

## TOX21\_ERa\_LUC\_VM7\_Agonist\_10nM\_ICI182780\_viability

### 1. General Information

- 1.1 Assay Title: Viability Assessment in the Tox21 VM7 Estrogen Receptor-alpha (ESR1) Agonism (10nM ICI182780) Luciferase Assay
- 1.2 Assay Summary: TOX21\_ERa\_LUC\_VM7\_Agonist\_10nM\_ICI182780 is a cell-based, single-readout assay that uses VM7, a human breast tissue cell line, with measurements taken at 22 hours after chemical dosing in a 1536-well plate. This is a secondary assay for specificity to TOX21\_ERa\_LUC\_VM7\_Agonist. See tox21-er-luc-bg1-4e2-agonist-p4. TOX21\_ERa\_LUC\_VM7\_Agonist\_10nM\_ICI182780\_viability is an assay readout measuring cellular ATP content and detected with CellTiter-Glo Luciferase-coupled ATP quantitation. TOX21\_ERa\_LUC\_VM7\_Agonist\_10nM\_ICI182780\_viability used a type of viability reporter where loss-of-signal activity can be used to understand changes in the cell viability. Furthermore, this assay endpoint can be referred to as a secondary readout, because this assay has produced multiple assay endpoints where this one serves a viability function. To generalize the intended target to other relatable targets, this assay endpoint is annotated to the cell cycle intended target family, where the subfamily is cytotoxicity.
- 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: Tox21 is an interagency agreement between the NIH, NTP, FDA and EPA. NIH Chemical Genomics Center (NCGC) is the primary screening facility running ultra high-throughput screening assays across a large interagency-developed chemical library.
- 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: This assay is not proprietary. The Tox21 qHTS robotic platform has a 1536-well per run capacity and is capable of fully-automated, hands-free execution (liquid dispensing and aspiration, plate centrifugation and incubation) and signal recording (plate readout). The GeneBLAzer System is publically available through Invitrogen.
- 1.9 Assay Throughput: 1536-well plate. Stably transfected VM7 cells are aliquoted into 1536-well microtiter plates and incubated with test compounds for 24 hours prior to monitoring luminescence resulting from ER gene expression.
- 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: Changes to bioluminescence signals produced from an enzymatic reaction catalyzed by luciferase between the key substrate [CellTiter-Glo] and the target cofactor [ATP] are correlated to the viability of the system.

The Tox21 VM7 estrogen receptor alpha agonism luciferase assay screened a library of diverse environmental compounds to probe for xenoestrogenic ligand-binding and potential to inhibit estrogen-dependent transcription, as monitored through luciferase reporter gene signal activity using an endogenous full-length ERalpha cell line of human breast tissue (VM7) that was stably transfected with an estrogen-responsive luciferase reporter gene plasmid (pGudLuc7ere). The assay is run in triplicate on 1536-well microplate and

bioluminescence was measured following 24 hour incubation of cells with test compounds and 30 min incubation of test system with ONE-Glo™ luciferase assay reagent. The bioluminescent signal was monitored using a Promega ViewLux plate reader to measure agonistic activity, this assay is performed with small amounts of an ER alpha antagonist (ICI182780) added to each well and each compound is evaluated against a known ER alpha agonist (beta-estradiol, E2) as a positive control (100 percent inhibition). Test compounds were assayed for cytotoxicity by measuring protease activity with Promega CellTiter-Fluor with tetraoctylammonium bromide as a positive control for cell death. Compound auto-fluorescence was monitored in various "TOX21\_AutoFluor" assays run at interfering wavelengths to allow for background artifact detection.

- 2.2 Scientific Principles: Luciferase reporter-gene assays are a commonly used bioluminescence assay. The construct includes a promoter region of a gene of interest followed by a luciferase gene. When this is introduced into a cell, luciferase is expressed in quantities that are proportional to the promoter activity. The luciferase (and hence the promoter activity) can then be quantified by the measurement of the luminescence produced when the enzyme substrate is added. In this way, the transcriptional activity of the gene of interest (i.e., its expression) can be measured in response to the effects of different modulators of the relevant signaling pathways. The luciferase reaction can also be used in combination with constitutively active promoters, to investigate cytotoxicity or transfection efficiency. Estrogen receptor (ER), a nuclear hormone receptor, plays an important role in development, metabolic homeostasis and reproduction. Endocrine disrupting chemicals (EDCs) and their interactions with steroid hormone receptors like ER causes disruption of normal endocrine function. Therefore, it is important to understand the effect of environmental chemicals on the ER signaling pathway. To identify ER agonists, VM7-Luc-4E2 cell line (provided by Dr. Michael Denison from University of California) has been used to screen the Tox21 library of diverse environmental compounds. VM7-Luc-4E2 cell line endogenously expresses full-length ER-alpha and is stably transfected with a plasmid containing four estrogen responsive elements (ERE) upstream of a luciferase reporter gene. The ERalpha\_LUC\_VM7 assays are qHTS format assays which measure the ability of a chemical to interact with estrogen receptor alpha (ERalpha) by monitoring modulation of fluorescence reporter gene signals. This assay utilized a human breast (VM7-Luc-4E2) cell line which expresses endogenous full-length ERalpha and a luciferase reporter gene (ER-luc) to quantify xenoestrogenic activity. This assay is intended for use as a part of an integrated testing strategy, to screen a large structurally diverse chemical library for compounds with the potential to interact with estrogen receptor alpha mediated pathways and potentially affect endocrine systems in exposed populations. There is strong evidence that estrogen receptor agonism is the Molecular Initiating Event (MIE) in an Adverse Outcome Pathway (AOP) leading to reproductive dysfunction in oviparous vertebrates, and there is some evidence that estrogen receptor activation is the MIE for putative AOPs leading to reduced survival due to renal failure and leading to skewed sex ratios due to altered sexual differentiation in males (all AOPs currently under development). Chemical-activity profiles derived from this assay can inform prioritization decisions for compound selection in more resource intensive in vivo studies to further investigate the involvement of ER agonism in pathways leading to hazardous outcomes in biological systems.
- 2.3 Experimental System: adherent VM7 cell line used. Michigan Cancer Foundation-7 (MCF-7) cells are a breast cancer cell line originating from a 69-year old woman in 1970 (Lee et al. 2015). This is an immortalized cell line which endogenously expresses receptors for estrogen (ESR1 and ESR2) and progesterone (Geisinger et al. 1989) as well as growth factors EGF and IGF-1 (Baldwin et al. 1998), and so provides an alternative to breast cell lines for estrogen-sensitive proliferation assays. VM7-Luc-4E2 cells are VMCF7 cells which are stably transfected with plasmid containing four estrogen responsive elements upstream of a luciferase reporter gene (Rogers and Denison 2000) to measure transactivation activity occurring along estrogen signaling pathways. This was previously the BG1-Luc estrogen receptor transactivation test method, but BG1Luc4E2 cells are being renamed VM7Luc4E2 cells because recent DNA analysis (STR) revealed that the original cell line used to generate the BG1Luc4E2 cells were not human ovarian carcinoma (BG-1) cells but a variant of human breast cancer (MCF7) cells.
- 2.4 Metabolic Competence: Xenobiotic biotransformation potential has not been characterized.

- 2.5 **Exposure Regime:** Quality Control Precautions: Maintain cells below 85-90% confluence in culture medium. For assay purpose, the cells should be cultured in assay medium with 10% charcoal stripped FBS for 5 days with alternate day medium changed to fresh medium. Especially while in assay culture, the cells should not reach more than 85% confluence as they would become harder to detach if they reach over confluence. Cell Thawing Method: Thaw a frozen vial of cells in 9ml of pre-warmed medium and seed them in T175 flask at 2 million cells. Cell Proliferation Method: Trypsinize cells from the flask and centrifuge. Add culture medium to the pellet and passage at 3-4 million per T-225 flask. Assay Protocol: Harvest cells from the 5-day culture in assay medium and resuspend cells in assay medium. Dispense 4000 cells/4uL/well into 1536-well tissue treated white/solid bottom plates using an 8 tip dispenser (Multidrop). Incubate the assay plates for 22hrs at 37C and 5% CO<sub>2</sub>. First 1uL of 10.0nM (final concentration) ICI-182,780 (ER-Antagonist) or assay buffer was added using two separate tips of a dispenser (BioRAPTR). Then transfer 23nL of compounds from the library collection and positive control to the assay plates by using a Pintool station. Incubate the assay plates for 22hrs at 37C and 5% CO<sub>2</sub>. After 21hrs of incubation, 1ul of CellTiter-Fluor(TM) Cell Viability Assay reagent was added using a single tip of a dispenser (BioRAPTR). Incubate the assay plates at 37C and 5% CO<sub>2</sub> for 1hr. Measure fluorescence signal by ViewLux plate reader (Exposure time = 1sec). Then add 4ul of ONE-Glo(TM) Luciferase reagent using a single tip of a dispenser (BioRAPTR). Incubate the plates at room temperature for 30min.

ASSAY DESIGN SUMMARY	
Nominal number of tested concentrations: 15	Target (nominal) number of replicates: 3
Standard minimum concentration tested: 0.0117795391705069 µM	Standard maximum concentration tested: 920.276497695853 µM
Key positive control: tetraoctylammonium bromide	Neutral vehicle control: DMSO
Baseline median absolute deviation for the assay (bmad): 1.88	
Response cutoff threshold used to determine hit calls: 20	
Detection technology used: CellTiter-Glo Luciferase-coupled ATP quantitation (Luminescence)	

- 2.6 **Response:** Agonism of the estrogen receptor (ERα) signaling pathway is measured by bioluminescence activity via an estrogen-responsive firefly luciferase reporter gene. Increased luciferase activity can be used to identify the compounds that induce xenoestrogenic ligand-binding and ERα agonism. The cytotoxicity of the compounds screened was tested in parallel by measuring the cell viability using CellTiter-Fluor assay in the same wells.
- 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:
- Cytotoxicity Burst: Assays used to define the cytotoxicity burst region

Additionally, this assay was annotated to the intended target family of cell cycle.

### 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: Cellular ATP content was detected with CellTiter-Glo Luciferase-coupled ATP quantitation. Decreased luminescence (loss-of-signal) was measured relative to tetraoctylammonium bromide (positive control) signal, using DMSO (neutral control) as a baseline for cell viability. Response was reported as a percent activity.

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;  $onesd = \sqrt{\sum((resp - \text{mean } resp)^2) / \text{sample size} - 1}$ . Onesd is used to establish BMR and therefore required for tcplfit2 processing.)

*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.), 6: bmad6 (Add a cutoff value of 6 multiplied by the baseline median absolute deviation (bmad). By default, bmad is calculated using test compound wells (wllt = t) for the endpoint.), 27: ow\_bidirectional\_loss (Multiply winning model hitcall (hitc) by -1 for models fit in the positive analysis direction. Typically used for endpoints where only negative 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 ( $\text{top} < 0$ ), flag if  $\text{count}(resp < -1 * \text{coff}) < 2 * \text{count}(resp > \text{coff})$ . If gain was winning directionality ( $\text{top} > 0$ ), flag if  $\text{count}(resp > \text{coff}) < 2 * \text{count}(resp < -1 * \text{coff})$ .), 6: singlept.hit.high (Flag single-point hit that's only at the highest conc tested, where series is an active hit call ( $\text{hitc} \geq 0.9$ ) 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 ( $\text{hitc} \geq 0.9$ ) 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 ( $\text{hitc} < 0.9$ ) 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 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);  $|top| \leq 1.2(coff)$  or  $|top| \geq 0.8(coff)$ .), 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;  $nconc \leq 4$ .), 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 ( $hitc \geq 0.9$ ) and efficacy values (e.g. top and maximum median response) less than 50 percent; intended for biochemical assays. If  $hitc \geq 0.9$  and  $coff \geq 5$ , then flag when  $top < 50$  or  $max\_med < 50$ . If  $hitc \geq 0.9$  and  $coff < 5$ , then flag when  $top < \log_2(1.5)$  or  $max\_med < \log_2(1.5)$ .), 18: ac50.lowconc (Flag series with an active hit call ( $hitc \geq 0.9$ ) if AC50 (concentration at 50 percent maximal response) is less than the lowest concentration tested; if  $hitc \geq 0.9$  and  $ac50 < 10^{\log_{10} c_{min}}$ , then flag.), 19: viability.gnls (Flag series with an active hit call ( $hitc \geq 0.9$ ) if denoted as cell viability assay with winning model is gain-loss (gnls); if  $hitc \geq 0.9$ ,  $modl == "gnls"$  and  $cell\_viability\_assay == 1$ , 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: 9667		Number of chemicals tested: 7871
ACTIVITY HIT CALLS		
Active hit count: $hitc \geq 0.9$ 843	Inactive hit count: $0 \leq hitc < 0.9$ 8103	NA hit count: $hitc < 0$ 721
WINING MODEL SELECTION		
Number of sample-assay endpoints with winning <i>hill</i> model:	478	
<i>gain-loss (gnls)</i> model:	547	
<i>power(pow)</i> model:	760	
<i>linear-polynomial (poly1)</i> model:	3278	
<i>quadratic-polynomial (poly2)</i> model:	695	
<i>exponential-2 (exp2)</i> model:	344	
<i>exponential-3 (exp3)</i> model:	106	
<i>exponential-4 (exp4)</i> model:	2877	
<i>exponential-5 (exp5)</i> model:	582	

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 `BMDEExpress2`. 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: <i>nmed</i>	0
Neutral control median absolute deviation, by plate: <i>nmad</i>	3.447
Coefficient of variation (CV%) in neutral control wells: $(nmad/nmed)*100$	Inf%
POSITIVE CONTROL (well type = "p")	
Positive control well median response value, by plate: <i>pmed</i>	-1.397
Positive control well median absolute deviation, by plate: <i>pmad</i>	3.496
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: $((pmed - nmed) / sqrt(pmad^2 + nmad^2))$	-0.264
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: <i>mmed</i>	NA
Negative control well median absolute deviation value, by plate: <i>mmad</i>	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: 582.

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:** Judson RS, Magpantay FM, Chickarmane V, Haskell C, Tania N, Taylor J, Xia M, Huang R, Rotroff DM, Filer DL, Houck KA, Martin MT, Sipes N, Richard AM, Mansouri K, Setzer RW, Knudsen TB, Crofton KM, Thomas RS. Integrated Model of Chemical Perturbations of a Biological Pathway Using 18 In Vitro High-Throughput Screening Assays for the Estrogen Receptor. *Toxicol Sci.* 2015 Nov;148(1):137-54. doi: 10.1093/toxsci/kfv168. Epub 2015 Aug 13. PMID: 26272952; PMCID: PMC4635633., Huang R, Sakamuru S, Martin MT, Reif DM, Judson RS, Houck KA, Casey W, Hsieh JH, Shockley KR, Ceger P, Fostel J, Witt KL, Tong W, Rotroff DM, Zhao T, Shinn P, Simeonov A, Dix DJ, Austin CP, Kavlock RJ, Tice RR, Xia M. Profiling of the Tox21 10K compound library for agonists and antagonists of the estrogen receptor alpha signaling pathway. *Sci Rep.* 2014 Jul 11;4:5664. doi: 10.1038/srep05664. PMID: 25012808; PMCID: PMC4092345.
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.