GenRA Virtual Training

Breakout Group Worksheet

This worksheet was developed for the Breakout Group session of the GenRA Virtual Training, hosted by the U.S. Environmental Protection Agency's Center for Computational Toxicology and Exposure on May 23, 2023.

For more information about the GenRA tool:

- Visit the GenRA at https://comptox.epa.gov/genra/
- Contact GenRA support at genra.support@epa.gov
- Read more in
    - Publication 2016: https://doi.org/10.1016/j.yrtph.2016.05.008
    - Publication 2022: https://doi.org/10.1016/j.comtox.2022.100258
    - GenRA Manual: https://www.epa.gov/chemical-research/generalized-read-across-genra-manual

---

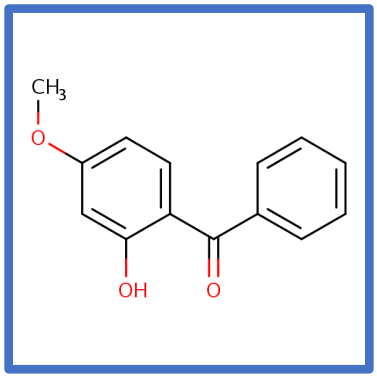
Goals

- **Hazard identification:** Use GenRA to evaluate the potential hazards of a substance found in a contaminated site. What can be determined based on the read-across workflow as related to a screening level human health hazard assessment?
- **Assessing toxicity:** Use of GenRA to infer the potential toxicity effect for a target chemical based on study-type and effect information. Based on the read-across predictions and their associated uncertainties, what might be the next steps?
    - E.g., Consumer product safety: Toxicity assessment of a given chemical used as an ingredient in a product. Based on the read-across predictions, what inferences could be made regarding product safety and what additional information might be needed?
- **Gaining an understanding of:**
    - How a read-across should be considered relative to the decision context.
    - What available Fingerprint types are in GenRA and what they represent.
    - How to navigate the Neighborhood Explorer to inform selection of analogues based on different fingerprint types. Additionally,
        - How to use the Physchem Data viewer, and
        - How to make binary toxicity effect predictions for a data-poor target chemical of interest.

Target Chemical

1. 2-HYDROXY-4-METHOXYBENZOPHENONE (HMB)

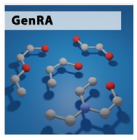
Oxybenzone or benzophenone-3 or BP-3 (trade names Milestab 9, Eusolex 4360, Escalol 567, KAHSCREEN BZ-3) is an organic compound. It is a pale-yellow solid that is readily soluble in most organic solvents. Oxybenzone belongs to the class of aromatic ketones known as benzophenones. It is widely used in plastics, toys, furniture finishes, and other products to limit UV degradation.



IUPAC Name: (2-Hydroxy-4-methoxyphenyl)(phenyl)methanone
CASRN: 131-57-7
DSSTox substance identifier (DTXSID): DTXSID3022405
Molecular Formula: $C_{14}H_{12}O_3$
SMILES: COC1=CC=C(C(=O)C2=CC=CC=C2)C(O)=C1
InChIKey: DXGLGDHPHMLXJC-UHFFFAOYSA-N

Scenario: Trace amounts of 2-hydroxy-4-methoxybenzophenone (HMB) were found in a drinking water source. The duration of exposure is unknown. After determining the concentration in drinking water, it was found to exceed the established Threshold of Toxicological Concern (TTC) values which necessitated a read-across approach.
Based on the information provided you would like to:

- Identify data gaps, and
- Evaluate potential toxicity effects for this contaminant.



Directions

1. Using the chemical information for HMB above, type the chemical name (or synonym), its Chemical Abstracts Service Registry (CASRN), and DSSTox Substance Identifier (DTXSID) below.

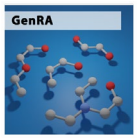
2. Access GenRA directly from https://comptox.epa.gov/genra/
3. Either use the Ketcher drawing palette or search box to search for HMB using the SMILES or enter DTXSID in the search box.
4. Identify potential source analogues for HMB using the default settings in Panel 1 which depicts the radial plot.
    a. How many source analogues are presented in the radial plot?

    b. Based on the Jaccard similarity metric, what is the most similar source analogue to the target chemical?

    c. Based on the Jaccard similarity metric, what is the least similar source analogue to the target chemical?

    d. Based on the Jaccard similarity metric, what is the most similar source analogue to the target chemical if the panel is updated using the ToxCast data filter?

    e. After this, change the Filter back to “ToxRef data”, which is the default.
    f. Evaluate the Physchem properties of each source analogue (based on the OPERA model estimates), access this information by clicking “Physchem Data”.
        i. Consider the likely physical state of the source analogues using the predicted melting and boiling points. How do they compare to the likely state of HMB? *[Note: Plot window is resizable!]*



ii. What is the predicted Log Kow of HMB?

iii. How many source analogues have Log Kow values above 2 and below 5?

g. Close the "Physchem Data" window. Now, explore the analogues by using "Neighborhood Exploration". Click on the target chemical (HMB) to zoom in. *[Reminder: The target chemical is indicated by a red circle.]*

i. Identify the top 3 source analogues based on Morgan Fingerprints.

ii. Identify the 3 source analogues based on ToxCast Fingerprints.

h. Identify the analogues on the basis of the hybrid option 50:50 Morgan to ToxCast.

5. Is the target chemical data poor? *[Hint: evaluate the Panel 2 on Step 3 of GenRA tool!]*

a. What data exists for the source analogs identified?

b. Do the data for the source analogues address the data gaps of interest for the target chemical? (Or, how feasible might a read-across be based on the quantity of data for the source analogues?)

6. Let's move on to Panel 3. Which types of toxicity effect information are available for the target chemical and across the analogues?

    a. Evaluate the binary predictions of toxicity data arising from ToxRefDB (default). What data gaps exist for the source analogues relative to the target substance, and which effects might be reasonably predicted by those analogues?

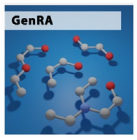
    b. Evaluate Chronic bone marrow assay end point for Benzophenone. Are there any experimental data?

    c. Are there any experimental data for source analogues for Chronic bone marrow assay end point?

7. Would Panels 2 and 3 provide any information on their potency (e.g., in mg/kg-day etc.) or hazard profile? *[Note: Click "Generate Data Matrix".]*

    a. Evaluate the suitability of the source analogues by looking at the data availability of the source analogues. Are there enough data for chronic and sub-chronic assay endpoints?

        i. If not, what are the next steps?



b. Find the dose for the analogue (Benzophenone) at which a toxicity effect for chronic body weight (CHR:liver) was observed. Note below if there are any observations exist for HMB for CHR:liver end point.

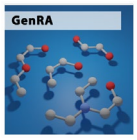
c. Consider the Physchem properties of each analogue and identify those with estimated Log Kow values within 1.5 of the estimate for the target chemical Log Kow (e.g., +1.5 and -1.5 of 3.785).
   i. De-select any analogues that are below 2 and above 5. How many analogues are now selected for Read-Across?

   ii. Once the desired source analogues are selected, click the "Run Read-Across" button to re-derive the GenRA predictions using the GenraPred engine. What is the prediction value?

d. Export and save the results as an Excel file.
   i. Does the file export the target predictions as well as all 10 analogues' data?

e. Is there any subsequent analysis after this supported by GenRA?

f. Filter column D, "pred_class", to show only positive effects.
g. Sort the pval from smallest to largest value.



Reflection

1. In what case example from your work environment would GenRA be useful?

2. What have you learned about the process and workflow used to find information in GenRA?

3. What challenges did you encounter, and how did you solve them?

WORKSHEET ANSWERS

Directions

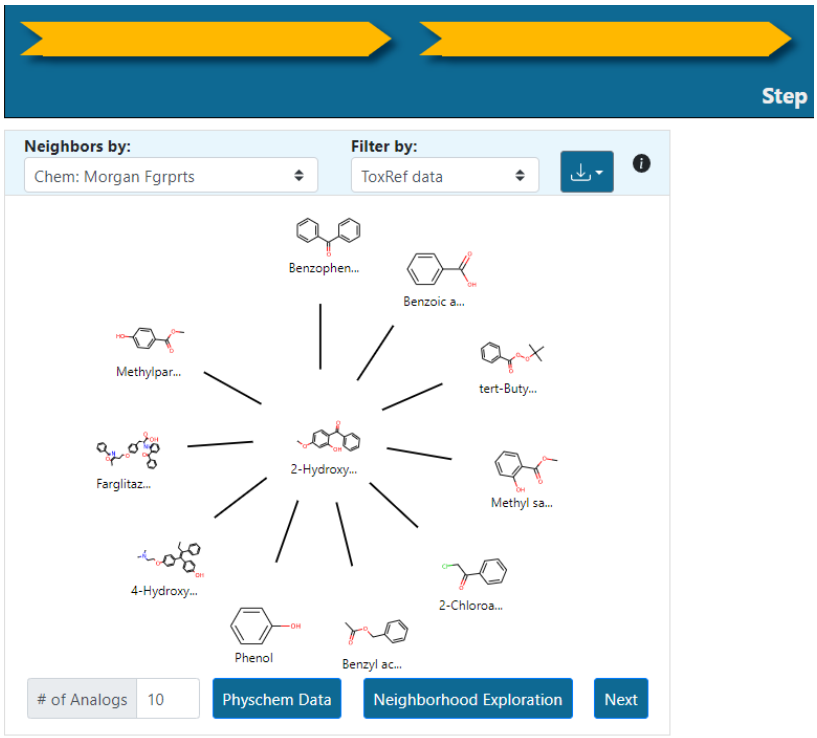
1. Using the chemical information for HMB above, type the chemical name (or synonym), its Chemical Abstracts Service Registry (CASRN) and DSSTox Substance Identifier (DTXSID) in the space below.

   2-HYDROXY-4-METHOXYBENZOPHENONE, 131-57-7, DTXSID3022405

2. Access GenRA directly from https://comptox.epa.gov/genra/
3. Either use the Ketcher drawing palette or search box to search for HMB using the SMILES or enter DTXSID in the search box.

   DTXSID3022405

4. Identify potential source analogues for HMB using the default settings in panel 1 which depicts the radial plot.



   a. How many source analogues are presented in the radial plot?

      (10, default)

   b. Based on the Jaccard similarity metric, what is the most similar source analogue to the target chemical?

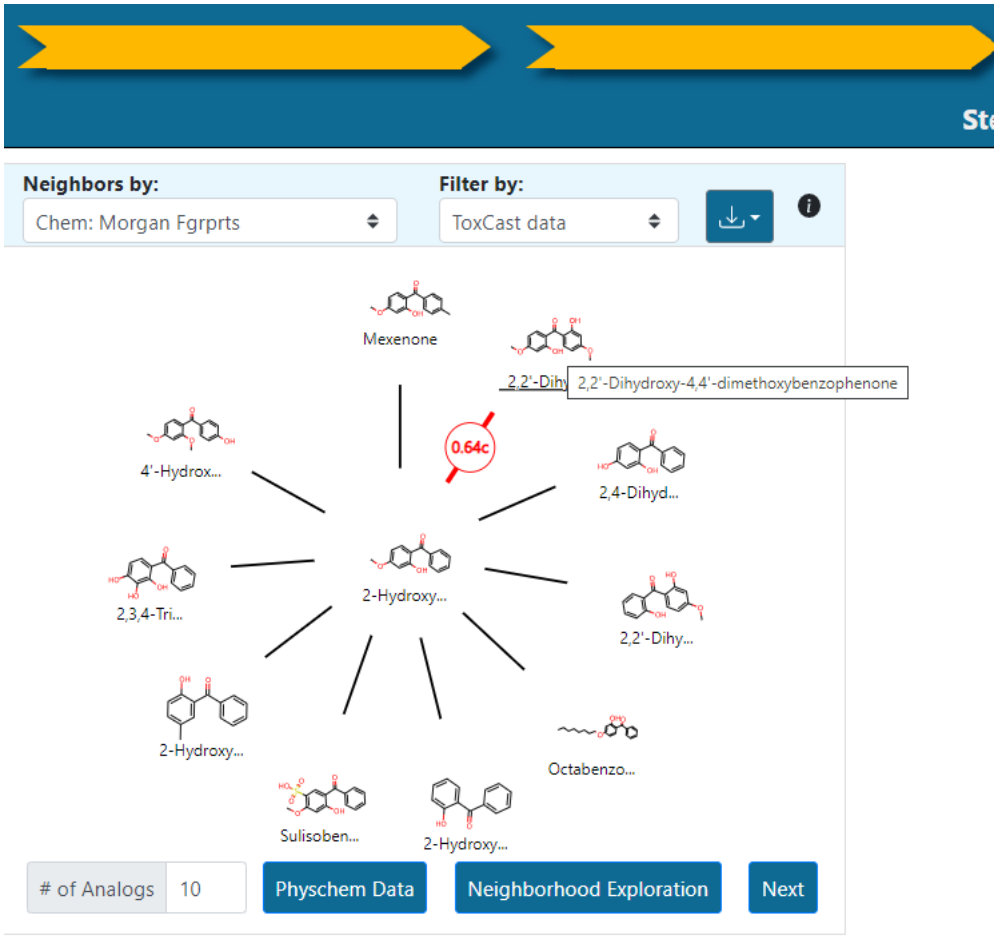
(Benzophenone, 0.38)

c. Based on the Jaccard similarity metric, what is the least similar source analogue to the target chemical?

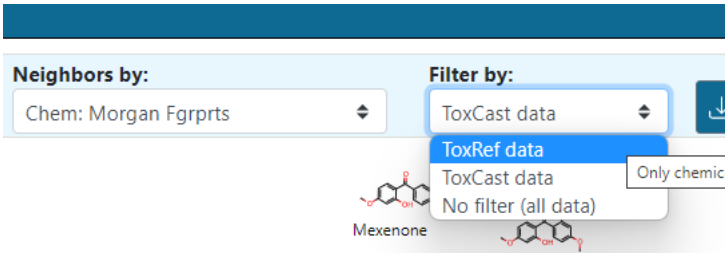
(Methyl paraben, 0.23)

d. Based on the Jaccard similarity metric, what is the most similar source analogue to the target chemical if the panel is updated using the ToxCast data filter?

(mexenone, 0.65)

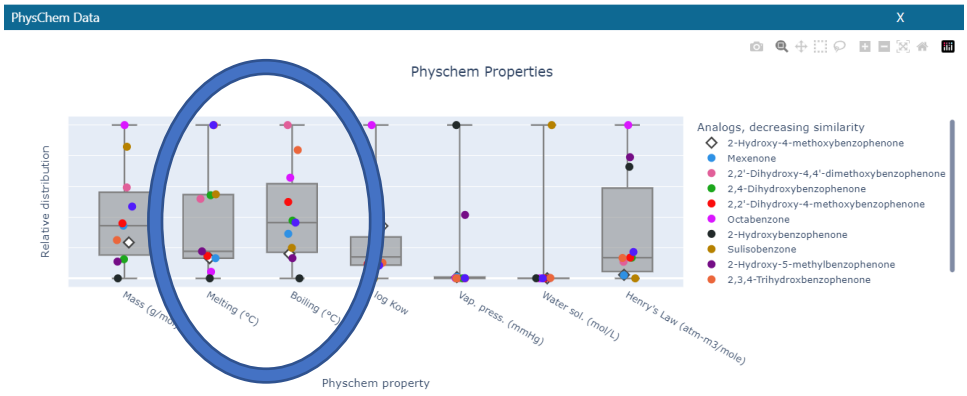


e. After this, change the Filter back to the “ToxRef data” default.



f. Evaluate the Physchem properties of each source analogue (based on the OPERA model estimates), access this information by clicking "Physchem Data".

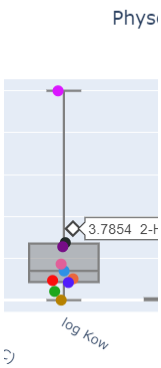
i. Consider the likely physical state of the source analogues using the predicted melting and boiling points. *[Note: Plot window is resizable!]*



HMB is known to be a solid – based on the description. This is consistent with its predicted MP (a value exceeding 25 deg C is indicative of a solid). Note that the example source analogues Benzyl acetate and Methyl salicylate have much lower MP values than our target substance. See https://www.epa.gov/sites/default/files/2015-05/documents/05-iad_discretes_june2013.pdf

ii. What is the predicted Log Kow of HMB?

(3.7854)



iii. How many source analogues have Log Kow values above 2 and below 5?

(Four (4): benzophenone, t-butyl perbenzoate, methyl salicylate, 2-chloroacetophenone)

g. Close the Physchem Data window. Now, explore the analogues by using "Neighborhood Exploration". Click on the target chemical (HMB) to zoom in. *[Reminder: The target chemical is indicated by a red circle.]*

i. Identify the top 3 source analogues based on Morgan Fingerprints.

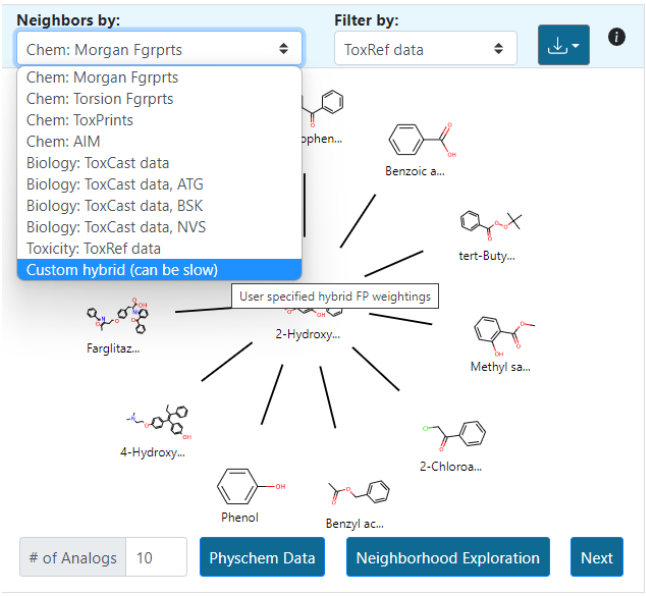
(Benzophenone, benzoic acid, and t-butyl perbenzoate)

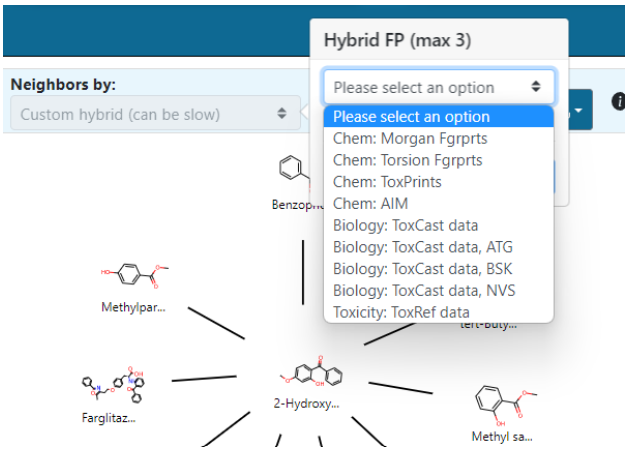
ii. Identify the 3 source analogues based on ToxCast Fingerprints.

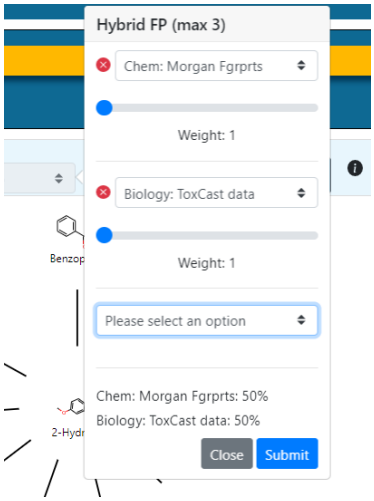
(Phenothrin, tris(methylphenyl)phosphate, 4,4'-dichlorophenylsulfone)

h. Identify the analogs on the basis of the hybrid option 50:50 Morgan to ToxCast.

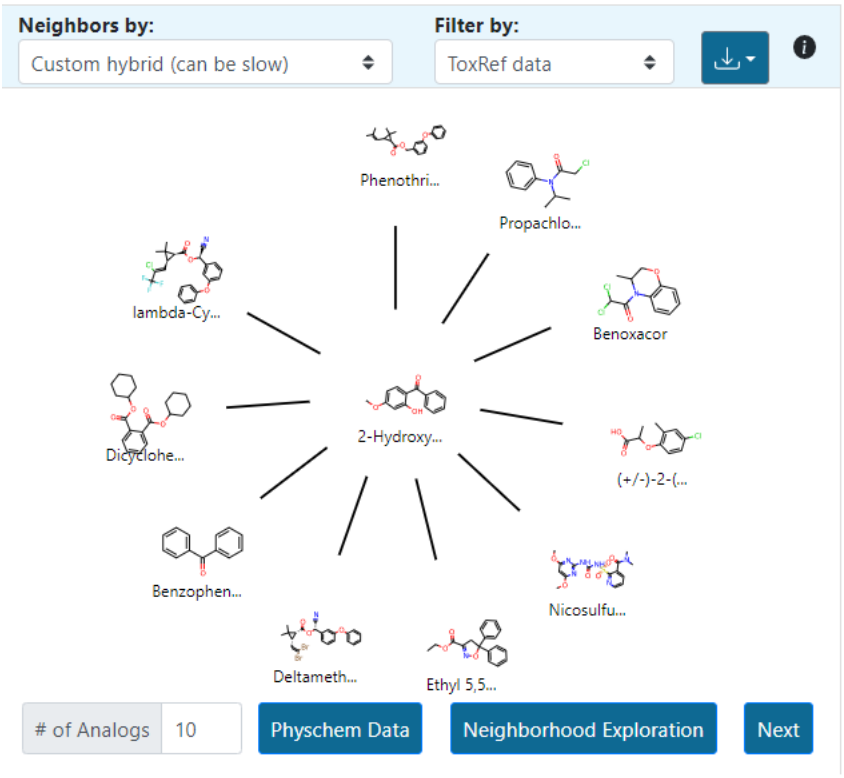
Close the neighborhood explorer and go back to panel one and select custom hybrid


Neighbors by:
Chem: Morgan Fgrprts
Filter by:
ToxRef data
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
Toxicity: ToxRef data
Custom hybrid (can be slow)
User specified hybrid FP weightings
Benzoic a...
tert-Buty...
Farglitaz...
2-Hydroxy...
Methyl sa...
4-Hydroxy...
2-Chloroa...
Phenol
Benzyl ac...
# of Analogs
10
Physchem Data
Neighborhood Exploration
Next


Hybrid FP (max 3)
Neighbors by:
Custom hybrid (can be slow)
Please select an option
Please select an option
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
Toxicity: ToxRef data
Methylpar...
Farglitaz...
2-Hydroxy...
Methyl sa...


Hybrid FP (max 3)
Chem: Morgan Fgrprts
Weight: 1
Biology: ToxCast data
Weight: 1
Please select an option
Chem: Morgan Fgrprts: 50%
Biology: ToxCast data: 50%
Close
Submit

Panel one will updated to reflect the below screen



5. Is the target chemical data poor? *[Hint: evaluate the Panel 2 on Step 3 of GenRA tool!]*

   The answer is Yes. Let's evaluate panels 2 and 3

PANEL 2

Summary Data Gap Analysis

Chemical	bio_txct	chm_ct	chm_httr	chm_mrgn	tox_txrf
2-Hydroxy-4-methoxybenzophenone	296	13	27	39	68
Phenothrin	270	15	44	63	54
Propachlor	266	12	21	33	180
Benoxacor	271	15	32	46	40
(+/-)-2-(4-Chloro-2-methylphenoxy)propio...	261	11	21	37	84
Nicosulfuron	257	16	50	63	75
Ethyl 5,5-diphenyl-2-isoxazoline-3-carboxy...	266	6	39	47	30
Deltamethrin	264	19	49	69	169
Benzophenone	350	9	18	19	198
Dicyclohexyl phthalate	279	6	35	34	55
lambda-Cyhalothrin	268	24	53	74	89

Rows: **11** Total Rows: **11**

Each column in Panel 2 corresponds to the fingerprint options from Panel 1. Each fingerprint is a binary bit vector reflecting the presence/absence of features (e.g., ToxPrints comprise 729 features, whereas Morgan fingerprints comprise 2048 features). The color density is scaled by fingerprint type from light to dark and reflects a measure of 'data availability'. The number of data records is reflected in each cell.

a. What data exists for the source analogs identified?

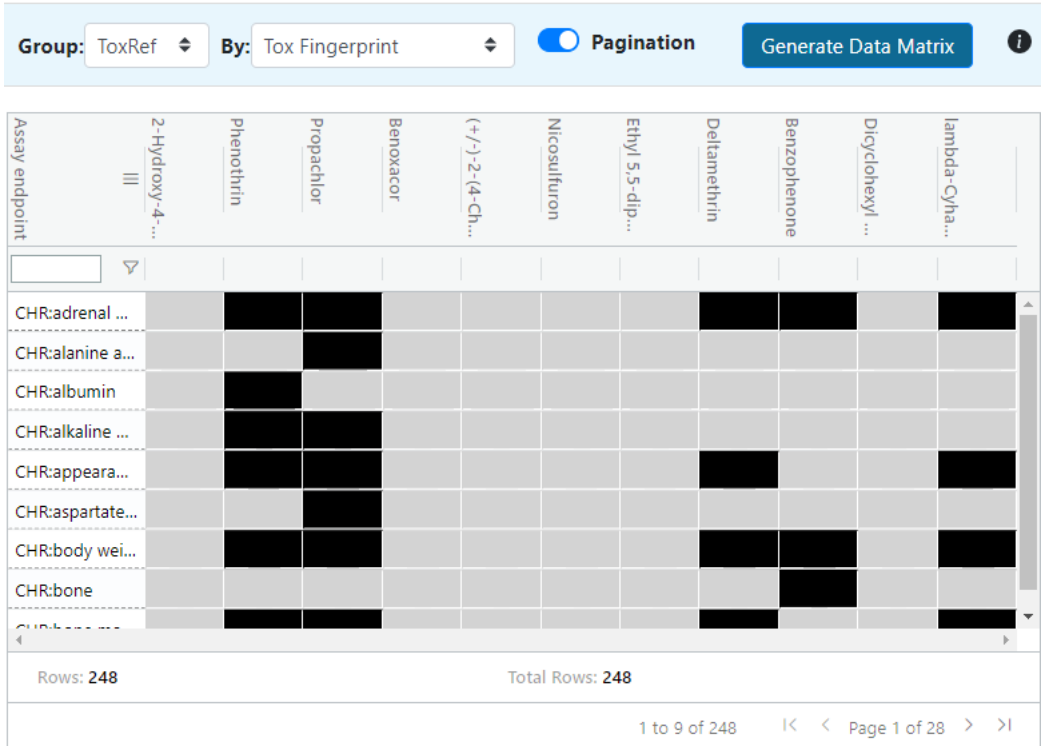
- (The data availability is segmented by data type – where bio_txct represents bioactivity hitcall data from ToxCast, e.g., chm_ct represents chemistry ToxPrint information and tox_txrf represents *in vivo* toxicity effect information from ToxRefDB v2.
- The color density represents a measure of 'data availability' for the target – from light to dark. The number of data records is reflected in the box itself.
- In the case of target HMB, there are moderate amount (tested in 296 assays) available Toxcast bioactivity hitcall data outcomes (bios_txct), but majority of the analogues have less data in comparison to HMB except for benzophenone (tested in 350 assays). There is a moderate amount of in vivo toxicity data available for HMB in ToxRefDB (tox_txrf).)
- However, analogues such as Benzophenone and Propachlor have a great deal more toxicity data which facilitates a read-across.

b. Do the data for the source analogues address the data gaps of interest for the target chemical? (Or, how feasible might a read-across be based on the quantity of data for the source analogues?)

This depends on the use case = it would be a question for discussion as say for a regulatory submission – it might be that you need to address a specific endpoint – so do the analogues identified possess experimental data for the endpoint you need… as well as whether you need to read-across for more than one endpoint to generate more of a predicted toxicity profile.

6. Let's move on to Panel 3. Which types of toxicity effect information are available for the target chemical and across the analogues? *[Hint: Evaluate the Panel 3 on Step 3 of GenRA tool.]*

PANEL 3



(This view provides a representation of the data quantity across different toxicity effects for the source analogues selected. This informs the data gap analysis step of the workflow. Panels 3 and 4 are duplicative in terms of the data presented… we can change the view in Panel 4 based on whether we choose the potency type or start Panel 1 with the use case of wanting to predict ToxCast assay hit calls and therefore filter for this accordingly.)

a. Evaluate the binary predictions of toxicity data arising from ToxRefDB (default). (What data gaps exist for the source analogues relative to the target substance, and which effects might be reasonably predicted by those analogues?)

Use Panel 3 to help profile the data availability across study-toxicity effect combination exists noting that the representation can be changed

if potency predictions are needed. Panel 4 differentiates the view presented in Panel 3.

b. Evaluate Chronic bone marrow assay end point for Benzophenone. Are there any experimental data?

(No, no data)

c. Are there any experimental data for source analogues for Chronic bone marrow assay end point?

(Yes)

7. Would Panels 2 and 3 provide any information on their potency (e.g., in mg/kg-day etc.) or hazard profile? [Note: Click "Generate Data Matrix".]

(No. To evaluate this, click, the 'Generate Data Matrix' button and to move to the next step of the workflow 'Run GenRA Prediction'. See below for the next view)

a. Evaluate the suitability of the source analogues by looking at the data availability of the source analogues. Are there enough data for chronic and sub-chronic assay endpoints?

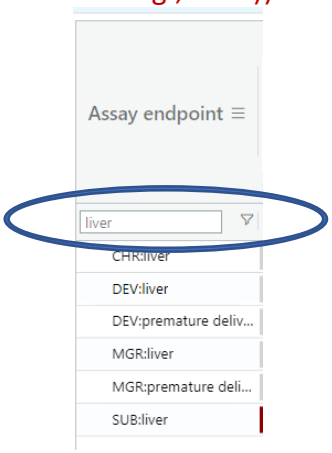
(YES/depends on the end point of interest)

i. If not, what are the next steps?

(If little data are available for the source analogs or they fail to address the data gaps of interest for the target chemical then this might lead the user to change the number of source analogues or select a different similarity context.)

b. Find the dose for the analogue (Benzophenone) at which a toxicity effect for chronic body weight (CHR:liver) was observed. Note if there are any existing observations for HMB for the CHR:liver end point.

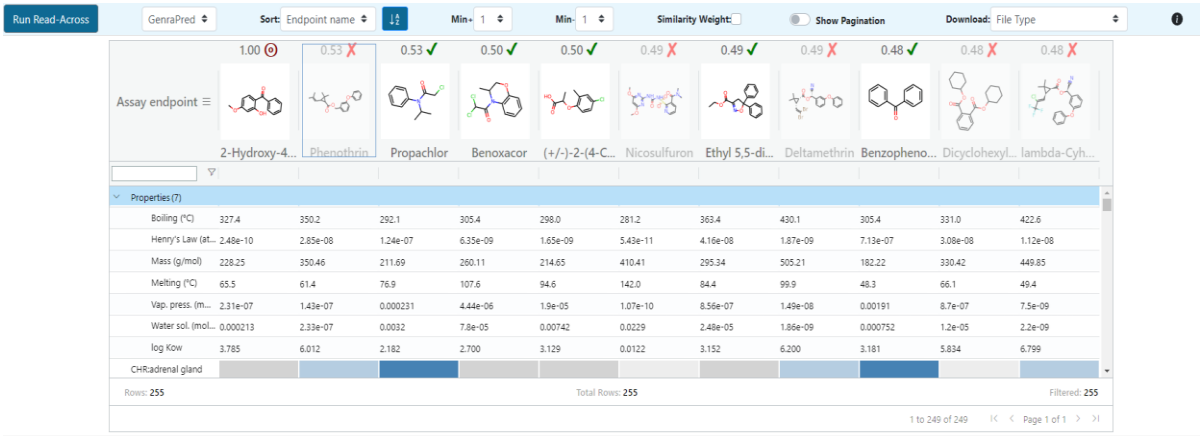
(Sort by 'positive obs', then toggle between ascending and descending obs.: 15 mg/kg/day) (filter by endpoint name e.g., Liver))



(Yes, and take a note of POS, ACT 0.613 generated for HMB after running Read-Across using the GenraPred engine which shows that the prediction is consistent with the empirical data available. The degree of cell opacity denotes the confidence associated with any prediction. The confidence is measured by 2 characteristics, the Area under the curve (AUC) and the p-value (see Shah et al., 2016 for further details). The higher the AUC and the lower the p-value then the more confident the prediction.)

c. Consider the Physchem properties of each analogue and identify those with estimated Log Kow values within 1.5 of the estimate for the target chemical Log Kow (e.g., +1.5 and -1.5 of 3.785)

   i. De-select any analogues that are below 2 and above 5. How many analogues are now selected for Read-Across?
   (5, Propachlor, benoxacor, 2,4-chloro-2-methylphenoxy) propionic acid, ethyl 5,5-diphenyl-2-isoxazoline-3-carboxylate, and benzophenone)



ii. Once the desired source analogues are selected, click the 'Run Read-across" button to re-derive the GenRA predictions using the GenraPred engine.

(The prediction is 1, still consistent with the POS empirical data for CHR-liver.)

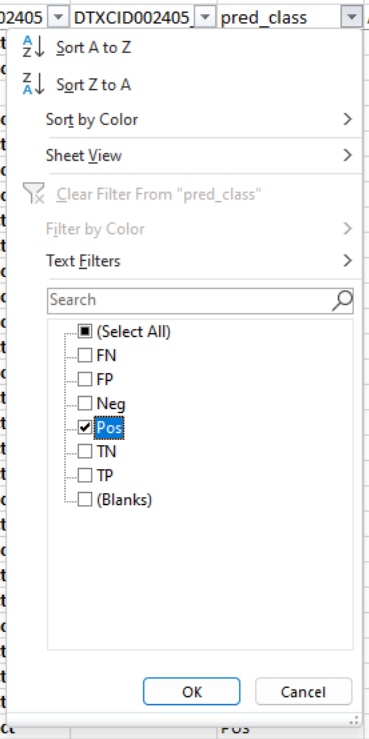
d. Export and save the results as an Excel file.
   i. Does the file export the target predictions as well as all 10 analogues' data?
   (NO, just target chemical predictions and selected 5 analogues)
   ii.

chem_id	DTXCID002405	DTXCID002405_ur	pred_class	ACT	AUC	pval	DTXCID504274	DTXCID504274_ur	DTXCID809572	DTXCID809572_ur	DTXCID904194	DTXCID904194_ur	DTXCID5020360	DTXCID5020360_u	DTXCID101961	DTXCID101961_units
role	target						analog		analog		analog		analog		analog	
preferred name	2-Hydroxy-4-methoxybenzophenone						Propachlor		Benoxacor		(+/-)-2-(4-Chloro-2-methylphenoxy		Ethyl 5,5-diphenyl-2-isoxazoline-3-		Benzophenone	
dsstox_sid	DTXSID3022405						DTXSID4024274		DTXSID3029572		DTXSID9024194		DTXSID7040360		DTXSID0021961	
dsstox_cid	DTXCID002405						DTXCID504274		DTXCID809572		DTXCID904194		DTXCID5020360		DTXCID101961	
similarity	1.00001						0.525478938		0.502310231		0.497943006		0.493418743		0.482142857	
Mass g/mol	228.247						211.69		260.11		214.65		295.338		182.222	
Melting °C	65.5044						76.9232		107.642		94.6045		84.3575		48.2661	
Boiling °C	327.437						292.105		305.371		297.982		363.43		305.417	
log Kow	3.7854						2.18155		2.69953		3.12858		3.15207		3.18141	
Vap. press. mmHg	2.30687E-07						0.000231259		4.43675E-06		1.90404E-05		8.55809E-07		0.00191301	
Water sol. mol/L	0.000212758						0.00320445		7.79988E-05		0.00741697		2.48317E-05		0.000751735	
Henry's Law atm-m3/mole	2.48418E-10						1.23831E-07		6.35321E-09		1.65471E-09		4.16158E-08		7.12881E-07	
Hydrogen Bond Donors																
Hydrogen Bond Acceptors																
chem_id	DTXCID002405	DTXCID002405	pred_class	ACT	AUC	pval	DTXCID504274	DTXCID504274	DTXCID809572	DTXCID809572	DTXCID904194	DTXCID904194	DTXCID502036(	DTXCID502036(	DTXCID101961	DTXCID101961_its
CHR:adrenal gland	no_effect		Neg	0	0	1	no_effect		no_data		no_data		no_data		no_effect	
CHR:alanine aminotransferas	pos_effect		Pos	1	0	1	65.5	mg/kg/day	no_data		no_data		no_data		no_data	
CHR:albumin	no_data						no_data		no_data		no_data		no_data		no_data	
CHR:alkaline phosphatase (al	pos_effect		Pos	1	0	1	125.3	mg/kg/day	no_data		no_data		no_data		no_data	
CHR:appearance and color	no_effect		Neg	0	0	1	no_effect		no_data		no_data		no_data		no_data	
CHR:aspartate aminotransfer	pos_effect		Pos	1	0	1	16.1	mg/kg/day	no_data		no_data		no_data		no_data	
CHR:body weight	pos_effect		Pos	1	0	1	6.25	mg/kg/day	no_data		no_data		no_data		30	mg/kg/day
CHR:bone	no_effect		Neg	0	0	1	no_data		no_data		no_data		no_data		no_effect	
CHR:bone marrow	no_effect		Neg	0	0	1	no_effect		no_data		no_data		no_data		no_data	
CHR:brain	pos_effect		Pos	0.522	0	0.405	222.9	mg/kg/day	no_data		no_data		no_data		no_effect	
CHR:cholesterol	pos_effect		Pos	1	0	1	19.3	mg/kg/day	no_data		no_data		no_data		no_data	
CHR:clinical signs	pos_effect		Pos	1	0	1	25	mg/kg/day	no_data		no_data		no_data		60	mg/kg/day
CHR:clitoral gland	no_effect		Neg	0	0	1	no_data		no_data		no_data		no_data		no_effect	
CHR:creatine phosphokinase	pos_effect		Pos	1	0	1	125.3	mg/kg/day	no_data		no_data		no_data		no_data	
CHR:ear	no_effect		Neg	0	0	1	no_effect		no_data		no_data		no_data		no_data	
CHR:epididymis	no_effect		Neg	0	0	1	no_effect		no_data		no_data		no_data		no_effect	
CHR:erythrocyte (rbc) count c	no_effect		Neg	0	0	1	no_effect		no_data		no_data		no_data		no_data	
CHR:esophagus	no_effect		Neg	0	0	1	no_effect		no_data		no_data		no_data		no_effect	
CHR:estrous cycle	pos_effect		Pos	1	0	1	0.625	mg/kg/day	no_data		no_data		no_data		no_data	
CHR:eye	no_effect		Neg	0	0	1	no_effect		no_data		no_data		no_data		no_effect	
CHR:food consumption	pos_effect		Pos	1	0	1	25	mg/kg/day	no_data		no_data		no_data		60	mg/kg/day
CHR:full gross necropsy	no_effect		Neg	0.478	0	0.51	no_effect		no_data		no_data		no_data		70	mg/kg/day
CHR:gallbladder	no_effect		Neg	0	0	1	no_data		no_data		no_data		no_data		no_effect	
CHR:gamma glutamyl transfe	pos_effect		Pos	1	0	1	53.6	mg/kg/day	no_data		no_data		no_data		no_data	
CHR:glucose	no_effect		Neg	0	0	1	no_effect		no_data		no_data		no_data		no_data	
CHR:harderian gland	no_effect		Neg	0	0	1	no_data		no_data		no_data		no_data		no_effect	
CHR:heart	no_effect		Neg	0	0	1	no_effect		no_data		no_data		no_data		no_effect	
CHR:hematocrit (hct)	pos_effect		Pos	1	0	1	292.1	mg/kg/day	no_data		no_data		no_data		no_data	
CHR:hemoglobin (hgb)	pos_effect		Pos	1	0	1	292.1	mg/kg/day	no_data		no_data		no_data		no_data	
CHR:intestine large	no_effect		Neg	0	0	1	no_effect		no_data		no_data		no_data		no_effect	
CHR:intestine small	no_effect		Neg	0	0	1	no_effect		no_data		no_data		no_data		no_effect	
CHR:kidney	pos_effect		Pos	1	0	1	125.3	mg/kg/day	no_data		no_data		no_data		15	mg/kg/day
CHR:lactic acid dehydrogenas	no_data						no_data		no_data		no_data		no_data		no_data	
CHR:larynx	no_effect		Neg	0	0	1	no_effect		no_data		no_data		no_data		no_effect	

e. Are there any subsequent analysis after this supported by GenRA?
(No, any subsequent analysis is contingent on the end user expertise. However, the excel file columns can be sorted and filter e.g., filtering by positive effect etc.)

f. Filter the column D (pred_class) to show only positive effects.
Have the excel file ready to be shared with the breakout room and filter based on only positive effects

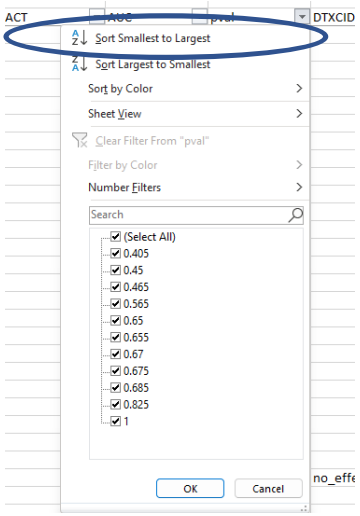
chem_id	DTXCID002405	DTXCID002405_ur	pred_class	ACT	Al
role	target				
preferred name	2-Hydroxy-4-methoxybenzophenone				
dsstox_sid	DTXSID3022405				
dsstox_cid	DTXCID002405				
similarity	1.00001				
Mass g/mol	228.247				
Melting °C	65.5044				
Boiling °C	327.437				
log Kow	3.7854				
Vap. press. mmHg	2.30687E-07				
Water sol. mol/L	0.000212758				
Henry's Law atm-m3/mole	2.48418E-10				
Hydrogen Bond Donors					
Hydrogen Bond Acceptors					
chem_id	DTXCID002405	DTXCID002405_	pred_class	ACT	Al
CHR:adrenal gland	no_effect			0	
CHR:alanine aminotransferas	pos_effec			1	
CHR:albumin	no_data				
CHR:alkaline phosphatase (al	pos_effec			1	
CHR:appearance and color	no_effect			0	
CHR:aspartate aminotransfer	pos_effec			1	
CHR:body weight	pos_effec			1	
CHR:bone	no_effect			0	
CHR:bone marrow	no_effect			0	
CHR:brain	pos_effec			0.522	
CHR:cholesterol	pos_effec			1	
CHR:clinical signs	pos_effec			1	
CHR:clitoral gland	no_effect			0	
CHR:creatine phosphokinase	pos_effec			1	
CHR:ear	no_effect			0	
CHR:epididymis	no_effect			0	
CHR:erythrocyte (rbc) count c	no_effect			0	
CHR:esophagus	no_effect			0	
CHR:estrous cycle	pos_effec			1	
CHR:eye	no_effect			0	
CHR:food consumption	pos_effec			1	
CHR:full gross necropsy	no_effect			0.478	
CHR:gallbladder	no_effect			0	
CHR:gamma glutamyl transfe	pos_effec			1	
CHR:glucose	no_effect			0	
CHR:harderian gland	no_effect			0	
CHR:heart	no_effect			0	
CHR:hematocrit (hct)	pos_effec			1	
CHR:hemoglobin (hgb)	pos_effect		Pos	1	



chem_id	DTXCID002405	DTXCID002405_ur	pred_class	ACT	AUC	pval	DTXC
role	target						anal
preferred name	2-Hydroxy-4-methoxybenzophenone						Prop
dsstox_sid	DTXSID3022405						DTXS
dsstox_cid	DTXCID002405						DTXC
similarity	1.00001						
Mass g/mol	228.247						
Melting °C	65.5044						
Boiling °C	327.437						
log Kow	3.7854						
Vap. press. mmHg	2.30687E-07						
Water sol. mol/L	0.000212758						
Henry's Law atm-m3/mole	2.48418E-10						
Hydrogen Bond Donors							
Hydrogen Bond Acceptors							
chem_id	DTXCID002405	DTXCID002405_	pred_class	ACT	AUC	pval	DTXC
CHR:alanine aminotransferas	pos_effect		Pos	1	0	1	
CHR:alkaline phosphatase (al	pos_effect		Pos	1	0	1	
CHR:aspartate aminotransfer	pos_effect		Pos	1	0	1	
CHR:body weight	pos_effect		Pos	1	0	1	
CHR:brain	pos_effect		Pos	0.522	0	0.405	
CHR:cholesterol	pos_effect		Pos	1	0	1	
CHR:clinical signs	pos_effect		Pos	1	0	1	
CHR:creatine phosphokinase	pos_effect		Pos	1	0	1	
CHR:estrous cycle	pos_effect		Pos	1	0	1	
CHR:food consumption	pos_effect		Pos	1	0	1	
CHR:gamma glutamyl transfe	pos_effect		Pos	1	0	1	
CHR:hematocrit (hct)	pos_effect		Pos	1	0	1	
CHR:hemoglobin (hgb)	pos_effect		Pos	1	0	1	
CHR:kidney	pos_effect		Pos	1	0	1	
CHR:leukocyte (wbc) count di	pos_effect		Pos	1	0	1	
CHR:liver	pos_effect		Pos	1	0	1	
CHR:lymphocyte	pos_effect		Pos	1	0	1	
CHR:ovary	pos_effect		Pos	0.522	0	0.465	
CHR:platelet	pos_effect		Pos	1	0	1	
CHR:stomach	pos_effect		Pos	0.522	0	0.45	
CHR:testes	pos_effect		Pos	1	0	1	
CHR:thyroid gland	pos_effect		Pos	1	0	1	
CHR:uterus	pos_effect		Pos	0.522	0	0.565	
DEV:aborted	pos_effect		Pos	0.671	0	0.67	
DEV:body weight	pos_effect		Pos	1	0	1	
DEV:bone	pos_effect		Pos	0.65	0	0.655	no_e
DEV:clinical signs	pos_effect		Pos	1	0	1	
DEV:food consumption	pos_effect		Pos	1	0	1	
DEV:kidney	pos_effect		Pos	1	0	1	no_c
DEV:liver	pos_effect		Pos	1	0	1	no_c
DEV:lung	pos_effect		Pos	1	0	1	no_c
DEV:mortality	pos_effect		Pos	1	0	1	
DEV:postimplantation loss	pos_effect		Pos	0.679	0	0.685	
DEV:resorptions	pos_effect		Pos	0.679	0	0.65	

GenRA Metadata

g. Sort the pval from smallest to largest value (The higher the AUC and the lower the p-value then the more confident the prediction.) Here again show the excel file and sort the pval.



ACT	AUC	pval	DTX
0.522	0	0.405	
0.522	0	0.45	
0.522	0	0.465	
0.522	0	0.565	
0.679	0	0.65	
0.65	0	0.655	
0.671	0	0.67	
0.669	0	0.675	
0.679	0	0.685	
0.512	0	0.825	
1	0	1	
1	0	1	
1	0	1	
1	0	1	
1	0	1	
1	0	1	
1	0	1	
1	0	1	
1	0	1	
1	0	1	
1	0	1	
1	0	1	

Reflection

1. In what case example from your work environment would GenRA be useful?

   [Open response]

2. What have you learned about the process and workflow used to find information in GenRA?

   [Open response]

3. What challenges did you encounter, and how did you solve them?

   [Open response]