Characterization of the Chemical Space of Genotoxicity Databases in Comparison to Pesticides and their Metabolites

Daniel Foil[1,2], Jeannette König[1], Kristin Herrmann[1], Johannes Kirchmair[2,3], Carsten Kneuer[1]

[1] *Toxicology of Active Substances and their Metabolites, Department Pesticides Safety, German Federal Institute for Risk Assessment (BfR), Max-Dohrn-Straße 8-10, 10589 Berlin, Germany*
[2] *Department of Pharmaceutical Sciences, Division of Pharmaceutical Chemistry, Faculty of Life Sciences, University of Vienna, 1090 Vienna, Austria*
[3] *Christian Doppler Laboratory for Molecular Informatics in the Biosciences, Department for Pharmaceutical Sciences, University of Vienna, 1090 Vienna, Austria*

Information about the genotoxic potential of substances present in plant protection products and biocides is crucial for appropriate regulatory decision-making, but this information is not always possible to obtain through the traditional experimental methods, especially in the case of low-abundance or unstable impurities and metabolites. *In silico* prediction of genotoxic effects, especially (Q)SAR, is increasingly accepted as an alternative to *in vitro* evaluation of genotoxicity, particularly for the Ames test [1]. However, the majority of the training data in the models is derived from pharmaceutical data, which may not adequately cover pesticides-relevant chemical space.

Herein, we compare the chemical space coverage of multiple datasets for two apical endpoints, gene mutation (in particular, Ames mutagenicity) and chromosome damage, to the chemical space of pesticides-relevant substances, as represented by datasets retrieved from regulatory authorities. As the scientific community seeks to move from a qualitative towards a quantitative characterization of genotoxicity, i.e., by regression modeling to support understanding of the dose-response relationship, we also analyzed the pesticide chemical space coverage by a dataset containing dose-response data [2].

Ames mutagenicity datasets cover a large proportion of pesticide chemical space, but only partially overlap with pesticide datasets. The chemical space coverage of chromosome damage data is smaller than that of gene mutation data, but still large enough for predictive modeling of some pesticide substances. Substances for which dose-response data are available cover a more modest subset of pesticide chemical space, but sufficient data exists for preliminary predictive modeling of the dose-response relationship of pesticides. Existing datasets provide a sound basis for the use of *in silico* techniques to predict the mutagenicity of pesticides, and chemical space coverage should already be sufficient to allow prediction of chromosome damage for a subset of pesticides. Moreover, the expansion of data sharing initiatives and generation of new experimental data, as well as increased availability of dose-response data, will lead to the maturation of predictive modeling of pesticide genotoxicity.

Bibliography

[1] R. Benigni et al. EFSA Supporting Publications. 16 (2019) 1598E.
[2] J. Menz et al. Archives of Toxicology. 93 (2023) 2303-2308.