Applied Microbiology and Biotechnology (2025) 109:50
https://doi.org/10.1007/s00253-025-13427-z

APPLIED MICROBIAL AND CELL PHYSIOLOGY

# Metabolomic and genomic insights into *Micromonospora carbonacea* subsp. *caeruleus* for anti-colorectal compound

**Tepakorn Kongsaya[1,2] · Nuttaporn Emthomya[1,2] · Chananan Ngamcharungchit[1,2] · Aiyada Aroonsri[3] · Umaporn Uawisetwathana[3] · Thapanee Pruksatrakul[3] · Jirayut Euanorasetr[4] · Bungonsiri Intra[1,2]**

Received: 23 July 2024 / Revised: 18 December 2024 / Accepted: 3 February 2025


## Abstract

Cancer is a predominant contributor to global morbidity and mortality, affecting populations worldwide. Marine *Micromonospora* species have been identified as significant sources of anticancer compounds. This work aimed to perform a polyphasic approach of isolated strain and conduct comparative metabolomic and genomic analyses to identify compounds with anticancer activity. The study utilized a polyphasic approach on isolated strains for anticancer compound identification. Taxonomic analysis of strain 2MTK254 revealed unique pigment and fatty acid patterns, designating it as a novel *Micromonospora carbonacea* subsp. *caeruleus*. Its crude extract displayed significant anti-colorectal activity (66.03% inhibition). Molecular network analysis classified metabolites into eight classes, highlighting a polycyclic tetramate macrolactams (PTMs) compound (P1, $C_{29}H_{38}N_2O_4$) with 99.31% inhibitory activity against HCT-116 cell lines ($IC_{50}$ at 0.125 µM). Genome analysis identified 32 biosynthetic gene clusters (BGCs), including unique PTMs BGCs (83% similarity) linked to the P1 compound. Thus, *M. carbonacea* subsp. *caeruleus* 2MTK254 holds promise as a source of novel PTMs with anti-colorectal cancer potential.

### Key points

- *A novel strain of Micromonospora carbonacea subsp. caeruleus 2MTK254 was isolated in Thailand*
- *A new polycyclic tetramate macrolactam (PTM) with anticancer activity was identified in 2MTK254*
- *The genome of 2MTK254 has unique secondary metabolite gene clusters*

**Keywords** Anticancer compound · Genome mining · Marine *Micromonospora* · Molecular networking · Polyphasic taxonomy

## Introduction

Cancer is a life-threatening condition that significantly affects human health in the twenty-first century (Baindara and Mandal 2020). In 2020, there were approximately 19.9 million new cancer cases and 9.7 million cancer-related deaths reported. This incidence is forecast to rise to 35 million deaths worldwide by 2050 (Ferlay et al. 2024). Colorectal cancer ranks as the third most prevalent cancer in terms of incidence and stands as the second leading cause of cancer-related mortality worldwide (Sung et al. 2021). Chemotherapy is a widely utilized treatment for various types of cancer. However, the majority of patients experience drug resistance due to genetic mutations and epigenetic modifications (Guimarães et al. 2020). Additionally, chemotherapy is associated with side effects caused by cytotoxicity affecting DNA and protein expression in both normal and cancer cells, leading to off-target problems such as hair loss,

✉ Bungonsiri Intra
Bungonsiri.int@mahidol.edu

1 Department of Biotechnology, Faculty of Science, Mahidol University, Bangkok 10400, Thailand

2 Mahidol University-Osaka University Collaborative Research Center for Bioscience and Biotechnology, Faculty of Science, Mahidol University, Bangkok 10400, Thailand

3 National Center for Genetic Engineering and Biotechnology (BIOTEC), National Science and Technology Development Agency (NSTDA), Pathum Thani 12120, Thailand

4 Laboratory of Biotechnological Research for Energy and Bioactive Compounds, Department of Microbiology, Faculty of Science, King Mongkut's University of Technology Thonburi, Bangkok, Thailand

Published online: 25 February 2025

cardiovascular issues, central nervous system toxicity, and bone marrow suppression (Amjad et al. 2024; Komarova and Wodarz 2005). As ineffective cytotoxic drugs often result in fatalities, the search for new drugs with improved efficacy remains a major goal of cancer research (Law et al. 2020).

Natural products are crucial source materials in drug discovery, particularly with regard to cancer treatment; indeed, more than 48% of drugs approved between January 1981 to September 2019 were anticancer or antiinfection agents based on natural products and their derivatives (Newman and Cragg 2020). Actinomycetota exhibit a wide range of biological activity and are important sources of natural products, with 70% of natural derivative compounds isolated from them. However, while most microbial metabolites from terrestrial sources have been thoroughly researched, making them not suitable for focus as novel sources for the discovery of secondary metabolites (Yan et al. 2022). Marine habitats are promising sources of natural products due to their unique environments and extreme conditions, including coastal areas, deep-sea sediments, seawater, marine organisms, and mangrove forests (Subramani and Sipkema 2019). Mangrove forests cover approximately 60–75% of the world's tropical habitats (Holguin et al. 2001). These forests have a unique tidal gradient that affects fluctuations in temperature, oxygen concentration, and salinity, which supports the production of extraordinary metabolites (Hong et al. 2009). Thus, marine Actinomycetota in mangrove forests have adapted to these environments, leading to the production of compounds with anticancer and other pharmaceutical activities not typically found in terrestrial organisms (Olano et al. 2009). In particular, the genus *Micromonospora* is a promising source of bioactive compounds with unique chemical diversity that demonstrate significant anticancer and antimicrobial activity. Consequently, this genus has been acknowledged as a valuable foundation for the discovery of novel compounds (Hifnawy et al. 2020). However, conventional methods frequently result in the rediscovery of known compounds and are highly time-consuming (Ossai et al. 2022). Therefore, a metabolomic approach is applied as a selection tool to obtain information on metabolites, helping to annotate and prioritize target compounds for isolation through the dereplication process. Liquid chromatography coupled to tandem mass spectrometry (LC–MS/MS) is commonly used for untargeted metabolomics, as it allows the determination of chemical structures through fragmentation patterns (Gaudêncio et al. 2023). Furthermore, in silico prediction of microbial secondary metabolites has become an excellent strategy for novel drug development (Amin et al. 2019). Genome mining, an in silico tool, has emerged to facilitate the examination and exploitation of natural product diversity through annotated biosynthetic gene clusters (BGCs) in the chromosome. antiSMASH (antibiotics and Secondary Metabolite Analysis Shell) is a computational tool that plays an important role in genome mining by performing comparative analyses of identified BGCs with the MIBiG database, which includes more than 1000 BGCs categorized by their chemical structure (Navarro-Muñoz et al. 2018). Genome mining has revolutionized the natural product discovery process by dereplicating known BGCs and identifying novel BGCs, thereby providing useful information for drug discovery and metabolomic analysis (Baltz 2021).

In this work, we aim to identify the uniqueness of the isolated marine *Micromonospora* using a polyphasic taxonomy approach. Additionally, we seek to comprehensively investigate this strain's distribution of metabolites and biosynthetic gene clusters through comparative metabolomic and genomic analyses. The results suggest that the target bioactive compound is likely classified in the polycyclic tetramate macrolactam group with anti-colorectal cancer activity.

## Materials and methods

### Marine Actinomycetota strain isolation

Strain 2MTK254 (= TBRC 19103) was isolated from mangrove sediment (Chanthaburi province, Thailand; GPS 12.53100, 102.0502). Marine sediment samples was performed tenfold serial dilution and then spread out on starch casein nitrate agar (10 g soluble starch, 0.3 g casein (vitamin free), 2 g $KNO_3$, 0.05 g $MgSO_4.7H_2O$, 2 g $K_2HPO_4$, 2 g NaCl, 0.02 g $CaCO_3$, 0.01 g $FeSO_4.7H_2O$, 15 g agar, and 1000 mL reverse osmosis (RO) water) supplemented with cycloheximide (50 µg/mL) and nalidixic acid (25 µg/mL) to inhibit the growth of fungi and Gram-negative bacteria, respectively. After incubating for 14 days at 30 °C, Actinomycetota-like colonies were obtained and stored at − 80 °C in tryptic soy broth (TSB; 17 g tryptone, 3 g soytone, 5 g NaCl, 2.5 g $K_2HPO_4$, 2.5 g glucose, and 1000 mL RO water) with 20% (v/v) glycerol.

### Genomic DNA extraction, 16S rRNA identification, and phylogenetic analysis

Genomic DNA of 2MTK254 was obtained using the commercial i-genomic™ BYF DNA Extraction Mini kit (iNtRON biotechnology) with modified manufacturer's instructions. In lysate preparation process, the mixture was supplemented with 3 µL of lysozyme and 10% SDS to enhance cell disruption efficiency.

For long-read sequencing, chromosomal DNA was extracted following the modified method of Saito and Mura. The cell biomass was physically disrupted using a mortar and pestle before cell lysis was conducted through lysozyme.

The 16S rRNA genes were amplified using universal primers 27F (5′-AGAGTTTGATCCTGGCTCAG-3′) and

1492R (5′-TACGGCTACCTTGTTACGACTT-3′) (Soergel et al. 2012) on a Perkin Elmer GeneAmp 2400 PCR system. The amplification protocol consisted of an initial denaturation at 95 °C for 5 min, followed by 30 cycles of 95 °C for 30 s, 52 °C for 30 s, and 72 °C for 2 min, with a final extension at 72 °C for 5 min. The obtained sequences were analyzed for similarity to available type strains using the EzBioCloud web server (https://www.ezbiocloud.net) (Na et al. 2018) for species identification. Multiple sequences aligned with type strains were utilized to construct a phylogenetic tree in MEGA software (version 11) employing CLUSTAL W (Thompson et al. 1994). The neighbor-joining tree was constructed based on Kimura's two-parameter model (Kimura 1980), while the minimum evolution tree was generated using the ordinary least squares method to minimize the sum of branch lengths (Rzhetsky and Nei 1993) Bootstrap analysis with 1000 replications was conducted to assess the statistical reliability of the tree (Felsenstein 1985).

## Polyphasic taxonomy analysis

### Genomic sequencing and genotypic characterization

The genomic DNA was submitted to Vishuo Biomedical, Thailand, for paired-end (PE150) sequencing using the Illumina HiSeq X Ten platform. For long-read sequencing, a library was prepared using the SQK-RBK114.24 kit (Oxford Nanopore Technologies, ONT, UK) and loaded into PromethION flow cell R10.4.1. The Illumina short read sequences were used to assemble the whole genome sequences through Medaka v1.4.3. The genomic DNA sequence of *M. carbonacea* DSM 43168 $^{T}$ (GenBank accession no. FMCT00000000) and *M. carbonacea* subsp. *aurantiaca* DSM 43815 (GenBank accession no. CP058322) were obtained from the NCBI database. The genome sequences were employed to construct a phylogenomic tree using the Genome BLAST Distance Phylogeny (GBDP) method with a trimming algorithm facilitated by the TYGS web server (https://ggdc.dsmz.de) (Meier-Kolthoff et al. 2013). Digital DNA-DNA hybridization (dDDH) values were calculated using Genome-to-Genome-Distance-Calculation version 3.0 (GGDC; https://ggdc.dsmz.de) (Meier-Kolthoff et al. 2013). Average nucleotide identity (ANI) was evaluated using ANIb and ANIm, which are accessible on the JSpeciesWS web server (https://jspecies.ribohost.com/jspeciesws) for pairwise comparisons (Richter et al. 2015).

### Phenotypic and chemotypic characterization

Cultural characteristics were investigated by incubating strain 2MTK254 in various International *Streptomyces* Project (ISP) media, including tryptone-yeast extract (ISP1), yeast extract-malt extract (ISP2), oatmeal (ISP3), inorganic salt-starch (ISP4), glycerol-asparagine (ISP5), peptone-yeast extract iron (ISP6), and tyrosine (ISP7) for 14 days at 30 °C (Shirling and Gottlieb 1966). The color of the substrate, aerial mycelium, and reverse-side pigment were recorded using the ISCC-NBS color charts (Balagurunathan et al. 2020). The spore structure of 2MTK254 was cultivated in soya flour mannitol agar (SFM: 20 g soya flour, 20 g mannitol, 15 g agar, and 1000 mL tap water) at 30 °C for 21 days. The spore structure was examined with a scanning electron microscope (JSM-6610LV, JEOL).

Tolerance testing of strain 2MTK254 was conducted across a range of temperatures (5, 20, 30, 40, and 50 °C), pH levels (3–11 units), and salt concentrations (1–5% w/v NaCl) on ISP6 media for 14 days. Single carbon source utilization in basal ISP9 medium with 1% w/v of carbon source was observed after incubating at 30 °C for 14 days (Shirling and Gottlieb 1966). Various biochemical tests were performed to assess starch hydrolysis, hydrogen sulfide production, melanin production, nitrate reduction, gelatin liquefaction, and milk coagulation and peptonization using ISP4, ISP6, ISP7, ISP8, peptone-glucose-gelatin medium (20 g glucose, 5 g peptone, 200 g gelatin, and 1000 mL RO water), and 10% skim milk (Difco) broth, respectively (Ngamcharungchit et al. 2023). Enzymatic activity was examined using a commercial API-ZYM 20E system (bioMérieux) following the manufacturer's protocol.

Chemotypic analysis was performed on freeze-dried cells that had been cultivated in TSB broth for 3 days at 30 °C. Diaminopimelic acid (DAP) isomers and whole-cell sugars were extracted and investigated on glass cellulose TLC plates following the method described by Staneck et al. (Staneck and Roberts 1974). Menaquinones (MK) were extracted and isolated according to a previously reported method (Collins et al. 1977) and then analyzed using reverse-phase HPLC. Phospholipids were extracted and separated on two-dimensional aluminum silica TLC plates following the protocol of Minnikin et al. (Minnikin et al. 1984). Cellular fatty acids were obtained and converted into methyl ester form. The mixture was identified using gas chromatography (GC) with the RTSBA6 database in accordance with the Microbial Identification System (MIDI, version 6.0) (Kämpfer and Kroppenstedt 1996).

### Production and extraction of secondary metabolites

*Micromonospora* strains (2MTK254, DSM 43168 $^{T}$, and DSM 43815) were precultured in 5 mL of 301 broth (1 g glucose, 24 g soluble starch, 3 g meat extract, 3 g peptone, 5 g yeast extract, 4 g $CaCO_3$, and 1000 mL RO water) in an incubator shaker (200 rpm) at 30 °C for 3 days. The seed culture was transferred into baffled flasks containing 100 mL of A3M medium (20 g soluble starch, 20 g glycerol, 15 g

pharmamedia, 10 g dianion HP-20, 5 g glucose, 3 g yeast extract, and 1000 mL RO water) as a production medium and incubated at 170 rpm and 30 °C for 7 days. The cell-free supernatant was obtained by centrifuging culture broth for 10 min at 4 °C and 8000 rpm. The supernatant was subjected to ethyl acetate extraction, and the solvent was evaporated using a rotary evaporator. Cell biomass was extracted with ethanol and concentrated into a dry powder. The crude extract was a combination of compounds from the supernatant and intracellular extracts (Intra et al. 2016). The secondary metabolite production of all strains was carried out with five replications. The production medium (A3M) was extracted and used as a blank in the metabolomic analysis.

## Untargeted metabolomic profiling

### LC–MS/MS analysis

Crude extracts were dissolved in methanol (LC–MS grade) to adjust final concentration at 1 mg/mL. Samples were analyzed with UHPLC coupled with an Orbitrap Fusion™ Tribrid™ mass spectrometer (Thermo Scientific™, Waltham, MA, USA). The UHPLC system utilized a reverse-phase BEH C18 column (2.1 × 100 mm, 1.7 µm particle size) (Waters, MA, USA) operating at a flow rate of 0.3 mL/min and an injection volume of 3 µL. Mobile phase contained water plus 0.02% formic acid (A) and acetonitrile plus 0.016% formic acid (B). The gradient elution condition was consisted of 5 to 100% B (0–25.0 min), 100% B (25.0–27.0 min), and 5% B (27.1–30.0 min), with the column temperature at 60 °C. MS data were acquired in positive and negative ion mode using electrospray ionization (ESI) with a voltage of 3500 V. The ion transfer and vaporizer temperatures were set to 333 °C and 317 °C, respectively. MS scans were collected over a range of *m/z* 150–2000, with a resolution of 120,000. MS/MS data were acquired in data-dependent mode for the five most abundant spectra per cycle at a resolution of 30,000 with %HCDs at 25, 35, and 45.

### MS data processing and molecular networking analysis

The MS data were converted from.RAW (Thermo) format to ".mzML" using MSConvert software from the ProteoWizard package (Chambers et al. 2012). Peak picking was performed with vendor algorithms for preferred filtering. The converted files were processed using MZmine 3 software (version 3.9.0) (Schmid et al. 2023). For mass detection, centroid mode was applied with a noise level set at 1E5 for MS level 1 and 9E4 for $MS^2$. The ADAP chromatogram parameters included minimum group size of four scans, a group intensity threshold, and minimum highest intensity of 2E5, with a 0.02 m/*z* tolerance. Chromatogram resolution was determined using a minimum absolute height of 2E5 and a peak top/edge ratio of 1.5. Isotopes were analyzed using the 13C isotope filter or isotope grouper with absolute retention time tolerance set at 0.3 min and intra-sample tolerance of 0.02 m/*z*. Features were aligned using the join aligner with a 0.02 m/*z* tolerance, an absolute RT tolerance of 0.25 min, and an RT weight of 25. The feature list rows filter was applied to filter features ranging from 150 to 2000 m/*z* and RT 0–30 min. Subsequently, gap filling was performed using the peak finder with an intensity tolerance of 20% and an *m/z* tolerance of 0.02. The processed feature list was then exported, generating a feature table (.CSV file) and an MS/MS spectral summary (.mgf file format).

The spectral data and feature table were uploaded into Global Natural Products Social Molecular Networking (GNPS) to construct a feature-based molecular network (Nothias et al. 2020) with precursor and fragment ion mass tolerance set at 0.02 Da and utilizing MZmine as table source. Molecular networking required library matching and edge cosine scores greater than 0.7. The network was visualized and illustrated using Cytoscape software (version 3.10.2) (Shannon et al. 2003).

### Dereplication, identification and chemical class annotation of secondary metabolites

The molecular network was dereplicated using the GNPS library, DEREPLICATOR +, and MolNetEnhancer (Ernst et al. 2019) workflow available on the GNPS platform. Annotation was performed following the default parameters. The MS/MS spectral summaries were uploaded into SIRIUS software (version 5.8.6) (Dührkop et al. 2019) to predict the chemical formulation of parent ions based on their isotope and compound fragmentation patterns. CANOPUS was used to forecast chemical compound classification through ClassyFire ontology (Dührkop et al. 2021). The feature table (CSV format) was analyzed in MetaboAnalyst (version 6.0) (https://new.metaboanalyst.ca) to conduct principal component analysis (PCA) and generate a heatmap with hierarchal cluster analysis. For heatmap analysis, all feature metabolites were subjected to an ANOVA test. The top 100 features were then selected for hierarchical clustering and generated as a heatmap.

### Anticancer activity

The anticancer activity of crude extracts was investigated using the MTT (3-(4,5-dimethylthiazol-2-yl)−2,5-diphenyltetrazolium bromide) assay (Mosmann 1983). Human colorectal cancer (HCT-116) cell lines (ATCC, Manassas, VA, USA) were seeded into 96-well microtiter plates at a concentration of $2 \times 10^3$ cells/mL and cultured in DMEM-high glucose medium (Cytiva, Marlborough, MA, USA) supplemented with 10% fetal bovine serum (FBS;

Merck, Schuchhardt, Darmstadt, Germany), 1X HEPES, and 1% penicillin/streptomycin (Thermo Fisher Scientific, Wilmington, DE, USA). The plates were incubated at 37 °C and 5% $CO_2$ for 24 h. The dried extracts were dissolved in dimethyl sulfoxide (DMSO; Merck, Schuchhardt, Darmstadt, Germany) to achieve a final concentration of 10 µg/mL and varying final concentrations (200, 2, 0.2, and 0.002 µg/mL) for $IC_{50}$. These solutions were then added to the cells for a 72-h treatment period. Doxorubicin served as a positive control, while DMSO was used as the negative control. Following the treatment, the cells were incubated with MTT solution (Abcam, Cambridge, UK) for 3 h. Formazan crystals were then solubilized in DMSO, and absorbance was measured using a microplate reader (ENVISION) at 570 nm. Additionally, another human colorectal cancer cell line (HT-29) was also used to investigate cytotoxicity activity via the MTT assay, with the final concentration of the extract applied at 10 µg/mL.

#### Cytotoxicity test in cancer cells

The normal cells cytotoxicity test was evaluated with human large intestine (CDD-841-CoN) cells by using MTT assay at ECDD center. The CDD-841-CoN was seeded at 1 × 104 cell/well per 96 well plate. Compounds were dissolved with DMSO to various final concentrations (100, 10, 0.1, and 0.00 µg/mL). Then, cells were treated with normal cells and incubated 24 h. MTT solution was applied to the cells. After 3 h, the MTT solution was discarded, and 100 µl of DMSO were added to dissolve formazan crystal. The change in color of a solution was measured the absorbance at 570 nm using a microplate reader.

#### Bioassay-guide fractionation and purification

The crude extracts were dissolved in methanol and partially purified using a 35 CC Sep-Pak® C18 Plus short cartridge (Water, Milford, MA, USA). The cartridge was equilibrated with 350 mL each of acetonitrile and type I water. Subsequently, the crude solution was loaded onto the cartridge and eluted using a stepwise gradient of acetonitrile–water system: 175 mL each of 20%, 40%, 60%, 80%, and 100% (v/v) acetonitrile, resulting in 5 fractions (Fr.1–Fr.5). Each fraction was concentrated using a rotary evaporator to obtain dry fractions.

The target compound was manually isolated through semi-preparative HPLC (Agilent 1260 infinity system) equipped with UV detector on a normal phase Cap cell Pak® $NH_2$ UG 80 Å column (4.6 × 250 mm, 5.0 µm particle size) (Shiseido, Tokyo, Japan). The sample injection volume was 100 µL. The mobile phase consisted of water with 0.1% formic acid (A) and acetonitrile (B). The elution solvent gradient was follows: 80–100% B (0–18.0 min), 80% B (18.01–20.0 min). The flow rate was set to 2.5 mL/min, and detection was carried out at a UV wavelength of 360 nm.

#### Genome mining

Genome annotation was performed with Rapid Annotation using Subsystem Technology (RAST) (version 2.0) (Aziz et al. 2008). The SEED viewer (Overbeek et al. 2014) was utilized to demonstrate the organism overview and subsystem category distribution. Secondary metabolite–related biosynthetic gene clusters (BGCs) were identified using the antiSMASH website (version 7.1.0) (Blin et al. 2023) with relaxed strictness detection. Various additional features such as ClusterBlast, Cluster Pfam analysis, KnownClusterBlast, and Pfam-based GO term annotation were activated for detection. The antiSMASH annotated GenBank files were explored, and BGCs were classified using BiG-SCAPE (version 1.1.5) with a 0.3 default cutoff (Navarro-Muñoz et al. 2020). However, the homologous biosynthetic gene clusters were aligned and visualized using Clinker (version 0.0.30) with minimum alignment sequences identity at value of 0.3 (Gilchrist and Chooi 2021).

## Results

### Polyphasic taxonomy study of an isolated marine Actinomycetota identified putative new subspecies of *Micromonospora carbonacea*

The completed 16S rRNA gene sequence (1515 bp) of 2MTK254 exhibited a relationship to members of the genus *Micromonospora*. This strain showed the highest similarity to *Micromonospora carbonacea* DSM 43168 $^T$, at 99.65%. The neighbor-joining and minimum evolution trees demonstrated that 2MTK254 clustered closely with *M. carbonacea* DSM 43168 $^T$, supported by high bootstrap values of 80 and 88, respectively (Fig. 1 and Fig. S1).

The phylogenomic tree based on TYGS showed a relationship between 2MTK254 and members of the genus *Micromonospora*. 2MTK254 clustered in the same clade with *M. carbonacea* subsp. *aurantiaca* DSM 43815 and *M. carbonacea* DSM 43168 $^T$, supported by a bootstrap value of 100% (Fig. 2a). The ANIb and ANIm values between this strain and its closest related species, *M. carbonacea* DSM 43168 $^T$ (97.54, 97.69) and *M. carbonacea* subsp. *aurantiaca* DSM 43815 (97.70, 97.82), were determined. Additionally, the in silico dDDH value between strain 2MTK254 and these two closest-related species ranged from 78.8 to 80%.

The morphology of 2MTK254 demonstrated abundant substrate mycelium development in all ISP media except for ISP2, where aerial mycelium was poorly produced, as illustrated in Table S1. Notably, soluble pigment was observed

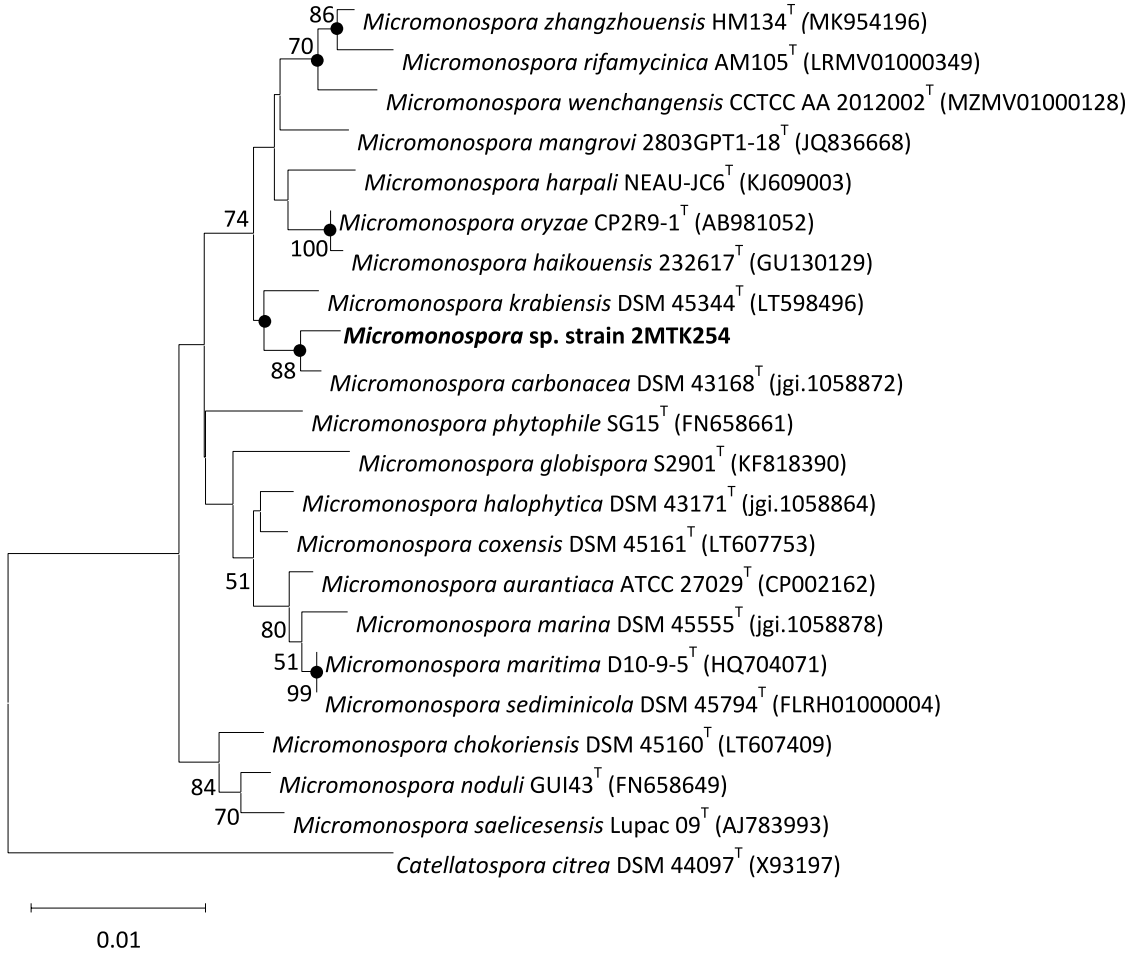


**Fig. 1** The neighbor-joining phylogenetic tree of 2MTK254 and other type strains within the genus *Micromonospora* was constructed based on the 16S rRNA gene sequence. *Catellatospora citrea* DSM 44097 T was used as the outgroup. Bootstrap values above 50% (based on 1000 replications) are indicated at branch points. The scale bar represents 0.01 substitutions per nucleotide. The solid dots indicate branches that were recovered in the minimum evolution tree (see Fig. S1)

in ISP2 (dark greenish blue) and ISP3 (greyish olive green) medium (Fig. 2b). Spore morphology of 2MTK254 demonstrated single ovoid spores with a warty surface morphology. The spores had a diameter of approximately 0.5–0.6 μm and formed at the ends of substrate mycelium (Fig. S2).

The cell wall of strain 2MTK254 contained *meso*-DAP and 3-OH-DAP. The whole-cell sugar composition included arabinose, glucose, mannose, and xylose, which corresponded to sugar pattern D (Lechevalier and Lechevalier 1970). The dominant menaquinone detected in strain 2MTK254 was MK-9($H_4$), at 90.79%, followed by MK-9($H_6$) at 4.17% and MK-9 at 5.04%. Diagnostic phospholipids identified were diphosphatidyl glycerol (DPG) and phosphatidylethanolamine (PE), corresponding to phospholipid type 1I (Lechevalier et al. 1977). The major fatty acid pattern of strain 2MTK254 was identified as type 3b (Kroppenstedt 1985), including *iso*-$C_{16:0}$ (36.63%), 10-methyl-$C_{17:0}$ (17.54%), *iso*-$C_{15:0}$ (7.75%), $C_{17:1}\omega8c$ (6.95%), and $C_{17:0}$ (6.01%). Chemotypic characteristic of 2MTK254 and their closely related strain were compared in Table S2.

The physiological properties of strain 2MTK254 and two closely related strains (*M. carbonacea* subsp. *aurantiaca* DSM 43815 and *M. carbonacea* DSM 43168 T) are summarized in Table S3. 2MTK254 and *M. carbonacea* DSM 43815 exhibited similar growth abilities at temperatures ranging from 20 to 40 °C and pH levels from 6 to 11 and had a NaCl tolerance of 1–2% (w/v). However, *M. carbonacea* DSM 43168 T demonstrated differences in tolerability. Strain 2MTK254 could utilize L-arabinose, D-mannose, sucrose, D-xylose, and D-glucose as a single carbon source. The starch hydrolysis, hydrogen sulfide production, melanin production, and milk coagulation and peptonization were positive. The enzymatic activities observed to be positive included acid phosphatase, alkaline phosphatase, cystine arylamidase, esterase (C4), esterase lipase (C8), leucine arylamidase, lipase (C14), N-acetyl-β-glucosaminidase,

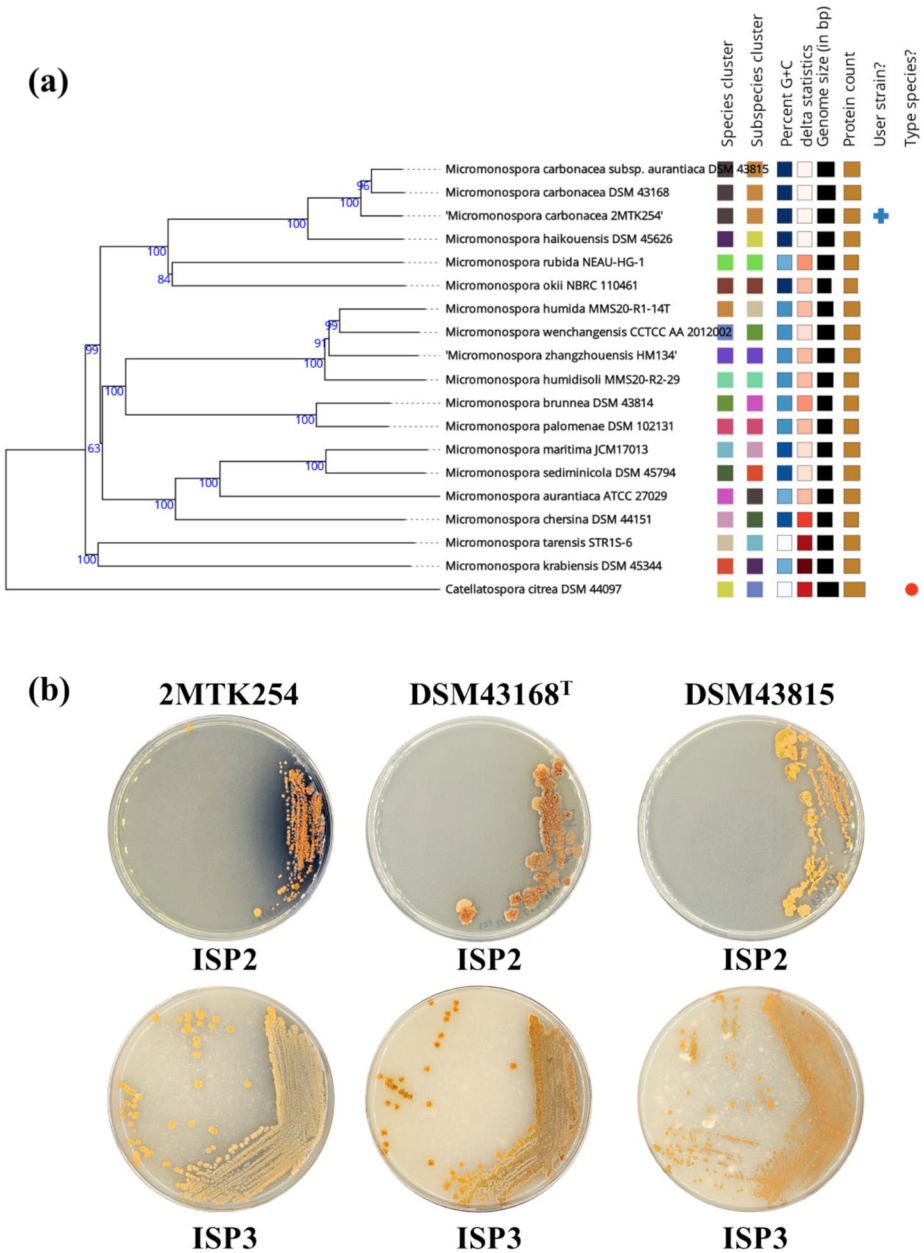


**Fig. 2** Genotypic and cultural characteristics. **a** Phylogenomic tree based on genome sequences in TYGS using the genome BLAST distance phylogeny (GBDP) method. The branch length is scaled according to GBDP distance formula d5. The numbers above branches are GBDP pseudo-bootstrap support values (> 60%) derived from 100 replications. **b** Colony morphology of *Micromonospora* strain 2MTK254 and the two closely related species in ISP2 and ISP3 medium

naphthol-AS-BI-phosphohydrolase, trypsin, valine arylamidase, α-chymotrypsin, α-galactosidase, α-glucosidase, β-galactosidase, and β-glucosidase.

## Comparative genomic and metabolomic analysis distinguished the novel subsp. *M. carbonacea* subspecies *caeruleus* 2MTK254 from two closely related strains

The overall distribution of the chemical diversity in A3M medium of strains 2MTK254, *M. carbonacea* DSM 43815, and *M. carbonacea* DSM 43168 $^{T}$ was evaluated using principal component analysis (PCA) (Fig. S3). The first- and second-dimensional axes contributed 36.4% and 17.7% to the analysis, respectively. Furthermore, the top 100 features across all strains were selected based on ANOVA results for heatmap analysis. This heatmap highlights the distinct metabolic signatures among these strains (Fig. 3). A notable separation among crude extracts was observed in 2MTK254, which showed distinct differences compared to the clustering of *M. carbonacea* DSM 43815 and *M. carbonacea* DSM 43168 $^{T}$.

The genomes of 2MTK254 and the two closely related strains were analyzed using RAST annotation. The analysis revealed that the genome of 2MTK254 is 7,520,400 bp in size (Fig. 4a) with highest G + C content of 74.1 mol%. The N50 and L50 values were calculated as 26,998 bp and 1, respectively. *Micromonospora*

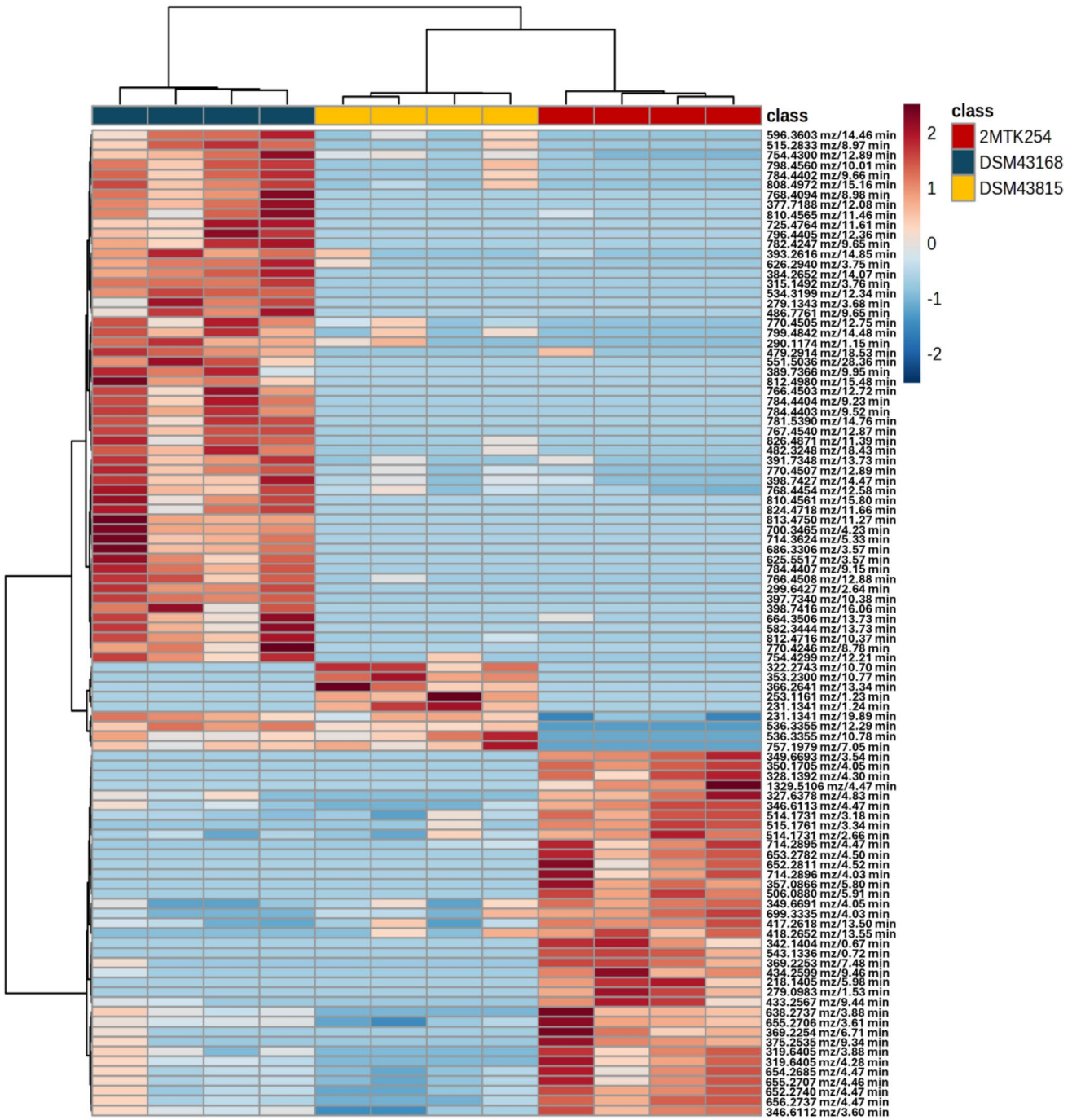


**Fig. 3** Heatmap analysis of 100 significant metabolites differentially abundant after the performed ANOVA test, with $p \leq 0.05$ in all metabolite features. Each block of four columns represents the different strains, with *Micromonospora* sp. strain 2MTK254 in red, *M. carbonacea* DSM 43168 $^{T}$ in blue, and *M. carbonacea* DSM 43815 in yellow. Rows represent the *m/z* and retention time of abundant metabolites. The color in the heatmap demonstrates the relative abundance of metabolites, with red signifying higher abundance relative to blue

sp. 2MTK254 contains the lowest total of 6535 coding sequences and 62 total RNA genes (Table S4). The RAST database grouped subsystems into 297 categories, with 18% subsystem coverage and 82% non-subsystem coverage (Fig. 4b). Furthermore, the genome analysis of strain 2MTK254 revealed that contig 1 in region 24 matched the *indC* gene (0.7 out of 1 similarity score), which is essential for blue pigment production observed in 2MTK254. The *indC* gene is primarily responsible for indigoidine synthase and contains three important

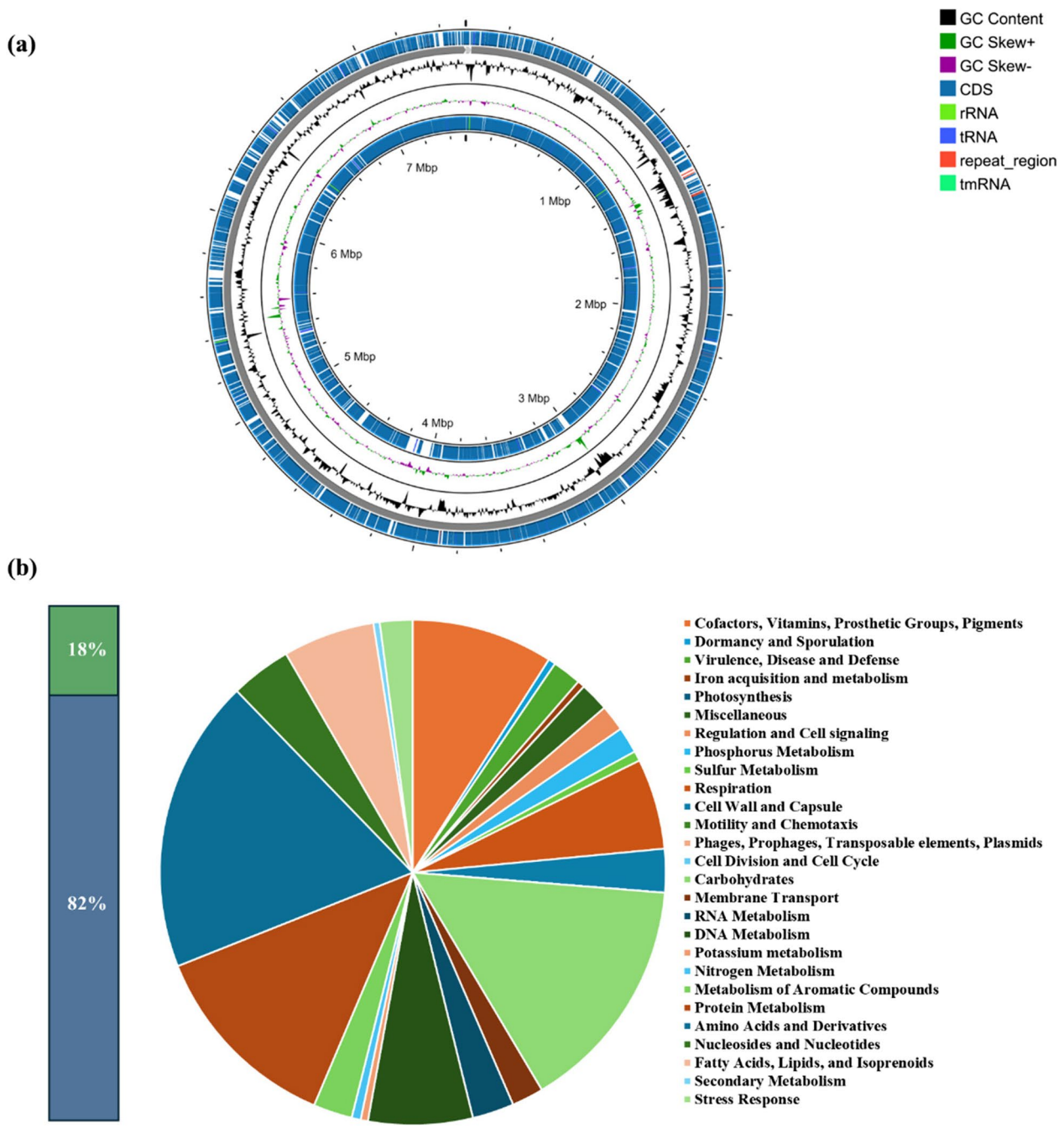


**Fig. 4** The whole-genome properties of strain 2MTK254. **a** Circular map of the total length, 7,520,400 bp, visualized by Proksee. **b** Subsystem category distribution of 2MTK254 annotated using the RAST server. The left bar chart displays the percentage of subsystem coverage (green) and non-subsystem coverage (blue). The pie chart illustrates the distribution of subsystem categories, with each feature category represented by a specific color

modules: the AMP-binding enzyme, the phosphopantetheine attachment site, and the thioesterase domain. Strain 2MTK254 includes all of these gene domains, supporting indigoidine production. In contrast, the other two strains lack the thioesterase domain necessary for manufacturing the blue pigment.

## Tandem mass spectrometry molecular networking identified secondary metabolites in *Micromonospora* species

The GNPS platform was utilized to identify compounds produced by strain 2MTK254 and two closely related *Micromonospora* species. The Venn diagram in Fig. 5A

shows that *M. carbonacea* DSM 43168 $^{T}$ exhibited the highest number of metabolic features, followed by DSM 43815 and strain 2MTK254. The three strains shared 761 similar metabolites, and 2MTK254 was revealed to have 7191 unique features. A molecular network was created using the dataset from positive electrospray ionization mode to detect the maximum number of compounds possible within a high mass detection range. The feature-based molecular network revealed 3651 nodes, 125 clusters with 3708 edge connections, and 3334 singleton nodes (Fig. 5B). Metabolites were classified using CANOPUS into seven natural product classes, including alkaloids, peptides, carbohydrates, fatty acids, polyketides, shikimates-phenylpropanoids, and terpenes. The molecular network indicated that the two largest clusters consisted of metabolites from all three strains, predominantly comprising fatty acids, alkaloids, and peptides. Strain-specific clusters such as the peptide cluster of *M. carbonacea* DSM 43168 $^{T}$ and alkaloids from strain 2MTK254 were also detected in the molecular network. Furthermore, using the GNPS library, brevianamide F (284.17 m/*z*), and cyclo(proline-leucine) (211.14 m/*z*) were annotated based on their mass fragmentation pattern similarity (Figs. S4–S5), with cosine scores equal to 0.98 and 0.75, respectively. The DEREPLICATOR + database matches revealed the highest similarity to octadecanal oxime, with a hit score of 16, followed by peptidolipin F and lysobacteramide A. Notably, lysobacteramide A was categorized into polycyclic tetramate macrolactams (PTM) groups with hybrid polyketide synthase-nonribosomal peptide synthase (hybrid PKS-NRPS) natural product classification at *m*/*z* 478.28 (Fig. 5).

## Bioassay-guided fractionation and anticancer activity of *Micromonospora carbonacea*

The primary screening demonstrated that the crude extract of strain 2MTK254 exhibited the stronger inhibition against human colorectal cancer HCT-116 cells compared to DSM 43168 $^{T}$ and DSM 43815 (Table S5). However, none of crude extracts demonstrated significant inhibition against the HT-29 cells. All crude extracts were fractionated in varying

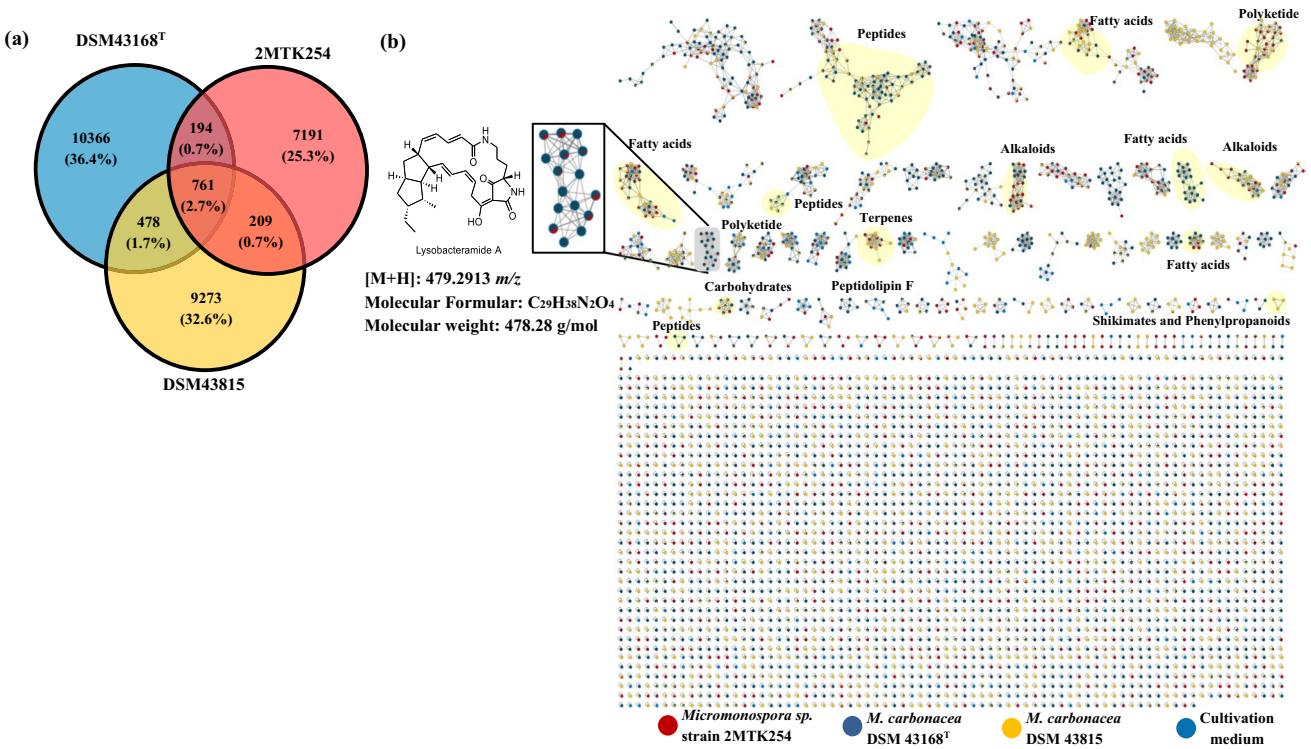


**Fig. 5** The metabolite diversity of *Micromonospora* strain 2MTK254 and closely related species. **a** Venn diagram showing the distribution of the numbers of *Micromonospora* metabolites, with *Micromonospora* sp. strain 2MTK254 in red, *M. carbonacea* DSM 43168 $^{T}$ in blue, and *M. carbonacea* DSM 43815 in yellow. **b** Feature-based molecular networking on crude extracts from three *Micromonospora* strains using the Global Natural Products Social Molecular Networking (GNPS) platform. Each node represents a unique compound from the strains. The node color corresponds to the strains, with *Micromonospora* sp. strain 2MTK254 in red, *M. carbonacea* DSM 43168 $^{T}$ in blue, *M. carbonacea* DSM 43815 in yellow, and production medium in light blue. The width of edges between nodes indicates compound similarity, represented by cosine scores. Compound annotation and natural product classification were performed based on mass fragmentation patterns using the GNPS platform and SIRIUS, as indicated within the light yellow circles. An enlarged insert was used to depict the structure of the lysobacteramide A cluster

concentrations of acetonitrile (ACN), and the resulted fractions were tested for inhibitory effects against HCT-116 and HT-29 cells. The 60%, 80%, and 100% ACN fractions of 2MTK254 (designed as 60%ACN-2MTK254, 80%ACN-2MTK254, and 100%ACN-2MTK254, respectively) exhibited over 80% inhibition against HCT-116 cells, with the 60%ACN-2MTK254 fraction showing the highest inhibition at 96.66%, though no activity was observed against HT-29 cells. For DSM 43168 $^{T}$, the 60%ACN and 80%ACN fractions exhibited 58.02% and 92.98% activity against HCT-116 cells, respectively, while the crude extract showed only 41.87% activity. Notably, the 80%ACN-DSM 43168 $^{T}$ fraction also showed activity against HT-29 cells (28.7%). This activity might be attributed to background interference from the A3M medium (10.55%). In comparison to other strains, all fractions of DSM 43815 demonstrated the lowest anticancer activity against both cancer cell lines (Table S5).

### The novel subspecies *M. carbonacea* subsp. *caeruleus* 2MTK254 produced anti-colorectal cancer compounds

To investigate compound from the 2MTK254 extract that exhibited anticancer activity against both HCT-116 and HT-29 cells lines, the 80%ACN-2MTK254 fraction was selected for compound purification. Based on high-performance liquid chromatography coupled to a UV detector (HPLC–UV) and LC–MS/MS analysis, along with the molecule dereplication, multiple peaks corresponding to compounds not found in the databases were identified. Hence, these PTM cluster were considered novel metabolites associated with ikarugamycin (IKA) and lysobacteramide A.

The putative new polycyclic tetramate macrolactam (P1), obtained as a white to light yellow powder from the 80%ACN-2MTK254 fraction using semi-preparative HPLC (Fig. 6b), exhibited an LC-HRMS signal with *m/z* 479.2904 $[M+H]^+$, corresponding to a chemical formula of $C_{29}H_{38}N_2O_4$. P1 displayed maximum ultraviolet absorption at 218 nm and 326 nm in methanol. Additionally, this compound demonstrated remarkable inhibitory activity of $84.84 \pm 16.3\%$ and $3.88 \pm 2.5\%$ at a final concentration of 10 µg/mL against HCT-116 and HT-29 cells, respectively. Moreover, P1 compound demonstrated potential $IC_{50}$ value of 0.125 µM and $CC_{50}$ of 12.641 µM. The selectivity index (SI) was calculated based on the $IC_{50}$ and $CC_{50}$ values. The result showed that the P1 compound had an SI index of 100.76, which was higher than the criteria for screening active compounds (Peña-Morán et al. 2016). However, a search of the SciFinder database (American Chemical Society) accessed on March 8, 2024, revealed two compounds with similar chemical formulae associated with anticancer activity. These included IKA analogs (isolated from *Streptomyces xiamenensis*) with anti-pancreatic cancer activity at 1300 nM (Xu et al. 2016) and epothilone derivatives (isolated from *Myxococcus xanthus*) exhibiting various activities against four cancer cell types (Ashley et al. 2003).

### Genome mining demonstrated biosynthetic gene clusters of polycyclic tetramate macrolactams compound

antiSMASH revealed the lowest number of predicted biosynthetic gene clusters (BGCs) in the genome of 2MTK254 compared to the two closely related strains. These 32 BGCs in the 2MTK254 genome were predominantly associated with hybrid PKS-NRPS, non-ribosomal peptide synthetase (NRPS), and terpene BGCs (Table 1). The BGCs showed the highest sequence similarity to known compounds in the MIBiG database such as everninomicin D (oligosaccharide; 100% similarity), sch 18,640 (thiopeptide; 100% similarity), sapB (lanthipeptide; 100% similarity), SGR polycyclic tetramate macrolactams (hybrid PKS-NRPS; 83% similarity), and loseolamycin A1 and A2 (T3PKS; 80% similarity).

Among all predicted BGCs in the genome of 2MTK254, a single PTM-BGC was annotated and found to be conserved with a cluster present in the genomes of DSM 43168 $^{T}$ and DSM 43815. However, the PTM-related BGCs of all strains shared the most similar match of biosynthetic gene cluster alignment with known secondary metabolites, as shown in Table S6. Interestingly, the SGR PTM BGCs displayed the highest similarity alignment, showing 83% gene similarity to the PTM-related BGCs in 2MTK254. Alignment of the biosynthetic gene clusters further demonstrated that 2MTK254 shares a similar core genes structure with the reference SGR PTMs, IKA, and two *Micromonospora carbonacea* strains. However, the BGCs of 2MTK254 exhibit unique patterns compared to the other strains, as shown in Fig. 8.

The BiG-SCAPE analysis classified the 108 BGCs identified across the three studied strains into 85 gene cluster families (GCFs), with the highest number of NRPS GCFs followed by other types and PKSother (T2 and T3PKS). The gene similarity network also confirmed the relationship of PTM-related BGCs of 2MTK254 and closely related species, as they clustered together but did not group with the reference SGR PTM BGCs (Fig. 7a). However, the IKA BGC was categorized into other GCF groups due to differences in gene functions, as shown in Fig. 7b.

## Discussion

In this study, the marine actinomycete strain 2MTK254, isolated from mangrove sediment in Chanthaburi province, Thailand, was classified within the *Micromonospora* genus based on morphological, physiological, and

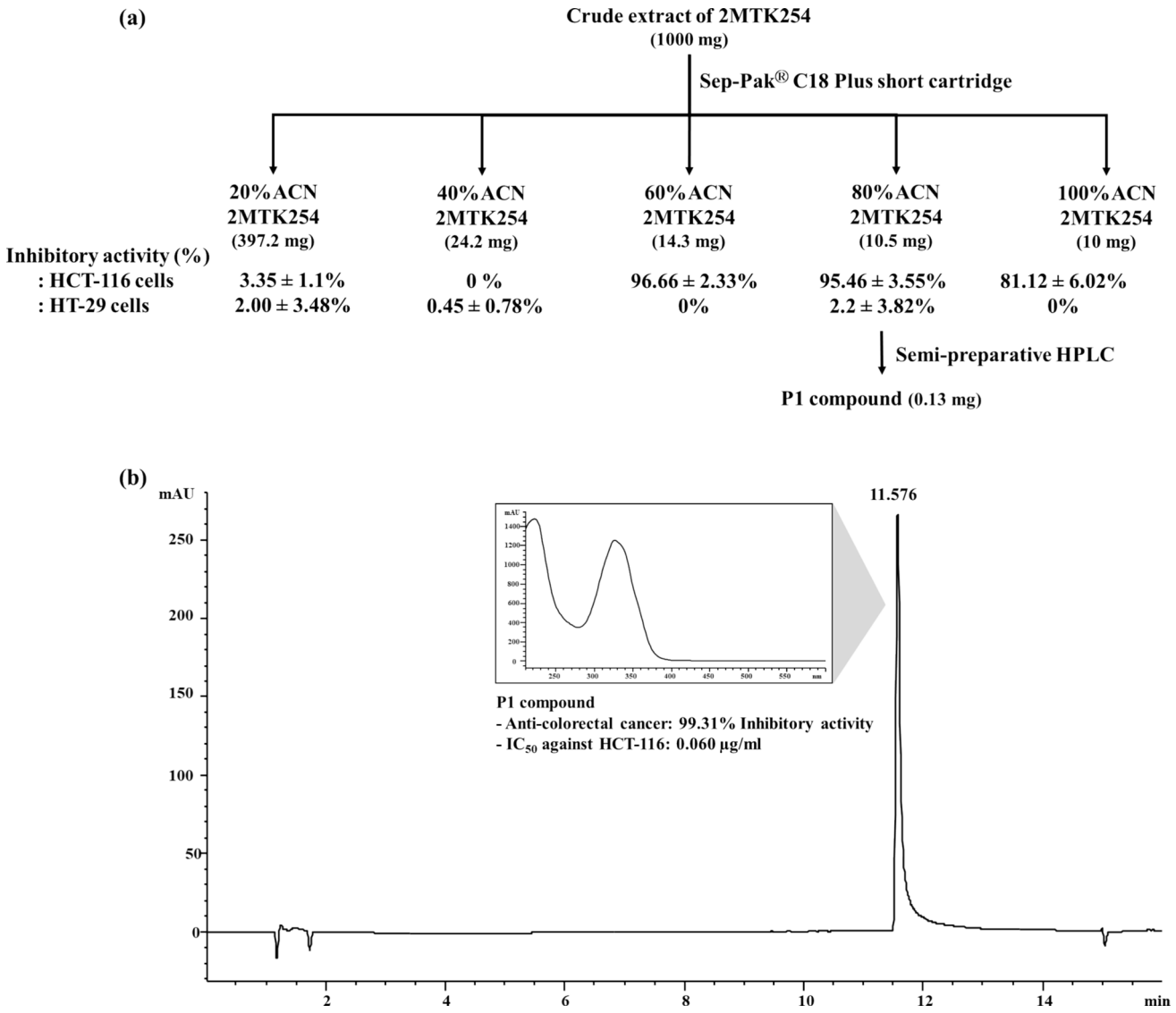


**Fig. 6** Schematic of the purification of crude extract from 2MTK254. **a** Fraction yields and inhibitory activity (%) against HCT-116 cells. **b** The HPLC chromatogram of the F4 fraction and UV spectrum of the P1 compound, showing their anti-colorectal cancer activity

chemotaxonomic characteristics (Genilloud 2015). Strain 2MTK254 exhibited distinct colony morphology and a unique fatty acid profile compared to two closely related strains (*M. carbonacea* subsp. *aurantiaca* DSM 43815 and *M. carbonacea* DSM 43168 $^{T}$). The chemotypic characteristics of all strains revealed that strain 2MTK254 has mannose as a distinct cell sugar and lacks phosphatidylinositol mannosides (PIM) lipids compared to the type species DSM 43168 $^{T}$ (Kittiwongwattana et al. 2015), as shown in Table S2. Notably, *M. carbonacea* DSM 43168 $^{T}$ was reported to be synonymous with *M. carbonacea* DSM 43815 in a taxonomic study (Genilloud 2015). Thus, the chemotypic characteristics of strain DSM 43168 $^{T}$ were used as a representative for *M. carbonacea*. Moreover, the whole genome of strain 2MTK254 showed ANIb values of 97.54–97.70% and ANIm values of 97.69–97.82%, which exceed the ANI threshold for species boundaries (Richter and Rosselló-Móra 2009) along with dDDH values of 78.8–79.9%. These genotypic characteristics strongly suggest that 2MTK254 belongs to the *Micromonospora* sp. However, previous studies have proposed that criteria for distinguishing subspecies should include ANI values lower than 98% and dDDH values in the range of 70–80% (Pearce et al. 2021; Venter et al. 2022). Therefore, strain 2MTK254 shows potential to be classified as a novel subspecies within *Micromonospora carbonacea*. However, terrestrial and marine habitats demonstrate distinct environmental conditions, such as temperature, oxygen levels, and salinity (Dharmaraj 2010). Mangrove forests are one of the marine habitats that exhibit unique environmental conditions, which lead

**Table 1** The number predicted secondary metabolites biosynthetic gene clusters of *Micromonospora* sp. 2MTK254, *M. carbonacea* DSM 43815 and *M. carbonacea* DSM 43168 $^{T}$ using antiSMASH version 7.1.0

| Type of BGCs | Strains | | |
|---|---|---|---|
| | *Micromonospora* sp. 2MTK254 | *M. carbonacea* DSM 43815 | *M. carbonacea* DSM 43168$^{T}$ |
| T1PKS | 2 | 3 | 7 |
| T2PKS | 1 | 1 | 1 |
| T3PKS | 1 | 1 | 1 |
| NRPS | 4 | 5 | 4 |
| Thioamide-NRP | 1 | 1 | 1 |
| Hybrid PKS-NRPS | 5 | 6 | 8 |
| RiPP-like | 1 | 2 | 1 |
| Thioamides/thiopeptide | 1 | 1 | 1 |
| Terpene | 3 | 4 | 3 |
| Lanthipeptide-class-I | 1 | 2 | 1 |
| Lanthipeptide-class-II | 0 | 1 | 1 |
| Lanthipeptide-class-III | 1 | 1 | 1 |
| Siderophore | 0 | 0 | 1 |
| Other | 11 | 6 | 8 |
| Summary | 32 | 34 | 39 |

microorganisms to adapt and survive in extreme environments. Furthermore, marine microorganisms must interact with various aquatic organisms, which induces the evolution of secondary metabolite pathways through chemical ecological signals between marine organisms. These unique adaptations, however, have been reported only in marine organisms (Jagannathan et al. 2021). Marine *Micromonospora* have been reported as candidates for producing growth-inhibiting and anticancer compounds. For example, Levantilide A, isolated from the deep Mediterranean Sea *Micromonospora* sp. strain M71-A77, showed antiproliferative activity against tumor cells with an $IC_{50}$ range of 20.7–52.4 µM (Gärtner et al. 2011). Tetrocarcins N and O were obtained from Micromonospora sp. 5–297, identified from Bohai Bay, China, and demonstrated MICs of 2 and 64 mg/mL against *Bacillus subtilis* (Tan et al. 2016). Lastly, lobsosamide A was produced by *Micromonospora* sp. RL09-050-HVF-A from Monterey Bay, CA. This compound demonstrated inhibitory activity against protozoan parasites with $IC_{50}$ values of 0.8 µM (Schulze et al. 2015).

The crude extract of strain 2MTK254 and its two closely related strains exhibited non-identical metabolites but shared some similarities in major chemical scaffolds, as observed in the results of the PCA and heatmap analyses. A feature-based molecular network was constructed to cluster structure-related metabolites based on the MS/MS fragmentation similarity of the compounds (Aron et al. 2020). GNPS provided a useful platform for annotating known

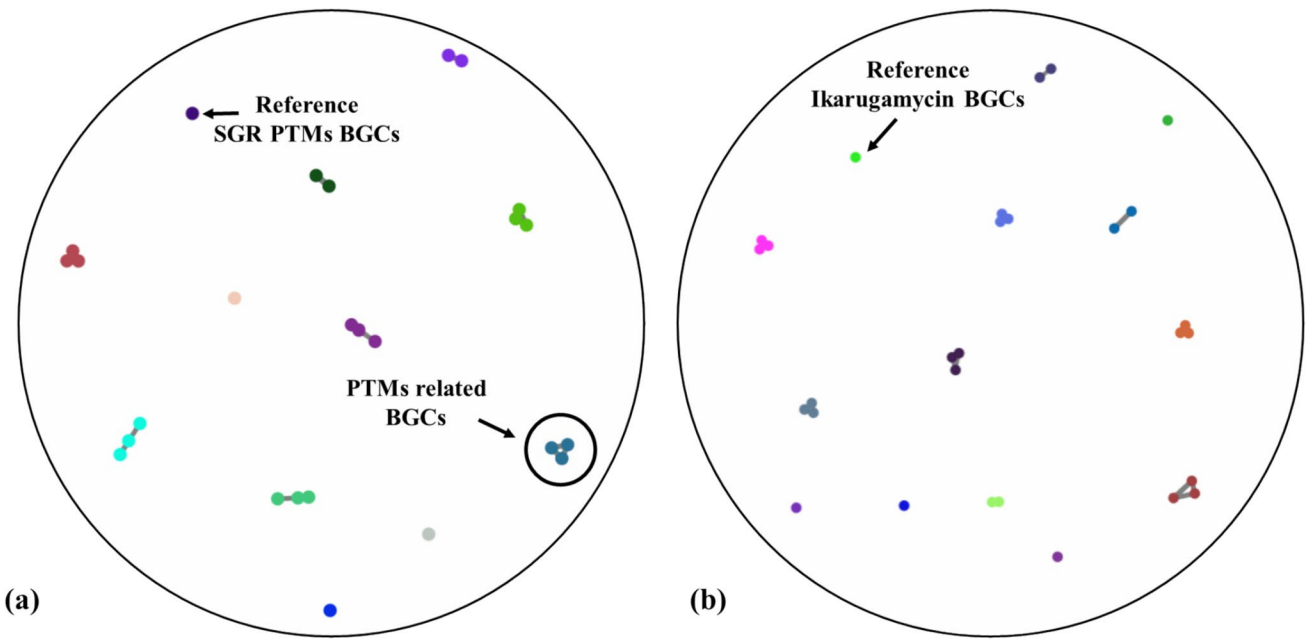


**Fig. 7** The biosynthetic gene cluster (BGC) network for 2MTK254, *M. carbonacea* DSM 43815, and *M. carbonacea* DSM 43168 $^{T}$ was created using BiG-SCAPE with a cutoff of 0.3. Nodes represent unique BGCs clustered into gene cluster families (GCFs). The black circle indicated PTM gene cluster family: **a** PKS gene cluster family group and **b** other cluster family group

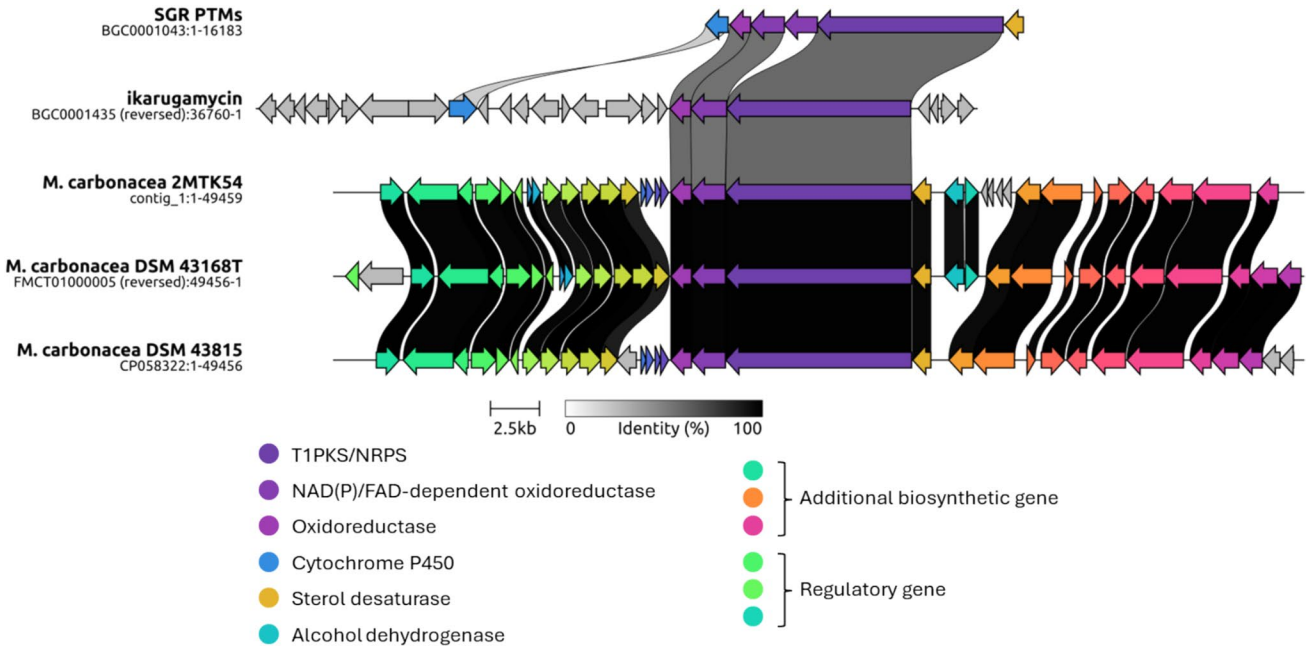


**Fig. 8** Alignment of biosynthetic gene clusters (BGCs) related to polycyclic tetramate macrolactam (PTM) based on amino acid sequences using Clinker. Genes encoding amino acids with more than 30% sequence identity are color-coded to highlight sequence homology. Genes shown in grey have low sequence similarity to others. The color legend indicates the predicted amino acid sequences, and the grey bars represent the percentage of amino acid identity

natural products and conducting dereplication processes. However, the GNPS spectral libraries cover only 2.5% of MS/MS reference spectra of known natural products (van der Hooft et al. 2020). Therefore, computational tools such as DEREPLICATOR + and SIRIUS can be integrated with molecular networking to improve the annotation process. The integrated molecular network revealed a high chemical diversity of alkaloids, followed by fatty acids and peptides, and allowed the prediction of some bioactive compounds. For example, brevianamide F is known as an antibacterial and antifungal agent previously isolated from *Micromonospora* and fungal species (Fernández et al. 2024; Mehetre et al. 2019). However, the polyketide cluster containing the predicted lysobacteramide A, with a low similarity score (score 13), was selected as the target cluster for purifying the anticancer compound. Lysobacteramide A ($C_{29}H_{38}N_2O_4$, 523.2549 m/z $[M+H+2Na]^+$) is a member of the polycyclic tetramate macrolactam (PTM) group, derived from *Lysobacter enzymogenes*. Lysobacteramide A has maximum ultraviolet absorption at 229 and 266 nm and demonstrates cytotoxic activity against various cancer cell lines, with $IC_{50}$ values ranging from 7.6 to 10.3 µM (Xu et al. 2015). Ikarugamycin ($C_{29}H_{38}N_2O_4$, 479.2919 m/z $[M+H]^+$) is also a member of the PTM group, appearing as a white to light yellow powder, with maximum UV absorption at 220 and 325 nm. IKA was reported to have an $IC_{50}$ at 1–10 µM against MCF-7, HMO2, HepG2, Huh7, and HL-60 cancer cells (Jiang et al. 2020). Based on these properties, compound P1 ($C_{29}H_{38}N_2O_4$, 479.2904 m/z $[M+H]^+$) exhibited chemical similarities with IKA in terms of molecular weight, mass error (−3.1296), UV spectrum, and core structure observed from the mass fragmentation pattern (Fig. S6). In addition, compound P1 demonstrated a lower $IC_{50}$ (0.125 µM) than previously reported, suggesting that P1 may be a derivative of IKA. The SI index also indicated that compound P1 should be effective and safe due to its high potential efficacy against cancer cells and lower toxicity against normal cells.

Furthermore, pothilone derivatives are known as potential sources of anticancer drugs against solid tumors (El-Sayed et al. 2024). For example, pothilone A ($C_{26}H_{39}NO_6S$, 494.2576 m/z $[M+H]^+$) can induce cell death through p53-independent apoptosis (Zhu and Panek 2000). Pothilone B ($C_{27}H_{41}NO_6S$, 508.2568 m/z $[M+H]^+$) induces cell apoptosis via the TRAIL and caspase 8-dependent pathway (Hardt et al. 2001). Ixabepilone ($C_{27}H_{42}N_2O_5S$, 507.281 m/z $[M+H]^+$), an aza-epothilone B, shows cytotoxicity against cisplatin- and paclitaxel-resistant ovarian cancer cell lines (Griffin et al. 2003). However, while epothilones A and B share the same maximum UV absorption at 211 and 249 nm, ixabepilone shows UV absorption at 211 and 247 nm (Hardt et al. 2001). The epothilone derivatives demonstrate quite different molecular weights and UV absorption compared to compound P1. Furthermore, comparing its molecular weight with those reported in the SciFinder database (American

Chemical Society) for compounds with cancer-related activity suggests that this compound may be a putative new PTM reported in *Micromonospora carbonacea*. Notably, marine *Micromonospora* sp. K310 was reported to produce butremycin, a 3-hydroxyl derivative of ikarugamycin (Kyeremeh et al. 2014).

The P1 compound is suspected to be related to IKA; therefore, P1 is likely to have a mode of action similar to IKA, targeting Hexokinase 2 (HK2) enzymes. HK2 is involved in the glycolysis pathway, which is highly expressed in cancer cells and tumor metabolism. Molecular docking studies demonstrated that IKA binds to hexokinase 2 via the tetramic acid within the macrolactam ring. Furthermore, the polycyclic ring of IKA resides in a pocket formed by HK2 residues through hydrophobic interactions (Jiang et al. 2020). According to metabolomic analysis, the P1 compound shares a similar core structure with IKA (Fig. S6). Therefore, the P1 compound likely contains both a tetramic acid ring and a polycyclic structure, which further supports its proposed mechanism of action.

Marine-derived *M. carbonacea* has been reported to produce the anticancer tetrocarcin antibiotic compounds AC6H, tetrocarcin N, and tetrocarcin Q, which exhibit moderate anticancer activity by inactivating the phosphatidylinositol-3 kinase (PI3K) pathway (Hifnawy et al. 2020; Nakajima et al. 2007). These compounds are categorized as macrolide natural products.

Genome mining and antiSMASH analysis revealed that 2MTK254 had the lowest number of BGCs related to secondary metabolites, followed by *M. carbonacea* DSM 43168 $^{T}$ and *M. carbonacea* subsp. *aurantiaca* DSM 43815. The dominant predicted secondary metabolites (hybrid PKS-NRPS, PKS, terpene) were consistent with those previously reported in *Micromonospora* sp. (Yan et al. 2022). Interestingly, strain 2MTK254 exhibited a distinct BGC pattern compared to the closely related strains. This strain showed the highest number of hybrid PKS-NRPS BGCs. Evernimomicin D and sch 18,640 are both reported to have antimicrobial activity against Gram-positive bacteria and have been identified in *M. carbonacea* and *Streptomyces aureus* (Foster and Rybak 1999; Puar et al. 2002). BiG-SCAPE grouped BGCs into GCFs based on the similarity of their Pfam domains. The FAM_00027 cluster was identified to be related to SGR RTMs BGCs, known for their diverse range of biological activity, such as antifungal, antibiotic, and antioxidant properties, which are mainly produced by *Streptomyces* spp. and *Lysobacter* spp. (Luo et al. 2013) but also reported in *Micromonospora* sp.

The polycyclic tetramate macrolactam (PTM) biosynthetic gene clusters (BGCs) typically contain hybrid polyketide synthase/non-ribosomal peptide synthase (PKS/NRPS) modules that initiate the synthesis of a polyene precursor, which serves as the core structure of the compound. The PTM is then cyclized through NAD(P)/FAD-dependent oxidoreductase and alcohol dehydrogenase (ADH) enzymes to form the polycyclic ring. However, the number of these two genes in the BGCs affects the number of rings in the polycyclic structures (Harper et al. 2024). Additionally, cytochrome P450 enzymes play a crucial role in the oxidative transformation of substrates, catalyzing multiple oxidative events (Munro et al. 2018). The predicted PTM BGCs of strain 2MTK254 contain hybrid type I PKS/NRPS, NAD(P)/FAD-dependent oxidoreductase, two alcohol dehydrogenases, sterol desaturase enzymes, and other regulatory genes. Therefore, the cyclization process in 2MTK254 is likely different from that in typical PTMs, with the possibility of forming a 5/6/5 ring structure, as observed in chlokamycin and ikarugamycin-like compounds (Fukuda et al. 2017). The hybrid type I PKS/NRPS in 2MTK254 (contig 2, locus 6109) exhibits 67.2% sequence identity to the *ikaA* gene, which is involved in the synthesis of IKA in *Streptomyces* sp. ZJ306 (Antosch et al. 2014). Contig 2, locus 6108, contains a flavin-containing amine oxidoreductase domain, which displays 66.72% similarity to the *ikaB* gene in IKA and 66.74% identity to the *SGR813/SGR812* genes in the SGR PTM pathway. However, 2MTK254 lacks cytochrome P450 and FAD-dependent oxidoreductases, which results in a distinct cyclization process compared to IKA and SGR PTM compounds (Fig. 8). These findings suggest that strain 2MTK254 possesses unique PTM BGCs. Furthermore, the comparison between the PTM BGCs of 2MTK254 and closely related strains shows that DSM 43815 has a distinct BGC because it lacks alcohol dehydrogenase. In contrast, the other two strains have similar PTM BGCs, but strain DSM 43168 $^{T}$ has additional luxR regulatory transcription factors and a histidine kinase regulator, which controls the expression of PTM-related BGCs in DSM 43168 $^{T}$. This result is supported by metabolomic data, as strain DSM 43168 $^{T}$ did not produce PTM-related compounds.

Evidence from metabolomic analysis suggests that the P1 compound is related to polycyclic tetramate macrolactam (PTM) and may be a derivative of IKA. Secondary metabolite predictions from antiSMASH revealed that contig 2, locus 32, in strain 2MTK254 is associated with PTM BGCs. This BGC shares a similar hybrid type I PKS/NRPS BGC to those involved in the synthesis of IKA. Furthermore, the 2MTK254 BGC also contains the same FAD-dependent oxidoreductase involved in the cyclization of the polycyclic ring, but it includes some unique genes in the biosynthetic cluster, which may be involved in constructing the PTM compound. Thus, this BGC is likely to produce a novel PTM that may be related to the IKA compound.

The in silico annotation required suitable databases and a significant amount of information to predict the compound. However, the open-source library for mass spectrometry annotation (GNPS library) was limited due to the

low number of marine microbial secondary metabolites in the database. The $MS^2$ mass fragmentation posed a challenge due to the low abundance of the target compound and matrix interference. Therefore, this process required partial purification or fractionation before LC-HRMS/MS analysis. However, P1 compound was predicted based on in silico tools combined with mass spectrometry data; therefore, its structure needs to be confirmed using NMR spectroscopy.

This work demonstrates that the novel subspecies *Micromonospora carbonacea* subsp. *caeruleus* 2MTK254 has the potential to produce putative novel polycyclic tetramate macrolactams (PTMs) based on the metabolomic and genomic data obtained, which revealed low-similarity annotated metabolites and BCGs. Additionally, the molecular weight, UV spectrum, and mass fragmentation of P1compound were correlated with those of PTM group. Thus, strain 2MTK254 possesses unique PTM BCGs that may lead to the production of putative novel PTM compounds not previously reported from *Micromonospora* sp.

In conclusion, strain 2MTK254, identified as a novel subspecies of the actinomycete genus *Micromonospora carbonacea*, was isolated from mangrove sediments in Chanthaburi province, Thailand. Through molecular networking integrated with computational tools, a novel compound named P1 ($[C_{29}H_{38}N_2O_4+H]^+$, 479.2904 m/*z*) was identified. P1 exhibited potent anticancer activity against human colorectal cancer cell lines (HCT-116) and high SI index value and was found to be related to unique biosynthetic gene clusters for polycyclic tetramate macrolactams. This study highlights the potential of *Micromonospora carbonacea* subspecies 2MTK254 as a valuable source of anticancer compounds for the pharmaceutical industry. The discovery of P1 underscores the importance of exploring novel microbial sources for bioactive compounds with therapeutic potential.

**Supplementary Information** The online version contains supplementary material available at https://doi.org/10.1007/s00253-025-13427-z.

**Acknowledgements** We are grateful to Assistant Professor Dr. Wongsakorn Phongsopitanun and Miss Pawina Kanchanasin for their technical support in the chemotaxonomic characterization study. The authors would like to thank Dr. Chanwit Suriyachadkun for the fatty acid and menaquinone analysis in the polyphasic taxonomy study and Ms. Boonsita Chanyong for guiding the fragmentation of the chemical structure of MS2 spectra. We also thank M. Iskander for critically proofreading the manuscript.

**Author contribution** TK: Conceptualization, investigation, methodology, validation, data curation, formal analysis, visualization, writing—original draft. NE and CN: Validation, writing—review and editing. UU and TP: Methodology, writing—review and editing. AA and JE: Conceptualization, writing—review and editing. BI: Conceptualization, supervision, writing—review and editing. All authors read and approved the final manuscript.

**Funding** This research project is supported by Mahidol University (Fundamental Fund: fiscal year 2023 by National Science Research and Innovation Fund (NSRF)). Additionally, partial support was provided by the Central Instrument Facility and Center of Nanoimaging at the Faculty of Science, Mahidol University. The author was partially supported by the Faculty of Graduate Studies and the Graduate Studies of Mahidol University Alumni Association for the year 2023.

**Data availability** Strain 2MTK254 was deposited at the Thailand Bioresource Research Center (TBRC) under the accession code TBRC 19103. The 16S rRNA gene sequence and genome sequence data of Micromonospora carbonacea subsp. 2MTK254 is available on the NCBI database under the accession number PP784756 and PRJNA1110966, respectively. The metabolomic data is available on GNPS, including the Feature-based Molecular Network (https://gnps.ucsd.edu/ProteoSAFe/status.jsp?task=c6b0489fab434409bd1177035b1c85f4), Dereplicator plus (https://gnps.ucsd.edu/ProteoSAFe/status.jsp?task=02419c5653c045ee8881936a6f4e6f16), and MolNetEnhancer (https://gnps.ucsd.edu/ProteoSAFe/status.jsp?task=17c23c684b7e410e937ea31648d5717a). The raw data is deposited in the MassIVE database and is available at MSV000094875.

## Declarations

**Ethics approval** This article does not contain any studies with human participants or animals performed by any authors.

**Competing interests** The authors declare no competing interests.



## References

Amin DH, Abolmaaty A, Borsetto C, Tolba S, Abdallah NA, Wellington EMH (2019) In silico genomic mining reveals unexplored bioactive potential of rare actinobacteria isolated from Egyptian soil. Bull Natl Res Cent 43(1):78. https://doi.org/10.1186/s42269-019-0121-y

Amjad MT, Chidharla A, Kasi A (2024) Cancer Chemotherapy StatPearls. StatPearls Publishing. Copyright © 2023, StatPearls Publishing LLC, Treasure Island (FL)

Antosch J, Schaefers F, Gulder TAM (2014) Heterologous reconstitution of ikarugamycin biosynthesis in E. coli. Angew Chem Int Ed 53(11):3011–3014. https://doi.org/10.1002/anie.201310641

Aron AT, Gentry EC, McPhail KL, Nothias L-F, Nothias-Esposito M, Bouslimani A, Petras D, Gauglitz JM, Sikora N, Vargas F, van der Hooft JJJ, Ernst M, Kang KB, Aceves CM,

Caraballo-Rodríguez AM, Koester I, Weldon KC, Bertrand S, Roullier C, Sun K, Tehan RM, Boya PCA, Christian MH, Gutiérrez M, Ulloa AM, Tejeda Mora JA, Mojica-Flores R, Lakey-Beitia J, Vásquez-Chaves V, Zhang Y, Calderón AI, Tayler N, Keyzers RA, Tugizimana F, Ndlovu N, Aksenov AA, Jarmusch AK, Schmid R, Truman AW, Bandeira N, Wang M, Dorrestein PC (2020) Reproducible molecular networking of untargeted mass spectrometry data using GNPS. Nat Protoc 15(6):1954–1991. https://doi.org/10.1038/s41596-020-0317-5

Ashley G, Arslanian RL, Carney J, Metcalf B, Tang L (2003) Preparation of epothilone derivatives for treatment of diseases or conditions characterized by cellular hyperproliferation, such as cancer. US20030045711

Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, Edwards RA, Formsma K, Gerdes S, Glass EM, Kubal M, Meyer F, Olsen GJ, Olson R, Osterman AL, Overbeek RA, McNeil LK, Paarmann D, Paczian T, Parrello B, Pusch GD, Reich C, Stevens R, Vassieva O, Vonstein V, Wilke A, Zagnitko O (2008) The RAST Server: rapid annotations using subsystems technology. BMC Genomics 9:75. https://doi.org/10.1186/1471-2164-9-75

Baindara P, Mandal SM (2020) Bacteria and bacterial anticancer agents as a promising alternative for cancer therapeutics. Biochim 177:164–189. https://doi.org/10.1016/j.biochi.2020.07.020

Balagurunathan R, Radhakrishnan M, Shanmugasundaram T, Gopikrishnan V, Jerrine J (2020) Characterization and Identification of Actinobacteria. In: Balagurunathan R, Radhakrishnan M, Shanmugasundaram T, Gopikrishnan V, Jerrine J (eds) Protocols in Actinobacterial Research. Springer Protocols Handbooks, Springer, US, New York, NY, pp 39–64

Baltz RH (2021) Genome mining for drug discovery: progress at the front end. J Ind Microbiol Biotechnol 48(9–10):kuab044. https://doi.org/10.1093/jimb/kuab044

Blin K, Shaw S, Augustijn HE, Reitz ZL, Biermann F, Alanjary M, Fetter A, Terlouw BR, Metcalf WW, Helfrich EJN, van Wezel GP, Medema MH, Weber T (2023) antiSmash 7.0: new and improved predictions for detection, regulation, chemical structures and visualisation. Nucleic Acids Res 51(W1):W46–W50. https://doi.org/10.1093/nar/gkad344

Chambers MC, Maclean B, Burke R, Amodei D, Ruderman DL, Neumann S, Gatto L, Fischer B, Pratt B, Egertson J, Hoff K, Kessner D, Tasman N, Shulman N, Frewen B, Baker TA, Brusniak M-Y, Paulse C, Creasy D, Flashner L, Kani K, Moulding C, Seymour SL, Nuwaysir LM, Lefebvre B, Kuhlmann F, Roark J, Rainer P, Detlev S, Hemenway T, Huhmer A, Langridge J, Connolly B, Chadick T, Holly K, Eckels J, Deutsch EW, Moritz RL, Katz JE, Agus DB, MacCoss M, Tabb DL, Mallick P (2012) A cross-platform toolkit for mass spectrometry and proteomics. Nat Biotechnol 30(10):918–920. https://doi.org/10.1038/nbt.2377

Collins SJ, Gallo RC, Gallagher RE (1977) Continuous growth and differentiation of human myeloid leukaemic cells in suspension culture. Nature 270(5635):347–349. https://doi.org/10.1038/270347a0

Dharmaraj S (2010) Marine Streptomyces as a novel source of bioactive substances. World J Microbiol Biotechnol 26(12):2123–2139. https://doi.org/10.1007/s11274-010-0415-6

Dührkop K, Fleischauer M, Ludwig M, Aksenov AA, Melnik AV, Meusel M, Dorrestein PC, Rousu J, Böcker S (2019) SIRIUS 4: a rapid tool for turning tandem mass spectra into metabolite structure information. Nat Methods 16(4):299–302. https://doi.org/10.1038/s41592-019-0344-8

Dührkop K, Nothias L-F, Fleischauer M, Reher R, Ludwig M, Hoffmann MA, Petras D, Gerwick WH, Rousu J, Dorrestein PC, Böcker S (2021) Systematic classification of unknown metabolites using high-resolution fragmentation mass spectra. Nat Biotechnol 39(4):462–471. https://doi.org/10.1038/s41587-020-0740-8

El-Sayed ASA, Shindia A, Ammar H, Seadawy MG, Khashana SA (2024) Bioprocessing of epothilone B from Aspergillus fumigatus under solid state fermentation: antiproliferative activity, tubulin polymerization and cell cycle analysis. BMC Microbiol 24(1):43. https://doi.org/10.1186/s12866-024-03184-w

Ernst M, Kang KB, Caraballo-Rodríguez AM, Nothias L-F, Wandy J, Chen C, Wang M, Rogers S, Medema MH, Dorrestein PC, van der Hooft JJJ (2019) MolNetEnhancer: enhanced molecular networks by integrating metabolome mining and annotation tools. Metabolites 9(7):144. https://doi.org/10.3390/metabo9070144

Felsenstein J (1985) Phylogenies and the Comparative Method. Am Nat 125(1):1–15

Ferlay J LM, Ervik M, Lam F, Colombet M, Mery L, Piñeros M, Znaor A, Soerjomataram I, Bray F (2024) Global Cancer Observatory: Cancer Tomorrow (version 1.1). PUblisher. https://gco.iarc.who.int/today Accessed 31 Mar 2024

Fernández S, Arnáiz V, Rufo D, Arroyo Y (2024) Current status of indole-derived marine natural products: synthetic approaches and therapeutic applications. Mar Drugs 22(3):126. https://doi.org/10.3390/md22030126

Foster DR, Rybak MJ (1999) Pharmacologic and bacteriologic properties of SCH-27899 (Ziracin), an investigational antibiotic from the everninomicin family. Pharmacotherapy 19(10):1111–1117. https://doi.org/10.1592/phco.19.15.1111.30576

Fukuda T, Takahashi M, Kasai H, Nagai K, Tomoda H (2017) Chlokamycin, a new chloride from the marine-derived Streptomyces sp. MA2–12. Nat Prod Commun 12(8):1934578X1701200818. https://doi.org/10.1177/1934578x1701200818

Gärtner A, Ohlendorf B, Schulz D, Zinecker H, Wiese J, Imhoff JF (2011) Levantilides A and B, 20-membered macrolides from a Micromonospora strain isolated from the Mediterranean deep sea sediment. Mar Drugs 9(1):98–108. https://doi.org/10.3390/md9010098

Gaudêncio SP, Bayram E, Lukić Bilela L, Cueto M, Díaz-Marrero AR, Haznedaroglu BZ, Jimenez C, Mandalakis M, Pereira F, Reyes F, Tasdemir D (2023) Advanced methods for natural products discovery: bioactivity screening, dereplication, metabolomics profiling, genomic sequencing, databases and informatic tools, and Structure Elucidation. Mar Drugs 21(5):308. https://doi.org/10.3390/md21050308

Genilloud O (2015) Micromonospora BMSAB. pp 1–28

Gilchrist CLM, Chooi Y-H (2021) clinker & clustermap.js: automatic generation of gene cluster comparison figures. Bioinform 37(16):2473–2475. https://doi.org/10.1093/bioinformatics/btab007

Griffin D, Wittmann S, Guo F, Nimmanapalli R, Bali P, Wang HG, Bhalla K (2003) Molecular determinants of epothilone B derivative (BMS 247550) and Apo-2L/TRAIL-induced apoptosis of human ovarian cancer cells. Gynecol Oncol 89(1):37–47. https://doi.org/10.1016/s0090-8258(03)00006-4

Guimarães TC, Gomes TS, Fernandes CD, Barros FD, Oliveira KV, Bilal M, Bharagava RN, Ferreira LFR, Hollanda LM (2020) Antitumor microbial products by actinomycetes isolated from different environments. In: Arora PK (ed) Microbial Technology for Health and Environment. Microorganisms for Sustainability. Springer Singapore, Singapore, pp 113–160

Hardt IH, Steinmetz H, Gerth K, Sasse F, Reichenbach H, Höfle G (2001) New natural epothilones from Sorangium cellulosum, strains So ce90/B2 and So ce90/D13: isolation, structure elucidation, and SAR studies. J Nat Prod 64(7):847–856. https://doi.org/10.1021/np000629f

Harper CP, Day A, Tsingos M, Ding E, Zeng E, Stumpf SD, Qi Y, Robinson A, Greif J, Blodgett JAV (2024) Critical analysis of polycyclic tetramate macrolactam biosynthetic gene cluster phylogeny and functional diversity. Appl Environ Microbiol 90(6):e0060024. https://doi.org/10.1128/aem.00600-24

Hifnawy MS, Fouda MM, Sayed AM, Mohammed R, Hassan HM, AbouZid SF, Rateb ME, Keller A, Adamek M, Ziemert N, Abdelmohsen UR (2020) The genus Micromonospora as a model microorganism for bioactive natural product discovery. RSC Adv 10(35):20939–20959. https://doi.org/10.1039/D0RA04025H

Holguin G, Vazquez P, Bashan Y (2001) The role of sediment microorganisms in the productivity, conservation, and rehabilitation of mangrove ecosystems: an overview. Biol Fertil Soils 33(4):265–278. https://doi.org/10.1007/s003740000319

Hong K, Gao AH, Xie QY, Gao H, Zhuang L, Lin HP, Yu HP, Li J, Yao XS, Goodfellow M, Ruan JS (2009) Actinomycetes for marine drug discovery isolated from mangrove soils and plants in China. Mar Drugs 7(1):24–44. https://doi.org/10.3390/md7010024

Intra B, Euanorasetr J, Nihira T, Panbangred W (2016) Characterization of a gamma-butyrolactone synthetase gene homologue (stcA) involved in bafilomycin production and aerial mycelium formation in Streptomyces sp. SBI034. Appl Microbiol Biotechnol 100(6):2749–2760. https://doi.org/10.1007/s00253-015-7142-8

Jagannathan SV, Manemann EM, Rowe SE, Callender MC, Soto W (2021) Marine actinomycetes. New Sour Biotechnol Prod 19(7):365

Jiang S-H, Dong F-Y, Da L-T, Yang X-M, Wang X-X, Weng J-Y, Feng L, Zhu L-L, Zhang Y-L, Zhang Z-G, Sun Y-W, Li J, Xu M-J (2020) Ikarugamycin inhibits pancreatic cancer cell glycolysis by targeting hexokinase 2. FASEB J 34(3):3943–3955. https://doi.org/10.1096/fj.201901237R

Kämpfer P, Kroppenstedt RM (1996) Numerical analysis of fatty acid patterns of coryneform bacteria and related taxa. Can J Microbiol 42(10):989–1005. https://doi.org/10.1139/m96-128

Kimura M (1980) A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. J Mol Evol 16(2):111–120. https://doi.org/10.1007/bf01731581

Kittiwongwattana C, Thanaboripat D, Laosinwattana C, Koohakan P, Parinthawong N, Thawai C (2015) Micromonospora oryzae sp. nov., isolated from roots of upland rice. Int J Syst Evol Microbiol 65(Pt11):3818–3823. https://doi.org/10.1099/ijsem.0.000500

Komarova NL, Wodarz D (2005) Drug resistance in cancer: principles of emergence and prevention. PNAS 102(27):9714–9719. https://doi.org/10.1073/pnas.0501870102

Kroppenstedt R (1985) Fatty acid and menaquinone analysis of actinomycetes and related organisms. Tech SocAppl Bacteriol:173–199

Kyeremeh K, Acquah KS, Sazak A, Houssen W, Tabudravu J, Deng H, Jaspars M (2014) Butremycin, the 3-hydroxyl derivative of ikarugamycin and a protonated aromatic tautomer of 5′-methylthioinosine from a Ghanaian Micromonospora sp. K310. Mar Drugs 12(2):999–1012

Law JW-F, Law LN-S, Letchumanan V, Tan LT-H, Wong SH, Chan K-G, Ab Mutalib N-S, Lee L-H (2020) Anticancer drug discovery from microbial sources: the unique mangrove streptomycetes. Molecules 25(22):5365. https://doi.org/10.3390/molecules25225365

Lechevalier MP, Lechevalier H (1970) Chemical composition as a criterion in the classification of aerobic actinomycetes. Int J Syst Evol Microbiol 20(4):435–443. https://doi.org/10.1099/00207713-20-4-435

Lechevalier MP, De Bievre C, Lechevalier H (1977) Chemotaxonomy of aerobic actinomycetes: phospholipid composition. Biochem Syst Ecol 5(4):249–260. https://doi.org/10.1016/0305-1978(77)90021-7

Luo Y, Huang H, Liang J, Wang M, Lu L, Shao Z, Cobb RE, Zhao H (2013) Activation and characterization of a cryptic polycyclic tetramate macrolactam biosynthetic gene cluster. Nat Commun 4(1):2894. https://doi.org/10.1038/ncomms3894

Mehetre GT, Vinodh JS, Burkul BB, Desai D, Santhakumari B, Dharne MS, Dastager SG (2019) Bioactivities and molecular networking-based elucidation of metabolites of potent actinobacterial strains isolated from the Unkeshwar geothermal springs in India. RSC Adv 9(17):9850–9859. https://doi.org/10.1039/C8RA09449G

Meier-Kolthoff JP, Auch AF, Klenk H-P, Göker M (2013) Genome sequence-based species delimitation with confidence intervals and improved distance functions. BMC Bioinformatics 14(1):60. https://doi.org/10.1186/1471-2105-14-60

Minnikin DE, O'donnell AG, Goodfellow M, Alderson G, Athalye M, Schaal A, Parlett JH (1984) An integrated procedure for the extraction of bacterial isoprenoid quinones and polar lipidS. J Microbiol Methods 2(5):233–241

Mosmann T (1983) Rapid colorimetric assay for cellular growth and survival: application to proliferation and cytotoxicity assays. J Immunol Methods 65(1–2):55–63. https://doi.org/10.1016/0022-1759(83)90303-4

Munro AW, McLean KJ, Grant JL, Makris TM (2018) Structure and function of the cytochrome P450 peroxygenase enzymes. Biochem Soc Trans 46(1):183–196. https://doi.org/10.1042/bst20170218

Na S-I, Kim YO, Yoon S-H, Ha S-m, Baek I, Chun J (2018) UBCG: Up-to-date bacterial core gene set and pipeline for phylogenomic tree reconstruction. J Microbiology 56(4):280–285. https://doi.org/10.1007/s12275-018-8014-6

Nakajima H, Sakaguchi K, Fujiwara I, Mizuta M, Tsuruga M, Magae J, Mizuta N (2007) Apoptosis and inactivation of the PI3-kinase pathway by tetrocarcin A in breast cancers. Biochem Biophys Res Commun 356(1):260–265. https://doi.org/10.1016/j.bbrc.2007.02.136

Navarro-Muñoz JC, Selem-Mojica N, Mullowney MW, Kautsar S, Tryon JH, Parkinson EI, Santos ELCDL, Yeong M, Cruz-Morales P, Abubucker S, Roeters A, Lokhorst W, Fernandez-Guerra A, Cappelini LTD, Thomson RJ, Metcalf WW, Kelleher NL, Barona-Gomez F, Medema MH (2018) A computational framework for systematic exploration of biosynthetic diversity from large-scale genomic data. bioRxiv:445270 https://doi.org/10.1101/445270

Navarro-Muñoz JC, Selem-Mojica N, Mullowney MW, Kautsar SA, Tryon JH, Parkinson EI, De Los Santos ELC, Yeong M, Cruz-Morales P, Abubucker S, Roeters A, Lokhorst W, Fernandez-Guerra A, Cappelini LTD, Goering AW, Thomson RJ, Metcalf WW, Kelleher NL, Barona-Gomez F, Medema MH (2020) A computational framework to explore large-scale biosynthetic diversity. Nat Chem Biol 16(1):60–68. https://doi.org/10.1038/s41589-019-0400-9

Newman DJ, Cragg GM (2020) Natural products as sources of new drugs over the nearly four decades from 01/1981 to 09/2019. J Nat Prod 83(3):770–803. https://doi.org/10.1021/acs.jnatprod.9b01285

Ngamcharungchit C, Matsumoto A, Suriyachadkun C, Panbangred W, Inahashi Y, Intra B (2023) Nonomuraea corallina sp. nov., isolated from coastal sediment in Samila Beach, Thailand: insights into secondary metabolite synthesis as anticancer potential. Front Microbiol 14:1226945. https://doi.org/10.3389/fmicb.2023.1226945

Nothias L-F, Petras D, Schmid R, Dührkop K, Rainer J, Sarvepalli A, Protsyuk I, Ernst M, Tsugawa H, Fleischauer M, Aicheler F, Aksenov AA, Alka O, Allard P-M, Barsch A, Cachet X, Caraballo-Rodriguez AM, Da Silva RR, Dang T, Garg N, Gauglitz JM, Gurevich A, Isaac G, Jarmusch AK, Kameník Z, Kang KB, Kessler N, Koester I, Korf A, Le Gouellec A, Ludwig M, Martin HC, McCall L-I, McSayles J, Meyer SW, Mohimani H, Morsy M, Moyne O, Neumann S, Neuweger H, Nguyen NH, Nothias-Esposito M, Paolini J, Phelan VV, Pluskal T, Quinn RA, Rogers S, Shrestha B, Tripathi A, van der Hooft JJJ, Vargas F, Weldon KC, Witting M, Yang H, Zhang Z, Zubeil F, Kohlbacher O, Böcker S, Alexandrov T, Bandeira N, Wang M, Dorrestein PC (2020) Feature-based molecular networking in the GNPS analysis

environment. Nat Methods 17(9):905–908. https://doi.org/10.1038/s41592-020-0933-6
Olano C, Méndez C, Salas JA (2009) Antitumor compounds from marine actinomycetes. Mar Drugs 7(2):210–248. https://doi.org/10.3390/md7020210
Ossai J, Khatabi B, Nybo SE, Kharel MK (2022) Renewed interests in the discovery of bioactive actinomycete metabolites driven by emerging technologies. J Appl Microbiol 132(1):59–77. https://doi.org/10.1111/jam.15225
Overbeek R, Olson R, Pusch GD, Olsen GJ, Davis JJ, Disz T, Edwards RA, Gerdes S, Parrello B, Shukla M, Vonstein V, Wattam AR, Xia F, Stevens R (2014) The SEED and the rapid annotation of microbial genomes using subsystems technology (RAST). Nucleic Acids Res 42(Databasee issue):D206-14. https://doi.org/10.1093/nar/gkt1226
Pearce ME, Langridge GC, Lauer AC, Grant K, Maiden MCJ, Chattaway MA (2021) An evaluation of the species and subspecies of the genus Salmonella with whole genome sequence data: proposal of type strains and epithets for novel S. enterica subspecies VII, VIII, IX X and XI. Genom 113(5):3152–3162. https://doi.org/10.1016/j.ygeno.2021.07.003
Peña-Morán OA, Villarreal ML, Álvarez-Berber L, Meneses-Acosta A, Rodríguez-López V (2016) Cytotoxicity, post-treatment recovery, and selectivity analysis of naturally occurring podophyllotoxins from Bursera fagaroides var. fagaroides on breast cancer cell lines. Molecules 21(8):1013. https://doi.org/10.3390/molecules21081013
Puar MS, Ganguly AK, Afonso A, Brambilla R, Mangiaracina P, Sarre O, MacFarlane RD (2002) Sch 18640. A new thiostrepton-type antibiotic. J Am Chem Soc 103(17):5231–5233. https://doi.org/10.1021/ja00407a047
Richter M, Rosselló-Móra R (2009) Shifting the genomic gold standard for the prokaryotic species definition. PNAS 106(45):19126–19131. https://doi.org/10.1073/pnas.0906412106
Richter M, Rosselló-Móra R, Oliver Glöckner F, Peplies J (2015) JSpeciesWS: a web server for prokaryotic species circumscription based on pairwise genome comparison. Bioinform 32(6):929–931. https://doi.org/10.1093/bioinformatics/btv681
Rzhetsky A, Nei M (1993) Theoretical foundation of the minimum-evolution method of phylogenetic inference. Mol Biol Evol 10(5):1073–1095. https://doi.org/10.1093/oxfordjournals.molbev.a040056
Schmid R, Heuckeroth S, Korf A, Smirnov A, Myers O, Dyrlund TS, Bushuiev R, Murray KJ, Hoffmann N, Lu M, Sarvepalli A, Zhang Z, Fleischauer M, Dührkop K, Wesner M, Hoogstra SJ, Rudt E, Mokshyna O, Brungs C, Ponomarov K, Mutabdžija L, Damiani T, Pudney CJ, Earll M, Helmer PO, Fallon TR, Schulze T, Rivas-Ubach A, Bilbao A, Richter H, Nothias L-F, Wang M, Orešič M, Weng J-K, Böcker S, Jeibmann A, Hayen H, Karst U, Dorrestein PC, Petras D, Du X, Pluskal T (2023) Integrative analysis of multimodal mass spectrometry data in MZmine 3. Nat Biotechnol 41(4):447–449. https://doi.org/10.1038/s41587-023-01690-2
Schulze CJ, Donia MS, Siqueira-Neto JL, Ray D, Raskatov JA, Green RE, McKerrow JH, Fischbach MA, Linington RG (2015) Genome-directed lead discovery: biosynthesis, structure elucidation, and biological evaluation of two families of polyene macrolactams against Trypanosoma brucei. ACS Chem Biol 10(10):2373–2381. https://doi.org/10.1021/acschembio.5b00308
Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T (2003) Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res 13(11):2498–2504. https://doi.org/10.1101/gr.1239303
Shirling EB, Gottlieb D (1966) Methods for characterization of Streptomyces species. Int J Syst Bacteriol 16(3):313–340. https://doi.org/10.1099/00207713-16-3-313
Soergel DA, Dey N, Knight R, Brenner SE (2012) Selection of primers for optimal taxonomic classification of environmental 16S rRNA gene sequences. Isme j 6(7):1440–1444. https://doi.org/10.1038/ismej.2011.208
Staneck JL, Roberts GD (1974) Simplified approach to identification of aerobic actinomycetes by thin-layer chromatography. Appl Microbiol 28(2):226–231. https://doi.org/10.1128/am.28.2.226-231.1974
Subramani R, Sipkema D (2019) Marine rare actinomycetes: a promising source of structurally diverse and unique novel natural products. Mar Drugs 17(5):249. https://doi.org/10.3390/md17050249
Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, Bray F (2021) Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA: Cancer J Clin 71(3):209–249. https://doi.org/10.3322/caac.21660
Tan Y, Hu Y, Wang Q, Zhou H, Wang Y, Gan M (2016) Tetrocarcins N and O, glycosidic spirotetronates from a marine-derived Micromonospora sp. identified by PCR-based screening. RSC Adv 6(94):91773–91778. https://doi.org/10.1039/C6RA17026A
Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res 22(22):4673–4680. https://doi.org/10.1093/nar/22.22.4673
van der Hooft JJJ, Mohimani H, Bauermeister A, Dorrestein PC, Duncan KR, Medema MH (2020) Linking genomics and metabolomics to chart specialized metabolic diversity. Chem Soc Rev 49(11):3297–3314. https://doi.org/10.1039/D0CS00162G
Venter SN, Palmer M, Steenkamp ET (2022) Relevance of prokaryotic subspecies in the age of genomics. New Microbes New Infect 48:101024. https://doi.org/10.1016/j.nmni.2022.101024
Xu L, Wu P, Wright SJ, Du L, Wei X (2015) Bioactive polycyclic tetramate macrolactams from Lysobacter enzymogenes and their absolute configurations by theoretical ECD calculations. J Nat Prod 78(8):1841–1847. https://doi.org/10.1021/acs.jnatprod.5b00099
Xu M, Zhang Z, Xu J, Yang X, He K, Yu H, Jiang S (2016) Preparation of polycyclic fused macrocyclic lactam compounds for treatment of pancreas cancer. CN106008531
Yan S, Zeng M, Wang H, Zhang H (2022) Micromonospora: a prolific source of bioactive secondary metabolites with therapeutic potential. J Med Chem 65(13):8735–8771. https://doi.org/10.1021/acs.jmedchem.2c00626
Zhu B, Panek JS (2000) Total synthesis of epothilone A. Org Lett 2(17):2575–2578. https://doi.org/10.1021/ol006104w