microLife, 2024, **5**, uqae023
DOI: 10.1093/femsml/uqae023
Advance access publication date: 19 November 2024
Short Review

Accessing microbial natural products of the past

Martin Klapper[1] and Pierre Stallforth[1,2,3,*]

[1]Leibniz Institute for Natural Product Research and Infection Biology—Hans Knöll Institute, Department of Paleobiotechnology, Beutenbergstraße 11a, D-07745 Jena, Germany
[2]Friedrich Schiller University Jena, Institute for Organic Chemistry and Macromolecular Chemistry, Humboldtstraße 10, D-07743 Jena, Germany
[3]Cluster of Excellence Balance of the Microverse, Friedrich Schiller University Jena, Fürstengraben 1, D-07743 Jena, Germany
*Corresponding author. Leibniz Institute for Natural Product Research and Infection Biology—Hans Knöll Institute, Department of Paleobiotechnology, Beutenbergstraße 11a, D-07745 Jena, Germany. E-mail: pierre.stallforth@leibniz-hki.de
Editor: [Axel Brakhage]

Abstract

Microbial natural products—low molecular weight compounds biosynthesized by microorganisms—form the foundation of important modern therapeutics, including antibiotics, immunomodulators, and anti-cancer agents. This perspective discusses and contrasts two emerging approaches for uncovering natural products of the past. On the one hand, ancestral sequence reconstruction allows recreating biosynthetic pathways that date back hundreds of millions of years. On the other hand, sequencing and *de novo* assembly of ancient DNA reveals the biosynthetic potential of ancient microbial communities up to 100 000 years. Together, these approaches unveil an otherwise hidden reservoir of functional and structural molecular diversity. They also offer new opportunities to study the biological function and evolution of these molecules within an archaeological context.

Keywords: ancient natural products; paleobiotechnology; phylogenetic analysis; metagenomics; heterologous expression; bacteria

Main text

Microbial natural products—low molecular weight compounds biosynthesized by microorganisms—display a vast structural and functional diversity. Often referred to as secondary or specialized metabolites, these natural products, while not always necessary for immediate survival, confer significant evolutionary advantages to their producers (Fischbach et al. 2008). They play critical physiological and ecological roles, from acting as inter- and intracellular signaling molecules to serving as defensive agents against potential competitors or predators. Many have been used as potent therapeutic agents; for instance, a vast majority of antibiotics are derived from microbial natural products (Newman and Cragg 2020). The quest for novel microbial natural products reached its zenith during the golden era of antibiotic discovery in the 1950s and 1960s, spearheaded by the exploration of natural products from taxonomically diverse microorganisms across different habitats.

These compounds are the result of dedicated biosynthetic machinery encoded within the producer's genome. The corresponding genes are often grouped into biosynthetic gene clusters (BGCs). The advent of next-generation sequencing (NGS) has resulted in an explosion of genomic data. This data can come from individual bacteria or from bacterial communities resulting in so-called metagenomic data (Paoli et al. 2022). Genome mining then enables the *in silico* identification of BGCs within these (meta)genomes. Tools such as antiSMASH are now commonly used for genome mining (Blin et al. 2023). In recent years, approaches to identify novel natural products have relied on savvy combinations of isolation techniques and genome mining to access novel natural products (Hemmerling and Piel 2022).

The nascent field of paleobiotechnology introduces a novel dimension to the discovery of microbial natural products: time. This perspective presents two recent breakthroughs in the exploration of ancestral or prehistoric natural products. These two distinct yet complementary strategies for accessing these age-old microbial natural products rely on an *in silico* reconstruction of ancestral BGCs and on the resurrection of BGCs from ancient DNA (aDNA).

Ancestral sequence reconstruction has been successfully pursued for almost two decades to access and functionally characterize ancient biomolecules (Thornton 2004). This strategy, for instance, was employed to obtain an ancestral dinosaur pigment (Chang et al. 2002), alcohol dehydrogenase from yeast (Thomson et al. 2005), a steroid receptor (Thornton et al. 2003), an elongation factor of the Tu family (Gaucher et al. 2003), and precambrian thioredoxin enzymes (Perez-Jimenez et al. 2011), amongst others. In a study published in 2023, this strategy was applied for a large BGC encoding a nonribosomal peptide synthetase (NRPS). These mega enzymes assemble peptidic natural products from amino acid building blocks. The aim was to obtain potentially bioactive natural products from the past, focusing on glycopeptide antibiotics (GPAs), which are important drugs of last resort (Hansen et al. 2023).

Ancestral sequence reconstruction relies on a substantial amount of high-quality database entries of related gene sequences to infer reliable phylogenies. While it is virtually impossible to directly show that the reconstructed sequences have actually existed in their predicted form, the advantage of ancestral sequence reconstruction lies in the large time span that can be covered (Thornton 2004). In the study, a putative paleomycin synthetase was predicted to have existed in the Paleozoic era

Received 7 September 2024; **revised** 9 October 2024; **accepted** 18 November 2024

(550–250 mya) (Hansen et al. 2023). In-depth phylogenetic analysis allowed retracing distinct evolutionary events that led to different GPA classes. From various related gene clusters, a phylogeny was constructed leading to an ancestral NRPS gene that would produce the hypothetical parental molecule termed paleomycin. Based on current database entries, it was shown that gene deletion and recombination progressively turned paleomycin into an extant natural product. To validate the prediction, the ancestral paleomycin NRPS gene was synthesized and heterologously expressed in an engineered host strain that would allow for various late-stage modifications. The latter were required for converting the precursor peptide into a bioactive GPA. While the core scaffold is assembled by the ancestral NRPS, the modifications were introduced by modern tailoring enzymes—homologs of which were predicted to be encoded in the ancestral BGC. The production of different paleomycin congeners was confirmed by liquid chromatography coupled with tandem mass spectrometry (LC–MS/MS) and molecular networking. Moreover, culture extracts of the heterologous host producing paleomycin showed the expected activity against *Bacillus subtilis*. Additionally, an evolutionary analysis of the adenylation domains (A domain), which are responsible for installing the amino acid building blocks in the NRP biosynthesis was conducted. The A domain substrate specificity of modern and ancestral GPA NRPS A domains was assessed, and structural information gained by X-ray crystallography. This allowed to demonstrate that the A domain evolution proceeded through point mutations leading to a shift in substrate specificity. Overall, evolutionary information for this distinct NRPS system could be obtained, which helps our general understanding of how NRPS BGCs evolve.

A complementary strategy, also published in 2023, relied on the reconstruction of microbial natural products from aDNA (Klapper et al. 2023). The genetic data stem from well-preserved sources of bacterial DNA, such as dental calculus (Fellows Yates et al. 2021). This calcified dental plaque contains exceptionally well preserved DNA of the bacterial community residing in the oral cavity (Warinner et al. 2014). Extensive archeological excavations have led to an impressive amount of samples that can now be found in the community curated database AncientMetagenomeDir (Fellows Yates et al. 2021). In order to obtain sequence information from ancient calculus, the entrapped DNA was extracted and sequenced using NGS. The study used dental calculus of 12 Neanderthals and 52 anatomically modern humans ranging from 100 000 years ago to the present (Klapper et al. 2023). In comparison to modern DNA, aDNA isolated from dental calculus is much shorter in length, which renders subsequent genome assembly difficult (Dalén et al. 2023). Increasing sequence depth and optimization of the assembly strategy gave access to long contigs that were binned into 459 bacterial metagenome-assembled genomes (MAGs). The ancientness of sequences was assessed using PyDamage, based on the quantification of age-related DNA damage (Borry et al. 2021). Notably, the MAG with the highest degree of completeness could be attributed to bacteria of the genus *Chlorobium*. Although these bacteria are not typically associated with the oral microbiome, they displayed the same age-related damage profile as *bona-fide* oral taxa from the same samples. Further analyses confirmed their presence as a genuine member of the paleolithic Neanderthal oral microbiome.

The MAGs and unbinned contigs were subjected to *in silico* analysis in order to identify biosynthetic genes. Specifically, antiSMASH (Blin et al. 2023) was used to identify an unprecedented BGC that is shared by seven Middle and Upper Paleolithic individuals. This analysis then enabled the heterologous production of a class of previously unknown metabolites that were named "paleofurans." Structure elucidation based on high-resolution MS and nuclear magnetic resonance spectroscopy as well as chemical synthesis allowed to determine the structure of the novel paleofurans.

In contrast to ancestral sequence reconstruction, the aDNA-based search is not limited to a specific BGC family that requires a reliable set of database entries. This allows assessing the comprehensive biosynthetic potential of individual ancient microorganisms. However, not all microbiomes preserve equally well. Dental calculus, permafrost, or in some cases paleofeces or coprolites allow for DNA to preserve well over time but for many microbiomes access to ancient DNA is not possible. In general, aDNA older than a few million of years would be difficult to sequence and to assemble due to the continuing degradation into very short DNA fragments.

Due to the highly fragmented nature of aDNA, it is difficult to obtain long sequences that would encode more complex biosynthetic machinery, such as NRPSs in their full length. Evaluation of the completeness of a potential BGC thus becomes a central part for functional downstream analyses. Complete BGCs are required for subsequent gene synthesis and construction of expression vectors or integration in the genomes of "modern" bacteria. Upon induction, these biosynthetic genes are then translated into enzymes that allow for the generation of natural products. Experimental validation of a variety of reconstructed BGCs—both, from ancestral sequence reconstruction or aDNA *de novo* assembly—will become more feasible with developments and decreasing costs in gene synthesis. Furthermore, automated approaches using well-established expression platforms may help to fasten the discovery process of ancient molecules in the future.

In summary, both approaches as well as a combination thereof will allow us to better understand how natural products have evolved over time, and to infer the ecological roles of ancient molecules. Importantly, paleobiotechnology will allow access to natural products of bacteria that may have gone extinct. It will be exciting to investigate the biological functions of many more of these ancient natural products once the throughput of these approaches has been increased.

Conflict of interest: None declared.

Funding

This research was supported by the Werner Siemens Foundation, the Leibniz Association, and the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy EXC 2051 (Project-ID 390713860, "Balance of the Microverse").

References

Blin K, Shaw S, Augustijn HE *et al.* antiSMASH 7.0: new and improved predictions for detection, regulation, chemical structures and visualisation. *Nucleic Acids Res* 2023;**51**:W46–W50. https://doi.org/10.1093/nar/gkad344.

Borry M, Hübner A, Rohrlach AB *et al.* PyDamage: automated ancient damage identification and estimation for contigs in ancient DNA de novo assembly. *PeerJ* 2021;**9**:e11845.

Chang BSW, Jönsson K, Kazmi MA *et al.* Recreating a functional ancestral archosaur visual pigment. *Mol Biol Evol* 2002;**19**:1483–9.

Dalén L, Heintzman PD, Kapp JD *et al.* Deep-time paleogenomics and the limits of DNA survival. *Science* 2023;**382**:48–53.

Fellows Yates JA, Andrades Valtueña A, Vågene ÅJ *et al.* Community-curated and standardised metadata of published ancient metagenomic samples with AncientMetagenomeDir. *Sci Data* 2021;**8**:31.

Fellows Yates JA, Velsko IM, Aron F *et al.* The evolution and changing ecology of the African hominid oral microbiome. *Proc Natl Acad Sci USA* 2021;**118**:e2021655118.

Fischbach MA, Walsh CT, Clardy J. The evolution of gene collectives: how natural selection drives chemical innovation. *Proc Natl Acad Sci* 2008;**105**:4601–8.

Gaucher EA, Thomson JM, Burgan MF *et al.* Inferring the palaeoenvironment of ancient bacteria on the basis of resurrected proteins. *Nature* 2003;**425**:285–8.

Hansen MH, Adamek M, Iftime D *et al.* Resurrecting ancestral antibiotics: unveiling the origins of modern lipid II targeting glycopeptides. *Nat Commun* 2023;**14**:7842.

Hemmerling F, Piel J. Strategies to access biosynthetic novelty in bacterial genomes for drug discovery. *Nat Rev Drug Discov* 2022;**21**:359–78.

Klapper M, Hübner A, Ibrahim A *et al.* Natural products from reconstructed bacterial genomes of the Middle and Upper paleolithic. *Science* 2023;**380**:619–24.

Newman DJ, Cragg GM. Natural products as sources of new drugs over the nearly four decades from 01/1981 to 09/2019. *J Nat Prod* 2020;**83**:770–803.

Paoli L, Ruscheweyh H-J, Forneris CC *et al.* Biosynthetic potential of the global ocean microbiome. *Nature* 2022;**607**:111–8.

Perez-Jimenez R, Inglés-Prieto A, Zhao Z-M *et al.* Single-molecule paleoenzymology probes the chemistry of resurrected enzymes. *Nat Struct Mol Biol* 2011;**18**:592–6.

Thomson JM, Gaucher EA, Burgan MF *et al.* Resurrecting ancestral alcohol dehydrogenases from yeast. *Nat Genet* 2005;**37**:630–5.

Thornton JW, Need E, Crews D. Resurrecting the ancestral steroid receptor: ancient origin of estrogen signaling. *Science* 2003;**301**:1714–7.

Thornton JW. Resurrecting ancient genes: experimental analysis of extinct molecules. *Nat Rev Genet* 2004;**5**:366–75.

Warinner C, Rodrigues JFM, Vyas R *et al.* Pathogens and host immunity in the ancient human oral cavity. *Nat Genet* 2014;**46**:336–44.

Received 7 September 2024; **revised** 9 October 2024; **accepted** 18 November 2024