National Institute of Environmental Health Sciences
Division of Translational Toxicology

Incorporating Ontologies into High-Throughput Screening Assay Annotations to Facilitate Interpretation and Increase Data Use

Bridgett Hill[1], J Todd Auman[1*], Aswani Unnikrishnan[1], Agnes L Karmaus[1*], Helena T Hogberg[2], Emily Reinke[1], Nicole Kleinstreuer[2]

[1]Inotiv, Research Triangle Park, NC; [2]NIH/NIEHS/DTT/NICEATM, Research Triangle Park, NC

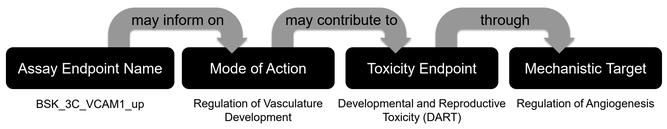
Background

- Abundant in vitro high-throughput screening (HTS) assay data are available to characterize effects of thousands of chemicals and have facilitated the development of computational approaches for testing and assessment.
- While HTS assays are valuable sources of mechanistic information, it is often not clear how these data relate to toxicological endpoints.
- Here we present annotation of HTS assay endpoints incorporating controlled vocabularies within an ontology framework to provide biological context and facilitate toxicological interpretation.
  - The annotation knowledge organization system applied also addresses two common challenges associated with large toxicology datasets: inconsistent terminologies and inconsistent reporting structure.

Annotations Knowledge Organization System

- HTS data from the U.S. Environmental Protection Agency's ToxCast database [1] were reviewed by NTP Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM) subject matter experts, generating a curated copy of these data referred to as the curated HTS (cHTS) data set.
- Annotations for "Mode of Action" and "Mechanistic Target" for all assays were also developed to support toxicological interpretation of assay endpoints. These are included in our knowledge organization structure (shown below) to define linkages between annotation terms and toxicity endpoints.

Knowledge Organization System (KOS) Structure



- cHTS assay data are available on the NICEATM Integrated Chemical Environment (ICE; https://ice.ntp.niehs.nih.gov/).

Mode of Action: pathways relating to toxicological endpoints	**Mechanistic Target**: biological processes
DART - Cell Process	Cellular Process
DART - Cytochrome P450 Activity	Cellular Stress Response
DART – Gene Expression Regulation	Cytochrome P450 Activity
DART – Nervous System Development	Energy Homeostasis

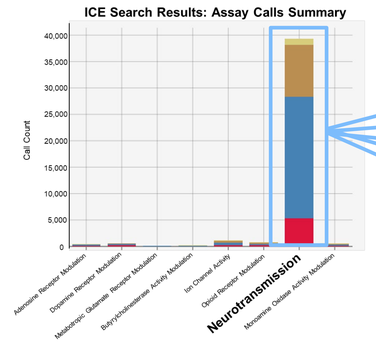
Integrated Chemical Environment

https://ice.ntp.niehs.nih.gov/

Ontologies Enable Refinement of Mode of Action and Mechanistic Target Terms

A

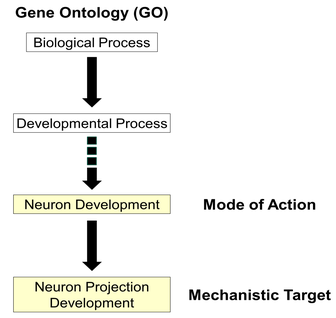
Historical NCI Metathesaurus "Neurotransmission" Mechanistic Target



Examples from the 20 OBO Foundry Ontologies Identified for this Mechanistic Target

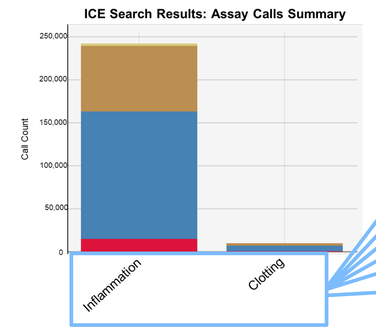
- **Glial Cell Migration GO:0008347**
- **Neural Precursor Cell Proliferation GO:0061351**
- **Neuron Projection Development GO:0031175**
- **Neuronal Signal Transduction GO:0023041**
- **Synapse Assembly GO:0007416**

OBO Foundry Ontologies Support Hierarchy for Mode of Action and Mechanistic Target Terms



B

Historical NCI Metathesaurus "Immune and Inflammatory Response" Mechanistic Targets



6 Additional Mechanistic Target Terms Identified using the OBO Foundry Ontologies

- **Cytokine Production Involved in Inflammatory Response GO:0002534**
- **Cytokine-mediated Signaling Pathway GO:0019221**
- **Regulation of Cytokine Production GO:0001817**
- **Regulation of Immune Response GO:0050776**
- **Regulation of Lymphocyte Activation GO:0051249**
- **Regulation of Tissue Remodeling GO:0034103**

OBO Foundry Ontologies Support Hierarchy for Mode of Action and Mechanistic Target Terms

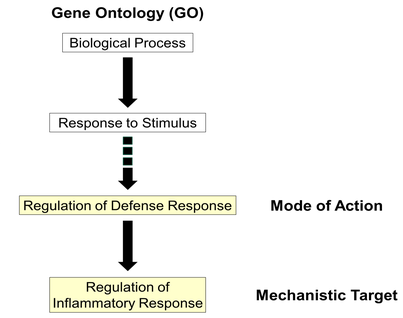


Figure 2: Expansion of annotation terms using OBO Foundry GO terms allow for broader categories with parent-child relationships defining hierarchical organization. Examples shown here are Neuronal Transmission (A) and Immune and Inflammatory Response (B) related annotations. Colors within the bars indicate the bioactivity or curation call: red (active), blue (inactive), gold (chemical QC-omit), yellow (curve/assay flag-omit).

Impact of Applying Controlled Vocabularies

- Controlled vocabularies from widely used, established terminologies were applied to facilitate data interoperability.
- Historical Mechanistic Target annotations in ICE were based on the National Cancer Institute's NCI Metathesaurus [2] terms, which emphasize biomedical annotation such that terms are focused on clinical- and cancer-centered terminology.
- Updated Mode of Action and Mechanistic Target annotations address the limitations posed by the scope of the NCI Methathesaurus terms by applying Open Biological and Biomedical Ontology (OBO) Foundry terminology [3], consisting of multiple knowledge areas to encompass a broader range of biological and toxicological processes (Figure 1). More specifically, annotation terms were organized according to ontologies primarily obtained from the Gene Ontology (GO) knowledgebase [4, 5].

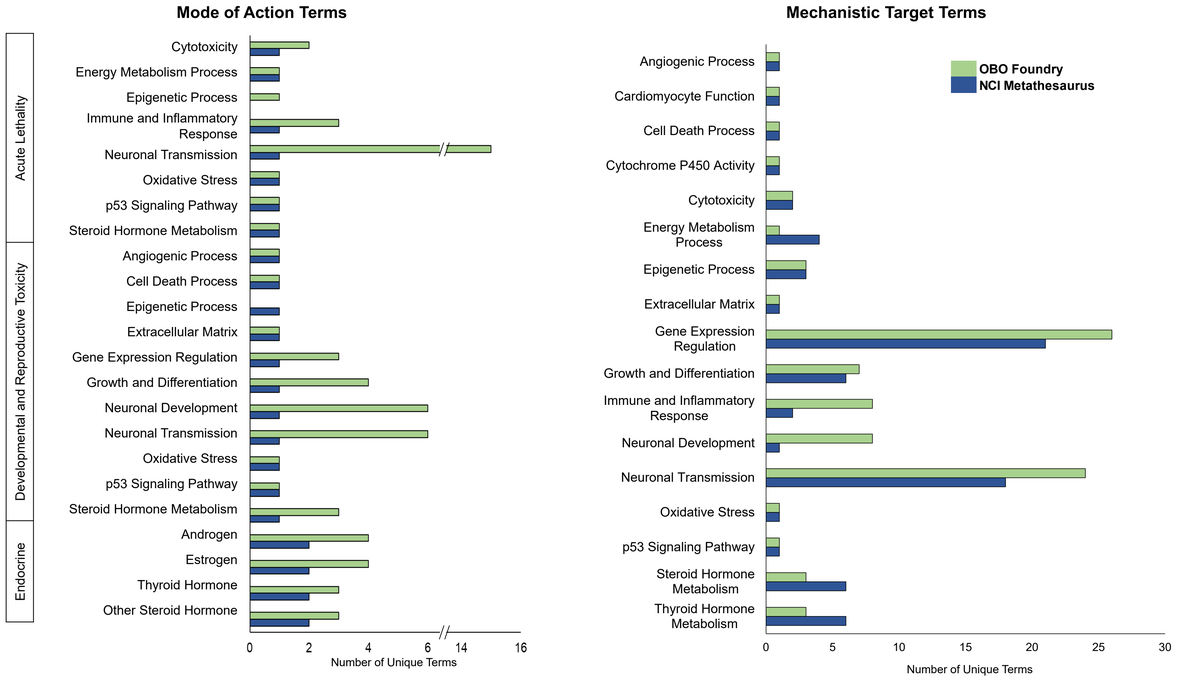


Figure 1: Comparison of cHTS assays annotated using both the NCI Metathesaurus (blue) and OBO Foundry (green) controlled vocabularies. Bars represent the number of terms applied to the cHTS dataset per category. Our expansion to ontology terms has increased the number of different annotation terms, providing more granularity to aid interpretation.

Summary and Future Directions

Summary

- Curation and mapping of the cHTS data set in ICE using ontologies facilitates data interpretation and supports our ongoing efforts to ensure that data are findable, accessible, interoperable, and reusable (FAIR).
  - Ontologies provide users with high-quality, structured, mechanistically relevant information in the broader data ecosystem available through the ICE interface.
  - Ontologies help with terminology standardization enabling interoperability with other resources such as the National Institute of Environmental Health Sciences' Chemical Effects in Biological Systems (CEBS) database.
- Increased accessibility, contextualization, and support with interpretation of cHTS data results in the following:
  - Aid in identifying data gaps.
  - Provide better characterization of chemical effects and hazard assessments.
  - Provide additional resources for investigations of toxicologically-relevant endpoints.

Future Directions

- NICEATM is aligning the ICE cHTS annotations with harmonized reporting efforts such as the Organisation for Economic Co-operation and Development's Harmonized Template 201 (OHT201).
- OHT201 currently includes picklists of ontology terms to represent Process and Object used to describe mechanistic observations (Table 1). We are aggregating current annotations to format inputs for these fields.

Table 1: Select Process and Object Picklist ontology terms from the OHT201 document

Process	List sup. (picklist with remarks) Display: Basic	**Picklist values:** - apoptotic process - [GO:0008219] - biosynthetic process - [GO:0009058] - catalytic activity - [GO:0003824] - cell activation - [GO:0001775] - cell death - [GO:0008219] - cell differentiation - [GO:0030154] - cell migration - [GO:0016477] - cell proliferation - [GO:0008283]
Object	List sup. (picklist with remarks) Display: Basic	**Picklist values:** - aldo-keto reductase family 1 member C2 (AKR1C2) - [PR:000003904] - androgen receptor - [PR:000004191] - CD54 molecule (intercellular adhesion molecule 1) - [PR:000001467] - CD86 molecule - [PR:000001412] - cytochrome P450 - [CHEBI:38559]

References

1. U.S. EPA. 2022. ToxCast & Tox21 Summary Files from invitrodb_v3.5 Retrieved from https://www.epa.gov/chemical-research/toxicity-forecaster-toxcasttm-data. Data released September 2022.
2. NCI Metathesaurus (NCIm). 2023. NCIm Version: 202308 (Browser Version 2.17). https://ncim.nci.nih.gov/ncimbrowser/.
3. Jackson et al. 2021. Open Biological and Biomedical Ontology (OBO) Foundry Database. https://doi.org/10.1093/database/baab069.
4. Ashburner et al. 2000. Gene ontology: tool for the unification of biology. Nat Genet. DOI: 10.1038/75556.
5. The Gene Ontology Consortium. 2023. The Gene Ontology knowledgebase in 2023. Genetics. DOI:10.1093/genetics/iyad031.

Acknowledgments

This project was funded with federal funds from NIEHS, NIH under Contract No. HHSN273201500010C. The views expressed above do not necessarily represent the official positions of any federal agency.

We would like to thank Catherine Sprankle and Elizabeth Farley-Dawson for their assistance with creating and reviewing this poster.

* A.L. Karmaus is currently employed at Syngenta Crop Protection, LLC, Greensboro, NC.

* J.T. Auman is currently employed at RTI International, Research Triangle Park, NC.

Contact the author: bridgett.hill@inotiv.com

Subscribe to the NICEATM News email list:

https://list.nih.gov/cgi-bin/wa.exe?SUBED1=niceatm-l&A=1

Check out other ICE-related presentations at ASCCT:

PSII-02: Unnikrishnan et al. Harmonizing Tox Data

PSII-18: Reisfeld et al. The Integrated Chemical Environment (ICE): Advancing Data Availability and Computational Tool Accessibility