Natural Resources for Human Health
eISSN: 2583-1194

Original Research

Received 10 April 2026
Revised 18 May 2026
Accepted 20 June 2026

Natural Resources for Human Health 2026, Vol. 6 (3s): 32–45

KEYWORDS:
Endophytic bacteria; Genome mining; antiSMASH; MiBIG; Biosynthetic gene clusters; Secondary metabolites; Natural products; Human health; Drug discovery; Ethnomedicinal plants.

Comparative Genome Mining Reveals Biosynthetic Potential of Endophytic Bacteria from Ethnomedicinal Plants of Manipur as Sustainable Natural Resources for Human Health Applications

Thokchom Shantirani[1], Mukherjee Saikat[2]

[1]Assistant Professor S-2, Microbial Genomics Laboratory, Department of Biochemistry, Manipur University, Manipur-795003, India
[2] Research Staff, Microbial Genomics Laboratory, Department of Biochemistry, Manipur University, Manipur-795003, India
Corresponding Author: Mukherjee Saikat
Email: rajivnandansahai@gmail.com

ABSTRACT:

Background: Plants have always been a great source of medicine for us, and even with new ways of making drugs, they still are [1]. Some bacteria, called endophytic bacteria, live inside these plants and can make lots of different helpful chemicals for human health [2]. These chemicals are often similar to the ones the plants themselves make, or they can work well together with them. This is really interesting because it means these tiny living things can help us find new medicines. They can even make chemicals that are just like the ones the plants make, which is really useful. So, scientists are paying more attention to these interesting endophytic bacteria and the plants they live in, because they might hold the key to new and exciting medicines

Objectives: Recent breakthroughs in sequencing and mining microbial genomes have changed the way we discover special compounds, called metabolites, by letting us predict what a microbe can make before we even try to isolate it. Our research looks at five types of bacteria that live inside plants used in traditional medicine in the Indo-Myanmar Biodiversity Hotspot and to sequence their whole genome and report [3]. We want to see what kinds of useful compounds these bacteria can produce and if they can be a new, sustainable source of natural products for improving human health.

Methods: To understand how certain bacteria can help us, we looked at the complete DNA sequences of five special kinds of bacteria that live inside plants. We used various computer programs such as antiSMASH [4] and MIBiG [5] to find groups of genes that work together to make useful compounds. These compounds might be able to help people feel better. We compared the biosynthetic potential found in these bacteria to see what kinds of useful compounds they might make. We were looking at bacteria called Cellulosimicrobium cellulans MUKR5 [3], Enterobacter hormaechei ST4 Bacillus subtilis ST2, Bacillus paralicheniformis ST3, and Bacillus velezensis ST5. These bacteria were isolated from plants used in traditional medicine in Manipur which lies in the Indo Myanmar Biodiversity hot Spot region. We already had sequenced the whole genome of these bacteria earlier through Next Generation Seqencing and were interested to look at what useful things these bacteria could make to be used for human disease therapy.

Results: The process of comparative genome mining has revealed a stunning array of biosynthetic gene clusters, including non-ribosomal peptide synthetases, type I and type III polyketide synthases[6], ribosomally synthesized and post-translationally modified peptides, terpene biosynthetic pathways, siderophores, betalactones, thiopeptides, and hybrid PKS-NRPS clusters. One particular isolate, Bacillus velezensis ST5, stands out for its exceptionally rich biosynthetic architecture, boasting twelve predicted biosynthetic regions that encode a range of clinically relevant metabolites, such as difficidin, bacillaene, macrolactin H, surfactin, fengycin, bacilysin, and bacillibactin [8-18]. Meanwhile, other isolates have been found to possess complementary biosynthetic capabilities, encoding metabolites like enteromycin, alkylresorcinol, lankacidin C, aerobactin, bacitracin, lichenysin, pulcherriminic acid, and carbapenem-related compounds [19-27]. MIBiG analysis has also identified a multitude of biosynthetic core enzymes, tailoring enzymes, regulatory proteins, and transporter proteins that underpin the complex process of secondary metabolite biosynthesis. This diverse array of biosynthetic gene clusters and associated enzymes highlights the remarkable potential of these isolates to produce a wide range of valuable compounds.

Conclusion: In short, the study shows that bacteria living inside plants used in traditional medicine are a great source of new medicines. By looking at the genes of these bacteria, we can find new ways to make drugs to fight diseases. This is a powerful tool for finding new medicines and can help us make new treatments for many diseases, including ones that are hard to treat. These bacteria are like a treasure chest of new medicines, and studying them can help us discover new ways to fight diseases like cancer, infections, and more.

INTRODUCTION

The increasing prevalence of antimicrobial resistance, emerging infectious diseases, cancer, inflammatory disorders and chronic metabolic diseases has intensified the search for novel therapeutic molecules from natural resources. Although synthetic chemistry has revolutionized modern drug development, nearly half of all approved pharmaceuticals are either natural products or derivatives inspired by naturally occurring compounds. Microbial secondary metabolites have historically contributed numerous clinically important drugs, including β-lactam antibiotics, aminoglycosides, tetracyclines, glycopeptides, statins and immunosuppressants [28, 29]. Consequently, microorganisms remain among the most promising biological reservoirs for the discovery of future medicines.

The world of microbes is full of surprises, and one of the most exciting areas of research is the study of endophytic bacteria. These tiny microorganisms live inside plant tissues, but instead of causing harm, they form mutually beneficial relationships with their hosts. Over time, they develop special ways of making things that help the plants grow strong and healthy, like compounds that fight off other microbes, grab onto iron, and calm down stress. What's really cool is that these compounds come in all shapes and sizes, and they have some amazing properties - they can fight off bacteria, fungi, and even viruses, and they might even be able to help us come up with new treatments for diseases like cancer. As we learn more about the microbes that live inside plants, we're starting to realize that they're a treasure trove of new and exciting things that could help us make new medicines. Plants and the tiny worlds of microbes that live inside them are like a renewable resource, just waiting to be explored and used to make our lives better.

The Indo-Myanmar region is home to a vast array of medicinal plants, making it one of the most valuable areas for discovering new medicines [1]. This area has a unique environment and history that has led to the development of highly diverse communities of microorganisms that live among plants [2]. These microorganisms have the potential to produce new and exciting compounds that could be used to create new medicines. Researchers have already found that some of these microorganisms can help plants grow, fight off diseases, and even produce compounds that could be used as medicines. However, the traditional way of searching for new medicines from these microorganisms is a long and expensive process that

often results in rediscovering compounds that are already known. This process involves growing the microorganisms in a lab, extracting their compounds, purifying them, and then characterizing their structure. Unfortunately, this approach often misses the hidden potential of these microorganisms, as some of their most valuable compounds may not be produced when they are grown in a lab.

Recent advances in whole-genome sequencing and computational genome mining have fundamentally transformed microbial natural-product research. Bioinformatic platforms such as antiSMASH [4] enable rapid identification of biosynthetic gene clusters (BGCs) responsible for the production of polyketides, non-ribosomal peptides, terpenes, ribosomally synthesized and post-translationally modified peptides (RiPPs), siderophores and numerous hybrid secondary metabolites. Complementary databases such as the Minimum Information about a Biosynthetic Gene cluster (MIBiG) facilitate functional annotation of biosynthetic enzymes, tailoring enzymes, regulatory proteins and transport systems by comparison with experimentally characterized gene clusters. Together, these computational approaches substantially accelerate the identification of promising microbial strains for downstream biochemical characterization and synthetic biology applications. Many important antibiotics, like surfactin, fengycin, bacitracin, difficidin, macrolactins, and bacillaene, are made through special pathways called modular non-ribosomal peptide synthetase and polyketide synthase. These pathways are controlled by complex groups of genes that work together to create these antibiotics. They have many different parts, including enzymes that help build the antibiotics, proteins that regulate the process, systems that transport the antibiotics, and enzymes that make final changes to the antibiotics. By studying how these pathways work, we can predict what kinds of antibiotics might be made and find new candidates for making new therapeutic molecules. This can be done through different methods, such as expressing the genes in different organisms, engineering the metabolic pathways, or combining different pathways to create new antibiotics. In the present investigation, five endophytic bacterial isolates recovered from ethnomedicinal plants of Manipur by the authors were comparatively evaluated through whole-genome mining using antiSMASH and MIBiG [4, 5].

The investigated organisms comprised *Cellulosimicrobium cellulans* strain MUKR5 [3], *Enterobacter hormaechei* strain ST4, *Bacillus subtilis* strain ST2, *Bacillus paralicheniformis* strain ST3 and *Bacillus velezensis* strain ST5. Genome mining identified diverse biosynthetic gene clusters encoding therapeutically important secondary metabolites, including polyketides, lipopeptides, siderophores, terpenes and RiPPs, together with their associated biosynthetic enzymes and regulatory machinery. Our study is different from others because it doesn't just look at one type of bacteria, it compares many types of bacteria that live inside medicinal plants. We want to see what kinds of helpful chemicals these bacteria can make. Our study also tries to understand how these chemicals might be used to improve human health. By comparing the genes of these bacteria and what they can do, our study creates a way to prioritize which bacteria might be most useful for making new medicines, and for other areas of research like synthetic biology and biomedical studies. This could lead to new discoveries and a better understanding of how to use these bacteria to help people. The goal is to find new ways to make medicines and to improve our health by using the helpful chemicals that these bacteria produce.

2. MATERIALS AND METHODS

2.1 Study Design

The present investigation employed a comparative genome-mining approach to evaluate the biosynthetic potential of endophytic bacteria isolated from ethnomedicinal plants of Manipur, India, a region belonging to the Indo-Myanmar Biodiversity Hotspot. The workflow consisted of bacterial isolation and characterization, whole-genome sequencing, genome assembly and annotation, computational prediction of biosynthetic gene clusters (BGCs), comparative analysis of biosynthetic pathways, and functional interpretation of the predicted metabolites with respect to their potential applications in human health. The study was designed to identify therapeutically relevant secondary metabolites encoded within the genomes of phylogenetically distinct endophytic bacteria and to compare their biosynthetic capabilities as sustainable natural resources for future pharmaceutical discovery.

2.2 Bacterial Isolates Included in the Comparative Study

We chose five special kinds of bacteria which we have have isolated earlier from tissues of plants used in traditional medicine in Manipur, a north eastern state of India which lies in Indo Burma Biodiversity Hot. These bacteria, called endophytic bacteria, reside within healthy parts of plants Spot. We chose these bacteria because they can fight off bad microbes, help

plants grow, and their whole genetic makeup is already known from our whole genome sequencing data. The five bacteria comprised of:

- *Cellulosimicrobium cellulans* strain MUKR5 (Draft Genome Accession No: **JBMYHP000000000**; Genbank, NCBI)
- *Enterobacter hormaechei* strain ST4 (Draft Genome Accession No: **CP196558**; Genbank, NCBI)
- *Bacillus subtilis* strain ST2 (Draft Genome Accession No: **JBPEOR000000000**; Genbank, NCBI)
- *Bacillus paralicheniformis* strain ST3 (Draft Genome Accession No: **JBPEOS000000000**; Genbank, NCBI)
- *Bacillus velezensis* strain ST5 (Draft Genome Accession No: **CP196559**; Genbank, NCBI)

These bacteria are different species and can make different things, which makes them perfect for comparing their genomes. Their DNA sequences were already known from earlier studies, so we used them for our computer analysis. We have discussed about the details of their genomes and what they can make in the results part.

2.3 Whole-Genome Sequencing and Assembly

We, the authors, at the Microbial Genomics Laboratory, Department of Biochemistry, Manipur University, had earlier isolated five types of bacteria from ethnomedicinal plants of Manipur. The whole genome sequence of these bacteria was performed using the Illumina platform, of BioKart India Private Limited, Bangalore, India. We did quality assessment of the raw reads and adapter trimming before we proceed for genome assembly. The genome assembly was carried out using the computational program Unicycler and circularization of bacterial genome was carried out using Circulator wherever needed.

2.4 Genome Annotation

To understand how an organism works, we need to figure out what each part of its genome does. This is called annotation. For this project, we used computer programs called the NCBI Prokaryotic Genome Annotation Pipeline (PGAP) and Prokka to look at the genome. We wanted to find out where the instructions for making proteins were, as well as other important parts like transfer RNAs, ribosomal RNAs, and non-coding RNAs. We also looked for pseudogenes, which are parts of the genome that don't seem to do anything. By understanding what each part of the genome does, we can start to see how the organism might make special chemicals, like antibiotics. All of the draft genome sequence raw and assembly data were sent to Genbank at the National Center for Biotechnology Information (NCBI

2.5 Prediction of Biosynthetic Gene Clusters

Genome mining was performed using the computational tool antiSMASH (Antibiotics and Secondary Metabolite Analysis Shell), one of the most comprehensive computational platforms for identifying microbial biosynthetic gene clusters. Each assembled genome was analyzed using the default parameters implemented in antiSMASH. The software predicted the genomic locations, biosynthetic classes and putative products of various biosynthetic gene clusters, including:

- Non-ribosomal peptide synthetases (NRPS)
- Type I polyketide synthases (T1PKS)
- Type III polyketide synthases (T3PKS)
- Hybrid PKS-NRPS clusters
- Ribosomally synthesized and post-translationally modified peptides (RiPPs)
- Terpene biosynthetic pathways
- Thiopeptides
- Lassopeptides
- Siderophores

- Metallophores
- β-Lactones
- Saccharide-associated clusters
- Other specialized metabolite pathways

For each biosynthetic region, antiSMASH predicted the genomic coordinates, cluster type, closest homologous biosynthetic gene cluster, percentage similarity and putative secondary metabolite.

2.6 Functional Annotation Using MiBIG

To improve functional interpretation of the identified biosynthetic gene clusters, each predicted BGC was compared with experimentally characterized biosynthetic pathways available in the **Minimum Information about a Biosynthetic Gene cluster (MiBIG)** repository.

MiBIG analysis enabled the prediction of:

- Core biosynthetic enzymes
- Tailoring enzymes
- Regulatory proteins
- Transporter proteins
- Biosynthetic domains
- Functional homologues
- Protein identity (%)
- Coverage
- BLAST score
- Expect (E) values

The integration of antiSMASH with MiBIG significantly enhanced functional annotation by linking predicted biosynthetic pathways with previously characterized natural products and biosynthetic mechanisms.

2.7 Comparative Analysis of Biosynthetic Diversity

The biosynthetic repertoires of all five bacterial genomes were compared to determine differences in:

- Total number of biosynthetic gene clusters
- Distribution of NRPS pathways
- Distribution of PKS pathways
- RiPP biosynthesis
- Terpene biosynthesis
- Siderophore pathways
- Hybrid biosynthetic clusters
- Predicted medically important metabolites

Rather than evaluating individual bacterial genomes independently, the present investigation adopted a comparative framework to identify bacterial strains possessing the highest biosynthetic richness and greatest pharmaceutical potential.

2.8 Therapeutic Annotation of Predicted Secondary Metabolites

The predicted metabolites obtained from antiSMASH and MiBIG analyses were manually curated to determine their reported biological activities and possible applications in human health.

Each metabolite was classified according to its known or predicted pharmacological activities, including:

- Antibacterial
- Antifungal
- Antiviral
- Anticancer
- Antioxidant
- Anti-inflammatory
- Immunomodulatory
- Iron-chelating activity
- Biofilm inhibition
- Quorum sensing regulation

To better understand the potential of these compounds, we looked at existing studies that showed how they worked in real-life experiments. It helped us connect the results of our genome mining to actual medical uses. By comparing the different compounds, we didn't just look at how varied they were in terms of how they were made, but also at how useful they could be for creating new treatments in the future.

2.9 Comparative Evaluation of Biosynthetic Machinery

Besides finding groups of genes that work together to make certain compounds, our study also looked at the enzymes that help create these metabolites, although the details of this part of the research are not included in this paper.

2.10 Data Presentation and Comparative Visualization

The results were synthesized into comparative tables and graphical summaries rather than presenting genome analyses individually. Comparative visualization included:

- Distribution of biosynthetic gene cluster classes
- Diversity of predicted secondary metabolites
- Comparison of biosynthetic enzymes
- Relationship between bacterial species and therapeutic metabolites
- Medical applications of predicted natural products

This strategy facilitated direct comparison among the five endophytic bacteria and highlighted the strains with the greatest biosynthetic potential for future drug discovery.

2.11 Statistical Analysis

The present investigation was primarily based on computational genome mining and comparative bioinformatic analyses. Quantitative comparisons were therefore descriptive in nature and involved enumeration of biosynthetic gene cluster classes,

predicted metabolites and functional proteins across bacterial genomes. Comparative frequencies and proportions were summarized using tables and graphical representations to facilitate interpretation of biosynthetic diversity. No inferential statistical tests were applied because the study focused on genome-derived biosynthetic predictions rather than experimentally measured biological responses.

3. RESULTS AND DISCUSSION

3.1 Comparative Genomic Landscape of Endophytic Bacteria

Genome mining has emerged as a transformative strategy for accelerating natural-product discovery for human health by identifying biosynthetic potential directly from microbial genomes. Unlike traditional metabolite screening, which is limited by culture conditions and the expression of cryptic pathways, comparative genomic analyses enable prediction of both known and previously uncharacterized biosynthetic gene clusters (BGCs). In the present investigation, comparative genome mining of five phylogenetically distinct bacterial endophytes isolated from ethnomedicinal plants revealed substantial variation in their capacity to encode therapeutically relevant secondary metabolites. The bacteria we studied came from different genera, including *Cellulosimicrobium, Enterobacter, and Bacillus*. These bacteria are found inside medicinal plants and can be either Gram-positive or Gram-negative. Even though they live in similar environments, their genetic material showed big differences in how they make certain compounds. This suggests that the variety of bacteria contributes a lot to the diversity of secondary metabolites. Because these bacteria have different abilities to make compounds, the collection of bacteria associated with medicinal plants becomes even more valuable as a source of new medicines. When looking at the different types of bacteria we studied, one called *Bacillus velezensis* ST5 stood out for having a lot of different things it could make. It had twelve areas in its genetic material that could produce different bioactive natural compounds. On the other hand, *Cellulosimicrobium cellulans* MUKR5 and *Enterobacter hormaechei* ST4 each had five big areas that could make bioactive natural things. Meanwhile, *Bacillus subtilis* ST2 and *Bacillus paralicheniformis* ST3 were in the middle when it came to how many different things they could make. This shows that just counting how many areas a bacteria has that can make things isn't enough to figure out if it could be useful for making medicines. We also need to look at the different types of things it can make and what those things are.

Table 1. Comparative biosynthetic diversity of the five endophytic bacterial genomes

Sl No.	Bacterial isolate	Major biosynthetic classes detected	Representative metabolites	Overall biosynthetic diversity
1.	*Cellulosimicrobium cellulans* MUKR5	RiPP, Terpene, T3PKS, Siderophore	Enteromycin; Alkylresorcinol; 7-Deoxypactamycin	Moderate
2	*Enterobacter hormaechei* ST4	NRPS, Thiopeptide, Redox cofactor, Siderophore, Arylpolyene	Enterobactin; Aerobactin; Lankacidin C	Moderate–High
3	*Bacillus subtilis* ST2	NRPS, PKS, RiPP, Terpene	Bacillaene; Fengycin; Surfactin; Bacilysin	High
4	*Bacillus paralicheniformis* ST3	NRPS, RiPP, Thiopeptide, Lassopeptide, Terpene	Bacitracin; Lichenysin; Pulcherriminic acid	High
5	*Bacillus velezensis* ST5	PKS, NRPS, Hybrid PKS–NRPS, Terpene, RiPP	Difficidin; Macrolactin H; Bacillaene; Fengycin; Surfactin; Bacilysin	Very High

3.2 Comparative Distribution of Biosynthetic Gene Cluster Classes

A total of multiple biosynthetic classes were identified across the five genomes, including non-ribosomal peptide synthetases (NRPS), type I and type III polyketide synthases (PKS), terpene pathways, RiPPs, siderophores, metallophores, thiopeptides, lassopeptides, β-lactones and hybrid PKS–NRPS clusters. The coexistence of these pathways within single bacterial genomes highlights the remarkable metabolic versatility of endophytic microorganisms. When we look at how living things make different compounds, we find that NRPS pathways are the most common. The products made by NRPS pathways, like antibiotics, lipopeptides, and anticancer compounds, are really important for medicine. What's interesting is that every bacterial genome we studied had at least one NRPS-related pathway, which suggests that making peptide-based natural products is something that many bacteria can do, especially those that live inside plants used for medicine. This is a key feature of these bacteria and helps us understand how they contribute to the health of the plants they live in. The bacteria *Bacillus velezensis* ST5 has a special set of genes called Hybrid PKS-NRPS clusters. These clusters are really good at making complex molecules that can be used to create new medicines. They do this by combining two different ways of making molecules, one that uses peptides and one that uses polyketides. This combination allows them to make a wide variety of molecules with unique properties, which is exciting for scientists who are looking for new drugs to discover. Having these hybrid pathways means that there are many more possibilities for creating new medicines, which could lead to new treatments for diseases. Terpene biosynthetic pathways were detected in four bacterial species. Although microbial terpenes remain less explored than plant terpenes, many possess antimicrobial, antioxidant, anti-inflammatory and anticancer activities, making them attractive targets for future metabolomic investigations. We found that certain pathways related to RiPPs were very important in two specific types of bacteria: *Cellulosimicrobium cellulans* MUKR5 and *Bacillus paralicheniformis* ST3. RiPPs are a group of peptides that are made by cells and then modified in various ways, often resulting in molecules that are very stable and have strong antimicrobial properties. This means they can effectively fight against other microorganisms. The fact that these pathways are prominent in these bacteria suggests that RiPPs play a significant role in their ability to interact with and influence their environment.

Table 2. Comparative distribution of biosynthetic gene cluster (BGC) classes among the five endophytic bacterial isolates

BGC class	MUKR5 (*Cellulosimicrobium cellulans*)	ST4 (*Enterobacter hormaechei)*	ST2 (*Bacillus subtilis*)	ST3 (*Bacillus paralicheniformis*)	ST5 (*Bacillus velezensis*)
NRPS	✓	✓	✓	✓	✓
Type I PKS	–	–	✓	–	✓
Type III PKS	✓	–	✓	✓	✓
Hybrid PKS–NRPS	✓	✓	✓	✓	✓
RiPP	✓	–	✓	✓	✓
Terpene	✓	–	✓	✓	✓
Siderophore	✓	✓	✓	✓	✓
Metallophore	–	✓	–	✓	✓
Thiopeptide	✓	✓	–	✓	–
Lassopeptide	–	–	–	✓	–
β-Lactone	–	–	✓	✓	✓

Presence (✓) indicates that the corresponding biosynthetic gene cluster class was predicted in the genome of the isolate based on the comparative antiSMASH analysis summarized from the manuscript raw data. Abbreviations: NRPS, non-ribosomal peptide synthetase; PKS, polyketide synthase; RiPP, ribosomally synthesized and post-translationally modified peptide.

3.3 Comparative Analysis of Predicted Therapeutically Important Secondary Metabolites

Our study made a really important discovery - it found a huge variety of metabolites that could be used to make new medicines in the genomes of five different types of bacteria that live inside other organisms. What's interesting is that instead of each bacterium making just one main metabolite, they each made a special mix of different things like antibiotics, molecules that grab onto iron, and peptides that are good for human. This suggests that these bacteria have different strengths when it comes to making these molecules, and that they could work well together. The genome of Cellulosimicrobium *cellulans* MUKR5 contains genes that produce several important compounds, including enteromycin, alkylresorcinol, and 7-deoxypactamycin. Enteromycin is a type of antibiotic that is effective against Gram-positive bacteria, which can cause a range of illnesses. Alkylresorcinols, on the other hand, are a type of phenolic lipid that have been shown to have both antimicrobial and antioxidant properties. The fact that MUKR5 is capable of producing both antibacterial and antioxidant compounds suggests that it may have a wide range of potential applications in the field of medicine, beyond just its use as an antimicrobial agent. This multifunctional potential makes MUKR5 a fascinating subject for further study, as it could potentially lead to the development of new and innovative treatments for a variety of diseases. *Enterobacter hormaechei* ST4 has a unique way of making things, which includes enterobactin, aerobactin, lankacidin C, and arylpolyenes. Enterobactin and aerobactin are like special helpers that grab iron from the host, even when the host is trying to limit iron availability. Meanwhile, arylpolyenes seem to protect against oxidative stress. Typically, siderophores are thought of as part of how microbes work, but their ability to bind iron has caught attention for use in medicine, such as fighting microbes by targeting their iron use or creating new antibiotics linked to siderophores. This area of research is interesting because it could lead to new ways to combat infections. By understanding how microbes like *Enterobacter hormaechei* ST4 make and use these compounds, scientists can explore new strategies for therapeutic applications. The three Bacillus species collectively encoded the richest diversity of bioactive metabolites. *Bacillus velezensis* ST5 contained biosynthetic pathways for difficidin, bacillaene, macrolactin H, fengycin, surfactin, bacilysin and bacillibactin, representing a broad spectrum of antibacterial, antifungal, antiviral and anticancer activities. The simultaneous presence of multiple lipopeptide and polyketide pathways indicates that ST5 may possess exceptional biosynthetic plasticity and therefore represents an outstanding candidate for future natural-product discovery.

Table 3. Predicted therapeutically important secondary metabolites and their biomedical relevance

Bacterial isolate	Major predicted metabolites	Predominant BGC class	Principal biomedical relevance
Cellulosimicrobium cellulans MUKR5	Enteromycin; Alkylresorcinol; 7-Deoxypactamycin	RiPP; T3PKS; Terpene; Siderophore-associated	Antibacterial; antioxidant; potential anticancer activity
Enterobacter hormaechei ST4	Lankacidin C; Enterobactin; Aerobactin; Arylpolyene [28, 29]	NRPS; Siderophore; Redox-cofactor; Arylpolyene	Antibacterial; iron acquisition; oxidative stress protection
Bacillus subtilis ST2	Bacillaene; Fengycin; Surfactin; Bacilysin; Subtilosin A [30]	NRPS; PKS; RiPP	Broad-spectrum antibacterial and antifungal activities
Bacillus paralicheniformis ST3	Bacitracin; Lichenysin; Pulcherriminic acid; Bacillibactin [19-27]	NRPS; RiPP; Siderophore	Antibiotic production; biosurfactant activity; iron sequestration
Bacillus velezensis ST5	Difficidin; Macrolactin H; Bacillaene; Fengycin; Surfactin; Bacilysin;	Hybrid PKS-NRPS; PKS; NRPS; RiPP	Antibacterial; antifungal; antiviral; anti-inflammatory; potential anticancer activity

	Bacillibactin [8-18]		

The table summarizes the major secondary metabolites predicted from comparative antiSMASH and MiBIG analyses and highlights their potential biomedical significance. Abbreviations: BGC, biosynthetic gene cluster; NRPS, non-ribosomal peptide synthetase; PKS, polyketide synthase; RiPP, ribosomally synthesized and post-translationally modified peptide.

3.4 Discussion

The study shows that the tiny organisms living inside medicinal plants are not just making individual compounds that fight off germs, but they have a wide range of genetic tools that can be used to create new medicines. What's really interesting is that these organisms have similar systems for making these compounds as the ones used to make many antibiotics that are already working well. This is important because it highlights the value of looking at the natural world for solutions to the growing problem of antibiotic resistance. By studying these tiny organisms, we can learn more about how to create new medicines that can help us fight off infections. The fact that these organisms have such a diverse range of genetic tools means that they have a lot of potential for helping us develop new treatments. Instead of just looking at the known compounds that are already present, the variety of BGC classes found in this study shows that there's a big potential for producing new and modified versions of these compounds through genetic engineering or by activating dormant gene clusters. This discovery makes a strong case for using genome-guided approaches in future natural product discovery programs, and it also highlights the need for further validation and characterization of these endophytic strains through metabolomic analysis. By doing so, we can unlock the full potential of these microorganisms and potentially discover new and innovative compounds. The fact that these strains have the ability to produce cryptic or structurally modified analogues is particularly exciting, as it opens up new avenues for research and development in the field of natural product discovery.

4. CONCLUSION

The study of the genomes of bacteria that live inside plants used in traditional medicine in the Indo-Myanmar region has shown that these bacteria are a rich source of new and varied natural products that could be useful for human health. By using advanced computer programs to analyze the entire genomes of these bacteria, we were able to identify many groups of genes that work together to produce compounds that have potential therapeutic benefits. When we compared the genomes of different bacteria, such as *Cellulosimicrobium cellulans*, *Enterobacter hormaechei*, *Bacillus subtilis*, *Bacillus paralicheniformis*, *and Bacillus velezensis*, we found a remarkable range of genetic pathways that produce different types of compounds, including peptides, polyketides, and terpenes. These findings suggest that the bacteria that live inside medicinal plants are incredibly versatile in their ability to produce a wide range of compounds, and that they could be a valuable source of new medicines. Our study used powerful tools like antiSMASH and MiBIG to analyze the genomes of these bacteria and identify the genes that are involved in producing these compounds. Overall, the study highlights the importance of exploring the genetic diversity of endophytic bacteria and their potential to produce new and useful natural products. The bacteria found in plants used for traditional medicine have a lot to offer when it comes to creating new medicines. We looked at five different types of bacteria and found that one, called *Bacillus velezensis* strain ST5, was especially good at making a variety of compounds that could be used as medicine. These compounds include difficidin, bacillaene, and surfactin, which are all important for human health. Two other types of bacteria, *Bacillus subtilis* ST2 and *Bacillus paralicheniformis* ST3, also made some valuable antibiotics and lipopeptides. The other two bacteria, *Cellulosimicrobium cellulans* MUKR5 and *Enterobacter hormaechei* ST4, made some unique compounds like enteromycin and lankacidin C. What's interesting is that each of these bacteria makes different things, so together they offer a wide range of possibilities for creating new medicines. This shows that the tiny organisms living in and around plants used for traditional medicine are a treasure trove of new ideas for pharmaceuticals. By studying these bacteria, scientists can discover new compounds and expand our knowledge of what's possible in the world of medicine. The study of the key enzymes and proteins involved in creating complex molecules has shown just how complicated and sophisticated the process is. This complexity gives us a lot of opportunities to use synthetic biology, engineer metabolism, and express hidden gene clusters in different organisms. What's more, using a tool called MiBIG, we found many more proteins that weren't predicted by another tool called antiSMASH. This means there are probably many biosynthetic pathways that we don't know much about yet, which increases the chances of discovering new and useful molecules in the future. By looking deeper into

these pathways, we might find completely new molecules with important biological activities. This is exciting because it could lead to new discoveries and a better understanding of how these complex molecules are made. The fact that we can now see the intricate details of these biosynthetic pathways opens up new possibilities for research and experimentation, and could potentially lead to breakthroughs in fields like medicine and biotechnology. The study of microbes has led to the discovery of many compounds that can fight off bacteria, fungi, viruses, and even cancer. These compounds have unique ways of working that are different from the medicines we have now. This is important because many germs are becoming resistant to our current medicines, which is a big threat to public health worldwide. By looking at the genetic material of certain bacteria that live inside plants, we can quickly and affordably find the ones that have the most potential to make new medicines. This approach brings together several fields of research, including the study of microbes, the chemistry of natural products, and biomedical research, to speed up the discovery of new treatments that come from sustainable sources. In conclusion, this investigation establishes endophytic bacteria associated with ethnomedicinal plants as promising renewable natural resources for future drug discovery. The comparative genome-mining strategy adopted in this study provides a robust framework for identifying high-value microbial strains and prioritizing biosynthetic gene clusters with significant therapeutic potential. Future studies should focus on activating cryptic biosynthetic pathways, integrating transcriptomics and metabolomics with genome mining, experimentally validating predicted metabolites through heterologous expression and advanced analytical chemistry, and evaluating their pharmacological efficacy using *in vitro* and *in vivo* disease models. Such multidisciplinary approaches will facilitate the translation of genome-derived biosynthetic predictions into clinically useful natural products and contribute to the development of novel antimicrobial, anticancer and other therapeutic agents for improving human health

ACKNOWLEDGEMENTS

We want to thank the Department of Biochemistry at Manipur University in Imphal, India. We're also grateful to the National Center for Biotechnology Information, antiSMASH, and the Minimum Information about a Biosynthetic Gene Cluster consortium (MIBiG) for letting us use their genomic databases and computer tools for free, which helped us compare and analyze the genomes in this study. Additionally, we appreciate BioKart India Private Limited, Bangalore, India for Illumina platform based next-generation sequencing.

AUTHOR CONTRIBUTIONS

Shantirani Thokchom: Conceptualization, Investigation., Methodology, Laboratory Analysis., Data Curation, and Analysis.

Dr. Saikat Mukherjee: Supervision, Writing of the manuscript, its Review & Editing.

CONFLICT OF INTERESTS

The authors declare that there is no conflict of interest.

ETHICS STATEMENT

Ethical approval was not required for this study.

DATA AVAILABILITY STATEMENT

The information that backs up this study's results can be found in the article itself. We had already isolated the five bacteria and did their whole genome seuqneicng. This data has been stored in the National Center for Biotechnology Information (NCBI) databases under GenBank accession numbers. For other results of our comparative genome minings and other supporting info, the corresponding author of this study can be contacted.

REFERENCES

[1] Newman, D.J., Cragg, G.M., 2020. Natural products as sources of new drugs over the nearly four decades from 01/1981 to 09/2019. Journal of Natural Products 83(3), 770–803. https://doi.org/10.1021/acs.jnatprod.9b01285.

[2] Gouda, S., Das, G., Sen, S.K., Shin, H.-S., Patra, J.K., 2016. Endophytes: A treasure house of bioactive compounds of medicinal importance. Frontiers in Microbiology 7, 1538. https://doi.org/10.3389/fmicb.2016.01538.

[3] Thokchom, S., Mukherjee, S., Ningthoujam, D.S., 2026. Draft genome sequence of plant growth-promoting bacterial endophyte Cellulosimicrobium cellulans strain MUKR 5 isolated from Iris laevigata Fisch. from Manipur, India. Microbiology Resource Announcements 15(2), e00892-25. https://doi.org/10.1128/mra.00892-25.

[4] Blin, K., Shaw, S., Augustijn, H.E., Reitz, Z.L., Biermann, F., Alanjary, M., Fetter, A., Terlouw, B.R., Metcalf, W.W., Helfrich, E.J.N., van Wezel, G.P., Medema, M.H., Weber, T., 2023. antiSMASH 7.0: New and improved predictions for detection, regulation, chemical structures and visualisation. Nucleic Acids Research 51(W1), W46–W50. https://doi.org/10.1093/nar/gkad344.

[5] Terlouw, B.R., Blin, K., Navarro-Muñoz, J.C., et al., 2023. MIBiG 3.0: A community-driven effort to annotate experimentally validated biosynthetic gene clusters. Nucleic Acids Research 51(D1), D603–D610. https://doi.org/10.1093/nar/gkac1049.

[6] Donadio, S., Monciardini, P., Sosio, M., 2007. Polyketide synthases and nonribosomal peptide synthetases: The emerging view from bacterial genomics. Natural Product Reports 24, 1073–1109. https://doi.org/10.1039/B514050C.

[7] Ongena, M., Jacques, P., 2008. Bacillus lipopeptides: Versatile weapons for plant disease biocontrol. Trends in Microbiology 16(3), 115–125. https://doi.org/10.1016/j.tim.2007.12.009.

[8] Zhen, C., Ge, X.-F., Lu, Y.-T., Liu, W.-Z., 2023. Chemical structure, properties and potential applications of surfactin, as well as advanced strategies for improving its microbial production. AIMS Microbiology 9(2), 195–217. https://doi.org/10.3934/microbiol.2023012.

[9] Deleu, M., Paquot, M., Nylander, T., 2005. Fengycin interaction with lipid monolayers at the air–aqueous interface: Implications for the effect of fengycin on biological membranes. Journal of Colloid and Interface Science 283(2), 358–365. https://doi.org/10.1016/j.jcis.2004.09.036.

[10] Wise, C., Falardeau, J., Hagberg, I., Avis, T.J., 2014. Cellular lipid composition affects sensitivity of plant pathogens to fengycin, an antifungal compound produced by Bacillus subtilis strain CU12. Phytopathology 104(10), 1036–1041. https://doi.org/10.1094/PHYTO-12-13-0336-R.

[11] Patel, P.S., Huang, S., Fisher, S., Pirnik, D., Aklonis, C., Dean, L., Meyers, E., Fernandes, P., Mayerl, F., 1995. Bacillaene, a novel inhibitor of procaryotic protein synthesis produced by Bacillus subtilis: Production, taxonomy, isolation, physico-chemical characterization and biological activity. Journal of Antibiotics 48(9), 997–1003. https://doi.org/10.7164/antibiotics.48.997.

[12] Zimmerman, S.B., Schwartz, C.D., Monaghan, R.L., Pelak, B.A., Weissberger, B., Gilfillan, E.C., Mochales, S., Hernandez, S., Currie, S.A., Tejera, E., et al., 1987. Difficidin and oxydifficidin: Novel broad spectrum antibacterial antibiotics produced by Bacillus subtilis. I. Production, taxonomy and antibacterial activity. Journal of Antibiotics 40(12), 1677–1681. https://doi.org/10.7164/antibiotics.40.1677.

[13] Romero-Tabarez, M., Jansen, R., Sylla, M., Lünsdorf, H., Häussler, S., Santosa, D.A., Timmis, K.N., Molinari, G., 2006. 7-O-Malonyl macrolactin A, a new macrolactin antibiotic from Bacillus subtilis active against methicillin-resistant Staphylococcus aureus, vancomycin-resistant enterococci, and a small-colony variant of Burkholderia cepacia. Antimicrobial Agents and Chemotherapy 50(5), 1701–1709. https://doi.org/10.1128/AAC.50.5.1701-1709.2006.

[14] Kim, D.H., Kim, H.K., Kim, K.M., Kim, C.K., Jeong, M.H., Ko, C.Y., Moon, K.H., Kang, J.S., 2011. Antibacterial activities of macrolactin A and 7-O-succinyl macrolactin A from Bacillus polyfermenticus KJS-2 against

vancomycin-resistant enterococci and methicillin-resistant Staphylococcus aureus. Archives of Pharmacal Research 34(1), 147–152. https://doi.org/10.1007/s12272-011-0117-0.

[15] Ozcengiz, G., Alaeddinoglu, N.G., Demain, A.L., 1990. Regulation of biosynthesis of bacilysin by Bacillus subtilis. Journal of Industrial Microbiology 6(2), 91–100. https://doi.org/10.1007/BF01576428.

[16] Rajavel, M., Mitra, A., Gopal, B., 2009. Role of Bacillus subtilis BacB in the synthesis of bacilysin. Journal of Biological Chemistry 284(46), 31882–31892. https://doi.org/10.1074/jbc.M109.014522.

[17] May, J.J., Wendrich, T.M., Marahiel, M.A., 2001. The dhb operon of Bacillus subtilis encodes the biosynthetic template for the catecholic siderophore 2,3-dihydroxybenzoate-glycine-threonine trimeric ester bacillibactin. Journal of Biological Chemistry 276(10), 7209–7217. https://doi.org/10.1074/jbc.M009140200.

[18] Stone, K.J., Strominger, J.L., 1971. Mechanism of action of bacitracin: Complexation with metal ion and C55-isoprenyl pyrophosphate. Proceedings of the National Academy of Sciences of the United States of America 68(12), 3223–3227. https://doi.org/10.1073/pnas.68.12.3223.

[19] Raymond, K.N., Dertz, E.A., Kim, S.S., 2003. Enterobactin: An archetype for microbial iron transport. Proceedings of the National Academy of Sciences of the United States of America 100(7), 3584–3588. https://doi.org/10.1073/pnas.0630018100.

[20] de Lorenzo, V., Bindereif, A., Paw, B.H., Neilands, J.B., 1986. Aerobactin biosynthesis and transport genes of plasmid ColV-K30 in Escherichia coli K-12. Journal of Bacteriology 165(2), 570–578. https://doi.org/10.1128/jb.165.2.570-578.1986.

[21] Johnston, I., Osborn, L.J., Markley, R.L., McManus, E.A., Kadam, A., Schultz, K.B., Nagajothi, N., Ahern, P.P., Brown, J.M., Claesen, J., 2021. Identification of essential genes for Escherichia coli aryl polyene biosynthesis and function in biofilm formation. npj Biofilms and Microbiomes 7, 56. https://doi.org/10.1038/s41522-021-00226-3.

[22] Arnaouteli, S., Matoz-Fernandez, D.A., Porter, M., Kalamara, M., Abbott, J., MacPhee, C.E., Davidson, F.A., Stanley-Wall, N.R., 2019. Pulcherrimin formation controls growth arrest of the Bacillus subtilis biofilm. Proceedings of the National Academy of Sciences of the United States of America 116(27), 13553–13562. https://doi.org/10.1073/pnas.1903982116.

[23] Miyanaga, A., Horinouchi, S., 2009. Enzymatic synthesis of bis-5-alkylresorcinols by resorcinol-producing type III polyketide synthases. Journal of Antibiotics 62(7), 371–376. https://doi.org/10.1038/ja.2009.44.

[24] Milke, L., Kabuu, M., Zschoche, R., Gätgens, J., Krumbach, K., Carlstedt, K.-L., Wurzbacher, C.E., Balluff, S., Beemelmanns, C., Jogler, C., Marienhagen, J., Kallscheuer, N., 2024. A type III polyketide synthase cluster in the phylum Planctomycetota is involved in alkylresorcinol biosynthesis. Applied Microbiology and Biotechnology 108, 239. https://doi.org/10.1007/s00253-024-13065-x.

[25] Coronel-León, J., Marqués, A.M., Bastida, J., Manresa, A., 2016. Optimizing the production of the biosurfactant lichenysin and its application in biofilm control. Journal of Applied Microbiology 120(1), 99–111. https://doi.org/10.1111/jam.12992.

[26] Grangemard, I., Bonmatin, J.M., Bernillon, J., Das, B.C., Peypoux, F., 1999. Lichenysins G, a novel family of lipopeptide biosurfactants from Bacillus licheniformis IM 1307: Production, isolation and structural evaluation by NMR and mass spectrometry. Journal of Antibiotics 52(4), 363–373. https://doi.org/10.7164/antibiotics.52.363.

[27] Jenny, K., Käppeli, O., Fiechter, A., 1991. Biosurfactants from Bacillus licheniformis: Structural analysis and characterization. Applied Microbiology and Biotechnology 36(1), 5–13. https://doi.org/10.1007/BF00164690.

[28] Sohn, M.-J., Zheng, C.-J., Kim, W.-G., 2008. Macrolactin S, a new antibacterial agent with FabG-inhibitory activity from Bacillus sp. AT28. Journal of Antibiotics 61, 687–691. https://doi.org/10.1038/ja.2008.98.

[29] Mondol, M.A.M., Tareq, F.S., Kim, J.H., Lee, M.A., Lee, H.-S., Lee, J.S., Lee, Y.-J., Shin, H.J., 2013. New antimicrobial compounds from a marine-derived Bacillus sp. Journal of Antibiotics 66, 89–95. https://doi.org/10.1038/ja.2012.102.

[30] Shelburne, C.E., An, F.Y., Dholpe, V., Ramamoorthy, A., Lopatin, D.E., Lantz, M.S., 2007. The spectrum of antimicrobial activity of the bacteriocin subtilosin A. Journal of Antimicrobial Chemotherapy 59(2), 297–300. https://doi.org/10.1093/jac/dkl495.