Biological Test Method: Toxicity Tests Using Aquatic Life Stages of Frogs (*Lithobates pipiens*)

Method Development and Applications Unit
Science and Technology Branch
Environment and Climate Change Canada
Ottawa, Ontario

Report
STB 1/RM/62
April 2024

Environment and Climate Change Canada
Environnement et Changement climatique Canada

Canada

Cat. No.: En83-10/1-62-2024E-PDF
ISBN: 978-0-660-70720-4
EC23235

Unless otherwise specified, you may not reproduce materials in this publication, in whole or in part, for the purposes of commercial redistribution without prior written permission from Environment and Climate Change Canada's copyright administrator. To obtain permission to reproduce Government of Canada materials for commercial purposes, apply for Crown Copyright Clearance by contacting:

Environment and Climate Change Canada
Public Information Centre
Place Vincent Massey building
351 St-Joseph boulevard
Gatineau Quebec K1A 0H3
Toll free: 1-800-668-6767
Email: enviroinfo@ec.gc.ca

Cover photos: © Environment and Climate Change Canada

© His Majesty the King in Right of Canada, represented
by the Minister of Environment and Climate Change, 2024

Aussi disponible en français

Biological Test Method:
Toxicity Tests Using Aquatic Life Stages of Frogs
(*Lithobates pipiens*)

Method Development and Applications Unit
Science and Technology Branch
Environment and Climate Change Canada
Ottawa, Ontario

Report STB 1/RM/62
April 2024

Readers' Comments

Comments regarding the content of this report should be addressed to:

Leana Van der Vliet, Manager
Method Development and Applications Unit
Biological Assessment and Standardization Section
Science and Technology Branch
Environment and Climate Change Canada
335 River Road
Ottawa, Ontario
K1A 0H3

General inquiries regarding this method can be addressed to:
methods@ec.gc.ca

Review Notice

This report has been reviewed by staff at the Science and Technology Branch, Environment and Climate Change Canada, and approved for publication. Mention of trade names or commercial products does not constitute endorsement by Environment and Climate Change Canada for use. Other products of similar value are available.

Abstract

This document provides procedures, conditions, and guidance on preparing for and conducting a biological test for measuring aquatic toxicity using the larval stage of the northern leopard frog (*Lithobates pipiens* [formerly *Rana pipiens*]).

Methods are given for two test options:

i) a 14-day test to assess the survival and growth of tadpoles, using test organisms that have recently hatched (GS 25); and

ii) a 42-day test to assess the survival, growth, and development of tadpoles, using test organisms that have just begun metamorphosis (GS 28/29).

Each test option is conducted as a static-renewal test, using samples of contaminated water or one or more concentrations of chemical(s) or chemical product(s) spiked in clean dilution water. The tests are conducted at a mean temperature of 23 ± 2 °C in glass aquaria, or other suitable vessels, containing a minimum of 7 L of test solution. The tests are initiated by placing 10 test organisms into each replicate vessel containing test solution or clean dilution water. The initial life stage of the organisms and the number of replicates prepared for each treatment depends on the method chosen and on specific objectives of the test. Observations of mortality, abnormal appearance or behaviour, and approximate developmental stage are recorded daily. At the end of each of the test options individual total length, wet weight, biomass, stage of development, and number of deformities are measured or calculated. The growth measurements are corrected for initial measurements before statistical endpoints are calculated. The mean of the replicates for each treatment is calculated and the percentage effect concentrations estimated for mortality and inhibition of growth (e.g., ICp). Changes in development for each treatment are calculated and compared with the control.

General or universal conditions and procedures are outlined for test preparation and performance. Additional conditions and procedures specific to the intended use of each test are stipulated. The biological test method described herein is suitable for measuring and assessing the toxicity of chemicals or contaminated waters. Instructions and requirements are included on culturing *L. pipiens* in the laboratory, facilities and water supply, handling and storage of samples, preparation of solutions, test conditions, observations to be made, endpoints with methods of calculations, and the use of positive control replicates or a reference toxicity test.

Foreword

This is one of a series of **recommended methods** for measuring and assessing the toxic effect(s) on single species of aquatic or terrestrial organisms caused by their exposure to samples of toxic or potentially toxic substances or materials under controlled and defined laboratory conditions. Recommended methods are those that have been evaluated by Environment and Climate Change Canada (previously Environment Canada) and are favoured:

- for use in Environment and Climate Change Canada environmental toxicity laboratories;
- for testing that is contracted out by Environment and Climate Change Canada or requested from outside agencies or industry;
- in the absence of more specific instructions, such as are contained in regulations; and
- as a foundation for the provision of very explicit instructions as might be required in a regulatory *protocol* or standard *reference method*.

The different types of tests included in this series were selected because of their acceptability for the needs of environmental protection and management programs carried out by Environment and Climate Change Canada. These reports are intended to provide guidance and to facilitate the use of consistent, appropriate, and comprehensive procedures for obtaining data on the toxicity to aquatic or terrestrial life of samples of specific test substances or materials destined for or within the environment. Depending on the biological test method(s) chosen and the environmental compartment of concern, substances or materials to be tested for toxicity could include samples of chemical or chemical product, effluent, elutriate, leachate, receiving water, sediment or similar particulate material, or soil or similar particulate material. Appendix A lists the biological test methods and supporting guidance documents published to date by Environment and Climate Change Canada as part of this series.

Words defined in the Terminology section of this document are italicized when first used in the body of the report according to the definition.

Table of Contents

Abstract ii
Foreword iii
List of Tables vii
List of Figures vii
List of Abbreviations and Chemical Formulae viii
Terminology ix
Acknowledgements xix

Section 1
Introduction 1
1.1 Background 1
1.2 Identification, Life History, and Distribution 3
1.3 Sensitivity of Frogs in Toxicity Tests 7
1.3.1 Comparison with Fish 7
1.3.2 Comparison Among Aquatic Life Stages 9
1.4 Use of Frogs in Toxicity Tests 10

Section 2
Test Organisms 14
2.1 Species and Life Stage 14
2.2 Source 19
2.3 Holding and Culturing of Embryos and Larvae 21
2.3.1 General 21
2.3.2 Facilities and Apparatus 22
2.3.3 Lighting 24
2.3.4 Temperature 24
2.3.5 Culture Water 25
2.3.6 Food and Feeding 28
2.3.7 Handling Organisms and Maintaining Cultures 29
2.3.8 Test Organism Health Criteria 29
2.3.9 Health, Quarantine, and Disease of Embryos and Larvae 30
2.4 Holding and Breeding of Adult Frogs 30
2.4.1 General 30
2.4.2 Holding of Adult Frogs 31
2.4.3 Hibernation and Breeding 34
2.4.4 Adult Frog Health, Quarantine, and Disease 34

Section 3
Test System 36
3.1 Facilities and Apparatus 36
3.2 Lighting 37
3.3 Preliminary and Definitive Tests 37
3.3.1 Preliminary Tests 37
3.3.2 Definitive Tests 37
3.4 Control/Dilution Water 38
3.5 Positive Control 39

Section 4
Universal Test Procedures 40
4.1 Preparing Test Solutions..41
4.2 Beginning the Test 46
4.3 Test Conditions and Procedures 47
4.3.1 Test Options..47
4.3.2 Test Type and Solution Renewal..47
4.3.3 Test Temperature and Lighting 48
4.3.4 Dissolved Oxygen and Aeration..48
4.3.5 pH 50
4.3.6 Food and Feeding 50
4.4 Test Observations and Measurements 51
4.5 Ending the Test..53
4.6 Test Endpoints and Calculations 55
4.6.1 Biological Endpoints 55
4.6.2 Multi-Concentration Tests..57
4.6.2.1 LCp..58
4.6.2.2 ICp..58
4.6.2.3 Development Endpoint..62
4.6.2.4 Power Analysis..62
4.6.3 Other Test Designs 63
4.6.3.1 Test Duration..64
4.6.3.2 Suborganismal or Non-apical Endpoints..66
4.7 Test Validity 67
4.8 Tests with a Reference Toxicant 68
4.8.1 Multi-concentration Sublethal Tests with a Reference Toxicant..70
4.8.2 Positive Control..73
4.8.3 Multi-concentration Acute Lethality Tests with a Reference Toxicant..74
4.8.4 Warning Charts..75
4.9 Animal Conservation and Welfare Considerations 76

Section 5
Specific Procedures for Testing Chemicals 78
5.1 Sample Properties, Labelling, and Storage 78
5.2 Preparing Test Solutions 78
5.3 Control/Dilution Water..79
5.4 Test Observations and Measurements 80
5.5 Test Endpoints and Calculations 81

Section 6
Specific Procedures for Testing Contaminated Water Samples 82
6.1 Sample Collection, Labelling, Transport, and Storage 83
6.2 Preparing Test Solutions 84
6.3 Control/Dilution Water..84
6.4 Test Observations and Measurements 85
6.5 Test Endpoints and Calculations 85
6.5.1 Variations in Design and Analysis 85

Section 7
Reporting Requirements..88
7.1 Minimum Requirements for a Test-specific Report..88

7.1.1 Test Substance or Material 88
7.1.2 Test Organisms 88
7.1.3 Test Facilities and Apparatus 89
7.1.4 Control/Dilution Water 89
7.1.5 Test Method 89
7.1.6 Test Conditions and Procedures 89
7.1.7 Test Results 90
7.2 Additional Reporting Requirements 90
7.2.1 Test Substance or Material 91
7.2.2 Test Organisms 91
7.2.3 Test Facilities and Apparatus 91
7.2.4 Control/Dilution Water 91
7.2.5 Test Method 92
7.2.6 Test Conditions and Procedures 92
7.2.7 Test Results 92

References 93

Appendix A
Biological Test Methods and Supporting Guidance Documents Published by Environment and Climate Change Canada's Method Development and Applications Unit 111

Appendix B
Environment and Climate Change Canada, Regional Environmental Testing Laboratories 114

Appendix C
Members of the Inter-Governmental Ecotoxicological Testing Group (as of June 2023) 115

Appendix D
Method Writers and Members of the Expert Peer Review Committee 118

Appendix E
Additional Procedures for Culturing *Lithobates pipiens* in the Laboratory and General Procedures for Frog Husbandry 119

Appendix F
Procedures for In-Laboratory Hibernation and Breeding for *Lithobates pipiens* Adults....124

Appendix G
Logarithmic Series of Concentrations Suitable for Toxicity Tests 130

List of Tables

1 Approximate timing of developmental stages of *Lithobates pipiens* at test temperature, based on ASTM (2022a) and data from inter-laboratory studies (Nautilus Environmental, 2020a, 2020b) 18

2 Checklist of required and recommended conditions and procedures for culturing *Lithobates pipiens* embryos and larvae for use in aquatic toxicity tests 23

3 Checklist of required and recommended conditions and procedures for quarantine, holding, and breeding *Lithobates pipiens* adult frogs to provide test organisms for use in aquatic toxicity tests 32

4 Checklist of required and recommended conditions and procedures for conducting toxicity tests using aquatic life stages of frogs 42

5 Example feeding rations for toxicity tests....51

6 Checklist of required and recommended conditions and procedures for conducting reference toxicity tests using aquatic life stages of frogs....71

List of Figures

1 Considerations for preparing and performing toxicity tests using aquatic life stages of frogs (*Lithobates pipiens*) and various types of test materials or substances....4

2.1 Developmental stages (GS 1 to GS 25) of anuran embryos and hatchlings as described by Gosner (1960) 16

2.2 Developmental stages (GS 26 to GS 46) of anuran larvae and metamorphs as described by Gosner (1960) 17

3 The general process for the statistical analysis and selection of the most appropriate model for quantitative toxicity data....60

List of Abbreviations and Chemical Formulae

ANOVA	analysis of variance
$CaCl_2$	calcium chloride
$CaCO_3$	calcium carbonate
C	carbon
Ca	calcium
Cl	chlorine
cm	centimetre(s)
CV	coefficient of variation
°C	degree(s) Celsius
d	day(s)
DEPC	diethyl pyrocarbonate
DO	dissolved oxygen
ERA	ecological risk assessment
g	gram(s)
GnRH-A	gonadotropin-releasing hormone agonist
GS	Gosner stage
h	hour(s)
H	hydrogen
HCl	hydrochloric acid
H&E	hematoxylin and eosin
HNO_3	nitric acid
HPG	hypothalamic-pituitary-gonadal
HPLC	high-performance liquid chromatography
HPT	hypothalamic-pituitary-thyroid
ICp	inhibiting concentration for a (specified) percent effect (e.g., IC25)
ICP-MS	inductively-coupled plasma – mass spectrometry
K	potassium
KCl	potassium chloride
kg	kilogram(s)
L	litre(s)
LC50	median lethal concentration
LCp	lethal concentration for a (specified) percent effect (e.g., LC25)
LC-MS	liquid chromatography – mass spectrometry
LED	light-emitting diode
LOEC	lowest-observed-effect concentration
m	metre(s)
M	mole(s) (concentration)
mM	millimole(s)
MET	metoclopramide hydrochloride
Mg	magnesium
$MgCl_2$	magnesium chloride
mg	milligram(s)
mL	millilitre(s)
mm	millimetre(s)
mS	millisiemen(s)
n	sample size
N	nitrogen
N	normality
Na	sodium
NaCl	sodium chloride
$NaHCO_3$	sodium bicarbonate
NaOH	sodium hydroxide
NH_3	ammonia
NO_2	nitrite
NOEC	no-observed-effect concentration
p	probability
QA/QC	quality assurance/quality control
SD	standard deviation
SVL	snout-vent length
s	second(s)
spp.	species (plural)
t	time
T4	thyroxine
TOC	total organic carbon
TRC	total residual chlorine
™	trademark
v:v	volume-to-volume
α	alpha, denotes Type I error
β	beta, denotes Type II error
µg	microgram(s)
µm	micrometre(s)
µmhos	micromhos
µmol	micromole(s)
®	registered trademark
$>$	greater than
$<$	less than
$\geq$	greater than or equal to
$\leq$	less than or equal to
%	percentage or percent
=	equals
+	plus
−	minus
±	plus or minus
×	times
÷	divided by
/	per; alternatively, “or” (e.g., GS 28/29)
≈	approximately equal to
~	approximately

Terminology

Note: All definitions are given in the context of the procedures in this report, and might not be appropriate in another context.

Grammatical Terms

Must is used to express an absolute requirement.

Should is used to state that the specified condition or procedure is recommended and ought to be met if possible.

May is used to mean "is (are) allowed to."

Can is used to mean "is (are) able to."

Might is used to express the possibility that something could exist or happen.

Technical Terms

Acclimation is physiological adjustment to a particular level of one or more environmental factors such as temperature. The term usually refers to the adjustment to controlled laboratory conditions.

Adult (*frog*) is a frog that is sexually mature. (See also *juvenile, metamorph, tadpole, larva, hatchling,* and *embryo*.)

Amplexus (Latin "embrace") is a copulatory embrace of anuran frogs in which the male grasps the female with his front legs, during which the male fertilizes the eggs being released by the female.

Batch means a single group of tadpoles taken from one (i.e., produced by a single male and female) or several egg masses, received from a supplier, or field-collected at a discrete time, in order to provide all of the test organisms intended for use in a discrete toxicity test (including any associated reference toxicity test). The larvae in a batch are normally derived from a single egg mass; however, if insufficient test organisms are available in a single egg mass, then a batch of test organisms may be made up of tadpoles from more than one egg mass. The term *batch* may also refer to a single group of adult frogs received from a supplier, or field-collected at a discrete time, in order to provide test organisms (i.e., to be spawned or cultured in the laboratory) intended for use in a discrete toxicity test (including any associated reference toxicity test).

Compliance means in accordance with governmental regulations or requirements for issuing a permit.

Conductivity is a numerical expression of the ability of an aqueous solution to carry an electric current. This ability depends on the concentrations of ions in solution, their valence and mobility, and on the solution's temperature. Conductivity is measured at 25 °C, and is reported in the SI unit of as millisiemens per metre (mS/m) or as micromhos per centimetre (µmhos/cm); 1 mS/m = 10 µmhos/cm.

Culture, as a noun, means the stock of organisms reared in the laboratory under defined and controlled conditions, to produce healthy test organisms. As a verb, it means to carry out the procedure of rearing healthy test organisms, under defined and controlled conditions.

Dispersant means a chemical substance that reduces the surface tension between water and a hydrophobic substance (e.g., oil), thereby facilitating the dispersal of the hydrophobic substance or material throughout the water as an emulsion.

Ecological risk assessment (ERA) is the process of evaluating the potential adverse effects on non-human organisms, populations, or communities in response to human-induced stressors. ERA entails the application of a formal framework, analytical process, or model to estimate the effects of human actions on natural organisms, populations, or communities, and interprets the significance of those effects in light of the uncertainties identified in each study component (FCSAP, 2019).

Egg is an encapsulated, spherical ovum, unfertilized or fertilized.

Egg mass is a cluster or large group of eggs laid at one time. Most amphibians lay their eggs in clusters or strings, often with several gelatinous envelopes surrounding them. *Lithobates pipiens* females lay their eggs in large spherical or elliptical clusters, and only produce one cluster of eggs during each breeding season; each egg mass typically contains several hundred to more than a thousand eggs, but may contain up to 7000 eggs (Dewey, 1999; Kendell, 2002; COSEWIC, 2009; Ontario Nature, 2016; Canadian Herpetological Society, 2020). For the purpose of this method, an egg mass is a single globular cluster of eggs (fertilized or unfertilized) laid by a female frog.

Embryo is an individual in any stage of development from fertilization until hatching (McDiarmid and Altig, 1999). In this method, it is used to denote the stages between fertilization of the egg and Gosner stage 20 (hatching of tadpoles). (See also: *hatchling, larva, tadpole, metamorph, juvenile,* and *adult).*

Emulsifier is a chemical substance that aids the fine mixing (in the form of small droplets) within the water, of an otherwise hydrophobic material.

Flocculation is the formation of a light, loose precipitate (i.e., a floc) from a solution.

Gametes are the eggs or sperm released from mature adult frogs during *spawning*.

Gosner stage (GS) is a staging system proposed by Gosner (1960) for recognizing certain morphological landmarks that are useful in comparing the sequence of events in the development of frogs from eggs to adults. This staging system allows for inter-species comparisons of organisms that are widely disparate in size and developmental period (McDiarmid and Altig, 1999). All references to developmental stages (GS) in this document refer to those of Gosner (1960), unless otherwise indicated.

Hatchling is an individual within a series of stages after embryo, but before tadpole (Gosner stages 21 to 24), used to distinguish individuals in these ecologically unique developmental stages from an embryo to a tadpole (McDiarmid and Altig, 1999). (See also*: embryo, tadpole, larva, metamorph, juvenile,* and *adult*.)

Hardness is the concentration of cations in water that will react with a sodium soap to precipitate an insoluble residue. In general, hardness is a measure of the concentration of calcium and magnesium ions in water, and is expressed as mg/L calcium carbonate or equivalent.

Hormesis is an observed stimulation of performance (e.g., growth) among organisms, compared to the control organisms, at low concentrations in a toxicity test.

Juvenile is a postmetamorphic frog up to the time of attainment of sexual maturity (McDiarmid and Altig, 1999). (See also *embryo, hatchling, larva, tadpole, metamorph,* and *adult.*)

Larva (plural, larvae) is a recently hatched organism that has physical characteristics other than those seen in the adult. For anuran frogs, the larval period begins with hatching of the embryo (Gosner stages 20) and lasts until metamorphosis is complete (Gosner stage 46). For anuran frogs, and for the purpose of this document, the term *larva* is synonymous with *tadpole*. (See also *embryo*, *hatchling, tadpole, metamorph, juvenile*, and *adult*.)

Lux is a unit of illumination based on units per square metre. One lux = 0.0929 foot-candles and one foot-candle = 10.76 lux. For conversion of lux to quantal flux [μmol/(m^2 · s)], the spectral quality of the light source must be known. Light conditions or irradiance are properly described in terms of quantal flux (photon fluence rate) in the photosynthetically effective wavelength range of approximately 400−700 nm. The relationship between quantal flux and lux or foot-candles is highly variable and depends on the light source, the light meter used, the geometrical arrangement, and the possibilities of reflections (see ASTM, 2022b). Approximate conversions between quantal flux and lux, however, are:

- for cool-white fluorescent light: 1 lux ≈ 0.014 μmol/(m^2 · s);
- for full-spectrum fluorescent light (e.g., Vita-Lux® by Duro-Test®): 1 lux ≈ 0.016 μmol/(m^2 · s); and
- for incandescent light: 1 lux ≈ 0.019 μmol/(m^2 · s) (Deitzer, 1994; Sager and McFarlane, 1997).

Metamorph is a tadpole that is undergoing the final transformation into juvenile frogs. For anuran frogs, the metamorph period begins with the emergence of forelimbs (Gosner stage 42) and lasts until the tail is reabsorbed and the juvenile frogs emerge from their aquatic habitat (Gosner stage 46). (See also *tadpole, juvenile,* and *adult*.)

Monitoring is the routine (e.g., daily, weekly, monthly, quarterly) checking of quality or collection and reporting of information. In the context of this report, it means either the periodic (routine) checking and measurement of certain biological or water-quality variables, or the collection and testing of samples of effluent, elutriate, leachate, or receiving water for toxicity.

Nuptial pads are a secondary sex characteristic present on some mature male frogs. Triggered by androgen hormones, this breeding pad appears as a spiked epithelial swelling on the forearm that aids with grip, used primarily by males to grasp females during amplexus (Kouba et al, 2012).

Operculum is the completed covering of the gill chamber of a tadpole provided by the development of the opercular fold (outgrowth from the hyoid arch in GS stages 20 to 24 that eventually forms a covering over the gills and associated structures and fuses with the body wall in patterns that produce the spiracle) (McDiarmid and Altig, 1999).

pH is the negative logarithm of the activity of hydrogen ions in gram equivalents per litre. The pH value expresses the degree or intensity of both acidic and alkaline reactions on a scale from 0−14, with 7 representing neutrality, numbers < 7 indicating increasingly greater acidic reactions, and numbers > 7 indicating increasingly basic or alkaline reactions.

Photoperiod is the duration of illumination and darkness within a 24-hour period.

Pollution is the addition of a substance or material, or a form of energy such as heat, to some component of the environment, in such an amount as to cause a discernible change that is deleterious to some organism(s) or to some human use of the environment. Some national and international agencies have formal definitions of pollution, which should be honoured in the appropriate contexts.

Precipitation means the formulation of a solid (i.e., precipitate) from some or all of the dissolved components of a solution.

Pretreatment means treatment of a sample or dilution thereof, before exposure of the test organisms.

Protocol is an explicit set of procedures for a test, formally agreed upon by the parties involved, and described precisely in a written document.

Quality assurance (QA) is a program within a laboratory, intended to provide precise and accurate results in scientific and technical work. It includes selection of proper procedures, sample collection, selection of limits, evaluation of data, *quality control*, and qualifications and training of personnel.

Quality control (QC) consists of specific actions within the program of *quality assurance*. It includes standardization, calibration, replication, control samples and statistical estimates of limits for the data.

Reference method refers to a specific *protocol* for performing a toxicity test, i.e., a biological test method with an explicit set of test procedures and conditions, formally agreed upon by the parties involved and described precisely in a written document. Unlike other multi-purpose (generic) biological test methods published by Environment and Climate Change Canada, the use of a reference method is frequently restricted to testing requirements associated with specific regulations.

Risk is the probability or likelihood that an adverse effect will occur.

Spawning means the release of eggs or sperm from mature adult frogs, or refers to behaviour related to the readiness of frogs to release gametes.

Spiracle is one of two openings of different shapes and positions for the exit of the water pumped into through the mouth and throat for respiration and feeding (McDiarmid and Altig, 1999).

Tadpole is a nonreproductive exotrophic larva of a frog between Gosner stages 25 and 41 (McDiarmid and Altig, 1999). (See *embryo, hatchling, tadpole, larva, metamorph, juvenile,* and *adult*.)

Turbidity is the extent to which the clarity of water has been reduced by the presence of suspended or other matter that causes light to be scattered and absorbed rather than transmitted in straight lines through the sample. It is generally expressed in terms of Nephelometric Turbidity Units.

Terms for Test Materials or Substances

Chemical is, in this report, any element, compound, formulation, or mixture of a substance that might be mixed with or found in association with water.

Control is a treatment in an investigation or study that duplicates all the conditions and factors that might affect results, except the specific condition being studied. In toxicity tests, the control must duplicate all the conditions of the exposure treatment(s), but must contain no contaminated test material or substance. The control is used as a check for the absence of measurable toxicity due to basic test conditions (e.g., quality of the dilution water, health of test organisms or effects due to their handling). Control is synonymous with *negative control*, unless indicated otherwise.

Control/dilution water means the water used for diluting the test material or substance, and for the *control* of a test.

Dechlorinated water means a chlorinated water (usually municipal drinking water) that has been treated to remove chlorine and chlorinated compounds from solution.

Dechloraminated water means a chloraminated water (usually municipal drinking water) that has been treated to remove chloramine and chloraminated compounds from solution.

Definitive (toxicity test) means decisive (as opposed to a preliminary, *range-finding test*). [See also *range-finding (test)*.]

Deionized water is water that has been purified by passing it through resin columns or a reverse osmosis system, for the purpose of removing ions such as Ca^{++} and Mg^{++}.

Dilution water is the water used to dilute a test substance or material in order to prepare different concentrations for the various toxicity test treatments.

Distilled water is water that has been passed through a distillation apparatus of borosilicate glass or other material, to remove impurities.

Effluent is any liquid waste (e.g., industrial, municipal) discharged to the aquatic environment.

Elutriate is an aqueous solution obtained after adding water to a solid material (e.g., sediment tailings, drilling mud, dredge spoil), shaking the mixture, then centrifuging or filtering it or decanting the supernatant.

Leachate is water or wastewater that has percolated through a column of soil or solid waste within the environment.

Material is the *substance* or substances from which something is made. A material would have more or less uniform characteristics. Effluent, leachate, elutriate, or surface water are materials. Usually, the material would contain several or many substances.

Negative control – see *control*.

Preliminary test refers to the two or more tests performed in a laboratory, using a reference toxicant, prior to performing *definitive toxicity tests*. Results from these tests can be used to demonstrate the laboratory's capability to meet test validity criteria and to obtain consistent toxicity results. The results can also be used to establish the reference toxicant concentration(s) to be used as positive control treatment(s) in conjunction with definitive toxicity tests. [See also *definitive (toxicity test)*.]

Product is a commercial formulation of one or more chemicals. (See also *chemical*.)

Range-finding (test) means a preliminary toxicity test, performed to provide an initial indication of the toxicity of the test material under defined conditions and to assist in choosing the range of concentrations to be used in a definitive multi-concentration test. [See also *definitive (toxicity test)*.]

Receiving water is surface water (e.g., in a stream, river, or lake) that has received a discharged waste, or else is about to receive such a waste (e.g., it is just upstream from the discharge point). Further descriptive information must be provided to indicate which meaning is intended.

Reconstituted water is deionized or glass-distilled water to which reagent-grade chemicals have been added. The resultant synthetic fresh water is free from contaminants and has the desired pH and hardness characteristics.

Reference toxicant is a standard chemical used to measure the sensitivity of the test organisms to establish confidence in the toxicity data obtained for a test material or substance. In most instances, a toxicity test with a reference toxicant is performed to assess the sensitivity of the organisms at the time the test material or substance is evaluated, and the precision and reliability of results obtained by the laboratory for that chemical.

Reference toxicity test is a test conducted using a *reference toxicant* in conjunction with a definitive toxicity test, to appraise the sensitivity of the organisms and the precision and reliability of results obtained by the laboratory for that chemical at the time the test material or substance is evaluated. Deviations outside an established normal range indicate that the sensitivity of the test organisms, and the performance and precision of the test are suspect and should be investigated as to the cause. A reference toxicity test with is performed with a standard chemical.

Solvent control is a sample of control/dilution water in the testing of insoluble chemicals, in which a solvent is required to solubilize the test chemical before testing it in water. The amount of solvent used when preparing the solvent control must contain the same concentration of solubilizing agent as that present in the highest concentration of the test chemical(s) to be tested. This concentration of solvent should not adversely affect the performance of test organisms during the test. Any test that uses a solvent other than water when preparing one or more concentrations of the test chemical must include a solvent control in the test.

Stock solution is a concentrated aqueous solution of the *substance* or *material* to be tested. The *substance* or *material* may be dissolved in water (e.g., *dilution water* or *deionized water*) and/or in solvent (see also *solvent control*). Measured volumes of a stock solution are added to *dilution water* in order to prepare the required strength of test solutions.

Substance is a particular kind of material having more or less uniform properties. The word substance has a narrower scope than *material*, and might refer to a particular chemical (e.g., an element) or chemical product.

Upstream (*water*) is surface water (e.g., in a stream, river, or lake) that is not influenced by the effluent (or other test material or substance), by virtue of being removed from it in a direction against the current or sufficiently far across the current.

Wastewater is a general term that includes effluents, leachates, and elutriates.

Statistical and Toxicological Terms

A priori literally refers to something that is independent of experience. In the context of test design and statistics, a priori tests are ones that have been planned before the data were collected. Test objectives and test design would influence the decisions on which a priori tests to select.

Acute means within a short period (seconds, minutes, hours, or a few days) in relation to the life span of the test organism and is generally used to describe the length of a test or exposure duration.

Acute toxicity is a discernible adverse effect (lethal or sublethal) induced in the test organisms within a short period (usually a few days) of exposure to *test solution(s).*

Bioassay is a test in which the strength or potency of a substance is measured by the response of living organisms. In standard pharmacological usage, a bioassay assesses the unknown potency of a given preparation of a drug, compared to the known potency of a standard preparation. *Toxicity test* is a more specific and preferred term for environmental studies.

Biomass means the total wet weight of living *L. pipiens* in a replicate, divided by the number of organisms that started in the replicate (typically 10). The biomass endpoint represents a combination of sublethal effect and mortality.

Chronic means occurring during a relatively long period of exposure (weeks, months, or years), usually a significant portion of the life span of the organism, and is generally used to describe the length of a test or exposure duration.

Chronic toxicity refers to discernible adverse effects observed during or after relatively long-term exposures to one or more contaminants, which are related to changes in growth, development, survival, or other biological variables (e.g., behaviour) being observed.

Coefficient of Variation (CV) is the standard deviation (SD) of a set of data divided by the mean of the data set, expressed as a percentage. It is calculated according to the following formula:
CV (%) = 100 × (SD ÷ mean).

Control – see definition in Terms for Test Materials or Substances.

Endpoint means the response(s) of the test organism that is measured (e.g., death or biomass), or the value(s) that characterize the results of a test (e.g., LC50, IC25).

Environmental toxicology is a branch of *toxicology* with the same general definition. However, the focus is on ecosystems, natural communities, and wild living species, without excluding humans as part of the ecosystems.

Flow-through describes tests in which solutions in test vessels are renewed continuously by the constant inflow of a fresh solution, or by a frequent intermittent inflow.

Geometric mean is the mean of repeated measurements, calculated logarithmically. It has the advantage that extreme values do not have as great an influence on the mean as is the case for an arithmetic mean. The geometric mean can be calculated as the nth root of the product of the *n* values, and it can also be calculated as the antilogarithm of the mean of the logarithms of the *n* values.

Heteroscedasticity refers herein to data showing heterogeneity of the residuals within a scatter plot (see EC, 2005). This term applies when the variability of the residuals changes significantly with that of the independent variables (i.e., the test concentrations or treatment levels). When performing statistical analyses and assessing residuals (e.g., using Levene's test), for test data demonstrating heteroscedasticity (i.e., non-homogeneity of residuals), there is a significant difference in the variance of residuals across concentrations or treatment levels. (See also *homoscedasticity* and *residual*.)

Homoscedasticity refers herein to data showing homogeneity of the residuals within a scatter plot (see EC, 2005). This term applies when the variability of the residuals does not change significantly with that of the independent variables (i.e., the test concentrations or treatment levels). When performing statistical analyses and assessing residuals (e.g., using Levene's test), for test data demonstrating homoscedasticity (i.e., homogeneity of residuals), there is no significant difference in the variance of residuals across concentrations or treatment levels. (See also *heteroscedasticity* and *residual*.)

ICp is the inhibiting concentration (e.g., 5 mg/kg) for a specified percent effect (p). It represents a point estimate of the concentration of test substance or material that causes a designated percent inhibition (p) compared to the control, in a *quantitative* (continuous) biological measurement such as biomass of test organisms at the end of the test (e.g., IC25 or IC50).

LC50 is the median lethal concentration, i.e., the concentration (e.g., % or mg/kg) of substance(s) or material(s) that is estimated to be lethal to 50% of the test organisms. The LC50 and its 95% confidence limits are usually derived by statistical analysis of percent mortalities in five or more test concentrations, after a fixed period of exposure. The duration of exposure must be specified (e.g., 14-day LC50). Depending on the study objectives, an LCp other than LC50 (e.g., an LC25) might be calculated instead of or in addition to the LC50.

Lethal means causing death by direct action. Death of test organisms is defined as the cessation of all visible signs of movement or other activity indicating life.

LOEC is the *lowest-observed-effect concentration*. This is the lowest concentration of a test substance or material for which a statistically significant adverse effect on the test organisms was observed, relative to the control.

NOEC is the *no-observed-effect concentration*. This is the highest concentration of a test substance or material at which no statistically significant adverse effect on the test organisms was observed, relative to the control.

Normality (or *normal distribution*) refers to a symmetric, bell-shaped array of observations. The array relates frequency of occurrence to the magnitude of the item being measured. In a *normal distribution*, most observations will cluster near the mean value, with progressively fewer observations toward the extremes of the range of values. The normal distribution plays a central role in statistical theory because of its mathematical properties. It is also central in biological sciences because many biological phenomena follow the same pattern. Many statistical tests assume that data are normally distributed, and therefore it can be necessary to test whether that is true for a given set of data.

Ordinal data in a toxicity test are those in which the measured effect indicates relative severity or level but not magnitude. In the context of this test method, *ordinal* is used to describe amphibian developmental stages. These are represented by numbers 1 to 46 as described by Gosner (1960), where larger numbers represent more advanced development; however, this development is defined by the presence or absence of specific physical characteristics, and is not otherwise quantifiable (Green *et al.*, 2018). That is, the difference between stages 21 and 22 is not an equivalent advancement in development compared to the difference between stages 41 and 42, and amphibians at stage 42 are not twice as developed as those at stage 21 (Green *et al.*, 2018; John W Green Ecostatistical Consulting, 2021; see also Figures 2.1 and 2.2 herein). *Ordinal* data are analyzed using statistical methods that rely only on the ordering indicated by the labels (i.e., stages), and are not misled by the numbers that are used only as labels (John W Green Ecostatistical Consulting, 2021) (see Section 4.6.2.3 herein).

Precision refers to the closeness of repeated measurements of the same quantity to each other, i.e., the degree to which data generated from replicate measurements are the same. It describes the degree of certainty around a result, or the tightness of a statistically derived endpoint such as an ICp.

Preliminary test – see definition in Terms for Test Materials or Substances.

Quantal effects in a toxicity test are those in which each test organism responds or does not respond. For example, an animal might live or die, or it might develop normally or abnormally. Generally, quantal effects are expressed as numerical counts or percentages thereof. (See also *quantitative*.)

Quantitative effects in a toxicity test are those in which the measured effect is continuously variable on a numerical scale. An example would be the weight attained by individual organisms at the end of the test. Generally, quantitative effects are determined and expressed as measurements. (See also *quantal*.)

Range-finding (test) – see definition in Terms for Test Materials or Substances.

Reference toxicity test – see definition in Terms for Test Materials or Substances.

Replicate (*treatment*, *test vessel* or *test unit*) refers to a single test vessel containing a prescribed number of organisms in either one concentration of the test material or substance, or in the control or reference treatment(s). A *replicate* of a treatment must be an independent test vessel (see Sections 3.3.2 and 4.2 herein, and Section 2.5 of EC, 2005).

Residual, in the context of Section 4.6.2.2, refers to the difference between the predicted estimate (based on the model) and the actual value observed, as determined by subtracting the former from the latter. (See also *heteroscedasticity* and *homoscedasticity*.)

Static describes a toxicity test in which the test solutions (or any chemical or chemical product therein) are not renewed or replaced during the test.

Static renewal describes a toxicity test in which test solutions are renewed (replaced) periodically, usually a minimum of three times weekly on non-consecutive days. Synonymous terms are "intermittent renewal", "semi-static", "renewed static", "static replacement", and "batch replacement".

Sublethal (toxicity) means detrimental to the organism, but below the concentration or level of contamination that directly causes death within the test period.

Sublethal effect is an adverse effect on an organism (e.g., reduced growth, change in development), resulting from exposure to the concentration or level of contamination below that which directly causes death within the test period.

Target effect size is the magnitude of adverse effect in a particular study that is deemed to be important. In this test method, the effect refers particularly to either: i) a reduction in growth, expressed as the percent reduction from the control, or ii) a reduction (i.e., delay) or an increase (i.e., acceleration) in development, expressed as the difference from the control (e.g., 4 Gosner stages). The target effect size can be linked to a policy statement, decided based on expert judgement, chosen to align with other effect sizes in a battery of toxicity tests, or derived through other means. The target effect size is selected before testing begins. Note that selecting a target effect size does not imply that adverse effects will be observed in a particular test; the selection of target effect size only links the number of *replicates* with the ability of the test to "detect" (in terms of statistical significance) an effect, if it does exist.

Toxic means poisonous. A toxic chemical or material can cause adverse effects on living organisms, if present in sufficient amounts at the right location (i.e., receptor/organ). *Toxic* is an adjective or adverb, and *toxicant* is the noun.

Toxicant is a toxic substance or material.

Toxicity is the inherent potential or capacity of a substance or material to cause adverse effect(s) on living organisms. These effect(s) could result from exposure to either *lethal* or *sublethal* concentrations of contaminants.

Toxicity test is a determination of the adverse effect(s) of a substance or material that results from exposure of a group of selected organisms, under defined conditions. An aquatic toxicity test usually measures (a) the proportions of organisms affected (*quantal*), and/or (b) the degree of effect observed (*quantitative* or graded), after exposure of the test organisms to a specific test substance or material (e.g., chemical or effluent) or specific concentrations thereof.

Toxicology is a branch of science that studies the toxicity of substances, materials, or conditions. There is no limitation on the use of various scientific disciplines, field or laboratory tools, or studies at various levels of organization, whether molecular, single species, populations, or communities. Applied toxicology would normally have a goal of defining the limits of safety of chemical or other agents. See also *environmental toxicology*.

Treatment is, in general, an intervention or procedure whose effect is to be measured. More specifically, in toxicity testing, it is a condition or procedure applied to the test organisms by an investigator, with the intention of measuring the effect(s) on those organisms. The treatment could be a specific concentration of a potentially toxic material or substance. Alternatively, a treatment might be a particular test material (e.g., a particular sample of sediment, chemical, effluent, elutriate, leachate, receiving water, or control water). Samples or subsamples of test material or substance representing a particular treatment are typically replicated in a toxicity test. See also *replicate*.

Warning chart is a graph used to follow changes over time in the endpoints for a reference toxicant. The date of the test is on the horizontal axis, and the statistical endpoint (e.g., LC50) is plotted on the vertical logarithmic scale.

Warning limit is plus or minus two standard deviations, calculated logarithmically, from a historic geometric mean of the endpoints from tests with a *reference toxicant*.

Acknowledgements

This document was prepared by Jennifer Miller (Miller Environmental Sciences Inc., Uxbridge, Ontario). Bonnie Lo and James Elphick (Nautilus Environmental Company Inc., Burnaby, BC) are thanked for their compilation of the method development research, and identification of data gaps and research recommendations that were used in the preparation of this method.

Leana Van der Vliet (Biological Assessment and Standardization Section, Environment and Climate Change Canada [ECCC]) acted as Scientific Authority for this method and provided technical input and direction throughout the work. Rick Scroggins, Lisa Taylor (retired from ECCC), and Carolyn Martinko (Biological Assessment and Standardization Section, ECCC) provided significant contributions to various sections of the document.

This method was based primarily on research conducted by ECCC's Atlantic Laboratory for Environmental Testing (ALET, Moncton, New Brunswick) with additional research related to breeding of amphibians conducted at the University of Ottawa under the supervision of Dr. Vance Trudeau (University of Ottawa, Ottawa, Ontario). Paula Jackman (retired from ECCC, ALET), Ken Doe (retired from Environment Canada, ALET), and Dr. Vance Trudeau are gratefully acknowledged for their many years of research effort. Dr. Andrea Edginton is thanked for her review of amphibian culturing and testing procedures.

The inter-laboratory studies undertaken to validate the two testing options of the test method described herein were coordinated by Leana Van der Vliet and Bonnie Lo (Nautilus Environmental Company Inc., Burnaby, BC) and performed by the following participating laboratories: ALET, Moncton, NB; Nautilus Environmental, Burnaby, BC; Nautilus Environmental, Calgary, AB; and AGAT Laboratories, Laval, QC. We gratefully acknowledge the contributions of all the inter-laboratory participants: Virginie Berube, Claude Brousseau-Fournier, Sebastien Dupuis, Caroline Leblanc-Houde, and Lydia Ranger (AGAT); Paula Jackman, Megan Bauer, Lison Hache, Charles Hopper, Dale Hughes, and Taylor Shaw (ALET); Bonnie Lo, Eric Cheung, Andy Diewald, Yvonne Lam, Jillian Sones (Nautilus Environmental, Burnaby); and Bonnie Lo, Jacklyn Poole, Stephanie Schiffer, Courtney Bogstie, Lindsey Clothier, Linda Fan, and Madison Lehti (Nautilus Environmental, Calgary).

We gratefully acknowledge the many useful comments provided by each of the members of the Expert Peer Review Committee, who were responsible for final reviews this report: Paula Jackman, Doug Fort, Melanie Gallant, Elissa Liu, Bonnie Lo, and Stacey Robinson. Contact information for each method writer, contributor, and reviewer is provided in Appendix D. Thanks also to Amélie Maure for her review of the French translation. The ongoing support of members of ECCC's Regional Environmental Testing Laboratories (Appendix B) and the Inter-Governmental Ecotoxicological Testing Group (Appendix C) is also acknowledged.

This project was supported by research funding from Environment and Climate Change Canada over many years.

Section 1

Introduction

1.1 Background

The global decline of many amphibian species is well documented. They are among the most threatened taxa in the current worldwide crisis of biodiversity loss (Houlahan *et al*., 2000; Cohen, 2001; Blaustein and Kiesecker, 2002; Stuart *et al*., 2004; McCallum, 2007; Wake and Vredenburg, 2008). Amphibians are considered to be valuable biological indicators of the quality of habitats and their decline is of great concern. The declining populations have been attributed to a number of anthropogenic activities including: *pollution*, pesticide use, habitat loss or modification, disease, increased UV radiation, and introduced species (Blaustein *et al*., 1994, 1998; Berger *et al*., 1998; Cohen, 2001; Blaustein and Kiesecker, 2002; Collins and Storfer, 2003; Beebee and Griffiths, 2005; Becker *et al*., 2007; Boone *et al*., 2007; Hayes *et al*., 2010; Brühl *et al.*, 2013). Many of the causes of amphibian declines, however, remain unknown, and research to better understand these causes is on the rise (Pauli *et al*., 2000).

Various international agencies have developed laboratory *toxicity tests* that measure the effects of aquatic exposures of *chemicals* and/or complex mixtures to aquatic life stages of amphibians, including: USEPA, 1975, 1996, 2009, 2015; Birge *et al*., 1985; ENSR International, 2004; ISO, 2006; OECD, 2009; and ASTM, 2019, 2022a, 2023a, 2023b. These methods vary in their design (i.e., exposure, test species, life stage used at test initiation, feeding regimes, test volume, etc.), duration (i.e., 96 hours to several weeks), and *endpoint* measurements (i.e., mortality, growth inhibition, development, basic morphometrics of tadpoles, behaviour, thyroid gland histology, gonadal histology, etc.), but have been increasingly used in regulatory frameworks to measure the impacts of anthropogenic materials (see Section 1.4). *Xenopus laevis* is the most commonly studied anuran species, and many of these toxicity test procedures for amphibians focus on the use of *X. laevis*. The relevance of this species to Canadian environments is limited, however, since *X. laevis* is native to Africa.

Although there is growing evidence of their sensitivity to contaminants, amphibian *toxicity* data are currently under-represented in risk assessments. Few standardized test methods are available, which contributes to this under-representation, and none of the available methods pair whole-organism *chronic* endpoints in an aquatic exposure with species that are relevant to Canadian environments. In the context of environmental contaminants, ranids were the most studied family of amphibians between 1996 and 2008 (Sparling *et al.*, 2010). Individual researchers use different methods and different species, which often prevents the direct comparison of highly studied contaminants (e.g., pesticides and metals). A standardized method using a commonly studied ranid species would facilitate these comparisons.

The most commonly studied anuran genera relevant to Canadian ecosystems are *Rana* and *Bufo* spp. Notably, *Lithobates pipiens* (formerly *Rana pipiens*) is a species whose natural range includes Canada and one that is commonly used in toxicity testing (Edginton *et al.*, 2004; Fridgen *et al.*, 2007, 2009; Jackman *et al*., 2007; Hogan *et al*., 2008; Fridgen, 2009; Melvin and Trudeau, 2012a; Melvin *et al*., 2013; Leduc *et al.*, 2016; Milotic *et al.*, 2018; Robinson *et al.*, 2019a, 2020, 2021; Young *et al.*, 2020; Gavel *et al.*, 2021). This test method document provides the first standardized procedure for culturing and conducting a chronic aquatic toxicity test with this species. *Lithobates pipiens* (commonly known as the northern leopard frog) occur in every province and territory except Yukon. They breed in shallow warm ponds and hibernate in deep, well-oxygenated water (CESCC, 2011). During the late 1970s, the northern leopard frog experienced rapid population declines in British Columbia, Alberta, Saskatchewan, and Manitoba, and although their populations have recovered somewhat, their conservation status varies from "Not at Risk" to "Endangered" for the various populations across the country (COSEWIC, 2009;

EC, 2013b; Species at Risk Committee, 2013; see Section 1.2).

The Method Development and Applications Unit (MDAU) of Environment and Climate Change Canada (ECCC) is responsible for the development, standardization, and publication (see Appendix A) of a series of biological test methods for measuring and assessing the *toxic* effect(s) on single species of aquatic or terrestrial organisms, caused by their exposure to samples of test *materials* or *substances* under controlled and defined laboratory conditions. Standardized biological testing procedures for *monitoring* and controlling toxic substances and complex mixtures are essential for the protection of the Canadian environment. Results from *toxicology* testing procedures give an overall integrated estimate of environmental hazard. Results from standardized toxicity tests can be used to determine the need for control of discharges, to set *effluent* standards and water quality guidelines, for the classification and risk assessment of chemicals, and as part of the *ecological risk assessment* of sites.

In 2000, Environment Canada's MDAU (formerly Method Development and Applications Section) commissioned a study to review current culturing and toxicity testing procedures using amphibians, to examine existing standardized guidelines, and to determine future method research needs (Edginton, 2001). This report and a survey of participants formed the foundation for a Learning Fund Workshop on Amphibian Culturing and Toxicity Testing Procedures held over a 2-day period in 2002. A report from this workshop was produced and included 18 recommendations upon which the Environment Canada research program was built.

For over 15 years, Environment Canada has sponsored research for the standardization of an amphibian toxicity test method using *Lithobates pipiens*. Method development research began at Environment Canada's Atlantic Laboratory for Environmental Testing (ALET). From 2010-2013, collaboration with Dr. Vance Trudeau at the University of Ottawa led to improvements in the induction of leopard frog *spawning* both during and outside of the natural breeding season. This focus on in-laboratory breeding *protocols* reflects MDAU's commitment to demonstrating that *L. pipiens* can be available for testing in all regions of Canada without the need to rely on wild populations. During the last 10 years, the focus has been on the development of standard methods for culturing and testing frogs (i.e., *L. pipiens*), specifically *sublethal*, chronic effects (i.e., growth, development) that manifest as a result of whole-organism, aqueous exposure. Based on the results of this research, together with the findings of a series on inter-laboratory method validation studies (Nautilus Environmental, 2020a, 2020b), Environment and Climate Change Canada proceeded with the preparation and publication of a biological test method for conducting aquatic toxicity tests that measure the survival, growth, and development of aquatic life stages of frogs (*L. pipiens*), as described in this method.

A group of national and international experts experienced with the design and implementation of toxicity tests using frogs served actively in providing a critical peer review of the fourth draft of this methodology document (see Appendix D).

Two options for test design are described herein. The 14-day test, designed primarily to assess the survival and growth of tadpoles prior to metamorphosis, starts with the exposure of young tadpoles (i.e., *Gosner stage* [GS] 25). The 42-day test, designed to capture changes in development leading to metamorphosis, starts with the exposure of tadpoles in later stages of development, beginning from the early stages of hindlimb bud development (GS 28/29) and covering stages of toe differentiation and development (GS 31 to 40). Both tests measure the survival (mortality), growth (length and weight), and development (Gosner stage) of aquatic life stages of *L. pipiens*. Deformities are also assessed at the end of each test.

Either of the two test options may be used to evaluate aquatic samples. Selection of the most suitable test option will depend on the study objectives and the nature of the substance being tested (see Section 2.1). Alternative test designs such as extended exposure durations (e.g., a 56-day test starting with GS 25) or the assessment of additional endpoints (e.g., histology, gene expression) may also be desirable, although these have not been standardized. Suggestions for other test designs and additional endpoints are provided in Section 4.6.3.

Universal procedures and conditions for preparing and conducting aquatic life stage tests using the northern leopard frog (*Lithobates pipiens*) are described in this method. Guidance is also provided for specific sets of test conditions and procedures that are required or recommended when using this biological test method for evaluating different types of *substances* or *materials* (e.g., samples of chemical or contaminated water; see Figure 1). The results from tests conducted using this standardized amphibian toxicity test method can be used in ecological risk assessments, the establishment of Canadian water quality guidelines, pesticide regulation for re-evaluation and registration for new use, the evaluation and registration of new and existing commercial substances, and contaminated-site clean-up decision making.

The assessment of aquatic toxicity is an integral part of environmental hazard and risk assessment of chemicals and has therefore been included in important environmental and chemical regulations and legislation around the world. Tests with a mortality endpoint using vertebrates (e.g., fish and amphibians), however, raise ethical and economic concerns. There is a movement towards the protection of animals used for scientific purposes, and an increasing demand for replacement, reduction, and refinement strategies and methods (EU, 2010; Scholz *et al*., 2013; Halder *et al*., 2014; Norberg-King *et al.*, 2018). As such, in formulating the biological test method, high priority was given to reducing the number of test organisms used, maximizing the number of test endpoints to enhance the relevance of the test, and the development of procedures for the provision of a supply of healthy test organisms with minimal impact on natural populations, thereby addressing issues related to animal conservation and animal welfare. See Section 4.9 for further details.

It is assumed that the user of this method has a certain degree of familiarity with aquatic toxicity tests and follows appropriate quality management protocols. Explicit instructions that might be required in a regulatory *protocol* or *reference method* are not provided, although the report is intended to serve as a guidance document useful for that and other applications.

For guidance on the implementation of this and other biological test methods, and on the interpretation and application of endpoint data for aquatic toxicity, the reader should consult Environment Canada's *Guidance Document on Application and Interpretation of Single-Species Tests in Environmental Toxicology* (EC, 1999).

1.2 Identification, Life History, and Distribution[1]

Lithobates pipiens (formerly *Rana pipiens*;[2] Schreber, 1782), commonly known as the northern leopard frog (and meadow or grass frog), is a member of the family Ranidae or "true frogs" (phylum, Chordata; subphylum, Vertebrata; class Amphibia; order Anura). It is a green or brown frog with large rounded dark brown- or olive-coloured, random, oval-shaped spots, outlined with light halos on its back, sides, and legs. It has two light-coloured dorsolateral ridges that line its back from behind each eye to the lower back and a white line extending from either side of the mouth, from the nose to the shoulder. It has a creamy white belly (occasionally with a yellowish or green tinge). *Adults* are usually 5 to 9 cm in body length; however, some may grow as large as 11 cm. Females are generally larger than males, but like many anurans, the males have paired vocal sacs that inflate over their shoulders as they call, significantly heavier forelimb muscles, and develop dark, swollen *nuptial pads* on the innermost fingers during breeding season (Dewey, 1999; COSEWIC, 2009).

[1] The details outlined in this section were compiled from several references including: Dewey, 1999; COSEWIC, 2009; Dodd, 2013; Environment Canada, 2013b; Ontario Nature, 2016; and Canadian Herpetological Society, 2020. The reader should consult these references for additional information on the identification, life history, and distribution of *L. pipiens*.

[2] Until recently, all North American ranid frogs were considered to belong to the single genus *Rana*. However, Frost *et al.* (2006) revised the genus *Rana*, placing most of the North American "true frogs" in the genus *Lithobates*. This taxonomic arrangement has been recognized by many (e.g., Che *et al*., 2007; Collins and Taggart, 2009; Frost *et al.*, 2017), but has been rejected by others (e.g., Hillis, 2007; Pauly *et al*., 2009).

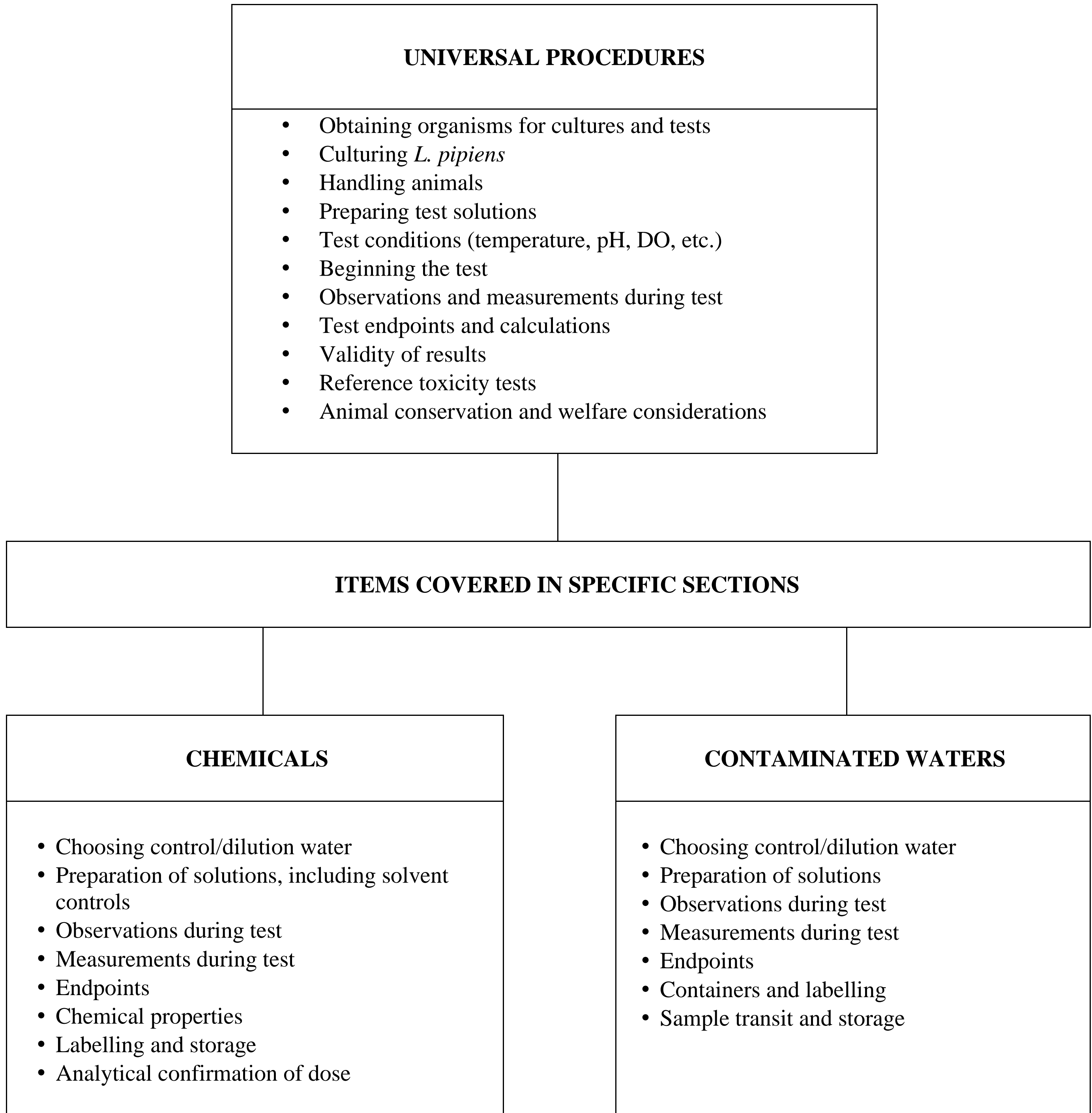


Figure 1 Considerations for preparing and performing toxicity tests using aquatic life stages of frogs (*Lithobates pipiens*) and various types of test materials or substances

In Canada, the pickerel frog (*Lithobates palustris*) is the species most similar to the northern leopard frog. The pickerel frog's spots are squarish or angular and usually arranged in two rows down the back, whereas the northern leopard frog's spots are rounded or oval and are in a more random pattern. The pickerel frog often has a yellow belly and bright yellow or orange coloration in the groin area, and is never green. Pickerel frogs are only found in eastern Canada (i.e., New Brunswick, Nova Scotia, Ontario, and Quebec).

In addition to the diagnostic features described above, DNA-based taxonomic identification (i.e., barcoding) is available for *L. pipiens*. The International Organization for Standardization (ISO) has published a standardized procedure for the identification of ecotoxicological test species using DNA barcoding (ISO, 2019). For *L. pipiens*, sequencing of the 5' region of mitochondrial Cytochrome Oxidase Subunit 1 and sequencing of the nuclear 18S Small Ribosomal Subunit from several specimens are available for comparison on the International Barcode of Life data portal (identified as *Rana pipiens*): www.boldsystems.org. The Barcode of Life Data Systems (BOLD) is one of several international databases that allows access to, and provides a platform for, analysis of DNA barcode sequences. Several studies have also sequenced the mitochondrial NADH Dehydrogenase Subunit 1 gene from Canadian and US northern leopard frog populations (Hoffman and Blouin, 2004; Wilson et al., 2008; O'Donnell and Mock, 2012; O'Donnell *et al.*, 2017).

At one time the leopard frog ranging throughout North America was considered to be a single species with phenotypic variation (Moore, 1944, 1975 in Dodd, 2013); however, it is now recognized as a complex of species based on variations in call structure, morphology, and genetic differentiation (COSEWIC, 2009). The leopard frog complex now consists of about 20 species ranging from northern Canada to Costa Rica (Dodd, 2013). *Lithobates pipiens* is the only member of the complex found in Canada (COSEWIC, 2009). DNA evidence indicates that there are distinct eastern and western clades with the boundary coinciding with the border between the Prairies and the Canadian Shield in Manitoba. There are two distinct haplotypes in the United States as well, with the boundary coinciding with the Mississippi River and the Great Lakes region. Investigators in the US are considering whether or not these two clades constitute two different species (Hoffman and Blouin, 2004; O'Donnell and Mock, 2012; O'Donnell *et al.*, 2017); however, the US Fish and Wildlife Service (2011) has concluded that the western population of northern leopard frogs is not considered discrete in relation to the other populations of northern leopard frogs.[3] In Canada, there is a considerable amount of genetic differentiation among populations of the eastern clade, but this tends to decrease in the western clade (Wilson *et al.,* 2008). The British Columbia populations are isolated and distinct from the other two haplotypes; as such, the Committee on the Status of Endangered Wildlife in Canada (COSEWIC, 2009) recognizes three designatable units (DU). These are:

[3] In the United States, several investigations have been conducted to determine the genetic diversity of northern leopard frog populations. Mitochondrial DNA evidence indicates that *L. pipiens* is split into populations containing discrete eastern and western haplotypes with the Mississippi River and Great Lakes region dividing the geographic ranges. The level of mitochondrial DNA variation is relatively high (3 to 4%) and comparable to the level of divergence observed in different ranid species (Hoffman and Blouin, 2004; O'Donnell and Mock, 2012). Microsatellite loci data corresponded well to the mitochondrial DNA results, with distinct differences between east and west populations and an abrupt change suggestive of a genetic boundary at the Mississippi River (O'Donnell and Mock, 2012). Nuclear DNA sequences also showed differing eastern and western lineages, but these shared a much larger zone of introgression (i.e., mixing of haplotypes). However, in response to a petition to list the western population of northern leopard frogs as a threatened distinct population segment (DPS), the US Fish and Wildlife Service (2011) concluded that the western population cannot be considered as markedly separate from other populations of the species and does not represent a valid DPS, based on the large zone of introgression and other available data. More recently, O'Donnell *et al.* (2017) found a majority of eastern haplotypes in a region where mostly western haplotypes had been found previously (Hoffman and Blouin, 2004), and raised concerns about the outbreeding of native haplotypes balanced with the benefits of increased genetic diversity.

- The Rocky Mountain DU ("Endangered"), which consists of populations in British Columbia;
- The Prairie/Western Boreal DU ("Special Concern"), containing populations in Alberta, Saskatchewan, the Northwest Territories, and Manitoba, west of the Canadian Shield; and
- The Eastern DU ("Not at Risk"), consisting of populations in Manitoba, east of the Canadian Shield, Ontario, Quebec, and eastern Canada.

In addition to the Species at Risk Act statuses listed above, the species is considered "Endangered" in British Columbia, "At Risk" in Alberta and Saskatchewan, "May be at Risk" in the Northwest Territories, and "Secure" in Manitoba (COSEWIC, 2009; EC, 2013b; Species at Risk Committee, 2013). These differences in population status across the country may have implications for field collection, importation, or testing of *L. pipiens*. At the time of writing, eastern and western haplotypes of *L. pipiens* are still recognized as one species (O'Donnell *et al.*, 2017), and differences in sensitivity between these populations of northern leopard frogs have not been investigated.

Northern leopard frogs are widely distributed in North America, from southcentral British Columbia to Labrador, and from the southcentral Northwest Territories down through the central and southwestern United States near to Mexico, northeast into Nebraska, east through West Virginia, and north into Maine; however, a recent survey suggests the species is disappearing from some historically inhabited regions of northeastern states (Schlesinger *et al.*, 2018). Northern leopard frogs are native to Canada, occurring in every Canadian province and territory except the Yukon. They were introduced to Vancouver Island and Newfoundland but are believed now to be extirpated from these areas. In British Columbia and Alberta, the populations are currently restricted to the southeastern portions of the provinces. *L. pipiens* are believed to be relatively widespread in southern Saskatchewan, southern Manitoba (Saskatchewan Prairie Conservation Action Plan, 2018), and eastern Canada, despite some regional declines. They are found in a wide variety of habitats including prairies, woodland, and tundra.

Lithobates pipiens are semi-terrestrial and use three distinct habitat types during their life cycle. During the winter (October to March), they hibernate in cold, well-oxygenated bodies of water that do not freeze solid, often in different ponds from those in which they breed. Breeding and *larvae* occur in pools, ponds, marshes, and lakes, or even slow-moving streams or creeks. During the summer, they can be found in moist upland meadows and native prairie habitats. Home ranges of up to 600 m^2 are maintained during the summer.

They emerge from hibernation shortly after the ice has melted in the early spring. Calling by males gathered at communal breeding ponds, indicating breeding activity, occurs as early as mid-April and can continue until June in more northerly regions. They breed in relatively permanent ponds, with or without fish, but most often wetlands without large fish. If successful, a male will hold a female in *amplexus* using his specialized thumbs, and will fertilize her *eggs* as they leave her body. Each female mates once annually, lays a single *egg mass*, and leaves the pond, whereas males probably mate more than once. Each female deposits between 300 to 7000 eggs (typically several hundred to more than a thousand) in a flattened spherical or oval mass that is often attached to submerged vegetation, on the pond bottom or floating at the water's surface. *L. pipiens* eggs average about 2.0 mm (range 1.3 to 2.3 mm) in diameter and are black on top and white on the bottom when released. Egg size depends on time since deposition, since eggs tend to swell after several days in water (Dewey, 1999; Dodd, 2013). The eggs are usually packed close together with two to three jelly envelopes surrounding each egg (Dodd, 2013). The eggs hatch in one to three weeks depending on the water temperature, but typically hatchlings emerge in 9 days and spend a couple of days clinging to the vegetation and the remnants of the egg mass before becoming free-swimming tadpoles.

When they first hatch, tadpoles range from 8 to 10 mm long. Tadpoles typically reach about 25 mm snout-vent length (total length 90 mm) before commencing metamorphic climax (i.e., before GS 41) (Species at Risk Committee, 2013). They are deep-bodied, dark brown or olive to gray dorsally often with gold speckles, and cream coloured on the underside with a bronze iridescence. Their ventral

side is light enough that viscera tend to be visible through the skin and the throat is translucent (Stebbins, 2003; Dodd, 2013). Tail fins may or may not be heavily marked with speckles or small spots, and the dorsal tail fin is rounded. In general, the tail fin is lighter than the body and the eyes are bronze (Dodd, 2013). Tadpoles are mainly herbivorous, feeding on plants and algae, although they may occasionally feed upon dead tadpoles or dead invertebrates. Their main predators include aquatic insect larvae, water birds, garter snakes, fish, and leeches (Species at Risk Committee, 2013). Tadpoles take approximately 2 to 3 months to reach metamorphosis, after which, as small frogs, they move to summer foraging habitat to feed on a variety of insects. Newly metamorphosed frogs are about 20 to 30 mm long (Dewey, 1999). *Juveniles* have the same coloration as adults (Species at Risk Committee, 2013). They will not become sexually mature until 1 to 3 years of age, and typically live for 3 to 5 years in the wild, although they may live up to 9 years in captivity (Dewey, 1999; Species at Risk Committee, 2013; Ontario Nature, 2016; Canadian Herpetological Society, 2020). Adult *L. pipiens* feed on almost anything they can catch, but primarily a variety of insects and other invertebrates including beetles, ants, flies, worms, spiders, snails, and slugs. Most feeding activity occurs at night, but they have been known to feed during the day if prey passes by their resting spot. The main predators for juvenile and adult *L. pipiens* include herons, waterfowl, raptors, snakes, turtles, fish, mammals (e.g., mink, raccoon, river otter, and fox), and larger frogs (COSEWIC, 2009; Species at Risk Committee, 2013; Ontario Nature, 2016). Humans have used these frogs in science (for educational purposes and for testing) and as a food item as well (Species at Risk Committee, 2013).

1.3 Sensitivity of Frogs in Toxicity Tests

Amphibians are a diverse group of non-target organisms that have traditionally received less attention in aquatic ecotoxicological assessments, when compared to fish and aquatic invertebrates (Relyea, 2004b; Jones *et al*., 2009; Brinkman and Johnston, 2012; Ortiz-Santaliestra *et al.*, 2018). Historically, amphibian toxicity data have not been considered when assessing the impacts of anthropogenic activities and the *risks* associated with chemicals being released into the environment. The reasons for their exclusion include: limited toxicity data; lack of standardized laboratory toxicity test protocols; lack of information on exposure parameters and life history; and difficulties in estimating exposure, particularly via food ingestion. However, with increased awareness of the potential sensitivity of amphibians, the global decline of amphibians, and their trophic importance in both aquatic and terrestrial ecosystems, there has been a substantial increase in research on the effects of anthropogenic substances on amphibians and a rise in the use of amphibian species in ecotoxicological investigations (Sparling *et al*., 2000, 2010; Edginton, 2001; Hopkins, 2007). Guidance on conducting ecological risk assessments for amphibians is now available in Module 6 of the Federal Contaminated Sites Action Plan (FCSAP) published by Environment and Climate Change Canada (FCSAP, 2019).

1.3.1 Comparison with Fish

North American frogs play a key role in aquatic and terrestrial habitats and are important indicators for ecotoxicology. Their biphasic lifestyle enables them to be used to monitor environmental contamination in both aquatic and terrestrial environments; however, their dependence on water for breeding and early life stage development, their highly permeable skin and gills, and the complex physiological and morphological changes that take place during metamorphosis mean that all three stages of development (i.e., *embryo*, *larva*, and *adult*) are susceptible to exposure to water contamination (Berrill *et al*., 1994; Mann *et al*., 2003; Howe *et al*., 2004; Melvin and Trudeau, 2012a).

Despite the complexities in amphibian life history traits, fish are often used as a surrogate for amphibians in risk assessment, with the assumption that regulatory decisions based on observed toxicity to fish will also be protective of the aquatic phases of the amphibian life cycle. To validate this assumption, a number of studies have evaluated the relative sensitivities of compounds in aquatic exposures to fish and amphibian species. Conclusions and scientific interpretations of the data have varied. There is some evidence that amphibians are more sensitive than fish to aquatic contaminants including certain metals (Birge *et al.*, 2000; Bridges

et al., 2002), phenols (Kerby *et al.*, 2010; Weltje *et al.*, 2013; Ortiz-Santaliestra *et al.*, 2018), and pesticides (Aldrich, 2009; Weltje *et al.*, 2013; Ortiz-Santaliestra *et al.*, 2018; Glaberman *et al.*, 2019), as well as perchlorate, a known thyroid disruptor (Weltje *et al.*, 2013). Other studies have shown that fish- or invertebrate-derived toxicity data may not be relevant to amphibians (Relyea 2004b, 2005b; Ortiz-Santaliestra and Brühl, 2014). This is not unexpected due to vast differences in their life history, physiology, respiration, immunology, endocrinology, etc. However, several studies have concluded that overall, fish (Weltje *et al.*, 2013; Ortiz-Santaliestra *et al.*, 2018) and invertebrates (Aldrich, 2009) are equivalently sensitive to, or more sensitive than, amphibians to *acute* chemical exposures, when the important test design features (duration, endpoint, level of effect) are held constant. Similar conclusions on sensitivity have been proposed for chronic chemical exposures (Glaberman *et al.*, 2019; Weltje *et al.*, 2013), although the relevance of fish data for chronic exposure warrants further discussion.

Birge *et al.* (2000) exposed embryo-larval life stages of various amphibian and fish species to 27 organic and 34 inorganic (including metals) compounds, from fertilization until 4 days post-hatch, and compared the resulting *LC50s*. Overall, approximately half of the 25 amphibian species were more tolerant to the tested compounds than rainbow trout (*O. mykiss*), the most sensitive fish species. Amphibians had lower LC50s than *O. mykiss* in 52% of 203 cases for metals and 36% of 44 cases for organic compounds. Some compounds elicited a wide range of responses across species, while for other compounds, multiple amphibian species exhibited similar sensitivity to *O. mykiss*.

Other studies have determined that amphibian and fish toxicity data have a strong positive correlation, and that amphibians would be protected in most instances by applying a 100× safety factor to fish data (Aldrich, 2009; Doe *et al.*, 2012; Weltje *et al.*, 2013; Ortiz-Santaliestra *et al.*, 2018; Glaberman *et al.*, 2019). Bridges *et al.* (2002) measured 96-h LC50s of five chemicals with different modes of actions in southern leopard frogs (*Rana sphenocephala*) and compared these to published 96-h LC50s from standardized fish toxicity tests. The results found that tadpoles had equal or greater tolerance than fish to all substances except copper, and that all differences in toxicity were <10-fold from rainbow trout and fathead minnow (*Pimephales promelas*) data. In a review of 24 pesticides by Aldrich (2009), the acute sensitivity of amphibians was less than or equivalent to that of invertebrates and fish, except in two cases (dimethoate and 2,4-D) where the 96-h LC50 for the most sensitive amphibian was lower than that of the most sensitive invertebrate or fish. A review by Weltje *et al.* (2013) also found that amphibians were not more sensitive than fish, and in some cases were much less sensitive. Amphibians had lower 96-h LC50s than fish for 16 of 55 substances, but this difference was less than 10-fold in 10 cases and was greater than 100-fold for only *p*-nonylphenol and dimethoate. Similarly, a comparative analysis of 96-h LC50s for 29 pesticide formulations tested at ALET concluded that northern leopard frog larvae were not more sensitive than juvenile *O. mykiss* under standardized conditions, and that *acute toxicity* data were strongly correlated for these species (Doe *et al.*, 2012; Martinko and Van der Vliet, 2021; ECCC, 2023). Lastly, amphibians were more sensitive than *O. mykiss* to 34.2% of 117 chemicals reviewed by Ortiz-Santaliestra *et al.* (2018), but not more sensitive to any particular group of substances. The study also found that applying a safety factor of 100 to *O. mykiss* 96-h LC50 data would be protective of amphibians for 94.9% of the assessed substances. Overall the evidence supports that fish, particularly *O. mykiss*, may be a suitable surrogate for larval amphibians, using a 96-hour exposure and lethality (LC50) as the endpoint.[4] It should be noted however that the Federal Contaminated Sites Action Plan recently stated that based on life history differences, they do not recommend the use of fish toxicity data as a surrogate for amphibians in risk assessment, unless amphibian data are not available (FCSAP, 2019). If toxicity data from fish are used as a surrogate, FCSAP (2019) recommends including additional lines of evidence that assess risk to amphibians more directly, as well as a comprehensive uncertainty analysis pertaining to the use of the surrogate species data.

[4] Most comparisons to date have used the LC50. As slopes of dose-response relationships might differ, a trend in LC50s between fish and amphibians does not indicate a trend in other mortality (e.g., LC10) endpoints.

Much of the published data has focused on short-term exposures that measure the effect of substances or materials on survival alone; however, there is a growing body of evidence that these short-term exposures seriously underestimate potential adverse effects on amphibians (Relyea, 2004b; Chen *et al.*, 2007). Compared with acute exposures, studies of chronic exposures in amphibians are relatively rare. Chronic tests prolong the exposure period, which is often more reflective of environmental exposures. In addition, chronic exposures enable measurement of *sublethal effects*, and can cover a complete reorganization of the physiological system (metamorphosis), which simply does not occur in fish. Comparing chronic exposures between fish and amphibians is more problematic than in the acute scenario, because duration, biological effect, and effect size vary in chronic studies. In the acute scenario, duration (96 hours), biological effect (mortality), and effect size (50% mortality) are defined and often consistent between studies. Chronic duration can mean any number of weeks or months. Biological effects in a chronic study can be growth, development, and/or mortality. Effect size is typically undefined or poorly defined because *quantitative* data resulting from chronic exposures are frequently analyzed using the no-observed-effect concentration (*NOEC*) and the lowest-observed-effect concentration (*LOEC*).

Weltje *et al.* (2013) concluded that amphibians were much less sensitive than fish to chronic exposures to 52 compounds, except for sodium perchlorate; however, their assessment compared tests of different durations and inconsistent NOEC endpoints, so the results must be interpreted with caution. Glaberman *et al.* (2019) was able to overcome some of these issues when comparing survival, body weight, and length endpoints from standardized tests with 21-day exposures (Amphibian Metamorphosis Assay [AMA] using *Xenopus laevis* and Fish Short-Term Reproduction Assay [FSTRA] using *Pimephales promelas*) for 45 pesticides. The authors found a strong positive correlation between AMA and FSTRA results, and amphibians were more sensitive than fish approximately half of the time. Amphibians exhibited >10-fold sensitivity in the different endpoints for up to 6 pesticides, and >100-fold sensitivity in survival for only the fungicide propiconazole. Glaberman *et al.* (2019) also highlighted the importance of testing the effects of potential thyroid-disrupting compounds in amphibians, and, like other authors (Ortiz-Santaliestra *et al.*, 2018), called for further research to generate chronic amphibian toxicity data. Finally, while *Xenopus laevis* toxicity data is widely available and frequently included in comparisons between fish and amphibians, it can be less sensitive than other amphibian species (Birge *et al.*, 2000; Kerby *et al.*, 2010; Ortiz-Santaliestra *et al.*, 2018). This emphasizes the importance of generating *chronic toxicity* data with more relevant, sensitive, and native amphibian species, such as *L. pipiens*.

1.3.2 Comparison Among Aquatic Life Stages

There are many unique aquatic life stages in amphibians. The sensitivity of these stages (i.e., embryonic, larval, or those involved in metamorphosis) may vary depending upon the life stage that is exposed, the substance under assessment, and/or the biological endpoint (Biga and Blaustein, 2013). Metamorphosis is a unique physiological change that may be particularly sensitive to disruption.

As aquatic breeding organisms, amphibian eggs and larvae are particularly vulnerable to chemicals in their environment (Blaustein *et al*., 2003). Unlike most taxa where the earliest life stage is considered the most sensitive, organisms with protective eggs, like amphibians, have several layers of jelly that protects them from a number of external disturbances (Marquis *et al*., 2006). The extent to which the jelly coat surrounding amphibian eggs protects the developing embryo from a chemical, however, is strongly dependent on both the chemical and the species examined (Greulich and Pflugmacher, 2004). Amphibian embryos are likely exposed to environmental pollutants as the jelly is filled with water shortly after being laid (Marquis *et al*., 2006), although this would be a difficult exposure scenario to capture in a standardized test method. In addition, uptake of waterborne contaminants has been observed in anuran eggs (Greulich and Pflugmacher, 2004).

Larval amphibian stages have been found to be at least as sensitive as, and often more sensitive to, chemical contaminants than embryos (Berrill *et al*., 1993, 1994, 1998; Edginton *et al*., 2003; Edginton *et al*., 2004; ALET, 2009; Wagner *et al*., 2013; Yu *et*

al., 2013) and *hatchlings* (Berrill *et al*., 1998; Howe *et al*., 2004). Some studies have shown that the early larval stages are often the most sensitive to chemical contaminants (Chen *et al*., 2007) whereas other studies have observed greater toxicity in older tadpoles (Earl and Whiteman, 2009; Melvin and Trudeau, 2012b; Martini *et al*., 2012). In keeping with these findings, the test options described in this test method start with larval amphibian stages (GS 25 and GS 28/29), thereby focusing exposure on the most sensitive and perhaps most relevant aquatic life stages of *L. pipiens* and avoiding the use of test organisms for low-impact tests (i.e., tests involving less sensitive life stages such as embryos; see Sections 2.1 and 4.6.3 and footnote 98).

1.4 Use of Frogs in Toxicity Tests

The demand for quality amphibian data for use in ecological risk assessments, the establishment of environmental quality guidelines, pesticide regulatory programs, and the evaluation and registration of new chemicals and substances is on the rise. The US (EDSP [Endocrine Disruptor Screening Program]), Japan (EXTEND [Extended Tasks on Endocrine Disruption]), and European Union (REACH [Registration, Evaluation and Authorization of Chemicals]) all have legislation requiring ecotoxicity data to screen chemicals for endocrine disrupting effects. Other EU regulations now require available terrestrial amphibian data to be considered during risk assessment of active substances, and if risk cannot be predicted from these data, then it must be addressed in another way (Ortiz-Santaliestra *et al.*, 2017). The European Food Safety Authority Panel on Plant Protection Products and their Residues (EFSA PPR Panel, 2018) is also encouraging the development of a pesticide risk assessment scheme for amphibians, including data on chronic exposures in aquatic and terrestrial environments.

One of the most detailed international protocols for aquatic amphibian toxicity testing is the OECD Guideline for the Testing of Chemicals: The Amphibian Metamorphosis Assay (AMA; OECD, 2009), which was designed as a *screening* assay to identify substances that interfere with the hypothalamic-pituitary-thyroid axis. This *bioassay* exposes *Xenopus laevis* tadpoles for 21 days, which is sufficient for control tadpoles to reach metamorphosis. The assay is recommended as part of the OECD's Conceptual Framework for Testing and Assessment of Endocrine Disrupting Chemicals. More recently (July 2015), the USEPA released a globally harmonized test method document entitled "Larval Amphibian Growth and Development Assay (LAGDA)" as part of a series of test guidelines to be used in the Endocrine Disruptor Screening Program that is also to be included in the OECD's Conceptual Framework for the Testing and Assessment of Endocrine Disrupters. These guidelines, established by the Office of Chemical Safety and Pollution Prevention (OCSPP), are for use in testing pesticides and chemical substances. The document provides guidelines for conducting toxicity tests using *X. laevis* from embryos through larval development, metamorphosis, and early juvenile development. The assay is not solely designed to detect effects of endocrine disruption, but is intended to evaluate the effects from exposure to contaminants, both through endocrine and non-endocrine mechanisms, and is expected to be sensitive to many chemicals as well as to other reproductive *toxicants*. Endpoints evaluated during the course of the exposure include mortality, abnormal behavior, and growth determinations (e.g., length and weight), as well as developmental endpoints (e.g., time to metamorphosis, developmental abnormalities) and histopathology endpoints (USEPA, 2015).

Other international protocols that have been developed and include limited information relevant to the method described herein include:

- Naval Facilities Engineering Command (NAVFAC) Technical Report TR-2245-ENV: This is a sediment toxicity test that exposes *L. pipiens* or *Bufo americanus* to sediment with overlying water in a 10-day exposure. Tadpoles ≤72 hours old are used. Organisms are fed after tadpoles reach Gosner stage 25. The endpoints are survival, and body length and width (ENSR International, 2004).
- ASTM-E2591-22: This protocol describes a method for whole sediment toxicity assays using *L. pipiens*. Test conditions and endpoints are very similar to those described by ENSR International (2004) (ASTM, 2022a).
- EPA-712-C-96-132 (OPPTS 850.1800): This protocol is a sediment subchronic toxicity assay

developed for *Rana catesbeiana*. This is a 30-day test initiated with tadpoles at Gosner stage 31-34 (USEPA, 1996).
- EPA 740-C-09-002 (OPPTS 890.1100): This method is consistent with the OECD Amphibian metamorphosis assay – No. 231 described earlier (USEPA, 2009).
- APHA 8930: This document describes acute toxicity testing procedures and some recommendations for chronic toxicity testing using *Xenopus laevis* or *Rana* (*Lithobates*) species (APHA *et al.*, 2011).
- ASTM-E1192-23: This document provides guidance for conducting acute toxicity tests on aqueous samples (i.e., waters, effluents, and leachates) with fish, macroinvertebrates, and amphibians. Testing guidelines are generic, and not amphibian-specific (ASTM, 2023a).
- ASTM-E729-23e1: This document provides guidance for conducting acute toxicity tests on test materials (i.e., chemicals) added to *dilution water* with fish, macroinvertebrates, and amphibians. Testing guidelines are generic, and not amphibian-specific (ASTM, 2023b).
- ASTM-E1439-12 (FETAX): The FETAX protocol was developed specifically for use with *Xenopus laevis*, a non-native species. The test duration is 96 hours and is initiated with de-jellied embryo at the mid blastula to early gastrula stage. Within the exposure period the organisms develop into tadpoles and head-to-tail length can be used as an endpoint (ASTM, 2019).
- GB/T 31270.18: This is an acute toxicity test that uses *Rana limnocharis* (native to parts of China and Southeast Asia) or *Xenopus laevis* larvae to assess pesticides. The duration is 96 hours and the endpoint is survival, but signs of toxicity including abnormal behaviour are also reported (ICAMA, 2014).
- ISO 21427-1 Evaluation of genotoxicity by measurement of induction of micro nuclei – Part 1: Evaluation of genotoxicity using amphibian larvae: This method is used to assess genotoxicity to *Xenopus laevis* and *Pleurodeles waltl* larvae after a 12-day aquatic exposure (ISO, 2006; Mouchet *et al.*, 2011).

The ASTM (2022a) and NAVFAC (ENSR International, 2004) test methods contain some guidance (e.g., holding/culturing, handling, staging, etc.) that was relevant to the test method described herein in that *L. pipiens* is the recommended test species.

Extensive reviews have been carried out summarizing the issues related to the ecotoxicology of amphibians including, but not limited to: Sparling *et al*., 2000, 2010; Mann *et al*., 2009; Egea-Serrano *et al.*, 2012; Wagner *et al*., 2013; Amaral *et al.*, 2019; Sievers *et al.*, 2019; and Trudeau *et al.*, 2020. The reader is directed to these textbooks and review papers, and the reference cited therein, for further detail.

Toxic effects resulting from chronic exposures of *L. pipiens* and other common anuran frog species to a wide range of environmental contaminants have been documented in laboratory studies involving samples of:

- **Pesticides** (Allran and Karasov, 2000; Howe *et al.*, 2004; Relyea, 2004a, 2004b, 2005a, 2005b; Orton *et al*., 2006; Jones *et al*., 2009; Shenoy *et al*., 2009; Williams and Semlitsch, 2010; Weir *et al*., 2012; Biga and Blaustein, 2013; Boone *et al*., 2013; Brühl *et al.*, 2013; Higley *et al*., 2013; Yahnke *et al*., 2013; Robinson *et al.*, 2019a, 2021; Gavel *et al.*, 2021);

- **Metals** (Baud and Beck, 2005; Gross *et al.*, 2009; Chen *et al*., 2007, 2009; Araújo *et al.*, 2014; Leduc *et al.*, 2016);

- **Per-/polyfluoroalkyl substances** (Hoover *et al.*, 2017; Brown *et al.*, 2021);

- **Solvents** (Young *et al.*, 2020);

- **Wastewater effluents** (Sowers *et al.*, 2009);

- **Other chemicals** (Mackenzie *et al.*, 2003; Ankley *et al*., 2004; Fraker and Smith, 2004; McDaniel *et al*., 2004; Sanzo and Hecnar, 2006; Hogan *et al*., 2008; Croteau *et al*., 2009; Earl and Whiteman, 2009, 2010; Paden *et al*., 2011; Melvin and Trudeau, 2012a, 2012b; Van Schmidt *et al*., 2012; Marlatt *et al.*, 2013; Tompsett *et al*., 2013; Stanley *et al*., 2015; Milotic *et al*., 2017, 2018; Pillard *et al.*, 2017; Robinson *et al.*, 2020); and

- **UV radiation** (Baud and Beck, 2005; Croteau *et al*., 2009).

Traditional endpoints commonly examined in chronic ecotoxicological investigations using frogs include:

- hatching success;
- survival (mortality);
- growth (length and weight; growth rate);
- development (% metamorphosis; time to metamorphosis or tail resorption, and/or developmental stage at the end of the test);
- morphological abnormalities; and
- histological analysis (e.g., thyroid gland).

More and more non-traditional endpoints are being used as alternatives to measure the effects of specific chemicals or to identify specific "mode-of action" effects. In addition, depending on test design features, alternative endpoints have the potential to contribute to replacement, reduction, and refinement strategies with vertebrate testing. Some of the non-traditional endpoints include:

- endocrine disruption, including cultured tadpole tail fin (C-fin) biopsy assay (Crump *et al*., 2002; Croteau *et al*., 2009; Hinther *et al*., 2010a, 2010b; Hersikorn and Smits, 2011; Smits *et al.*, 2012; Tompsett *et al*., 2013; Wojnarowicz *et al*., 2013);
- physiology (Goulet and Hontela, 2003; Smits *et al*., 2012; Melvin *et al*., 2013; Gavel *et al.*, 2021; Robinson *et al.*, 2021);
- behaviour, including swimming activity and avoidance (Fraker and Smith, 2004; Chen *et al*., 2007; Sowers *et al.*, 2009; Storrs Méndez *et al*., 2009; Paden *et* al., 2011; Biga and Blaustein, 2013; Araújo *et al*., 2014; Stanley *et al.*, 2015; Heerema *et al.*, 2018);
- gene expression (Howe *et al.*, 2004; Higley *et al*., 2013);
- the EcoToxChip qPCR array (Basu *et al.*, 2019; Crump *et al.*, 2023; www.ecotoxchip.ca/);
- transcriptomics (Jackman *et al.*, 2018);
- metabolomics (Melvin *et al.*, 2017);
- sex differentiation (Mackenzie *et al.*, 2003; Orton *et al*., 2006; Hogan *et al.*, 2008; Sowers *et al.*, 2009; van Schmidt *et al.*, 2012; Robinson *et al.*, 2020; Young *et al.*, 2020); and
- predatory or parasite stress (Relyea, 2004b, 2005b; Budischak *et al*., 2009; Milotic *et al.*, 2017, 2018; Robinson *et al.*, 2019a; Brown *et al.*, 2021; Gavel *et al.*, 2021).

Xenopus laevis (African clawed frog) is the most commonly studied anuran species, and published toxicity test procedures for amphibians have historically focused on the use of *X. laevis* (ISO, 2006; OECD, 2009; USEPA, 2009, 2015; ASTM, 2019). It has received a disproportionate focus due to widespread knowledge and ease of culturing in the laboratory at any time of the year. The development, growth, and biology of *X. laevis* under laboratory conditions have been extensively studied and are well documented. However, as described in Section 1.1, the relevance of this species to Canadian environments is low since it is fully aquatic and native to southern Africa. In addition, sensitivity can vary dramatically among different amphibian species, and therefore tests with *X. laevis* may tell us little about the effects of chemicals on North American frogs (Relyea, 2004b).

There are a number of *Lithobates* (formerly *Rana*), *Bufo*, *Pseudacris,* and *Hyla* species whose natural ranges include Canada and are commonly used in toxicity testing. These include *Lithobates pipiens*, *Lithobates sylvaticus*, *Lithobates clamitans*, *Lithobates septentrionalis*, *Lithobates catesbeiana*, *Lithobates palustris, Bufo boreas*, *Bufo americanus*, *Pseudacris crucifer*, *Pseudacris regilla*, *Hyla versicolor*, and *Hyla chrysoscelis*. The most commonly studied native Canadian anuran is *L. pipiens* (Edginton, 2001).

Extensive work with amphibians has been done at two ECCC laboratories. Research at the Pacific and Yukon Laboratory for Environmental Testing (PYLET) facility has primarily been performed using American bullfrogs (*Lithobates catesbeiana*), while ALET has studied northern leopard frogs (*Lithobates pipiens*) and wood frogs (*Lithobates sylvaticus*). Studies at PYLET, in collaboration with the University of Victoria, have investigated the effects of acute or chronic exposures to atrazine (Gunderson *et al.*, 2011), ibuprofen (Veldhoen *et al.*, 2014), nanometals (Hinther *et al.*, 2010b), triclosan (Veldhoen *et al.*, 2006), thyroid hormones, estradiol, and municipal wastewater effluent (Heerema *et al.*, 2018; Jackman *et al.*, 2018) on *L. catesbeiana*. The

effects of these substances on amphibian thyroid or olfactory systems have been assessed using both traditional endpoints and gene expression. In another study, the Amphibian Metamorphosis Assay (see OECD, 2009) was adapted for the Pacific tree frog (*Pseudacris regilla*), and changes in gene expression were compared to morphological endpoints (Marlatt *et al.*, 2013). Studies at ALET, in collaboration with the University of Ottawa, have assessed the effects of pulsed or chronic exposures to glyphosate-based herbicides on wood frogs (*Lithobates sylvaticus*; Lanctôt *et al.*, 2014; Navarro-Martín *et al.*, 2014). As mentioned in Section 1.3.1, ALET assessed the acute toxicity of various pesticides to *L. pipiens*, *L. sylvaticus*, and *L. catesbeiana* tadpoles as well as non-amphibian species over a twelve-year study period (Doe *et al.*, 2012). ALET also contributed to the development of a published artificial hibernation and breeding protocol for *Lithobates pipiens* (Trudeau *et al.*, 2010), which was refined throughout the development of this test method (ALET, 2004, 2006, 2009, 2013, 2018; see Appendix F). ALET has developed many amphibian husbandry practices and test method procedures since 2002, including: optimizing culturing, holding, food source, loading density, and chronic toxicity test conditions for different life stages; disease prevention and treatment; assessing effects of water quality (e.g., *hardness*, ammonia); evaluating potential reference toxicants; and providing test organisms for, as well as participating in, three rounds of inter-laboratory studies (ALET, 2004, 2006, 2009, 2013, 2015, 2016a, 2016b, 2018; P. Jackman, Environment Canada, Moncton, NB, personal communication, 2007; Nautilus Environmental, 2020a, 2020b).

Section 2

Test Organisms

2.1 Species and Life Stage

This biological method must be performed using *Lithobates pipiens* (formerly *Rana pipiens*; Schreber, 1782) for each of the two test options (14-day or 42-day test) described herein (see Section 1.1 and 4.3.1).[5] The identification, life history, and distribution of *L. pipiens*, commonly known as the northern leopard frog, is summarized in Section 1.2. Confirmation and documentation[6] of species for each *batch* of organisms to be used for breeding or testing received from a supplier or collected from the wild must be made by qualified personnel (i.e., taxonomist) experienced with identifying anuran species. Species identification can be made using distinguishing taxonomic features described and illustrated in taxonomic keys, or using DNA-based taxonomic identification (i.e., barcoding) (see Section 1.2).[7] Organisms that are purchased from a commercial supplier should be supplied with certification of the organisms' species identification, and the taxonomic reference or name(s) of the taxonomic expert(s) consulted. After the initial taxonomic identification of the species provided by a given supplier, confirmation of the species of organisms in a shipment from that same supplier can be conducted by the testing laboratory. All information needed to properly identify the organisms transported to a testing laboratory must be provided with each shipment. Records accompanying each batch of organisms purchased from a biological supplier or another laboratory must include, at a minimum: the estimated quantity and source of organisms in each shipment, supplier's name, date of shipment, date of arrival at the testing laboratory, arrival condition (i.e., appearance, mortality, temperature, and if shipped in water, DO and pH) and species identification. For field-collected organisms, records must also include date and time of collection and location, and should include conditions at collection site.

Different clades or designatable units (DU; see Section 1.2) of the same species might have different sensitivities, although this has not been investigated. For greater standardization, the same genetic group should be used over time within the laboratory.[8, 9]

The toxicity test described herein must be started using *larvae* (i.e., *tadpoles*) at Gosner stage 25 for the 14-day test and Gosner stage 28 to 29 (hereafter referred to as 28/29) for the 42-day test (see Section 2.3.8). The different stages of development in the early life of frogs are important and integral to the test option to be chosen. The first test option is a 14-day test. This shorter test option starts with newly

---

[5] Although this test method could be used to test other anuran species (see Section 4.6.3), only *L. pipiens* was validated for use with this biological test method (Nautilus Environmental, 2020a, 2020b).

[6] Acceptable forms of documentation include identification of laboratory specimens by a qualified taxonomist, and identification of laboratory specimens by molecular analysis (such as DNA barcoding).

[7] Species verification based on molecular methods is possible (Hoffman and Blouin, 2004; Wilson *et al.*, 2008); however, its availability to commercial labs is unknown and DNA barcoding may not be readily accessible.

[8] The taxonomy of *L. pipiens* is complicated by the existence of distinct clades or designatable units in both Canada and the US. This can be complicated further by the potential genetic differentiation between local populations separated by short distances (≥45 km) (Wilson *et al.*, 2008). Another consideration in the case of field-collected organisms is the conservation of local populations. Over-collection of wild populations from local sources might deplete distinct populations. Also, it may be difficult to acquire test organisms representing the western clade due to their conservation status (see Section 1.2). Extra care should be taken when purchasing *L. pipiens* from US suppliers to ensure test organisms are indeed *L. pipiens*, and information as to which clade (east or west) they belong should be obtained, if possible.

[9] This method has been developed using test organisms from both the eastern and western populations of *L. pipiens*.

feeding tadpoles (GS 25) and covers the beginning of hindlimb bud development (GS 26–29). These early developmental stages also include a period of general growth of the larvae.[10] The 14-day test is particularly useful for measuring effects on growth; however, it incorporates a relatively short portion of the aquatic life stage of *L. pipiens*. The second test option is a 42-day test. This longer test option starts with tadpoles at GS 28/29 (i.e., with developed hindlimb buds) and covers stages of toe differentiation and development (GS 31–40).[11] This test includes a period of metamorphic changes that includes the growth and development of external limbs. These metamorphic changes involve several endocrine processes, including but not limited to the hypothalamic-pituitary-thyroid (HPT) and the hypothalamic-pituitary-gonadal (HPG) systems. The 42-day test is a much longer exposure and is designed to characterize adverse effects of chemicals on these systems and others involved (directly or indirectly) in the metamorphic process. Alternative test designs that describe test initiation with different life stages, longer-term tests (i.e., a 56-day test that is essentially a combination of the 14- and 42-day test options), tests that continue for longer (i.e., through metamorphosis), or inclusion of additional endpoints are described in Section 4.6.3; these are more investigative in nature and have not been validated for standardized use in this test method document.

The generalized development of anuran frog species from *egg* to *metamorph* is described by Gosner (1960) and is shown in Figures 2.1 and 2.2. The classification includes 46 stages, with the first 25 based on a scheme developed by Shumway (1940). Healthy eggs are black on top and white at the bottom when released, and have a gel coating. If the eggs are fertilized, the black area will increase in size until the eggs are completely black. Thereafter, the eggs will begin to elongate, and movement will be seen in the egg mass. The gel coating will eventually breakdown, and the embryos will hatch as *hatchlings* at approximately stage 20 (GS 20). Hatchlings are normally dark in colour. GS 25 (tadpoles) can be identified by the complete loss of external gills (see Figure 2.1; left *operculum* closes last), and it is at this point that tadpoles begin to feed. From GS 25 until the beginning of metamorphosis (GS 42) is the longest part of the larval period, and the stages are generally identified by growth and limb development. Hindlimb development is visible from GS 26 to 40, whereas forelimbs develop internally (forelimbs become apparent under the skin at GS 41; see Figure 2.2), erupt suddenly, and typically erupt asynchronously (i.e., one forelimb emerges hours or days before the second one). Gosner stage 42 (GS 42) is marked by the appearance of the second forelimb (see Figure 2.2). In the later stages of metamorphosis (GS 43–46), the organisms undergo reabsorption of the tail and an increase in mouth size, with replacement of larval feeding structures by adult jaws and tongue (see Figure 2.2). Also, forelimbs and hindlimbs become functional (McDiarmid and Altig, 1999). The tests described herein are designed to determine effects on *L. pipiens* at various stages of development, ranging from the beginning of tadpole feeding (GS 25) through to limb development and toe differentiation, depending on the test option selected (see Section 4.3.1). To assist laboratory staff in planning for the beginning of the two test options, approximate time for development to GS30 at test temperature has been calculated (Table 1).

[10] In the first round of ECCC's inter-laboratory method validation study involving four participating laboratories, the median Gosner stage in control organisms at the end of a 14-day exposure ranged from 27.3 to 29. In addition, the wet weight of control tadpoles increased significantly (i.e., by 7- to 10-fold) during the 14-day exposure. Total body length and snout-vent length also increased, but less dramatically, with increases of 24-50% for total length and 4–33% for snout-vent length across the four participating laboratories (Nautilus Environmental, 2020a).

[11] In the third round of ECCC's inter-laboratory method validation study involving three participating laboratories, the median Gosner stage in control organisms at the end of a 42-day exposure ranged from 35 to 41. Two laboratories had 10% and 37% of control tadpoles, respectively, reach Gosner stage 42 (marked by the appearance of the forelimbs). Although growth was observed during the 42-day exposure, it was not as pronounced as that observed in the 14-day test carried out in round 1 of the inter-laboratory investigation. These data suggest that increases in wet weight are more pronounced in the earlier stages (i.e., GS 25 to GS 29) of development (Nautilus Environmental, 2020b).

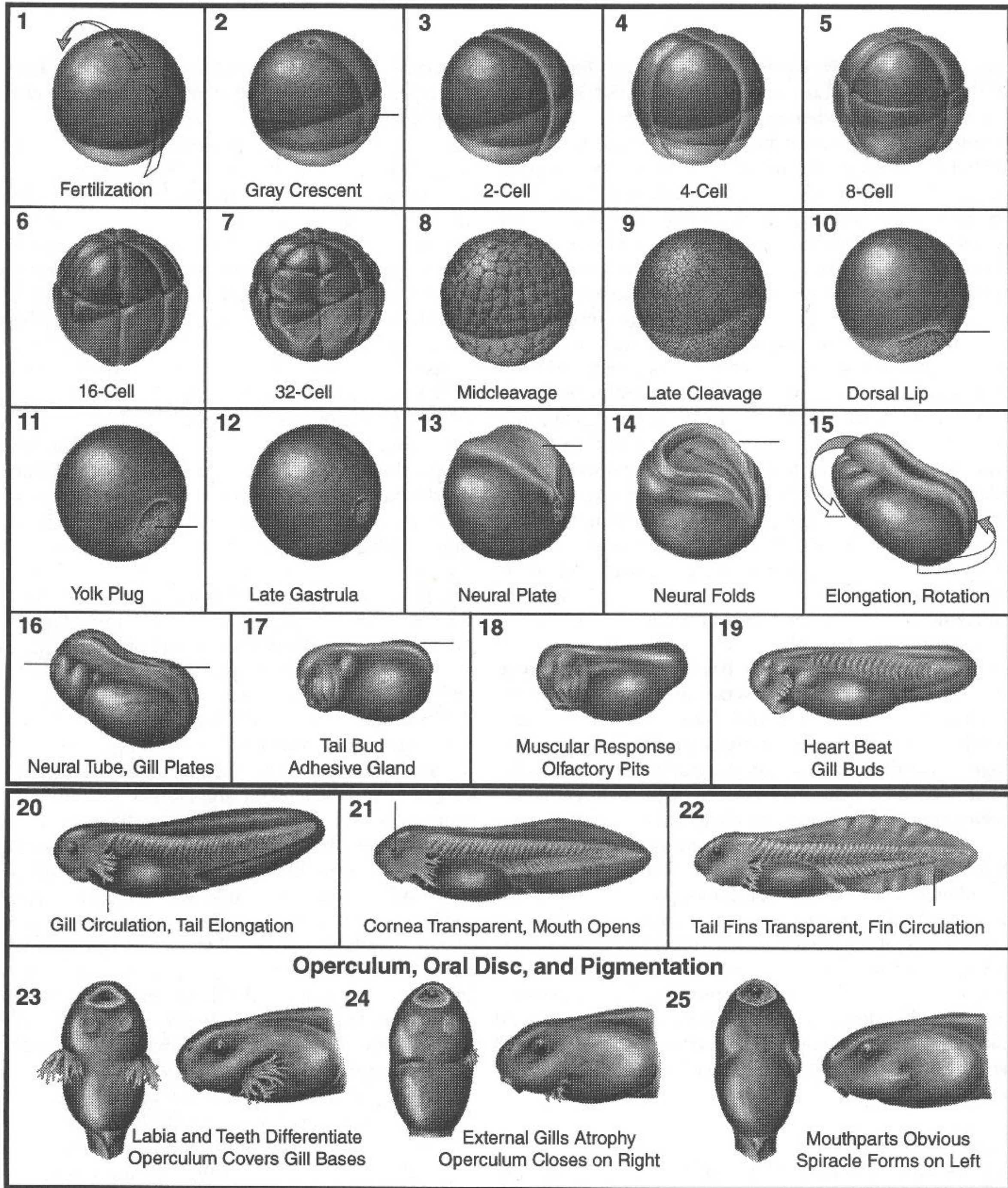


Figure 2.1 **Developmental stages (GS 1 to GS 25) of anuran embryos and hatchlings as described by Gosner (1960). (Reproduced from McDiarmid and Altig, 1999, Figure 2.1, with permission from University of Chicago Press).**

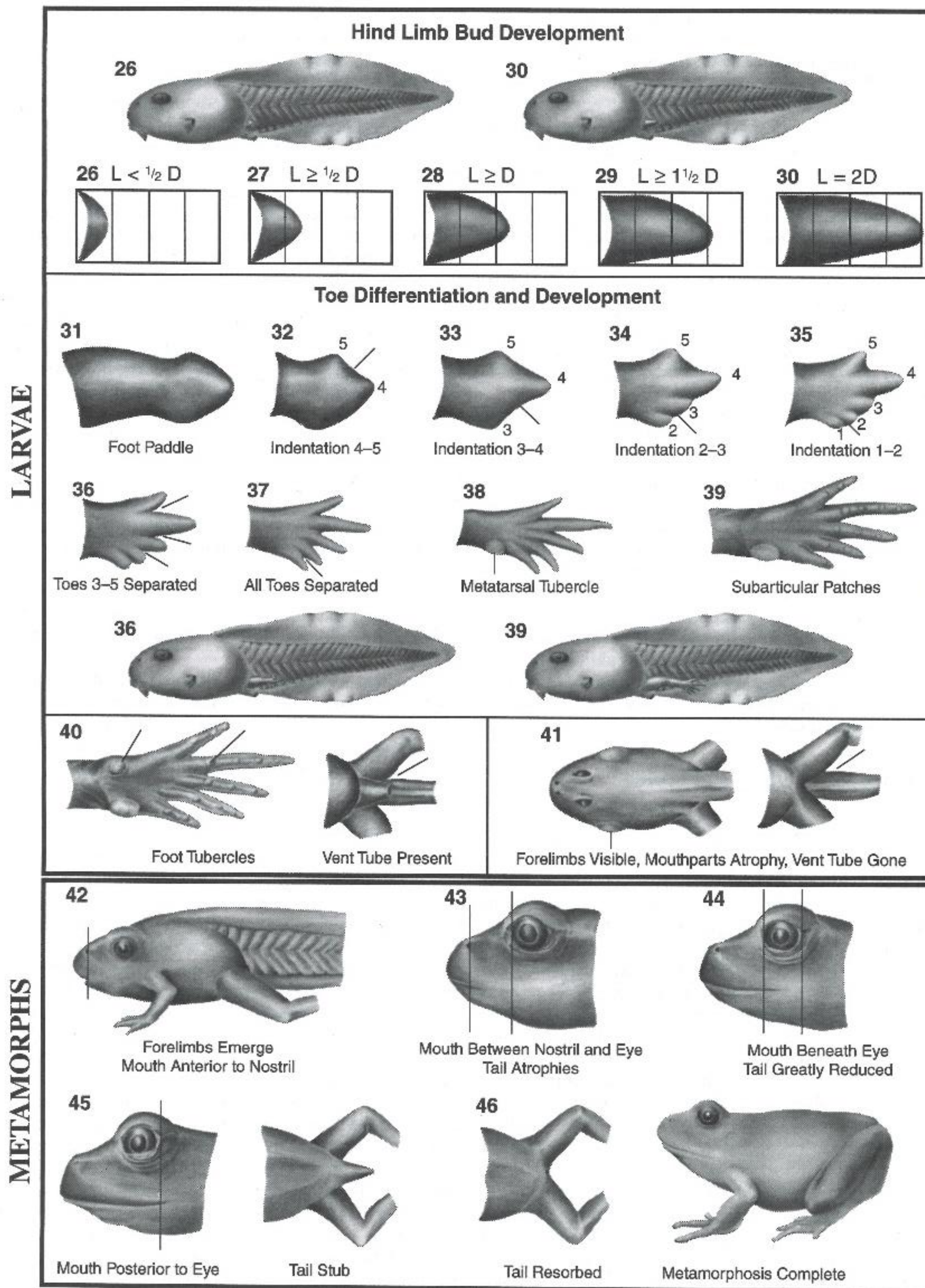


Figure 2.2 Developmental stages (GS 26 to GS 46) of anuran larvae and metamorphs as described by Gosner (1960). (Reproduced from McDiarmid and Altig, 1999, Figure 2.1, with permission from University of Chicago Press).

Table 1 Approximate timing of developmental stages of *Lithobates pipiens* at test temperature, based on ASTM (2022a) and data from inter-laboratory studies (Nautilus Environmental, 2020a, 2020b)[12]

Stage	~Age (d) at 23 ± 1 °C	Major Characteristics/Formations of the Stage
1	0	Prior to fertilization
9/10	1	Late cleavage/ Appearance of dorsal lip of blastopore
13	2	Neural plate, blastopore forming slit
16	3	Neural tube
18	4	"Tadpole" shape becoming distinct; muscular response to stimulation
19	5	Heartbeat; external gill buds; hatching begins
20	6–7	Complete hatching; swimming upon physical stimulation; capillary circulation of first gill
21	7	Mouth open; transparent cornea; tail length approximately equal to length of head and body
22	8	Transparent epidermis; capillary circulation in tail; asymmetrical appearance from dorsal aspect; left gill filaments more apparent
23	8–9	Opercular fold apparent; asymmetrical from ventral aspect
24	9–10	Operculum covering right external gills; external gills on left side still apparent; sucker represented by two small prominences
25	10–12	Operculum complete; no external gill filaments; sucker represented by two pigmented patches; begin feeding; gut clearly visible
26–30	≥ 14–21	Hindlimb buds appear and grow progressively larger; spiracle present on left side (most North American tadpoles)

[12] Development of *L. pipiens* progresses at a much slower rate at lower temperatures. For example, during round 1 of inter-laboratory testing, three laboratories held eggs/larvae at 10–15 °C for 3–4 weeks to delay development before beginning a test with GS 25 tadpoles; this was successful and no observations of increased mortality or stress were observed, except in one laboratory that experienced a chiller failure (Nautilus Environmental, 2020a). In addition, during round 3 of inter-laboratory testing, two laboratories held egg masses at 15–16 °C for 28 or 35 days to delay hatching (Nautilus Environmental, 2020b).

2.2 *Source*

Organisms to be *cultured* for use in testing can be acquired as embryos or larvae. Alternatively, mature and gravid adults can be obtained for immediate spawning, or for hibernation and future spawning, in order to provide test organisms. Embryo, larvae, or adult *L. pipiens* may be acquired from a commercial biological supplier, or government or private laboratories that are known to have disease-free stock. Less desirably, *L. pipiens* embryos or adults may be acquired by collection in the field, but careful identification is required to separate this species from similar ones (see Section 1.2).[13] In addition, the genetic diversity of different sources of *L. pipiens* (e.g., eastern vs. western clades, diversity in local populations) should be taken into consideration when obtaining organisms for breeding or testing (see Section 2.1). Parasites and disease are not uncommon in embryos and larvae, but are more likely in adult frogs collected from the wild. All field-collected organisms should be carefully examined, quarantined, and acclimated as necessary prior to use in tests or prior to providing test organisms for use in tests. Any site from which field-collected organisms are taken should be known to be clean and free of any sources of chemical contamination during the past five years or longer.

Laboratory-cultured test organisms are preferable for a number of reasons, including: the taxonomy is known, they are known to be disease-free, and the impact to wild populations is minimized. However, the successful maintenance of a breeding population of frogs requires a tremendous amount of resources and is relatively uncommon.[14]

Commercial sources for *L. pipiens* include:[15]

Boreal Science (Canadian supplier for Ward's® tadpole food)
399 Vansickle Road
St. Catharines, ON
L2S 3T4 Canada
Tel.: 1-800-387-9393
Fax: 1-800-668-9106
Website: www.boreal.com/store/
Email: borealcs@vwr.com
Stage: adults: (available Sept. to July)

Carolina Biological Supply Company
International Sales Department
PO Box 6010
Burlington, NC 27216-6010
USA
Tel.: 1-336-586-4399
Website: www.carolina.com/
Email: internationalsales@carolina.com
Stage: embryos: (available Jan. to June)
adults: (available Oct. to June)

For current information on suppliers for *L. pipiens*, contact:

Method Development and Applications Unit
Science and Technology Branch
Environment and Climate Change Canada
335 River Road
Ottawa, ON K1A 0H3
Email: methods@ec.gc.ca

All tadpoles used in a toxicity test must be derived from the same *batch* and source. Embryos or larvae received from commercial suppliers or collected in the wild should be handled as minimally as possible

---

[13] Since it may be difficult to distinguish between the eggs of *L. pipiens* and related anuran species, collectors should be well trained in species' habitats and identification (ASTM, 2013). Those collecting test organisms must comply with all federal, provincial, and/or regional/municipal regulations and be in possession of relevant collecting permits, if required.

[14] The culturing of metamorphs and long-term holding of sub-adults and adults to provide a breeding population has been successfully carried out in the laboratory by the Atlantic Laboratory for Environmental Testing (ALET). ALET faced significant challenges in maintaining a *L. pipiens* population during all parts of their life cycle, including food types and feeding rates, water quality maintenance, and disease control. Several documents are available that provide guidance and procedures for maintaining laboratory cultures of *L. pipiens* and include: ALET 2004, 2006, 2009, 2013, 2018. In addition, general guidance on frog husbandry may prove useful for establishing and maintaining cultures of frogs in the laboratory (Wright and Whitaker, 2001).

[15] The list of commercial suppliers was current as of the date of publication of this method.

(see Section 2.3.7).[16] Suppliers generally package and ship embryos or larvae in water-filled sealed bags or other containers that have been injected with oxygen (dissolved oxygen levels should be maintained above 4 mg/L to avoid stressing the test organisms). Upon receipt, the temperature and appearance of the organisms must be recorded (see Section 2.3.9). In addition, the DO and pH of the shipping water and the tank water into which they will be placed should be measured and recorded. The embryos or larvae can be transferred to gently aerated holding tanks (see Section 2.3.2), and their water temperature should be adjusted gradually (e.g., ≤ 3 °C per day) to the temperature to be used during culturing (see Section 2.3.4) or the temperature to be used for testing (see Section 4.3.3). To minimize the impact of sudden changes in water chemistry, shipping water and laboratory water should be gradually mixed (Bradfield, 2010; ASTM, 2022a).[17] Egg masses can be separated into small sections by cutting the connecting thread with a scalpel or scissors, if required. Ensure embryos are kept under water during any manipulations. Guidance for handling *L. pipiens* embryos and larvae is provided in Section 2.3.7 and should be followed when transferring organisms from an outside source to holding containers (Section 2.3.2). Other conditions during this interim holding period for *acclimation* of embryos or larvae to laboratory conditions should be as similar as possible to those used for maintaining test organisms (Section 2.3).

Adult frogs purchased or collected to supply test organisms should be transported to the laboratory in a cooler lined with wet moss to prevent the frogs from drying out during shipment. The cooler should have holes in it to allow air flow. Several frogs can be placed in a loosely tied-up pillowcase, which is used to prevent the frogs from jumping and injuring themselves, and then placed on the moss. Breeding-ready males should be separated from gravid females (e.g., using pillowcases) during shipping. Shipping and transport containers should be insulated to minimize changes in temperature during transit. Depending on the time of year the frogs were collected or are being shipped, heat packs or ice packs may be placed below the moss.[18] Live organisms should be transported quickly to ensure their prompt (i.e., within 24 h) delivery. Excessive crowding of animals during shipment or transport should be avoided to minimize stress in transit.

Adult frogs can be collected in the spring prior to spawning (i.e., during migration to spawning ponds), and maintained in the laboratory briefly prior to hormone injections and breeding for in-season breeding. Alternatively, adult frogs can be collected in the fall (during fall migrations to hibernation ponds) or in the winter from various hibernacula (Fred Schueler, Bishop Mills Natural History Centre, Bishops Mills, ON, personal communication, 2015), artificially hibernated in the laboratory, and then injected with a neurotransmitter hormone mixture to induce spawning for out-of-season breeding (see Section 2.4.3 and Appendix F). Upon arrival at the laboratory, a batch of adult frogs should be held separately from any existing laboratory population in order to determine the general health of the frogs. It may be necessary to maintain adult *L. pipiens* in quarantine prior to being used to obtain eggs, and to treat adult frogs for disease if present or suspected (see Section 2.4.4 and Appendix E). Further guidance on holding, handling, hibernation, breeding, and quarantine of adult frogs is provided in Section 2.4 and Appendices E and F.

Moving animals from one location to another raises serious questions of introducing non-native species or transporting diseases and parasites. Any proposed procurement, shipment, and transfer of frogs, larvae, embryos, or eggs must be approved, if required by regional, provincial, or federal authorities. For frogs purchased from the USA, procedures from the International Air Transport Association (IATA) must be followed for air shipment. Provincial governments (typically, the Ministry of Natural Resources, e.g., Ontario Introduction and Transfers Committee) might require a permit to collect or import amphibians or their embryos whether or not the species is native to the area. Alternatively,

[16] Hatching success is higher if handling of eggs is minimized (ASTM, 2022a).

[17] Note that large shifts in pH can occur when the shipping bags are opened and degas (Bradfield, 2010).

[18] A max-min thermometer can be included in any shipment of live organisms to confirm that appropriate temperatures were maintained during shipping.

movement or collection of amphibian species might be controlled by a Federal-Provincial Introductions and Transplant Committee. Application for a permit to the appropriate provincial or federal agency might be required, depending on procedures in place locally. A laboratory might be required to provide proof of a veterinarian inspection of the amphibians from the collector or supplier prior to obtaining a permit. For further information on possible Environment and Climate Change Canada permit requirements, contact the inquiry centre: enviroinfo@ec.gc.ca.

The laboratory might be required to establish and use a quarantine section within their facilities where imported organisms can be isolated and all equipment and fluids that come in contact with the organisms can be sterilized and disposed of according to provincial or federal regulations.

In addition, the Canadian Council on Animal Care (CCAC) and various provincial agencies may require annual reports on animal use using live animals for testing.

2.3 Holding and Culturing of Embryos and Larvae

Guidance is provided in this section for holding, culturing, and acclimating embryos and larvae for use in toxicity tests at the required developmental stage.

2.3.1 General

The recommended and required conditions for culturing, holding, and acclimating *Lithobates pipiens* embryos and larvae, summarized in Table 2, are intended to allow some degree of flexibility within a laboratory. While guidance and recommendations are provided herein, explicit directions regarding many aspects of holding/culturing, including choice of holding containers, number of organisms per container, and water-renewal conditions, are left to the discretion and experience of laboratory personnel. Much of Section 2.3 is derived from procedures developed through research conducted by Canadian laboratories; however, general procedures for the husbandry of amphibians (see references in Appendix E) can be consulted and followed if further details are required. In addition, it is recommended that laboratory technicians handle all organisms according to the CCAC (2021) guidelines for amphibians, and/or according to guidance from veterinarians and animal care committees when applicable. Proper handling and care of amphibians is essential to the “refinement” portion of the 3Rs (see Section 4.9).

When initially setting up to perform toxicity tests with *L. pipiens, preliminary tests* (typically, tests using a reference toxicant(s), and/or tests using *control/dilution water*; see Section 3.3.1) can be carried out by the laboratory prior to any *definitive* toxicity tests. The results from these tests can be used to demonstrate the laboratory’s capability to meet test validity criteria and to obtain consistent toxicity results. The results can also be used to establish the reference toxicant concentration(s) to be used as positive control *treatment*(s) in conjunction with definitive toxicity tests (see Section 4.8.2).

Performance-based indices[19] are used to evaluate the suitability of test organisms and the acceptability of the test results. Organisms being held for use in the test must have acceptably low mortality rates (see Section 2.3.8). Organisms must appear healthy and behave and feed normally (see Section 2.3.8). Section 2.3.9 provides guidance on monitoring the health of test organisms pre- and post-hatch. Additionally, those used as *controls* in the test must have acceptably low mortality rates and meet all criteria for a valid toxicity test (see Sections 2.3.9 and 4.7). The acceptability of the *culture* must also be demonstrated by the performance of individuals from the batch of test organisms used to start a toxicity test in one (or more) concentration(s) of *reference toxicant* known to have an adverse effect on the test organisms (i.e., a positive control conducted concurrently with the test, or a multi-concentration *reference toxicity test*; see Sections 2.3.9 and 4.8). If desired, a portion of the batch of organisms can be held at a reduced temperature

[19] Performance-based indices include those related to the survival and condition of cultured organisms intended for use in the test (see Section 2.3.8), criteria that must be met by control organisms for a test to be valid (see Section 4.7), and the performance of groups of animals in reference toxicant tests (see Section 4.8).

(e.g., 10 to 15 °C) to delay development and provide organisms for use in testing at a later date (see Section 2.3.4 and footnote 12). If a batch of organisms fails to meet these performance-based criteria, its cause should be investigated.

A checklist of required and recommended conditions and procedures for culturing, holding, and acclimating *L. pipiens* to produce organisms at the required developmental stage for use in toxicity tests is given in Table 2. Further details on the procedures for the holding, hibernation, and injection of adult frogs to obtain *gametes* for the generation of test organisms is provided in Section 2.4 and Appendices E and F.

2.3.2 Facilities and Apparatus

Embryos and larvae must be held in a controlled-temperature laboratory facility. Equipment for temperature control (i.e., an incubator or a room with constant temperature) must be adequate to maintain temperature within the recommended limits (Section 2.3.4). The culturing area should be isolated from any testing, sample storage, or sample-preparation areas, to avoid contamination from these sources. It must be designed and constructed to prevent contamination of cultures (e.g., elimination of copper or galvanized piping or fixtures that could drip metal-contaminated condensation).

All equipment, vessels, and accessories that might contact the organisms within the culturing facility must be clean, rinsed as appropriate, and made of non-toxic materials (e.g., glass, Teflon™, type 316 stainless steel, nylon, Nalgene™, porcelain, polyethylene, polypropylene). Toxic materials including copper, zinc, brass, galvanized metal, lead, and natural rubber must not come in contact with this apparatus and equipment, or the culturing water.

Embryos and larvae may be hatched and cultured in aquaria, troughs, or tanks made of non-toxic materials such as those listed previously.[20] The depth of water must be sufficient for egg masses to remain fully submerged, and for larvae to move vertically (e.g., 20 cm or more). Substrates are not needed or recommended. The loading density of embryos or larvae in each holding/culturing vessel must be recorded, and should be restricted to prevent overcrowding and any resulting adverse effects on growth and organism health (Gromko *et al.*, 1973; Newman 1986; Denver, 1997; Glennemeier and Denver, 2002). The recommended loading densities for holding/culturing vessels are 1 egg mass per 17 to 20 L, no more than approximately 6 organisms/L for hatchlings up to GS 25, and no more than approximately 3 organisms/L for tadpoles ≥GS 25.

Embryos and larvae being held at the laboratory should be observed frequently (i.e., minimum three times per week; preferably daily). Ideally, records should be maintained documenting details on the maintenance of the animals and observations on their health (see Sections 2.3.7, 2.3.8, and 2.3.9). Mortality among hatchlings and tadpoles must be recorded (preferably daily), and dead test organisms must be removed when observed. The tanks are filled with culture water (Section 2.3.5) and should be equipped with some form of temperature monitoring and aeration. The tanks can also be equipped with an ammonia monitor.[21]

The choice of size and numbers of holding/culturing containers required for *L. pipiens* embryos or larvae might be influenced by the number of organisms being held at the laboratory and the number of test organisms required by the testing facility for one or more series of toxicity tests. Additional holding/culturing containers should be set up as needed to maintain the recommended loading densities.

---

[20] Tanks for culturing tadpoles should provide a large surface area for gas exchange plus aeration.

[21] If tadpoles are being held for longer periods of time, a platform must be added to the tank or the tank tilted on an angle once they near Gosner stage 42 (i.e., front leg emergence; see Section 2.1) to allow the newly metamorphosed frogs to escape from the water and prevent drowning. Alternatively, metamorphs should be transferred to tanks with mixed terrestrial and aquatic habitat as described for adult frogs (see Section 2.4.2).

Table 2 Checklist of required and recommended conditions and procedures for culturing *Lithobates pipiens* embryos and larvae for use in aquatic toxicity tests

Source of *Lithobates pipiens*	– disease-free test organisms (embryos/larvae) from biological suppliers, another laboratory, or field collection (embryos only); must be positively identified to species; all information needed to properly identify the organisms collected or transported to a testing laboratory must be obtained for each batch or shipment; must include, as a minimum: o for organisms purchased from a biological supplier or another laboratory: the quantity and source of embryos or larvae in each shipment; supplier's name; date of shipment; date of arrival at the testing laboratory; arrival condition; species identification o for field-collected organisms: date and time of collection; location; approximate number of embryos collected; conditions at collection site; date of shipment; date of arrival at testing laboratory; arrival condition; species identification
Lithobates pipiens seasonal availability	– for commercial suppliers: January to June (embryos); for feral animals: April to June, generally mid-April to mid-May but extending as late as June further to the north
Acclimation	– temperature, DO, and pH of water in which organisms were shipped and appearance of embryos/larvae recorded upon arrival; transfer gently into holding tanks and gradually acclimate to holding or test temperature (≤3 °C/day)
Holding containers and conditions	– embryos: aquaria, tanks, or troughs with appropriate water depth; no substrate; 1 egg mass per 17 to 20 L; loading density is recorded – larvae: aquaria, tanks, or troughs with appropriate water depth; no substrate; ≤ ~6 tadpoles/L for < GS 25; ≤ ~3 tadpoles/L for ≥GS 25; loading density is recorded
Water source	– uncontaminated ground, surface, or dechlorinated or dechloraminated municipal water; surface water, if used, is filtered and/or UV-sterilized; hardness 10 to 230 mg/L as $CaCO_3$
Water renewal	– static-renewal or flow-through; for static-renewal: 50% renewal ≥ 3 times per week, organisms remain submerged during renewals; tanks are siphoned regularly to remove debris
Water quality	– temperature measured daily; DO and pH for each tank measured regularly (e.g. three times per week, or before and after water renewal); ammonia (target value ≤ 0.2 mg/L un-ionized ammonia), nitrite (target value ≤ 1 mg/L nitrite), conductivity, and total residual chlorine (if dechlorinated or dechloraminated municipal water is used) measured regularly (e.g., weekly or more frequently); hardness, alkalinity, total organic carbon, suspended solids, total dissolved gases, nitrate, metals, and pesticides as necessary to document water quality; flow rate to each holding tank is monitored, preferably daily, if flow-through is used
Temperature	– water temperature of 20 ± 2 °C as daily average, and 20 ± 3 °C as instantaneous; fertilized egg masses may be held at lower temperatures (10 to 15 °C) for up to 5 weeks to slow development if organisms are not intended for immediate use in a test; acclimate gradually to holding or test temperature (≤3 °C/day); larvae held at 23 ± 2 °C for a minimum of 24–36 hours prior to use in a test
Oxygen/aeration	– DO 80 to 100% saturation; maintained by gentle and continuous aeration with filtered, oil-free air
pH	– 6.5 to 8.5
Lighting	– fluorescent, incandescent, or LED; intensity of 100 to 500 lux at the water surface in the holding container; 16-h light:8-h dark photoperiod; 15- to 30-minute transition period between light and dark recommended

Feeding	–	larvae: feed *ad libitum* after GS 25 is reached; recommend providing a 4:1 blended mixture (slurry) of cooked or thawed frozen kale and dried tadpole food (e.g., Ward's®), three times per week; example feeding rates are 2 g kale : 0.5 g tadpole food per 100 tadpoles at GS 25, and 3 g kale : 0.75 g tadpole food per 100 tadpoles at GS 26–29, however, amount is adjusted as necessary (i.e., increase if food is completely consumed each day and decrease if food remains after 2 days); recommend providing dried algal pellets (e.g., Hikari mini algae wafers) at a target rate of 3% of the total body weight of all tadpoles, once per week as a fourth feeding
Organism handling	–	gloves are worn when handling embryos and larvae; organisms handled minimally; large bore transfer pipettes or small plastic spoons used to transfer embryos; small fish net or small cup used to transfer tadpoles
Stage for test	–	14-day test: GS 25
	–	42-day test: GS 28/29
Culture health	–	monitor tanks for mortality at least 3 times per week (preferably daily); assess animal welfare (e.g., feeding and other behaviour, skin colour, and appearance); cumulative mortality rate ≤10% (preferably ≤5%) in the 5 days preceding use in a toxicity test (see footnote 42 in Section 2.3.8); remove dead eggs and dead, diseased, or moribund tadpoles

The information in this table is for summary purposes only. Definitive requirements and recommendations of this test method are contained in the main body of this document.

2.3.3 Lighting

Embryos and tadpoles should be illuminated with a fixed daily *photoperiod* of 16-h light and 8-h dark, using full-spectrum fluorescent, incandescent, or light-emitting diode (LED) lights.[22] Light intensity adjacent to the surface of the solution should be 100 to 500 *lux* with a photoperiod of 16-h light:8-h dark.

A 15- to 30-minute transition period between light and dark is recommended for all life stages of frogs being held in the laboratory.[23]

2.3.4 Temperature

Cultures of embryos and larvae should be maintained at water temperatures of 20 ± 2 °C, as a daily average (instantaneous temperature should be 20 ± 3 °C).[24] Test organisms must be acclimated to 23 ± 2 °C and maintained at that temperature for a

[22] Although excess UV light can cause damage to amphibian skin and eyes, reasonable quantities of UV-A and UV-B may be important for amphibian health (i.e., production of vitamin D3 for calcium metabolism), reproduction, and immunity (Adkins *et al*. 2003; Pough, 2007; Ferrie *et al.*, 2014; CCAC, 2021). LED lights provide low to no UV light (Chang *et al.*, 2012; ASTM, 2022b), which could potentially impact amphibian health. However, there are now specialty LED lights available that emit UV-A and UV-B (e.g., ZooMed products), which could be used if there are concerns about providing appropriate levels of UV light to amphibians during culturing and testing. Note that LED lights were not used during the inter-laboratory studies.

[23] A "dawn/dusk" transition period is recommended because abrupt changes in lighting intensity startle and stress frogs. Automated control systems are available for dimming and brightening the intensity of fluorescent lights, although they are costly. Alternatively, a secondary incandescent light source, regulated by time clock and automated rheostat, may be used to provide the transition period.

[24] The preferred body temperature (i.e., PBT, defined as the temperature larvae select when placed in a thermal gradient in the laboratory) depends, in part, on the temperature to which the tadpoles are acclimated, with higher temperatures more often preferred in organisms acclimated to higher temperatures. For *L. pipiens*, tadpoles appear to have similar PBT to adults, and the daytime PBT does not appear to differ from the nighttime PBT; however, seasonal differences in PBT have been observed (Ultsch *et al.*, 1999 and relevant citations therein). The influence of temperature on differentiation and growth is significant, with increasing rates of both as the temperature is increased, until an inhibition temperature is reached. For *L. pipiens*, this temperature is 23 °C (Smit-Gill and Berven, 1979 cited in Ultsch *et al.*, 1999).

minimum of 24 to 36 hours prior to use in a test.[25] Fertilized egg masses may be held at lower temperatures (e.g., 10 to 15 °C) for up to 5 weeks to slow down development if they are not to be used immediately in a toxicity test (P. Jackman, Environment and Climate Change Canada, Moncton, NB, personal communication, 2015; Nautilus Environmental, 2016, 2020a, 2020b). When preparing for use of organisms in a test, they must be acclimated to test temperature at a rate not exceeding 3 °C/day. Changes in temperature during this acclimation period must be recorded.

2.3.5 Culture Water

Sources of water for holding and culturing *L. pipiens* can be "uncontaminated" groundwater, surface water, or *dechlorinated* or *dechloraminated* municipal drinking water.[26] The water supply should previously have been demonstrated to consistently and reliably support good survival, health, and growth of *L. pipiens*. Monitoring and assessment of variables such as residual chlorine (if municipal water is used), *pH*, *hardness*, alkalinity, total organic carbon, *conductivity*,[27] suspended solids, dissolved oxygen, total dissolved gases, temperature, ammonia, nitrite, nitrate, metals, and pesticides should be performed as frequently as necessary to document water quality. Dechlorinated water or dechloraminated water may be used with caution, since its quality is often variable and it could contain unacceptably high concentrations of chlorine, chloramines, fluoride, perchlorate, chlorate, copper, lead, zinc, or other contaminants.[28] Notwithstanding, certain laboratories routinely use dechlorinated or dechloraminated municipal water for culturing/holding *L. pipiens* and as test water with no apparent problems (Nautilus Environmental 2020a, 2020b). If municipal drinking water is to be used, effective dechlorination[29] or dechloramination[30] must

---

[25] Although the time required for anuran larvae to fully acclimate to a new temperature has not been documented (Ultsch *et al.*, 1999), 24 to 36 hours is typically used as the minimum time required for measuring the final thermal preferendum (i.e., the temperature ultimately selected by organisms, regardless of previous thermal experience) (Ultsch *et al.*, 1999 and relevant citations therein).

[26] *Reconstituted water* may be used as dilution and control water depending on study objectives (see Sections 5.3 and 6.3); however, due to the large volumes of water required for testing, this may not be practical. If reconstituted water is to be used as dilution and control water, it should be introduced at the start of culturing, i.e., on receipt of the egg mass in the laboratory. Recipes for reconstituted water were not evaluated during method development, and as a result, instructions on preparation and use of reconstituted water are not provided.

[27] If surface water is used for holding, culturing, or testing, it may be particularly important to monitor conductivity because of its relationship to salinity, which can have toxic effects on amphibian embryos and tadpoles (e.g., premature hatching, reduced survival) (Karraker *et al.*, 2008; Haramura, 2016; S. Robinson, Environment and Climate Change Canada, Ottawa, ON, personal communication, 2023).

[28] Copper is particularly toxic to larval amphibians. A concentration of <0.05 mg/L (Odum and Zippel, 2011) is recommended. Analyses of dechlorinated or dechloraminated city water for background levels of fluoride, perchlorate, and chlorate (byproduct of drinking water disinfection) should also be included as these anions are substrates of the iodine transporter of the thyroid gland, and elevated levels of these anions might affect growth and metamorphosis results (Sparling and Harvey, 2006; OECD, 2009). For the culturing of *Xenopus laevis*, ASTM (2019) recommends maximum concentrations of the following metals: cadmium 10 µg/L, lead 5 µg/L, mercury 0.144 µg/L, nickel 25 µg/L, selenium 140 µg/L, and zinc 70 µg/L. Characteristics of acceptable dilution water for testing with *Xenopus laevis* as recommended by OECD (2015) and USEPA (2015) include limit concentrations for the following: particulate matter 5 mg/L, TOC 2 mg/L, un-ionized ammonia 1 µg/L, residual chlorine 10 µg/L, total organophosphorous pesticides 50 ng/L, total organochlorine pesticides plus polychlorinated biphenyls 50 ng/L, total organic chlorine 25 ng/L, aluminum, arsenic, chromium, cobalt, copper, iron, lead, nickel, and zinc 1 µg/L, cadmium 100 ng/L, mercury 100 ng/L, and silver 100 ng/L.

[29] Vigorous aeration of the water can be applied to strip out part of the volatile chlorine gas. That could be followed by the use of activated carbon (bone charcoal) filters and subsequent ultraviolet radiation (Armstrong and Scott, 1974) for removing most of the residual chloramines and other chlorinated organic compounds. Aging the water in aerated holding tanks for one or two days might be of further benefit.

[30] Unlike chlorine, chloramines do not off-gas or evaporate, and are not reduced by aging water; instead, chloramines can be removed using sodium thiosulfate or carbon-based filters (CCAC, 2021). Alternatively, laboratories may treat municipal water with conditioner

rid the water of any harmful concentration of residual chlorine or chloramines, and regular (i.e., minimum weekly, recommend daily) monitoring for total residual chlorine (TRC) is required.[31] In addition to measurement of chlorine, monitoring of egg and tadpole health can provide evidence of satisfactory water.

Hard water can cause skin problems (e.g., lesions) in some species of amphibians by disrupting normal osmotic regulation. Hard water (364 mg/L as $CaCO_3$) has also been shown to cause spinal malformations (Budischak *et al.*, 2009). Most amphibians show a preference for soft water, but this is species-specific (Whitaker, 2001; Odum and Zippel, 2011). Regardless of the source of water, the hardness of culture/holding water for *L. pipiens* should be between 10 and 230 mg/L as $CaCO_3$ (see Section 6.3).[32]

The pH of the water used for holding and culturing *L. pipiens* should be within the range of 6.5 to 8.5.[33] The dissolved oxygen (DO) content of the water within holding and culturing tanks should be 80 to 100% air saturation. Gentle and continuous aeration of the tanks (e.g., 6.5 ± 1 mL/min·L) should be carried out using filtered, oil-free compressed air. Air to culture/holding tanks should be dispensed through disposable airline tubing and disposable glass or plastic pipettes, or for large volume tanks, aquarium supply airstones. Vigorous aeration should be avoided.

If surface water is used for holding/culturing, it should be filtered (e.g., ≤60 µm) to remove potential predators and competitors of embryos and larvae. A conventional sand filter or commercial in-line filter (e.g., 0.45 to 5 µm) would also be suitable for finer filtration. Ultraviolet (UV) sterilization is recommended to

---

(e.g., Seachem® Prime®) and age for 2 days before using in holding or testing; if this approach is used, it is recommended to conduct chemical analyses with untreated and treated water prior to definitive toxicity testing to confirm the conditioner would not bind the test chemical or material (S. Robinson, Environment and Climate Change Canada, Ottawa, ON, personal communication, 2023).

[31] The guideline value for total residual chlorine for the protection of freshwater aquatic life is 0.5 µg/L (CCME, 1999). Values greater than 0.5 µg/L might risk interaction of chlorine/chloramine toxicity with the contaminant(s) being tested. CCAC recommends values <10 µg/L (CCAC, 2021). The limit of detection for the analytical technique used to measure residual chlorine or chloramines in the treated supply of dechlorinated or dechloraminated water should ideally be low enough to assure that residual chlorine is ≤0.5 µg/L; however, this might be unrealistic for methods used in the laboratory for routine measurments. Using equipment that can measure down to 20 µg/L is acceptable and achievable (P. Jackman, Environment and Climate Change Canada, Moncton, NB, personal communication, 2022).

[32] ALET (2006) investigated the effects of four different culture waters ranging in hardness from 20 to 150 mg/L as $CaCO_3$ on egg hatching, survival, growth, and metamorphosis of northern leopard frog embryos. There was no significant difference between the four waters used for any of the endpoints measured, indicating that waters within the hardness range of 20 to 150 mg/L are appropriate for culturing and testing *L. pipiens*. During the inter-laboratory study, test organisms were successfully cultured in water with hardness values as low as 10 mg/L as $CaCO_3$ and as high as 230 mg/L as $CaCO_3$. These data agree with general guidelines for maintaining and culturing amphibians (Whitaker, 2001; Odum and Zippel, 2011).

[33] The objective of holding and culturing is to provide conditions that are favourable for the amphibians. Although most North American ranids are relatively tolerant of acidic conditions (Lacoul *et al*., 2011), the northern leopard frog is particularly sensitive to acidification (Freda and Dunson, 1984 and Freda and Taylor, 1992 cited in Rowe and Freda, 2000; Freda and McDonald, 1990). Its eggs cannot develop normally at pH ≤5.8, sperm has been shown to have decreased motility at pH <6.5, and it has been shown to have a lethal threshold (pH>4.0) that is higher than other North American ranid species (Schlichter, 1981; Freda, 1986; Lacoul *et al*., 2011). *L. pipiens* has been shown to avoid water with a pH of 4.0, but not pH 4.5 (Freda and Taylor, 1992 cited in Row and Freda, 2000). Between pH 6.8 and 7.5, there was no difference between the developmental success of eggs exposed to various pH levels for 48 hours (Schlichter, 1981). In another study, embryos, hatchlings, and 3-week-old tadpoles showed low mortality in controls at pH 6.5 when exposed for 4 or 5 days (Freda and McDonald, 1990). Generally, amphibian embryos are the most sensitive to low pH, with tolerance increasing with larval age (Pierce, 1985 cited in Horne and Dunson, 1995; Freda, 1986; Freda and McDonald, 1990). The pH limits recommended herein for holding and culturing *L. pipiens* are in keeping with these findings and the general recommendations for culturing amphibians (Whitaker, 2001; Odum and Zippel, 2011; CCAC, 2021).

reduce the possibility of introducing pathogens to the colony of frogs.

The water in vessels containing embryos, larvae, or adults must be renewed using *static-renewal* or *flow-through* systems to prevent a build-up of metabolic wastes.[34] For static-renewal (i.e., no filter and water quality maintained simply through water changes), 50% of the water in the culture tanks should be siphoned out and replaced with clean culture water a minimum of three times per week, or more often if there are water quality problems. Embryos and larvae must remain submerged during renewals. For flow-through, tadpoles and holding vessels should be carefully observed for potential challenges associated with water flow rates that may be too high (e.g., stunted growth due to increased energy required to move around the vessel; fouling due to food accumulation in specific areas of the vessel) (M. Gallant, Nautilus Environmental Inc., Burnaby, BC, personal communication, 2023). Flow rate should be monitored, preferably daily. Ammonia and nitrite should be measured frequently to check that they do not reach harmful levels. Target values for holding/culturing *L. pipiens* recommended herein are ≤0.2 mg/L of un-ionized ammonia[35] and ≤1 mg/L nitrite.[36] Water entering the holding/culturing containers should not be supersaturated with gases. Remedial measures must be taken (e.g., use of aeration columns or vigorous aeration in an open reservoir) if dissolved gases exceed 100% saturation.

In order for normal metamorphosis to occur, iodide must be available to frog larvae in sufficient quantities for thyroid hormone synthesis. This can be made available to the larvae through water or dietary sources, or both (USEPA, 2015). Currently there are no empirically derived guidelines for minimum iodide concentrations in either food or water to ensure proper development (USEPA, 2015). Iodide levels should be measured in culture/control/dilution water (and, if desired, tadpole food) to monitor exposure levels. Levels of iodide ranging from 1 to 3.3 µg/L have been measured in control/dilution waters used by several

---

[34] A static-recycled system may be used for holding embryos or larvae; however, this type of system was not evaluated during the development of this method nor during the inter-laboratory studies. As a result, no guidance on static-recycled systems for holding embryos or larvae can be provided. For the static-recycled system, a filter suitable for removing metabolic wastes is used and recycled water is filtered to remove solid waste and biobeads. An ammonia remover is often added to the filter to control ammonia and nitrite concentrations in the water (CCAC, 2005; Timmons *et al*., 2018).

[35] The recent CCAC guidelines for amphibians (2021) recommend un-ionized ammonia levels of <0.02 mg/L. The CCME (2010) guidelines recommend a value of 0.019 mg/L of un-ionized ammonia based on the most sensitive freshwater study identified, which was the toxicity of ammonia to rainbow trout (*Oncorhynchus mykiss*). In a study carried out by Jofre and Karasov (1999), un-ionized ammonia concentrations of >1.5 mg/L had negative effects on *L. pipiens*, negatively affecting hatching and percent deformities, growth, and development in embryos exposed for 5 days. The authors indicated that anurans may not be as sensitive to ammonia as some species of fish, including rainbow trout, with effects on embryos seen at concentrations 10- and 100-times lower than those observed in their study for *L. pipiens*. In a more recent investigation into the effects of ammonia on *L. pipiens* larvae, GS 27 tadpoles were exposed to varying concentrations of ammonia in a 21-day test (ALET, 2015). Results showed that mean un-ionized ammonia levels of 0.23 mg/L (ranging as high as 0.33 mg/L) had no effect on the survival, growth (wet weight, snout-vent length, and tail length) and development (Gosner stage) of tadpoles. Un-ionized ammonia values in the controls remained ≤0.01 mg/L for the duration of the 21-day exposure, with 10 tadpoles per 6-L replicate and three-times weekly water renewals (ALET, 2015). The recommended target of ≤ 0.2 mg/L is in keeping with these findings and the general recommendations for culturing amphibians recommended by Odum and Zippel (2011).

[36] The CCME guideline value for nitrite is 0.06 mg/L (CCREM, 1987). In a study examining the effect of nitrite on the cascades frog (*Rana cascadae*), GS 39–40 tadpoles exposed to a nitrite ($N\text{-}NO_2^-$) concentration of 3.5 mg/L for 14 days developed more slowly and emerged at an earlier developmental stage. In addition, they occupied shallow water more frequently. There was no effect on time at emergence or snout-vent length at emergence at the same exposure level (Marco and Blaustein, 1999). The recent CCAC guidelines for amphibians recommend nitrite levels of <1 mg/L and nitrate levels of <50 mg/L (CCAC, 2021). The recommended value of ≤1 mg/L nitrite is in keeping with these findings and the general recommendations for culturing amphibians recommended by Whitaker (2001) and Odum and Zippel (2011).

Canadian laboratories during tests with *L. pipiens*.[37] These levels should be used as guidelines to ensure that adequate iodide is available to allow for appropriate functioning of the thyroid gland, and therefore allow for proper development.

The quality of water in culture tanks must be monitored and recorded routinely. For embryos and larvae, water temperature must be measured daily. Dissolved oxygen and pH must be measured at regular intervals in order to document water quality.[38] Regular (e.g., weekly or more frequently if necessary) monitoring of levels of ammonia, nitrite, conductivity, and total residual chlorine (if municipal water source) is recommended. Culture water hardness and alkalinity should be measured as frequently as necessary to document water quality. It is recommended that these variables be measured at least once during the period of culturing tadpoles as well as on the day before the start of a test.

2.3.6 Food and Feeding

Various types of food and feeding regimes have been used for culturing tadpoles in preparation for toxicity tests (ALET, 2004; APHA *et al*., 2011; Nautilus Environmental, 2014). A detailed feeding experiment was conducted, comparing four different diets, beginning at Gosner stage 25. A mixture of kale and dried tadpole food delivered three times a week and one or more algal pellets delivered once a week was determined to be the optimal diet (Nautilus Environmental, 2016). This diet has been successfully used at ALET and was successfully used to feed tadpoles during the inter-laboratory studies.

Tadpoles must be fed *ad libitum* once they reach Gosner stage 25 (see Section 2.1). The recommended food type for *L. pipiens* is a 4:1 mixture of kale and dried tadpole food (e.g., Ward's® Food for *Xenopus* Tadpoles, available from Boreal Science or VWR International, catalogue no. 470030-346).[39] The kale (i.e., curly) is de-stemmed and the leaves boiled or steamed for several minutes to soften. The kale is then drained or patted dry, and cooled prior to feeding; this prepared kale can be frozen in small aliquots for up to 1 year. Alternatively, organic kale purchased frozen can be used without steaming. Frozen kale is thawed before use and most of the water squeezed out prior to weighing. A slurry of kale and tadpole food can be prepared by coarsely blending the mixture with a small amount of control water.[40] Tadpoles should be fed the kale/tadpole food diet three times per week. The amount will vary depending on the number and size of tadpoles in each culture tank; example feeding rates are 2 g kale : 0.5 g tadpole food per 100 tadpoles at GS 25, or 3 g kale : 0.75 g tadpole food per 100 tadpoles at GS 26–29, per feeding. The amount can be increased if the food is entirely consumed each day, however, if food remains after 2 days, the amount should be decreased. Once per

---

[37] During ECCC's inter-laboratory study, participating laboratories monitored iodide levels in clean control/dilution water and in control replicates just prior to water changes (i.e., after 2 days of static exposure with feeding). These levels ranged from 1.0 to 2.3 µg /L iodine in the clean water and from 2.1 to 3.3 µg /L iodine in the 2-day-old water (Nautilus Environmental, 2020a). These data indicate that food and/or variability in levels found in control/dilution water (i.e., dechlorinated tap water) can impact the iodine levels to which the tadpoles are exposed, as these values can fluctuate over time.

[38] Even a one-day unexpected increase in temperature due to equipment malfunctioning can impact test organism health (L. Van der Vliet, Environment and Climate Change Canada, Ottawa, ON, personal communication, 2021). Because short-term unexpected changes have been shown to impact test organism health, temperature must be monitored daily. In addition, temperature may be adjusted daily when acclimating egg masses, and these daily changes must be recorded (see Section 2.3.4). Examples of regular intervals for monitoring for dissolved oxygen and pH include three times a week, or before and after water renewal.

[39] ALET investigated several tadpole diets, including frozen spinach (thawed), boiled romaine lettuce, Ward's dry tadpole food, Tetramin, trout chow, algal pellets, newly hatched *Artemia*, and boiled kale. During feeding experiments, they determined that the kale and tadpole food mixture resulted in 62.5% of tadpoles reaching metamorphosis compared to ≤37.5% of tadpoles in other treatments, and thus it was chosen as the primary tadpole diet (ALET, 2004).

[40] It is important to not over-blend the kale tadpole food mixture, as the leaves of kale are useful for providing a visual indicator of *ad libitum* feeding (Nautilus Environmental, 2020a).

week (the fourth feeding per week) the tadpoles should be fed dried algal pellets (e.g., Hikari mini algae wafers) instead of the kale and tadpole food mixture. Dried algal pellets are also fed *ad libitum*, and estimated amounts normally start at a target rate of 3% of the total body weight of all tadpoles in a given tank. The algae pellets may be broken up or partially ground prior to feeding.

2.3.7 Handling Organisms and Maintaining Cultures

Embryos and larvae should be handled as little as possible, to avoid damage and undue stress. When handling is necessary, it should be done gently, carefully, and quickly to minimize stress to the animals. Large bore transfer pipettes or small plastic spoons can be used to transfer embryos. Small fish nets or small cups can be used to transfer tadpoles. Gloves must be worn when handling embryos or larvae in order to prevent the potential transmission of harmful materials or pathogens to the organisms. The use of rinsed, powder-free gloves is recommended, and latex gloves must not be used unless proven to be non-toxic in a laboratory investigation (see Section 2.4.4 and Appendix E.2.1 for additional details).

It is recommended that the contents of each holding container be inspected just before each feeding, to determine the apparent condition of the test organisms and the holding tanks. Organism welfare should be assessed at this time (e.g., changes in feeding or other behaviour, social interactions such as aggression, and appearance such as skin colour); recommendations of welfare indicators are available in Section 9.1 of the CCAC guidelines for amphibians (CCAC, 2021), and additional information on abnormal tadpole behaviour and appearance is described in Section 4.4. Records should be kept of the apparent condition of the culture (organisms and holding/culture tanks) noted during each observation period (see Section 2.3.2). The number of embryos and hatchlings must be estimated and recorded, as well as the number of dead, diseased, or moribund organisms removed from each holding/culture tank. When removing dead hatchlings, extreme care should be taken not to bump or damage adjacent embryos since they are extremely delicate and sensitive until Gosner stage 19 or 20 (ASTM, 2022a). Suggested procedures for the assessment of individual tadpoles (i.e., developmental stage, length, and weight) prior to test initiation, during interim test measurements, and at the end of the test are provided in Section 4.2.

Tanks should be siphoned regularly to remove solid wastes and debris. The water in the embryonal and larval holding/culturing vessels should be renewed as described in Section 2.3.5.

2.3.8 Test Organism Health Criteria

Tadpoles being cultured or held for use in a test must be checked daily. Individuals that appear unhealthy (e.g., discoloured, bent, abnormal appearance), inactive, stressed, or dead when gently prodded must not be used for testing.[41] Organisms being held for use in the test should have ≤ 5% cumulative mortality, and must have ≤ 10% cumulative mortality, in the 5 days before the start of the test (ECCC, 2023).[42]

---

[41] It is possible for organisms to appear healthy prior to and during testing at earlier life stages, but then exhibit low survival and high deformity rates at later stages of development (e.g., > GS 33); this is of particular concern for field-collected egg masses, where environmental conditions such as temperature, UV exposure, fertilization rate, and disease exposure are out of the laboratory's control (M. Gallant, Nautilus Environmental, Calgary, AB, personal communication, 2023). If multiple egg masses are available, it may be beneficial to use 2–3 egg masses for a test in order to protect against this possibility (see Section 4.2).

[42] Sample calculation: On Day 0, the test is started. At the end of Day -6, there are approximately 500 hatchlings.

Day	Number of dead organisms
Day -5	5
Day -4	3
Day -3	5
Day -2	2
Day -1	3
Day 0	2
Total	**20**

In total, 20/500 organisms were found dead in the 5 days preceding the test, so the cumulative mortality rate is 4%. Note that mortality can be difficult to accurately assess because dead tadpoles might be eaten prior to assessment (ECCC, 2023). Tadpoles that are removed from the holding/culturing vessel for reasons other than mortality (e.g., tadpoles ≥GS 30 if only GS 28/29 tadpoles are required) are not counted as mortalities in this calculation.

Tadpoles at Gosner stage 25 are actively swimming and feeding, and are characterized by the following physical benchmarks: operculum complete; no external gill filaments; *spiracle* on left side (see Figure 2.1). At test initiation (i.e., for the 14-day test), tadpoles should be within a range of expected measurements for the GS 25 developmental stage. These include: 0.01–0.04 g wet weight, 3–6 mm for snout-vent length, and 10–14 mm for total length.

Tadpoles at GS 28/29 are actively swimming and feeding and have hindlimb buds with a length ≥ the depth of the limb bud (for GS 28) or 1.5 times the depth of the limb bud (for GS 29). At test initiation (i.e., for the 42-day test), tadpoles at these stages of development should be 0.16–0.48 g wet weight, 9–13 mm snout-vent length, and 24–34 mm total length. Tadpoles of similar size should be used to start a test.

Since growth and development rates can vary greatly within a given batch of *L. pipiens*, a number of tadpoles in excess of those needed for use in a test should be cultured at one time in order to provide enough test organisms at the required life stage. For the 42-day test (requiring GS 28/29 tadpoles at test initiation), it is recommended that laboratories culture as many as four times the number of test organisms needed for use in the test (Nautilus Environmental, 2020b).

2.3.9 Health, Quarantine, and Disease of Embryos and Larvae

Tanks containing eggs or tadpoles should be checked daily, during which time culture performance must be monitored and recorded (see Sections 2.3.2, 2.3.6, 2.3.7, and 2.3.8). Procedures and conditions used to maintain each culture should be evaluated routinely, and adjusted as necessary to maintain or restore the health of the culture. If the culture appears unhealthy or atypical during any check, it should be checked more frequently to make sure that "cascade mortality" (i.e., rate of death increasing exponentially over time) is not occurring. Additional information on common diseases in *Lithobates pipiens* are provided in Appendix E.

The appearance of healthy, fertilized eggs is described in Section 2.1. Dead eggs appear white in colour and should only be removed if there are large sections on the edges of the egg mass (see Section 2.3.7). Healthy tadpoles should appear dark in colour, and upon hatching, tadpoles will initially fall to the bottom of the tank with little movement, but will swim actively and attach to the sides of the tank within a day or two (ALET, 2020). Tadpoles should be swimming actively after Gosner stage 20 (hatching) and actively feeding after Gosner stage 25.

Additional measures of organism health can be achieved through use of reference toxicity testing and monitoring of control performance. There are two test options for meeting minimum requirements using a known reference substance (e.g., sodium chloride or thyroxine) and using a portion of the same batch of organisms used to start the definitive toxicity test (see Section 4.8). All tests with the reference toxicant(s) should be performed using the conditions and procedures outlined in Section 4.8. Test-related criteria used to judge the validity of a particular toxicity test (and, indirectly, the health of the culture), based on the performance of test organisms in the *control/dilution water*, are given in Section 4.7.

A laboratory that routinely (e.g., several times per year or more) performs toxicity tests with amphibians might find it useful to monitor the data on survival, growth, and development in control/dilution water, as a measure of the health of a given batch of test organisms and their performance. The results, plotted over time (i.e., performance control charts) are useful for monitoring the acceptability of the test system and environmental conditions, the proficiency of the technician performing the test, and the performance and health of the test organisms.

2.4 Holding and Breeding of Adult Frogs

Guidance is provided in this section for holding and acclimating adults for use in breeding.

2.4.1 General

The recommended and required conditions for quarantining, holding, acclimating, hibernating, and breeding adult *Lithobates pipiens*, summarized in Table 3, are intended to allow some degree of flexibility within a laboratory. While guidance and recommendations are provided herein, explicit

directions regarding many aspects of holding, including choice of holding containers, number of organisms per container, and water-renewal conditions, are left to the discretion and experience of laboratory personnel. Much of Section 2.4 is derived from procedures developed through research conducted by Canadian laboratories; however, general procedures for the husbandry of amphibians (see references in Appendix E) can be consulted and followed if further details are required. In addition, it is recommended to handle all organisms according to the CCAC (2021) guidelines for amphibians, and/or according to guidance from veterinarians and animal care committees when applicable, to incorporate the 3R of "refinement" into animal handling procedures (see Section 4.9).

2.4.2 Holding of Adult Frogs

Small groups of adult male and female *L. pipiens* may be held in containers provided with appropriate habitat and environmental cues. Adult frogs may be maintained in a variety of containers made of non-toxic materials such as glass, stainless steel, porcelain, fibreglass-reinforced polyester, perfluorocarbon plastics (Teflon™), acrylic, polyethylene, or polypropylene. Rectangular vessels that are 1 m wide × 3 m long × 30–50 cm high are suitable for holding 12–20 adult *L. pipiens.* These holding containers should provide both aquatic and terrestrial habitat (see Figures E.1 and E.2 in Appendix E). The containers should be filled with culture water (see Section 2.3.5) at a level where the animals can submerge completely (i.e., ≥20 cm for adults), and equipped with some form of temperature monitoring and aeration. Terrestrial habitat may consist of river rock and gravel, Plexiglass, or Styrofoam covered by potting soil and damp moss. Potting soil used as substrate must be fertilizer- and pesticide-free. Sheet, sphagnum, or Spanish moss are recommended for use as substrate. Moss should be soaked in culture water before use and may need to be buffered (refer to guidance in footnote 11 in EC, 2014) to ensure its pH is suitable before use (see footnote 33 in Section 2.3.5). Additional materials such as logs, plastic plants, or half-buckets should be included in holding containers to provide additional cover and hiding spaces. Each holding container should have a perforated lid (e.g., mesh) to prevent the frogs from escaping and to allow air exchange and light to enter. Humidity is a consideration for holding amphibians, and CCAC (2021) recommends that facilities be capable of maintaining humidity around 50%. Tanks and equipment should be cleaned and disinfected before introducing a new batch of adult frogs.

Sources of water for holding *L. pipiens* adults can be "uncontaminated" groundwater, surface water, or dechlorinated or dechloraminated municipal drinking water, with hardness of 10 to 230 mg/L as $CaCO_3$, pH of 6.5 to 8.5, and DO of 80 to 100% saturation, as described in Section 2.3.5. The water in the adult holding vessels should be either flow-through with continuous low-flow or static-renewal with ≥50% renewal at least once per week. If a flow-through system is used, the flow rate should be monitored, preferably daily, and if necessary, water can be siphoned regularly to remove solid waste and debris.

Adult *L. pipiens* should be maintained with overhead lighting using full-spectrum fluorescent, incandescent, or LED light, combined with specialty amphibian/reptile light bulbs that provide higher amounts of UV-A and UV-B (e.g., EXO Terra Repti Glo 5.0). Light intensity should be 100–500 lux at the water/terrestrial surface. The photoperiod should normally be 16-h light:8-h dark, with a 15- to 30-minute transition between light and dark is recommended (see footnote 23 in Section 2.3.3). However, during hibernation and breeding (Section 2.4.3), the photoperiod is adjusted periodically according to a schedule, such as that provided in Appendix F (Table F.1). In addition, a red light (e.g., Zoo Med Nightlight Red™ Reptile Bulb, 60 Watt) positioned over an area provided for basking should remain lit at all times.[43]

[43] A red amphibian/reptile basking light not only provides a warm area for basking, which is said to aid digestion, but it also simulates moonlight during dark hours of the photoperiod.

Table 3 Checklist of required and recommended conditions and procedures for quarantine, holding, and breeding *Lithobates pipiens* adult frogs to provide test organisms for use in aquatic toxicity tests

Source of *Lithobates pipiens*	−	disease-free adult *Lithobates pipiens* (to supply test organisms) from biological suppliers, another laboratory, or field collection; must be positively identified to species; all information needed to properly identify the organisms collected or transported to a testing laboratory must be obtained for each batch or shipment; must include, as a minimum: o for organisms purchased from a biological supplier or another laboratory: the quantity and source of adults in each shipment; supplier's name; date of shipment; date of arrival at the testing laboratory; arrival condition; species identification o for field-collected organisms: date and time of collection; location; number of adults collected; conditions at collection site; date of shipment; date of arrival at testing laboratory; arrival condition; species identification
Lithobates pipiens spawning season	−	for feral animals: generally mid-April to mid-May, and extending as late as June further to the north; alternatively, during fall migration or in winter from hibernacula; for commercial suppliers: September to July
Acclimation/quarantine for adult frogs	−	tanks and equipment disinfected prior to introducing new adult frogs; new adult frogs may be held in quarantine for 2 weeks, or longer (i.e., 6 to 8 weeks)
Holding containers/conditions	−	1 m wide × 3 m long × 30–50 cm tall fibreglass tanks for 12–20 adult frogs per tank; terrestrial and aquatic habitat (see Appendix E.1); covered to prevent escape but allow light and air to enter enclosure; water depth ≥20 cm; plastic plants should be disinfected and moss in adult tanks should be replaced once every two months and after any mortalities, and terrestrial substrate cleaned and disinfected or replaced as necessary
Water source	−	uncontaminated ground, surface, or dechlorinated or dechloraminated municipal water with hardness of 10 to 230 mg/L as $CaCO_3$, as described in Section 2.3.5
Water renewal	−	flow-through or static-renewal replacement; flow-through: continuous low flow, equivalent to ≥50% renewal a minimum of once per week; static-renewal: ≥50% renewal a minimum of once per week (preferably twice); if necessary, water is siphoned to remove debris
Water quality	−	temperature measured daily; DO and pH for each tank measured regularly (e.g. three times per week, or before and after water renewal); ammonia, nitrite, conductivity, and total residual chlorine (if dechlorinated or dechloraminated municipal water is used) measured regularly (e.g., weekly or more frequently); hardness, alkalinity, total organic carbon, suspended solids, total dissolved gases, nitrate, metals, and pesticides as necessary to document water quality; flow rate to each holding tank is monitored, preferably daily, if flow-through is used
Temperature	−	acclimated upon receipt to 20 ± 2 °C as daily average, 20 ± 3 °C as instantaneous; achieved at a rate of ≤3 °C/day
Oxygen/aeration	−	DO 80 to 100% saturation; maintained by continuous aeration with filtered, oil-free air
pH	−	6.5 to 8.5
Lighting	−	full-spectrum fluorescent, incandescent, or LED lighting combined with specialty lights for amphibians and reptiles with higher UV-A and UV-B lighting (e.g., Exo Terra Repti Glo 5.0); 100 to 500 lux at water/terrestrial surface; 16-h light:8-h dark photoperiod; recommend 15- to 30-minute transition period between light and dark; additional red basking light on 24 h per day (e.g., Zoo Med Nightlight Red™ Reptile Bulb 60 Watt)

Feeding	– feed 3 times/week a variety of live food including crickets, earthworms, and mealworms; each feeding is the equivalent of 3 insects per frog (i.e., can be a combination of crickets and worms); at least twice per week insects are lightly dusted with a 1:4 mixture of vitamins (e.g., Reptivite™) and $CaCO_3$ prior to use for feeding; any uneaten food accompanied by visible mould or fungi should be removed before feeding
Organism handling	– gloves are worn when handling frogs; frogs handled minimally
Culture health	– monitor holding containers at least three times weekly just before feeding; assess animal welfare (e.g., feeding and other behaviour, skin colour, and appearance); remove dead frogs, and quarantine and treat diseased or moribund adult frogs – maintain records for each frog including: source, weight, date received, disease/medication dosages and dates, hibernation, hormone injections, breeding attempts, and photo of spot pattern to identify individual (alternatively, tagging organisms may be acceptable) – if treated for disease prevention or control, allow at least 2 weeks before collecting eggs for use in toxicity tests
Hibernation	– hibernate frogs in temperature- and photoperiod-controlled chamber (see Appendix F.2); place frogs in plastic containers half-filled with culture water or a 1:20 Ringer's solution, up to a maximum of 6 frogs/20 L; males and females are kept in separate hibernation tanks; adjust temperature and photoperiod as necessary according to a schedule (e.g., Table F.1); flow-through with continuous low flow or static renewal with ≥50% replacement daily for the first two weeks and three times per week thereafter; frogs minimally disturbed – measure air and water temperature daily; measure DO and pH at the time of renewal if static-renewal, or at least 3 times/week if flow-through
Breeding	– frogs are given a priming dose of gonadotropin-releasing hormone agonist (GnRH-A) and then a combined dose of GNRH-A and metoclopramide hydrochloride (MET); hormones are delivered through intraperitoneal injections (see Appendix F.3) – following hormone injections, adults are moved to covered large breeding tanks or aquaria with a terrestrial surface and egg laying substrate (e.g., plastic plants) in the water column; water depth ≥20 cm; up to 10 frogs per tank; male:female ratio of 3:2 is recommended; breeding tanks are monitored daily and any incidences of amplexus or spawning recorded; frogs minimally disturbed

The information in this table is for summary purposes only. Definitive requirements and recommendations of this test method are contained in the main body of this document.

Adults should be held in a facility with an air temperature of 20 ± 2 °C, as a daily average (instantaneous temperature of the facility should be 20 ± 3 °C). However, during hibernation and breeding (Section 2.4.3), the temperature is adjusted periodically according to a schedule, such as that provided in Appendix F (Table F.1).

Success in holding adult *L. pipiens* has been achieved using a variety of live food. As a general rule, live prey must be available and must be as varied as possible (Mattison, 1993). Earthworms and mealworms can be purchased or cultured in the laboratory, and crickets can be purchased from a local pet store (ALET, 2004, 2006, 2009). Frogs should be fed three times per week and each feeding should be the equivalent of three insects per frog (i.e., a combination of crickets and worms may be used). Crickets and worms should be dusted lightly with a 1:4 mixture of vitamins (e.g., Reptivite™) and calcium carbonate at least twice per week before being fed to the frogs. During culture feeding, any old food accompanied by mould or fungi should be removed and discarded. See Appendix E for further details on obtaining and holding insects for feeding adult frogs.

The monitoring of hardness, pH, DO, aeration, and other parameters of water used for holding *L. pipiens* adults should be as described in Section 2.3.5. For adult frogs, the water and air temperatures as well as the dissolved oxygen and pH of the water in each

holding tank should be measured at least three times weekly. Ammonia levels in each adult frog holding tank should be measured at least weekly. Once every two months and after any mortalities occur, the moss in each tank should be replaced and the plastic plants in each tank should be disinfected. The other terrestrial substrate in each adult holding vessel should be cleaned and disinfected or replaced as required.

2.4.3 Hibernation and Breeding

Environmental cues (i.e., temperature and photoperiod manipulations) that stimulate hibernation, gonad development and maturation, and physiological changes in adult frogs that result in gamete production can be simulated in the laboratory in order to prolong the potential supply of test organisms beyond the seasonal breeding period of *L. pipiens*.[44] Breeding is induced with hormone injections (gonadotropin-releasing hormone agonist and metoclopramide hydrochloride), fertilized egg masses are collected, and organisms are cultured for use in testing. Procedures for in-laboratory hibernation and breeding of *L. pipiens* to obtain test organisms are described in Appendix F.

To simulate hibernation, male and female frogs are placed into separate tanks in temperature- and photoperiod-controlled chambers half-filled with culture water or a 1:20 Ringer's solution with a maximum of 6 frogs/20 L. A specific schedule of temperature and photoperiod changes is followed to initiate hibernation (see Appendix F.2). During hibernation, solutions should be renewed by flow-through with continuous low flow or by static renewal with ≥50% replacement daily for the first two weeks and three times per week thereafter. Frogs should be minimally disturbed during in-laboratory hibernation and breeding.

It is recommended that records be maintained for each adult *L. pipiens* being hibernated or held to provide tadpoles for use in toxicity tests. Records should include source, weight, date received, incidents of disease, medication dosages and dates (if applicable), as well as breeding attempts (including hibernation and hormone injections details). Individuals may be identified by photographs of spot patterns, or alternatively through tagging (CCAC, 2021).

Temperature (air and water) should be measured and recorded daily during hibernation. DO and pH should be measured and recorded daily or at the time of renewal for static-renewal, or at least 3 times weekly if flow-through, for each hibernation tank. Each hibernation tank should be monitored daily for mortality and signs of stress or disease. Further details on maintaining frogs for hibernation and to provide embryos through hormone injections is provided in Section 2.4.2 and Appendix E.

Adult female and male frogs to be spawned shortly after arrival at the laboratory (i.e., without hibernation) are kept in separate tanks. The frogs should be acclimated, housed, and fed as previously described (see Section 2.4.2 and Appendix E). The protocol for administering breeding hormones, delivered through intraperitoneal injections, is described in detail in Appendix F.3. This includes a priming dose of gonadotropin-releasing hormone agonist (GnRH-A) followed by a combined dose of GNRH-A and metoclopramide hydrochloride (MET). Following hormone injections, adult frogs are transferred to large, covered breeding tanks with a small terrestrial surface and substrate appropriate for egg laying (e.g., plastic plants), as described in Appendix F. The recommended ratio of male:female frogs in these breeding tanks is 3:2. Each breeding tank should be monitored daily and any incidences of amplexus or spawning are recorded.

2.4.4 Adult Frog Health, Quarantine, and Disease

Adult frogs should be handled as little as possible, to avoid damage and undue stress. Gloves must be worn when handling frogs. The use of rinsed, powder-free gloves is recommended, and latex gloves must not be used unless proven to be non-toxic in a laboratory investigation (see Appendix E for additional details). When handling is necessary, it should be done gently, carefully, and quickly to minimize stress to the animals. Each holding container containing adult frogs should be checked at least three times per week (e.g., before feeding),

[44] It has been reported that under natural conditions, oocyte growth and yolk deposition begin in the late spring and early summer, followed by maturation of eggs in the fall and increase in size in winter, with ovulation during the spring breeding season (Mizell, 1964).

during which time the health of the organisms must be monitored and recorded. It is recommended that the contents of each holding container be inspected just before each feeding, to determine the apparent condition of the organisms and the holding tanks. Organism welfare should be assessed at this time (e.g., changes in feeding or other behaviour, social interactions such as aggression, and appearance such as skin colour or excessive skin sloughing); recommendations of welfare indicators are available in Section 9.1 of the CCAC guidelines for amphibians (CCAC, 2021). The number of frogs must be counted and recorded, and any dead, diseased, or moribund organisms counted and removed from each holding tank. If necessary, adult frogs can be treated for disease prevention or control in consultation with a veterinarian. Records of any disease and treatment of adult frogs should be kept on file. If adult frogs are treated for disease prevention or control, a minimum two-week period should follow before collecting eggs from the adult(s) for use in toxicity tests.

Adult frogs should be active and green to brown in colour with no visible signs of disease (e.g., abrasions, redness, skin sloughing, not feeding). Procedures and conditions used to maintain each culture should be evaluated routinely, and adjusted as necessary to maintain or restore the health of the culture. If the culture appears unhealthy or atypical during any check, it should be checked more frequently to make sure that "cascade mortality" (i.e., rate of death increasing exponentially over time) is not occurring. Any frogs showing signs of disease should be isolated, quarantined, and treated if recommended by a veterinarian (see Appendix E). The dates and dosages of any treatment provided should be recorded for each frog.

New batches of adult frogs received by the laboratory should be held in quarantine for a period of at least 2 weeks, or longer (6–8 weeks).[45] The tanks must be kept separate from any current laboratory frog tanks (i.e., preferably in a separate room). No transfer of frogs, water, or materials between tanks is permitted. Gloves should be changed following contact with a potentially infected frog, and all equipment should be disinfected.

Guidance on prophylactic treatment, fecal sampling, and quarantining adult frogs, as well as information on common diseases in *Lithobates pipiens*, is provided in Appendix E.

[45] Quarantine is recommended when a new batch of adult frogs is obtained, particularly for laboratories that are already holding amphibians in-house, in order to prevent the transfer of disease between organism batches. In addition, quarantine may help prevent the use of diseased organisms in testing. However, if a laboratory is obtaining frogs to be used for immediate breeding or immediate hibernation followed by breeding, then the laboratory may choose to accept the risk of disease and not quarantine the new batch of adult frogs.

Section 3

Test System

3.1 Facilities and Apparatus

Tests must be performed in an environmental chamber or equivalent facility having acceptable temperature and lighting control (see Section 4.3.3). The test facility should be well ventilated to prevent exposure of personnel to harmful fumes, and it should be isolated from physical disturbances or any contaminants that might affect the test organisms. The area used to prepare substances or materials in preparation for tests should contain a fume hood and be properly ventilated.

The test facility should be isolated from the area where embryos, tadpoles, and frogs are cultured (see Sections 2.3.3 and 2.3.4) to avoid potential contamination. Additionally, the test facility should be removed from places where samples are stored or prepared, to prevent the possibility of contamination of test vessels and their contents from these sources. Based on the experience of Canadian laboratories, tadpoles and adult frogs are more susceptible to disease than commonly used vertebrates, such as *Pimephales promelas* (fathead minnow) and *Oncorhynchus mykiss* (rainbow trout). When laboratory staff are performing routine laboratory operations, gloves should be changed when moving between different experiments. Instruments and surfaces should be cleaned thoroughly between use. The ventilation system should be designed, inspected, and operated to prevent air within the testing facility from contaminating the culturing facilities. Return air from sample handling and storage facilities or those where chemicals are processed or tested should not be circulated to the area of the laboratory where tests are conducted.

Any construction materials that might contact the organisms, water, or test vessels within this facility must be non-toxic (see Section 2.3.2) and should minimize sorption of chemicals. Borosilicate glass, nylon, high-density polyethylene, high-density polycarbonate, fluorocarbon plastics, Teflon™, Nalgene™, porcelain, fibreglass, and type 316 stainless steel should be used whenever possible to minimize chemical sorption and leaching. The use of toxic materials including copper, zinc, brass, galvanized metal, lead, and natural rubber must be avoided.

Compressed air used within the test facility for aerating water must be free of oil and fumes. Oil-free air pumps should be used wherever possible. Any oil or particulate in the air supply should be removed by filters, which are replaced as required to ensure their effectiveness.

The test facility must have the instruments required to monitor the basic variables of water quality (e.g., temperature, conductivity, dissolved oxygen, pH) and must be prepared to undertake prompt and accurate analysis of other variables such as hardness, alkalinity, ammonia, and (in instances where dechlorinated or dechloraminated municipal water is used as culture or control/dilution water) residual chlorine. Safety apparatus including gloves, laboratory clothing, and glasses for eye protection are required when preparing *stock solutions* and/or test substances or materials.

All test vessels, equipment, and supplies that might contact test samples, control/dilution water, stock solutions, or test solutions must be clean and rinsed with control/dilution water, *deionized water*, or *distilled water* before use. All non-disposable materials should be washed after use. The following cleaning procedure is recommended (EC, 2013a):

1. soak in tap water (with or without detergent added) for 15 minutes, then scrub with detergent or clean in an automatic dishwasher;

2. rinse twice with tap water;

3. rinse carefully with fresh, dilute (10%, v:v[46]) nitric (HNO_3) or hydrochloric acid (HCl) (metal-free grade) to remove scale, metals, and bases;

[46] To prepare a 10% solution of acid, carefully add 10 mL of concentrated acid to 90 mL of deionized water.

4. rinse twice with deionized water (or equivalent);

5. rinse once with full-strength, pesticide-grade acetone to remove organic compounds and with reagent-grade (e.g., HPLC-grade, ≥ 98.5% purity) hexane for oily residues (use a fume hood);[47]

6. allow organic solvent to volatilize from dishware in fume hood and rewash with detergent (scrub if necessary); and

7. rinse three times with deionized water (or equivalent).

This cleaning procedure assumes the identity of contaminants is unknown, and so provides proper cleaning for metal and organic contaminants. If contaminants are known (e.g., metal salt), cleaning procedure may be scaled back (e.g., removal of steps 5 to 7).

3.2 Lighting

All test vessels should receive full-spectrum (e.g., fluorescent or equivalent) illumination from directly overhead, at an intensity sufficient to provide 100 to 500 lux adjacent to the surface of the water. Illumination should be as uniform as possible for all test vessels. Photoperiod must be regulated at 16-h light and 8-h dark. A 15- to 30-minute transition period between light and dark is recommended (see Section 2.3.3 and footnote 23).

3.3 Preliminary and Definitive Tests

3.3.1 Preliminary Tests

It is the laboratory's responsibility to demonstrate its ability to obtain consistent, precise results using the test methods described herein before *definitive* toxicity tests are performed for the first time. To meet this responsibility when a laboratory is inexperienced with a biological test method, personnel should conduct a minimum of two or more multi-concentration tests or single-concentration (i.e., positive control) tests using a reference toxicant(s) (see Section 4.8) and the methods defined in Section 4 for a definitive test. In addition, if the reference toxicant test does not use the same duration or exposure as the planned definitive test (e.g., 96-h acute lethality test for the 14-day definitive test, or a 14-d positive control using thyroxine for the 42-day definitive test), then personnel should conduct a minimum of two or more 14-day or 42-day tests using control/dilution water only and the methods defined in Section 4 for a definitive test. These preliminary tests are recommended to confirm that acceptable performance of the test species (*L. pipiens*) can be achieved in control/dilution water used by that laboratory and the culturing/holding conditions and procedures specified in this report (see Section 2.3). In addition, these tests will enable the laboratory to establish the concentration(s) of a reference toxicant(s) to be used as a positive control and/or in a multi-concentration reference toxicant test (see Sections 3.5 and 4.8).

The conditions and procedures used to perform these preliminary tests should be identical and according to Section 4. Each preliminary test should ideally be performed using a different batch of test organisms.

Control performance data from these preliminary tests must show that the criteria for test validity (see Section 4.7) can be met using the control/dilution water intended for use in each of the two definitive toxicity test options, assuming the laboratory intends to use both test options. Data from these preliminary tests should be examined with the intent of choosing a single concentration or concentration series to be used as a positive control or multi-concentration reference toxicant, respectively.

3.3.2 Definitive Tests

Test vessels to be used in definitive tests must be inert to test and reference substances or *contaminant* mixtures (i.e., the test or reference substances, or mixtures thereof, should not adhere to or react in any way with the test vessel). The volume of the vessel must be sufficiently large to accommodate a minimum of 7 L of control/dilution water or test solution.[48] Glass aquaria are recommended. Each

[47] Rinsing Plexiglas™ or any plastic equipment or vessels with acetone or hexane is **not** recommended, since plastic can become pitted and etched by these solvents and can turn from transparent to opaque.

[48] During the inter-laboratory study, participating

test vessel must be cleaned thoroughly before and after use, and should be rinsed well with deionized, distilled, or control/dilution water before use (see Section 3.1). Vessels can be loosely covered to prevent debris (e.g., dust from lab air) from entering the test vessel.

The renewal of each test solution (minimum of three times weekly) is achieved either by siphoning out ~80% of the old solution and replacing it immediately with fresh (new) test solution (i.e., static-renewal test), or by the continuous addition of fresh solution to the test chamber (i.e., flow-through test). Section 4.3.2 outlines details of flow-through test systems. The test vessel should be adapted to accommodate either static-renewal or flow-through conditions, depending on the requirements and objectives of the test.[49]

3.4 Control/Dilution Water

Depending on the test design and intent (Sections 5 and 6), the control/dilution water may be: "uncontaminated" groundwater or surface water from a stream, river, or lake; laboratory water adjusted to a desired pH and hardness (e.g., simulating that of the *receiving water*); a sample of receiving water collected *upstream* of the source of contamination, or adjacent to the source but removed from it; or dechlorinated or dechloraminated municipal water (see Section 2.3.5). The water supply used as control/dilution water is frequently the same as that used for holding/culturing the tadpoles/frogs (see Section 2.3.5), although it may come from another source. For instance, the use of receiving or "upstream" water, or laboratory water adjusted to the pH and hardness of water at a collection site, might prove a good choice (see Section 6.3). The quality of control/dilution water is extremely important; this water must have been demonstrated to allow acceptable survival, growth, and development of test organisms in preliminary tests (see Section 3.3.1) before it is used in definitive toxicity tests. When surface water (including receiving or "upstream" water) is used as control/dilution water, a second set of controls must be prepared using a supply (source) of laboratory water shown previously by the testing laboratory to routinely enable valid test results.[50] If a surface water sample contains debris or indigenous organisms that might be confused with or attack the test organisms, the sample must be filtered before use (see Section 2.3.5). Use of surface water can pose a substantial risk to amphibians, due to the presence of disease-causing organisms (i.e., pathogens).[51] The risk is more pronounced than in other standardized test methods because the amphibian test is longer duration, providing more time for potential disease to manifest itself. The size

---

laboratories used aquaria with a total capacity of either 8 L or 17 L. The study design used 6 L of exposure solution; however, in some labs during the final weeks of the 42-day exposure, ammonia levels were slightly elevated. A loading density of 10 tadpoles per 7 L will help buffer increasing ammonia concentrations, aligns well with literature values, and is below loading densities that are correlated with increased control mortality (Melvin and Houlahan, 2012; ECCC, 2023).

[49] With many types of test substances, static tests with 12- or 24-h renewal of test solutions, when done properly, can be as sensitive and accurate as flow-through tests (Sprague, 1973). Static-renewal tests with more frequent renewals might also be desirable or necessary when degradation products of the test substance are of concern. High chemical or biochemical oxygen demand, volatility, or instability of certain substances might necessitate the use of a flow-through test.

[50] If the intent of the test is to measure the extent to which a particular receiving or "upstream" water might modify the toxicity of the test substance or material due to its physicochemical characteristics (e.g., hardness, pH, turbidity, etc.) and/or the presence of other contaminants, the investigator might choose to use the receiving or "upstream" water as control/dilution water. A comparison of controls for this water with those for the controls held in laboratory water will identify toxic effects that might be contributed by the receiving or "upstream" water. A clearer understanding of the differing influence of each type of control/dilution water on the toxicity of the test substance or material can be achieved by undertaking a side-by-side comparison of toxic effects using each type of water to prepare test treatments.

[51] For example, the prevalence of *Batrachochytrium dendrobatidis* (a waterborne fungus causing chytrids disease; see Appendix E.2) infection in wild populations of *L. pipiens* has been observed to be as high as 18.6% in British Columbia (Voordouw *et al.*, 2010), 25% in Prince Edward Island (Forzán *et al.*, 2010), and 25.7% in Maine (Longcore *et al.*, 2007).

range of the bacteria and fungi known to cause disease in amphibians (see Appendix E) is approximately 0.3 µm (*Aeromonas hydrophila*) to 800 µm, and ranaviruses are approximately 150 nm. It may therefore be desirable to filter surface water to be used as control/dilution water through a conventional sand filter or a commercial in-line filter (e.g., 0.45 to 5 µm). UV sterilization is also recommended to reduce the risk of pathogens (see Section 2.3.5). Note that the risk of *Ranavirus* infection may not be mitigated by filtration alone due to the small size of the viruses.

Control/dilution water must be adjusted to the test temperature (23 ± 2 °C) before use. The dissolved oxygen content of the water should be 90 to 100% of the air saturation value at this temperature. As necessary, the required volume of water should be aerated vigorously (oil-free compressed air passed through airstones) immediately before use, and its dissolved oxygen content checked to confirm that 90 to 100% saturation has been achieved. The pH of the water should be measured and stable before use.

3.5 Positive Control

The use of one or more samples of positive control is recommended for inclusion in each definitive toxicity test described herein to assist in interpreting the test results. The intent is to select one or two concentrations of a reference toxicant that will elicit a response in the test organism (for the test option chosen) that is predictable based on earlier toxicity tests with the same material. The positive control sample(s) is control/dilution water spiked with a reference toxicant for which historic data are available, and/or have been established in the laboratory, on its toxicity to *L. pipiens* using specified test conditions and procedures. These positive control samples provide an alternative option to the multi-concentration reference toxicant tests traditionally required in Environment and Climate Change Canada biological test methods. For both the 14- and 42-day test options described herein, concentrations of a reference toxicant must be used either in a separate multi-concentration test, or as *replicates* of a positive control (at one or two specific concentrations) included with each definitive test, when appraising the sensitivity of the test organisms and the *precision* and reliability of results obtained by the laboratory for that material (see Section 4.8).

Section 4

Universal Test Procedures

General procedures and conditions described in this section for toxicity tests with *Lithobates pipiens* larvae apply to all tests of samples of *chemical*, chemical *product*, or contaminated water, and also apply to their associated reference toxicity tests. More specific procedures for conducting tests with chemicals or chemical products are given in Section 5. Guidance and specific procedures for conducting tests with contaminated water samples (e.g., impacted wetlands or *receiving waters*, *effluents*, *elutriates*, or *leachates*) are described in Section 6. All aspects of the test system described in Section 3 must be incorporated into these universal test procedures. Those conditions and procedures described in Section 2 for culturing and/or acclimating *L. pipiens* in preparation for toxicity tests also apply. The summary checklist in Table 4 describes required and recommended conditions and procedures to be universally applied to each test, as well as those for testing specific types of test materials or substances.

This biological test method measures the effects of exposure to various materials and substances on the survival, growth, and development of early aquatic life stages of frogs. The test species is *L. pipiens* (see Section 1.2). The test is conducted as static-renewal or flow-through and the duration depends on the study objectives and the test option chosen (see Sections 1.1, 2.1, and 4.3.1). Tadpoles (Gosner stages 25 and higher) are fed *ad libitum* a 4:1 mixture of kale and commercial dry tadpole food three days per week, and dried algal pellets/wafers one day per week. Feeding rates are adjusted as needed based on observations of available food.

This definitive test method was applied and validated by several participating laboratories in three rounds of concurrent tests using *L. pipiens* (Nautilus Environmental, 2020a, 2020b).[52]

---

[52] Four laboratories participated in the first round of the inter-laboratory validation tests. Two 14-day static-renewal tests were conducted by each laboratory. The first was exposure of Gosner stage 25 tadpoles to 6 concentrations of NaCl (0.80 to 6.0 g/L NaCl), and the second was an exposure of Gosner stage 28/29 tadpoles to 5 concentrations of thyroxine (0.074 to 6.0 µg/L T4) and a solvent (NaOH). For the NaCl tests, all laboratories reported good control survival with averages at individual laboratories ranging from 82.5 to 100% and tadpoles advancing ~ 2 to 4 Gosner stages (mean GS at test end ranged from 27.3 to 29). All laboratories observed a decrease in survival with increasing NaCl concentration. The mean LC50 was 4.2 g/L, with values ranging from 3.5 to 4.9 g/L. The inter-laboratory variability expressed as the coefficient of variation (CV) was 16%. Adverse effects on growth were also observed. The average IC50 for biomass was 3.5 g/L NaCl and the CV was 26%. The calculated IC50 for length was >4 g/L NaCl with two labs showing a 50% reduction in length at 4 g/L NaCl (i.e., IC50 = 4 g/L NaCl). A decrease of 1.4 to 1.9 Gosner stages was observed at the highest NaCl concentration in three of the four laboratories (Nautilus Environmental, 2020a). For the T4 tests, one laboratory had a disease outbreak affecting some of their replicates, which were excluded from the summarized data. Survival in the solvent controls for the three laboratories ranged from 83 to 100%. At the end of the 14-day exposure there were no effects on survival up to 2 µg/L T4 and up to 6 µg/L T4 in two out of three laboratories. The mean IC50 for decrease in wet weight was 1.5 µg/L T4, with values ranging from 1.4 to >2 µg/L T4. As expected from its mode of action, higher concentrations of T4 also resulted in increased development of tadpoles. After 2 weeks of exposure, all surviving tadpoles in the highest treatement (6 µg/L) had reached GS 41 or higher relative to tadpoles in the solvent control in which the mean GS at the end of the test ranged from 30.4 to 32.0. There were significant increases in tadpole development at one laboratory in the lowest test concentration and in all three laboratories at the 0.67 µg/L T4 test concentration (Nautilus Environmental, 2020a).

Two laboratories participated in the second round of the inter-laboratory validation tests. For this round, tadpoles at GS 25 were exposed to 6 concentrations of triclosan (ranging from 18.8 to 600 µg/L) and a solvent (NaOH) and the test was scheduled to continue until 60 to 80% of the control tadpoles reached GS 42. One of the laboratories had to terminate the test early (after 82 days) due to the appearance of disease symptoms and related mortalities. The second laboratory terminated the test after 105 days. Up to the eighth week of testing, survival was high in both the laboratory water and solvent controls in the two laboratories, ranging from 80 to 97%. The

4.1 Preparing Test Solutions

Each test vessel (see Section 3.3.2) placed within the test facility must be clearly coded or labelled to enable identification of the sample and (if diluted) its concentration. The test vessels should be positioned such that observations and measurements can be made easily, and must be positioned randomly within the test facility.

For any test that is intended to estimate the *LC50* or other *LCp* for survival (see Section 4.6.2.1), estimate the *ICps* for growth endpoints (i.e., total length, wet weight, and *biomass*; see Section 4.6.2.2), and detect significant effects on development (i.e., change in Gosner stage in treatment(s) compared to control(s) at test end; see Section 4.6.2.3), a minimum of seven test concentrations plus a control solution (100% dilution water) must be prepared; however, more (i.e., ≥ 8 plus control[s]) are recommended to improve the likelihood of bracketing each endpoint sought.[53]

In certain cases, it may be permissible to use six test concentrations. For example, the investigator may have run a *range-finding test* before beginning the *definitive test*, and this data can be used to select test concentrations. Any reduction in number of test concentrations must be recorded and reported, with the appropriate rationale.

Concentrations should be chosen to span a wide range, including a low concentration that evokes no adverse effects (e.g., similar to that for the *negative control* treatment), and a high concentration that results in "complete" or severe effects. If the anticipated endpoint is bracketed with a closely

---

highest concentration of triclosan tested (600 µg/L) was lethal, resulting in complete mortality within 4 days of test initiation. The remaining tadpoles did not exhibit concentration-related effects for any of the endpoints, likely due to the rapid degradation of triclosan in the exposure concentrations throughout the test. Tadpole development varied greatly, with some reaching GS 42 within 6 weeks, whereas up to 32% of other tadpoles remained at the early stages of tadpole development (GS 25 to GS 29) (Nautilus Environmental, 2020a). This varied rate of development and therefore lengthy and unpredictable test duration contributed to ECCC's move towards changing the test design for determining developmental effects from one with a development-based termination to one with a definitive exposure duration.

Three laboratories participated in the third round of the inter-laboratory validation tests. For this final round, tadpoles at GS 28/29 were exposed to 6 concentrations of perchlorate (ranging from 9.22 to 900 mg/L) for 6 weeks (42 days). All labs exhibited good control survival throughout the exposure with survival ranging from 87 to 100%. Control organisms advanced 6 to 11 Gosner stages with the average at test end ranging from 34.3 to 39.6. Dose-dependent adverse effects on growth and development were observed. Growth endpoints were most sensitive (i.e., lower IC50s) at the 14- and 28-day interim assessments with mean IC50s for biomass of 185.2 and 208.6 mg/L, respectively, and mean IC50s for length of 61.8 and 178.0 mg/L, respectively. The average IC50s for biomass and length at the end of the test were 270.4 and 244.2 mg/L respectively. Inter-laboratory variability for these endpoints, expressed as CV, ranged from 37.6% to 84.1%. Statistically significant decreases in tadpole development (based on Gosner stage) were observed in all but the lowest test concentration (9.21 mg/L perchlorate) at the end of the 42-day exposure (Nautilus Environmental, 2020b).

Two pilot studies involving the exposure of GS 28/29 tadpoles to a single concentration of T4 (0.67 µg/L) in rounds two and three of the inter-laboratory investigation showed similar results. Participating laboratories in both rounds demonstrated very consistent acceleration of tadpole development (i.e., 5.0 and 4.9 stages, relative to controls) after two weeks of exposure. These results are consistent with the accelerated development of tadpoles by 4.6 Gosner stages relative to controls observed at the same exposure concentration (0.67 µg/L T4) in round one of the inter-laboratory tests (Nautilus Environmental, 2020a, 2020b).

[53] The use of eight or more test concentrations plus the control solution(s) is recommended to improve the likelihood of attaining each endpoint sought and to enable calculations of ICps for growth using regression analyses as well as detecting significant effects on tadpole development. The large number of test treatments is needed to show the shape of the concentration-response relationship and to choose the appropriate linear or non-linear regression model (see Section 4.6.2.2).

Table 4 Checklist of required and recommended conditions and procedures for conducting toxicity tests using aquatic life stages of frogs

Test type	– static-renewal or flow-through* – options for test durations include: o 14-day test: start with tadpoles (Gosner stage 25), and end test after 14 days of exposure; and o 42-day test: start with tadpoles (Gosner stage 28/29), and end test after 42 days of exposure
Control/dilution water	– clean ground or surface water, or dechlorinated or dechloraminated municipal water (hardness 10 to 230 mg/L as $CaCO_3$); "upstream" or receiving water to assess toxic impact at specific location**; DO content 90% to 100% saturation at time of use in a test; pH 6.5 to 8.5
Organisms	– *Lithobates pipiens* (formerly *Rana pipiens*; northern leopard frog) tadpoles (Gosner stage 25 for 14-day test, and Gosner stage 28/29 for 42-day test); ≥10 organisms per test vessel
Organism handling	– small fish net or small cup used to transfer tadpoles; care taken to avoid touching tadpoles; at the end of the test, animals are removed gently using a net and euthanized by a humane procedure (e.g., transferred to a solution of buffered, veterinary-grade tricaine methanesulfonate (MS-222))
Test design	– for multi-concentration test: minimum of 3 replicates required (recommend 4) for 14-day test; minimum of 4 replicates per treatment and 8 replicates for the control(s) required for 42-day test – for single-concentration test: number of replicates to be determined based on project goals
Test vessel and solution	– glass aquaria or other appropriate vessel, containing ≥7 L of test or control solution; vessels may be covered to prevent contamination or debris
Number of test concentrations	– ≥ 7, plus control(s); recommend more (i.e., 8), plus control(s)
Test solution renewal	– minimum 3 times weekly on non-consecutive days (e.g., Monday, Wednesday, and Friday); if static-renewal, remove 80% of overlying water using siphon (remove waste/debris from tank as necessary); replace test solution with as little agitation as possible; if flow-through, recommend flow rate that completely renews test solution a minimum of 3 times weekly
Temperature	– daily mean 23 ± 2 °C, recommend 23 ± 1 °C; instantaneous temperature 23 ± 3 °C
Oxygen/aeration	– if DO < 60% or > 100% air saturation, pre-aerate test solutions for 30 minutes (e.g., 6.5 ± 1 mL/min·L); if necessary, continue pre-aeration until 60% to 100% saturation is achieved or for no more than an additional 90 minutes; continuous gentle aeration (e.g., 6.5 ± 1 mL/min·L) is provided throughout the test (required for static-renewal, optional for flow-through)
pH	– no adjustment if pH of test solutions is in the range of 6.0 to 8.5***; a second (pH-adjusted) test might be required or appropriate for pH beyond that range
Lighting	– overhead full-spectrum (fluorescent, incandescent, or LED); intensity ≥100 lux, recommend 100 to 500 lux at surface of test solution; normally 16-hour light:8-hour dark; recommend gradual transition

Feeding	– *ad libitum* (example feeding rates provided in Table 5 in Section 4.3.6); feed tadpoles three times/week on non-consecutive days with a 4:1 mixture (slurry) of kale:dried tadpole food; feed tadpoles dried algae pellet(s) one day per week as a fourth feeding; feeding rates are adjusted as needed based on observations of available food in dilution-water control
Duration	– duration is dependent on which test option is chosen; 14 or 42 days
Observations	– initial observations: individual wet weight, total length, and (optionally) snout-vent length for 20 representative organisms at the appropriate stage of development – daily observations: record any unusual appearance (e.g., abnormalities) or behaviour, as well as approximate developmental stage, and mortality; dead organisms or organisms reaching GS 42 prior to test end must be removed – number and percent of surviving tadpoles/metamorphs, total length, individual wet weight, biomass based on wet weight, stage of development, and occurrence of deformities or asynchronous development at test end; interim non-destructive measurements every 14 days recommended for the 42-day test
Measurements of water quality	– temperature daily or continuously; DO and pH at test start and end and at least three times weekly in one replicate of representative treatments (at least low, medium, and high test concentrations and controls), before and after each solution renewal; ammonia at test start and end in one replicate of representative treatments (low, medium, and high test concentrations and controls), and once weekly before and after solution renewal; recommend conductivity at test start and at each test solution renewal (new solution only) in representative treatments; recommend hardness and/or alkalinity of control/dilution water and highest test concentration at start of test and once per week.
Endpoints	– mean (± SD) percent survival and growth (total length, wet weight, and biomass based on wet weight, all corrected for initial measurements); median developmental stage (Gosner stage) – if multi-concentration test, LCp is calculated for survival and ICps are calculated for growth endpoints; for 42-day test, significant difference in developmental stage in treatments versus control is calculated – optional additional endpoints include snout-vent length, tissue/organ histology, and gene expression analysis
Test with reference toxicant	– reference toxicity test is performed for each batch of test organisms used in a definitive toxicity test; multiple acceptable options for reference toxicant testing (see Section 4.8), including single-concentration and multi-concentration test designs, depending on test option used for definitive testing; sodium chloride and/or thyroxine recommended
Test validity	– for 14-day test: invalid if mean mortality of control organisms is >20%; invalid if median Gosner stage of control organisms is < 27 – for 42-day test: invalid if mean mortality of control organisms is >20%; invalid if median Gosner stage of control organisms is < 33

Growth benchmarks for control organisms	– for 14-day test: ≥14.4 mm mean corrected total length; ≥4.2 mm mean corrected snout-vent length; ≥0.18 g mean corrected wet weight; ≥0.16 g mean corrected biomass (based on wet weight) – for 42-day test: ≥13.2 mm mean corrected total length; ≥4.8 mm mean corrected snout-vent length; ≥0.44 g mean corrected wet weight; ≥0.42 g mean corrected biomass (based on wet weight)
Chemicals	
Characterization of chemical(s) or chemical product(s)	– information on concentration of active ingredients, water solubility, vapour pressure, stability, dissociation constants, absorption coefficients, toxicity to humans and aquatic organisms, and biodegradability of chemical(s) or chemical product(s) spiked into water should be known beforehand; recommend additional information include structural formulae, degree of purity, nature and percentage of significant impurities, presence and amounts of additives, and n-octanol:water partition coefficient
Solvents	– uncontaminated or deionized laboratory water is the preferred solvent; if another solvent is used, the maximum concentration is 0.1 mL/L or 100 mg/L (see Section 5.2), and the test must include a solvent control containing the same concentration of solubilizing agent as in the most concentrated solution of the test chemical, in addition to a dilution-water control
Concentration	– recommended measurements are at beginning and end test in low, medium, and high concentrations, and control(s); beginning and end of renewal period, if necessary; if concentration declines ≥20%, and depending on the study objectives, re-test with more frequent renewals or flow-through methods
Control/dilution water	– as specified and/or depending on intent; receiving water if concerned with local toxic impact; otherwise, uncontaminated laboratory water demonstrated to meet test validity criteria
Contaminated Waters	
Sample requirement	– either multiple subsamples from a single sampling, or multiple separate samples collected (or prepared, if elutriate) and handled as indicated in Section 6.1; ~150 L for single-concentration 14-day test and routine sample analysis, and more needed for multi-concentration test or 42-day test designs (see Section 6)
Transport and storage	– if sample > 7 °C, cool to 1 to 7 °C (ice or frozen gel packs); transport in dark at 1 °C to 7 °C (preferably 4 ± 2 °C); store in dark at 4 ± 2 °C; samples must not freeze or partially freeze during transport or storage; holding times minimized as much as possible
Control/dilution water	– as specified and/or depends on intent; laboratory water demonstrated to meet test validity criteria or "upstream" receiving water** for monitoring and remediation

The information in this table is for summary purposes only. Definitive requirements and recommendations of this test method are contained in the main body of this document.

* Special situations (e.g., volatile or unstable chemicals in solution) might require the use of flow-through tests.
** For this option, there must be an additional control using a separate water supply (natural, dechlorinated or dechloraminated municipal, or reconstituted) that has been shown by the testing laboratory to routinely achieve valid test results in previous chronic tests with *L. pipiens*.
*** If pH is outside this range, results might reflect toxicity due to biologically adverse pH.

spaced series of concentrations, all may turn out to be either too low or too high.

To keep the wide range of concentrations and also obtain the important mid-range effects, it might be necessary to use additional treatments in order to split the selected range more finely.

In cases of appreciable uncertainty about sample toxicity, it is beneficial to run a range-finding or screening test for the sole purpose of choosing concentrations for the definitive test. Conditions and procedures for running the test can be relaxed. A range-finding test normally covers a broader concentration range, incorporates fewer replicates (i.e., 1 or 2) and organisms, and may be shorter in duration.[54]

A geometric dilution series should be used in which each successive concentration is about 50% of the previous one (e.g., 100, 50, 25, 12.5, 6.3, 3.1, 1.6, etc.). Test concentrations may also be selected from other appropriate dilution series (e.g., 100, 75, 56, 42, 32, 24, 18, 13, 10, 7.5; see Column 7 in Appendix H). If a high rate of mortality is observed within the first few days of the test, and extra test organisms are available, additional dilutions (i.e., at lower concentrations) may be added and these organisms exposed for the remainder of the test duration. A dilution factor as low as 30% (e.g., concentrations 100, 30, 9, etc.) is not recommended for routine use because of poor precision of the estimate of toxicity, however, it might be used if there is considerable uncertainty about the range of concentrations likely to be toxic. See EC (2005) for additional guidance on selecting test concentrations that apply here. Volume requirements for tests will vary according to the test option (14- or 42-day) used (see Sections 4.3.1, 5.2, 5.3, and 6.1).

For both the multi-concentration and single-concentration test designs, each treatment including the control(s) must include multiple replicate test vessels. The number of replicates varies depending on which test option (14 or 42-day), and type (multi-concentration or single concentration) is being used (see Sections 4.2 and 4.6).

For each definitive test, control solution(s) must be prepared at the same time as the experimental treatments. Any dilution water used to prepare test concentrations must also be used for preparing one set of controls. Each test solution must be mixed well using a glass rod, Teflon™ stir bar, or other device made of non-toxic material. Temperatures must be adjusted as required to 23 ± 2 °C. Dissolved oxygen and pH of representative test concentrations must be measured and recorded (see Section 4.4). It might be necessary to adjust the pH of the sample of test material or the test solution (see Section 4.3.5), or to provide preliminary aeration of the test solutions (see Section 4.3.4). Conductivity of test solutions should be measured before dispensing it to the test vessels, as this can serve as a check on correct preparation of test concentrations (see Section 4.4).

When receiving water from upstream of the discharge is used as control/dilution water (see Sections 5.3 and 6.3), a second control solution must be prepared using a supply (source) of laboratory water shown previously by the testing laboratory to routinely enable valid test results in a definitive test (see Section 3.4).

Following the addition of a measured volume of test solution to each test vessel (i.e., ≥7 L for 10 test organisms; see Section 3.3.2), gentle aeration must be provided throughout the test. Pre-aeration (30 minutes

[54] The design of a range-finding test depends on the study objectives, the test option chosen, and the availability of test organisms, while keeping in mind the balance between getting the information needed and the practical and ethical aspects of using test organisms to conduct range-finding test. Egg masses maintained in the laboratory at cooler temperatures (e.g., 10 to 15 °C) will delay the onset of embryo development (see Section 2.3.4), thereby providing the opportunity to use a portion of the egg mass to provide test organisms for a range-finding test in advance of the of the definitive test. Options for range-finding tests may include running a 96-h lethality test, a 14-day exposure starting with young tadpoles (using a portion of the test organisms to be used in the definitive test), or a full test on an additional egg mass carried out prior to the definitive test. To reduce the number of test organisms used, fewer replicates and test concentrations can be used. Alternatively, it might be feasible to use a 96-h fish acute lethality test as a "surrogate" range-finding test for the 14-day test described in this method, since it has been shown that in acute exposures, fish may have similar sensitivities to larval amphibians (see Section 1.3.1).

at a rate of 6.5 ± 1 mL/min·L) can be provided, depending on DO levels (see Section 4.3.4). The temperature, dissolved oxygen, and pH of one replicate of each treatment must be measured and recorded, after any pre-aeration but before organisms are added to the test vessels.

4.2 Beginning the Test

The day that animals are initially exposed to samples of test materials or substances is designated Day 0. A minimum of ten organisms are required per replicate, with an equal number in each vessel. For the 14-day multi-concentration test, a minimum of three replicates per treatment (concentration), including the control treatment(s), must be included in each test, and four replicates per treatment are recommended. For the 42-day multi-concentration test, a minimum of four replicates per treatment (concentration) and a minimum of eight replicates for the control treatment(s) must be included in each test.[55]

For single-concentration tests (e.g., sample tested at 100% concentration only, or a particular concentration of a test chemical), a minimum of 8 replicate test vessels per treatment must be used for both the 14- and 42-day test options.

A test with seven concentrations plus a control and with three replicates per treatment (concentration), requires at least 240 test organisms. A number in excess of those required for the test can be collected from the holding/acclimation vessels in a large clean container (e.g., pail) containing culture water using a small fish net. Tadpoles should be assessed and then counted into a series of small beakers or holding vessels with >1 L of dilution water. Either GS 25 or GS 28/29 tadpoles, depending on the test option chosen, must be used to initiate the test. The assessment of the developmental stage (i.e., Gosner stage) of each test organism prior to introduction into the test vessel must be carried out using a dissecting microscope. This can be achieved by gently placing a tadpole into a petri dish, ensuring there is just enough water to cover the tadpole. Orienting the tadpole for a side view, the Gosner stage should be assessed based on specific physical benchmarks (see Figures 2.1 and 2.2, and footnote 68 in Section 4.5). Any tadpoles possibly damaged or injured during assessment of Gosner stage or during transfer must be discarded. Tadpoles that appear abnormal in any way (e.g., any visible defects or damaged bodies, or are inactive) must not be selected for the test. Once enough organisms at the appropriate developmental stage have been collected and assessed, they must be randomly distributed into each test vessel. An additional 20 organisms at the appropriate stage of development that are (visually) representative of those used in the exposure are then selected, euthanized, and measured for wet weight, total length, and optionally, snout-vent length. Using a ruler, calipers, or digital imagery, the total length (tip of snout to tip of tail) of each tadpole must be measured. Using the same measuring tools, the snout-vent length (tip of snout to vent) can also be measured. Blotting excess water from the tadpole, the wet weight of each tadpole must then be measured. The developmental stage, length and weight of the tadpoles are recorded and are used as the initial measurements for the test. Dip nets or small plastic spoons can be used to transfer the test organisms into temporary holding vessels and then into the test vessels (see Section 3.3.2).

Tadpoles used in a given test should ideally all be from the same egg mass (i.e., batch). If, however, insufficient organisms are available from a single egg mass, several egg masses are available for use, there are concerns about organism health (see footnote 41 in Section 2.3.8), or the study objectives include representing a larger range of genetic diversity, then organisms can be taken from more than one egg mass. In this case, test organisms should be approximately equally represented from

---

[55] Three or more replicates are beneficial for point estimates of ICp as an endpoint. For the developmental stage endpoint, power analysis (see Section 4.6.2.4) indicated that an uneven test design with eight replicates per control and four replicates per treatment is necessary to provide sufficient (80%) power to detect significant treatment effects on Gosner stage (Green, 2021). This uneven test design is also used in the USEPA (2015) LAGDA test method, which recommends similar statistics for the development endpoint (see Sections 1.4 and 4.6.2.3).

each egg mass.[56] Alternatively, if several egg masses are available for use, investigators can choose the "best" egg mass based on monitoring/health criteria collected during hatching (see Section 2.3.8).

After transferring the test organisms into each test vessel, the number of tadpoles should be recounted to ensure that the required number is present and to make any necessary adjustments. Final observations of all tadpoles in each test vessel should also be made at this time and any tadpoles appearing atypical in size, shape, colour, or behaviour must be discarded and replaced. The date and time when the test is started must be recorded, either directly on the labels or on separate data sheets dedicated to the test.

4.3 Test Conditions and Procedures

4.3.1 Test Options

Depending on the study objectives (see Section 1.1), one or more of the following two test options may be used: i) a 14-day test, initiated with free-swimming tadpoles (Gosner stage 25), designed primarily to assess the survival and growth of tadpoles prior to any metamorphic changes; and ii) a 42-day test, initiated with tadpoles in later stages of development (Gosner stage 28/29), designed to capture changes in development and growth leading to metamorphosis.

Test designs for multi-concentration tests are provided in detail, based on extensive laboratory experience, statistical power, and best scientific judgement. Analysis of single-concentration tests would require hypothesis testing for each biological endpoint (development, growth, mortality). It is not possible to optimize number of replicates for all biological endpoints.[57] Should single-concentration tests be necessary, investigators are encouraged to follow all the steps outlined for a multi-concentration test, but to revise the number of replicates to suit the goals of the investigation.

Both the 14-day and 42-day test options may be used to evaluate samples of chemical or contaminated water, depending on the objective of the test. Survival (% mortality), growth (total length, wet weight, and biomass), and developmental stage (Gosner stage) must be measured at the end of the test for both test options. Observations of any deformed test organisms must also be made and reported (see footnote 75 in Section 4.6.1). Snout-vent length may also be measured at the end of the test.

4.3.2 Test Type and Solution Renewal

Tests may be run either as static-renewal or flow-through. For some substances having high chemical or biochemical oxygen demand, volatility, or instability, use of a flow-through test with rapid replacement of test solutions might be necessary.

In static-renewal tests, solutions are changed a minimum of three times per week on non-consecutive days (e.g., Monday, Wednesday, and Friday) or more frequently if required (i.e., if the test

[56] An attempt must be made to achieve "homogeneity of the experimental units" to avoid any differences among vessels that are related to sensitivity differences between batches of tadpoles. This can be achieved in two different ways (prof. J. Hubert, Dept. of Mathematics and Statistics, University of Guelph, personal communication, 1991; EC, 1998, 2011). In the first method, tadpoles from different egg masses that have been held separately may be combined (pooled) before introducing the organisms to the test vessels. In the second method, tadpoles from a given egg mass may be divided evenly among all replicates of all concentrations, and then tadpoles from another egg mass are similarly allotted evenly to all incubation units or test vessels, to make up the full number per replicate. The second method requires more care and effort in culturing and handling. However, it should reduce the "noise" of the variation between replicates at the same concentration and avoid the chance that exists in the first method of getting high proportions of insensitive or highly sensitive test organisms in a particular vessel, if such spawning-related variation exists.

[57] For example, in order to have adequate power for detecting differences in wet weight, approximately 6 replicates would be required if using a *t*-test. However, using 6 replicates when evaluating mortality could lead to an overly sensitive test, such that 15–20% mortality would be declared statistically significant (L. Van der Vliet, Environment and Climate Change Canada, Ottawa, ON, personal communication, 2022). This level of mortality is permissible in the controls, and so is considered to be within experimental error. However, with a statistically significant statistical test result, investigators may erroneously conclude the result is biologically significant.

substance degrades, if DO levels decrease, and/or ammonia levels increase), depending on the experimental design. To conduct a solution renewal, almost all (i.e., 80%) of the test solution is removed from each test vessel and replaced with freshly prepared solution. Siphoning (i.e., siphon with the opening covered with a small piece of Nitex™ or netting) is the usual procedure. Any waste, uneaten food, or other detritus should be siphoned from the bottom of each vessel and care should be taken to avoid touching the tadpoles. Tadpoles must remain submerged in test solution at all times during the renewal process. New test solution must be added to the original total volume of test solution in each vessel. The entire process must be done as cautiously as possible and with as little agitation as possible to avoid injuring the tadpoles. The solution that is siphoned out or otherwise removed must be examined to ensure that no tadpoles have been accidentally removed. Such test organisms are likely to be injured and must be discarded and noted on the benchsheet as accidentally removed at that time; and the results of the test must be analyzed as if the discarded test organisms had not been present.

Flow-through tests require a system that continually delivers a series of pre-mixed concentrations of the test substance or material to the test vessels at a controlled rate. Various devices might create successive dilutions of a stock solution or test substance by means of metering pumps or proportional diluters. A flow rate that produces a complete turnover of the solution volume at least 3 times weekly is recommended as a minimum, although depending on the test substance and study objectives, a more frequent replacement of solution volume might be warranted. The flow rates of test solutions, or stock solutions and control/dilution water, should be checked daily throughout the test, and should not vary by more than 10%. For flow-through exposures, fresh (new) test solutions should be such that the desired volume of test solution is replaced daily with minimal agitation and/or disturbance of the tadpoles. For further guidance on flow-through system designs, rates, and procedures, the reader is referred to APHA *et al*., 2011; ASTM, 2023a, 2023b.

4.3.3 Test Temperature and Lighting

The test must be conducted at a daily mean temperature of 23 ± 2 °C and should be 23 ± 1 °C.[58] Additionally, the instantaneous temperature must always be 23 ± 3 °C. Sample/solution temperature must be adjusted as required to attain an acceptable value for each solution. Samples or test solutions must not be heated by immersion heaters, since this could alter chemical constituents and toxicity. Temperature must be measured and recorded daily in one replicate of each treatment including the controls. On the days of test solution renewal, measurements must be made in both the fresh test solution and the used solution just before or just after it has been changed.

Test vessels must be illuminated with a fixed daily *photoperiod* of 16-h light and 8-h dark, and should use full-spectrum fluorescent, incandescent, or LED lights. Light intensity adjacent to the surface of the solution in each test vessel should be 100 to 500 lux, and must be at least 100 lux as a minimum. A gradual transition between light and dark is recommended (see Section 2.3.3).

4.3.4 Dissolved Oxygen and Aeration

The dissolved oxygen content (DO) of the control/dilution water used for preparing test solutions should be 90 to 100% saturation before its use, and, if necessary, the water should be aerated vigorously to achieve this.

Pre-aeration (before exposure of test organisms) of each test solution might be required or appropriate, depending on the test substance, type, and objectives (see Sections 4.1, 5.2, and 6.2). Apparatus for exposing tadpoles to test solutions with aeration is described in Section 3.1.

[58] The rate of growth and differentiation of anuran larvae is highly influenced by temperature. As temperature increases, rates of differentiation and growth both increase in *L. pipiens* until a plateau is reached at 23°C (Smit-Gill and Berven, 1979 cited in Ultsch *et al.*, 1999). Differences as small as 1 °C among replicate tanks caused differences in developmental rate in *Xenopus laevis*, which resulted in these replicates being removed as outliers (Lutz *et al.*, 2008). Similarly, a 1 °C difference in temperature between two labs conducting the same exposure with *X. laevis* was identified as a “prime candidate” to explain notable differences in developmental rate between these two labs (Lutz *et al.*, 2008).

If, and only if, the DO is <60% or >100% of air saturation in one or more test solutions when they have been freshly prepared for test initiation or solution renewal, then each test solution including the control(s) should be pre-aerated before the tadpoles are exposed to it. To achieve this, test solutions should be pre-aerated[59] for 30 minutes at a rate of 6.5 ± 1 mL/min·L. Immediately thereafter, the dissolved oxygen content of the sample or solutions must be measured. If, and only if, the measured value in one or more solutions is <60% or >100% of air saturation, then pre-aeration of either sample or all test solutions (including the control) should be continued at the same rate (i.e., 6.5 ± 1 mL/min·L) for an additional period not to exceed 90 minutes. This additional period of pre-aeration must be restricted to the lesser of 90 minutes and attaining 60% saturation in the highest test concentration (or 100% saturation, if supersaturation is evident).[60] Immediately thereafter, tadpoles must be exposed to each test solution, regardless of whether 60% to 100% saturation was achieved in the sample or all test solutions. Any pre-aeration must be reported, including the duration and rate (Section 7.1.6).

For a static-renewal test (see Section 4.3.2) each test solution, including the controls, must be aerated continuously throughout the test using gentle and controlled aeration (e.g., 6.5 ± 1 mL/min·L). Aeration must be provided using conventional air-control valves and aeration apparatus (e.g., oil-free compressed air dispensed through a narrow bore pipette, capillary tubing, or an airstone; see Section 2.3.5).

Dissolved oxygen must be measured and recorded at the beginning and end of the test and before each test solution renewal period (i.e., a minimum of 3 times per week on non-consecutive days) in representative test concentrations (i.e., at least low, medium, and high test concentrations, and controls), of the freshly prepared test solutions, which should meet the requirements of ≥60 to ≤100% saturation described earlier in this section. Measurements must also be made in used test solutions for each representative concentration to establish the extent of oxygen depletion that occurred prior to renewal. Measurements must be made in a way that minimizes the risk of accidentally harming tadpoles (e.g., before tadpoles are added to test solutions at the start of the test, on portions of new or used solutions during each renewal period, or after tadpoles are removed from the test solutions at the end of the test).

Oxygen in the test vessels should not fall below 60% saturation (5 mg/L at 23 °C). If it does, the investigator should be aware that the test is not measuring the toxic quality, *per se*, of the material or substance being tested. Rather, such a test would measure the total effect of the material (e.g., effluent) or substance (chemical) inducing its deoxygenating influence. The low-level aeration used in this test method should maintain DO levels ≥60% saturation.[61, 62]

---

[59] A volume of sample or of each test solution adequate to prepare or renew all replicate groups should be pre-aerated in a non-toxic container of suitable size. Pre-aeration should be oil-free compressed air dispensed through a narrow-bore pipette, capillary tubing, or a commercial air diffuser.

[60] Aeration can strip volatile chemicals from solution or can increase the rate of chemical oxidation and degradation to other substances. However, aeration of test solutions before exposure of test organisms could be necessary due to the oxygen demand of the test material (oxygen depleted in the sample during storage). If it is necessary to pre-aerate or aerate any test solution, *all* solutions are to be aerated in the manner stipulated in Section 4.3.4.

[61] Throughout inter-laboratory testing, dissolved oxygen was generally ≥80% saturation; however, some laboratories occasionally observed levels of 68–79% saturation (Nautilus Environmental, 2020a, 2020b). In addition, during round 1 of testing, the DO dropped to 50% in one laboratory, but after water renewal the DO increased to 96%; this single instance of low DO did not impact survival (Nautilus Environmental, 2020a). Note that northern leopard frog tadpoles are known to tolerate water with lower dissolved oxygen, as ASTM (2022a) recommends a minimum dissolved oxygen level of 3.0 mg/L at 23 °C (~35% saturation). ASTM (2022a) states that DO < 3.0 mg/L can stress the organisms and may induce mortality during holding and testing of *L. pipiens* tadpoles.

[62] Natural waters, site waters, or the occasional high concentration of a chemical might have high oxygen demand and deplete dissolved oxygen in the test water,

A flow-through test can be performed with or without aeration of the test solutions since the continuous flow of fresh solution to each test vessel provides an ongoing exchange of solution across the developing tadpoles. The nature of the test substance (e.g., volatility, oxygen demand, stability) should be considered when deciding if a flow-through setup is appropriate and whether or not to aerate. Depending on the oxygen demand, gentle aeration of each test solution might be necessary during flow-through tests to maintain dissolved oxygen at adequate levels of 60 to 100% saturation. If aeration is used, each replicate solution (including the controls) must be aerated at a similar and controlled rate, as previously described. Alternatively or additionally, more rapid renewal of solutions might be required to maintain DO at 60 to 100% of saturation.

4.3.5 pH

The pH must be measured in representative test concentrations (i.e., at least low, medium, and high test concentrations, and controls) at the beginning and end of the test, as well as at the beginning and end of each renewal period (i.e., in the fresh test solution and the used solution). Measurements must be made in a way that minimizes the risk of accidentally harming tadpoles (e.g., before tadpoles are added to test solutions at the start of the test, on portions of new or used solutions during each renewal period, or after tadpoles are removed from the test solutions at the end of the test).

Toxicity tests should normally be carried out without adjustment of pH. However, if the sample of test material or substance causes the pH of any test solution to be outside the range of 6.0 to 8.5, and it is desired to assess toxic chemical(s) rather than the deleterious or modifying effect of pH, then the pH of the solutions or sample should be adjusted before adding tadpoles, or a second, pH-adjusted test should be conducted concurrently using a portion of the sample.[63]

For an adjusted test, the initial pH of the sample or of each test solution could, depending on objectives, be adjusted to within ± 0.5 pH units of that of the control/dilution water, before exposure of the larvae. Another acceptable approach for this second, pH-adjusted test is to adjust each test solution, including the control, upwards to pH 6.0 (if the solution has pH <6.0), or downwards to pH 8.5 (if the solution has pH >8.5). Solutions of hydrochloric acid (HCl) or sodium hydroxide (NaOH) at strengths ≤1 *N* should normally be used for all pH adjustments. Some situations (e.g., effluent samples with highly buffered pH) might require higher strengths of acid or base. Abernethy and Westlake (1989) provide useful guidelines for adjusting pH.

4.3.6 Food and Feeding

Tadpoles begin to feed once they are actively swimming, a few days after hatch, at Gosner stage 25. For both tests (14-day and 42-day tests), tadpoles must be fed *ad libitum*, at least four times per week.

Three of the feedings occur on non-consecutive days (e.g., Monday, Wednesday, and Friday) immediately following test solution renewal. The recommended food during these feedings for tadpoles is a 4:1 combination of a steamed or thawed kale and a dried

---

which might reduce the dissolved oxygen in the treatments, even with aeration.

[63] The justification for not adjusting the pH of sample or solution is that pH might have a strong influence on the toxicity of a substance or material being tested. Thus, for the (generally) low concentrations of waste found in receiving water after dilution, any change from the natural pH, with concomitant modification of toxicity, should be accepted as part of the pollution "package". That leads to the rationale that the pH should not be adjusted in tests, and that is the requirement to be followed in most instances if test solutions are in the pH range of 6.0 to 8.5. Effects on the survival, growth, and development of *L. pipiens* have been found in 96-h exposures at pH <4.4 (Freda and McDonald, 1990); and no effects were apparent at pH 8.7 in a 5-day exposure (Jofre and Karasov, 1999). Although Schlichter (1981) found that sperm motility was decreased at pH <6.5 and egg development was negatively affected at pH <6.3, he determined that eggs fertilized at near-neutral pH and then transferred to low pH had an improved rate of survival. This is justification for the different pH levels recommended for culturing test organisms versus those recommended for testing herein. The pH ranges recommended for testing take into consideration the long-term exposures (i.e., 14-day to 42–day), included in this test method protocol, the stage of development at which the test begins, and experience during the inter-laboratory study. During three rounds of the inter-laboratory study, four participating labs successfully tested tadpoles in waters with pH ranging from 6.1 to pH 8.4.

commercial tadpole food (e.g., Ward's), which can be prepared and delivered as a homogeneous slurry (i.e., ensure kale is sufficiently blended into small pieces; see Section 2.3.6). The amount of water used to prepare the slurry should be minimized as much as is practical, to prevent dilution of the exposure solution. The slurry should be prepared fresh on each feeding day. A volume of slurry of about 10 mL per tank works well for routine feeding during the test. The food should be kept in suspension throughout feeding (e.g., swirl occasionally) so that the amount of food delivered is equal among tanks. Food must be provided *ad libitum*; however, example starting points for the amount of food to be provided during the test are based on a target of 10 % of the total body weight of tadpoles in a given replicate. The weight of food provided therefore increases during the test to accommodate the increased needs of the growing test organisms. Interim non-destructive measurements of tadpole wet weight can be carried out using the guidance provided in Section 4.4. Example feeding rations are shown in Table 5.

Table 5 Example feeding rations for toxicity tests*

Week	Kale (g)	Dried tadpole food (g)
14-Day Test		
1	0.2	0.05
2	0.3	0.075
42-Day Test		
1	0.3	0.075
2	0.4	0.1
3	0.5	0.125
4	0.6	0.15
5	0.7	0.175
To test end	0.8	0.2

* Feeding rates represent examples of food rations that can be provided as a starting point for each test vessel containing 10 tadpoles starting with GS 25 for the 14-day test and GS 28/29 for the 42-day test.

The feeding rations in Table 5 are examples only for tadpoles starting at GS 25 for the 14-day test and tadpoles starting at GS 28/29 for the 42-day test. The rations must be adjusted based on daily observations of available food (see Section 2.3.6). Although test organisms are fed every-other day, food must be available to them at all times. If during daily observations it is noticed that all of the food is gone in over half of the test vessels, the food ration must be increased. Likewise, feeding rations must be reduced if food is remaining in the dilution-water control treatment at test solution renewal (i.e., two days after being placed in the test vessel).[64] The same quantity of food must be added to each test vessel at a given feeding.

The kale/dried tadpole diet must be supplemented one day per week (i.e., a fourth feeding day per week, usually on the weekend for convenience) with algae pellets/wafers (e.g., Hikari mini algae wafers), delivered *ad libitum* (see Section 2.3.6). The weight of algae pellets/wafers is based on a target of 3% of the total body weight of tadpoles in a given replicate. For the 14-day test, this is typically 1–2 algal pellets per tank. For the 42-day test, the first feeding is typically 2 algal pellets per tank, and gradually increases over the course of the test by 1–2 algal pellets per week. The algae pellets may be broken up or partially ground prior to feeding. The feeding rations must be adjusted based on daily observations of available food, as described in the previous paragraph.

Detailed record of the food type and ration added to each test vessel must be made on each feeding occasion. Observations of the feeding behavior, and uneaten food in each test vessel should also be recorded during daily observations.

4.4 Test Observations and Measurements

The condition, appearance, and number of live tadpoles transferred to each test vessel on Day 0 must be observed and recorded.

Test vessels must be examined daily throughout the test. Observations and records should be made at this time regarding feeding activity, unusual appearance of test solution, and the presence and quantity of any uneaten food. Any mortality, unusual appearance

[64] Reduced feeding can be caused by exposure to toxicants or the cessation of feeding during metamorphosis. Feeding rations may also be reduced to compensate for mortalities or if metamorphs are removed from a replicate (i.e., lower numbers of test organisms in the vessel). Therefore, the reduction of feeding rations is based on observations of the dilution-water control treatment.

(i.e., grossly visible malformations or lesions) or behaviour of test organisms as well as approximate developmental stage must be recorded. Any dead tadpoles must be removed as soon as they are noted, and their numbers recorded.

For the 42-day test, any test organisms that reach Gosner stage 42 (i.e., characterized by the emergence of both forelimbs) must be removed from the test vessel, euthanized, and test-end measurements (i.e., wet weight, length, biomass, developmental stage, and any deformities) performed and recorded (see Section 4.5).

Mortality is determined when an organism does not appear to have any respiratory functions or movement and does not respond to gentle prodding stimuli using a glass stir rod or upon removal from the water.

Tadpoles in early stages (Gosner 25–30) are typically active only intermittently, and normally rest at the bottom of the tank. Tadpoles in later stages (Gosner 30–42) are still only active intermittently, although show a stronger startle response (e.g., during siphoning associated with water changes) than younger tadpoles. When food is added to the tank, tadpoles at both early and late stages do not respond with immediate feeding. Except for feeding on dead or dying test organisms, tadpoles do not display aggression to other tadpoles. Given their general lower level of activity in comparison to some fish species (e.g., rainbow trout), toxicity manifested via abnormal behavior is likely to not be apparent. Abnormal behaviour, however, could include uncoordinated swimming, hyperventilation, and loss of equilibrium. Gross malformations could include limb deformities, scoliosis, lesions, edema, tail necrosis, and/or visible bacterial or fungal infections.[65]

Temperature of the test solutions (Section 4.3.3) must be measured daily (e.g., using a maximum/minimum thermometer) or continuously (e.g., using a continuous recording device). Dissolved oxygen and pH must be measured as described in Sections 4.3.4 and 4.3.5. Conductivity should also be measured in fresh test solutions, at least at the test start and at the beginning of each water renewal period. Total ammonia levels must be measured at a minimum in representative replicates in the low, medium, and high concentrations, and in the control(s) at the test start and end, and at the beginning and end of one water renewal per week. However, if ammonia levels before solution renewal are acceptable, ammonia measurements in fresh test solution may be omitted. More frequent renewals of test solutions should be made if pH levels drift to <6.0 or >8.5, and ammonia concentrations rise to >0.2 mg/L un-ionized NH_3-N (see Sections 2.3.5 and 4.3.5). Additionally, hardness and/or alkalinity of the control/dilution water and, as a minimum, the highest test concentration should be measured and reported at the test start and once per week.

Dissolved oxygen, conductivity, and pH may be measured using probes and calibrated meters. Ammonia may be measured using an ion-specific electrode or by extracting an aliquot of test solution for this analysis. Any probe inserted in a test vessel must be rinsed in deionized or distilled water between samples to minimize cross-contamination. For measurements of hardness, alkalinity, and ammonia requiring sample aliquots, samples of test solution should be taken directly from the test vessel or used solution.

In addition to the observations described above, there are certain additional observations and measurements to be made during tests with chemicals (see Section 5.4). If chemicals are to be measured, sample aliquots should be taken from at least the low, medium, and high test concentrations

[65] There are several deformities that commonly occur in *L. pipiens* tadpoles that do not appear to be related to exposure to a toxicant and are not necessarily detrimental to the organism. These include scoliosis, unbent leg, and some toe, mouth, and eye malformations. Scoliosis was observed in the third round of ECCC's inter-laboratory study with percentages of tadpoles with scoliosis ranging from 3.6% to 22.6%. The incidence of scoliosis was not related to perchlorate dose, and no obvious patterns (e.g., across replicates) were observed (Nautilus Environmental, 2020b). The impact of scoliosis on the health of these test organisms from an ecological perspective is unknown. However, within the context of a laboratory-based measurements, scoliosis in tadpoles might affect the accuracy of total length measurements.

and the control(s) at the beginning and end of the test, as a minimum (see Section 5.4).

Interim, non-destructive measurements of tadpole development (Gosner stage), gross deformities, asynchronous development, wet weight, and total length (and optional snout-vent length) should be conducted every 14 days during the 42-day test.[66] For wet weight measurements, individually net a tadpole, tap against paper towel to remove most of the water, and tip the tadpole from the net into a tared dish. A small amount of water can then be added to the dish to keep the tadpole moist during the remaining measurements. Staging should be carried out following the guidelines provided in Section 4.2, ensuring tadpoles are handled with extreme care, remain covered with water at all times (except during weight measurement), and measurements are made as quickly as possible to avoid overheating or unnecessary stress. The length measurement(s) of each live tadpole can be made at the same time using a ruler, calipers, or digital imagery. Experience with handling, staging, and taking measurements on live tadpoles is recommended prior to carrying out these non-destructive, interim measurements in a definitive test.

4.5 Ending the Test

At the end of both tests (Day 14 or Day 42, depending on the test option), observations of the number and percentage of surviving tadpoles and/or metamorphs must be recorded in each replicate. Live tadpoles or metamorphs are removed from the test vessel and humanely euthanized (e.g., by immersion in a solution of tricaine methanesulfonate [MS-222], buffered to approximately pH 7 using sodium bicarbonate).[67] Once mortality is confirmed (see

[66] Interim measurements in the third round of ECCC's inter-laboratory study with perchlorate demonstrated more sensitive growth endpoints after 14 and 28 days of exposure as compared to the end of the test (42 days). In addition, there were no observations of mortality directly following these non-destructive interim measurements, suggesting that handling stress of individual tadpoles was minimal. Interim measurements might be less accurate than final measurements taken on euthanized individuals since working quickly with live organisms can be challenging. Consistency in methods used for handling the tadpoles and taking the measurements might improve accuracy. In addition, for assessing the developmental stage in particular, accuracy might be improved if the same observer conducts all of the interim and final assessments (Nautilus Environmental, 2020b).

[67] According to CCAC guidelines on the euthanasia of amphibians (CCAC 2010, 2021), acceptable methods include immersion in or injection of buffered MS-222 or benzocaine, topical application of benzocaine gel, injection of barbiturates into the lymph sac, or overdose of inhalant anesthetics. Amphibian larvae should be euthanized using the same methods as adults (AVMA, 2020; CCAC, 2021). Immersion in buffered MS-222 followed by a physical method is recommended; freezing in liquid nitrogen is acceptable, but larger larvae should be anesthetized before using this method (AVMA, 2020; CCAC, 2021).

The inhalant method must be followed by a physical method of euthanasia (i.e., decapitation or pithing), but inhalants are generally not effective for amphibians, and are not acceptable for *Xenopus* species (CCAC, 2021). For the immersion method, CCAC (2021) refers to the AVMA (2020) guidelines, which recommend prolonged immersion in buffered MS-222 with a concentration of 5–10 g/L for at least one hour. The immersion and topical application methods should be followed by a secondary physical method to ensure death (CCAC, 2021). The immersion dose recommended for *L. pipiens* tadpoles in the ASTM (2022a) sediment toxicity test method is 1 mL of a 2 g/L MS-222 solution, buffered to pH 7, in 10 to 20 mL of water containing the test organisms. Robinson *et al.* (2019a, 2020, 2021) anesthesized *L. pipiens* tadpoles by immersion in 0.01–0.02% buffered MS-222 and then euthanized tadpoles by immersion in 0.2% buffered MS-222. Robinson *et al.* (2019b) euthanized GS46 *L. pipiens* frogs by direct immersion in 1% buffered MS-222. Allran and Karasov (2000) anesthesized *L. pipiens* metamorphs in 0.5 g/L MS-222 followed by euthanization via decerebration. OECD (2009) and USEPA (2015) recommend 150 to 200 mg/L and 0.03% (w/v) of buffered MS-222, respectively, for euthanizing *X. laevis*. Torreilles *et al.* (2009), however, determined that higher doses than those recommended by OECD (2009) were required for euthanasia of *X. laevis*. Immersion of frogs in 5 g/L MS-222 resulted in deep anesthesia within 4 minutes, but at least 1 hour of immersion at this concentration was required to reliably euthanize 100% of *X. laevis* test organisms. Clove oil (eugenol) has been shown to provide a safer and more economical alternative to MS-222, and has been used in some studies for euthanizing tadpoles of North American amphibians (McDaniel *et al.*, 2004) and fish (Holloway *et al.*, 2004). However, clove oil is not currently approved in

determination of mortality in Section 4.4), the animals are then removed from the MS-222, rinsed, and placed on paper towel to remove excess moisture. The stage of development (i.e., Gosner stage) for each tadpole/metamorph must be determined under a dissecting stereomicroscope and recorded. Figures 2.1 and 2.2 provide guidance for assigning the appropriate Gosner stage.[68] Any determination of asynchronous development must also be reported.[69] If asynchronous development is observed, the tadpole developmental stage should be based on the most advanced feature.[70] Each tadpole is then placed into a tared weigh dish, weighed individually on an electronic balance (to the nearest 0.01 g), and the wet weight must be recorded. The total body length (and optionally, the snout-vent length, i.e., length from the tip of the snout to the anus)[71] for each individual tadpole must then be measured along the center line of the body (to the nearest 0.5 mm) using a metric ruler, calipers, or digital imagery and recorded.[72] Regardless of the method used for measuring tadpole lengths and weights, all measurements must be made using the same method. Instrument and verification checks for accuracy and regular maintenance are part of the

---

Canada for use on amphibians (CCAC, 2021).

[68] At Gosner stage 25, the spiracle should be present, and the external gills have atrophied. From GS 26–30, the primary distinguishing feature is the length of the hindlimb bud. From GS 31–35, foot paddle/indentation changes take place. From GS 36–39, toe separation occurs with the development of metatarsal tubercles and subarticular patches. At GS 40, foot tubercles and a vent tube are present. At GS 41, the forelimbs are visible and under the skin, and the vent tube is gone. At GS 42, both forelimbs have emerged (one forelimb may emerge before the other) (Nautilus Environmental, 2020a, 2020b).

[69] Developmental stage data can be used to determine if development is accelerated, delayed, or unaffected. Acceleration or delay of development is determined by making a comparison between the median stage achieved by the control and treated groups. Asynchronous development can be reported when tissues are examined and are not malformed or abnormal, but relative timing of the development of different tissues is disrupted within a single tadpole (OECD, 2009). For ECCC's inter-laboratory investigation, hindlimb length and development was used with the developmental stage of the tadpole body to determine if development was asynchronous. Asynchronous development was reported when the relative timing of the morphogenensis or development of different parts of the body was disrupted. Asynchronous development was observed in tadpoles exposed to 0.67 µg/L of thyroxine in round 2 of ECCC's inter-laboratory study. In cases where asynchronous development is observed, professional judgement is needed to definitively assign Gosner stage to an individual tadpole (Fort Environmental Laboratories Inc., 2018).

[70] For example, if both forelimbs have emerged but the hindlimbs lack a metatarsal tubercle, this tadpole would be recorded as GS 42.

[71] During ECCC's inter-laboratory investigation, snout-vent length and total length measurements showed equivalent sensitivity to the toxicants tested in all three rounds (NaCl, triclosan, and perchlorate). Therefore, total length is the preferred and required measurement for length as it is easier to identify compared to the location of the vent tube required for the snout-vent length measurement (Nautilus Environmental, 2020a, 2020b). The snout-vent length may be measured as an optional additional endpoint; an extra length measurement can be useful for identifying outliers or transcription errors of total length measurements.

[72] Digital photography and digitizing software can be used to carry out length measurements at the end of the test and is recommended for use in both the OECD (2009) and USEPA (2009) Amphibian Metamorphosis Assay (AMA) methods. However, the accuracy of measurements made using digital photography relies on tadpole/metamorph position and placement relative to the dissecting scope platform and camera lens (Coady *et al*., 2014). Coady *et al.* (2014) found that measuring to 0.1 mm was reasonable based on the precision of the measure due to variability in placement of the microscope for imaging. Snout-vent length may be challenging to measure since the vent can be difficult to see clearly in digital photographs (Coady *et al*., 2014). In ECCC's inter-laboratory investigation, total body length measurements made by one laboratory using a ruler were verified by an independent consultant using computer digitization. On average, the digital measurements were 2.4 mm less than those recorded using a ruler. Since computer digitization allows for the measurement of the curvature of the body, these results were suprising. The differences were attributed to the greater resolution of the tip of the snout and tail using computer digitization when compared to the manual measurement (Fort Environmental Laboratories, Inc., 2018).

quality management system of the laboratory (ISO, 2017).

Although it was the intention of Environment and Climate Change Canada to use corrected total length, corrected snout-vent length (SVL), corrected wet weight, and corrected biomass (see Section 4.6.1) as additional test validity criteria for definitive tests, there were insufficient data at the time of publication on which to base minimum requirements for control organisms. The lowest mean values obtained for growth endpoints during ECCC method development testing can, however, be used as benchmarks for achievable growth for control organisms.

These values are as follows for the 14-day test:

- mean corrected total length ≥ 14.4 mm
- mean corrected SVL ≥ 4.2 mm
- mean corrected wet weight ≥ 0.18 g
- mean corrected biomass ≥ 0.16 g

and as follows for the 42-day test:

- mean corrected total length ≥ 13.2 mm
- mean corrected SVL ≥ 4.8 mm
- mean corrected wet weight ≥ 0.44 g
- mean corrected biomass ≥ 0.42 g

Any unusual appearance or morphological abnormalities must also be recorded. Gross malformations could include limb deformities, scoliosis (see footnote 65 in Section 4.4), lesions, edema, bent tail, tail necrosis, and/or bacteria or fungal infections, etc. (Rostand, 1958; Bantle *et al.*, 1991; Ouellet, 2000).

Unless test organisms are to be used for histological, tissue, or gene expression analysis, the dry weights of the test organisms can be measured as an additional optional endpoint. Alternatively, depending on study objectives, test organisms can be placed in pre-labelled vials (either as a group per replicate or treatment, or individually) containing a fixative solution for future[73] or further endpoint determination (see Section 4.6.3.2).

Tissue samples may also be collected at the end of the test, prior to drying and/or preservation, for histology analysis, or analysis of gene expression. Biopsies can be collected from live tadpoles. Different tissue or whole-organism preservation techniques are required depending on the study objectives. Careful attention to these techniques and the additional endpoints sought in a given study is advised (see Section 4.6.3.2).

4.6 Test Endpoints and Calculations

4.6.1 Biological Endpoints

The required biological endpoints for both test options described herein include the adverse effects of test materials or substances on the survival, growth (total length, wet weight, and biomass),[74] and development (delayed or accelerated) of *L. pipiens* tadpoles. Gross deformities and asynchronous development must also be assessed and recorded.[75] Snout-vent length is an additional optional endpoint for both tests.

At the end of the exposure, the number of tadpoles alive and number dead are recorded for each replicate of the control and the various concentrations of the test material or substance. For each test, the mean (± SD) percent survival for all replicate groups (*n* = 3, 4, or 8) of tadpoles

[73] Tadpoles can be preserved following euthanization in 10% neutral buffered formalin in approximately a 1:10 ratio of tissue to preservative (Nautilus Environmental, 2020b). Alternatively, Davidson's fixative was successfully used by one inter-laboratory participant to preserve tadpoles for the assessment of development endpoints and additional histology endpoints at a later date (Nautilus Environmental, 2020a).

[74] The usual assumption within a toxicity framework is a decrease in growth, and method development work focused on observations and analysis of growth inhibition. However, substances that disrupt thyroid function may elicit an increase in growth (M. Gallant, Nautilus Environmental, Burnaby, BC, personal communication, 2023). In lieu of standardized guidance, investigators are encouraged to interpret and analyze increased growth on a case-by-case basis.

[75] Gross deformities and asynchronous development are less likely to be observed in the 14-day test because the test is shorter in duration (i.e., less time for tadpoles to develop) and begins with younger test organisms (i.e., tadpole stages are more focused on growth rather than notable morphological changes).

exposed to each treatment for the test duration must be calculated and reported.

The growth endpoint for these tests is based on the mean total length (tip of snout to tip of tail) and the mean wet weight of each individual tadpole/metamorph surviving in each treatment at the end of the test as well as the biomass for each replicate. Weight and length measurements must be corrected using initial weights and lengths (i.e., the means of the weight and length measurements made on the 20 organisms representative of the test organisms used to start the test (see Section 4.2) are subtracted from the values measured at the end of the test to give corrected measurements of weight and length) prior to any statistical calculations.[76] A substantial reduction in the size or weight is considered indicative of an adverse toxic effect of the treatment on the growth of surviving tadpoles/metamorphs.[77] The biomass must be calculated as the corrected total wet weight per replicate divided by the number of tadpoles that were placed in the replicate at the start of the test (presumably 10 tadpoles). The biomass endpoint represents a combination of sublethal effect (i.e., reduced total wet weight of surviving test organisms in each replicate at the end of the test) and mortality[78] an approach currently used in the fathead minnow larval growth test in both Canada (EC, 2011) and the United States (USEPA, 2002). Since it integrates effects on survival with those on growth, the biomass endpoint has the potential to show a greater sensitivity to toxic samples than the growth endpoint based on sublethal effect alone. Section 8.2 of Environment Canada's guidance document on statistical methods for environmental toxicity tests (EC, 2005) describes the use of this endpoint as one of three options for measuring growth as a quantitative sublethal effect. When making any calculation for survival or biomass, test organisms in any replicate that were accidentally killed or accidentally removed during the renewal of test solutions should be deducted from the initial number of test organisms for that replicate at the start of the test, as if they had not been in the test.

The development endpoint for these tests is based on the Gosner stage reached by each tadpole at the end of the test.[79] The median Gosner stage for each replicate must be calculated and reported. The median Gosner stage for each treatment, including the control treatment(s), must also be calculated,

---

[76] Using corrected values isolates for growth of the tadpole during the test exposure. Correcting for starting wet weight and length improved comparison of inter-laboratory results (Nautilus Environmental, 2020a, 2020b). As it can correct for slight variances in the organisms used to start the test, it is expected to also reduce intra-laboratory variability over time.

[77] It has been noted that a delay in development might occur in tadpoles exhibiting reduced growth. This was observed in the first round of ECCC's inter-laboratory study in tests with NaCl. Care must be taken in interpreting results in these cases, where the impact on development may not be due to thyroid disruption, but simply an indirect effect due to impact on growth (e.g., as with NaCl).

[78] Occasionally, dual effects may be observed in which the growth of survivors increases when partial mortality is observed (Nautilus Environmental, 2020a). When some organisms die it might allow more test solution, space, and/or food for the remaining organisms, thereby possibly affecting the growth or health of the survivors. These possibilities should be considered when interpreting test results (EC, 2005). Data should be examined graphically to determine the presence or absence of a relationship between growth and survival. Although no statistical correction can be made for such interactions in standard regression analyses, the effects should be reported if apparent and potential measures (e.g., lower test organism density) taken in future tests to ensure that procedures are in place to minimize the potential for these effects (EC, 2005). The calculation of biomass (instead of growth directly) can sometimes be effective in addressing dual effect. It may also be possible to incorporate dual effects by including "Count_Survival" in general linear mixed models using R software (Robinson *et al.*, 2017; S. Robinson, Environment and Climate Change Canada, Ottawa, ON, personal communication, 2023).

[79] In ECCC's inter-laboratory evaluation, it was determined that the measurement error of assigning the stage of development is approximately 1 Gosner stage (Nautilus Environmental, 2020b). This is a result of the subjectivity involved with assigning the developmental stage based on semi-quantitative observations during some stages (e.g., ratio of the length to width of limb bud in GS 26–30) or in some cases due to the different rates of development on one side of the organism compared to the other (e.g., toe indentation may not match on both sides of a given tadpole).

based on replicate medians (n = 3, 4, or 8), and reported. Depending on experimental objectives, investigators may choose further distribution-free techniques[80] for summary statistics (e.g., inter-quartile range, maximum, minimum) for either individual replicates or pooled replicates.

The presence of any abnormal tadpoles (i.e., tadpoles with gross deformities) must be noted and reported at the end of the test. Any incidence of asynchronous development (see Section 4.5) must also be reported if observed.

Any tadpoles that reach GS 42 (i.e., emergence of both forelimbs) prior to the end of the test must be removed from the test vessel and the date of their removal recorded and reported. These organisms (metamorphs) must be euthanized, and final measurements made as described above.[81] For statistical analyses, they are treated as if they survived until the end of the test, even if they were removed prior to the test end.

The most common test design for the methods described herein involves multiple concentrations of a toxicant, achieved by testing a sample of chemical/chemical product (Section 5.2), at a range of concentrations to calculate point estimates for survival and growth, and adverse effects on development. Samples of effluent, elutriate or leachate (see Section 6.2) can also be assessed in multi-concentration exposures using this method. For a multi-concentration test, the *LC50* or other LCp[82] for survival, *ICps* for growth (total length, wet weight, and biomass), and significant effects on development (Gosner stage) must be calculated and reported (data permitting).

Single-concentration tests (samples tested at full strength only) can be conducted using the methodologies described herein and are discussed elsewhere (see Section 6.5.1).

4.6.2 Multi-concentration Tests

In a multi-concentration test, the statistical endpoints are: (i) an LCp and its 95% confidence limits for the mortality of tadpoles, and (ii) ICps and their 95% confidence limits for growth (i.e., total length, wet weight, and biomass all corrected for initial measurements) of surviving test organisms at test end,[83] and for the 42-day test only, (iii) significant effects on development (Gosner stage) of surviving tadpoles in any treatment(s) relative to the control(s) at the end of the test. Environment Canada (2005) provides direction and advice for calculating LCp and ICp endpoints, which should be followed;

[80] Gosner stage is a measurement that is not normally distributed. As such, summary statistics that rely on the normal distribution (e.g., mean, standard deviation) are not appropriate.

[81] Metamorphs are removed from the test vessel to prevent them from drowning, as they require a terrestrial surface that the tadpole test vessel does not have. In addition, this biological test method is designed to measure effects on tadpoles up until the point of metamorphosis, but not beyond that point. Metamorphs are euthanized to complete more accurate final measurements of biological endpoints. Section 4.6.3 provides additional guidance on alternative test designs, such as extending test duration, if study objectives include assessing effects on metamorphs.

[82] For a typical lethality test, investigators choose exposure concentrations that are expected to result in partial and complete morality, with the goal of ensuring a robust LC50 calculation. In contrast, this test method is focused on chronic, sublethal endpoints. As such, investigators are encouraged to choose concentrations of toxicants that are expected to result in less than 50% mortality. As a result, LC25 may be a more appropriate endpoint. Alternatively, if the goal is to focus on sublethal effects only, lethality analysis would be optional.

[83] Historically, investigators have frequently analyzed *quantitative* sublethal data from multi-concentration tests by calculating the no-observed-effect concentration (NOEC) and the lowest-observed-effect concentration (LOEC). Disadvantages of these statistical endpoints include their dependence on the test concentrations chosen, failure to adequately describe the exposure-response curve, and the inability to provide any indication of precision (i.e., no 95% or other confidence limits can be derived) (NERI, 1993; EC, 2005; Landis and Chapman, 2011). Given these disadvantages, ICp is the required statistical endpoint for growth data derived in a multi-concentration test using early aquatic life stages of frogs. Regression-based approaches were also explored for the developmental endpoint. However, due to the ordinal nature of developmental data (Green *et al*., 2018), and for ease of interpretation, MDAU decided to use a robust and sensitive hypothesis test for the developmental endpoint, rather than a point estimate.

Sections 4.6.2.1 and 4.6.2.2 give further guidance in this regard. Section 4.6.2.3 provides guidance on determining developmental impacts based on Gosner stage at the end of the 42-day test. In statistical tests used to derive endpoints, concentrations can be entered as logarithms. Initially, regression techniques (see Section 4.6.2.2) must be applied to multi-concentration data intended for calculation of an ICp.[84] In the event that the data do not lend themselves to calculating the ICps for growth using the appropriate regression analysis, linear interpolation of these data using the program ICPIN should be applied in an attempt to derive an ICp (see Section 4.6.2.2).

An initial plot of the raw data (percent mortality, corrected data for length, wet weight, and biomass and developmental stage) against the logarithm of concentration is highly recommended, both for a visual representation of the data, and to check for reasonable results by comparison with later statistical computations. Any major disparity between the approximate graphic LCp or ICp and the subsequent computer derived LCp or ICp must be resolved. The graph would also show whether a logical relationship was obtained between log concentrations (or, in certain instances, concentration) and effect, a desirable feature of a valid test (EC, 2005).

4.6.2.1 LCp

When a multi-concentration test for effects of exposure of aquatic life stages of *L. pipiens* to test substances or materials is conducted (either the 14- or 42-day test option), the *quantal* mortality data must be used to calculate (data permitting) the appropriate LCp.[85] To estimate an LCp, mortality data at the specified period of exposure is tabulated for each replicate, and data for each replicate is entered directly for statistical analysis. Any LCp that is calculated and reported must include the 95% confidence limits.

The guidance provided by Environment Canada (2005) on choosing statistical test methods to be applied to quantal (e.g., LCp) data should be consulted when choosing the statistical test to be applied to such data for toxicity tests using aquatic life stages of frogs.

4.6.2.2 ICp

When a multi-concentration test for effects of exposure of aquatic life stages of *L. pipiens* to test substances or materials is conducted (either the 14- or 42-day test option), the *quantitative* data representing growth (total length, wet weight, and biomass, corrected for measurements at test initiation; see Section 4.6.1) must be used to calculate ICps. The ICp is a quantitative estimate of the concentration causing a fixed percent reduction in the length, wet weight, or biomass of tadpoles during the test.

The ICp is calculated as a specified percent reduction (e.g., the IC25 and/or IC20, which represent 25% and 20% inhibition, respectively). The desired value of *p* is selected by the investigator, typically 25% or 20%. Any ICp that is calculated and reported must include the 95% confidence limits.

These calculations must be made using the appropriate linear or non-linear regression analyses (described in this section). If, however, regression analyses fail to provide meaningful ICps for the three growth endpoints, then ICPIN analyses (described in this section) should be applied to the

[84] Regression is the method of choice for estimating an ICp. It involves fitting the data mathematically to a selected model and then calculating the statistical endpoint using the model that best describes the exposure-concentration-response relationship. Non-linear regression techniques were originally recommended by Stephenson *et al.* (2000) for several reasons including: the relationship that exists between exposure concentration and the test organism response is typically non-linear; the *heteroscedasticity* of the data is rarely reduced by transformation; the more standard bootstrap simulation technique has several limitations for these types of data; and non-linear regression can fit effect distributions showing *hormesis*. By using standard mathematical techniques, a regression can be well-described in terms that convey useful information to others, effects at high and low concentrations can be predicted, and confidence intervals can be estimated. Deficiencies of the smoothing and interpolation method can be largely remedied (EC, 2005).

[85] Investigators are encouraged to develop study objectives that are focused on sublethal effects and choose exposure concentrations accordingly. As such, the test might not include a sufficient number of high (lethal) concentrations to enable the calculation of an LC50.

corresponding data. Any procedures applied to the data, details regarding any transformation of the data, and the statistical method used for the calculation of ICp must be reported.

Use of regression analysis. Upon completion of a definitive multi-concentration test, ICps (including its 95% confidence limits) for the growth (total length, wet weight, and biomass) of *L. pipiens* must be calculated using regression analysis, provided that the assumptions below are met. A number of models are available to assess the data (using quantitative statistical tests) via regression analysis. The proposed models for application consist of one linear model, and the following four non-linear regression models: exponential, Gompertz, logistic and logistic adjusted to accommodate *hormesis*[86] (see Section 6.5.8 in EC, 2005). Use of regression techniques requires that the data meet assumptions of *normality* and *homoscedasticity*. The reader is strongly advised to consult EC (2005) for additional guidance on the general application of linear and non-linear regression for the analysis of *quantitative* toxicity data.[87]

The general process for the statistical analysis and selection of the most appropriate regression model (linear or non-linear) for quantitative toxicity data is outlined in Figure 3. The selection process begins with an examination of a scatter plot or line graph of the test data to determine the shape of the concentration-response curve. The shape of the curve is then compared to available models so that one or more appropriate model(s) that best suits the data is (are) selected for further examination (refer to Figure O.1, Appendix O, in EC, 2005 for an example of five potential models).

Once the appropriate model(s) is (are) selected for further consideration, assumptions of *normality* and *homoscedasticity* of the *residuals* are assessed. If the regression procedure for one or more of the examined models meets the assumptions, the data (and regression) are examined for the presence of outliers. If an outlier has been observed, the test records and experimental conditions should be scrutinized for human error. If there are one or more outliers present, the analysis should be performed with and without the outlier(s), and the results of the analyses compared to examine the effect of the outlier(s) on the regression. Thereafter, a decision must be made as to whether the outlier(s) should be removed from the final analysis. The decision should take into consideration natural biological variation, and biological reasons that might have caused the apparent anomaly. Additional guidance on the presence of outliers and unusual observations is provided in Section 10.2 of EC (2005). If there are no outliers present or none are removed from the final analysis, the model that demonstrates the smallest residual mean square error is selected as the model of best choice.[88] Additional guidance from a statistician familiar with dealing with outlier data is also advised.

Normality should be assessed using the *Shapiro-Wilk's test* as described in EC (2005). A normal probability plot of the *residuals* may also be used during the regression procedure, but is not recommended as a stand-alone test for normality, as the detection of a "normal" or "non-normal"

[86] A hormetic response (i.e., *hormesis*) might be observed at one or more of the lowest, sublethal concentration(s), i.e., performance at such concentration(s) is enhanced relative to that in the negative control (see Section 10.3 in EC, 2005). For instance, enhanced tadpole growth in samples at low test concentrations has been observed relative to those in the control treatment (Nautilus Environmental, 2020a, 2020b). To calculate the ICp when this phenomenon occurs, the data should be analyzed using the hormesis model. The hormetic effects are included in the regression, but do not bias the estimate of the ICp. An estimated IC25 would still represent a 25% reduction in performance from that of the control. Hormesis has not been observed in the development endpoint for tadpoles to date.

[87] Some of the specific guidance provided in EC (2005) refers to the use of a general-purpose statistical package (i.e., SYSTAT); however, CETIS (a software package designed for environmental toxicology) contains the models described herein for regression analysis. The latest version of SYSTAT is available for purchase; see website: www.systatsoftware.com/systat/. The latest version of CETIS is available for purchase by contacting Tidepool Scientific Software, P.O. Box 2203 McKinleyville, CA 95519, USA; Phone/Fax 707-839-5174; email: sales@tidepool-scientific.com; website: www.tidepool-scientific.com/Cetis/CetisStats.html.

[88] The Akaike Information Criterion (or an equivalent, such as the Bayesian Information Criterion) is another option for determining best model fit.

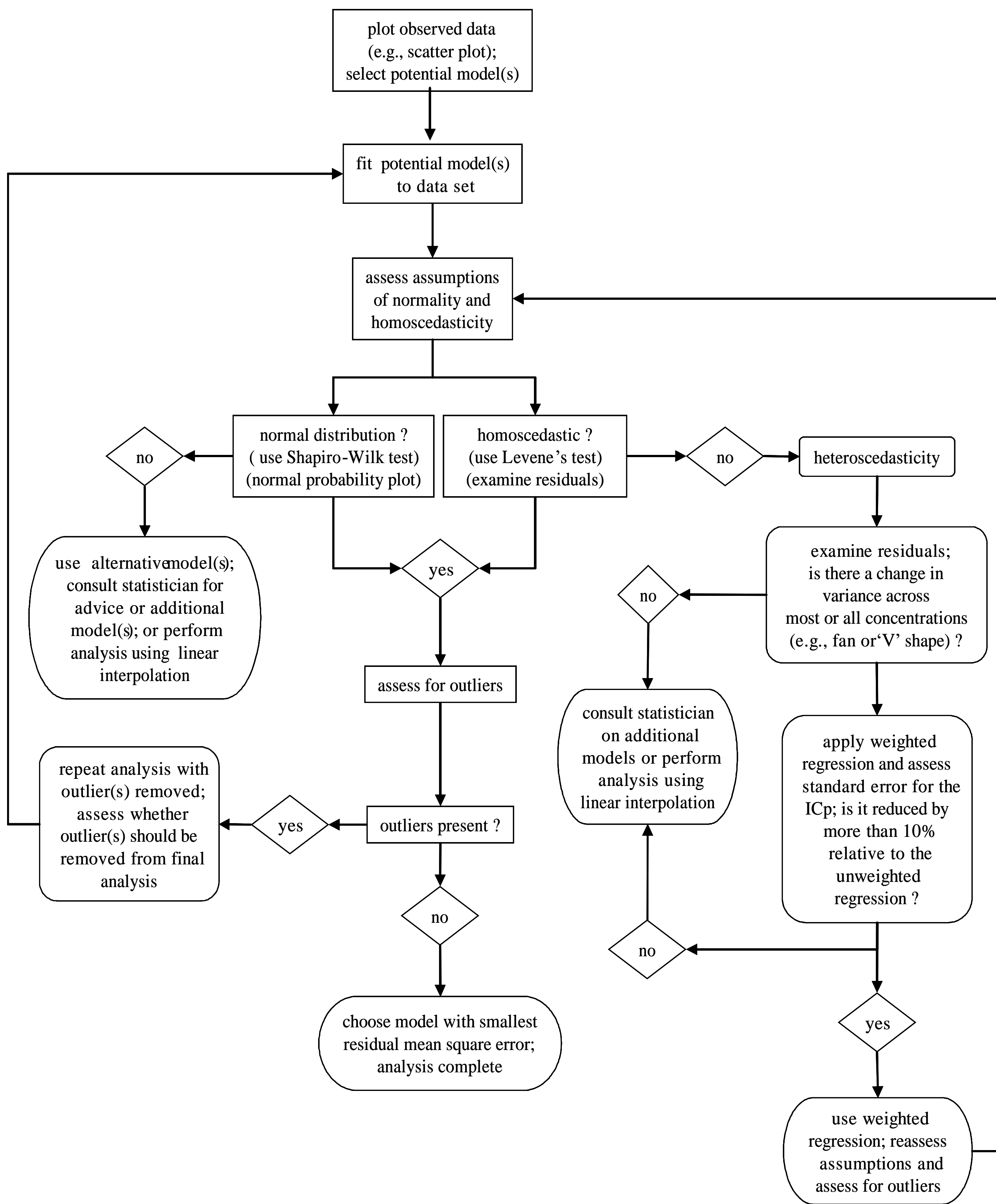


Figure 3 The general process for the statistical analysis and selection of the most appropriate model for quantitative toxicity data (adapted and modified from Stephenson *et al.*, 2000)

distribution is dependent upon the subjective assessment of the user. If the data are not normally distributed, then the user is advised to try another model, consult a statistician for further guidance on model selection, or perform the less-desirable linear interpolation (using ICPIN, as described in this section) method of analysis.

Homoscedasticity of the *residuals* should be assessed using *Levene's test* as described in EC (2005), and by examining the graphs of the residuals against the actual and predicted (estimated) values. Levene's test provides a definite indication of whether the data are homogeneous (e.g., as in Figure O.2A of Appendix O in EC, 2005) or not. If the data (as indicated by Levene's test) are *heteroscedastic* (i.e., not homogeneous), then the graphs of the residuals should be examined. If there is a significant change in the variance and the graphs of the residuals produce a distinct fan or "V" pattern (refer to Figure O.2B, Appendix O in EC, 2005 for an example), then the data analysis should be repeated using weighted regression. Traditionally, the data have been weighted by dividing by the inverse of the variance; however, other options are available. Before choosing the weighted regression, the standard error of the ICp is compared to that derived from the unweighted regression. If there is a difference of greater than 10% between the two standard errors,[89] then the weighted regression is selected as the regression of best choice. However, if there is less than a 10% difference in the standard error between the weighted and unweighted regressions, then the user should consult a statistician for the application of additional models, given the test data, or the data could be re-analyzed using the less-desirable linear interpolation (using ICPIN, as described in this section) method of analysis. This comparison between weighted and unweighted regression is completed for each of the selected models while proceeding through the process of final model selection (i.e., model and regression of best choice). Some non-divergent patterns might be indicative of an inappropriate or incorrect model (refer to Figure O.2C, Appendix O in EC, 2005, for an example), and the user is again urged to consult a statistician for further guidance on the application of additional models.

Endpoints generated by regression analysis must be bracketed by test concentrations; extrapolation of endpoints beyond the highest test concentration is not an acceptable practice (EC, 2005).

Linear interpolation using ICPIN. If regression analyses of the endpoint data (see preceding text) fail to provide an acceptable ICp for growth (i.e., assumptions of normality and homoscedasticity cannot be met), linear interpolation using the computer program called *ICPIN* should be applied. This program (Norberg-King, 1993; USEPA, 2002) is not proprietary, is available from the USEPA, and is included in most computer software for *environmental toxicology,* including TOXSTAT (1996) and CETIS (2020). The original instructions for ICPIN from the USEPA are clearly written and make the program easy to use (Norberg-King, 1993).[90] An earlier version was called BOOTSTRP.

Analysis by ICPIN does not require equal numbers of replicates in different concentrations. The ICp is estimated by smoothing of the data as necessary, then using the two data-points adjacent to the selected ICp (USEPA, 2002, Appendix M). The ICp cannot be calculated unless there are test concentrations both lower and higher than the ICp; both those concentrations should have an effect reasonably close to the selected value of p, preferably within 20% of it. At present, the computer program does not use a logarithmic scale of concentration, and so Canadian users of the program must enter the concentrations as logarithms. Some commercial computer packages have the logarithmic transformation as a general option, but investigators should make sure that it is actually retained when proceeding to ICPIN. ICPIN estimates confidence

[89] The value of 10% is only a "rule-of-thumb" based upon experience. Objective tests for the improvement due to weighting are available, but beyond the scope of this document. Weighting should be used only when necessary, as the procedure can introduce additional complications to the modelling procedure. A statistician should be consulted when weighting is necessary.

[90] The instructions in Norberg-King (1993) are sometimes misleading on the identity of "replicates." The term is used in such a way that it would apply to numbers of individual organisms within the same vessel. This slip of wording does not affect the functioning of the program. Some commercial programs have been less user-friendly for entry of data and analysis.

limits by a special "bootstrap" technique because usual methods would not be valid. Bootstrapping performs many resamplings from the original measurements. The investigator must specify the number of resamplings, which can range from 80 to 1000. At least 400 is recommended here, and 1000 would be beneficial.[91] Graphs of reduction in total length, wet weight, and biomass should be plotted against the logarithm of concentration, to check the mathematical estimations and to provide visual assessments of the nature of the data (EC, 2005). If the ICPIN program is used when there is a hormetic effect, an inherent smoothing procedure could change the control value and bias the estimate of ICp. Accordingly, before statistical analysis, hormetic values at low concentration(s) should be systematically replaced by the control value. This is considered a temporary expedient until a superior approach is established (see Option 4 in Section 10.3.3 in EC, 2005). The correction is applied for any test concentration in which the average effect (i.e., the geometric average of the replicate means) is higher ("better") than the average for the control. To apply this correction, replace the observed growth measurements of the replicates in the hormetic concentration(s), with the means of replicates in the control. The geometric average for that/those concentration(s) will then be the same as that for the control.

4.6.2.3 Development Endpoint

When a 42-day multi-concentration test for estimating the effects on aquatic life stages of *L. pipiens* is conducted, the *ordinal* data representing development (Gosner stage) must be used to test for effects on development in the treatment(s) compared to the control(s). The outcome of these tests will be the identification of which treatments are significantly different from the control(s). The median Gosner stage for each replicate is entered directly for statistical analysis. The recommended statistical test for treatment effects on Gosner stage is the step-down application of the Jonckheere-Terpstra test, and preferably the multi-quantile Jonckheere-Terpstra test from the 20th to the 80th percentiles to account for changes to the distribution profile (EC, 2005; OECD, 2009; Green *et al.*, 2018; John W Green Ecostatistical Consulting, 2021).[92] While other statistical analyses are available for developmental data, ECCC has experience with the Jonckheere-Terpstra test. In addition, power analysis determined that the test design described in Section 4.2 for the 42-day test has acceptable power (i.e., ≥80%) to detect a difference of ≥2 Gosner stages between treatment(s) and control(s)[93] using the Jonckheere-Terpstra test or the multi-quantile Jonckheere-Terpstra test (John W Green Ecostatistical Consulting, 2021; see Section 4.6.2.4).

4.6.2.4 Power Analysis

An important factor to consider in the analysis of the results for toxicity tests is the potential for declaring false positives (i.e., declaring clean water contaminated; Type I error) or false negatives (i.e., declaring contaminated water clean; Type II error). Scientists are usually cautious in choosing the level of significance for tolerating false positive results (Type I error), and usually set it at $p = 0.05$ or 0.01. Commonly, scientists following a specified test design will never consider the relationship between power, variability, and effect size, leaving the Type II error (β) completely unspecified. There are

---

[91] ICPIN has some deficiencies, which is why it is recommended herein only in cases where the use of regression fails to provide an acceptable ICp. Its interpolation method is an inefficient use of data and is sensitive to peculiarities of the two concentrations used. The program fails to adopt logarithm of concentration, which would introduce a slight bias towards a higher value of ICp. A modification of the bootstrap method has now remedied a problem of overly narrow confidence limits; however, regression analyses provide more accurate methods of estimating the ICp and its 95% confidence limits (EC, 2005) (see Section 4.6.2.2).

[92] At time of publication, the Jonckheere-Terpstra Step-Down Test was available as a Group Comparison test option in CETIS (versions 1.9.7 and 2.1), the software package that is commonly used at toxicology laboratories in Canada.

[93] During test method development, analysts noted that differences of ± 1 Gosner stage can be expected due to organism placement or interanalyst variability when staging test organisms at certain developmental stages. To eliminate the possibility of declaring a significant treatment effect that could be explained by experimental error, differences of ≥2 Gosner stages were chosen for the power analysis. While biologically significant effects may be chosen to align with project goals, investigators are encouraged to consider changes of ≥2 Gosner stages biologically significant.

several factors that influence statistical power, including:

- variability of replicate samples representing the same treatment;
- α (i.e., the probability of making a Type I error);
- effect size (i.e., the magnitude of the true effect for which you are testing, such as a difference of ≥2 Gosner stages); and
- *n* (i.e., the number of replicates used in a test).

Environment Canada's guidance document on statistical methods for environmental toxicity tests (EC, 2005) provides further information and guidance on Types I and II errors.

In research-based science, power analysis is most useful as part of a preliminary test design (Hoenig and Heisey, 2001; Lenth, 2007; Newman, 2008). In this case, a preliminary experiment is run to determine the approximate standard deviation (variability), and to troubleshoot the execution of the experiment in general. Other factors in power analysis, such as effect size and number of replicates, can then be considered along with the variability so that the final test design is optimized (e.g., number of replicates needed to detect a certain effect size is determined).

In the development of standardized test methods, the purpose of employing power analysis remains the optimization of test design or at least estimating the power of the current test design.[94] However, instead of a single estimate for variability and effect size, there would typically be a much richer data set to consider. For example, test method experts could collect a number of estimates of variability across different laboratories and different contaminant scenarios (Thursby *et al*., 1997; Van der Hoeven, 1998; Denton *et al*., 2011, 2019).

Data from inter-laboratory testing and ALET were used to estimate power for detecting a change in development (i.e., Gosner stage) in the 42-day amphibian toxicity test. Power analysis was not performed for the 14-day test, as determining significant effects on development is not a required endpoint. Power analysis was performed using the 42-day test design described in Section 4.2, with ten tadpoles per replicate, four replicates per concentration, and eight replicates per control (John W Green Ecostatistical Consulting, 2021). Variability estimates were collected from 10 tests with chemicals, which included thyroxine (3 tests), perchlorate (3 tests), triclosan (2 tests), a pesticide (VisionMAX™, active ingredient: glyphosate; 1 test), and an unknown chemical (1 test). In the simulation studies, different data distributions were evaluated, including distributions with differences in the amount of variability, differences in expected developmental stage at the end of the test, and differences in data truncation. Alpha was maintained at 0.05 throughout, and a *target effect size* of ≥ 2 Gosner stages was used. Under these conditions, the properties of three different statistical tests (the Jonckheere-Terpstra test, the multi-quantile Jonckheere-Terpstra test, and the Rao-Scott Cochran Armitage by Slices test) were evaluated. All power analysis was performed in SAS (John W Green Ecostatistical Consulting, 2021).

The Jonckheere-Terpstra test or the multi-quantile Jonckheere-Terpstra test are the recommended statistical tests, based on ease of use, ease of interpretation, and acceptable power properties. Both of these tests could achieve 80% power in most of the simulations[95] used in the simulation study. The multi-quantile Jonckheere-Terpstra test is the preferred option because it is capable of evaluating changes at several quantiles in the data, whereas the Jonckheere-Terpstra test can only evaluate changes in the median, i.e., the 50th quantile. In using the multi-quantile Jonckheere-Terpstra test,

[94] In 2010, the USEPA introduced a data analysis approach termed the test of significant toxicity approach (TST; USEPA, 2010). The TST is a hypothesis testing approach based on bioequivalence, which is extensively used in pharmaceutical development and evaluation. It is included in the discussion here because power analysis and the TST share some similar goals (e.g., a priori statement of Type I and Type II error) and because of the similar context (application of standardized testing).

[95] Some simulations assumed high variability in the data, and in these cases, the statistical tests did not achieve 80% power. However, this high level of variability was not frequently observed in ECCC data sets (John W Green Ecostatistical Consulting, 2021).

investigators are able to base statistical conclusions on a broader array of data. It should be noted that increasing the number of test organisms per replicate (e.g., 20 or 40 organisms per replicate) increases the power of the test (John W Green Ecostatistical Consulting, 2021).

4.6.3 Other Test Designs

Although this test method document describes standardized methods for measuring the survival, growth, and development of tadpoles exposed to samples of water in 14- and 42-day exposures, ECCC recognizes the value in additional endpoints that can be added to the current test design, as described in this section. In addition, this section describes alternate test designs that can be used with *L. pipiens* and that may be more relevant to other study objectives. Additional endpoints (e.g., histopathology or gene expression) and alternate test designs (e.g., extended test duration or modified exposure vessels[96]) need not necessarily conflict with the main test design described in this test method document.[97] Rather, they are described here to allow a researcher or regulator to add on to, or slightly modify the test design to address specific study objectives. Through incorporating additional endpoints into the standardized test design or slightly modifying test design, investigators can avoid repeating exposures with *L. pipiens*, and in doing so, apply the 3Rs principle of "reduction" into test objectives (see also Section 4.9).

Investigators might also want to focus on impacts to another anuran species of ecological relevance to a specific contaminated location. Other species of frogs could be used with either of the test options described in this test method document, or alternate test designs described in this section, although modifications to certain procedures and/or conditions (e.g., test volumes, test temperature, test duration) would have to be confirmed. One or more preliminary tests (i.e., using control/dilution water) with necessary procedural adjustments are recommended in order to demonstrate that the test conditions are suitable and that the criteria for test validity described herein are achievable before it is used for measuring the toxicity of a test substance or potentially contaminated material according to either of the test options described herein (with or without modification). This method has only been validated using *L. pipiens*, however, both the wood frog (*Lithobates sylvaticus*) and the North American bullfrog (*Lithobates catesbeiana*) have been used at two ECCC laboratories (ALET and PYLET) for research and during the development of this method (see Section 1.4).

4.6.3.1 Test Duration

Two options for test duration (14-day and 42-day) are described in this test method document and selection of the most suitable option depends on the study objectives and the nature of the substance or material being tested (see Section 4.3.1). These tests are designed to determine effects on *L. pipiens* at various stages of development, and have been validated using GS 25 (for the 14-day test) and GS 28/29 (for the 42-day test) tadpoles at test initiation (see Section 2.1). A longer chronic exposure that covers more early life stages of *L. pipiens* might include a combination of the 14-day and 42-day tests for a 56-day test starting with GS 25 tadpoles. This 56-day exposure would cover most of the larval frog stages, starting with tadpoles that are actively feeding (GS 25), and ending with some tadpoles at GS 42, marked by the emergence of forelimbs (see Section 2.1 and Figure 2.2).

An alternative to conducting a 56-day test starting with GS 25 tadpoles is to conduct the 14-day test starting with GS 25 tadpoles, and also conduct the 42-day test using new organisms staged at GS 28/29. The advantage of this approach over the continuous 56-day exposure described previously is that the initial Gosner stage of the test organisms are in

---

[96] In an experiment where individual tadpole responses were of interest, ALET placed each tadpole in a separate test vessel. The advantage of this approach is that the individual responses of each tadpole can be monitored; however, this type of exposure is very time-and labour-intensive and may or may not provide better information than having multiple tadpoles per tank (P. Jackman, Environment and Climate Change Canada, Moncton, NB, personal communication, 2020).

[97] ECCC had some experience with all of the test designs described in this section during the development of the 14- and 42-day tests described in this test method document.

effect "reset" or "re-zeroed" prior to the second phase of exposure (42-day test). Some tadpoles lag in their development and might show no signs of developing beyond GS 25 or developing to metamorphosis. This lag in development can create challenges with data interpretation, especially in longer exposures. Using this approach, the variability created by the differential rates of tadpole development can be reduced by "re-setting" the developmental stage of the test organisms prior to the 42-day portion of the test. Another advantage is to "reset" control mortality if it is approaching the acceptable limits by the end of the 14-day test. This approach also has the advantage of providing an additional time point for assessing growth endpoints (i.e., after 14 days and after 42 days).

If an investigation warrants, the test could be started with embryos[98] or earlier larval stages (tadpoles) and the exposure time could vary depending on the study objectives. Chronic investigations covering all of the early life stages of *L. pipiens* might start with embryos or newly hatched tadpoles (i.e., Gosner stage 19 or 20), and continue until the completion of metamorphosis and tail resorption (i.e., Gosner stage 46; ~ 90 to 120 days after test initiation), or at least until the appearance of forelimbs (i.e., Gosner stage 42; ~ 42-90 days after test initiation). A test that includes these later stages of metamorphosis (Stages 43 to 46), would see the organisms undergo reabsorption of the tail and an increase in mouth size with replacement of larval feeding structures by adult jaws and tongue (see Figure 2.2). Also, forelimbs and hindlimbs become functional in these later stages of development (McDiarmid and Altig, 1999). If the test is to be continued until tail resorption, it would be necessary to add a substrate or platform (to mimic terrestrial land) to the tanks to prevent drowning of the metamorphs. As an alternative, water could be lowered, and the test vessels tipped at an angle to allow the metamorphs to climb out of the water. For shorter exposures focussed on earlier stages of development (i.e., depending on study objectives), a test might be initiated with embryos (i.e., (<48 h old, Gosner stage 10 to 12) and continue just until tadpoles are actively swimming (Gosner stage 25; ~7 days after test initiation). This exposure would incorporate the very early larval stages not covered in the 14- or 42-day test exposures described in this test method document (see footnote 98 for more discussion on the use of embryos in toxicity tests).

Tests can be terminated after a fixed duration of exposure, as described herein for the 14- and 42-day tests. Alternatively, test termination can be stage-based and a test terminated when a certain percentage (e.g., 80%) of control tadpoles reach a specific stage of development (e.g., GS 42 or GS 46). Depending on study objectives, this can be advantageous where rate of development is of

---

[98] If beginning the test using embryos is desired, test incubation units similar to those used in Environment Canada's early life stage rainbow trout test (EC, 1998) can be used to contain the embryos inside the test vessels, to allow for flow of test solution around the embryos, and to allow for renewal of test solutions with minimal disturbance (Yee *et al*., 1996). The incubation unit is made from an 800-mL (or larger) Tri-Pour™ plastic beaker with slightly tapered sides. A series of horizontal slits are cut in the sides, near the bottom, to allow the circulation of test solutions within the beaker. A circular hole is drilled in the centre of the bottom of the beaker, and a removable "pressure-fit" 5-cm long standpipe, cut from a standard-supply 10-mL disposable polystyrene volumetric pipette, is inserted through the hole (EC, 1998). The incubation unit is suspended in a test vessel, and the embryos are maintained in the test incubation units until they become actively swimming tadpoles (GS 25). At this point the tadpoles may swim out of the incubation units on their own; however, once all of the eggs have hatched (approximately 7 days after the eggs are introduced to the incubation units), the tadpoles are gently tipped out of the incubation unit into the test vessel (see Section 4.2), and the incubation units are removed. Care is taken to avoid unnecessary handling of embryos, or bumping or dropping them as they are transferred to the incubation units (see Section 2.3.7). Within the units, embryos need adequate space to ensure sufficient oxygen exchange and removal of metabolic wastes. The embryos should be distributed as evenly as possible on the bottom of each incubation unit. During this counting procedure, the incubation unit may be raised gently to just below the surface of the test solution if this is necessary for observation, but the embryos should remain in the test solution at all times. For embryos, daily observations of the number hatched and the Gosner stage are made and recorded.

interest as an endpoint (e.g., for endocrine disruption studies) or where there is interest in comparing endpoints in organisms where developmental stage is constant.[99] These tests can be quite long, as rates of development vary greatly, and the unpredictability of test duration using this approach can be challenging for laboratories from the perspective of planning and resource management.

Exposure durations might be lengthened (or shortened), depending on suborganismal endpoints that are of interest, ideally matching the stage of organisms used at test start and the test duration to the effect or effects being investigated. For example, a study focused on effects on organogenesis would focus on young embryos (i.e., GS 1-25), but for studying effects on the thyroid or metamorphosis, the test would need to include the stage where the thyroid is becoming active and responsive (i.e., GS 29) (D. Fort, Fort Environmental Laboratories Inc., Stillwater, OK, personal communication, 2020). An investigator would also want to ensure that specific organs of interest for histology or gene expression endpoints are well developed in the test tadpoles before ending the exposure. For example, sex is difficult to determine in tadpoles that are < GS 36 (Robinson et al., 2020), so exposure durations would need to be adjusted to ensure that tadpoles had well-developed gonads and testes if sexing the tadpoles or gonadal histology was of interest.

4.6.3.2 Suborganismal or non-apical endpoints

It is becoming more common to see non-traditional endpoints being used as alternatives to measuring the effects of toxicants on test organisms or to identify specific mode-of-action effects. Section 1.4 summarizes examples of amphibian tests that use alternative endpoints (e.g., behaviour, physiology, gene expression, endocrine disruption) to assess impact or as a tool to screen for or predict impact.

As there is no requirement for dry weight measurements at the end of the tests described herein (see Sections 4.5 and 4.6.1), tissue samples can be collected at the end of the 14- or 42-day tests (or during the test using additional test organisms) for histology analysis and/or analysis of gene expression, depending on study objectives. Care must be taken and study objectives clearly understood when processing tadpoles at the end of the exposure as different treatments and preservation techniques are likely necessary and endpoint dependent.[100, 101]

[99] Stage-based endpoints can be added on to the standard 42-day test. For example, if a laboratory wishes to continue exposures until 80% of control organisms reach GS 42, they can do an "interim" test end at 42 days by carrying out all of the standard endpoint measurements on live tadpoles and assessing test validity described in this test method document. Any tadpoles that have not reached GS 42 could then be put back into the exposure solutions until 80% of their control tadpoles reach GS 42.

[100] In a study conducted by Holloway *et al.* (2004), the effects of two anaesthetic or euthanizing compounds, MS-222 (tricaine methanesulfonate) and clove oil (eugenol), on blood hormone profiles were compared using rainbow trout (*Oncorhynchus mykiss*). The study found significant differences in some blood hormones depending on the compound used, which highlights the importance of investigating the potential effects of euthanizing compounds on blood plasma parameters if blood and tissue analyses are to be incorporated as study endpoints.

[101] The use of preservation techniques allows test organisms to be preserved and alternative endpoints to be assessed at a later time as an add-on to the standardized test method. Different test endpoints measured on different tissues or whole organisms would likely require different preservation techniques, depending on study objectives. For thyroid histology, tadpoles have been preserved in Davidson's fixative, rinsed, and transferred to 10% neutral buffered formalin according to the procedures outlined in OECD (2007) "Guidance Document on Amphibian Thyroid Histology Part 1: Technical Guidance for Morphologic Sampling and Histological Preparation" (Experimental Pathology Laboratories, Inc., 2019). For gonadal histology, tadpoles have been slit from the anus up the abdomen using fine-tipped sharp scissors and placed in CALEX® for 24 hours, and then transferred to alcohol (ALET, 2009). For steroid or hormone measurement, tadpoles have been flash frozen in liquid nitrogen and stored at -84 °C (ALET, 2009). For genomic analysis, tadpoles have been dissected to remove the organ of interest. Organs were preserved in RNAlater®, stored at 4 °C for 24 hours, and then transferred to -84 °C for longer storage. All tools were washed in a solution of 1% diethyl pyrocarbonate (DEPC), then 5% hydrogen peroxide in DEPC, and again in DEPC (ALET, 2009). Metamorphs have been placed in formalin or CALEX® for 24 hours and then transferred to alcohol for preservation of tissues for analysis at a later

These types of tests can provide valuable insight into the mode-of-action behind any observed apical endpoints, or provide indication of suborganismal effects where overt survival, growth, or developmental effects were not observed. For example, changes in metamorphosis are not always directly related to thyroid disruption (D. Fort, Fort Environmental Laboratories Inc., Stillwater, OK, personal communication, 2020). Pairing accelerated or delayed development with thyroid histopathology, RNA activity, or another endpoint better designed for determining mode-of-action, can provide a much clearer picture.

Several ECCC studies have been carried out where the test design for exposure described in this test method document was paired with histology endpoints of interest including heart blood smears collected for leukocyte profiles, kidneys collected for parasite assessment, livers collected for determination of liver somatic index and oxidative stress, and gonads collected for sex ratio and gonadal histology (Robinson *et al.*, 2019a, 2020, 2021). Some of these studies showed that signs of physiological stress observed at the suborganismal level did not necessarily result in overt, whole-organism effects (e.g., survival, growth, development, morphology).

In some cases, observed effects using traditional apical endpoints might not result in expected suborganismal-level changes as well. In a recent study, ECCC retained Experimental Pathology Laboratories to perform a histopathologic evaluation of the thyroid glands collected from unsexed *L. pipiens* tadpoles that had been exposed to 0.67 µg/L thyroxine for two weeks in a static-renewal exposure.[102] Based on Gosner staging (apical endpoint), thyroxine exposure was clearly associated with accelerated metamorphic development as expected; however, there was no visible evidence of thyroid glandular atrophy, as expected. As thyroxine is known to enhance thyroid activity and accelerate metamorphosis, these results were likely due to the young age of the exposed tadpoles (Experimental Pathology Laboratories Inc., 2019).

Whole organisms or tissue samples collected at the end of the test (or during the test using additional test organisms) can also be used for analysis of gene expression. In several ECCC investigations, North American bullfrog (*L. catesbeiana*) and Pacific tree frog (*Pseudacris regilla*) tadpoles were exposed to a variety of chemicals, after which organs (e.g., brain and liver) and tissues, (e.g., olfactory and tail) were collected for gene transcript analysis (i.e., quantitative reverse-transcription polymerase chain reaction [RT-qPCR]) (see Section 1.4). Morphometric and behavioural endpoints were also collected following tadpole exposures in order to link whole organism responses with gene expression endpoints (Veldhoen *et al.*, 2006, 2014; Gunderson *et al.*, 2011; Heerema *et al.*, 2018). Biopsies can be collected from live tadpoles to allow the direct tissue assessment of contaminants of concern (e.g., in vitro cultured tail fin assays using *L. catesbeiana* tadpoles for the C-fin assay: Hinther *et al.*, 2010b; Veldhoen *et al.*, 2014). In addition, whole organisms or tissues can also be collected and contaminants of concern measured in body-burden analyses (Marlatt *et al.*, 2013; Robinson *et al.*, 2020).

Tissue samples can also be collected for EcoToxChips, a next-generation toxicogenomics tool that can be used to identify specific molecular

---

date (ALET, 2004, 2006). For EcoToxChip analysis, the recommended preservation techniques are: i) preserve whole tadpoles by flash freezing in liquid nitrogen and storing at -80 °C, or ii) for larger test organisms, perform dissection and preserve tissues of interest in RNAlater® at -80 °C, -20 °C, or room temperature (D. Crump, Environment and Climate Change Canada, Ottawa, ON and N. Hogan, University of Saskatchewan, personal communication, 2020).

[102] The thyroid glands were evaluated in whole body sections. The tadpoles were preserved in Davidson's fixative, rinsed, and transferred to 10% neutral buffered formalin, following the procedures outlined in the OECD (2007) publication "Guidance Document on Amphibian Thyroid Histology Part 1: Technical Guidance for Morphologic and Histological Preparation". Tadpoles were embedded in paraffin, and 4–6-micron thick sections were obtained in the horizontal longitudinal plane at 4–10-micron intervals, depending on the size of the tadpole. Sections were mounted on glass slides with glass cover strips and stained with hematoxylin and eosin (H&E), according to routine methods used for *Xenopus laevis* (OECD, 2007; Grim *et al.*, 2009). Sections were evaluated by a pathologist using bright-field microscopy (Experimental Pathology Laboratories, Inc., 2019).

mechanisms of action for the assessment of chemicals of concern and environmental management (Basu *et al.*, 2019; Crump *et al.*, 2023). The EcoToxChip is a RT-qPCR microplate containing numerous gene targets and quality control measures. It is designed with an emphasis on analysing liver tissue or whole organisms, depending on organism size. The other key component of the EcoToxChip project is the EcoToxXplorer, a data-evaluation tool for analyzing and interpreting EcoToxChip results (Soufan *et al.*, 2022). Species-specific EcoToxChips have been developed for both laboratory model species and native species including *L. pipiens*, and methods to calculate the transcriptomic point of departure from EcoToxChip data are under development. In a recent study, GS 24/25 *L. pipiens* were exposed for 96 hours to six concentrations of thyroxine, triiodothyronine, and perchlorate, and whole larvae tissues were preserved (see footnote 101) for future EcoToxChip analysis (N. Hogan, University of Saskatchewan, personal communication, 2022, 2023). This technique was designed using early life stage exposures for use as a high-throughput screening tool and has not been validated for use at the end of the 14-day or 42-day test options described in this ECCC method. The EcoToxChip website (www.ecotoxchip.ca) can be consulted for more information.

Investigators can contact Environment and Climate Change Canada's Method Development and Application Unit for further guidance on other test designs and endpoints.

4.7 Test Validity

For the results of the 14-day test described in this biological test method document to be considered valid, each of the following two criteria must be achieved: [103, 104]

i) the mean survival for tadpoles held in control water for 14 days must be ≥80% at the end of the test, and
ii) the median Gosner stage for tadpoles held in control water for 14 days must be ≥27 at the end of the test.

For the results of the 42-day test described in this biological test method document to be considered valid, each of the following two criteria must be achieved: [96, 97]

i) the mean survival for tadpoles held in control water for 42 days must be ≥80% at the end of the test, and
ii) the median Gosner stage for tadpoles held in control water for 42 days must be ≥33 at the end of the test.

4.8 Tests with a Reference Toxicant

The routine use of a reference toxicant or toxicants is used to assess, under standardized test conditions, the relative sensitivity of a portion of the population or batch of organisms within a particular culture (Section 2.3.1) from which test organisms are selected for use in one or more definitive toxicity tests. Tests with a reference toxicant also serve to demonstrate the precision and reliability of data produced by the laboratory for that reference toxicant, under standardized test conditions, as well as technical proficiency of the laboratory staff conducting the test (EC, 1990).

Sensitivity of larvae to the reference toxicant(s) must be evaluated at the time that each 14-day or 42-day definitive test is performed, using a portion of the same batch of organisms used to start that test (EC, 1998, 2011). Batch-to-batch biological variability

[103] The test validity criteria presented here are based on control data generated in several studies carried out during the development of the method (ALET, 2016a, 2016b; Nautilus Environmental, 2016), values compiled from the literature (S. Melvin, Queensland University, Australia, personal communication, 2013; S. Robinson, Environment and Climate Change Canada, Ottawa, ON, personal communication, 2021), and on three rounds of an inter-laboratory study performed with this method (mean survival 80–100%; Nautilus Environmental 2020a, 2020b).

[104] Other standardized amphibian methods include considerations for scenarios where validity criteria are not met in a test (ASTM, 2022a; USEPA, 2015). In these scenarios, it may still be possible to gain valuable information from a test. If a test performed according to the method described herein does not meet validity criteria, MDAU may be contacted for further guidance and interpretation of results.

can be high, even from the same breeding event; therefore a portion of each batch of test organisms (i.e., egg mass or masses) used in definitive testing must be assessed with a reference toxicity test.[105] The reference toxicity test must be performed under the same experimental conditions as those used with the test sample(s), except when otherwise indicated in Sections 4.8.1 to 4.8.3. The initial developmental stage used for the reference toxicity test must be matched to that used for the corresponding definitive test (i.e., Gosner stage 25 for the 14-day definitive test and Gosner stage 28/29 for the 42-day definitive test). Reference toxicity and definitive tests can be set up on different days if a sufficient number of test organisms from the same batch are not available for use on the same day.[106] Testing with a reference toxicant, conducted according to the procedures and conditions described herein, must be performed according to one of the following two regimes for the 14-day definitive test:

i) a 14-day multi-concentration sublethal reference toxicity test using organisms taken from the batch of *L. pipiens* that is being cultured for use in the definitive test(s) (Section 2.3.1);[107] or a *static* 96-hour multi-concentration acute lethality test (see Section 4.8.3) using organisms taken from the batch of *L. pipiens* that is being cultured for use in the definitive test(s) (Section 2.3.1) assuming that sufficient justification is provided for this short test option[108]; or

ii) a positive control concentration run concurrently with each test;

and must be performed according to one of the following two regimes for the 42-day definitive test:

i) a 42-day multi-concentration sublethal reference toxicity test using organisms taken from the batch of *L. pipiens* that is being cultured for use in the definitive test(s) (Section 2.3.1); or

ii) a positive control concentration run concurrently with each test.

If a laboratory is performing both the 14-day and 42-day definitive tests using the same batch of organisms, then it is acceptable to only perform one of the reference toxicity test options available for the 42-day test, instead of performing one option for each definitive test type. Alternatively, provided there is adequate justification, a 96-hour multi-concentration test alone may be acceptable as a reference toxicant test (see footnote 108).

A series of test concentrations should be chosen to enable the calculation of the appropriate endpoints (i.e., growth and/or mortality for 14-day test, and development for 42-day test). The test concentration(s) used should be based on the data provided herein as well as the concentration-response curves resulting from preliminary tests conducted by the laboratory itself (see Section

[105] Differences in test performance were observed between different egg masses during test method development and inter-laboratory studies. These included differences in egg fertilization rate as well as tadpole development rate, size (Nautilus Environmental, 2020b), and weight (Nautilus Environmental, 2020a).

[106] In practice, there may be a period of 4 to 5 days when a sufficient number of test organisms of the required developmental stage are available from the same batch (L. Van der Vliet, Environment and Climate Change Canada, Ottawa, ON, personal communication, 2021). Alternatively, a portion of the fertilized egg mass to be used for definitive testing may be held at a lower temperature (10 to 15 °C) for up to 5 weeks to delay development and enable later testing (see Section 2.3.4 and footnote 12). For these reasons, the flexibility in timing of ± 14 days that ECCC test methods typically provide to perform a reference toxicity test before/after the definitive test is not applicable for this test method.

[107] ECCC typically includes monthly multi-concentration reference toxicity tests as the option for routine testing (EC, 1990b); however, due to the lack of availability of test organisms on a year-round basis, and in an attempt to reduce the number of tadpoles used for testing, this is not the case for this test method.

[108] Although a 96-hour lethality test is listed as an acceptable reference toxicity test option for the 14-day definitive test, it is intended for use when the laboratory does not have the capacity (i.e., resources, sufficient test organisms) to perform the other more intensive reference toxicity test options. An acute lethality test is not the preferred reference toxicity test option as it cannot confirm if adverse effects on the sublethal endpoint of interest (i.e., growth) are within the expected range.

3.3.1). These preliminary multi-concentration definitive tests are carried out using a reference toxicant with the laboratory's own test system and test organisms to demonstrate their proficiency in carrying out the test, to determine that the laboratory's own test system is adequate for the methods described herein, and to establish the reference toxicant concentration(s) to be used in conjunction with each definitive test (see Sections 3.3 and 3.5). The procedures and conditions to be applied to these preliminary tests should be consistent with those described in Section 4 herein. Any endpoint data (i.e., mean or median responses at each test concentration) should be compared with values obtained in the past by the same laboratory and with results published herein established by Environment and Climate Change Canada for the same reference toxicants.

Appropriate criteria for selecting the reference toxicant to be used in conjunction with a definitive toxicity test using *L. pipiens* include the following (EC, 1990, 2011):

- chemical readily available in pure form;
- stable (long) shelf life of chemical;
- highly soluble in water;
- stable in aqueous solution;
- minimal hazard posed to user;
- good concentration-response curve for test organism;
- known influence of pH on toxicity of chemical to test organism;
- known influence of water hardness on toxicity of chemical to test organism; and
- concentration easily analyzed with precision.

Each test concentration should be made up according to the guidance in Sections 4.1 and 5.2. Concentrations of reference toxicant (including single concentrations used as a positive control) in all stock solutions should be measured chemically using appropriate analytical methods (APHA *et al.*, 2017) (e.g., ICP-MS for sodium, ion chromatography for chloride, or conductivity as a proxy for sodium chloride if compared to historical data; LC-MS for thyroxine). Upon preparation of test concentrations of the reference toxicant, aliquots should be taken from at least the control and the low, middle, and high concentrations or the single concentration used as the positive control.[109] Each aliquot should either be analyzed directly, or stored for future analysis (i.e., at the end of the test) if the reference toxicity test endpoint(s) or positive control response based on nominal concentrations was found to be outside the *warning limits*. If stored, sample aliquots must be held in the dark at 4 ± 2 °C. Stored aliquots requiring chemical measurement should be analyzed promptly upon completion of testing with the reference toxicant. The statistical endpoints for the multi-concentration reference toxicity test or positive control should be calculated based on the measured concentrations if they are appreciably (i.e., ≥ 20%) different from nominal ones and if the accuracy of the chemical analyses is satisfactory.

A summary of the test conditions and validity criteria for each reference toxicity test option is provided in Table 6 and described in detail in Sections 4.8.1 to 4.8.3.

Reference toxicity tests and positive controls can be used to monitor consistency over time (i.e., similar means between tests) and precision over time (i.e., overlapping ranges between tests). Identifying outliers in test organism response or extreme variability in response for individual tests must be used to trigger investigations into potential causes such as culture sensitivity, culture health, environmental/facility conditions, and technician performance. Data obtained from negative controls, multi-concentration reference toxicity tests or positive controls, and culture health data should be monitored over time (i.e., by trend analysis) to proactively identify changes in the organism response.

4.8.1 Multi-concentration Sublethal Tests with a Reference Toxicant

Procedures and conditions to be followed when performing multi-concentration sublethal reference toxicity tests in conjunction with definitive toxicity

[109] If the statistical endpoints for each reference toxicity test are to be based on measured concentrations, it is recommended that one or more aliquots of each test concentration be collected and analyzed. However, if the statistical endpoints for each test are based on nominal concentrations, then sampling and analysis of aliquots from at least the low, middle, and high test concentrations are recommended.

Table 6 Checklist of required and recommended conditions and procedures for conducting reference toxicity tests using aquatic life stages of frogs

Definitive test	**14-day**			**42-day**	
Reference toxicity test option	**i) Multi-concentration sublethal**	**ii) Positive control**	**iii) Multi-concentration acute lethality**	**i) Multi-concentration sublethal**	**ii) Positive control**
Organisms*	*Lithobates pipiens*, Gosner stage 25; ≥10 organisms per test vessel			*Lithobates pipiens*, Gosner stage 28/29; ≥10 organisms per test vessel	
Replicates per treatment	≥3	≥3	1	≥4 for test concentrations; ≥8 for control(s)	≥4
Concentrations	≥ 7, or ≥ 5**	1	≥ 7, or ≥ 5**	≥ 7, or ≥ 5**	1
Recommended chemical	NaCl	N/A	NaCl	N/A	T4
Exposure duration	14 days	14 days	96 hours	42 days	42 days; 14 days if T4 is used***
Endpoints	survival; growth (total length, wet weight, biomass)	survival; growth (total length, wet weight, biomass)	survival	survival; growth (total length, wet weight, biomass); development (Gosner stage)	survival; growth (total length, wet weight, biomass); development (Gosner stage)
Requirements in the negative or solvent control(s) at test end for valid test	≥ 80% mean survival; GS ≥ 27 median development	≥ 80% mean survival; GS ≥ 27 median development	≥ 90% mean survival	mean survival ≥ 80%; GS ≥ 33 median development	≥ 80% mean survival; GS ≥ median developmental stage defined by laboratory depending on duration
Requirements for positive control	N/A	target effect size is pre-defined by the laboratory; change in growth falls within acceptability limits (see Appendix H in ECCC, 2022 for example)	N/A	N/A	target effect size (e.g., increase of 4 GS) is pre-defined by the laboratory; significant difference in development from negative or solvent control
Warning chart	growth (ICp for total length, wet weight, or biomass)	growth (ICp for total length, wet weight, or biomass)	survival (LC50)	development (LOEC for median change in development [Gosner stage] compared to negative or solvent control); in addition to a warning chart, can track consistency of dose-response over time	development (median change in development [Gosner stage] compared to negative or solvent control)

The information in this table is for summary purposes only. Definitive requirements and recommendations of this test method are contained in the main body of this document.

* The same batch of test organisms used for the definitive test must be used for the reference toxicant test.

** Fewer test concentrations (≥ 5) may be used if the laboratory has existing evidence to show that a reduced concentration series can reliably capture the target effects.

*** If a reference toxicant other than T4 is used that similarly advances development, the laboratory should investigate an appropriate exposure duration for the positive control.

tests using *L. pipiens* are described in this section. These procedures also apply to tests for assessing the acceptability and suitability of batches of test organisms to be used in definitive toxicity tests. They should be applied to assess intra-laboratory precision when a laboratory is inexperienced with the biological test method defined in this document and during initial test setup (see Sections 2.3.1, 2.3.9, and 3.3.1).

Multi-concentration sublethal reference toxicity tests must be run using organisms from the same batch used for the definitive test, and may be run concurrently with a definitive toxicity test if resources and the number of organisms available at the required developmental stage allows (see Section 2.3.4 and footnote 12).

Sodium chloride (NaCl) is recommended for the multi-concentration sublethal reference toxicity test option performed in conjunction with the 14-day definitive test option described herein.[110] Thyroxine (T4) would be suitable for use in a multi-concentration format in conjunction with a 42-day test, as T4 was successfully used in the inter-laboratory study with a 14-day test duration.[111] However, at time of publication, ECCC had not fully derived criteria to evaluate results from a multi-concentration test design. Section 4.8.2 provides guidance on the use of T4 in a single-concentration test design.

The sublethal reference toxicity test must be conducted as a static-renewal or flow-through multi-concentration test, with survival and growth endpoints for a 14-day test or survival, growth, and development endpoints for a 42-day test. The test conditions and procedures described herein for performing the 14-day or 42-day definitive test options must be applied to each reference toxicity test. Additional conditions and procedures described in Section 4 for performing a multi-concentration test with water samples apply to each reference toxicity test. Procedures described in Section 5 for the preparation and testing of chemicals and chemical products also apply here.

If this option is chosen, the multi-concentration sublethal reference toxicity test must be performed using the same test vessels as those used for definitive tests (see Section 3.3.2), with the same volume of test solution. A minimum of ≥ 7 concentrations plus control(s) must be used; however, fewer test concentrations (≥ 5) may be used if the laboratory has existing evidence to show that a reduced concentration series can reliably capture the target effects. For reference toxicity tests conducted in conjunction with the 14-day test, the number of replicate test vessels per reference toxicant concentration and control must be ≥ 3. For those reference toxicity tests performed in conjunction with the 42-day test, the number of replicate vessels per reference toxicant concentration must be ≥4 and the number of replicate vessels for the control must be ≥8. The number of tadpoles per test vessel must be ≥10 as described in Section 4.2.

Procedures for starting and ending a reference toxicity test must be consistent with those described in Sections 4.2 and 4.5. Test conditions described in Section 4.3 must be applied. Test organisms must be fed as described in Section 4.3.6. Test observations and measurements described in Section 4.4 must be performed.

The validity criteria for reference toxicity tests are the same as those described for definitive tests of the same duration (see Section 4.7 and Table 6).

---

[110] The average 14-day LC50 (± SD) of NaCl for *L. pipiens* was 4.2 ± 0.7 g/L during round 1 of inter-laboratory testing. The concentrations used were 0, 0.80, 1.2, 1.8, 2.7, 4.0, and 6.0 g/L. The average 14-day IC50 (± SD) for biomass (based on wet weight) from this study was 3.5 ± 0.9 g/L, and the IC50 for length could not be calculated, but was likely close to 4.0 g/L based on observations at two of the four labs (Nautilus Environmental, 2020a).

[111] ECCC gained some experience using T4 as a reference toxicant in the inter-laboratory investigation. During round 1 of testing, 14-day exposures to T4 produced a dose-response where increasing advancements in development were observed with increasing T4 concentration at all participating laboratories. The concentrations used were 0, 0.074, 0.22, 0.67, 2.0, and 6.0 µg/L. Significant effects on development compared to the control were observed at concentrations ≥0.67 µg/L in all labs, and at lower concentrations in two labs. The average 14-day IC50 (± SD) for wet weight from this study was 1.5 ± 0.1 g/L, and the IC50 for length could not be calculated (Nautilus Environmental, 2020a).

Endpoints must be calculated as described in Section 4.6.2 and should be expressed as mg or g reference chemical/L of test solution.

4.8.2 Positive Control

Environment and Climate Change Canada's Biological Assessment and Standardization Section is introducing the use of positive control replicates, included with each definitive toxicity test, as an alternative to multi-concentration reference toxicity testing. For the positive control reference toxicant, a single concentration of a known toxicant, which elicits a consistent partial response, must be used (as compared to traditional reference toxicity tests conducted using multiple concentrations to capture a range of effects, e.g., complete effect on growth or development to no effect on growth or development). This approach has a biological advantage, as the same endpoints are measured as in the definitive test, as well as an economical and ethical advantage, as the use of positive control replicates reduces the effort, resources, and number of organisms required. Positive controls are defined as an exposure of test organisms to conditions similar to a negative control (i.e., same number of replicates, number of organisms per replicate, vessels, test conditions, etc.), except they are exposed to a single concentration of a known toxicant (see Section 3.5). This option could be more feasible and practical for longer term sublethal- and life cycle-type toxicity tests, such as the 42-day definitive test with *L. pipiens* described in this test method document.

For the 14-day test option, ECCC could not recommend a single-concentration reference toxicant at the time of publication of this method;[112] however, a laboratory may choose to investigate an appropriate reference toxicant and test concentration for use as a positive control. Appendix H in ECCC (2022) describes a worked example of choosing a single concentration of a reference toxicant for assessing quantitative endpoints, and can be referred to for guidance.

Thyroxine (T4) is recommended herein for use as a single-concentration positive control in conjunction with the 42-day definitive test. T4 exposure results in accelerated development, with tadpoles responding consistently over time after only 14 days of exposure.[113] If T4 is chosen as the reference toxicant to be run concurrently with the 42-day definitive test, the positive control and *solvent control* must only be run for 14 days because a longer duration would likely lead to most tadpoles reaching Gosner stage 42 prior to the end of the definitive 42-day test. If another reference toxicant that similarly advances development is chosen, the laboratory should investigate an appropriate exposure duration. Procedures described in Section 5 for the preparation and testing of chemicals and chemical products also apply here.

The positive control replicates must be prepared using the same test vessels as those used for definitive tests (Section 3.3.2), with the same volume of test solution (i.e., ≥7 L). The number of replicate test vessels per positive control concentration must be ≥3 when used in conjunction with the 14-day definitive test; the number of replicate test vessels per positive control concentration must be ≥4 when used in conjunction with the 42-day definitive test. The number of tadpoles per test vessel must be ≥10 as described in

[112] Sodium chloride (NaCl) proved to be a very effective reference toxicant for use in multi-concentration reference toxicity tests; however, no single concentration gave a reproducible partial response for growth (Nautilus Environmental, 2020a). As well, the effects of a single concentration of thyroxine (T4) on growth were not reproducible in inter-laboratory testing (Nautilus 2020b; L. Van der Vliet, Environment and Climate Change Canada, Ottawa, ON, personal communication, 2021). Therefore, ECCC cannot recommend NaCl or T4 for use as a positive control in conjunction with the 14-day definitive test at this time.

[113] In 3 rounds of ECCC's inter-laboratory investigation, tadpoles in one laboratory responded consistently to a 14-day exposure to 0.67 µg/L T4 with accelerated development of 4.6 to 5.0 Gosner stages relative to the solvent controls (3 mg/L NaOH) (Nautilus Environmental, 2020a, 2020b). In four tests performed at three other laboratories, tadpoles responded to this exposure regime with accelerated development of 3.0 to 10.8 Gosner stages relative to the dilution-water controls (Nautilus Environmental, 2020a). These data suggest that a single concentration of T4 can reliably induce an increase in tadpole development and is therefore suitable for use as a positive control in a test that is focused on endocrine disruption and development endpoints.

Section 4.2. The positive control concentration should be made up according to the guidance in Sections 4.1 and 5.2, and the procedures and conditions for testing must be consistent with those used in the definitive test, described in Sections 4.2 to 4.5.

For positive controls used in conjunction with the 14-day definitive test, the required endpoint is the mean corrected growth (i.e., measured as individual total length, wet weight, or biomass) in the positive control concentration subtracted from the mean in the negative control, divided by the mean negative control response and multiplied by 100 to provide a percent change. For the positive controls used in conjunction with the 42-day definitive test, the required endpoint is the median developmental stage (i.e., Gosner stage) for the positive control concentration subtracted from the median for the negative control, to provide the change in developmental stage (in the case of thyroxine, it is accelerated development).

If selecting this option, the positive control response (i.e., target effect size) must be defined and include acceptability limits for each endpoint(s). Acceptability limits for the purposes of this method are synonymous with *warning limits* and must be operationally defined at each laboratory with variability limits that are fit for purpose. ECCC test methods usually provide guidance on how to derive warning limits for positive controls; however, the data for developmental stage is not normally distributed, so the typical measures (e.g., coefficient of variation, standard deviation) are not appropriate here. Laboratories might expect positive control results for T4 to fall within the ranges observed during inter-laboratory testing (see footnote 111 in Section 4.8.1). MDAU can be contacted for further advice on deriving acceptability limits for developmental stage data.

4.8.3 Multi-concentration Acute Lethality Tests with a Reference Toxicant

Under certain conditions, another alternative is provided for testing with a reference toxicant for the 14-day definitive test (see footnote 108). For this option, a 96-hour multi-concentration acute lethality reference toxicity test must be performed, using survival as the endpoint. Multi-concentration acute lethality reference toxicity tests must be run using organisms from the same batch as those used for the definitive test, and may be run concurrently with a definitive toxicity test if resources and the number of organisms available at the required developmental stage allows. Sodium chloride (NaCl) is recommended as the reference toxicant for this 96-hour test.[114]

If this option is chosen, the test conditions and procedures described herein must be applied. These procedures are based on standard operating procedures developed by ALET (P. Jackman, Environment and Climate Change Canada, Moncton, NB, personal communication, 2021). The multi-concentration reference toxicity test must be performed using the same test vessels as those used for definitive tests (see Section 3.3.2), with the same volume of test solution. The tests are static with no renewal. A minimum of ≥ 7 concentrations plus control(s) must be used; however, fewer test concentrations (≥ 5) may be used if the laboratory has existing evidence to show that a reduced concentration series can reliably capture the target effect. Procedures described in Section 5 for the preparation and testing of chemicals, chemical substances, and chemical mixtures also apply here. A minimum of one replicate per treatment, including the control(s), must be included in the test. The number of tadpoles per test vessel must be ≥10 as described in Section 4.2. *Lithobates pipiens* tadpoles that are at Gosner stage 25 and from the same batch as those used for the definitive 14-day test, must be used to start the reference toxicant test. Test

---

[114] ALET (2006) has successfully used test concentrations of 0, 1000, 1800, 3200, 5600 and 10 000 mg/L NaCl in 96-h acute lethality reference toxicity tests with larval *L. pipiens*. The mean 96-hour LC50 of NaCl for *L. pipiens* at ALET using this concentration series was 5120 mg/L (*n* = 12 tests; P. Jackman, Environment and Climate Change Canada, Moncton, NB, personal communication, 2022). During round 1 of inter-laboratory testing, two labs observed no or almost no mortality at the highest concentration of NaCl tested (6.0 g/L) after 96 hours, but two labs observed mortality consistent with results from ALET (i.e., mean mortality of 13% to 27% at 4.0 g/L NaCl, and 87% to 100% at 6.0 g/L NaCl [Nautilus Environmental, 2020a]), with 96-h LC50s of 4.3 and 4.9 g/L NaCl.

organisms are fed at the start of the test (e.g., 0.2 g kale and 0.05 g dried tadpole food per test vessel, assuming 10 tadpoles per replicate, provided at t = 0; see Table 5 in Section 4.3.6), but are not fed for the remainder of the test. The temperature and lighting requirements described in Section 4.3.3 must be followed. Test observations and measurements described in Section 4.4 must be performed. Procedures for ending a reference toxicity test must be consistent with those described in Section 4.5. The test must be ended after 96 hours, and the number of surviving organisms in each treatment must be recorded. The 96-hour LC50 must be calculated (see Section 4.6.2.1). Growth endpoints such as mean corrected total length, mean corrected wet weight, and mean corrected biomass in each treatment at test end may also be measured and recorded (see Section 4.6.1). Results for a reference toxicity test should be expressed as mg or g reference chemical/L of test solution. The test is valid if survival is ≥ 90% in the control(s).

4.8.4 Warning Charts

For both multi-concentration reference toxicity tests and positive controls, once sufficient data (e.g., minimum of five data points) are available (EC, 1990, 2005), all comparable endpoints (i.e., LC50s, ICps, or LOEC for median change in development [Gosner stage] compared to control, for a particular reference toxicant derived from multi-concentration reference toxicity tests; percent change in growth endpoints or change in developmental stage relative to control for a single concentration of reference toxicant tested as positive controls) must be plotted successively on a *warning chart*. For multi-concentration reference toxicity tests, the warning chart should plot logarithm of concentration on the vertical axis against date of the test or test number on the horizontal axis. For positive control concentrations, the warning chart should plot the change in response (percent change in growth, or median change in developmental stage) compared to the dilution-water or solvent control on the vertical axis against the date of the test or test number on the horizontal axis (ECCC, 2020, 2022). Each new data point for the reference toxicant should be examined to determine whether it falls within the warning limits (± 2 SD of values obtained in previous comparable tests using the same reference toxicant and test procedure) (ECCC, 2020, 2022). A separate warning chart must be prepared and updated for each dissimilar procedure (e.g., differing test duration, differing reference toxicant) and endpoint. For reference toxicity tests or positive controls run in conjunction with the 14-day test, a warning chart must be prepared using at least one of the growth endpoints (total length, wet weight, or biomass) corrected for initial measurements. For reference toxicity tests or positive controls run in conjunction with the 42-day test, a warning chart must be prepared using the development endpoint (Gosner stage). Each new data point for the reference toxicant should be compared with established limits of the chart; the reference toxicant result is acceptable if it falls within the warning limits. Typical procedures used to develop warning limits for other test designs (e.g., ± 2 SD) may not be appropriate for this situation. Reasonable warning limits may be derived based on expert judgement.

For multi-concentration reference toxicity tests, the logarithm of concentration must be used in all calculations of mean and standard deviation, and in all plotting procedures. This represents continued adherence to the assumption by which each LC50 or ICp was estimated based on the logarithms of concentrations. The warning chart can be constructed by plotting the mean and ± 2 SD as the logarithms, or by converting them to arithmetic values and plotting them on a logarithmic scale of concentration. Different approaches to creating a warning chart (e.g., Levey-Jennings, moving average) are acceptable. For the 42-day multi-concentration reference toxicity test, in addition to the warning chart, the consistency of the dose-response relationship for the development endpoint can be monitored over time. For positive control concentrations, the warning chart can be constructed by plotting the mean and ± 2 SD for percent change in growth. MDAU can be contacted for advice on constructing warning charts for the development endpoint, as the distribution of developmental stage data is not normal, so the typical measures (e.g., SD) are not appropriate.

The mean of the available endpoint values, together with the upper and lower warning limits (± 2 SD), should be recalculated with each successive endpoint for the reference toxicant until the statistics stabilize (ECCC, 2020, 2022). Warning charts can be used to detect trends over time. Examples of trends that might be observed include an increasing or

decreasing trend, several successive points on one side of the mean, changes that are observed at different times of the year, and successive data points outside the ± 2 SD warning limits. If a particular data point fell outside the warning limits, the sensitivity of the test organisms and the performance and precision of the test are suspect. Since this might occur 5% of the time due to chance alone, an outlying data point would not necessarily indicate abnormal sensitivity of the test organisms, nor unsatisfactory precision of toxicity data. Rather, it provides a warning that this might be the case. A thorough check of all culture and test conditions and procedures, as well as technical proficiency, is required at this time. Depending on the findings, it might be necessary to repeat the reference toxicity test or positive control concentration or obtain a new batch of test organisms before undertaking further toxicity tests.

Results that fall within the warning limits do not necessarily indicate that a laboratory is generating consistent results. A laboratory that produced extremely variable historic data for a reference toxicant would have wide warning limits; a new datum-point could be within the warning limits but still represent an undesirable variation in results obtained in the test. A *coefficient of variation* (CV) of no more than 30%, and preferably 20% or less, has been suggested as a reasonable limit by Environment Canada (1990, 2005) for the mean of the available endpoint values. For this biological test method, the CV for mean historic data derived for multi-concentration reference toxicity tests performed using sodium chloride should not exceed 20% for survival and 30% for growth endpoints.

If an LC50, ICp, or positive control result fell outside the control limits (mean ± 3 SD), it would be highly probable that the test was unacceptable and should be repeated, with all aspects of the test being carefully scrutinized. If endpoints fell between the control and warning limits more than 5% of the time, a deterioration in precision would be indicated, and again the most recent test should be repeated with careful scrutiny of procedures, conditions, and calculations.

4.9 Animal Conservation and Welfare Considerations

The 3Rs framework was developed by Russell & Burch (1959) and is applicable to all vertebrate animal testing. It describes "replacement", to avoid or replace the use of animals in toxicity testing; "reduction", to minimize the number of animals used per toxicity test; and "refinement", to use animal handling and testing procedures that minimize pain and distress. ECCC recognizes and agrees with the push to avoid the use of vertebrates in toxicity testing, and has therefore incorporated the 3Rs into this test method during its development (see Sections 1.1 and 1.3).

To incorporate "replacement", the option for a definitive acute lethality test (96-h LC50) for *L. pipiens* was not included, as fish acute lethality data can be used as a surrogate (see Section 1.3.1; Martinko and Van der Vliet, 2021; ECCC, 2023), and this endpoint is unlikely to provide sufficient information for risk assessment (FCSAP, 2019). In addition, gene expression analyses such as the EcoToxChip assay can be used for screening chemicals to estimate potential hazard or as an additional endpoint for the method described herein (see Section 4.6.3.2).

To incorporate "reduction", the number of biological endpoints that can be measured in each test was maximized. For example, wet weight is a required endpoint, but dry weight is not. Test organisms can therefore be preserved or prepared as necessary for alternative test endpoints such as gene expression or histology in order to gain as much relevant information as possible from a single test exposure (see Section 4.6.3.2). In addition, the option to replace multi-concentration reference toxicant testing with a single concentration positive control treatment (see Section 4.8.2) is provided, which can reduce the number of animals used from ≥210 to ≥30. Another consideration for reduction was using power analysis to optimize the test design for the 42-day test option with a focus on development endpoints (see Section 4.6.2.4). This optimization defined the number of organisms necessary to detect developmental effects with sufficient power in statistical analysis. Using this approach avoids wasting organisms (i.e., using more than are

necessary, or fewer than are necessary to detect an effect). Lastly, method users are strongly encouraged to run a range-finding test prior to definitive testing. The range-finding test uses fewer replicates to not only help define more closely the concentrations to be used in a definitive test, but to potentially reduce the number of concentrations needed to capture the required endpoints (see Section 4.1). This ultimately results in the use of fewer test organisms in a definitive test, and reduces the risk of needing to repeat the testing.

To incorporate "refinement", the test method was designed to use the more sensitive early life stages of amphibians (i.e., tadpoles instead of embryos; see Section 1.3.2). Guidance for obtaining healthy test organisms from hormone-induced breeding in the laboratory is provided, to minimize impact on natural populations (see Section 2.4.3 and Appendix F). Guidance for organism handling and for the quarantine, diagnosis, and treatment of adult breeder frogs showing signs of disease is also provided, to avoid unnecessary euthanization (Appendix E). For laboratories that have existing amphibian cultures, quarantine for new batches of test organisms is recommended to prevent disease transfer between organisms (see Sections 2.3.9, 2.4.4, and Appendix E; CCAC, 2021). Filtration of natural waters used in testing (contaminated or clean) is encouraged to eliminate potential pathogens (see Section 3.4). Laboratories are also encouraged to seek advice from local animal care committee(s) and/or veterinarians as necessary to maintain the health of the test organisms. Users are encouraged to refer to the recently published guidelines from CCAC (2021), which describe animal welfare considerations for amphibians and recommend the following resources for more information: Burghardt (2013), Kuppert (2013), Michaels *et al.* (2014), and Morgan and Tromborg (2007). Additional references are provided in Appendix E of this method.

Section 5

Specific Procedures for Testing Chemicals

This section gives specific instructions for testing individual *chemicals*, chemical substances (e.g., formulated products), or chemical mixtures (i.e., water samples amended with a test substance), in addition to the procedures described in Section 4.

5.1 Sample Properties, Labelling, and Storage

Information should be obtained about the properties of the chemical, formulated product, or chemical mixture to be tested, including the concentration of major ingredients (active ingredients), water solubility, vapour pressure, chemical stability, dissociation constants, adsorption coefficients, toxicity to humans and aquatic organisms, and biodegradability. Data sheets on safety aspects of the test substance(s) (e.g., Safety Data Sheets) should be consulted, if available. Where aqueous solubility is in doubt or problematic, acceptable procedures used previously for preparing aqueous solutions of the chemical should be obtained and reported and/or chemical solubility in control/dilution water should be determined experimentally. Other available information, such as structural formulae, degree of purity, nature and percentage of significant impurities, presence and amounts of additives, and n-octanol:water partition coefficient, should be obtained and recorded.[115] An acceptable analytical method for measuring the chemical in water at concentrations intended for the test should also be known, together with data indicating the precision and accuracy of the analysis.

Chemical containers must be sealed and coded or labelled upon receipt. Required information (i.e., chemical name, supplier, date received) must be indicated on the label and/or recorded on a separate data sheet dedicated to the sample, as appropriate. Storage conditions (e.g., temperature, protection from light) are frequently dictated by the nature of the chemical. Standard operating procedures for chemical handling and storage should be followed.

5.2 Preparing Test Solutions

An estimate of the lowest concentration of test substance or substances that causes sublethal effects in larval northern leopard frogs is useful in predicting chemical concentrations appropriate for the chronic toxicity tests. A 14-day range-finding test with reduced replicates and treatments is recommended for this purpose (see Section 4.1). Alternatively, the results of a 96-h static LC50 test (see Section 4.8.3), conducted at 23 ± 2 °C using the control/dilution water intended for the 14- or 42-day test, might provide this information (see Section 4.1 and footnote 54 therein).

Test solutions of the chemical to be tested are usually prepared by adding aliquots of a stock solution made up in control/dilution water. Alternatively, for strong solutions or large volumes, weighed (using an appropriate balance) quantities of the chemical can be added to the control/dilution water to give the nominal strengths for testing. If stock solutions are used, the concentration and stability of the test chemical in the solution should be determined before the test. Stock solutions subject to photolysis should be shielded from light. Unstable stock solutions must be newly prepared as necessary to maintain consistent concentrations for

[115] Knowledge of the properties of the chemical will assist in determining any special precautions and requirements necessary while handling and testing it (e.g., testing in a well-ventilated facility, need for solvent, frequency of solution renewal). Information regarding chemical solubility and stability in freshwater will also be of use in interpreting test results. Biodegradability and stability are particularly important chemical properties to understand before beginning a definitive test with a long exposure duration, such as those described in this method. In round 2 of the inter-laboratory study, triclosan rapidly degraded during definitive tests such that the measured concentrations were reduced to ≤5% of nominal within 48 hours at both participating laboratories (Nautilus Environmental, 2020a). Range-finding tests are highly recommended to verify the stability of a chemical before beginning the definitive test.

each renewal of test solutions. Stock solutions should be prepared by dissolving the chemical in control/dilution water. For chemicals that do not dissolve readily in water, guidance provided in the OECD's document on the aquatic toxicity testing of difficult test chemicals (OECD, 2019) should be followed. *Emulsifiers* or *dispersants* should not be used to increase chemical solubility except in instances in which these substances might be formulated with the test chemical for its normal commercial purposes. The use of a solvent other than water should be avoided if possible. An organic or other inorganic solvent may be used for the dissolution of the test substance in dilution water where no other acceptable method of test solution preparation is available. If used, an additional control solution must be prepared containing the control/dilution water and the same concentration of solubilizing agent as that present in the most concentrated solution of the test chemical (i.e., solvent control). Such agents should be used sparingly (i.e., using the minimum volume necessary to dissolve or suspend the test substance in dilution water) and should not exceed the concentration that affects the survival, growth, or development of *L. pipiens* or a maximum of 0.1 mL/L or 100 mg/L in any test solution; generally solvent concentrations ≤20 µL/L are recommended (Hutchinson *et al*., 2006; Green and Wheeler, 2013; OECD, 2019; Young *et al.*, 2020). In the ECCC inter-laboratory studies, 3 mg/L sodium hydroxide (NaOH) solvent controls were successfully used in conjunction with thyroxine (round 1) and triclosan (round 2) tests; the solvent controls performed well, with survival ≥80% and similar increases in development and growth compared to the negative controls, indicating minimal solvent effects (Nautilus Environmental, 2020a, 2020b). A recent study investigated the acute or chronic sublethal effects of three solvents on *L. pipiens* larvae at concentrations of 10–100 µL/L (Young *et al.*, 2020). The study suggests that dimethyl sulfoxide (DMSO) and ethanol should be used at concentrations ≤ 20 µL/L for amphibians, and that acetone should be used at concentrations ≤ 50 µL/L for *L. pipiens* but at concentrations ≤ 10 µL/L for other amphibians based on the literature (Young *et al*., 2020). Another study assessed the acute lethality effects of several solvents to *Rana temporaria* ([European] common frog) larvae, and suggested that acetone, ethanol, and methylene chloride should be used at concentrations ≤10 µL/L for these organisms, and that methanol should not be used due to mortalities occurring at concentrations as low as 1 µL/L (Marquis *et al.*, 2006). If information on the effects of a solvent on the test organism is unknown, then a preliminary solvent-only test, using various concentrations of the solvent, should be conducted to determine the threshold-effect concentration of the particular solvent being considered for use in the definitive test. If solvents are used, the following are preferred based on the above information: DMSO, acetone, NaOH, and ethanol.[116] For the solvent control, a minimum of 3 replicates (recommend 4) are required for each 14-day test and a minimum of 8 replicates are required for each 42-day test (see Section 4.2).

Upon preparation of test solutions including the control(s), the dissolved oxygen content should be measured. Thereafter, either tadpoles should be introduced and the test initiated (see Section 4.4), or each test solution should be pre-aerated and then the tadpoles added. In most instances, the pre-aeration of test solutions is not necessary nor warranted (see footnote 60 in Section 4.3.4). For those situations in which pre-aeration is appropriate (i.e., if, upon preparation, the DO content of one or more test solutions is $< 60\%$ or $> 100\%$ of air saturation), the guidance for pre-aeration of solutions given in Section 4.3.4 should be followed.

5.3 Control/Dilution Water

Control/dilution water may be one of the following: "uncontaminated" groundwater or surface water from a stream, river, or lake; *reconstituted water* of a desired pH and hardness (e.g., simulating that of the receiving water); a particular sample of receiving

[116] Solvents recommended by OECD (2019) for use in aquatic toxicity testing are acetone, ethanol, methanol, tertiary-butyl alcohol, acetonitrile, dimethyl formamide, dimethyl sulfoxide, and triethylene glycol. Amphibian data is lacking for some of these solvents, therefore the recommendations provided herein are more selective. OECD (2019) also notes that acetone, ethanol, and methanol can lead to substantial bacterial growth and subsequent oxygen depletion in static-renewal aquatic test systems, and that the use of flow-through systems should be considered when using these solvents.

water if there is special interest in a local situation; or dechlorinated municipal water (see Sections 2.3.5 and 3.4). The water supply used as control/dilution water is frequently the same as that used for holding/culturing the tadpoles/frogs (see Section 2.3.5), although it may come from another source. The choice of control/dilution water to be used depends on the intent of the test.

If the toxic effect of a chemical on a particular receiving water is to be assessed, sample(s) of the receiving water could be taken from an area that was isolated from influences of the chemical and used as the control/dilution water. Examples of such situations would include appraisals of the toxic effects of chemical spills (real or potential) or intentional applications of a chemical (e.g., spraying of a pesticide) on a particular water body. If a sample of receiving water is to be used as control/dilution water, a separate control solution must be prepared using the culture/control/dilution water that is normally used for the *L. pipiens* 14-day or 42-day toxicity test and is able to achieve valid test results on a routine basis (see Sections 4.1 and 4.7). Difficulties and costs associated with the collection and shipment of receiving water samples for use as control/dilution water, as well as the risks surface water can pose to amphibian tests due to the presence of potential pathogens, should also be considered (see Section 3.4).

The laboratory supply of uncontaminated natural water may also be used to appraise the toxic effect of a chemical on a particular receiving environment, especially where logistical or cost constraints make the collection and use of receiving water impractical, or if there is already an interfering toxicity in the receiving water. The laboratory's normal water supply is also appropriate for use as control/dilution water in other instances (e.g., preliminary or intra-laboratory assessment of chemical toxicity).

5.4 Test Observations and Measurements

In addition to the observations on toxicity described in Section 4.4, there are other observations and measurements to be made during testing with chemicals.

During solution preparation and at each of the prescribed observation periods during the test, each test solution should be examined for evidence of chemical presence and change (e.g., odour, colour, opacity, *precipitation*, or *flocculation* of chemical). Any observations should be recorded.

It is desirable and recommended that aliquots of test solutions be analyzed to determine the concentrations of chemicals to which test organisms are exposed.[117] If chemicals are to be measured, samples should be taken from the high, medium, and low test concentrations, and the control solution(s) at the beginning and end of the test as a minimum. Additional samples can be collected for analyses at the beginning and end of a renewal period if there is concern about the stability of the chemical. These samples should be preserved, stored, and analyzed according to the best proven, validated methods with acceptable detection limits available for determining the concentration of the particular chemical in an aqueous solution.

If chemical measurements indicate that the concentrations declined by more than 20% during the renewal period or test period, the toxicity of the chemical should be re-evaluated by a test in which solutions are renewed more frequently than ≥3 times per week, or in a test using a flow-through system, depending on the study objectives. Toxicity results for any tests in which concentrations are measured should be calculated and expressed in terms of those measured concentrations, unless there is good reason

[117] Such analyses need not be undertaken in all instances due to cost, analytical limitations, or previous technical data indicating chemical stability in solution under conditions similar to those in the test. Chemical analyses are particularly advisable if (USEPA, 1985): the test solutions are aerated; the test substance is volatile, insoluble, or precipitates out of solution; the test chemical is known to sorb to the material(s) from which the test vessels are constructed; or a flow-through system is used. Some situations (e.g., testing of pesticides for registration purposes) might require the measurement of chemical concentrations in test solutions.

to believe that the chemical measurements are not accurate. In making these calculations, each test solution should be characterized by the *geometric mean* measured concentration to which the test organisms were exposed.

5.5 Test Endpoints and Calculations

The endpoints for tests performed with chemicals will usually be the LC50 or other LCp at the end of the test, the ICp for growth endpoints (wet weight, biomass, total length), and significant effects for development endpoints (Gosner stage) (see Sections 4.5 and 4.6).

If additional controls (e.g., solvent and/or other) are used, the results must be examined to determine if they independently meet the test validity criteria (Sections 4.7). The test is rendered invalid if > 20% of the control organisms exhibit mortality (see Section 4.7) in any additional control or in the untreated dilution-water control. The test is also invalid if the combined (for all replicates of the same treatment) median final developmental stage (Gosner stage) of the surviving control organisms is < GS 27 for the 14-day test and < GS 33 for the 42-day test (see Section 4.7) in any additional control or in the untreated dilution-water control. If solvents are used to prepare test solutions, only the data from the solvent control should be used for calculating any statistical endpoints involving comparisons of the findings for each set of test concentrations versus those for control solutions. Existing guidance has demonstrated that other techniques for handling data from dilution-water controls and solvent controls will result in a beneficial increase in power (Green *et al*., 2018). However, the 42-day test has already been optimized for power, and using just the solvent controls will maintain the false positive rate (Green *et al*., 2018).

Section 6

Specific Procedures for Testing Contaminated Water Samples

This section gives specific instructions for the collection, preparation, and testing of contaminated waters (e.g., impacted wetlands or receiving waters, effluents, elutriates, leachates) in addition to the procedures listed in Section 4. In particular, guidance for contaminated site risk assessors and managers is provided to align with previously published guidance on the use of amphibian testing for ecological risk assessment (FCSAP, 2010, 2019). In a recent review of ecological risk assessment methods for amphibians, Johnson *et al*. (2017) encouraged focused testing using standardized methods with laboratory-cultured native amphibian species, including exposures designed to measure important sublethal effects before and after metamorphosis.

Testing of samples of contaminated water might use either the 14-day or 42-day test options (Section 4.3.1) depending on the study objectives, targeted endpoints, and suspected contaminants of concern. Before one of these test options is adopted for periodic or frequent use in measuring the toxicity of contaminated site waters or industrial/municipal wastewaters, comparative assessment of these test options is recommended in order to identify which endpoints and duration (survival, growth, development, 14-day or 42-day) are most sensitive and relevant. Combined with practical limitations, such as resources to perform the test, the comparative assessment can assist in selecting a test option and test design. Either of the test options might be conducted as either static-renewal or continuous-flow assays, depending on the objectives, nature of the sample, volume needed, etc.

The requirements for sample volume for testing should be given serious consideration before undertaking any testing program. Large amounts of sample would be required for both static-renewal and flow-through tests, and the amount differs considerably for the two different test options.[118] Given the requirements for large volumes of sample for these tests, investigators might consider a tiered approach to an investigation. Single-concentration tests (a test that compares full-strength solution to a control; see Section 6.5.1) can be used to screen potentially contaminated water samples in order to select those for which more definitive testing (multi-concentration tests; see Section 4.6.2) might be necessary. In addition, the 14-day test may prove useful to screen samples for positive effects, to inform decisions on further toxicological testing (e.g., 42-day test), or as a range-finding test prior to a longer-term 42-day test, in cases where developmental effects are not anticipated.[119]

[118] For a 14-day test, with 7 L of test solution per replicate (i.e., for 10 test organisms), and three replicates per test concentration, the requirement for volume to set up replicates of full-strength contaminated water would be ~23 L (allowing for extra sample for spillage and physicochemical analyses). Seven concentrations plus a control in a geometric series including full strength (e.g., 100, 50, 25, 12.5, 6.25, 3.13, 1.56, 0%) would require approximately twice as much test sample as for the 100% concentration alone and thus the test would require ~46 L of contaminated water for set-up. An additional volume of ~37 L would be required for each 80% water renewal (a minimum of 3 times weekly for 2 weeks).

For a 42-day test, with 7 L of test solution per replicate (i.e., for 10 test organisms), and four replicates per test concentration, the requirement for volume to set up replicates of full-strength contaminated water would be ~30 L (allowing for extra sample for spillage and physicochemical analyses). Seven concentrations plus a control in a geometric series including full strength (e.g., 100, 50, 25, 12.5, 6.25, 3.13, 1.56, 0%) would require approximately twice as much test sample as for the 100% concentration alone and thus the test would require ~60 L of contaminated water for set-up. An additional volume of ~48 L would be required for each 80% water renewal (a minimum of 3 times weekly for 6 weeks).

[119] The use of the 14-day test as a screening or range-finding test for the 42-day test should be used with caution. The two tests differ in the stage of tadpoles being used to initiate the test as well as the test duration, and therefore are designed to capture different endpoints. The 14-day test primarily assesses the survival and growth of tadpoles prior to metamorphosis. In contrast, the 42-day test covers a period of metamorphic changes and is

6.1 *Sample Collection, Labelling, Transport, and Storage*

Containers for transportation and storage of water samples must be made of non-toxic material. Collapsible polyethylene or polypropylene containers manufactured for transporting drinking water (e.g., Reliance™ plastic containers) are recommended. The volume of these containers can be reduced to fit into a cooler for transport, and the air space within kept to a minimum when portions are removed in the laboratory for the toxicity test or for chemical analyses. The containers must either be new or thoroughly cleaned, and rinsed with control/dilution water or deionized or distilled water before use. They should also be rinsed with the sample to be collected. Each sample container should be filled completely, to exclude air.

Most tests with contaminated water will be performed in a controlled laboratory facility. Due to the large volumes of sample required for use in these tests, the collection of samples will be a compromise between logistical and practical constraints (e.g., time, effort, and cost). Contaminated water samples for use in either single-concentration tests or full definitive multi-concentration tests can be collected as a single sample for use throughout the test or collected periodically on several occasions during the test period, depending on many factors, including but not limited to, the objectives of the study, the practicality and resource availability for sampling, and/or the known or anticipated stability of the samples. If a single sample is used throughout the test, it might be advantageous to divide it into several separate containers (e.g., several subsamples) upon collection or receipt at the laboratory. Each subsample would remain sealed until used for a pre-determined number of test solution renewals. This approach is offered as an option in several longer-term ECCC methods requiring solution renewals, in order to minimize the likelihood that the sample is degrading over time due physical and chemical processes such as volatilization, oxidation, and photochemical or microbial degradation (EC, 2007, 2011, 2017).

In instances where the toxicity of the contaminated water is known or anticipated to change significantly if stored for up to 6 weeks before use, fresh samples can be collected on several occasions using evenly spaced sampling intervals. Sampling intervals can be shortened (i.e., more frequent sampling) for contaminated water known or anticipated to be particularly unstable, or lengthened if stability is of a lesser concern.

Testing of contaminated water should commence as soon as possible following sample collection. Typically, ECCC recommends that testing begin within 1 day of sampling whenever possible, and requires that testing commence no later than 3 days after sampling, however, due to logistical and practical constraints inherent with the volumes of sample potentially required for this test, the recommendation herein is that sample holding times be minimized as much as possible, and should be based on the project objectives, the stability of the sample, and the contaminants of interest. Samples of sediment or other solid material might also be collected for extraction and subsequent testing. These samples should also be tested as soon as possible. Procedures given in Environment Canada (1994) for the preparation of elutriates should be followed.

Generally, a 150 L sample is adequate for a 14-day single-concentration test, with 3 weekly water renewals, associated routine chemical analysis, and any necessary adjustments or allowance for spillage/rinsing. For 14-day multi-concentration tests or for 42-day tests, sample volume requirements will be much higher (see footnote 118). Immediately after filling, each sample container must be sealed, and labelled or coded. Labelling and accompanying records made at this time must include at least a code or description that identifies sample type, source, precise location (e.g., water body, latitude, and longitude), replicate number, and date of collection; and should include the name and signature of sampler(s). Unlabelled or uncoded containers arriving at the laboratory should not be tested, nor should samples arriving in partially filled containers be routinely tested, because volatile

designed to capture potential impacts on the HPT and HPG systems (see Sections 1.1 and 2.1). The value of the 14-day test as a screening or range-finding tool for the 42-day test might therefore be questionable where thyroid-disrupting compounds are of interest, for example.

toxicants escape into the air space. However, if it is known that volatility is not a factor, such samples might be tested at the discretion of the investigator.

An effort must be made to keep all[120] samples cool (1 to 7 °C, preferably 4 ± 2 °C) throughout their period of transport. Upon collection, warm (>7 °C) samples should be cooled to 1 to 7 °C with regular ice (not dry ice) or frozen gel packs. As necessary, gel packs, regular ice, or other means of refrigeration should be included in the transport container in an attempt to maintain sample temperature within 1 to 7 °C during transit. Samples must not freeze or partially freeze during transport or storage.

The date of receipt of the sample(s) at the laboratory must be recorded, and the temperature of the sample upon receipt must be measured and recorded. Samples required for testing at that time may be adjusted immediately or overnight to the test temperature and used in the test. Any samples or remaining portion(s) of sample held for future use or possible additional testing must be stored in darkness in sealed containers, without air headspace, at 4 ± 2 °C.

6.2 Preparing Test Solutions

Each sample or subsample in a collection container must be agitated thoroughly just before pouring, to ensure the re-suspension of settleable solids and their homogeneity. Depending on the nature of the sample and the objectives of the test, homogenization of samples might or might not be required before testing. If mixing is carried out, it must be thorough. If using static renewal, only the amount of sample required at that time to initiate the test or perform solution renewal should be removed from storage and adjusted to the test temperature. The dissolved oxygen content and pH of each sample must be measured just before its use. As necessary, each test solution should be pre-aerated (see Section 4.3.4) before the test solutions are distributed to replicate test vessels.

Filtration of samples is normally not required nor recommended. Filtration could remove suspended or settleable solids that are characteristic of the sample and might otherwise contribute to part of the toxicity or modify the toxicity. However, if a contaminated water sample contains debris or indigenous organisms that might be confused with or attack the test organisms, the sample must be filtered (e.g., ≤60 µm) before use (USEPA, 1994). As well, to reduce the risk of introducing pathogens that may be present in any contaminated water, samples (e.g., wetland, pond, upstream, or receiving water) can be further filtered using a finer filter (0.45 to 5 µm pore size) (Sections 2.3.5 and 3.4). Note that the risk of *Ranavirus* infection may not be mitigated by filtering due to the small size of the viruses (~150 nm) (see Section 3.4 and Appendix E). In instances where concern exists regarding the effect of this filtration on sample toxicity, a second test could be conducted concurrently using an unfiltered portion of the sample.

6.3 Control/Dilution Water

Tests conducted with samples of contaminated water for monitoring and contaminant mitigation purposes should use, as the control/dilution water, a supply (source) of the laboratory water shown previously by the testing laboratory to routinely enable valid test results. If there is a special interest in a local situation, either a sample of the receiving water or "upstream" water or laboratory water adjusted to the pH and hardness of the collection site (i.e., reconstituted water) can be used as the control/dilution water (see Sections 2.3.5, 3.4, and 4.1, and footnote 27 in Section 2.3.5). The choice of control/dilution water depend on the intent of the test. Because results could be different for the two sources of water, the objectives of the test must be decided before a choice is made. Difficulties and costs associated with the collection and shipment of receiving-water samples for use as control/dilution water, as well as the risks surface water can pose to amphibian tests due to the presence of potential pathogens should also be considered (see Section 3.4).

The use of uncontaminated receiving or "upstream" water (e.g., wetland, pond, or upstream river water) as the control/dilution water can be desirable if site-

[120] This applies to contaminated wetland or receiving water, effluent, leachate, and elutriate, as well as sediment or solid waste samples that will undergo extraction in the laboratory.

specific information is required on the potential toxic impact of contaminated water on a particular receiving environment (see rationale in Section 5.3). Conditions for the collection, transport, and storage of such receiving-water samples should be as described in Section 6.1. Any sample of receiving or "upstream" water used as the control/dilution water for testing contaminated water samples should be filtered according to the recommendations for natural control/dilution water, described in Section 3.4. If a sample of receiving water or "upstream" water is used as control/dilution water, a separate control solution must be prepared using the laboratory control water that is normally used for performing tests with *Lithobates pipiens* tadpoles (i.e., holding water or other suitable laboratory water; see Section 4.1.). The survival, growth, and development of the tadpoles (see Section 4.6.1) in the laboratory control water must be compared to that in the sample of receiving water.

6.4 Test Observations and Measurements

Survival, growth, and development at the end of the exposure must be determined, as described in Section 4.6.

Colour, *turbidity*, odour, homogeneity (i.e., the presence of floatable or settleable solids), and the presence of indigenous organisms (i.e., other organisms that might pose a threat or compete with the test organisms), should be observed in the sample of contaminated wetland or impacted surface water, receiving water, effluent, leachate, or elutriate at the time of preparing test solutions. A record should be made of any reactions or overt changes of test solutions upon dilution with water or during the test, such as precipitation, flocculation, foaming, odour, and change in colour or turbidity.

For water samples having appreciable solids content, it is desirable to measure total suspended and settleable solids (APHA *et al*., 2017) upon receipt, as part of the overall description of the sample, and its characteristics that might influence the results of the toxicity test.

6.5 Test Endpoints and Calculations

Endpoints for tests performed with samples of contaminated water or industrial wastewater will usually be the standard ones described in Section 4.6.

Tests for assessing and managing the risk of contaminated water at sites might be multi-concentration tests or single-concentration tests. If a multi-concentration design is used and the LCp, ICps, and developmental impacts are to be calculated, guidance on number of replicates and test concentrations provided in Sections 4.1 and 4.2 for this test design should normally be followed. Single-concentration tests are often more cost-effective for determining the presence or absence of measurable toxicity, or as a method for screening a large number of contaminated water samples for relative toxicity. Hypothesis testing is often the first and/or only choice for investigators in the analysis of single-concentration tests. However, the test design described herein, specifically the number of replicates, has not been optimized for all types of hypothesis testing. Investigators should consider the endpoints targeted (e.g., mortality, growth, development), the magnitude of biological effect expected, the objectives of the investigation, the statistical analyses to be used, and statistical power before beginning the experiment. These sources of information can be used to determine number of replicates. In general, for a single-concentration test, all procedures described in this test method would apply; however, the number of replicates of the test sample could vary between 4 and 8. For single-concentration tests, the number of replicates chosen must be justified *a priori*. Further guidance on the use of single concentration tests is provided in Section 6.5.1.

6.5.1 Variations in Design and Analysis

Toxicity tests might be restricted to a single concentration (e.g., undiluted test sample or a prescribed concentration of a test chemical) and a control. Single-concentration tests are often cost-effective for determining the presence or absence of measurable toxicity or as a method for screening a large number of samples for relative toxicity. Statistical analyses and endpoints for these tests would depend on the objectives of the study but

could include an arbitrary "pass" or "fail" ratings, or percentage effect relative to controls at a specified concentration. Environment Canada (2005) provides detailed statistical guidance on the analysis of quantal data (i.e., survival) in various test designs that examine multiple sampling locations. Choice of a specific statistical test depends on several considerations, including but not limited to:

- the type of comparison that is sought (e.g., complete series of pairwise comparisons between all sampling locations, or compare the response from each sampling location only with that for the control);
- if a chemical and/or biological response gradient is expected;[121] and
- the level and type (laboratory or field) of replication.

Environment Canada (2005) has also provided detailed statistical guidance on the analysis of quantitative measurements,[122] which can be readily applied to measurements of amphibian growth (i.e., length and weight of tadpoles at the end of the test) in a multiple sampling location scenario. If test results at a single test sampling location are to be compared with test results at a control site or with a control water, a *t*-test[123] is normally the appropriate statistical test (see Section 3.2 in EC, 2005). In situations where more than one test sampling location (treatment) is under study, and the investigator wishes to compare multiple sampling locations with the control, or compare sampling locations with each other, a variety of ANOVA and multiple comparison tests (and nonparametric equivalents) exist (Section 3.3 in EC, 2005). Choice of a specific test depends on the three conditions described above for quantal tests, in addition to assumptions of *normality* and *homoscedasticity* being met.

A very preliminary survey might have only one sample of test water (i.e., contaminated or potentially contaminated site water) and one sample of water from a control location, without replication. Simple inspection of the results might provide guidance for designing more extensive studies. A preliminary evaluation might conceivably be conducted with samples from many stations, but without either field replicates or laboratory (within-sample) replicates. The objective might be to identify a reduced number of sampling stations deserving of more detailed and further study. In this case, opportunities for statistical analysis would be limited (EC, 2005).

A more usual survey of a contaminated site would involve the collection of replicate samples from several places by the same procedures, and their comparison with replicate samples of water from a control location and/or laboratory control water. There are several pathways for analysis, depending on the type and quality of data. In these multi-location surveys, the type of replication would influence the interpretation of results (i.e., field replicates or laboratory replicates, or both). If both replicate samples (i.e., field replicates) and replicate vessels/units (i.e., laboratory replicates) have been tested, a statistician should be consulted for analysis options. If only laboratory replicates and no field replicates were tested, it is difficult to make statistically robust conclusions regarding differences between sampling stations (locations) within a site or between sites. The laboratory replicates would only show any differences in the samples that were greater than the baseline variability in the within-laboratory procedures for setting up and running the test. Sample variability due to location would not

---

[121] In this case, the expected gradient is determined during the experimental design phase (a priori), not after the data has been collected. Section 3.3 in EC (2005) provides guidance on cases where a gradient effect is expected. If necessary, a statistician should be consulted for further guidance on analyses of data where a gradient is expected.

[122] Sections 3.2 and 3.3 in EC (2005) provide guidance on the analysis of quantitative measurements for a single location and quantitative measurements for multi-locations, respectively, and should be consulted for the analysis of growth data. Section 7.5 in EC (2005) provides additional guidance on multiple-comparison tests for hypothesis testing, and should be consulted for additional detail; however, the calculation of NOEC/LOEC is not recommended herein.

[123] The *t*-test assumes equal variance between groups; however, modification of the *t*-test that can accommodate unequal variance is also available (EC, 2005).

really be assessed in the statistical analysis, except that it would contribute to any difference in test results associated with sampling location.

If it were desired to compare the test results for the replicate samples from each sampling location with those for samples from a control location, a number of tests are recommended, depending on whether the samples show a gradient and depending on whether there is an even or uneven number of replicates (see Section 3 in EC, 2005).

In a multi-location survey, an investigator might wish to know which of the samples from various sampling locations showed results that differed statistically from the others, as well as knowing which ones were different from the control location and/or laboratory control sample(s). Such a situation might involve sampling from a number of locations at progressively greater distances from a point source of contamination or contaminated site, in which instance the investigator might want to know which sampling locations provided samples that had significantly higher toxicity than others, and thus which locations were particularly deserving of cleanup. Sections 3.1, 3.3, and 7.5 in EC (2005) provide further details, alternate tests, and nonparametric options, and the guidance therein should be followed.

Section 7

Reporting Requirements

Each test-specific report must indicate if there has been any deviation from any of the *must* requirements delineated in Sections 2 to 6 and, if so, provide details of the deviation(s). The reader must be able to establish from the test-specific report whether the conditions and procedures preceding and during the test rendered the results valid and acceptable for the use intended.

Section 7.1 provides a list of items that must be included in each test-specific report. A list of items that must either be included in the test-specific report, provided separately in a general report, or held on file for a minimum of five years, is found in Section 7.2. Specific monitoring programs, related test protocols, or regulations might require selected test-specific items listed in Section 7.2 (e.g., details about the test material and/or explicit procedures and conditions during sample collection, handling, transport, and storage) to be included in the test-specific report, or might relegate certain test-specific information as *data to be held on file*.

Procedures and conditions common to a series of ongoing tests (e.g., routine toxicity tests for monitoring or *compliance* purposes) and consistent with specifications in this document, may be referred to by citation or by attachment of a general report that outlines standard laboratory practice.

Details on the procedures, conditions, and findings of the test, which are not conveyed by the test-specific report or general report, must be kept on file by the laboratory for a minimum of five years so that the appropriate information can be provided if an audit of the test is required (Section 7.2).

7.1 Minimum Requirements for a Test-specific Report

The following items must be included in each test-specific report.

7.1.1 Test Substance or Material

- brief description of sample type (e.g., chemical, chemical product, effluent, leachate, elutriate, receiving water) or coding, as provided to the laboratory personnel;
- information on labelling or coding for each sample or subsample;
- date and time of sample/subsample collection; date and time sample(s)/subsample(s) received at test facility;
- for contaminated or receiving waters, measurement of temperature of sample or, if multiple subsamples, one only of these subsamples, upon receipt at test facility;
- measurement of dissolved oxygen and pH of sample or subsample of contaminated or receiving waters, just before its preparation and use in toxicity test;
- for samples of elutriate or any liquid extracted from sediments or similar solids, dates for sample generation and use; description of procedure for preparation; and
- dates or days during test when individual samples or subsamples used, if applicable.

7.1.2 Test Organisms

- species and source of test organisms and breeding stock, if applicable;
- age (days since fertilized, if available) or time since collection, and developmental stage of tadpoles at start of test;
- developmental stage (GS), total body length, and wet weight (mean ± SD) of representative organisms at start of test;

- any unusual appearance, behaviour, or treatment of the organisms, before their use in the test; and

- percent cumulative mortality (must be ≤ 10%; Section 2.3.8), for any batch of tadpoles in the 5 days before start of test.

7.1.3 Test Facilities and Apparatus

- name and address of test laboratory;

- person(s) performing the test (or each component of the test) and verifying results; and

- brief description of test vessels (size, shape, type of material).

7.1.4 Control/Dilution Water

- type(s) and source(s) of water used as control and dilution water;

- measured characteristics of control or dilution water, before and/or at start of test; and

- type and quantity of any chemical(s) added to control or dilution water.

7.1.5 Test Method

- citation of biological test method used (i.e., as per this document);

- brief mention and description of test options chosen (e.g., 14-day or 42-day test; static-renewal or flow-through test);

- brief description of procedure(s) in those instances in which a sample, subsample, or test solution has been filtered, settled, and decanted, or adjusted (e.g., for hardness or pH);

- design and description if specialized procedure (e.g., renewal of test solutions other than three times weekly, or manner and rate of exchange of test solutions, if flow-through; preparation and use of elutriate; preparation and use of solvent and, if so, solvent control);

- brief description of frequency and type of all observations and all measurements made during test; and

- name and citation of program(s) and methods used for calculating statistical endpoints.

7.1.6 Test Conditions and Procedures

- design and description of any deviation(s) from or exclusion of any of the procedures and conditions specified in this document;

- number, concentration, volume, and depth of test solutions, including controls;

- number of organisms per test vessel, and number of replicates per treatment;

- brief statement (including procedure, rate, and duration) if any pre-aeration of test solutions;

- brief statement concerning aeration (including rate and duration) of test solutions during exposure of test organisms;

- dates when test was started and ended;

- frequency and rate of test solution renewal;

- food types, as well as feeding regime and ration during the test;

- all required (see Section 4.4) measurements of temperature, pH, dissolved oxygen (mg/L and percent saturation), conductivity, and ammonia in test solutions (including controls), made during the test; and

- brief statement indicating date and type of testing using a reference toxicant; whether the reference toxicity test or positive control was performed under the same experimental conditions as those used with the test sample(s); and description of any deviation(s) from or exclusion(s) of any of the procedures and conditions specified for reference toxicity testing in this document.

7.1.7 Test Results

- for each replicate test solution (including each of the control replicates): the number and percent of mortalities in each test vessel, as recorded during each observation period over the duration of the test; the number of tadpoles that reached GS 42 prior to test end and the date of their removal from the test; the median Gosner stage at the end of the test;

- for each treatment (i.e., each concentration, including the control treatment): mean (± SD) percent mortality, at the end of the test;

- for each control treatment: the combined and cumulative (over time) mean (± SD) percentage of test organisms that died at each period of observation and at the end of the test, as used for the survival criterion for test validity; the median Gosner stage at the end of the test, as used for the development criterion for test validity;

- for each treatment, including the control treatment(s): mean (± SD) corrected total length, corrected wet weight, and corrected biomass (expressed on a wet-weight basis) at the end of the test, as used for ICp calculation;

- for each treatment, including the control treatment(s): median Gosner stage at the end of the test, based on replicate medians; any observations of abnormal or asynchronous development; any unusual appearance (i.e., grossly visible malformations or lesions) or behaviour of test organisms as well as approximate developmental stage recorded during any observation period;

- any LC50 (including the associated 95% confidence limits, quantal method used, and, if calculated, the slope) determined; any additional LCp (e.g., LC25) calculated;

- any ICps (including the associated 95% confidence limits) determined for the data on corrected total length, corrected wet weight, and corrected biomass; details regarding any transformation of data, and indication of quantitative statistical method used or procedures applied to the data;

- for 42-day multi-concentration tests: results of any tests for treatment effects on development (Gosner stage), including any treatment(s) that are significantly different from the control(s); indication of statistical method used or procedures applied to the data;

- any outliers, and the justification for their removal;

- for a test with chemicals, indication as to whether results are based on nominal or measured concentrations of chemical(s) or chemical product(s); all values for measured concentrations;

- results and duration of any toxicity tests performed with the reference toxicant in conjunction with the definitive toxicity test; if applicable, geometric mean value (± 2 SD) for the same reference toxicant, as derived at the test facility in previous tests using the procedures and conditions for reference toxicity tests described herein; and

- anything unusual about the test, any problems encountered, any remedial measures taken.

7.2 Additional Reporting Requirements

The following list of items must be either included in the test-specific report or the general report, or held on file for a minimum of five years. Filed information might include:

- a record of the chain of custody for field-collected or other samples tested for monitoring or remediation purposes;

- a copy of the record of acquisition for the sample(s);

- chemical analytical data on the sample(s) not included in the test-specific report;

- bench sheets for the observations and measurements recorded during the test;

- bench sheets and warning chart(s) for the reference toxicity tests; and

- information on the calibration of equipment and instruments.

Original data sheets must be signed or initialled, and dated by the laboratory personnel conducting the tests.

7.2.1 Test Substance or Material

- name and signature of person(s) who collected and/or provided the sample(s) or subsample(s);
- records of sample/subsample log-entry sheets; all information (e.g., code, sample description, date/time of sampling) affixed to label(s) on sample container(s); description of sample container (size and material);
- volume of sample(s) or subsample(s);
- transport and storage conditions (e.g., times, in sealed container, in darkness; temperature during storage at the laboratory; indication if samples/subsamples frozen or partially frozen on arrival);
- appearance (e.g., odour, colour) and conditions (e.g., temperature, in darkness, in sealed container) of samples/subsamples upon receipt and during storage;
- any additional records obtained for field samples (e.g., field records provided or maintained during sample collection) or chemical samples (impurities, additives, structural formulae, etc.).

7.2.2 Test Organisms

- records of taxonomic confirmation of species, including name of person(s) or facility identifying the organisms and the taxonomic guidelines or method used to confirm species;
- history, transport conditions, and age of any egg masses, tadpoles, or breeding adult frogs used to provide test organisms;
- description of culture conditions and procedures, including temperature, lighting, water source and quality and details on its renewal, type and quality of substrate, density of organisms, records of health, disease treatment, hibernation, breeding attempts, and performance indices; and any acclimation conditions and procedures (e.g., temperature), including rate of change;
- procedures used to count, handle, and transfer animals; and those to determine their mortality, condition, appearance, and behaviour;
- source and composition of food, procedures used to prepare and store food, feeding method(s), feeding frequency and ration;
- if test organisms are imported for immediate use in tests: all supplier records provided with each shipment, including species, life stage, age, and number of test organisms shipped, as well as date and time of shipment; temperature and dissolved oxygen concentration of water in shipment container(s) when shipped and upon arrival; and
- any permits required for the collection and use of the test species as required by regional, provincial, and/or federal authorities as well as the Canadian Council on Animal Care.

7.2.3 Test Facilities and Apparatus

- all results for preliminary tests with control/dilution water and reference toxicant, undertaken by the laboratory previously inexperienced with performing the biological test method described herein in advance of any reporting of definitive test results (see Section 3.3);
- description of systems for providing lighting and for regulating temperature within the test facility;
- description of any system for providing air and regulating air flow to test vessels; and
- description of procedures used to clean or rinse test apparatus.

7.2.4 Control/Dilution Water

- sampling and storage details if control/dilution water was receiving water;

- details regarding any water *pre-treatment* (e.g., filtration, sterilization, chlorination, dechlorination, dechloramination; adjustment for pH, temperature, and/or hardness; degassing, aeration rate and duration);

- any ancillary water-quality variables (e.g., dissolved metals, ammonia, pesticides, suspended solids, residual chlorine, iodide; see Section 2.3.5) measured before and/or during the toxicity test; and

- storage conditions (if any) and duration before use.

7.2.5 Test Method

- description of the laboratory's previous experience (e.g., preliminary tests, control performance history, etc.) with this biological test method for measuring toxicity using *Lithobates pipiens*;

- procedure used in preparing and storing stock and/or test solutions of chemicals; description and concentration(s) of any solvent used;

- methods used (with citations) for chemical analyses of sample or test solutions; details concerning sampling, sample/solution preparation, and storage before chemical analyses; and

- use and description of preliminary or range-finding test(s).

7.2.6 Test Conditions and Procedures

- photoperiod, light source, and measurements of light intensity adjacent to surface of test solutions;

- procedure for adding test organisms to test vessels;

- description of food source, type, and ration (quantity and frequency of feeding);

- conditions, procedures, and frequency for toxicity tests with reference toxicant(s);

- chemical analyses of concentrations of reference toxicant in stock solution(s), and if measured, in test solutions;

- appearance of sample (or mixture thereof) or test solutions in test vessels; changes in appearance noted during the test;

- water quality measurements for water supply used as culture/control/dilution water, and for water in aquaria or tanks containing adults (see Sections 2.3.5 and 2.4.2);

- description of any terrestrial habitat provided for adult frogs;

- total hardness and/or alkalinity of the control/dilution water and at least the highest test concentration at the start of the test; and

- any other chemical measurements on the sample, stock solutions, or test solutions (e.g., chemical concentration, suspended solids content, conductivity, hardness, alkalinity) before and/or during the test.

7.2.7 Test Results

- results for any range-finding test(s) conducted;

- results for any statistical analyses conducted both with outliers and with outliers removed; for regression analyses or analysis of significant effects, information indicating sample size (e.g., number of replicates per treatment), parameter estimates with variance, any ANOVA table(s) generated, plots of fitted and observed values of any models used, results of outlier tests, results of tests for normality and homoscedasticity, and the output provided by the statistical program;

- warning chart showing the most recent and historical results for reference toxicity tests or positive control concentrations with the reference toxicant; if applicable, CV for mean historical data derived for reference toxicity tests or positive control concentrations performed using the reference toxicant; and

- graphical presentation of data.

References

Abernethy SG, and Westlake GF. 1989. Guidelines for pH Adjustment of Effluent Samples for Toxicity Testing, Ontario Ministry of Environment, Rexdale, Ontario. 11 p.

Aldrich AP. 2009. Sensitivity of Amphibians to Pesticides, *Agrarforschung*, 16(11-12):466–471.

ALET (Atlantic Laboratory for Environmental Testing). 2004. Progress Report on Leopard Frog Culturing and Test Method Development, Report prepared for Method Development and Applications Section, Environment Canada, Ottawa, Ontario. 19 p.

ALET. 2006. Progress Report on Culturing and Test Method Development of *Rana pipiens* (Leopard Frogs), Report prepared for Method Development and Applications Section, Environment Canada, Ottawa, Ontario. 29 p.

ALET. 2009. Progress Report on Culturing and Test Method Development of Amphibians at the Atlantic Laboratory for Environmental Testing Up to Fiscal Year 2008-9, Report prepared for Method Development and Applications Unit, Environment Canada, Ottawa, Ontario. 46 p.

ALET. 2013. Progress Report on Culturing and Breeding Amphibians at the Atlantic Laboratory for Environmental for Fiscal Year 2012-13, Report prepared for Method Development and Applications Unit, Environment Canada, Ottawa, Ontario. 10 p.

ALET. 2015. *Lithobates pipiens* Ammonia Toxicity Test Results, Report prepared for Method Development and Applications Unit, Environment Canada, Ottawa, Ontario. 12 p.

ALET. 2016a. *Lithobates pipiens* 21-Day Chronic Feeding Study, Report prepared for Method Development and Applications Unit, Environment and Climate Change Canada, Ottawa, Ontario. 22 p.

ALET. 2016b. *Lithobates pipiens* 21-Day Chronic Sodium Chloride Toxicity Test Results, Report prepared for Method Development and Applications Unit, Environment and Climate Change Canada, Ottawa, Ontario. 29 p.

ALET. 2018. Review of Amphibian Culturing Progress at ALET 2017-18, Report prepared for Method Development and Applications Unit, Environment and Climate Change Canada, Ottawa, Ontario. 9 p.

Allran JW, and Karasov WH. 2000. Effects of Atrazine and Nitrate on Northern Leopard Frog (*Rana pipiens*) Larvae Exposed in the Laboratory from Posthatch Through Metamorphosis, *Environmental Toxicology and Chemistry*, 19(11):2850–2855.

Alworth LC, and Harvey SB. 2007. IACUC Issues Associated with Amphibian Research, *ILAR Journal*, 48(3):278–289.

Amaral DFD, Guerra V, Motta AGC, e Silva DM, and Rocha TL. 2019. Ecotoxicity of Nanomaterials in Amphibians: A Critical Review, *Science of the Total Environment*, 686:332–344.

Ankley GT, Kuehl DW, Kahl MD, Jensen KM, Butterworth BC, and Nichols JW. 2004. Partial Life-Cycle Toxicity and Bioconcentration Modeling of Perfluorooctanesulfonate in the Northern Leopard Frog (*Rana pipiens*), *Environmental Toxicology and Chemistry*, 23(11):2745–2755.

APHA, AWWA, and WEF (American Public Health Association, American Water Works Association, and Water Environment Federation). 2011. 8930 Amphibians (Proposed), In: *Standard Methods for the Examination of Water and Wastewater, 22nd ed.*, Rice EW, Baird RB, Eaton AD, and Clesceri LS (eds.), Washington, DC. 10 p.

APHA, AWWA, and WEF. 2017. *Standard Methods for the Examination of Water and Wastewater, 23rd ed.*, Baird RB, Eaton AD, and Rice EW (eds.), Washington, DC. 1504 p.

Araújo CVM, Shinn C, Moreira-Santos M, Lopes I, Espíndola ELG, and Ribeiro R. 2014. Copper-Driven Avoidance and Mortality in Temperate and Tropical Tadpoles, *Aquatic Toxicology*, 146:70–75.

ASTM. 2019. Standard Guide for Conducting the Frog Embryo Teratogenesis Assay-Xenopus (FETAX), ASTM International, West Conshohocken, Pennsylvania, ASTM E1439-12(2019). 16 p.

ASTM (American Society for Testing and Materials). 2022a. Standard Guide for Conducting Whole Sediment Toxicity Tests with Amphibians, ASTM International, West Conshohocken, Pennsylvania, ASTM E2591-22(2022). 19 p.

ASTM. 2022b. Standard Guide for Use of Lighting in Laboratory Testing, ASTM International, West Conshohocken, Pennsylvania, ASTM E1733-95(2022). 15 p.

ASTM. 2023a. Standard Guide for Conducting Acute Toxicity Tests on Aqueous Ambient Samples and Effluents with Fishes, Macroinvertebrates, and Amphibians, ASTM International, West Conshohocken, Pennsylvania, ASTM E1192-23(2023). 15 p.

ASTM. 2023b. Standard Guide for Conducting Acute Toxicity Tests on Test Materials with Fishes, Macroinvertebrates, and Amphibians, ASTM International, West Conshohocken, Pennsylvania, ASTM E729-23e1(2023). 23 p.

AVMA (American Veterinary Medical Association). 2020. AVMA Guidelines for the Euthanasia of Animals: 2020 Edition, AVMA, Schaumburg, Illinois. 121 p.

Basu N, Crump D, Head J, Hickey G, Hogan N, Maguire S, Xia J, and Hecker M. 2019. EcoToxChip: A Next-Generation Toxicogenomics Tool for Chemical Prioritization and Environmental Management, *Environmental Toxicology and Chemistry*, 38(2):279–288.

Baud DR, and Beck ML. 2005. Interactive Effects of UV-B and Copper on Spring Peeper Tadpoles (*Pseudacris crucifer*), *Southeastern Naturalist*, 4(1):15–22.

Becker CG, Fonseca CR, Haddad CFB, Batista R, and Prado PI. 2007. Habitat Split and the Global Decline of Amphibians, *Science*, 318(5857):1775–1777.

Beebee TJC, and Griffiths RA. 2005. The Amphibian Decline Crisis: A Watershed for Conservation Biology?, *Biological Conservation*, 125(3):271–285.

Berger L, Speare R, Daszak P, Green DE, Cunningham AA, Goggin CL, Slocombe R, Ragan MA, Hyatt AD, McDonald KR, Hines HB, Lips KR, Marantelli G, and Parkes H. 1998. Chytridiomycosis Causes Amphibian Mortality Associated with Population Declines in the Rain Forests of Australia and Central America, *Proceedings of the National Academy of Sciences of the United States of America*, 95(15):9031–9036.

Berrill M, Bertram S, Wilson A, Louis S, Brigham D, and Stromberg C. 1993. Lethal and Sublethal Impacts of Pyrethroid Insecticides on Amphibian Embryos and Tadpoles, *Environmental Toxicology and Chemistry*, 12(3):525–539.

Berrill M, Bertram S, McGillivray L, Kolohon M, and Pauli B. 1994. Effects of Low Concentrations of Forest-Use Pesticides on Frog Embryos and Tadpoles, *Environmental Toxicology and Chemistry*, 13(4):657–664.

Berrill M, Coulson D, McGillivray L, and Pauli B. 1998. Toxicity of Endosulfan to Aquatic Stages of Anuran Amphibians, *Environmental Toxicology and Chemistry*, 17(9):1738–1744.

Biga LM, and Blaustein AR. 2013. Variations in Lethal and Sublethal Effects of Cypermethrin Among Aquatic Stages and Species of Anuran Amphibians, *Environmental Toxicology and Chemistry*, 32(12):2855–2860.

Birge WJ, Black JA, and Westerman AG. 1985. Short-Term Fish and Amphibian Embryo-Larval Tests for Determining the Effects of Toxicant Stress on Early Life Stages and Estimating

Chronic Values for Single Compounds and Complex Effluents: Complex Mixtures, *Environmental Toxicology and Chemistry*, 4(6):807–821.

Birge WJ, Westerman AG, and Spromberg JA. 2000. Comparative Toxicology and Risk Assessment of Amphibians, In: *Ecotoxicology of Amphibians and Reptiles*, Sparling DW, Linder G, and Bishop CA (eds.), SETAC Technical Publication Series, SETAC Press, Pensacola, Florida, pp. 727–791.

Blaustein AR, Hoffman PD, Hokit DG, Kiesecker JM, Walls SC, and Hays JB. 1994. UV Repair and Resistance to Solar UV-B in Amphibian Eggs: A Link to Population Declines?, *Proceedings of the National Academy of Sciences of the United States of America*, 91(5):1791–1795.

Blaustein AR, Kiesecker JM, Chivers DP, Hokit DG, Marco A, Belden LK, and Hatch A. 1998. Effects of Ultraviolet Radiation on Amphibians Field Experiments, *American Zoologist*, 38(6):799–812.

Blaustein AR, and Kiesecker JM. 2002. Complexity in Conservation: Lessons from the Global Decline of Amphibian Populations, *Ecology Letters*, 5(4):597–608.

Blaustein AR, Romansic JM, Kiesecker JM, and Hatch AC. 2003. Ultraviolet Radiation, Toxic Chemicals and Amphibian Population Declines, *Diversity and Distributions*, 9(2):123–140.

Boone MD, Semlitsch RD, Little EE, and Doyle MC. 2007. Multiple Stressors in Amphibian Communities: Effects of Chemical Contamination, Bullfrogs, and Fish, *Ecological Applications*, 17(1):291–301.

Boone MD, Hammond SA, Veldhoen N, Youngquist M, and Helbing CC. 2013. Specific Time of Exposure During Tadpole Development Influences Biological Effects of the Insecticide Carbaryl in Green Frogs (*Lithobates clamitans*), *Aquatic Toxicology*, 130–131:139–148.

Bradfield K. 2010. Managing Water Quality for Amphibians in Captivity, Herpetology Department, Durrell Wildlife Conservation Trust. 28 p.

Bridges CM, Dwyer FJ, Hardesty DK, and Whites DW. 2002. Comparative Contaminant Toxicity: Are Amphibian Larvae More Sensitive Than Fish?, *Bulletin of Environmental Contamination and Toxicology*, 69:562–569.

Brinkman SF, and Johnston WD. 2012. Acute Toxicity of Zinc to Several Aquatic Species Native to the Rocky Mountains, *Archives of Environmental Contamination and Toxicology*, 62:272–281.

Brown SR, Flynn RW, and Hoverman JT. 2021. Perfluoroalkyl Substances Increase Susceptibility of Northern Leopard Frog Tadpoles to Trematode Infection, *Environmental Toxicology and Chemistry*, 40(3):689–694.

Browne RK, Odum RA, Herman T, and Zippel K. 2007. Facility Design and Associated Services for the Study of Amphibians, *ILAR Journal*, 48(3):188–202.

Browne RK, and Zippel K. 2007. Reproduction and Larval Rearing of Amphibians, *ILAR Journal*, 48(3):214–234.

Brühl CA, Schmidt T, Pieper S, and Alscher A. 2013. Terrestrial Pesticide Exposure of Amphibians: An Underestimated Cause of Global Decline?, *Nature: Scientific Reports*, 3:1–4.

Budischak SA, Belden LK, and Hopkins WA. 2009. Relative Toxicity of Malathion to Trematode-Infected and Noninfected *Rana palustris* Tadpoles, *Archives of Environmental Contamination and Toxicology*, 56:123–128.

Burggren WW, and Warburton S. 2007. Amphibians as Animal Models for Laboratory Research in Physiology, *ILAR Journal*, 48(3):260–269.

Burghardt GM. 2013. Environmental Enrichment and Cognitive Complexity in Reptiles and Amphibians: Concepts, Review, and Implications for Captive Populations, *Applied Animal Behaviour Science*, 147:286−298.

Canadian Herpetological Society. 2020. Northern Leopard Frog (*Lithobates pipiens*). URL: http://canadianherpetology.ca/species/species_page.html?cname=Northern%20Leopard%20Frog. Last accessed on December 2, 2020.

Cashins SD, Alford RA, and Skerratt LF. 2008. Lethal Effects of Latex, Nitrile, and Vinyl Gloves on Tadpoles, *Herpetological Review*, 39(3):298–301.

CCAC (Canadian Council on Animal Care). 2005. Guidelines on the Care and Use of Fish in Research, Teaching and Testing, Ad Hoc Subcommittee on Fish of the Canadian Council on Animal Care Guidelines Committee, Ottawa, Ontario. 94 p.

CCAC. 2010. CCAC Guidelines On: Euthanasia of Animals Used in Science, Ad Hoc Subcommittee on Euthanasia of the Canadian Council on Animal Care Guidelines Committee, Ottawa, Ontario. 36 p.

CCAC. 2021. CCAC Guidelines: Amphibians, Canadian Council on Animal Care, Ottawa, Ontario. 98 p.

CCME (Canadian Council of Ministers of the Environment). 1999. Canadian Water Quality Guidelines for the Protection of Aquatic Life: Reactive Chlorine Species, In: *Canadian Environmental Quality Guidelines 1999*, Canadian Council of Ministers of the Environment, Winnipeg, Manitoba, Publication No. 1299. 9 p.

CCME. 2010. Canadian Water Quality Guidelines for the Protection of Aquatic Life: Ammonia, In: *Canadian Environmental Quality Guidelines 2010*, Canadian Council of Ministers of the Environment, Winnipeg, Manitoba, Publication No. 1299. 8 p.

CCREM (Canadian Council of Resource and Environmental Ministers). 1987. Canadian Water Quality Guidelines, Task Force on Water Quality Guidelines, Environment Canada, Ottawa, Ontario. 1484 pp.

CESCC (Canadian Endangered Species Conservation Council). 2011. Wild Species 2010: The General Status of Species in Canada, National General Status Working Group, Ottawa, Ontario. 321 p.

Che J, Pang J, Zhao H, Wu G-f, Zhao E-m, and Zhang Y-p. 2007. Phylogeny of Raninae (Anura: Ranidae) Inferred from Mitochondrial and Nuclear Sequences, *Molecular Phylogenetics and Evolution*, 43(1):1–13.

Chen T-H, Gross JA, and Karasov WH. 2007. Adverse Effects of Chronic Copper Exposure in Larval Northern Leopard Frogs (Rana pipiens), *Environmental Toxicology and Chemistry*, 26(7):1470–1475.

Chen T-H, Gross JA, and Karasov WH. 2009. Chronic Exposure to Pentavalent Arsenic of Larval Leopard Frogs (Rana pipiens): Bioaccumulation and Reduced Swimming Performance, *Ecotoxicology*, 18:587–593.

Coady KK, Lehman CM, Currie RJ, and Marino TA. 2014. Challenges and Approaches to Conducting and Interpreting the Amphibian Metamorphosis Assay and the Fish Short-Term Reproduction Assay, *Birth Defects Research (Part B)*, 101(1):80–89.

Cohen MM Jr. 2001. Frog Decline, Frog Malformations, and a Comparison of Frog and Human Health, *American Journal of Medical Genetics*, 104(2):101–109.

Collins JP, and Storfer A. 2003. Global Amphibian Declines: Sorting the Hypotheses, *Diversity and Distributions*, 9(2):89–98.

Collins JT, and Taggart TW. 2009. *Standard Common and Current Scientific Names for North American Amphibians, Turtles, Reptiles, and Crocodilians, 6th ed.*, Center for North American Herpetology, Lawrence, Kansas. 53 p.

COSEWIC (Committee on the Status of Endangered Wildlife in Canada). 2009. COSEWIC Assessment and Update Status Report on the Northern Leopard Frog *Lithobates pipiens*, Rocky Mountain Population, Western Boreal/Prairie Populations, and Eastern Populations in Canada, Committee on

the Status of Endangered Wildlife in Canada, Ottawa, Ontario. 76 p.

Croteau MC, Davidson M, Duarte-Guterman P, Wade M, Popesku JT, Wiens S, Lean DRS, and Trudeau VL. 2009. Assessment of Thyroid System Disruption in *Rana pipiens* Tadpoles Chronically Exposed to UVB Radiation and 4-*tert*-octylphenol, *Aquatic Toxicology*, 95(2):81–92.

Crump D, Werry K, Veldhoen N, Van Aggelen G, and Helbing CC. 2002. Exposure to the Herbicide Acetochlor Alters Thyroid Hormone-Dependent Gene Expression and Metamorphosis in *Xenopus laevis*, *Environmental Health Perspectives*, 110(12):1199–1205.

Crump D, Hickey G, Boulanger E, Masse A, Head JA, Hogan N, Maguire S, Xia J, Hecker M, and Basu N. 2023. Development and Initial Testing of EcoToxChip, a Novel Toxicogenomics Tool for Environmental Management and Chemical Risk Assessment, *Environmental Toxicology and Chemistry*, 42(8):1763-1771.

Deitzer G. 1994. Spectral Comparisons of Sunlight and Different Lamps, In: *Proceedings of International Lighting in Controlled Environments Workshop*, Tibbits TW (ed.), Madison, Wisconsin, pp. 197–199.

Densmore CL, and Green DE. 2007. Diseases of Amphibians, *ILAR Journal*, 48(3):235–254.

Denton D, Diamond J, and Zheng L. 2011. Test of Significant Toxicity: A Statistical Application for Assessing Whether an Effluent or Site Water is Truly Toxic, *Environmental Toxicology and Chemistry*, 30:1117–1126.

Denton D, Diamond J, and Stuber R. 2019. Comparison of False-Positive Rates of 2 Hypothesis-Test Approaches in Relation to Laboratory Toxicity Test Performance, *Environmental Toxicology and Chemistry*, 38:511–523.

Dewey T. 1999. *Lithobates pipiens* (Northern Leopard Frog), in Animal Diversity Web. URL: https://animaldiversity.org/accounts/Lithobates_pipiens/. Last accessed on December 8, 2020.

Dodd CK Jr. 2013. *Lithobates pipiens* (Schreber, 1782), Northern Leopard Frog/Grenouille Leopard, In: *Frogs of the United States and Canada, Vol. 1.*, John Hopkins University Press, Baltimore, Maryland, pp. 578–608.

Doe K, Jackman P, and Pauli BD. 2012. Acute Toxicity of Thirty Four Pesticide Formulations to Amphibians, Fish, Invertebrates, Bacteria, and Algae, and a Comparative Analysis of Their Relative Sensitivity, Environment Canada, Moncton, New Brunswick. 149 p.

Earl JE, and Whiteman HH. 2009. Effects of Pulsed Nitrate Exposure on Amphibian Development, *Environmental Toxicology and Chemistry*, 28(6):1331–1337.

Earl JE, and Whiteman HH. 2010. Evaluation of Phosphate Toxicity in Cope's Gray Treefrog (*Hyla chrysoscelis*) Tadpoles, *Journal of Herpetology*, 44(2):201–208.

Edginton AN. 2001. Review of Amphibian Culturing and Toxicity Testing Procedures, Report prepared for Method Development and Applications Section, Environment Canada, Ottawa, Ontario. 70 p.

Edginton AN, Stephenson GR, Sheridan PM, Thompson DG, and Boermans HJ. 2003. Effect of pH and Release® on Two Life Stages of Four Anuran Amphibians, *Environmental Toxicology and Chemistry*, 22(11):2673–2678.

Edginton AN, Sheridan PM, Stephenson GR, Thompson DG, and Boermans H. 2004. Comparative Effects of pH and Vision® Herbicide on Two Life Stages of Four Anuran Amphibian Species, *Environmental Toxicology and Chemistry*, 23(4):815–822.

EC (Environment Canada). 1990. Guidance Document on the Control of Toxicity Test Precision Using Reference Toxicants, Environment Canada, Ottawa, Ontario, Report EPS 1/RM/12. 90 p.

EC. 1994. Guidance Document on Collection and Preparation of Sediments for Physicochemical Characterization and Biological Testing,

Environment Canada, Ottawa, Ontario, Report EPS 1/RM/29. 170 p.

EC. 1998. Biological Test Method: Toxicity Tests Using Early Life Stages of Salmonid Fish (Rainbow Trout), Environment Canada, Ottawa, Ontario, Report EPS 1/RM/28 2nd ed. 122 p.

EC. 1999. Guidance Document on Application and Interpretation of Single-Species Tests in Environmental Toxicology, Environment Canada, Ottawa, Ontario, Report EPS 1/RM/34. 226 p.

EC. 2005. Guidance Document on Statistical Methods for Environmental Toxicity Tests, Environment Canada, Ottawa, Ontario, Report EPS 1/RM/46, including June 2007 amendments. 283 p.

EC. 2007. Biological Test Method: Test of Reproduction and Survival Using the Cladoceran *Ceriodaphnia dubia*, Environment Canada, Ottawa, Ontario, Report EPS 1/RM/21 2nd ed. 100 p.

EC. 2011. Biological Test Method: Test of Larval Growth and Survival Using Fathead Minnow, Environment Canada, Ottawa, Ontario, Report EPS 1/RM/22 2nd ed. 100 p.

EC. 2013a. Biological Test Method: Test for Growth in Contaminated Soil Using Terrestrial Plants Native to the Boreal Region, Environment Canada, Ottawa, Ontario, Report EPS 1/RM/56. 137 p.

EC. 2013b. Management Plan for the Northern Leopard Frog (*Lithobates pipiens*), Western Boreal/Prairie Populations, in Canada, Species at Risk Act Management Plan Series, Environment Canada, Ottawa, Ontario. 33 p.

EC. 2014. Test for Measuring Survival and Reproduction of Springtails Exposed to Contaminants in Soil, Environment Canada, Ottawa, Ontario, Report EPS 1/RM/47 2nd ed. 184 p.

EC. 2017. Biological Test Method: Test for Survival, Growth and Reproduction in Sediment and Water Using the Freshwater Amphipod *Hyalella azteca*, Environment and Climate Change Canada, Ottawa, Ontario, Report EPS 1/RM/33 3rd ed. 203 p.

ECCC (Environment and Climate Change Canada). 2020. Biological Test Method: Test for Measuring Reproduction of Oribatid Mites Exposed to Contaminants in Soil, Environment and Climate Change Canada, Ottawa, Ontario, Report STB 1/RM/61. 135 p.

ECCC. 2022. Biological Test Method: Tests for Measuring Avoidance Behaviour or Reproduction of Earthworms (*Eisenia andrei* or *Dendrodrilus rubidus*) Exposed to Contaminants in Soil, Environment and Climate Change Canada, Ottawa, Ontario, Report STB 1/RM/43 2nd ed. 180 p.

ECCC. 2023. Technical Report on Northern Leopard Frogs (*Lithobates pipiens*), Report prepared by Environment and Climate Change Canada, Ottawa, Ontario. 19 p.

EFSA PPR Panel (EFSA Panel on Plant Protection Products and their Residues), Ockleford C, Adriaanse P, Berny P, Brock T, Duquesne S, Grilli S, Hernandez-Jerez AF, Bennekou SH, Klein M, Kuhl T, Laskowski R, Machera K, Pelkonen O, Pieper S, Stemmer M, Sundh I, Teodorovic I, Tiktak A, Topping CJ, Wolterink G, Aldrich A, Berg C, Ortiz-Santaliestra M, Weir S, Streissl F, and Smith RH. 2018. Scientific Opinion on the State of the Science on Pesticide Risk Assessment for Amphibians and Reptiles, *EFSA Journal*, 16(2):5125. 301 p.

ENSR International. 2004. Development of a Standardized Approach for Assessing Potential Risks to Amphibians Exposed to Sediment and Hydric Soils, Technical Report TR-2245-ENV, Report prepared for Naval Facilities Engineering Command (NAVFAC), Port Hueneme, California. 340 p.

Egea-Serrano A, Relyea RA, Tejedo M, and Torralva M. 2012. Understanding of the Impact of Chemicals on Amphibians: A Meta-Analytic Review, *Ecology and Evolution*, 2(7):1382–1397.

EU (European Union). 2010. Directive 2010/63/EU of the European Parliament and of the Council of

22 September 2010 on the Protection of Animals Used for Scientific Purposes, *Official Journal of the European Union*, L 276:33–79.

Experimental Pathology Laboratories Inc. 2019. Histology and Pathology of Northern Leopard Frogs Exposed to Thyroxine: Pathology Report, Report prepared for Environment and Climate Change Canada, 14 p.

FCSAP (Federal Contaminated Sites Action Plan). 2010. Ecological Risk Assessment Guidance – Module 1: Toxicity Test Selection and Interpretation, Environment and Climate Change Canada, Ottawa, Ontario. 85 p.

FCSAP. 2019. Ecological Risk Assessment Guidance – Module 6: Ecological Risk Assessment for Amphibians on Federal Contaminated Sites, Version 1.0, Environment and Climate Change Canada, Ottawa, Ontario. 145 p.

Ferrie GM, Alford VC, Atkinson J, Baitchman E, Barber D, Blaner WS, Crawshaw G, Daneault A, Dierenfeld E, Finke M, Fleming G, Gagliardo R, Hoffman EA, Karasov W, Klasing K, Koutsos E, Lankton J, Lavin SR, Lentini A, Livingston S, Lock B, Mason T, McComb A, Morris C, Pessier AP, Olea-Popelka F, Probst T, Rodriguez C, Schad K, Semmen K, Sincage J, Stamper MA, Steinmetz J, Sullivan K, Terrell S, Wertan N, Wheaton CJ, Wilson B, and Valdes EV. 2014. Nutrition and Health in Amphibian Husbandry, *Zoo Biology*, 33(6):485–501.

Fort Environmental Laboratories Inc. 2018. Evaluation of Northern Leopard Frog Tadpoles, Report prepared for Environment and Climate Change Canada, 10 p.

Forzán MJ, Vanderstichel R, Hogan NS, Teather K, and Wood J. 2010. Prevalence of *Batrachochytrium dendrobatidis* in Three Species of Wild Frogs on Prince Edward Island, Canada, *Diseases of Aquatic Organisms*, 91:91–96.

Fraker SL, and Smith GR. 2004. Direct and Interactive Effects of Ecologically Relevant Concentrations of Organic Wastewater Contaminants on *Rana pipiens* Tadpoles, *Environmental Toxicology*, 19(3):250–256.

Freda J. 1986. The Influence of Acidic Pond Water on Amphibians: A Review, *Water, Air, and Soil Pollution*, 30:439–450.

Freda J, and McDonald DG. 1990. Effects of Aluminum on the Leopard Frog, *Rana pipiens*: Life Stage Comparisons and Aluminum Uptake, *Canadian Journal of Fisheries and Aquatic Sciences,* 47:210–216.

Fridgen CM, Hogan NS, Trudeau VL, Berrill M, Doe K, and Pauli BD. 2007. Developmental Profiles of the Northern Leopard Frog: Are There Sensitive Periods for Disruption of the Thyroid and Reproductive Axis?, SETAC – 17th European Annual Meeting, Porto, Portugal.

Fridgen CM. 2009. Laboratory Controlled Normal and Atrazine Exposed Development of Northern Leopard Frog, *Rana pipiens* Tadpoles. Master's thesis – Trent University, Peterborough, Ontario. 126 p.

Fridgen CM, Xenopoulos M, and Metcalfe C. 2009. TiO_2 Nanoparticles Increase the Toxicity of Roundup WeatherMax® in Exposed Frog Tadpoles, SETAC – 30th North American Annual Meeting, New Orleans, Louisiana.

Frost DR, Grant T, Faivovich J, Bain RH, Haas A, Haddad CFB, de Sá RO, Channing A, Wilkinson M, Donnellan SC, Raxworthy CJ, Campbell JA, Blotto BL, Moler PE, Drewes RC, Nussbaum RA, Lynch JD, Green DM, and Wheeler W. 2006. The Amphibian Tree of Life, *Bulletin of the American Museum of Natural History*, 297:1–291.

Frost DR, Moriarty Lemmon E, McDiarmid RW, and Mendelson JR III. 2017. Anura – Frogs, In: *Scientific and Standard English Names of Amphibians and Reptiles of North America North of Mexico, with Comments Regarding Confidence in Our Understanding, 8th ed.*, B.I. Crother (ed.), Society for the Study of Amphibians and Reptiles, SSAR Herpetological Circular No. 43, pp. 6–24.

Gavel MJ, Young SD, Blais N, Forbes MR, and Robinson SA. 2021. Trematodes Coupled with Neonicotinoids: Effects on Blood Cell Profiles of a Model Amphibian, *Parasitology Research*, 120(6):2135–2148.

Gentz EJ. 2007. Medicine and Surgery of Amphibians, *ILAR Journal*, 48(3):255–259.

Glaberman S, Kiwiet J, and Aubee CB. 2019. Evaluating the Role of Fish as Surrogates for Amphibians in Pesticide Ecological Risk Assessment, *Chemosphere*, 235:952–958.

Glennemeier KA, and Denver RJ. 2002. Role for Corticoids in Mediating the Response of *Rana pipiens* Tadpoles to Intraspecific Competition, *Journal of Experimental Zoology*, 292(1):32–40.

Gosner KL. 1960. A Simplified Table for Staging Anuran Embryos and Larvae with Notes on Identification, *Herpetologica*, 16(3):183–190.

Goulet BN, and Hontela A. 2003. Toxicity of Cadmium, Endosulfan, and Atrazine in Adrenal Steroidogenic Cells of Two Amphibian Species, *Xenopus laevis* and *Rana catesbeiana*, *Environmental Toxicology and Chemistry*, 22(9):2106–2113.

John W Green Ecostatistical Consulting LLC. 2021. Comparison of Three Models for Statistical Analysis of Amphibian Developmental Stage, Report prepared for Method Development and Applications Unit, Environment and Climate Change Canada, Ottawa, Ontario. 172 p.

Green J, and Wheeler JR. 2013. The Use of Carrier Solvents in Regulatory Aquatic Toxicology Testing: Practical, Statistical, and Regulatory Considerations, *Aquatic Toxicology*, 144–145:242–249.

Green JW, Springer TA, and Holbech H. 2018. *Statistical Analysis of Ecotoxicity Studies*. John Wiley & Sons Inc., Hoboken, New Jersey. 390 p.

Greer AL, Schock DM, Brunner JL, Johnson RA, Picco AM, Cashins SD, Alford RA, Skerratt LF, and Collins JP. 2009. Guidelines for the Safe Use of Disposable Gloves with Amphibian Larvae in Light of Pathogens and Possible Toxic Effects, *Herpetological Review*, 40(2):145–147.

Greulich K, and Pflugmacher S. 2004. Uptake and Effects on Detoxication Enzymes of Cypermethrin in Embryos and Tadpoles of Amphibians, *Archives of Environmental Contamination and Toxicology*, 47(4):489–495.

Gromko MH, Mason FS, and Smith-Gill SJ. 1973. Analysis of the Crowding Effect in *Rana pipiens* Tadpoles, *Journal of Experimental Zoology*, 186(1):63–71.

Gross JA, Johnson PTJ, Prahl LK, and Karasov WH. 2009. Critical Period of Sensitivity for Effects of Cadmium on Frog Growth and Development, *Environmental Toxicology and Chemistry*, 28(6):1227–1232.

Gunderson MP, Veldhoen N, Skirrow RC, Macnab MK, Ding W, van Aggelen G, and Helbing CC. 2011. Effect of Low Dose Exposure to the Herbicide Atrazine and its Metabolite on Cytochrome P450 Aromatase and Steroidogenic Factor-1 mRNA Levels in the Brain of Premetamorphic Bullfrog Tadpoles (*Rana catesbeiana*), *Aquatic Toxicology*, 102(1-2):31–38.

Gutleb AC, Bronkhorst M, van den Berg JHJ, and Murk AJ. 2001. Latex Laboratory-Gloves: An Unexpected Pitfall in Amphibian Toxicity Assays with Tadpoles, *Environmental Toxicology and Pharmacology*, 10(3):119–121.

Halder M, Kienzler A, Whelan M, and Worth A. 2014. EURL ECVAM Strategy to Replace, Reduce and Refine the Use of Fish in Aquatic Toxicity and Bioaccumulation Testing, Joint Research Centre Science and Policy Reports, European Union, Luxembourg, Report EUR 26973. 34 p.

Haramura T. 2016. Hatching Plasticity in Response to Salinity Levels in a Rhacophorid Frog Inhabiting a Coastal Area, *Journal of Zoology*, 299(2):125-131.

Hayes TB, Falso P, Gallipeau S, and Stice M. 2010. The Cause of Global Amphibian Declines: A Developmental Endocrinologist's Perspective, *Journal of Experimental Biology*, 213(6):921–933.

Heerema JL, Jackman KW, Miliano RC, Li L, Zaborniak TSM, Veldhoen N, van Aggelen G, Parker WJ, Pyle GG, and Helbing CC. 2018. Behavioral and Molecular Analyses of Olfaction-

Mediated Avoidance Responses of *Rana* (*Lithobates*) *catesbeiana* Tadpoles: Sensitivity to Thyroid Hormones, Estrogen, and Treated Municipal Wastewater Effluent, *Hormones and Behavior*, 101:85–93.

Hersikorn BD, and Smits JEG. 2011. Compromised Metamorphosis and Thyroid Hormone Changes in Wood Frogs (*Lithobates sylvaticus*) Raised on Reclaimed Wetlands on the Athabasca Oil Sands, *Environmental Pollution*, 159(2):596–601.

Higley E, Tompsett AR, Giesy JP, Hecker M, and Wiseman S. 2013. Effects of Triphenyltin on Growth and Development of the Wood Frog (*Lithobates sylvaticus*), *Aquatic Toxicology*, 144-145:155–161.

Hillis DM. 2007. Constraints in Naming Parts of the Tree of Life, *Molecular Phylogenetics and Evolution*, 42(2):331–338.

Hinther A, Domanski D, Vawda S, and Helbing CC. 2010a. C-Fin: A Cultured Frog Tadpole Tail Fin Biopsy Approach for Detection of Thyroid Hormone-Disrupting Chemicals, *Environmental Toxicology and Chemistry*, 29(2):380–388.

Hinther A, Vawda S, Skirrow RC, Veldhoen N, Collins P, Cullen JT, van Aggelen G, and Helbing CC. 2010b. Nanometals Induce Stress and Alter Thyroid Hormone Action in Amphibia at or Below North American Water Quality Guidelines, *Environmental Science and Technology*, 44(21):8314–8321.

Hoenig JM, and Heisey DM. 2001. The Abuse of Power: The Pervasive Fallacy of Power Calculations for Data Analysis, *The American Statistician*, 55:19–24.

Hoffman EA, and Blouin MS. 2004. Evolutionary History of the Northern Leopard Frog: Reconstruction of Phylogeny, Phylogeography, and Historical Changes in Population Demography from Mitochondrial DNA, *Evolution*, 58(1):145–159.

Hogan N, Duarte P, Wade MG, Lean DRS, and Trudeau VL. 2008. Estrogenic Exposure Affects Metamorphosis and Alters Sex Ratios in the Northern Leopard Frog (*Rana pipiens*): Identifying Critically Vulnerable Periods of Development, *General and Comparative Endocrinology*, 156:515–523.

Holloway AC, Keene JL, Noakes DG, and Moccia RD. 2004. Effects of Clove Oil and MS-222 on Blood Hormone Profiles in Rainbow Trout *Oncorhynchus mykiss*, Walbaum, *Aquaculture Research*, 35(11):1025–1030.

Hoover GM, Chislock MF, Tornabene BJ, Guffey SC, Choi YJ, De Perre C, Hoverman JT, Lee LS, and Sepulveda MS. 2017. Uptake and Depuration of Four Per/Polyfluoroalkyl Substances (PFASS) in Northern Leopard Frog *Rana pipiens* Tadpoles, *Environmental Science and Technology Letters*, 4(10):399–403.

Hopkins WA. 2007. Amphibians as Models for Studying Environmental Change, *ILAR Journal*, 48(3):270–277.

Horne MT, and Dunson WA. 1995. Effects of Low pH, Metals, and Water Hardness on Larval Amphibians, *Archives of Environmental Contamination and Toxicology*, 29:500–505.

Houlahan JE, Findlay CS, Schmidt BR, Meyer AH, and Kuzmin SL. 2000. Quantitative Evidence for Global Amphibian Population Declines, *Nature*, 404(6779):752–755.

Howe CM, Berrill M, Pauli BD, Helbing CC, Werry K, and Veldhoen N. 2004. Toxicity of Glyphosate-Based Pesticides to Four North American Frog Species, *Environmental Toxicology and Chemistry*, 23(8):1928–1938.

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

ICAMA (Institute for the Control of Agrochemicals, Ministry of Agriculture). 2014. Test Guidelines on Environmental Safety Assessment for Chemical Pesticides – Part 18: Amphibian Acute Toxicity Test, GB/T 31270.18-2014, Hangzhou, China. 13 p.

ISO (International Organization for Standardization). 2006. Water Quality – Evaluation of Genotoxicity by Measurement of the Induction of Micronuclei – Part 1: Evaluation of Genotoxicity Using Amphibian Larvae, Geneva, Switzerland, ISO 21427-1:2006. 15 p.

ISO. 2017. General Requirements for the Competence of Testing and Calibration Laboratories, Geneva, Switzerland, ISO/IEC 17025:2017. 38 p.

ISO. 2019. Soil Quality – Identification of Ecotoxicological Test Species by DNA Barcoding, Geneva, Switzerland, ISO 21286:2019. 28 p.

Jackman P, Doe K, Pauli B, and Scroggins R. 2007. The Development of a Standardized Acute Toxicity Test with the Leopard Frog (*Rana pipiens*) and Comparison of the Results with Standard Test Species. Aquatic Toxicity Workshop – 34th Annual Meeting, Halifax, Nova Scotia.

Jackman KW, Veldhoen N, Miliano RC, Robert BJ, Li L, Khojasteh A, Zheng X, Zaborniak TSM, van Aggelen G, Lesperance M, Parker WJ, Hall ER, Pyle GG, and Helbing CC. 2018. Transcriptomics Investigation of Thyroid Hormone Disruption in the Olfactory System of the *Rana* [*Lithobates*] *catesbeiana* Tadpole, *Aquatic Toxicology*, 202:46–56.

Jofre MB, and Karasov WH. 1999. Direct Effect of Ammonia on Three Species of North American Anuran Amphibians, *Environmental Toxicology and Chemistry*, 18(8):1806–1812.

Johnson ML, Berger L, Philips L, and Speare R. 2003. Fungicidal Effects of Chemical Disinfectants, UV Light, Desiccation and Heat on the Amphibian Chytrid, *Batrachochytrium dendrobatidis*, *Diseases of Aquatic Organisms*, 57(3):255–260.

Johnson MJ, Aubee C, Salice CJ, Leigh K, Liu E, Pott U, and Pillard D. 2017. A Review of Ecological Risk Assessment Methods of Amphibians: Comparative Assessment of testing Methodologies and Available Data. Integrated Environmental Assessment and Management, Volume 13, Number 4, pp. 601–613.

Jones DK, Hammond JI, and Relyea RA. 2009. Very Highly Toxic Effects of Endosulfan Across Nine Species of Tadpoles: Lag Effects and Family-Level Sensitivity, *Environmental Toxicology and Chemistry*, 28(9):1939–1945.

Karraker NE, Gibbs JP, and Vonesh JR. 2008. Impacts of Road Deicing Salt on the Demography of Vernal Pool-Breeding Amphibians, *Ecological Applications*, 18(3):724-734.

Kendell K. 2002. Survey Protocol for the Northern Leopard Frog, Alberta Sustainable Resource Development, Fish and Wildlife Division, Edmonton, Alberta, Alberta Species at Risk Report No. 43. 42 p.

Kerby JL, Richards-Hrdlicka KL, Storfer A, and Skelly DK. 2010. An Examination of Amphibian Sensitivity to Environmental Contaminants: Are Amphibians Poor Canaries?, *Ecology Letters*, 13(1):60–67.

Kouba A, Vance C, Calatayud N, Rowlison T, Langhorne C, and Willard S. 2012. Assisted Reproductive Technologies (ART) for Amphibians, In: *Amphibian Husbandry Resource Guide, Edition 2.0*, Poole VA, and Grow S, eds., Association of Zoos and Aquariums, Silver Spring, Maryland, pp. 60–118.

Kuppert S. 2013. Providing Enrichment in Captive Amphibians and Reptiles: Is it Important to Know Their Communication?, *Smithsonian Herpetological Information Service*, 142:1–42.

Lacoul P, Freedman B, and Clair T. 2011. Effects of Acidification on Aquatic Biota in Atlantic Canada, *Environmental Reviews*, 19:429–460.

Landis WG, and Chapman PM. 2011. Well Past Time to Stop Using NOELs and LOELs, *Integrated Environmental Assessment and Management*, 7(4):vi–viii.

Leduc J, Echaubard P, Trudeau V, and Lesbarrères D. 2016. Copper and Nickel Effects on Survival and Growth of Northern Leopard Frog (*Lithobates*

pipiens) Tadpoles in Field-Collected Smelting Effluent Water, *Environmental Toxicology and Chemistry*, 35(3):687–694.

Lenth RV. 2007. Statistical Power Calculations, *Journal of Animal Science*, 85(E. Suppl.):E24–E29.

Longcore JR, Longcore JE, Pessier AP, and Halteman WA. 2007. Chytridiomycosis Widespread in Anurans of Northeastern United States, *Journal of Wildlife Management*, 71(2):435–444.

Lutz I, Kloas W, Springer TA, Holden LR, Wolf JC, Krueger HO, and Hosmer AJ. 2008. Development, Standardization and Refinement of Procedures for Evaluating Effects of Endocrine Active Compounds on Development and Sexual Differentiation of *Xenopus laevis*, *Analytical and Bioanalytical Chemistry*, 390(8):2031–2048.

Mackenzie CA, Berrill M, Metcalfe C, and Pauli BD. 2003. Gonadal Differentiation in Frogs Exposed to Estrogenic and Antiestrogenic Compounds, *Environmental Toxicology and Chemistry*, 22(10):2466–2475.

Mann RM, Bidwell JR, and Tyler MJ. 2003. Toxicity of Herbicide Formulations to Frogs and the Implications for Product Registration: A Case Study from Western Australia, *Applied Herpetology*, 1(1–2):13–22.

Mann RM, Hyne RV, Choung CB, and Wilson SP. 2009. Amphibians and Agricultural Chemicals: Review of the Risks in a Complex Environment, *Environmental Pollution*, 157(11):2903–2927.

Marco A, and Blaustein AR. 1999. The Effects of Nitrite on Behavior and Metamorphosis in Cascades Frogs (*Rana cascadae*), *Environmental Toxicology and Chemistry*, 18(5):946–949.

Marlatt VL, Veldhoen N, Lo BP, Bakker D, Rehaume V, Vallée K, Haberl M, Shang D, van Aggelen GC, Skirrow RC, Elphick JR, and Helbing CC. 2013. Triclosan Exposure Alters Postembryonic Development in a Pacific Tree Frog (*Pseudacris regilla*) Amphibian Metamorphosis Assay (TREEMA), *Aquatic Toxicology*, 126:85–94.

Marquis O, Millery A, Guittonneau S, and Miaud C. 2006. Solvent toxicity to amphibian embryos and larvae, *Chemosphere*, 63(5):889–892.

Martini F, Tarazona JV, and Pablos MV. 2012. Are Fish and Standardized FETAX Assays Protective Enough for Amphibians? A Case Study on *Xenopus laevis* Larvae Assay with Biologically Active Substances Present in Livestock Wastes, *The Scientific World Journal*, 2012:605804.

Martinko C, and Van der Vliet L. 2021. Incorporating the 3Rs Into the Development of a Standardized Test Method Using a Native Amphibian Species, SETAC North America – 42nd Annual Meeting, Portland, Oregon.

Mattison C. 1993. *Keeping and Breeding Amphibians: Caecilians, Newts, Salamanders, Frogs and Toads*. Blandford Publishing, London, England. 224 p.

McCallum ML. 2007. Amphibian Decline or Extinction? Current Declines Dwarf Background Extinction Rate, *Journal of Herpetology*, 41(3):483–491.

McDaniel TV, Martin PA, Ross N, Brown S, Lesage S, and Pauli BD. 2004. Effects of Chlorinated Solvents on Four Species of North American Amphibians, *Archives of Environmental Contamination and Toxicology*, 47(1):101–109.

McDiarmid RW, and Altig R. 1999. *Tadpoles: The Biology of Anuran Larvae*, University of Chicago Press, Chicago, Illinois. 458 p.

Melvin SD, and Houlahan JE. 2012. Tadpole Mortality Varies Across Experimental Venues: Do Laboratory Populations Predict Responses in Nature?, *Oecologia*, 169(4):861–868.

Melvin SD, and Trudeau VL. 2012a. Growth, Development and Incidence of Deformities in Amphibian Larvae Exposed as Embryos to Naphthenic Acid Concentrations Detected in the Canadian Oil Sands Region, *Environmental Pollution*, 167:178–183.

Melvin SD, and Trudeau VL. 2012b. Toxicity of Naphthenic Acids to Wood Frog Tadpoles (*Lithobates sylvaticus*), *Journal of Toxicology and Environmental Health, Part A*, 75(3):170–173.

Melvin SD, Lanctôt CM, Craig PM, Moon TW, Peru KM, Headley JV, and Trudeau VL. 2013. Effects of Naphthenic Acid Exposure on Development and Liver Metabolic Processes in Anuran Tadpoles, *Environmental Pollution*, 177:22–27.

Melvin SD, Habener LJ, Leusch FDL, and Carroll AR. 2017. ^{1}H NMR-Based Metabolomics Reveals Sub-Lethal Toxicity of a Mixture of Diabetic and Lipid-Regulating Pharmaceuticals on Amphibian Larvae, *Aquatic Toxicology*, 184:123–132.

Michaels CJ, Downie JR, and Campbell-Palmer R. 2014. The Importance of Enrichment for Advancing Amphibian Welfare and Conservation Goals: A Review of a Neglected Topic, *Amphibian & Reptile Conservation (General Section)*, 8(1):7–23.

Milotic D, Milotic M, and Koprivnikar J. 2017. Effects of Road Salt on Larval Amphibian Susceptibility to Parasitism Through Behavior and Immunocompetence, *Aquatic Toxicology*, 189:42–49.

Milotic M, Milotic D, and Koprivnikar J. 2018. Exposure to a Cyanobacterial Toxin Increases Larval Amphibian Susceptibility to Parasitism, *Parasitology Research*, 117(2):513–520.

Mizell S. 1964. Seasonal Differences in Spermatogenesis and Oogenesis in *Rana pipiens*, *Nature*, 202:875–876.

Morgan KN, and Tromborg CT. (2007) Sources of Stress in Captivity, *Applied Animal Behaviour Science*, 102(3–4):262–302.

Mouchet F, Landois P, Datsyuk V, Puech P, Pinelli E, Flahaut E, and Gauthier L. 2011. International Amphibian Micronucleus Standardized Procedure (ISO 21427-1) for *In Vivo* Evaluation of Double-Walled Carbon Nanotubes Toxicity and Genotoxicity in Water, *Environmental Toxicology*, 26(2):136–145.

Nautilus Environmental. 2014. Amphibian Test Method Development: Review and Recommendations, Report prepared for Method Development and Applications Unit, Environment Canada, Ottawa, Ontario. 95 p.

Nautilus Environmental. 2016. Amphibian Test Method Development: Fall 2015-Spring 2016, Report prepared for Method Development and Applications Unit, Environment Canada, Ottawa, Ontario. 42 p.

Nautilus Environmental. 2020a. Amphibian Test Method Development using *Lithobates pipiens*: Inter-Laboratory Study: Round 1 and 2, Report prepared for Method Development and Applications Unit, Environment and Climate Change Canada, Ottawa, Ontario. 91 p.

Nautilus Environmental. 2020b. Amphibian Test Method Development using *Lithobates pipiens*: Inter-Laboratory Study: Round 3, Report prepared for Method Development and Applications Unit, Environment and Climate Change Canada, Ottawa, Ontario. 180 p.

Navarro-Martín L, Lanctôt C, Jackman P, Park BJ, Doe K, Pauli BD, and Trudeau VL. 2014. Effects of Glyphosate-Based Herbicides on Survival, Development, Growth and Sex Ratios of Wood Frogs (*Lithobates sylvaticus*) Tadpoles. I: Chronic Laboratory Exposures to VisionMax®, *Aquatic Toxicology*, 154:278–290.

NERI (National Environmental Research Institute). 1993. Manual of SECOFASE - Development, Improvement and Standardization of a Test System for Assessing Sublethal Effects of Chemicals on Fauna in the Soil Ecosystem, Report from a Workshop held in Silkeborg January 18-19, 1993, Silkeborg, Denmark. 41 p.

Newman MC. 2008. “What Exactly are You Inferring?” A Closer Look at Hypothesis Testing, *Environmental Toxicology and Chemistry*, 27:1013–1019.

Nolan MW, and Smith SA. 2007. Amphibian Resources on the Internet, *ILAR Journal*, 48(3):290–296.

Norberg-King TJ. 1993. A Linear Interpolation Method for Sublethal Toxicity: The Inhibition Concentration (ICp) Approach (Version 2.0), National Effluent Toxicity Assessment Center, Environmental Research Laboratory, Duluth, Minnesota, Technical Report 03-93. 25 p.

Norberg-King TJ, Embry MR, Belanger SE, Braunbeck T, Butler JD, Dorn PB, Farr B, Guiney PD, Hughes SA, Jeffries M, Journel R, Léonard M, McMaster M, Oris JT, Ryder K, Segner H, Senac T, Van Der Kraak G, Whale G, and Wilson P. 2018. An International Perspective on the Tools and Concepts for Effluent Toxicity Assessments in the Context of Animal Alternatives: Reduction in Vertebrate Use, *Environmental Toxicology and Chemistry*, 37(11):2745–2757.

O'Donnell RP, and Mock KE. 2012. Two Frog Species or One? A Multi-Marker Approach to Assessing the Distinctiveness of Genetic Lineages in the Northern Leopard Frog, *Rana pipiens*, *Conservation Genetics*, 13(5):1167–1182.

O'Donnell RP, Drost CA, and Mock KE. 2017. Cryptic Invasion of Northern Leopard Frogs (*Rana pipiens*) Across Phylogeographic Boundaries and a Dilemma for Conservation of a Declining Amphibian, *Biological Invasions*, 19:1039–1052.

Odum RA, and Zippel KC. 2008. Amphibian Water Quality: Approaches to an Essential Environmental Parameter, *International Zoo Yearbook*, 42(1):40–52.

Odum RA, and Zippel K. 2011. *Water Quality*, Amphibian Ark. 30 p.

OECD (Organisation for Economic Co-Operation and Development). 2007. Guidance Document on Amphibian Thyroid Histology, Series on Testing and Assessment, No. 82, OECD Publishing, Paris, France. 77 p.

OECD. 2009. OECD Guideline for the Testing of Chemicals, The Amphibian Metamorphosis Assay, No. 231, OECD Publishing, Paris, France. 33 p.

OECD. 2019. Guidance Document on Aqueous-Phase Aquatic Toxicity Testing of Difficult Test Chemicals, 2nd edition, Series on Testing and Assessment, No. 23, OECD Publishing, Paris, France. 81 p.

Ontario Nature. 2016. Northern Leopard Frog (*Lithobates pipiens*). URL: https://ontarionature.org/programs/community-science/reptile-amphibian-atlas/northern-leopard-frog/. Last accessed on July 19, 2021.

Ortiz-Santaliestra ME, and Brühl CA. 2014. Identification of Endpoints Useful to Characterize the Impact of Pesticides on Amphibian Aquatic Stages, SETAC – 24th European Annual Meeting, Basel, Switzerland.

Ortiz-Santaliestra ME, Maia JP, Egea-Serrano A, Brühl CA, and Lopes I. 2017. Biological Relevance of the Magnitude of Effects (Considering Mortality, Sub-Lethal and Reproductive Effects) Observed in Studies with Amphibians and Reptiles in View of Population Level Impacts on Amphibians and Reptiles, EFSA Supporting Publications, 14(7):1251E. 151 p.

Ortiz-Santaliestra ME, Maia JP, Egea-Serrano A, and Lopes I. 2018. Validity of Fish, Birds and Mammals as Surrogates for Amphibians and Reptiles in Pesticide Toxicity Assessment, *Ecotoxicology*, 27(7):819–833.

Orton F, Carr JA, and Handy RD. 2006. Effects of Nitrate and Atrazine on Larval Development and Sexual Differentiation in the Northern Leopard Frog *Rana pipiens*, *Environmental Toxicology and Chemistry*, 25(1):65–71.

Paden NE, Smith EE, Maul JD, and Kendall RJ. 2011. Effects of Chronic 2,4,6,-Trinitrotoluene, 2,4-Dinitrotoluene, and 2,6-Dinitrotoluene Exposure on Developing Bullfrog (*Rana catesbeiana*) Tadpoles, *Ecotoxicology and Environmental Safety*, 74(4):924–928.

Pauli BD, Perrault JA, and Money SL. 2000. *RATL: A Database of Reptile and Amphibian Toxicology Literature*, Environment Canada, Hull, Quebec, Canadian Wildlife Service Technical Report Series Number 357. 494 p.

Pauly GB, Hillis DM, and Cannatella DC. 2009. Taxonomic Freedom and the Role of Official Lists of Species Names, *Herpetologica*, 65(2):115–128.

Poole VA, and Grow S (eds.). 2012. *Amphibian Husbandry Resource Guide, Edition 2.0*, Association of Zoos and Aquariums, Silver Spring, Maryland. 238 p.

Pough FH. 2007. Amphibian Biology and Husbandry, *ILAR Journal*, 48(3):203–213.

Pillard DA, Eck WS, Johnson MS, and Packard S. 2017. Effects of 3-Nitro-1,2,4-triazol-5-one on Survival, Growth and Metamorphosis in the Northern Leopard Frog, *Lithobates pipiens*, *Ecotoxicology*, 26(9):1170-1180.

Relyea RA. 2004a. Growth and Survival of Five Amphibian Species Exposed to Combinations of Pesticides, *Environmental Toxicology and Chemistry*, 23(7):1737–1742.

Relyea RA. 2004b. Synergistic Impacts of Malathion and Predatory Stress on Six Species of North American Tadpoles, *Environmental Toxicology and Chemistry*, 23(4):1080–1084.

Relyea RA. 2005a. The Lethal Impacts of Roundup and Predatory Stress on Six Species of North American Tadpoles, *Archives of Environmental Contamination and Toxicology*, 48(3):351–357.

Relyea RA. 2005b. The Lethal Impact of Roundup on Aquatic and Terrestrial Amphibians, *Ecological Applications*, 15(4):1118–1124.

Robinson SA, Richardson SD, Dalton RL, Maisonneuve F, Trudeau VL, Pauli BD, and Lee-Jenkins SSY. 2017. Sublethal Effects on Wood Frogs Chronically Exposed to Environmentally Relevant Concentrations of Two Neonicotinoid Insecticides, *Environmental Toxicology and Chemistry*, 36(4):1101-1109.

Robinson SA, Gavel MJ, Richardson SD, Chlebak RJ, Milotic D, Koprivnikar J, and Forbes MR. 2019a. Sub-Chronic Exposure to a Neonicotinoid Does Not Affect Susceptibility of Larval Leopard Frogs to Infection by Trematode Parasites, via Either Depressed Cercarial Performance or Host Immunity, *Parasitology Research*, 118(9):2621–2633.

Robinson SA, Richardson SD, Dalton RL, Maisonneuve F, Bartlett AJ, de Solla SR, Trudeau VL, and Waltho N. 2019b. Assessment of Sublethal Effects of Neonicotinoid Insecticides on the Life-History Traits of 2 Frog Species, *Environmental Toxicology and Chemistry*, 38(9):1967-1977.

Robinson SA, Young SD, Brinovcar C, McFee A, and De Silva AO. 2020. Ecotoxicity Assessment and Bioconcentration of a Highly Brominated Organophosphate Ester Flame Retardant in Two Amphibian Species, *Chemosphere*, 260:127631.

Robinson SA, Chlebak RJ, Young SD, Dalton RL, Gavel MJ, Prosser RS, Bartlett AJ, and de Solla SR. 2021. Clothianidin Alters Leukocyte Profiles and Elevates Measures of Oxidative Stress in Tadpoles of the Amphibian, *Rana pipiens*, *Environmental Pollution*, 284:117149.

Rocchini RJ, Servizi JA, Watts RG, Horvath S, Clark MJR, Singleton HJ, Young RH, Sholund A, Jordan AJ, and McLeay DJ. 1982. Provincial Guidelines and Laboratory Procedures for Measuring Acute Lethal Toxicity of Liquid Effluents to Fish, British Columbia Ministry of Environment, Victoria, British Columbia. 25 p.

Rowe CL and Freda J. 2000. Effects of Acidification on Amphibians at Multiple Levels of Biological Organization, In: *Ecotoxicology of Amphibians and Reptiles*, Sparling DW, Linder G, and Bishop CA (eds.), Society of Environmental Toxicology and Chemistry, Pensacola, Florida, pp. 545–571.

Russell WMS, and Burch RL. 1959. *The Principles of Humane Experimental Technique*, Methuen & Co. Ltd., London, England.

Sager JC and McFarlane C. 1997. Radiation, In: Plant Growth Chamber Handbook, Langhans RW and Tibbitts TW (eds.), North Central Regional Research Publication No. 340, Iowa Agriculture and Home Economics Experiment Station Special Report No. 90, Iowa State University of Science and Technology, Ames, Iowa, pp. 1–30.

Sanzo D, and Hecnar SJ. 2006. Effects of Road De-Icing Salt (NaCl) on Larval Wood Frogs (*Rana sylvatica*), *Environmental Pollution*, 140(2):247–256.

Saskatchewan Prairie Conservation Action Plan. 2018. Guide to Managing for Optimal Habitat Attributes: Northern Leopard Frog (*Lithobates pipiens* – Western Boreal/Prairie Populations). 33 p.

Schlesinger MD, Feinberg JA, Nazdrowicz NH, Kleopfer JD, Beane JC, Bunnell JF, Burger J, Corey E, Gipe K, Jaycox JW, Kiviat E, Kubel J, Quinn DP, Raithel C, Scott PA, Wenner SM, White EL, Zarate B, and Shaffer HB. 2018. Follow-Up Ecological Studies for Cryptic Species Discoveries: Decrypting the Leopard Frogs of the Eastern U.S., *PLoS ONE*, 13(11):e0205805.

Schlichter LC. 1981. Low pH Affects the Fertilization and Development of *Rana pipiens* Eggs, *Canadian Journal of Zoology*, 59(9):1693–1699.

Scholz S, Sela E, Blaha L, Braunbeck T, Galay-Burgos M, García-Franco M, Guinea J, Klüver N, Schirmer K, Tanneberger K, Tobor-Kapłon M, Witters H, Belanger S, Benfenati E, Creton S, Cronin MTD, Eggen RIL, Embry M, Ekman D, Gourmelon A, Halder M, Hardy B, Hartung T, Hubesch B, Jungmann D, Lampi MA, Lee L, Léonard M, Küster E, Lillicrap A, Lukenbach T, Murk AJ, Navas JM, Peijnenburg W, Repetto G, Salinas E, Schüürmann G, Spielmann H, Tollefsen KE, Walter-Rhode S, Whale G, Wheeler JR, and Winter MJ. 2013. A European Perspective on Alternatives to Animal Testing for Environmental Hazard Identification and Risk Assessment, *Regulatory Toxicology and Pharmacology*, 67(3):506–530.

Shenoy K, Cunningham BT, Renfroe JW, and Crowley PH. 2009. Growth and Survival of Northern Leopard Frog (*Rana pipiens*) Tadpoles Exposed to Two Common Pesticides, *Environmental Toxicology and Chemistry*, 28(7):1469–1474.

Shumway W. 1940. Stages in the Normal Development of *Rana pipiens*, I. External Form, *The Anatomical Record*, 78(2):139–147.

Sievers M, Hale R, Parris KM, Melvin SD, Lanctôt CM, and Swearer SE. 2019. Contaminant-Induced Behavioural Changes in Amphibians: A Meta-Analysis, *Science of the Total Environment*, 693:133570.

Skerratt LF, Berger L, Speare R, Cashins S, McDonald KR, Phillott AD, Hines HB, and Kenyon N. 2007. Spread of Chytridiomycosis Has Caused the Rapid Global Decline and Extinction of Frogs, *EcoHealth*, 4:125–134.

Smith SA. 2007. Compendium of Drugs and Compounds Used in Amphibians, *ILAR Journal*, 48:297–300.

Smith SA, and Stoskopf MK. 2007. The Art of Amphibian Science, *ILAR Journal*, 48(3):179–182.

Smits JEG, Hersikorn BD, Young RF, and Fedorak PM. 2012. Physiological Effects and Tissue Residues from Exposure of Leopard Frogs to Commercial Naphthenic Acids, *Science of the Total Environment*, 437:36–41.

Sobotka JM, and Rahwan RG. 1994. Lethal Effect of Latex Gloves on *Xenopus laevis* Tadpoles, *Journal of Pharmacological and Toxicological Methods*, 32(1):59.

Soufan O, Ewald J, Zhou G, Hacariz O, Boulanger E, Alcaraz AJ, Hickey G, Maguire S, Pain G, Hogan N, Hecker M, Crump D, Head J, Basu N, and Xia J. 2022. EcoToxXplorer: Leveraging Design Thinking to Develop a Standardized Web-Based Transcriptomics Analytics Platform for Diverse Users, *Environmental Toxicology and Chemistry*, 41(1):21-29.

Sowers AD, Mills MA, and Klaine SJ. 2009. The Developmental Effects of a Municipal Wastewater Effluent on the Northern Leopard Frog, *Rana pipiens*, *Aquatic Toxicology*, 94:145–152.

Sparling DW, Linder G, and Bishop CA (eds.). 2000. *Ecotoxicology of Amphibians and Reptiles*,

Society of Environmental Toxicology and Chemistry, Pensacola, Florida. 904 p.

Sparling DW, and Harvey G. 2006. Comparative Toxicity of Ammonium and Perchlorate to Amphibians, *Bulletin of Environmental Contamination and Toxicology*, 76(2):210–217.

Sparling DW, Linder G, Bishop CA, and Krest SK (eds.). 2010. *Ecotoxicology of Amphibians and Reptiles, 2nd ed.*, CRC Press, Pensacola, Florida. 944 p.

Species at Risk Committee. 2013. Species Status Report for Northern Leopard Frog (*Lithobates pipiens*) in the Northwest Territories, Species at Risk Committee, Department of Environment and Natural Resources, Yellowknife, Northwest Territories. 66 p.

Sprague JB. 1973. The ABC's of Pollutant Bioassay Using Fish, In: *Biological Methods for the Assessment of Water Quality*, Cairns J Jr., and Dickson K (eds.), American Society for Testing and Materials, Philadelphia, Pennsylvania, Special Technical Publication 528, pp. 6–30.

Stanley JK, Lotufo GR, Biedenbach JM, Chappell P, and Gust KA. 2015. Toxicity of Conventional Energetics TNT and RDX Relative to New Insensitive Munitions Constituents DNAN and NTO in *Rana pipiens* Tadpoles, *Environmental Toxicology and Chemistry*, 34(4):873–879.

Stebbins RC. 2003. *A Field Guide to Western Reptiles and Amphibians, 3rd ed.*, Houghton Mifflin Harcourt, New York, New York. 560 p.

Stephenson GL, Koper N, Atkinson GF, Solomon KR, and Scroggins RP. 2000. Use of Nonlinear Regression Techniques for Describing Concentration-Response Relationships of Plant Species Exposed to Contaminated Site Soils, *Environmental Toxicology and Chemistry*, 19(12):2968–2981.

Storrs Méndez SI, Tillitt DE, Rittenhouse TAG, and Semlitsch RD. 2009. Behavioral Response and Kinetics of Terrestrial Atrazine Exposure in American Toads (*Bufo americanus*), *Archives of Environmental Contamination and Toxicology*, 57(3):590–597.

Stuart SN, Chanson JS, Cox NA, Young BE, Rodrigues ASL, Fischman DL, and Waller RW. 2004. Status and Trends of Amphibian Declines and Extinctions Worldwide, *Science*, 306(5702):1783–1786.

Thursby GB, Heltshe J, and Scott KJ. 1997. Revised Approach to Toxicity Test Acceptability Criteria Using a Statistical Performance Assessment, *Environmental Toxicology and Chemistry*, 16:1322–1329.

Timmons MB, Guerdat T, and Vinci BJ. 2018. *Recirculating Aquaculture, Fourth Edition*, Ithaca Publishing Company LLC, New York.

Tompsett AR, Wiseman S, Higley E, Giesy JP, and Hecker M. 2013. Effects of Exposure to 17α-Ethynylestradiol During Larval Development on Growth, Sexual Differentiation, and Abundances of Transcripts in the Liver of the Wood Frog (*Lithobates sylvaticus*), *Aquatic Toxicology*, 126:42–51.

Torreilles SL, McClure DE, and Green SL. 2009. Evaluation and Refinement of Euthanasia Methods for *Xenopus laevis*, *Journal of the American Association for Laboratory Animal Science*, 48(6):512–516.

TOXSTAT. 1996. Version 3.5, Lincoln Research Associates Inc., PO Box 4276, Bisbee, Arizona 85603, USA, Phone 520-432-4092, email danlra@msn.com. [Programs on disk, with printed user's manual].

Trudeau VL, Somoza GM, Natale GS, Pauli B, Wignall J, Jackman P, Doe K, and Schueler FW. 2010. Hormonal Induction of Spawning in 4 Species of Frogs by Coinjection with a Gonadotropin-Releasing Hormone Agonist and a Dopamine Antagonist, *Reproductive Biology and Endocrinology*, 8(36):1–9.

Trudeau VL, Schueler FW, Navarro-Martin L, Hamilton CK, Bulaeva E, Bennett A, Fletcher W, and Taylor L. 2013. Efficient Induction of Spawning of Northern Leopard Frogs (*Lithobates*

pipiens) During and Outside the Natural Breeding System, *Reproductive Biology and Endocrinology*, 11(14):1–9.

Trudeau VL, Thomson P, Zhang WS, Reynaud S, Navarro-Martin L, and Langlois VS. 2020. Agrochemicals Disrupt Multiple Endocrine Axes in Amphibians, *Molecular and Cellular Endocrinology*, 513:110861.

Ultsch GR, Bradford DF, and Freda J. 1999. Physiology: Coping with the Environment, In: *Tadpoles: The Biology of Anuran Larvae*, McDiarmid RW, and Altig R (eds.), University of Chicago Press, Chicago, Illinois, pp. 189–214.

US Fish and Wildlife Service. 2011. Endangered and Threatened Wildlife and Plants; 12-Month Finding on a Petition to List the Northern Leopard Frog in the Western United States as Threatened; Proposed Rule, 50 CFR Part 17 [Docket No. FWS-R2-ES-2009-0030; 92210-1111-FY08-B2], Federal Register, 76(193):61896–61931.

USEPA (United States Environmental Protection Agency). 1975. Methods for Acute Toxicity Tests with Fish, Macroinvertebrates, and Amphibians, United States Environmental Protection Agency, National Environmental Research Centre, Office of Research and Development, Corvallis, Oregon, EPA-660/3-75-009. 69 p.

USEPA. 1996. Ecological Effects Test Guidelines OPPTS 850.1800 Tadpole/Sediment Subchronic Toxicity Test, United States Environmental Protection Agency, Prevention, Pesticides and Toxic Substances, Washington, DC, EPA 712-C-96-132. 15 p.

USEPA. 2002. Short-Term Methods for Estimating the Chronic Toxicity of Effluents and Receiving Waters to Freshwater Organisms, United States Environmental Protection Agency, Office of Water, Washington, DC, EPA-821-R-02-013 4th ed. 350 p.

USEPA. 2009. Endocrine Disruptor Screening Program Test Guidelines OPPTS 890.1100: Amphibian Metamorphosis (Frog), United States Environmental Protection Agency, Prevention Pesticides and Toxic Substances, Washington, DC, EPA 740-C-09-002. 50 p.

USEPA. 2015. Endocrine Disruptor Screening Program Test Guideline OCSPP 890.2300: Larval Amphibian Growth and Development Assay (LAGDA), United States Environmental Protection Agency, Office of Chemical Safety and Pollution Prevention, Washington, DC, EPA Pub. No. 740-C-15-001. 89 p.

van Dam RA, Harford AJ, and Warne MSJ. 2012. Time to Get Off the Fence: The Need for Definitive International Guidance on Statistical Analysis of Ecotoxicity Data, *Integrated Environmental Assessment and Management*, 8(2):242–245.

Van der Hoeven N. 1998. Power Analysis for the NOEC: What is the Probability of Detecting Small Toxic Effects on Three Different Species using the Appropriate Standardized Test Protocols?, *Ecotoxicology*, 7:355–361.

Van der Vliet L, Taylor LN, and Scroggins R. 2012. NOEC: Notable Oversight of Enlightened Canadians: A Response to van Dam *et al.* (2012), *Integrated Environmental Assessment and Management*, 8(3):397–398.

Van Schmidt ND, Cary TL, Ortiz-Santaliestra ME, and Karasov WH. 2012. Effects of Chronic Polybrominated Diphenyl Ether Exposure on Gonadal Development in the Northern Leopard Frog, *Rana pipiens*, *Environmental Toxicology and Chemistry*, 31(2):347–354.

Veldhoen N, Skirrow RC, Osachoff H, Wigmore H, Clapson DJ, Gunderson MP, van Aggelen G, and Helbing CC. 2006. The Bactericidal Agent Triclosan Modulates Thyroid Hormone-Associated Gene Expression and Disrupts Postembryonic Anuran Development, *Aquatic Toxicology*, 80(3):217–227.

Veldhoen N, Skirrow RC, Brown LLY, van Aggelen G, and Helbing CC. 2014. Effects of Acute Exposure to the Non-steroidal Anti-inflammatory Drug Ibuprofen on the Developing North American Bullfrog (*Rana catesbeiana*) Tadpoles,

Environmental Science and Technology, 48(17):10439–10447.

Voordouw MJ, Adama D, Houston B, Govindarajulu P, and Robinson J. 2010. Prevalence of the Pathogenic Chytrid Fungus, *Batrachochytrium dendrobatidis*, in an Endangered Population of Northern Leopard Frogs, *Rana pipiens*, *BMC Ecology*, 10(6):1–10.

Vu M, Weiler B, and Trudeau VL. 2017. Time- and Dose-Related Effects of a Gonadotropin-Releasing Hormone Agonist and Dopamine Antagonist on Reproduction in the Northern Leopard Frog (*Lithobates pipiens*), *General and Comparative Endocrinology*, 254:86–96.

Wagner N, Reichenbecher W, Teichmann H, Tappeser B, and Lötters S. 2013. Questions Concerning the Potential Impact of Glyphosate-Based Herbicides on Amphibians, *Environmental Toxicology and Chemistry*, 32(8):1688–1700.

Wake DB, and Vredenburg VT. 2008. Are We in the Midst of the Sixth Mass Extinction? A View from the World of Amphibians, *Proceedings of the National Academy of Sciences of the United States of America*, 105:11466–11473.

Weir SM, Yu S, and Salice CJ. 2012. Acute Toxicity of Herbicide Formulations and Chronic Toxicity of Technical-Grade Trifluralin to Larval Green Frogs (*Lithobates clamitans*), *Environmental Toxicology and Chemistry*, 31(9):2029–2034.

Weltje L, Simpson P, Gross M, Crane M, and Wheeler JR. 2013. Comparative Acute and Chronic Sensitivity of Fish and Amphibians: A Critical Review of Data, *Environmental Toxicology and Chemistry*, 32(5):984–994.

Whitaker BR. 2001. Water Quality. In: *Amphibian Medicine and Captive Husbandry*, Wright KM, and Whitaker BR (eds.), Kreiger Publishing Company, Malabar, Florida, pp. 147–157.

Williams BK, and Semlitsch RD. 2010. Larval Responses of Three Midwestern Anurans to Chronic, Low-Dose Exposures of Four Herbicides, *Archives of Environmental Contamination and Toxicology*, 58(3):819–827.

Wilson GA, Fulton TL, Kendell K, Scrimgeour G, Paszkowski CA, and Coltman DW. 2008. Genetic Diversity and Structure in Canadian Northern Leopard Frog (*Rana pipiens*) Populations: Implications for Reintroduction Programs, *Canadian Journal of Zoology*, 86(8):863–874.

Wojnarowicz P, Ogunlaja OO, Xia C, Parker WJ, and Helbing CC. 2013. Impact of Wastewater Treatment Configuration and Seasonal Conditions on Thyroid Hormone Disruption and Stress Effects in *Rana catesbeiana* Tailfin, *Environmental Science and Technology*, 47(23):13840–13847.

Wright KM, and Whitaker BR (eds.). 2001. *Amphibian Medicine and Captive Husbandry*, Krieger Publishing Company, Malabar, Florida. 499 p.

Yahnke AE, Grue CE, Hayes MP, and Troiano AT. 2013. Effects of the Herbicide Imazapyr on Juvenile Oregon Spotted Frogs, *Environmental Toxicology and Chemistry*, 32(1):228–235.

Yee SG, McLeay DJ, and Fennel M. 1996. Recent Laboratory Studies Related to Improving Environment Canada's Biological Test Method EPS 1/RM/28 Using Rainbow Trout Embryos ("E" Toxicity-test Option)", Report prepared by McLeay Environmental Ltd. for Environment Canada, Ottawa, Ontario.

Young SD, Gavel MJ, Gutierrez-Villagomez JM, Forbes MR, and Robinson SA. 2019. Assessment of Sublethal Ecotoxicity of Solvents on Larvae of a Model Native Amphibian (*Lithobates pipiens*), *Journal of Applied Toxicology*, 40(4):483–492.

Yu S, Wages MR, Cai Q, Maul JD, and Cobb GP. 2013. Lethal and Sublethal Effects of Three Insecticides on Two Developmental Stages of *Xenopus laevis* and Comparison with Other Amphibians, *Environmental Toxicology and Chemistry*, 32(9):2056–2064.

Appendix A

Biological Test Methods and Supporting Guidance Documents Published by Environment and Climate Change Canada's Method Development and Applications Unit[a]

Title of Biological Test Method or Guidance Document	Report Number	Publication Date	Applicable Amendments
A. Generic (Universal) Biological Test Methods			
Acute Lethality Test Using Rainbow Trout	EPS 1/RM/9	July 1990	May 1996, May 2007
Acute Lethality Test Using *Daphnia* spp.	EPS 1/RM/11	July 1990	May 1996
Test of Reproduction and Survival Using the Cladoceran *Ceriodaphnia dubia*	EPS 1/RM/21 2nd Edition	February 2007	–
Test of Larval Growth and Survival Using Fathead Minnows	EPS 1/RM/22 2nd Edition	February 2011	–
Toxicity Test Using Luminescent Bacteria	EPS 1/RM/24	November 1992	–
Growth Inhibition Test Using a Freshwater Alga	EPS 1/RM/25 2nd Edition	March 2007	–
Acute Test for Sediment Toxicity Using Marine or Estuarine Amphipods	EPS 1/RM/26	December 1992	October 1998
Fertilization Assay Using Echinoids (Sea Urchins and Sand Dollars)	EPS 1/RM/27 2nd Edition	February 2011	–
Toxicity Tests Using Early Life Stages of Salmonid Fish (Rainbow Trout)	EPS 1/RM/28 2nd Edition	July 1998	–
Test for Survival and Growth in Sediment Using the Larvae of Freshwater Midges (*Chironomus tentans* or *Chironomus riparius*)	EPS 1/RM/32	December 1997	–
Test for Survival, Growth, and Reproduction in Sediment and Water Using the Freshwater Amphipod *Hyalella azteca*	EPS 1/RM/33 3rd Edition	September 2017	–
Test for Measuring the Inhibition of Growth Using the Freshwater Macrophyte, *Lemna minor*	EPS 1/RM/37 2nd Edition	January 2007	–

[a] These documents are available for purchase from the Publication Catalogue, Environment and Climate Change Canada, Ottawa ON K1A 0H3, Canada. Printed copies can also be requested by email from methods@ec.gc.ca. These documents are available free of charge in PDF format at the following website: www.canada.ca/en/environment-climate-change/services/wildlife-research-landscape-science/biological-test-method-publications.html. For further information or comments, contact the Manager, Method Development and Applications Unit, Environment and Climate Change Canada, Ottawa ON K1A 0H3.

Title of Biological Test Method or Guidance Document	**Report Number**	**Publication Date**	**Applicable Amendments**
A. Generic (Universal) Biological Test Methods (continued)			
Test for Survival and Growth in Sediment Using Spionid Polychaete Worms (*Polydora cornuta*)	EPS 1/RM/41	December 2001	–
Tests for Measuring Avoidance Behaviour or Reproduction of Earthworms (*Eisenia andrei* or *Dendrodrilus rubidus*) Exposed to Contaminants in Soil	STB 1/RM/43 2nd Edition	August 2022	–
Tests for Measuring Emergence and Growth of Terrestrial Plants Exposed to Contaminants in Soil	EPS 1/RM/45	February 2005	June 2007
Test for Measuring Survival and Reproduction of Springtails Exposed to Contaminants in Soil	EPS 1/RM/47 2nd Edition	February 2014	–
Test for Growth in Contaminated Soil Using Terrestrial Plants Native to the Boreal Region	EPS 1/RM/56	August 2013	–
Test for Measuring Reproduction of Oribatid Mites Exposed to Contaminants in Soil	STB 1/RM/61	September 2020	–
Toxicity Tests Using Aquatic Life Stages of Frogs (*Lithobates pipiens*)	STB RM/62	April 2024	–
B. Reference Methods[b]			
Reference Method for Determining Acute Lethality Using Threespine Stickleback	EPS 1/RM/10 2nd Edition	December 2017	–
Reference Method for Determining Acute Lethality of Effluents to Rainbow Trout	EPS 1/RM/13 2nd Edition	December 2000	May 2007, February 2016, December 2023
Reference Method for Determining Acute Lethality of Effluents to *Daphnia magna*	EPS 1/RM/14 2nd Edition	December 2000	February 2016
Reference Method for Determining Acute Lethality of Sediment to Marine or Estuarine Amphipods	EPS 1/RM/35	December 1998	–
Reference Method for Determining the Toxicity of Sediment Using Luminescent Bacteria in a Solid-Phase Test	EPS 1/RM/42	April 2002	–
Reference Method for Measuring the Toxicity of Contaminated Sediment to Embryos and Larvae of Echinoids (Sea Urchins or Sand Dollars)	EPS 1/RM/58	July 2014	–
Reference Method for Determining Acute Lethality Using *Acartia tonsa*	STB 1/RM/60	June 2019	–

[b] For this series of documents, a reference method is defined as a specific biological test method for performing a toxicity test, i.e., a toxicity test method with an explicit set of test instructions and conditions that is described precisely in a written document. Unlike other generic (multipurpose or "universal") biological test methods published by Environment and Climate Change Canada, the use of a reference method is frequently restricted to testing requirements associated with specific regulations.

Title of Biological Test Method or Guidance Document	Report Number	Publication Date	Applicable Amendments
C. Supporting Guidance Documents			
Guidance Document on Control of Toxicity Test Precision Using Reference Toxicants	EPS 1/RM/12	August 1990	–
Guidance Document on Collection and Preparation of Sediment for Physicochemical Characterization and Biological Testing	EPS 1/RM/29	December 1994	–
Guidance Document on Measurement of Toxicity Test Precision Using Control Sediments Spiked with a Reference Toxicant	EPS 1/RM/30	September 1995	–
Recommended Procedure for the Importation of Test Organisms for Sublethal Toxicity Testing	–	September 1999	–
Guidance Document on Application and Interpretation of Single-Species Tests in Environmental Toxicology	EPS 1/RM/34	December 1999	–
Revised Procedures for Adjusting Salinity of Effluent Samples for Marine Sublethal Toxicity Conducted Under Environmental Effects Monitoring (EEM) Programs	–	December 2001	–
Guidance Document for Testing the Pathogenicity and Toxicity of New Microbial Substances to Aquatic and Terrestrial Organisms	EPS 1/RM/44 2nd Edition	December 2016	–
Guidance Document on Statistical Methods for Environmental Toxicity Tests	EPS 1/RM/46	March 2005	June 2007
Procedure for pH Stabilization During the Testing of Acute Lethality of Wastewater Effluent to Rainbow Trout	EPS 1/RM/50	March 2008	–
Supplementary Background and Guidance for Investigating Acute Lethality of Wastewater Effluent to Rainbow Trout	–	March 2008	–
Guidance Document on the Sampling and Preparation of Contaminated Soil for Use in Biological Testing	EPS 1/RM/53	February 2012	–
Procedure for pH Stabilization During the Testing of Acute Lethality of Pulp and Paper Effluent to Rainbow Trout	STB 1/RM/59	March 2018	–
Supplementary Guidance for Investigating Acute Lethality of Pulp and Paper Mill Effluents due to Ammonia	–	March 2018	–

Appendix B

Environment and Climate Change Canada, Regional Environmental Testing Laboratories

Atlantic Laboratory for Environmental Testing
Environmental Science Building
443 Université Avenue, Université de Moncton
Moncton, New Brunswick
E1A 3E9

Pacific and Yukon Laboratory for Environmental Testing
Pacific Environmental Science Centre
2645 Dollarton Hwy
North Vancouver, British Columbia
V7H 1B1

Québec Laboratory for Environmental Testing
105 McGill Street
Montréal, Quebec
H2Y 2E7

Prairie and Northern Laboratory for Environmental Testing
Northern Forestry Building
5320 122 St NW
Edmonton, Alberta
T6H 3S5

Soil Toxicology Laboratory
335 River Road
Ottawa, Ontario
K1A 0H3

For current regional laboratory contact information please contact:

Method Development and Applications Unit
Science and Technology Branch
Environment and Climate Change Canada
335 River Road
Ottawa, Ontario
K1A 0H3
Email: methods@ec.gc.ca

Appendix C

Members of the Inter-Governmental Ecotoxicological Testing Group (as of June 2023)

Federal, Environment and Climate Change Canada

Suzanne Agius
Marine Protection Programs Section
Gatineau, Quebec

Adrienne Bartlett
Aquatic Contaminants Research Division
Burlington, Ontario

Megan Bauer
Atlantic Laboratory for Environmental Testing
Moncton, New Brunswick

Lee Beaudette
Wildlife Toxicology Research Section
Ottawa, Ontario

Rene Beaulieu
Prairie & Northern Laboratory for Environmental Testing
Edmonton, Alberta

Patrick Boyd
Biological Assessment & Standardization Section
Ottawa, Ontario

Julia Brydon
Marine Protection Programs Section
Gatineau, Quebec

Craig Buday
Pacific & Yukon Laboratory for Environmental Testing
North Vancouver, British Columbia

Sheena Campbell
Aquatic Contaminants Research Division
Burlington, Ontario

Marshneil Chandra
Prairie & Northern Laboratory for Environmental Testing
Edmonton, Alberta

Ajith Dias Samarajeewa
Biological Assessment & Standardization Section
Ottawa, Ontario

Heather Dillon
Prairie & Northern Laboratory for Environmental Testing
Edmonton, Alberta

Ken Doe (Emeritus)
Atlantic Laboratory for Environmental Testing
Moncton, New Brunswick

Richard Frank
Aquatic Contaminants Research Division
Burlington, Ontario

Christopher Fraser
Priority Assessments Section
Ottawa, Ontario

François Gagné
Aquatic Contaminants Research Division
Montréal, Quebec

Patty Gillis
Aquatic Contaminants Research Division
Burlington, Ontario

Ève Gilroy
Aquatic Contaminants Research Division
Burlington, Ontario

Christina Heise
Prairie & Northern Laboratory for Environmental Testing
Edmonton, Alberta

Charles Hopper
Atlantic Laboratory for Environmental Testing
Moncton, New Brunswick

Natasha Hostal
Prairie & Northern Laboratory for Environmental Testing
Edmonton, Alberta

Paula Jackman (Emeritus)
Atlantic Laboratory for Environmental Testing
Moncton, New Brunswick

Heather Jovanovic
Aquatic Contaminants Research Division
Burlington, Ontario

Hufsa Khan
Aquatic Contaminants Research Division
Burlington, Ontario

Stephanie Kvas
Biological Assessment & Standardization Section
Ottawa, Ontario

Chris Le
Pacific & Yukon Laboratory for Environmental Testing
North Vancouver, British Columbia

Heather Lemieux
Biological Assessment & Standardization Section
Ottawa, Ontario

Michelle Linssen-Sauvé
Pacific & Yukon Laboratory for Environmental Testing
North Vancouver, British Columbia

Sue Ellen Maher
Marine Protection Programs Section
Ottawa, Ontario

Bill Martin
National Guidelines and Standards Office
Ottawa, Ontario

Carolyn Martinko (Co-chair)
Biological Assessment & Standardization Section
Ottawa, Ontario

Danielle Milani
Aquatic Contaminants Research Division
Burlington, Ontario

Alicia O'Neill
Aquatic Contaminants Research Division
Burlington, Ontario

Joanne Parrott
Aquatic Contaminants Research Division
Burlington, Ontario

Linda Porebski
Marine Protection Programs Section
Gatineau, Quebec

Juliska Princz
Biological Assessment & Standardization Section
Ottawa, Ontario

Rick Scroggins
Biological Assessment & Standardization Section
Ottawa, Ontario

David Taillefer
Marine Protection Programs Section
Gatineau, Quebec

Sylvain Trottier
Quebec Laboratory for Environmental Testing
Montréal, Quebec

Leana Van der Vliet
Biological Assessment & Standardization Section
Ottawa, Ontario

Jessica Velicogna
Biological Assessment & Standardization Section
Ottawa, Ontario

Brian Walker
Quebec Laboratory for Environmental Testing
Montréal, Quebec

Peter Wells (Emeritus)
Environmental Conservation Service
Dartmouth, Nova Scotia

Federal, Health Canada

Ellyn Ritchie
Pest Management Regulatory Agency,
Environmental Assessment Directorate
Gatineau, Quebec

Federal, Natural Resources Canada

Morgan King
CanmetMINING, Green Mining Innovation
Ottawa, Ontario

Carrie Rickwood
CanmetMINING, Green Mining Innovation
Ottawa, Ontario

Provincial

Melanie Appleton
Ontario Ministry of Environment, Conservation and Parks
Etobicoke, Ontario

Lisa Kennedy (Co-chair)
Ontario Ministry of Environment, Conservation and Parks
Etobicoke, Ontario

Heather Osachoff
British Columbia Ministry of the Environment and Climate Change Strategy
Victoria, British Columbia

David Poirier (Emeritus)
Ontario Ministry of Environment, Conservation and Parks
Etobicoke, Ontario

Kathleen Stevack
Ontario Ministry of Environment, Conservation and Parks
Etobicoke, Ontario

Éloïse Veilleux
Centre d'expertise en analyse environnementale du Québec
Ste. Foy, Quebec

Kirstin Webster
British Columbia Ministry of the Environment and Climate Change Strategy
Victoria, British Columbia

Appendix D

Method Writers and Members of the Expert Peer Review Committee

MDAU Method Coordinator and Contributors

Leana Van der Vliet
Method Coordinator
Biological Assessment and Standardization Section
Science and Technology Branch
Environment and Climate Change Canada
335 River Road
Ottawa, Ontario K1V 1C7

Rick Scroggins
Biological Assessment and Standardization Section
Science and Technology Branch
Environment and Climate Change Canada
335 River Road
Ottawa, Ontario K1V 1C7

Lisa Taylor (retired)
Biological Assessment and Standardization Section
Science and Technology Branch
Environment and Climate Change Canada
335 River Road
Ottawa, Ontario K1V 1C7

Carolyn Martinko
Biological Assessment and Standardization Section
Science and Technology Branch
Environment and Climate Change Canada
335 River Road
Ottawa, Ontario K1V 1C7

External Method Writer

Jennifer Miller
Miller Environmental Sciences Inc.
18600 Marsh Hill Road
Uxbridge, Ontario L9P 1R3

ECCC Technical Reviewer

Paula Jackman (retired)
Science and Technology Branch
Environment and Climate Change Canada
University of Moncton Campus
443 University Avenue
Moncton, New Brunswick E1A 3E9

Expert Peer Reviewers

Doug Fort
Fort Environmental Laboratories Inc.
515 South Duncan Street
Stillwater, Oklahoma 74074

Melanie Gallant
Nautilus Environmental Company Inc.
10823 27 Street SE
Calgary, Alberta T2Z 3V9

Elissa Liu
Chemicals Management Division
Environmental Protection Branch
Environment and Climate Change Canada
101-401 Burrard Street
Vancouver, British Columbia V6C 3R2

Bonnie Lo
Simon Fraser University
8888 University Dr W
Burnaby, BC V5A 1S6

Stacey Robinson
Wildlife Toxicology Research Section
Science and Technology Branch
Environment and Climate Change Canada
1125 Colonel By Drive
Ottawa, Ontario K1S 5B6

Appendix E

Additional Procedures for Culturing *Lithobates pipiens* in the Laboratory and General Procedures for Frog Husbandry

Many details on the source of adult frogs for spawning or hibernation, timing of collection and transportation to the laboratory, as well as guidance on holding, hibernation, breeding, and quarantine, are provided in Sections 2.1, 2.2, and 2.4. This appendix contains additional information related to the culturing of adult *Lithobates pipiens* in the laboratory and general guidance on their husbandry.

E.1 Culturing of Adult Frogs: Additional Guidance and Resources

E.1.1 Additional Acknowledgements and Resources

The guidance described in Section 2.4 and this appendix were developed at Environment and Climate Change Canada's Atlantic Laboratory for Environmental Testing (ALET), with input from Andrea Edgington (University of Guelph), Dr. Vance Trudeau (University of Ottawa), and the Vancouver Aquarium. Northern leopard frogs (*Lithobates pipiens*) have been successfully raised and held in the ALET laboratory throughout all life stages (i.e., from egg masses to tadpoles, metamorphs, and adults; ALET, 2004, 2006, 2009, 2013, 2018). These methods have also been successfully used to hold adult *Lithobates pipiens* at a private laboratory (Nautilus Environmental, 2016). The reader is directed to the following references for additional information on amphibian husbandry and breeding: "Amphibian Medicine and Captive Husbandry" (Wright and Whitaker, 2001) and "Amphibian Husbandry Resource Guide, Edition 2.0" (Poole and Grow, 2012), which provide useful information about diseases, pharmaceuticals, and hormone manipulation; "CCAC Guidelines: Amphibians" (CCAC, 2021); and the special edition of the Institute for Laboratory Animal Research Journal entitled "Use of Amphibians in the Research, Laboratory, or Classroom Setting" (Alworth and Harvey, 2007; Browne and Zippel, 2007; Browne *et al.*, 2007; Burggren and Warburton, 2007; Densmore and Green, 2007; Gentz, 2007; Hopkins, 2007; Nolan and Smith, 2007; O'Rourke, 2007; Pough, 2007; Smith, 2007; Smith and Stoskopf, 2007). Examples of adult frog holding containers are shown in Figures E.1 and E.2.

E.1.2 Culturing Food for Adult Frogs

Crickets purchased from a pet store can be held for up to 2 weeks. To obtain small crickets, place a substrate (i.e., container with a mixture of sand and soil) in the cricket-holding aquarium to allow the adults to lay eggs; keep the container covered for about 6 weeks until pinhead crickets have hatched. These young crickets can also be purchased from local suppliers. Crickets are fed commercial cricket food and oatmeal. To provide moisture, wet sponges, wet paper towel, a commercial cricket water gel, or vegetables (i.e., potatoes or carrots) are also provided. Egg cartons and/or paper towel cardboard tubes are placed in the tanks as hiding spaces. Frog feeding can be supplemented with newly hatched mealworms when they are available. Mealworms cultured in the laboratory are held in small vessels containing a mixture of bran and organic spelt flour. The holding vessels are covered with brown paper towel. Pieces of potatoes are added to the holding vessels as necessary. Cultures should be thinned and substrates replaced approximately once every 6 weeks, when the bran has disappeared and the media appears very sandy. Older mealworm beetles are not used for feeding, as the exoskeleton is too difficult to digest. Mealworms have routinely and successfully been used for feeding various ages of frogs at ALET (especially small frogs). Earthworms (*Eisenia andrei*) from laboratory test cultures (see ECCC, 2022a) can also be used to feed the frogs. However, there are reports in the literature of insects used as a food source, including mealworms, earthworms, and fruit flies, having low levels of calcium or poor calcium:phosphorus ratios; therefore, it is recommended to dust insects with a calcium-rich powder before feeding (Densmore and Green, 2007) (recommend a 1:4 mixture of vitamins and calcium carbonate at least twice per week; see Section 2.4.2).

E.1.3 Identifying Adult Frogs

Identification of each adult northern leopard frog can be maintained by knowing tank identification and using photographic identification of each frog (see Figure E.3 for examples). The "Frog Log" can include: a photograph of spot pattern, weights, source, date received, parentage (if applicable), dosages and dates of medications (if applicable), and breeding attempts (including hibernation and hormone injection details; see Appendix F). All the frogs at ALET have been successfully identified in this manner; frogs as small as 2 grams have been photographed and identified at a later date by their spot pattern. Alternatively, frogs can be tagged for identification purposes (see Appendix 2 in CCAC, 2021).

E.2 Quarantine and Disease: Additional Guidance

Prophylactic treatment can be considered upon receipt of new batches of adult frogs, and should be discussed with a veterinarian. At ALET, frogs have previously been placed in a tetracycline bath (prepared in 0.6% sodium chloride with a tetracycline dosage of 10 µg/mL) for 20 minutes prior to quarantine; however, more recently, fecal samples have been collected and analyzed upon receipt of a new batch of frogs prior to performing any treatment (P. Jackman, Environment and Climate Change Canada, Moncton, NB, personal communication, 2022). Guidance for duration of quarantine is provided in Section 2.4.4. Quarantine tanks must be kept separate from any current laboratory frog tanks (Section 2.4.4). Any mortality in the new tank should be investigated by sending samples for pathology testing to determine if disease is present and if any treatments are required. Fecal samples can be collected from new batches of frogs and sent for parasite analysis; if a batch of frogs is held in the lab for a prolonged period, it is recommended that fecal samples be collected and analyzed on a regular basis (e.g., every 3 months). Swab samples can also be taken to check for the presence of *Ranavirus*. See Appendix E.2.2 for additional details on sampling for disease. If disease is suspected, the frog should be isolated. Gloves (see Appendix E.2.1) should be changed following contact with a potentially infected frog, and all equipment should be disinfected with an appropriate solution (e.g., Wescodyne™). If positive results are obtained, treatments are given, and fecal samples are repeated four weeks after treatment. A veterinarian should be consulted during diagnosis of disease and throughout the course of any treatment.

If mortality is low and fecal results are negative, the frogs can be removed from quarantine and moved into the main culturing facility. Any positive fecal or pathology results should be discussed with a veterinarian or animal disease specialist for possible treatments or further testing before quarantine is removed.

In some cases, it may be necessary to euthanize a frog (see Section 4.5 and CCAC, 2010, 2021). This may be required in cases of severe disease, poor condition (i.e., emaciation), abnormal development affecting the health of the animal, evident distress or pain, or when directed by a veterinarian.

E.2.1 Gloves

Gloves must be worn when handling frogs (see Sections 2.3.9 and 2.4.4), and no transfer of frogs, water, or materials between tanks is permitted during quarantine. The use of rinsed, powder-free gloves is recommended, and nitrile gloves have been used successfully at ALET (P. Jackman, Environment and Climate Change Canada, Moncton, NB, personal communication, 2022). Latex gloves must not be used unless proven to be non-toxic in a laboratory investigation. All brands and types of gloves are potentially toxic, however, and should not be used until they have been tested with the species intended for use. In addition, variation in production within a given brand of glove can occur, so intermittent testing on new batches of gloves for safe use should also be carried out (Cashins *et al.*, 2008; Greer *et al.*, 2009). The use of gloves is required herein when handling all developmental stages of *L. pipiens* to prevent the inadvertent transmission of potential pathogens or other materials that may be harmful to the organisms. Latex and nitrile gloves have been shown to be extremely toxic to various species of larval amphibians even when used in the general cleaning and maintenance of tanks being used to hold organisms (Sobotka and Rahwan, 1994; Gutleb et al., 2001). The findings of Cashins *et al.* (2008) with respect to latex and nitrile gloves were similar, however, they also reported that vinyl gloves, if unrinsed prior to use, could cause

mortality in tadpoles, even after short-term exposures during routine handling. Based on these results, well-rinsed, powder-free nitrile or vinyl gloves are recommended for use herein when cleaning aquaria or handling frogs, provided they prove to be non-toxic in a laboratory investigation.

E.2.2 Sampling for Disease

Any mortality in a new tank should be investigated by sending samples for pathology testing to determine if disease is present and if any treatments are required (see Section 2.4.4).

Fecal samples can be collected from new batches of frogs and sent for parasite analysis. These can be collected from the culture tank water using a wide bore glass rod with a bulb attached, or a disposable plastic pipette with the end cut off. The feces and some culture water are drawn into the pipette and then expelled into a plastic sample container. As much feces as possible should be collected, with all pieces from the same tank going into the same sample container. Each culture tank should have its own sample container with multiple pieces of feces. Each sample container should be labelled with the tank number, date collected, and name of the person who collected it. Once sampling is complete, the containers should be placed in a cooler on ice and shipped for parasite analysis. Instructions on sampling and shipping as provided by the analytical laboratory should be followed if they differ from the recommendations listed here.

Swab samples may also be taken and analyzed by PCR for the fungus causing chytridiomycosis (see Appendix E.2.2). Polyester swabs with plastic sticks should be used to collect samples to be analyzed for the chytrids fungus by PCR. Swabs should be placed into a dry plastic container. One swab can be used per frog, but samples from the same culture tank may be combined into one sample container to determine if the chytrids fungus is present in a culture tank. The sample containers should be placed in a cooler on ice and shipped for analysis. Instructions on sampling and shipping as provided by the analytical laboratory should be followed if they differ from the recommendations listed here.

E.2.3 Common Diseases Affecting Adult Frogs and Tadpoles

There are a number of pathogens causing infectious diseases in *L. pipiens* that threaten their health and survival both in the wild and in captivity. These include viruses, fungi, bacteria, water mould, and parasites. A thorough review of common amphibian diseases is provided in Densmore and Green (2007).

Ranaviral disease in amphibians is caused by multiple “species” of closely related viruses placed in the genus *Ranavirus*. Ranaviruses are highly infectious since inoculating doses can be very low. Clinical signs of acute ranaviral disease seen in tadpoles, metamorphs, juveniles, and adults include decreased activity, abnormal body posture or swimming behavior, ascites, skin ulcerations, focal hemorrhages, and death. *Ranavirus* vaccines are not currently available for amphibians and there is no known treatment or cure.

Chytridiomycosis, or chytrids, is a disease that is on the rise and is caused by the highly transmissible fungal pathogen *Batrachochytrium dendrobatidis* (*Bd*); this disease has caused rapid population decline or extinction of up to 200 species of frogs globally (Skerratt *et al.*, 2007). The fungus is transmitted via zoospores that require water as a medium. The fungus targets keratinized tissues, including the mouthparts of tadpoles and the skin of metamorphosed frogs (Berger *et al.*, 1998; Voordouw *et al.*, 2010). Tadpoles can be readily infected but do not always suffer obvious clinical effects. Clinical signs of chytridiomycosis in post-metamorphic amphibians can include hyperkeratosis (thickening of the outer skin layer), excessive skin peeling or sloughing, and sudden death (Berger *et al.*, 1998; Voordouw *et al.*, 2010). Other clinical signs in juveniles and adults can include abnormal postural changes (e.g., hind legs held out from flanks), loss of righting reflex, behavioural changes (e.g., lethargy, nocturnal frog sitting out in daylight, unresponsive to stimuli, lack of flee response, or absence of fear when handled), redness, and vascularisation of extremities. The prevalence of *Bd* infection in *L. pipiens* populations has been observed to be as high as 18.6% in British Columbia (Voordouw *et al.*, 2010), 25% in Prince Edward Island (Forzán *et al.*, 2010), and 25.7% in Maine (Longcore *et al.*, 2007). The fungus cannot grow when air temperatures ≥28 °C, so disease prevalence is usually lowest in the summer and highest during spring and fall when

temperatures are cooler and amphibians enter breeding or overwintering habitat (Voordouw *et al.*, 2010). The risk of *Bd* infection is also higher for species that hibernate in aquatic habitats rather than terrestrial habitats (Longcore *et al.*, 2007), so it is important to monitor frogs for signs of disease during artificial hibernation in the laboratory. *Bd* infection can be detected by analyzing swab samples of the abdomen or feet by PCR. Antifungal agents can kill *Bd* in culture, but the effect in the infected tadpole, juvenile, and adult is variable in terms of cure (Johnson *et al*., 2003; Densmore and Green, 2007).

Another disease found among *L. pipiens* is known as red leg or red-legged disease (see Figure E.4), which is most commonly associated with a bacterial pathogen called *Aeromonas hydrophila*. The most common symptom is erythema; however, other clinical signs may include anorexia, swelling, edema, or coelomic effusions, as well as epidermal erosions, ulcers, sloughing, or necrosis. This disease is most often fatal and may present as sudden death with few or no clinical signs (Densmore and Green, 2007). Red-legged disease can be treated with broad-spectrum antibiotics.

Saprolegniasis, or common water mould disease, is associated with *Saprolegnia ferax* or *S. parasitica*. Saprolegniasis can result in significant egg mortality, and in larvae it presents as the external appearance of fungal colonies that appear fluffy or cotton-like in texture. Erythematous or ulcerated skin may also be visible, most commonly affecting the tail, hindlimbs, gills, and oral mucous membranes (Densmore and Green, 2007). Treatment with various antifungal agents has proven effective.

Helminth parasites are also common among northern leopard frogs. *Ribeiroia ondatrae* causes limb deformities and mortality at various stages of development (COSEWIC, 2009). The most commonly seen parasites in fecal samples from ALET are protozoa and nematodes. Protozoa usually do not require treatment, but the presence of nematodes usually requires the use of a deworming agent.

E.2.4 History of Disease Outbreaks, Adult Health Issues, and Treatment at ALET[124]

On several occasions, including two hibernation events, a disease outbreak was observed at ALET. During treatment of diseased frogs, frogs with the least symptoms are always handled first.

- Symptoms similar to red-legged disease were observed and mortality occurred in some of these animals. Chytridiomycosis (*Bd* causative agent) was identified during pathological examination in one case. Treatment with injections of amikacin into the hind leg were performed for 7 days, but less than half of the frogs survived. In an unrelated later event, there was a mass die-off of *L. pipiens* metamorphs due to chytrids. Mortality occurred approximately six weeks after the animals metamorphosed. These animals had been raised in the laboratory since they were young tadpoles, however, tadpoles may not show clinical signs of chytrids due to a lack of keratin. A treatment of an itraconazole bath for 5 minutes for 10 consecutive days was provided and proved successful; future swab tests were negative, and mortality ceased.

- *Aeromonas hydrophila* was identified as being present on several occasions. On one occasion, mass mortality occurred with new metamorphs after transfer to the mixed habitat. Intramuscular injections of enrofloxacin (Baytril®) at 10 µg/g for 7 consecutive days was a successful treatment.

- *Mycobacterium marinum* was identified in a tank with high mortality at ALET. Since there is no treatment for this disease, the remainder of the test animals were euthanized. *Ranavirus* was also identified in the frogs that tested positive for *Mycobacterium marinum.*

[124] The information in this section was sourced from ALET (2004, 2006, 2009).

Two incidents of mild cloacal prolapse (prolapse of rectal or cloacal tissue) occurred in an adult frog and were repaired by laboratory staff according to the procedures outlined in Wright and Whitaker (2001). This may have been caused by hypocalcemia, gastrointestinal impact, or obstruction; the cause was not determined. This is a known problem with amphibians, but occurrence in the lab is rare.

Figure E.1 Adult frog culture tank with river rocks as terrestrial platform (ALET, 2004)

Figure E.2 Adult frog culture tank with Plexiglas terrestrial platform, during feeding

Figure E.3 Examples of spot patterns of adult northern leopard frogs (ALET, 2004)

Figure E.4 Frog with red-legged disease (ALET, 2004). The red coloration on the legs and abdomen is much brighter than the slight redness associated with handling frogs.

Appendix F

Procedures for In-Laboratory Hibernation and Breeding for *Lithobates pipiens* Adults

F.1 Background

The procedures described in this appendix are based on those developed by Dr. Vance Trudeau, and further refined at ALET. Dr. Trudeau's procedure involves an artificial hibernation period followed by a combined hormone-neurotransmitter injection of gonadotropin-releasing hormone agonist (GnRH-A) and metoclopramide hydrochloride (MET; a dopamine antagonist), referred to as the AMPHIPLEX method (Trudeau *et al.*, 2010, 2013; Vu *et al.*, 2017). A second iteration of the method incorporated a single priming dose of GnRH-A given to each frog 24 hours before the combined GnRH-A/MET injection (Trudeau *et al.*, 2013). This resulted in a high fertilization rate and viability of eggs outside of the natural breeding season, a result that has been replicated in two laboratories as part of ECCC test method validation (Nautilus Environmental, 2016; ALET, 2018). The AMPHIPLEX method has been used to successfully induce breeding in *L. pipiens* and provide test organisms for a number of published studies (Melvin and Trudeau, 2012a; Leduc *et al.*, 2016; Milotic *et al.*, 2017; Robinson *et al.*, 2019; Young *et al.*, 2020). Other artificial hibernation and breeding procedures, or variations of those provided herein, may also be used to produce larval *L. pipiens* test organisms for this ECCC test method provided that the test organisms meet the health criteria requirements described in Section 2.3.8. The procedures documented in this appendix closely follow those that were successfully used by ALET and Nautilus Environmental.

F.2 Hibernation Protocol

1. Male and female adult *L. pipiens* may be collected from the field or ordered from a commercial supplier. More males than females should be used, if possible; a ratio of 3 males : 2 females is recommended (P. Jackman, Environment and Climate Change Canada, Moncton, NB, personal communication, 2022).[125] Do not feed adults for 48 hours prior to transferring to hibernation tanks, in order to reduce metabolic waste in the tanks.

2. Prepare hibernation tanks in a temperature-controlled room or growth chamber by adding plastic plants and rocks or inverted glass dishes to 20-L plastic containers (Figure F.1). Fill containers halfway with dechlorinated water or 1:20 diluted Ringer's solution,[126] and add a commercial aquarium airstone and thermometer to each.

---

[125] Breeding trials have been performed with varying ratios of male to female frogs; however, more male than female frogs are usually used to induce environmental cues (i.e., competition). Breeding trials performed during ECCC method development produced two viable egg masses using 8 female and 19 male northern leopard frogs (Nautilus Environmental, 2016), and five viable egg masses using 11 female and 17 male northern leopard frogs (ALET, 2018). Results from artificial breeding described in the literature include 11 egg masses produced from 12 female and 16 male northern leopard frogs (Trudeau *et al.*, 2013).

[126] Trudeau *et al.* (2010) held frogs in dilute Ringer's solution (1:20 dilution of 0.1 M NaCl, 1.8 mM KCl, 2 mM $CaCl_2$, 1 mM $MgCl_2$, and 300 mg/L $NaHCO_3$ in water) during hibernation and breeding; however, further AMPHIPLEX method development was performed using dechlorinated water (Trudeau *et al.*, 2013; Vu *et al.*, 2017). Dechlorinated water has been successfully used for hibernation and breeding in two labs during method development (Nautilus Environmental, 2016; ALET, 2018).

3. Transfer adult frogs to hibernation tanks, with a maximum of 6 frogs per 20-L container. Males and females should be kept in separate tanks throughout hibernation. Frogs are not fed during hibernation.

4. Initiate the light and temperature schedule to start hibernation (Table F.1). The recommended 62-day schedule has been successfully used by ALET (2018); however, hibernation schedules of varying lengths (30, 38, 41, 50, 62, and 76 days) have also been successfully used in different laboratories (Trudeau *et al.*, 2010, 2013; Nautilus Environmental, 2016; ALET, 2018).

5. Static-renewal or flow-through systems may be used during hibernation. If using static-renewal, perform ≥50% renewal of solution in hibernation tanks daily for the first two weeks of hibernation, and three times per week for each week after. Minimize disruption to frogs during solution renewal. If using static-renewal, place aerated renewal solution inside the temperature-controlled room or growth chamber to ensure it is at the correct temperature before it is added to the hibernation tanks. If using flow-through, provide continuous low flow in each tank, with solution replacement rate approximately equal to that recommended for static renewal; if necessary, siphon regularly to remove solid waste and debris. Maintain continuous, gentle aeration in each tank throughout hibernation. Measure air and solution temperature daily. Measure dissolved oxygen and pH of new and old solutions for each hibernation tank at the time of renewal, or at least three times per week if using a flow-through system. Monitoring ammonia is optional; as frogs are not fed during hibernation, metabolic waste is assumed to be minimal. Check daily for mortality and signs of stress or disease in each tank.

6. Adjust the photoperiod and temperature as required by the hibernation schedule.

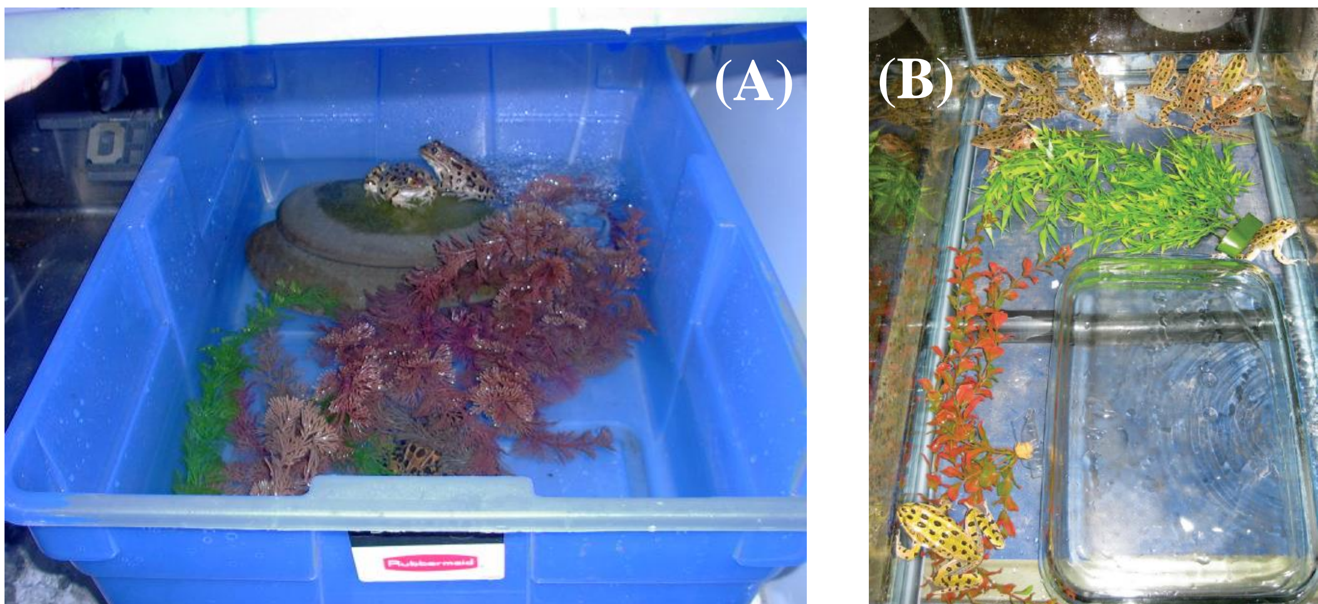


Figure F.1 Examples of hibernation tanks with plastic plants, inverted glass dishes or rocks, and dechlorinated water (A: ALET, 2004; B: Nautilus Environmental, 2016 [note that a higher loading density was used in this study than what is recommended in the final protocol]).

Table F.1 Hibernation light and temperature cycle (ALET, 2018).

Day	Temperature (°C)	Light on	Light off	Light (h):Dark (h)
1	15	6:30	19:30	13:11
2	13	6:30	19:30	13:11
3	11	7:00	19:00	12:12
4	9	7:00	19:00	12:12
5	8	7:00	19:00	12:12
6	7	7:00	19:00	12:12
7	5	7:30	18:30	11:13
8-9	3	7:30	18:30	11:13
10-12	3	8:00	18:00	10:14
13-15	3	8:00	17:00	9:15
16-52	3	8:00	17:30	9.5:14.5
53	6	8:00	18:00	10:14
54	6	7:30	18:00	10.5:13.5
55	6	7:00	19:00	12:12
56-58	8	6:30	20:00	13.5:10.5
59-60	10	6:00	21:00	15:9
61	13	6:00	22:00	16:8
62*	15	6:00	22:00	16:8

*Frogs are primed with GnRH-A on this day. Frogs are taken out of hibernation the next day, injected with GnRH-A/MET, and transferred to breeding tanks.

F.3 Breeding Injection Protocol

All injections described in this protocol are intraperitoneal and delivered using 26-gauge needles with disposable 1 mL syringes (Figure F.2); each needle should only be used once, as they dull quickly (CCAC, 2021). All injections and animal handling are performed in accordance with animal care guidelines (CCAC, 2021).

Chemicals:

- GnRH-A (des-Gly^{10}, D-Ala^{6}, Pro-NHEt^{9}-LHRH acetate salt; Bachem Product No. 4012028)
- MET (metoclopramide hydrochloride; Sigma-Aldrich Product No. M0763)
- Saline (sterile, IV-grade)

1. Prepare breeding tanks in a temperature-controlled room or growth chamber (Figure F.3). Breeding tanks should be large (e.g., 150 L) with deep water (minimum ≥ 20 cm) and a lid. Breeding tanks should have a terrestrial surface, such a Plexiglas platform, covering ~25% of the surface, as well as sufficient substrate for egg-laying sites, such as plastic plants covering ~80% of the water surface and additional plants weighed down to provide substrate at different depths of the tank. Plastic plants can be attached to string and tied to a platform on one end and onto a rock on the other end, so that the plants are suspended from the top and bottom of the water column. A red light (see Section 2.4.2) should be placed above a surface to allow for basking.

2. Prepare stock solutions of GnRH-A in saline on the day of use (e.g., on Day 62 of the hibernation schedule in Table F.1): for example, add 2 mg of GnRH-A to 1 mL of saline to prepare a 2000 µg/mL stock solution, then add 0.1 mL of this stock solution to 19.9 mL of saline to prepare a 10 µg/mL solution. 24 hours before the end of hibernation,[127] after the temperature has been increased to 15 °C, weigh each male and female frog and prime each frog with an injection of GnRH-A at 0.04 µg/g body weight (e.g., for a 50 g frog, inject 0.2 mL of 10 µg/mL stock solution). Keep males and females separated after priming. Record any observed mating calls or behaviour changes after priming.

3. The next day (i.e., at the end of hibernation, 24 hours after priming), prepare a combined hormone solution by dissolving 1.6 mg of GnRH-A and 40 mg of MET in 20 mL of saline; this is equivalent to 80 µg/mL GnRH-A and 2000 µg/mL MET. Inject each frog with the combined hormone solution, such that each frog receives 0.4 µg GnRH-A/g body weight and 10 µg MET/g body weight, to induce breeding. To determine the amount to inject into each frog, use the following equation:

$$\text{injection volume} = \frac{\text{weight of frog in grams} \times 0.4\ \mu\text{g/g GnRHA}}{80\ \mu\text{g/ml GnRHA}}$$

   For example, a 50 g frog would require a 0.25 mL injection. Since the combined hormone solution also contains 2000 µg/mL of MET, this frog would receive 500 µg of MET, which is equivalent to the desired concentration of 10 µg/g body weight.

4. After injections are complete, transfer frogs to breeding tanks in the temperature-controlled room or chamber, with a maximum of 10 frogs per tank; a ratio of 3 males : 2 females is recommended (P. Jackman, Environment and Climate Change Canada, Moncton, NB, personal communication, 2022). The temperature is maintained at 15 °C, with a photoperiod of 16-h light : 8-h dark. Frogs are not fed during breeding.

F.4 Breeding and Egg Collection Protocol

1. Frog behaviour in breeding tanks is observed daily. Observed incidences of mating calls, number of pairs in amplexus (Figure F.4), and number of egg masses (Figures F.5 and F.6) are recorded for each tank daily. Avoid disturbances and excessive monitoring, as frogs seem to prefer low-activity environments for breeding.

2. After amplexus is no longer observed and all egg masses are laid (typically 2–5 days after the final hormone injection), egg masses are transferred to aerated holding vessels (see Sections 2.3.2 and 2.3.5, and Figure F.5). Fertilized egg masses may be maintained at cooler temperatures (e.g., 10 to 15 °C) for up to 5 weeks to delay development, if they are not to be used immediately in a toxicity test (see Section 2.3.4).

3. Record the approximate number of eggs in each egg mass (typically several hundred to more than a thousand eggs per egg mass). If using static renewal, replace 50% of the water in each tank at least three times per week; egg masses must remain submerged during renewals. Measurements of temperature, pH, and dissolved oxygen should be measured during water renewals. Egg masses should be monitored daily for signs of disease, fungus, or stress. The approximate fertilization rate (%) should be recorded. When first laid, eggs are

[127] If collected during breeding season and adults are in breeding-ready condition, northern leopard frogs can also be spawned shortly after arrival at the laboratory (i.e., without hibernation). Adult females and males are kept in separate tanks. Frogs should be quarantined, acclimated, housed, and fed as previously described. Following hormone injections, frogs are transferred to large breeding tanks appropriate for egg laying, as described in Appendix F.3. Trudeau *et al.* (2013) describes an example of successful in-season breeding induced by hormone injections in the laboratory using field-collected *L. pipiens*.

about half black (animal hemisphere) and half white (vegetal hemisphere); if the eggs are fertilized, the black will move into the white and the eggs will be completely black in about 24 hours (Figure F.6). It is not atypical for a portion of the egg mass to remain unfertilized.

4. Once tadpoles are hatched and swimming, they are transferred to aerated holding vessels (see Sections 2.3.2 and 2.3.5) and the approximate number of tadpoles is recorded. The temperature can then be increased at a rate of ≤3 °C/day until it reaches test temperature (23 ± 2 °C). Other culturing conditions (feeding, water quality, handling, etc.) are described in Section 2.3.

Figure F.2 Intraperitoneal injection of *L. pipiens* adult frog with GnRH-A/MET to induce breeding (ALET, 2018).

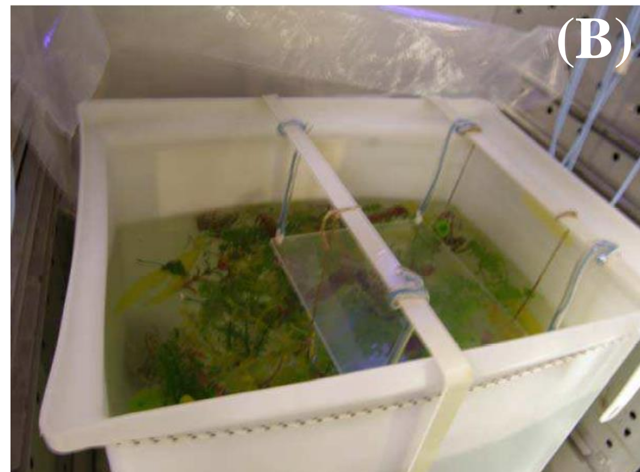
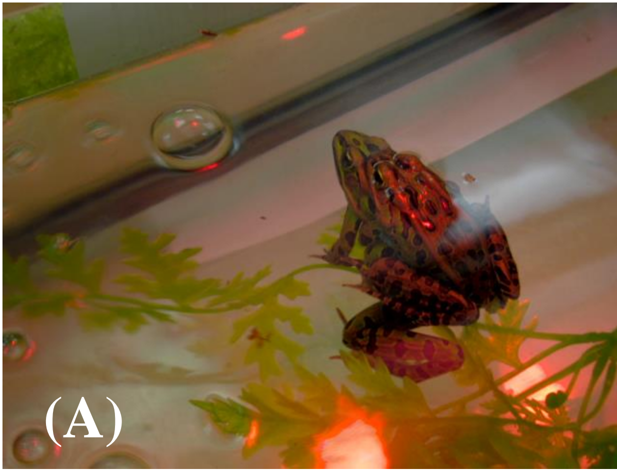


Figure F.3 Examples of 150-L breeding tanks for *L. pipiens*, with raised terrestrial platforms and plastic plants at different levels of the water column to act as breeding substrate (A: ALET, 2009, B: ALET, 2018).



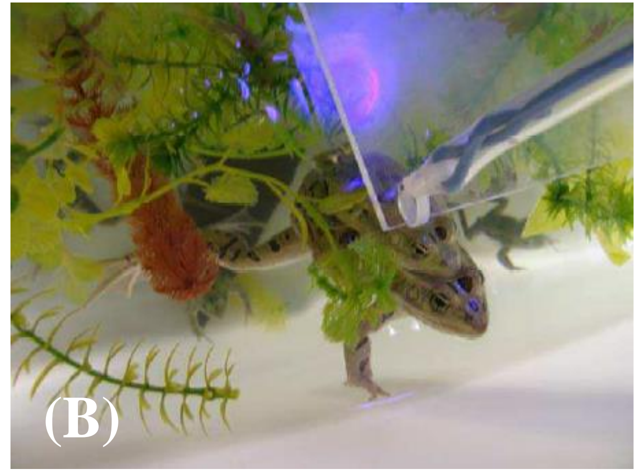


Figure F.4 Frogs in amplexus (A: ALET, 2004, B: ALET, 2018).

Figure F.5 *L. pipiens* egg masses in holding containers (ALET, 2009).

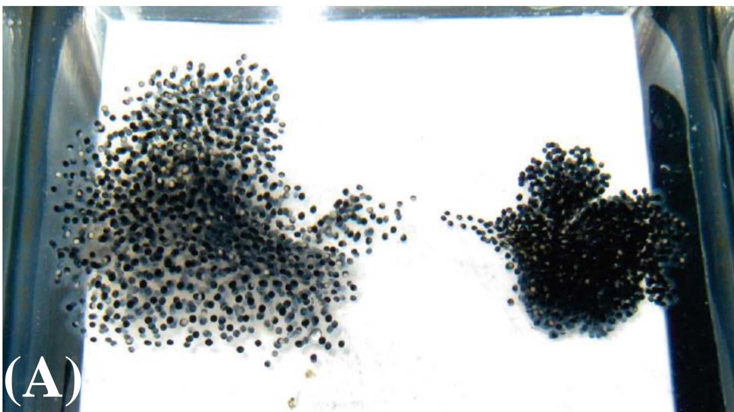


Figure F.6 *L. pipiens* egg masses (A: Nautilus Environmental, 2016; B: ALET, 2004).

Appendix G

Logarithmic Series of Concentrations Suitable for Toxicity Tests[128]

Column (Number of concentrations between 10.0 and 1.00, or between 1.00 and 0.10)[129]

1	2	3	4	5	6	7
10.0	10.0	10.0	10.0	10.0	10.0	10.0
3.2	4.6	5.6	6.3	6.8	7.2	7.5
1.00	2.2	3.2	4.0	4.6	5.2	5.6
0.32	1.00	1.8	2.5	3.2	3.7	4.2
0.10	0.46	1.00	1.6	2.2	2.7	3.2
	0.22	0.56	1.00	1.5	1.9	2.4
	0.10	0.32	0.63	1.00	1.4	1.8
		0.18	0.40	0.68	1.00	1.3
		0.10	0.25	0.46	0.72	1.00
			0.16	0.32	0.52	0.75
			0.10	0.22	0.37	0.56
				0.15	0.27	0.42
				0.10	0.19	0.32
					0.14	0.24
					0.10	0.18
						0.13
						0.10

[128] Modified from Rocchini *et al.* (1982).

[129] A series of successive concentrations may be chosen from a column. Midpoints between concentrations in column (x) are found in column (2x + 1). The values listed can represent concentrations expressed on a percentage by weight (e.g., mg/kg) or weight-to-volume (e.g., mg/L) basis. As necessary, values can be multiplied or divided by any power of 10. Column 2, which spans two orders of magnitude in concentration, might be used if there was considerable uncertainty about the degree of toxicity. More widely spaced concentrations should not be used, since such usage gives poor resolution of the confidence limits surrounding any threshold-effect value calculated. The finer gradations of columns 4 to 7 might occasionally be useful for testing chemicals that have an abrupt threshold of effect.