**RRES Press Release 4 July 2023 New Tools for Disease Detectives**

*New curation will accelerate the study of interactions between species*

Researchers struggling to sift through mountains of apparently unconnected information in scientific publications now have a new set of tools at their disposal that can match up information on completely unrelated species.

This will be especially useful for scientists studying interactions between pathogens and hosts that lead to disease, as it will allow them to more quickly investigate how such interactions are taking place right down to the molecular level.

Researchers at Rothamsted Research, in collaboration with the University of Cambridge, developed a set of software tools that allow researchers to select information from a scientific publication, collect that information in one place (such as a database) and ensure that the information is represented using standard terminology. The team tested the framework using the [Pathogen–Host Interactions Database](http://www.phi-base.org/) (PHI-base) as a case study. The team also created a new concept of multispecies genotype – the metagenotype – to help capture changes in both the pathogen’s ability to cause disease, and the host’s ability to resist disease.

“The amount of data being produced in some areas of genomics is increasing dramatically each month,” said Dr Alayne Cuzick, who led the study. “The sheer quantity and complexity of this deluge creates huge challenges for researchers, particularly if they are looking for interactions between completely unrelated species, as is often the case when studying disease-causing pathogens and their hosts.”

The researchers found that existing software tools for curating peer-reviewed literature in the life sciences were designed solely for a single species, or closely-related species (for example, fruit flies). No tools were available to curate interactions between multiple different species, particularly pathogens and their hosts. Therefore, there was no support for databases like PHI-base, which curates knowledge from the text, tables and figures published in over 200 journals.

“The vast amount of data exploring interactions between species is dispersed across hundreds of different journals – many of which require expertise in highly specialised terminologies and concepts,” said Dr Cuzick. “Furthermore, the data are often represented in non-standard formats making it difficult for both researchers and machine learning systems to access. However, across the research community we are fully committed to making this data FAIR: Findable, Accessible, Interoperable and Reusable. This new tool will help us to achieve just that.”

It is also hoped that these new tools could be used by researchers in other disciplines to compare and contrast interactions across multiple species at different scales (microscopic and macroscopic).

“Ultimately, this should assist the development of new approaches to reduce the impact of pathogens on humans, livestock, crops and ecosystems, thereby reducing disease, whilst increasing food security and biodiversity,” said co-author Dr Kim Hammond-Kosack.

Publication

Alayne Cuzick,James Seager, Valerie Wood, Martin Urban, Kim Rutherford, Kim E Hammond-Kosack (2023) A framework for community curation of interspecies interactions literature eLife 12:e84658. https://doi.org/10.7554/eLife.84658