National Institute of Environmental Health Sciences
Division of Translational Toxicology

Integration of Technological Interference into Curated HTS Data

V. Hull[1], A. Borrel[1], A.L. Karmaus[1]*, D.G. Allen[1], N. Kleinstreuer[2]

[1]Inotiv, RTP, NC, USA; [2]NIH/NIEHS/DTT/NICEATM, RTP, NC, USA

Introduction

- Highly curated, robust data are essential for building confidence in new approach methodologies. However, large quantities of data generated by high-throughput screening (HTS) assays such as the U.S. Environmental Protection Agency's (EPA's) Tox21/ToxCast programs can be difficult to interpret.
- To increase confidence in HTS bioactivity calls, the National Toxicology Program Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM) has derived curated HTS (cHTS) data from EPA's invitrodb v3.5 (Feshuk et al. 2022).
  - This data set incorporates curation flags for chemical quality control, curve fit, and assay performance, as well as additional levels of expert review.
  - cHTS data are available through NICEATM's Integrated Chemical Environment (ICE; https://ice.ntp.niehs.nih.gov/).
- Technological interference refers to an assay response that is driven by the interaction of chemicals with the assay technology rather than by true bioactivity.
  - Chemical structure can result in false signals for fluorescence through quenching or autofluorescence, and for luminescence through luciferase inhibition.
  - Many assays in Tox21/ToxCast use either luminescence or fluorescence readouts which are susceptible to this type of interference.
- This poster presents the approach to integrate technological interference into the ICE cHTS curation pipeline and ultimately into ICE Tools.
  - The ICE interference flags will allow users to make more informed decisions when interpreting bioactivity calls.

Summary

- Technological interference is an important consideration in the interpretation of fluorescence and luminescence assays.
- Flags identifying such interference are being incorporated into the ICE cHTS pipeline to highlight potential limitations in data interpretation.
- We searched EPA's invitrodb for luminescence and fluorescence assay detection technologies in Tox21 and ToxCast data. Tox21 interference-specific assays and InterPred predictions were used to determine interferent chemicals in these assays.
- We created eight new warning flags for potentially problematic chemical/assay endpoint pairs and incorporated them into the cHTS dataset. These flags allows users to more critically interpret bioactivity calls that could be false positives.
- Warning flags for technological interference will soon appear in ICE cHTS dataset downloads and the ICE Curve Surfer tool.

Methods: Sourcing Interferent Chemicals

- To identify potential interferent chemicals, we used data from Tox21 assays specifically designed to test luciferase and auto-fluorescence interference (Figure 1; Borrel et al. 2020a).
- These assays were used to test chemicals in the Tox21 10k library for luciferase inhibition and autofluorescence. The autofluorescence assays evaluated red, blue, and green fluorescence across HepG2 and HEK293 cell lines.
- InterPred is an open-access quantitative structure-activity relationship (QSAR) modeling workflow developed to predict luciferase inhibition and autofluorescence under blue, green, and red wavelengths.
- For ~500 ToxCast chemicals that were not tested in the Tox21 interference assays, we used the InterPred tool to predict luciferase and fluorescence interference (Figure 2; Borrel et al. 2020b; https://sandbox.ntp.niehs.nih.gov/interferences/).

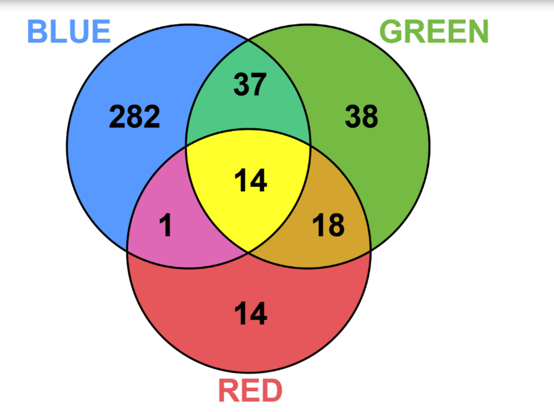

Figure 1: Venn Diagram of interferent chemicals by fluorescence color. Adapted from Borrel et al. 2020a.

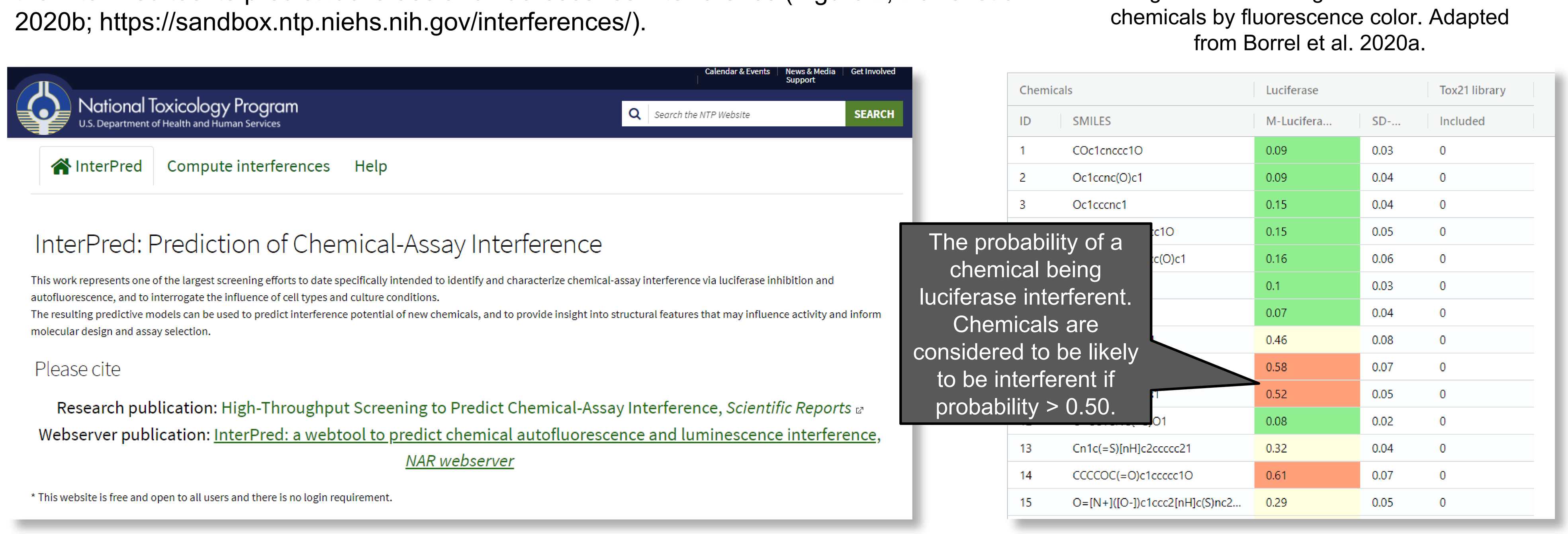

Figure 2: Left panel shows home page of the InterPred tool. Right panel shows InterPred output, which provides chemical structure in Simplified Molecular-Input Line-Entry System (SMILES) format ("SMILES"), the chemical's probability of being interferent ("M-Luciferase"), the standard deviation of probabilities from multiple random forest models ("SD"), and whether or not a chemical is present in the Tox21 library ("Included").

Methods: Workflow for Flag Generation

Step 1: Retrieved Tox21 and ToxCast data

- Data from EPA's invitrodb v3.5 (Feshuk et al. 2022) using EPA's tcpl v.2.1.0 R package (Filer et al. 2017).

Step 2: Identified assay endpoints that used luciferase and fluorescence detection technologies

- Reviewed "detection technologies" descriptions within invitrodb v3.5 to identify assays that used luminescence and fluorescence to determine readouts.
- Reviewed additional methods descriptions within invitrodb v3.5, assay information from the National Institutes of Health Tripod website (https://tripod.nih.gov/tox/), and relevant references to determine which assays used luciferase and the wavelength/color of fluorescence. Assays that used fluorescence other than green, red, or blue were excluded from flagging.

Step 3: Determined which chemicals were tested in assay endpoints from Step 2

- Used tcpl to pull chemical lists from invitrodb v3.5.

Step 4: Identified which chemicals from Step 3 were potentially interferent

- Cross-referenced chemical lists with chemicals identified as luciferase-inhibiting or autofluorescent in the Tox21 interference assays.
- Generated InterPred predictions for ToxCast chemicals that were not tested in the Tox21 interference assays.

Step 5: Assigned interference flags to chemical-assay endpoint pairs

- If a chemical was interferent for a technology and the assay endpoint used that technology, a flag was created and incorporated into the cHTS data.

Results: Integrating Interference Flags into ICE Tools

- Future ICE Curve Surfer Tool results will integrate interference flags (Figure 3). This tool allows users to explore cHTS concentration-response curves, under which text fields provide users details pertaining to the results, including interference warning flag(s).

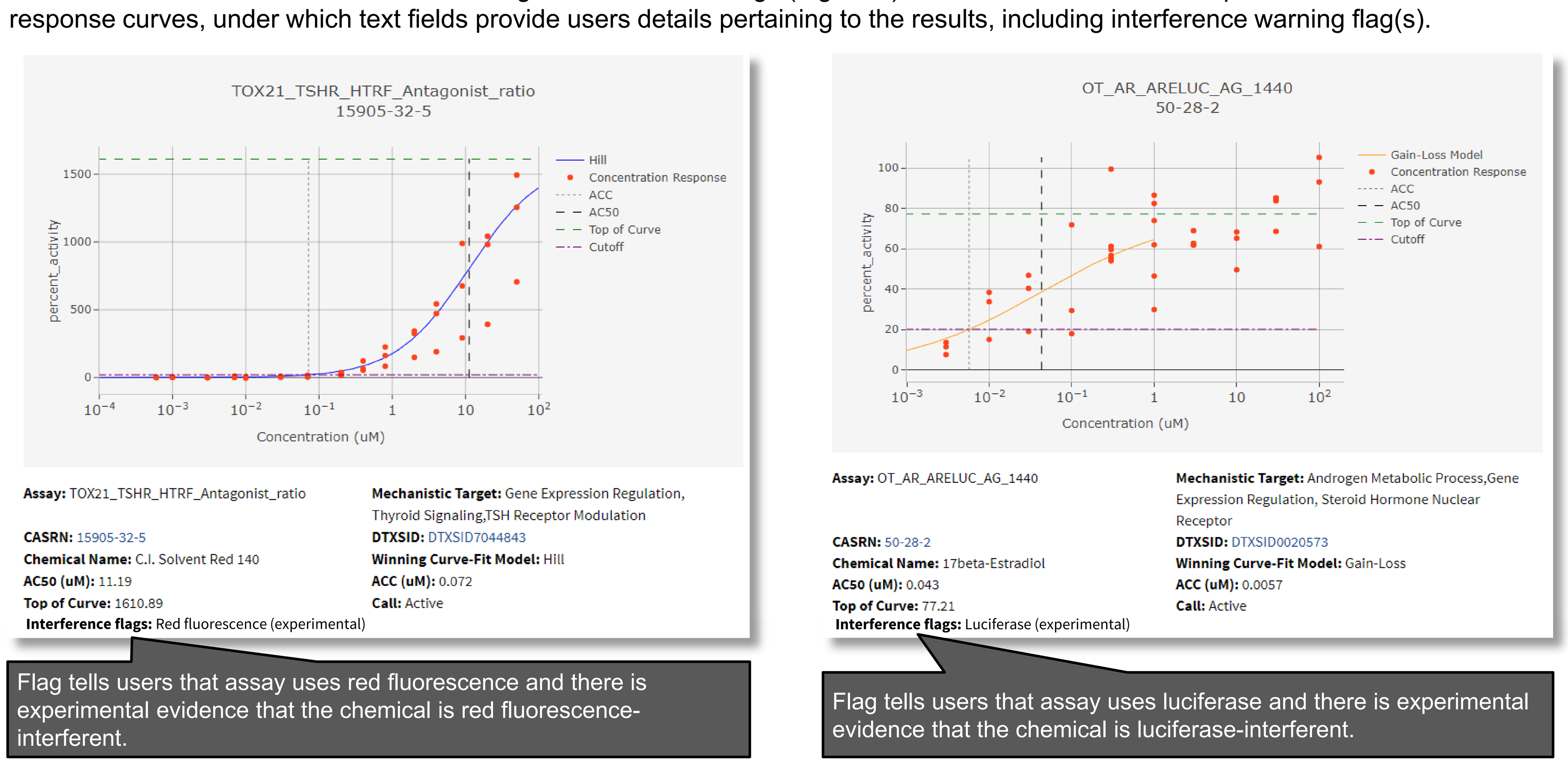

Figure 3: Example of how interference flags will be incorporated into future updates of the ICE Curve Surfer tool.

Results: Number of Interferent Assays and Chemicals

Detection Technology Type	Number of Assay Endpoints	Number of Interferent Chemicals	Number of Positive Calls
Luciferase	120	567	10,669
Blue fluorescence only	301	330	1,375
Green fluorescence only	277	109	865
Red fluorescence only	24	45	42
Blue and green fluorescence	68	55	260
Red and green fluorescence	17	35	36

- Blue fluorescence was the most prevalent detection technology and had the most interferent chemicals. Red fluorescence was the least prevalent technology and had the fewest interferent chemicals.
- We created eight interference flags:
  1. Blue fluorescence interferent (experimental)
  2. Blue fluorescence interferent (predicted)
  3. Green fluorescence interferent (experimental)
  4. Green fluorescence interferent (predicted)
  5. Red fluorescence interferent (experimental)
  6. Red fluorescence interferent (predicted)
  7. Luciferase interferent (experimental)
  8. Luciferase interferent (predicted)
- Interference flags will be provided as warning flags and do not affect bioactivity calls.
  - This differs from the ICE QC-Omit and Flag-Omit flags, which override the bioactivity calls and recommend that the assay endpoint-chemical pair be omitted.

Results: Interference Flags

- Technological Interference will be added as an endpoint to the cHTS download file from ICE in the future.
  - The cHTS download file is available from the ICE Data Sets page (Figure 4; https://ice.ntp.niehs.nih.gov/DATASETDESCRIPTION).

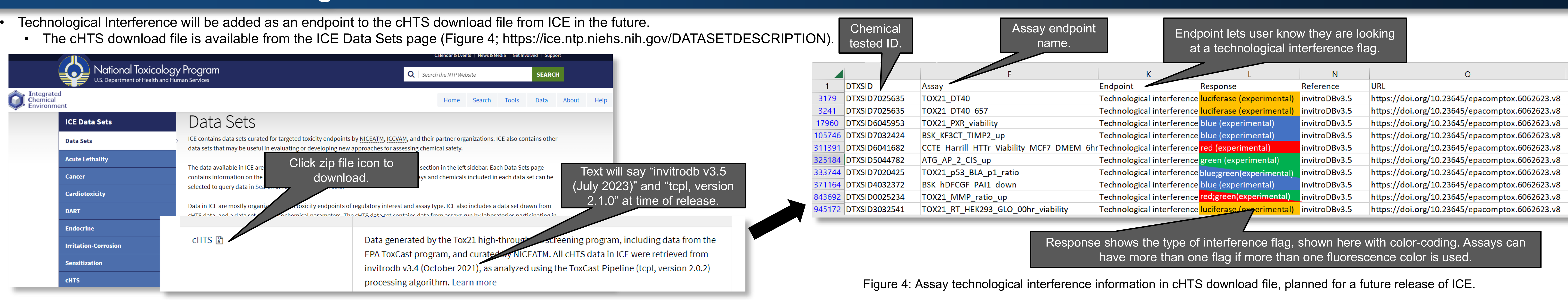

Figure 4: Assay technological interference information in cHTS download file, planned for a future release of ICE.

References

Borrel et al. 2020a. Nucleic Acids Res. 48(W1):W586-W590. doi: 10.1093/nar/gkaa378.

Borrel et al. 2020b. Sci Rep. 10(1):3986. doi: 10.1038/s41598-020-60747-3.

Feshuk et al. 2022. Invitrodb version 3.5 release. EPA, Washington, DC. https://doi.org/10.23645/epacomptox.6062623.v8

Filer et al. 2017. tcpl: ToxCast Data Analysis Pipeline. R package version 2.1.0 (2022), https://CRAN.R-project.org/package=tcpl.

Acknowledgments

This project was funded with federal funds from NIEHS, NIH, under Contract No. HHSN273201500010C. Since the poster was written as part of the official duties of the authors, it can be freely copied. We thank Catherine Sprankle, Inotiv, for editorial input.

*A.L. Karmaus is currently affiliated with Syngenta, Greensboro, NC

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