

## Center for Integrative Bioinformatics (CIBI Tübingen)

### Tools for the Reproducible Analysis of Mass Spectrometric Data

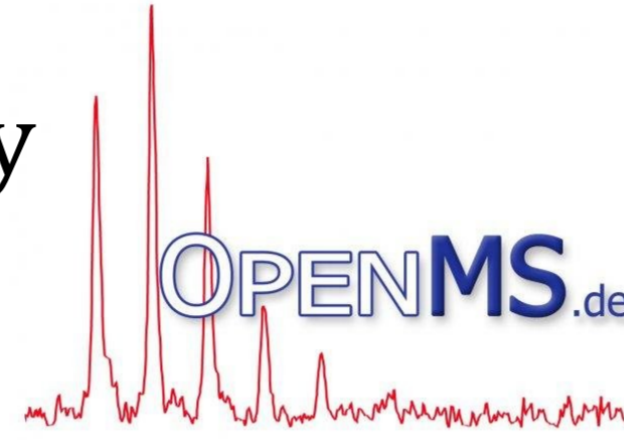
FKZ: 31A535A

Oliver Alka, Felix Bartusch, Leon Bichmann, Benjamin Gläßle, Maximilian Hanussek,  
Julianus Pfeuffer, Oliver Kohlbacher, Jens Krüger, Timo Sachsenberg

#### Short description of the project

##### OpenMS:

- Cross-platform mass spectrometry data analysis framework and programming library
- More than 180 tools for proteomics and metabolomics analysis tasks
- Tools can be chained into complex analysis workflows
- Workflows for high-throughput data analysis on HPC and cloud resources
- Consulting, training, and support for users and developers



##### de.NBI cloud site Tübingen:

- State-of-the-art technologies
- Access to more than 1,650 compute cores, 15 TB RAM, 13 PB storage
- Easy application and fast setup of own projects



#### de.NBI services

##### OpenMS:

- Consulting and software support for both users and developers
- Integration of external code, data repositories, or software contributions into OpenMS
- Train users and external developers
- Improvements to software documentation and teaching material
- Two software releases per year, including: binary installer, container, KNIME plugins, python packages, bioconda packages, and integration into commercial platforms
- Maintenance of build and test infrastructure

#### de.NBI cloud activities

- Established de.NBI cloud site in Tübingen
- Upcoming extension with 30 high-memory nodes
- Network infrastructure upgrade (Infiniband, 40GE)
- Currently more than 20 active projects with more than 60 individual users
- Two computationally intensive mass spectrometry projects started
  - KNIME in the cloud
  - *de novo* Metabolite identification

#### Staff

- 3 FTE for CIBI-1 and de.NBI Cloud paid through de.NBI and an additional 3 FTE own contribution
- All positions are currently filled

#### Progress report

##### OpenMS:

##### Infrastructure development with focus on cloud

- Automatic container builds for cloud computing
- KNIME and Nextflow cloud workflows (nf-core)
- Increased interoperability with public repositories (PRIDE)
- Consulting and software support

##### Elixir activities

- Involved in four implementation studies
- Training coordination (O. Kohlbacher ELIXIR-DE TC)
- Involvement in EXCELERATE WP (Proteomics Beacon)
- Communities: proteomics (coord.), metabolomics



##### Training and education

- Revised training material
- Developer, user meetings, and conferences
- Online lectures and tutorials
- Training coordination (O. Kohlbacher de.NBI TC)

##### Translation to industry

- STORM Therapeutics, detection of modified RNAs
- Workflow integration in *Thermo Fisher* ProteomeDiscoverer

##### Interaction w/ other activities

- Close interaction with Medical Informatics Initiative (DIFUTURE consortium) on omics data analysis
- Interaction with GO-FAIR initiative and Personal Health Train implementation network

##### de.NBI Cloud services

- Support of bioinformatics tool integration and service
- Specific data mirroring for proteomics and proteogenomics (CPTAC, TCGA, ICGC)

#### de.NBI training and education

##### 10 training events in 2018

Jan 2018	(Ghent, BE):	EUBIC Developer Meeting
Mar 2018	(Berlin, DE):	OpenMS workshop at the KNIME Spring Summit Workshop 2018
Mar 2018	(Wittenberg, DE):	de.NBI Winter School on Computational Metabolomics
Apr 2018	(Troia, PT):	OpenMS Developer Meeting
May 2018	(Boston, US):	May Institute Proteomics and Metabolomics with OpenMS
Jul 2018	(Konstanz, DE):	Data Science for Life Sciences
Sep 2018	(Konstanz, DE):	De.NBI CIBI KNIME Cluster and Cloud Hackathon
Sep 2018	(Athens, GR):	Computational Mass Spectrometry with OpenMS (ECCB)
Sep 2018	(Berlin, DE):	2 <sup>nd</sup> CIBI User Meeting
Sep 2018	(Vienna, AT):	Label-free quantification with OpenMS (GCB)

#### Selected Publications

OpenMS: a flexible open-source software platform for mass spectrometry data analysis  
HL Röst, T Sachsenberg, S Aiche, C Bielow, H Weisser, et al. – *Nat. Methods*

OpenMS-A platform for reproducible analysis of mass spectrometry data  
Pfeuffer, J, Sachsenberg, T, Alka, O., Walzer, M., Fillbrunn, A., Nilse, L., Schilling O, Reinert K, Kohlbacher, O. *Journal of Biotechnology*

BioContainers: An open-source and community-driven framework for software standardization.  
Leprevost FD, Grüning BA, Alves Aflitos S, Röst HL, Uszkoreit J, Barsnes H, Vaudelet M, Moreno P, Gatto L, Weber J, Bai M, Jimenez RC, Sachsenberg T, Pfeuffer J, Vera Alvarez R, Griss J, Nesvizhskii AI, Perez-Riverol Y. *Bioinformatics*

Recognizing millions of consistently unidentified spectra across hundreds of shotgun proteomics datasets  
J Griss, Perez-Riverol, Y., Lewis, S, Tabb, DL, Dienes, JA, del-Toro, N, Rurik, M, Walzer, MM, Kohlbacher, O, Hermjakob, H, Wang, R, and Vizcalno, JA – *Nat. Methods*

Ultrastructural alterations in the retinal pigment epithelium and photoreceptors of a Stargardt patient and three Stargardt mouse models: indication for the central role of RPE melanin in oxidative stress. Tsubitiz T, Tschulakov AV, Tikhonov M, et al. ed. *PeerJ*. 2018.

Containerization of Galaxy Workflows increases Reproducibility  
F. Bartusch, M. Hanussek, and J. Krüger. Proceedings of the bwHPC Symposium 2017, Tübingen, p 16-19, dx.doi.org/10.15496/publikation-25200.