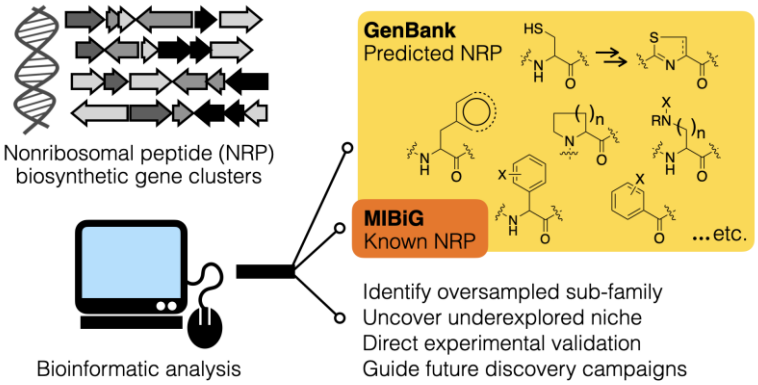
Bioinformatic analysis reveals both oversampled and underexplored biosynthetic diversity in nonribosomal peptides

John Chu*
Department of Chemistry, National Taiwan University, Taipei City 10617, Taiwan
E-mail: johnchu@ntu.edu.tw

The traditional natural product discovery approach has accessed only a fraction of the chemical diversity in nature. The use of bioinformatic tools to interpret the instructions encoded in microbial biosynthetic genes has the potential to circumvent existing methodological bottlenecks and greatly expand the scope of discovery. Structural prediction algorithms for nonribosomal peptides (NRP), the largest family of microbial natural products, lie at the heart of this new approach. To understand the scope and limitation of the existing prediction algorithms, we evaluated their performances on NRP synthetase biosynthetic gene clusters. Our systematic analysis shows that the NRP biosynthetic landscape is uneven. Phenylglycine and its derivatives as a group of NRP building blocks (BB), for example, have been over sampled, reflecting an extensive historical interest in the glycopeptide antibiotics family. In contrast, the benzoyl BB, including 2,3-dihydroxybenzoate (DHB), have been the most underexplored, hinting at the possibility of a reservoir of as yet unknown DHB containing NRP with functional roles other than siderophore. Our results also suggest that there is still vast unexplored biosynthetic diversity in nature, and the analysis presented herein shall help guide and strategize future natural product discovery campaigns. We also discuss possible ways bioinformaticians and biochemists could work together to improve existing prediction algorithms.



References

[1] Jian, B.-S.; Chiou, S.-L.; Hsu, C.-C.; Ho, J.; Wu, Y.-W.; Chu, J. *ACS. Chem. Biol.* **2023**, *18*, 476-483.
[2] Chu, J.; Koirala, B.; Forelli, N.; Vila-Farres, X.; Ternei, M. A.; Ali, T.; Colosimo, D. A.; Brady, S. F. *J. Am. Chem. Soc.* **2020**, *142*, 14158-14168.
[3] Chu, J.; Vila-Farres, X.; Brady, S. F. *J. Am. Chem. Soc.* **2019**, *141*, 15737-15741.
[4] Chu, J.; Vila-Farres, X.; Inoyama, D.; Ternei, M. A.; Cohen, L. J.; Gordon, E. A.; Reddy, B. V. B.; Charlop-Powers, Z.; Zebroski, H. A.; Gallardo-Macias, R.; Jaskowski, M.; Satish, S.; Park S.; Perlin, D. S.; Freudlich, J. S.; Brady, S. F. *Nat. Chem. Biol.* **2016**, *12*, 1004-1006.

John Chu. National Taiwan University (BS, 2004), The Scripps Research Institute (Ph.D., 2010, M. Reza Ghadiri), Yale University (Postdoc, 2011-2014, Alanna Schepartz), Rockefeller University (Postdoc, 2014-2018, Sean F. Brady), National Taiwan University (Assistant Professor, 2019-present). While natural products have and continue to inspire the design of new drugs, we see them more broadly as part of the chemical language of nature, whose function could help us better understand how living organisms communicate with each other and interact with the environment. My research group takes unconventional approaches towards natural product discovery and the study of their bio- and chemical syntheses. [Field of research] Chemical biology; natural products; biosynthetic gene clusters