CSS Chemical Safety for Sustainability Program

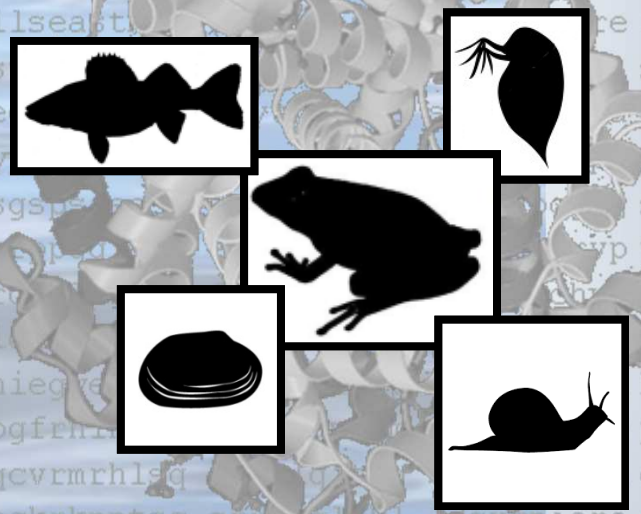
EPA

Cross-species Evaluation of Molecular Target Sequence and Structural Conservation as a Line of Evidence for Identification of Susceptible Taxa to Inform Toxicity Testing

Carlie A. LaLone, Ph.D.

Office of Research and Development

U.S. Environmental Protection Agency



Outline

- **Aquatic Life Criteria – Required data**
- **New tool to inform criteria: Sequence Alignment to Predict Across Species Susceptibility**
- **Demonstrate Application**
  - **Existing Criteria**
  - **Emerging chemicals**

Current Data Requirements

Data from a Prescribed List of Species

Freshwater Species

Acute Tests:

- Salmonidae
- Rec/Comm important warm-water fish
- Other vertebrate
- Planktonic crustacean
- Benthic crustacean
- Insect
- Phylum not Arthropoda or Chordata
- Another phylum or insect

Chronic Tests: Life-cycle, Partial life-cycle, Early life-stage

- Fish
- Invertebrate
- Sensitive fresh water

Saltwater Species

Acute Tests:

- 2 species from phylum Chordata
- Phylum not Arthropoda or Chordata
- Mysidae or Penaeidae family
- 3 families not Chordata or used above
- Any other family

Chronic Tests: Life-cycle, Partial life-cycle, Early life-stage

- Fish
- Invertebrate
- Sensitive saltwater species

Key Questions for Deriving Criteria

- **Does data exist to develop criteria?**
  - **If data gaps exist are they necessary/unnecessary to fill?**

- **If data available for some species are they representative of others needing protection?**

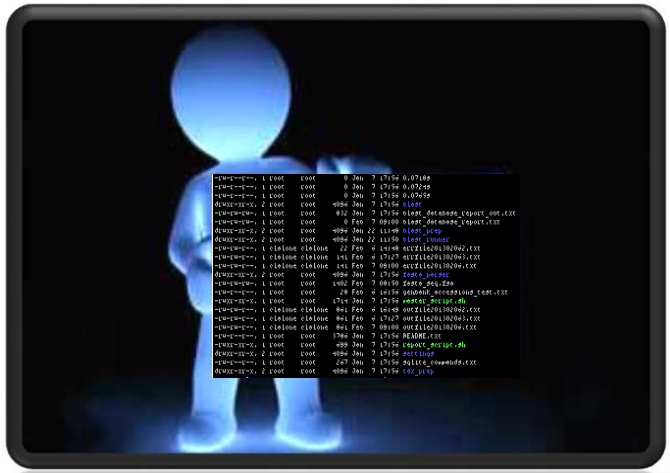
- **With emerging chemicals, sucn as pnarmaceuticals, limited if any toxicity data exists across taxa**
  - **What data would be necessary for deriving criteria?**

Proposed use of New Computational Tool: Guide chemical specific species requirements based on conservation of molecular target(s) as a line of evidence for species susceptibility

CSS Chemical Safety for Sustainability Program

EPA

Sequence Alignment to Predict Across Species Susceptibility (SeqAPASS)

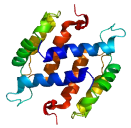


Understanding Protein Conservation

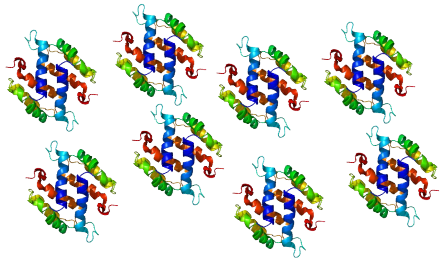
Sequence Alignment to Predict Across Species Susceptibility

- **Computational Assessment of Protein Similarity: Quantitative Metrics**
  - **Must know the molecular target (e.g., pharmaceuticals, pesticides)**
  - **Must identify target species or have knowledge of sensitive species**

Chemical Molecular Target in Target Species



≅



Compare to Millions of Proteins From Thousands of Species

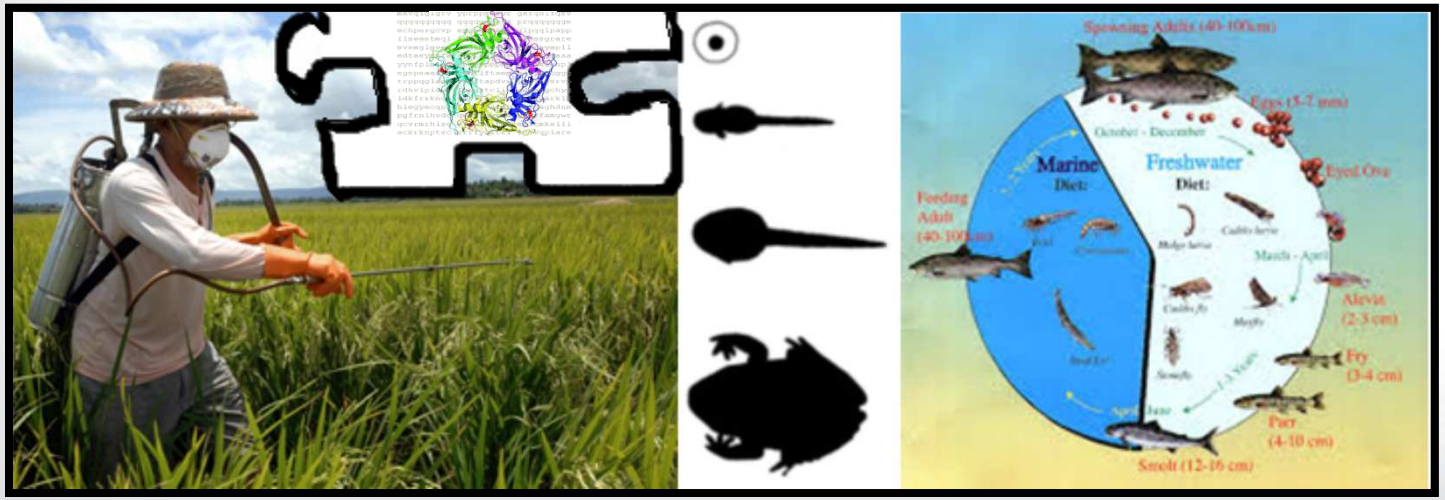
Greater similarity = Greater likelihood that chemical can act on the protein

Line of Evidence: Predict Potential Chemical Susceptibility Across Species

CSS Chemical Safety for Sustainability Program
EPA

Predict Relative Intrinsic Susceptibility

- **Intrinsic susceptibility can be defined as the vulnerability (or lack thereof) of an organism to chemical insult due to its inherent biological composition**
  - **Receptor/enzyme (protein) available for the chemical to act upon**
- **Relative: based on comparisons to a query protein**
  - **Molecular target conservation is but a component of multiple determinants of species susceptibility**

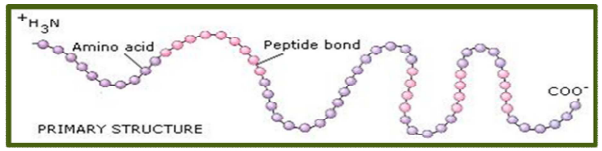


Strategic Automated Approach for Assessing Protein Similarity

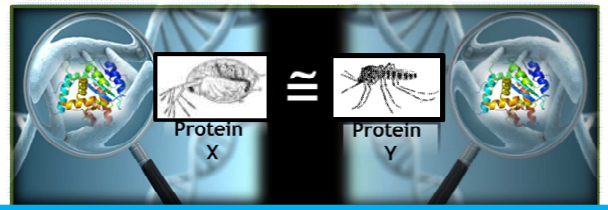
Ortholog: Sequences that descend from the same ancestral sequence – Often share similar function

Level 1

Primary Amino Acid Sequence Alignments

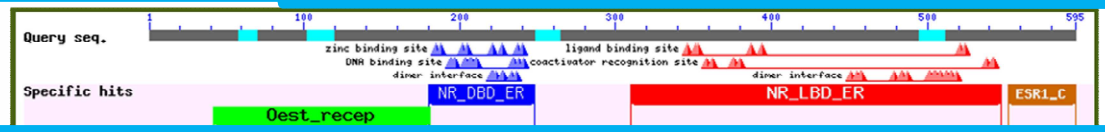


Ortholog Candidate Identification (RBH)



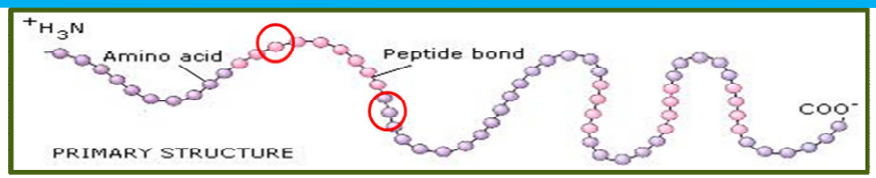
Level 2

Conserved Functional Domain Alignments



Level 3

Individual Amino Acid Residue Queries



Tertiary Protein Structure Considerations

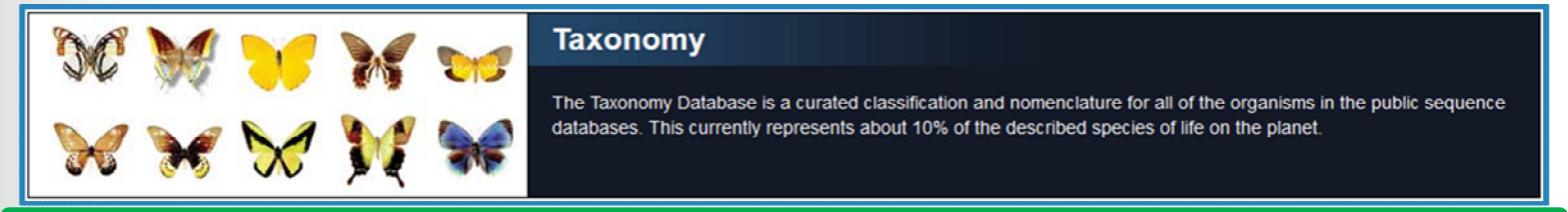
Low Level of Complexity

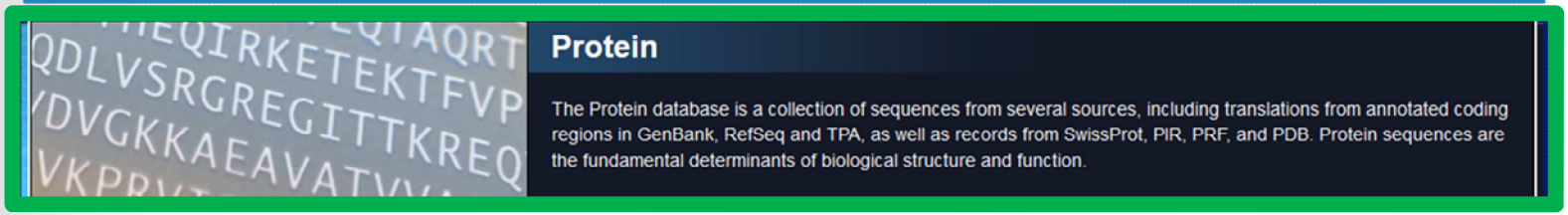
High level of Complexity

Developed with both researchers and risk assessors in mind

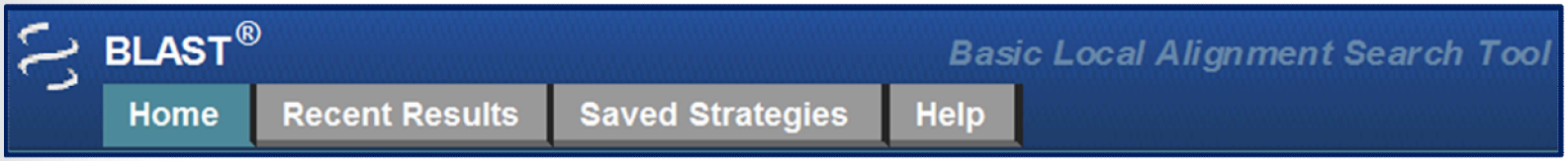
Available Databases and Tools

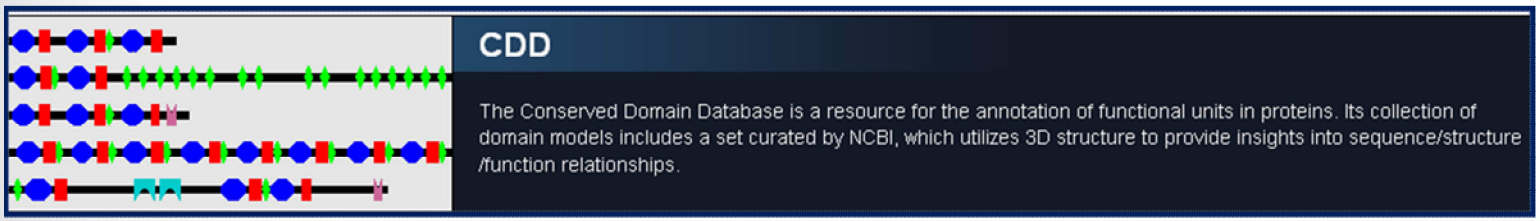
- **National Center for Biotechnology Information**
- **Established in 1988: a division of National Library of Medicine at NIH**





As of last week
~52 million Proteins
~55 thousand Species






DEMONSTRATE APPLICATIONS FOR SeqAPASS

Evaluation of Acetylcholinesterase Similarity

- **Molecular Target (Russom et al. 2014)**
  - **Acetylcholinesterase**
- **Sensitive species**
  - **Daphnid**

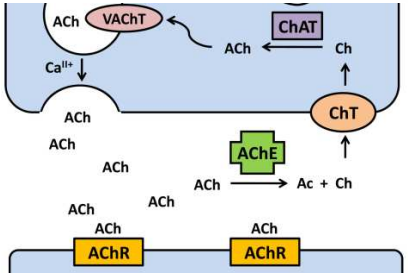
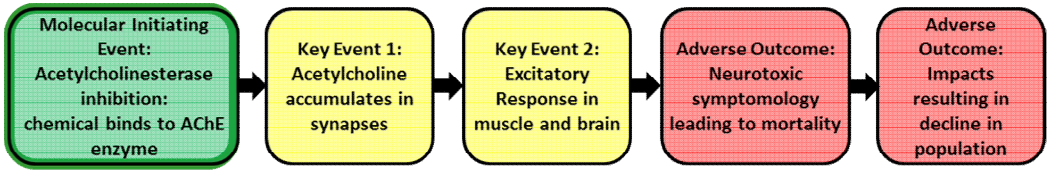


Figure adapted from: Pohanka et al. 2012



- **Chemicals with Aquatic Life Criteria**
  - **Carbaryl (2012)**
  - **Chloropyrifos (1986)**
  - **Diazinon (2005)**
  - **Malathion (1986)**
  - **Parathion (1995)**

Sequence Alignment to Predict Across Species Susceptibility (SeqAPASS)

Logout

Home | Request SeqAPASS Run | View SeqAPASS Reports | Susceptibility Cutoff Formulas

Request SeqAPASS Run

Choose Search Type: By Species

Query Species Search: Daphnia pulex (taxid:6669)

Add Query Species

Selected Query Species: Daphnia pulex (taxid:6669)

Query Protein Search: DAPPUDRAFT_3925

Filter Protein

http://www.ncbi.nlm.nih.gov/protein

Query Proteins:

[EFX89932.1] hypothetical protein DAPPUDRAFT_39257 [Daphnia pulex

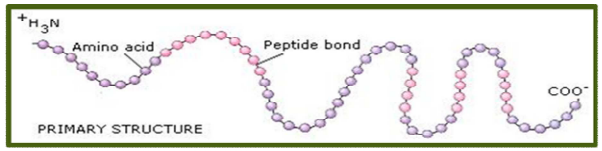
Final Query Protein(s) for SeqAPASS Run: [EFX89932.1] acetylcholinesterase 1 [Daphnia pulex]

Request Run | Clear

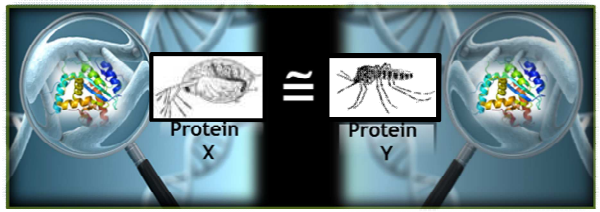
Strategic Approach for Assessing Protein Similarity

Level 1

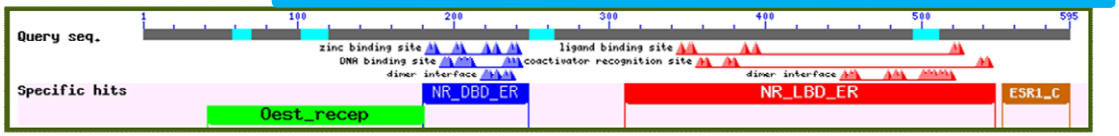
Primary Amino Acid Sequence Alignments



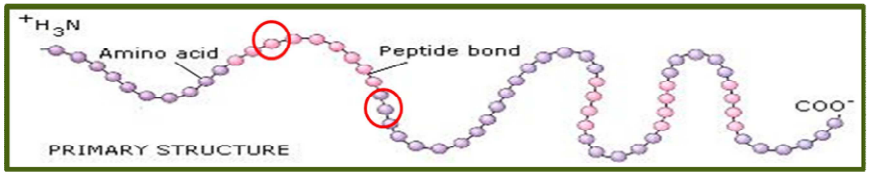
Ortholog Candidate Identification (RBH)



Conserved Functional Domain Alignments



Individual Amino Acid Residue Queries



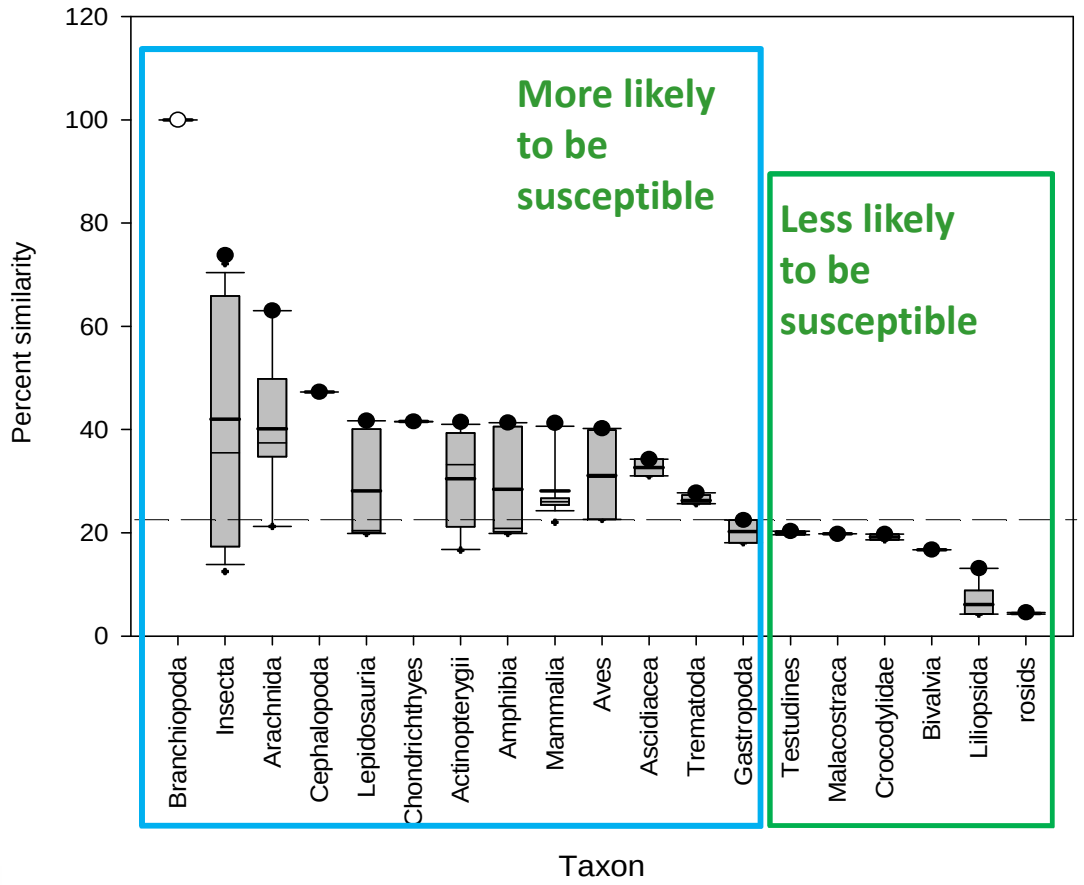
Tertiary Protein Structure Considerations

Low Level of Complexity

High level of Complexity

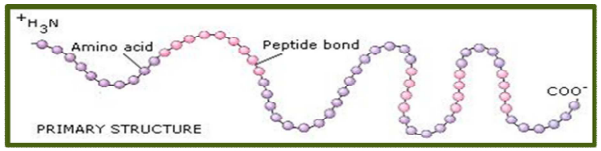
Level I SeqAPASS Results

Primary Amino Acid Sequence Alignment

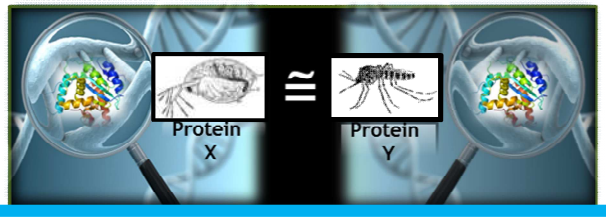


Strategic Approach for Assessing Protein Similarity

Primary Amino Acid Sequence Alignments

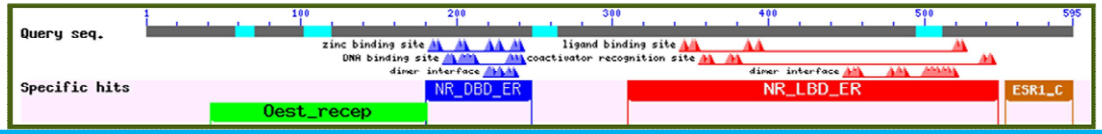


Ortholog Candidate Identification (RBH)

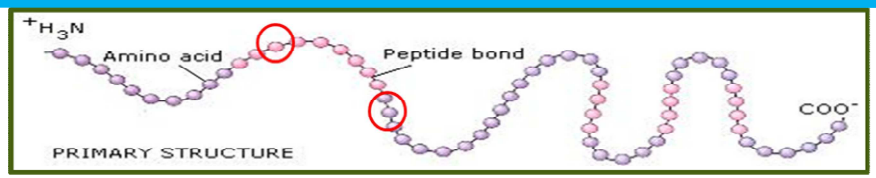


Level 2

Conserved Functional Domain Alignments



Individual Amino Acid Residue Queries



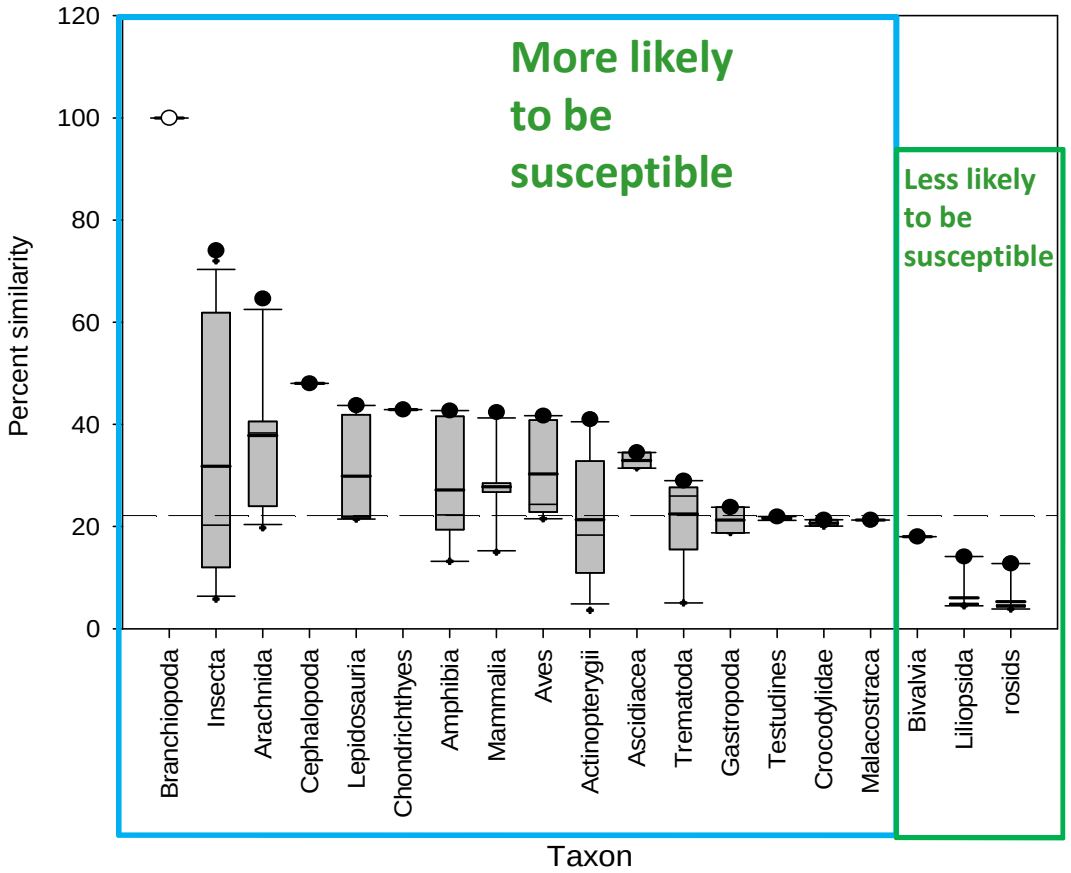
Tertiary Protein Structure Considerations

Low Level of Complexity

High level of Complexity

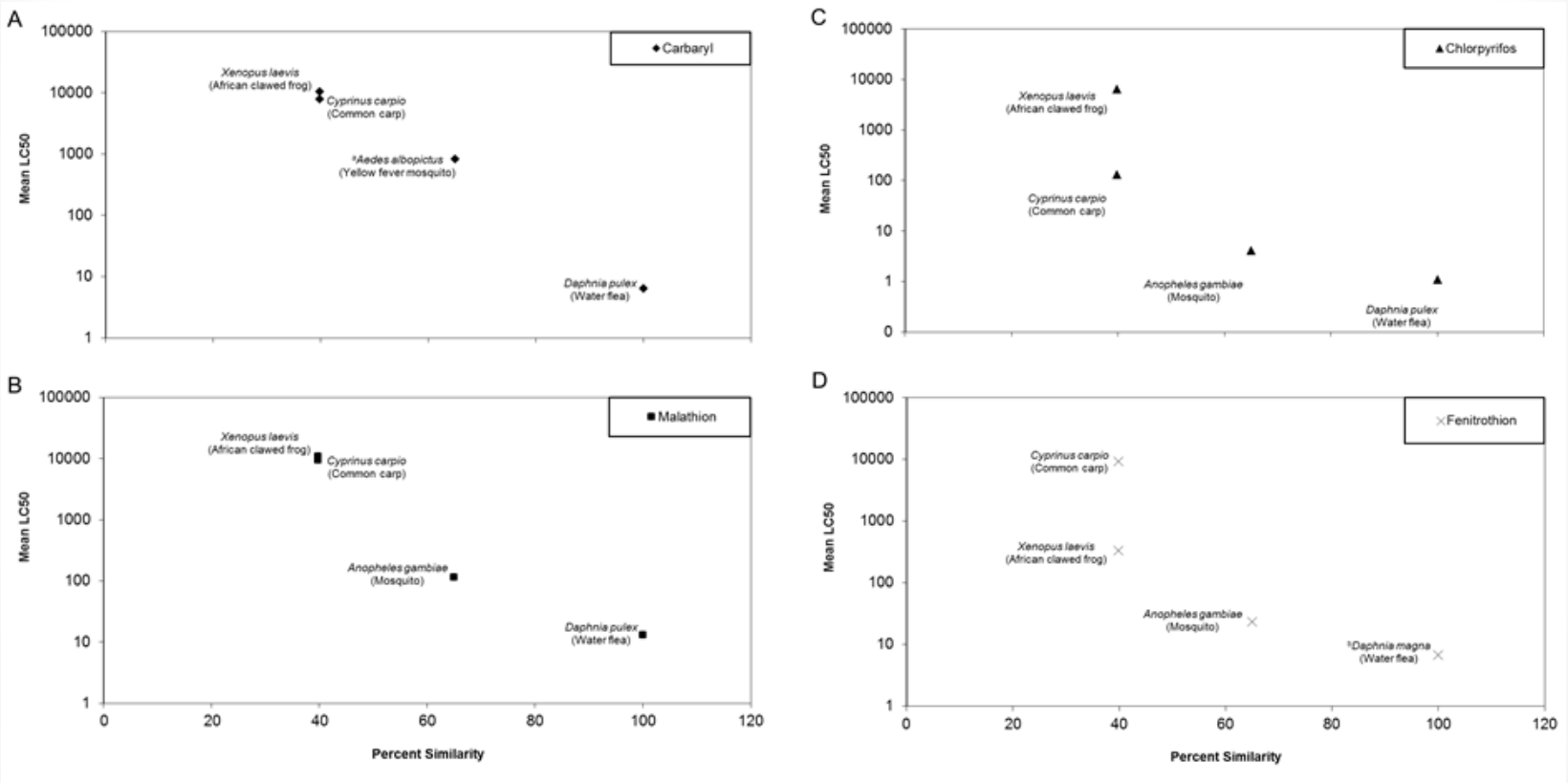
Level 2 SeqAPASS Results

Functional Domain Alignment: Esterase_lipase



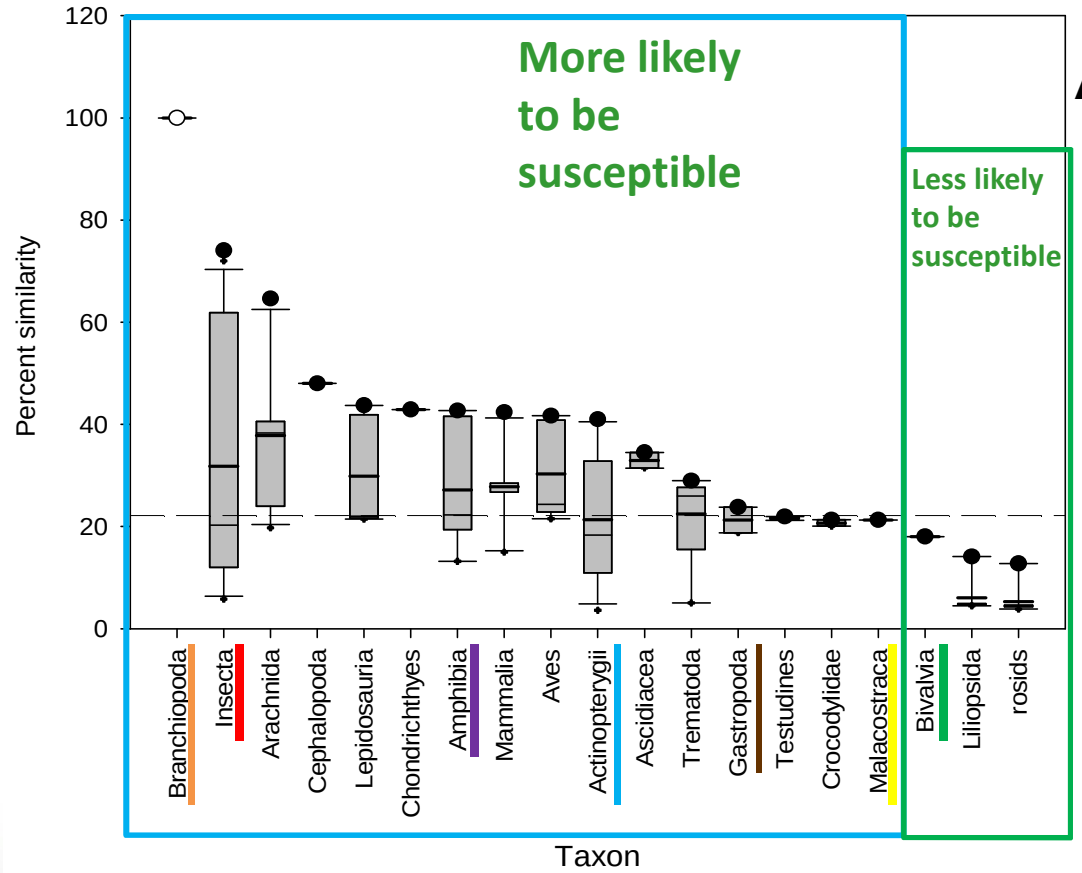
CSS Chemical Safety for Sustainability Program
EPA

Predictions vs. Empirical Toxicity Data



Evidence that quantitative measures of sequence similarity can be used to predict susceptibility

Level 2 SeqAPASS Results



ALC Recommendation

Priority Data

- Branchiopoda
- Insecta
- Amphibia
- Actinopterygii
- Gastropoda
- Malacostraca

Low Priority Data

- Bivalvia
- Liliopsida

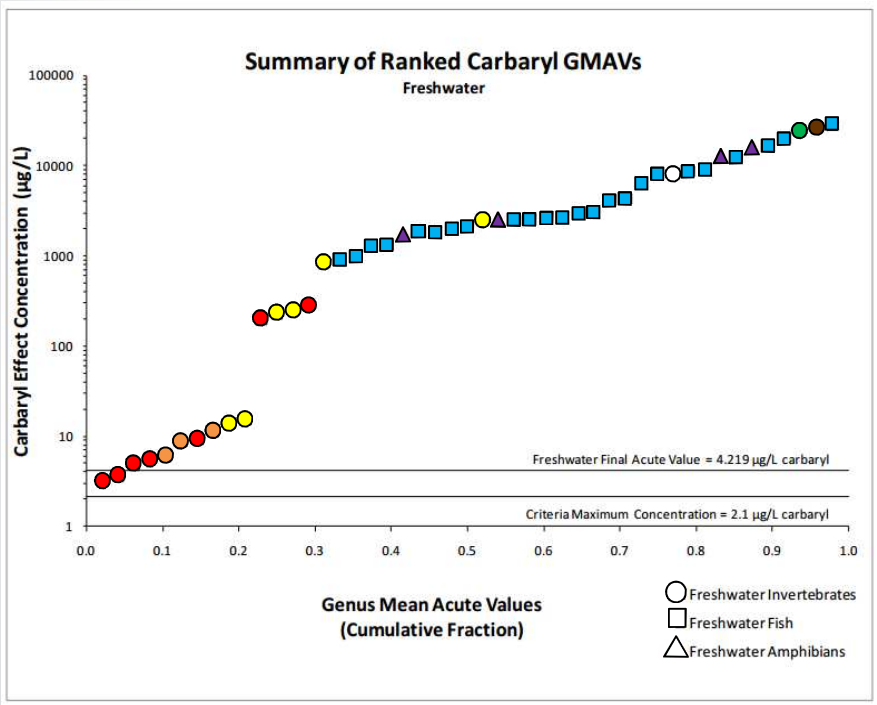
SeqAPASS vs Carbaryl Summary from ALC

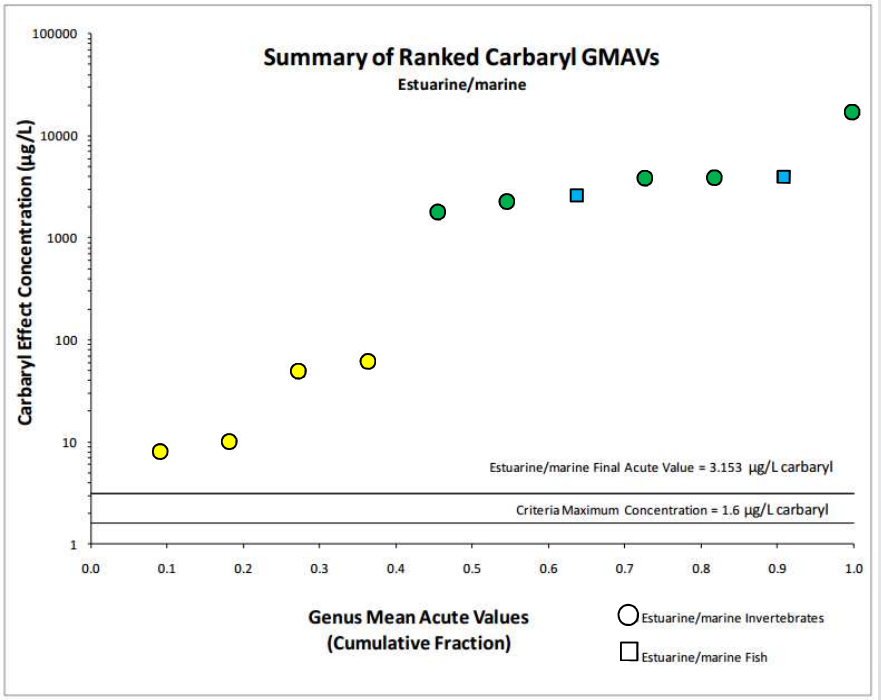
Priority Data

- Branchiopoda
- Insecta
- Amphibia
- Actinopterygii
- Gastropoda
- Malacostraca

Low Priority Data

- Bivalvia
- Liliopsida





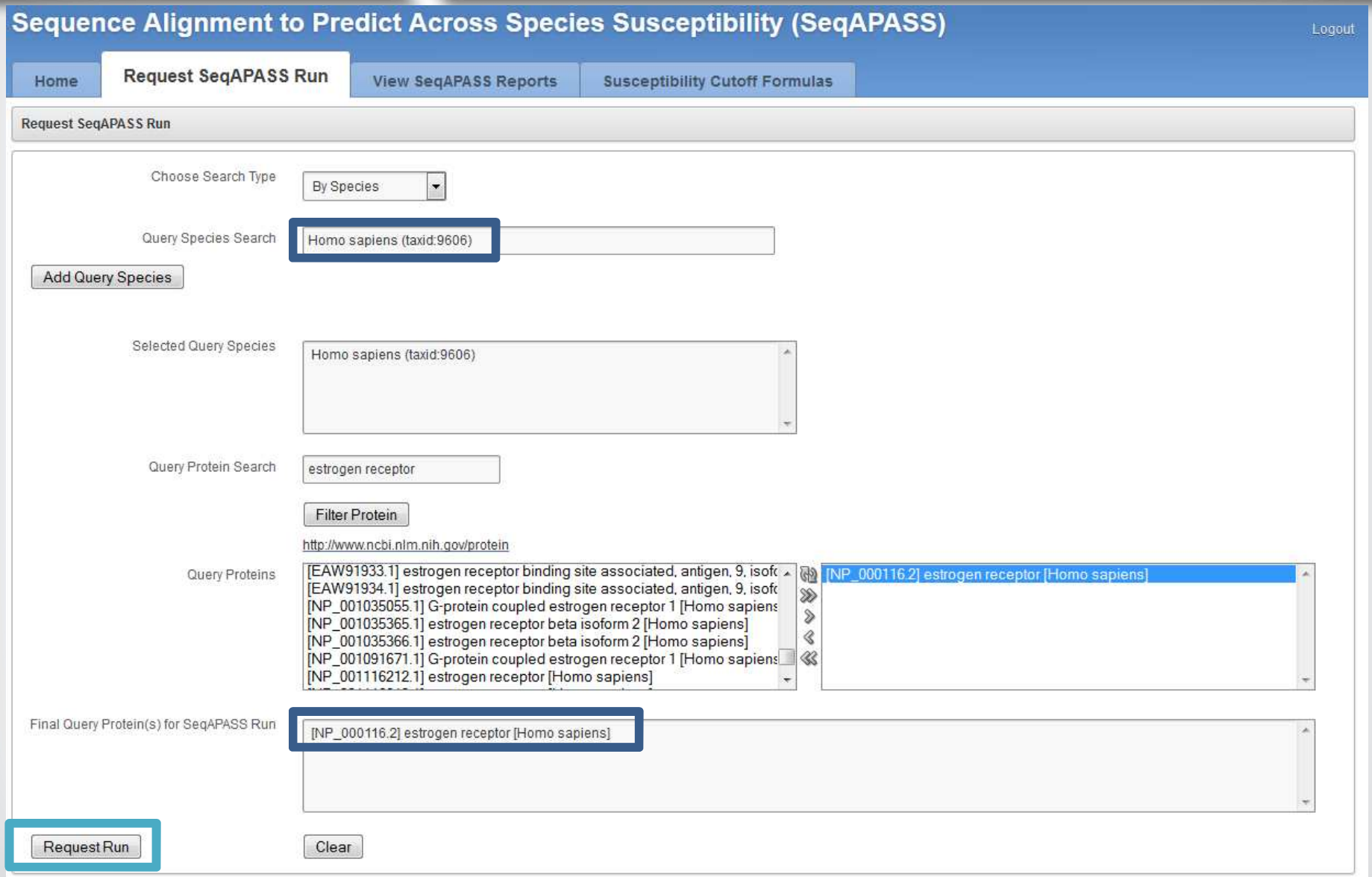
ALC: Emerging Chemicals

SeqAPASS Example: Ethinyl Estradiol

Target Species: Human

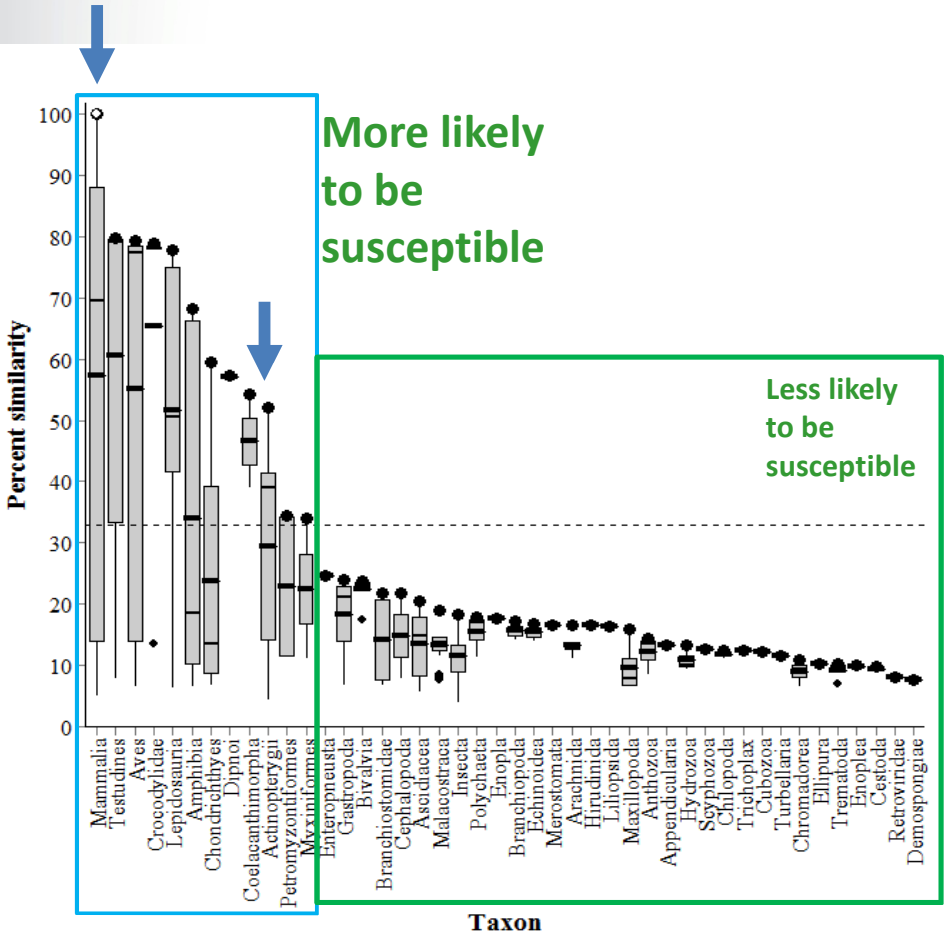
Molecular Target: Estrogen receptorα

Evaluation of Protein Conservation

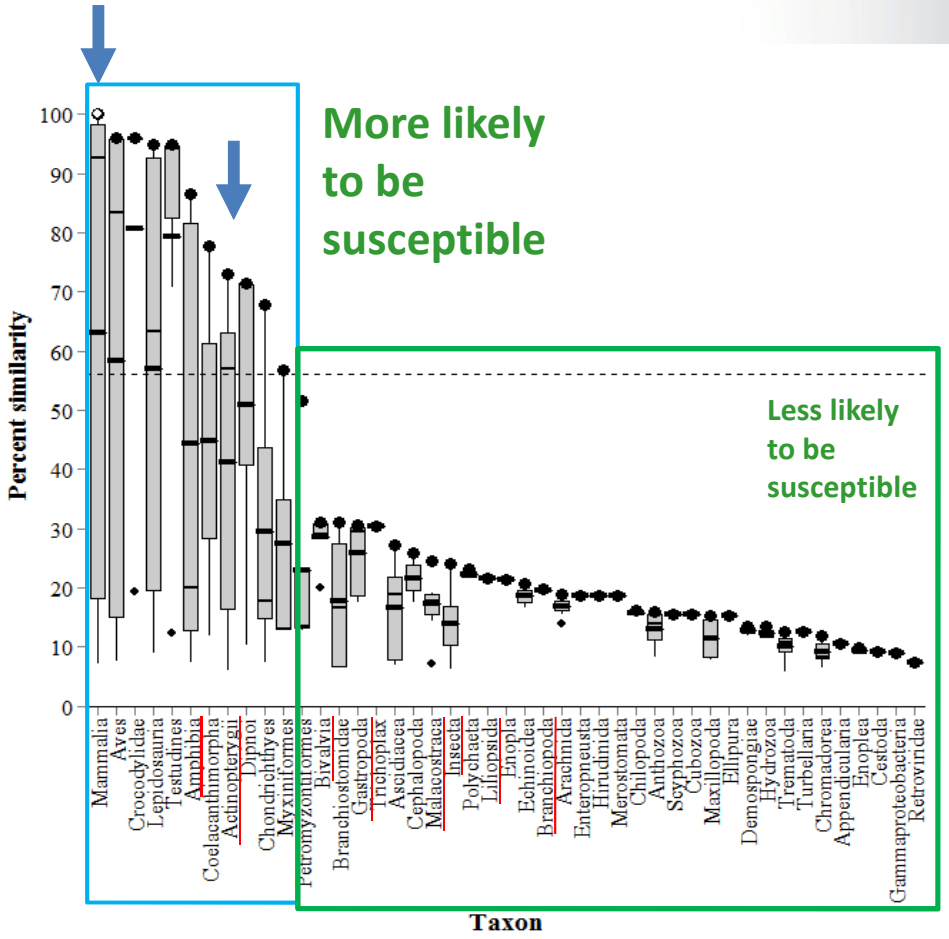


SeqAPASS Level 1 and 2

Primary Amino Acid Sequence

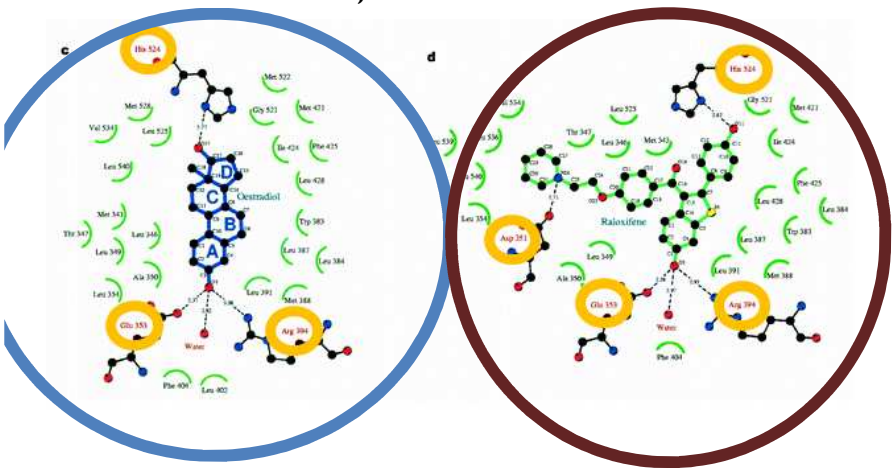


Ligand Binding Domain

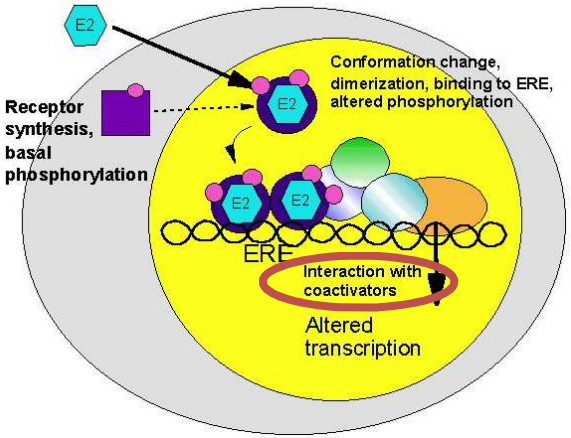


SeqAPASS Level 3: Focus on fish as an example

Brzozowski et al., 1997. *Nature* 389



McInerney et al., 1996. *Mol Endocrin*



Residues:

Histidine 524
Glutamic acid 353
Arginine 394
Aspartic acid 351
Lysine362
Valine 364

Fish Species

Scientific Name	Common Name	Protein Name	D351	E353	K362	V364	R394	H524
Homo sapiens	human	Estrogen receptor	D	E	K	V	R	H
Atractosteus tropicus	tropical gar	estrogen receptor alpha	D	E	K	V	R	H
Acipenser schrenckii	Amur sturgeon	estrogen receptor alpha1	D	E	K	V	R	H
Zoarces viviparus	viviparous blenny	estrogen receptor alpha	D	E	K	I	R	H
Pimephales promelas	fathead minnow	estrogen receptor alpha	D	E	K	L	R	H
Tanichthys albonubes	white cloud mountain minnow	estrogen receptor alpha	D	E	K	V	R	H
Epinephelus coioides	orange-spotted grouper	estrogen receptor alpha	D	E	K	V	R	H
Rutilus rutilus	roach minnow	estrogen receptor alpha	D	E	K	L	R	H
Cyprinus carpio	common carp	estrogen receptor alpha	D	E	K	V	R	H
Sebastes schlegelii	Korean rockfish	estrogen receptor alpha	D	E	K	V	R	H
Odontesthes bonariensis	pejerrey	estrogen receptor alpha	D	E	K	L	R	H
Salmo salar	Atlantic salmon	estrogen receptor	D	E	K	L	R	H
Poecilia reticulata	guppy	estrogen receptor alpha	D	E	K	V	R	H
Kryptolebias marmoratus	mangrove rivulus	estrogen receptor alpha	D	E	K	L	R	H
Halichoeres tenuispinis	Chinese wrasse	estrogen receptor alpha	D	E	K	L	R	H
Gasterosteus aculeatus	three-spined stickleback	estrogen receptor alpha	D	E	K	L	R	H
Ictalurus punctatus	channel catfish	estrogen receptor type alpha	D	E	K	L	R	H
Clarias gariepinus	North African catfish	estrogen receptor alpha	D	E	K	V	R	H
Oryzias latipes	Japanese medaka	Estrogen receptor	D	E	K	L	R	H
Oncorhynchus mykiss	rainbow trout	Estrogen receptor	D	E	K	V	R	H
Paralichthys olivaceus	Japanese flounder	estrogen receptor alpha	D	E	K	L	R	H
Oncorhynchus masou	cherry salmon	estrogen receptor alpha	D	E	K	V	R	H
Gambusia affinis	western mosquitofish	estrogen receptor alpha	D	E	K	L	R	H
Micropterus salmoides	largemouth bass	estrogen receptor alpha	D	E	K	L	R	H
Melanotaenia fluviatilis	Murray River rainbowfish	estrogen receptor alpha	D	E	K	L	R	H
Candidia barbatus	Lake Candidus dace	putative estrogen receptor alpha	D	E	K	V	R	H
Spinibarbus denticulatus	bony fishes	estrogen receptor alpha	D	E	K	V	R	H
Paramisgurnus dabryanus	bony fishes	estrogen receptor alpha	D	E	K	V	R	H
Oryzias javanicus	javanese ricefish	estrogen receptor alpha	D	E	K	L	R	H
Micropogonias undulatus	Atlantic croaker	Estrogen receptor	D	E	K	L	R	H
Lepomis macrochirus	bluegill	estrogen receptor alpha1	D	E	K	L	R	H
Pseudolabrus japonicus	bony fishes	estrogen receptor alpha	D	E	K	L	R	H
Dicentrarchus labrax	European seabass	oestrogen receptor alpha	D	E	K	L	R	H
Sparus aurata	gilthead seabream	estrogen receptor	D	E	K	L	R	H
Acanthopagrus schlegelii	black porgy	estrogen receptor alpha	D	E	K	L	R	H
Fundulus heteroclitus	mummichog	estrogen receptor a	D	E	K	L	R	H
Oreochromis niloticus	Nile tilapia	estrogen receptor	D	E	K	L	R	H
Carassius auratus	goldfish	estrogen receptor alpha	D	E	K	V	R	H
Pagrus major	red seabream	Estrogen receptor	D	E	K	L	R	H
Varicorhinus barbatulus	bony fishes	putative estrogen receptor alpha protein	D	E	K	V	R	H
Gobiocypris rarus	bony fishes	estrogen receptor alpha	D	E	K	V	R	H
Acheilognathus yamatsutae	Korean striped bitterling	estrogen receptor alpha	D	E	K	V	R	H
Haplochromis burtoni	bony fishes	estrogen receptor alpha	D	E	K	L	R	H
Danio rerio	zebrafish	Estrogen receptor 1	D	E	K	V	R	H
Cirrhinus molitorella	mud carp	estrogen receptor a, partial	D	E	K	V	R	H
Oncorhynchus kisutch	coho salmon	estrogen receptor alpha 1	D	E	K	V	R	H
Halichoeres trimaculatus	threespot wrasse	estrogen receptor, partial	D	E	K	L	R	H
Perca flavescens	yellow perch	estrogen receptor alpha	D	E	K	L	R	H
Oreochromis mossambicus	Mozambique tilapia	estrogen receptor type I	D	E	K	L	R	H
Mugil cephalus	flathead mullet	estrogen receptor alpha	D	E	K	L	R	H
Acanthogobius flavimanus	bony fishes	estrogen receptor alpha	D	E	K	L	R	H
Oreochromis aureus	blue tilapia	Estrogen receptor	D	E	K	L	R	H
Amphiprion melanopus	fire clownfish	estrogen receptor beta 1, partial	D	E	K	I	R	H
Lepisosteus oculatus	spotted gar	estrogen receptor alpha	D	E	K	V	R	H
Porichthys notatus	plainfin midshipman	estrogen receptor alpha	D	E	K	L	R	H
Cynoglossus semilaevis	tongue sole	estrogen receptor alpha, partial	D	E	K	L	R	H
Oncorhynchus mykiss irideus	bony fishes	estrogen receptor, partial	D	E	K	V	R	H
Solea solea	common sole	estrogen receptor beta	D	E	K	I	R	H
Anguilla japonica	Japanese eel	Estrogen receptor beta	D	E	K	I	R	H
Hippoglossus hippoglossus	Atlantic halibut	estrogen receptor alpha, partial	D	E	K	L	R	H
Larimichthys crocea	large yellow croaker	estrogen receptor beta	D	E	K	I	R	H
Conger myriaster	whitespotted conger	estrogen receptor beta	D	E	K	I	R	H
Thunnus thynnus	northern bluefin tuna	estrogen receptor a	D	E	K	L	R	H
Salmo trutta	brown trout	estrogen receptor beta	D	E	K	I	R	H
Acanthopagrus latus	yellowfin seabream	estrogen receptor beta1, partial	D	E	K	I	R	H
Misgurnus anguillicaudatus	oriental weatherfish	estrogen receptor alpha	D	E	K	V	R	-
Morone saxatilis	striped sea-bass	estrogen receptor type B	D	E	K	I	R	-
Ameiurus nebulosus	brown bullhead	estrogen receptor alpha	D	E	K	V	R	H
Oryzias melastigma	Indian medaka	estrogen receptor beta, partial	D	E	K	I	R	-
Monopterus albus	swamp eel	estrogen receptor alpha	D	E	K	L	R	-
Mugil curema	white mullet	estrogen receptor alpha, partial	D	E	K	L	R	H
Pelteobagrus fulvidraco	yellow catfish	estrogen receptor EcR1	D	E	K	I	R	-
Gambusia holbrooki	eastern nosquitofish	estrogen receptor beta-like 2	D	E	K	I	R	-
Alburnus tarichi	tarek	estrogen receptor alpha	D	E	K	V	R	-
Chelon labrosus	thicklip grey mullet	estrogen receptor alpha	K	L	K	L	R	H
Micropterus dolomieu	smallmouth bass	estrogen receptor alpha	D	E	-	-	-	-
Merluccius merluccius	European hake	estrogen receptor alpha	-	L	K	V	R	-
Thalassoma bifasciatum	bluehead	estrogen receptor beta a	D	E	K	I	R	-
Cichlasoma dimerus	bony fishes	estrogen receptor alpha	-	-	-	-	-	-
Oncorhynchus nerka	sockeye salmon	estrogen receptor alpha, partial	K	L	K	V	R	-
Oncorhynchus tshawytscha	Chinook salmon	estrogen receptor alpha, partial	K	L	K	V	R	-
Channa punctata	spotted snakehead	estrogen receptor alpha	E	E	-	-	-	-
Gobius niger	black goby	estrogen receptor beta	-	-	-	-	-	H

Priority Data for ALC

Priority Data

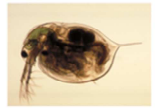
- ▲ Amphibia
- ■ Actinopterygii

Low Priority Data

- ● Branchiopoda
- ● Insecta
- ● Gastropoda
- ● Malacostraca
- ● Bivalvia
- - Liliopsida

Estrogen Receptor (OW/ORD Emerging Contaminants Workgroup, 2008)

Animal Kingdom	Genus	Common name	Chronic value (ng/L)
Vertebrates, aquatic	*Danio*	Zebrafish	<1.1
	Pimephales	Fathead minnow	1.5
	Oryzias	Medaka	3.2
	Oncorhynchus	Rainbow trout	<16
Invertebrates, aquatic	*Potamopyrgus*	Snail	50
	Gammarus	Freshwater shrimp	>7,600
	Daphnia	Water flea	45,000
	Tisbe	Copepod	>100,000
	Chironomus	Midge	320,000
	Brachionus	Rotifer	800,000



Proposed use of SeqAPASS: Guide chemical specific species data priorities based on conservation of molecular target(s) as a line of evidence for species susceptibility

Hypothetical Minimum Data Requirements

Acute:

Species	Requirement	Biology	Chemistry	Total	Basis
Rainbow Trout	Salmonid	$4,000.00	$500.00	$4,500.00	96-h FT
Bluegill	Rec/comm important warmwater fish	$4,000.00	$500.00	$4,500.00	96-h FT
Fathead minnow	Other vertebrate	$4,000.00	$500.00	$4,500.00	96-h FT
Daphnid	Planktonic crustacean	$1,500.00	$250.00	$1,750.00	48-h R
Hyalella	Benthic crustacean	$2,500.00	$500.00	$3,000.00	96-h R
Midge	Insect	$2,500.00	$500.00	$3,000.00	96-h R
Lumbriculus	Phylum not Arthropoda or Chordata	$2,500.00	$500.00	$3,000.00	96-h R
Physa	Another phylum or another insect order	$2,500.00	$500.00	$3,000.00	96-h R

Chronic:

Species	Requirement	Biology	Chemistry	Total	Basis
Fathead Minnow ELS	Fish	$28,500.00	$1,500.00	$30,000.00	28-d ELS; FT
Daphnid	Invert and sensitive FW (assumed)	$5,000.00	$1,000.00	$6,000.00	21-d LC; R
Hyalella	Something else	$10,500.00	$1,500.00	$12,000.00	42-d PLC; R

Total Estimated Cost: $**75,250.00**

Cost to run SeqAPASS to Prioritize Tests: $0

Acknowledgements

- U.S. Environmental Protection Agency
  - Daniel Villeneuve
  - Gerald Ankley
- Computer Sciences Corporation
  - Henry Helgen
  - Derek Lane
  - Steve Walata

Thank you

LaLone.Carlie@epa.gov