Integrating Untargeted Metabolomics and Genome Mining to Unlock the Natural Product Potential of *Actinoallomurus*

Alessa Sauter[a], Andrea Gentile[a], Sonia I. Maffioli[a], Matteo Simone[a], David Faurdal[b], Oliwia Vuksanovic[b], Troels Ø. Hansen[b], Tue S. Jørgensen[b], Tilmann Weber[b], Stefano Donadio[a], Margherita Sosio[a]

[a]NAICONS Srl, Milan, Italy
[b]The Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark, Kgs.Lyngby, Denmark

Keywords: genome-metabolome integration, paired-omics approach, data-driven compound prioritization, rare Actinomycetes, query-based compound detection

The growing threat of antibiotic resistance and the limited pace of new antibiotic discovery highlight the urgent need for novel natural products spanning medical, agricultural, and other relevant applications. Actinomycetes are among the most prolific producers of specialized metabolites, yet a large fraction of their biosynthetic potential remains unexplored.
To access this hidden chemical diversity, a paired omics approach was applied, by integrating untargeted metabolomics and genome mining, to a rare and underexplored genus of Actinomycetes, *Actinoallomurus*. This integrative strategy offers key advantages: it decreases rediscovery rates, improves strain prioritization, and accelerates the identification of novel compounds.

Metabolomic analysis of 170 *Actinoallomurus* strains generated HRLC–MS/MS data from 340 extracts, which were processed using the Micro4all pipeline[1]. This analysis revealed nearly 12,000 molecular features, more than 90% of which lacked known annotations.
Complementary genomic analysis of 143 of these strains uncovered more than 3,600 BGCs via antiSMASH[2] and over 650 GCFs using BiG-SCAPE[3]. Most showed low similarity to known clusters, again indicating a high degree of novelty.
The integration of both datasets enabled targeted prioritization, purification, and structural characterization of a novel polyketide compound.

In addition, a new query-based approach was developed that combines high-throughput mass spectrometry and comparative metabolomics to reliably detect sulfur-rich compounds directly in high-resolution MS data. This method offers a complementary strategy for integrating genomic and metabolomic information, facilitating genome-guided discovery of sulfur-rich peptides and laying the foundation for more efficient, data-driven compound isolation and characterization.

References

[1] Simone M., Iorio M., Monciardini P., et al., (2024). J. Nat. Prod. 87(11): 2615–2628.
[2] Blin K., Shaw S., Augustijn H.E., Reitz Z.L., Biermann F., Alanjary M., Fetter A., Terlouw B.R., Metcalf W.W., Helfrich E.J.N., van Wezel G.P., Medema M.H., Weber T. (2023). Nucleic Acids Res. 51(W1).
[3] Navarro-Muñoz J.C., Selem-Mojica N., Mullowney M.W., et al., (2020). Nat. Chem. Biol. 16(1): 60-68.