Tanguay_ZF_120hpf_PFIN_legacy

1. General Information

1.1 Assay Title: (Legacy) Oregon State University Tanguay Lab 120 Hour Post-fertilization Zebrafish Assay for Pectoral Fin Malformation

1.2 Assay Summary: Tanguay_ZF_120hpf_legacy is a whole embryo, multiplexed endpoint assay using zebrafish larvae exposed for 120 hours post fertilization on a 96-well plate. Tanguay_ZF_120hpf_PFIN_legacy is one of one assay component(s) measured or calculated from the Tanguay_ZF_120hpf_legacy assay. It is designed to make measurements of zebrafish development as detected with brightfield microscopy of developing zebrafish embryos. Data from the assay component Tanguay_ZF_120hpf_PFIN_legacy was analyzed into 1 assay endpoints. This assay endpoint, Tanguay_ZF_120hpf_PFIN_legacy, was analyzed in the positive analysis fitting direction relative to DMSO as the negative control and baseline of activity. Using a type of morphology reporter, gain-of-signal activity can be used to understand changes in developmentals as they relate to the whole embryo. To generalize the intended target to other relatable targets, this assay endpoint is annotated to the zebrafish development intended target family, where the subfamily is pectoral fin morphogenesis.

1.3 Date of Document Creation: September 05 2024

1.4 Authors and Contact Information:

US Environmental Protection Agency (EPA) Center for Computational Toxicology and Exposure (CCTE)
109 T.W. Alexander Drive (Mail Code D143-02)
Research Triangle Park, NC 27711

1.5 Assay Source: The Tanguay Lab, based at the Oregon State University Sinnhuber Aquatic Research Laboratory, uses zebrafish as a systems toxicology model.

1.6 Date of Assay Development: For date of assay development, see *Section 6: Bibliography*.

1.7 References: For complete list of references, see *Section 6: Bibliography*.

1.8 Proprietary Elements: Assay is non-proprietary; observations were made using a custom photomotor response analysis tool (PRAT), Viewpoint Zebralab, and Zebrafish acquisition and analysis program (ZAAP).

1.9 Assay Throughput: 96-well plate. The assay is conducted on 96-well plates with each plate containing 1, six hour post-fertilization dechorionated embryo using an automated embryo placement system.

1.10 Status: The assay is fully developed, and data are publicly available in ToxCast's invitroDB.

1.11 Abbreviations:

AIC: Akaike Information Criterion
AOP: Adverse Outcome Pathway
CV: Coefficient of Variation
DMSO: Dimethyl Sulfoxide
ToxCast: US EPA's Toxicity Forecaster Program
tcpl: ToxCast Data Analysis Pipeline R Package
SSMD: Strictly Standardized Mean Difference

2. Test Method Description

2.1 Purpose: Morphology is measured by light microscopic examination of developing zebrafish embryos.

Zebrafish (Danio rerio) is popular species in embryology, pharmacology and biomedical research and is particularly amenable to large-scale screening of chemical libraries. These animals easy to rear and maintain and they mature rapidly (6 days). Zebrafish are also are small enough for sustaining in 96-well microtiter plates. These assays screened embryonic responses to chemical exposures by visually assessing multiple phenotypic indicators of developmental interference, including malformations, failure to hatch, and mortality. There are scientific advantages to assessing zebrafish as a prototype for delineating the functional activity of specific biological pathways and their regulatory controls. Many key developmental signaling pathways and their regulatory mechanisms are conserved between fish and mammals, making zebrafish toxicity assays a unique integrative model of embryogenesis and highly adaptable to a medium throughput toxicity screening platform.

2.2 Scientific Principles: The utilization of simultaneously measured endpoints means that the entire system serves as a robust biological sensor for chemical hazard. The experimental design enables the description of

global patterns of variation across tested compounds, evaluation the concordance of the available in vitro and in vivo data, can highlight specific mechanisms and novel biology related development, and demonstrate that the developmental zebrafish detects adverse responses that would be missed by less comprehensive testing strategies.

2.3 Experimental System: suspension NA whole embryo used. Dechorionated tropical 5D wild-type zebrafish (Danio rerio) embryos placed 1 embryo per well in a 96-well plate. The parental fish were tropical 5D wild-type zebrafish were housed at Oregon State University's Sinnhuber Aquatic Research Laboratory (SARL, Corvallis, OR) in densities of 1000 fish per 100-gallon tank according to the Institutional Animal Care and Use Committee protocols (Barton et al., 2016). Fish were maintained at 28C on a 14:10 h light/dark cycle in recirculating filtered water, supplemented with Instant Ocean salts. Adult, larval and juvenile fish were fed with size-appropriate GEMMA Micro food 2–3 times a day (Skretting). Spawning funnels were placed in the tanks the night prior, and the following morning, embryos were collected and staged (Kimmel et al., 1995, Westerfield, 2007). Embryos were maintained in embryo medium (EM) in an incubator at 28C until further processing. EM consisted of 15 mM NaCl, 0.5 mM KCl, 1 mM MgSO4, 0.15 mM KH2PO4, 0.05 mM Na2HPO4, and 0.7 mM NaHCO3 (Westerfield, 2000).

2.4 Metabolic Competence: Zebrafish provide a rapidly developing and easily maintained test organism which is visually transparent through much of its embryonic development and has an elevated xenobiotic biotransformation potential when compared to other commonly used models of developmental toxicity. Zebrafish embryos (Danio rerio) were obtained from tropical 5D wild-type adult zebrafish were housed in at an approximate density of 1000 per 100 gallon tank at the Sinnhuber Aquatic Research Laboratory, Oregon State University, Corvallis, OR. Zebrafish are a good model in which to study metabolism because they possess all the key organs required for metabolic control in humans, from the appetite circuits that are present in the hypothalamus, through to the pancreas and insulin-sensitive tissues [liver, muscle and white adipose tissue (WAT)].

2.5 Exposure Regime: Zebrafish husbandry: Tropical 5D wild-type adult zebrafish were housed in at an approximate density of 1000 per 100 gallon tank at the Sinnhuber Aquatic Research Laboratory, Oregon State University, Corvallis, OR. Each tank was kept at standard laboratory conditions of 28C on a 14-h light/10-h dark photoperiod in fish water consisting of reverse osmosis water supplemented with a commercially available salt (Instant Ocean). Spawning funnels were placed into the tanks the night prior, and embryos were collected and staged. To increase bioavailability, the chorion was enzymatically removed using pronase (63.6mg/ml, ≥ 3.5U/mg, Sigma-Aldrich: P5147) at 4 hours post fertilization (hpf) using a custom automated dechorionator. Chemical exposures: Six hpf dechorionated embryos were placed 1 embryo per well in a 96-well plate prefilled with 90 ul of EM using automated embryo placement systems (AEPS). Ten microliters of each row of dilution plate 2 was added to 2 exposure plates. The final DMSO concentration used was 0.64% (vol/vol). Thirty-two embryos were also exposed to 5uM trimethyltin chloride (positive control). Plates were sealed to prevent evaporation and foil covered to reduce light exposure and kept in a 28C incubator. Embryos were statically exposed until 120 hpf. At 24 hpf, embryos were assessed for photomotor response using a custom photomotor response analysis tool (PRAT) and for developmental toxicity endpoints (MO24: mortality at 24 hpf, DP: developmental progression, SM: spontaneous movement, and NC: notochord distortion). At 120 hpf, locomotor activity was measured using Viewpoint Zebralab and assessed for 18 endpoints. Zebrafish acquisition and analysis program (ZAAP), a custom program designed to inventory, acquire, and manage zebrafish data, was used to collect developmental endpoints as either present or absent (i.e., binary responses were recorded). If mortality occurred for an embryo (at either 24 or 120 hpf), the non-mortality endpoints were not measured.

ASSAY DESIGN SUMMARY	
Nominal number of tested concentrations:	Target (nominal) number of replicates:
5	2
Standard minimum concentration tested:	Standard maximum concentration tested:
0.0095 μM	95.12 μM

Key positive control:	Neutral vehicle control:
NA	DMSO

Baseline median absolute deviation for the assay (bmad): 2.183

Response cutoff threshold used to determine hit calls: 20

Detection technology used: light microscopy (Microscopy)

2.6 Response: The raw data from the larval assessments consisted of an assigned binary (0 or 1) response to every larvae observed for each of the eighteen endpoints. Responses across all these endpoints were collapsed into a singular binary (0 or 1) morphology endpoint named 'ANY'.

2.7 Quality and Acceptance Criteria: Each assay may utilize different acceptance criteria and quality assurance methods as it pertains to the individual assay platform and implementation. Pre-processing transformations may indicate issues in plates or wells by setting well quality (wllq) values to 0. Analytical QC calls per sample and substance can also be considered to understand the applicability domain of the chemicals for screening.

2.8 Technical Limitations: ToxCast data can provide initial (screening) information about the capacity for a chemical to illicit a biological response; caution is advised with extrapolation of these results to organism-level responses. The potential for a chemical to elicit adverse health outcomes in living systems is a function of multiple factors, and this assay is not intended to provide predictive details regarding long term or indirect adverse effects in complex biological systems but can aid in the prioritization of compound selection for more resource intensive toxicity studies. See *Section 4.4*. for more information on the chemical applicability of the assay.

2.9 Related Assays: For related assays, consult the following assay lists or intended target families. This assay is present in the following assay lists:

Developmental Toxicity: Assays associated with developmental toxicity, Non-mammalian Vertebrate: Assays associated with non-mammalian vertebrate species

Additionally, this assay was annotated to the intended target family of zebrafish development.

3. Data Interpretation

The ToxCast Data Analysis Pipeline (tcpl) R package includes processing functionality for two screening paradigms: (1) single-concentration ("SC") and (2) multiple-concentration ("MC") screening. SC screening consists of testing chemicals at one concentration, often for the purpose of identifying potentially active chemicals to test in the multiple-concentration format. MC screening consists of testing chemicals across a concentration range, such that the modeled activity can give an estimate of potency, efficacy, etc. MC data is the focus of this documentation, with SC data processing metrics to be incorporated in the future.

3.1 Responses captured in prediction model: See *Section 2.6* for additional information on responses measured.

3.2 Data Analysis: All statistical analysis was performed using code developed in R. The data used were binary incidences recorded for each endpoint from Zebrafish acquisition and analysis program (ZAAP), plus associated plate and well-location information. This information was used to test for confounding plate, well, and chemical effects across all controls and to identify outliers. Outliers were defined as chemicals having an incidence rate greater than 3 SDs from the mean rate in controls across multiple endpoints. The lowest effect level (LEL) in micromolar is computed as the concentration at which the incidence exceeded a significance threshold over the background (control) incidence rate. Because the endpoints are binary and replicates are measured in separate wells, the 0/1 responses for each chemical-endpoint-concentration-replicate combination translate to a series (n = 32) of Bernoulli trials, or "coin-flips." Therefore, the LEL significance threshold was estimated using a binomial test, which provided a straightforward method to adjust for plate and/or chemical effects and the pooling/separation of controls. Given the experimental design, the binomial maximized power versus a typical logistic/curve-fit approach by accounting for the falsely "nonmonotonic" responses occurring when the MORT endpoint led to missing specific endpoint measurements at higher concentrations. Because background incidence rate varied slightly across chemicals and endpoints, the significance threshold (x) was determined independently from the binomial distribution function for each chemical-endpoint pair.

Prior to the data processing, all the data must go through pre-processing to transform the heterogeneous data into a uniform format before it can be loaded into a database. Level 0 pre-processing is done in R by

vendor/dataset-specific scripts with all manual transformations to the data documented with justification. Common examples of manual transformations include fixing a sample ID typo or changing well quality value(wllq) to 0 after identifying problems such a plate row/column missing an assay reagent.

Once data is loaded into the database, tcpl utilizes generalized processing functions provided to process, normalize, model, qualify, and visualize the data. To promote reproducibility, all method assignments must occur through the database and should come from the available list of methods for each processing level. Assigned multiple concentration processing methods include:

Level 2: Component-specific corrections include:

1: none (Use corrected response value (cval) as is; cval = cval. No additional mc2 methods needed for component-specific corrections.)

Level 3: Endpoint-specific normalization include:

1: none (Set the corrected response value (cval) as the normalized response value (resp); cval = resp. No additional mc3 methods needed for endpoint-specific normalization.)

Level 4: Baseline and required tcplFit2 parameters defined by:

1: bmad.aeid.lowconc.twells (Calculate the baseline median absolute value (bmad) as the median absolute deviation of normalized response values (rep) for test compound wells (wllt = t) with concentration index (cndx) equal to 1 or 2. Calculate one standard deviation of the normalized response for test compound wells (wllt = t) with a concentration index (cndx) of 1 or 2; onesd = sqrt(sum((resp - mean resp)^2)/sample size - 1). Onesd is used to establish BMR and therefore required for tcplfit2 processing.), 6: no.unbounded.models (Exclude unbounded models and only fit data to bounded models (hill, gnls, exp4 and exp5).)

Level 5: Possible cutoff thresholds, where higher value for endpoint is selected, include:

2: pc20 (Add a cutoff value of 20. Typically for percent of control data.), 5: bmad5 (Add a cutoff value of 5 multiplied the baseline median absolute deviation (bmad). By default, bmad is calculated using test compound wells (wllt = t) for the endpoint.), 28: ow_bidirectional_gain (Multiply winning model hitcall (hitc) by -1 for models fit in the negative analysis direction. Typically used for endpoints where only positive responses are biologically relevant.)

Level 6: Cautionary flagging include:

5: modl.directionality.fail (Flag series if model directionality is questionable, i.e. if the winning model direction was opposite, more responses (resp) would have exceeded the cutoff (coff). If loss was winning directionality (top < 0), flag if count(resp < -1*coff) < 2*count(resp > coff). If gain was winning directionality (top > 0), flag if count(resp > coff) < 2*count(resp < -1*coff).), 6: singlept.hit.high (Flag single-point hit that's only at the highest conc tested, where series is an active hit call (hitc >= 0.9) with the median response observed above baseline occurring only at the highest tested concentration tested.), 7: singlept.hit.mid (Flag single-point hit that's not at the highest conc tested, where series is an active hit call (hitc >= 0.9) with the median response observed above baseline occurring only at one concentration and not the highest concentration tested.), 8: multipoint.neg (Flag multi-point miss, where series is an inactive hit call (hitc < 0.9) with multiple median responses observed above baseline.), 9: bmd.high (Flag series if modeled benchmark dose (BMD) is greater than AC50 (concentration at 50 percent maximal response). This is indicates high variability in baseline response in excess of more than half of the maximal response.), 10: noise (Flag series as noisy if the quality of fit as calculated by the root mean square error (rmse) for the series is greater than the cutoff (coff); rmse > coff.), 11: border (Flag series if borderline activity is suspected based on modeled top parameter (top) relative to cutoff (coff); |top| <= 1.2(coff) or |top| >= 0.8(coff).), 13: low.nrep (Flag series if the average number of replicates per concentration is less than 2; nrep < 2.), 14: low.nconc (Flag series if 4 concentrations or less were tested; nconc <= 4.), 15: gnls.lowconc (Flag series where winning model is gain-loss (gnls) and the gain AC50 is less than the minimum tested concentration, and the loss AC50 is less than the mean tested concentration.), 17: efficacy.50 (Flag low efficacy hits if series has an active hit call (hitc >= 0.9) and efficacy values (e.g. top and maximum median response) less than 50 percent; intended for biochemical assays. If hitc >= 0.9 and coff >= 5, then flag when top < 50 or max_med < 50. If hitc >= 0.9 and coff < 5, then flag when top < log2(1.5) or max_med < log2(1.5).), 18: ac50.lowconc (Flag series with an active hit

call (hitc >= 0.9) if AC50 (concentration at 50 percent maximal response) is less than the lowest concentration tested;if hitc >= 0.9 and ac50 < 10^logc_min, then flag.), 20: no.med.gt.3bmad (Flag series where no median response values are greater than baseline as defined by 3 times the baseline median absolute deviation (bmad); nmed_gtbl_pos and nmed_gtbl_neg both = 0, where nmed_gtbl_pos/_neg is the number of medians greater than 3*bmad/less than -3*bmad.)

The following is an aggregate endpoint summary of the number of samples and chemicals tested, as well as active or inactive hit calls (hitc) and predicted winning models for all samples tested in this endpoint.

SAMPLE AND CHEMICAL COVERAGE	
Number of samples tested: 1078	Number of chemicals tested: 1060

ACTIVITY HIT CALLS		
Active hit count: hitc≥0.9 58	Inactive hit count: 0≤hitc<0.9 889	NA hit count: *hitc<0* *131*

WINING MODEL SELECTION	
Number of sample-assay endpoints with winning *hill* model:	21
gain-loss (gnls) model:	114
power(pow) model:	0
linear-polynomial (poly1) model:	0
quadratic-polynomial(poly2) model:	0
exponential-2 (exp2) model:	0
exponential-3 (exp3) model:	0
exponential-4 (exp4) model:	877
exponential-5 (exp5) model:	52

For each concentration series, several point-of-departure (POD) estimates are calculated for the winning model. The major estimates include: (1) the activity concentration at the specified benchmark response (BMR) (bmd), (2) the activity concentration at 50% of the maximal response (ac50), (3) the activity concentration at the efficacy cutoff (acc), (4) the activity concentration at 10% of the maximal response, and (5) the concentration at 5% of the maximal response.

3.3 Prediction Model: All statistical analyses were conducted using R programming language, employing the tcpl package to generate model parameters and confidence intervals. Each chemical concentration response series is fit to ten predictive models, encoded by the dependency package tcplfit2. The models include the constant, Hill, gain-loss, two polynomials (i.e. linear and quadratic), power, and four exponential variants. The polynomials, power, and exponential models are all based on BMDExpress2. The winning model (modl) is selected based on the lowest AIC value and is used to determine the activity (or hit call) for the concentration series. If two models have equal AIC values, then the simpler model (i.e. model with fewer parameters) wins. In invitrodb, levels 4 and 5 capture model fit information. mc4 captures summary values calculated for each concentration series, whereas mc4_param stores the estimated model parameters for all models fit to data in long format. mc5 captures the winning model selected and the activity hit call, whereas mc5_param stores the estimated model parameters for the selected winning model in long format. Activity for each concentration-response series is determined by calculating a continuous hit-call for the winning model, which is the product of three proportional weights. The first weight reflects whether there is at least one median response outside the efficacy cutoff band. Second, the top (or maximal change in the predicted response) is larger than the cutoff. The last weight reflects whether the AIC of the winning model is less than the constant model, i.e. the winning model is better fit than a flat line.

The continuous hit call value (hitc), fit category (fitc), and cautionary flags (mc6) can be used to understand the goodness-of-fit, enabling the user to decide the stringency with which to filter and interpret results. Hitc may be further binarized into active or inactive, depending on the level of stringency required by the user; herein, hitc greater than or equal to 0.90 is active, hitc between 0 and 0.90 is inactive, and hitc less than 0 is not applicable, but different thresholds may be used. Fitc was summarize curve behavior relative activity, efficacy, and potency comparisons between the AC50 and the concentration range screened. Cautionary flags on fitting were developed in previous versions of tcpl and have been stored at level 6. These flags are programmatically generated and indicate characteristics of a curve that need extra attention or potential anomalies in the curve or data. Users may review these filtered groupings to understand high-confidence curves.

3.4 Software: The ToxCast Data Analysis Pipeline (tcpl) is an R package that manages, curve-fits, plots, and stores ToxCast data to populate its linked MySQL database, invitrodb. Data for invitrodb v4.2 was processed using the tcpl v3.2. See *Section 7: Supporting Information* on the ToxCast program and tcpl R package.

4. Test Method Performance

4.1 Robustness: The following assay performance metrics surmise the robustness of the method i.e. the reliability of the experimental results and the prediction capability of the model used.

NEUTRAL CONTROL (well type = "n")	
Neutral control well median response value, by plate: *nmed*	0.437
Neutral control median absolute deviation, by plate: *nmad*	0
Coefficient of variation (CV%) in neutral control wells: *(nmad/nmed)*100*	0%
POSITIVE CONTROL (well type = "p")	
Positive control well median response value, by plate: *pmed*	NA
Positive control well median absolute deviation, by plate: *pmad*	NA
Z Prime Factor for median positive and neutral control across all plates: *(1 - ((3 * (pmad + nmad)) / abs(pmed - nmed))*	NA
Strictly standardized mean difference (SSMD) for positive compared to neutral control wells: *((pmed - nmed) / sqrt($pmad^2$ + $nmad^2$)*	NA
Positive control signal-to-noise: *((pmed-nmed)/nmad)*	NA
Positive control signal-to-background: *(pmed/nmed)*	NA
NEGATIVE CONTROL (well type = "m")	
Negative control well median, by plate: *mmed*	NA
Negative control well median absolute deviation value, by plate: *mmad*	NA
Z Prime Factor for median negative and neutral control across all plates: *(1 - ((3 * (mmad + nmad)) / abs(mmed - nmed))*	NA
Strictly standardized mean difference (SSMD) for negative compared to neutral control wells: *((mmed - nmed) / sqrt($mmad^2$ + $nmad^2$)*	NA
Signal-to-noise (median across all plates, using negative control wells): *((mmed-nmed)/nmad)*	NA
Signal-to-background (median across all plates, using negative control wells): *(mmed/nmed)*	NA

4.2 Reference Chemical Information: Reference chemical curation is ongoing, and this section will be updated as more information becomes available.

4.3 Performance Measures and Predictive Capacity: The performance and predictivity for a given assay may be evaluated with a variety of performance statistics but is dependent upon available data. Predictive capacity (i.e. false negative, false positive rates) will be assessed when reference chemical information is available. Ideally, assays will have sufficient data on reference chemicals (i.e. positive and negative controls) to enable estimation of accuracy statistics, such as sensitivity and specificity.

ToxCast targets may align to a range of event types in the Adverse Outcome Pathway (AOP) framework, however each assay technology may have specific limitations, which may require user discretion for more complex interpretations of the data.

The median root mean squared error (RMSE) across all winning models for active hits was calculated as: 52.

4.4 Chemical Library Scope and Limitations: The ToxCast Chemical Library was designed to capture a large spectrum of structurally and physicochemically diverse compounds. This chemical inventory incorporates toxicity data-rich chemicals, chemicals spanning major use-categories, and chemicals with exposure potential, including but not limited to pesticides, antimicrobials, fragrances, green chemistry alternatives, food additives, toxicity reference compounds and failed pharmaceuticals. In addition to environmental or exposure concerns, chemical selection criteria also consider practical constraints, such as commercial availability, dimethyl sulfoxide (DMSO) solubility and stability, and suitability for testing in automated or semi-automated systems (e.g., low volatility and moderate LogP values). Under these constraints, there were three major, interrelated drivers for chemical selection: availability of animal toxicity data or mechanistic knowledge, exposure potential, and EPA regulatory interest. The first driver would provide the necessary *in vivo* and mechanistic data to anchor and validate subsequent prediction modeling efforts, whereas the latter two were intended to provide coverage of the chemical landscape to which humans and ecosystems are potentially exposed and for which toxicity data are mostly lacking. Analytical QC calls per sample and substance should be considered to understand the applicability domain of the chemicals for screening.

5. Potential Regulatory Applications

5.1 Context of Use: Examples of end use scenarios could include, but are not limited to:

- *Support Category Formation and Read-Across:* The outcomes from the assay could be used to substantiate a hypothesis for grouping substances together for the purposes of read-across,
- *Priority Setting:* The assay might help prioritize substances within an inventory for more detailed evaluation,
- *Screening Level Assessment of a Biomarker or Mechanistic Activity or Response:* The screening level assessment may be sufficient to identify a hazard and provide a gauge of potency; or
- *Integrated approaches to testing and assessment (IATA):* The assay may form one component of an IATA.

6. **Bibliography:** Truong L, Reif DM, St Mary L, Geier MC, Truong HD, Tanguay RL. Multidimensional in vivo hazard assessment using zebrafish. Toxicol Sci. 2014 Jan;137(1):212-33. doi: 10.1093/toxsci/kft235. Epub 2013 Oct 17. PubMed PMID: 24136191; PubMed Central PMCID: PMC3871932.

7. **Supporting Information:**

More information on the ToxCast program can be found at: https://www.epa.gov/chemical-research/toxicity-forecasting. The most recent version of downloadable data can be found at: https://www.epa.gov/chemical-research/exploring-toxcast-data-downloadable-data. The ToxCast Data Analysis Pipeline (tcpl) R package is available on CRAN or GitHub. Check out tcpl's vignette for comprehensive documentation describing ToxCast data processing, retrieval, and interpretation.