EPA United States Environmental Protection Agency

Sequence Alignment to Predict Across Species Susceptibility (SeqAPASS) to Understand Chemical Susceptibility Across Bee Species

Donovan Blatz

Carlie A. LaLone, Ph.D.

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

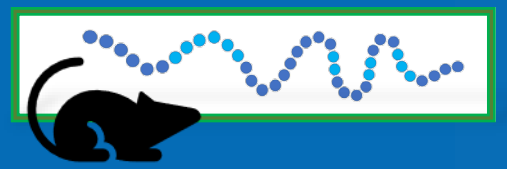
TOXICOLOGICAL SCIENCES, 153(2), 2016, 228–245

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

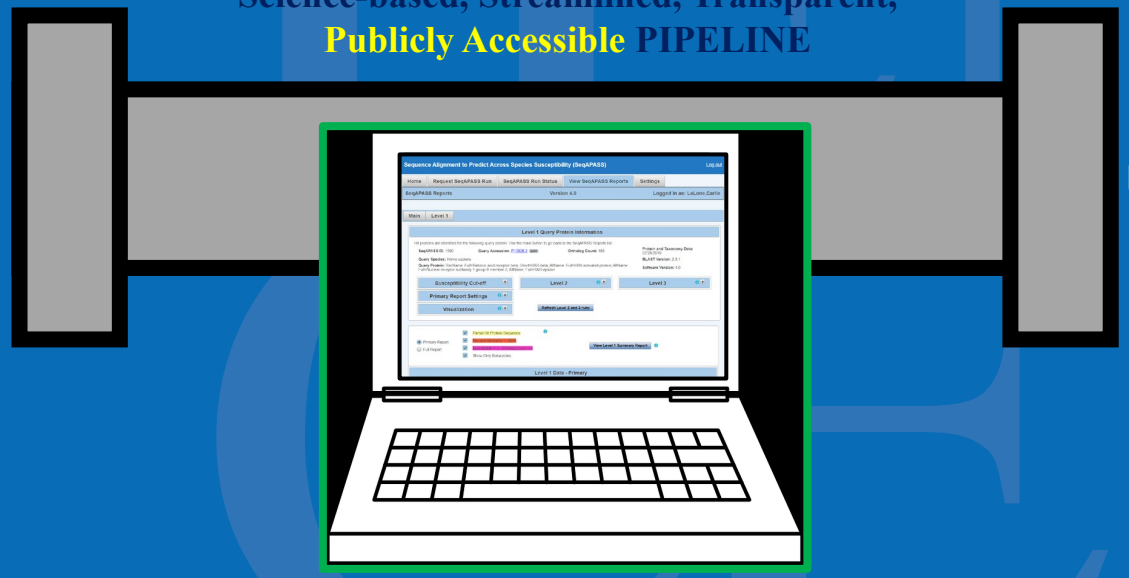
Sequence Alignment to Predict Across Species Susceptibility (SeqAPASS): A Web-Based Tool for Addressing the Challenges of Cross-Species Extrapolation of Chemical Toxicity

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*

Available Knowledge



Science-based, Streamlined, Transparent, Publicly Accessible PIPELINE



Predict Susceptibility

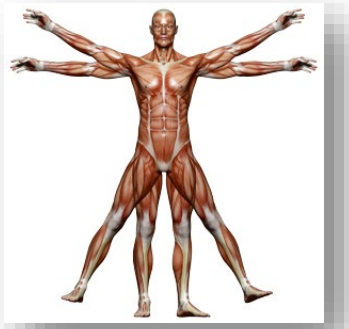
Office of Research and Development
Center for Computational Toxicology and Exposure, Great Lakes Toxicology and Ecology Division

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

Transformation of Toxicity Testing

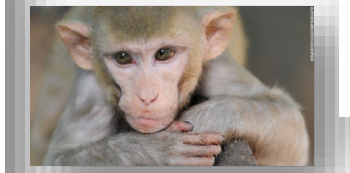
Assumed that sensitivity of species to a chemical is a function of their relatedness

- **Human Health Risk Assessment**



Cannot Test

≅

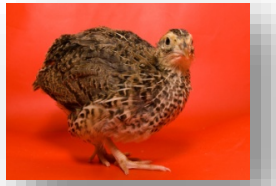


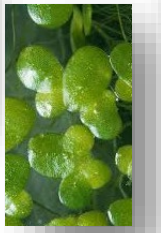
Use of Surrogates

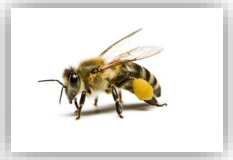
- **Ecological Risk Assessment**

Cannot Test

≅







Representative species across a diversity of organism classes

EPA United States Environmental Protection Agency

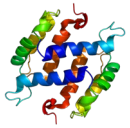
Understanding Protein Conservation

Sequence Alignment to Predict Across Species Susceptibility

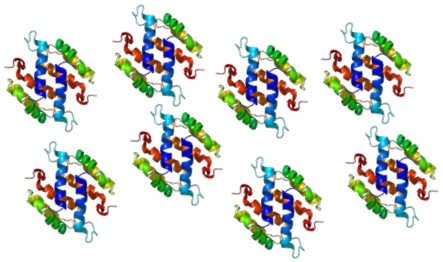
- Computational Assessment of Protein Similarity: Quantitative Metrics
  - A comparison of species at the molecular level

1. Must know the molecular target (e.g., pharmaceuticals, pesticides)
2. Must identify target species or have knowledge of sensitive species

Chemical Molecular Target in Target Species



≅

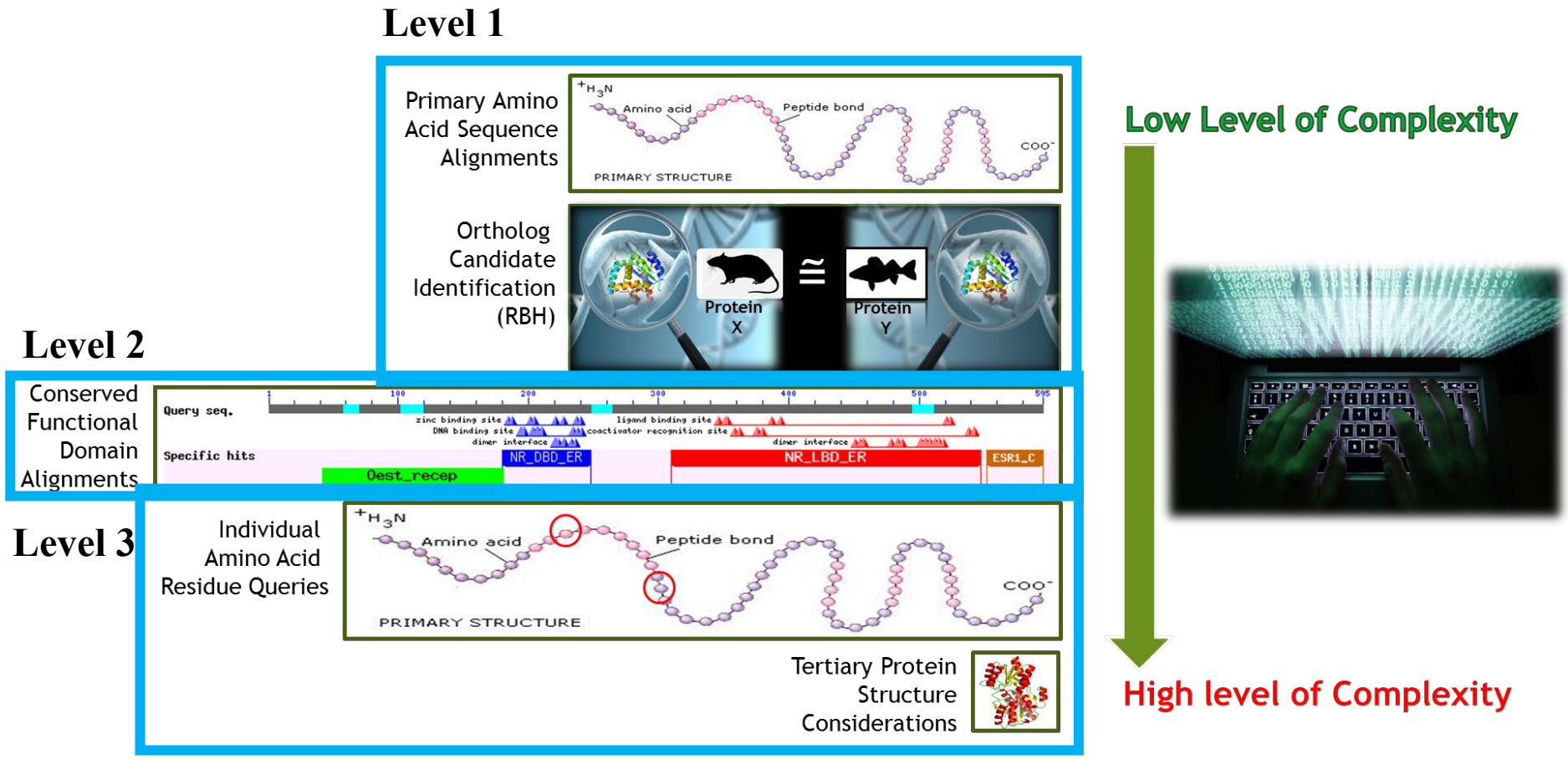


Compare to Millions of Proteins From Thousands of Species

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

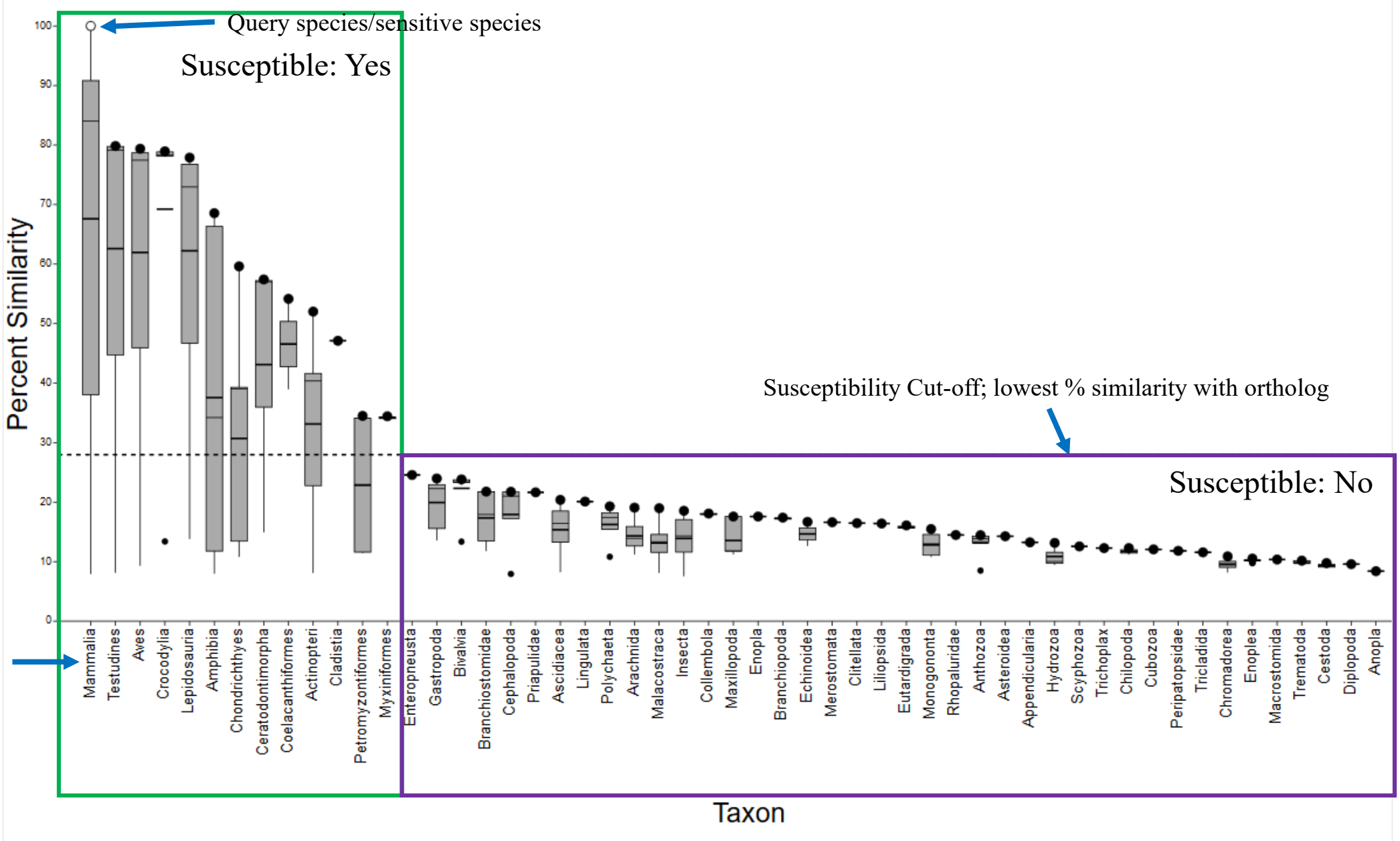
Line of Evidence: Predict Potential Chemical Susceptibility Across Species

Strategic Automated Approach for Assessing Protein Similarity



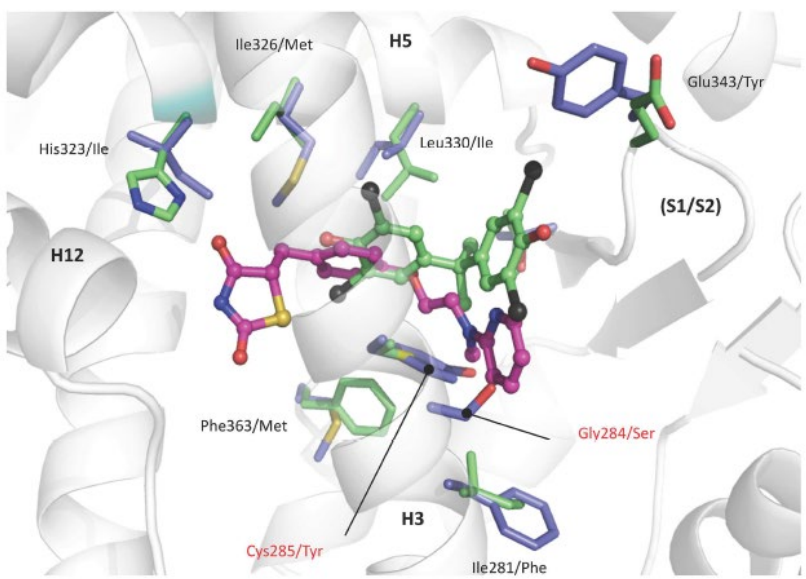
Developed with both researchers and risk assessors in mind

Result from Level 1 and Level 2: Predictions for Hundreds of Species Rapidly



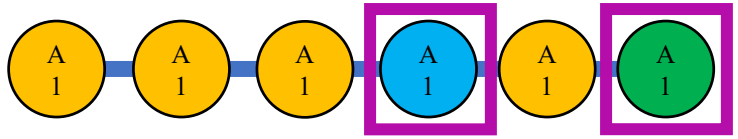
Level 3: Individual Amino Acid Residue Comparison

Crystal structure of protein – Ligand bound



Grimaldi et al 2015

Amino acid residues that interact with the chemical



- Types of studies to identify critical amino acids:
  - Site-directed mutagenesis
  - Field resistance (pesticides)
  - Studies of x-ray crystallography
  - Molecular docking

Amino Acid info

ID	Name	Side Chain	Size
A	Alanine	Aliphatic	89.094
C	Cysteine	Sulfur-Containing	121.154
D	Aspartic Acid	Acidic	133.104
E	Glutamic Acid	Acidic	147.131
F	Phenylalanine	Aromatic	165.192
G	Glycine	Aliphatic	75.067
H	Histidine	Basic	155.156
I	Isoleucine	Aliphatic	131.175
K	Lysine	Basic	146.189
L	Leucine	Aliphatic	131.175
M	Methionine	Sulfur-Containing	149.208
N	Asparagine	Amidic	132.119
P	Proline	Aliphatic	115.132
Q	Glutamine	Amidic	146.146
R	Arginine	Basic	174.203
S	Serine	Hydroxylic	105.093
T	Threonine	Hydroxylic	119.119
V	Valine	Aliphatic	117.148
W	Tryptophan	Aromatic	204.228
X	Unknown	Unknown	-100.0
Y	Tyrosine	Aromatic	181.191

- **Likelihood of an amino acid difference causing a difference in protein-chemical interaction is automatically evaluated by SeqAPASS based on 2 simple rules:**

1) **Same side chain classification as Template (Y/N)**
2) **Molecular weight 30g/mol or less from Template (Y/N)**

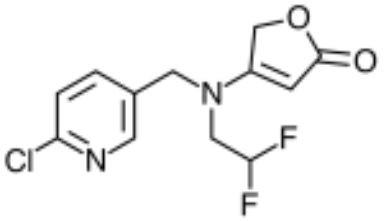
Level 3 Example Output

Similar Susceptibility as Template	Position 1	Amino Acid 1	Direct Match 1	Side Chain 1	Side Chain Match 1	MW 1	MW Match 1	Total Match 1	Position 2	Amino Acid 2	Direct Match 2	Side Chain 2	Side Chain Match 2	MW 2	MW Match 2	Total Match 2
Y	14	T	Y	Hydroxylic	Y	119.119	Y	Y	221	F	Y	Aromatic	Y	165.192	Y	Y
Y	12	T	Y	Hydroxylic	Y	119.119	Y	Y	219	F	Y	Aromatic	Y	165.192	Y	Y
Y	14	T	Y	Hydroxylic	Y	119.119	Y	Y	221	F	Y	Aromatic	Y	165.192	Y	Y
Y	18	S	N	Hydroxylic	Y	105.093	Y	Y	227	F	Y	Aromatic	Y	165.192	Y	Y
Y	14	S	N	Hydroxylic	Y	105.093	Y	Y	221	F	Y	Aromatic	Y	165.192	Y	Y
Y	12	T	Y	Hydroxylic	Y	119.119	Y	Y	220	F	Y	Aromatic	Y	165.192	Y	Y
N	12	A	N	Aliphatic	N	89.094	N	N	220	F	Y	Aromatic	Y	165.192	Y	Y
N	12	G	N	Aliphatic	N	75.067	N	N	220	F	Y	Aromatic	Y	165.192	Y	Y
N	12	A	N	Aliphatic	N	89.094	N	N	220	F	Y	Aromatic	Y	165.192	Y	Y
N	12	A	N	Aliphatic	N	89.094	N	N	220	F	Y	Aromatic	Y	165.192	Y	Y

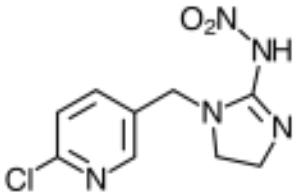
Amino Acid info

ID	Name	Side Chain	Size
A	Alanine	Aliphatic	89.094
C	Cysteine	Sulfur-Containing	121.154
D	Aspartic Acid	Acidic	133.104
E	Glutamic Acid	Acidic	147.131
F	Phenylalanine	Aromatic	165.192
G	Glycine	Aliphatic	75.067
H	Histidine	Basic	155.156
I	Isoleucine	Aliphatic	131.175
K	Lysine	Basic	146.189
L	Leucine	Aliphatic	131.175
M	Methionine	Sulfur-Containing	149.208
N	Asparagine	Amidic	132.119
P	Proline	Aliphatic	115.132
Q	Glutamine	Amidic	146.146
R	Arginine	Basic	174.203
S	Serine	Hydroxylic	105.093
T	Threonine	Hydroxylic	119.119
V	Valine	Aliphatic	117.148
W	Tryptophan	Aromatic	204.228
X	Unknown	Unknown	-100.0
Y	Tyrosine	Aromatic	181.191

Flupyradifurone

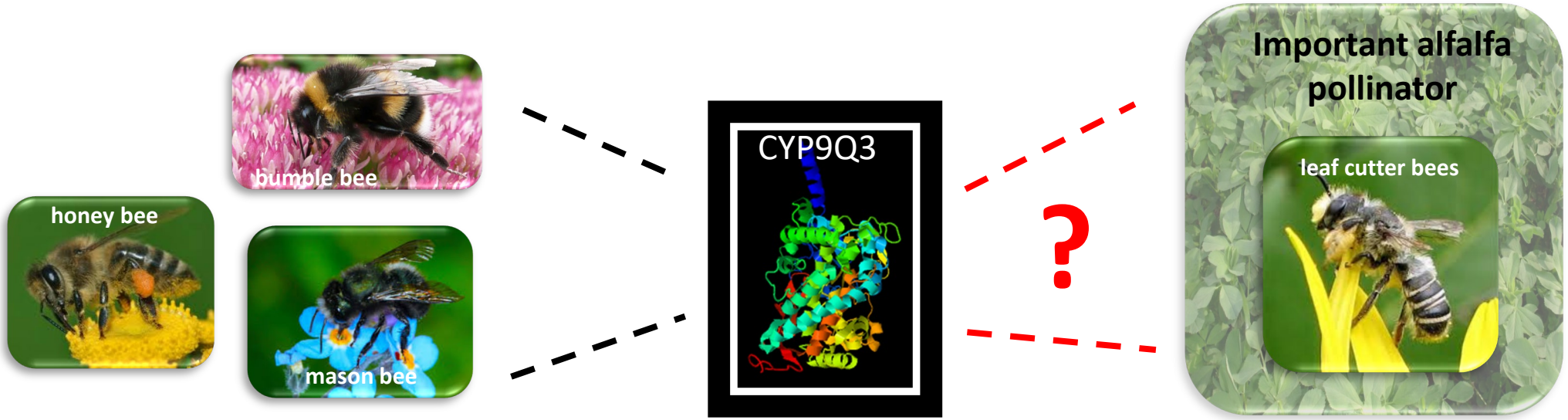


Flupyradifurone

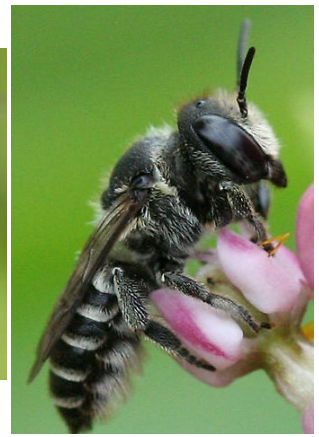


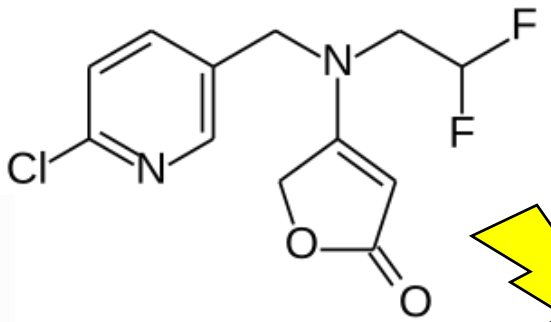
Imidacloprid

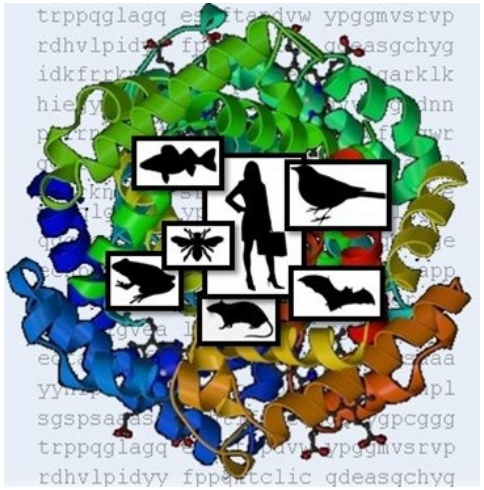
- Pesticide is in the same MOA category as neonicotinoids and sulfoxaflors
  - Has a very similar chemical structure to neonicotinoids
  - Acts on sucking pest species – aphids, stink bugs, and white flies
- Honey bee, mason bees, and bumble bees have similar sensitivity
- Very toxic to leaf cutter bees (*Megachile rotundata*)



SeqAPASS to evaluate conservation of CYP9Q across species







Pesticide Interaction

Mao et al. 10.1073/pnas.1109535108

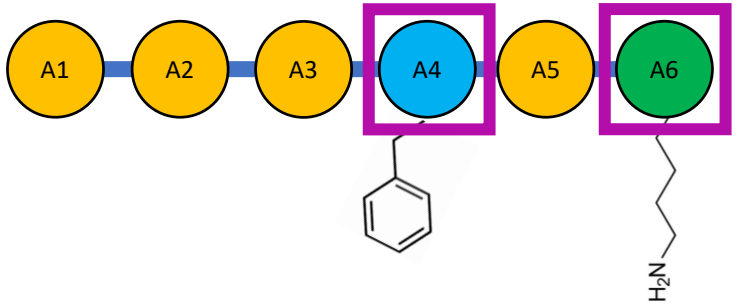
- Mao et al. 2011 found a critical amino acid that would bind to the pesticide for metabolism.
- CYP9Q3 has been shown to bind bulkier molecules and demonstrate a broader substrate specificity suggesting a more accessible/non-selective active site.
- There is a catalytic pocket formed by phenylalanine residues

SeqAPASS Level 3

Bee CYP9Q3

Catalytic pocket: **F123, F305, F374, I491**

Negative charge contact Pesticide: **K219**



Side Chain Classification: acidic, basic, aromatic, etc.

MW as surrogate for size: > 30g/mol different size

Susceptibility different than template = Both Class and Size Differ

EPA United States Environmental Protection Agency

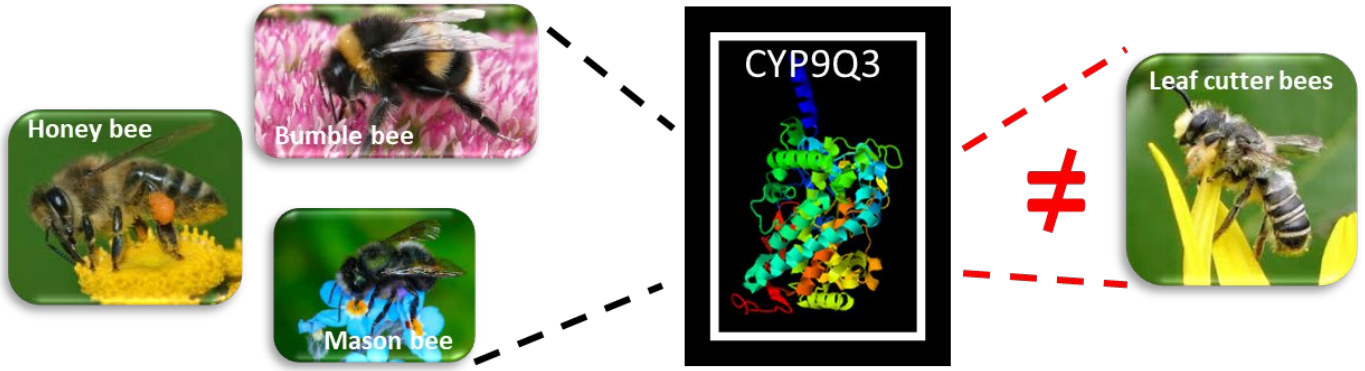
SeqAPASS Level 3 output

Protein Name	Common Name	Similar Susceptibility	Amino Acid and Position
cytochrome P450 9Q3	Asiatic honeybee	Y	219K
cytochrome P450 9e2	Honey bee	Y	219K
PREDICTED: cytochrome P450 9e2-like isoform X1	Giant honeybee	Y	219K
Cytochrome P450 9e2	Stingless bees	Y	454K
cytochrome P450 9e2	Buff-tailed bumblebee	Y	220K
PREDICTED: cytochrome P450 9e2-like	Orchid bees	Y	219K
uncharacterized protein LOC100740972	Common eastern bumble bee	Y	220K
PREDICTED: cytochrome P450 9e2-like	Little honeybee	Y	216K
cytochrome P450 9e2-like	Carpenter bees	Y	218K
cytochrome P450 mono-oxygenase	Red mason bee	Y	219T
PREDICTED: cytochrome P450 9e2-like	Digger bees	Y	217L
PREDICTED: cytochrome P450 9e2-like	Bees	Y	219V
PREDICTED: cytochrome P450 9e2-like	Alfalfa leafcutting bee	N	215W

Total Match
Partial Match
Not a Match

Y Yes
N No

Conclusions



- Limited sequence information available for bee species
  - Advocate for genome sequencing and annotation of key species
    - Predictive approaches require sequence data
- Limited data on species-pesticide interactions that highlight key amino acids, which can be utilized by SeqAPASS
  - SeqAPASS requires previous knowledge/literature of critical residues to help inform a level 3 amino acid comparisons between species
- There is evidence of conservation of CYP9Q across bee species
  - Level 3 critical individual amino acid comparisons can be run on SeqAPASS
    - Species-specific differences between *Apis/Osmia/Bombus* and *Megachile rotundata* @ K219, indicating that the pesticide-CYP9Q interaction may differ