Adverse outcome pathways coming to life

Summary

Risk assessors struggle to keep up with the growing number of chemicals requiring testing. The aim is to shift from costly and ethically challenging animal experimentation to more efficient and humane *in vitro* methods, *in silico* models, and human data for risk assessments and human safety promotion. However, the transition involves the challenges that come with the development of novel techniques as alternatives to animal testing.

In order to support risk assessments, the Adverse Outcome Pathway (AOP) approach has been introduced to capture and organize literature-derived mechanistic knowledge of toxicological processes to guide the paradigm shift in risk assessments toward alternative models. It does this by separating the cascade of biological perturbations upon stressor interaction into smaller, measurable effects called Key Events (KE). Despite the increasing number of AOPs and the growing momentum of the use of AOPs in risk assessments, there are challenges in validating and incorporating *in vitro* targeted assays and large-scale omics datasets into the testing strategies. Although a promising tool in many fields of biomedical research to study molecular processes such as the understanding of toxicological responses, the production and use of transcriptomic data in risk assessment face barriers related to reproducibility, reliability, and acceptance within the risk assessment community. This is why this thesis had the two aims of improving AOP usability and establishing a method to analyse and interpret transcriptomic data that utilises and extends AOPs to facilitate better insights into KE activation, with the ultimate goal to increase the overall acceptance of transcriptomic data in risk assessments.

Improving AOP usability

Before being able to make a link between transcriptomic data and AOPs, the work presented in this thesis involved the exploration of the overall usability of AOPs which are generally stored in the AOP-Wiki and seek interoperability with the established molecular pathway database called WikiPathways to expand KEs with molecular entities and processes. This has led to an introductory description of WikiPathways and highlighting the strengths of community-driven developments and methods of accessing data in Chapter 2. This was followed by investigating the level of coverage of KEs in AOP-Wiki as molecular pathways in WikiPathways in Chapter 3. This has shown that the majority of early KEs in AOP-Wiki have corresponding molecular pathways in WikiPathways, and that opportunities exist to make the AOP-Wiki more linked to other biological databases by using the ontological annotations of KEs and molecular entities captured in their description. To further increase the usability of AOPs in the AOP-Wiki, this thesis has resulted in a more Findable, Accessible, Interoperable, and Reusable (FAIR) version of the AOP-Wiki by employing semantic web technologies and producing a Resource Description Framework (RDF) version of the data, as described in Chapter 4. As an extension to the AOP-Wiki RDF, various tables of the AOP-DB have been modelled into RDF format using the same principles, which was presented in Chapter 5. The producing of the AOP-Wiki RDF and AOP-DB RDF and loading these into SPARQL endpoints allow the exploration and integration of the data with external resources and allow computational querying of the contents, which has been illustrated in Chapter 6 as a Jupyter notebook. The flexible, reproducible workflow that is presented

accesses and uses a range of public services to find and analyse data to support an AOP of interest, showing the utility of the semantic web versions of AOP-Wiki, AOP-DB, and WikiPathways.

Extending AOPs with molecular pathways

With the establishment of the AOP-Wiki and WikiPathways, Chapter 7 presents the establishment of an analysis method for transcriptomic data, utilising WikiPathways as an integrative platform between AOPs and molecular pathways with an intermediate model called the molecular AOP. It was expected that integrating these biological databases with experimental data holds promise for improving transcriptomic data usability by enabling data interpretation and links to KEs to measure biological processes and KE activation. This integration potentially enables the use of transcriptomics data to support AOP-based risk assessment strategies by providing measurements and visualizations of KE activity. As illustrations and proof of principles, case studies were performed on a liver steatosis AOP network and on an AOP that initiates with mitochondrial complex I inhibition in neuronal cells. This has shown us that there is value in the model to analyse and interpret transcriptomic data and generate hypotheses on KE activation. However, the case studies have shown the challenges of modelling molecular AOPs and their use with more extensive datasets, and require more comprehensive testing to define their domain of applicability and technology readiness level.

Impact of this research

The overall goal of this thesis was to make better use of existing mechanistic knowledge in AOPs and utilise large-scale omics approaches based on *in vitro*, to drive the transition away from animal testing for the risk assessment of chemicals and nanomaterials. The increased accessibility and interoperability of the AOP-Wiki and AOP-DB could lead to more effective use of AOP knowledge, leading to a more efficient establishment of knowledge-driven Integrated Approaches to Testing and Assessment (IATA). Ultimately, this can facilitate better and faster risk assessment approaches to ensure human and environmental safety. Regarding the proposed method of analysing transcriptomic data by utilising molecular AOPs, this novel method can aid the integration and utility of transcriptomic data in risk assessment, providing a clear analysis and interpretation model to assess KE activation.

Conclusion

This thesis project aimed to enhance the usability of AOPs and transcriptomic data for risk assessment. First, this thesis explored integrating the molecular pathway database of WikiPathways with AOP-Wiki to allow the establishment of a connection between transcriptomic data and AOPs by linking KEs to biological processes represented in molecular pathways. This thesis also emphasized the importance of making the data in AOP-Wiki more interoperable and FAIR. This facilitated an automated workflow in a Jupyter Notebook that could find and analyze experimental data to support a specific AOP of

interest. Finally, the introduced molecular AOP approach allows for the visualization and reproducible analysis of transcriptomic data to identify KE activation, which can bridge the gap between big data approaches and the risk assessment community.