



Your One-Stop Platform for Information on Marine Natural Products, Their Gene Clusters and Producing Organisms

Cedric Hermans*, Maarten L. De Mol, Marieke Mispelaere, Anne-Sofie De Rop, Jeltien Rombaut, Tesneem Nusayr, Rebecca Creamer, Sofie L. De Maeseneire, Wim K. Soetaert and Paco Hulpiau

* Presenter, Bioinformatics Knowledge Center (BiKC), Campus Brugge Station, Howest University of Applied Sciences, Belgium

Abstract

Marine environments harbor a vast diversity of **microorganisms** that produce a variety of **natural products** with diverse chemical structures and biological activities. These natural products are valuable resources for Belgium and Europe, as they have potential applications in **biotechnology, medicine, agriculture and other fields**. However, identifying and characterizing these compounds and their gene clusters (BGCs) is a challenging task that requires advanced bioinformatic tools and skills.

To address this challenge, we developed **MariClus**, an online platform that integrates the results of **antiSMASH 7**, the most widely used open-source tool for BGC prediction and analysis, for over 500 manually curated high-quality genome assemblies of marine prokaryotes. MariClus also provides additional information

from **public databases** such as NCBI, MIBiG and PubChem, as well as relevant publications for each species, cluster and molecule. MariClus allows users to easily search, filter and visualize the data on the **web portal**, as well as to perform meta-analyses on the taxonomic distribution, cluster types and molecule diversity of the marine prokaryotes in the database. Furthermore, MariClus enables several case studies related to **microbiology, gene cluster identification and natural product discovery**, such as finding marine bacteria that degrade oil in cold environments, comparing gene clusters of abyssomicin and related compounds, and linking predicted molecules to their bioactivities and patents. MariClus is **freely available** at <https://www.mariclus.eu> and aims to provide a user-friendly and comprehensive resource for marine natural product research.

Methods and results

Data collection:

Over 500 marine prokaryote genome assemblies were collected and curated from NCBI Genome datasets and information on their taxonomy, isolation source, publications and culture collection number were extracted from various sources.

Genome mining:

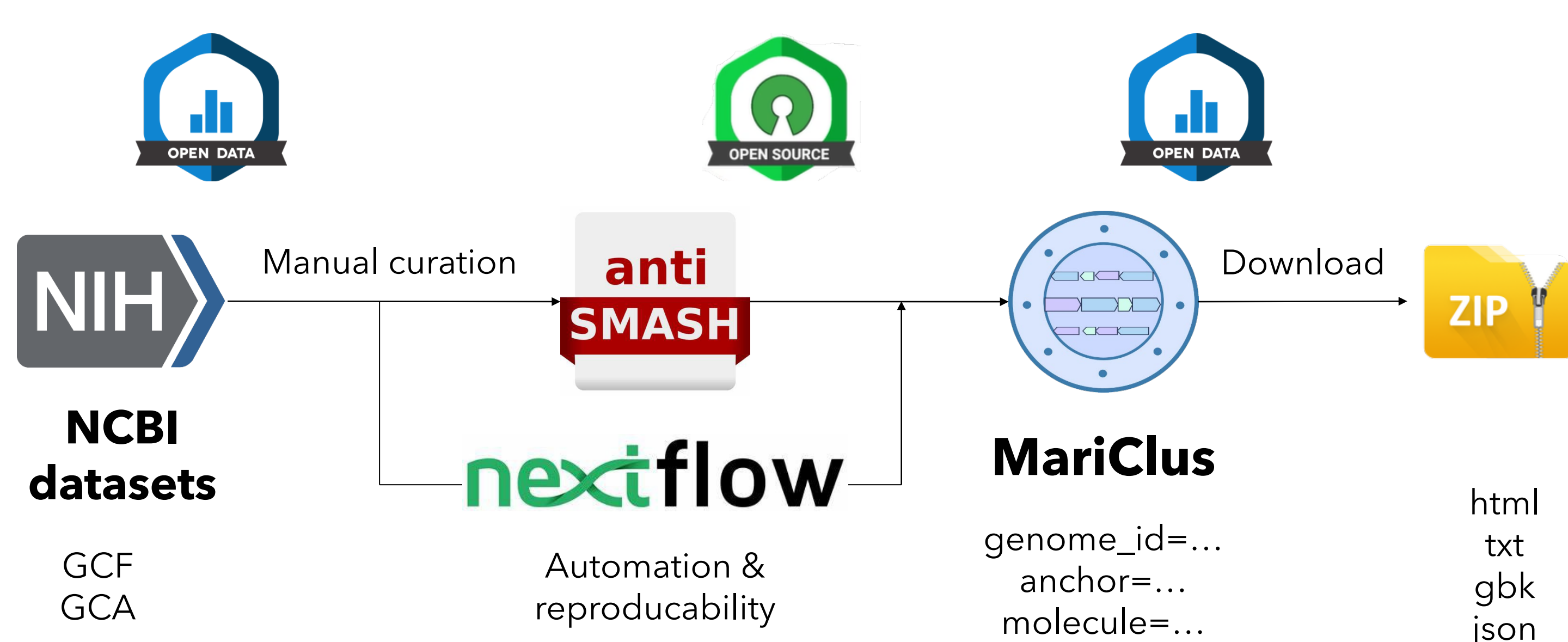
The genome assemblies were analyzed using antiSMASH 7 to identify biosynthetic gene clusters and predict their products. Additional features from the antiSMASH results, such as cluster type, contig, length, number of genes and cumulative BLAST score, were also extracted.

Database integration:

All the data were integrated into a MariaDB database and a web interface was developed to display and query the data on marine species, gene clusters and molecules. Links to external resources, such as NCBI, PubMed, PubChem, MIBiG and antiSMASH, were also provided.

Open science and FAIR:

MariClus uses open-source tools and public databases. Our platform allows for easy downloading of the antiSMASH results and sharing of specific clusters through unique URLs.



Examples

Examples of the search results for salinosporamide

Salinosporamide A (Marizomib) is a potent proteasome inhibitor being studied as a **potential anticancer agent**. The MariClus platform returns information about species, gene clusters, molecules and the titles of possible relevant articles. It links directly to NCBI and PubMed references. Each result also contains a link to the details page for the linked species containing more information

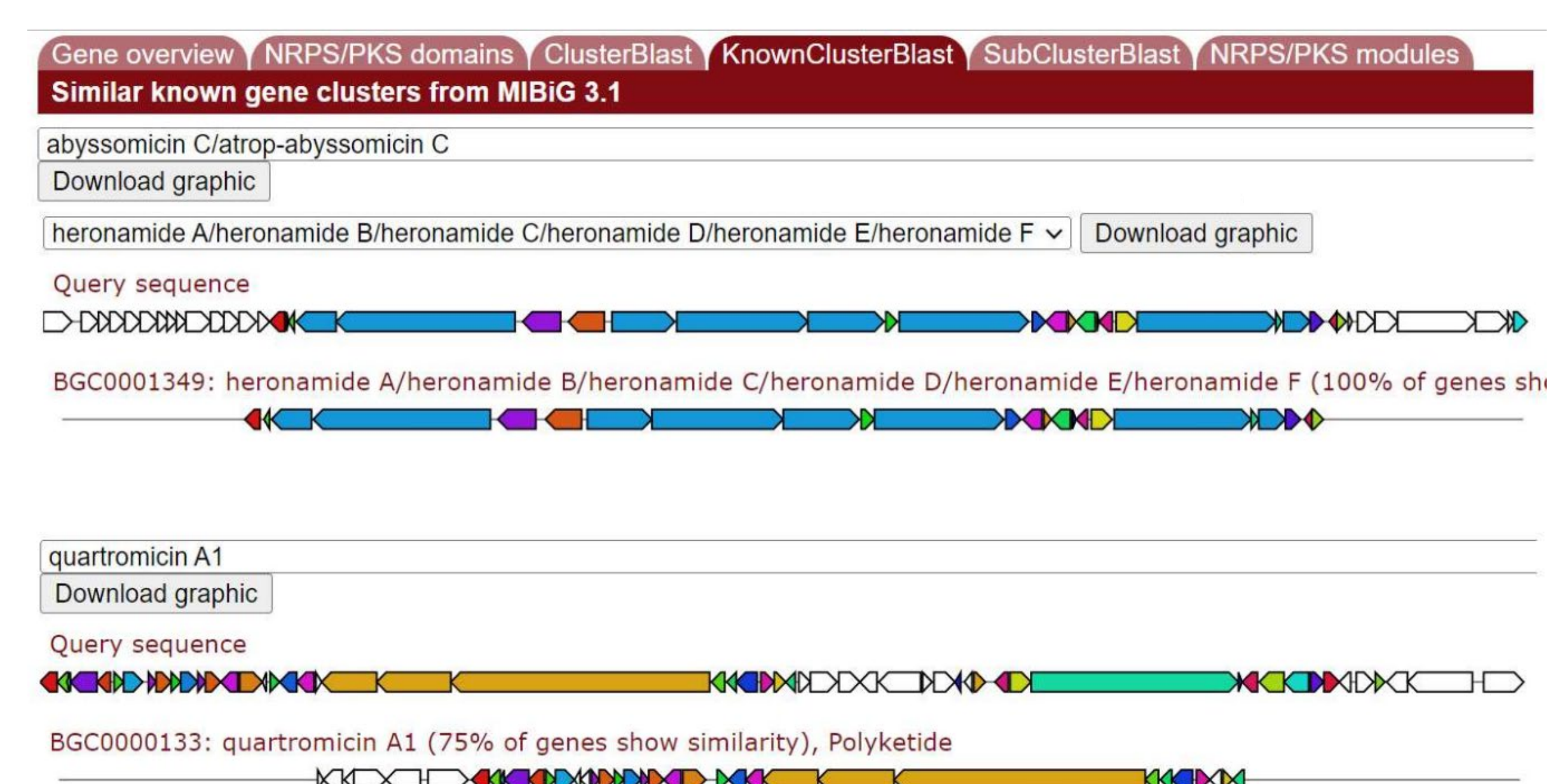
Species	Description	Accession number	Pubmed ID
Salinispora tropica	This species produces the chemicals salinosporamide A and B which may be useful for cancer treatment.	GCF_000016425.1	17563368, 19590008

Species	Contig ID	Type	Predicted product
Salinispora fenicalii	NZ_AZWQ01000016.1	transAT-PKS-like	salinosporamide A
Salinispora pacifica	NZ_KB905359.1	NRPS-like	salinosporamide A
Salinispora tropica	NC_009380.1	transAT-PKS-like	salinosporamide A

Species	PMID	Title
Salinispora tropica	19590008	Biosynthesis of the salinosporamide A polyketide synthase substrate chloroethylmalonyl-coenzyme A from S-adenosyl-L-methionine

Examples of the antiSMASH results

KnownClusterBlast results showing the overlap between detected clusters and known clusters from the MIBiG database



Summary

MariClus is a novel and user-friendly resource that facilitates **marine natural product discovery** by integrating **genome mining** results from antiSMASH 7 with data from public databases and publications for over 500 marine prokaryote species. MariClus provides comprehensive and transparent information on the **genome assemblies, gene clusters, predicted molecules and relevant publications** for each species. Users can easily explore and

compare the data on the web portal, as well as access the antiSMASH pages for each cluster and molecule. MariClus also enables meta-analyses and case studies on the **marine biodiversity, microbial ecology, biosynthetic pathways, enzyme characterization, metabolic engineering and drug development**. MariClus is a valuable tool for researchers interested in marine natural products and their source organisms and gene clusters.

Funding: This research was funded by the Research Foundation Flanders (FWO), grant number S001422N, and by a PWO grant with acronym MetaTec, provided by Howest University College and the Flemish Government.

More information: <https://www.mariclus.eu>

cedric.hermans@howest.be

 Paper: <https://doi.org/10.3390/md21080449>

