United States
Environmental Protection
Agency

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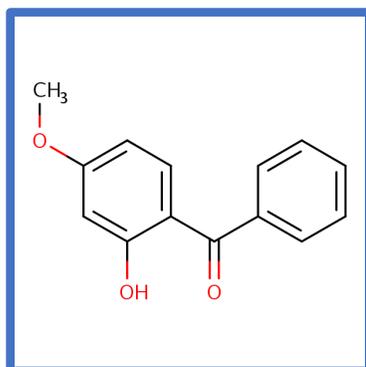
IUPAC Name: (2-Hydroxy-4-methoxyphenyl)(phenyl)methanone
CASRN: 131-57-7
DSSTox substance identifier (DTXSID): DTXSID3022405
Molecular Formula: $C_{14}H_{12}O_3$
SMILES: COC1=CC=C(C(=O)C2=CC=CC=C2)C(O)=C1
InChIKey: DXGLGDHPHMLXJC-UHFFFAOYSA-N

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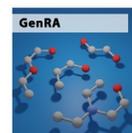
b. Find the dose for the analogue (Benzophenone) at which a toxicity effect for chronic body weight (CHR:liver) was observed. Note below if there are any observations exist for HMB for CHR:liver end point.

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?