Integrating Structural and Mechanistic Evidence to Select Relevant Analogs for Read-Across Assessment

TEXAS A&M UNIVERSITY
Veterinary Medicine & Biomedical Sciences

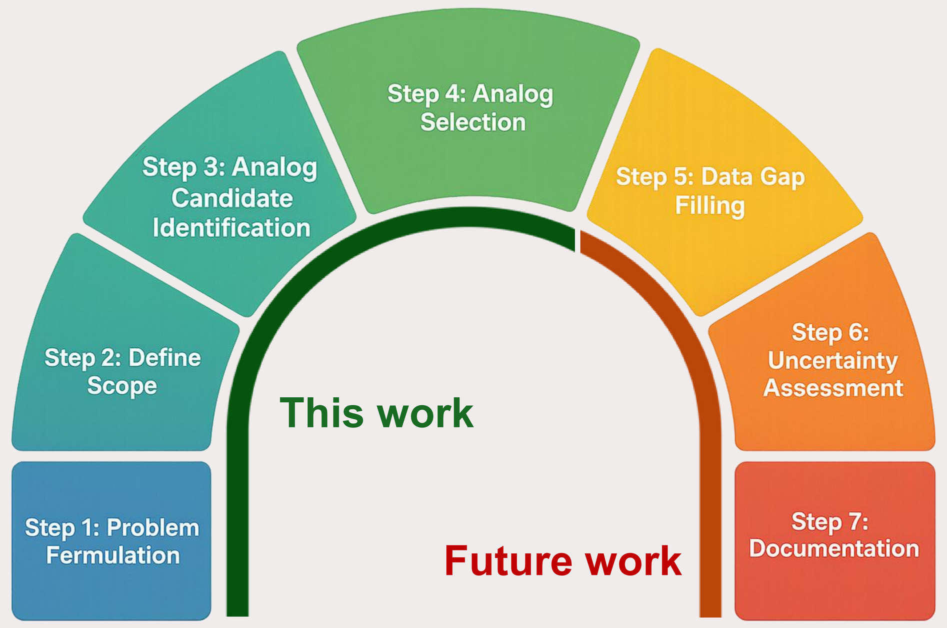
PETA SCIENCE CONSORTIUM INTERNATIONAL e.V.

Chi Lam (Thomas) Chan[1], Gina Hilton[2], Amber Goetz[3], Agnes Karmaus[3]
[1]Texas A&M University, College Station, TX; [2]PETA Science Consortium International e.V.;
[3]Syngenta Crop Protection, LLC, Greensboro, NC

BACKGROUND AND OBJECTIVE

Enhancing the quality of assessments and reducing animal testing for pesticides requires robust read-across methods—especially considering more than 10,000 animals are used to bring one pesticide ingredient to market.

This study presents a framework for analog selection and inclusion criteria specifically for the ACCase inhibitor class. By integrating structural and biological similarity, we provide a precautionary assessment strategy that reduces animal dependency without compromising safety.



METHODS

Step 1 Problem formulation

There is sufficient knowledge about the target compound to warrant leveraging read-across to fill a data gap for carcinogenicity data:

- Known mode of action (ACCase)
- Known structure for comparison
- Unique feature (no PPAR activation)

Step 2 Define scope

Compile pesticide inventories as these chemicals would have the relevant exposures:

- Fungicides (FRAC): 212 chemicals
- Herbicides (HRAC): 325 chemicals
- Insecticide (IRAC): 261 chemicals

Step 3 Analog candidate identification

Database development with relevant parameters:

- Chemical structure (fingerprints)
- Pesticidal mode of action
- Toxicity data (90-day subchronic and/or 2-year chronic)
- Key features (PPAR transactivation data)

Step 4 Analog selection

Evaluate data relevance and reliability

Calculate similarity
- Feature reduction
- Tanimoto similarity

Document justification and selection

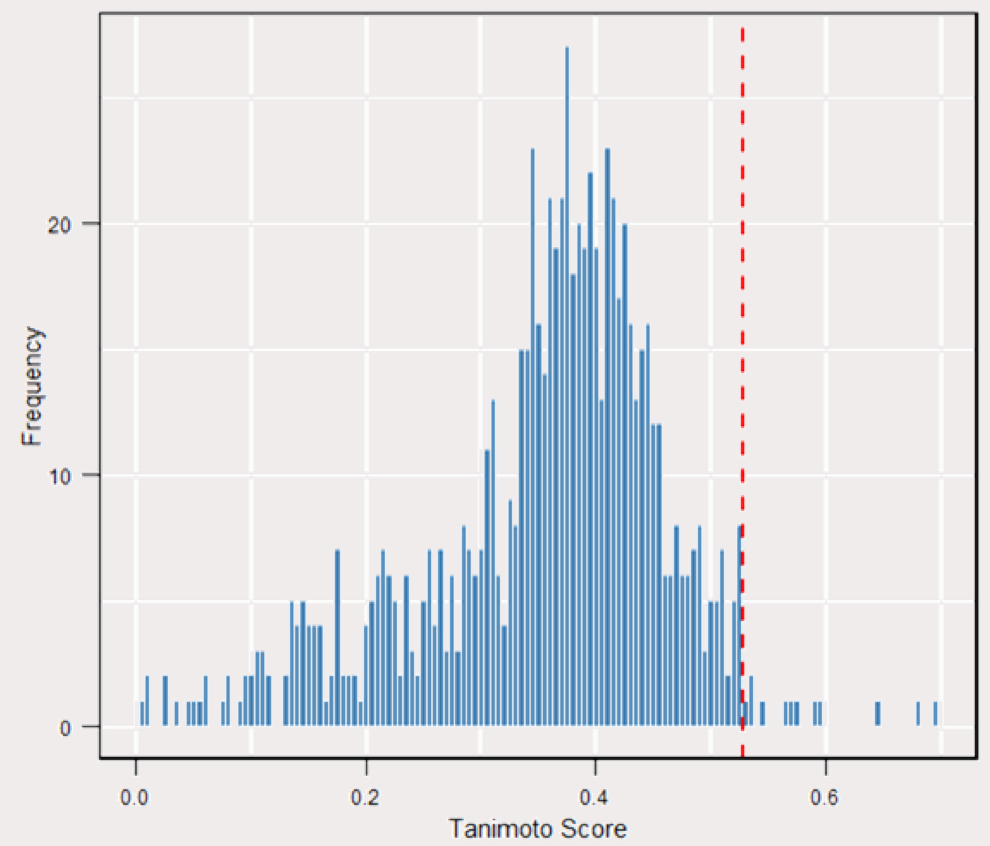
Fingerprint Matrix Subset

	ToxPrints			MACCS			PubChem		
CASRN	bit 1	bit 2	bit 3	bit 1	bit 2	bit 3	bit 1	bit 2	bit 3
144-21-8	0	0	0	0	1	1	0	0	0
79540-50-4	0	0	0	1	1	1	0	0	0
59682-52-9	0	0	0	1	1	1	0	0	0
42609-73-4	0	0	0	1	0	1	0	0	0
7287-36-7	0	0	0	1	0	1	1	0	0
2163-80-6	0	0	0	0	1	1	0	0	0
112-80-1	1	0	1	0	1	1	0	0	0
112-05-0	1	0	1	0	1	1	0	0	0
88678-67-5	0	0	0	1	1	1	0	0	0

Three complementary chemical fingerprinting libraries were used in aggregate: ToxPrints, MACCS, and PubChem. ToxPrints (730 bits) is tailored for toxicological modeling but may omit non-toxicity-related structures; MACCS (166 bits) comprises predefined fragments with limited coverage; PubChem (641 bits) provide broad structural representation but are not toxicity-specific. **Combining these fingerprints (3DB) maximized toxicological relevance and structural diversity.** Structural similarity was quantified using Tanimoto coefficients calculated from an aggregate fingerprint after feature reduction, which reduced the total number of bits from 1,537 to 1,017. ToxPrints data were obtained from the CompTox Chemicals Dashboard. MACCS and PubChem fingerprints were generated using the rcdk v3.8.1 R package.

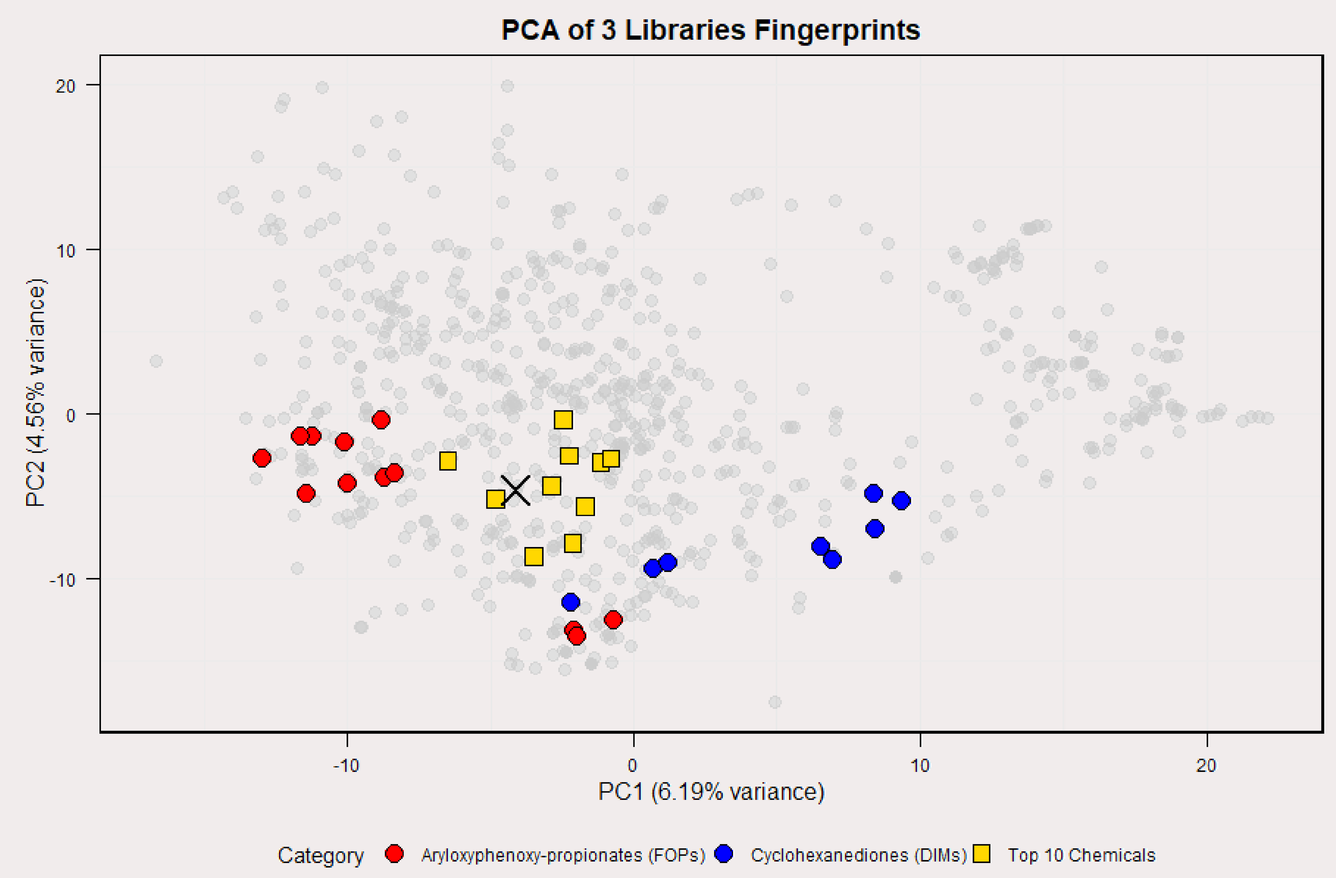
RESULTS: ANALOG CANDIDATE IDENTIFICATION

Tanimoto Score Distribution



Similarity score cutoff was set at the inflection point (drop shoulder) of the high-score region to identify the topmost ranked analogs. This approach improves transparency and objectivity while avoiding reliance on arbitrary thresholds.

Principal Component Analysis (PCA)



PCA based on aggregate fingerprints was used to evaluate analog distribution. Despite low variance explained by the leading principal components, substantial clustering among the top-ranked analogs supported their structural relevance.

PPAR Activity Table

Biological similarity assessment based on PPAR activation was used to refine highest ranking analogs. This approach was anchored on identifying analogs that do not activate PPAR. Seventeen ToxCast PPAR assay endpoints were evaluated (ICE v4.1.1); pesticides showing less than 10 percent PPAR activity across tested assay endpoints were retained.

PPAR Assay Name (ToxCast)	94-74-6	39196-18-4	55406-53-6	104-55-2	534-52-1	114369-43-6	30560-19-1	135410-20-7	34256-82-1	135158-54-2
OT_PPARg_PPARgSRC1_0480	0	NA	0	NA	0	0	0	0	0	0
OT_PPARg_PPARgSRC1_1440	0	NA	0	NA	0	0	0	0	0	0
ATG_PPARg_TRANS_up	0	NA	0	NA	1	0	0	0	1	0
ATG_PPARa_TRANS_up	0	NA	0	NA	0	0	0	0	0	0
TOX21_PPARd_BLA_agonist_ratio	0	0	0	0	0	0	0	0	0	0
TOX21_PPARd_BLA_antagonist_ratio	0	0	1	0	0	0	0	0	1	0
TOX21_PPARg_BLA_antagonist_ratio	0	0	1	0	1	0	0	0	1	0
TOX21_PPARg_BLA_Agonist_ratio	0	0	1	0	0	0	0	0	0	0
NVS_NR_hPPARa	NA	NA	0	NA	0	NA	NA	NA	NA	NA
ATG_PPARd_TRANS_up	0	NA	0	NA	0	0	0	0	0	0
NVS_NR_hPPARg	1	NA	0	NA	0	NA	0	0	NA	NA
ATG_mPPARg_XSP2_up	NA	NA	1	NA	NA	NA	NA	NA	1	NA
ATG_hPPARg_XSP2_up	NA	NA	1	NA	NA	NA	NA	NA	1	1
ATG_mPPARg_XSP1_up	NA	NA	1	NA	NA	NA	NA	NA	1	NA
ATG_hPPARg_XSP1_up	NA	NA	1	NA	NA	NA	NA	NA	1	NA
ATG_zfPPARg_XSP2_up	NA	NA	0	NA	NA	NA	NA	NA	0	0
ATG_zfPPARg_XSP1_up	NA	NA	0	NA	NA	NA	NA	NA	0	NA
Total PPAR Assays Active	1	0	7	0	2	0	0	0	7	1
Total PPAR Assays Tested	11	5	18	5	12	10	11	11	16	12
Percent of PPAR Assays Active	**9%**	**0%**	**39%**	**0%**	**17%**	**0%**	**0%**	**0%**	**44%**	**8%**

RESULTS: ANALOG SELECTION

Analog Selection Table

CASRN	ToxPrints	MACCS	PubChem	3DB	Pesticide Type	PPAR Activity %	Chemical Class	Pesticidal Mode of Action	90-day rat NOAEL (mg/kg/day)	90-day rat LOAEL (mg/kg/day)	90-day dog NOAEL (mg/kg/day)	90-day dog LOAEL (mg/kg/day)	Chronic Dietary (All Populations; mg/kg/day)
1229023-00-0	1.00	1.00	1.00	1.00	Insecticide	0%	Tetronic and tetramic acid derivatives	Inhibitors of acetyl-CoA carboxylase	31.5/36.1 (M/F)	159/110 (M/F)	15	30	NOAEL = 15
57646-30-7	0.36	0.40	0.52	0.48	Fungicide	0%	Phenylamide fungicides	Nucleic acids metabolism	N/A	N/A	N/A	N/A	N/A
57837-19-1	0.33	0.44	0.57	0.51	Fungicide	0%	Phenylamide fungicides	Nucleic acids metabolism	17	83	7.80/7.41 (M/F) (Chronic)	30.63/32.36 (M/F) (Chronic)	NOAEL= 7.41
32809-16-8	0.54	0.46	0.51	0.50	Fungicide	7%	Dicarboximides	Signal transduction	7.5 (6 months)	25 (6 months)	100	500	NOAEL = 3.5
50471-44-8	0.26	0.50	0.51	0.48	Fungicide	0%	Dicarboximides	Signal transduction	4.4	24	2.5 (6-month)	7.5 (6-month)	N/A
156052-68-5	0.29	0.43	0.59	0.51	Fungicide	0%	Benzamide	Cytoskeleton and motor proteins	1509	N/A	281	N/A	NOAEL = 48
77732-09-3	0.32	0.58	0.56	0.55	Fungicide	0%	Phenylamide fungicides	Nucleic acids metabolism	N/A	N/A	N/A	N/A	N/A
134605-64-4	0.28	0.61	0.47	0.49	Herbicide	0%	N-phenyl-imides	Inhibitors of protoporphyrinogen oxidase	6.1/7.1 (M/F)	18.8/20.6 (M/F)	200	N/A	N/A
81777-89-1	0.38	0.62	0.61	0.59	Herbicide	0%	Isoxazolidinone	Inhibitors of deoxy-D-xylulose phosphate synthase	135.2/160.9 (M/F)	273/319.3 (M/F)	>= 1038/1012 (Chronic)	> 1038/1012 (Chronic)	NOAEL = 84
87546-18-7	0.33	0.65	0.50	0.52	Herbicide	9%	N-phenyl-imides	Inhibitors of protoporphyrinogen oxidase	67.75	659	10	100	N/A
243973-20-8	0.36	0.67	0.50	0.53	Herbicide	0%	Phenylpyrazoline	Inhibitors of acetyl CoA carboxylase	466/537 (M/F)	900/965 (M/F)	100	250	N/A
87820-88-0	0.26	0.46	0.55	0.49	Herbicide	0%	Cyclohexanediones (DIMs)	Inhibitors of acetyl CoA carboxylase	20.5/23.0 (M/F)	204.8/219.3 (M/F)	0.5	5	N/A
112226-61-6	0.26	0.36	0.63	0.51	Insecticide	0%	Diacylhydrazines	Ecdysone receptor agonists	N/A	N/A	N/A	N/A	N/A
161050-58-4	0.30	0.41	0.55	0.49	Insecticide	0%	Diacylhydrazines	Ecdysone receptor agonists	69.3/72.4 (M/F)	353.4 (M/F)	197.5/208.8 (M/F)	> 197.5/208.8 (M/F)	NOAEL = 10.2
732-11-6	0.32	0.51	0.57	0.52	Insecticide	0%	Organophosphates	Inhibitors of acetylcholinesterase	1.69	N/A	N/A	N/A	N/A
148477-71-8	0.29	0.53	0.57	0.53	Insecticide	0%	Tetronic and tetramic acid derivatives	Inhibitors of acetyl-CoA carboxylase	N/A	N/A	7.7/8.4 (M/F)	26.6/28.0 (M/F)	NOAEL = 1.5
283594-90-1	0.32	0.44	0.54	0.49	Insecticide	0%	Tetronic and tetramic acid derivatives	Inhibitors of acetyl-CoA carboxylase	31.7	204	1.85	9.2 (F)	NOAEL = 3.8
203313-25-1	0.52	0.70	0.73	0.69	Insecticide	0%	Tetronic and tetramic acid derivatives	Inhibitors of acetyl-CoA carboxylase	148	616	5 (1-year)	19 (1-year)	N/A
112410-23-8	0.32	0.33	0.57	0.47	Insecticide	0%	Diacylhydrazines	Ecdysone receptor agonists	14.35	140	2.1	20.1	N/A

A curated database of 18 candidate analogs was developed to support read-across. Analog suitability was evaluated using five criteria: (1) structural similarity, based on Tanimoto scores (ToxPrints, MACCS, PubChem, and 3DB; ~0.5–0.7 for closest analogs); (2) PPAR activity, capturing potential biological relevance across PPARα/δ/γ; (3) chemical class, including pesticide type and sub-classification (4) pesticide mode of action, enabling mechanistic grouping; and (5) 90-day toxicity data, incorporating Point of Departure (POD) values from rat and dog studies.

CONCLUSIONS

- By incorporating a reproducible computational workflow supporting candidate analog identification integrating multiple types of evidence, this approach enhances transparency to bolster confidence in the underlying data identified for the read-across assessment.
- Overall, this approach demonstrates a systematic, evidence-based method for selecting suitable analogs for read-across applications.

ACKNOWLEDGEMENTS

This work was funded by PETA Science Consortium International e.V., and Texas A&M University, and NIEHS Training Grant (T32 ES026568).

REFERENCES

- Durant JL, et al. (2002). Reoptimization of MDL keys for use in drug discovery. *J Chem Inf Comput Sci*, 42(6):1273–1280. doi:10.1021/ci010132r
- Fungicide Resistance Action Committee (FRAC). FRAC Code List: Fungicides sorted by mode of action. https://www.frac.info
- Guha R. (2007). Chemical informatics functionality in R. *J Stat Softw*, 18(5):1–16.
- Herbicide Resistance Action Committee (HRAC). Classification of herbicides according to mode of action. https://hracglobal.com
- Insecticide Resistance Action Committee (IRAC). IRAC Mode of Action Classification Scheme. https://irac-online.org
- Kim S, et al. (2019). PubChem 2019 update: improved access to chemical data. *Nucleic Acids Res*, 47(D1):D1102–D1109. doi:10.1093/nar/gky1033
- National Toxicology Program's Integrated Chemical Environment version 4.1.1, released February 2025. https://ice.ntp.niehs.nih.gov/
- Steinbeck C, et al. (2006). The Chemistry Development Kit (CDK): An open-source Java library for chemo- and bioinformatics. *J Chem Inf Comput Sci*, 46(3):1043–1047. doi:10.1021/ci050400b
- Yang C, et al. (2015). ToxPrints: A chemoinformatics approach to association of chemical features with toxicological endpoints. *J Chem Inf Model*, 55(11):2360–2373. doi:10.1021/acs.jcim.5b00386