

---

# ToxCast Owner's Manual – Guidance for Exploring Data

## Resources

- [Toxicity Forecaster \(ToxCast\) Fact Sheet](#)
- [ToxCast Publications](#)
- [About ToxCast](#)

ToxCast is a multi-year effort launched in 2007 that uses automated chemical screening technologies called high-throughput screening assays to expose living cells, isolated proteins, or other biological molecules to chemicals. The cells or proteins are then screened for changes in biological activity that may suggest potential toxic effects.

These innovative methods have the potential to limit the number of required laboratory animal-based toxicity tests while quickly and efficiently screening thousands of chemicals for potential health effects.

## Generating ToxCast Data

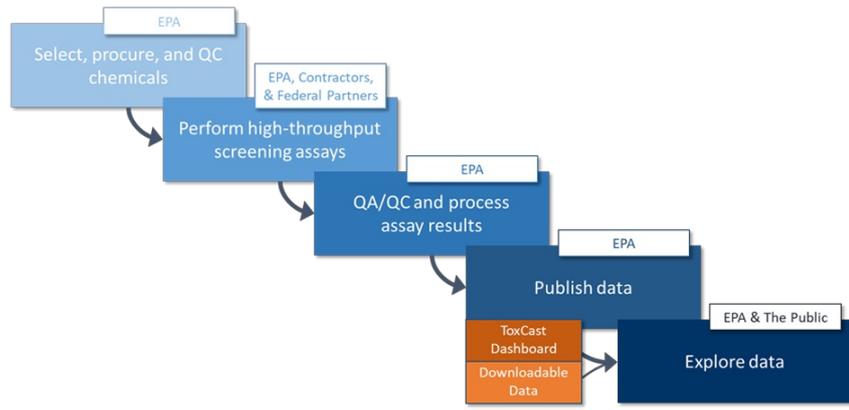
- [Chemicals](#)
- [Assays](#)
- [Results Processing and Analysis](#)
- [Publicly Available Data](#)

## Exploring ToxCast Data

- [ToxCast Data](#)
- [Accessing ToxCast Data and Scenarios for Exploring Data](#)
- [Citations for ToxCast Data](#)

The figure below provides a visual overview of the process ToxCast follows to create an accessible database of high-throughput screening information on thousands of chemicals.

**Click on the boxes in the figure below to link to more information.**



# ToxCast Chemicals

## Resources

- [Chemical Inventory](#)
- [Chemical Landscape paper](#)
- [About ToxCast](#)

EPA's ToxCast research effort has a chemical library of approximately 4,400 unique chemicals (as of December 2017). The library includes a broad range of chemicals with potential for human and ecosystem exposure which have heightened regulatory concern yet are lacking data on potential health effects of these chemicals.

The chemicals in the library have been screened in hundreds of high-throughput screening assays to generate biological activity data. The chemical library is organized into three phases.

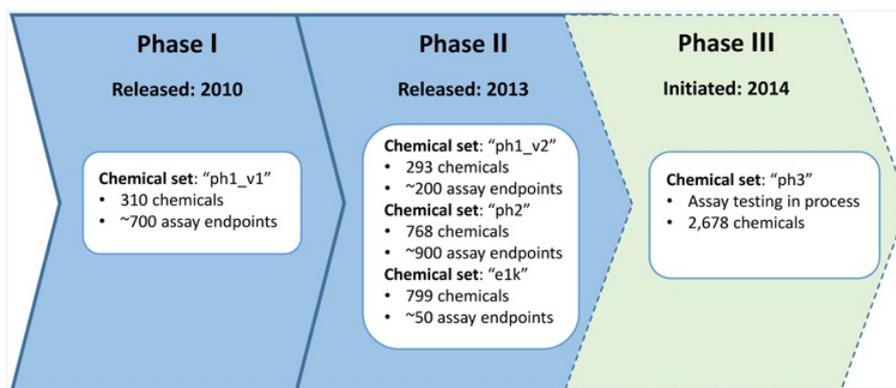
In 2007, **Phase I** was launched, screening 310 chemicals (mostly pesticides) across hundreds of assay endpoints. **Phase II** data increased the chemical inventory to about 1,800 chemicals. In 2014, **Phase III** was initiated, increasing the library to over 4,500 chemicals.

Below are links to information about the various ToxCast chemical lists and the workflow EPA follows to select chemicals, procure chemicals, register chemicals, conduct a quality review of the chemicals, and prepare the chemicals for high-throughput screening.

## Links

- [Chemical Lists](#)
- [Chemical Workflow](#)

## ToxCast Phases, Timing, and Chemical and Assay Endpoint Counts



---

# ToxCast Data Generation: Chemical Lists

## Resources

- [Chemical Inventory](#)
- [Chemical Landscape paper](#)
- [CompTox Chemistry Dashboard](#)
- [About ToxCast](#)

The list of EPA's ToxCast chemicals with publicly available high-throughput screening data can be exported from the ToxCast Dashboard. EPA's ToxCast Chemical lists that includes chemicals currently in process and chemicals already screened can be downloaded from the FTP site.

## FTP Download

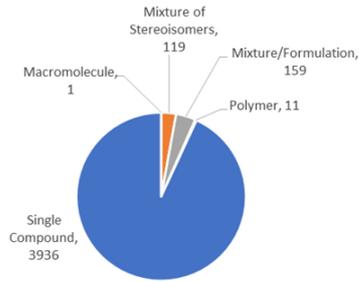
The file labeled DSSTox\_TOXCST\_20160129.zip contains readme files as well as an Excel file containing chemical information and ToxCast phase information for each of the substances screened.

## ToxCast Dashboard

Descriptions of the chemicals, their structures, and physicochemical properties are available via the ToxCast dashboard. Users can select chemicals of interest using a number of filters. Chemical data filters include Chemical Abstract Service Registry Number (CASRN), chemical name, chemical category, use category, and physicochemical properties such as the octanolwater partition coefficient (log P). For each chemical, the ToxCast Dashboard summarizes chemical information.

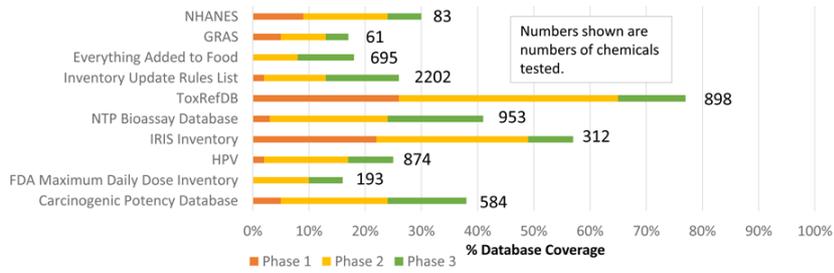
## Chemical Summary Charts

### Chemicals by Substance Type

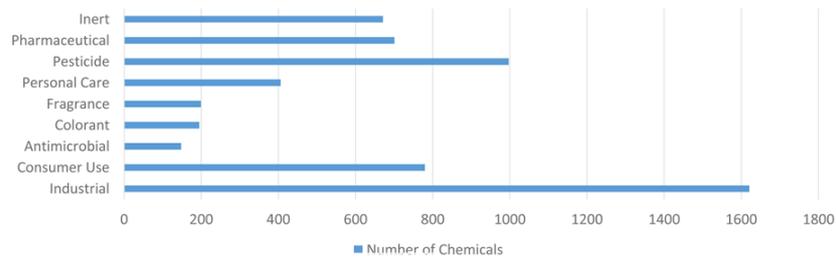


### Chemicals by Inventories of Interest

Coverage of Databases of Regulatory Interest



Numbers of Tested Chemicals in Use Categories of Interest



## Links

- [Chemical Workflow](#)
- [ToxCast Assays](#)

# ToxCast Data Generation: Chemical Workflow

This page describes the process EPA follows to select chemicals, procure chemicals, register chemicals, conduct a quality review of the chemicals, and prepare the chemicals for high-throughput screening. There is also information about each phase of testing.

## Resources

- [Chemical Inventory](#)
- [Chemical Landscape paper](#)
- [About ToxCast](#)

## On This Page

- |  |   |
|--|---|
| <ul style="list-style-type: none"><li>• <a href="#">Chemical Selection and Lists</a><ul style="list-style-type: none"><li>◦ <a href="#">Phase I</a></li><li>◦ <a href="#">Phase II</a></li><li>◦ <a href="#">Phase III</a></li></ul></li><li>• <a href="#">Procuring and Registering Chemicals</a></li></ul> | <ul style="list-style-type: none"><li>• <a href="#">Quality Review of Chemicals</a></li><li>• <a href="#">Chemical Preparation Process and Timeline</a></li><li>• <a href="#">Links</a></li></ul> |
|--|---|

## Chemical Selection and Lists

### Phase I

#### Chemical Types

ToxCast Phase I screened 310 unique chemicals, mostly pesticides. The pesticides in Phase I included diverse chemical structures and a broad range of pesticidal mechanisms. The chemicals already had significant toxicity data (approximately 275 compounds had near complete [guideline data coverage](#)) and a variety of chemical reactivity features and mechanistic diversity.

Pesticides were selected to be screened first because of EPA's extensive animal toxicity data from testing pesticides. The ToxCast screening results from Phase I were compared to these animal toxicity study results as proof-of-concept for using ToxCast to evaluate chemicals for potential health effects.

Thirty non-pesticidal environmental chemicals of research or regulatory interest were also evaluated in Phase I, including:

- Perfluorinated compounds - Perfluorooctanoic acid (PFOA), a surfactant used in fluoropolymers such as Teflon; and Perfluorooctanesulfonic acid (PFOS), used in the semiconductor industry, ScotchGuard™ formulations, and flame retardant foams.
- Bisphenol A (BPA) and a set of phthalate alternatives used as plasticizers.
- 10 toxicologically active metabolites of included phthalates and pesticides such as monobutyl phthalate (a metabolite of dibutyl phthalate) and diazoxon (a metabolite of diazinon).

## ToxCast Screening

The original 310 chemicals tested in ToxCast Phase I are referred to as “ph1\_v1”. A listing of the ToxCast ph1\_v1 chemicals, including modified or discontinued compounds labeled and issues annotated, is provided in the ToxCast chemical inventory file (DSSTox\_TOXCST, date stamped 20160129) available from the EPA DSSTox Data download page.

## Toxicity Data

An extensive amount of toxicity data was already available for most Phase I chemicals and included data for subchronic, chronic, multigenerational reproductive, and developmental endpoints across several species (rat, mouse, dog and rabbit). The data was available in data evaluation records from the US EPA's Office of Pesticides Programs' review of registrant study data. EPA researchers extracted the data from [EPA's Data Evaluation Records](#) and systematically captured these data in a database of animal toxicity data called the [Toxicity Reference Database](#) (ToxRefDB).

[Top of Page](#)

## Phase II

### Chemical Types

Phase II retained most of Phase I and expanded the chemicals screened to include data-poor chemicals of research and regulatory interest. 293 unique chemicals from Phase I (ph1\_v2) were chosen, excluding chemicals which we learned from Phase I were not suitable for high-throughput screening. These excluded chemicals were 14 sulfurons determined to undergo rapid hydrolysis in DMSO, 3 chemicals deemed insufficiently soluble in DMSO, and 1 chemical originally tested in parent form that was instead tested in a salt form. 768 unique new ToxCast chemicals identified through a selection process were added in Phase II (ph2). ToxCast Phase II also includes a list of 799 chemicals of interest to EPA's Endocrine Disruption Screening Program (e1k). Specifically, the e1k library contains many known estrogen receptor (ER) and androgen receptor (AR) active reference chemicals. The e1k list is specifically identified because the chemicals were not screened in the entire set of Phase II high-throughput assays, but a specific subset for evaluation of potential endocrine-related activity.

The majority of Phase II chemicals were purchased from commercial suppliers and approximately 150 chemicals were donated by research collaborators external to EPA. Some chemicals were donated by the chemical industry and the U.S. Food and Drug Administration's National Center for Toxicological Research, (“green” plasticizer alternatives and reference liver toxicants, respectively). Pharmaceutical companies donated the remaining 136 “failed pharma”

compounds (i.e., drug candidates discontinued due to toxicity in preclinical or clinical trials). A complete list of the donated pharmaceuticals (“donated pharma” chemical list) is available on the ToxCast chemical library file (DSSTox\_TOXCST, date-stamped 20160129), available from the [EPA DSSTox Data download page](#).

## **Toxicity Data**

The donation of failed drug compounds, along with some preclinical and clinical data, introduced to ToxCast a set of chemicals designed to be bioactive at specific human (or veterinary) targets. All compatible donated data on these failed pharmaceuticals were added to [EPA’s Toxicity Reference database](#).

## **ToxCast Screening**

ToxCast Phase II encompassed screening of the newly added ph2 inventory in the majority of the original Phase I ToxCast assays, as well as testing of the ph2 and reprocurd ph1\_v2 inventory in newly acquired ToxCast and [Tox21](#) assays. In addition, the e1k chemicals were screened in a limited subset of Phase II endocrinelrelated assays, including many assays run using the Tox21 robotics technology. Phase II concluded in 2015 with the [public release of ToxCast assay data for 1,863 unique ToxCast chemicals](#) screened to that point (1,878 if discontinued ph1\_v1 chemicals are included in the total), including data for the ph1\_v1, ph1\_v2, ph2, and e1k chemical sets.

[Top of Page](#)

## **Phase III**

### **Chemical Types**

ToxCast Phase III launched in late 2014. Phase III added new assays, endpoints, and chemicals. Five hundred chemicals and a small set of donated mixtures and water samples were added to ToxCast. Phase III also added selected chemicals from [Tox21](#) and e1K to test in the new assays. Unique Phase III chemicals include flame retardants and chemicals of interest to EPA’s [Endocrine Disruption Screening Program](#) (EDSP).

With the addition of Phase III chemicals, the total combined set of chemicals comprising ToxCast is over 4,200 unique chemicals. The screening and analysis of data from Phase III chemicals is currently underway (Note: this number of chemicals does not include chemicals only tested in Tox21).

### **ToxCast Screening**

Phase III chemicals are not being run through all ToxCast assays. Using the results from Phase I and Phase II, EPA identified the assays most appropriate for the Phase III chemicals selected.

The chemical list for ToxCast Phase III is abbreviated as “ph3” chemicals and does not overlap with earlier chemical inventories having full or partial testing coverage in Phases I and II (ph1\_v1, ph1\_v2, ph2, and e1k). Each of these inventories, as well as the testing phase in which the chemical was screened (I, II, and/or III) is labeled within the ToxCast chemical inventory file (DSSTox\_TOXCST, date-stamped 20160129), available from the [EPA DSSTox Data download page](#).

# Procuring and Registering Chemicals

EPA follows a standard process for procuring and registering chemicals which includes purchasing chemicals from a supplier and working closely with a contractor to register and catalogue the chemicals. The majority of ToxCast chemicals are purchased from commercial suppliers.

1. First, EPA provides the supplier a list of generic CAS and chemical names that need to be purchased. EPA may also provide chemical structure information in the form of [SMILES](#) to the contractor to help with structure searching of larger aggregated commercial chemical services, such as [eMolecules](#) [EXIT](#) or [ChemNavigator](#) [EXIT](#).
2. EPA purchases two bottles of analytical grade samples (> 98% purity) from a supplier. A majority of samples are then shipped directly from the supplier to the contractor in pre-sterilized, barcoded vials. One of two bottles is designated for solubilization and the other stored in neat powder form. The chemical supplier also provides a chemical structure file and molecular weight for defined pure compounds (SMILES or Structure Data Format [SDF] file).
3. When the contractor receives the chemicals, bottles are scanned, weighed, registered into the contractor's chemical management tracking system, and either immediately solubilized, or stored in a -20°C freezer under inert conditions until a solubilization order is placed. The contractor's chemical management tracking system records the Supplier, Catalog number, Lot number, the Contractor shipment sample code, SMILES, molecular weight (usually derived from the structure file), physical form of the chemical received (solid or liquid), quantity (ul or mg), and the date the sample was registered in the system. In cases where a chemical name, CAS or structure are not provided, the contractor attempts to fill in this information through reference to the original orders or supplier website catalogs, or through internal database matches to the compound structures provided.

A small percentage of the chemicals deviate from the standard process (less than 10% of Phase II and Tox21 chemicals). Most common deviations from this process are the use of specialty suppliers and procurement of larger or smaller quantities for hard-to-locate chemicals.

EPA requires, wherever possible, a Certificate of Analysis (COA) and Material Safety Data Sheet (MSDS) from the chemical supplier. The chemical supplier provides the COA and MSDS to the contractor and the contractor provides it to EPA. Occasionally samples are accepted by EPA without documentation when the supplier has problems obtaining the documentation (COAs are missing for approximately 15% of unique lot/batch chemical samples).

## Quality Review of Chemicals

The goal of chemical quality check (QC) is to have accurate and reliable information on every chemical undergoing ToxCast screening. There are practical limits as well as a balance that must be struck between cost/time/efficiency and the larger objectives of ToxCast. In the absence of perfect knowledge and certainty, the [ToxCast chemical QC process](#) minimizes controllable sources of errors, particularly in the chemical information QC review and registration process, but also in handling and storage procedures. At the same time, the ToxCast analytical QC processes attempt to detect, understand, document, and communicate actual errors and problems impacting particular chemicals under testing conditions.

In the QC process, EPA may request its chemical management services contractor to analyze chemical samples to confirm the supplier's identification of the chemical sample and to estimate sample purity (ideally >90% for commercial grade and >99% for pharmaceutical or pestinal analytical grade samples) using appropriate analytical methods. Analytical QC provides an experimental standard of verification and is required to confirm the chemical identity and purity in the plated DMSO solutions undergoing testing at the time of plating, as well as at later time points (to assess sample stability over time).

The QC analysis requires the availability of analytical methods suitable to analyzing the various types of chemical compounds derived from the chemical domains identified. The standard analysis would consist of high-throughput, low-resolution liquid chromatography mass spectrometry (LCMS), using UV and ELS detection methods run in positive and negative ion mode, of solutions in microtiter 96 well plates.

The types of documentation QC review include:

- Certificate of Analysis (COA) Chemical Validation – To establish chemical identity (chemical name, CAS, MW) from the supplier provided COA.
- DSSTox Chemical Information Review – To ensure accurate and consistent substance (CAS, name, description) and structure annotation of the generic chemical (independent of supplier, lot, batch) as part of the DSSTox chemical registration process.

The results of the analytical QC review include:

- An analytical determination and report supplied to EPA in electronic format.
- All spectra and supporting information, along with summary purity and identity confidence score.

[Top of Page](#)

## **Chemical Preparation Process and Timeline**

EPA has developed a step-by-step process to prepare chemicals for sending to EPA contract laboratories and to those requesting the chemical library for

research purposes. This process ensures that the neat chemical or DMSO solution stocks are provided at quantities and concentrations primarily designed for high-throughput testing applications, to be run in assays deemed of sufficient value to EPA's ToxCast mission.

The process is as follows:

1. Requestors submit an EPA ToxCast Chemical Request Form outlining the number of chemicals, quantities, and concentrations needed; background information on the proposed assay testing; pertinent plating details; and the recipient of the shipment and desired shipment date.
2. The National Center for Computational Toxicology (NCCT) Chemical Management reviews the request, prepares the order, and provides it to the NCCT chemical services contractor.
3. Based on the order, the contractor aliquots, packages, and ships chemical samples in various formulations (neat, dried down, or plated solutions), with sufficient quantities of dry ice (to ensure no thawing), to the designated laboratories or investigators within or outside the U.S.

It takes an average of 2-6 weeks (depending on whether chemical stocks are available for plating) to process a chemical order request from receipt of a request to shipment of the samples.

The chemical library is typically requested by contractors hired by EPA to screen the chemicals, EPA research collaborators with a signed agreement and EPA staff. External organizations requesting the chemical library must agree to NCCT's policy of only providing blinded plate maps during assay testing. These organizations should provide EPA NCCT raw or processed screening results prior to the unblinding of the plates. However, when collaborators have a legitimate need for unblinding earlier in testing (e.g., for method development only) or additional details, such as actual concentration, NCCT will provide either partially or fully unblinded plate maps.

[Top of Page](#)

## Links

- [Chemical Lists](#)
- [ToxCast Assays](#)

[Top of Page](#)

---

# ToxCast Data Generation: ToxCast Assays

## Resources

- [Toxicity Forecaster \(ToxCast\) Fact Sheet](#)
- [ToxCast Publications](#)
- [About ToxCast](#)

ToxCast uses automated chemical screening technologies, called high-throughput screening assays, to expose living cells or biological macromolecules such as proteins to [chemicals](#). The cells or isolated proteins are then screened for changes in biological activity that may suggest potential health hazards.

- **Cell-based in vitro assays** use intact cells to measure cellular changes in response to exposure with the test substances. Example cells include human primary cells and cell lines and rat primary liver cells. An example is [ACEA\\_T47D](#).
- **Biochemical in vitro assays** measure the activity of a biological macromolecule, either a purified protein or a nucleic acid. An example is [NVS\\_ADME\\_hCYP19A1](#).

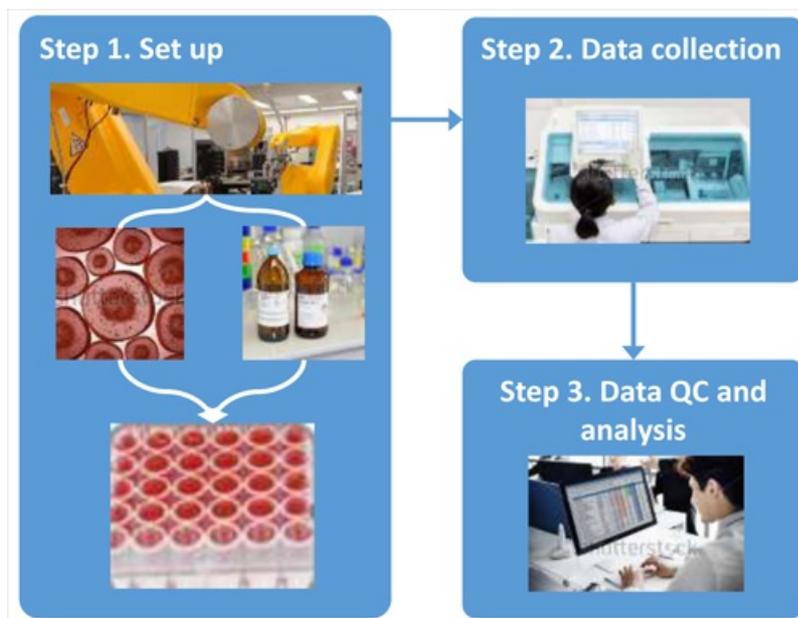
ToxCast uses more than 700 high-throughput assays that cover a range of high-level cell responses and approximately 300 signaling pathways.

Click links below for an overview of ToxCast assays and the platform sources, descriptions, and protocols to test chemical activity in high-throughput screening.

## Assay Information

- [Overview of ToxCast Assays](#)
- [ToxCast Assay Platform Sources, Descriptions, and Protocols](#)

## Assay Process Overview



**Step 1.** High-throughput technology perform repetitive motions of loading the test chemicals, positive/negative control chemicals, dilution solvents, reagents, and cells (if applicable) into plates containing numerous wells (e.g., a 384-well plate).

**Step 2.** The well plates are incubated for different time durations depending on the assay. Detection technologies translate the biological or physical process into a signal (e.g., fluorescence). The signals are measured and compiled into databases.

**Step 3.** Results are quality control checked, normalized, analyzed, and formatted for release. For more information visit [Results Processing and Analysis](#).

---

# ToxCast Data Generation: Overview of ToxCast Assays

## Resources

- [Toxicity Forecaster \(ToxCast\) Fact Sheet](#)
- [ToxCast Publications](#)
- [About ToxCast](#)

There are more than 15 commercial or federal government [platform sources](#) used for ToxCast assays. The ToxCast assays range from simple, single-readout to complex, multiplexed and multi-parametric technologies.

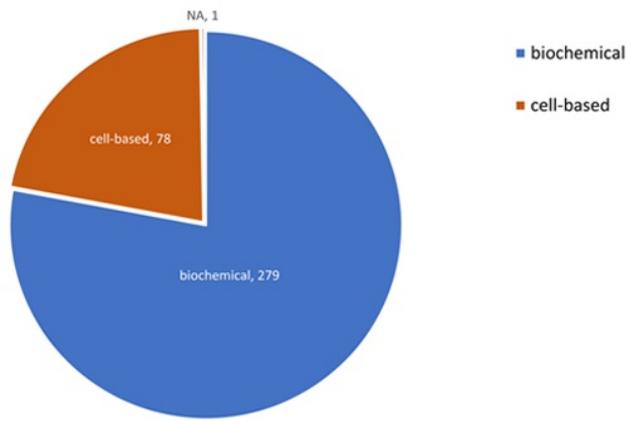
Each assay can generate one or more measured entities or technological targets (called assay components) and one or more assay endpoints (e.g., readouts such as the concentration at 50% of maximum activity or the AC50). With few exceptions, the complete ToxCast Phase I and II chemical sets have been tested across 700+ assay endpoints of data.

The activity of a chemical in a specific assay does not necessarily mean that it will cause toxicity or an adverse health outcome. There are many factors that determine whether a chemical will cause a specific adverse health outcome. Careful review is required to determine the use of the data in a particular decision context.

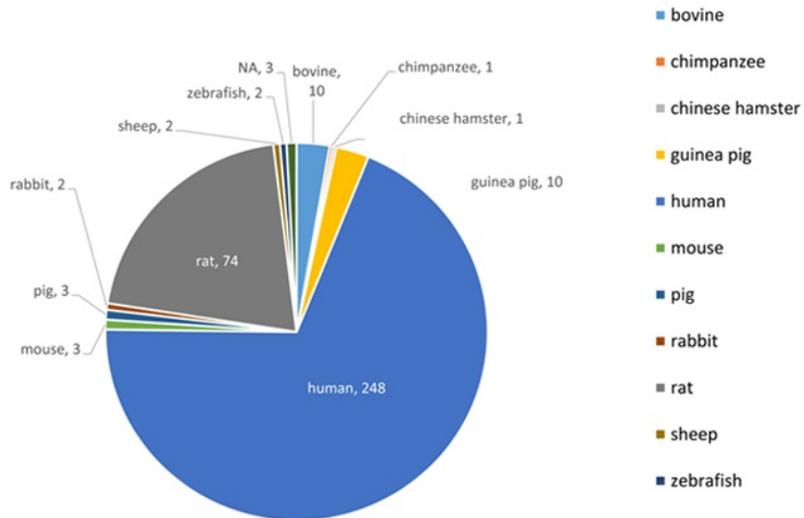
The charts and tables on this page summarize characteristics of the ToxCast assays by the following categories:

- [Assay by Assay Function Type](#)
- [Assays by Organism Cell Type](#)
- [Assay Design Type](#)
- [Assay by Biological Activity Type](#)
- [Cell Based Assays by Cell Line](#)

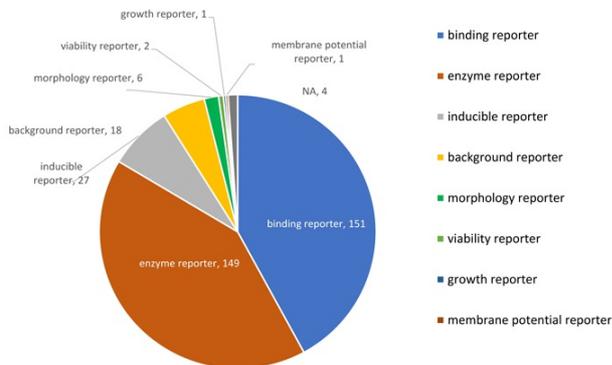
### Assays by Assay Function Type



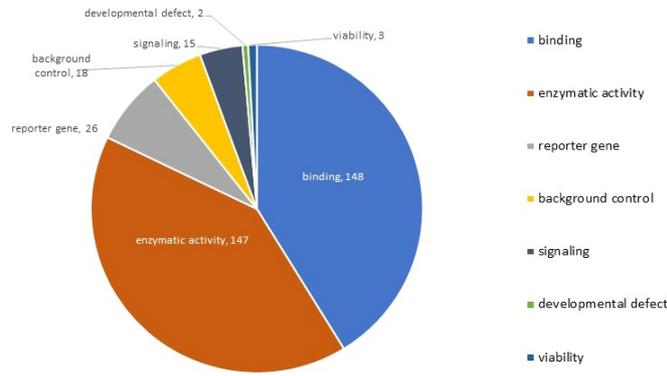
### Assays by Organism Cell Type



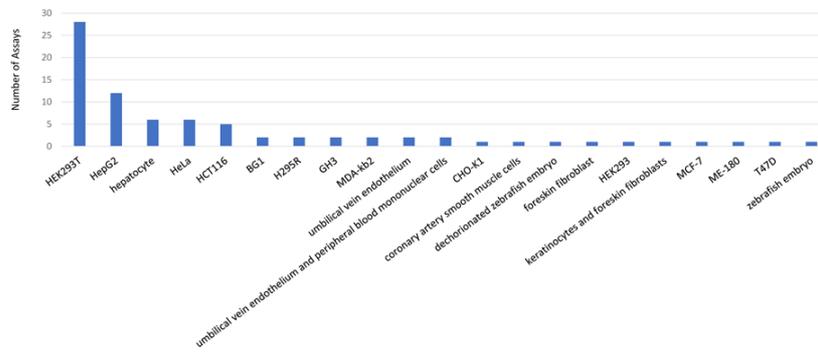
### Assays by Assay Design Type



### Assays by Biological Activity Type



### Cell Based Assays by Cell Line



## Links

- [ToxCast Assay Platform Sources, Descriptions, and Protocols](#)

# ToxCast Data Generation: Assay Platform Sources, Descriptions, and Protocols

## Resources

- [Toxicity Forecaster \(ToxCast\) Fact Sheet](#)
- [ToxCast Publications](#)
- [About ToxCast](#)

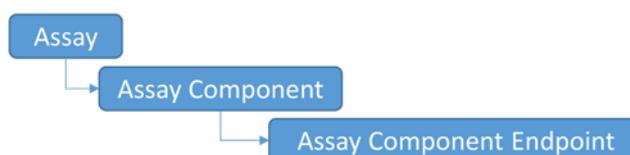
## Assay Platform Sources

Download [ToxCast Assay Platform Source Description](#) document to learn more about the platform sources of the ToxCast assays.

## Assay Descriptions

Descriptions of the ToxCast assays, their components, and endpoints are available via the ToxCast dashboard or a downloadable CSV file. This information is essential for proper interpretation of screening results and facilitates higher-level analyses.

- In the [ToxCast Dashboard](#), click on Choose a View on the top left, select Assays. Click on the plus sign symbol in front of the assay name to expand the entry to show the short description of the assay. To see additional assay information, click on the assay to highlight it, and click on the Assay Summary tab on the right side of the dashboard.
- [Assay Summary Information](#) provides descriptions of each ToxCast assay, which includes its measured readouts and analyzed results.
- [OECD Test 211 ToxCast Assay Documentation](#) provides descriptions and guidelines for ToxCast endocrine-related assays in format outlined by the OECD Guidance Document 211 for describing non-guideline in vitro test methods.



[EPA's ToxCast Assay Annotation Version 1.0 Data User Guide](#) [Version 2.0 Coming Soon] documents the annotations to assist ToxCast data users in understanding how to use the annotation terms within the ToxCast screening data. It explains

annotation concepts and terminologies.

# Assay Protocols

## **Tox21**

[EPA's standard laboratory protocols](#) for [Tox21](#) assays including descriptions, protocols (reference, quality control, procedures and performance), and assay data.

National Institutes of Health Tox21 assay descriptions and protocols are also available on the National Center for Biotechnology Information's [PubChem Bioassay database](#).

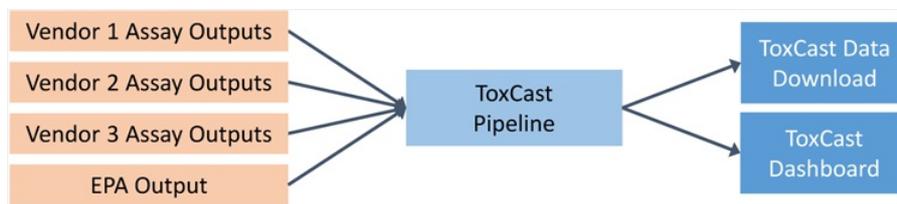
# ToxCast Data Generation: Processing and Analysis

## Resources

- [ToxCast Pipeline \(tcpl\)](#)
- [How to setup pipeline](#)
- [Overview of R package](#)
- [ToxCast Pipeline R Package](#)
- [ToxCast pipeline article](#)
- [About ToxCast](#)
- [ToxCast Data](#)

Once the [testing](#) is complete, the labs send the resulting data back to EPA for processing and quality assurance/quality checking. The diversity of the ToxCast data received from the numerous vendors led to the development of a flexible [high-throughput screening \(HTS\) data analysis pipeline](#) capable of efficiently processing and storing large volumes of data. The prescribed procedure EPA uses to process this data is outlined below, and the pipeline used to process data is [available as an R package \(tcpl\)](#). While developed primarily for ToxCast, EPA has attempted to make the tcpl package generally applicable to the community for processing any high-throughput chemical screening data.

1. The data, received in unique formats from each vendor, are transformed to a standard computable format and loaded into the ToxCast database (invitrodb) by vendor-specific R scripts.
2. Once data is loaded into the database, the ToxCast Pipeline (tcpl) uses the generalized processing functions implemented in the package to process, normalize, model, qualify, flag, inspect, and visualize the data.
3. The data resulting from the processing is available for download and in the ToxCast Dashboard.



---

# ToxCast Data Generation: ToxCast Pipeline (tcpl)

### Resources

- [ToxCast Pipeline R Package](#)
- [ToxCast Pipeline Article](#)
- [About ToxCast](#)

The ToxCast pipeline (tcpl) is an R package (R Core Team, 2016) that provides storage, normalization, dose-response modeling and visualization solutions for HTS screening efforts. The tcpl package provides functionality for two screening paradigms:

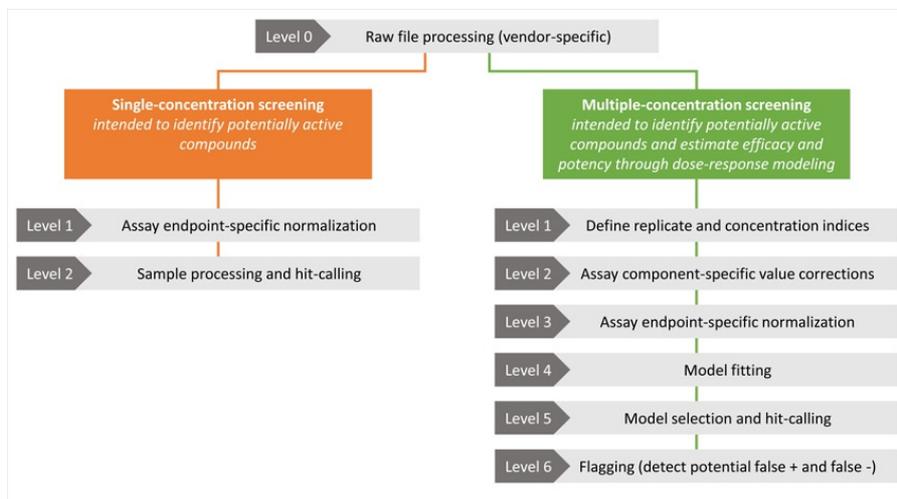
1. single-concentration screening, intended to only identify potentially active compounds, and
2. multiple-concentration screening intended to identify potentially active compounds and estimate the efficacy and potency through dose-response modeling

Based on the screening paradigm, the package processes data through a set of Levels as outlined in the figure below. The tcpl database stores the same raw information at Level 0 for each screening paradigm. To address both data storage and reproducibility issues the tcpl package interacts with a MySQL relational database to store all data and processing decisions made by the user. In addition to storing the data at every level of processing, the accompanying database stores chemical and assay annotations to facilitate further analyses and disseminating results.

The final processed data, including hit calls, can be found on the [ToxCast Dashboard](#).

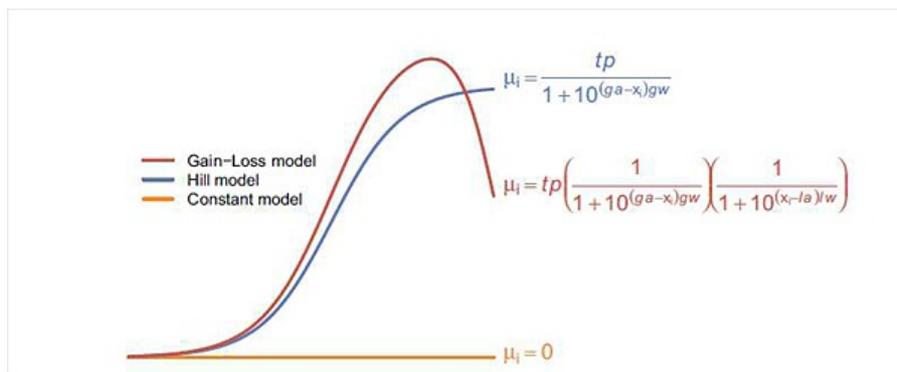
For more details on downloading and using the tcpl, please [click here](#). Detailed information about all processing steps is included in the package vignette.

## Data Analysis Pipeline Overview



## Dose-Response Modeling

The dose-response modeling procedure was designed to handle the outliers and cytotoxicity-related loss of signal common to HTS efforts. The modeling methodology evaluates each data series with three models: (i) a constant model at zero (orange in figure), (ii) a constrained three-parameter Hill model (blue) and (iii) a constrained five-parameter gain-loss model (red).



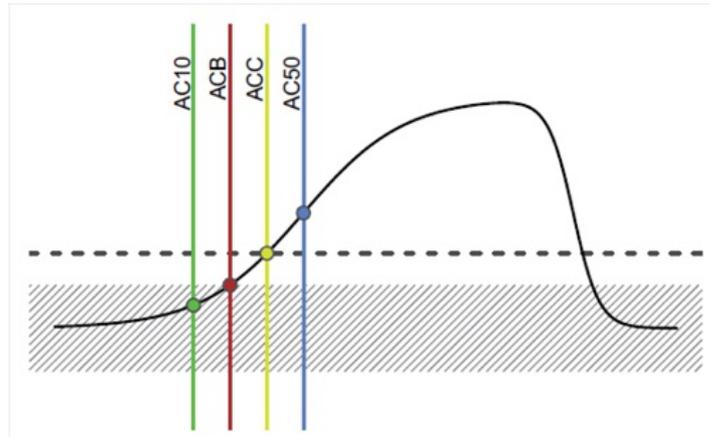
To facilitate simple cross-experiment comparisons and reduce the parameter space, dose-response modeling is constrained to a zero-centered, positive response paradigm. Therefore, negative response data requires an inverse transformation during the normalization process. To obtain robust results without removing any data, we define the log-likelihood using a Student's t-distribution with 4 degrees of freedom (Lange et al., 1989). Utilizing the Student's t-distribution and the gain-loss model provide robust solutions to the problems of outliers and the cytotoxicity-related signal loss that commonly occurs at high concentrations. After fitting each model, the model with the lowest Akaike information criterion value is selected as the 'winning' model.

## Hit-Calling and Point-of-Departure Estimates

A dose-response series must meet three criteria to have an active hit call: (i) the Hill or Gain-Loss model must win and (ii) the modeled curve fit top must exceed the efficacy cutoff, and (iii) for at least one concentration the median response value must also exceed the efficacy cutoff. The efficacy cutoff is a user-defined value selected from an expandable list of methods at Level 5.

In addition to the standard AC50 (activity concentration at 50% of maximal activity) provided for the Hill and Gain-Loss models, the tcpl package provides

three point-of-departure (POD) estimates for the winning model. The AC10, or activity concentration at 10%, is derived solely from the model parameters. Conversely, the ACB (activity concentration at baseline) and ACC (activity concentration at cutoff) are based on levels of noise and significance, respectively.



The package estimates the noise of an assay by calculating the median absolute deviation over all response values given by the first two concentrations (bmad). The baseline region is then defined as  $0 \pm 3bmad$ , and the ACB is the concentration at which the model first reaches  $3bmad$ . Similarly, the ACC is the concentration at which the model first reaches the user-defined cutoff value for a data-series to be considered active. The four POD estimates are illustrated in the figure.

No POD estimates are calculated when the constant model is the selected model winner, because the POD estimates do not apply.

## Links

- [Publicly Available Data](#)

# Exploring ToxCast Data: Overview of tcpl R Package

## Resources

- [Toxicity Forecaster \(ToxCast\) Fact Sheet](#)
- [ToxCast Publications](#)
- [ToxCast Citation](#)
- [About ToxCast](#)

The ToxCast pipeline (tcpl) is an R package that serves as an interface for dose-response modeling and visualization for high-throughput screening efforts based on the data generated by ToxCast. For advanced users who would like to use the tcpl package to process their own data or review the ToxCast data processing, the R package is [available for download](#) `EXIT`. Users will need experience with R to use these files. The R package includes:

- A reference manual.
- A vignette describing the package in detail.
- Exercises to help users work through and understand the vignette.
- A list of package functions.

The user can choose to use the tcpl package (v2.0) with a [MySQL database](#) `EXIT` or with comma-delimited (\*.csv) files that do not require MySQL. A full description of the tcpl configuration options is available in the R manual for tcpl v2.0.

See [How to Set up R and MySQL](#) if needed before using.

The package includes the following sections to describe the functionality of the pipeline.

1. Introduction
2. Register and Upload New Data
3. Data Processing and the tcplRun Function
4. Data Normalization
5. Single-concentration Screening
  - a. Level1
  - b. Level2
6. Multiple-concentration Screening
  - a. Level1
  - b. Level2
  - c. Level3
  - d. Level4
  - e. Level5
  - f. Level6
7. Field Explanation/Database Structure

- a. Single-concentration tables
  - b. Multiple-concentration tables
  - c. Auxiliary annotation tables
8. B Level 0 Pre-processing
  9. C Burst Z-Score Calculation

## Links

- [How to Set Up R and MySQL](#)

# Exploring ToxCast Data: How to Set Up R and MySQL

## Resources

- [Toxicity Forecaster \(ToxCast\) Fact Sheet](#)
- [ToxCast Publications](#)
- [ToxCast Citation](#)
- [About ToxCast](#)

The following links exit the site [EXIT](#)

It is helpful to have a working knowledge of R and MySQL to use the ToxCast [tcpl R package](#). For further details as to how to use R and MySQL, please reference these instructions and external resources.

## Setting up R

1. Download the latest version of the R package from [the CRAN](#) according to your operating system (Linux, Mac, Windows).
  - a. Review the FAQs and instructions for installing R for the first time ([Windows](#)) ([Mac](#)) ([Linux](#)).

## Setting up MySQL

1. Download the latest version of the [MySQL community server](#)
  - a. Select the appropriate installer for your operating system
    - i. Windows: MSI installer
    - ii. MAC: DMG installer
  - b. The installer will walk you through the installation
  - c. **During the installation, be sure to copy the temporary root password. You will need it later.**
  - d. For Windows, MySQL should automatically be added to your PATH.
  - e. For MAC, if MySQL was not added to your path automatically then add it manually
    - i. Open terminal and type:

```
echo 'export PATH=/usr/local/mysql/bin:$PATH' >>  
~/.bash_profile  
</>
```

2. Open the command line (Windows) or terminal (MAC) to login to the MySQL server with the command:

```
mysql -u root -p  
</>
```

3. Enter the temporary root password when prompted for a password.
4. Change the root password
  - a. Full documentation can be found at [MySQL Website](#)
5. Create the database, select it as the default database, and load the dump file:

```
mysql> CREATE DATABASE IF NOT EXISTS invitrodb_v2;  
mysql> USE invitrodb_v2;  
mysql> source prod_external_invitrodb_v2.all.sql  
</>
```

- a. Further documentation can be found at [MySQL Website](#)
- b. Be patient; this can take anywhere from 3 to 8 hours.

## Links

- [Overview of R Package](#)

# ToxCast Data Generation: Publicly Available Data

## Resources

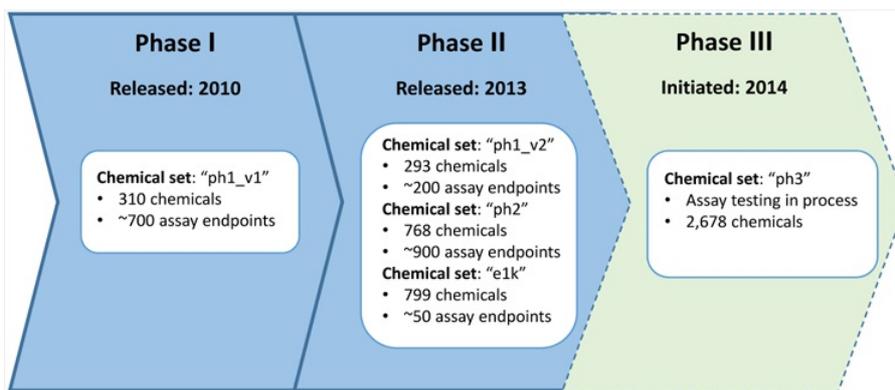
- [Toxicity Forecaster \(ToxCast\) Fact Sheet](#)
- [ToxCast Publications](#)
- [ToxCast Citation](#)
- [About ToxCast](#)

## Data Releases

The ToxCast effort has been developed in three phases: **Phase I**, **Phase II** and **Phase III** (in progress). Starting with the release of Phase I data in 2010, each phase has added new data for unique chemicals and assay endpoints to the library.

There is currently data available for **more than 1800 unique chemicals and 700 assay endpoints** (Note: not all chemicals are subjected to testing in all assays). The graphic below presents a breakdown by phase of the chemical set names, release dates, number of unique chemicals assessed and number of assay endpoints.

The current data files are available for download [here](#).



## Previous ToxCast Data Releases

Data files from previously published ToxCast data releases are available [here](#). These files are NOT recommended for use in new analyses, but are provided for users who may need access to previous data files for ongoing analyses.

## Additional Information on Data Releases

[Richard, A. M., Judson, R. S., Houck, K. A., Grulke, C. M., Volarath, P., Thillainadarajah, I., ... & Knudsen, T. B. \(2016\). ToxCast chemical landscape: Paving the road to 21st century toxicology. \*Chemical Research in Toxicology\*, 29\(8\), 1225-1251.](#)

## **Links**

- [ToxCast Data](#)
- [Downloadable Data](#)

# Exploring ToxCast Data

## Resources

- [Toxicity Forecaster \(ToxCast\) Fact Sheet](#)
- [Downloadable ToxCast Data](#)
- [ToxCast Dashboard](#)
- [ToxCast Publications](#)
- [ToxCast Citation](#)
- [About ToxCast](#)
- [Example Use Cases](#)

ToxCast data, once it has been generated by labs and processed by EPA, can be downloaded from our website and is also available in the ToxCast Dashboard.

## Download ToxCast Data

For more advanced users, the data files themselves are available for download and use, as well as the data processing pipeline used by EPA (tcpl R package).

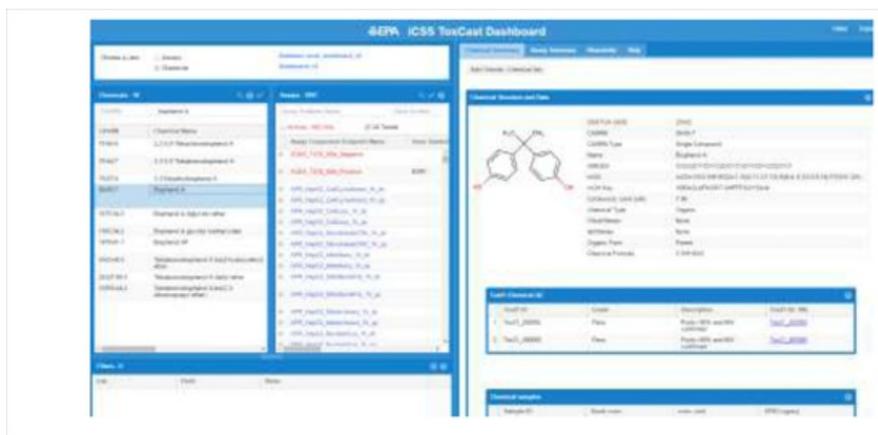
- [Downloadable Data](#)

## ToxCast Dashboard

The ToxCast Dashboard was developed by EPA to help users examine high-throughput assay data to inform chemical safety decisions. All of the publicly released ToxCast data are accessible via the dashboard in an intuitive and easy-to-use format. EPA expects that most users will be able to find and explore data of interest using this format. The Dashboard includes Help annotation and instructions for use.

- [ToxCast Dashboard](#)

*Note: For the best user experience, we recommend using Google Chrome or Firefox*



# Accessing ToxCast Data and Options for Exploring Data

Click [here](#) for examples of how ToxCast data can be accessed and explored by users (through the Dashboard or the tcpl) or how ToxCast has been used to inform decision making.

# Exploring ToxCast Data: Accessing ToxCast Data and Scenarios for Exploring Data

ToxCast data can be accessed, explored and used in different ways and for different purposes. This page highlights examples of ways to access and explore the data.

## Resources

- [Toxicity Forecaster \(ToxCast\) Fact Sheet](#)
- [ToxCast Publications](#)
- [ToxCast Citation](#)
- [About ToxCast](#)

## On This Page

- |  |   |
|--|---|
| <ul style="list-style-type: none"><li>• <a href="#">ToxCast Dashboard</a><ul style="list-style-type: none"><li>◦ <a href="#">Explore the dashboard</a></li><li>◦ <a href="#">Explore biological activity</a></li><li>◦ <a href="#">Explore chemical activity</a></li></ul></li></ul> | <ul style="list-style-type: none"><li>• <a href="#">R-Package TCPL Scenarios</a><ul style="list-style-type: none"><li>◦ <a href="#">Export data</a></li><li>◦ <a href="#">Register new data</a></li><li>◦ <a href="#">Process my own data</a></li></ul></li><li>• <a href="#">Endocrine Disruptor Screening Program</a></li><li>• <a href="#">Links</a></li></ul> |
|--|---|

## ToxCast Dashboard Exploration Examples

The [ToxCast Dashboard](#) provides three useful tutorials to demonstrate possible use cases.

### **I would like to explore the dashboard in general.**

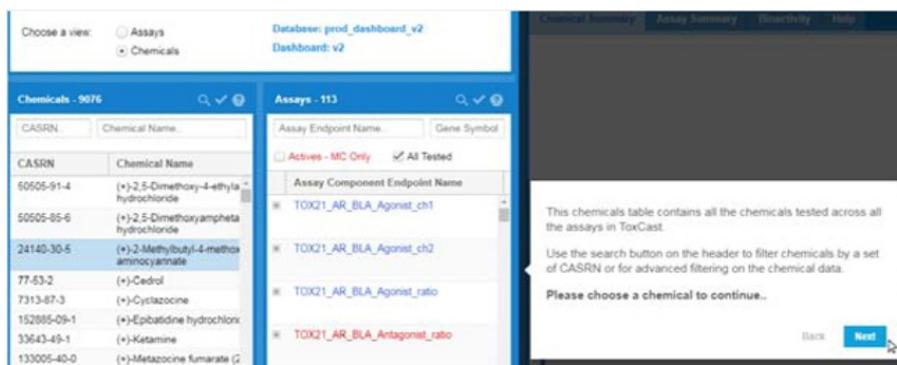
*This tutorial is located under the Chemical Summary tab of the Dashboard and provides introductory information showing users how to select chemicals in the Dashboard. Users can follow a similar process to select assays in the Dashboard.*

1. There are four tabs in the top right half of the ToxCast Dashboard home page, Chemical Summary, Assay Summary, Bioactivity, and Help.
2. Under the Chemical Summary tab, scroll down and click on the “Start

Tutorial – Chemical Tab” button.



3. A pop up window will appear with instructions. Follow each instruction before clicking the “Next” button.



4. There is no tutorial under the Assays tab, but users can follow the process outlined in Chemical tutorial to explore on their own. Be sure to change the “Choose a view” option in the top left and click over to the Assay Summary tab to see what type of information is available. For detailed information on these assays, please see the [Assays page](#).

[Top of Page](#)

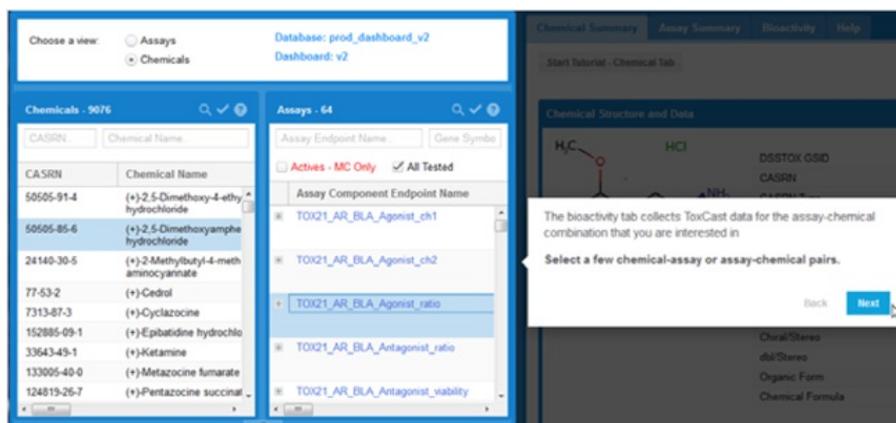
## I would like to explore biological activity.

*This tutorial can be found under the Bioactivity tab of the Dashboard. This tutorial assumes that users are familiar with selecting assays and chemicals in the Dashboard.*

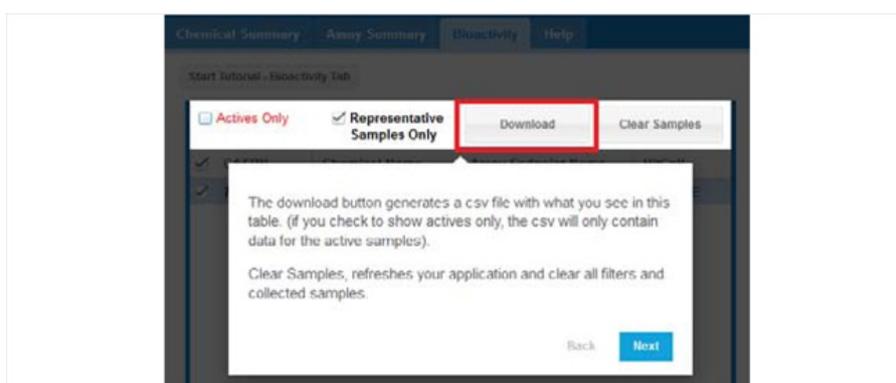
1. Under the Bioactivity tab, click on the “Start Tutorial – Bioactivity Tab” button.



2. A pop up window will appear with instructions. Follow each instruction before clicking the “Next” button.



3. Users can start by selecting either an assay or a chemical, depending on their interest.
4. Users can also download data into a .csv file.



[Top of Page](#)

## I would like to explore chemical activity.

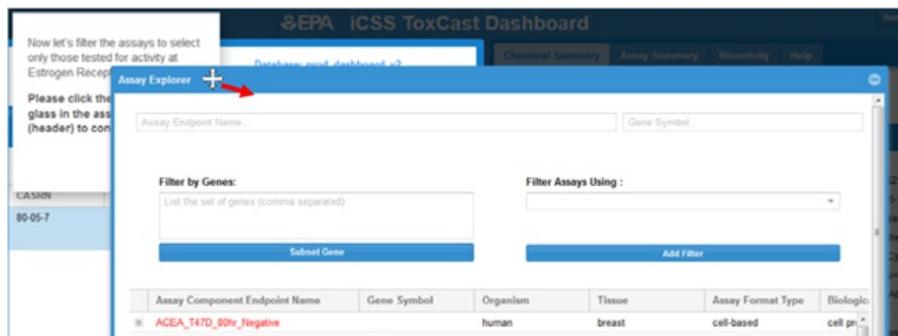
### Example – Activity in flame retardants.

*This tutorial is located under the Help tab of the Dashboard and explores the data relevant to Bisphenol A and all Estrogen Receptor (ER) assays.*

1. Under the Help tab, expand the “How to Use the iCSS Dashboard” window.
2. Click on the “Use Case Tutorial” button. Maximize your browser window.



3. A pop up window will appear with instructions. Follow each instruction before clicking the “Next” button. Please note that when you click on the magnifying glass in the assay window, a new Assay Explorer window will pop up and may block your instructions; if this happens, simply move the new window slightly to the right and down and press the “Next” button on the instruction pop up.



4. After adding the Gene Symbol filters, click “Next” and close the Assay Explorer window to reveal the next instruction. To close the Assay Explorer window, click on the minus sign in the top right corner of the window.
5. Continue to follow the instructions until the end of the tutorial.

[Top of Page](#)

## R–Package TCPL Scenarios

### **I would like to export data from ToxCast and review it.**

*This example assumes that users are familiar with R and the tcpl R Package.*

This example is geared towards users who would like to review the ToxCast data at a processing level prior to that available in the ToxCast Dashboard (Level 6).

This example is in development.

### **I would like to register new high–throughput screening data for in vitro db.**

*This example assumes that users are familiar with R and the tcpl R Package.*

This example, available in the [tcpl Vignette](#) `EXIT`, explains how to register and upload new data into the tcpl database using a small subset of ToxCast data showing changes in intracellular cortisol hormone. The subset of data comes from an assay measuring steroidogenesis through cellular levels of multiple steroid hormones.

The tcpl package provides three functions for adding new data:

1. tcplRegister to register a new assay or chemical ID
2. tcplUpdate to change or add additional information for existing assay or chemical IDs, and
3. tcplWriteLvl0 for loading data.

Before writing any data to the tcpl database, the user has to register the assay and chemical information.

See the [tcpl Vignette](#) `EXIT` for further details.

## I would like to process my own data using the tcpl pipeline.

This example assumes that users are familiar with R and the tcpl R Package and have registered their data (see previous example for details).

This example, available in the [tcpl Vignette](#) `EXIT`, explains how the data are processed through the [levels of the pipeline](#) for single-concentration and multiple-concentration screening.

All processing in the tcpl package occurs at the assay component or assay endpoint level. There is no capability within either screening paradigm to do any processing which combines data from multiple assay components or assay endpoints. Any combining of data must occur before or after the pipeline processing.

The processing is sequential, and every level of processing requires successful processing at the antecedent level. The processing requirements vary by screening paradigm and level. The Vignette includes the details, but in general, many of the processing steps require specific methods to accommodate different experimental designs or data processing approaches.

See the [tcpl Vignette](#) `EXIT` for further details.

[Top of Page](#)

## ToxCast Data Relevant to EPA's Endocrine Disruptor Screening Program

Based on scientific advances, EPA intends to implement the use of high throughput screening assays and computational models to evaluate, and to a significant extent, screen chemicals.

- [Endocrine Disruptor Screening Program](#). The in vitro high throughput and computational model alternatives provide an accurate quantitative measure of specific endocrine receptor binding bioactivity and mechanisms that can serve as alternatives to the current Tier 1 estrogen receptor (ER) binding, ER transactivation (ERTA) and uterotrophic assays.
- [Prioritizing Chemicals for Potential Endocrine Disruption](#). EPA researchers have been developing new rapid methods that can quickly screen thousands of chemicals.
- [High-throughput Screening Data for Estrogen Receptor Model](#). Estrogen receptor model data from the manuscript titled *Integrated Model of Chemical Perturbations of Biological Pathways Using 18 In Vitro High-Throughput Screening Assays for the Estrogen Receptor* (Judson et al.) published in *Toxicological Sciences*.

## Links

- [ToxCast Dashboard](#)

- [Downloadable Data](#)

[Top of Page](#)

---

# Exploring ToxCast Data: Downloadable Data

## Resources

- [Toxicity Forecaster \(ToxCast\) Fact Sheet](#)
- [ToxCast Publications](#)
- [ToxCast Citation](#)
- [About ToxCast](#)

The results after processing through the Pipeline are available on the [ToxCast Dashboard](#), and for most users EPA recommends accessing the data there.

- [ToxCast Chemicals](#)
- [ToxCast Assays](#)

## ToxCast Data and Information

- [ToxCast & Tox21 Summary Files](#). Data for a single chemical endpoint pair for thousands of chemicals and assay endpoints for 20 variables such as the activity or hit call, activity concentrations, whether the chemical was tested in a specific assay, etc.
- [ToxCast & Tox21 Data Spreadsheet](#). A spreadsheet of EPA's analysis of the chemicals screened through ToxCast and the Tox21 collaboration which includes EPA's activity calls from the screening of over 1,800 chemicals.
- [ToxCast & Tox21 Concentration Response Plots](#). Concentration response plots for all of the ToxCast and Tox21 assays.
- [Collaborative Estrogen Receptor Activity Prediction Project Data](#). Data and supplemental files from CERAPP (A large-scale modeling project) which demonstrated the efficacy of using predictive computational models trained on high-throughput screening data to evaluate thousands of chemicals for estrogen-related activity. CERAPP combined multiple models developed in collaboration with 17 groups in the United States and Europe to predict ER activity of a common set of 32,464 chemical structures. *Quantitative structure-activity relationship* models and docking approaches were employed, to build a total of 40 categorical and 8 continuous models for binding, agonist, and antagonist ER activity.
- [High-throughput Screening Data for Estrogen Receptor Model](#). Estrogen receptor model data from the manuscript titled [Integrated Model of Chemical Perturbations of Biological Pathways Using 18 In Vitro High-throughput Screening Assays for the Estrogen Receptor](#) (Judson et al) published in *Toxicological Science*.
- [High-throughput Screening Data for Androgen Receptor Model](#). Androgen receptor activity model data.
- [Previous ToxCast Data Releases](#). Data files from previously published ToxCast data releases. We DO NOT recommend using this data for new analyses, but

are providing these files in case users need them for ongoing analyses.

- [Animal Toxicity Studies: Effects and Endpoints](#). The Toxicity Reference Database (ToxRefDB) provides results from across the thousands of animal toxicity studies including No Effect Level (NEL)/Lowest Effect Level (LEL) and/or No Observed Adverse Effect Level (NOAEL)/Lowest Observed Adverse Effect Level (LOAEL).

## ToxCast Results and Processing

Descriptions of how data are processed through can be found on the [Results Processing and Analysis](#).

### Links

- [Overview of R Package](#)
- [How to Set Up R and MySQL](#)

---

# Exploring ToxCast Data: Citing ToxCast Data

## Resources

- [Toxicity Forecaster \(ToxCast\) Fact Sheet](#)
- [ToxCast Publications](#)
- [About ToxCast](#)

The following citation format can be used when citing ToxCast data:

### General Citation Suggestion

U.S. EPA. Data download year. Data set name from database version. Retrieved from <https://www.epa.gov/chemical-research/toxicity-forecaster-toxcasttm-data> on date retrieved. Data release date.

### Example Citation

U.S. EPA. 2015. ToxCast & Tox21 Summary Files from invitrodb\_v2. Retrieved from <https://www.epa.gov/chemical-research/toxicity-forecaster-toxcasttm-data> on October 28, 2015. Data released October 2015.