Mining Genomes of Three Marine Sponge-Associated Actinobacterial Isolates for Secondary Metabolism

Hannes Horn, Ute Hentschel, Usama Ramadan Abdelmohsen

Department of Botany II, Julius-von-Sachs Institute for Biological Sciences, University of Würzburg, Würzburg, Germany; Department of Marine Microbiology, GEOMAR Helmholtz Centre for Ocean Research, RD3 Marine Microbiology and Christian-Albrechts University of Kiel, Kiel, Germany

Here, we report the draft genome sequences of three actinobacterial isolates, Micromonospora sp. RV43, Rubrobacter sp. RV113, and Nocardiopsis sp. RV163 that had previously been isolated from Mediterranean sponges. The draft genomes were analyzed for the presence of gene clusters indicative of secondary metabolism using antiSMASH 3.0 and NapDos pipelines. Our findings demonstrated the chemical richness of sponge-associated actinomycetes and the efficacy of genome mining in exploring the genomic potential of sponge-derived actinomycetes.

Actinomycetes are known for their unprecedented ability to produce novel lead compounds of clinical and pharmaceutical importance (1–4). Among the many actinobacterial genera, Streptomyces, Micromonospora, Nocardiopsis, and Rhodococcus are the most prolific producers of secondary metabolites, which display broad chemical diversity and diverse pharmacologically and medically relevant bioactivities (5–8). Recent genomic sequencing data have revealed the presence of a plethora of putative biosynthetic gene clusters on the genomes of actinomycetes encoding for secondary metabolites that are not observed under standard fermentation conditions (9–13). In the present study, draft genomes of three actinobacterial isolates, Micromonospora sp. RV43, Rubrobacter sp. RV113, and Nocardiopsis sp. RV163 that had previously been cultivated from the Mediterranean sponges Aplysina aerophoba (RV43 and RV113) and Dysidea avara (RV163) (14), were established.

The genomic DNA of the isolates was extracted from 5-day-old ISP2 cultures. Paired-end, 2 × 250-bp libraries were prepared with the Nextera XT kit (Illumina, Inc.). Sequencing was performed on an Illumina MiSeq device. A total of 5,900,702 raw reads were produced for Micromonospora sp. RV43, 2,206,732 raw reads for Rubrobacter sp. RV1113 and 4,851,980 raw reads for Nocardiopsis sp. RV163. Reads were adapter clipped, quality trimmed and length filtered (15). Initial contigs were generated using SPAdes (16) and only contigs ≥1000 bp were maintained. A further clean-up of contigs was performed using G+C content, coverage, and taxonomic assignments (17). For ab initio gene prediction, prodigal was applied (18) and functional annotation of the predicted protein sequences was performed with the RAST webserver (19). Secondary metabolite gene clusters and possible encoded compounds were predicted with antiSMASH (20) and NapDos (21).

A number of 101 (RV43), 33 (RV1113), and 82 (RV163) secondary metabolite gene clusters were detected with antiSMASH. For strain RV43, 5 terpene clusters, 4 type 1 PKS clusters, 2 lantipeptides, 1 type 2 PKS cluster, 1 siderophore, 1 NRPS cluster, and 1 bacteriocin were found. For strain RV113, 3 terpene clusters, 1 fatty acid, and 1 mixed type 3 PKS-fatty acid cluster were found. The draft genome sequence of strain RV163 showed homologies to 7 NRPS clusters, 4 terpene gene clusters, 2 type 1 PKS clusters, 2 ectoines, 2 bacteriocins, 1 phenazine, 1 butyrolactone, 1 type 2 PKS, and 1 siderophore.

For Micromonospora sp. RV43, NaPDos predicted the presence of gene clusters encoding for compounds such as leinamycin, kirromycin, aclacinomycin, and tetronomycin. For Nocardiopsis sp. RV163, compounds such as alnumycin, avermectin, and neo-carzinostatin were predicted. For Rubrobacter sp. RV113, only gene clusters encoding for fatty acids synthesis were found. These results highlight the genomic potential of at least two of three inspected isolates for natural products discovery.

Nucleotide sequence accession numbers. This whole-genome shotgun project was deposited in DDBJ/ENA/GenBank under the accession numbers LEKG00000000, LEKH00000000, and LEKI00000000. The versions described in this paper are the first versions LEKG01000000, LEKH01000000, and LEKI01000000.

ACKNOWLEDGMENTS
We gratefully acknowledge C. Gernert for technical assistance in the laboratory. We thank the Group of Molecular Biodiversity/Zoology 3 for library preparation and the Department of Human Biology for access to the sequencing device (both University of Würzburg).

Financial support was provided by the DFG (SFB 630 TP A5) to U.H.

REFERENCES

Address correspondence to Usama Ramadan Abdelmohsen, usama.ramadan@uni-wuerzburg.de.


