

## First public resource for secondary metabolites searches

## August 6 2015



A chart of ecosystem sources for the 10 most common biosynthetic cluster enzymatic types. Generating these data in the IMG-ABC database is made possible by the integration of the portal with the IMG and GOLD databases, which allow access to metadata annotations. Credit: Hadjithomas M et al.

The wealth of genomic and metagenomic datasets for microbes, particularly from previously unstudied environments, within the Integrated Microbial Genomes (IMG) system is being applied in a new public database to the search for novel secondary metabolites that could be used in a wide range of applications from bioenergy to health.

This public database allows researchers to more efficiently harness the genomic data generated by advanced sequencing technologies to identify



novel small molecules relevant to DOE missions in bioenergy and environment.

Secondary metabolites are organic compounds that aren't involved in functions such as growth or reproduction but which may still play key roles in regulation, immunity, or other vital pathways in an organism. Many secondary metabolites from plants and microbes also have been shown to have commercial applications; for example, terpenes from a variety of plants including eucalyptus have bioenergy applications.

In a study published July 14, 2015 in mBio, researchers from the DOE Joint Genome Institute, a DOE Office of Science User Facility, the Biosciences Computing Group at Lawrence Berkeley National Laboratory, and the University of California, San Francisco, describe a new <u>public database</u> that allows researchers to search for biosynthetic gene clusters and secondary metabolites. The database is known as IMG-ABC because it is an Atlas of Biosynthetic gene Clusters within the Integrated Microbial Genomes (IMG) system and includes information from single-cell and metagenome data from the DOE JGI's Microbial Dark Matter project.

"For the first time," the team reported, "IMG-ABC links information regarding genomic pathways for the biosynthesis of <u>secondary</u> <u>metabolites</u> with chemical structure information on a scale of several thousand data sets. With careful efforts for quality control, it combines the predictive power of state-of-the-art computational tools... with the exhaustive analysis framework offered by the IMG family of systems. This combination delivers a powerful punch, predicting both familiar and novel biosynthetic gene pathways in thousands of cultured isolates, single cells, and metagenomes."

The team conducted a pilot "proof-of-principle" study in which they searched through more than 25,000 isolate <u>microbial genomes</u> for



biosynthetic clusters that contain at least six of the seven core genes required to produce phenazine, a compound that acts as a shuttle for electrons in metabolism, is found in bacteria that can process nutrients from wastes, and from which dyes and antibiotics are also derived. The search yielded nearly 1,000 hits and identified potentially novel phenazine pathways in Gammaproteobacteria, Betaproteobacteria, and Actinobacteria, and even, unexpectedly, in a root-nodulating bacterium.

Processing vast amounts of genomic and metagenomic data requires a high performance computing environment, and in this case, the DOE JGI relies on high-efficiency resources from the National Energy Research Scientific Computing Center (NERSC) to support the demand for computational access.

DOE JGI Prokaryote Super Program head Nikos Kyrpides alluded to the development of IMG-ABC when he discussed future plans for the continuing improvement of the IMG database on its 10th anniversary. "There's a huge amount of functionality in IMG already, but we certainly need to continue adding more," he said. "The two main directions in the near future include adding more functionality and efficient supporting data/size growth."

**More information:** "IMG-ABC: A Knowledge Base To Fuel Discovery of Biosynthetic Gene Clusters and Novel Secondary Metabolites." *mBIO* 6:4 e00932-15; Published 14 July 2015, DOI: <u>10.1128/mBio.00932-15</u>

## Provided by DOE/Joint Genome Institute

Citation: First public resource for secondary metabolites searches (2015, August 6) retrieved 28 April 2024 from <u>https://phys.org/news/2015-08-resource-secondary-metabolites.html</u>



This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.