EPA United States Environmental Protection Agency

The Sequence Alignment to Predict Across Species Susceptibility (SeqAPASS) Tool: Catalyzing a Change in Species Extrapolation

Carlie A. LaLone, Ph.D.
Research Bioinformaticist

Office of Research and Development
Center for Computational Toxicology and Exposure, Great Lakes Toxicology and Ecology Division

The views expressed in this presentation are those of the authors and do not necessarily reflect the views or policies of the US EPA

Sept. 28th, 2021

Chemicals make up the world around us – necessary for our modern society

Toxicity Testing to Understand Chemical Safety

- Regulatory decision-making

US EPA Examples:

Clean Air Act
Clean Water Act
Resource Recovery Act
Endangered Species Act
Food Quality Protection Act
Endocrine Disruptor Screening Program
Federal Insecticide, Fungicide, and Rodenticide Act
Frank R. Lautenberg Chemical Safety for the 21st Century Act
Comprehensive Environmental Response, Compensation, and Liability Act
Guidelines for Deriving Numerical National Water Quality Criteria for the Protection of Aquatic Organisms and Their Uses

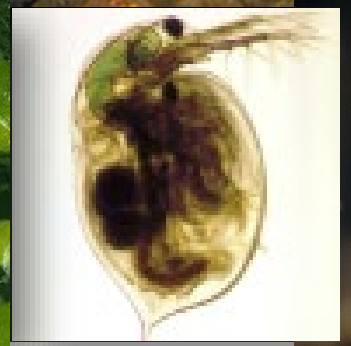
cheap and readily available

short lifespans and rapid life cycles

easy maintenance and good breeding capabilities

ability to control diet and surroundings

requires least space and time-consuming care



Species Extrapolation

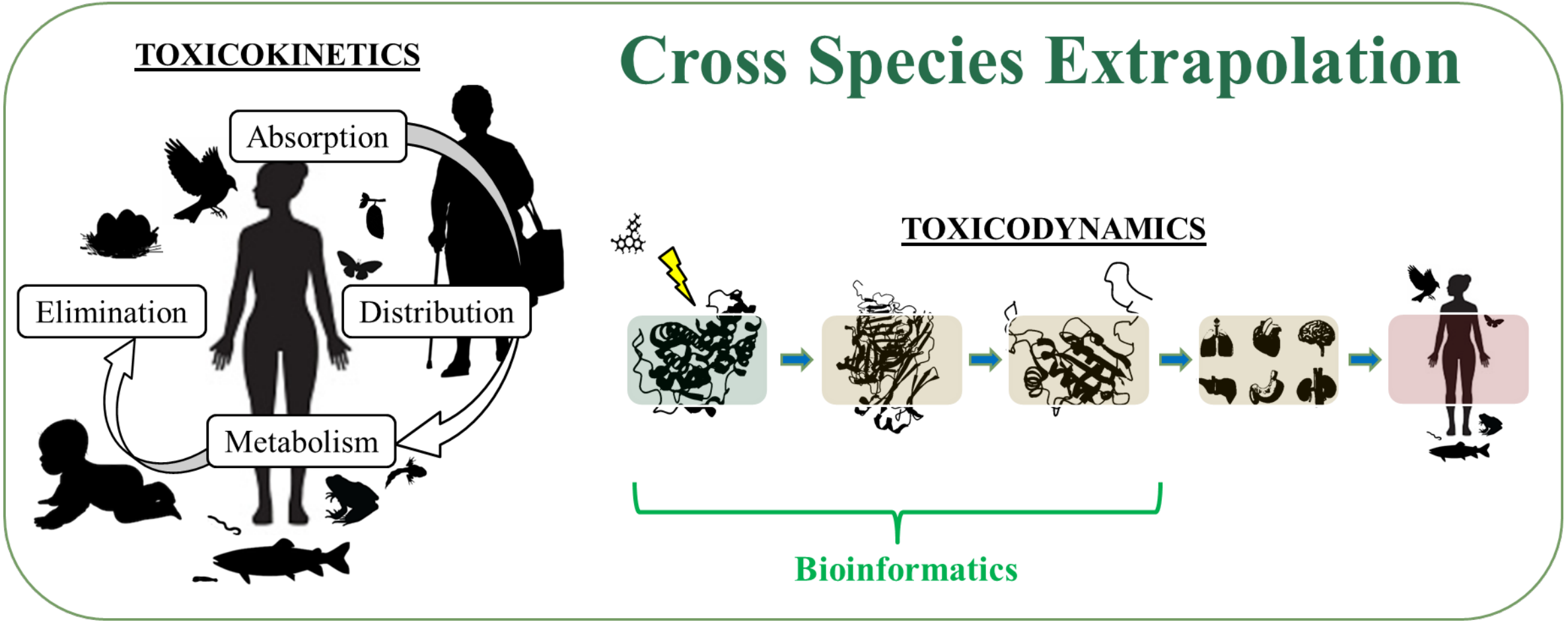
What is it?

- Using existing knowledge about one species to estimate, predict, project, or infer the effect, impact, or trajectory of another species
  - For chemical safety typically dealing with toxicity

Why is it important:

- Limited or no toxicological data for the animal or plant species of interest – reliance on surrogate (model organisms)
  - Impractical to generate new data for all species
- Testing resources are limited
  - International interest to reduce animal use
  - Ever-increasing demand to evaluate more chemicals in a timely and sometimes expedited manner
- Sensitivity of species must be estimated based on scientifically-sound methods of cross-species extrapolation
  - Immense diversity of species in the wild
  - Important challenge for species listed under the Endangered Species Act

Sensitivity to Chemical Perturbation

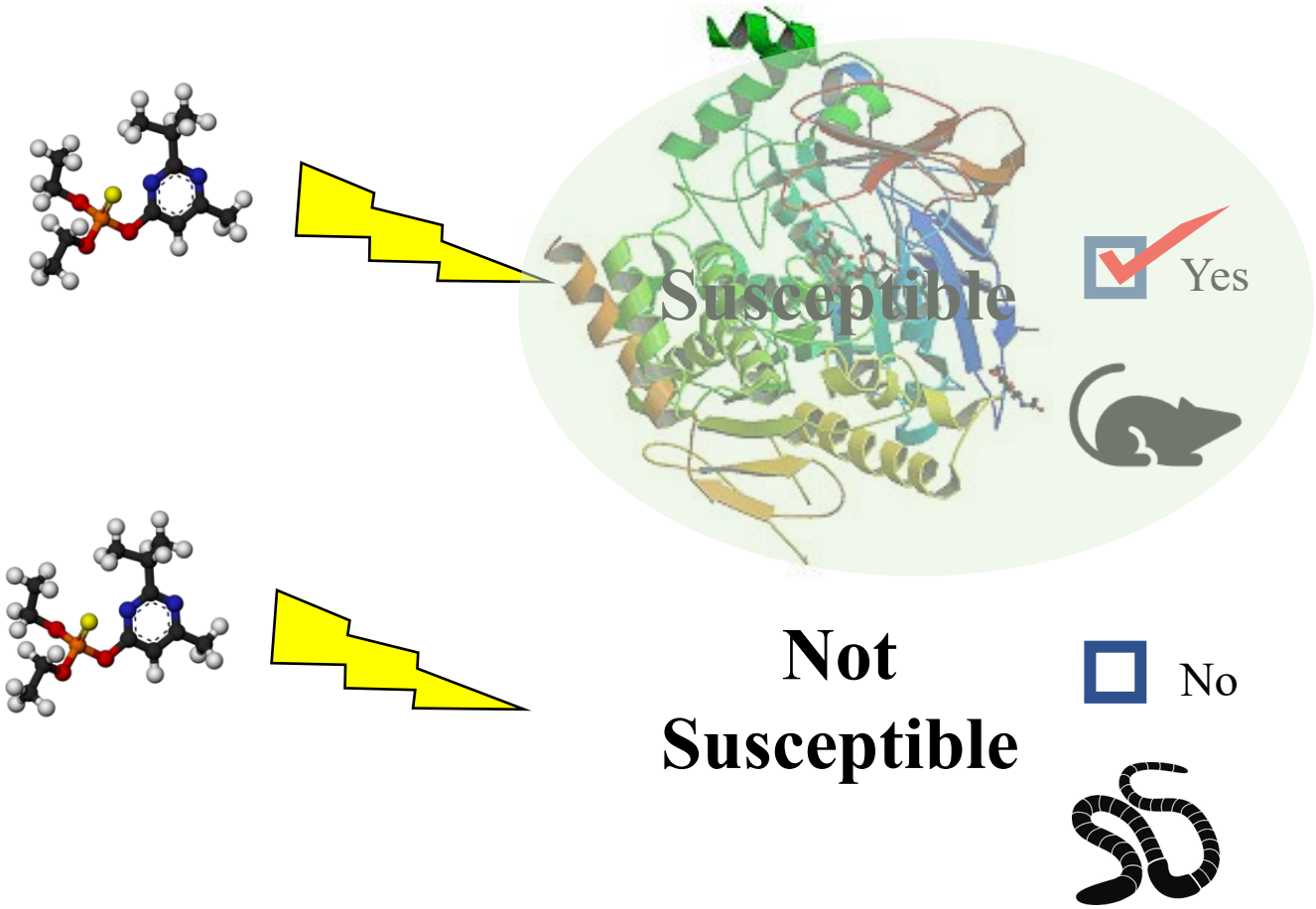


Simplify Complexity

Considering chemical sensitivity?

Factors that make a species sensitive

- Exposure
- Dose
- ADME
- Target receptor availability
- Life stage
- Life history
- etc.
- etc.



Simple question to address:

Is the known chemical target available in a species for a chemical to act upon?

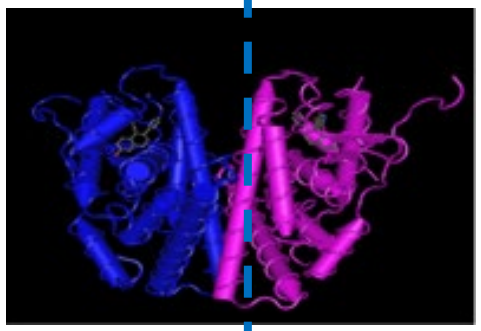
Yes or No

Likely susceptible or Not likely susceptible (at least through the known mechanism)

Start simple

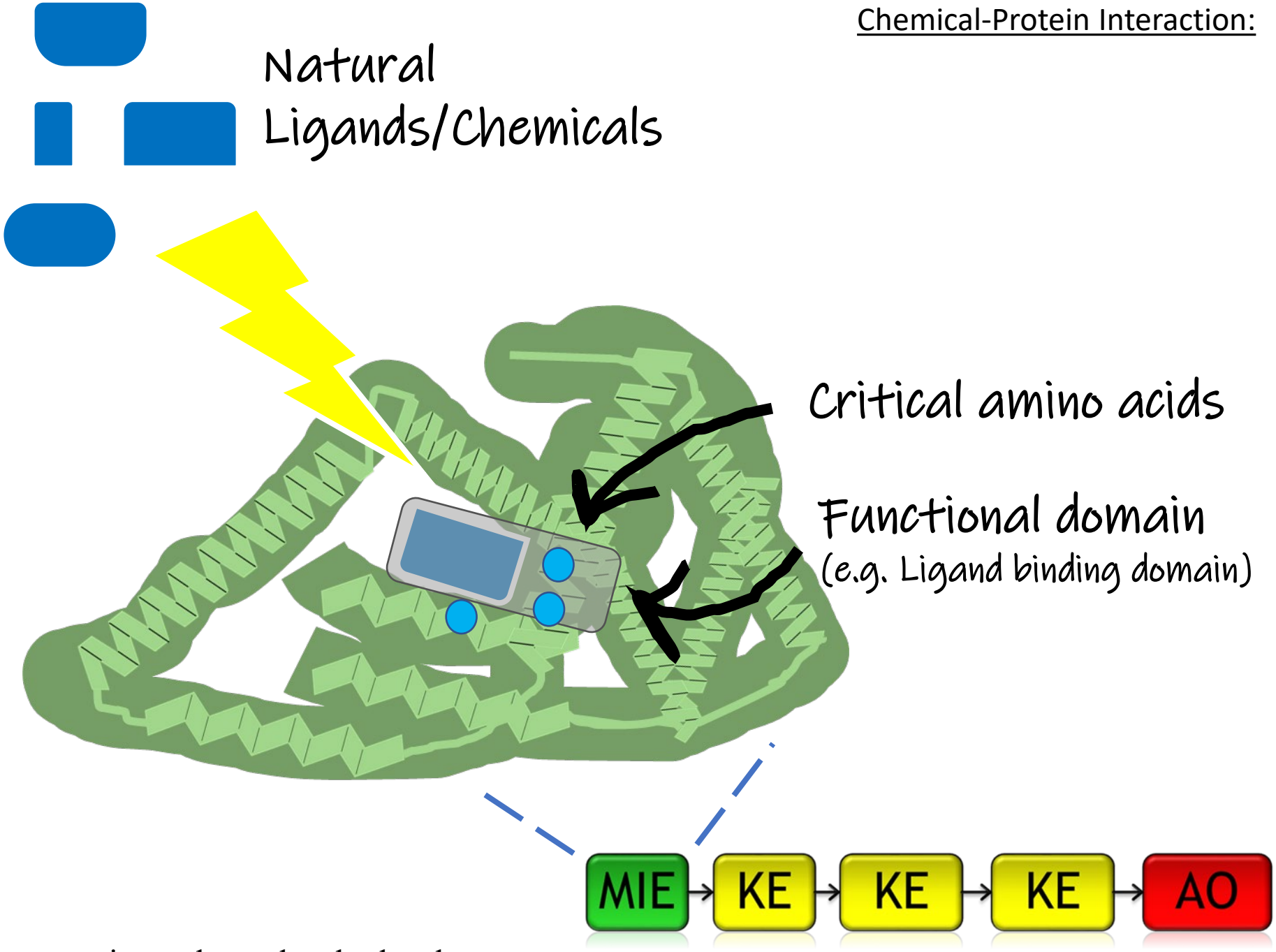
Sequence

MTMTLHTKASGMALLHQIQGNELEPLNRPQLKIPLERPLGE
VYLDSSKPAVYNYPEGAAYEFNAAAAANAQVYGQTGLPYG
PGSEAAAFGSNGLGGFPPLNSVSPSPLMLLHPPPQLSPFLQ
PHGQQVPYYLENEPSGYTVREAGPPAFYRPNSDNRRQGGR
ERLASTNDKGSMAMESAKETRYCAVCNDYASGYHYGVWSC
EGCKAFFKRSIQGHNDYMCPATNQCTIDKNRRKSCQACRLR
KCYEVGMMKGGIRKDRRGGRMLKHKRQRDDGEGRGEVG
SAGDMRAANLWPSPLMIKRSKKNSLALSLTADQMVSALLA
EPPILYSEYDPTRPFSEASMMGLLTNLADRELVHMINWAKV
PGFVDLTLHDQVHLLECAWLEILMIGLVWRSMEHPGKLLFA
PNLLLDRNQGKCVEGMVEIFDMLLATSSRFRMMNLQGEEF
VCLKSIILLNSGVYTFLSSTLKSLEEKDHIHRVLDKITDTLIHLM



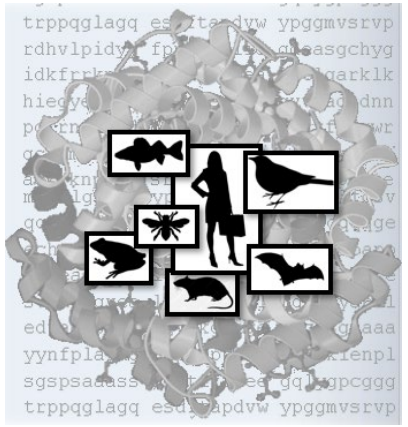
Bioinformatics

Chemical-Protein Interaction:



Similarity across species at the molecular level

EPA United States Environmental Protection Agency



TOXICOLOGICAL SCIENCES, 153(2), 2016, 228–245

doi: 10.1093/toxsci/kfw119
Advance Access Publication Date: June 30, 2016
Research article

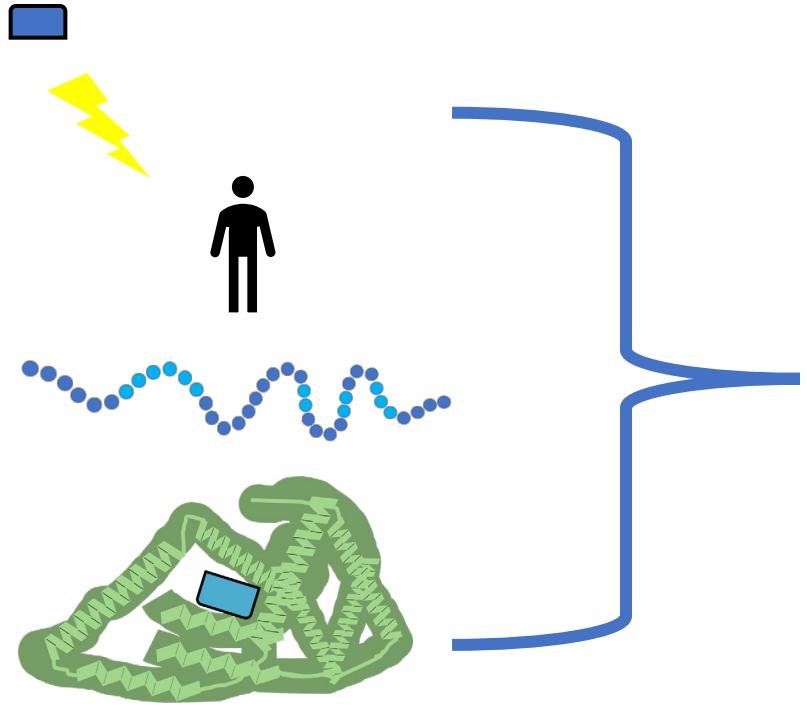
OXFORD SOT Society of Toxicology
www.toxsci.oxfordjournals.org

https://seqapass.epa.gov/seqapass/

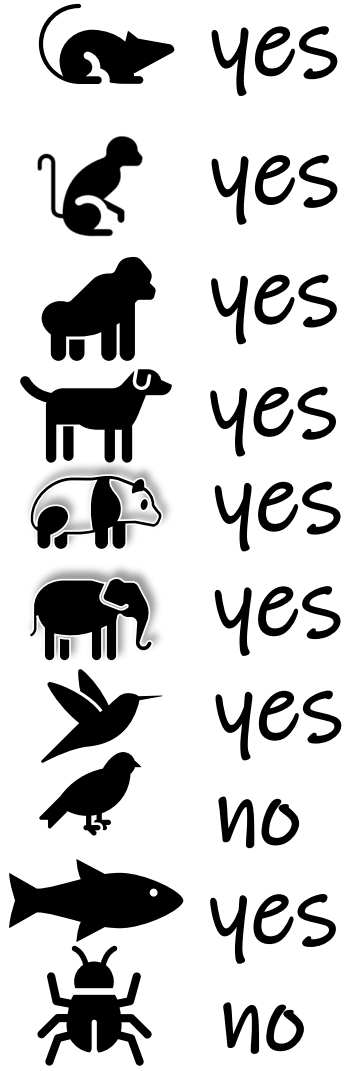
Sequence Alignment to Predict Across Species Susceptibility (SeqAPASS)

Carlie A. LaLone,*,[1] Daniel L. Villeneuve,* David Lyons,† Henry W. Helgen,‡ Serina L. Robinson,§,[2] Joseph A. Swintek,¶ Travis W. Saari,* and Gerald T. Ankley*

SeqAPASS Predicts Likelihood of Similar Susceptibility based on Sequence Conservation:



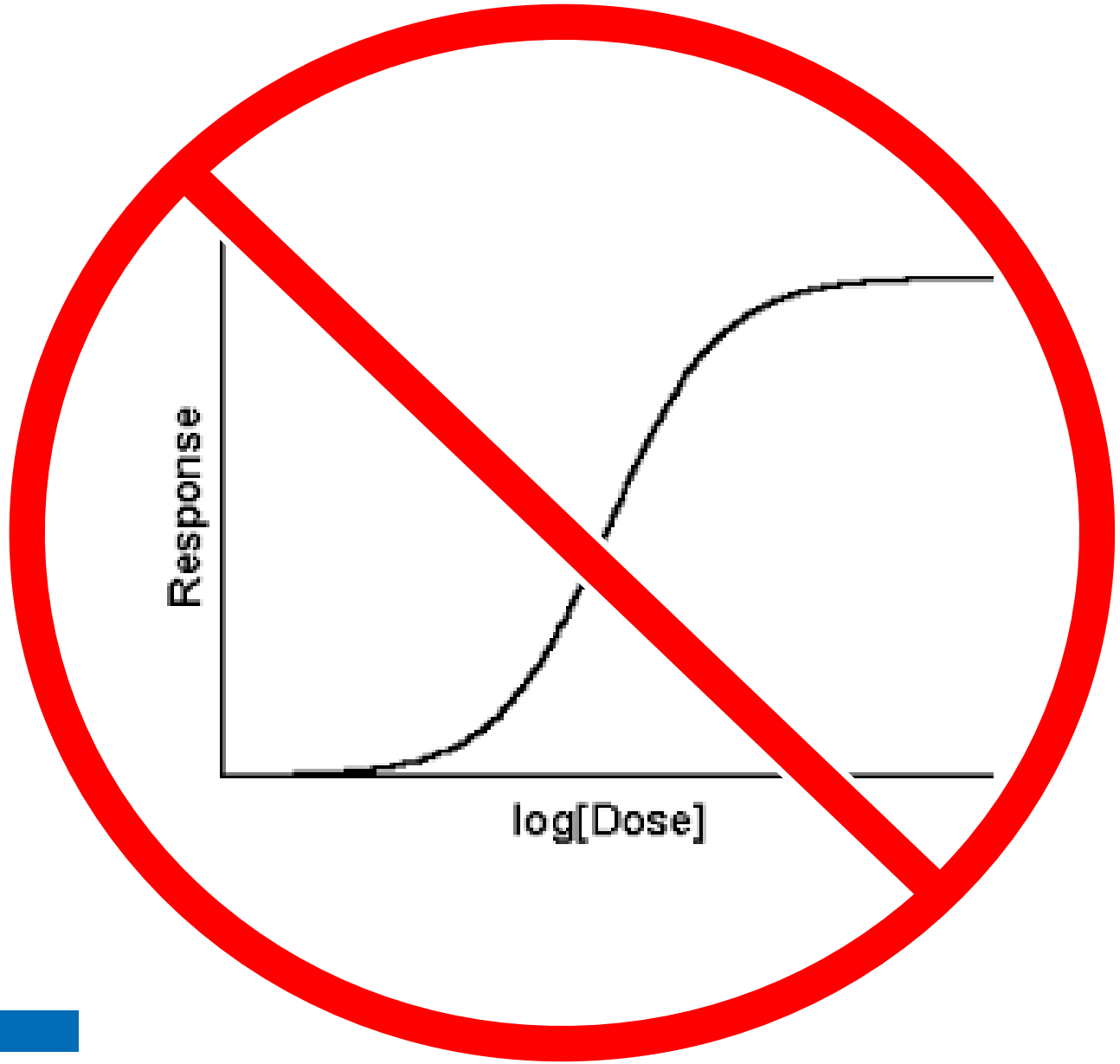
yes



Line(s) of evidence indicate

- The protein is conserved
- The protein is NOT conserved

SeqAPASS DOES NOT predict the degree of sensitivity/susceptibility:



Factors that make a species sensitive

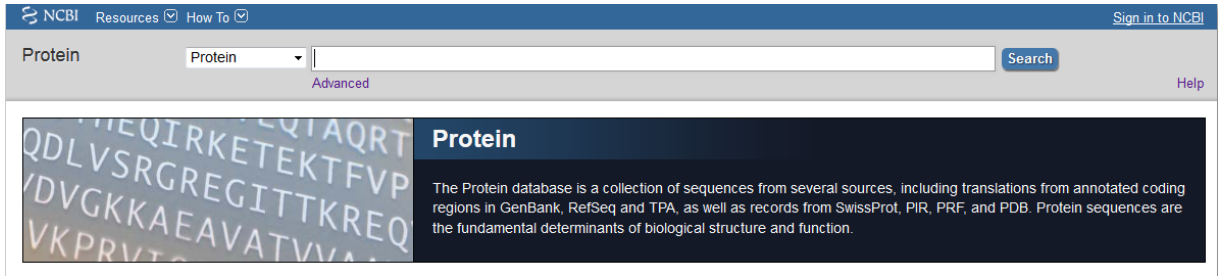
- Exposure
- Dose
- ADME
- Target receptor availability
- Life stage
- Life history
- etc.
- etc.

Strengths of SeqAPASS

New tools and technologies

- Improved sequencing technologies
- Large databases of sequence data

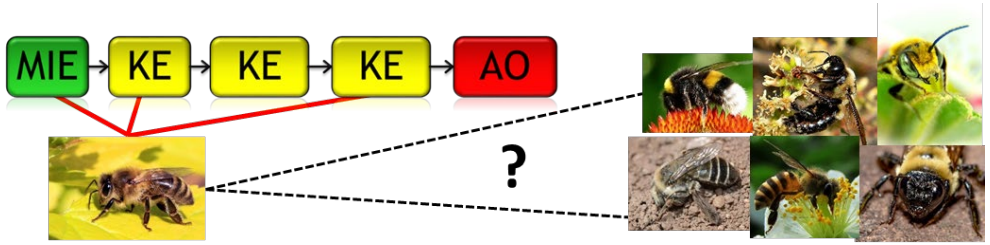
NCBI: 197,232,209 Proteins representing 108,257 Organisms



- Publicly available to all
- Lines of evidence for conservation for 100s-1000s of species rapidly
- Takes advantage of well-established tools and databases
- Streamlined, consistent, transparent, and published methods
  - Case examples to demonstrate applications
- Guides users to appropriate input
- Evolves as bioinformatics approaches become more user friendly
  - Smart automation or semi-automation

Applications of Bioinformatics: Case Studies

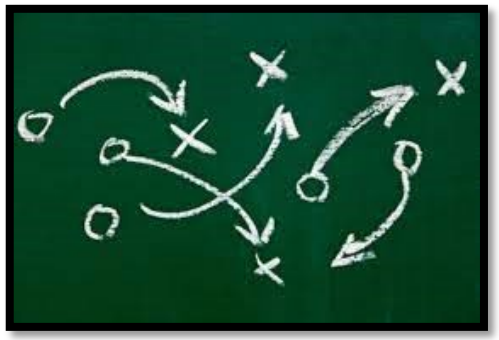
- **Extrapolate adverse outcome pathway knowledge across species**
  - Define the taxonomic relevance: Apis vs Non-Apis bees



- **Extrapolate high throughput screening data**
  - Chemicals that target human estrogen receptor alpha, androgen receptor, steroidogenic enzymes, thyroid axis proteins
  - All ToxCast Assay targets
- **Predict relative intrinsic susceptibility**
  - Pesticides
  - Endangered Species Act
  - Derivation of Aquatic Life Criteria
- **Predict chemical bioaccumulation across species**
  - Chemicals of concern: PFAS
- **Generate research hypotheses** Strobilurin fungicides
- **Prioritization strategies** Pharmaceuticals

Where do we go from here?

- Connect bioinformatics with general informatics including systematic methods
  - ECOTOX Knowledgebase
  - Systematic literature review for WOE
- Advance structural evaluations (computing power and storage)
  - Move from docking to virtual screening
- Specific laboratory studies to support bioinformatics
  - Site-directed mutagenesis



Government

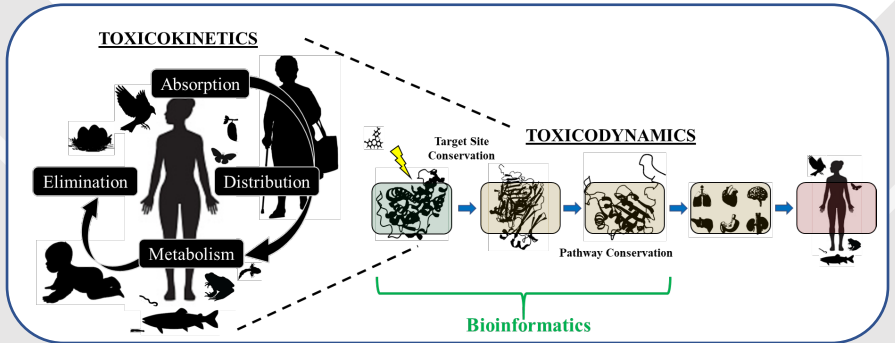
Industry

International Consortium to Advance Cross Species Extrapolation in Regulation ICACSER

Steering Committee:
Carlie LaLone (US EPA)
Geoff Hodges (Unilever)
Nil Basu (McGill U)
Steve Edwards (RTI)
Fiona Sewell (NC3Rs)
Michelle Embry (HESI)
Patience Browne (OECD)

1. Define the taxonomic domain of applicability
2. Define the global regulatory landscape/need
3. Develop a bioinformatics toolbox
4. Communicate a shared scientific vision

Interested in Learning more or Joining: Contact LaLone.Carlie@epa.gov or Geoff.Hodges@unilever.com

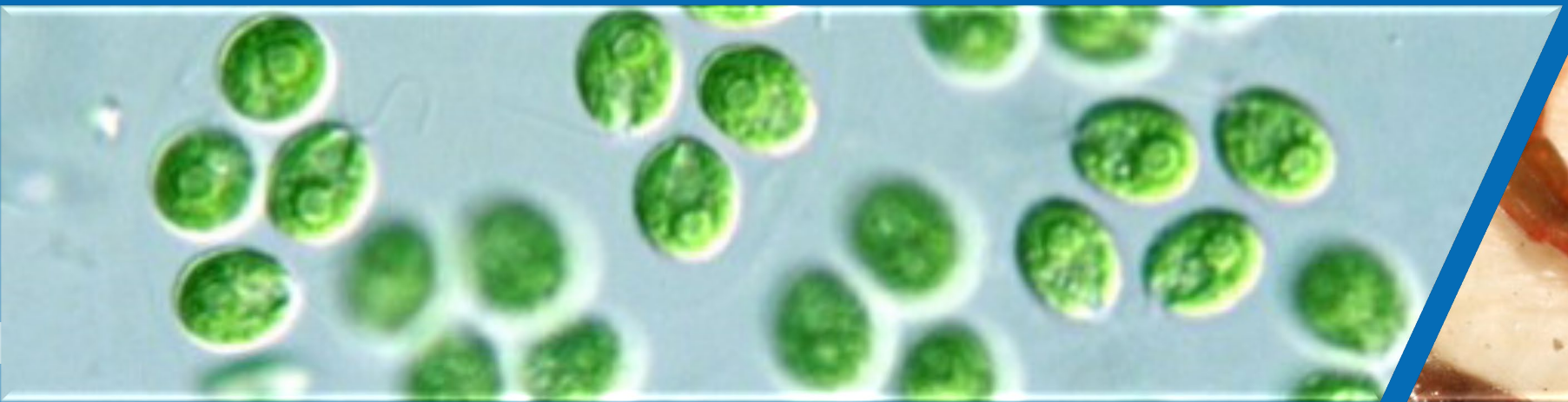


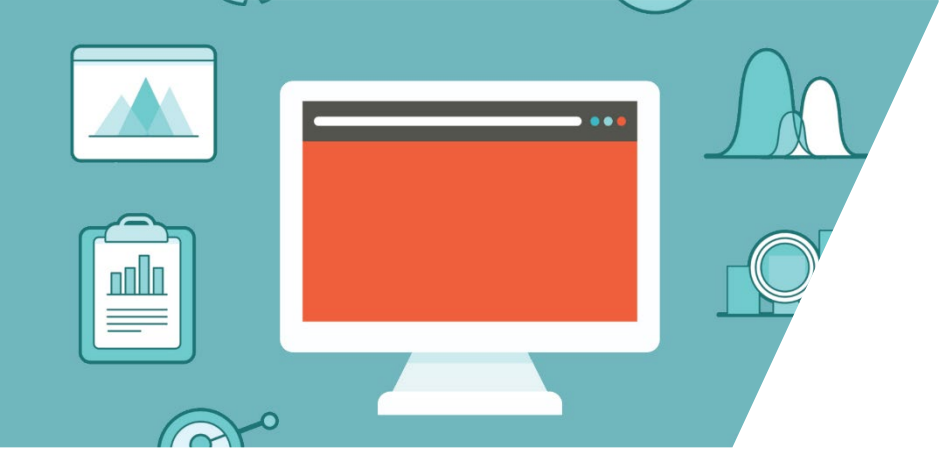
Academia

NGO

EPA
United States
Environmental Protection
Agency

Other ORD Research: High throughput transcriptomics for Ecotoxicology

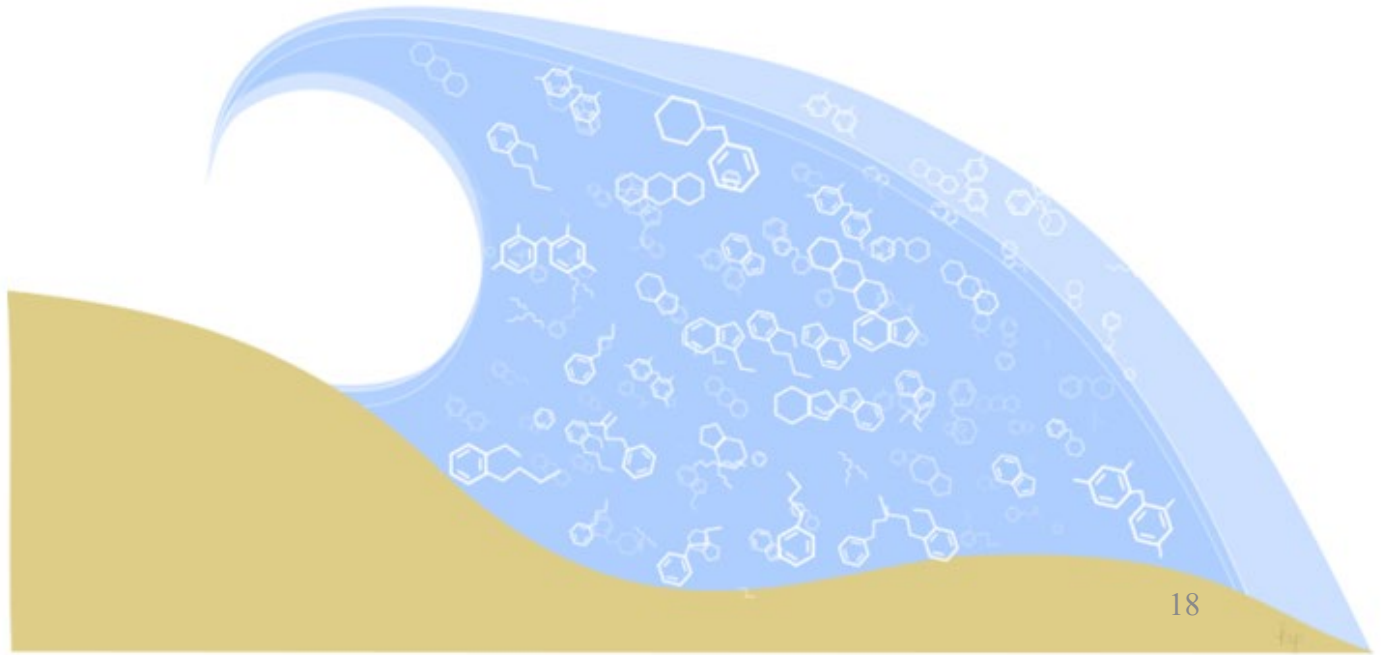




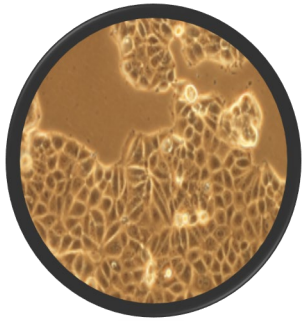
Problem Statement:

Tens of thousands of chemicals are currently in use and hundreds more are introduced to the market every year. Only a small fraction has been thoroughly evaluated for potential risks to human health and the environment.

- Need for high-throughput assays to evaluate chemical safety.
- Provide adequate coverage of relevant pathways through which chemicals can cause toxicity.



Toxicogenomic Approach



Whole human transcriptome

NTP
National Toxicology Program
U.S. Department of Health and Human Services

NTP Research Report on
National Toxicology Program Approach to Genomic Dose-Response Modeling

NTP RR 5

APRIL 2018

- Number of mammalian studies have shown short-term transcriptomics-based PODs are predictive of apical potency.
- Generally, within ½ log.
- Health protective points of departure.

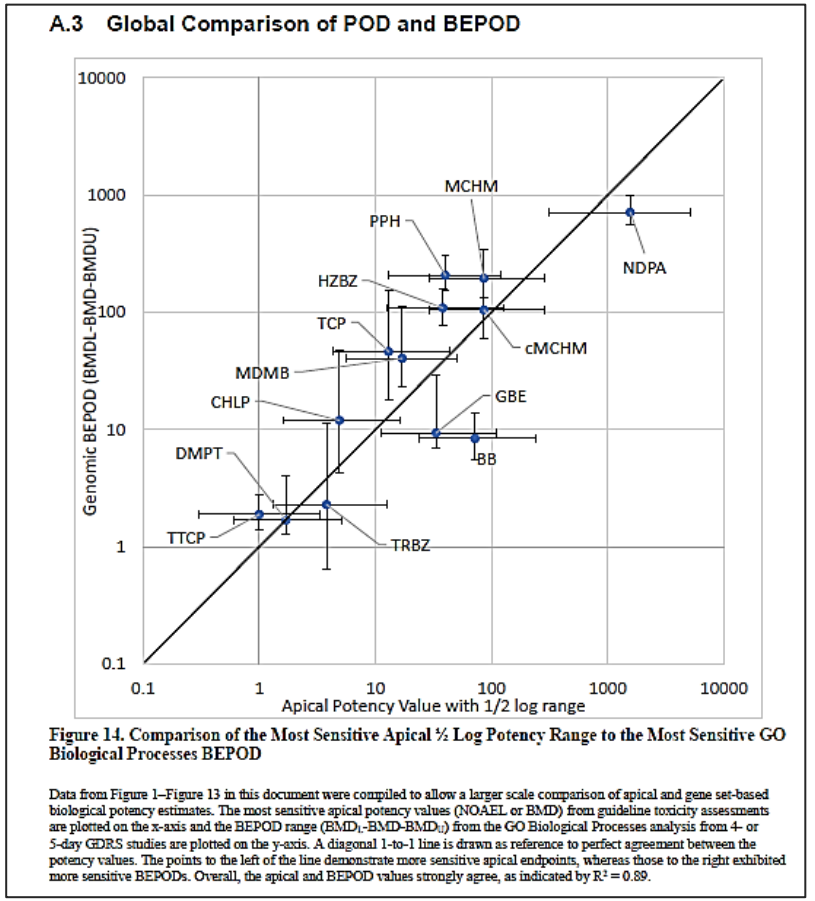


Figure 14. Comparison of the Most Sensitive Apical ½ Log Potency Range to the Most Sensitive GO Biological Processes BEPOD

Data from Figure 1–Figure 13 in this document were compiled to allow a larger scale comparison of apical and gene set-based biological potency estimates. The most sensitive apical potency values (NOAEL or BMD) from guideline toxicity assessments are plotted on the x-axis and the BEPOD range (BMD_L-BMD-BMD_U) from the GO Biological Processes analysis from 4- or 5-day GDRS studies are plotted on the y-axis. A diagonal 1-to-1 line is drawn as reference to perfect agreement between the potency values. The points to the left of the line demonstrate more sensitive apical endpoints, whereas those to the right exhibited more sensitive BEPODs. Overall, the apical and BEPOD values strongly agree, as indicated by $R^2 = 0.89$.

OXFORD SOT Society of Toxicology
academic.oup.com/toxsci

TOXICOLOGICAL SCIENCES, 181(1), 2021, 68–89

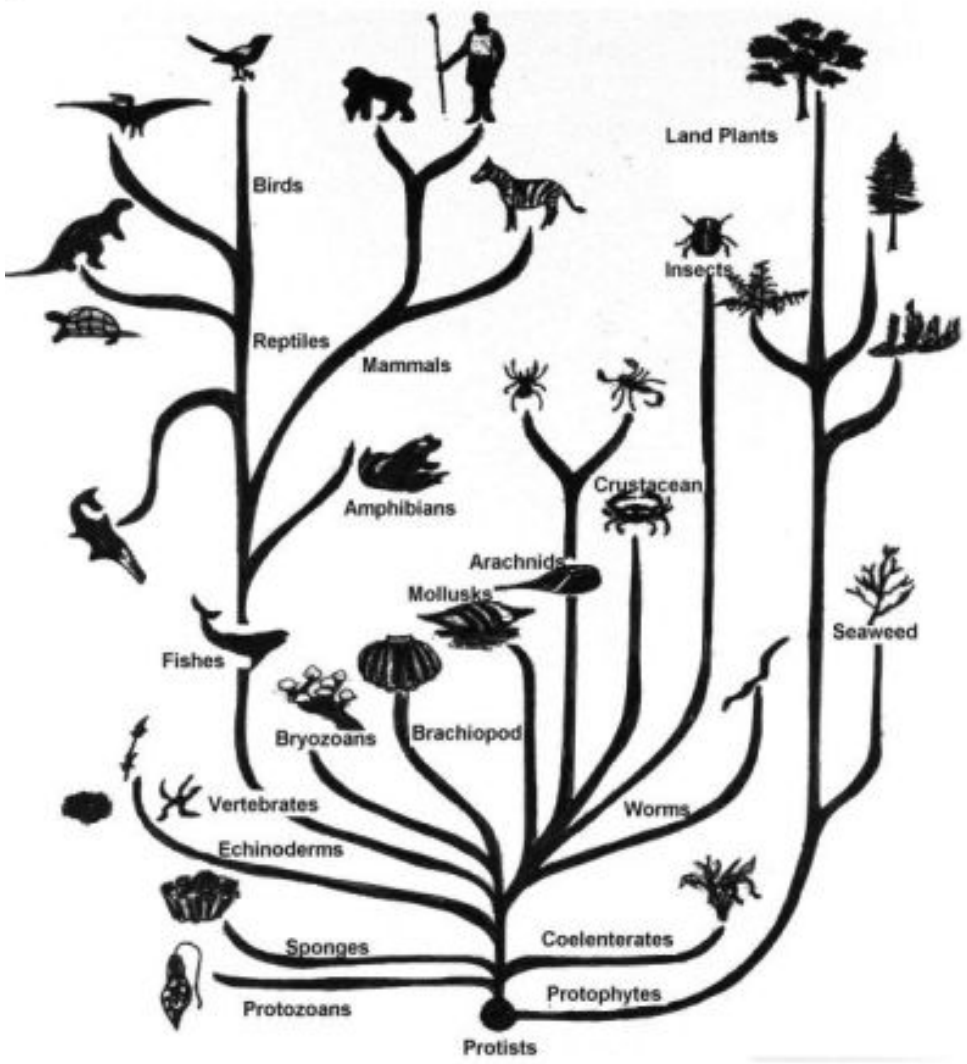
doi: 10.1093/toxsci/kfab009
Advance Access Publication Date: 4 February 2021
Research Article

High-Throughput Transcriptomics Platform for Screening Environmental Chemicals

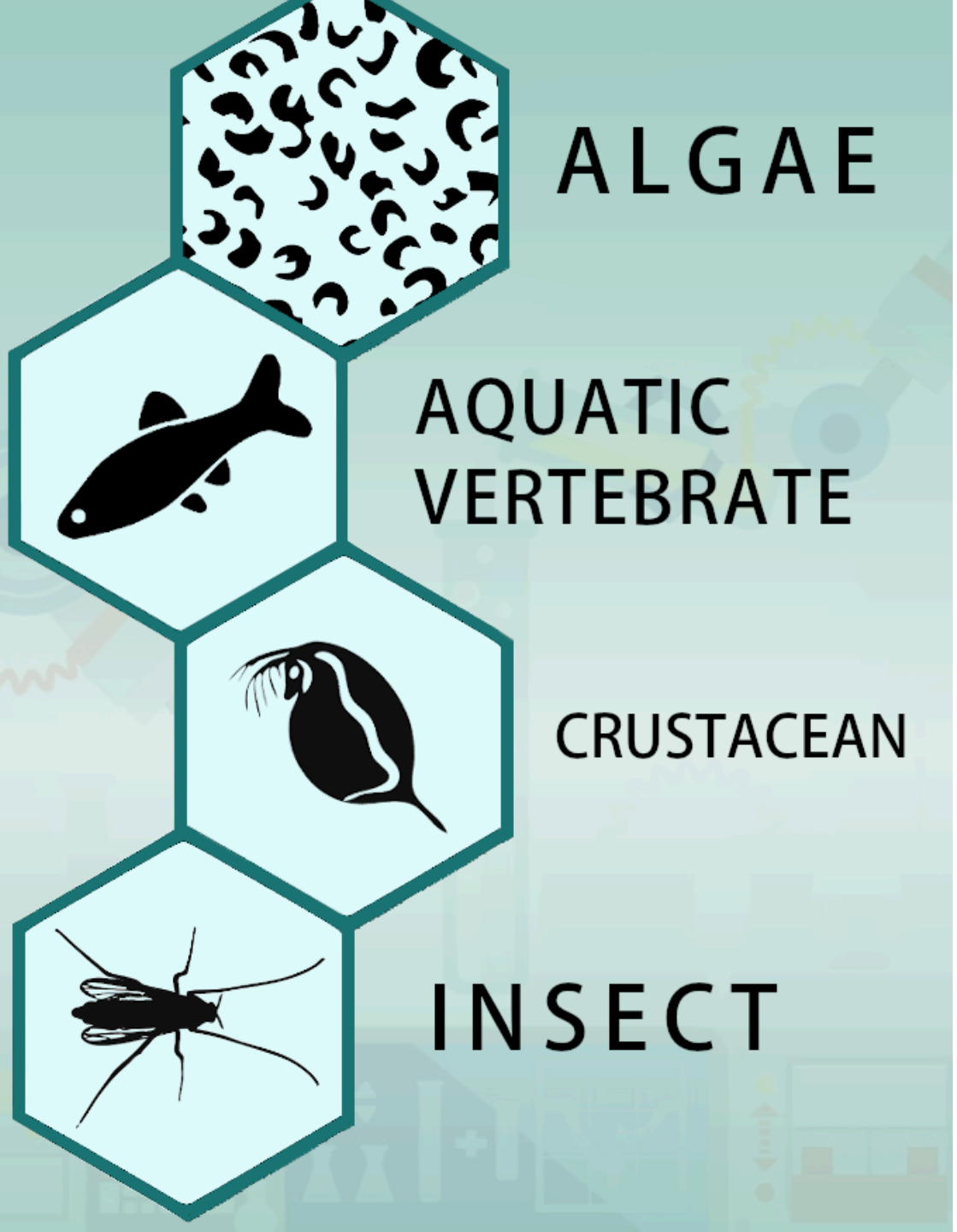
Joshua A. Harrill,*,1 Logan J. Everett,* Derik E. Haggard,*,† Thomas Sheffield,*,† Joseph L. Bundy,* Clinton M. Willis,*,‡ Russell S. Thomas,* Imran Shah,* and Richard S. Judson*

Toxicological Sciences, Volume 181, Issue 1, May 2021, Pages 68–89, https://doi.org/10.1093/toxsci/kfab009

Ecotoxicology Perspective



- Humans are just a tiny fraction of the biological diversity we are charged to protect.
- Many genes/pathways are conserved
- Unique physiology in other kingdoms, phyla, classes…
- How do we assure those pathways are covered?



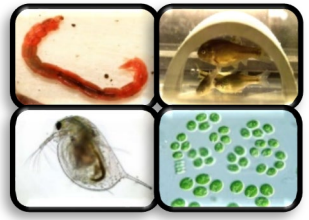
High throughput assays for three major trophic levels of aquatic ecosystems

- Primary producers (e.g., algae)
- Primary consumers (e.g., zooplankton, aquatic inverts)
- Secondary consumers (e.g., fish)

Commonly used for GHS classification and labeling of chemicals for environmental hazard

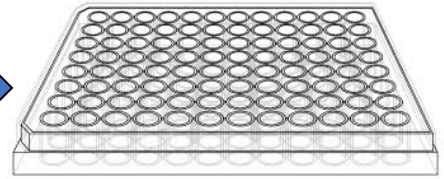
Aquatic organisms highly vulnerable to exposure

Incorporating transcriptomics as assessment endpoint



Load organisms

24 h exposure

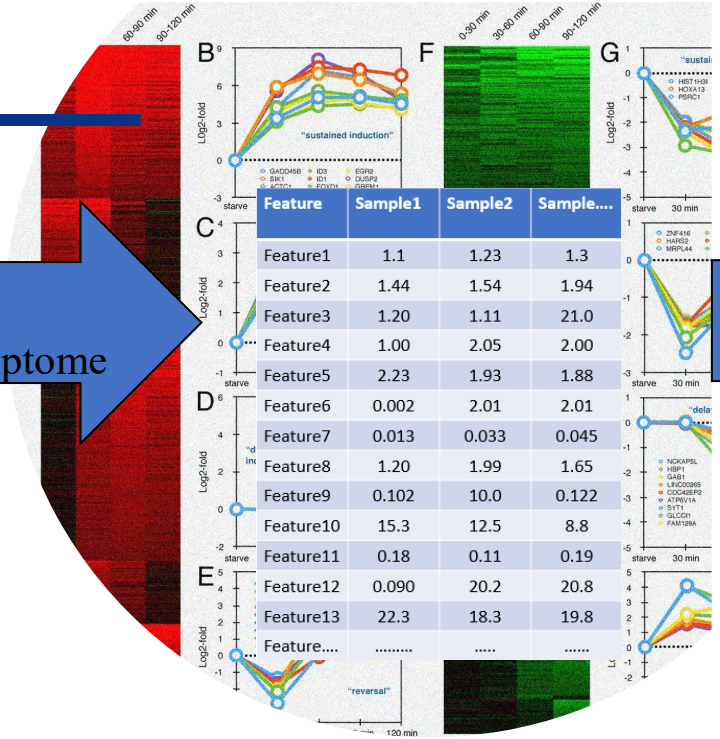


Phenotypic anchoring

Homogenate /extract

RNA

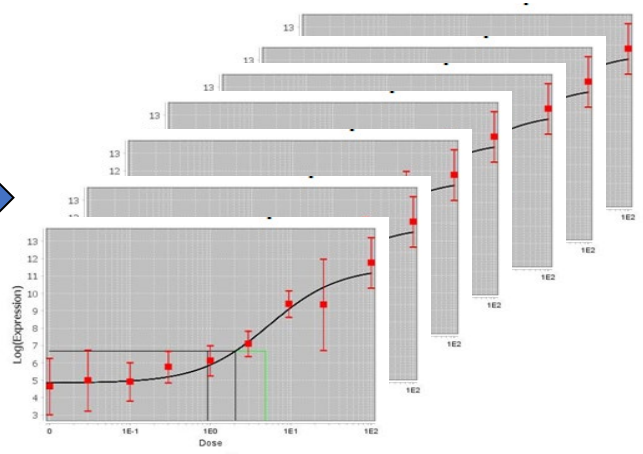
Whole transcriptome



Feature	Sample1	Sample2	Sample....
Feature1	1.1	1.23	1.3
Feature2	1.44	1.54	1.94
Feature3	1.20	1.11	21.0
Feature4	1.00	2.05	2.00
Feature5	2.23	1.93	1.88
Feature6	0.002	2.01	2.01
Feature7	0.013	0.033	0.045
Feature8	1.20	1.99	1.65
Feature9	0.102	10.0	0.122
Feature10	15.3	12.5	8.8
Feature11	0.18	0.11	0.19
Feature12	0.090	20.2	20.8
Feature13	22.3	18.3	19.8
Feature....	.....	.....	.....

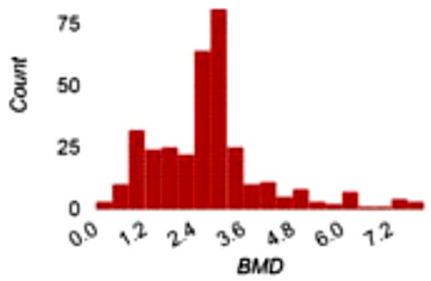
Curv

ve fitting



BMD Express

Distribution of BMDs



10th centile BMD

Point of departure

[tPOD]

Rough estimate per chemical ~$5K for one species

Evaluating the approach

[tPOD] ≤ [Most sensitive chronic endpoint]

[tPOD] > [Most sensitive chronic endpoint]

[tPOD] <<<< [Most sensitive chronic endpoint]

Approach is being explored with cell-lines for eco species as well

APCRA Case study: Transcriptomics-based PODs for Ecotoxicology

1. Generate transcriptomic PODs for ≈ 20 chemicals
   - Initial focus on fathead minnow
   - Parallel assays with additional taxa – for future analyses

2. Compare tPODs with available acute and chronic toxic toxicity data

3. Compare tPODs with in vitro-derived PODs

Acknowledgements

U.S. EPA, ORD

Dan Villeneuve
Kevin Flynn
Adam Biales
Donovan Blatz (ORISE)
Sara Vliet (ORISE)
Sally Mayasich (ORISE)
Marissa Jensen (Univ. Minnesota Duluth)

GDIT

Thomas Transue
Cody Simmons
Audrey Wilkinson

University of Pittsburgh

Carla Ng
Weixiao Cheng

Newly Released: SeqAPASS v6.0

LaLone.Carlie@epa.gov

https://seqapass.epa.gov/seqapass/