SEQUENCE ALIGNMENT TO PREDICT ACROSS SPECIES SUSCEPTIBIILTIY (SeqAPASS): CROSS SPECIES EXTRAPOLATION IN THE 21ST CENTURY

[1]Carlie A. LaLone

[1]Daniel L. Villeneuve

[2]David Lyons

[1]Gerald T. Ankley

[1]United States Environmental Protection Agency, Office of Research and Development, National Health and Environmental Effects Laboratory, Duluth, MN, United States

[2]United States Environmental Protection Agency, Office of Research and Development, National Center for Computational Toxicology, Research Triangle Park, NC, United States

LaLone.Carlie@epa.gov

INTRODUCTION

An objective of the U.S. Environmental Protection Agency is to develop innovative approaches and decision support tools for the safety assessment of current and emerging chemicals in relation to their impact on the environment. To support this goal, there is a compelling need for novel, quantitative, scientifically-based approaches to evaluate and predict species differences in susceptibility to various types of chemical contaminants particularly when testing resources are limited and/or rapid turn-around is required. Within the Agency, high-throughput screening (e.g., EPA's ToxCast program) is enhancing the efficiency of toxicity testing. Therefore, with this innovative technology it is necessary to have methods in place to define the taxonomic domain of applicability as a means to extrapolate data across species and to aid in understanding relevant molecular initiating events across species as they relate to the adverse outcome pathway (AOP) conceptual framework.

METHODS

This research was conducted to evaluate and develop a computational method to assess molecular target similarity, determined by comparing protein sequence/structural information across species (primary amino acid sequence, functional domain, and individual residues known to interact with a chemical or necessary for maintaining conformation) as a basis to predict relative species intrinsic susceptibility to chemicals with known modes of action (e.g., drugs and pesticides). Specifically, this research 1) applied quantitative metrics of protein sequence similarity to predict the relative intrinsic susceptibility of a broad range of species to a number of specifically-acting chemicals (e.g., human and veterinary drugs; and pesticides with known molecular targets); 2) developed a web-based tool, Sequence Alignment to Predict Across Species Susceptibility (SeqAPASS), that efficiently performs automated protein comparisons based on the defined sequence/structural similarity metrics, and allowed for rapid predictions of chemicals susceptibility across taxa; 3) tested those predictions with currently available and newly derived empirical toxicity data to ascertain pathway conservation; and 4) utilized the results to

refine the quantitative metrics and improve their predictive utility and determine the appropriate domain of applicability. The species extrapolation approach is predicated on the assumption that the molecular targets in target species are the same as those that occur in non-target species. Predictions about the potential intrinsic susceptibility of the non-target species to the chemical were made based on protein alignment relative to the target species (e.g., humans, livestock, bacteria, or viruses for human/veterinary drugs). Closer phylogeny between the molecular target of a non-target species and that of the animal treated with the chemical would be indicative of greater potential for adverse effects.

RESULTS/DISCUSSION

Case studies focused on the action of organophosphates/carbamates on the daphnid acetylcholinesterase enzyme, estrogens/anti-estrogens on the human estrogen receptor alpha, and neonicotinoids on the honey bee nicotinic acetylcholine receptor were used as a proof of concept for this predictive methodology. Briefly, sequence similarity analyses were conducted to predict relative species intrinsic susceptibility and compared to available empirical toxicity data relevant to the chemical of concern providing evidence that this strategy has merit [1]. Other work in support of validating this method demonstrated how knowledge of conservation of molecular targets, in conjunction with defined adverse outcome pathways can be used to guide empirical toxicity testing in the laboratory [2]. The SeqAPASS tool will be publically released to the internet for rapid evaluations of protein conservation across taxa relevant for researchers and risk assessors alike.

CONCLUSIONS

We envision applications of this methodology for risk assessment purposes, particularly in instances where there is crude or otherwise limited empirical toxicity information available (e.g., as a complement to species sensitivity distributions) or when a higher level of uncertainty is acceptable for screening or prioritizing chemicals for testing or monitoring. When used in this manner the methodology can be utilized as a broad class level prediction of relative species intrinsic susceptibility. Therefore, a relevant application may be to inform test designs that currently exist for human and veterinary drug environmental impact assessments or applicable to current registration of pesticides by assisting in species and endpoint selection for required toxicity testing. In the context of a proposed 21st century approach to toxicity testing, sequence similarity-based tools could help define the taxonomic domain of applicability for mechanism- or toxicity pathway-based in vitro screening data, such as those currently being generated by EPA's ToxCast program. Thus, one can envision sequence-based species extrapolation tools as an important component of green chemistry, sustainable product design, efficient prospective toxicity testing, and environmental monitoring.