



Project no.: 722493

Project acronym: NaToxAq

Marie Skłodowska-Curie Action

Start date of project: 1-January 2017

Duration: 4 years

Deliverable reference number and title

D6.16 Upload of data to databases and archives 6

Due date of deliverable: Month 26 (29 February 2020)

Actual submission date: 28 February 2020

Organisation name of lead beneficiary for this deliverable:

Helmholtz Centre for Environmental Research (UFZ) supported by Federal Department of Economic Affairs, Education and Research (WBF)

Project co-funded by the European Commission within the Horizon 2020 Programme		
Dissemination Level		
PU	Public	X
PP	Restricted to other programme participants (including the Commission Services)	
RE	Restricted to a group specified by the consortium (including the Commission Services)	
CO	Confidential, only for members of the consortium (including the Commission Services)	

1. Introduction and Objective(s)

In agreement with the DOW, NaToxAq successfully uploaded several mass spectra of natural toxins to the European MassBank (<https://massbank.eu/MassBank>) allowing the share, exchange and free access of the scientific community to these data and supporting and stimulating the collaboration and exchange within the project. To facilitate a broader awareness of natural toxins in monitoring and target screening, a first list of natural toxins in MassBank was submitted to NORMAN Suspect List Exchange (<https://www.norman-network.com/?q=node/236>). The list is not processed yet, but part of the Appendix of this deliverable. The upload of raw mass spectral data to GNPS (<https://gnps.ucsd.edu/ProteoSAFe/static/gnps-splash.jsp>) is planned and will be performed until finalization of the NaToxAq project.

2. Strategy for solving the task

The lead beneficiary UFZ is collecting standard substances of other NaToxAq partners for LC-HRMS analysis and processing of MassBank spectra in order to facilitate the task. UFZ started first discussions with GNPS maintainers how to integrate the raw mass spectral data of NaToxAq in GNPS.

3. Delays and difficulties

There are no delays and difficulties in deriving and uploading mass spectral records in MassBank. The NaToxAq partners Agroscope, Recetox, UCPH and UFZ provided analytical standards to UFZ which are measured, processed and uploved to MassBank in batches.

4. Result

In this deliverable we report the upload of more than 1500 mass spectra to the specific NaToxAq database in MassBank (<https://massbank.eu/MassBank>, NaToxAq database: <https://goo.gl/joofrG>). Due to the new release process, the data is not yet uploaded to MassBank, but available in the MassBank consortium repository in Github (<https://github.com/MassBank/MassBank-data/tree/dev/NaToxAq>). The data was committed to the development branch with commit 026ba9f529eace6e3d0f73a459cad583a78213b4 and will be released and deployed to MassBank with the next data release end of February 2020.

The data includes mass spectra of analytical standards kindly provided by Agroscope (Switzerland) related to the publication of Günthardt et al. [1].

After upload, the records can be reached by the Record Index (<https://massbank.eu/MassBank/RecordIndex>) list or via the Search (<https://massbank.eu/MassBank/Search>). The database is directly reachable via <https://goo.gl/joofrG>. The landing page of each record contains metadata regarding the chemical identity, instruments, analytical conditions and the mass spectral data itself. The records were automatically derived from the raw data files using RMassBank 2.14.1 [2; <https://github.com/MassBank/RMassBank>].

In total, 3756 mass spectra of 130 natural toxins are available in MassBank, which are a great outcome of the efforts of NaToxAq to collate information about natural toxins and to provide analytical open access to improve the identification of those compounds in the environment.

In order to facilitate the awareness of natural toxins, the NaToxAq list of natural toxins in MassBank was submitted to the responsible person to be uploaded to the NORMAN Suspect List Exchange (<https://www.norman-network.com/?q=node/236>).

See <https://natoxaq.ku.dk/toxins/> for an excel sheet of MassBank accessions.

References

1. Günthardt, B.F., et al., *Comprehensive Toxic Plants–Phytotoxins Database and Its Application in Assessing Aquatic Micropollution Potential*. Journal of Agricultural and Food Chemistry, 2018. **66**(29): p. 7577-7588.
2. Stravs, M.A., et al., *Automatic recalibration and processing of tandem mass spectra using formula annotation*. Journal of Mass Spectrometry, 2013. **48**(1): p. 89-99.