WEP Warwick Evans Publishing

International Journal of Public Health and Medical Research

ISSN: 3006-7707 (Print), ISSN: 3006-3752 (Online) | Volume 6, Number 6, Year 2026
DOI: https://doi.org/10.62051/ijphmr.v6n6.09
Journal homepage: https://wepub.org/index.php/IJPHMR/index

Mining of Antibacterial Related BGC and Analysis of Antibacterial Mechanism At Protein Level

Haohong Tu

Faculty of Medicine, Chinese university of Hong Kong, Hong Kong, China

ABSTRACT

Antibacterial biosynthetic gene clusters (BGCs) are rich in new natural products, and therefore, genome mining has been used to discover them in bacteria and fungi. This paper investigates the following two problems: how to extract antibacterial BGCs efficiently, and how to understand the antibacterial mechanisms of these genes at the protein level. Based on the above analysis, a single set of discovery methods is not suitable for identifying important genes and proteins. Antibacterial activity at the protein level is generally mediated by enzymes, precursor peptides, transporters, tailoring proteins, immunity proteins and regulatory factors that regulate biosynthesis and modes of action. Based on recent studies of antiSMASH-based mining, BGC classification, proteomining and fungal and bacterial cluster characterization, a practical workflow for discovery and mechanistic interpretation has been established in this paper. Therefore, BGC mining should also investigate the origin of antibacterial functions at the protein level via biosynthetic systems, not only cluster prediction.

KEYWORDS

Biosynthetic gene cluster; Genome mining; Antibacterial natural products; Proteomics; Protein mechanism; Antismash

1. INTRODUCTION

Most of the new antibacterial compounds now start with a search in the genomes of microorganisms. Due to the emergence of antibiotic resistance and the identification of numerous silent or cryptic biosynthetic gene clusters (BGCs) encoding valuable natural products in many microorganisms [1, 2], there has been a change. Rather than randomly screening for compounds, now researchers are mining genomes to predict the biosynthetic potential of a substance and then activating or validating the most promising gene clusters.

Genome mining has been used to find new antibacterial drugs, and many of the BGCs produced by these drugs are involved in inhibiting the growth of microorganisms. These clusters may produce nonribosomal peptides, polyketides, ribosomally synthesized and post-translationally modified peptides (RiPPs), terpenes, or hybrid products with antimicrobial activity [3-4]. However, prediction is not all there is. In order to know why a cluster is antibacterial, one also needs to study the proteins it encodes and the biochemical logic of the pathway.

Therefore, the two related problems in this paper are the mining of antibacterial-related BGCs and the analysis of antibacterial mechanisms at the protein level. First, the discussion will present mining strategies and tools; next, it will explain how proteomics and protein-level analysis are used to reveal biosynthetic functions and antibacterial mechanisms; finally, some practical applications of discovery pipelines will be introduced.

2. GENOME MINING

2.1. Mining Strategies

The beginning of BGC mining is a whole-genome sequence and then automatic annotation. Tools such as antiSMASH and its corresponding database can be used to identify known cluster classes, compare them across genomes, and determine whether candidate regions are novel [5-6]. Previous studies on methodology have also pointed out that to avoid both false positives and missing clusters, an accurate definition of the boundary, family classification and a reference comparison are required [1].

Many mining modes can be employed. Signature-gene searches are suitable for well-studied groups such as NRPS, PKS and bacteriocins; comparative genomics can be used to identify lineage-specific clusters; and resistance-guided mining can find clusters associated with self-protection or target modification [7-8]. At the same time, rule-based detection and prioritization based on biology, ecology and resistance logic are also used in practice.

2.2. Prioritization of Antibacterial Clusters

Not all BGCs are suitable for antibacterial discovery. Clusters should be given priority if they are (i) taxonomically unusual, (ii) associated with resistance or immunity genes, (iii) highly expressed under relevant conditions, or (iv) structurally related to known antibacterial scaffolds [3, 7]. Antibacterial compounds in microbial competition often need to have their own self-immunity proteins and transport systems, which can be used as important clues for priority.

Recent genome-mining studies have shown that many bacteria and gut strains possess a large number of candidate secondary metabolite clusters, including NRPS, bacteriocin, arylpolyene and terpene classes [9-10]. Therefore, the search space is still relatively large. Therefore, the previous sorting will help to focus the research funds on the better antibacterial-gene clusters.

2.3. Silent Clusters and Induction

Most antibacterial BGCs are not expressed in a standard laboratory environment, and therefore difficult to study. One way to deal with this is to use co-culture, pathway-specific regulatory modification or environmental induction to induce expression [11]. Turn on a silent cluster, and then conduct comparative studies of transcriptomics, metabolomics, and proteomics to link the gene cluster with its products.

At this time, discovery shifts from being a prediction to being evidence. A cluster predicted by bioinformatics is more likely to be real if its proteins are expressed, its product is identified, and an association can be established between the antibacterial effect and the biosynthetic pathway.

3. PROTEIN-LEVEL ANALYSIS

3.1. Biosynthetic Enzymes

Protein-level analysis shows how a BGC is realised at the protein level. The core biosynthetic enzymes are adenylation domains, thiolation domains, condensation domains, ketosynthases, methyltransferases, oxidoreductases and glycosyltransferases; they determine the chemical skeleton and functional groups of the final product. Many antibacterial pathways regulate the choice of substrate and chain assembly and post-translational modification by these enzymes [3,4].

The antibacterial mechanism often begins with biosynthesis rather than target action. Tailored enzymes are employed to introduce halogens and other unusual oxidation states to enhance the

membrane-disrupting ability, binding affinity at the target site, or product molecule stability of the product. Therefore, we can determine the reason for the antibacterial effect of this metabolite at the protein level.

3.2. Immunity and Transport Proteins

Many antibacterial BGCs also have immunity genes to protect themselves from self-poisoning. These proteins may bind the compound, alter the target, or export the metabolite from the cell [7, 12]. Transporters have dual functions and help with secretion, but they may also be associated with resistance. Often, their presence can be traced back to an ecologically active antibacterial product encoded by BGC.

Immunity proteins are relatively more informative and can provide some hints about the target class of the compound. If a cluster codes for a protein that is similar to a known target-protection system, one can determine whether the antibacterial compound targets the ribosome, cell wall synthesis, membrane integrity, DNA metabolism, etc. Thus, protein analysis can reveal the mechanism earlier in the research process before full biochemical characterization.

3.3. Regulatory Proteins

Regulatory proteins are needed to decide when a BGC should be expressed by controlling transcription factors and other activators. Studies of transcriptional architecture have recently identified several BGCs that are regulated by clustered regulators in response to changes in the environment [13]. Antibacterial production is often induced by the environment and circumstances. Without knowledge of regulations, a highly potent BGC may be wrongly determined to be inactive because it does not show activity in the lab.

Regulatory proteins thus make up the mechanism. They connect the environmental signal with biosynthetic products, so protein-level analysis should also consider structural genes and regulatory elements. Regulation in antibacterial systems is often referred to as ecological competition; thus, the cluster is expressed when microbes detect competitors or are under stress.

4. MECHANISTIC INTERPRETATION

Antibacterial mechanisms at the protein level can be divided into three categories. First, biosynthetic proteins build the chemical structure of the metabolite. Second, immunity and transport proteins ensure that the producer can live with its own compound. Thirdly, the target interaction or pathway disruption in the susceptible organism. Together, these three layers show why a BGC is antibacterial [4, 12].

Gene mining at the protein level can find a cluster of genes, but it does not offer the particular substance behind them. Proteomics can identify the expressed proteins; then, comparative bioinformatics can be used to find conserved functional motifs, and targeted gene disruption experiments can reveal which proteins are required for antibacterial activity. Proteomics studies of Streptomyces have identified active BGCs by associating protein expression with changes in metabolite production [14].

This way has genomics and chemistry. Genes and proteins will be regarded as an integrated biosynthetic system rather than as separate analytic areas. Then this system will be linked to the observed antibacterial effects more convincingly than by prediction alone.

5. INTEGRATED WORKFLOW

A feasible discovery workflow generally has five stages. Step 1 is genome sequencing and antiSMASH annotation. Step two is cluster optimisation by resistance genes, novelty, phylogeny and product class. Step three is the activation or induction of silent clusters by culture or regulatory modification. Step four is proteomics and transcriptomics for the expressed proteins. Step five is the function test of metabolite isolation, antibacterial assay and, when feasible, target-level mechanism study [5-6, 11, 14].

This workflow is more robust than any individual method, has a reduced risk of false annotations, and can uncover mechanisms at multiple levels of biology. Increase the likelihood that the new cluster will be associated with actual antibacterial activity and not remain a result of computational prediction.

6. CHALLENGES AND PROSPECTS

Some difficulties remain. Many BGCs are fragmented in draft genomes and difficult to delineate. Some clusters are only produced under very specific conditions and are difficult to reproduce experimentally. Protein annotation is also restricted for biosynthetic enzymes with new domains or low-homology sequences [1, 9, 10]. Antibacterial mechanisms may involve the collaboration of multiple proteins and are not carried out by a single enzyme.

Future research will therefore combine long-read genomics, enhanced cluster databases, high-resolution proteomics, and machine-learning-based prioritisation. Resistance-guided mining and proteomics can be used to explore biosynthetic potential in conjunction with functional analysis. The final goal is a discovery pipeline that can move from a genome to a protein and then to the mechanism of antibacterial action without ambiguity [13-14].

7. CONCLUSION

Antimicrobial biosynthetic gene cluster mining can be conducted more efficiently in conjunction with genome annotation and protein-level mechanistic studies. Genome mining tools can find a cluster, but protein analysis is needed to explain the biosynthesis, immunity, regulation and antibacterial effect [5, 6, 14]. First of all, this cluster of antibacterial genes is not directly antibacterial itself; rather, the proteins in this cluster and their modifications or interactions with other cellular and environmental factors result in an antibacterial effect.

In short, the above indicates that discovery pipelines should move away from cluster detection towards integrated mechanistic interpretation. Only then will antibacterial BGC mining reliably contribute to the discovery of natural products, the development of antibiotics and synthetic biology.

REFERENCES

[1] Du, C., & van Wezel, G. P. (2017). Mining bacterial genomes for secondary metabolite gene clusters. Methods in Molecular Biology, 1520, 23–47. https://doi.org/10.1007/978-1-4939-6688-4_2

[2] Cimermancic, P., Medema, M. H., Claesen, J., Kerezsi, B., Cimermancic, V., Donia, M. S., et al. (2014). Insights into secondary metabolism from a global analysis of biosynthetic gene clusters in the human microbiome. Cell, 158(2), 412–421. https://doi.org/10.1016/j.cell.2014.06.034

[3] Tietz, J. I., & Mitchell, D. A. (2016). Using genome mining to discover natural products. Natural Product Reports, 33(10), 1094–1103. https://doi.org/10.1039/C6NP00048A

[4] Ziemert, N., Alanjary, M., & Weber, T. (2016). The evolution of genome mining in microbes – a review. Natural Product Reports, 33(8), 988–1005. https://doi.org/10.1039/C6NP00025A

[5] Blin, K., Shaw, S., Kloosterman, A. M., Charlop-Powers, Z., van Wezel, G. P., Medema, M. H., & Weber, T. (2021). antiSMASH 6.0: improving cluster detection and comparison capabilities. Nucleic Acids Research, 49(W1), W29–W35. https://doi.org/10.1093/nar/gkab335

[6] Blin, K., Wolf, T., Chevret, E., Lu, X., Schwalen, C. J., Kautz, P., et al. (2019). The antiSMASH database version 2: A comprehensive resource on secondary metabolite biosynthetic gene clusters. Nucleic Acids Research, 47(D1), D634–D639. https://doi.org/10.1093/nar/gky1122

[7] Cox, R. J., Zhang, T., Dorival, A., Vior, E. G., & Challis, G. L. (2022). Resistance-guided mining of bacterial genotoxins defines a distinct family of biosynthetic gene clusters. Nature Chemical Biology, 18, 1–11. https://doi.org/10.1038/s41589-021-00923-8

[8] Yin, Y., Chen, H., Zhang, M., Cao, L., & Chen, H. (2022). Mining of gene clusters for biosynthesis of secondary metabolites and analysis of genes encoding antibiotic resistance and virulence in 4,644 representative human gut strains. Biotechnology Journal, 17(12), 2200220. https://doi.org/10.1002/biot.202200220

[9] Li, X., Wang, Y., Liu, J., Zhou, L., & Wang, H. (2023). Genome mining to identify valuable secondary metabolites in actinobacteria. Archives of Microbiology, 205(4), 127. https://doi.org/10.1007/s00203-023-03487-2

[10] Singh, K., Pal, A., Kaur, J., & Sharma, S. (2024). Phylogenetic classification of natural product biosynthetic gene clusters. Nature Communications, 14, 7961. https://doi.org/10.1038/s41467-024-52389-2

[11] Routledge, M., O'Brien, J. S., & Widdick, D. A. (2016). Antibiotic dialogues: induction of silent biosynthetic gene clusters by co-culture. FEMS Microbiology Reviews, 41(1), 19–33. https://doi.org/10.1093/femsre/fuw035

[12] Takita, M. A., Silva, V. M., & Oliveira, V. M. (2019). Functional analysis of bacteriocin biosynthetic gene clusters. Journal of Bioscience and Bioengineering, 127(1), 1–9. https://doi.org/10.1016/j.jbiosc.2018.06.010

[13] Anonymous. (2025). The transcriptional architecture of bacterial biosynthetic gene clusters. Journal of Natural Products. https://doi.org/10.1021/acs.jnatprod.5c00241

[14] Gubbens, J., de Vries, M., van der Horst, S., & van Wezel, G. P. (2022). Proteomining-based elucidation of natural product biosynthetic pathways in Streptomyces. Frontiers in Microbiology, 13, 913756. https://doi.org/10.3389/fmicb.2022.913756