a binocular microscope and using a transfer pipet which tip is cut off to avoid damaging the eggs, isolate each fertilized egg in a new 100 ml Petri dish containing MMR 0.1X + gentamycin 50 µg/L. For description of eggs that can be found in the dish, see Annex 7 : sorting fertilized eggs.

Annotate the Petri dish : line number, mating date, number of eggs per dish. Incubate the Petri dishes in the dark at 21 ± 1°C.

Wednesday*: (sort and change medium)*

- Remove the Petri dishes from incubator.
- Sort the embryos, removing dead ones as on previous day using a transfer pipet with a cut off tip.
- Renew half of the medium, and resupply with MMR 0.1X + gentamycin 50 µg/L.
- Incubate the Petri dishes at 21ºC in the dark at 21 ± 1°C.

Thursday*: (prepare for experiment or ship to another laboratory)*

Option 1: continue the culture until the start of the experiment

- Fill ¼ of an aquarium with MMR 0.1X.
- Annotate the aquarium with line number, date of reproduction.
- Gently pour the embryos into the aquarium without mixing spawns, place 200 embryos/L maximum.
- Place the aquarium in a 21ºC incubator in the dark for the weekend.

Option 2: Ship the embryos to another laboratory

Materials needed for shipping:

- Sterile 100 mL container (e.g., tissue culture flask with screw cap lid)
- Transfer pipets with the tips cut off
- Pipets (50 mL)
- Pipetaid

Please note that the embryos should be packed just before the shipment to reduce the stress.

- Annotate the 100 mL container (both the cap and the container itself) with:
  - « Name of line »
  - « Date of reproduction »
  - « Number of embryos/container »
- Add 50 mL of test medium to the container.
- Place 100 live embryos in the container.
- Close the container only when it is ready to be sent.
- Place the container in a sealable freezer bag.
- Close the bag.
- Place the container in a polystyrene box and firmly secure it inside the box so that it does not move.
- Close and tape the box carefully.
- Mail to receiving laboratory by overnight mail.

ANNEX 7: SOME CHEMICAL CHARACTERISTICS OF AN ACCEPTABLE DILUTION WATER

Alkalinity 10 – 250 mg/L Ca CO_3

Hardness = 75-150 mg/L

pH = 6.5-8.5

Salinity = 0.4 mg/L

Conductivity = 50 -2000 µS/cm

Non-ionised ammonia (NH_3) < 0.02 mg/L

Nitrite (NO_2^-) < 1 mg/L

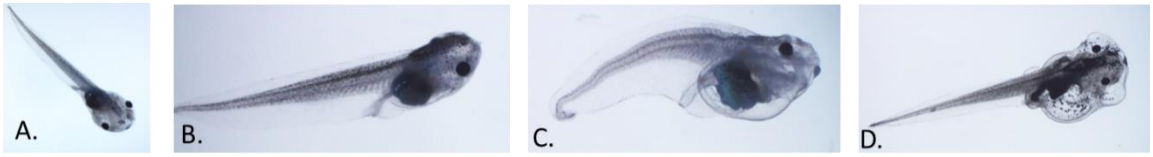
Nitrate (NO_3^-) < 50 mg/L

Residual Chlorine <10 µg/L

Dissolved oxygen content > 80% saturation

CO_2 < 5 mg/L
Total organophosphorous pesticides <50 ng/L
Total organochlorine pesticides plus polychlorinated biphenyls <50 ng/L
Total organic chlorine <25 ng/L
Aluminium, Arsenic, Chromium, Cobalt, Copper, Iron, Lead,Nickel, Zinc. Each < 1 µg/L
Cadmium, Mercury, Silver. Each < 100 ng/L

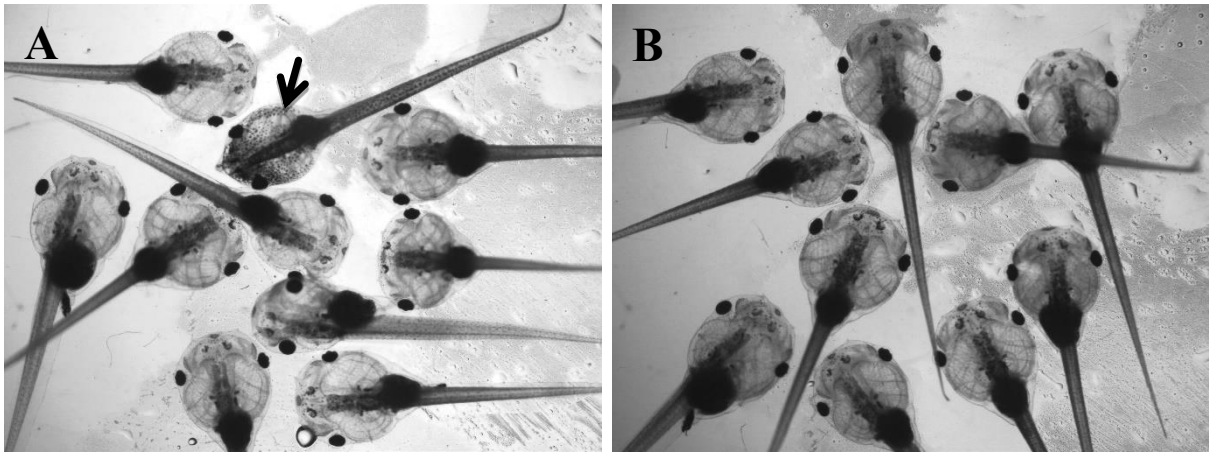
ANNEX 8: PHOTOGRAPHIC GUIDANCE FOR IDENTIFICATION OF NORMAL VERSUS ABNORMAL ELEUTHEROEMBRYO



Normal eleutheroembryo (A). Abnormal eleutheroembryo : pigmentation and microcephaly (B), oedema and several malformations (C and D).

ANNEX 9: SORTING ELEUTHEROEMBRYOS: UNUSUAL PIGMENTATION

Eleutheroembryos presenting unusual appearance of melanophores are often found in *X. laevis* spawns and should be discarded from the test organisms as the pigmentation could interfere with the fluorescence quantification. Such eleutheroembryos should be euthanised using the procedure described in paragraph 44.



Picture A is an example of a heterogeneous clutch of eleutheroembryos in terms of pigmentation. An eleutheroembryo presents an unusual appearance due to melanophores (arrow). Picture B shows an example of a clutch homogenous for pigmentation.

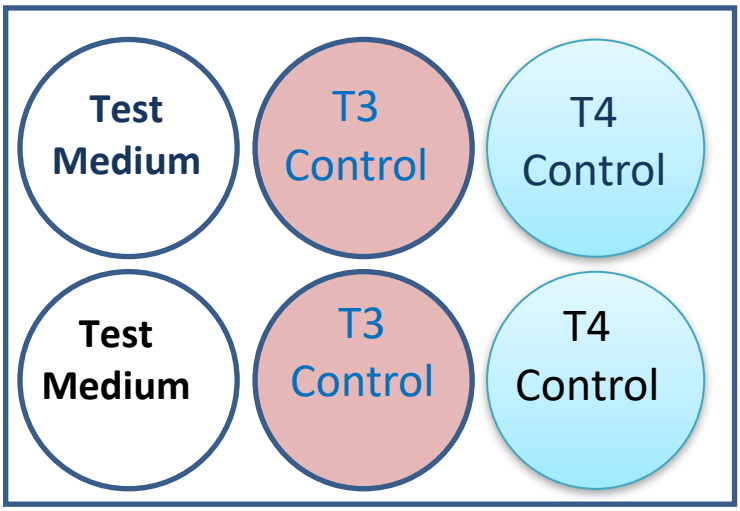
ANNEX 10: SAMPLE SIX-WELL PLATE SET-UP

Sample layout of 6-well plate exposure for one run:

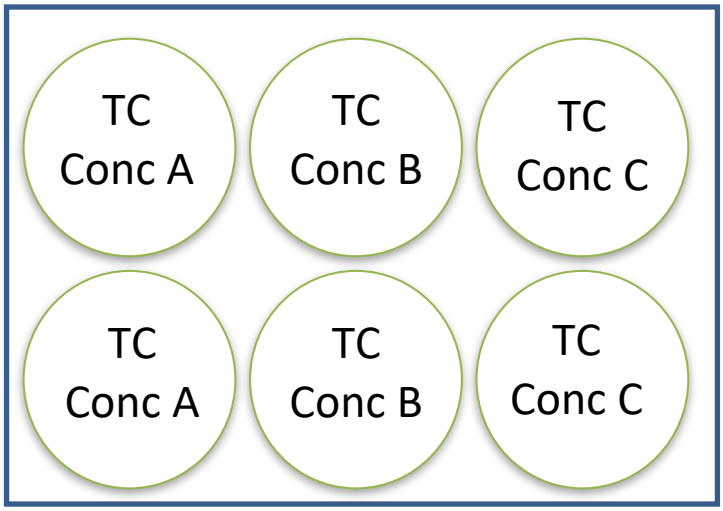
- 10 eleutheroembryos per well
- 3 plates in total are required to test one test chemical

The XETA test design uses 180 eleutheroembryos per run, i.e. 540 for the three runs

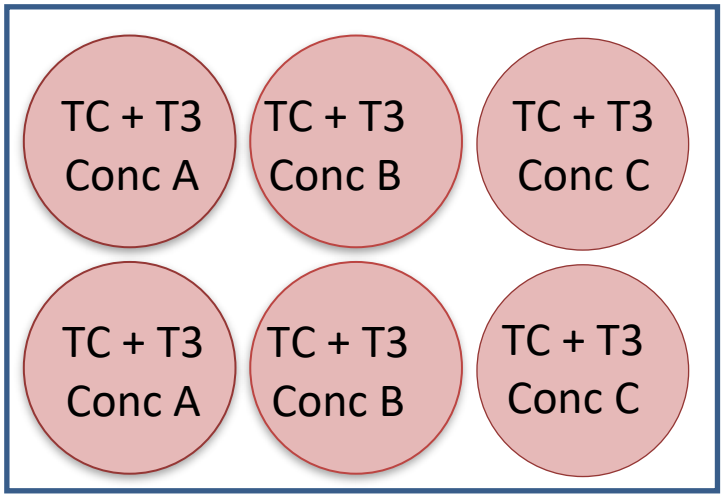
Control Plate (1 plate): Three controls



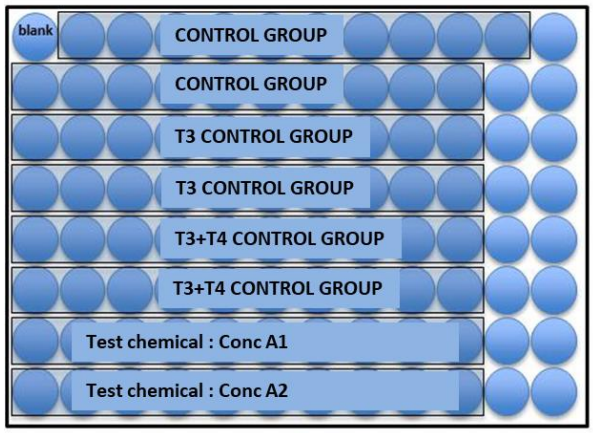
Test Chemical (TC) 1 plate – one chemical, 2 x 3 concentrations)

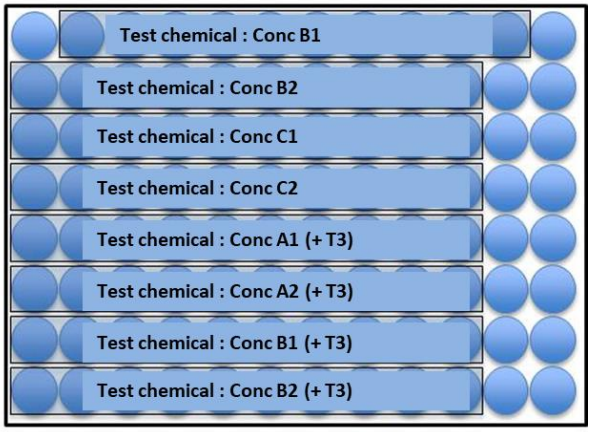


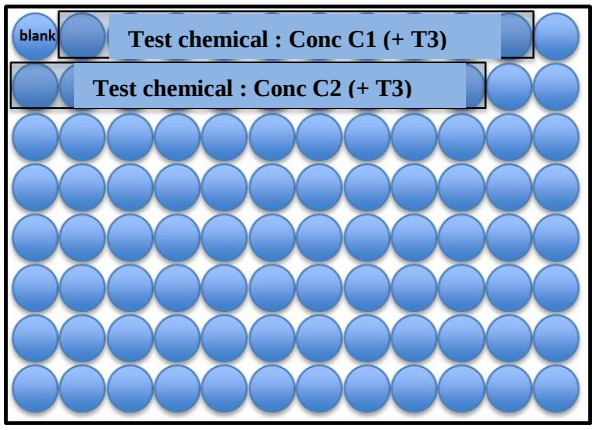
Test Chemical + T3 (1 plate – one chemical + T3, 2 x 3 concentrations)



ANNEX 11: SCHEME OF 96-WELL PLATES FOR FLUORESCENCE READING



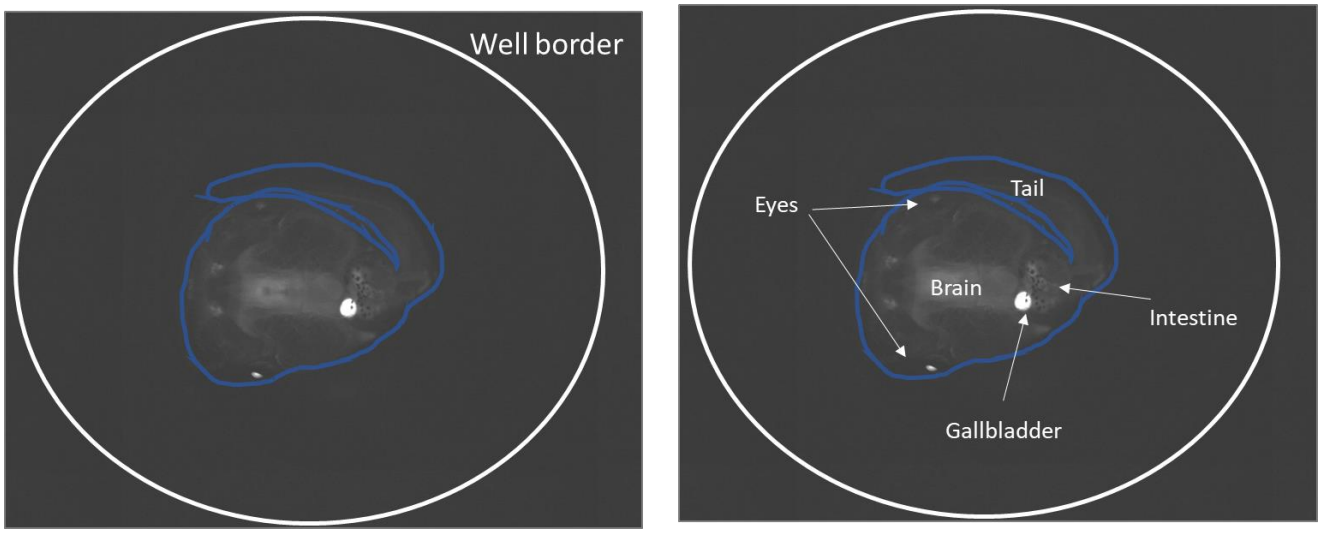




ANNEX 12: ELEUTHEROEMBRYO POSITIONING

The figure below is showing the expected positioning of the eleutheroembryo in the well of the 96-well plate.

Figure 3. Image of a THb/ZIP-GFP eleutheroembryo using GFP fluorescence imaging in a well of a 96-well plate.



Note: The image shows the typical pattern from a spiked T3 control.

ANNEX 13: METHODS FOR THE STATISTICAL ANALYSIS OF XETA DATA

This Annex describes a possible way for statistical analyses of the data obtained in the XETA.

Perform a 10% trim of each test concentration group, omitting the highest 10% and lowest 10% of the fluorescence values from each group. A trim less than 10% may be used based on expert judgement. This step is optional and should be preferably replaced by an image quality check if the fluorescence quantification is done by imagery.

Do a mixed effects ANOVA with one fixed factor corresponding to treatment and random factors corresponding to Run, Run-by-Treatment interaction, and well nested within Run-by-Treatment. Compute the variance of each difference in mean responses between each Treatment and the Control. Use those variances and the pooled degrees of freedom from the ANOVA in an otherwise standard Williams test. A program to carry out these calculations is provided in [Green JW, Springer TA, Holbech H 2018. Statistical Analysis of Ecotoxicity Studies. Wiley. ISBN: 978-1-119-48881-1].

Construct a histogram or Quantile-Quantile (QQ)-plot or stem-and-leaf plot of the residuals and examine them for visual consistency with a normal distribution and outlier identification.

Ideally, a histogram of the residuals should resemble a bell-shaped curve, symmetric around zero and tapering off in both directions. A QQ plot should roughly follow a straight line. Expert judgment can be used to assess whether a histogram or QQ plot is consistent with normality or indicates the presence of one or more outliers. Alternatively, compute the Shapiro-Wilk or Anderson-Darling test for normality at the 0.01 significance level. Be aware that a large dataset may give rise to a significant test, i.e. indicate non-normality, even though there is no need for concern.

Plot residuals vs. treatment to identify possible variance heterogeneity or give further indication of outliers. Alternatively, compute Levene's test (or some alternative) at the 0.01 significance level.

If the plots or formal tests indicate non-normality or variance heterogeneity that cannot be eliminated by omitting a few outliers, then repeat the above steps with a log- or square-root transform of the data.

Once a normalizing, variance stabilizing transformation has been found, do Williams' test for increasing trend and separately for a decreasing trend based on the ANOVA indicated in the third paragraph of this Annex.
Examine the pooled treatment means, if any, from Williams' test. If three or more adjacent test concentration are combined by the PAVA process underlying Williams' test, examine the (un-amalgamated) treatment means for consistency with a monotone concentration-response. If the data are not consistent with monotonicity, then Williams' test is not appropriate and Dunnett's test should be used. Note: If Dunnett's test is used, then the indicated ANOVA should include Dunnett's test to compare treatments to control. A two-sided test should be used unless there is scientific justification to expect only a change in one direction. Note: Dunnett's test is done at the 0.05 significance level regardless of the ANOVA F-test.

Use the Tukey outlier rule to identify outliers from the ANOVA. If outliers are found and there is evidence of non-normality or variance heterogeneity, repeat the ANOVA and Dunnett test with outliers omitted. If the normality and variance heterogeneity issues disappear and Dunnett's test identifies the same NOEC as

with the full data, then accept the NOEC. Note: William's test by default establishes a NOEC, this NOEC has not to be considered as the result from the XETA but is to be used for identifying active concentrations and classify the test chemicals into potentially thyroid active or inactive (see paragraphs 45-52).
Figure 4 captures the broad outline of this flow chart.

Figure 4. Flow Chart for Fluorescence Statistical Analysis.