Proposals for NORMAN Joint Programme of Activities 2024

Title	**Integration of computational toxicity driver prioritization tools to support non-target screening workflows in high-throughput effect-directed analysis**
Type of activity	Pilot study
Leader	UFZ
Topic / activities	***Background / Justification for the proposed activity:*** In the past few years, effect-directed analysis has made significant advancement in the direction of getting more high-throughput. These changes have been particularly marked by the automation and miniaturisation of fractionation, as well as the optimization and downscaling of bioassays. It is true that high resolution mass spectrometry coupled to liquid chromatography and consequently non-target screening enabled us to identify novel risk driver compounds in aquatic environments. However, the current identification approaches have two main challenges two overcome: (i) prioritization of features to be subjected to identification efforts among a list that comprises thousands of features, (ii) prioritization of generated candidate structures of these selected features by in-silico fragmentation tools, which can reach hundred possible structures depending on the assigned molecular formula. Recently, several computational tools have been developed supported by machine learning approaches in order to enhance toxicity prediction based on different available information such as ms2 spectra, candidate structure, ionization behaviour...etc. We believe that many of the developed tools for toxicological endpoint prediction are perfectly suited to toxicity driver prioritisation and have not yet been explored in the context of EDA. **Aim:** With this JPA we aim to implement the most promising tools in order to improve the efficiency and success rates of current compound identification workflows/protocols and adapt them to HT-EDA applications. ***Description of the proposed activity and expected outcomes for 2024:*** Activities will be organised around three objectives: **Activity 1:** Evaluate computational tools that are established to prioritize candidate structures of toxicity drivers within non-target screening workflows adapted in EDA studies (supported by the info gathered for EDA review paper currently being written by the lead of WG3). **Activity 2:** Computationally associate the structures in the NORMAN Suspect List to potential effects on the gene/pathway level for environmentally relevant endpoints that are suitable for EDA studies such as estrogenicity, androgenicitiy, and neurotoxicity by using *deepFPlearn* (https://doi.org/10.1093/bib/bbac257). **Activity 3:** Validate the efficiency of selected tools for prioritizing potential toxicity driver features/structures using the data obtained in the SOURCES project as the basis of the research. Mainly, we will assess if the prioritised compounds with these tools explain successfully the toxicity of the selected SOURCES samples. **Activity 4:** Establish optimal endpoint-tailored toxicant identification protocols for HT-EDA that incorporates these tools and make them accessible to the rest of the community. Expected outcomes: - The results of Activity 1 will be used in the critical review article “Effect Direct Analysis for identification of toxic emerging pollutants” carried out by WP3 - The results of Activity 2 will allow us to use the Norman SLE more efficiently in EDA studies by having endpoint-tailored and thus prioritized suspect lists - The results of Activity 3 will enable the identification of source-specific toxicity drivers in selected SOURCES samples -The results of Activities 3 and 4 will be part of a series of publications where endpoint specific drivers will be identified using HT-EDA on the SOURCES samples

	Added value / Link with other NORMAN activities and / or other projects This activity will be directly linked to the NORMAN Sources project as a supporting tool in the implementation of HT-EDA for the identification of source-specific risk drivers. It will expand the utilization of the NORMAN SLE and facilitate its integration in EDA investigations. It will strengthen the cooperation between WG3, WG2 (bioassays) and Cross-Working Group Activity Non-target Screening. The results will be beneficial for all NORMAN community, among others especially for WG1: Prioritization of emerging substances.
Participants	*UFZ* *UPV/EHU* *VU* *BFG* *EAWAG* *Goethe University Frankfurt*
Proposed in-kind contribution	All practical work for chemical analysis and biotesting will be done as in-kind contribution. Biotesting consumables will be financed by the SOURCES 2023 JPA budget. Support is needed to train the *deepFPlearn* models for additional endpoints with the updated TOX21 data and for buying standard reference materials for confirmation of the prioritized compounds to be identified.
Contribution needed from NORMAN Association[1]	Total: 9,2k€ (50% non-academic staff member for 3 months, 2,400k€ per month, total 7,200 k€ and 2,000 k€ reference standards)

1