ToxCast Assay Network (TCAN) Viewer: A Visualization Tool for High-throughput Assay Chemical Data

Parth Kothiya[1], Jimmy Phuong[2], Matt Martin[2]

(1) ORISE Fellow at the U.S. EPA, Research Triangle Park, NC, USA, (2) National Center for Computational Toxicology, U.S. EPA, Research Triangle Park, NC, USA

Parth Kothiya | Kothiya.parth@epa.gov | 919-541-4240

Introduction

- US EPA's ToxCast program has generated high-throughput screening (HTS) data on the bioactivity for thousands of chemicals. ToxCast provides descriptions and annotations for the HTS assays with respect to assay design and target information (e.g., gene target). Recent stakeholder and use feedback highlighted a need for improved and expanded assay information, as well as improved data visualization tools.
- To foster easier data exploration, we developed the ToxCast Assay Network (TCAN) Viewer, a web application that is based on R software and the Shinyapps library in R. Here, we have added additional assay annotations (i.e. gene targets and assay vendors) and associated quantitative parameters (e.g., activity concentration and efficacy).
- TCAN is currently equipped with two primary views: animated spider plots and assay-gene network.

Objectives

- Enable users to select chemicals and annotation features of interest, and to use these selections to aggregate and visualize trends in big data sets.
- Display chemical-centrical (chemical explorer) or assay-centrical (assay explorer) views in a user-friendly way that allows hypotheses generation, development of prioritization schemes, or creation of analysis.

Methods

TCAN Selection tool

- The Shiny package in R-studio was used to create the selection tool and link to the different view options
- Input variables can be an assay endpoint or a chemical, or a user-customized set by an assay type or a chemical set of interest.

Annotation network

- The assay-gene network view displays connections across assays based on common gene targets and allows multilayered circles for each node to quickly indicate the response of a particular chemical across the tested concentration range.
- Each network can be centered on an assay annotation (or a group of them) connected to 'assay endpoints'.
- Each layer of the circle is a log concentration on based 10.
- Each layer is shaded according to rescaled response-value calculated as: the reciprocal of bmad times the response value for each given concentration, where the higher responses receive darker shades.
- All the data has pulled from the ToxCast data processing pipeline.

Spider plot

- Animated spider plots display the modeled rescaled response across a set of assays with the variable of concentration being controlled by the user.
- The user can dynamically illustrate concentration responses for a set of related assays and chemicals demonstrating concordance across assays or a specific response profile as concentration changes.

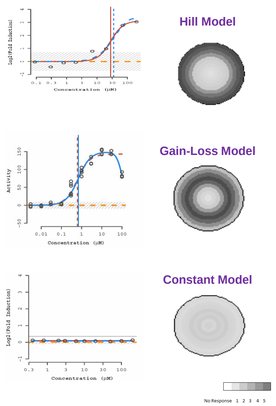


Fig 1.
Representation of concentration response data with a multilayer circle.

Annotation network view

Fig 2.
Example of intended target annotation network for ATG CIS and TRANS 'gene-targeting' assay endpoints.

Gene Symbols are linked to each assay endpoint, and grouped by gene families (orange).

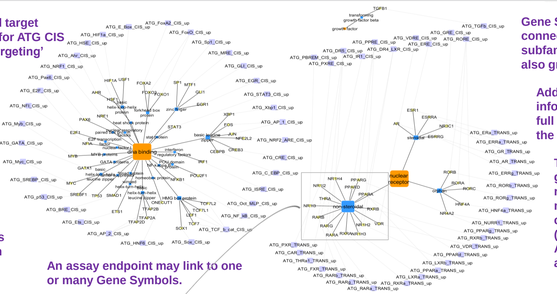
Assay endpoints may be grouped using different levels of target information (e.g. non-steroidal nuclear receptors).

An assay endpoint may link to one or many Gene Symbols.

Gene Symbols are connected to intended target subfamilies (blue), which also group to target.

Additional gene information such as gene full names are available in the assay annotations.

The node sizes for gene targets are representative of how many assay endpoints can be mapped to them (here, it is restricted to ATG CIS and TRANS assay endpoints)



Chemical explorer using the Assay network view

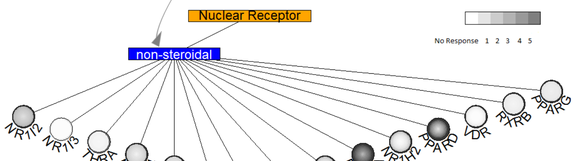


Fig 3.
TCAN output for 'Tetrabutyltin' to view the bioactivity across assay endpoints as selected and grouped by an assay annotation (i.e. non-steroidal nuclear receptors).

Multilayer circle view

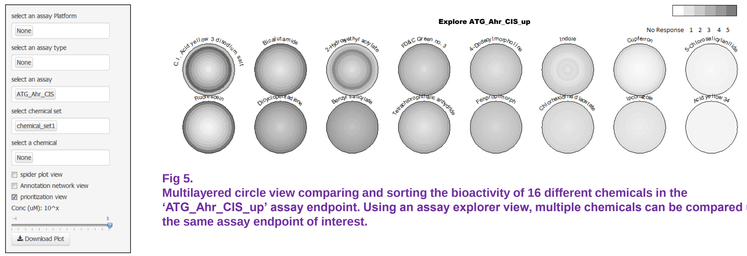


Fig 5.
Multilayered circle view comparing and sorting the bioactivity of 16 different chemicals in the 'ATG_Ahr_CIS_up' assay endpoint. Using an assay explorer view, multiple chemicals can be compared under the same assay endpoint of interest.

Assay annotation options

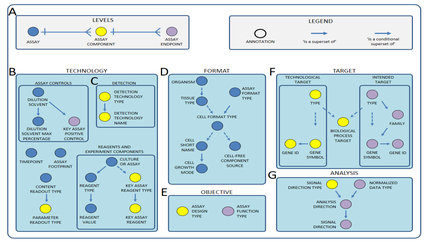


Fig 4. ToxCast Assay Annotations that are captured for all ATG CIS and TRANS assay endpoints and at-option to identify assay endpoints.

Spiderplot view

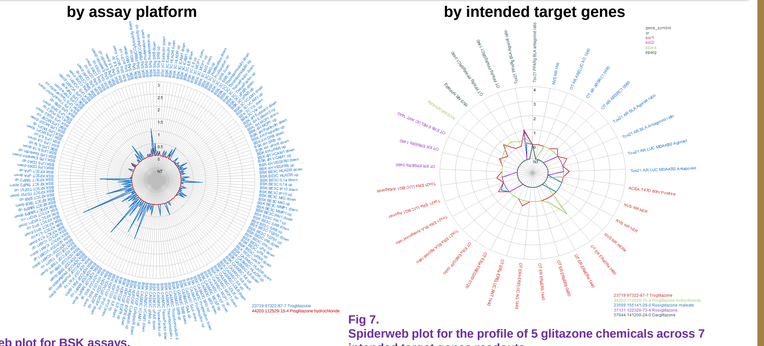


Fig 6.
Spiderweb plot for BSK assays.

The axis shown uses fold change response data (e.g. 1.5 fold difference from the negative control) for 'Troglitazone' and 'Proglitazone hydrochloride' at the 100 uM test concentration.

Fig 7.
Spiderweb plot for the profile of 5 glitazone chemicals across 7 intended target genes readouts.

The axis shown uses percentage activity given by the chemical's 100 uM response data. Assay endpoints that use fold-change response units are converted into percent activity of the negative control. The most bioactive chemicals will have a line with a longer peak.

Conclusions

- These graphical enhancements improves access, analyses and visual display of the large and complex ToxCast data.
- TCAN enables rapid decision making to generate hypotheses or inform chemical prioritization schemes.
- The R Shiny package enables the TCAN viewer to be interfaced via a desktop application or web browser.

Future Steps

- Integrate TCAN viewer to ToxCast Dashboard and test out the different views. Some implementation has been tested using tcan.shinyapps.io
- Incorporate more search functions for the annotation network view.