Environmental Toxicology and Chemistry—Volume 42, Number 1—pp. 71–87, 2023
Received: 5 May 2022 | Revised: 30 June 2022 | Accepted: 12 October 2022

Environmental Toxicology

Defining the Biologically Plausible Taxonomic Domain of Applicability of an Adverse Outcome Pathway: A Case Study Linking Nicotinic Acetylcholine Receptor Activation to Colony Death

Marissa A. Jensen,[a,b] Donovan J. Blatz,[c] and Carlie A. LaLone[b,*]

[a]Department of Biology, Swenson College of Science and Engineering, University of Minnesota Duluth, Duluth, Minnesota, USA
[b]US Environmental Protection Agency, Center for Computational Toxicology and Exposure, Great Lakes Toxicology and Ecology Division, Duluth, Minnesota, USA
[c]Oak Ridge Institute for Science and Education, Duluth, Minnesota, USA

Abstract: For the majority of developed adverse outcome pathways (AOPs), the taxonomic domain of applicability (tDOA) is typically narrowly defined with a single or a handful of species. Defining the tDOA of an AOP is critical for use in regulatory decision-making, particularly when considering protection of untested species. Structural and functional conservation are two elements that can be considered when defining the tDOA. Publicly accessible bioinformatics approaches, such as the Sequence Alignment to Predict Across Species Susceptibility (SeqAPASS) tool, take advantage of existing and growing databases of protein sequence and structural information to provide lines of evidence toward structural conservation of key events (KEs) and KE relationships (KERs) of an AOP. It is anticipated that SeqAPASS results could readily be combined with data derived from empirical toxicity studies to provide evidence of both structural and functional conservation, to define the tDOA for KEs, KERs, and AOPs. Such data could be incorporated in the AOP-Wiki as lines of evidence toward biological plausibility for the tDOA. We present a case study describing the process of using bioinformatics to define the tDOA of an AOP using an AOP linking the activation of the nicotinic acetylcholine receptor to colony death/failure in *Apis mellifera*. Although the AOP was developed to gain a particular biological understanding relative to *A. mellifera* health, applicability to other *Apis* bees, as well as non-*Apis* bees, has yet to be defined. The present study demonstrates how bioinformatics can be utilized to rapidly take advantage of existing protein sequence and structural knowledge to enhance and inform the tDOA of KEs, KERs, and AOPs, focusing on providing evidence of structural conservation across species. *Environ Toxicol Chem* 2023;42:71–87. © 2022 The Authors. *Environmental Toxicology and Chemistry* published by Wiley Periodicals LLC on behalf of SETAC. This article has been contributed to by U.S. Government employees and their work is in the public domain in the USA.

Keywords: Adverse outcome pathway; ecotoxicology; species extrapolation; predictive toxicity; SeqAPASS; pollinator; bioinformatics

INTRODUCTION

The adverse outcome pathway (AOP) framework provides a means to organize existing knowledge and data from the literature to understand causal linkages connecting a molecular initiating event (MIE) to an adverse outcome (AO) at a biological level of organization relevant to risk assessment (Ankley et al., 2010). Key events (KEs) that represent measurable changes at a given level of biological organization are connected via KE relationships (KERs) that capture the evidence for a causal relationship between one KE and the next (Villeneuve et al., 2014). The AOP framework also captures weight of evidence (WoE) for the causal linkages through considering empirical evidence and the biological plausibility of KERs (Becker et al., 2017; Villeneuve et al., 2014). Typically, AOPs are developed considering one or a handful of species for which

This article contains online-only Supporting Information.
This is an open access article under the terms of the Creative Commons Attribution-NonCommercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.
* Address correspondence to LaLone.Carlie@epa.gov
Published online 20 October 2022 in Wiley Online Library (wileyonlinelibrary.com).
DOI: 10.1002/etc.5501

© 2022 The Authors.

empirical data describing the biological pathway exist. Although developers tend to assume broader species coverage, based on biological plausibility, in AOPs actual species-specific evidence supporting the taxonomic domain of applicability (tDOA) remains relatively narrow, limiting confidence in application across species. The tDOA of an AOP is an important consideration in development and use of an AOP in regulatory decision-making, particularly when an understanding of the appropriateness of surrogate species, relative to broader species representation, may be important. To date, the tDOA is defined by the specific species used in the studies describing the KEs; and in some cases, text descriptions assume broader taxonomic coverage with limited documented evidence to do so. To define the tDOA of a KE, conservation of structure and function are the two primary considerations (Organisation for Economic Co-operation and Development [OECD], 2018). Evaluating structure can include determining if a biological entity (e.g., gene, protein, organ, tissue) can be measured or observed, is present, and/or is conserved in the taxa of interest (OECD, 2018). Evaluating function can include determining whether the biological object plays the same role in other taxa of interest (OECD, 2018). To enhance the tDOA descriptions for the KEs, the KERs, and the overall AOP, the objective of the present study was to demonstrate how bioinformatics approaches can be used to more thoroughly define the tDOA by extrapolating existing knowledge across species. Evaluation of conserved biology is of interest particularly when limited empirical data exist for the majority of species that may be exposed to chemical stressors in the environment.

As previously described, the US Environmental Protection Agency's (USEPA's) Sequence Alignment to Predict Across Species Susceptibility (SeqAPASS) tool can be used in a hierarchical framework alongside a collection of in vitro and in vivo toxicity test data to determine if structural conservation and functional conservation exist across the AOP in other species (Ankley et al., 2016). The publicly available web-based SeqAPASS tool evaluates cross-species protein sequence and structural similarities and differences through three levels of evaluation to aid inference of chemical susceptibility across species (LaLone et al., 2016). Level 1 compares the primary amino acid sequence of a molecular target based on sequence similarity and identifies orthologs or sequences that are likely to have diverged from a speciation event and maintained similar function. Level 2 evaluates conservation of selected functional domains. Level 3 compares critical individual amino acid residues that are important for protein–ligand interactions, protein–protein interactions, or protein function (Doering et al., 2018; LaLone et al., 2016). The SeqAPASS tool has many applications, but importantly, it provides lines of evidence toward structural conservation, which is an element used to define the tDOA of an AOP (LaLone et al., 2016; OECD, 2018).

We describe how results from a computational tool, SeqAPASS, can be used at multiple levels of organization and coupled with available empirical evidence to define the biologically plausible tDOA for KEs, KERs, and the entire AOP. The objective is to demonstrate the utility of this bioinformatics approach in a manner consistent with the available and developing features within the AOP-Wiki (https://aopwiki.org/). Further application of SeqAPASS using a case example with an AOP extracted from an existing AOP network highlights the utility of these approaches and the challenges.

FIGURE 1: Illustration of the adverse outcome pathway (AOP) network and the selected AOP used for the case study to define the plausible taxonomic domain of applicability. The selected AOP links the activation of the nicotinic acetylcholine receptor to the adverse outcome of colony death/failure (AOP 89; LaLone, 2021). Out of the six AOPs developed by LaLone et al. (2017; AOPs 88, 89, 77, 87, 79, and 178), AOP 89 was chosen to use as the case example. ID = identifier; KE = key event; MIE = molecular initiating event; AO = adverse outcome; nAChR = nicotinic acetylcholine receptor; KER = KE relationship.

© 2022 The Authors.

In 2017, an AOP network was described as connecting activation of the nicotinic acetylcholine receptor (nAChR) to the AO of colony death/failure (LaLone et al., 2017). The AOP network was developed to provide a better understanding of the causal linkages that may ensue from interactions with prototypical stressors, neonicotinoids, and the MIE of nAChR activation in honey bees (*Apis mellifera*). Honey bees have been at the forefront of scientific and public concern in recent years because of increased colony death and failure. Although there are many possible stressors that may contribute, chemical stressors, such as neonicotinoid insecticides, have been drawing significant attention (LaLone et al., 2017; Potts et al., 2010). Neonicotinoids are one of the most widely used insecticides around the world, comprising a market share of >25% of the total global insecticide market in 2014 (Bass et al., 2015). They act on the nAChR, causing neurotoxicity in pest insects, such as those found in the Hemiptera, Coleoptera, and Lepidoptera orders (Jeschke et al., 2013).

Although there are concerns for *Apis* bee health, declines in non-*Apis* bee populations, such as bumble bees and solitary bees, are also concerning (Blacquière et al., 2012). Further, recent research suggests that non-*Apis* bees may also be vulnerable to pesticides based on analysis of traits related to pesticide exposure and intrinsic sensitivity across species, such as foraging range and body length (Schmolke et al., 2021). Therefore, the AOP network describing nAChR activation leading to colony death/failure (focused on *A. mellifera*) was selected with specific attention on one of the described AOPs (AOP 89; LaLone, 2021; Figure 1) as a case example for demonstration of how SeqAPASS could aid in defining the tDOA. Specifically, the tDOA for KEs (guided by existing empirical studies as support) and the biologically plausible tDOA for KEs and KERs (plausible in that knowledge is extrapolated using additional lines of scientific evidence from bioinformatics or other computational approaches), along with the overall AOP, were evaluated to determine whether there are lines of evidence to support broader applicability to other bee species (LaLone et al., 2017). Motivation to incorporate tDOA descriptions across AOPs that have remained relatively static in the AOP-Wiki has grown over the years for application purposes. The present study aimed to define the taxonomic relevance of the AOP using SeqAPASS and to lay the foundation of an approach that expands the taxonomic space of an AOP with rapid computational methods that can aid in identifying where structural conservation across species is most likely to exist.

METHODS

From the AOP network, a single AOP was selected linking nAChR activation to colony death/failure (AOP 89, LaLone, 2021), to demonstrate the utility of computational approaches in expanding on and defining the tDOA for KEs, KERs, and the entire AOP (Figure 1; LaLone et al., 2017). First, the empirical tDOA was identified based on which species were cited in the text of the KER descriptions providing support for the relationship or in the text of a KE indicating that the event had been measured in a given species. To expand on the tDOA for the KEs, which captures existing empirical data, and to define the biologically plausible tDOA for the KEs and in turn KERs using computational methods, nine proteins were identified to be involved in this AOP (Table 1). These nine proteins served as the query proteins to submit to the SeqAPASS tool. The resulting computationally derived data on protein similarity across a range of species help to add scientific evidence to support definition of the biologically plausible tDOA beyond the scope of specific species cited in the KE descriptions and/or the empirical evidence supporting the KERs (LaLone et al., 2016). For each of the nine proteins identified as being relevant to the AOP, SeqAPASS analyses at Levels 1, 2, and 3 were conducted. Evaluations at Levels 1 and 2 of protein conservation included all taxonomic groups, aligning sequences across invertebrates, vertebrates, plants, and so forth. However, to simplify and demonstrate the utility of the bioinformatics approach, the Level 3 evaluations focused on the specific question of whether the AOP developed with a focus on honey bee (*A. mellifera*) may be relevant for non-*Apis* bees (e.g., bumble bees and solitary bees).

TABLE 1: Identified proteins in the key events (KEs), respective KE identifiers of adverse outcome pathway 89 (LaLone, 2021), and associated KEs from Figure 1

Protein	AOP-Wiki KE ID	Associated KE
Nicotinic acetylcholine receptor	559,663	MIE, KE1
Calmodulin	1243	KE2
Adenylyl cyclase	1243	KE2
Protein kinase A	1243	KE2
Calcium–calmodulin-dependent protein kinase II	1243	KE2
cAMP-responsive element-binding protein	1243	KE2
Vitellogenin precursor	1108	KE5
Juvenile hormone acid *O*-methyltransferase	1108	KE5
Methyl farnesoate epoxidase	1108	KE5

AOP = adverse outcome pathway; ID = identifier; cAMP = cyclic 3′5′-adenosine monophosphate; MIE = molecular initiating event.

To submit a SeqAPASS job, a query species is selected. Because the AOP network was originally developed for *A. mellifera*, all SeqAPASS Level 1 analyses used *A. mellifera* as the query species (Table 2). For Level 2 SeqAPASS analyses, functional domains were selected for each respective protein from the National Center for Biotechnology Information's (NCBI's) Conserved Domain Database or from the literature, and sequence similarity for those specific domains was evaluated in SeqAPASS (Table 2). For Level 3 SeqAPASS analyses, an extensive literature search was performed to determine the critical amino acids for each relevant protein function, protein–chemical interaction, or protein–ligand interaction. The terms used for the literature search in Google Scholar included *insect *protein name* amino acids*, *insect *protein name* binding site*, *insect *protein name**, *honeybee *protein name**, **protein name* structure*, and *Apis mellifera *protein name**. If the search was for amino acids

 © 2022 The Authors.

TABLE 2: SeqAPASS Level 1, Level 2, and Level 3 inputs for all early key event proteins evaluated

Targets (KE ID)	Level 1 query species (protein accession)	Level 2 domain (domain accession)	Level 3 template species (protein accession)	Level 3 amino acids	Amino acid functions	References
nAChR α1 subunit (559,663)	*Apis mellifera* (NP_001091690.1)	Neurotransmitter-gated ion-channel ligand-binding domain (pfam02931)	*Drosophila melanogaster* (NP_524481.2)	R57, E78, K140, W170, Y220, Y227	Neonicotinoid binding	Matsuda et al. (2020); Matsuda (2020)
				S221	Neonicotinoid binding	Matsuda (2020)
nAChR α2 subunit (559,663)	*A. mellifera* (AJE70260.1)	Neurotransmitter-gated ion-channel ligand-binding domain (pfam02931)	*D. melanogaster* (NP_524482.1)	P242	Neonicotinoid binding	Shimomura et al. (2004)
				W190, Y241, Y248	Neonicotinoid binding	Wang et al. (2016); Shimomura et al. (2004)
nAChR β1 subunit (559,663)	*A. mellifera* (AAY87897.1)	Neurotransmitter-gated ion-channel ligand-binding domain (pfam02931)	*D. melanogaster* (NP_523927.2)	R81, L141, V143	Neonicotinoid binding	Matsuda et al. (2020); Wang et al. (2016)
Calmodulin (1243)	*A. mellifera* (XP_006565317.1)	Calcium-binding motif (cd00051)	*Rattus norvegicus* (P0DP29.1)	D21, D23, D25, T27, E32	Calcium binding	Babu et al. (1988)
Adenylyl cyclase (1243)	*A. mellifera* (CCE60554.1)	Catalytic domain (pfam00211)	*A. mellifera* (CCE60554.1)	D381, D425	Mg^{2+} binding	Balfanz et al. (2012)
				K1117, D1193	ATP binding	
CaMKII (1243)	*A. mellifera* (NP_001128422.1)	Catalytic domain (cd14086)	*A. mellifera* (NP_001128422.1)	K43	ATP binding	Coultrap and Bayer (2012);
				T254, T287, T306, T307, S315	Phosphorylation	Wayman et al. (2011)
PKA (1243)	*A. mellifera* (CAG27571.1)	Cyclic nucleotide–binding domain (pfam00027)	*A. mellifera* (CAG27571.1)	G194, E195, A197, R204, A205	cAMP binding, PBC domain A	Leboulle and Müller (2004); Berman et al. (2005)
				G317, E318, A320, R327, A328	cAMP binding, PBC domain B	
CREB (1243)	*A. mellifera* (XP_006570112.1)	pKID domain (pfam02173)	*Rattus norvegicus* (P15337.1)	R130, R131	PKA recognition	Gonzalez et al. (1989); Montminy (1997); Gonzalez et al. (1991); Mayr and Montminy (2001); Kandel (2012)
				P132, Y134	Part of PKA recognition site	
				S133	Phosphoacceptor site	

The query species for all SeqAPASS Level 1 evaluations was *Apis mellifera*. The information for Level 3 template species, amino acids, and amino acid functions was gathered from the respective references listed. Protein and domain accessions correspond to those reported in the National Center for Biotechnology Information (https://www.ncbi.nlm.nih.gov/protein).
AOP = adverse outcome pathway; ATP = adenosine triphosphate; cAMP = cyclic 3′5′-adenosine monophosphate; CaMKII = calcium–calmodulin-dependent protein kinase II; CREB = cAMP-responsive element binding protein; KE ID = key event identifier from the AOP-Wiki (https://aopwiki.org/); nAChR = nicotinic acetylcholine receptor; PBC = phosphate-binding cassette; PKA = protein kinase A; SeqAPASS = Sequence Alignment to Predict Across Species Susceptibility. Amino acid abbreviations: A = alanine; D = aspartic acid; E = glutamic acid; G = glycine; K = lysine; L = leucine; P = proline; R = arginine; S = serine; T = threonine; V = valine; W = tryptophan; Y = tyrosine.

© 2022 The Authors.

critical for a specific function, search terms that included the specific function were used (e.g., *protein kinase A [PKA] 3′5′-adenosine monophosphate [cAMP] binding* was used for PKA). The Level 3 template sequences and critical amino acids were selected from those reported in the literature (Table 2). All *Apis* and non-*Apis* bees were chosen from the Level 1 output and aligned in Level 3 with the template sequence. Any aligned protein accessions annotated as "partial" or "low-quality" were included but were noted as such for later review. No sequences annotated as "hypothetical" were chosen for Level 3 evaluations comparing *Apis* to non-*Apis* species. Importantly, Level 3 SeqAPASS results allow for species-specific comparisons of sequence similarity. In our study, the scope of the Level 3 evaluation was reduced

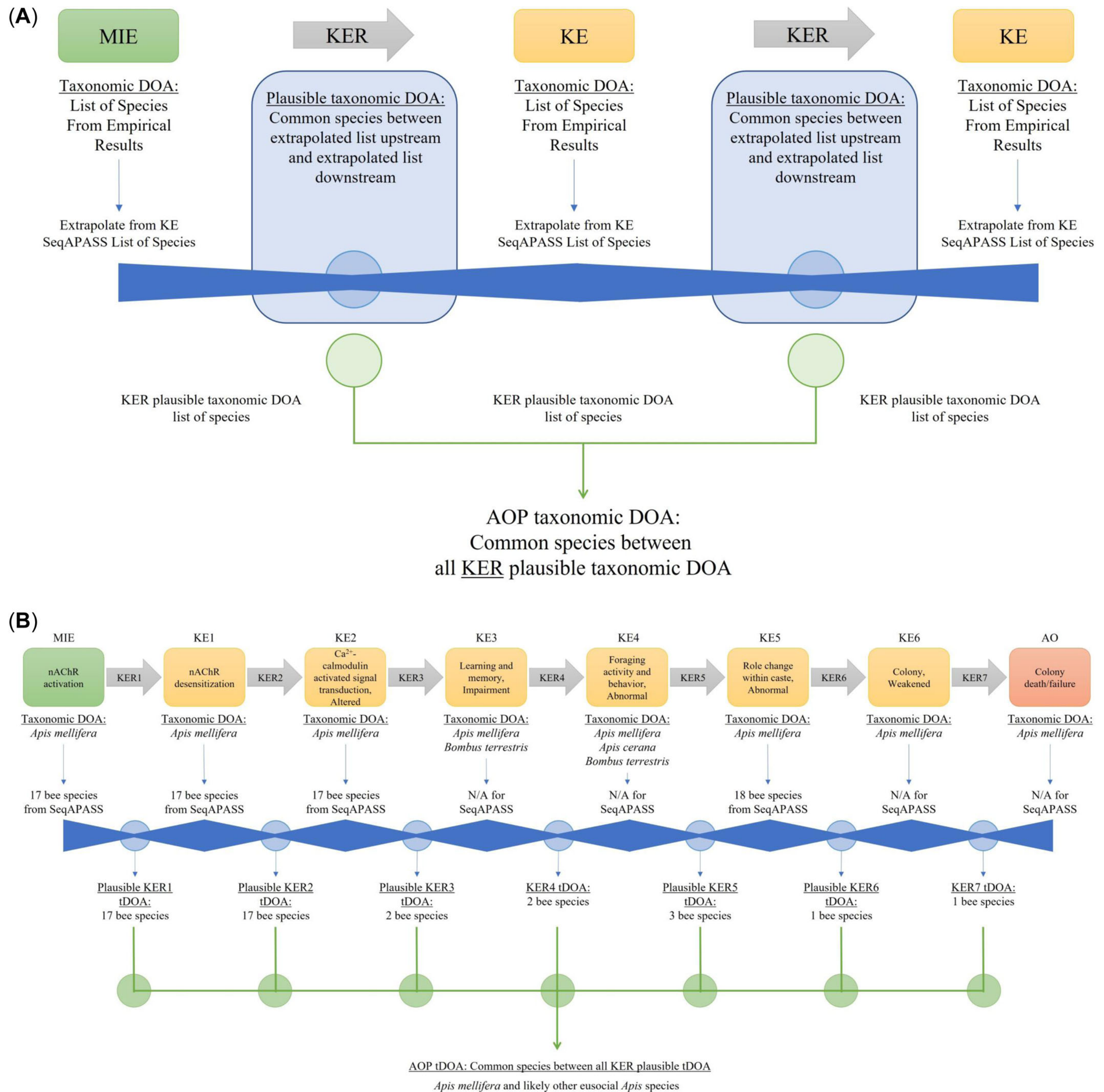

FIGURE 2: (**A**) Determining the taxonomic domain of applicability (tDOA) for key events, key event relationships, and the adverse outcome pathway (AOP). (**B**) An example of determining the biologically plausible tDOA for an AOP using nicotinic acetylcholine receptor activation leading to colony death/failure as a case study. In this case study, the biologically plausible tDOA for the AOP was determined to be *Apis mellifera*. In this figure, *plausible taxonomic DOA* refers to the biologically plausible tDOA. MIE = molecular initiating event; KER = key event relationship; KE = key event; SeqAPASS = Sequence Alignment to Predict Across Species Susceptibility; AO = adverse outcome; nAChR = nicotinic acetylcholine receptor; N/A = not applicable.

 © 2022 The Authors.

to bee species only. Results from Levels 1, 2, and 3 were combined for each protein, with the exception of three proteins from KE5 (vitellogenin [Vg], juvenile hormone acid *O*-methyltransferase, and methyl farnesoate epoxidase) where only data from Levels 1 and 2 were generated.

To define the biologically plausible tDOA for the AOP, the tDOA for each KE and KER must be defined first (Figure 2A). To define the biologically plausible tDOA for each KE, the list of species collected from empirical data describing the KE was combined with the SeqAPASS-derived list of species. When the KE had more than one protein involved in the description (e.g., KE2, KE5) or a protein had more than one subunit involved (e.g., MIE, KE1), the combined SeqAPASS output across species and proteins/subunits was used to define the list of species that would be incorporated for the determination of the biologically plausible tDOA (Figures 3 and 4). These species were identified by taking the narrowest point of overlapping species conservation between all the proteins/subunits evaluated for that KE (Figures 3 and 4). Once biologically plausible tDOAs were defined for all KEs, the information was used to define the biologically plausible tDOAs for the KERs by comparing the upstream and downstream tDOA for the KEs from both the empirical evidence and SeqAPASS-derived species lists (if available), identifying the overlapping species between each KE pair (Figure 2A). The biologically plausible tDOA for the KER, therefore, can have no broader species coverage than the more restrictive of the two KEs. After providing evidence for defining the biologically plausible tDOAs for each KER, the biologically plausible tDOA for the entire AOP can be determined. This is accomplished by comparing the lists of species defining the biologically plausible tDOA of all KERs in the AOP and identifying the overlapping species among all KERs (Figure 2A). From this tDOA comparison among KERs, a list of species was determined to define the biologically plausible tDOA for the AOP.

RESULTS

To evaluate the tDOA for the AOP describing the causal linkages from nAChR activation to colony death/failure, the protein targets involved in KEs were evaluated for structural conservation using SeqAPASS (Table 1). The following subsections describe the evaluation of structural conservation for the protein targets involved in the KEs, including description of the biological role of each protein in the KEs and the rationale for choosing certain protein subunits for SeqAPASS evaluation when applicable. Further, analyses of SeqAPASS Levels 1 and 2, the rationale for choosing the critical amino acids for Level 3 analyses, and the SeqAPASS Level 3 analyses are thoroughly described in subsequent sections.

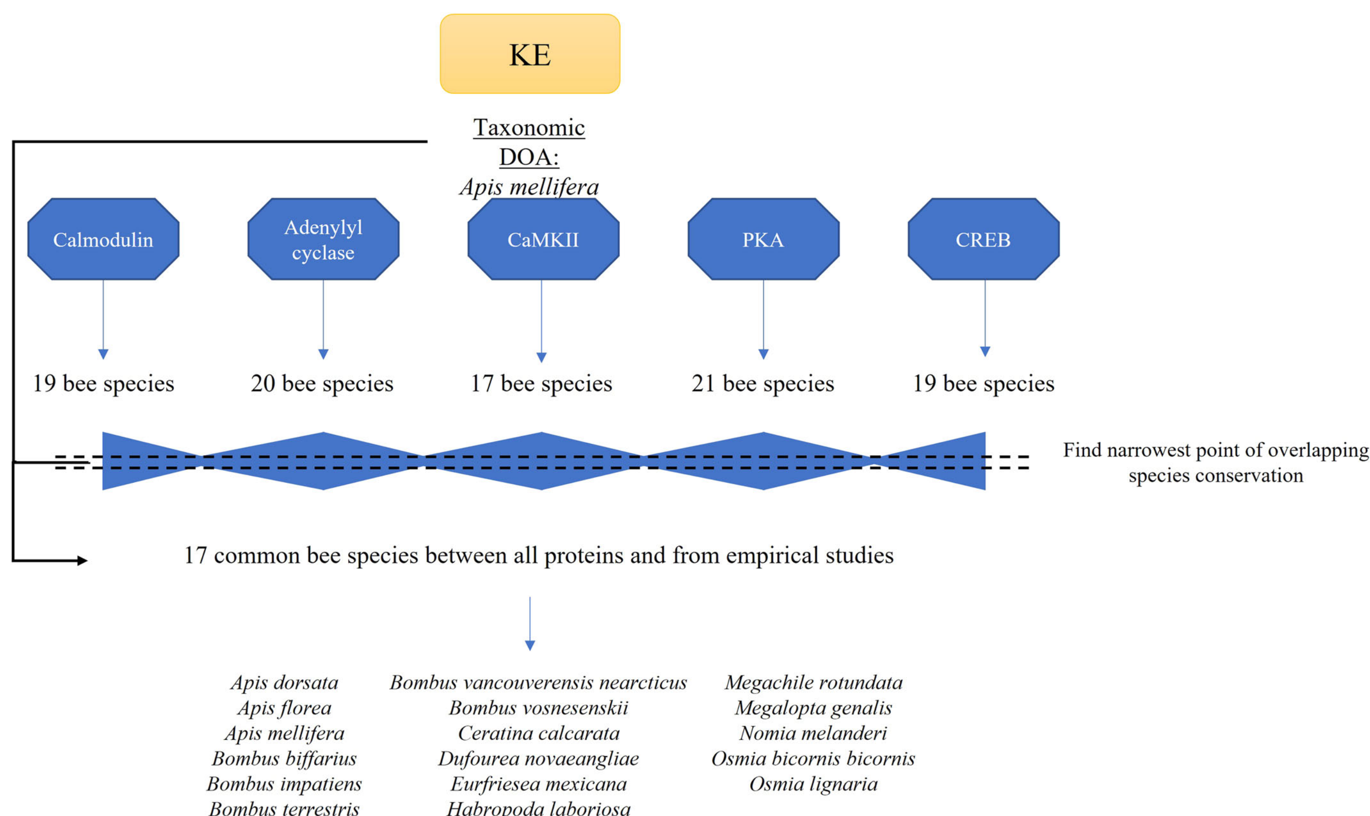

FIGURE 3: Determining biologically plausible taxonomic domain of applicability (tDOA) for a key event (KE) that contains multiple proteins. The nicotinic acetylcholine receptor activation leading to the colony death/failure adverse outcome pathway was used as a case study, focusing on KE2 (from Figure 1) as an example, where multiple proteins were involved in the KE. In this case study, 17 bee species were determined to represent the biologically plausible tDOA for KE2, in which five proteins were involved. CaMKII = calcium–calmodulin-dependent protein kinase II; PKA = protein kinase A; CREB = cyclic 3′5′-adenosine monophosphate–responsive element binding protein.

© 2022 The Authors.

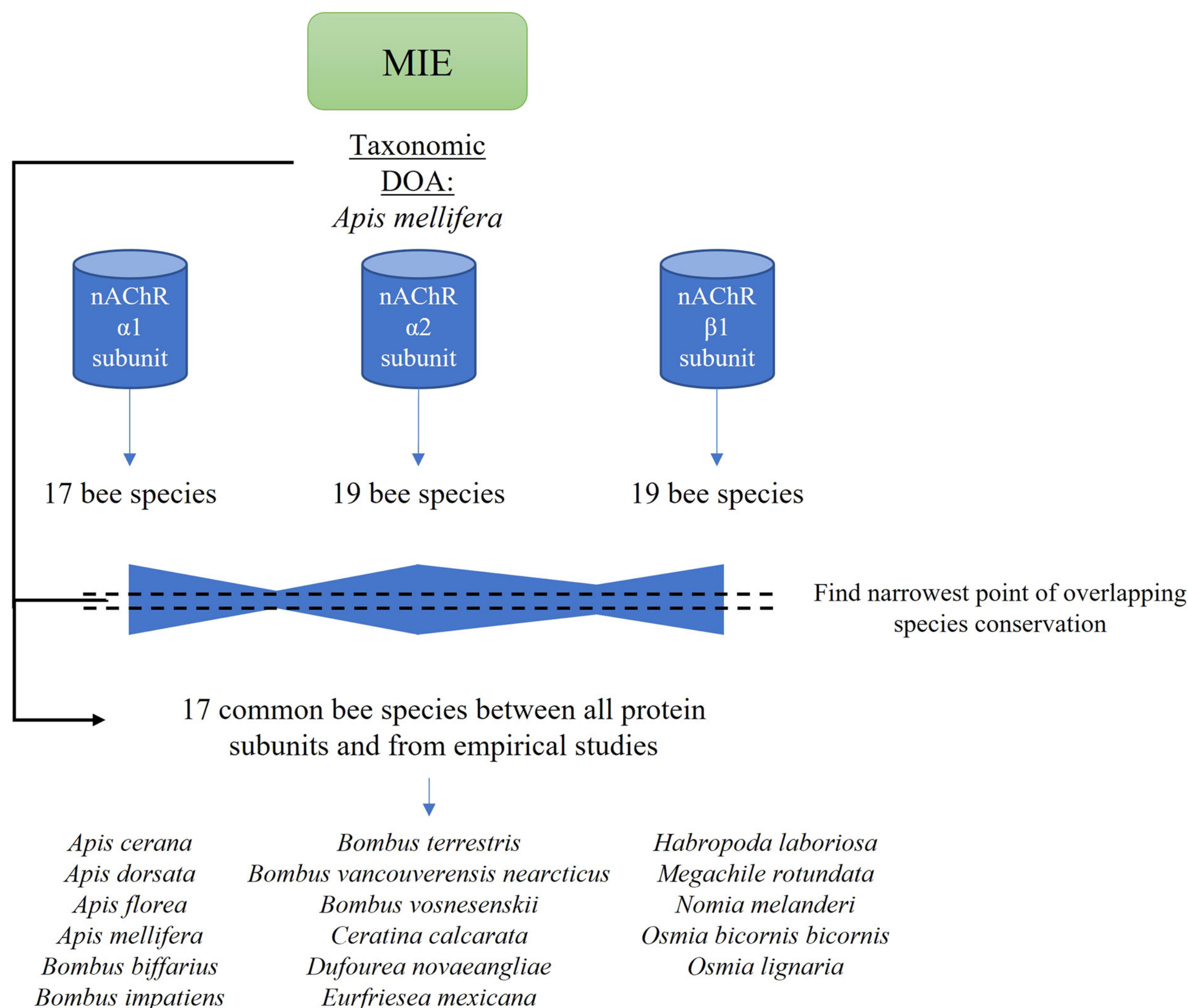

FIGURE 4: Determining biologically plausible taxonomic domain of applicability (tDOA) for a key event that contains a protein with multiple subunits. The nicotinic acetylcholine receptor (nAChR) activation leading to the colony death/failure adverse outcome pathway was used as a case study, focusing on the molecular initiating event (MIE) as an example, where multiple subunits were evaluated for the nAChR. In this case study, 17 bee species were determined to represent the biologically plausible tDOA for the MIE, in which three nAChR subunits were evaluated.

Evidence toward biologically plausible tDOA for MIE and KE1

The MIE and KE1 in the case example are nAChR activation and nAChR desensitization, respectively. The nAChR is the only protein involved in the MIE and KE1; therefore, the SeqAPASS-derived list of bee species as well as the list of species from empirical results for the MIE and KE1 are identical. The honey bee (*A. mellifera*) was the species identified from the AOP development process and empirical results to define the tDOA for the MIE and KE (LaLone et al., 2017).

nAChR. The nAChR is the molecular target for stressors such as neonicotinoid pesticides (Jeschke et al., 2013; Matsuda et al., 2020). Therefore, as the MIE, neonicotinoid pesticides and other stressors bind to the nAChR, initiating the biological pathway. As members of the cys-loop ligand-gated ion-channel superfamily, nAChRs are composed of combinations of five subunits and are thought to function as heteromers consisting of α- and non-α-subunits. However, it is unclear which specific subunits assemble to form nAChRs in insects (Matsuda, 2020; Matsuda et al., 2020; Perry et al., 2021). Under normal conditions, acetylcholine (ACh) binds at an interface between two subunits of the pentameric nAChR (Corringer et al., 2000; Perry et al., 2021). Six loops make up this orthosteric site (i.e., sites for binding of substrates/agonists), with the α-subunit contributing loops A, B, and C to the principal face and the adjacent α-subunit or β-subunit contributing loops D, E, and F to the complementary face (Corringer et al., 2000; Matsuda et al., 2020; Perry et al., 2021).

Insects have many different subunits that can assemble to form the pentameric nAChR. For example, *Drosophila* contain 10 genes that encode nAChR subunits (seven α, three β), and *A. mellifera* contain 11 genes that encode nAChR subunits (nine α, two β; Jones et al., 2006; Sattelle et al., 2005). Due to the many different subunit types, there are many possible combinations that could assemble to create the pentameric nAChR, which may differ physiologically (Matsuda, 2020; Perry et al., 2021). Multiple

 © 2022 The Authors.

studies in *Drosophila melanogaster* have shown that various combinations of the α1-, α2-, β1-, and α8/β2-subunits of the nAChR are responsive to several neonicotinoids (Ihara et al., 2020; Perry et al., 2021). The *A. mellifera* α8-subunit shares most similarity with the β2-subunit in other insects and is known to assemble with the α1- and β1-subunits to form a neonicotinoid-targeted nAChR (Ihara et al., 2020). However, in this assembly, it is proposed that the α1- and α8-subunits would form an orthosteric site where the α8-subunit would contribute loops D, E, and F (Matsuda, 2020). To our knowledge, critical amino acids for binding neonicotinoids specific to the α8-subunit complementary face have not been elucidated. In addition, other α-subunits were not found to be major targets for neonicotinoids but perhaps could just have low binding affinity or be limited in expression (Perry et al., 2021).

It is possible that subunits not evaluated in the present study differ across species in the role of neonicotinoid binding; however, it should be noted that all nAChR subunits identified, using *A. mellifera* as the query species, were evaluated using SeqAPASS Levels 1 and 2 by LaLone et al. (2016). For the present study, which focuses on Level 3 evaluations, the α1-, α2-, and β1-subunits were selected for evaluation in SeqAPASS because those subunits had the most information on the amino acids that interact with neonicotinoids.

Critical amino acids were chosen for Level 3 SeqAPASS analyses because evidence suggests that they are important for neonicotinoid binding (Table 2). It was determined through molecular modeling and site-directed mutagenesis techniques in *D. melanogaster* that in the α1-/β1-subunit interface, the α1-subunit amino acids tryptophan at position 170 (W170) and tyrosine at positions 220 (Y220) and 227 (Y227) and the β1-subunit amino acids arginine at position 81 (R81), leucine at position 141 (L141), and valine at position 143 (V143) interact with imidacloprid (Matsuda et al., 2020). In the α1-/α1-subunit interface, the amino acids contributed by the α1-subunit are W170, Y220, serine at position 221 (S221), and Y227; but the loop D, E, and F amino acids from the adjacent complimentary α1-subunit are R57, glutamate at position 78 (E78), and K140 (Matsuda, 2020; Matsuda et al., 2020). Therefore, these amino acids were chosen for Level 3 SeqAPASS analyses for the α1- and β1-subunits (Table 2). The importance of the arginine, or another basic amino acid, in loop D (R57 and R81 in α1- and β1-subunits, respectively) has been well studied for its role in binding neonicotinoids; and mutation to threonine in this position is known to

Total Match; Partial Match; Not a Match; Susceptible Yes; Susceptible No

Scientific Name	Similar Susceptibility	Amino Acid 1	Amino Acid 2	Amino Acid 3	Amino Acid 4	Amino Acid 5	Amino Acid 6	Amino Acid 7
Drosophila melanogaster	Y	57R	78E	140K	170W	220Y	221S	227Y
Apis mellifera	Y	53R	74E	136K	166W	216Y	217I	223Y
Apis cerana	Y	79R	100E	162K	192W	242Y	243I	249Y
Apis florea	Y	79R	100E	162K	192W	242Y	243I	249Y
Habropoda laboriosa	Y	53R	74E	136K	166W	216Y	217I	223Y
Osmia bicornis bicornis	Y	79R	100E	162K	192W	242Y	243I	249Y
Osmia lignaria	Y	79R	100E	162K	192W	242Y	243I	249Y
Bombus bifarius	Y	79R	100E	162K	192W	242Y	243I	249Y
Bombus vancouverensis nearcticus	Y	79R	100E	162K	192W	242Y	243I	249Y
Bombus vosnesenskii	Y	79R	100E	162K	192W	242Y	243I	249Y
Bombus terrestris	Y	79R	100E	162K	192W	242Y	243I	249Y
Megachile rotundata	Y	79R	100E	162K	192W	242Y	243I	249Y
Dufourea novaeangliae	Y	79R	100E	162K	192W	242Y	243I	249Y
Bombus impatiens	Y	79R	100E	162K	192W	242Y	243I	249Y
Nomia melanderi	Y	79R	100E	162K	192W	242Y	243I	249F
Eufriesea mexicana	Y	79R	100E	162K	192W	242Y	243I	249Y
Megalopta genalis	N	--	8E	70K	100W	150Y	151I	157F
Apis dorsata	Y	57K	78E	140K	170W	220Y	221T	227Y
Ceratina calcarata	Y	61K	82E	144K	174W	224Y	225T	231Y

FIGURE 5: Heatmap representing Sequence Alignment to Predict Across Species Susceptibility Level 3 output aligning critical amino acids for nicotinic acetylcholine receptor α1 subunit. Of all the species represented in this level of evaluation, *Megalopta genalis* is the only species where the subunit was considered to not be conserved based on evaluation of conservation of amino acids known to be critical for binding neonicotinoids. The species in first row was selected as the template sequence based on knowledge of critical amino acids. Y = yes; N = no; – = gap.

© 2022 The Authors.

confer resistance in target pests (Bass et al., 2011; Ihara et al., 2020; Matsuda et al., 2020; Shimada et al., 2020).

In comparison, for the α2-subunit, site-directed mutagenesis studies revealed the importance of proline in loop C in selective insect actions with imidacloprid (Shimomura et al., 2004). Similar to the α1-subunit, two tyrosine residues in loop C and a tryptophan residue in loop B were also postulated to interact with neonicotinoids (Shimomura et al., 2004; Wang et al., 2016). Therefore, those four amino acid residues, P242, W190, Y241, and Y248, were chosen for Level 3 SeqAPASS analyses (Table 2).

Level 1 results revealed that the nAChR subunits' primary amino acid sequences were conserved across both invertebrates and vertebrates, displaying conservation across 35, 45, and 45 taxonomic groups (790, 848, and 844 species) and with 110, 228, and 296 ortholog candidates being identified for the α1-, α2-, and β1-subunits, respectively (see Level 1 summary in Supporting Information, Results 1, 2, 3). Further, Level 2 results also revealed conservation of the neurotransmitter-gated ion-channel ligand binding domain in the nAChR subunits across 46, 45, and 45 taxonomic groups (850, 849, and 834 species) for the α1-, α2-, and β1-subunits, respectively (see Level 2 summary in Supporting Information, Results 1, 2, 3). To specifically evaluate the question of whether the AOP is relevant to other bee species, the Level 3 SeqAPASS results for the α1-, α2-, and β1-subunits showed conservation of all amino acids evaluated for the α2- and β1-subunits across the bee species evaluated (see Level 3 results in Supporting Information, Results 2, 3). For the α1-subunit, however, Level 3 SeqAPASS results provided evidence that all bee species aligned in SeqAPASS contain an arginine residue in loop D, or another basic residue in that position, except for *Megalopta genalis*, where part of the α1-subunit sequence has not been elucidated (Figure 5). Of note, S221 was only partially conserved across all 18 bee species evaluated (Figure 5). Nevertheless, because isoleucine and threonine are considered partial matches to serine in SeqAPASS (LaLone et al., 2016), the species still share similar susceptibility predictions (Figure 5).

Overall, the targeted Level 3 SeqAPASS results showed conservation of critical amino acids for neonicotinoid binding in all three subunits across four *Apis* and 13 non-*Apis* bee species (Figure 4 and Table 3; Supporting Information, Results 1, 2, 3). Therefore, because there are lines of evidence for structural conservation of nAChR subunits across *Apis* and non-*Apis* bees, based on the narrowest point of conservation between the subunits evaluated in SeqAPASS and the already defined species listed from empirical results, the biologically plausible tDOA for the MIE and KE includes four *Apis* species and 13 non-*Apis* bee species (Table 3 and Figure 4).

Evidence toward biologically plausible tDOA for KE2

The next KE, KE2, is Ca^{2+}–calmodulin-activated signal transduction, altered. There are multiple proteins involved in KE2 including calmodulin, adenylyl cyclase (AC), PKA, calcium–calmodulin-dependent protein kinase II (CaMKII), and cAMP-responsive element-binding protein (CREB). It is recognized that this KE involves a complex signaling pathway, as

TABLE 3: SeqAPASS Level 3 early key event protein summary

		Molecular initiating event key event 1			Key event 2				
	Scientific name	nAChR α1	nAChR α2	nAChR β1	Calmodulin	Adenylyl cyclase	CaMKII	PKA	CREB
Apis bees (six species)	*Apis mellifera*	y	y	y	y	y	y	y	y
	Apis cerana	y	y	y	y	y	n	y	y
	Apis florea	y	y	y	y	y	y	y	y
	Apis dorsata	y	y	y	y	y	y	y	y
	Apis cerana cerana			y	y	y		y	y
	Apis mellifera carnica							y	
Non-*Apis* bees (15 species)	*Habropoda laboriosa*	y	y	y	y	y	y	y	y
	Osmia bicornis bicornis	y	y	y	y	y	y	y	y
	Osmia lignaria	y	y	y	y	y	y	y	y
	Bombus vancouverensis nearcticus	y	y	y	y	y	y	y	y
	Bombus bifarius	y	y	y	y	y	y	y	y
	Bombus vosnesenskii	y	y	y	y	y	y	y	y
	Bombus terrestris	y	y	y	y	y	y	y	y
	Megachile rotundata	y	y	y	y	y	y	y	y
	Dufourea novaeangliae	y	y	y	y	y	y	y	y
	Bombus impatiens	y	y	y	y	y	y	y	y
	Nomia melanderi	y	y	y	y	y	y	y	y
	Eufriesea mexicana	y	y	y	y	y	y	y	y
	Megalopta genalis	n	y	y	y	y	y	y	y
	Ceratina calcarata	y	y	y	y	y	y	y	y
	Melipona quadrifasciata		y			y		y	n

Results were summarized from separate SeqAPASS evaluations for each protein from the molecular initiating event/key event 1 and key event 2. All species, separated into *Apis* and non-*Apis*, represented in SeqAPASS analyses are listed in the first column, with *y* representing conservation of the respective proteins in each species. Cells highlighted in red with *n* indicate a protein that was not conserved in that species.
CaMKII = calcium–calmodulin-dependent protein kinase II; CREB = cAMP-responsive element binding protein; nAChR = nicotinic acetylcholine receptor; PKA = protein kinase A; SeqAPASS = Sequence Alignment to Predict Across Species Susceptibility.

 © 2022 The Authors.

described in LaLone et al. (2017); however, these proteins are those specifically defined in the description for KE2. To provide evidence to define the biologically plausible tDOA for this KE, the narrowest point of conservation for species across all proteins evaluated was identified, in addition to *A. mellifera*, the only species that had already been defined from the empirical data in the AOP development process (Figure 3).

Calmodulin. At the start of the signaling cascade in KE2, Ca^{2+} enters neurons through the nAChR and subsequently binds to calmodulin, which activates proteins described later as the signaling cascade continues (LaLone et al., 2017). Therefore, the role of calmodulin in the AOP is binding of Ca^{2+}. Level 1 SeqAPASS evaluation showed that the primary sequence of calmodulin was conserved across 24 vertebrate and invertebrate taxonomic groups (1295 species) including Insecta, Mammalia, and Aves, among others; and 229 ortholog candidates were identified. Level 2 SeqAPASS evaluations showed that the calcium binding motif was conserved across 162 invertebrate and vertebrate taxonomic groups (6492 species), and 229 ortholog candidates were identified (see Levels 1 and 2 summary in Supporting Information, Results 4). The five critical amino acids directly involved with Ca^{2+} binding within the Ca^{2+} binding regions were elucidated from the Protein Data Bank (www.rcsb.org, idenitifier 3CLN) structure and evaluated in SeqAPASS Level 3 (Table 2; Babu et al., 1988). Level 3 results, where the scope was reduced to bee species only, show that the amino acids important for Ca^{2+} binding are conserved across *Apis* and non-*Apis* bees (Table 3; Supporting Information, Results 4). Considering SeqAPASS Level 3 results for calmodulin, evidence suggests that the protein is conserved across five *Apis* and 14 non-*Apis* bees (Supporting Information, Results 4).

Adenylyl cyclase. Adenylyl cyclase (AC) is a protein activated by calmodulin in the signaling cascade described in KE2 (Figure 1; LaLone et al., 2017). When AC is activated by the Ca^{2+}–calmodulin complex, it catalyzes the conversion of adenosine triphosphate (ATP) to cAMP, which then goes on to interact with downstream proteins in the signaling cascade described for this KE (Balfanz et al., 2012; LaLone et al., 2017). Seven ACs have been identified in *A. mellifera*, with only AmAC8 and AmAC2t being functionally characterized (Balfanz et al., 2012). It has been suggested that AmAC8 has a role in *A. mellifera* learning behavior due to its sequence similarity and expression patterns similar to *D. melanogaster* AC78C and that it is likely a member of a family of ACs that are known to be stimulated by the Ca^{2+}–calmodulin complex (Balfanz et al., 2012). Because of this, AC8 was identified as the appropriate protein to represent KE2 and used for the SeqAPASS evaluation.

The SeqAPASS Level 1 results provided evidence that AC8 is conserved across 17 taxonomic groups (657 species), including invertebrates and vertebrates; and 22 ortholog candidates were identified. Evidence was collected from the SeqAPASS Level 2 results indicating that the catalytic domain was also conserved across 39 taxonomic groups (782 species), including Insecta, Amphibia, and Actinopteri, among others (see Levels 1 and 2 summary in Supporting Information, Results 5). When AmAC8 was functionally characterized, the amino acids critical for binding ATP and Mg^{2+}, as well as for other functions, were elucidated (Balfanz et al., 2012). Those four amino acids, two aspartate residues critical for Mg^{2+} binding and the lysine and aspartate residues critical for ATP binding, were evaluated in SeqAPASS Level 3 (Table 2). For some *Apis* and non-*Apis* species, it should be noted that proteins annotated as AC78C were chosen for evaluation because they aligned with the highest sequence similarity in the SeqAPASS Levels 1 and 2 output, though their annotation was not specifically "adenylyl cyclase type 8" (confirmed by NCBI BLAST runs; see Supporting Information, Results 5). Level 3 SeqAPASS results showed that amino acids critical for ATP and Mg^{2+} binding in AC8 are conserved across *Apis* and non-*Apis* bees (Table 3; Supporting Information, Results 5). Considering all levels of the SeqAPASS evaluation, results suggest that AC is conserved across five *Apis* and 15 non-*Apis* bees (Supporting Information, Results 5).

PKA. In KE2, PKA, otherwise known as *cAMP-dependent protein kinase*, is a protein activated by cAMP (Figure 1; LaLone et al., 2017). It is a tetrameric enzyme made up of two regulatory and two catalytic subunits (R_2C_2). Dissociation of the catalytic (C) subunits, mediated by cAMP binding to the regulatory (R) subunits, activates the enzyme (Berman et al., 2005; Leboulle & Müller, 2004; Taylor et al., 1990). Although cAMP-binding domains exist in a variety of proteins, cAMP binds to two of these domains in the R subunits of PKA, termed *A* and *B*, each containing a phosphate-binding cassette (Berman et al., 2005; Leboulle & Müller, 2004; Taylor et al., 1990). In addition, it was found that cAMP activates the RII subunit in the honey bee *A. mellifera carnica* (Leboulle & Müller, 2004). The amino acids directly interacting with cAMP were elucidated in *A. mellifera carnica* PKA-RII in both phosphate-binding cassettes, and similar findings in the Norway rat (*Rattus norvegicus*) supported the importance of these amino acids (Berman et al., 2005; Diller et al., 2001; Leboulle & Müller, 2004). Because there is evidence for such interactions, PKA-RII was identified as the query sequence for SeqAPASS analyses (Table 2).

Level 1 evaluation showed that the PKA-RII amino acid sequence was conserved across 125 taxonomic groups (2389 species), including invertebrates and vertebrates; and 837 ortholog candidates were identified. Level 2 SeqAPASS evaluation showed that the cAMP-binding domain was also conserved across 144 taxonomic groups (2441 species), including Insecta, Collembola, and Merostomata, among others; and 839 ortholog candidates were identified (see Levels 1 and 2 summary in Supporting Information, Results 6). Level 3 SeqAPASS evaluation revealed full conservation of the amino acids that directly interact with cAMP in both domains A and B across *Apis* and non-*Apis* bees (Table 3; Supporting Information, Results 6). Considering all levels of the SeqAPASS evaluation, the evidence suggests that PKA is conserved across six *Apis* and 15 non-*Apis* bees (Supporting Information, Results 6).

CaMKII. As mentioned before (*Adenylyl cyclase*), the Ca^{2+}–calmodulin complex can activate AC; however, it can

© 2022 The Authors.

also activate CaMKII (Coultrap & Bayer, 2012; LaLone et al., 2017; Wayman et al., 2011). The CaMKII protein is well studied, and it has been shown that stimulation of CaMKII by the Ca^{2+}–CaM complex leads to autophosphorylation of distinct sites and that one of them, T286, generates Ca^{2+}-independent kinase activity (Coultrap & Bayer, 2012; Wayman et al., 2011). Five amino acids important for phosphorylation and one amino acid important for ATP binding were chosen for SeqAPASS Level 3 evaluation (Table 2; for review, see Coultrap & Bayer, 2012; Wayman et al., 2011).

The SeqAPASS Level 1 results showed that the primary amino acid sequence of CaMKII was conserved across 30 taxonomic groups (697 species), including invertebrates and vertebrates; and 191 ortholog candidates were identified. Level 2 SeqAPASS evaluation showed that the catalytic domain of CaMKII was also conserved across 46 taxonomic groups (757 species), including Insecta, Branchiopoda, and Merostomata, among others; and 193 ortholog candidates were identified (see Levels 1 and 2 summary in Supporting Information, Results 7). Level 3 results showed full conservation of the critical amino acids, except for *Apis cerana* (Table 3; Supporting Information, Results 7). Importantly, *A. cerana* CaMKII has two omitted amino acids in its sequence in comparison to the template sequence, although the primary amino acid sequence and catalytic domain were conserved (Supporting Information, Results 7). Considering all levels of the SeqAPASS evaluation, CaMKII is suggested to be conserved across three *Apis* and 14 non-*Apis* bees (Supporting Information, Results 7).

CREB. As the last protein involved in KE2, CREB, once activated through phosphorylation, binds its coactivator, CREB-binding protein (CBP), and recognizes and activates transcription of genes with the cAMP response element (CRE) in their promoters (Figure 1; Kandel, 2012; LaLone et al., 2017; Mayr & Montminy, 2001; Montminy, 1997). The kinases PKA, CaMKII, and mitogen-activated protein kinase (MAPK) activate CREB1 through phosphorylation at a serine residue within the kinase-inducible domain (Gonzalez et al., 1989, 1991; Kandel, 2012; Mayr & Montminy, 2001). In addition, the serine residue is located within the known PKA recognition site, a site composed of five amino acids starting with two arginine residues critical for PKA recognition (Gonzalez et al., 1989, 1991; Kandel, 2012; Mayr & Montminy, 2001). Thus, the kinase-inducible domain and the amino acids critical for PKA recognition and phosphorylation were chosen for Levels 2 and 3 SeqAPASS evaluation, respectively (Table 2).

Levels 1 and 2 SeqAPASS results showed conservation of the primary amino acid sequence and the kinase-inducible domain for CREB1 across 35 taxonomic groups (712 species), including invertebrates and vertebrates; and 423 ortholog candidates were identified. Level 2 SeqAPASS results showed conservation of the kinase-inducible domain across 34 taxonomic groups (662 species), including Insecta, Malacostraca, and Branchiopoda, among other taxa; and 420 ortholog candidates were identified (see Levels 1 and 2 summary in Supporting Information, Results 8). Of note, the kinase-inducible domain was not conserved in *Melipona quadrifasciata*. Level 3 SeqAPASS results showed that the critical amino acids, specifically the phosphoacceptor site S133, are conserved across *Apis* and non-*Apis* bees (Table 3; Supporting Information, Results 8). However, CREB1 in the stingless bee, *M. quadrifasciata*, again, was not conserved. Three out of the five amino acids were not conserved, including the crucial S133 needed for CREB1 activation; and the other two amino acids were only partial matches to the template sequence (Table 3; Supporting Information, Results 8). This suggests that *M. quadrifasciata* CREB1 may not share the same structure or function of other bee CREB1 proteins. However, through reviewing all levels of the SeqAPASS evaluation, the evidence suggests that CREB is conserved across five *Apis* and 14 non-*Apis* bees (Supporting Information, Results 8).

Overall, SeqAPASS Levels 1, 2, and 3 evaluations suggest that all proteins in this KE are conserved across most *Apis* and non-*Apis* bees. Through bringing together the species from empirical results for this KE, only *A. mellifera* defines the tDOA. Computational results from SeqAPASS identify conservation of the KE in three *Apis* and 14 non-*Apis* bee species and can be used to further define biological plausibility for KERs (Table 3 and Figure 3).

Evidence toward tDOA for KE3 and KE4

The next two KEs are KE3, learning and memory, impairment, and KE4, foraging activity and behavior, abnormal. Because neither of these KEs includes the specific action of proteins that would allow for computationally generated evidence through SeqAPASS for the tDOAs, the tDOAs were derived solely from the empirical data used to develop the KEs. The species defined for the tDOA based on empirical data for KE3 are *A. mellifera* and *Bombus terrestris*, and those for KE4 are *A. mellifera*, *A. cerana*, and *B. terrestris*.

Evidence toward biologically plausible tDOA for KE5

The next KE, KE5, is an abnormal role change within caste. The proteins involved in this KE are Vg, juvenile hormone acid *O*-methyltransferase, and methyl farnesoate epoxidase. The proteins juvenile hormone acid *O*-methyltransferase and methyl farnesoate epoxidase were used to be representative of the juvenile hormone biosynthesis pathway (Aurori et al., 2020). The species defined for the tDOA using empirical data during AOP development is *A. mellifera*.

In *A. mellifera*, Vg has many different functions, including playing a role in embryo nourishment, oxidative stress resistance, cell-based immunity, life span, and, specifically, foraging onset (Amdam et al., 2012; Havukainen et al., 2013). Together with juvenile hormone, Vg acts in a feedback loop to control the onset of foraging (Amdam & Omholt, 2003). It has been shown that these proteins have similar physiological roles in the division of labor of other bee species as well but may not have the same relationship that is found in *A. mellifera*. For example, Vg titers were found to be involved in the division of labor in

 © 2022 The Authors.

the facultatively eusocial sweat bee, *M. genalis*, where queens had higher Vg titers than workers (Kapheim et al., 2011). It has also been found that Vg is associated with caste (queen vs. worker) and social context (Amsalem et al., 2014) and that juvenile hormone is involved in caste determination in *B. terrestris* (Cnaani et al., 2000), but the two seem to be uncoupled in that species (compared with the negative correlation seen in *A. mellifera*; Amsalem et al., 2014).

To our knowledge, the specific amino acids important to these protein functions are unknown; the proteins involved in KE5 were evaluated using SeqAPASS Levels 1 and 2 only. In addition, the proteins juvenile hormone acid *O*-methyltransferase and methyl farnesoate epoxidase were used in SeqAPASS evaluations to be representative of juvenile hormone because they are involved in the juvenile hormone biosynthesis pathway in *A. mellifera* (Aurori et al., 2020). For Vg, SeqAPASS Levels 1 and 2 evaluations showed conservation of the primary amino acid sequence and the lipoprotein amino terminal region across 14 and 15 taxonomic groups (392 and 556 species); and 304 and 310 ortholog candidates were identified, respectively (Supporting Information, Results 9). For methyl farnesoate epoxidase, SeqAPASS Levels 1 and 2 evaluations showed conservation of the primary amino acid sequence and the cytochrome P450 domain across 46 and 53 taxonomic groups (1161 and 1191 species); and 258 and 263 ortholog candidates were identified, respectively (Supporting Information, Results 10). For juvenile hormone acid *O*-methyltransferase, both SeqAPASS Levels 1 and 2 evaluations showed conservation of the primary amino acid sequence and the methyltransferase domain across 37 taxonomic groups (652 and 657 species); and 434 and 438 ortholog candidates were identified, respectively (Supporting Information, Results 11). With a specific focus on bee species only, SeqAPASS Levels 1 and 2 evaluations provided evidence that the key proteins were conserved in five *Apis* and 13 non-*Apis* bees (Figure 2B; Supporting Information, Results 9, 10, 11). Because SeqAPASS results showed that the proteins are present and contain the respective conserved domains across species, species-specific factors and protein functions can be considered to further determine the tDOA for KE5 and/or plausible differing KEs, KERs, and AOs. Taken together, this KE is an example of how SeqAPASS Levels 1 and 2 results and species-specific factors could help to further define the tDOA at downstream KEs and KERs.

Evidence toward tDOA for KE6

The last KE in the AOP is KE6, colony weakened. It does not contain any proteins that would allow for computationally generated evidence through SeqAPASS for biological plausibility for the tDOA. Therefore, the tDOA was derived from empirical data used to develop the KE, which is *A. mellifera* only and could be assumed applicable to other species that have similar colony structure.

Biologically plausible tDOA for KER1

KER1 links the MIE and KE1 of nAChR activation and nAChR desensitization, respectively. Because the nAChR was the only protein identified for MIE and KE1, the narrowest point of conservation across the nAChR subunits evaluated was identified (or the narrowest point of conservation used for determining the biologically plausible tDOA for MIE and KE1; they are the same in this case). Therefore, the biologically plausible tDOA for KER1 includes *A. mellifera*, *A. cerana*, *Apis dorsata*, *Apis florea*, *Bombus bifarius*, *Bombus vosnesenskii*, *B. terrestris*, *Bombus impatiens*, *Bombus vancouverensis nearcticus*, *Ceratina calcarata*, *Dufourea novaeangliae*, *Eufriesea mexicana*, *Habropoda laboriosa*, *Megachile rotundata*, *Nomia melanderi*, *Osmia bicornis*, and *Osmia lignaria* (Figure 2B).

Biologically plausible tDOA for KER2

Key event relationship 2 links KE1 of nAChR desensitization and KE2 of Ca^{2+}–calmodulin-activated signal transduction, altered. To define the biologically plausible tDOA for KER2, the narrowest point of conservation between the list of species for the biologically plausible tDOA for KE1 and KE2 was identified; and those species are *A. mellifera*, *A. dorsata*, *A. florea*, *B. bifarius*, *B. vosnesenskii*, *B. terrestris*, *B. impatiens*, *B. vancouverensis nearcticus*, *C. calcarata*, *D. novaeangliae*, *E. mexicana*, *H. laboriosa*, *M. rotundata*, *N. melanderi*, *O. bicornis*, and *O. lignaria* (Figure 2B).

Biologically plausible tDOA for KER3, KER4, KER5, KER6, and KER7

Key event relationship 3 links KE2 of Ca^{2+}–calmodulin-activated signal transduction, altered and KE3 of learning and memory, impairment; KER4 links KE3 of learning and memory, impairment and KE4 of foraging activity and behavior, abnormal. To define the biologically plausible tDOA for KERs, the narrowest point of conservation between the species for the tDOA of the respective upstream and downstream KEs was identified. Although the biologically plausible tDOA for KE2 contained a list of species generated from SeqAPASS evaluations, KE3 did not have any proteins involved that would allow for computationally generated evidence for the biologically plausible tDOA. Therefore, the biologically plausible tDOA for KER3 was defined by the narrowest point of conservation between the biologically plausible tDOA for KE2 and the tDOA for KE3 based on empirical data, which was *A. mellifera* and *B. terrestris* (Figure 2B). Similarly, KE4 did not have any proteins involved that would allow for computationally generated evidence for the biologically plausible tDOA, so the tDOA for KER4 was derived from empirical data used to develop KE3 and KE4 only. Therefore, the bee species that define the tDOA for KER4 are *A. mellifera* and *B. terrestris* (Figure 2B).

Key event relationship 5 links KE4 of foraging activity and behavior, abnormal and KE5 of role change within caste, abnormal. The narrowest point of conservation between the list of species from empirical data for KE4 and the list of species from empirical data as well as the SeqAPASS list of species for KE5 are *A. mellifera*, *A. cerana*, and *B. terrestris* (Figure 2B). Key event relationship 6 links KE5 of role change within caste,

© 2022 The Authors.

abnormal and KE6 of colony, weakened. The list of species from empirical data as well as the SeqAPASS list of species for KE5 and the empirical list of species for KE6 were used to identify the narrowest point of conservation between the two lists of species, which was *A. mellifera* (Figure 2B). Key event relationship 7 links KE6 of colony, weakened to the AO of colony death/failure. Because some proteins are not involved in either the KE or the AO, the list of species from empirical data for KE6 and the AO were used. The species that define the tDOA for KER7 is *A. mellifera* (Figure 2B).

Biologically plausible tDOA for AOP

The biologically plausible tDOA of an AOP is defined by the overlapping species among all KERs (OECD, 2018). Therefore, to define the biologically plausible tDOA for the entire AOP, the narrowest point of conservation of species across the KERs could be identified. From that, the species that define the biologically plausible tDOA for the AOP is defined by available SeqAPASS results and through other lines of evidence of structural and functional evidence potentially expanding beyond *A. mellifera* (Figure 2B). It should be noted that this definition of biologically plausible tDOA is derived specifically from instances where there is supporting evidence of structural and functional conservation from the SeqAPASS results and/or the empirical data used to generate the AOP description. Biological plausibility can be expanded on using inference for tDOA determinations; however, gathering lines of evidence toward conservation and understanding the data streams supporting those lines of evidence are important for application.

DISCUSSION

It is recognized that the majority of AOPs in the AOP-Wiki are limited in their descriptions of tDOA. Based on global AOP network metrics (Pollesch et al., 2019) for the AOP-Wiki measured in April 2021, out of 1149 KEs in the AOP-Wiki, only 301 (26%) provided any specification of tDOA. An even smaller percentage of KERs (22%) were annotated with respect to tDOA. Likewise, among the 333 user-defined AOPs available in the AOP-Wiki at that time, only 83 (25%) included any description of the overall DOA of the AOP (D. Villeneuve, USEPA, personal communication, April 2021).

To enhance the utility of the AOP framework for decision-making applications and guide users as to which species the AOP can reasonably be applied, it is important to consider structural and functional conservation of the biology across species (OECD, 2018), while taking advantage of available and scientifically defensible data streams. Opportunity exists to apply new approach methods and bioinformatics to inform tDOA. The purpose of the present study was to describe a process for determining the biologically plausible tDOA for KEs, KERs, and entire AOPs using both available toxicity data and computational information that allows for the extrapolation of existing data to other untested species through an understanding of structural conservation. The present study lays the foundation for other bioinformatics approaches to be incorporated into the existing AOP framework for broadening the definition of the tDOA beyond the model organisms used in the assays described within the KEs of an AOP. Specifically, the SeqAPASS tool (i.e., computational predictive approach) was used to generate data as evidence for structural conservation of proteins involved in the KEs and to further demonstrate how such information can readily be incorporated into AOP descriptions based on available data and methods for cross-species extrapolation. When upstream and downstream KE tDOAs were combined to determine species overlap, the biologically plausible tDOA for each KER was further defined. This approach allows for description of hundreds of untested organisms to be considered in the tDOA when incorporating data streams that assess structural conservation and allows for a focus on specific groups of species, if warranted. The other key element to tDOA is functional conservation, which can be gleaned from in vitro and in vivo studies to support and refine computational data streams considering structural conservation (Ankley et al., 2016). In the future, helpful tools to provide functional lines of evidence could be something that comprises phylogenetic or trait-based information and capitalizes on advances in systematic methods for literature review to more rapidly locate, evaluate, and synthesize critical information into the lines of evidence for tDOA.

It is important to understand that the SeqAPASS tool can only be applied when there are proteins specifically identified to be essential in the KE descriptions. In the present study, the AOP linking nAChR activation to colony death/failure was used to demonstrate the process for incorporating bioinformatics to evaluate structural conservation of proteins identified within the AOP. This example also demonstrated how empirical tDOA would be defined in instances where proteins are not specifically identified by the AOP developer to be essential in KEs (e.g., KE3 or KE4 in the AOP in the present study) and how empirical data only, with consideration of biological plausibility, are then used to define the tDOA for the KERs. The present study intended to lay the foundation for the utilization of bioinformatics approaches to more thoroughly and rapidly evaluate structural conservation to readily improve descriptions of tDOA for KERs. When such information has been generated, other existing empirical data and methods can be used to help further define the functional conservation, adding strength to the WoE collected to support the biologically plausible tDOA.

A large part of AOP development is determining the WoE for KERs in the AOP, which can include biological plausibility, empirical support, and/or evidence for the essentiality of the KEs (OECD, 2013, 2018). When using the biologically plausible tDOA of the KERs to then determine the tDOA for the whole AOP, the WoE of the developed KERs can be used to add confidence to the biologically plausible tDOA. For example, in application of WoE to AOP descriptions, developers typically describe the evidence supporting the description as weak, moderate, or strong. There are specific guidelines that assist developers to make this evidence call (OECD, 2013, 2018). It is envisioned that AOP development guidance would adopt language to include data from computational approaches to

 © 2022 The Authors.

provide an evidence call for structural conservation of "weak" WoE, whereas inclusion of in vitro and/or in vivo results in tDOA descriptions for functional conservation could enhance evidence to "moderate" and "strong" WoE, respectively. Overall, transparency as to which data stream is being used to generate the evidence for both structural and functional conservation should be accounted for. Advances in AOP-Wiki development could incorporate methods to readily capture and display both computational and empirical data to support WoE calls. Such developments would enhance the utility of AOPs for a variety of research and regulatory applications, particularly when advances can be made to collect more detailed species information that defines the tDOA for downstream KEs more broadly.

Overall, the biologically plausible tDOA for the AOP that links nAChR activation to colony death/failure was limited to *A. mellifera* to illustrate the use of bioinformatics and empirical evidence. However, it may be biologically plausible to consider extrapolation of existing knowledge to other bee species; the evidence that exists to support such extrapolation should be documented. From the case example with the nAChR activation AOP, the narrow tDOA coverage across the entire AOP points to the need for new tools/methods to extrapolate or further define species relevance for KEs that do not specifically describe genes/proteins as measurable or observable for the KE. For early KEs, especially those at the molecular and cellular levels of biological organization, it is more likely that genes/proteins are important for the KE descriptions and therefore logical that a larger number of species could be derived from current bioinformatics approaches to define the biologically plausible tDOA. As for downstream KEs describing apical endpoints, typically with KEs that are relatively taxon- or trait-specific, it is logical that the biologically plausible tDOA for those KERs will be relatively narrow. It is envisioned that, especially for apical KEs, SeqAPASS results should be used as a line of evidence for conservation in combination with consideration of general species-specific factors. For example, KE5 describing role change within caste, abnormal and KE6 describing colony, weakened would not be applicable to species that do not form bee-like colonies or that differ in colony structure(s) from *A. mellifera*, such as solitary bee species that completely lack colony structure. Therefore, these more apical KEs are expected to differ across bee species depending on factors such as colony structure or lack thereof or foraging strategies. In such instances early events may be linked to different AOs for certain types of species.

Regardless of the data stream for evidence of conservation to further define and broaden the tDOA for KEs, a decision tree, as illustrated in Figure 6A, can be applied. Through the use of a decision tree, AOP developers could ask specific questions to understand whether species or other broader taxonomic lineage designations could be included in the biologically plausible tDOA. For example, with each KE the developer could ask whether it is biologically plausible that it is conserved in other species/taxa and seek evidence to support that claim. If the answer to the question posed in the KE is "yes," then the species/taxa should be included in the tDOA. Although a comprehensive review for all species that would be considered biologically plausible could be considered in the decision tree, the focus in the present study was to be illustrative and provide two straightforward examples using a decision tree to define the tDOA. The examples shown are for *A. mellifera*, a eusocial insect (Figure 6B). and *O. lignaria* (Figure 6C), a solitary bee species. It is demonstrated that KEs are conserved across the whole AOP for *A. mellifera*, and it is anticipated that other eusocial bee species would broaden the biologically plausible tDOA for the case study AOP through the use of the decision tree. By comparison, the KEs are only applicable through KE3 with the

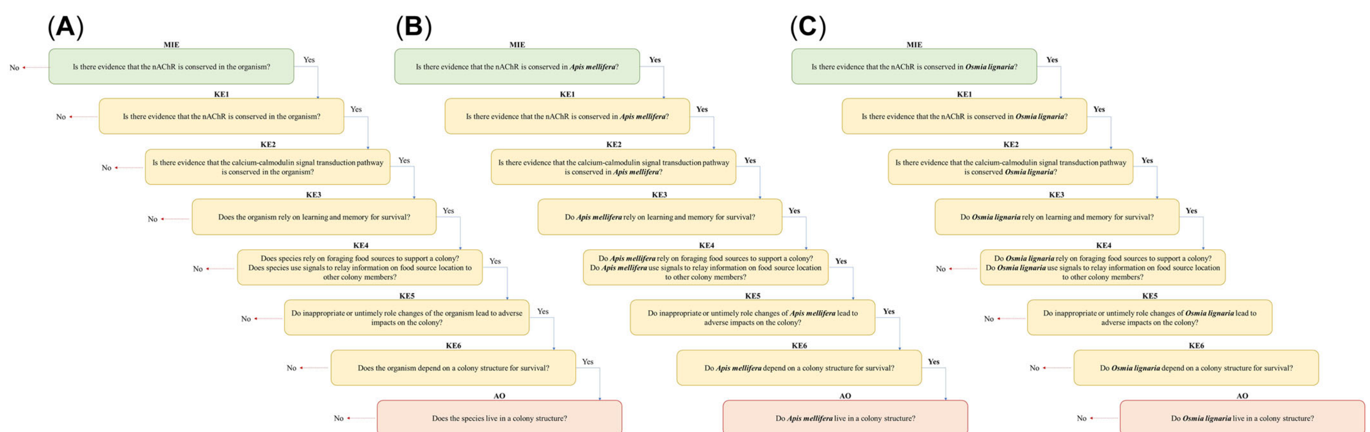

FIGURE 6: (**A**) Example of how a decision tree can be applied to inform the taxonomic domain of applicability (tDOA) for the nicotinic acetylcholine receptor activation leading to the colony death/failure adverse outcome pathway (AOP). Such a decision tree can be developed for any AOP, where a specific biological question is posed based on the description for the molecular initiating event (MIE), each key event (KE), and the adverse outcome. This process can be repeated for any desired species or taxa to decide if inclusion in the biologically plausible tDOA is warranted. (**B**) Example of how the decision tree can be used with *Apis mellifera*, where all events in the AOP are conserved (LaLone et al., 2017). (**C**) Example of how the decision tree can be used with *Osmia lignaria*. Only the early KEs, through KE3, are conserved in this solitary bee species. This is because Sequence Alignment to Predict Across Species Susceptibility (SeqAPASS) results from this case study show conservation of the proteins evaluated for the MIE/KE1 and KE2 in this species (see Table 3), and *Osmia lignaria* do rely on learning and memory for survival (Amaya-Márquez et al., 2008); however, they do not live in a colony structure (Amaya-Márquez et al., 2008). nAChR = nicotinic acetylcholine receptor; AO = adverse outcome.

© 2022 The Authors.

solitary bee species because they do not live in a colony structure as defined by the AOP. Therefore, it is likely that the solitary bee may be adversely impacted though perturbation of the biology upstream, whereas downstream KEs and AOs may differ. The use of a decision tree introduces additional lines of evidence that can be considered to define the tDOA. When considering the example presented using *O. lignaria* in Figure 6C, SeqAPASS results showed conservation of the three proteins evaluated for KE5 in that species (Supporting Information, Results 9, 10, 11); however, based on the decision tree, the species would not be included in the tDOA for KE5. This example demonstrates how considerations of both structural conservation (e.g., SeqAPASS data) and functional conservation (which may include functional traits and other complex characteristics) should be used to complement one another to pragmatically define the biologically plausible tDOA.

Nevertheless, even though the outcomes may differ across species, the tDOA knowledge of early KEs is useful in understanding the biology of bee species that are typically untested. Such understanding of shared structural and, when data exist, functional conservation at early molecular- and cellular-level events can lead to the development of additional AOPs for unique outcomes that may occur in other species. Databases exist (e.g., https://animaldiversity.org/, https://opentraits.org/datasets.html) that capture characteristics, traits, organization, and life-history information that may fill some of these knowledge gaps. Future work may focus on the development of tools to rapidly mine such information to understand species similarities and differences that could be used to further define tDOAs in downstream KEs.

It is anticipated that lines of evidence generated from SeqAPASS for species extrapolation in defining biologically plausible tDOAs will continue to be enhanced as more genomes are sequenced across the diversity of species with improved annotation capabilities. Although sequence information and protein information were available for many *Apis* and non-*Apis* bees, some knowledge gaps still arose for a few species within the KEs (Table 3). For example, many of the proteins evaluated in the present study have not been elucidated for *M. quadrifasciata* or *A. cerana cerana* (Table 3). Because of this, the determination of the biologically plausible tDOA for many of the KEs did not include these species; and therefore, it is unknown whether structural conservation exists for those species. In addition to those gaps, there are many bee species not represented in these analyses because the sequence information has not been generated to date. It has been estimated that there are approximately 17 000–20 000 bee species in the world (Michener, 2000), although genome assemblies with annotation only exist for 24 bee species in NCBI (date searched September 7, 2021). In the present study, 21 of those species were present in the SeqAPASS results. However, because of increased efficiencies and lower costs for generating whole-genome sequence data, as well as continually improved bioinformatics approaches, public repositories of gene and protein sequence information continue to expand rapidly.

Although the use of bioinformatics approaches, like SeqAPASS, can provide useful information for defining tDOAs, such extrapolations are reliant on the sequence information available and the quality of those data. For example, in the SeqAPASS results for nAChR α1-subunit, the α1-subunit for *M. genalis* was predicted to be not conserved because the computationally predicted sequence was incomplete, and the first amino acid residue evaluated in SeqAPASS Level 3 was not present in that sequence. Specifically, the nAChR α1-subunit accession for that species (XP_033338174.1) is approximately 80 amino acids shorter in length compared to the nAChR α1-subunit sequence for *A. mellifera*. Also, within the nAChR α1-subunit SeqAPASS results, S221 was only partially conserved across all 18 bee species evaluated (Figure 5). That serine residue in loop C of the α1-subunit has been well studied and recognized as important for neonicotinoid binding; however, mutations to isoleucine or threonine have not been evaluated specifically (Shimada et al., 2020). However, it was found that mutations to alanine and glutamine (which could be representative of changes to isoleucine or threonine as seen in bee species, due to side-chain classifications) minimally affected the affinity of neonicotinoids, and it was suggested that the serine residue does not play a major role in binding neonicotinoids when the basic residues critical for interactions are present (Shimada et al., 2020). It would be of interest to evaluate the effects of neonicotinoid binding when the serine is mutated to isoleucine or threonine to gain an even better understanding of *Apis* and non-*Apis* bee nAChRs. These examples point to the fact that as more sequence information for species is elucidated and updated and more molecular modeling and site-directed mutagenesis studies are performed, predictions reliant on these data will continue to improve and expand.

A goal of the present study was to demonstrate how bioinformatics approaches, such as SeqAPASS, can be used to provide lines of evidence for structural conservation as development moves from upstream molecular KEs, containing gene/protein information, to downstream apical KEs as a means to determine how broadly an AOP can be extrapolated across species, which can aid in AOP development and use. It is important to note that this process can be performed for any AOP to extrapolate beyond the model organisms typically defined during AOP development. Although the AOP evaluated in the present study was created based on an MIE known to be modulated by neonicotinoid pesticides and the purpose of determining the biologically plausible tDOA was to provide insight as to whether neonicotinoids could potentially be impacting *Apis* and non-*Apis* bees similarly, AOPs are not chemical-specific. Therefore, this application of bioinformatics for determining the plausible tDOA could be defined even more broadly (e.g., across many taxa) depending on the AOP of interest and its intended application.

Supporting Information—The Supporting Information is available on the Wiley Online Library at https://doi.org/10.1002/etc.5501.

Acknowledgment—We thank D. Villeneuve and J. O'Brien for providing comments on the present study. The manuscript was

 © 2022 The Authors.

reviewed in accordance with the requirements of the USEPA Office of Research and Development.

Conflict of Interest—The authors declare no conflict of interest.

Disclaimer—The views expressed in the present study are those of the authors and do not necessarily reflect the views or policies of the USEPA, nor does the mention of trade names or commercial products constitute endorsement or recommendation for use.

Author Contributions Statement—**Marissa A. Jensen**: Data curation; Formal analysis; Investigation; Visualization; Writing—original draft; Writing—review & editing. **Donovan J. Blatz**: Data curation; Formal analysis; Investigation; Visualization; Writing—review & editing. **Carlie A. LaLone**: Conceptualization; Supervision; Visualization; Writing—original draft; Writing—review & editing.

Data Availability Statement—All data are provided in the Supplemental Data. For all other enquiries, please contact the corresponding author (LaLone.Carlie@epa.gov).

REFERENCES

Amaya-Márquez, M., Hill, P. S. M., Barthell, J. F., Pham, L. L., Doty, D. R., & Wells, H. (2008). Learning and memory during foraging of the blue orchard bee, *Osmia lignaria* Say (Hymenoptera: Megachilidae). *Journal of the Kansas Entomological Society*, *81*(4), 315–327. https://doi.org/10.2317/JKES801.29.1

Amdam, G. V., Fennern, E., & Havukainen, H. (2012). Vitellogenin in honey bee behavior and lifespan. In C. Galizia, D. Eisenhardt, & M. Giurfa (Eds.), *Honeybee neurobiology and behavior* (pp. 17–29). Springer. https://doi.org/10.1007/978-94-007-2099-2_2

Amdam, G. V., & Omholt, S. W. (2003). The hive bee to forager transition in honeybee colonies: The double repressor hypothesis. *Journal of Theoretical Biology*, *223*(4), 451–464. https://doi.org/10.1016/s0022-5193(03)00121-8

Amsalem, E., Malka, O., Grozinger, C., & Hefetz, A. (2014). Exploring the role of juvenile hormone and vitellogenin in reproduction and social behavior in bumble bees. *BMC Evolutionary Biology*, *14*(1), Article 45. https://doi.org/10.1186/1471-2148-14-45

Ankley, G. T., Bennett, R. S., Erickson, R. J., Hoff, D. J., Hornung, M. W., Johnson, R. D., Mount, D. R., Nichols, J. W., Russom, C. L., Schmieder, P. K., Serrrano, J. A., Tietge, J. E., & Villeneuve, D. L. (2010). Adverse outcome pathways: A conceptual framework to support ecotoxicology research and risk assessment. *Environmental Toxicology and Chemistry*, *29*(3), 730–741. https://doi.org/10.1002/etc.34

Ankley, G. T., LaLone, C. A., Gray, E., Villeneuve, D. L., & Hornung, M. W. (2016). Evaluation of the scientific underpinnings for identifying estrogenic chemicals in nonmammalian taxa using mammalian test systems. *Environmental Toxicology and Chemistry*, *35*(11), 2806–2816. https://doi.org/10.1002/etc.3456

Aurori, C. M., Giurgiu, A. I., Conlon, B. H., Kastally, C., Dezmirean, D. S., Routtu, J., & Aurori, A. (2020). Juvenile hormone pathway in honey bee larvae: A source of possible signal molecules for the reproductive behavior of varroa destructor. *Ecology and Evolution*, *11*(2), 1057–1068. https://doi.org/10.1002/ece3.7125

Babu, Y. S., Bugg, C. E., & Cook, W. J. (1988). Structure of calmodulin refined at 2.2 Å resolution. *Journal of Molecular Biology*, *204*(1), 191–204. https://doi.org/10.1016/0022-2836(88)90608-0

Balfanz, S., Ehling, P., Wachten, S., Jordan, N., Erber, J., Mujagic, S., & Baumann, A. (2012). Functional characterization of transmembrane adenylyl cyclases from the honeybee brain. *Insect Biochemistry and Molecular Biology*, *42*(6), 435–445. https://doi.org/10.1016/j.ibmb.2012.02.005

Bass, C., Denholm, I., Williamson, M. S., & Nauen, R. (2015). The global status of insect resistance to neonicotinoid insecticides. *Pesticide Biochemistry and Physiology*, *121*, 78–87. https://doi.org/10.1016/j.pestbp.2015.04.004

Bass, C., Puinean, A. M., Andrews, M., Cutler, P., Daniels, M., Elias, J., Paul, V. L., Crossthwaite, A. J., Denholm, I., Field, L. M., Foster, S. P., Lind, R., Williamson, M. S., & Slater, R. (2011). Mutation of a nicotinic acetylcholine receptor β subunit is associated with resistance to neonicotinoid insecticides in the aphid *Myzus persicae*. *BMC Neuroscience*, *12*, Article 51. https://doi.org/10.1186/1471-2202-12-51

Becker, R. A., Dellarco, V., Seed, J., Kronenberg, J. M., Meek, B., Foreman, J., Palermo, C., Kirman, C., Linkov, I., Schoeny, R., & Dourson, M. (2017). Quantitative weight of evidence to assess confidence in potential modes of action. *Regulatory Toxicology and Pharmacology*, *86*, 205–220. https://doi.org/10.1016/j.yrtph.2017.02.017

Berman, H. M., Ten Eyck, L. F., Goodsell, D. S., Haste, N. M., Kornev, A., & Taylor, S. S. (2005). The camp binding domain: An ancient signaling module. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(1), 45–50. https://doi.org/10.1073/pnas.0408579102

Blacquière, T., Smagghe, G., van Gestel, C. A., & Mommaerts, V. (2012). Neonicotinoids in bees: A review on concentrations, side-effects and risk assessment. *Ecotoxicology*, *21*(4), 973–992. https://doi.org/10.1007/s10646-012-0863-x

Cnaani, J., Robinson, G. E., & Hefetz, A. (2000). The critical period for caste determination in *Bombus terrestris* and its juvenile hormone correlates. *Journal of Comparative Physiology A*, *186*(11), 1089–1094. https://doi.org/10.1007/s003590000163

Corringer, P.-J., Novère, N. L., & Changeux, J.-P. (2000). Nicotinic receptors at the amino acid level. *Annual Review of Pharmacology and Toxicology*, *40*(1), 431–458. https://doi.org/10.1146/annurev.pharmtox.40.1.431

Coultrap, S. J., & Bayer, K. U. (2012). CaMKII regulation in information processing and storage. *Trends in Neurosciences*, *35*(10), 607–618. https://doi.org/10.1016/j.tins.2012.05.003

Diller, T. C., Madhusudan, Xuong, N. H., & Taylor, S. S. (2001). Molecular basis for regulatory subunit diversity in cAMP-dependent protein kinase: Crystal structure of the type IIβ regulatory subunit. *Structure*, *9*(1), 73–82, https://doi.org/10.1016/s0969-2126(00)00556-6

Doering, J. A., Lee, S., Kristiansen, K., Evenseth, L., Barron, M. G., Sylte, I., & LaLone, C. A. (2018). In silico site-directed mutagenesis informs species-specific predictions of chemical susceptibility derived from the Sequence Alignment to Predict Across Species Susceptibility (SeqAPASS) tool. *Toxicological Sciences*, *166*(1), 131–145. https://doi.org/10.1093/toxsci/kfy186

Gonzalez, G. A., Menzel, P., Leonard, J., Fischer, W. H., & Montminy, M. R. (1991). Characterization of motifs which are critical for activity of the cyclic AMP-responsive transcription factor CREB. *Molecular and Cellular Biology*, *11*(3), 1306–1312. https://doi.org/10.1128/mcb.11.3.1306-1312.1991

Gonzalez, G. A., Yamamoto, K. K., Fischer, W. H., Karr, D., Menzel, P., Biggs, W., III, Vale, W. W., & Montminy, M. R. (1989). A cluster of phosphorylation sites on the cyclic AMP-regulated nuclear factor CREB predicted by its sequence. *Nature*, *337*, 749–752. https://doi.org/10.1038/337749a0

Havukainen, H., Münch, D., Baumann, A., Zhong, S., Halskau, Ø., Krogsgaard, M., & Amdam, G. V. (2013). Vitellogenin recognizes cell damage through membrane binding and shields living cells from reactive oxygen species. *Journal of Biological Chemistry*, *288*(39), 28369–28381. https://doi.org/10.1074/jbc.m113.465021

Ihara, M., Furutani, S., Shigetou, S., Shimada, S., Niki, K., Komori, Y., Kamiya, M., Koizumi, W., Magara, L., Hikida, M., Noguchi, A., Okuhara, D., Yoshinari, Y., Kondo, S., Tanimoto, H., Niwa, R., Sattelle, D. B., & Matsuda, K. (2020). Cofactor-enabled functional expression of fruit fly, honeybee, and bumblebee nicotinic receptors reveals picomolar neonicotinoid actions. *Proceedings of the National Academy of Sciences of the United States of America*, *117*(28), 16283–16291. https://doi.org/10.1073/pnas.2003667117

Jeschke, P., Nauen, R., & Beck, M. E. (2013). Nicotinic acetylcholine receptor agonists: A milestone for modern crop protection. *Angewandte Chemie International Edition*, *52*(36), 9464–9485. https://doi.org/10.1002/anie.201302550

Jones, A. K., Raymond-Delpech, V., Thany, S. H., Gauthier, M., & Sattelle, D. B. (2006). The nicotinic acetylcholine receptor gene family of the honey bee, *Apis mellifera*. *Genome Research*, *16*(11), 1422–1430. https://doi.org/10.1101/gr.4549206

© 2022 The Authors.

Kandel, E. R. (2012). The molecular biology of memory: CAMP, PKA, CRE, CREB-1, CREB-2, and CPEB. *Molecular Brain*, *5*(1), Article 14. https://doi.org/10.1186/1756-6606-5-14

Kapheim, K. M., Smith, A. R., Ihle, K. E., Amdam, G. V., Nonacs, P., & Wcislo, W. T. (2011). Physiological variation as a mechanism for developmental caste-biasing in a facultatively eusocial sweat bee. *Proceedings of the Royal Society B: Biological Sciences*, *279*(1732), 1437–1446. https://doi.org/10.1098/rspb.2011.1652

LaLone, C. A., Villeneuve, D. L., Lyons, D., Helgen, H. W., Robinson, S. L., Swintek, J. A., Saari, T. W., & Ankley, G. T. (2016). Editor's highlight: Sequence Alignment to Predict Across Species Susceptibility (SeqAPASS): A web-based tool for addressing the challenges of cross-species extrapolation of chemical toxicity. *Toxicological Sciences*, *153*(2), 228–245. https://doi.org/10.1093/toxsci/kfw119

LaLone, C. A. (2021). Nicotinic acetylcholine receptor activation followed by desensitization contributes to abnormal foraging and directly leads to colony loss/failure. AOP: 89. AOP Wiki. https://aopwiki.org/aops/89

LaLone, C. A., Villeneuve, D. L., Wu-Smart, J., Milsk, R. Y., Sappington, K., Garber, K. V., Housenger, J., & Ankley, G. T. (2017). Weight of evidence evaluation of a network of adverse outcome pathways linking activation of the nicotinic acetylcholine receptor in honey bees to colony death. *Science of the Total Environment*, *584–585*, 751–775. https://doi.org/10.1016/j.scitotenv.2017.01.113

Leboulle, G., & Müller, U. (2004). Synergistic activation of insect cAMP-dependent protein kinase A (type II) by cyclicAMP and cyclicGMP. *FEBS Letters*, *576*(1–2), 216–220. https://doi.org/10.1016/j.febslet.2004.08.079

Matsuda, K. (2020). Robust functional expression of insect nicotinic acetylcholine receptors provides new insights into neonicotinoid actions and new opportunities for pest and vector control. *Pest Management Science*, *77*(8), 3626–3630. https://doi.org/10.1002/ps.6182

Matsuda, K., Ihara, M., & Sattelle, D. B. (2020). Neonicotinoid insecticides: Molecular targets, resistance, and toxicity. *Annual Review of Pharmacology and Toxicology*, *60*(1), 241–255. https://doi.org/10.1146/annurev-pharmtox-010818-021747

Mayr, B., & Montminy, M. (2001). Transcriptional regulation by the phosphorylation-dependent factor CREB. *Nature Reviews Molecular Cell Biology*, *2*, 599–609. https://doi.org/10.1038/35085068

Michener, C. D. (2000). *The bees of the world*. Johns Hopkins University Press.

Montminy, M. (1997). Transcriptional regulation by cyclic AMP. *Annual Review of Biochemistry*, *66*, 807–822. https://doi.org/10.1146/annurev.biochem.66.1.807

Organisation for Economic Co-operation and Development. (2013). *Guidance document on developing and assessing adverse outcome pathways* (OECD Series on Testing and Assessment, No. 184).

Organisation for Economic Co-operation and Development. (2018). *Users' handbook supplement to the guidance document for developing and assessing adverse outcome pathways* (OECD Series on Adverse Outcome Pathways, No. 1). https://doi.org/10.1787/5jlv1m9d1g32-en

Perry, T., Chen, W., Ghazali, R., Yang, Y. T., Christesen, D., Martelli, F., Lumb, C., Bao Luong, H. N., Mitchell, J., Holien, J. K., Parker, M. W., Sparks, T. C., & Batterham, P. (2021). Role of nicotinic acetylcholine receptor subunits in the mode of action of neonicotinoid, sulfoximine and spinosyn insecticides in *Drosophila melanogaster*. *Insect Biochemistry and Molecular Biology*, *131*, Article 103547. https://doi.org/10.1016/j.ibmb.2021.103547

Pollesch, N. L., Villeneuve, D. L., & O'Brien, J. M. (2019). Extracting and benchmarking emerging adverse outcome pathway knowledge. *Toxicological Sciences*, *168*(2), 349–364. https://doi.org/10.1093/toxsci/kfz006

Potts, S. G., Biesmeijer, J. C., Kremen, C., Neumann, P., Schweiger, O., & Kunin, W. E. (2010). Global pollinator declines: Trends, impacts and drivers. *Trends in Ecology and Evolution*, *25*(6), 345–353. https://doi.org/10.1016/j.tree.2010.01.007

Sattelle, D. B., Jones, A. K., Sattelle, B. M., Matsuda, K., Reenan, R., & Biggin, P. C. (2005). Edit, cut and paste in the nicotinic acetylcholine receptor gene family of *Drosophila melanogaster*. *BioEssays*, *27*(4), 366–376. https://doi.org/10.1002/bies.20207

Schmolke, A., Galic, C., Feken, M., Thompson, H., Sgolastra, F., Pitts-Singer, T., Elston, C., Pamminger, T., & Hinarejos, S. (2021). Assessment of the vulnerability to pesticide exposures across bee species. *Environmental Toxicology and Chemistry*, *40*(9), 2640–2651. https://doi.org/10.1002/etc.5150

Shimada, S., Kamiya, M., Shigetou, S., Tomiyama, K., Komori, Y., Magara, L., Ihara, M., & Matsuda, K. (2020). The mechanism of loop C-neonicotinoid interactions at insect nicotinic acetylcholine receptor α1 subunit predicts resistance emergence in pests. *Scientific Reports*, *10*, Article 7529. https://doi.org/10.1038/s41598-020-64258-z

Shimomura, M., Yokota, M., Matsuda, K., Sattelle, D. B., & Komai, K. (2004). Roles of loop C and the loop B–C interval of the nicotinic receptor α subunit in its selective interactions with imidacloprid in insects. *Neuroscience Letters*, *363*(3), 195–198. https://doi.org/10.1016/j.neulet.2003.12.115

Taylor, S. S., Buechler, J. A., & Yonemoto, W. (1990). cAMP-dependent protein kinase: Framework for a diverse family of regulatory enzymes. *Annual Review of Biochemistry*, *59*, 971–1005. https://doi.org/10.1146/annurev.bi.59.070190.004543

Villeneuve, D. L., Crump, D., Garcia-Reyero, N., Hecker, M., Hutchinson, T. H., LaLone, C. A., Landesmann, B., Lettieri, T., Munn, S., Nepelska, M., Ottinger, M. A., Vergauwen, L., & Whelan, M. (2014). Adverse outcome pathway (AOP) development I: Strategies and principles. *Toxicological Sciences*, *142*(2), 312–320. https://doi.org/10.1093/toxsci/kfu199

Wang, N. X., Watson, G. B., Loso, M. R., & Sparks, T. C. (2016). Molecular modeling of sulfoxaflor and neonicotinoid binding in insect nicotinic acetylcholine receptors: Impact of the Myzus B1 R81T mutation. *Pest Management Science*, *72*(8), 1467–1474. https://doi.org/10.1002/ps.4220

Wayman, G. A., Tokumitsu, H., Davare, M. A., & Soderling, T. R. (2011). Analysis of Cam-kinase signaling in cells. *Cell Calcium*, *50*(1), 1–8. https://doi.org/10.1016/j.ceca.2011.02.007

 © 2022 The Authors.