

Description of Additional Supplementary Files

File Name: Supplementary Data 1

Description: Complete catalog of biosynthetic domains (a), enzymatic tailoring reactions (b), substrates (c), and sugars (d) comprising the PRISM 4 chemical structure prediction engine.

File Name: Supplementary Data 2

Description: Structures of known products and predicted chemical structures, and Tanimoto coefficients between them, for the 1,281 characterized BGCs in the 'gold standard' set.

File Name: Supplementary Data 3

Description: Predicted chemical structures from PRISM 4 and antiSMASH 5 for BGCs identified in 3,759 complete bacterial genomes (a) and 6,362 metagenome-associated genomes (b).