National Institute of Environmental Health Sciences
Division of Translational Toxicology

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

B.N. Hill[1], E.N. Reinke[1], J.T. Auman[1], A. Unnikrishnan[1], D.G. Allen[1]*, A.L. Karmaus[1]*, H. Hogberg[2], N. Kleinstreuer[2],
[1]Inotiv, RTP, NC; [2]NIH/NIEHS/DTT/NICEATM, RTP, NC;

Abstract 4385
Poster P172

Background

- In vitro high-throughput screening (HTS) assay data have increased our understanding of chemically mediated effects for thousands of chemicals and 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 an annotation scheme for HTS assays that incorporates controlled vocabularies within the framework of ontologies. Our annotations provide biological context and facilitate toxicological interpretation. The annotation scheme also addresses two common challenges associated with large toxicology datasets: inconsistent terminologies and inconsistent reporting structure.

Annotation Methods

HTS data from the U.S. Environmental Protection Agency's ToxCast database (invitrodb v3.5) were reviewed by subject matter experts who assigned annotations for modes of action (based on pathways relating to toxicological endpoints) and mechanistic targets (based on biological processes).

Curated HTS (cHTS) assay data are available on NTP Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM) Integrated Chemical Environment (ICE; https://ice.ntp.niehs.nih.gov/).

Mode of Action

- Mode of Action
  - DART - Growth and Differentiation
  - DART - Angiogenic Process
  - DART - Cell Death Process
  - DART - Epigenetic Process
  - DART - Extracellular Matrix
  - DART - Gene Expression Regulation

Mechanistic Target

- cHTS
  - Malformation
  - Vascularization
  - Cellular Process
  - Cellular Stress Response
  - Endocrine-Related Processes
  - Energy Metabolism Process

Controlled Vocabularies

Controlled vocabularies from widely used and established terminologies were chosen to facilitate data interoperability. Existing Mechanistic Target annotations in ICE are based on the National Cancer Institute's NCI Metathesaurus (https://ncim.nci.nih.gov) terms, which encompass a wide range of biomedical space. However, these terms are focused on clinical- and cancer-centered terminology.

To address the limitations inherent in the NCI Methathesaurus terms, we are updating Mode of Action and Mechanistic Target annotations using the Open Biological and Biomedical Ontology (OBO) Foundry (http://obofoundry.org/), a harmonized and interoperable database consisting of multiple knowledge areas to encompass a broader range of biologic and toxicologic processes (Figure 1).

We organized our annotations according to ontologies primarily obtained from the Gene Ontology (GO) knowledgebase.

Ontologies Enable Refinement of Mode of Action and Mechanistic Target Terms

Figure 1: Comparison between cHTS NCI Metathesaurus and OBO Foundry Ontologies

The bar graphs below represent cHTS assays categorized by broader regulatory relevant toxicological endpoints. Numbers in parentheses indicate the number of cHTS assays within each category. The same assays were annotated using both the NCI Metathesaurus and OBO Foundry controlled vocabularies. Bars represent the number of terms per category. Our transition to ontology terms has resulted in an increase in the number of different annotation terms, providing more granularity to facilitate interpretation.

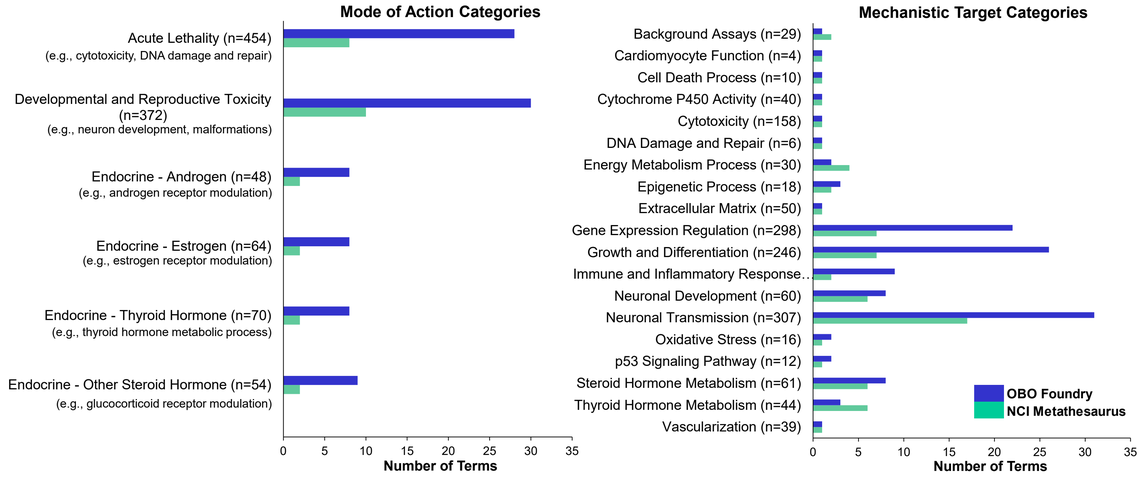
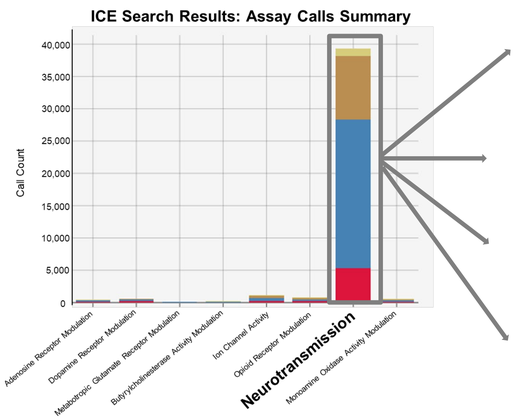


Figure 2: Example Using Neuronal Transmission Related Annotations: Expansion of Broader Categories with Parent-Child Linkages

NCI Metathesaurus "Other Neurotransmission" Mechanistic Target



OBO Foundry: Four Ontologies Identified for NCI Metathesaurus Mechanistic Target Annotation:

Neuron Projection Development GO:0031175

Neural Precursor Cell Proliferation GO:0061351

Neuron Apoptotic Process GO:0051402

Synapse Assembly GO:0007416

Gene Ontology (GO)

Biological Process

↓

Cellular Process

↓

Cell Population Proliferation — **Mode of Action**

↓

Neural Precursor Cell Proliferation — **Mechanistic Target**

Summary and Discussion

Curation and mapping of the HTS data using ontologies supports our ongoing efforts to ensure that data objects are findable, accessible, interoperable, and reusable (FAIR).

- These ontologies are intended to provide users with high-quality, structured, mechanistically-relevant information in the broader data ecosystem available through the ICE interface.
- Ontologies will 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 and interpretation of HTS data will aid in identifying data gaps, better inform characterization of chemical effects and hazard assessments, and provide additional resources for investigations of toxicologically-relevant endpoints.

Integrated Chemical Environment

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

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.

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 and Acknowledgments

Ashburner et al. 2000. Gene ontology: tool for the unification of biology. Nat Genet. DOI: 10.1038/75556.

The Gene Ontology Consortium. 2023. The Gene Ontology knowledgebase in 2023. Genetics. DOI:10.1093/genetics/iyad031.

Jackson et al. 2021. Open Biological and Biomedical Ontology (OBO) Foundry Database, https://doi.org/10.1093/database/baab069.

NCI Metathesaurus (NCIm). 2023. NCIm Version: 202308 (Browser Version 2.17). https://ncim.nci.nih.gov/ncimbrowser/.

Learn more about ICE and its assay annotations at SOT 2024! Demonstrations at the NIEHS booth #617 (10:00 AM - 12:00 PM).

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. *D.G. Allen's current affiliation is with the International Collaboration on Cosmetics Safety, New York, NY. A.L. Karmaus' current affiliation is with Syngenta, Greensboro, NC.

Contact the author: bridgett.hill@inotivco.com

To subscribe to the NICEATM News email list visit: https://list.nih.gov/cgi-bin/wa.exe?SUBED1=niceatm-l&A=1