Gathering Evidence of Endocrine Pathway Conservation for Cross-Species Extrapolation Using New Approach Methods

Sara M. Vliet, Ph.D., Scott G. Lynn, Kristan Markey, Carlie A. LaLone

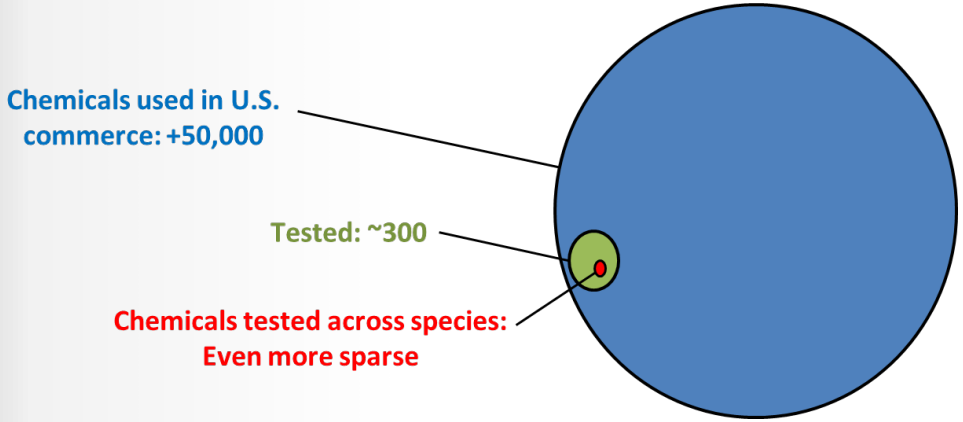
ORISE Postdoctoral Fellow
Great Lakes Toxicology & Ecology Division
Office of Research and Development &
Office of Science Coordination and Policy
U.S Environmental Protection Agency

SETAC North America 41st Annual Meeting- SciCon2
November 15th- 19th, 2020

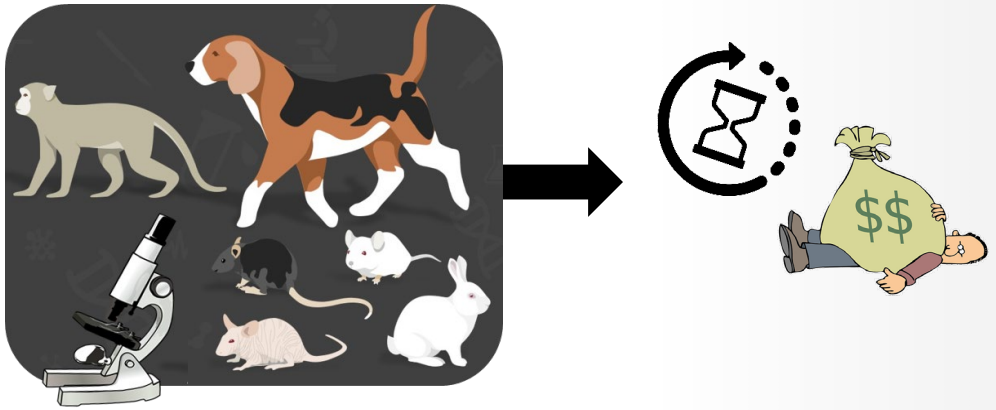
The views expressed in this presentation are those of the author and do not necessarily reflect the views or policies of the US EPA.

The Need for New Approach Methods (NAMs)

- EPA & the endocrine disruptor screening program (EDSP) is tasked with evaluating thousands of chemicals for potential endocrine bioactivity to protect **human health and wildlife**

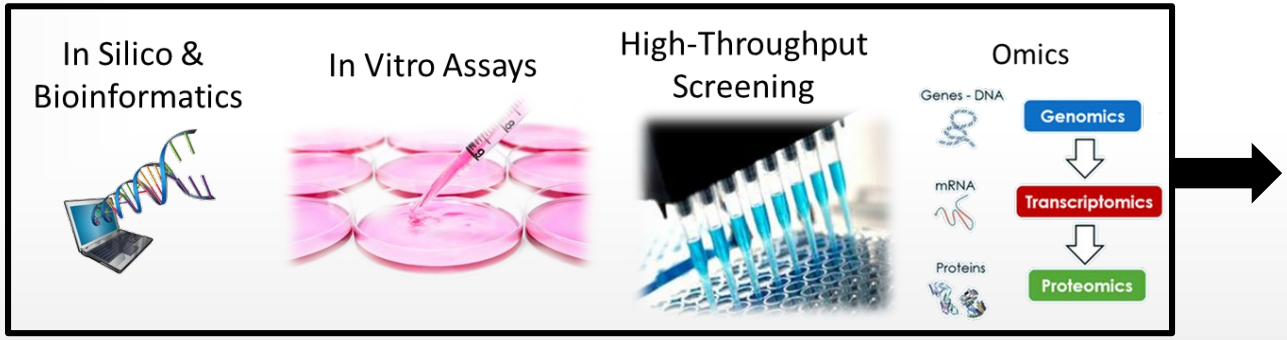


Whole Animal Testing Methods



- Limited data for many compounds, limited resources for traditional toxicity testing, and international efforts to reduce animal use all necessitate the development of **new approach methods (NAMs)**

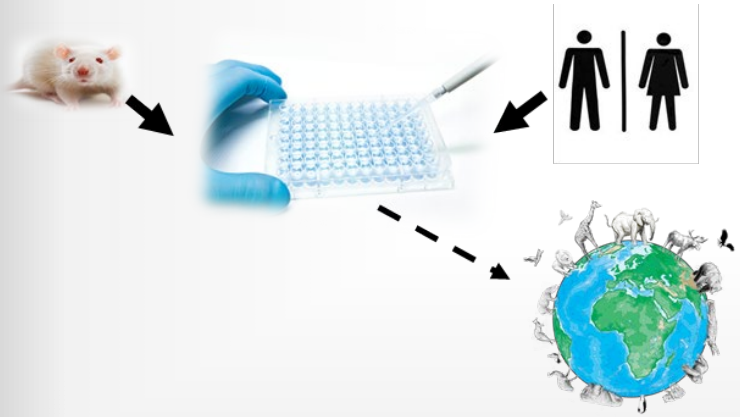
New Approach Methods (NAMs)



Surrogate Species in Toxicity Testing

- In whole animal testing, it is assumed that the sensitivity of species to a chemical is a function of their relatedness

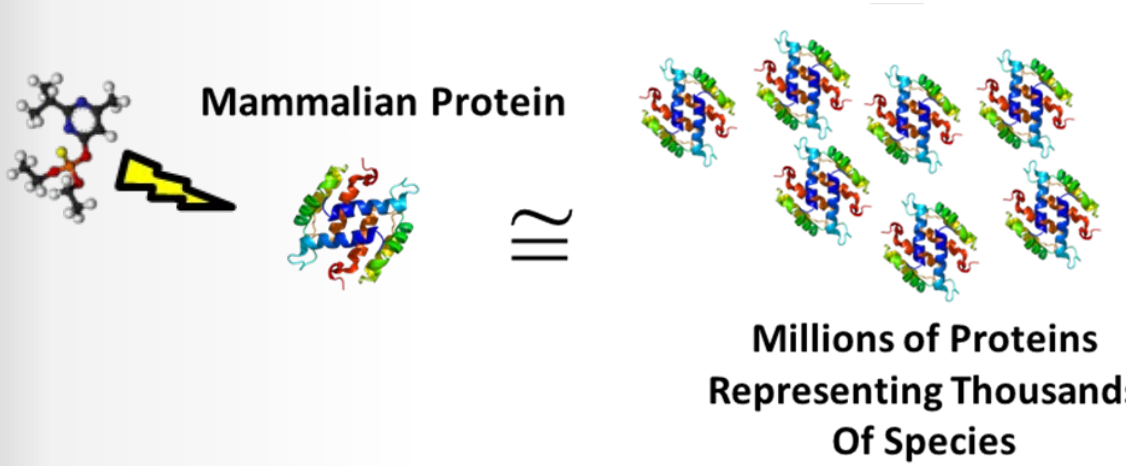
- High throughput screening assays (US EPA ToxCast) rapidly test chemicals, identify those most likely to be endocrine disruptors, and help inform putative molecular targets for chemicals using **mammalian cells**
- Knowledge of the molecular target be linked to an adverse outcome of regulatory concern



- Extrapolation from mammalian species to other species of concern is essential
- **Can we expect chemicals that interact with mammalian receptors to also interact with receptors of other species we want to protect?**

SeqAPASS: Sequence Alignment to Predict Across Species Susceptibility

- Online, publicly available tool for understanding **target conservation** across thousands of diverse species
- Facilitates rapid and quantitative assessment of protein similarity and provides a foundation for predicting the taxonomic domain of applicability
- Developed with both researchers and risk assessors in mind



SeqAPASS Applications

- Extrapolate high throughput screening data
- Extrapolate biological pathway knowledge across species
- Predict relative intrinsic susceptibility
- Generate research hypotheses
- Prioritize testing efforts

Sequence Alignment to Predict Across Species Susceptibility (SeqAPASS): A Web-Based Tool for Addressing the Challenges of Cross-Species Extrapolation of Chemical Toxicity

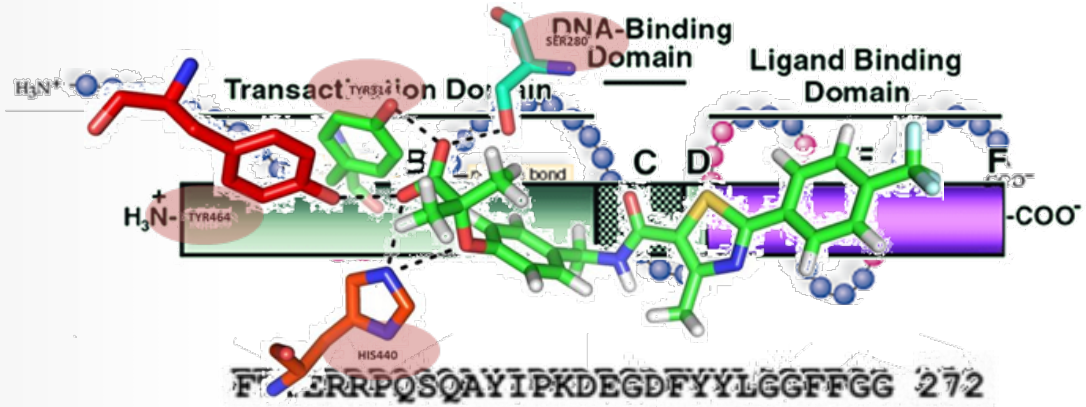
Carlie A. LaLone,[*,1] Daniel L. Villeneuve,[*] David Lyons,[†] Henry W. Helgen,[‡] Serina L. Robinson,[§,2] Joseph A. Swintek,[¶] Travis W. Saari,[*] and Gerald T. Ankley[*]

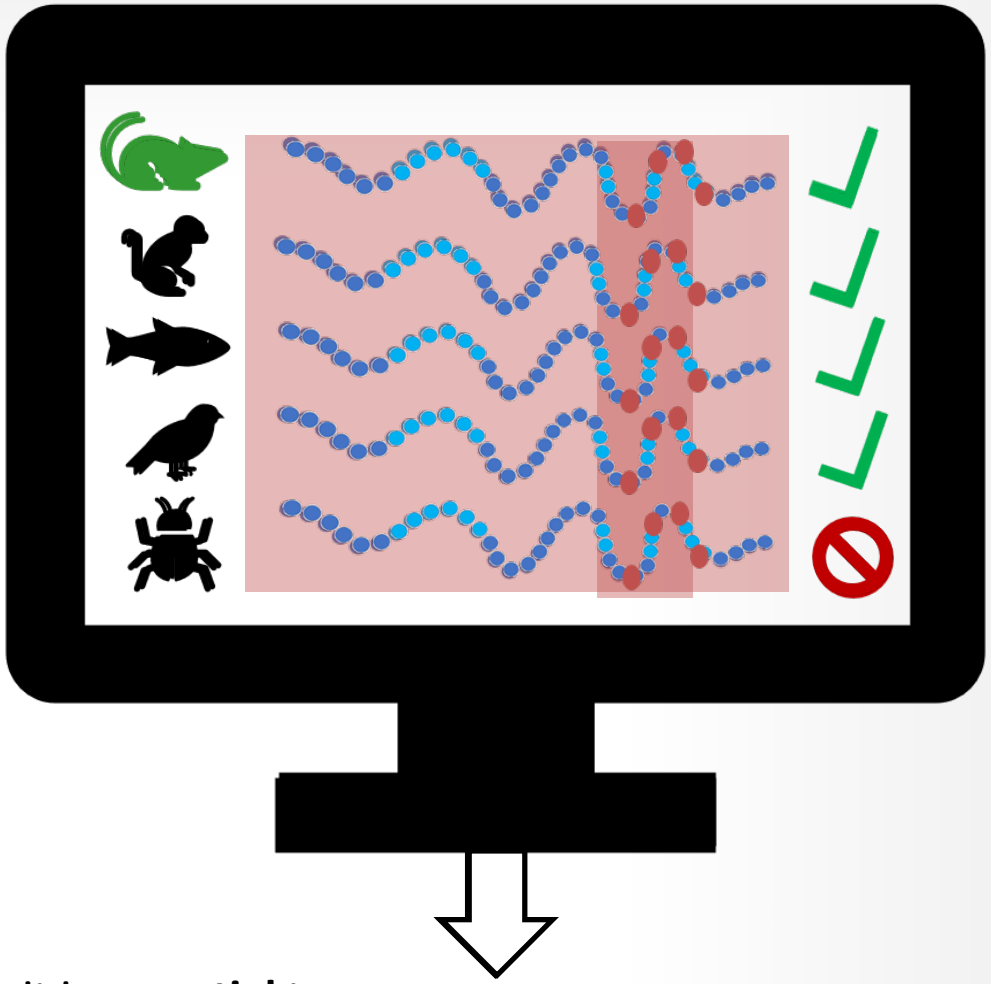
https://seqapass.epa.gov/seqapass/

SeqAPASS: The Basics

Flexible Analysis Based On Available Data

- **Level 1** Primary Amino Acid Sequence Alignments
- **Level 2** Conserved Functional Domain Alignments
- **Level 3** Critical (Close Contact) Amino Acid Conservation





For SeqAPASS to be used in a regulatory context and the EDSP pipeline, it is **essential** to understand how our computational predictions relate to empirical data across species

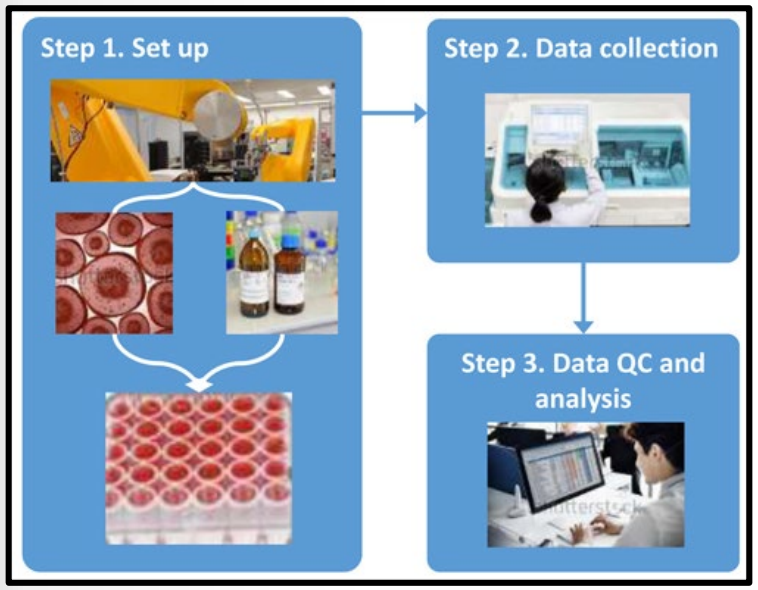
1. Evaluating **existing data** and **literature**
2. Conducting **in vitro** molecular biology studies
3. Conducting further **in silico** work (E.g. molecular docking, etc.)

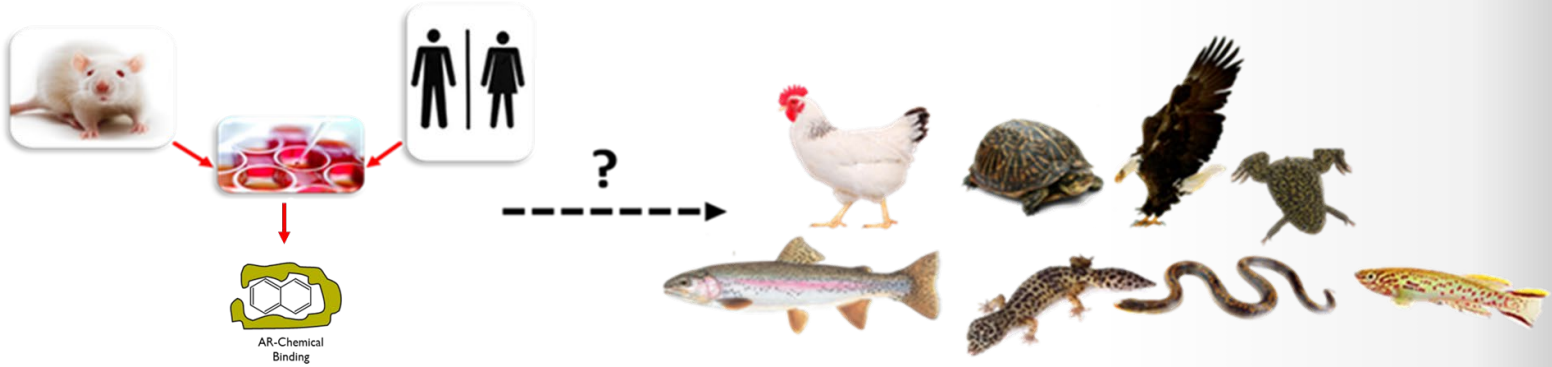
Case Study #1: Evaluating Existing Data to Extrapolate High-Throughput Androgen Receptor Screening Data Across Species

US EPA ToxCast Program: Uses mammalian cell-based assays to screen chemicals and identify putative molecular targets

US EPA Endocrine Disruptor Screening Program: Tasked with assessing thousands of chemicals for potential endocrine activity

- The androgen receptor (AR) is an important endocrine target for many environmental chemicals



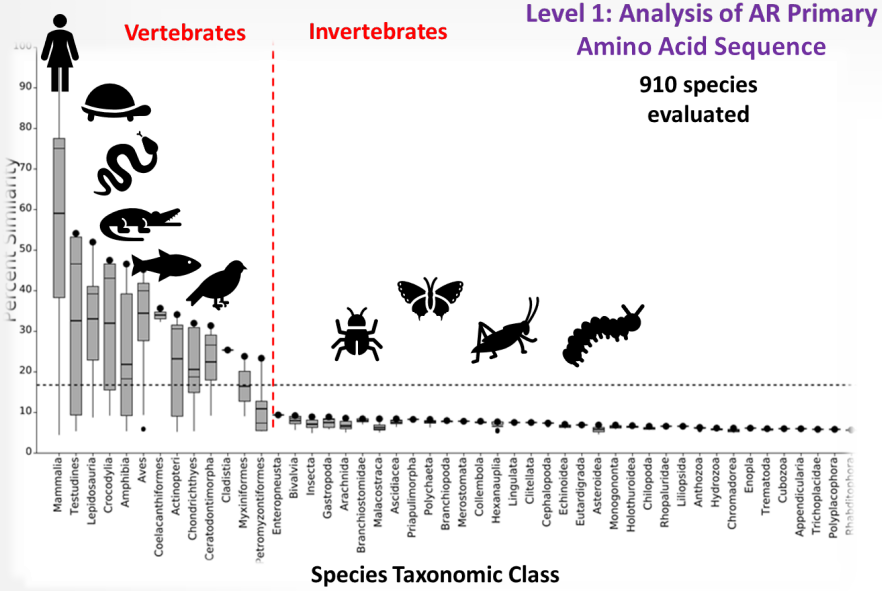


Guiding Question:
Can we expect chemicals that interact with AR in mammalian screening models to reflect potential toxicity across ecologically-relevant species?

Assessing AR Conservation Across Species Using the SeqAPASS Tool

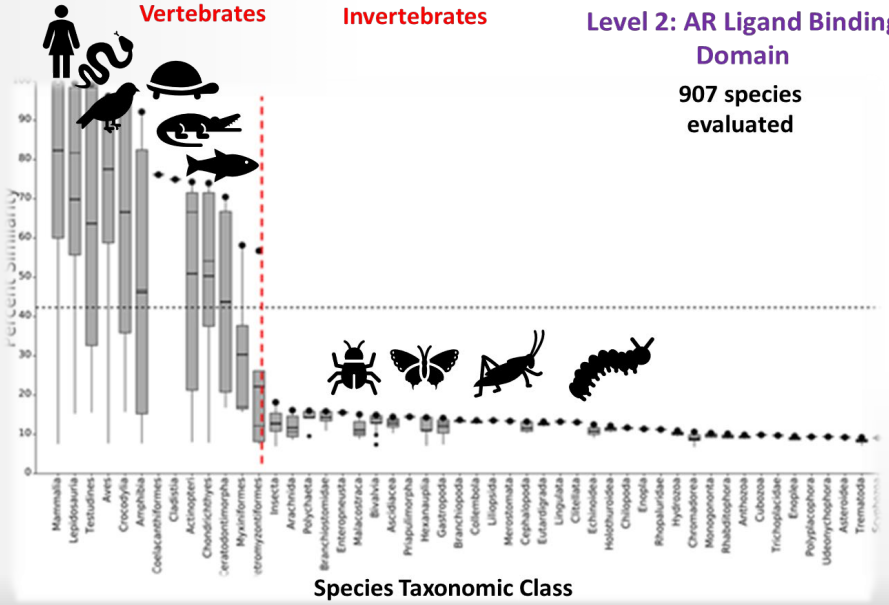
1. **Level 1: Analysis of AR Primary Amino Acid Sequence**

910 species evaluated



2. **Level 2: AR Ligand Binding Domain**

907 species evaluated



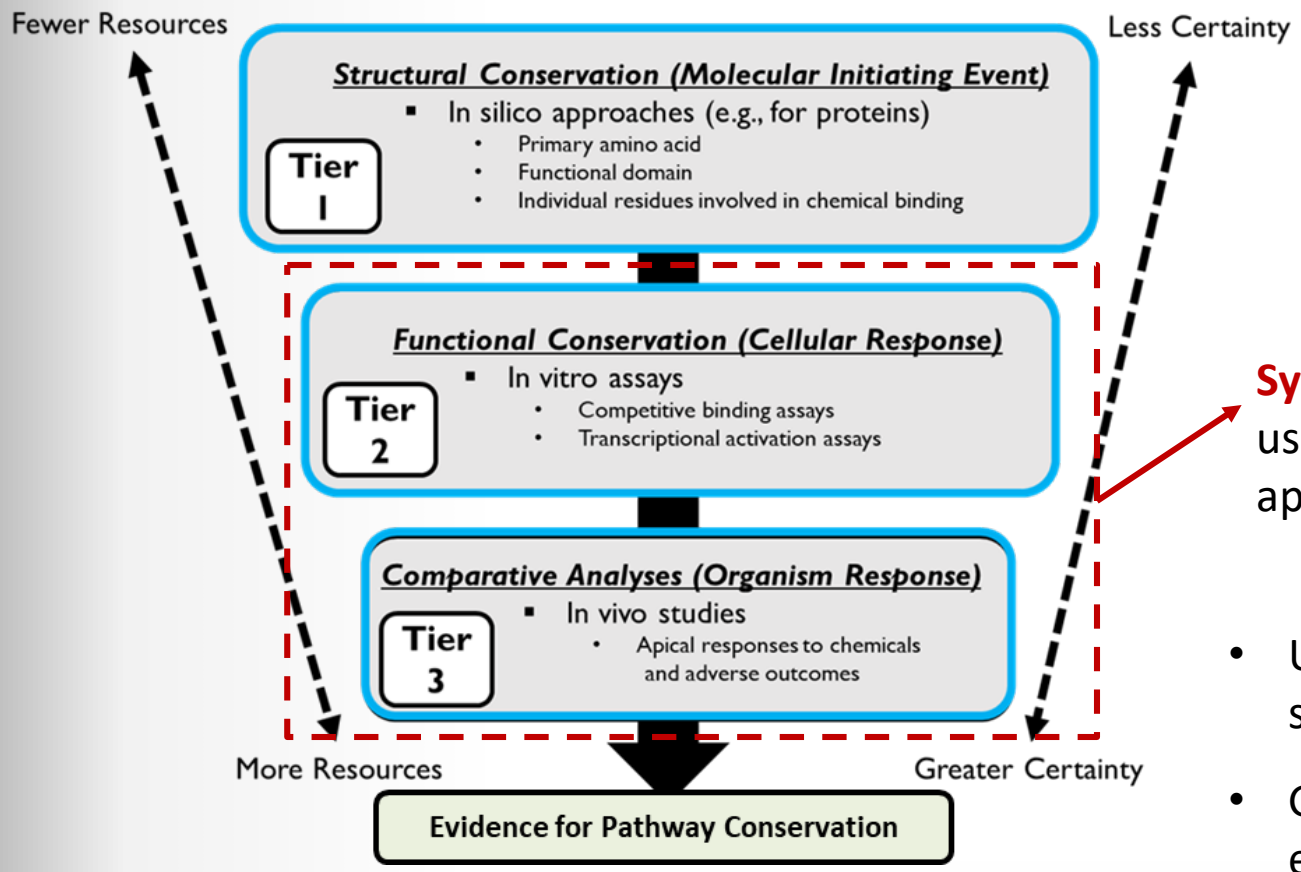
3. **Level 3: Analysis of Conservation of Individual Amino Acid Residues**

250 species evaluated

Taxonomic Group	# of Spp.	Shared Susceptibility
Mammals	117/1	Yes/No
Lizards, Snakes	11	Yes
Turtles	3	Yes
Birds	58	Yes
Crocodiles, Alligators	4	Yes
Amphibians	13	Yes
Coelacanths	2	Yes
Eel-shaped	1	Yes
Bony Fish	87/1	Yes/No
Sharks, Rays	4	Yes
Lungfish	2	Yes

- Across all three levels, SeqAPASS results suggest conservation of AR across vertebrate species
- Overall, these predictions suggest that chemicals that bind and activate AR in mammalian-based assays, are likely to interfere with AR in other vertebrate species
- Line of evidence for pathway conservation

Evaluating Existing Data to Extrapolate High-Throughput Androgen Receptor Screening Data Across Species



Environ Toxicol Chem. 2016 Nov;35(11):2806-2816. doi: 10.1002/etc.3456. Epub 2016 Jun 28.

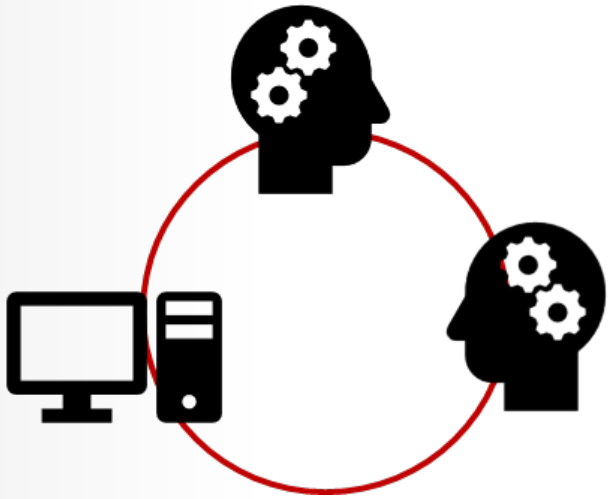
Evaluation of the scientific underpinnings for identifying estrogenic chemicals in nonmammalian taxa using mammalian test systems

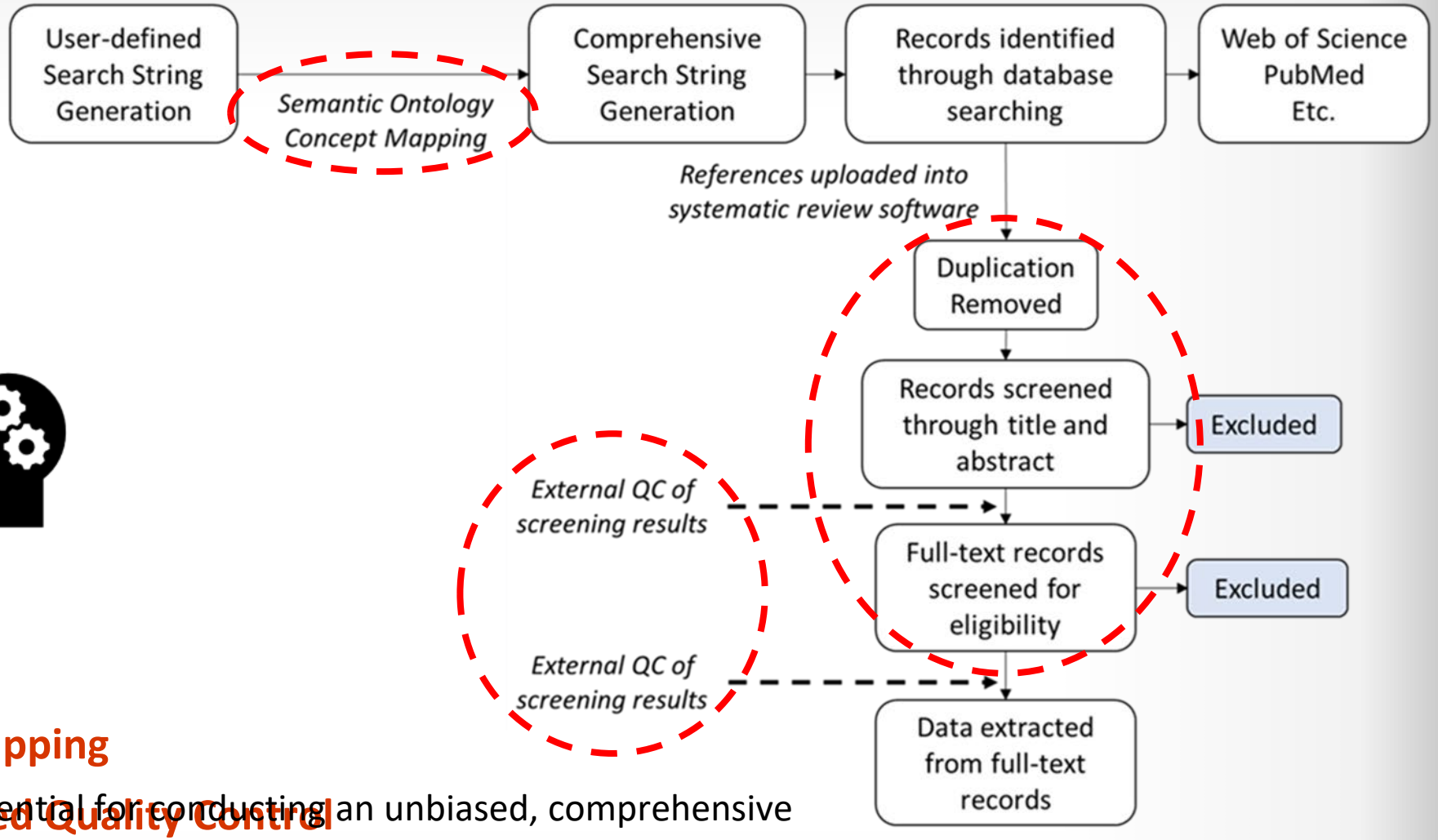
Gerald T Ankley [1], Carlie A LaLone [2], L Earl Gray [3], Daniel L Villeneuve [2], Michael W Hornung [2]

Systematic Literature Review: A type of literature review that uses systematic methods to collect secondary data, critically appraise research studies, and synthesize findings

- Using existing evidence (literature), we can evaluate the scientific basis of our cross-species predictions
- Gathering in vivo and in vitro data from vertebrate species exposed to known androgenic compounds provides **additional lines of evidence** for the conservation of the biological pathway across species

Incorporation of Technical Advances and Tools for Improved Systematic Review





Semantic Ontology Concept Mapping

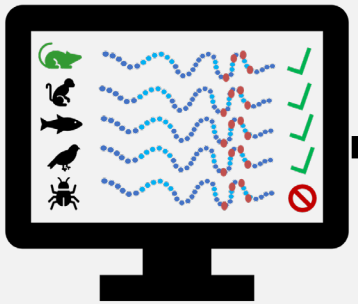
- Using accurate search terms is essential for conducting an unbiased, comprehensive survey of the literature landscape
- Scientific language is often non-standard, redundant, and mischaracterized
- Semantic ontology mapping approaches can develop comprehensive literature search strings by expanding vocabulary based on knowledge of related concepts

Machine Learning for Automated Quality Control in Systematic Review Software

- Development of machine learning models to screen through identified articles and identify possible errors (compared with external review by scientists)
- Collaborative web-based systematic literature review software provides a platform for unbiased article evaluation and data collection
- e.g. Does the "computer" agree with the human that this article is relevant/irrelevant
- Platform facilitates Quality Assurance/ Quality Control, leading to transparent and reproducible reviews

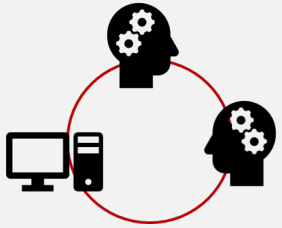
Evaluating Existing Data to Extrapolate High-Throughput Androgen Receptor Screening Data Across Species

SeqAPASS Evaluation of Structural Conservation Across Species

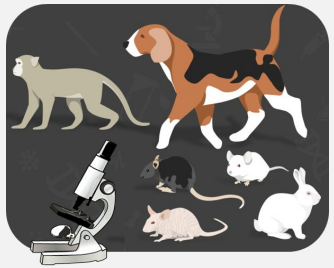


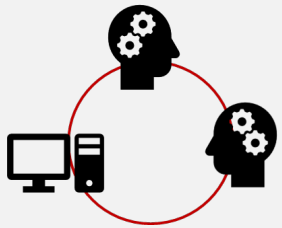
Conserved?
Yes
No

Systematic Evaluation of In Vitro Cross-Species Data



Systematic Evaluation of In Vivo Cross-Species Data





Weight of Evidence for Pathway Conservation Across Species for Defined Risk Assessment Applications

- Apply pathway to other targets of interest
- Repeat process to account for the emergence of new information

Case Study #2: Expanding and Evaluating In Silico Predictions Through In Vitro Laboratory Techniques

Site-directed mutagenesis: Change single amino acids to test the effect of amino acid change on protein-chemical binding

Default Rules in SeqAPASS: Shared Susceptibility with the Template = No **only** if both side chain and size are No

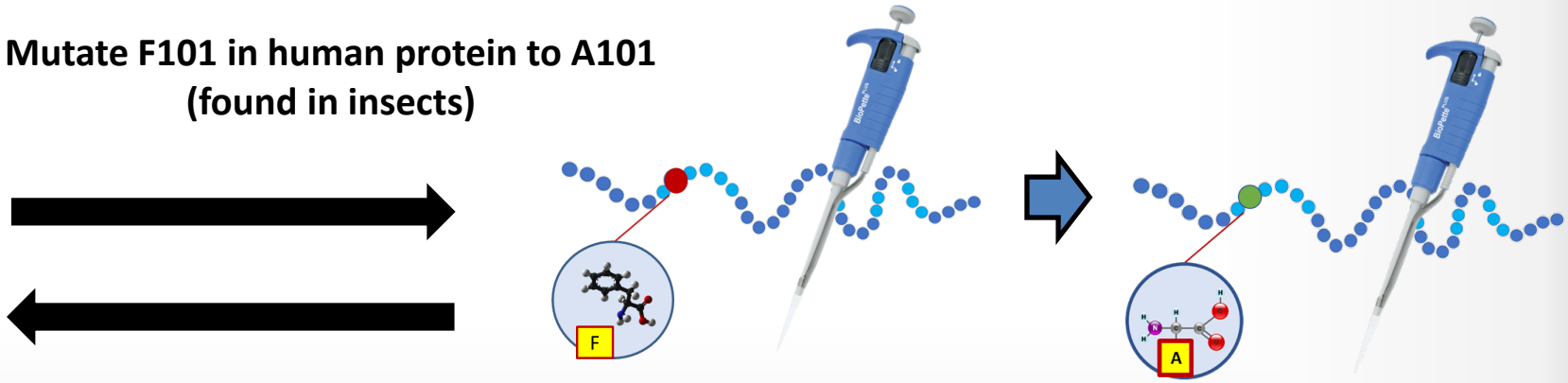
- Same side chain class as query (Y/N)?
- Size 30g/mol or less from query (Y/N)?

Guiding Question:
Can we use site-directed mutagenesis to confirm SeqAPASS predictions based on default settings

SeqAPASS Predictions

Protein	Shared Susceptibility	Position 1	Residue
Mouse protein	Yes	101	F
Human protein	Yes	101	F
Bird protein	Yes	101	Y
Turtle protein	Yes	101	R
Frog protein	No	101	D
Fish protein	No	101	N
Insect protein	No	101	A

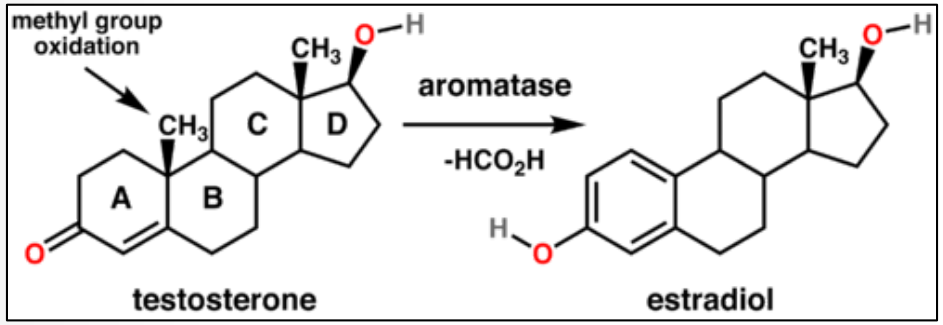
Mutate F101 in human protein to A101 (found in insects)



- **Direct** comparisons of amino acid substitutions
- Will **inform** and help **refine** current SeqAPASS settings

Case Study #2: Expanding and Evaluating In Silico Predictions Through In Vitro Laboratory Techniques: Aromatase

- Aromatase catalyzes biosynthesis of estrogens to androgens via hydroxylation and is an **important endocrine target**
- Environmental chemicals inhibiting aromatase can lead to endocrine disruption
- The structural components important in aromatase activity are well characterized, and assays measuring aromatase activity are readily available
- Pipeline can be applied to AR and other protein targets of interest



Key Amino Acids Involved in Aromatase-Ligand Interactions

- **Asp309** makes a **strong hydrogen bond** with the 3-keto group of ASD
- The 17-keto oxygen of ASD forms **a strong hydrogen bond** with **Met374**.
- A **weak hydrogen bond** between the 17-keto oxygen of ASD and **Arg115**.
- The residues that are directly involved in **catalysis** are **Pro308, Asp309** and **Thr310**.

- Level 3 SeqAPASS analysis identifies several amino acid positions as candidates for site-directed mutagenesis

Taxon	Species Name	Similar?	hCYP19A1 Pos. 1	hCYP19A1 Pos. 2	hCYP19A1 Pos. 3	hCYP19A1 Pos. 4	hCYP19A1 Pos. 5
Mammalia	Human	Y	R	P	D	T	M
Actinopteri	Channel catfish	N	R	P	D	T	Y

Case Study #2: Expanding and Evaluating In Silico Predictions Through In Vitro Laboratory Techniques: Aromatase Workflow

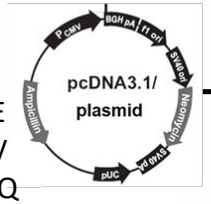
1. Critical residues identified through SeqAPASS and literature review

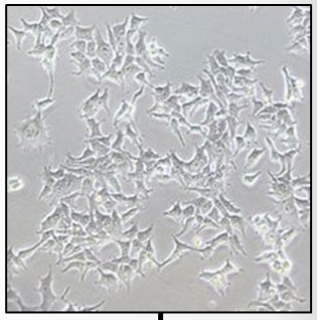
```
    10        20        30        40        50
MVLEMLNPIH YNITSIVPEA MPAATMPVLL LTGLFLLVWN YEGT
    60        70        80        90        100
GYCMGIGPLI SHGRFLWMGI GSACNYYNRV YGEFMRVWIS GE
    110       120       130       140       150
SSMFHIMKHN HYSSRFGSKL GLQCIGMHEK GIIFNNNPEL WKT
    160       170       180       190       200
ALSGPGLVRM VTVCAESLKT HLDRLEEVTN ESGYVDVLTL LRRV
    210       220       230       240       250
TLFLRIPLDE SAIVVKIQGY FDAWQALLIK PDIFFKISWL YKKYEKS
    260       270       280       290       300
LKDAIEVLIA EKRRRISTEE KLEECMDFAT ELILAEKRGD LTRENV
    310       320       330       340       350
LEMLIAAPDT MSVSLFFMLF LIAKHPNVEE AIIKEIQTVI GERDIKI
    360       370       380       390       400
QKLKVMENFI YESMRYQPVV DLVMRKALED DVIDGYPVKK GTNIILNIGR
    410       420       430       440       450
MHRLEFFPKP NEFTLENFAK NVPYRYFQPF GFGPRGCAGK YIAMVMMKAI
    460       470       480       490       500
LVTLLRRFHV KTLQGQCVES IQKIHDLSLH PDETKNMLEM IFTPRNSDRC

LEH
```

2. Plasmids designed and optimized with single aromatase mutations

3. Plasmids with mutated sequences synthesized and transfected into cells

Gene
hCyp19A1_WT
hCyp19A1_R115E
hCyp19A1_P308V
hCyp19A1_ P308Q
hCyp19A1_D309Y
hCyp19A1_M374Y
hCyp19A1_ M374A





Lyse/sonicate cells

4. Activity of mutated aromatase enzyme assessed via aromatase inhibition assay

- How do our amino acid substitutions affect Aromatase activity?
- Do these results support our computational predictions?
- How can this knowledge inform and refine future computational approaches?

Gathering Evidence of Endocrine Pathway Conservation for Cross-Species Extrapolation Using New Approach Methods

Wrapping it up

- The US EPA SeqAPASS tool is a New Approach Methods that can be used to computationally examine biological pathway conservation across taxa and predict chemical susceptibility across diverse species
- Using systematic literature review techniques and technological advances in data curation science, we demonstrated a framework for the evaluation of existing in vitro and in vivo data to anchor SeqAPASS-derived in silico predictions of species susceptibility
- This pipeline provides weight of evidence for the extrapolation of androgen and estrogen responses across species
- Using aromatase, we demonstrated current efforts to expand and validate in silico predictions through laboratory techniques, a pipeline that can be applied to future targets of interest
- Overall, we provide a framework for addressing the conservation of endocrine targets across species and understanding the degree to which mammalian-based NAMs can accurately reflect chemical interactions with non-mammalian targets.

Acknowledgements
Carlie LaLone, EPA, ORD
Kristan Markey, EPA, OCSPP
Scott Lynn, EPA, OCSPP
Battelle Contract Support
ILS Contract Support

Thanks!

Any questions?

SeqAPASS v4.0

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

Anyone can use SeqAPASS to help inform their own research questions! If you are interested in using SeqAPASS we are happy to help!

Vliet.Sara@epa.gov
LaLone.Carlie@epa.gov
Blatz.Donovan@epa.gov