

UC San Diego

UC San Diego Previously Published Works

Title

Bacterial diversity and chemical ecology of natural product-producing bacteria from Great Salt Lake sediment.

Permalink

<https://escholarship.org/uc/item/6s6771rp>

Journal

ISME Communications, 4(1)

Authors

Bring Horvath, Elijah

Brazelton, William

Kim, Min

et al.

Publication Date

2024

DOI

10.1093/ismeco/ycae029

Copyright Information

This work is made available under the terms of a Creative Commons Attribution License, available at <https://creativecommons.org/licenses/by/4.0/>

Peer reviewed

Bacterial diversity and chemical ecology of natural product-producing bacteria from Great Salt Lake sediment

Elijah R. Bring Horvath ^{1,2,3}, William J. Brazelton ², Min Cheol Kim ⁴, Reiko Cullum ⁴, Matthew A. Mulvey ^{2,5}, William Fenical ^{4,6,7}, Jaclyn M. Winter ^{1,*}

¹Department of Pharmacology and Toxicology, University of Utah, Salt Lake City, UT 84112, United States

²School of Biological Sciences, University of Utah, Salt Lake City, UT 84112, United States

³Department of Medicinal Chemistry, University of Utah, Salt Lake City, UT 84112, United States

⁴Center for Marine Biotechnology and Biomedicine, Scripps Institution of Oceanography, University of California at San Diego, CA 92093, United States

⁵Henry Eyring Center for Cell and Genome Science, University of Utah, Salt Lake City, UT 84112, United States

⁶Skaggs School of Pharmacy and Pharmaceutical Sciences, University of California at San Diego, CA 92093, United States

⁷Moore Comprehensive Cancer Center, University of California at San Diego, CA 92037, United States

*Corresponding author: Jaclyn M. Winter, Department of Pharmacology and Toxicology, University of Utah, 30 S 2000 E, SKH 201, Salt Lake City, UT 84112, United States. Email: jaclyn.winter@utah.edu

Abstract

Great Salt Lake (GSL), located northwest of Salt Lake City, UT, is the largest terminal lake in the USA. While the average salinity of seawater is ~3.3%, the salinity in GSL ranges between 5% and 28%. In addition to being a hypersaline environment, GSL also contains toxic concentrations of heavy metals, such as arsenic, mercury, and lead. The extreme environment of GSL makes it an intriguing subject of study, both for its unique microbiome and its potential to harbor novel natural product-producing bacteria, which could be used as resources for the discovery of biologically active compounds. Though work has been done to survey and catalog bacteria found in GSL, the Lake's microbiome is largely unexplored, and little to no work has been done to characterize the natural product potential of GSL microbes. Here, we investigate the bacterial diversity of two important regions within GSL, describe the first genomic characterization of Actinomycetota isolated from GSL sediment, including the identification of two new Actinomycetota species, and provide the first survey of the natural product potential of GSL bacteria.

Keywords: Great Salt Lake, halophiles, natural products, Actinomycetota, chemical ecology

Introduction

Great Salt Lake (GSL), located in the state of UT, is the largest saltwater body in the USA and represents one of the most hypersaline and extreme environments in the world [1, 2]. Its area spans five counties, including Weber, Box Elder, Salt Lake, Tooele, and Davis, and covers ~4400 km², though its area is rapidly shrinking due to climate change and water diversion [1, 3, 4]. Due to its location in the Great Basin of the Intermountain West, GSL experiences temperature fluctuations ranging from -5°C in the winter months to >35°C in the summer months. As GSL is a terminal lake, meaning that water flowing into it only leaves by evaporation, and the minerals, ions, and salts that enter the lake are retained and concentrated. A railroad causeway constructed in the 1950s physically divides GSL into two main regions: the North and South Arms. During the snowpack melt in the spring, most of the fresh water is brought into the South Arm and because of this, the salinity can range from 5% to >15% in the South Arm and often reaches ≥27% in the North Arm, which receives little to no fresh water inflow [5].

The chemical makeup of GSL closely resembles that of typical ocean water, with sodium and chloride being the primary ions, followed by sulfate, magnesium, calcium, and potassium [1]. Although the North Arm exhibits higher ion abundance than the South Arm, their ion compositions are similar [6, 7]. However, the considerable differences in salinity significantly influence the microbiota between the two arms, creating two sub ecosystems within the lake [1, 2, 7, 8]. A broader microbial diversity is observed in the South Arm compared to the North Arm, and this is primarily attributed to the differences in salinity [9, 10]. Interestingly, early studies found that the majority of halophiles isolated from GSL habitats were obligate halophiles rather than halotolerant transplants [11]. Media studies using nutrient agar with varying concentrations of GSL water, Pacific Ocean water, or distilled water revealed a notable decrease in colony-forming units as the percentage of GSL water was reduced in the media [12]. These findings suggest that despite similar ion concentrations between GSL and typical ocean water, the microbiology between these two environments is quite different.

Received: 28 November 2023. Revised: 20 January 2024. Accepted: 27 February 2024

© The Author(s) 2024. Published by Oxford University Press on behalf of the International Society for Microbial Ecology.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

Aside from being a hypersaline environment, GSL also exhibits elevated levels of heavy metals and metalloids, including mercury, arsenic, cadmium, and lead [13]. Due to the extreme environments of GSL, microorganisms surviving and thriving in the North and South Arms have been described as polyextremophiles, as they have adapted to not only toxic levels of heavy metals and extreme osmotic stress but also to extreme seasonal temperature fluctuations [1, 9]. To thrive in these harsh environments, GSL microorganisms may evolve and employ different methods for survival, such as heavy metal efflux, degradation of toxic compounds [14], and production of specialized small molecules. Of particular interest to us is the natural product potential of GSL bacteria.

Natural products are small molecules produced in nature and represent some of the most important pharmaceutical agents used in human health care [15–17]. This especially holds true with anti-infective agents, as many of the current clinically significant antibiotics are natural products or derivatives thereof. However, antibiotic resistance is on the rise and it has been recognized by the World Health Organization as a leading global health issue [18, 19], which needs to be urgently addressed. Antibiotic discovery over the last few decades has relied heavily on chemically modifying the scaffolds of known antibiotic agents. Therefore, the identification of bioactive natural products possessing new chemical scaffolds is a promising approach for the discovery of antibiotics with novel modes of action. Unique environments can influence the chemical diversity of natural products, and microorganisms isolated from extreme environments serve as ideal resources for drug discovery efforts. Furthermore, because GSL is a terminal lake and serves as the endpoint of wastewater treatment runoff from a major metropolitan center, we expect pathogenic microbes residing there to be a reservoir of antibiotic resistance genes. Therefore, native GSL microbes may be producing novel antibiotic agents to compete with these pathogens, which could make them an ideal resource for the identification of new natural products with novel scaffolds and new mechanisms of action.

To prioritize strains for downstream fermentation studies and streamline the labor-intensive natural product discovery pipeline when working with large numbers of bacterial isolates, robust bioinformatic and genomic tools can complement the bioactivity-guided isolation efforts [20–22]. In the producing organism, the genes encoding the enzymatic machinery used to assemble small molecules are typically clustered together within a chromosome or on extra-chromosomal genetic elements. By correlating genetic information to protein function, chemical logic can be used to connect a natural product to its respective biosynthetic gene cluster (BGC) [23]. However, little information on bacterial populations and even less on bacterial genomes has been reported from GSL [1, 8, 9]. Between the end of the 19th century and early 20th century, culture-dependent approaches were used to identify bacteria within GSL, with most research focusing on planktonic communities from the water column [11, 24–29] rather than sediment-derived microbial populations [10]. However, after construction of the causeway in the 1950s, culture-independent approaches became the predominant method for monitoring the composition of planktonic communities in the water column, as perturbations within the community can negatively affect the brine shrimp industry [8, 10, 27, 29, 30]. While the culture-dependent and culture-independent approaches were used to study and understand the microbial ecology of GSL, no genomic studies have been published to date and little is known about the natural product potential of the microorganisms residing in the hypersaline environment. Thus, it is imperative that we further characterize this unique and understudied ecosystem,

with a focus on expanding our knowledge of what sediment-dwelling microbes are present using both culture-dependent and independent methods. Just as importantly, we must also assess GSL bacteria for their potential to produce novel natural products. Due to severe drought-related shrinkage of the Lake, it is imperative that we perform these studies before this invaluable resource is gone [1, 3]. Here, we investigate the microbial diversity in the South Arm of GSL as well as characterize the natural product potential of Actinomycetota strains isolated from the GSL sediment.

Materials and methods

Collection and isolation of bacteria strains

Sediment samples were collected in sterile Whirl-Pak bags or sterile Falcon tubes from the South Arm of GSL at four Black Rock Beach (BRB) and four Marina sites during the summer of 2017. For each collection site, samples were taken ca. 2.5 m apart in a linear sampling pattern. *Saccharomonospora* sp. GSL17-019 and *Streptomyces* sp. GSL17-113 were isolated from the Marina sediment that was desiccated for 72 h in a biological safety cabinet. Desiccated sediment, ~0.5 g, was added to yeast-peptone-mannitol (YPM) agar plates (per liter: 4 g mannitol, 2 g yeast extract, 2 g peptone, 18 g agar, and 42 g Instant Ocean Aquarium Sea Salt Mixture, Spectrum Brands, USA). Plates were incubated at 30°C for up to 90 days, and bacterial colonies were subcultured on YPM media until pure isolates were obtained. The 16S rRNA gene sequencing was used to identify *Saccharomonospora* sp. GSL17-019, *Streptomyces* sp. GSL17-113, and *Streptomyces* sp. GSL17-111 using NCBI Blast/Blast+ [31] via multiBLAST (<https://github.com/ERBringHorvath/multiBLAST>) (GenBank accession for the *Escherichia coli* reference 16S rRNA sequence is CP082835.1).

Environmental sequencing of 16S ribosomal ribonucleic acid genes

DNA preparations of GSL sediment were carried out using a FastDNA Spin Kit (MP Biomedicals, Cat. No.: 116540-600). Sequencing of amplicons generated from 16S rRNA genes was performed at the Genomics Core Facility at Michigan State University on an Illumina MiSeq instrument using dual-indexed Illumina fusion primers targeting the V4 region of the 16S rRNA gene [32]. Amplicon sequence variants (ASVs) were inferred from 16S rRNA amplicon sequences with DADA2 v. 1.10.1 [33] after the removal of primer sequences with cutadapt v. 1.15 [34]. Taxonomic classification of ASVs was performed with DADA2 using the SILVA reference alignment (SSURfv132) and taxonomy outline [35, 36]. The ordination plot was generated with PhyloSeq v. 1.26.1 [22] using principal coordinate analysis (PCoA) ordination of Bray-Curtis dissimilarity values.

Statistical analysis and data visualization

Multivariate plot and DESeq2 analyses were carried out using PhyloSeq v.1.42.0 [21, 22] in RStudio [37] using R v.4.3.1. DESeq2 v.1.38.3 was used to compare ASV data from the four BRB and four Marina sampling sites and graph differential abundance. During the DESeq2 analysis, each of the four collection sites from both BRB and the Marina were treated as replicates of the primary collection region to obtain a BRB versus Marina comparison. Significance was determined using a Wald test with a Benjamini and Hochberg adjusted *P*-value. Only ASVs meeting the minimal cut-off of an adjusted *P*-value of $P > .01$ were included in the results. Bar plots were created in RStudio using ggplot2 v.3.4.2 and ggbreak v.0.1.1 [38]. PhyloSeq-generated taxonomy classifications were updated to reflect current literature using the NCBI

taxonomy browser. Phylogenetic trees were constructed with autoMLST using IQ-TREE Ultrafast Bootstrap analysis of 1000 replicates. The autoMLST-generated phylogenetic tree was exported as a Newick file and was used to generate final figures in R using ggtree v.3.8.0 [39].

Genome sequencing and genome mining of Great Salt Lake genomes

For *Saccharomonospora* sp. GSL17-019 and *Streptomyces* sp. GSL17-113, construction of genomic libraries and sequencing were performed at the High-Throughput Genomics Center in the Huntsman Cancer Institute at the University of Utah. A PCR-free NEB-Next Ultra II DNA library was generated and sequenced using NovaSeq S4 Reagent Kit v1.5 (150 × 150 bp and 2500 M read-pairs/lane). Adapter sequences and PhiX were removed from all reads with BBDuk, part of the BBtools suite, V35.85 [40]. Quality trimming was performed with seq-qc as previously described [41], and sequences were assembled using SPAdes [42]. Contigs shorter than 200 bp were trimmed using a custom Python script. CheckM [43] was used to assess assembly completeness. The GSL17-113 assembly consists of 1051 contigs with an average coverage depth of 1219×, an N50 value of 314 799, L50 value of 7, 99.64% completeness, and a contamination score of 0.53. The GSL17-019 assembly consists of 6254 contigs with an average coverage depth of 1331×, an N50 value of 363 195, L50 value of 4, 100% genome completeness, and a contamination score of 0. For *Streptomyces* sp. GSL17-111, sequencing and genome assembly were performed by Plasmidsaurus (1850 Millrace Drive, Suite 200 Eugene, Oregon 97 403). The GSL17-111 assembly consists of five contigs with an average coverage of 62×, an N50 value of 5 998 596, L50 value of 1, 97.6% genome completeness, and a contamination score of 0.71. BGCs were initially identified using antiSMASH v.7.0 [20], and clusters of interest were manually annotated and characterized using NCBI Blast/Blast+. BGCs identified in our GSL genomes that exhibited ≥85% similarity to characterized BGCs in the MIBiG v.3 database [44] were classified as “known” in our analyses.

Comparative genomic and phylogenomic analyses of Great Salt Lake genomes

Phylogenomic analyses, including average nucleotide identity (ANI) and digital DNA–DNA hybridization (dDDH) were conducted using autoMLST [45] and the Type Strain Genome Server [46, 47]. Functional genomics were performed using the Anvi’o v.7.1 [48] suite of phylogenomic tools. Comparative genomic analyses were performed using the Anvi’o Pangenomic’s Workflow (<https://merenlab.org/2016/11/08/pangenomics-v2/>). For the soil- and marine-derived *Streptomyces* comparison, eight soil-derived *Streptomyces* spp. genomes were used (accession numbers: GCF_024752535.1, GCF_000717025.1, GCF_013307045.1, GCF_014655295.1, GCF_000010605.1, GCF_000739105.1, GCF_0018 86595.1, and GCF_014649995.1) and eight marine-derived *Streptomyces* spp. genomes were used (accession numbers: GCF_0034437 35.1, PRJNA175192, CP077658.1, GCF_014779715.1, CP054920, GCF_019303475.1, GCF_016803985.1, and GCF_004305975.1). For the closest related strain analysis for *Streptomyces* sp. GSL17-111 and *Saccharomonospora* GSL17-019, seven strains were used for each analysis (*Streptomyces* accession numbers: GCF_000718985.1, GCF_000813365.1, GCF_001493375.1, GCF_900110735.1, GCF_900111245.1, GCF_900114215.1, and GCF_900116145.1; *Saccharomonospora* accession numbers: GCF_000 231035.2, GCF_000244955.1, GCF_000383775.1, GCF_000383795.1, GCF_000430445.1, GCF_000719975.1, and GCF_002077655.1). After the Anvi’o pangenomes were constructed, amino acid sequences

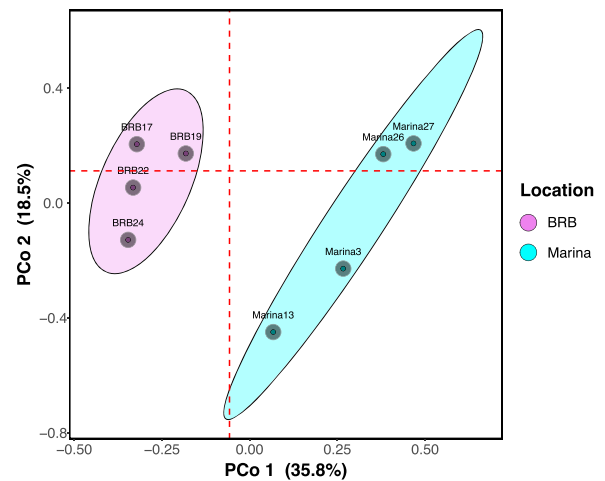


Figure 1. Bacterial community composition of sediment samples from BRB and the Marina regions, as measured by 16S rRNA gene amplicon sequencing; the PCoA plot illustrates overall community similarities and differences between four BRB and four Marina collection sites in the South Arm of GSL; dashed lines represent medians, and ellipses represent 75% confidence intervals around the samples from each region.

of predicted genes unique to each GSL genome were extracted using the “anvi-get-sequences-for-gene-clusters” call, with a minimum functional homogeneity index of 0.75, and setting both “min-” and “max-num-genomes-gene-cluster-occurs” to 1 to extract only singleton genes. This resulted in a multi-FASTA file, from which sequences associated with GSL strains were extracted using a custom Python script. Sequences were then annotated using the eggNOG-mapper v.2 [49] web server using default settings. Distributions of genes and their associated Clusters of Orthologous Genes (COGs) categories were visualized in RStudio using the following packages: dplyr v.2.3.4, readr v.2.1.4, stringr v.1.5.0, ggplot2 v.3.4.2, purrr v.1.0.1, forcats v.1.0.0, and ggbreak v.0.1.2.

Results

Microbial diversity of Great Salt Lake sediments in the south arm

We collected eight sediment samples from two different geographic regions within the South Arm of GSL—BRB and the Marina, which are only ca. 180 m apart from each other. From each region, four sediment samples were collected and assessed for the presence of bacteria using 16S rRNA gene amplicon sequencing, which resulted in the identification of 748 251 ASVs (BRB = 441 628 and Marina = 306 623). Though we expected to see little variation in the bacterial community composition between our two geographically close collection sites, we found that the regions did exhibit significant differences in their taxonomic composition and structure (Fig. 1; Bray-Curtis, $P < .05$, PERMANOVA). Upon phylogenomic analysis, we found that the ASVs comprised 53 phyla and 421 genera (Fig. 2 and Fig. S1 and Table S1 and S2). We identified Pseudomonadota (49% of total ASVs) and Bacteroidota (22% of total ASVs) as the most abundant phyla (Fig. 2A and Fig. S1A, and Table S2), and Gammaproteobacteria (34% of total ASVs) and Bacteroidia (17% of total ASVs) as the most abundant classes (Fig. 2B and Fig. S1B). Among Gammaproteobacteria, *Sulfurimonas* and *Marinobacter* spp. were the most abundant (Table S1).

Upon a closer inspection of our datasets, we found that the relative abundances of ASVs associated with specific organisms were more abundant in one of the two regions, either more

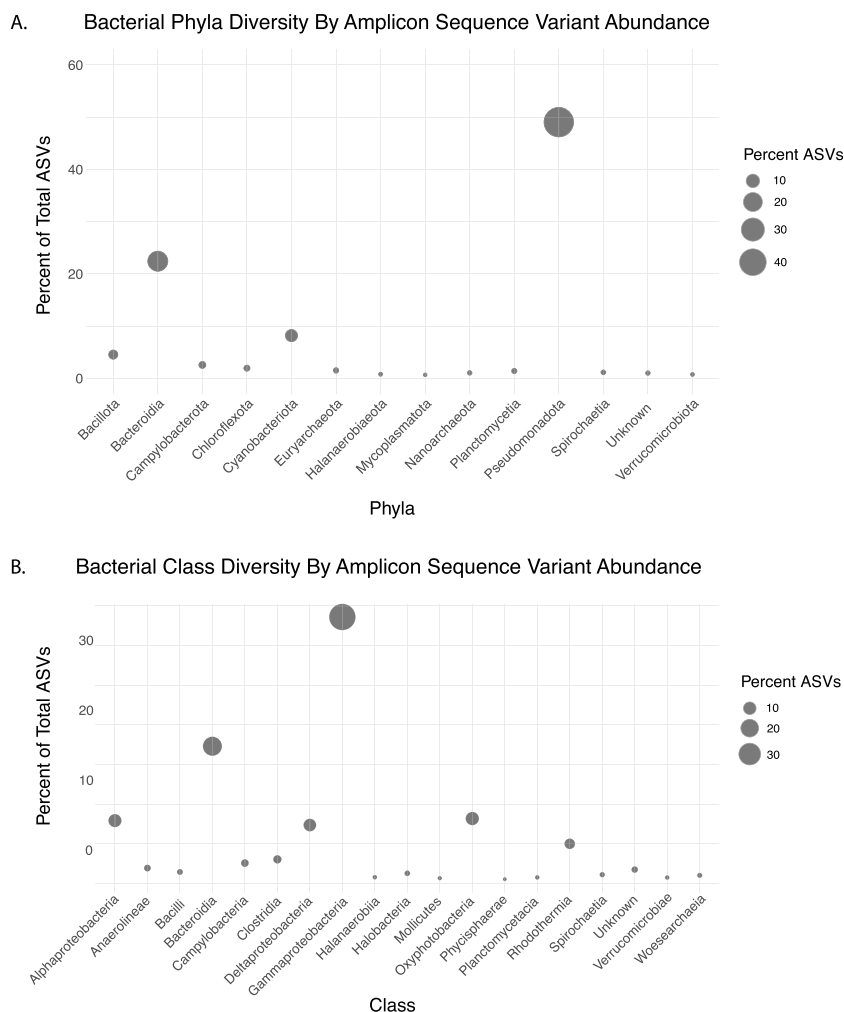


Figure 2. Diversity of GSL bacteria comprising at least 0.5% of total ASVs; (A) bubble chart illustrating diversity of 14 bacterial phyla identified by 16S ASV; (B) bubble chart illustrating diversity of 19 bacterial classes identified by 16S ASV; bacteria comprising <0.5% of total ASVs can be found in Fig. S1.

abundant in BRB sediment or more abundant in Marina sediment. For example, ASVs associated with several Planctomycetota genera were found at significantly higher levels in the Marina sediment compared to BRB sediment (*Pirellula*, $P=3.68e-13$, $P=2.9e-4$; *Rubripirellula*, $P=1.25e-14$, Wald test) (Fig. 3). In total, ASVs comprising 24 distinct phyla and 60 genera exhibited significantly different abundance between our two collection regions (Fig. 3 and Fig. S2). Interestingly, we observed differential abundances of ASVs representing two *Sulfurimonas* spp., with one found at much higher abundance in our BRB sample ($P=5.16e-11$) and the other from the Marina sediment ($P=9.22e-11$). Actinomycetota, which are common natural product producers, were found in both regions (0.31% of total ASVs), and an unknown *Saccharomonospora* species was found at a higher abundance in the Marina sediment compared to BRB ($P=5.24e-12$) (Fig. 3 and Fig. S2).

Taxonomic classification of Great Salt Lake Actinomycetota strains and comparative genomics studies

Using a culture-dependent approach, we isolated and sequenced three Actinomycetota strains, *Saccharomonospora* sp. GSL17-019, *Streptomyces* sp. GSL17-113, and *Streptomyces* sp. GSL17-111 from the sediment collected at the Marina. To determine the taxonomic position of these three strains, we conducted

whole-genome phylogenomic analyses. ANI and dDDH analyses were conducted on each genome (Table S3) [45–47], and maximum-likelihood phylogenetic trees were assembled based on multilocus sequence analysis [45] (Figs S4–S6). Interestingly, while *Streptomyces* sp. GSL17-113 was closely related to *Streptomyces albus*, *Streptomyces* sp. GSL17-111 exhibited low relatedness to the reference *Streptomyces* strains. Indeed, ANI analysis revealed only 82.5% similarity to the closest related reference strain, *Streptomyces pini*. This falls far short of the 95%–96% species threshold. Additionally, GSL17-111 exhibited a dDDH range of 53.3%–58.8% identity (dDDH, d4) compared to its closest type strain, *Streptomyces chumphonensis* KK1-2, which is below the species threshold of 70%. Taken together, this strongly suggests that GSL17-111 represents a new *Streptomyces* species. *Saccharomonospora* sp. GSL17-019 also did not exhibit significant relatedness with any known *Saccharomonospora* species, with its closest related type strain ranging from 51.8% to 57.2% identity (dDDH, d4) and 93.5% ANI (Table S3). As these dDDH and ANI values are below the species threshold, we posit that strain GSL17-019 represents a new *Saccharomonospora* species that could be specific to GSL.

To better understand the uniqueness of *Streptomyces* spp. GSL17-111 and GSL17-113, we independently carried out genomic comparisons of our GSL strains against eight soil- and eight marine-derived *Streptomyces* spp. using the Anvi'o [48]

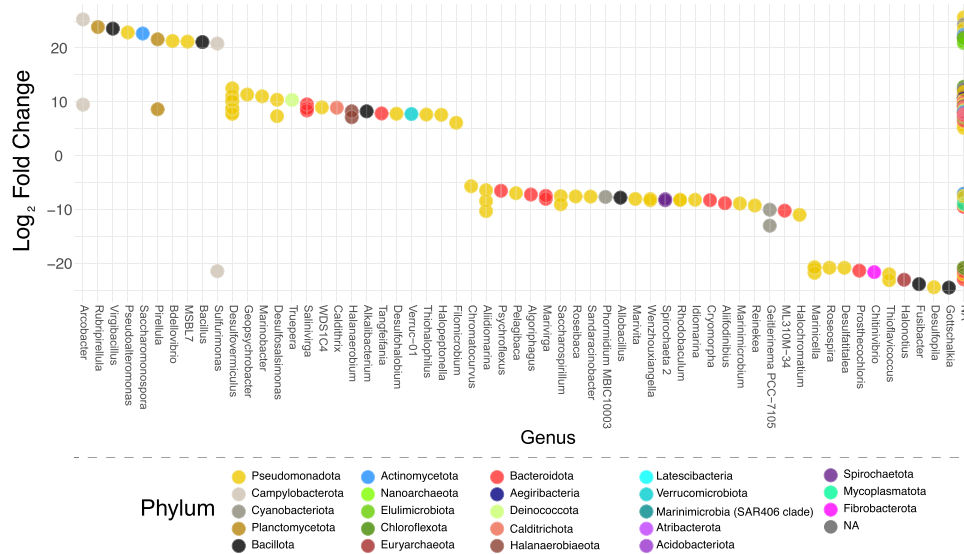


Figure 3. DESeq2 plot showing \log_2 -fold change of phyla on the y-axis and genera on the x-axis, illustrating differential abundance of microorganisms between BRB and the Marina based on ASVs; a positive \log_2 -fold-change indicates a significantly higher abundance at the Marina, while a negative \log_2 -fold-change indicates a significantly higher abundance at BRB; analysis was generated using an adjusted P -value cutoff of $<.01$; NA/NA values represent unknown phylum/unknown genus and potentially represent uncharacterized bacteria.

phylogenomic platform. Using a functional homogeneity index cutoff of 0.75, we identified and extracted all amino acid sequences with a minimum length of 40 residues unique to either GSL17-111 or GSL17-113. These sequences were then annotated using the COG database [50] (Fig. 4). When compared to marine-derived *Streptomyces* spp., we identified 799 genes unique to GSSL17-111. From the 799 identified genes, 655 sequences were annotated with a putative function using the eggNOG-mapper [49] annotation pipeline (Fig. 4A). Encoded gene products included transporters, regulators, and enzymes associated with primary and secondary metabolic processes, posttranslational modifications, and cell wall biogenesis. The most abundant proteins from each COG subcategory included those involved in signal transduction mechanisms, transcription, and secondary metabolite biosynthesis, transport, and catabolism (Fig. 4B). The majority of the annotated sequences, however, were predicted as hypothetical or proteins of unknown function, which was consistent for both *Streptomyces* spp. GSL17-111 and GSL17-113 across both comparative analyses. For the comparison of *Streptomyces* sp. GSL17-113 to marine-derived *Streptomyces* spp., 1776 unique genes were identified and of these 1565 were functionally annotated (Fig. 4A). The COG subcategories were similar to those identified in GSL17-111, although enzymes associated with transcription were far more abundant in GSL17-113 (Fig. 4B).

When compared to soil-derived *Streptomyces* spp., we identified only 89 genes unique to *Streptomyces* sp. GSL17-113. Of these, only 50 were functionally annotated using the eggNOG mapper (Fig. 4C). This is unsurprising, given GSL17-113 exhibits significant similarity to the soil-dwelling *S. albus*. Although enzymes associated with transcription and replication, recombination, and repair were relatively abundant, the majority of the predicted gene products were characterized as hypothetical or proteins of unknown function (Fig. 4D). Unlike GSL17-113, when we compared *Streptomyces* sp. GSL17-111 to soil-derived *Streptomyces* spp., we identified 1475 unique genes. Of those identified, 1261 were functionally annotated (Fig. 4C). Similar to the marine comparison with GSL17-111, the most abundant COG subcategories included proteins involved

in signal transduction mechanisms, transcription, and secondary metabolite biosynthesis, transport, and catabolism. Additional high-abundance subcategories included enzymes associated with cell wall/membrane/envelope biogenesis; replication, recombination, and repair; amino acid transport and metabolism; and inorganic ion transport and metabolism (Fig. 4D). Taken together, these results suggest that *Streptomyces* sp. GSL17-113 is more closely related to soil-derived *Streptomyces* spp., whereas *Streptomyces* sp. GSL17-111 is more closely related to marine-derived *Streptomyces* spp. However, it should be noted that *Streptomyces* sp. GSL17-111 is indeed quite unique and exhibited an overall low genomic similarity to both marine-derived and terrestrial *Streptomyces* spp. (Fig. 4 and Fig. S6 and Table S3).

To further evaluate whether *Streptomyces* sp. GSL17-111 and *Saccharomonospora* sp. GSL17-019 represent new species, we again employed a comparative genomics approach. Based on the maximum-likelihood phylogenetic trees constructed using multilocus sequence analysis, we chose seven of the most closely related strains for both GSL17-111 and GSL17-019 for additional comparisons. For GSL17-111, all seven strains were *Streptomyces* spp., while the seven closest related strains to GSL17-019 were either *Saccharomonospora* or *Actinopolyspora* spp. and included the closest related strain identified in the dDDH analysis (*Saccharomonospora iraqiensis* subsp. *paurometabolica* YIM90007, accession number GCF_000231035.2). Again, we identified and extracted all amino acid sequences with a minimum length of 40 residues unique to either GSL17-111 or GSL17-019 using a functional homogeneity index cutoff of 0.75 and used eggNOG-mapper for putative annotations (Fig. 5). From this analysis, we identified 1342 genes unique to *Streptomyces* sp. GSL17-111. Of these, 1139 were functionally annotated (Fig. 5A). Aside from the uncharacterized/hypothetical categories, genes encoding proteins associated with metabolism represented the most abundant COG category. This included enzymes involved with secondary metabolite biosynthesis, inorganic ion transport, and amino acid transport and metabolism. Enzymes associated with transcription represent the most abundant COG subcategory and proteins involved with cellular processes, and signal transduction

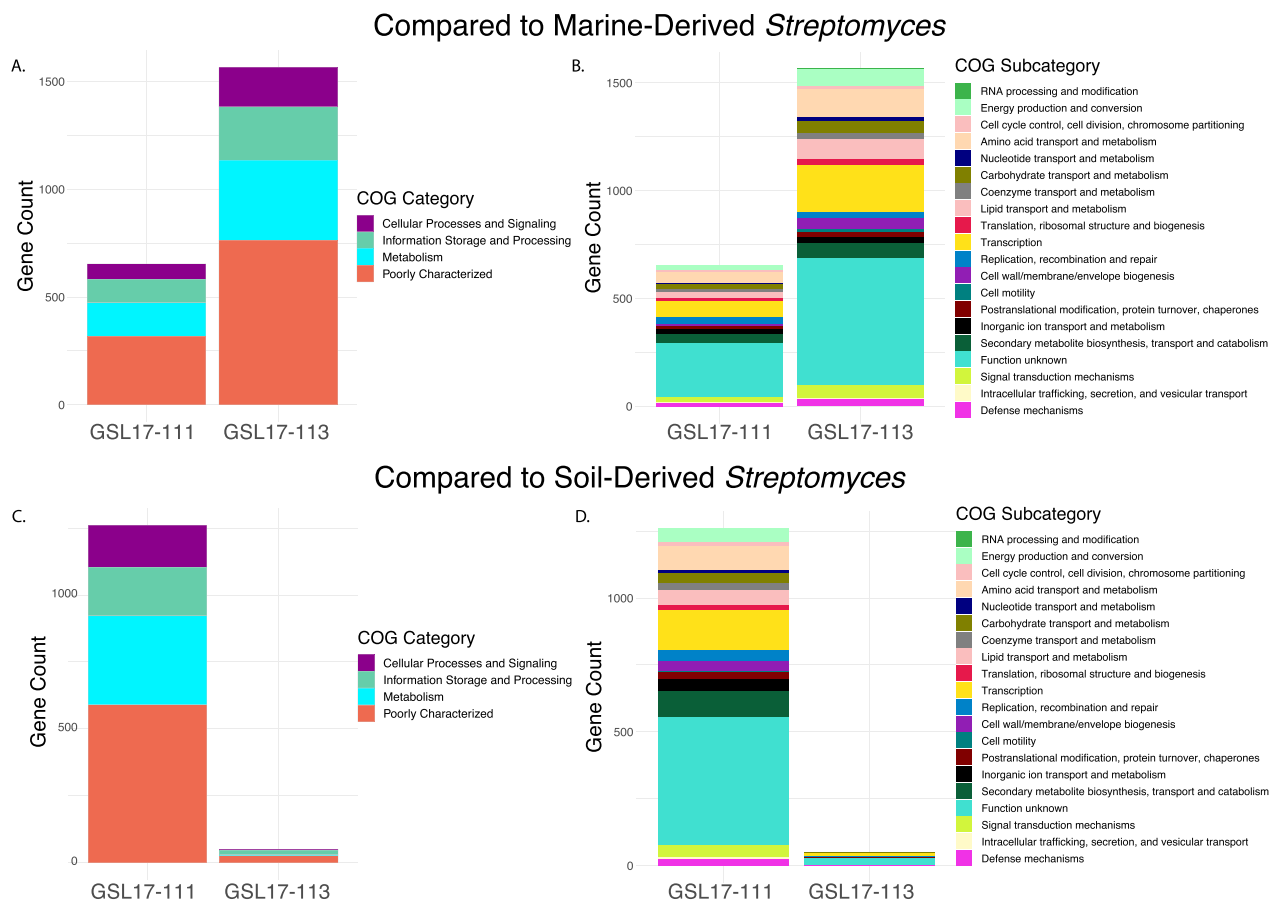


Figure 4. COG annotations of GSL *Streptomyces* with no homologs in soil- or marine-derived *Streptomyces* spp.; (A and B) genomic comparison of GSL *Streptomyces* spp. to marine-derived *Streptomyces* spp.; (A) functional classification of the 655 proteins encoded in *Streptomyces* sp. GSL17-111 and 1565 proteins identified in *Streptomyces* sp. GSL17-113; (B) distribution and classification by COG subcategories; (C and D) genomic comparison of GSL *Streptomyces* spp. to soil-derived *Streptomyces* spp.; (C) functional classification of the 1261 proteins encoded in *Streptomyces* sp. GSL17-111 and 50 proteins encoded in *Streptomyces* sp. GSL17-113; (D) distribution and classification by COG subcategories; a minimum functional homogeneity index cutoff of 0.75 was used to identify homologs in soil- or marine-derived *Streptomyces* spp.

mechanisms, cell wall/membrane/envelope biogenesis, and defense mechanism represent the most abundant subcategories (Fig. 5B). Proteins associated with cell cycle control, cell division and chromosomal partitioning, cell motility, carbohydrate transport and metabolism, posttranslational modification, protein turnover, and chaperones were also identified, illustrating a range of enzymes involved with many important cellular functions that may be necessary to thrive in the extreme environment of GSL.

When compared to the closest related reference *Saccharomonospora* spp., we identified 636 genes unique to *Saccharomonospora* sp. GSL17-019. Of these, 449 sequences were annotated with a putative COG function (Fig. 5A). Aside from genes encoding hypotheticals and proteins with unknown function, the most abundant category was associated with metabolism and included proteins involved with secondary metabolite biosynthesis, coenzyme transport and metabolism, and lipid transport and metabolism. Genes encoding enzymes associated with transcription, replication, recombination, and repair, as well as proteins involved with cell wall/membrane/envelope biogenesis and signal transduction mechanisms were also identified (Fig. 5B). Given the large number of unique genes, even when compared to the closest related reference strains, these data further support the uniqueness of *Streptomyces* sp. GSL17-111 and *Saccharomonospora* sp. GSL17-019.

Natural product potential of Great Salt Lake bacteria

We investigated the potential for specialized metabolite production by our sequenced isolates through the identification and annotation of natural product BGCs encoded in each genome (Fig. 6 and Tables S4–S6). We identified 20 putative BGCs in *Saccharomonospora* sp. GSL17-019 (Table S4) and 27 putative BGCs in *Streptomyces* sp. GSL17-113 (Table S5), representing 19 different BGC classes. These classes included more common polyketide synthase and nonribosomal peptide synthetase-containing clusters as well as less common ribosomally synthesized and post-translationally modified peptide clusters predicted to produce lassopeptide [51] and ranthipeptide [51, 52] natural products (Fig. 6). From *Streptomyces* sp. GSL17-111, we identified 22 BGCs (Fig. 6 and Table S6). Of the combined BGCs identified, only 11 strongly correlated ($\geq 85\%$ predicted similarity) to characterized clusters in public databases that have been associated with specific natural products (Fig. 7A). To better assess the natural product potential of our GSL isolates, *Streptomyces* sp. GSL17-113 was subjected to small-scale cultivation studies. From the culture extract, we identified tambjamine BE-18591 (Fig. 7B and C and Figs S7 and S8 and Table S7). Tambjamine BE-18591 was originally reported from *Streptomyces* sp. BA18591, a plant-derived isolate collected in Japan [53, 54]. The tambjamines possess antimicrobial

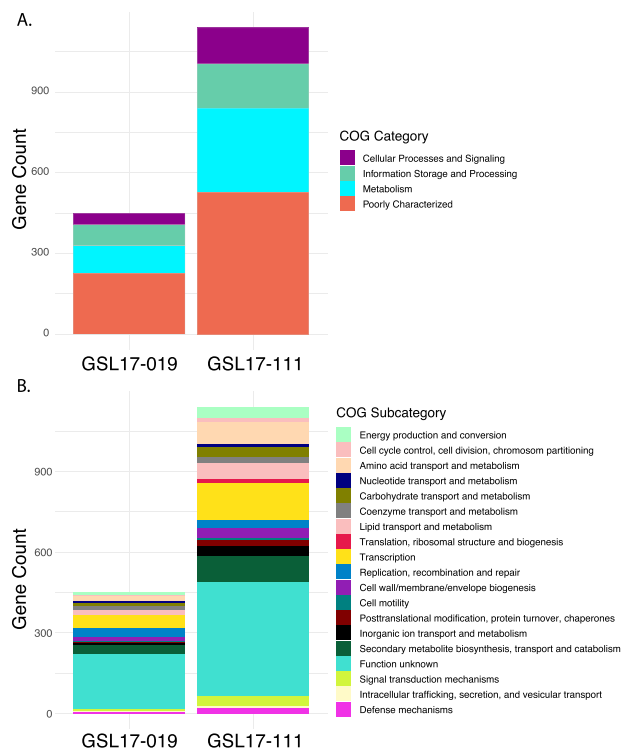


Figure 5. COG annotations encoded by GSL *Streptomyces* found to have no homologs in their closest-related reference strains; (A) functional classification of the 1139 proteins encoded in *Streptomyces* sp. GSL17-111 and 449 proteins identified in *Saccharomonospora* sp. GSL17-019; (B) distribution and classification by COG subcategories; taxonomic position was determined using the autoMLST [45] phylogenomic pipeline and a phylogenetic tree was constructed for each strain; based on the maximum-likelihood trees (Figs S4 and S6), the closest seven strains from each analysis were used for comparative genomics.

and cytotoxic activities [53, 55], and tambjamine BE-18591 has been reported to possess antimicrobial activity against both fungi and bacteria, including *Candida albicans*, *Malassezia furfur*, *E. coli*, and *Staphylococcus aureus* [53, 55]. Further, tambjamine BE-18591 displays broad antitumor activity against leukemia, melanoma, colorectal, and glioblastoma cell lines [55], as well as inhibition of immunoproliferation and gastritis in rabbits [56]. Our genomic analysis also confirmed the presence of the tambjamine BE-18591 BGC in *Streptomyces* sp. GSL17-113 (Fig. 7B and Tables S5 and S8), emphasizing the importance and utility of genome mining when prioritizing strains for downstream fermentation studies and aiding in the dereplication process. Though our initial fermentation study resulted in the isolation of a known compound, >80% of the identified gene clusters (>50 BGCs) annotated in our three datasets had weak to no similarity with characterized BGCs, indicating that the corresponding biosynthetic machinery could be synthesizing new natural product scaffolds.

Discussion

In this study, we have highlighted the diversity of GSL's unique microbiome, including the genomic characterization of a potentially new *Saccharomonospora* and *Streptomyces* species, and emphasized the natural product potential of GSL bacteria to produce bioactive compounds. Previously, we isolated the bonnevillamides from a GSL bacterium, *Streptomyces* sp. GSL-6B. The bonnevillamides are linear heptapeptides containing a distinctive 3-(3,5-dichloro-4-hydroxyphenyl)-2-methoxypropenoic acid moiety, and bonnevillamide A harbors an additional unprecedented

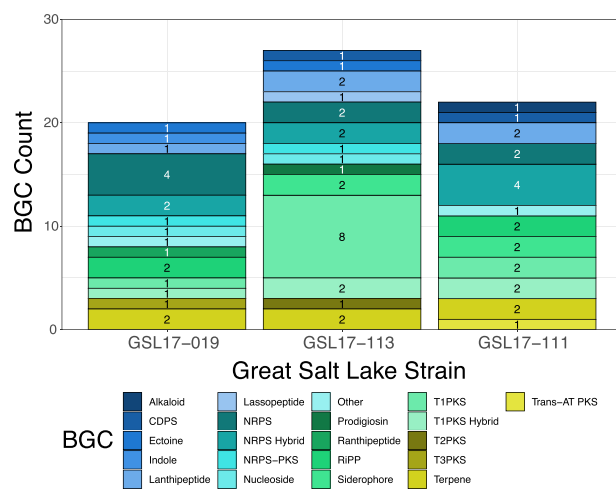


Figure 6. Number and type of BGCs identified in *Saccharomonospora* sp. GSL17-019, *Streptomyces* sp. GSL17-113, and *Streptomyces* sp. GSL17-111; abbreviations: cyclodipeptide synthase, CDPS; nonribosomal peptide synthetase, NRPS; polyketide synthase, PKS; ribosomally synthesized and post-translationally modified peptide, RIPP; acyltransferase, AT; Type I polyketide synthase, T1PKS; Type II polyketide synthase, T2PKS; Type III polyketide synthase, T3PKS.

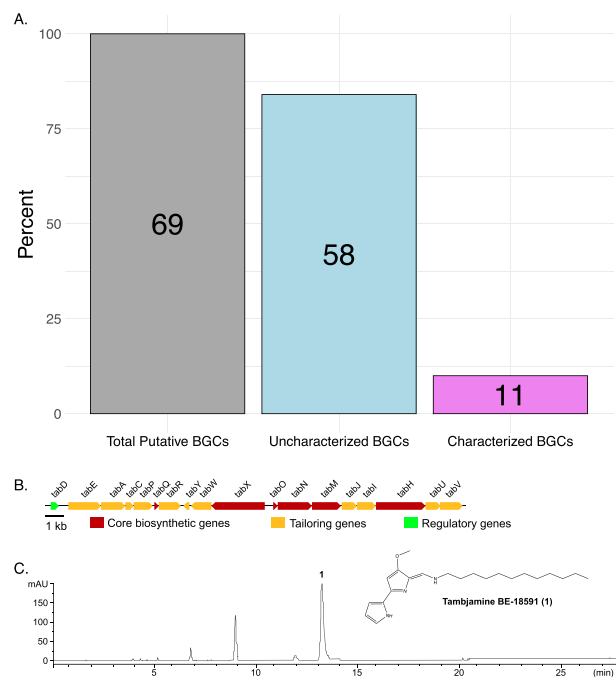


Figure 7. Natural product potential of *Saccharomonospora* sp. GSL17-019, *Streptomyces* sp. GSL17-113, and *Streptomyces* sp. GSL17-111; (A) a total of 69 BGCs were identified from the three GSL strains; of those, 11 were strongly associated with characterized BGCs deposited in MIBiG v.3 ($\geq 85\%$ predicted similarity) and included BGCs responsible for ectoine, coprisamide, xenematide, desferrirosamine, tambjamine, geosmin, isorenieratene, and piericidine biosynthesis; the remaining 58 BGCs were not strongly associated with known gene clusters; BGC-associated compounds were identified through gene cluster comparison using both antiSMASH v.7.0 and manual annotation; (B) organization of the tambjamine BGC identified in *Streptomyces* sp. GSL17-113; (C) high-performance liquid chromatography analysis (390 nm) of GSL17-113 and identification of tambjamine BE-18591 (1).

4-methyl-azetidino-2-carboxylic acid methyl ester chemical motif [57, 58]. We also reported on the isolation and structure elucidation of the salinipeptins, ribosomally synthesized and post-translationally modified peptides containing rare D-amino

acids, a highly functionalized N-terminus, and a C-terminal aminovinyl-cysteine residue [59]. Importantly, salinipeptin A displayed moderate activity against Group A *Streptococcus pyogenes* as well as glioblastoma and colon cancer cell lines. Although tambjamine BE-18591 is a known natural product, its isolation from *Streptomyces* sp. GSL17-113 represents an additional bioactive natural product isolated from a GSL bacteria. Importantly, when we queried the tambjamine BE-18591 BGC through the NCBI database, only three *Streptomyces* spp. displayed query coverage above 62% (Table S9), suggesting that tambjamine BE-18591 is an uncommon natural product. When taken with the comparative and phylogenomic analyses, our research supports GSL as an underexplored environment harboring uncharacterized bacteria potentially rich in bioactive natural products. Additional fermentation studies with both *Streptomyces* sp. GSL17-111 and *Saccharomonospora* sp. GSL17-019 are currently underway to further characterize the natural product potential of these strains.

In parallel with the isolation and characterization of novel natural products, we intend to explore potential resistance mechanisms utilized by GSL microbes. As GSL is a terminal lake and serves as the endpoint for wastewater treatment runoff, we anticipate it to be a significant source of antibiotic-resistant pathogenic microbes, especially as the environment can serve as a reservoir for antibiotic-resistant bacteria [60–65]. Furthermore, from an ecological standpoint, as GSL water levels fluctuate, it is important that we continue to investigate the Lake's microbiome and how the changing water levels, drought, and inflow affect microbial diversity. From this study, we not only provide an initial insight into the diversity of GSL sediment-derived bacteria but also provide the first genomic characterization of Actinomycetota isolated from GSL. This is of particular interest to our group, as Actinomycetota are the dominant producers of therapeutic compounds essential for human health [63, 66–68].

Bioprospecting of GSL bacteria led to the identification of the bonnevillamides, salinipeptins, and in this study, the antitumor and antibiotic compound, tambjamine BE-18591. Given the numerous unidentified BGCs (Fig. 7A and Tables S4–S6), it is evident that GSL represents a largely untapped resource of natural products. Though our primary interest is in characterizing known natural product producing microbes, the high number of unknown phylum/unknown genus values we observed in our sediment samples (Fig. 3) suggests that GSL is home to many as-of-yet characterized bacteria. Though several common phyla identified in this study were similar to those identified in a similar report investigating marine sediment-derived bacteria [69], there were several differences in the microbial composition of GSL sediment. The most abundant phyla identified from marine sediment, in the order of abundance, included Pseudomonadota, Cyanobacteriota, unclassified species, Bacteroidia, and Desulfobacteriota [69]. In our study, the most abundant phyla identified included Pseudomonadota, Bacteroidia, Cyanobacteriota, Bacillota, and Campylobacterota (Fig. 2A and Fig. S1A and Table S2), indicating a differential microbial composition compared to marine sediment and further supporting GSL is unique compared to marine or terrestrial environments. GSL's uniqueness is additionally emphasized by the discovery and genomic characterization of the potentially new *Streptomyces* sp. GSL17-111 and *Saccharomonospora* sp. GSL17-019 isolated from the GSL sediment. Taken together, our results give an overview of GSL's large microbial and chemical diversity as well as differences in microbial populations between geographically close regions and revealed GSL's potential to harbor new species of natural product-producing bacteria.

Acknowledgements

We thank Abby Scott, University of Utah, for assistance with processing Great Salt Lake sediment and helpful discussions and would like to thank the three reviewers for their insightful comments and suggestions.

Author contributions

Elijah R. Bring Horvath (Conceptualization [equal], Bioinformatics [equal], Culturing GSL strains [equal], Data analysis [equal], writing—original draft [lead], writing—review & editing [equal]), William J. Brazelton (Conceptualization [in part], Bioinformatics [equal], Bacterial isolation, sequencing, and genome assembly [equal], Data analysis [equal] writing—original draft [equal], writing—review & editing [equal]), Min Cheol Kim (Culturing GSL strains [equal], Isolation of natural products and NMR [lead], Data analysis [equal]), Reiko Cullum (Culturing GSL strains [equal], Mass spectrometry analysis [lead], Data analysis [supporting]), Matthew A. Mulvey (Data analysis [supporting], writing—review & editing [supporting]), William Fenical (Conceptualization [equal], Data analysis [supporting], writing—review & editing [supporting]), and Jaclyn M. Winter (Conceptualization [lead], project administration [lead], Bacterial isolation, sequencing, genome assembly [equal], Data analysis [equal], writing—original draft [equal], writing—review & editing [lead]).

Supplementary material

Supplementary material is available at ISME Communications online.

Conflicts of interest

None declared.

Funding

This work was supported by a 3iii Graduate Research Fellowship to E.R.B.H.; by the National Institutes of Health (1R01AI155694 to W.F. and J.M.W.); by the University of Utah Research Foundation to J.M.W.; by the Margolis Foundation to W.J.B., M.A.M., and J.M.W.; and in part by the Department of Defense award (W81XWH-22-1-0800 and SC210103) to W.J.B., J.M.W., and M.A.M.

Data availability

All relevant code and data, including genes identified in the comparative genomic analysis, are available at <https://github.com/ERBringHorvath/Great-Salt-Lake-Microbial-Diversity>. 16S amplicon sequences are available under BioProject accession number PRJNA975952. The genomes of *Streptomyces* sp. GSL17-113 and *Saccharomonospora* sp. GSL17-019 are deposited under BioProject accession number PRJNA1066849. The genome of *Streptomyces* sp. GSL17-111 is deposited under BioProject accession number PRJNA1077303. The tambjamine BE-18591 biosynthetic cluster from *Streptomyces* sp. GSL17-113 is deposited under the GenBank accession number PP179500.

Ethics approval statement

The research described in this study was undertaken with approval from the University of Utah Institutional Biosafety Committee (266-20), and J.M.W. was issued a sediment sampling permit through the State of Utah Division of Forestry, Fire & State Lands (FFSL # 420-00313).

References

- Baxter BK. Great Salt Lake microbiology: a historical perspective. *Int Microbiol* 2018;**21**:79–95. <https://doi.org/10.1007/s10123-018-0008-z>
- Belovsky GE, Stephens D, Perschon C et al. The Great Salt Lake ecosystem (Utah, USA): long term data and a structural equation approach. *Ecosphere* 2011;**2**:art33–40. <https://doi.org/10.1890/ES10-00091.1>
- Carlowicz M. Record Low for Great Salt Lake. NASA Earth Observatory, 2021. <https://earthobservatory.nasa.gov/images/148700/record-low-for-great-salt-lake> (31 October 2023, date last accessed).
- Yeung P. Salt lake confronts a future without a lake. In: *Bloomberg*. 2022. https://www.bloomberg.com/news/features/2022-07-08/drought-leaves-salt-lake-city-with-a-looming-water-crisis?utm_source=website&utm_medium=share&utm_campaign=copy (31 October 2023, date last accessed).
- Merck MF, Tarboton DG. The salinity of the Great Salt Lake and its deep brine layer. *Water* 2023;**15**:1488. <https://doi.org/10.3390/w15081488>
- Post FJ. The microbial ecology of the Great Salt Lake. *Microb Ecol* 1977;**3**:143–65. <https://doi.org/10.1007/BF02010403>
- Shope CL, Angerth CE. Calculating salt loads to Great Salt Lake and the associated uncertainties for water year 2013; updating a 48 year old standard. *Sci Total Environ* 2015;**536**:391–405. <https://doi.org/10.1016/j.scitotenv.2015.07.015>
- Almeida-Dalmet S, Sikaroodi M, Gillevet PM et al. Temporal study of the microbial diversity of the north arm of Great Salt Lake, Utah, U.S. *Microorganisms* 2015;**3**:310–26. <https://doi.org/10.3390/microorganisms3030310>
- Baxter BK, Zalar P. The extremophiles of Great Salt Lake: Complex microbiology in a dynamic hypersaline ecosystem. In: Seckbach J, Rampelotto P. (eds.), *Model Ecosystems in Extreme Environments*. Elsevier, Amsterdam, 2019, 57–99.
- Lindsay MR, Anderson C, Fox N et al. Microbialite response to an anthropogenic salinity gradient in Great Salt Lake, Utah. *Geobiology* 2017;**15**:131–45. <https://doi.org/10.1111/gbi.12201>
- ZoBell CE. Direct microscopic evidence of an autochthonous bacterial flora in Great Salt Lake. *Ecology* 1937;**18**:453–8. <https://doi.org/10.2307/1931221>
- ZoBell CE, Anderson DQ, Smith WW. The bacteriostatic and bactericidal action of Great Salt Lake water. *J Bacteriol* 1937;**33**:253–62. <https://doi.org/10.1128/jb.33.3.253-262.1937>
- Ekrami J, Nemati Mansour S, Mosaferi M et al. Environmental impact assessment of salt harvesting from the salt lakes. *J Environ Health Sci Eng* 2021;**19**:365–77. <https://doi.org/10.1007/s40201-020-00609-2>
- Shuaib M, Azam N, Bahadur S et al. Variation and succession of microbial communities under the conditions of persistent heavy metal and their survival mechanism. *Microb Pathog* 2021;**150**:104713. <https://doi.org/10.1016/j.micpath.2020.104713>
- 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. <https://doi.org/10.1021/acs.jnatprod.9b01285>
- Dias DA, Urban S, Roessner U. A historical overview of natural products in drug discovery. *Meta* 2012;**2**:303–36. <https://doi.org/10.3390/metabo2020303>
- Hutchings MI, Truman AW, Wilkinson B. Antibiotics: past, present and future. *Curr Opin Microbiol* 2019;**51**:72–80. <https://doi.org/10.1016/j.mib.2019.10.008>
- Darby EM, Trampari E, Siasat P et al. Molecular mechanisms of antibiotic resistance revisited. *Nat Rev Microbiol* 2022;**21**:280–95. <https://doi.org/10.1038/s41579-022-00820-y>
- Langford BJ, Soucy J-PR, Leung V et al. Antibiotic resistance associated with the covid-19 pandemic: a systematic review and meta-analysis. *Clin Microbiol Infect* 2023;**29**:302–9. <https://doi.org/10.1016/j.cmi.2022.12.006>
- Blin K, Shaw S, Steinke K et al. Antismash 5.0: updates to the secondary metabolite genome mining pipeline. *Nucleic Acids Res* 2019;**47**:W81–7. <https://doi.org/10.1093/nar/gkz310>
- Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with deseq2. *Genome Biol* 2014;**15**:550. <https://doi.org/10.1186/s13059-014-0550-8>
- McMurdie PJ, Holmes S. Phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. *PLoS One* 2013;**8**:e61217. <https://doi.org/10.1371/journal.pone.0061217>
- Biermann F, Helfrich EJM. Hidden treasures: microbial natural product biosynthesis off the beaten path. *mSystems* 2021;**6**:e0084621. <https://doi.org/10.1128/mSystems.00846-21>
- Daniels LL. On the flora of Great Salt Lake. *Am Nat* 1917;**51**:499–506. <https://doi.org/10.1086/279627>
- Packard AS. The sea-weeds of salt lake. *Am Nat* 1879;**13**:701–3.
- Vorhies CT. Notes of the fauna of the Great Salt Lake. *Am Nat* 1917;**51**:494–9. <https://doi.org/10.1086/279626>
- Meuser JE, Baxter BK, Spear JR et al. Contrasting patterns of community assembly in the stratified water column of Great Salt Lake. *Utah Microb Ecol* 2013;**66**:268–80. <https://doi.org/10.1007/s00248-013-0180-9>
- Parnell JJ, Rompato G, LCT L et al. Functional biogeography as evidence of gene transfer in hypersaline microbial communities. *PLoS One* 2010;**5**:e12919. <https://doi.org/10.1371/journal.pone.0012919>
- Tazi L, Breakwell DP, Harker AR et al. Life in extreme environments: microbial diversity in Great Salt Lake, Utah. *Extremophiles* 2014;**18**:525–35. <https://doi.org/10.1007/s00792-014-0637-x>
- Parnell AC, Inger R, Bearhop S et al. Source partitioning using stable isotopes: coping with too much variation. *PLoS One* 2010;**5**:e9672. <https://doi.org/10.1371/journal.pone.0009672>
- Camacho C, Coulouris G, Avagyan V et al. Blast+: architecture and applications. *BMC Bioinform* 2009;**10**:421. <https://doi.org/10.1186/1471-2105-10-421>
- Kozich JJ, Westcott SL, Baxter NT et al. Development of a dual-index sequencing strategy and curation pipeline for analyzing amplicon sequence data on the MiSeq Illumina sequencing platform. *Appl Environ Microbiol* 2013;**79**:5112–20. <https://doi.org/10.1128/AEM.01043-13>
- Callahan BJ, McMurdie PJ, Rosen MJ et al. Dada2: high-resolution sample inference from Illumina amplicon data. *Nat Methods* 2016;**13**:581–3. <https://doi.org/10.1038/nmeth.3869>
- Marcel M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet* 2011;**17**:10–2. <https://doi.org/10.14806/ej.17.1.200>
- Quast C, Pruesse E, Yilmaz P et al. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res* 2013;**41**:D590–6. <https://doi.org/10.1093/nar/gks1219>
- Yilmaz P, Parfrey LW, Yarza P et al. The SILVA and “all-species living tree project (ltp)” taxonomic frameworks. *Nucleic Acids Res* 2014;**42**:D643–8. <https://doi.org/10.1093/nar/gkt1209>

37. RStudio Team. *Rstudio: Integrated Development Environment for R*. Posit Software, PBC, Boston, MA, 2023. <https://www.rstudio.com/>
38. Xu S, Chen M, Feng T et al. Use ggbreak to effectively utilize plotting space to deal with large datasets and outliers. *Front Genet* 2021;**12**:774846. <https://doi.org/10.3389/fgene.2021.774846>
39. Yu G, Smith D, Zhu H et al. Ggtree: an r package for visualization and annotation of phylogenetic trees with their covariates and other associated data. *Methods Eco Evol* 2017;**8**:28–36. <https://doi.org/10.1111/2041-210X.12628>
40. Bushnell B, Rood J, Singer E. Bbmerge - accurate paired shotgun read merging via overlap. *PLoS One* 2017;**12**:e0185056. <https://doi.org/10.1371/journal.pone.0185056>
41. Thornton CN, Tanner WD, VanDerslice JA et al. Localized effect of treated wastewater effluent on the resistome of an urban watershed. *Gigascience* 2020;**9**:giaa125. <https://doi.org/10.1093/gigascience/giaa125>
42. Bankevich A, Nurk S, Antipov D et al. Spades: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol* 2012;**19**:455–77. <https://doi.org/10.1089/cmb.2012.0021>
43. Parks DH, Imelfort M, Skennerton CT et al. Assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res* 2014;**25**:1043–55. <https://doi.org/10.1101/gr.186072.114>
44. Terlouw BR, Blin K, Navarro-Munoz JC et al. Mibig 3.0: a community-driven effort to annotate experimentally validated biosynthetic gene clusters. *Nucleic Acids Res* 2023;**51**:D603–10. <https://doi.org/10.1093/nar/gkac1049>
45. Alanjary M, Steinke K, Ziemert N. Automlst: an automated web server for generating multi-locus species trees highlighting natural product potential. *Nucleic Acids Res* 2019;**47**:W276–82. <https://doi.org/10.1093/nar/gkz282>
46. Meier-Kolthoff JP, Goker M. Tygs is an automated high-throughput platform for state-of-the-art genome-based taxonomy. *Nat Commun* 2019;**10**:2182. <https://doi.org/10.1038/s41467-019-10210-3>
47. Meier-Kolthoff JP, Auch AF, Klenk HP et al. Genome sequence-based species delimitation with confidence intervals and improved distance functions. *BMC Bioinform* 2013;**14**:60. <https://doi.org/10.1186/1471-2105-14-60>
48. Eren AM, Kiefl E, Shaiber A et al. Community-led, integrated, reproducible multi-omics with anvio. *Nat Microbiol* 2020;**6**:3–6. <https://doi.org/10.1038/s41564-020-00834-3>
49. Cantalapiedra CP, Hernandez-Plaza A, Letunic I et al. EggNOG-mapper v2: functional annotation, orthology assignments, and domain prediction at the metagenomic scale. *Mol Biol Evol* 2021;**38**:5825–9. <https://doi.org/10.1093/molbev/msab293>
50. Galperin MY, Wolf YI, Makarova KS et al. Cog database update: focus on microbial diversity, model organisms, and widespread pathogens. *Nucleic Acids Res* 2021;**49**:D274–81. <https://doi.org/10.1093/nar/gkaa1018>
51. Montalban-Lopez M, Scott TA, Ramesh S et al. New developments in ripp discovery, enzymology and engineering. *Nat Prod Rep* 2021;**38**:130–239. <https://doi.org/10.1039/D0NP00027B>
52. Chen Y, Wang J, Li G et al. Current advancements in sactipeptide natural products. *Front Chem* 2021;**9**:595991. <https://doi.org/10.3389/fchem.2021.595991>
53. Kojiri KNS, Suzuki H, Okura A et al. A new antitumor substance, be-18591, produced by a streptomycete. I. Fermentation, isolation, physicochemical and biological properties. *J Antibiot* 1993;**12**:1799–803
54. Nakajima SKK, Suda H. A new antitumor substance, be-18591, produced by a streptomycete ii. Structure determination. *J Antibiot* 1993;**46**:1894–6.
55. Pinkerton DM, Banwell MG, Garson MJ et al. Antimicrobial and cytotoxic activities of synthetically derived tambjamins c and e–j, be-18591, and a related alkaloid from the marine bacterium *Pseudoalteromonas tunicata*. *Chem Biodivers* 2010;**7**:1311–24. <https://doi.org/10.1002/cbdv.201000030>
56. Tanigaki K, Sato T, Tanaka Y et al. Be-18591 as a new H⁺/Cl⁻ symport ionophore that inhibits immunoproliferation and gastritis. *FEBS Lett* 2002;**524**:37–42. [https://doi.org/10.1016/S0014-5793\(02\)02996-4](https://doi.org/10.1016/S0014-5793(02)02996-4)
57. Wu G, Nielson JR, Peterson RT et al. Bonnevillamides, linear heptapeptides isolated from a Great Salt Lake-derived streptomycetes sp. *Mar Drugs* 2017;**15**:195. <https://doi.org/10.3390/md15070195>
58. Shin YH, Ban YH, Shin J et al. Azetidine-bearing non-ribosomal peptides, bonnevillamides d and e, isolated from a carrion beetle-associated actinomycete. *J Org Chem* 2021;**86**:11149–59. <https://doi.org/10.1021/acs.joc.1c00360>
59. Shang Z, Winter JM, Kauffman CA et al. Salinipeptins: integrated genomic and chemical approaches reveal unusual d-amino acid-containing ribosomally synthesized and post-translationally modified peptides (ripps) from a Great Salt Lake *Streptomyces* sp. *ACS Chem Biol* 2019;**14**:415–25. <https://doi.org/10.1021/acscchembio.8b01058>
60. Bleichenbacher S, Stevens MJA, Zurfluh K et al. Environmental dissemination of carbapenemase-producing enterobacteriaceae in rivers in Switzerland. *Environ Pollut* 2020;**265**:115081. <https://doi.org/10.1016/j.envpol.2020.115081>
61. Antunes P, Machado J, Sousa JC et al. Dissemination of sulfonamide resistance genes (sul1, sul2, and sul3) in Portuguese salmonella enterica strains and relation with integrons. *Antimicrob Agents Chemother* 2005;**49**:836–9. <https://doi.org/10.1128/AAC.49.2.836-839.2005>
62. Centers for Disease Control and Prevention US Department of Health and Human Services. *Antibiotic Resistance Threats in the United States*. CDC Atlanta, GA, 2019.
63. D'Costa VM, King CE, Kalan L et al. Antibiotic resistance is ancient. *Nature* 2011;**477**:457–61. <https://doi.org/10.1038/nature10388>
64. de Vries LE, Valles Y, Agero Y et al. The gut as reservoir of antibiotic resistance: microbial diversity of tetracycline resistance in mother and infant. *PLoS One* 2011;**6**:e21644. <https://doi.org/10.1371/journal.pone.0021644>
65. Allen HK, Donato J, Wang HH et al. Call of the wild: antibiotic resistance genes in natural environments. *Nat Rev* 2010;**8**:251–9. <https://doi.org/10.1038/nrmicro2312>
66. Wainwright M. Streptomycin: discovery and resultant controversy. *Hist Phil Life Sci* 1991;**13**:97–124.
67. Lenz KD, Klosterman KE, Mukundan H et al. Macrolides: from toxins to therapeutics. *Toxins* 2021;**13**:347. <https://doi.org/10.3390/toxins13050347>
68. Maiese WM, Lechevalier MP, Lechevalier HA et al. Calicheamincins, a novel family of antitumor antibiotics: taxonomy, fermentation and biological properties. *J Antibiot* 1989;**42**:558–63. <https://doi.org/10.7164/antibiotics.42.558>
69. Chase AB, Bogdanov A, Demko AM et al. Biogeographic patterns of biosynthetic potential and specialized metabolites in marine sediments. *ISME J* 2023;**17**:976–83. <https://doi.org/10.1038/s41396-023-01410-3>