Incorporating new features for rapid interpretation of results from the Sequence Alignment to Predict Across Species Susceptibility (SeqAPASS) Tool

Donovan J. Blatz[1], Cody Simmons[2], Audrey Wilkinson[2], Tom Transue[2], Carlie A. LaLone[3]
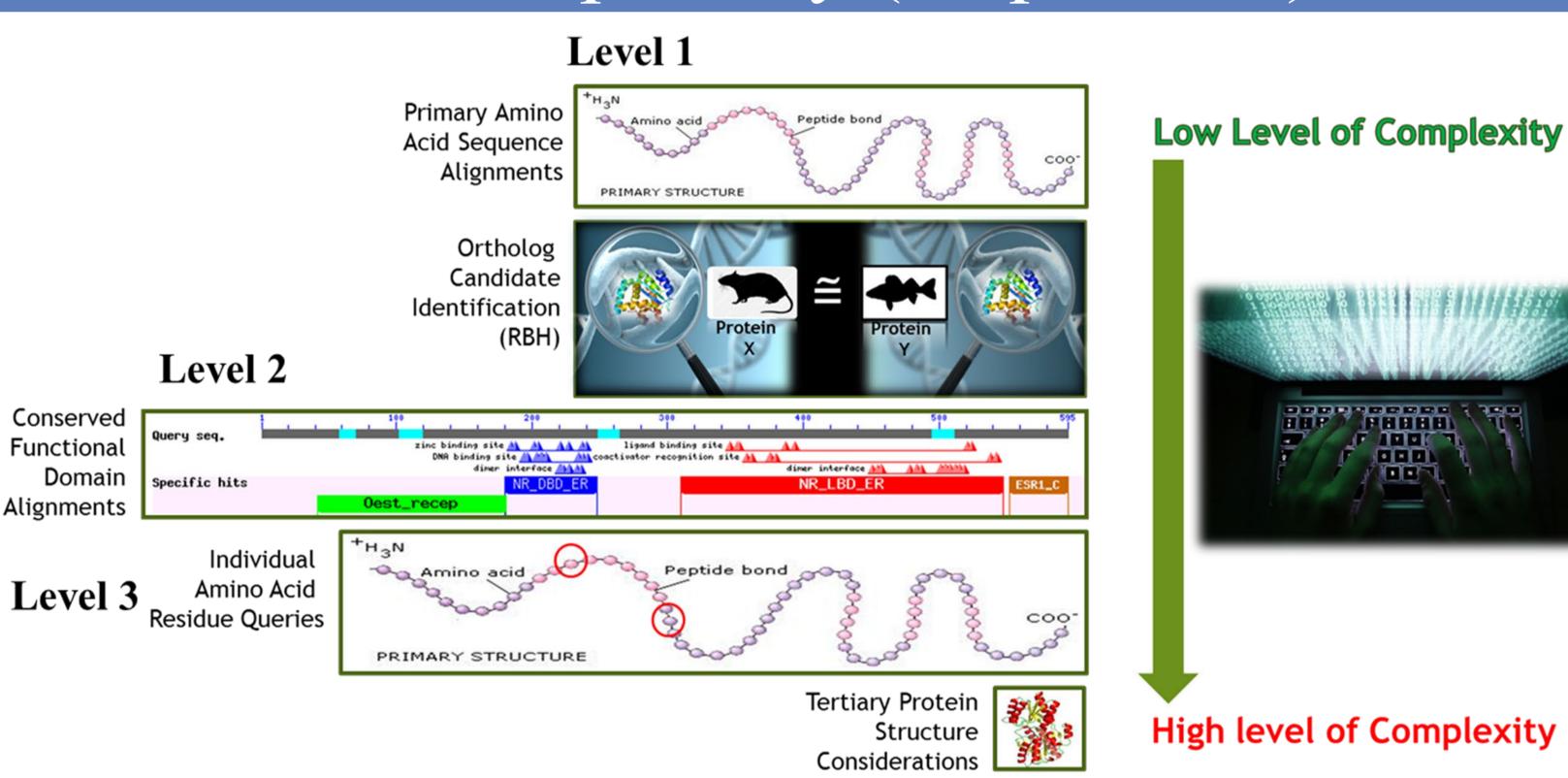
1: ORISE Fellow, U.S. Environmental Protection Agency, Office of Research and Development, Great Lakes Toxicology and Ecology Division, Duluth, MN 55804
2: General Dynamics Information Technology, Research Triangle Park, NC 27711
3 U.S. Environmental Protection Agency, Office of Research and Development, Great Lakes Toxicology and Ecology Division, Duluth, MN 55804

www.epa.gov

Introduction

- The US EPA is evaluating alternatives to animal testing, making computational tools that evaluate chemical safety and cross species susceptibility both important and necessary
- Two major updates have been integrated in the Sequence Alignment to Predict Across Species Susceptibility (SeqAPASS) v5.0 pipeline, including a heat map visualization of individual amino acid comparisons (Level 3) and a Decision Summary Report that allows for a combination of susceptibility predictions across the Levels (1,2, and 3) of the SeqAPASS evaluations
- A case example will demonstrate the utility of the new SeqAPASS features focused on evaluating an individual amino acid residue across bee species

Sequence Alignment to Predict Across Species Susceptibility (SeqAPASS)

Level 1 — Primary Amino Acid Sequence Alignments / Ortholog Candidate Identification (RBH) — Low Level of Complexity

Level 2 — Conserved Functional Domain Alignments

Level 3 — Individual Amino Acid Residue Queries — Tertiary Protein Structure Considerations — High level of Complexity

- SeqAPASS was developed to create a strategic and automated approach for assessing protein similarity and exploring taxonomic domain of applicability for biological pathways. We utilize available information, from publicly accessible databases, for our SeqAPASS evaluations. This allows us to be flexible in the approach for each SeqAPASS run

- Level 1 allows for comparison of the full-length protein sequence from a known sensitive species to all other species with sequence information available
- Level 2 evaluates conservation of functional domains, such as a ligand binding domain, across species
- Level 3 requires previous knowledge of critical amino acid residues that are important for the chemical-protein interaction for comparisons across species
- Each Level of the SeqAPASS evaluation provides an additional line of evidence toward conservation of a protein and can be used to predict chemical susceptibility

Understanding Protein Conservation

1. Must know the molecular target (e.g. pesticides)
2. Must identify a 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

New Features in SeqAPASS v5.0

Heat Map Visualization

- The Heat Map is a feature that allows the user to have a visual representation of the chosen amino acid(s) for a single Level 3 run
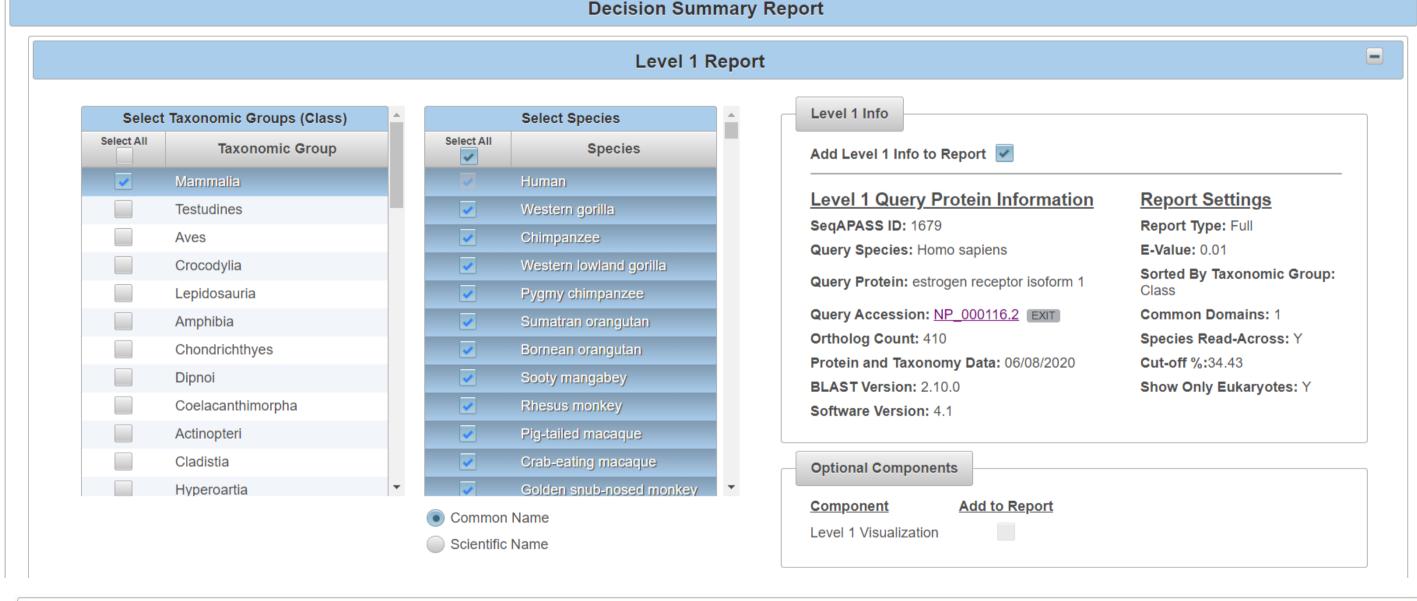
Common Name	Similar Susceptibility	Amino Acid 1	Amino Acid 2	Amino Acid 3	Amino Acid 4
Human	Y	133E	135E	137S	584I
Western painted turtle	Y	127E	129D	131S	577V
Lappet-faced vulture	N	125E	127E	129G	575I
Nile crocodile	Y	127E	129D	131S	575I
Split-tongued squamates	Y	127E	129E	131N	576V
Japanese giant salamander	N	126E	128E	130G	573L
West African lungfish	Y	131E	133E	135S	580G

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

Total Match / Partial Match / Not a Match

Figure 1: An example of a SeqAPASS Level 3 Heat Map simple report showing amino acid susceptibility predictions utilizing colors to denote Total Match, Partial Match, Not a Match

Decision Summary Report

- The Decision Summary (DS) Report is a feature that gives the user options to design a single custom output page to concisely view results from all Levels of the SeqAPASS evaluation for completed jobs
- The output is customizable to include visualizations and susceptibility predictions that can be downloaded in a PDF format

Figure 2: An example set up for the Decision Summary Report showing all three levels

Case Example

- A case example showing the set up of the DS Report and Heat Map to display a possible critical amino acid utilized in binding of pesticides for key pollinators
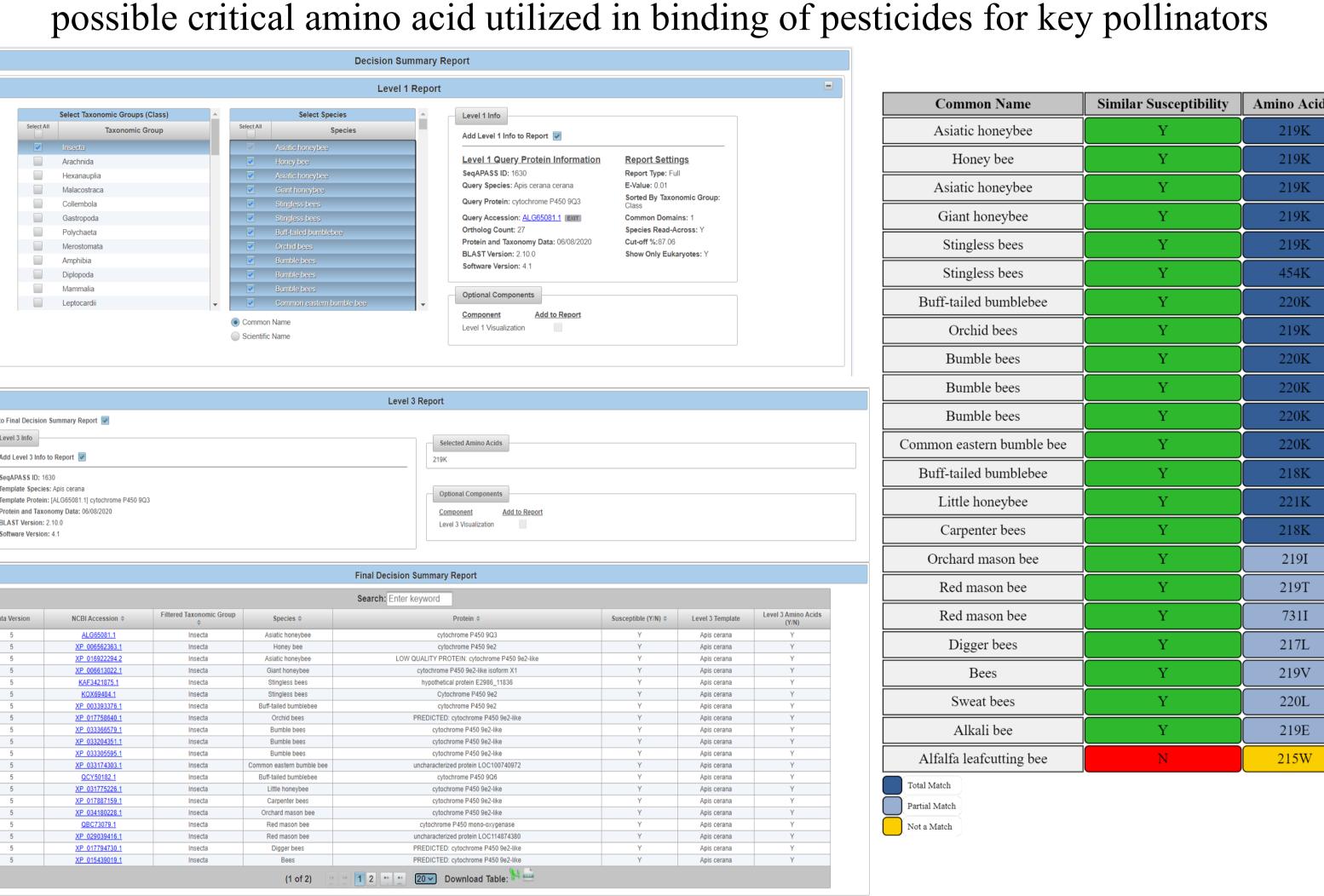

Common Name	Similar Susceptibility	Amino Acid 1
Asiatic honeybee	Y	219K
Honey bee	Y	219K
Asiatic honeybee	Y	219K
Giant honeybee	Y	219K
Stingless bees	Y	219K
Stingless bees	Y	454K
Buff-tailed bumblebee	Y	220K
Orchid bees	Y	219K
Bumble bees	Y	220K
Bumble bees	Y	220K
Common eastern bumble bee	Y	220K
Buff-tailed bumblebee	Y	218K
Little honeybee	Y	221K
Carpenter bees	Y	210K
Orchard mason bee	Y	219I
Red mason bee	Y	219T
Red mason bee	Y	731I
Digger bees	Y	217L
Bees	Y	219V
Sweat bees	Y	220L
Alkali bee	Y	219E
Alfalfa leafcutter bee	N	215W

Total Match / Partial Match / Not a Match

Figure 3a: The set up for the case study in the Decision Summary Report tab. Level 1 and Level 3 runs that were pushed to the DS Report and then added to be displayed in the Final Decision Summary Report table. **3b:** The Level 3 visualization (Heat Map) displaying the Yes (Y)/No (N) susceptibility prediction based off the template specie's amino acid. Amino acid comparison is based on side chain classification and MW as surrogate for size.

Summary and Conclusions

- SeqAPASS gives decision makers and risk assessors the ability to quickly interpret important susceptibility predictions and with the addition of the Heat Map and Decision Summary Report, it allows for quicker discernment of important results
- The Heat Map uses color to denote differences in between the amino acid(s) to the template specie
- The DS Report's customizable features give users the option to design a downloadable report to concisely show all important information in an easy viewable format

References

1. **LaLone et al. 2016**, SeqAPASS: A Web-Based Tool for Addressing the Challenges of Cross-Species Extrapolation of Chemical Toxicity. *Toxicol Sci*, 153 (2): 228-245 (2016)
2. **Moa et al. 2011**, CYP9Q-mediated detoxification of acaricides in the honey bee (Apis mellifera). *PNAS*, 108 (31): 12657-12662 (2011)

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