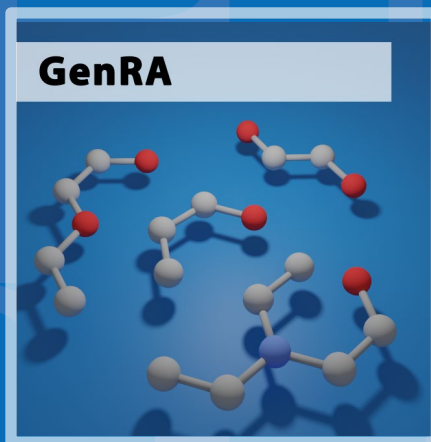


EPA Center for Computational Toxicology and Exposure (CCTE)

# Generalized Read-Across (GenRA) Virtual Training



Grace Patlewicz, U.S. EPA CCTE

# Meeting Logistics



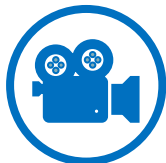
If you have a question for the presenters, use the Zoom Q&A feature.



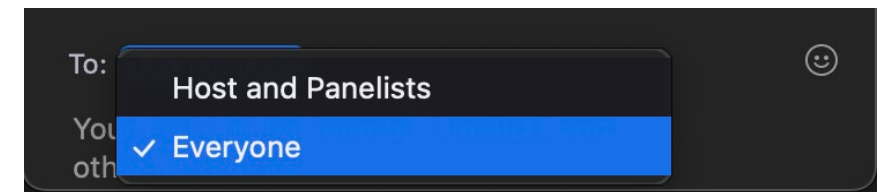
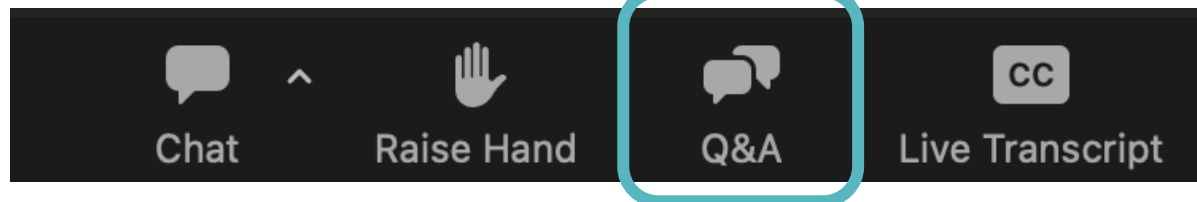
If you prefer to call in by phone, please refer to the Meeting Invite to obtain the phone number and passcode.



Please update your Zoom name to your first and last name.



This webinar will be recorded and posted to the CCTE NAMs Training website.



# EPA NAMs Pilot Training Program

- New Approach Methodologies (NAMs) Training Program is a deliverable in the Agency's Work Plan, first released in 2019 and updated in 2021.
  - Previous trainings include ECOTOX, CompTox Chemicals Dashboard
- Goal: Develop, implement and maintain an engaging training program.
  - Interactive case studies to encourage active learning
  - Train the trainer
  - Obtain feedback
- More trainings are being planned
- The EPA NAMs training website includes existing training resources, including recordings and guidance documents.



# AGENDA

- Welcome and Introductions
- Background of GenRA
- Functionalities and Features of GenRA
- A Walk Through Example
- Summary
- Breakout Session
  - A walk through with a data poor chemical
  - Practical applications of GenRA
  - Read-across predictions and interpretations
- Debrief Session

# Conflict of Interest Statement

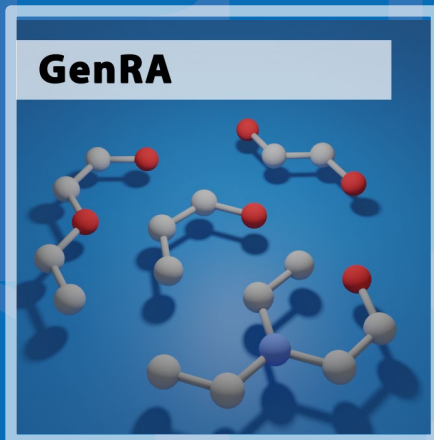
- No conflicts of interest to declare.

- Disclaimer:

The views expressed herein are those of the presenter and do not necessarily reflect the views or policies of the U.S. EPA

EPA Center for Computational Toxicology and Exposure (CCTE)

# Generalized Read-Across (GenRA) Virtual Training



Grace Patlewicz, U.S. EPA CCTE

# GenRA Project Team

## **Facilitators**

Dr. Grace Patlewicz

Dr. Imran Shah

Dr. Louis Groff

## **GenRA Development Team**

Terry Brown\*

Kenta Baron-Furuyama

Freddie Valone\*

Sean Hamilton\*

PK Do

## **Contributors**

George Helman

Tia Tate

Willysha Jenkins

Mark Nelms

Matthew Adams\*

Brett Hagan\*

\*current

# Generalized Read-Across (GenRA)

- Predicting toxicity as a similarity-weighted activity of nearest neighbors based on chemistry and bioactivity descriptors (Shah et al, 2016)
- **Goal:** To establish an objective performance baseline for read-across and quantify the uncertainty in the predictions made

$$y_i^{\beta, \alpha} = \frac{\sum_j^k S_{ij}^{\alpha} x_j^{\beta}}{\sum_j^k S_{ij}^{\alpha}}$$

Jaccard similarity:

$$S_{ij} = \frac{\sum_l (x_{il} \wedge x_{jl})}{\sum_l (x_{il} \vee x_{jl})}$$

$$\alpha \equiv \{chm, bio, bc\}$$

$$\beta \equiv \{bio, tox\}$$

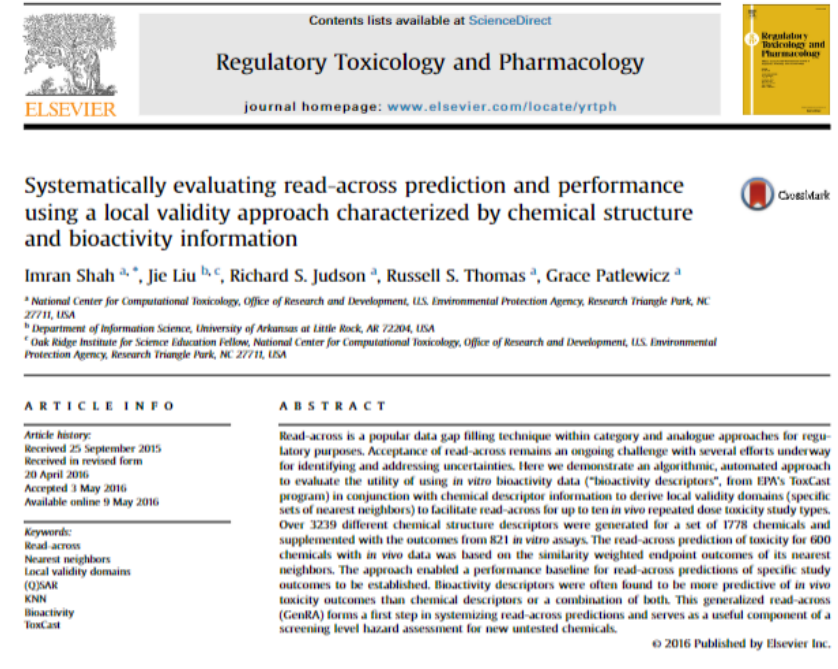
$y_i$  = predicted activity of chemical ( $c_i$ )

$x_j^{\beta}$  = activity of  $c_j$  in  $\beta$

$S_{ij}^{\alpha}$  = Jaccard similarity between  $x_i^{\alpha}$ ,  $x_j^{\alpha}$

$k$  = up to  $k$  nearest neighbours

Regulatory Toxicology and Pharmacology 79 (2016) 12–24





Navigating through the minefield of read-across tools: A review of in silico tools for grouping

Grace Patlewicz<sup>a,\*</sup>, George Helman<sup>a,b</sup>, Prachi Pradeep<sup>a,b</sup>, Imran Shah<sup>a</sup>



(Patlewicz et al., 2017)

**Table 1**  
Summary of key features of selected publicly available read-across tools.

	AIM	ToxMatch	Ambit	OECD Toolbox	CBRA	ToxRead	CIIPro
Development timeline	Java based version is dated 2012. Initial development of web version was 2005.	First public version released in Dec 2006	Original AMBIT tool was developed in 2004–2005	Proof of concept released in 2008	Implementation of the Low et al. [27]	Implementation of Gini et al. [22]	Implementation described in Russo et al. [45]
Type of Tool	Standalone	Standalone	Web-based and standalone	Standalone or Client/Server	Standalone	Standalone	Web-based
Latest Version	1.01 (Nov 2013) Static	1.07 (Jan 2009) Static	3.0.3 Ongoing Enhanced in 2013–2015	3.4 (July 2016) Version 4 released April 2017 Ongoing	0.75 First release	0.11 BETA Ongoing	First release
Developed by	SRC Inc	Ideaconsult Ltd	Ideaconsult Ltd	LMC, Bourgas	Fourches Lab at North Carolina State University	Istituto di Ricerche Farmacologiche Mario Negri	Zhu Research Group at Rutgers University
Available from	<a href="https://www.epa.gov/tsca-screening-tools/analogue-identification-methodology-aim-tool">https://www.epa.gov/tsca-screening-tools/analogue-identification-methodology-aim-tool</a>	<a href="https://eurl-ecvam.jrc.ec.europa.eu/laboratories-research/predictive_toxicology/qsar_tools/toxmatch">https://eurl-ecvam.jrc.ec.europa.eu/laboratories-research/predictive_toxicology/qsar_tools/toxmatch</a>	<a href="http://cefic-lri.org/lri_toolbox/ambit/">http://cefic-lri.org/lri_toolbox/ambit/</a>	<a href="http://www.qsartoolbox.org">www.qsartoolbox.org</a>	<a href="http://www.fourches-laboratory.com/software">http://www.fourches-laboratory.com/software</a>	<a href="http://www.toxread.eu/">http://www.toxread.eu/</a>	<a href="http://ciipro.rutgers.edu/">http://ciipro.rutgers.edu/</a>
Accepted Chemical Input	CAS, Name, SMILES, structure drawing/import	CAS, Name, SMILES, InChI	Name, identifiers, SMILES, InChI	CAS, Name, SMILES, structure drawing, MOL, sdf	Mol file, descriptors as txt	SMILES	PubChem CID, CAS, IUPAC, SMILES, InChI
Endpoint Coverage	N/A	Any based on user input	IUCLID <sup>a</sup> 5-supported endpoints (43 total)	Any as per the regulatory endpoints	Any based on user input	Mutagenicity and Bioconcentration Factor (BCF)	Any based on user input
Analogue Identification Approach	Fragment matching	Distance and correlation based similarity indices based on descriptors or fingerprints	Substructure or similarity searching using structure, name, SMILES, InChI	Category definition followed by subcategorisations	Tanimoto distance using chemical and biological descriptors	VEGA similarity algorithm	Weighted Estimated Biological Similarity
Neighbour Selection	Automatic	Automatic	Manual	Automatic + Manual Filter	Automatic	Automatic	Automatic + Manual Filter
Data Source	Tool provides inventory index	User provided or tool provided	User and tool provided	User provided or tool provided	User provided	Tool provided as a result of the EU ANTARES project	User provided but tool provides PubChem <i>in vitro</i> data
Quantitative vs Qualitative	N/A	Both	User determined - Qualitative	Both	Qualitative	Qualitative for mutagenicity, quantitative for BCF	Qualitative
Visualisation	None	Standard 2D plots, histograms and similarity matrix	None	Standard 2D Plots	Radial plot of neighbours	Interactive Neighbour plot	Activity Plot
Output/Export	Output reports in the form of HTML, pdf or Excel	sdf or txt files of data, image files of plots	Assessment report as docx or xlsx, data matrix as xlsx	IUCLID format, pdf and rtf files of prediction report, text files of data, image files of plots etc	NA	Image file of plot	Tabulation of predictions and image of similarity plot

G. Patlewicz et al. / Computational Toxicology 3 (2017) 1–18

<sup>a</sup> IUCLID stands for International Uniform Chemical Information Database. IUCLID is a software program for the administration of data on chemical substances first developed to fulfill EU information requirements under REACH.

# GenRA – Overall goal



Quantify the contribution that different similarity contexts play in toxicity prediction and how that differs depending on the toxicity endpoint of interest, the chemical of interest and whether it mirrors expert driven read-across



Quantify level of confidence for prediction made



=> objective, reproducible read-across assessments

# GenRA Summary

- **Read-across** is a commonly used technique to fill data gaps.
- **Read-across** is an expert-driven approach which is challenging from a reproducibility and scalability perspective.
- **GenRA** is an algorithmic approach to permit objective and reproducible read-across predictions.
  - Presents opportunities for how NAM data can be incorporated.
- **GenRA v1.0 (2018)** established a baseline in performance.
  - The approach relied on chemical descriptors to predict binary toxicity values. Our ongoing research programme aims to characterize other contexts of similarity (e.g., mechanistic, reactivity, and metabolism) and quantify their contribution in predicting toxicity outcomes.
- Subsequent versions through to the current Version 3.2 have added a number of significant new features whilst addressing bug fixes.

# GenRA Milestones

## Released March 2021

- A python package (genra-py) to facilitate batch processing using user specific datasets.

## Released December 2021

- A standalone web app linked to the CCD

## Released September 2022

- New visualization tool: Neighborhood explorer graph
- Physical property and e-fate visualization tool
- Potency prediction using ToxRefDB data
- Hit call outcome prediction of ToxCast assays

**Version  
1**

## Released September 2018

- Integrated into the EPA CompTox Chemicals Dashboard (CCD)

**Version  
2**

**Version  
3**

## Released February 2022

- UI rebuilt using AG Grid
- Custom fingerprints
- Ketcher drawing palette

**Version  
3.1**

## Released March 2023

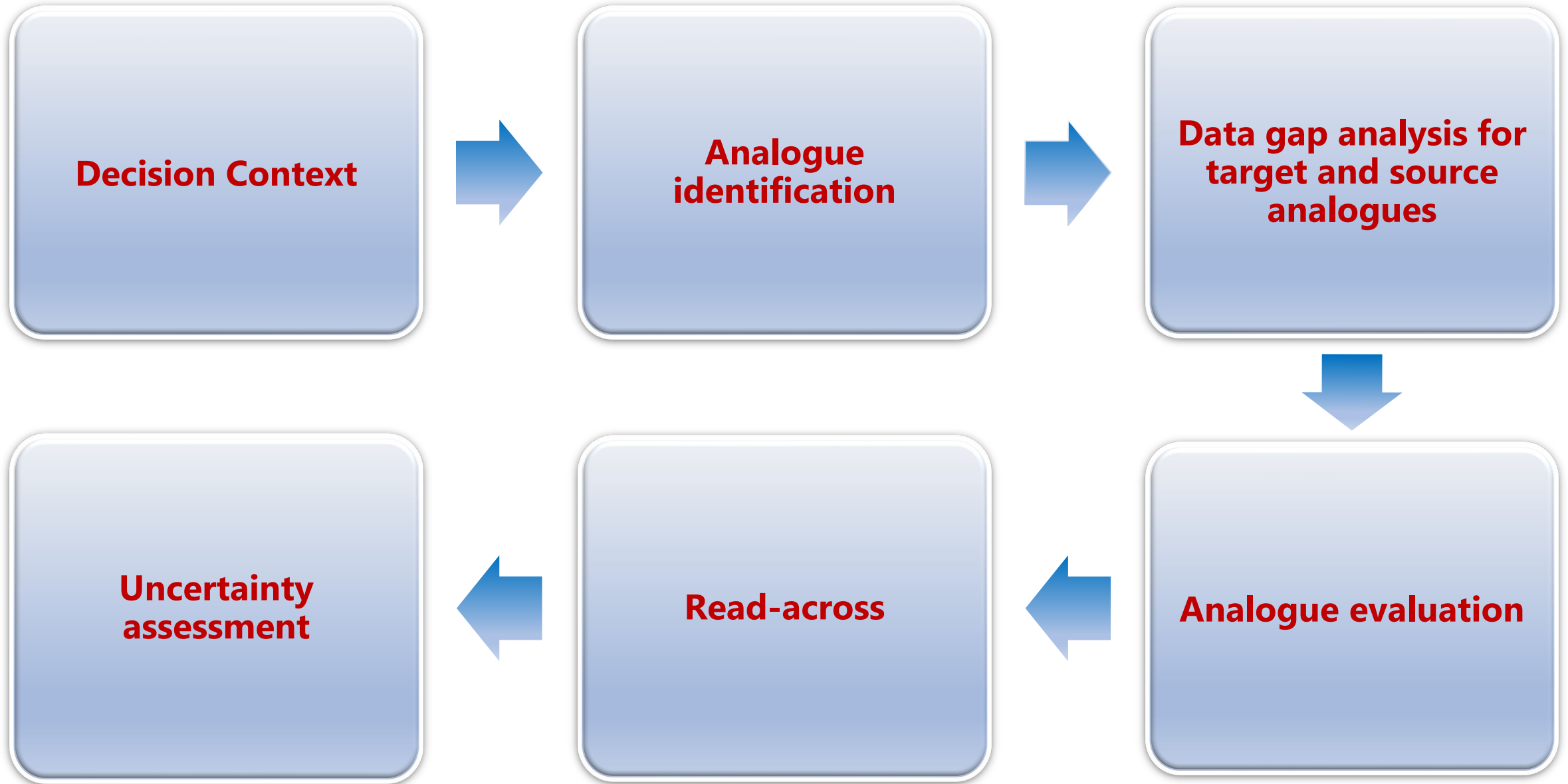
- Major speed up
- New download and sorting options
- New AIM chemical fingerprints

**Version  
3.2**

# GenRA Functionalities

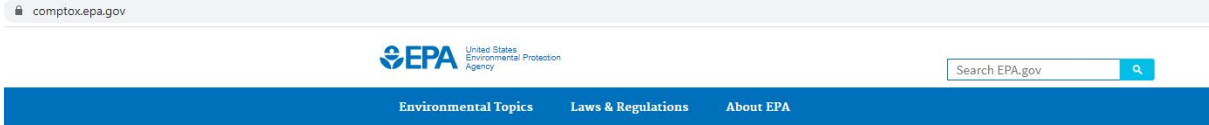
- **Search** for candidate source analogues based on different identifiers e.g. SMILES, CAS, Name etc. or introduce a chemical structure/SMILES using the Ketcher drawing pad;
- **Identify** candidate source analogues on the basis of chemical, bioactivity fingerprints or a combination of both using the custom hybrid option;
- **Download** the top 100 candidate source analogues, their pairwise similarities and their chemical/bioactivity fingerprint matrices;
- **View** chemical neighborhoods via a neighborhood explorer graph visualization tool and filter on the basis of ToxCast or ToxRefDB data;
- **Compare** the distribution of relevant physicochemical properties across candidate source analogues;
- **Make binary *in vitro* predictions** of ToxCast assay outcomes or binary *in vivo* toxicity predictions on the basis of study type-toxicity effect using ToxRefDB data;
- **Make potency-based predictions** of study type-toxicity effects using ToxRefDB data;
- **Sort predictions** based on number of positive/negative toxicity effects or on the basis of the prediction confidence. Download the predictions in a spreadsheet format for subsequent review and assessment.

# Conceptual read-across workflow in GenRA



# Accessing GenRA

# Main Entry Point

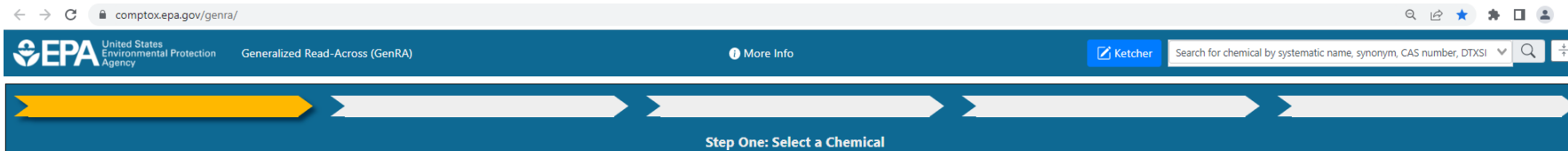


**GenRA main entry point is from the comptox portal**

**[comptox.epa.gov](https://comptox.epa.gov)**

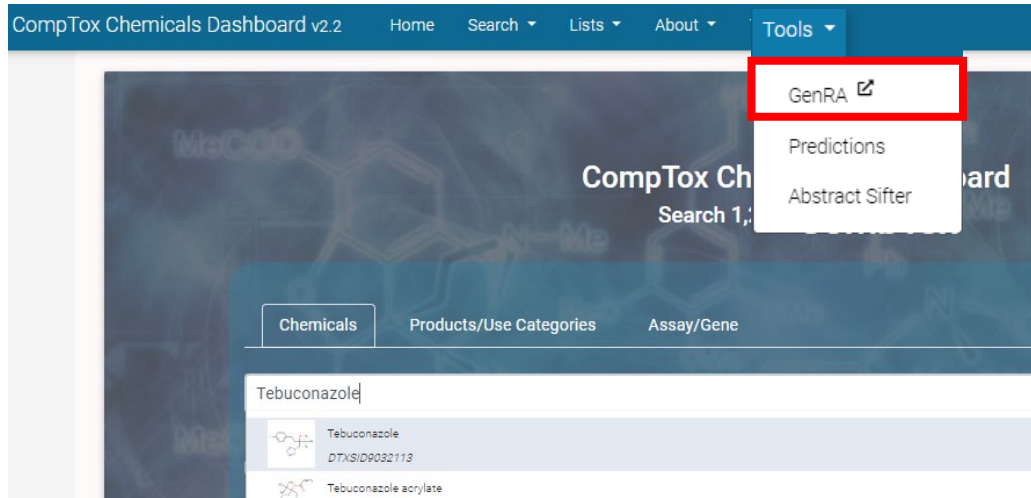


**GenRA entry point at [comptox.epa.gov/genra/](https://comptox.epa.gov/genra/)**





# Alternative Entry Points



CompTox Chemicals Dashboard v2.2

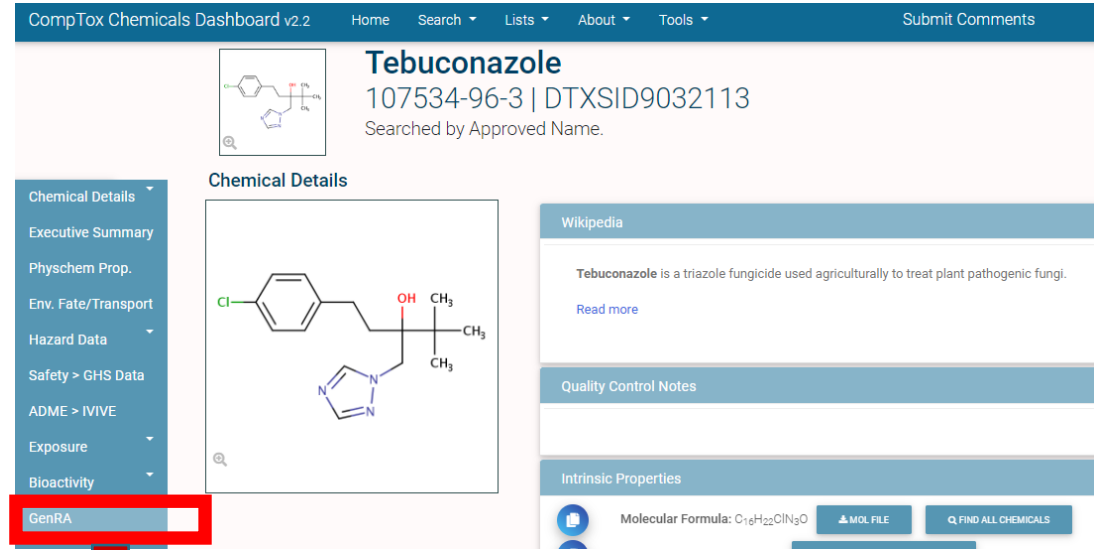
Home Search Lists About Tools

GenRA  
Predictions  
Abstract Sifter

Chemicals Products/Use Categories Assay/Gene

Tebuconazole  
DTXSID9032113

Tebuconazole acrylate



CompTox Chemicals Dashboard v2.2

Home Search Lists About Tools Submit Comments

**Tebuconazole**  
107534-96-3 | DTXSID9032113  
Searched by Approved Name.

Chemical Details

Chemical Details  
Executive Summary  
Physchem Prop.  
Env. Fate/Transport  
Hazard Data  
Safety > GHS Data  
ADME > IVIVE  
Exposure  
Bioactivity

GenRA

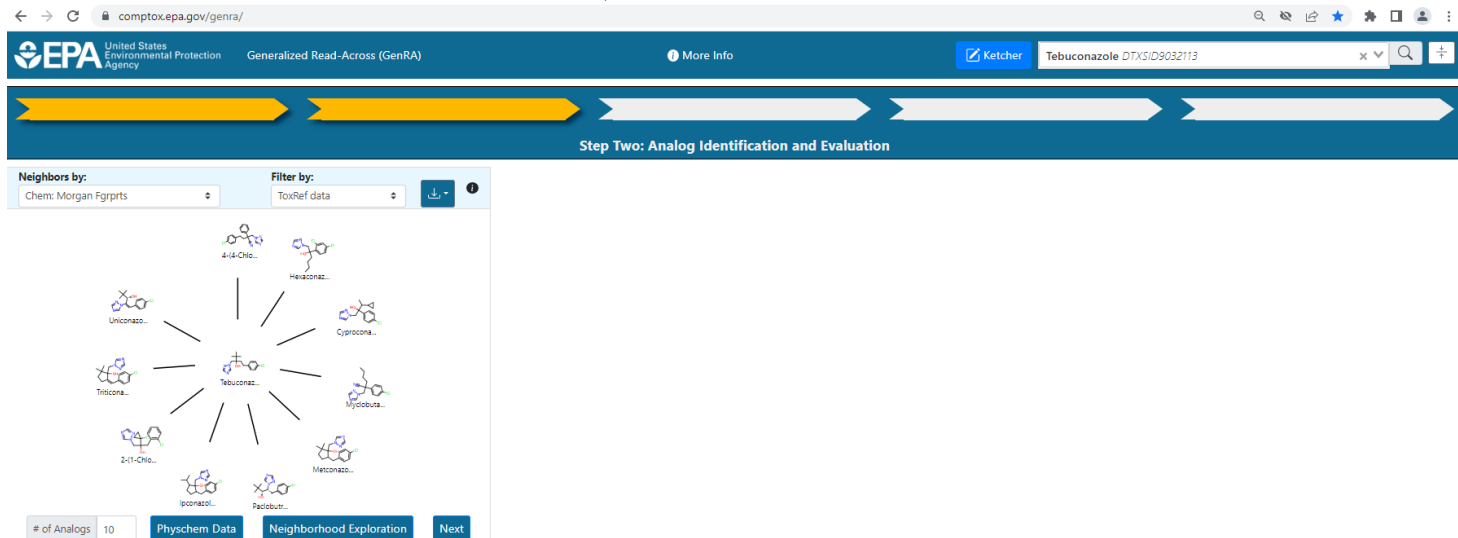
Wikipedia

Tebuconazole is a triazole fungicide used agriculturally to treat plant pathogenic fungi.  
[Read more](#)

Quality Control Notes

Intrinsic Properties

Molecular Formula: C<sub>16</sub>H<sub>22</sub>ClN<sub>3</sub>O [MOL FILE](#) [FIND ALL CHEMICALS](#)



comp tox.epa.gov/genra/

EPA United States Environmental Protection Agency Generalized Read-Across (GenRA) More Info Ketcher Tebuconazole DTXSID9032113

Step Two: Analog Identification and Evaluation

Neighbors by: Chem: Morgan Fgprfts Filter by: ToxRef data

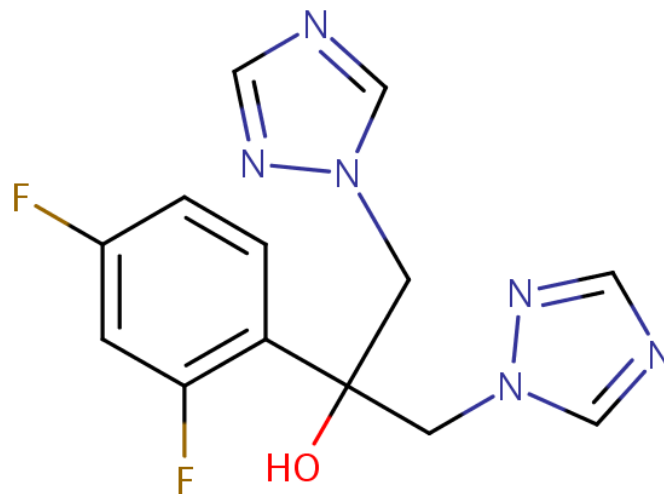
10 of Analogs

Physchem Data Neighborhood Exploration Next

# **Demonstration of GenRA (Part 1)**

## **Example Walkthrough**

# Fluconazole (A Data Rich Chemical)



- **IUPAC Name:** 2-(2,4-Difluorophenyl)-1,3-bis(1H-1,2,4-triazol-1-yl)propan-2-ol
- **CAS RN:** 86386-73-4
- **DSSTox substance identifier (DTXSID):** DTXSID3020627
- **Molecular Formula:** C<sub>13</sub>H<sub>12</sub>F<sub>2</sub>N<sub>6</sub>O
- **SMILES:** OC(CN1C=NC=N1)(CN1C=NC=N1)C1=C(F)C=C(F)C=C1
- **InChIKey:** RFHAOTPXVQNOHP-UHFFFAOYSA-N

# GenRA Workflow Overview

**Step Five: Filter by Endpoint or Analog**

**Panel 1**

Neighbors by: Chem Morgan Fgprts | Filter by: ToxRef data

# of Analogs: 10

Physchem Data | Neighborhood Exploration

Run Read-Across | GenraPred | Sort: Endpoint name

**Panel 2**

Summary Data Gap Analysis

	Fluconazole	Hexaconazole	Tebuconazole	Flusilazole	Cyproconazole	2-(1-Chloro-4-Chlorophenyl)-2-phenyl-2-(1H-1,2,4-triazol-1-yl)ethanol	Myclobutanil	4-(4-Chlorophenyl)-2-phenyl-2-(1H-1,2,4-triazol-1-yl)ethylbutyrate	Epoxiconazole	Tetraconazole	Metconazole
Fluconazole	340	15	43	43	0						
Hexaconazole	449	18	35	55	185						
Tebuconazole	412	19	32	49	83						
Flusilazole	458	9	34	39	179						
Cyproconazole	442	16	31	53	225						
2-(1-Chloro-4-Chlorophenyl)-2-phenyl-2-(1H-1,2,4-triazol-1-yl)ethanol	210	18	41	54	87						
Myclobutanil	457	15	34	53	198						
4-(4-Chlorophenyl)-2-phenyl-2-(1H-1,2,4-triazol-1-yl)ethylbutyrate	435	17	41	59	194						
Epoxiconazole	192	11	50	60	40						
Tetraconazole	450	20	35	59	186						
Metconazole	252	15	46	58	41						

Rows: 11 | Total Rows: 11

Min: 1 | Min: 1 | Similarity W

**Panel 3**

Group: ToxRef | By: Tox Fingerprint | Pagination

	Fluconazole	Hexaconazole	Tebuconazole	Flusilazole	Cyproconazole	2-(1-Chloro-4-Chlorophenyl)-2-phenyl-2-(1H-1,2,4-triazol-1-yl)ethanol	Myclobutanil	4-(4-Chlorophenyl)-2-phenyl-2-(1H-1,2,4-triazol-1-yl)ethylbutyrate	Epoxiconazole	Tetraconazole	Metconazole
adrenal gland											
alanine aminotransf...											
albumin											
alkaline phosphatas...											
aminopyrine-n-de...											
anilicytosis											
appearance and col...											
blood clotting											
blood vessel											

Rows: 353 | Total Rows: 353

1 to 9 of 353 | Page 1 of 40

Hide Pagination | Download: File Type

**Panel 4**

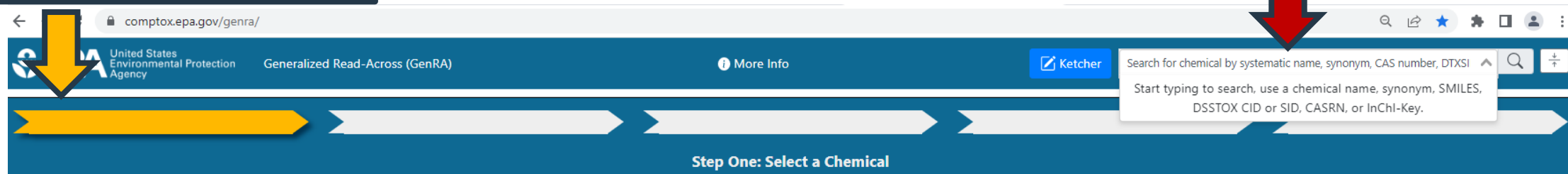
Assay endpoint

	Fluconazole	Hexaconazole	Tebuconazole	Flusilazole	Cyproconazole	2-(1-Chloro-4-Chlorophenyl)-2-phenyl-2-(1H-1,2,4-triazol-1-yl)ethanol	Myclobutanil	4-(4-Chlorophenyl)-2-phenyl-2-(1H-1,2,4-triazol-1-yl)ethylbutyrate	Epoxiconazole	Tetraconazole	Metconazole
CHR:albumin											
CHR:alkaline phosp...											
CHR:aminopyrine-n...											
CHR:anilicytosis											
CHR:appearance an...											
CHR:blood clotting											
CHR:blood vessel											
CHR:body weight											
CHR:bone											
CHR:bone marrow											
CHR:brain											
CHR:calcium											
CHR:cholesterol											
CHR:clinical signs											
CHR:cytochrome p4...											
CHR:ear											
CHR:erythrocyte											

# Basic Search

## Step 1: Select a Chemical

**Search for a chemical of interest  
(target) using the search box**



The screenshot shows the EPA GenRA website interface. A yellow arrow points to the EPA logo in the top left. A red arrow points to the search box in the top right. The search box contains the text: "Search for chemical by systematic name, synonym, CAS number, DTXSI" and "Start typing to search, use a chemical name, synonym, SMILES, DSSTOX CID or SID, CASRN, or InChI-Key." Below the search box is a blue button labeled "Ketcher". The main navigation bar includes the EPA logo, "United States Environmental Protection Agency", "Generalized Read-Across (GenRA)", and "More Info". A large yellow arrow points from the "Step 1: Select a Chemical" header to the search box. The bottom of the page features a blue banner with the text "Step One: Select a Chemical".

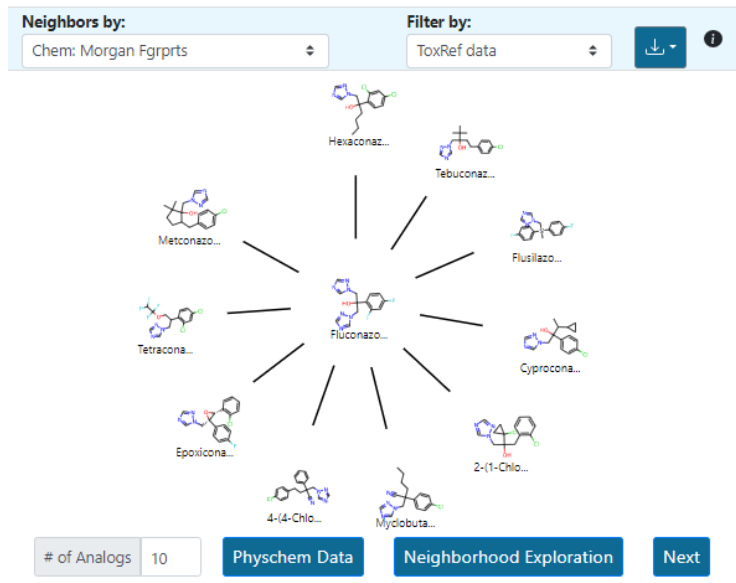
# Basic Search

## Step 1: Select a Chemical

Search for a chemical of interest  
(target) using the search box

The screenshot shows the EPA GenRA website interface. At the top, there is a search bar with a dropdown menu that lists search criteria: "Search for chemical by systematic name, synonym, CAS number, DTXSID", "Start typing to search, use a chemical name, synonym, SMILES, DSSTOX CID or SID, CASRN, or InChI-Key." Below the search bar, the text "Fluconazole DTXSID3020627" is entered and highlighted with a red oval. The page title is "Generalized Read-Across (GenRA)" and there is a "More Info" link. A yellow arrow points from the search bar to the search results section below.

## Step 2: Analogue Identification and Evaluation



- Radial plot with target in the center and source analogues (similar) ordered clockwise by decreasing similarity (Jaccard)

# Analogue Identification and Evaluation

## Step 2: Analogue Identification and Evaluation

EPA United States Environmental Protection Agency Generalized Read-Across (GenRA) More Info

Fluconazole DTXSID3020627

Step Two: Analog Identification and Evaluation

Neighbors by: Chem: Morgan Fgrprts Filter by: ToxRef data

# of Analogs 10 Physchem Data Neighborhood Exploration Next

- Radial plot with target in the center and source analogues (similar) ordered clockwise by decreasing similarity (Jaccard)

**Downloads**  
**Top ~100 nearest neighbors (CSV)**  
**Radial Plot Image (png)**

**NEW**

# Analogue Identification and Evaluation

## Step 2: Analogue Identification and Evaluation

EPA United States Environmental Protection Agency Generalized Read-Across (GenRA) More Info Ketcher Fluconazole DTXSID3020627

Step Two: Analog Identification and Evaluation

Neighbors by: Chem: Morgan Fgrprts Filter by: ToxRef data

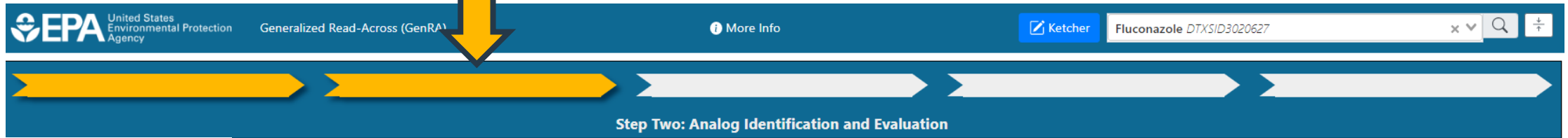
# of Analogs 10 Physchem Data Neighborhood Exploration Next

- Radial plot with target in the center and source analogues (similar) ordered clockwise by decreasing similarity (Jaccard)
- Default 10 analogues based on Morgan chemical fingerprints and prefiltered based on ToxRefDB data



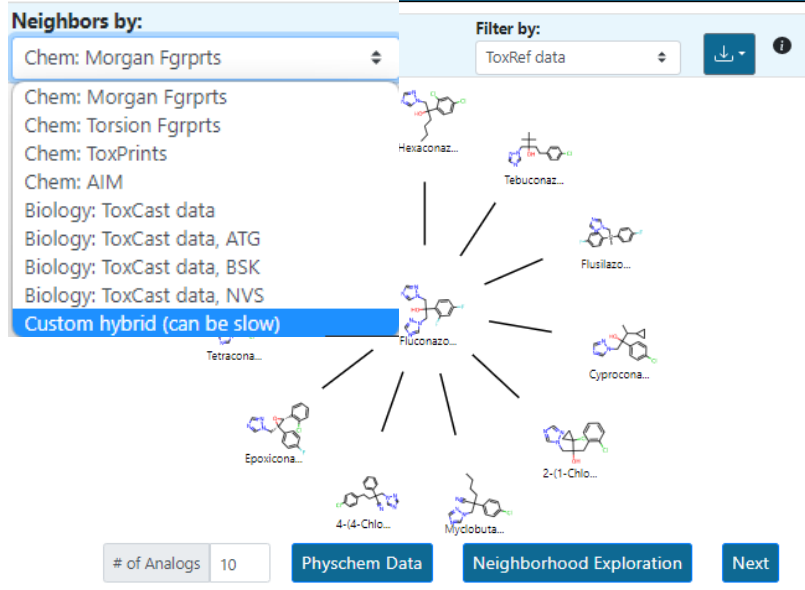
# Analogue Identification and Evaluation

## Step 2: Analogue Identification and Evaluation



EPA United States Environmental Protection Agency Generalized Read-Across (GenRA) More Info Ketcher Fluconazole DTXSID3020627

Step Two: Analog Identification and Evaluation



Neighbors by:

- Chem: Morgan Fgrprts
- Chem: Morgan Fgrprts
- Chem: Torsion Fgrprts
- Chem: ToxPrints
- Chem: AIM
- Biology: ToxCast data
- Biology: ToxCast data, ATG
- Biology: ToxCast data, BSK
- Biology: ToxCast data, NVS
- Custom hybrid (can be slow)

Filter by: ToxRef data

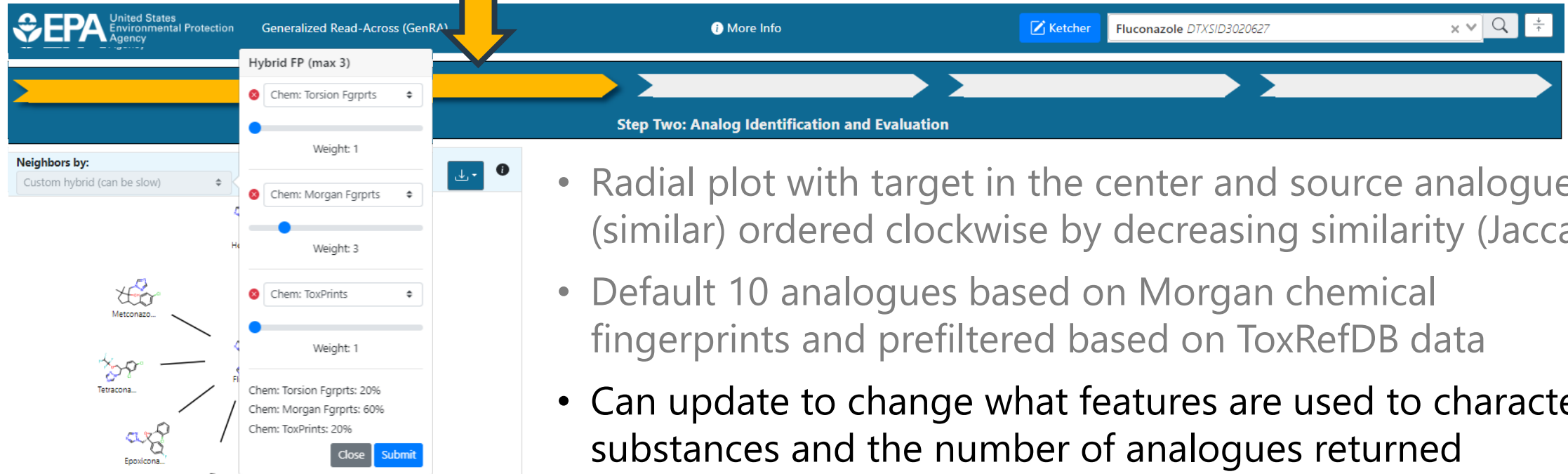
Hexaconaz... Tebuconaz... Fluconazo... Cyprocona... 2-(1-Chlo... Myclobuta... 4-(4-Chlo... Epoxiconaz... Tetracona...

# of Analogs 10 Physchem Data Neighborhood Exploration Next

- Radial plot with target in the center and source analogues (similar) ordered clockwise by decreasing similarity (Jaccard)
- Default 10 analogues based on Morgan chemical fingerprints and prefiltered based on ToxRefDB data
- Can update to change what features are used to characterize substances and the number of analogues returned

# Analogue Identification and Evaluation

## Step 2: Analogue Identification and Evaluation



Generalized Read-Across (GenRA) More Info Ketcher Fluconazole DTXSID3020627

Step Two: Analog Identification and Evaluation

Neighbors by: Custom hybrid (can be slow)

Hybrid FP (max 3)

- Chem: Torsion Fgrprts Weight: 1
- Chem: Morgan Fgrprts Weight: 3
- Chem: ToxPrints Weight: 1

Chem: Torsion Fgrprts: 20%  
Chem: Morgan Fgrprts: 60%  
Chem: ToxPrints: 20%

Close Submit

Metconazo...  
Tetracona...  
Epoxicona...

- Radial plot with target in the center and source analogues (similar) ordered clockwise by decreasing similarity (Jaccard)
- Default 10 analogues based on Morgan chemical fingerprints and prefiltered based on ToxRefDB data
- Can update to change what features are used to characterize substances and the number of analogues returned

### Custom Hybrid

Choose up to 3 fingerprints  
e.g. 20% Torsion Fingerprints, 60% Morgan  
Fingerprints, and 20% ToxPrints

# Analogue Identification and Evaluation

## Step 2: Analogue Identification and Evaluation

Step Two: Analog Identification and Evaluation

Neighbors by: Chem: Morgan Fgrprts

Filter by: ToxRef data, ToxRef data, ToxCast data, No filter (all data)

# of Analogs: 10

Physchem Data Neighborhood Exploration Next

- Radial plot with target in the center and source analogues (similar) ordered clockwise by decreasing similarity (Jaccard)
- Default 10 analogues based on Morgan chemical fingerprints and prefiltered based on ToxRefDB data
- Can update to change what features are used to characterize substances and the number of analogues returned
- Update radial plot to return analogues irrespective of data
  - ToxRefDB data filter is for *in vivo* data
  - ToxCast data filter is for *in vitro* data
  - No filter – returns the most similar analogues irrespective of data

# Analogue Identification and Evaluation

## Step 2: Analogue Identification and Evaluation

EPA United States Environmental Protection Agency Generalized Read-Across (GenRA) More Info Ketcher Fluconazole DTXSID3020627

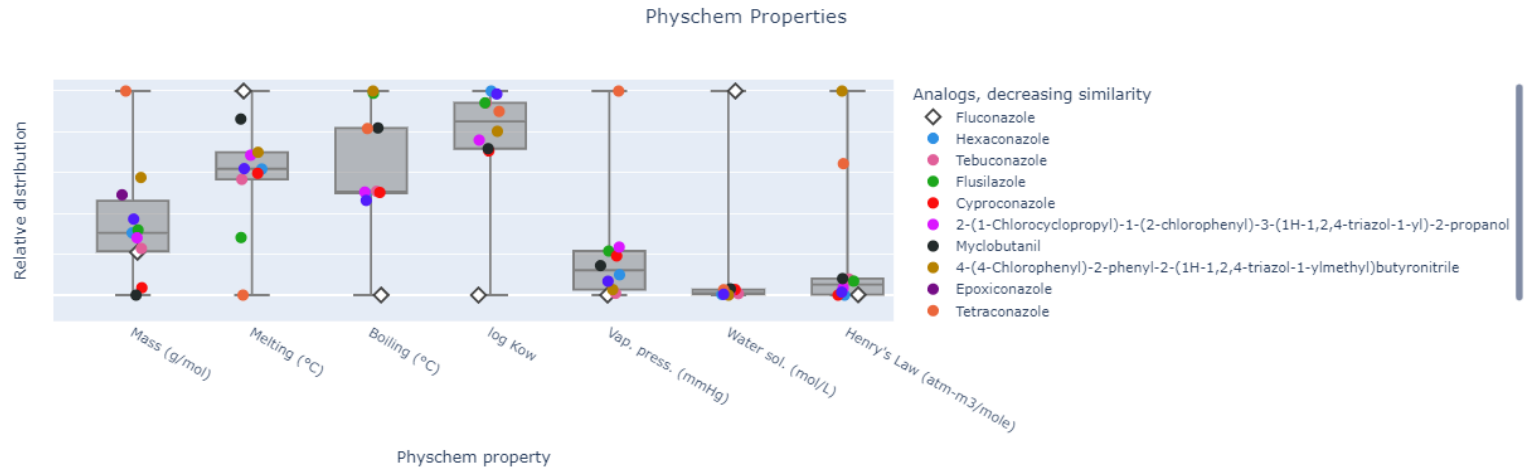
Step Two: Analog Identification and Evaluation

Neighbors by: Chem: Morgan Fgrpts Filter by: ToxRef data

# of Analogs 10

**Physchem Data** Neighborhood Exploration Next

### PhysChem Data



# Analogue Identification and Evaluation

## Step 2: Analogue Identification and Evaluation

EPA United States Environmental Protection Agency Generalized Read-Across (GenRA) More Info Ketcher Fluconazole DTXSID3020627

Step Two: Analog Identification and Evaluation

Neighbors by: Chem: Morgan Fgrprts Filter by: ToxRef data

# of Analogs 10 **Physchem Data** Neighborhood Exploration Next



# Analogue Identification and Evaluation

## Step 2: Analogue Identification and Evaluation

EPA United States Environmental Protection Agency Generalized Read-Across (GenRA) More Info Ketcher Fluconazole DTXSID3020627

Step Two: Analog Identification and Evaluation

Neighbors by: Chem: Morgan Fgrprts Filter by: ToxRef data

# of Analogs 10

**Physchem Data** Neighborhood Exploration Next



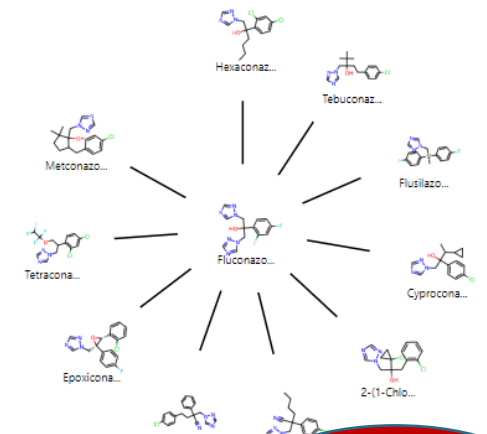
# Analogue Identification and Evaluation

## Step 2: Analogue Identification and Evaluation

EPA United States Environmental Protection Agency Generalized Read-Across (GenRA) More Info Ketcher Fluconazole DTXSID3020627

Step Two: Analog Identification and Evaluation

Neighbors by: Chem: Morgan Fgprpts Filter by: ToxRef data

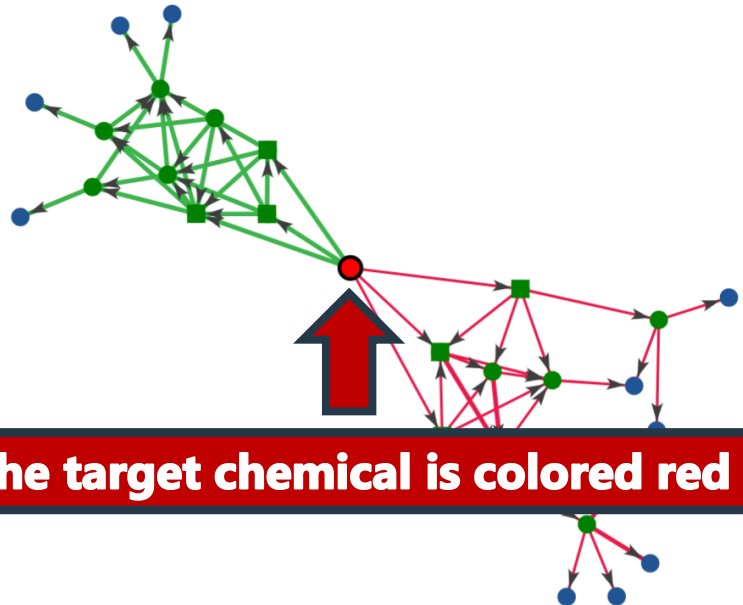


Neighbors: Hexaconaz..., Tebuconaz..., Metconazo..., Flusilazo..., Tetracona..., Cyprocona..., Epoxicon..., 2-(1-Chlo..., 4-(4-Chlo...

# of Analogs: 10 Physchem Data Neighborhood Exploration Next

**NEW**

Neighborhood Exploration Graph



**The target chemical is colored red**

Chemical Information

Name: Fluconazole  
Mol. Weight: 306.277  
DTXCID: DTXCID10627  
DTXSID: DTXSID3020627

Focus Chemical

Fingerprints

- Chem: Morgan Fgprpts
- Chem: Torsion Fgprpts
- Chem: ToxPrints
- Chem: AIM
- Biology: ToxCast data
- Biology: ToxCast data. ATG
- Biology: ToxCast data. BSK
- Biology: ToxCast data. NVS

Graph Type: all-paths  
Filter By: ToxRef data  
Update

# Analogue Identification and Evaluation

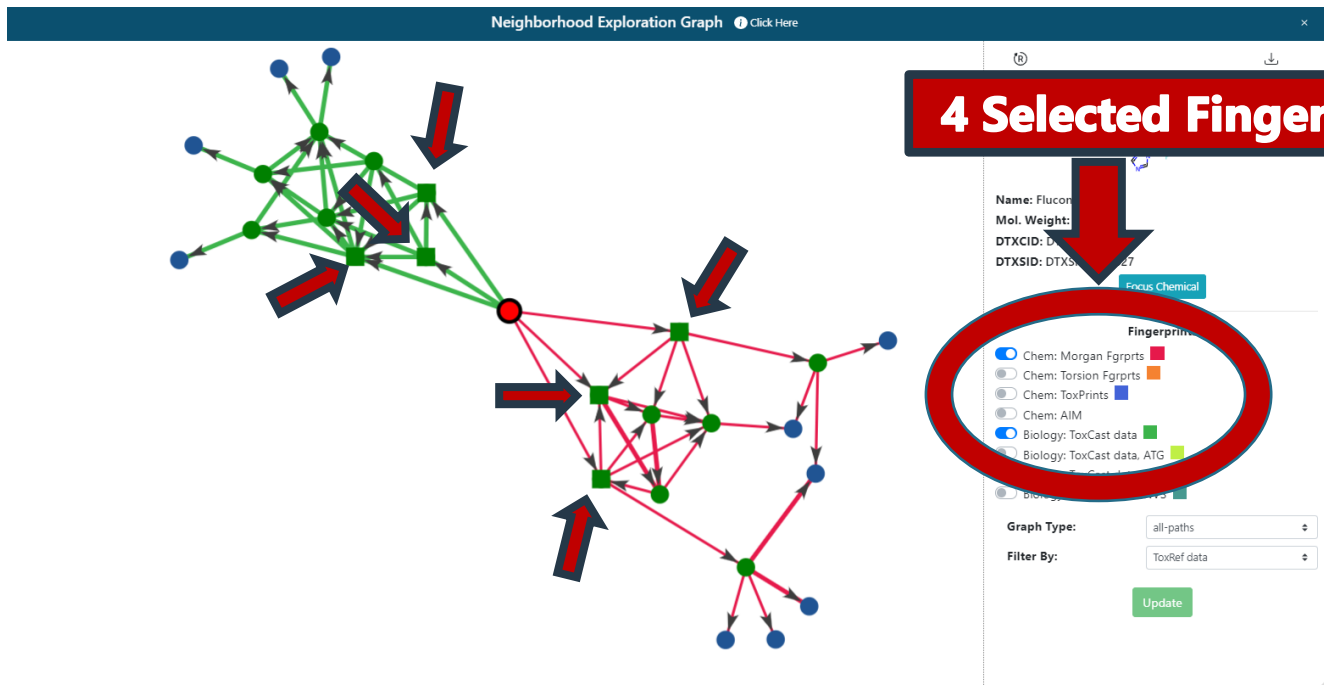
## Step 2: Analogue Identification and Evaluation

EPA United States Environmental Protection Agency Generalized Read-Across (GenRA) More Info Ketcher Fluconazole DTXSID3020627

Step Two: Analog Identification and Evaluation

Neighbors by: Chem: Morgan Fgprpts Filter by: ToxRef data

# of Analogs 10 Physchem Data Neighborhood Exploration Next



### 4 Selected Fingerprint Types

Name: Fluconazole  
Mol. Weight: 306.34  
DTXCID: DTXCID3020627  
DTXSID: DTXSID3020627

Fingerprint

- Chem: Morgan Fgprpts
- Chem: Torsion Fgprpts
- Chem: ToxPrints
- Chem: AIM
- Biology: ToxCast data
- Biology: ToxCast data, ATG

Graph Type: all-paths Filter By: ToxRef data Update

**NEW**

- Line thickness proportional to similarity between analogues (thicker=more similar)
- Squares are top 3 highest similarity analogues to your target for the chosen fingerprint type



# Analogue Identification and Evaluation

## Step 2: Analogue Identification and Evaluation

EPA United States Environmental Protection Agency Generalized Read-Across (GenRA) More Info Ketcher Fluconazole DTXCID3020627

Step Two: Analog Identification and Evaluation

Neighbors by: Chem: Morgan Fgprpts Filter by: ToxRef data

# of Analogs 10 Physchem Data Neighborhood Exploration Next

Neighborhood Exploration Graph Click Here

Chemical Information  
DTXCID: DTXCID10627  
DTXSID: DTXSID3020627  
Focus Chemical

Fingerprints

- Chem: Morgan Fgprpts
- Chem: Torsion Fgprpts
- Chem: Torsion Fgprpts
- Chem: AIM
- Biology: ToxCast data
- Biology: ToxCast data, ATG
- Biology: ToxCast data, BSK
- Biology: ToxCast data, NVS

Graph Type: all-paths  
Filter By: ToxRef data  
Update

**Successive neighbors are not shown**

**NEW Chemical Fingerprint AIM**

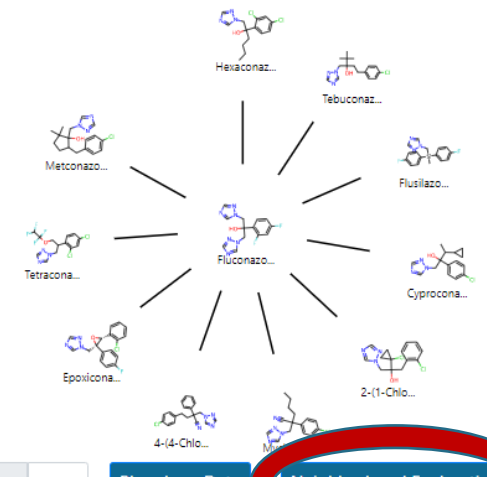
# Analogue Identification and Evaluation

## Step 2: Analogue Identification and Evaluation

EPA United States Environmental Protection Agency Generalized Read-Across (GenRA) More Info Ketcher Fluconazole DTXSID3020627

Step Two: Analog Identification and Evaluation

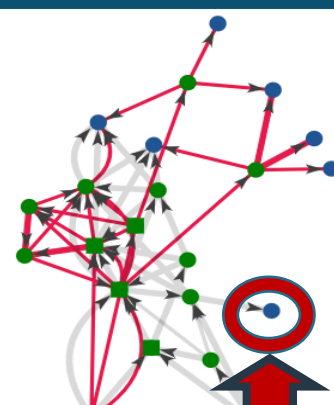
Neighbors by: Chem: Morgan Fgrpts Filter by: ToxRef data



# of Analogs 10 Physchem Data Neighborhood Exploration Next

- Green markers indicate all neighbors shown
- Blue markers indicate neighbors not shown

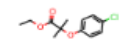
Neighborhood Exploration Graph Click Here



**Click on blue neighbor(s)**

**1. Expand  
2. Focus Chemical  
3. GenRA**

Chemical Information



Name: Clofibrate  
Mol. Weight: 242.7  
DTXCID: DTXCID70336  
DTXSID: DTXSID3020336

Expand Focus Chemical GenRA

Chem: Morgan Fgrpts  
Chem: Torsion  
Chem: ToxPrints  
Chem: AIM

Update

# Data Gap Analysis

## Step 3: Data Gap Analysis & Generate Data Matrix

EPA United States Environmental Protection Agency Generalized Read-Across (GenRA) More Ketcher Fluconazole DTXSID3020627

Step Two: Analog Identification and Evaluation

Neighbors by: Chem: Morgan Fgrprts Filter by: ToxRef data

# of Analogs 10 Physchem Data Neighborhood Exploration Next

# Data Gap Analysis

## Step 3: Data Gap Analysis & Generate Data Matrix

EPA United States Environmental Protection Agency Generalized Read-Across (GenRA) [More Info](#) [Ketcher](#) Fluconazole DTXSID3020627

### Step Three: Data Gap Analysis & Generate Data Matrix

**Neighbors by:** Chem: Morgan Fgrpts

# of Analogs: 10

[Physchem Data](#) [Neighborhood Exploration](#)

**Filter by:** ToxRef data

**Summary Data Gap Analysis**

Chemical	bio_toxt	chm_ct	chm_htr	chm_mrgn	tox_txf
Fluconazole	340	15	43	45	0
Hexaconazole	449	18	36	55	185
Tebuconazole	412	19	32	49	85
Flusilazole	458	9	34	39	179
Cyproconazole	442	16	39	53	225
2-(1-Chlorocyclopropyl)-1-(2-chlorophenyl)-3-(1H-...	210	18	41	54	87
Myclobutanil	457	15	34	53	198
4-(4-Chlorophenyl)-2-phenyl-2-(1H-1,2,4-triazol-1-yl)-...	435	17	41	59	194
Epoxiconazole	192	11	50	60	40
Tetraconazole	450	20	39	59	186
Metconazole	250	15	46	58	41

Rows: 11 Total Rows: 11

**Group:** ToxRef **By:** Tox Fingerprint  **Pagination** [Generate Data Matrix](#)

Assay endpoint	Fluconazole	Hexaconazole	Tebuconazole	Flusilazole	Cyproconazole	2-(1-Chlorocyclopropyl)-1-(2-chlorophenyl)-3-(1H-...	Myclobutanil	4-(4-Chlorophenyl)-2-phenyl-2-(1H-1,2,4-triazol-1-yl)-...	Epoxiconazole	Tetraconazole	Metconazole
CHR:adrenal gla...											
CHR:alanine ami...											
CHR:albumin											
CHR:alkaline ph...											
CHR:aminopyrin...											
CHR:anisocytosis											
CHR:appearanc...											
CHR:blood clott...											
CHR:blood vessel											

Rows: 353 Total Rows: 353

1 to 9 of 353 Page 1 of 40

# Data Gap Analysis

## Step 3: Data Gap Analysis & Generate Data Matrix

EPA United States Environmental Protection Agency Generalized Read-Across (GenRA) [More](#) [Ketcher](#) Fluconazole DTXSID3020627

Step Three: Data Gap Analysis & Generate Data Matrix

Neighbors by: Chem: Morgan Fgrpts Filter by: ToxRef data

Summary Data Gap Analysis

Group: ToxRef By: Tox Fingerprint  Pagination [Generate Data Matrix](#)

Chemical	bio_txt	chm_ct	chm_ht	chm_mrgn	tox_txt
Fluconazole	340	15	43	45	0
Hexaconazole	449	18	36	55	185
Tebuconazole	412	19	32	49	85
Flusilazole	458	9	34	39	179
Cyproconazole	442	16	39	53	225
2-(1-Chlorocyclopropyl)-1-(2-chlorophenyl)-3-(1H-1,2,4-triazol-1-yl)-1H-1,2,4-triazole	210	18	41	54	87
Myclobutanil	457	15	34	53	198
4-(4-Chlorophenyl)-2-phenyl-2-(1H-1,2,4-triazol-1-yl)-1H-1,2,4-triazole	435	17	41	59	194
Epoxiconazole	192	11	50	60	40
Tetraconazole	450	20	39	59	186
Metconazole	250	15	46	58	41

Rows: 11 Total Rows: 11

Assay endpoint	Fluconazole	Hexaconazole	Tebuconazole	Flusilazole	Cyproconazole	2-(1-Chlorocyclopropyl)-1-(2-chlorophenyl)-3-(1H-1,2,4-triazol-1-yl)-1H-1,2,4-triazole	Myclobutanil	4-(4-Chlorophenyl)-2-phenyl-2-(1H-1,2,4-triazol-1-yl)-1H-1,2,4-triazole	Epoxiconazole	Tetraconazole	Metconazole
CHR:adrenal gla...											
CHR:alanine ami...											
CHR:albumin											
CHR:alkaline ph...											
CHR:aminopyrin...											
CHR:anisocytosis											
CHR:appearanc...											
CHR:blood clotti...											
CHR:blood vessel											

Rows: 353 Total Rows: 353

1 to 9 of 353 Page 1 of 40

- How data poor is my target and what data exists for the source analogues identified?
- Do the data for the source analogues address the data gaps of interest for the target chemical?

# Data Gap Analysis

## Step 3: Data Gap Analysis & Generate Data Matrix

EPA United States Environmental Protection Agency Generalized Read-Across (GenRA) More Ketcher Fluconazole DTXSID3020627

Step Three: Data Gap Analysis & Generate Data Matrix

Neighbors by: Chem: Morgan Fgrpts Filter by: ToxRef data

Summary Data Gap Analysis Group: ToxRef By: Tox Fingerprint  Pagination Generate Data Matrix

Chemical	bio_text	chm_ct	chm_htr	chm_mrgn	tox_txf
Fluconazole	340	15	43	45	0
Hexaconazole	449	18	36	55	185
Tebuconazole	412	19	32	49	85
Flusilazole	458	9	34	39	179
Cyproconazole	442	16	39	53	225
2-(1-Chlorocyclopropyl)-1-(2-chlorophenyl)-3-(1H-...	210	18	41	54	87
Myclobutanil	457	15	34	53	198
4-(4-Chlorophenyl)-2-phenyl-2-(1H-1,2,4-triazol-1-...	435	17	41	59	194
Epoxiconazole	192	11	50	60	40
Tetraconazole	450	20	39	59	186
Metconazole	250	15	46	58	41

Rows: 11 Total Rows: 11

	Fluconazole	Hexaconazole	Tebuconazole	Flusilazole	Cyproconazole	2-(1-Chlorocyclopropyl)-1-(2-chlorophenyl)-3-(1H-...	Myclobutanil	4-(4-Chlorophenyl)-2-phenyl-2-(1H-1,2,4-triazol-1-...	Epoxiconazole	Tetraconazole	Metconazole
CHR:adrenal gla...											
CHR:alanine ami...											
CHR:albumin											
CHR:alkaline ph...											
CHR:aminopyrin...											
CHR:anisocytosis											
CHR:appearanc...											
CHR:blood clotti...											
CHR:blood vessel											

Rows: 353 Total Rows: 353

1 to 9 of 353 Page 1 of 40

- Summary of the data availability associated with the target chemical and its identified source analogues.
- The quantity of data associated with a given chemical is indicated by the number of records and the darkness of the associated box.



# Data Gap Analysis

## Step 3: Data Gap Analysis & Generate Data Matrix

The screenshot shows the EPA GenRA interface for Step 3: Data Gap Analysis & Generate Data Matrix. The interface includes a navigation bar with the EPA logo and the text "Generalized Read-Across (GenRA)". A search bar contains "Fluconazole DTXSID3020627". Below the navigation bar, there are four yellow arrows pointing right, with the text "Step Three: Data Gap Analysis & Generate Data Matrix" centered below them.

The main interface is divided into several sections:

- Neighbors by:** Chem: Morgan Fgrpts
- Filter by:** ToxRef data
- Summary Data Gap Analysis:** This section shows a central chemical structure of Fluconazole surrounded by other chemical structures. Below this, there are buttons for "# of Analogs" (10), "Physchem Data", and "Neighborhood Exploration".
- Column Configuration:** A red box highlights a menu with options: "Pin Column", "Autosize This Column", "Autosize All Columns", and "Reset Columns". Below this menu is a list of columns with checkboxes:
  - Chemical
  - bio\_bxt
  - bio\_bxt\_ATG
  - bio\_bxt\_BSK
  - bio\_bxt\_NVS
  - chm\_aim
  - chm\_ct
  - chm\_htr
  - chm\_mrgn
  - tox\_trf
- Data Matrix:** A table showing assay endpoints for various chemicals. The columns are: Fluconazole, Hexaconazole, Tebuconazole, Fusilazole, Cyproconazole, 2-(1-Chlorocyclopropyl)-1H-imidazole, Myclobutanol, 4-(4-Chlorophenyl)-2-methyl-1H-imidazole, Epoxiconazole, Tetraconazole, and Metconazole. The rows are assay endpoints such as "CHR:adrenal gla...", "CHR:alanine ami...", "CHR:albumin", "CHR:alkaline ph...", "CHR:aminopyrin...", "CHR:anisocytosis", "CHR:appearanc...", "CHR:blood clotti...", and "CHR:blood vessel".
- Group:** ToxRef
- By:** Tox Fingerprint
- Pagination:** 1 to 9 of 353, Page 1 of 40

- To see more data streams click the ☰ icon at the top of the Chemicals column, then the ≡ icon.
- This panel starts to provide a perspective to inform data gaps.



# Data Gap Analysis

## Step 3: Data Gap Analysis & Generate Data Matrix

**Neighbors by:** Chem: Morgan Fgrprts

**Filter by:**

**Group:** ToxRef **By:** Tox Fingerprint **Pagination** **Generate Data Matrix**

**Chemicals:** Hexaconaz..., Metconazo..., Tetracona..., Epoxicona..., 4-(4-Chlo..., Myclobuta..., Fluconazo..., Flusilazo..., Cyprocona...

**ToxCast Fingerprints Table:**

Assay endpoint	Fluconazole III	Hexaconazole	Tebuconazole	Fusilazole	Cyproconazole	2-(1-Chloro...	Myclobutani...	4-(4-Chlorop...	Epoxiconazole	Tetraconazole	Metconazole
CHR:adrenal gla...											
CHR:alanine ami...											
CHR:albumin											
CHR:blood vessel											

**Summary:** Rows: 11, Total Rows: 11

**Matrix Summary:** Rows: 353, Total Rows: 353

**Page:** 1 to 9 of 353 | Page 1 of 40

**Callout:** Hover over the fingerprint(s) for more information

**Tooltip:** ToxCast fingerprints are a non-directional fingerprint representation of assay hit calls represented by the assay component name of the associated platform. Hit calls are taken from Level 5 of the associated invitrodb. This fingerprint collection comprises both ToxCast and Tox21 assay outcomes.

- To see more data streams click the ☰ icon at the top of the Chemicals column, then the ||| icon.
- This panel starts to provide a perspective to inform data gaps.

# Data Gap Analysis

## Step 3: Data Gap Analysis & Generate Data Matrix

EPA United States Environmental Protection Agency Generalized Read-Across (GenRA) More Ketcher Fluconazole DTXSID3020627

**Step Three: Data Gap Analysis & Generate Data Matrix**

Neighbors by: Chem: Morgan Fgrprts Filter by: ToxRef data Summary Data Gap Analysis Group: ToxRef By: Tox Fingerprint Generate Data Matrix

Chemical	bio_toxt	dmn_ct	dmn_hlt	dmn_mrgn	tox_db
Fluconazole	340	15	43		
Tebuconazole	449	18	36		
Flusilazole	412				
Cyproconazole	458	9	34		
2-(1-Chloro-4-(4-chlorophenyl)-1H-1,2,4-triazol-5-yl)ethanol	442	16	39		
Epoxiconazole	210	18	41		
Tetraconazole	152	11	30		
Metconazole	450	20	39		
Metconazole	250	15	46		

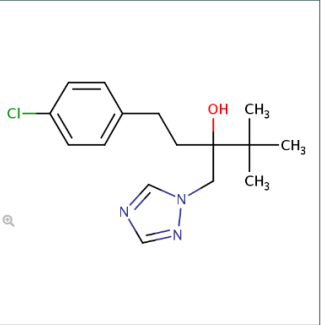
**OR Click on the neighbor(s) on the radial plot**

**Click on the neighbor(s)**

CompTox Chemicals Dashboard v2.2

**Tebuconazole**  
107534-96-3 | DTXSID9032113  
Searched by DTXSID9032113.

**Chemical Details**



Wikipedia  
Tebuconazole is a triazole fungicide used agriculturally to treat plant pathogenic fungi.  
[Read more](#)

Quality Control Notes

Intrinsic Properties

- Molecular Formula: C<sub>16</sub>H<sub>22</sub>ClN<sub>3</sub>O MOL FILE Q FIND ALL CHEMICALS
- Average Mass: 307.82 g/mol ISOTOPE MASS DISTRIBUTION
- Monoisotopic Mass: 307.14514 g/mol

Structural Identifiers

# Data Gap Analysis

## Step 3: Data Gap Analysis & Generate Data Matrix

EPA United States Environmental Protection Agency Generalized Read-Across (GenRA) [More](#) [Ketcher](#) Fluconazole DTXSID3020627

Step Three: Data Gap Analysis & Generate Data Matrix

Neighbors by: Chem: Morgan Fgrpts Filter by: ToxRef data

Summary Data Gap Analysis Group: ToxRef By: Tox Fingerprint  Pagination [Generate Data Matrix](#)

Chemical	bio_text	dmn_ct	dmn_htr	dmn_mrgn	tox_text
Fluconazole	340	15	43	45	0
Hexaconazole	449	18	36	55	185
Tebuconazole	412	19	32	49	85
Flusilazole	458	9	34	39	179
Cyproconazole	442	16	39	53	225
2-(1-Chlorocyclopropyl)-1-(2-chlorophenyl)-3-(1H-...	210	18	41	54	87
Myclobutanil	457	15	34	53	198
4-(4-Chlorophenyl)-2-phenyl-2-(1H-1,2,4-triazol-1-...	435	17	41	59	194
Epoxiconazole	192	11	50	60	40
Tetraconazole	450	20	39	59	186
Metconazole	250	15	46	58	41

Rows: 11 Total Rows: 11

Assay endpoint	Fluconazole	Hexaconazole	Tebuconazole	Flusilazole	Cyproconazole	2-(1-Chlorocyclopropyl)-1-(2-chlorophenyl)-3-(1H-...	Myclobutanil	4-(4-Chlorophenyl)-2-phenyl-2-(1H-1,2,4-triazol-1-...	Epoxiconazole	Tetraconazole	Metconazole
CHR:adrenal gla...											
CHR:alanine ami...											
CHR:albumin											
CHR:alkaline ph...											
CHR:aminopyrin...											
CHR:anisocytosis											
CHR:appearanc...											
CHR:blood clotti...											
CHR:blood vessel											

Rows: 353 Total Rows: 353

1 to 9 of 353 Page 1 of 40

- The toxicity information by default is an expanded grid view of which types of toxicity effect information are available across the analogues.

# Data Gap Analysis

## Step 3: Data Gap Analysis & Generate Data Matrix

EPA United States Environmental Protection Agency Generalized Read-Across (GenRA) [More](#) [Ketcher](#) Fluconazole DTXSID3020627

Step Three: Data Gap Analysis & Generate Data Matrix

Neighbors by: Chem: Morgan Fgrpts Filter by: ToxRef data

Summary Data Gap Analysis

Group: ToxRef ToxRef ToxCast

Chemical	bio_text	chm_ct	chm_htr	chm_mrgn	tox_text
Fluconazole	340	15	43	45	0
Hexaconazole	449	18	36	55	185
Tebuconazole	412	19	32	49	85
Flusilazole	458	9	34	39	179
Cyproconazole	442	16	39	53	225
2-(1-Chlorocyclopropyl)-1-(2-chlorophenyl)-3-(1H-...	210	18	41	54	87
Myclobutanil	457	15	34	53	198
4-(4-Chlorophenyl)-2-phenyl-2-(1H-1,2,4-triazol-1-yl)-...	435	17	41	59	194
Epoxiconazole	192	11	50	60	40
Tetraconazole	450	20	39	59	186
Metconazole	250	15	46	58	41

Rows: 11 Total Rows: 11

Assay endpoint: CHR:adrenal gla... CHR:alanine ami... CHR:albumin CHR:alkaline ph... CHR:aminopyrin... CHR:anisocytosis CHR:appearanc... CHR:blood clotti... CHR:blood vessel

Rows: 353 Total Rows: 353

1 to 9 of 353 Page 1 of 40

- Default binary representations of toxicity data arising from ToxRefDB
- Consistent with the filter by option in the Panel 1

# Data Gap Analysis

## Step 3: Data Gap Analysis & Generate Data Matrix

EPA United States Environmental Protection Agency Generalized Read-Across (GenRA) More Ketcher Fluconazole DTXSID3020627

Step Three: Data Gap Analysis & Generate Data Matrix

Neighbors by: Chem: Morgan Fgrpts Filter by: ToxRef data

# of Analogs: 10 Physchem Data Neighborhood Exploration

Summary Data Gap Analysis

Chemical	bio_txt	dmn_ct	dmn_htr	dmn_mrgn	tox_txt
Fluconazole	340	15	43	45	0
Hexaconazole	449	18	36	55	185
Tebuconazole	412	19	32	49	85
Flusilazole	458	9	34	39	179
Cyproconazole	442	16	39	53	225
2-(1-Chlorocyclopropyl)-1-(2-chlorophenyl)-3-(1H-...	210	18	41	54	87
Myclobutanil	457	15	34	53	198
4-(4-Chlorophenyl)-2-phenyl-2-(1H-1,2,4-triazol-1-...	435	17	41	59	194
Epoxiconazole	192	11	50	60	40
Tetraconazole	450	20	39	59	186
Metconazole	250	15	46	58	41

Rows: 11 Total Rows: 11

Group: ToxRef By: Tox Fingerprint

Tox Fingerprint  
Tox Fingerprint  
Tox Fingerprint Dosage

Assay endpoint	Fluconazole	Hexaconazole	Tebuconazole			
CHR:adrenal gla...						
CHR:alanine ami...						
CHR:albumin						
CHR:alkaline ph...						
CHR:aminopyrin...						
CHR:anisocytosis						
CHR:appearanc...						
CHR:blood clott...						
CHR:blood vessel						

Rows: 353 Total Rows: 353

**Total Rows: 353**

- Tox fingerprint reflects the toxicity effects within each study type.
- Over 300 different study type-toxicity combinations

# Data Gap Analysis

## Step 3: Data Gap Analysis & Generate Data Matrix

The screenshot displays the EPA GenRA interface for Step 3: Data Gap Analysis & Generate Data Matrix. The search bar contains 'Fluconazole DTXSID3020627'. The 'Neighbors by' section shows 'Chem: Morgan Fgrpts' and 'Filter by: ToxRef data'. A central diagram shows Fluconazole at the center with various chemical structures around it. The 'Summary Data Gap Analysis' table shows the following data:

Chemical	bio_txt	dmn_ct	dmn_hlt	dmn_mrgn	tox_txt
Fluconazole	340	15	43	45	0
Hexaconazole	449	18	36	55	185
Tebuconazole	412	19	32	49	85
Flusilazole	458	9	34	39	179
Cyproconazole	442	16	39	53	225
2-(1-Chlorocyclopropyl)-1-(2-chlorophenyl)-3-(1H-...	210	18	41	54	87
Myclobutanil	457	15	34	53	198
4-(4-Chlorophenyl)-2-phenyl-2-(1H-1,2,4-triazol-1-...	435	17	41	59	194
Epoxiconazole	192	11	50	60	40
Tetraconazole	450	20	39	59	186
Metconazole	250	15	46	58	41

The 'Data Matrix' table shows toxicity effects for various assay endpoints across different chemicals. A 'Pagination' callout points to the 'Rows: 11' and 'Total Rows: 11' information at the bottom of the matrix.

- Tox fingerprint reflects the toxicity effects within each study type.
- Over 300 different study type-toxicity combinations
- Pagination option to scroll through the toxicity effects



# Analogue Evaluation

EPA United States Environmental Protection Agency Generalized Read-Across (GenRA) More Info Ketcher Fluconazole DTXSID3020627

Step Three: Data Gap Analysis & Generate Data Matrix

Neighbors by: Chem: Morgan Fgrpts Filter by: ToxRef data

# of Analogs 10 Physchem Data Neighborhood Exploration

Summary Data Gap Analysis

Chemical	tox_txf	chl_mrgn	chl_htr	chl_ct	bio_txf
Fluconazole	340	15	43	45	0
Hexaconazole	449	18	36	55	185
Tebuconazole	412	19	32	49	85
Flusilazole	458	9	34	39	179
Cyproconazole	442	16	39	53	225
2-(1-Chlorocyclopropyl)-1-(2-chlorophenyl)-3-(1H-1,2,4-triazol-1-yl)ethyl-1H-imidazole	210	18	41	54	87
Myclobutanil	457	15	34	53	198
4-(4-Chlorophenyl)-2-phenyl-2-(1H-1,2,4-triazol-1-yl)ethyl-1H-imidazole	435	17	41	59	194
Epoxiconazole	192	11	50	60	40
Tetraconazole	450	20	39	59	186
Metconazole	250	15	46	58	41

Group: ToxRef By: Tox Fingerprint Pagination Generate Data Matrix

Assay endpoint	Fluconazole	Hexaconazole	Tebuconazole	Flusilazole	Cyproconazole	2-(1-Chlorocyclopropyl)-1-(2-chlorophenyl)-3-(1H-1,2,4-triazol-1-yl)ethyl-1H-imidazole	Myclobutanil	4-(4-Chlorophenyl)-2-phenyl-2-(1H-1,2,4-triazol-1-yl)ethyl-1H-imidazole	Epoxiconazole	Tetraconazole	Metconazole
CHR:adrenal gla...											
CHR:alanine ami...											
CHR:albumin											
CHR:alkaline ph...											
CHR:aminopyrin...											
CHR:anisocytosis											
CHR:appearanc...											
CHR:blood clotti...											
CHR:blood vessel											

Rows: 11 Total Rows: 11 Rows: 353 Total Rows: 353 1 to 9 of 353 Page 1 of 40

- Evaluating the suitability and relevance of the source analogues identified
- Data gap analysis across the source analogues and the target
- Little data for the source analogues or they fail to address the data gaps of interest
  - Change the number of neighbors or select a different similarity context.



# Analogue Evaluation

EPA United States Environmental Protection Agency Generalized Read-Across (GenRA) More Info Ketcher Fluconazole DTXSID3020627

Step Three: Data Gap Analysis & Generate Data Matrix

Neighbors by: Chem: Morgan Fgrpts Filter by: ToxRef data

Summary Data Gap Analysis Group: ToxRef By: Tox Fingerprint Pagination Generate Data Matrix

Chemical	toxtxt	chl_mrgn	chl_htr	chl_ct	bio_txt
Fluconazole	340	15	43	45	0
Hexaconazole	449	18	36	55	185
Tebuconazole	412	19	32	49	85
Flusilazole	458	9	34	39	179
Cyproconazole	442	16	39	53	225
2-(1-Chlorocyclopropyl)-1-(2-chlorophenyl)-3-(1H-...	210	18	41	54	87
Myclobutanil	457	15	34	53	198
4-(4-Chlorophenyl)-2-phenyl-2-(1H-1,2,4-triazol-1-yl)-...	435	17	41	59	194
Epoxiconazole	192	11	50	60	40
Tetraconazole	450	20	39	59	186
Metconazole	250	15	46	58	41

Assay endpoint	Fluconazole	Hexaconazole	Tebuconazole	Flusilazole	Cyproconazole	2-(1-Chlorocyclopropyl)-1-(2-chlorophenyl)-3-(1H-...	Myclobutanil	4-(4-Chlorophenyl)-2-phenyl-2-(1H-1,2,4-triazol-1-yl)-...	Epoxiconazole	Tetraconazole	Metconazole
CHR:adrenal gla...											
CHR:alanine ami...											
CHR:albumin											
CHR:alkaline ph...											
CHR:aminopyrin...											
CHR:anisocytosis											
CHR:appearanc...											
CHR:blood clotti...											
CHR:blood vessel											

# of Analogs 10 Physchem Data Neighborhood Exploration

Rows: 11 Total Rows: 11 Rows: 353 Total Rows: 353 1 to 9 of 353 Page 1 of 40

- Both Panels 2 and 3 provide a context of available data for the source analogues
  - the quantity of data and its type (Panel 2)
  - across study type on the basis of the toxicity effects (Panel 3)

# Analogue Evaluation

**Step 4: Run GenRA Prediction**

**Step Three: Data Gap Analysis & Generation**

Neighbors by: Chem: Morgan Fgrpts | Filter by: ToxRef data

**Summary Data Gap Analysis**

Chemical	bio_txt	dm_ct	dm_ht	dm_mg	tox_txt
Fluconazole	340	15	43	45	0
Hexaconazole	449	18	36	55	185
Tebuconazole	412	19	32	49	85
Flusilazole	458	9	34	39	179
Cyproconazole	442	16	39	53	225
2-(1-Chlorocyclopropyl)-1-(2-chlorophenyl)-3-(1H-...	210	18	41	54	87
Myclobutanil	457	15	34	53	198
4-(4-Chlorophenyl)-2-phenyl-2-(1H-1,2,4-triazol-1-yl)-...	435	17	41	59	194
Epoxiconazole	192	11	50	60	40
Tetraconazole	450	20	39	59	186
Metconazole	250	15	46	58	41

Rows: 11 | Total Rows: 11

**Generate Data Matrix**

Endpoint	Fluconazole	Hexaconazole	Tebuconazole	Flusilazole	Cyproconazole	2-(1-Chlorocyclopropyl)-1-(2-chlorophenyl)-3-(1H-...	Myclobutanil	4-(4-Chlorophenyl)-2-phenyl-2-(1H-1,2,4-triazol-1-yl)-...	Epoxiconazole	Tetraconazole	Metconazole
CHR:adrenal gla...											
CHR:alanine ami...											
CHR:albumin											
CHR:alkaline ph...											
CHR:aminopyrin...											
CHR:anisocytosis											
CHR:appearanc...											
CHR:blood clotti...											
CHR:blood vessel											

Rows: 353 | Total Rows: 353

- More detailed evaluation of the source analogues
  - Click the Generate Data Matrix button
  - Evaluate source analogues' concordance and consistency within and across the study types

# Data Gap Analysis/Data Matrix View

**Step 4: Run GenRA Prediction**

EPA United States Environmental Protection Agency Generalized Read-Across (GenRA) More Info Fluconazole DTXSID3020627

**Step Four: Run GenRA Prediction**

Neighbors by: Chem: Morgan Fgrpts Filter by: ToxRef data

**Summary Data Gap Analysis**

Group: ToxRef By: Tox Fingerprint

# of Analogs: 10 Physchem Data Neighborhood Exploration

Chemical	bio_tact	dmn_ct	dmn_hlt	dmn_mrgn	tox_txt
Fluconazole	340	15	43	45	0
Hexaconazole	449	18	36	55	185
Tebuconazole	412	19	32	49	85
Flusilazole	458	9	34	39	179
Cyproconazole	442	16	39	53	225
2-(1-(1-Chlorocyclopropyl)-1-(2-chlorophenyl)-3-(1H-1,2,4-triazol-1-yl)methyl)-4-(4-chlorophenyl)-2-phenyl-2-(1H-1,2,4-triazol-1-yl)methyl-1H-1,2,4-triazole	210	18	41	54	87
Myclobutanil	457	15	34	53	198
4-(4-(4-Chlorophenyl)-2-phenyl-2-(1H-1,2,4-triazol-1-yl)methyl)-1H-1,2,4-triazole	435	17	41	59	194
Epoxiconazole	192	11	50	60	40
Tetraconazole	450	20	39	59	186
Metconazole	250	15	46	58	41

Assay endpoint	Fluconazole	Hexaconazole	Tebuconazole	Flusilazole	Cyproconazole	2-(1-(1-Chlorocyclopropyl)-1-(2-chlorophenyl)-3-(1H-1,2,4-triazol-1-yl)methyl)-4-(4-chlorophenyl)-2-phenyl-2-(1H-1,2,4-triazol-1-yl)methyl-1H-1,2,4-triazole	Myclobutanil	4-(4-(4-Chlorophenyl)-2-phenyl-2-(1H-1,2,4-triazol-1-yl)methyl)-1H-1,2,4-triazole	Epoxiconazole	Tetraconazole	Metconazole
CHR:adrenal gland											
CHR:alanine amin...											
CHR:albumin											
CHR:alkaline phos...											
CHR:aminopyrine...											
CHR:anisocytosis											
CHR:appearance a...											
CHR:blood clotting											

Rows: 11 Total Rows: 11

Rows: 353 Total Rows: 353

1 to 9 of 353 Page 1 of 40

- This creates a DATA MATRIX VIEW in Panel 4

# Data Gap Analysis/Data Matrix View

**Step 4: Run GenRA Prediction**

**Filter by:**

- ToxRef data
- ToxRef data
- ToxCast data
- No filter (all data)

**The toxicity effects**

- Default:** ToxRefDB (*in vivo* studies)
- ToxCast data filter (*in vitro* studies)
- No filter

**Target**

**Data Matrix View**

**Similarity Index**

Assay endpoint	Fluconazole	Hexaconazole	Tebuconazole	Flusilazole	Cyproconazole	2-(1-Chloro...	Myclobutanil	4-(4-Chloro...	Epoxico...
Properties (7)									
CHR:adrenal gland									
CHR:alanine amin...									
CHR:albumin									
CHR:alkaline phos...									
CHR:aminopyrine...									

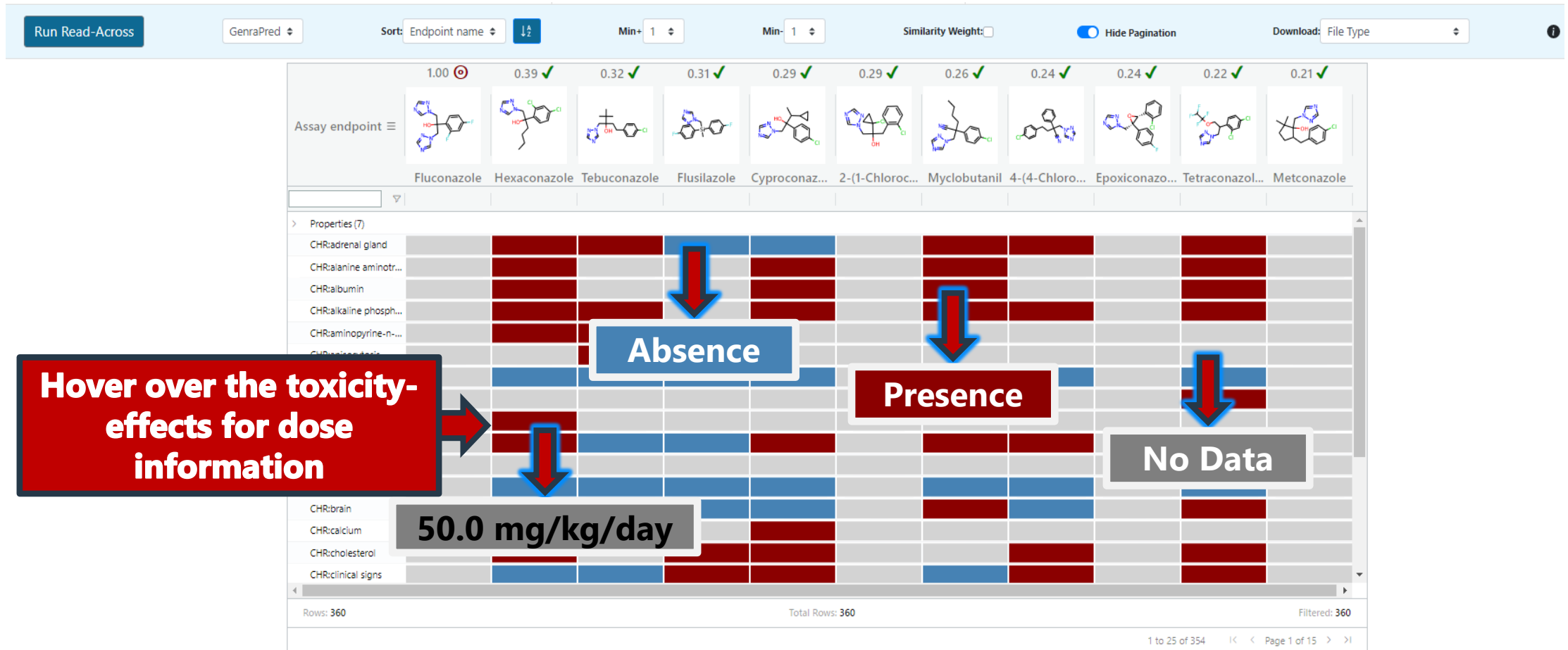
# Data Gap Analysis/Data Matrix View



**Toxicity Effects**

- The source analogues are ordered by their pairwise similarity relative to the target chemical
- The toxicity effects (*in vivo* studies from ToxRefDB) are populated across the analogues
- ToxCast data filter is for *in vitro* data and no filter is without considering any data
- Binary data / predictions and Potency data / predictions

# Data Gap Analysis/Data Matrix View



- Blue indicates 'Absence' of toxicity effects
- Red indicates 'Presence' of toxicity effects
- Grey represents 'No Data'
- Endpoints are measured in mg/kg/day=minimum dose for observed effects

# Data Gap Analysis/Data Matrix View

Run Read-Across    GenraPred    Sort: Endpoint name    Min- 1    Min- 1    Similarity Weight:    Hide Pagination    Download: File Type

Assay endpoint    1.00    0.39 ✓    0.32 ✓    0.31 ✓    0.29 ✓    0.29 ✓    0.26 ✓    0.24 ✓    0.24 ✓    0.22 ✓    0.21 ✓

Fluconazole    Hexa...    Tebuconazole    Flusilazole    Cyproconaz...    2-(1-Chloroc...    Myclobutanil    4-(4-Chloro...    Epoxiconazo...    Tetraconazol...    Metconazole

Properties (7)

Boiling (°C)	293.1	321.2	321.7	348.8	321.4	321.5	339.2	349.4	N/A	339.1	319.2
Henry's Law (at...)	7.66e-09	5.47e-10	3.1e-07	2.7e-07	5.48e-10	1.37e-07	3.18e-07	3.93e-06	N/A	2.53e-06	5.76e-08
Mass (g/mol)	306.28	314.21	307.82	315.40	291.78	312.19	288.78	336.82	329.76	372.14	319.83
Melting (°C)	176.1	111.1	102.5	54.0	107.5	122.6	152.8	125.1	N/A	6.1	111.5
Vap. press. (m...	8.72e-10	1.36e-07	1.29e-08	2.93e-07	2.61e-07	3.19e-07	1.96e-07	3.74e-08	N/A	1.35e-06	9.27e-08
Water sol. (mol...	0.0158	5.58e-05	0.000117	0.000149	0.000429	0.000114	0.000481	1.13e-06	N/A	0.000417	6.7e-05
log Kow	0.501	3.899	3.700	3.701	2.902	3.081	2.939	3.229	N/A	3.560	3.850
CHR:adrenal gland											
CHR:alanine aminotr...											
CHR:blood clotting											
CHR:blood vessel											
CHR:body weight											
CHR:bone											
CHR:bone marrow											
CHR:brain											
CHR:calcium											
CHR:cholesterol											
CHR:clinical signs											

Rows: 360    Total Rows: 360    Filtered: 360

1 to 25 of 354    Page 1 of 15

- Click on the 'Properties' arrow to evaluate (OPERA predicted) physical property and e-fate information across analogues

# Data Gap Analysis/Data Matrix View

Run Read-Across    GenraPred    Sort: Endpoint name    Min+ 1    Min- 1    Similarity Weight:    Hide Pagination    Download: File Type

Assay endpoint	1.00	0.39	0.32	0.31	0.29	0.26	0.24	0.24	0.22	0.21
Fluconazole										
Properties (7)										
CHR:adrenal gland										
CHR:alanine aminotr...										
CHR:albumin										
CHR:alkaline phosph...										
CHR:aminopyrine-n...										
CHR:anisocytosis										
CHR:appearance an...										
CHR:blood clotting										
CHR:blood vessel										
CHR:body weight										
CHR:bone										
CHR:bone marrow										
CHR:brain										
CHR:calcium										
CHR:cholesterol										
CHR:clinical signs										

Rows: 360    Total Rows: 360    Filtered: 360    1 to 25 of 354    Page 1 of 15

- Removing Analogues within the Data Matrix:
- De-select directly from analogue bar



# Data Gap Analysis/Data Matrix View

The screenshot shows the Data Matrix View interface with several callouts highlighting key features:

- Sort:** A dropdown menu showing options: Endpoint name, Endpoint name, Positive obs., Negative obs., and Observations. A red arrow points to this menu.
- Toggle between Ascending and Descending Endpoint Name:** A red box with white text pointing to the sort dropdown.
- Export to CSV or excel:** A red box with white text pointing to the 'Download: File Type' dropdown in the top right.
- Similarity Weight:** A checkbox labeled 'Similarity Weight:' in the top right, with a red arrow pointing to it.
- Assay endpoint details:** A callout box showing a detailed view of an assay endpoint, including chemical structures and a grid of red and blue cells representing data points.

- Toggle between ascending and descending endpoint name
- The data displayed can be filtered by specific effects (Sort by Observations)
- Similarity Weight checkbox modifies the size of the box to reflect the pairwise similarity metric.

# Data Gap Filling

**Generate Predictions  
Click on  
'Run Read-Across'**

Run Read-Across GenraPred GenraPy

Min+ 1 Min- 1 Similarity Weight: Hide Pagination Download: File Type

Assay endpoint	Fluconazole	Hexaconazole	Tebuconazole	Flusilazole	Cyproconaz...	2-(1-Chloroc...	Myclobutanil	4-(4-Chloro...	Epoxiconazo...	Tetraconazol...	Metconazole
Properties (7)											
CHR:adrenal gland											
CHR:alanine aminotr...											
CHR:albumin											
CHR:alkaline phosph...											
CHR:aminopyrine-n...											
CHR:anisocytosis											
CHR:appearance an...											
CHR:blood clotting											
CHR:blood vessel											
CHR:body weight											
CHR:bone											
CHR:bone marrow											
CHR:brain											
CHR:calcium											
CHR:cholesterol											
CHR:clinical signs											

Rows: 360 Total Rows: 360 Filtered: 360

1 to 25 of 354 Page 1 of 15

- This step generates GenRA predictions.
  - GenraPred: Legacy prediction engine, binary only, reports confidence
  - GenraPy: Binary and continuous prediction, no confidence reporting (Yet!)

# Data Gap Filling Predictions

**Step 5: Filter by Endpoint or Analogue**

**EPA** United States Environmental Protection Agency | Generalized Read-Across (GenRA) | [More Info](#) | [Ketcher](#) | Fluconazole DTXSID3020627

---

**Step Five: Filter by Endpoint or Analogs**

**Neighbors by:** Chem: Morgan Fgprts

# of Analogs: 10

[Physchem Data](#) | [Neighborhood Exploration](#)

**Filter by:** ToxRef data

**Summary Data Gap Analysis**

Chemical	bio_tct	dmn_ct	dmn_htr	dmn_mgn	tox_tct
Fluconazole	340	15	43	45	0
Hexaconazole	449	18	36	55	185
Tebuconazole	412	19	32	49	85
Flusilazole	458	9	34	39	179
Cyproconazole	442	16	39	53	225
2-(1-Chlorocyclopropyl)-1-(2-chlorophenyl)-3-(1H-...	210	18	41	54	87
Myclobutanil	457	15	34	53	198
4-(4-Chlorophenyl)-2-phenyl-2-(1H-1,2,4-triazol-1-...	435	17	41	59	194
Epoxiconazole	192	11	50	60	40
Tetraconazole	450	20	39	59	186
Metconazole	250	15	46	58	41

Rows: 11 | Total Rows: 11

**Group:** ToxRef | **By:** Tox Fingerprint |  **Pagination**

Assay endpoint	Fluconazole	Hexaconazole	Tebuconazole	Flusilazole	Cyproconazole	2-(1-Chloro...	Myclobutanil	4-(4-Chlorop...	Epoxiconazole	Tetraconazole	Metconazole
CHR:adrenal gl...											
CHR:alanine a...											
CHR:albumin											
CHR:alkaline p...											
CHR:aminopyri...											
CHR:anisocytosis											
CHR:appearanc...											
CHR:blood clot...											
CHR:blood ves...											

Rows: 353 | Total Rows: 353

1 to 9 of 353 | Page 1 of 40

---

[Run Read-Across](#) | **GenraPred** | **Sort:** Endpoint name | **Min+** 1 | **Min-** 1 | **Similarity Weight:** |  **Hide Pagination** | **Download:** File Type

Assay endpoint	Fluconazole	Hexaconazole	Tebuconazole	Flusilazole	Cyproconazo...	2-(1-Chloro...	Myclobutanil	4-(4-Chloro...	Epoxiconazo...	Tetraconazol...	Metconazole
	1.00	0.39	0.32	0.31	0.29	0.29	0.26	0.24	0.24	0.22	0.21
CHR:adrenal gland											
CHR:alanine aminotr...											
CHR:albumin											

# Data Gap Filling Predictions



The screenshot displays the GenraPred web application interface. At the top, there are navigation and control elements: a 'Run Read-Across' button, a 'GenraPred' dropdown, a 'Sort: Endpoint name' dropdown, a 'Priority Weight' input field, a 'Hide Pagination' toggle, and a 'Download: File Type' dropdown. The main content area shows a list of assay endpoints with their corresponding chemical structures and predicted values. The endpoints include Fluconazole (1.00), Hexaconazole (0.39), Tebuconazole (0.32), Myclobutanil (0.26), 4-(4-Chloro... (0.24), Epoxiconazo... (0.24), Tetraconazol... (0.22), and Metconazole (0.21). Below this, a heatmap displays predictions for various assay endpoints across different chemical structures. The heatmap cells are colored red (high confidence) or blue (low confidence). A red box highlights a specific prediction cell, and a red arrow points to it from a red callout box labeled 'Hover over the Predictions'. Another red box labeled 'Prediction Panel' points to the heatmap area. A tooltip is visible over a prediction cell, displaying the text: 'Pos; ACT=0.704; AUC=0; pval=0.96'. At the bottom of the interface, there are statistics: 'Rows: 360', 'Total Rows: 360', and 'Filtered: 360'. The page number 'Page 1 of 15' is also visible.

- The opacity of the predictions reflects the confidence in the prediction
- A faint colored prediction will denote lower confidence in the prediction.

# Data Gap Filling Predictions

The screenshot displays the GenraPred interface with a data table. The table has columns for Assay endpoint (with chemical structures and AUC values like 1.00, 0.39), Properties (7), and Predictions (represented by red and blue cells). A 'Run Read-Across' button is at the top left. A 'Sort: Endpoint name' dropdown is highlighted with a red box and labeled 'Sort by Endpoint'. A 'Sort: AUC, pval' dropdown is also highlighted with a red box and labeled 'Sort by AUC, pval'. A 'Download: File Type' dropdown is highlighted with a red box and labeled 'Export to CSV or excel'. A 'Prediction Panel' is highlighted with a red box on the left side of the table. The table footer shows 'Rows: 360', 'Total Rows: 360', and 'Filtered: 360'.

- Endpoints can be sorted
- Predictions can be exported by clicking on the 'Download' button

# Exporting Results

chem_id	DTXCID10627	DTXCID10627_uni	pred_class	ACT	AUC	pval	DTXCID2014653	DTXCID2014653_u	DTXCID7012113	DTXCID7012113_u	DTXCID704235	DTXCID704235_ur	DTXCID8012601	DTXCID8012601_u
role	target						analog		analog		analog		analog	
preferred name	Fluconazole						Hexaconazole		Tebuconazole		Flusilazole		Cyproconazole	
dsstox_sid	DTXSID3020627						DTXSID4034653		DTXSID9032113		DTXSID3024235		DTXSID0032601	
dsstox_cid	DTXCID10627						DTXCID2014653		DTXCID704235		DTXCID704235		DTXCID8012601	
similarity	1.00001						0.388888		0.3125		0.3125		0.289473684	
Mass g/mol	306.277						314		315.399		315.399		291.78	
Melting °C	176.083						111		54,0469		54,0469		107.543	
Boiling °C	293.102						321		348.793		348.793		321.422	
log Kow	0.501414						3.85		3.70142		3.70142		2.90198	
Vap. press. mmHg	8.71859E-10						1.35751E-09		2.9314E-07		2.9314E-07		2.60598E-07	
Water sol. mol/L	0.0158043						5.57587E-09		0.000149013		0.000149013		0.000429129	
Henry's Law atm-m3/mole	7.66143E-09						5.47037E-09		2.70159E-07		2.70159E-07		5.47602E-10	
Hydrogen Bond Donors														
Hydrogen Bond Acceptors														
chem_id	DTXCID10627	DTXCID10627	pred_class	ACT	AUC	pval								
CHR:adrenal gland	pos_effect		Pos	1	0	1								
CHR:alanine aminotransferase	pos_effect		Pos	1	0	1								
CHR:albumin	pos_effect		Pos	1	0	1								
CHR:alkaline phosphatase (al)	pos_effect		Pos	1	0	1								
CHR:aminopyrine-n-demethyl	pos_effect		Pos	1	0	1								
CHR:anisocytosis	pos_effect		Pos	1	0	1	no_data							
CHR:appearance and color	no_effect		Neg	0	0	1	no_effect							
CHR:blood clotting	pos_effect		Pos	1	0	1	no_data							
CHR:blood vessel	pos_effect		Pos	1	0	1								
CHR:body weight	pos_effect		Pos	0.687	0	0.96								
CHR:bone	pos_effect		Pos	1	0	1	no_data							
CHR:bone marrow	no_effect		Neg	0	0	1	no_effect							
CHR:brain	no_effect		Neg	0.236	0	0.945	no_effect							
CHR:calcium	pos_effect		Pos	1	0	1								
CHR:cholesterol	pos_effect		Pos	1	0	1								
CHR:clinical signs	pos_effect		Pos	0.523	0	0.955	no_effect							
CHR:cytochrome p450, nos	pos_effect		Pos	1	0	1	no_data							
CHR:tear	no_effect		Neg	0	0	1	no_effect							
CHR:epididymis	no_effect		Neg	0.269	0	0.965	no_effect							
CHR:erythrocyte (rbc) count	no_effect		Neg	0.344	0	0.9	no_effect							
CHR:esophagus	no_effect		Neg	0	0	1	no_effect							
CHR:eye	no_effect		Neg	0.159	0	0.89	no_effect							
CHR:food consumption	no_effect		Neg	0.46	0	0.975								
CHR:full gross necropsy	no_effect		Neg	0	0	1	no_effect							
CHR:gamma glutamyl transferase	pos_effect		Pos	1	0	1	no_data							
CHR:globulins	pos_effect		Pos	1	0	1	no_data							
CHR:glucose	no_effect		Neg	0.11	0	0.825	no_effect							
CHR:glutathione-s-transferase	pos_effect		Pos	1	0	1	no_data		no_data		no_data		no_data	3.2 mg/kg/day
CHR:heart	no_effect		Neg	0.11	0	0.84	no_effect		no_effect		no_effect		no_effect	
CHR:hematocrit (hct)	pos_effect		Pos	1	0	1	no_data		no_data		no_data		no_data	
CHR:hemoglobin (hgb)	pos_effect		Pos	1	0	1	no_data		no_data		no_data		no_data	

- Snapshot of a sample prediction output produced in the Excel (xlsx) export file
- The **AUC** performance measure and the **p-value** that is associated with the prediction are provided
- P values can be sorted. (**High p-values = non-significant confidence in the predictions**)
- The actual experimental data for the source analogues is also reflected.

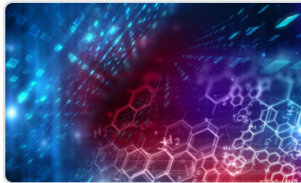
# GenRA Summary

- GenRA is an attempt to move towards an objective read-across approach where uncertainties and performance can be quantified.
- Provides opportunities for different NAM data to be incorporated.
- GenRA v1.0 established a baseline in performance. Research has continued to evaluate performance using other contexts of similarity e.g. physicochemical similarity, targeted transcriptomic similarity, metabolic similarity.
- New features coming include defined categories and/or user specific neighbourhoods, other similarity contexts and other *in vivo* toxicity data streams.

Related Topics: [Safer Chemicals Research](#)

[CONTACT US](#)

## Generalized Read-Across (GenRA)

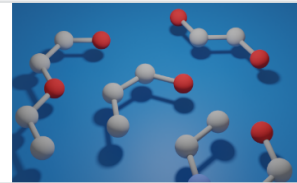


### Access GenRA

[GenRA | Web Application](#)

### About

[GenRA | About](#)



### User Manual

Learn to navigate and utilize GenRA.

[GenRA | Manual](#)

### Release Notes

Access information on the latest and previous versions of GenRA.

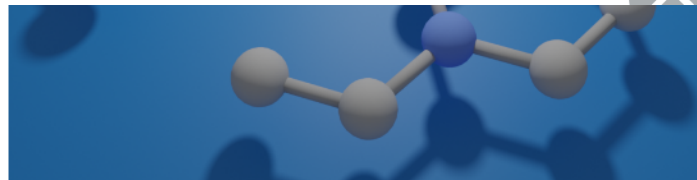
[GenRA | Release Notes](#)

### Training

Explore training materials for GenRA and other tools through the New Approach Methods (NAMs) Training Site.

[NAMs Training](#)

perspective. GenRA is an algorithmic approach to permit objective and reproducible read-across predictions of *in vivo* toxicity and *in vitro* bioactivity.



### More Resources

- [GenRA: Resource Hub](#)
- [Access GenRA](#)
- [GenRA: About](#)
- [GenRA: User Manual](#)
- [GenRA: Release Notes](#)
- [New Approach Methods Training](#)



# Q&A

# Thank You!

Contact us if you have

- GenRA questions: [genra.support@epa.gov](mailto:genra.support@epa.gov)
- Future NAMs training questions: [NAM@EPA.gov](mailto:NAM@EPA.gov)

**Save the date for our next training!**

**HTTK**

**November 8<sup>th</sup>, 2023**