Genomic study and lipidomic bioassay of *Leeuwenhoekiella parthenopeia*: A novel rare biosphere marine bacterium that inhibits tumor cell viability

Giuliano Gattoni[1], Rafael R. de la Haba[2], Jesús Martín[3], Fernando Reyes[3], Cristina Sánchez-Porro[2], Antonia Feola[1], Candida Zuchegna[1], Shaday Guerrero-Flores[4], Mario Varcamonti[1], Ezio Ricca[1], Nelly Selem-Mojica[4], Antonio Ventosa[2] and Paulina Corral[1,2]*

[1]Department of Biology, University of Naples Federico II, Naples, Italy, [2]Department of Microbiology and Parasitology, Faculty of Pharmacy, University of Sevilla, Sevilla, Spain, [3]Fundación MEDINA, Granada, Spain, [4]Centro de Ciencias Matemáticas, Universidad Nacional Autónoma de México (UNAM), Morelia, Mexico

The fraction of low-abundance microbiota in the marine environment is a promising target for discovering new bioactive molecules with pharmaceutical applications. Phenomena in the ocean such as diel vertical migration (DVM) and seasonal dynamic events influence the pattern of diversity of marine bacteria, conditioning the probability of isolation of uncultured bacteria. In this study, we report a new marine bacterium belonging to the rare biosphere, *Leeuwenhoekiella parthenopeia* sp. nov. Mr9[T], which was isolated employing seasonal and diel sampling approaches. Its complete characterization, ecology, biosynthetic gene profiling of the whole genus *Leeuwenhoekiella*, and bioactivity of its extract on human cells are reported. The phylogenomic and microbial diversity studies demonstrated that this bacterium is a new and rare species, barely representing 0.0029% of the bacterial community in Mediterranean Sea metagenomes. The biosynthetic profiling of species of the genus *Leeuwenhoekiella* showed nine functionally related gene cluster families (GCF), none were associated with pathways responsible to produce known compounds or registered patents, therefore revealing its potential to synthesize novel bioactive compounds. *In vitro* screenings of *L. parthenopeia* Mr9[T] showed that the total lipid content (lipidome) of the cell membrane reduces the prostatic and brain tumor cell viability with a lower effect on normal cells. The lipidome consisted of sulfobacin A, WB 3559A, WB 3559B, docosenamide, topostin B-567, and unknown compounds. Therefore, the bioactivity could be attributed to any of these individual compounds or due to their synergistic effect. Beyond the rarity and biosynthetic potential of this bacterium, the importance and novelty of this study is the employment of sampling strategies based on

ecological factors to reach the hidden microbiota, as well as the use of bacterial membrane constituents as potential novel therapeutics. Our findings open new perspectives on cultivation and the relationship between bacterial biological membrane components and their bioactivity in eukaryotic cells, encouraging similar studies in other members of the rare biosphere.

1. Introduction

The exploration of the uncultivated marine microbiota is a promising target in the search for alternative drugs to meet the urgent therapeutic needs where the current pharmacotherapy is ineffective (García-Davis et al., 2021; Zhou et al., 2021). Every year, new cases of cancer are registered worldwide, and many of them are resistant forms that do not respond to available treatments (Maeda and Khatami, 2018). This low therapeutic efficacy reduces life expectancy, as cancer is among the main leading causes of death (Siegel et al., 2021). In this scenario, the isolation of uncultured bacteria is a current trend, in addition to culture-independent studies, since there is a need to explore the true metabolic and intrinsic characteristics both in their habitat and *in vitro* (Pascoal et al., 2020; Sanz-Sáez et al., 2020). Most of the discovered bioactive compounds were derived from only 1% of the total microbial diversity known so far (Baker et al., 2021). An outstanding example is a salinosporamide A, an agent in phase III clinical trial for the treatment of glioblastoma produced by the rare marine actinobacteria of the genus *Salinispora* (Fenical et al., 2009; Roth et al., 2020; Boccellato et al., 2021). Despite this successful precedent, the huge biosynthetic potential remains in the great prokaryotic diversity not yet cultivated as predicted by microbiome studies (Paoli et al., 2022). The low rate of cultivation of new taxa is even more challenging when targeting the very low-abundant fraction (0.1 to 0.01% of relative abundance) known as the rare biosphere (Sogin et al., 2006; Pedrós-Alió, 2007). To overcome this limitation, prokaryotic diversity is being addressed by considering environmental variabilities that affect the microbial population distribution, such as biomass migration and seasonality (Shade et al., 2014; Cram et al., 2015; Mestre et al., 2020).

With the aim to cultivate novel and rare bacteria with anticancer potential, here, we employed sampling strategies considering seasonality and natural dynamic processes in the marine environment rather than culture techniques *per se*. The largest biomass dynamic event in the ocean is the diel vertical migration (DVM) where a part of the inhabitants, especially zooplankton, ascends from the depths to the surface waters during the night or light absence and moves downward during

the day (Bosch and Rowland Taylor, 1973; Omand et al., 2021). In the same way, seasonality influences the pattern of microbial diversity as the water column in the ocean stratifies during the warm period and mixes during the cold months (Giovannoni and Vergin, 2012; Haro-Moreno et al., 2018; Mena et al., 2020). In this study, cold seasons and night were identified as the best sampling time since the water column is mixed and DVM brings planktonic particulate as drivers of microbial transfer from the depths during the night (Gilbert et al., 2012; Fuhrman et al., 2015; Kelly et al., 2019). In both cases, the microbial community is redistributed, increasing the possibility of reaching uncommon bacteria. Furthermore, it has been demonstrated that surface and deep-sea prokaryotic communities are strongly connected in the oceans through the transport of sinking particles (Mestre et al., 2018; Wenley et al., 2021). By employing this approach, a novel and rare marine bacterium, that we have designated as *Leeuwenhoekiella parthenopeia* strain Mr9[T], was isolated from reef seawater in the Gulf of Naples, Italy, from night sampling series during the late autumn. Herein, we fully characterized this strain and inferred its specialized metabolism through the genome by predicting unknown biosynthetic gene clusters (BGCs) of pharmacological relevance (Agrawal et al., 2017; Belknap et al., 2020; Scherlach and Hertweck, 2021). Similarly, the biosynthetic potential of all described species of the genus *Leeuwenhoekiella* was genomically assessed by determining the functional homologies based on gene cluster families (GCFs) (Navarro-Muñoz et al., 2020). Considering that most BGCs of the secondary metabolism are not expressed *in vitro* or are not released outside the cell (Letzel et al., 2017), such as in the case of *L. parthenopeia* Mr9[T], we investigated the structural constituents of the bacterial membrane, specifically the lipidome. The bacterial lipid composition variates to adapt in response to environmental changes maintaining the optimal functional properties of the membrane (Angelini et al., 2012; Chwastek et al., 2020). The lipid profile is used as a chemotaxonomic marker in the characterization of bacteria and archaea and can be different among species of the same genus (de la Haba et al., 2018), but there are no studies on bacterial membrane lipid composition and its effect on eukaryotic cells as such. Our study demonstrates that the

total lipid extract (TLE) from *L. parthenopeia* Mr9^T affects tumor cell viability of prostate adenocarcinoma DU-145, and glioblastoma U-87 MG cell lines with a lower effect on the normal cell line HaCaT. The genus *Leeuwenhoekiella*, the object of this study, includes chemoorganotrophic Gram-negative bacteria, producing non-diffusible yellow-pigmented colonies with gliding motility (Nedashkovskaya et al., 2005; Tahon et al., 2020). Currently, this genus includes only seven species (Parte et al., 2020) derived from marine or sea-related habitats, and no bioactive compounds that impair tumor cell viability produced by strains of these species have been reported to date.

2. Materials and methods

2.1. Sampling and isolation

Serial weekly nightly samplings were carried out in the reef of the Gulf of Naples, Italy 40°49′48.9″N 14°13′31.8″E, during the fall and winter of 2020–2021. The selection of diel and cold seasons was based on previous studies (Gilbert et al., 2012; Fuhrman et al., 2015; Kelly et al., 2019) that demonstrate that the water column is mixed during these periods and DVM brings planktonic particulate acting as drivers of microbial transfer from the depths during the night, thus suggesting fall, winter, and night as the best sampling time. The salinity, pH, conductivity, particulate (total dissolved solids, TDS), and temperature were measured *in situ* with a refractometer and a digital multiparameter pH/EC/TDS/TEMP, Hanna HI-9812-51.

Five liters of seawater were collected from shallow waters (∼1 m) each time and homogenized by magnetic stirring. Aliquots of 40 mL of each sample were vortexed for 10 min, and 1 mL was serially diluted 10-fold in 9 mL of sterile seawater to a 10^{-6} dilution. A total of 100 μL of each dilution was inoculated on four different solid minimal media designed to cultivate oligotrophic and heterotrophic bacteria, formulated with filter-sterilized natural seawater as the base component and varying the mineral composition. The plates were incubated on the laboratory bench to follow the light day cycle at room temperature (∼21°C) and aerobic conditions. The growth was checked daily, and the smallest colonies were picked and streaked on the same medium until obtaining axenic cultures.

The strain Mr9^T was isolated from the night sampling series collected in November 2020 at 1-m depth after 4 days of incubation in seawater oligotrophic medium (SWOM) with the following composition (g/L of natural seawater): casamino acids, 0.5; yeast extract, 1.0; tryptone, 1.0; and agar, 15.0. This medium was supplemented with glycerol, 0.1 M; mineral solution, 500 μL/L; and vitamin solution, 500 μL/L. All supplement solutions were previously sterilized by filtration and added after autoclaving.

The mineral solution compositions are as follows (mg/L): *Na-EDTA, 250.0; ZnSO$_4$ 7H$_2$O, 1095.0; FeSO$_4$ 7 H$_2$O, 500; MnSO$_4$ H$_2$O, 154; CuSO$_4$ 5H$_2$O, 39.2; Co(NO$_3$)$_2$ 6H$_2$O, 24.8;

Na$_2$B$_4$O$_7$ 10H$_2$O, 17.7. Distilled water up to 1,000 mL. *Dissolve the EDTA and add a few drops of concentrated H$_2$SO$_4$ to retard the precipitation of the heavy metal ions.

Vitamin solution composition (mg/L): Biotin, 4.0; folic acid, 4.0; pyridoxine-HCl, 20.0; riboflavin, 10.0; thiamine-HCl 2H$_2$O, 10.0; nicotinamide, 10.0; D-Ca-pantothenate, 10.0; vitamin B$_{12}$, 0.20. Distilled water up to 1,000 mL. Stored in dark at 4°C.

After primary isolation and purification in SWOM, no growth was observed in synthetic medium marine agar (MA) 2260 (Difco), but the strain Mr9^T was adapted to grow on MA medium by adding the same supplement solutions as those used for SWOM, where a poor growth was visible after 10 days of incubation. After successive cultivation, the strain Mr9^T was well adapted to routine growth in MA without supplemental solutions, establishing the optimal growing condition to 4 days of incubation at 25°C. For long-term conservation, strain Mr9^T was cultivated in marine broth (MB) for 72 h and stored at −80°C in cryovials with 40% (v/v) of glycerol.

2.2. Characterization of new taxa

2.2.1. Strain identification by amplification and sequencing of 16S rRNA gene

For the identification of strain Mr9^T, the 16S rRNA gene was amplified using universal primers: 27F (5′-AGAG TTTGATCCTGGCTCAG-3′) and 1492R (5′-GGTTACCTT GTTACGACTT-3′). The PCR amplification was performed using a single colony of a 48-h culture of Mr9^T and resuspending it in 40 μL of a master mix as follows: *Reaction buffer 5×, 10 μL; primer forward 27F (20 μM), 1 μL; primer reverse 1492R (20 μM), 1 μL; *Taq* DNA polymerase MyTaq (Bioline), 1 μL; and H$_2$O nuclease-free, 27 μL. *Reaction buffer 5× contains dNTPs, MgCl$_2$, and enhancers. The reaction was performed using a thermal cycler T100 (BioRad, CA, USA) with the following settings: Initial denaturation for 5′ at 95°C, 25 cycles of 30″ at 94°C, 30″ at 50°C, 90″ at 72°C, and final extension for 10′ at 72°C.

The amplification was verified by 0.8% agarose gel electrophoresis in TBE and stained with SYBR Safe DNA Gel Stain 10,000× in dimethyl sulfoxide (DMSO) (Invitrogen, MA, USA), and the size was checked using the 1 Kb Plus DNA Ladder (Invitrogen, MA, USA).

The PCR product was purified using a silica-based membrane GeneJET PCR purification kit (Invitrogen, MA, USA) following the manufacturer's protocol and sequenced by Stab Vida (Portugal) using the Sanger method. Sequence quality analysis and trimming were accomplished using the software 4Peaks V1.8 and contig assembly of the whole gene with ChromasPro [Research Resource Identifiers (RRIDs)]: (RRID: SCR_000229) v.2.1.10.

Sequence similarity search was performed against the EzBioCloud database (Yoon et al., 2017) with the 16S-based ID application, and with the NCBI BLAST (RRID: SCR_004870)

against NCBI nr/nt Nucleotide (RRID: SCR_004860) database restricted to type-sequenced taxa and then also including uncultured/environmental sample sequences.

2.2.2. Genome sequencing, assembly, annotation, and relatedness indexes

Genomic DNA was extracted using the DNeasy UltraClean kit (QIAGEN, Hilden, Germany) following the protocol for Gram-negative bacteria. DNA purity and integrity were checked by agarose gel electrophoresis (0.8%) and using NanoDrop 2000 Spectrophotometer (Thermo Scientific, MA, USA). The DNA was quantified employing the Qubit assay kit for double-stranded DNA broad range (dsDNA BR, Invitrogen) and measured using the Qubit 4 Fluorometer (Invitrogen, MA, USA).

The genome of strain Mr9^T was sequenced using Illumina NovaSeq 6000 platform (2 × 150 bp paired-end sequencing reads) at Novogene (Genomic sequencing center, Cambridge, UK). Quality control of raw reads was checked with FastQC (RRID: SCR_014583) v.0.11.9 and reads containing adapters and low-quality bases ($Q \leq 20$) were removed with Trimmomatic (RRID: SCR_011848) v.0.36 (Bolger et al., 2014). De novo assembly of the reads was performed using SPAdes (RRID: SCR_000131) v.3.15.4 (Bankevich et al., 2012) and SOAP3 (RRID: SCR_005502) (Liu et al., 2012). The quality of final contigs was assessed using CheckM (RRID: SCR_016646) v.1.0.5 (Parks et al., 2015) and QUAST (RRID: SCR_001228) v.5.1.0rc1 (Gurevich et al., 2013). The genome sequence was annotated using the NCBI Prokaryotic Genomes Automatic Annotation Pipeline (PGAP) (RRID: SCR_021329) (Tatusova et al., 2016).

To assess the taxonomic status of strain Mr9^T as a new taxon, the genome sequence similarities among closely related species were estimated by computing the overall genome relatedness index (OGRI) (Chun and Rainey, 2014). The average nucleotide identity (ANI) was calculated using the OrthoANIu tool (RRID: SCR_022562) (Liu et al., 2016). The digital DNA–DNA hybridization (dDDH) was calculated using the Genome-to-Genome Distance Calculator (GGDC) website, with formula 2 (Meier-Kolthoff et al., 2013). This genomic comparison was between strain Mr9^T and type species of validly published species names as well as two genomes of non-type cultivated species of the genus Leeuwenhoekiella. The results were evaluated based on the cutoff for species boundaries; ANI, 95–96% (Konstantinidis and Tiedje, 2005; Goris et al., 2007; Richter and Rosselló-Móra, 2009; Chun and Rainey, 2014), and dDDH, 70% (Stackebrandt and Goebel, 1994; Auch et al., 2010).

2.2.3. Phylogenetic and phylogenomic analysis

To infer the 16S rRNA gene-based phylogeny, the respective sequences of strain Mr9^T and its close relative species of the genus Leeuwenhoekiella were downloaded from DDBJ/ENA/GenBank databases. After sequence alignment,

phylogenetic trees were constructed using ARB v.6.0.5 software package (Westram et al., 2011) with neighbor-joining (NJ) (Saitou and Nei, 1987), maximum-parsimony (MP) (Fitch, 1971), and maximum-likelihood (ML) (Felsenstein, 1981) algorithms. The distance matrix was corrected by applying the Jukes–Cantor model of DNA evolution (Jukes and Cantor, 1969). To infer ML phylogeny, the general time-reversible model (Tavaré and Miura, 1986) with gamma distribution and proportion of invariant sites to estimate rate heterogeneity over sites (GTR + Γ + I) were used. Branch support was assessed by 1,000 bootstrap pseudoreplicates (Felsenstein, 1985).

Since phylogenies based on the 16S rRNA gene sequence comparison have been demonstrated not to be totally reliable in determining evolutionary relationships (Corral et al., 2018; de la Haba et al., 2018; Infante-Domínguez et al., 2020), a more robust genome-based phylogeny was inferred by constructing a phylogenomic tree. For this purpose, 11 genome sequences of cultivated species of the genus Leeuwenhoekiella were retrieved from the NCBI GenBank database (**Supplementary Table 1**). Of those genomes, nine corresponded to type strains comprising seven Leeuwenhoekiella species (given the fact that two genomes from each of the type strains of L. palythoae and Leeuwenhoekiella marinoflava were available in the public databases, although with different culture collection designation key); the other two out of 11 genomes corresponded to cultivated non-type strains but closely related to Leeuwenhoekiella sp. strain Mr9^T.

Pan- and core-genome datasets were determined using an all-vs.-all BLAST comparisons among the coding sequences (CDS) and translated CDS features of the annotated genomes under study, as previously described (de la Haba et al., 2019). Then, single-copy core gene and amino acid sequences were individually aligned with muscle (Edgar, 2004) and concatenated into a super-nucleotide or super-protein alignment, which was further analyzed to generate the phylogenomic tree using the appropriate ML algorithm implemented in FastTreeMP v.2.1.8 (Price et al., 2010). Generalized time-reversible (GTR) model of nucleotide evolution (Tavaré and Miura, 1986) and Jones–Taylor–Thornton (JTT) model of amino acid evolution (Jones et al., 1992), both with a single rate for each site (CAT), were applied for nucleotide- and amino acid-based phylogenomic reconstructions, respectively. Tree branch support was inferred using the Shimodaira–Hasegawa test (Shimodaira and Hasegawa, 1999). For the display, annotation, and management of 16S rRNA gene-based and phylogenomic trees, the online tool iTOL (RRID: SCR_018174) v.6.5.2 (Letunic and Bork, 2021) was used.

2.2.4. Phenotypic and physiological characterization

The macroscopic features of Leeuwenhoekiella sp. strain Mr9^T were based on the colony morphology in solid media

SWOM and MA during incubation from the 4th to the 30th day. The colony appearance such as shape, pigmentation, and consistency was visually evaluated. Cell morphology, size, and motility were observed by phase-contrast microscopy (Olympus BX41, Tokio, Japan) from a 72-h liquid culture incubated at 25°C with agitation. The optimal growth conditions and tolerance to temperature, salinity (NaCl), and pH were determined in the following ranges: temperature, 4°C–55°C, intervals of 10°C; NaCl, 0–20% (w/v), intervals of 0.5%; and pH, 3.0–12.0, intervals of 1.0.

The basal medium for salinity and pH tests contained 0.1% (w/v) of yeast extract. For the salinity, the basal medium was prepared in distilled water, and for the pH test, the salinity was adjusted to 3.5% (w/v) of NaCl (sea salts, Sigma-Aldrich, MO, USA) and then adjusted with MES (pH 5.0–6.0), MOPS (pH 6.5–7.0), Tris (pH 7.5–8.5), or CHES (pH 9.0) at a final concentration of 50 mM.

Marine broth (MB) and seawater oligotrophic medium (SWOM) broth were used for temperature testing, both with original formulations to observe the difference in temperature among synthetic and natural media. For each test, 2 mL of the basal medium was inoculated with 100 μL of a 48-h liquid culture (0.5 McFarland scale = 1.5×10^8 CFU/mL). All tubes were incubated with agitation at 25°C, except those for the temperature tolerance test.

The capacity of strain Mr9^T to grow under anaerobic conditions was determined by inoculating MA and SWOM agar plates and incubation at 25°C for 72 h in an anaerobic jar with an anaerobic gas generator and anaerobic indicator (Oxoid). Oxidase activity was examined by using oxidase sticks (PanReac AppliChem, Darmstadt, Germany) containing tetramethyl-p-phenylenediamine. Catalase activity was assessed by adding a 3% (w/v) H_2O_2 solution to colonies on a solid medium.

The biochemical and physiological tests were evaluated following the methods for characterization tests described by Barrow and Feltham (1993). Hydrolysis of Tween 80, gelatin, and DNA, urease, indole production from tryptophan, methyl red and Voges–Proskauer tests, Simmons's citrate, production of H_2S, and anaerobic growth in the presence of nitrate, arginine, or DMSO were evaluated supplementing the medium with 3.5% (w/v) of NaCl. The physiological profiling was approached by testing the assimilation of a battery of carbohydrates, alcohols, organic acids, and amino acids as a sole source of carbon and energy. The basal medium contained 0.05% (w/v) of yeast extract and 3.5% (w/v) of NaCl supplemented with 1% (w/v) of the tested substrate.

For comparative purposes, the type strain of the closest species to *Leeuwenhoekiella* sp. strain Mr9^T according to the 16S rRNA gene sequence comparison, *Leeuwenhoekiella nanhaiensis* KCTC 42729^T, and the type strain of the type species of the genus, *L. marinoflava* Laboratory of Microbiology of Ghent University (LMG) 1345^T, were purchased from the Korean Collection for Type Cultures

(KCTC) and from the Laboratory of Microbiology of Ghent University/Belgian Coordinated Collections of Microorganisms (BCCM), respectively. All phenotypic and physiological tests were determined in triplicate including the reference strains and under the same laboratory conditions. The results were reported after 48–72 h, based on the controls (without inoculum).

2.2.5. Chemotaxonomic analysis

The polar lipid analysis of *Leeuwenhoekiella* sp. strain Mr9^T was carried out by DSMZ services, Leibniz Institute DSMZ-German Collection of Microorganisms and Cell Cultures (Germany), from the deposited strain *L. parthenopeia* Mr9^T (DSM 112950). In brief, polar lipids were extracted using the modified Bligh and Dyer's (1959) method and separated by two-dimensional silica gel thin-layer chromatography. Total lipids were revealed by spraying molybdate-phosphoric acid and specific reagents to detect defined functional groups.

The composition of cellular fatty acids methyl esters was also determined following the protocol recommended by the MIDI Microbial Identification System (Sasser, 1990). This determination was carried out at the CECT, Spanish Type Culture Collection (Valencia, Spain). The biomass of strain Mr9^T for the fatty acid determination was obtained from a culture on MA after 72 h at 30°C of incubation. The cellular content of fatty acids was analyzed by gas chromatography with an Agilent 6850 gas chromatograph and identified according to the TSBA6 method using the Microbial Identification Sherlock software package (MIDI, 2012).

2.3. Relative abundance of *Leeuwenhoekiella* sp. strain Mr9^T in marine habitats

To explore the occurrence of the *Leeuwenhoekiella* sp. strain Mr9^T in marine microbial communities from different habitats, the genome of this strain was mapped against 41 metagenomes of the Mediterranean Sea. The mapping involved two metagenomic collections from Mediterranean Sea datasets (Supplementary Table 2) derived from the global survey Tara Oceans expedition (Pesant et al., 2015; Sunagawa et al., 2015) and the metagenomic profile study at different depths during seasonal stratification and mixing period along the water column (Haro-Moreno et al., 2017, 2018, 2021; López-Pérez et al., 2017).

The genome sequence of *Leeuwenhoekiella* sp. strain Mr9^T was recruited from the metagenomic datasets using BLASTn (RRID: SCR_001598) search, considering only the reads that matched the genome with a nucleotide similarity ≥99% over a minimum alignment length of ≥50 bp,

and discarding all recruitments with less than three reads per kilobase of genome per gigabase of metagenome (RPKG) (Haro-Moreno et al., 2018).

Relative abundance was calculated and plotted using the abundance scale from 0.0 to 0.01% for each of the five categories of metagenomic collections: winter; early autumn incubation samples; early autumn depth profile stratification; summer and early autumn; and Tara Oceans Expedition.

2.4. Comparative genomic analysis of secondary metabolite biosynthetic gene clusters

The potential of species of the genus *Leeuwenhoekiella* to produce novel compounds was predicted by exploring the secondary metabolite biosynthetic gene clusters (BGCs) encoded in their genomes using antiSMASH (RRID: SCR_022060) v.6.0 (Blin et al., 2021). All genome sequences derived from type strains of *Leeuwenhoekiella*, including *L. parthenopeia* strain Mr9^T, were analyzed. Cluster sequence similarities were obtained by comparison with experimentally characterized genes encoding the biosynthesis of known chemical molecules present in the Minimum Information about a Biosynthetic Gene cluster (MIBiG) database (Kautsar et al., 2019). BGC relationships among species were circularly displayed using Circos (RRID: SCR_011798) Table Viewer visualization software v.0.63-9 (Krzywinski et al., 2009). In addition to antiSMASH, the antimicrobial potential was evaluated by predicting gene clusters responsible for producing post-translationally modified peptides (RiPPs) and other bacteriocins using the BAGEL 4 web server (van Heel et al., 2018). To determine the BGC homologies with known and patented proteins, each sequence of core biosynthetic gene cluster was translated into amino acids and further searched against non-redundant proteins and patented protein sequences database (pataa) using BLASTP (RRID: SCR_001010). Assignation of proteins to families was carried out using the Pfam (RRID: SCR_004726) database (Mistry et al., 2021).

Biosynthetic gene clusters diversity and evolution across the genus *Leeuwenhoekiella* were assessed by calculating the Jaccard index (JI), adjacency index (AI), and domain sequence similarity (DSS) of each BGC by employing the software BiG-SCAPE (Biosynthetic Gene Similarity Clustering and Prospecting Engine) (RRID: SCR_022561) and CORASON (CORe Analysis of Syntenic Orthologs to prioritize Natural Product Biosynthetic Gene Clusters) (Navarro-Muñoz et al., 2020). Based on sequence similarity networks that encode the biosynthesis of highly similar or identical molecules, BGC sequences from each genome were linked to enzyme phylogenies to create gene cluster families (GCFs). This approach allowed us to obtain a global biosynthetic profile to determine the potential of all species of this genus to produce novel compounds.

2.5. Lipidomic bioassay of *Leeuwenhoekiella parthenopeia* Mr9^T on tumor cell viability

2.5.1. Total lipid extraction of bacterial cell membrane

As several antitumor compounds have been obtained from marine bacteria, we evaluate the *in vitro* activity of *L. parthenopeia* Mr9^T extracts on eukaryotic cells. Since the non-diffusible yellow pigment of strain Mr9^T is retained in the cells and low released to the medium, a poor yield of extract not displaying any activity was obtained from the free-cell supernatant. Therefore, the extraction of the total lipidic content of the cell membrane (lipidome) was performed following a modified Bligh and Dyer protocol (Bligh and Dyer, 1959). This protocol allows the extraction of total lipids constituents of cell matrix such as lipopolysaccharides, phospholipids, glycolipids, lipoproteins, carotenoid pigment content, and other medium-to-low polarity compounds included in the cells. As carotenoid pigments are liposoluble, these are extracted together as lipids. In detail, strain Mr9^T was cultivated in a mat over SWOM solid medium, and after 7 days, when the pigmentation was more intense, the cell biomass was collected and washed by resuspending it in sterile seawater or 3.5% (w/v) NaCl sterile solution and lately centrifuged.

The pellet was resuspended with 3.5% NaCl solution in equal proportion 1:1 (w/v), that is, to 1 g of pellet, 1 mL of 3.5% NaCl was added to obtain an aqueous cell suspension. A monophasic mixture of cell suspension, methanol, and chloroform was created in a ratio of 0.8:2:1 (v/v). This ratio ensures that all membrane components independent of their affinity (hydrophobic or hydrophilic) are solubilized in a monophase. After gently mixing by inversion for 1 h, the mixture was centrifuged, and the supernatant was separated from the colorless pellet. The extraction from the pellet was repeated and collected from both supernatants. The monophase of supernatants was disrupted by adding 150 µL of KCl 0.2 M. Bilayer phases were obtained after centrifugation, and the lower phase of chloroform was collected in a weighted empty glass vial and dried under a fume hood. A total of 200 µL of chloroform was added to the residual aqueous layer, mixed, and centrifuged again until residual pigments were removed from the suspension. The dried extracts were weighed and stored at −20°C. For screenings, a stock of the extract was prepared by dissolving it in DMSO at a final concentration of 100 mg/mL.

2.5.2. Cell line culture

The lipid extract from *L. parthenopeia* Mr9^T was tested against two human tumor cell lines: metastatic prostate adenocarcinoma DU-145 (ATCC HTB-81^{TM}) and glioblastoma U-87 MG (ATCC HTB-14^{TM}) of moderate

and high malignancy, respectively. As a normal cell line, immortalized human keratinocyte HaCaT (RRID: CVCL_0038) was used. All cell lines were cultured in Dulbecco's Modified Eagle Medium (DMEM) with high glucose, supplemented with inactivated 10% (v/v) of fetal bovine serum (FBS), 100 IU/mL of penicillin G, 100 mg/mL of streptomycin, and 2 mM L-glutamine, and incubated at 37°C in a 5% CO_2 humidified atmosphere (Zuchegna et al., 2020).

2.5.3. Antiproliferative assay and statistical analysis

The cell viability of tumor and normal cells after treatment with the lipid extract of *L. parthenopeia* Mr9[T] was assessed by colorimetric MTT [3-(4,5-dimethylthiazolyl)-2,5-diphenyl tetrazolium bromide] assay. In detail, all cell lines were seeded with 3×10^3 cells/well in 96-well microtiter plates and incubated for 24 h at 37°C for cell adhesion. After this time, the medium was replaced with fresh medium with increasing concentration (3.5 to 1,000 µg/mL) of lipid extract of strain Mr9[T].

For dose–response curve, 9-point serial dilutions (1:2) were performed. The vehicle DMSO without lipid extract (zero drug: 0 µg/mL) was used as a control. Plates were incubated for 24, 48, and 72 h. After each test time, the medium was removed and 100 µL of a 0.5-mg/mL MTT solution was added to each well and incubated at 37°C for 3 h. Then, the MTT solution was gently removed and 100 µL of DMSO was added to dissolve the formazan crystals. The plate was gently mixed for about 2 min under light protection, and the optical density in each well was measured using a microplate spectrophotometer (TECAN Infinite® 200 PRO) at 570 nm. Three individual replicates of the experiment were performed, each time in triplicate.

Statistical analyses were performed using GraphPad Prism Software V9.4.1 (RRID: SCR_002798) v.9.4.1. Data are expressed as a percentage normalized with controls. Values are represented as the mean ± standard deviation (SD) of three independent experiments in triplicate (N 9). Statistical significance between groups was determined by two-way analysis of variance (ANOVA) (mixed model) followed by Dunnett's multiple comparisons test.

In addition, statistical significance between the different cell lines for each biological replicate (time tested) was determined by two-way ANOVA followed by Tukey's multiple comparisons test with a single pooled variance [*p*-value in detail: 0.002 (**), 0.0002 (***), <0.0001 (****)]. A *p*-value < 0.05 is considered statistically significant and a *p*-value < 0.0001 is considered statistically highly significant. The IC$_{50}$ value of the total lipid extract (lipidome) was determined by non-linear regression/dose–response inhibition [curve fit (Inhibitor) vs. normalized response].

2.6. Dereplication: HPLC/UV/HRMS analysis

Analysis of the composition of the lipidic extract of *L. parthenopeia* Mr9[T] was performed by liquid chromatography hyphenated to UV detection and high-resolution tandem mass spectrometry (HPLC-UV-HRMS) using a previously reported analytical method (Martín et al., 2014). Nuclear magnetic resonance (^{1}H NMR) spectroscopy and heteronuclear single quantum coherence (HSQC) analyses were performed in a mixture of CD_3OD–$CDCl_3$ at 24°C on a Bruker Avance III spectrometer equipped with a 1.7-mm micro-cryoprobe.

3. Results

3.1. Strain Mr9[T] is a novel species of the phylum *Bacteroidota*

The use of sampling strategies that consider seasonal and diel factors was the premise of this study to reach the hidden microbiota. From night serial samplings carried out in late autumn 2020, a collection of isolates, including the strain Mr9[T], was obtained from reef seawater in the gulf of Naples, Italy.

The physicochemical parameters of seawater in the sampling point registered a temperature of 13.8°C, pH 8.17, TDS of 2,515 ppm, and salinity of 3.4%. The strain Mr9[T] was first isolated on a natural seawater oligotrophic medium (SWOM) producing smooth and yellow colonies after 4 days of incubation at room temperature (**Supplementary Figure 1**). According to the 16S rRNA gene sequence comparison, the strain Mr9[T] was affiliated with the genus *Leeuwenhoekiella*, most closely related to the species *L. nanhaiensis* G18[T], sharing 98.6% sequence similarity. The genomic similarities inferred by measurements of OGRIs established that strains Mr9[T] and *L. nanhaiensis* G18[T] have 88.0% of average nucleotide identity (OrthoANIu) and 33.9% of digital DNA–DNA hybridization (dDDH) (**Figure 1A**). Species delineation based on a 98.7% for the standard taxonomic marker 16D rRNA gene sequence (Chun et al., 2018), on a 95% ANI (Konstantinidis and Tiedje, 2005; Goris et al., 2007; Richter and Rosselló-Móra, 2009; Chun and Rainey, 2014), and 70% dDDH (Stackebrandt and Goebel, 1994; Auch et al., 2010), strain Mr9[T] represents a new species of the genus *Leeuwenhoekiella*, since it exhibited percentages to known species below these established thersholds.

The phylogenomic tree (**Figure 1B**), based on the alignment of 2,103 translated single-copy core genes, clearly showed that *Leeuwenhoekiella* sp. Mr9[T] forms a separate branch, clustering with *L. nanhaiensis* and it is not related to any other species affiliated with this genus. A similar result was observed in the phylogenetic tree inferred from the alignment of 16S rRNA gene sequences (**Supplementary Figure 2**). The genome-to-genome relatedness obtained from dDDH and OrthoANIu
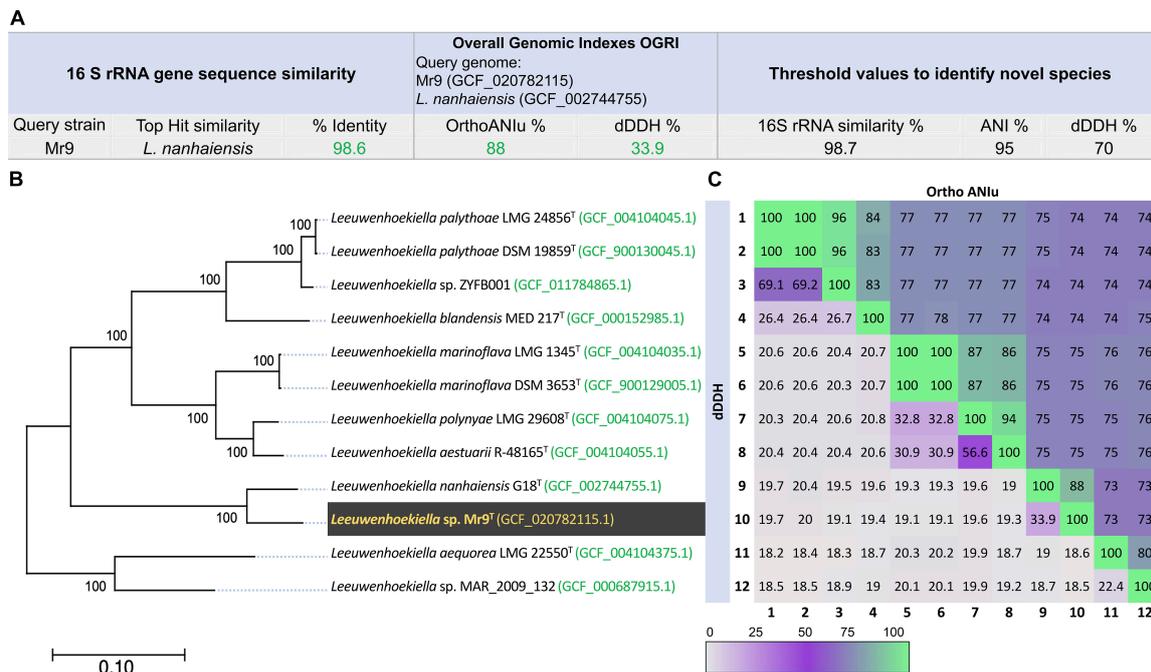
FIGURE 1
(A) 16S rRNA gene sequences similarity and overall genome relatedness index (OGRI) values among strain Mr9^T and *Leeuwenhoekiella nanhaiensis* G18^T. **(B)** Approximately maximum-likelihood phylogenomic tree based on the alignment of 2,103 translated single-copy core genes, showing the relationship among strain *Leeuwenhoekiella* sp. Mr9^T and the members of the genus *Leeuwenhoekiella*. The genome sequence accession numbers are shown in parentheses. Bootstrap percentages are shown above the branches. Bar, 0.10 substitutions per nucleotide position. **(C)** Heatmap plot of genome-to-genome analysis by OrthoANIu (upper triangle) and dDDH (lower triangle) showing the percentage of similarity among the genus under study. The colors are displayed according to the color key, with a 100% value when converging in the matrix with the same species.

analysis against the 11 genomes of species of the genus *Leeuwenhoekiella* (**Figure 1C**) showed that *Leeuwenhoekiella* sp. Mr9^T is not related to any described species of this genus, as previously confirmed by the limits for species delineation. Taxonomically, the genus *Leeuwenhoekiella* belongs to the family *Flavobacteriaceae* (Reichenbach, 1992), order *Flavobacteriales*, class *Flavobacteriia* (Bernardet, 2015), and phylum *Bacteroidota* (Oren and Garrity, 2021). Currently, this genus constituted only seven species (Parte et al., 2020), all derived from marine or sea-related habitats.

The differential phenotypic and physiological features of *Leeuwenhoekiella* sp. Mr9^T with respect to the closest related species *L. nanhaiensis* and the type species of the genus, *L. marinoflava*, are shown in **Table 1**. The polar lipid profile obtained by two-dimensional thin-layer chromatography (2D TLC) is shown in **Supplementary Figure 3**.

Based on the results obtained from the 16S rRNA gene sequences similarity, OGRIs, phylogenomic reconstruction, and physiological features, the strain Mr9^T constitutes unequivocally a new species of the genus *Leeuwenhoekiella*, for which the name *L. parthenopeia* sp. nov. is proposed. The type strain Mr9^T was deposited in three public culture collections: Spanish Type Culture Collection (CECT); Laboratory of Microbiology of

Ghent University LMG/Belgian Coordinated Collections of Microorganisms (BCCM); and German Collection of Microorganisms and Cell Cultures GmbH (DSMZ). *L. parthenopeia* Mr9^T is available in these culture collections as CECT 30318^T, DSM 112950^T, and LMG 32428^T, respectively.

3.1.1. Description of *Leeuwenhoekiella parthenopeia* sp. nov.

Leeuwenhoekiella parthenopeia (par.the.no.pei'a. L. fem. adj. *parthenopeia*, of or belonging to Naples) cells are Gram-stain-negative rods with 0.6 × 1.0 μm; colonies are circular, entire, yellow-to-orange pigmented, with 2–4 mm in diameter in medium MA incubated at 25°C for 7 days; able to grow in oligotrophic medium with 3.0–8.0% (w/v) salts (optimal at 3.5%), in the pH range of 6.5–9.0 (optimal at pH 8.0) and from 10.0 to 37.0°C (optimal at 25°C); chemoorganotrophic and aerobic; and catalase positive and oxidase negative. Nitrate is reduced, but nitrite is not. Indole and H$_2$S are not produced. Voges–Proskauer and methyl red are negative. Aesculin is hydrolyzed, but casein, DNA, gelatin, Tween 80, and starch are not hydrolyzed. Acid is produced from L-arabinose, D-galactose, D-glucose, D-maltose, D-sucrose, and D-xylose but

TABLE 1   Differential features of strains: Mr9[T], *Leeuwenhoekiella marinoflava* LMG 1345[T], and *L. nanhaiensis* KCTC 42729[T].

Feature	Strain Mr9[T]	*L. nanhaiensis* KCTC 42729[T]	*L. marinoflava* LMG 1345[T]
Cell size (width × length; μm)	0.6 × 1.0	0.4–07 × 1.4–4.1	1.0 × 1.5
Colony pigmentation	Yellow/orange	Orange	Yellow
Oxidase	−	−	+
Nitrate reduction	+	−	−
Hydrolysis of:			
Aesculin	+	+	−
Casein	−	−	+
DNA	−	+	−
Tween 80	−	+	+
Starch	−	+	−
Acid production from:			
L-Arabinose	+	+	−
D-Galactose	+	−	+
Glycerol	−	−	+
D-Glucose	+	+	−
D-Maltose	+	−	−
D-Sucrose	+	−	−
D-Xylose	+	+	−
DNA G + C (mol%, genome)	42.2	42.1	37.5

All data are from this study. +, positive reaction; −, negative reaction.

not from glycerol. Major fatty acids (>10%) are $C_{17:0}$ iso 3-OH and iso-$C_{15:0}$, followed by iso-$C_{15:1}$ G, iso-$C_{17:1}$ ω9c/$C_{16:0}$ 10-methyl, and $C_{16:1}$ ω6c/$C_{16:1}$ ω7c. The polar lipid profile consists of two aminolipids, one glycolipid, three lipids of undetermined composition, and a phospholipid with similar mobility to phosphatidylethanolamine (PE) as major lipid. The DNA G + C content is 42.2 mol% (genome).

The type strain is Mr9[T] (=CECT 30318[T] = DSM 112950[T] = LMG 32428[T]), isolated from seawater in the Gulf of Naples, Italy. The GenBank/EMBL/DDBJ accession numbers for the 16S rRNA and the complete genome of the type strain are MW785572 and JAJGMW000000000, respectively.

3.2. *Leeuwenhoekiella parthenopeia* belongs to the rare biosphere

The occurrence of *L. parthenopeia* Mr9[T] in marine habitats was mapped against 41 Mediterranean Sea metagenomes categorized into five groups by depths and seasons.

As observed in the chart (Figure 2), *L. parthenopeia* showed a relatively higher abundance as follows: 0.0026% at 40-m depth during winter; 0.0066% at 60-m depth after 14 h of incubation of the sample from early autumn, October; 0.0047% at 15-m depth in a non-incubated sample from early autumn depth profile; and 0.0090% at 3,000-m depth across summer and early autumn. Finally, in the TARA Oceans expedition dataset, the higher abundance was 0.0065% at 5-m depth during December.

The results obtained from the metagenomic mapping across depths and seasons showed that *L. parthenopeia* is present in the microbial communities of 41 metagenomes investigated with a maximum detected relative abundance of 0.0090%. The average relative abundance of *L. parthenopeia* Mr9[T] from all categories of metagenomic data from the Mediterranean Sea was as low as 0.0029%, revealing the scarce occurrence of this bacterium in the total microbial communities analyzed.

3.3. Biosynthetic potential of the genus *Leeuwenhoekiella*

3.3.1. Biosynthetic gene cluster analysis of *Leeuwenhoekiella parthenopeia* Mr9[T]

The potential of *L. parthenopeia* Mr9[T] to synthesize specialized compounds was inferred by predicting the secondary metabolite biosynthetic gene clusters (BGCs) encoded in its genome. Through antiSMASH analysis, the biosynthetic profile of *L. parthenopeia* Mr9[T] consisted of five BGCs belonging to four typologies of secondary metabolite clusters: two terpenes; one type III polyketide synthase (T3-PKS); one NRPS; and one aryl polyene (APE). As shown in Table 2, just one BGC belonging to the terpene region exhibited a low homology of 28% with a carotenoid biosynthetic gene cluster from *Myxococcus xanthus*, while the rest of the BGC types do not show similarities with any experimentally characterized gene encoding the synthesis of known chemical molecules present in the MIBiG repository (Kautsar et al., 2019). Nevertheless, according to the protein family database Pfam, the amino acid sequence of core enzyme-coding genes of each BGC of *L. parthenopeia* Mr9[T] was related with multifunctional enzymes as follows: terpene—squalene/phytoene synthase and lycopene cyclase protein; T3-PKS—chalcone and stilbene synthases; NRPS—AMP-binding enzyme/phosphopantetheine attachment site; APE—beta-ketoacyl synthase. Besides the BGCs predicted and their relatedness with protein families, the mining of gene clusters involved in the biosynthesis of bacteriocins, and RiPPs obtained with BAGEL 4 web server revealed two areas of interest in the genome with homologies higher than 97% with gene clusters related with the biosynthesis of Zoocin A (bacteriocin) and Sactipeptides (RiPP), both involved in the synthesis of natural complex products.

Furthermore, amino acid sequence homology with patented protein sequences (pataa) showed that the biosynthetic gene
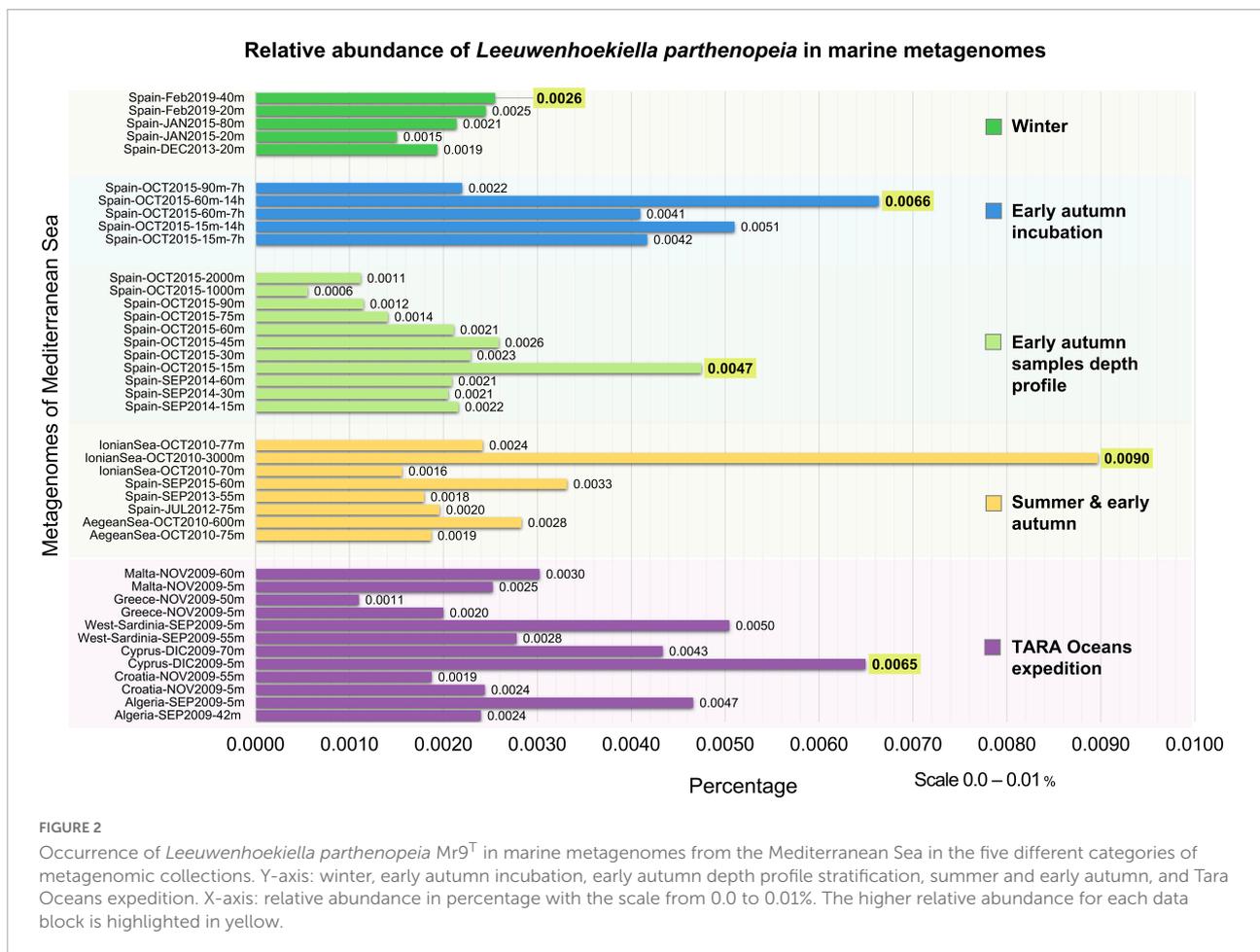
FIGURE 2

Occurrence of *Leeuwenhoekiella parthenopeia* Mr9[T] in marine metagenomes from the Mediterranean Sea in the five different categories of metagenomic collections. Y-axis: winter, early autumn incubation, early autumn depth profile stratification, summer and early autumn, and Tara Oceans expedition. X-axis: relative abundance in percentage with the scale from 0.0 to 0.01%. The higher relative abundance for each data block is highlighted in yellow.

content of *L. parthenopeia* Mr9[T] does not exhibit full similarity to any registered patent for various applications, showing a frequent match with patent US 8119385 that falls into the categories of terpene (carotenoids), T3-PKS, and bacteriocin. The identity percentages with five different patents ranged from 22.39 to 67.08%, thus revealing the high potential of *L. parthenopeia* Mr9[T] for the synthesis of novel polypeptides for biotechnological application.

3.3.2. Diversity and distribution of BGCs across *Leeuwenhoekiella*

In the same way as the profile of *L. parthenopeia* Mr9[T], the specialized metabolite potential of all species of the genus *Leeuwenhoekiella*, including *L. parthenopeia* Mr9[T], was assessed by predicting the BGC content and diversity in each of the eight species through antiSMASH. A content of three to four BGCs was identified in each species, except *L. parthenopeia* Mr9[T], which harbors five BGCs as previously detailed (**Supplementary Figure 4**). A total of 30 BGCs (29 + 1 hybrid) were identified across the genus and, like the BGC profile of *L. parthenopeia* Mr9[T], only one of the two BGCs terpenes typologies exhibited 28% homology to a carotenoid gene cluster of *M. xanthus*

(**Supplementary Table 3**). Moreover, this type of terpene is the unique gene cluster present in all species that shows homology with known BGCs; however, its similarity is very low and probably would not code for the synthesis of the same carotenoid as that of *M. xanthus*. The rest of the BGCs types showed no similarities to any experimentally characterized gene encoding the synthesis of known chemical molecules present in the MIBiG repository.

To determine how the gene cluster profile varies across *Leeuwenhoekiella*, a comparative genomic study was performed to show the correlation between species and categories of BGC. Through the chord diagram of **Figure 3**, it can be observed that three types of gene clusters are conserved in all eight species of the genus: two terpenes and one T3-PKS. It can also be evidenced by the different BGC contents between *L. parthenopeia* Mr9[T] and the most closely related species *L. nanhaiensis* G18[T], where in the latter, the gene clusters NRPS and APE are absent. The exclusive characteristic of *L. parthenopeia* Mr9[T] BGC pattern with respect to the other species of the genus is the presence of the NRPS gene cluster, which is absent in the rest of the species. Although an NRPS is also present in *L. polynyae*, it constitutes a hybrid

TABLE 2   Biosynthetic gene cluster types of *Leeuwenhoekiella parthenopeia* Mr9[T] showing its relatedness with known molecules, protein families hits, and patented proteins based on sequence similarity.

BGC's content			
Prediction tool: AntiSMASH			
Type	Molecule related similarity (MIBiG)	Protein family identity (Pfam hit)	Patented protein sequences (pataa) similarity
Terpene	Carotenoid 28%	PF00494.19 Squalene/phytoene synthase	US 8119385 48.02% (Polynucleotides and polypeptides for pharmaceutical, agricultural, cosmetic, and nutraceutical context)
	Unknow	PF05834.12 Lycopene cyclase protein	US 7630836 34.69% (Polynucleotide array)
Aryl polyene	Unknow	PF00109.26 and PF02801.22 Beta-ketoacyl synthase, N and C terminal domain	US 6673910 34.95% (Aminoacidic sequences for diagnostics and therapeutics of *Moraxella catarrhalis*)
T3-PKS	Unknow	PF00195.19 and PF02797.15 Chalcone and stilbene synthases, N and C terminal domain	US 8119385 67.08%
NRPS-like	Unknow	PF00501.28 AMP-binding enzyme PF00550.25 Phosphopantetheine attachment site	US 7319142 22.39% (Insecticidal proteins for controlling insect infestation)
Bacteriocins and RiPPs content			
Prediction tool: BAGEL4			
Class Gene name	Function	Protein homology redundant	Patented protein sequences (pataa) similarity
(RiPP) Sactipeptides BmbF	GTP 3′,8-cyclase Peptoclostridium acidaminophilum	GTP 3′,8-cyclase MoaA 97.21%	US 11078247 43.34% (Insecticidal proteins for controlling insect infestation)
(Bacteriocin) Zoocin A	Zoocin A	Peptidase M23 97.36%	US 8119385 50.66%

The best score for amino acidic homology with patented protein sequences is marked in orange.

NRPS-like + T1-PKS cluster, since both BGCs overlap in the same region of the genome, therefore counting separately within its category.

3.3.3. BGC similarity networking and evolution of the genus *Leeuwenhoekiella*

To evaluate the diversity and evolutionary history of BGC repertoires across members of the genus *Leeuwenhoekiella*, a comparative interspecies analysis of conserved-coding genes was performed by means of sequence similarity networking and BGC phylogenetic reconstruction *via* BiG-SCAPE/CORASON (Navarro-Muñoz et al., 2020).

The similarity networking groups BGCs of each species into gene cluster families (GCF) according to the presence, copy number, organization, and sequence homology of the protein domains on antiSMASH-predicted regions. As shown in the similarity matrix calculation obtained through BiG-SCAPE

(Figure 4A), the clustering of the total 30 BGC content of *Leeuwenhoekiella* gave rise to nine GCFs affiliated with four types of BGCs, the terpene category being the most represented. It is observed in the matrix that all eight species share the same core enzyme-coding genes that form two families: FAM_00027 and FAM_00026. These GCFs are affiliated with the BGC types terpene and T3-PKS, respectively. Conversely, three singletons were identified: FAM_00016 belonging to terpenes corresponding to *L. marinoflava*; FAM_00003 belonging to NRPS which was exclusive of *L. parthenopeia*, and FAM_00028 presents in *L. polynyae;* the latter corresponds to the chemical hybrid NRPS-like-T1-PKS, which explains the two times computing and its assignation to two categories, NRPS and PKS. These singletons result when the domains present in the BGC do not reach the thresholds that capture sequence similarity, synteny, or presence–absence of the Jaccard index against BGCs from other species to form a network, thus, a family of shared

FIGURE 3
Circle plot showing the relationships among the eight species of *Leeuwenhoekiella* (left half of the circle) and the biosynthetic gene cluster (BGC)'s diversity by typologies (right half of the circle). Each genome is represented by different colors, which are segmented according to the number of BGC content in their genomes. The chords connect both sides indicating, according to their width, the number of BGCs belonging to each typology. The outer rings in the gray scale represent the percentage of presence of BGC and in the color scale the proportion of genomes per BGC.

gene groups could not be defined, remaining as an individual member. In all cases, whether as clusters or singletons, after the GCF's comparison with MIBiG reference BGCs, none of the nine GCFs match any pathways responsible to produce known compounds.

Evolutionary relationships between BGCs and within GCF were initially inferred by CORASON inside the GCF defined for BiG-SCAPE for each family. CORASON establishes phylogenies based on BGC's conserved core of common domains. The clades resulting from the phylogenetic reconstruction may be responsible for the biosynthesis of different types of chemistry due to the association of specific types of genes encoding additional enzymes. In correspondence with the families in the presence–absence matrix, the phylogenetic reconstructions

of all BGCs from *Leeuwenhoekiella* (**Figure 4B**) generated six multi-locus approximately maximum-likelihood trees, and the three singletons (FAM_00016, FAM_00003, and FAM_00028) do not form clades. The featured conserved core domains and their related enzymes used to define the GCFs are specified for each tree. In addition, the genes that encode functionally uncharacterized proteins (domains of unknown function, DUF) were identified to see their cryptic potential of novel structures and functions.

To further investigate possible distant phylogenetic relationships among these families, CORASON was run outside the BiG-SCAPE-predicted BGC families. Instead of restricting the analysis to anti-SMASH-predicted regions, full *Leeuwenhoekiella* genomes were used as CORASON
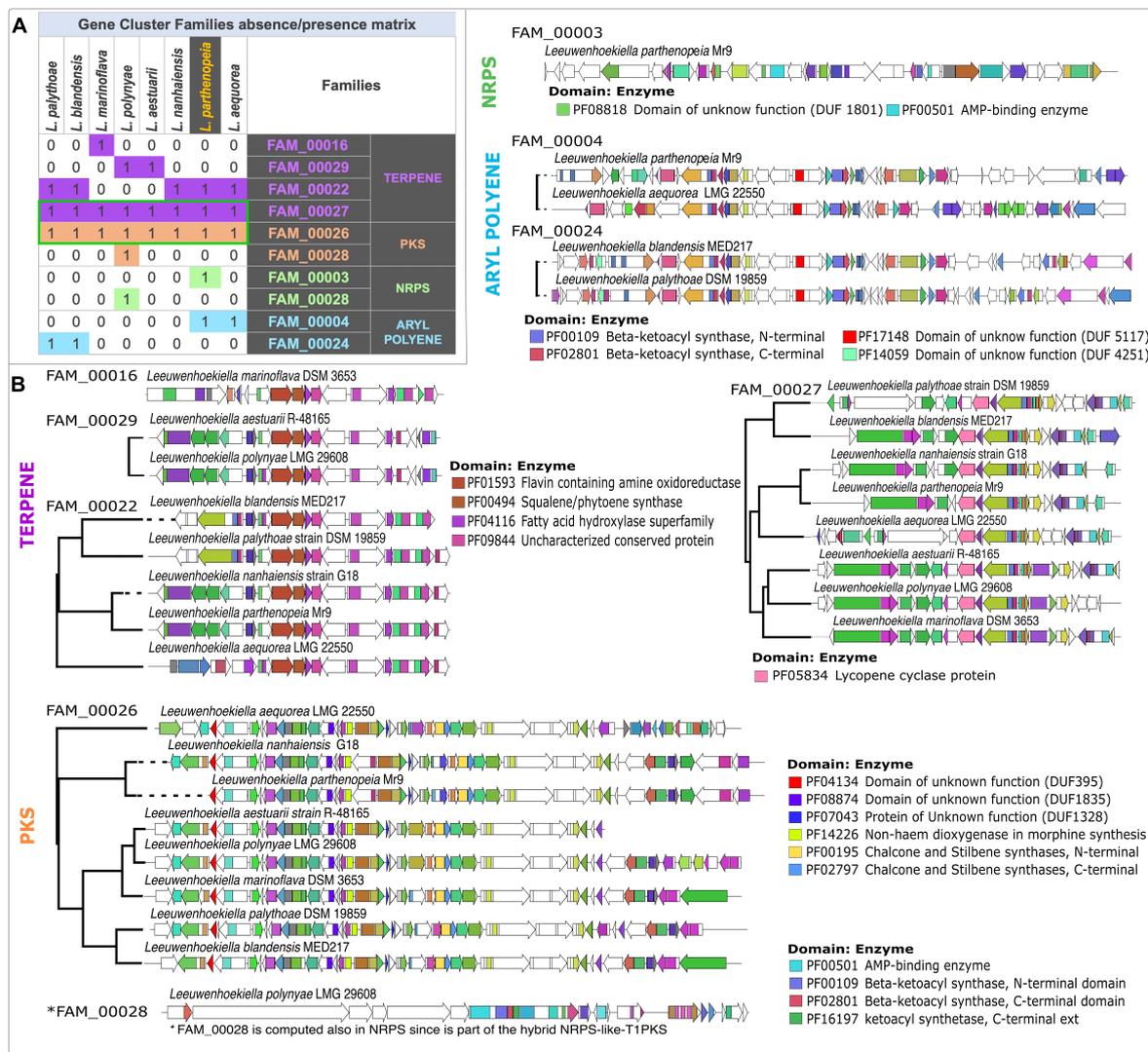
FIGURE 4

(A) Similarity network matrix based on the absence and presence of gene cluster families from each species of the genus *Leeuwenhoekiella*. Green boxes, gene cluster families (GCFs) are present in all members of *Leeuwenhoekiella*. **(B)** Multi-locus approximately maximum-likelihood trees based on the conserved core common domain.

inputs. As a result, families FAM_00016 (*L. parthenopeia*), FAM_00022 (*L. nanhaiensis, L. marinoflava, L. blandensis, L. palythoae, and L. aequorea*), and FAM_00029 (*L. polynyae and L. aestuarii*) were clustered in a single phylogenetic tree (**Supplementary Figure 5**).

With this observation, it is shown that a new superfamily that comprises a terpene BGC in each of the species is studied here. It is possible that due to the variation in gene content and sequence similarity, BiG-SCAPE split this superfamily into the three families shown in the presence–absence matrix, but these families shared a common evolutionary history as can be seen in **Figure 4B**. The same is true for aryl polyene families FAM_00004 and FAM_00024, and they share a common ancestor. In fact, though not all the genomes have

an antiSMASH region predicted, the beta-ketoacyl synthase genomic vicinity is conserved in all eight genomes as can be seen in the **Supplementary Figure 5** produced by CORASON.

According to the featured conserved core domains linked to enzymes, within the terpene category, the GCF: FAM_00016 (singleton), FAM_00029, and FAM_00022 encode the biosynthesis of highly similar or identical molecules to flavin, squalene, chalcones–stilbenes, and an unknown compound associated with an uncharacterized conserved protein. This category highlights the FAM_00027 which is composed of all species of *Leeuwenhoekiella* revealing the potential of members of this genus to synthesize carotenoid pigments as lycopene.

In the category PKS, the FAM_00026 is also clustered with all species of *Leeuwenhoekiella* and encodes the biosynthesis of

carotenoids, chalcone, and stilbenes. The singleton FAM_00028 (*L. polynyae*) also falls in this group, which is duplicated in the category NRPS, since *L. polynyae* contains the hybrid gene cluster NRPS-like-T1_PKS, computing in both groups. The featured conserved domains of this singleton are linked to the family protein AMP-binding enzyme involved in the synthesis of several bioactive compounds and beta-ketoacyl synthase involved in the synthesis of fatty acids. In both cases, given the high capacity to synthesize a wide spectrum of molecules, it is not possible to precise the most probable-related molecule. Regarding the NRPS category, the singleton FAM_00003 (*L. parthenopeia* Mr9^T) possesses the domain of the unknown function (DU1801) and a major facilitator superfamily domain; in both cases, the possible synthesis of a determined molecule is unknown.

Finally, in the category of aryl polyene (APE), two GCFs, FAM_00004 and FAM_00024, possess common conserved domains of unknown functions (DUF 5117 and DUF 4251) for which the eventually synthesized molecules remain unknown, and the beta-ketoacyl synthase is involved in the synthesis of fatty acids. The GCF analysis was performed, allowing us to detect the frequent presence of domains of unknown function (DUFs) in all GCFs by categories, strongly suggesting that this genus harbors the potential to synthesize novel compounds not described nor chemically tested yet.

3.4. Total lipid extract from *Leeuwenhoekiella parthenopeia* affects viability of tumor cells

The potential antiproliferative effect of total lipid extract (TLE) from *L. parthenopeia* Mr9 was evaluated using the colorimetric MTT assay, which measures cellular metabolic activity as an indicator of cell viability. Two human tumor cell lines: metastatic prostate adenocarcinoma DU-145 and glioblastoma U-87 MG of moderate and high malignancy, respectively, were tested. As a normal cell line, it was used to immortalize human keratinocyte HaCaT.

After the treatment of the cell lines with several doses of TLE from 3.5 to 1,000 µg/mL for 24, 48, and 72 h, it was possible to observe that the viability of tumor cells DU-145 and U-87 MG decreases gradually in time- and dose-dependent manner with lower effect on HaCaT (**Figure 5A**). In particular, the cell death of DU-145 and U-87 MG above 50% occurred with 120 µg/mL of extract after 24 h, while the cell death of HaCaT happens only with the highest concentration (1,000 µg/mL). After 48 h, DU-145 and U-87 MG are inhibited above 50% with 60 µg/mL, while HaCaT requires the highest concentration of the extract. In all cell lines, both tumor and normal, the greatest inhibitory effect was observed after 72 h of treatment, where the cell viability of DU-145 and U-87 MG decreased significantly, requiring a concentration of around 15 µg/mL to inhibit the

50% of the cells, while HaCaT requires more than 250 µg/mL. Between the two tumor cell lines, DU-145 and U-87 MG, it is observed that TLE is less effective on U-87 MG at doses lower than 60 µg/mL at 24 and 48 h of treatment, and successively at higher concentration, the difference of the inhibitory effect on both tumor cell line is reduced.

According to the cell viability results and the inhibitory concentration value IC$_{50}$ (**Figure 5B**), the lipid extract of *L. parthenopeia* Mr9^T possesses inhibitory activity against metastatic prostate adenocarcinoma DU-145 and glioblastoma U-87 MG with greater effectiveness on DU-145. The minor effect on immortalized human keratinocyte HaCaT suggests a selective effect on tumor cells.

3.5. Dereplication: Total lipid composition

The chemical composition of the lipid extract from *L. parthenopeia* Mr9^T was analyzed by HPLC-UV-HRMS, and the molecular formulae of major components were determined. **Figure 6** shows the LC/UV trace obtained in the analysis. Searches of molecular formulae of the peaks in the figure against the Dictionary of Natural Products Database (DNP) tentatively identified some compounds and other molecules present in the extract with no matches in the database.

Following the above spectra, 20 peaks were identified, and their molecule formula and putative component are detailed in **Table 3**. From these peaks, nine molecular formulae do not match in the DNP, five corresponds to fermentation medium component (FMC), and six corresponds to known molecules as follows: sulfobacin-like component, sulfobacin A, WB 3559A, WB 3559B, docosenamide, and topostin B-567.

Searches of molecular formula of the peaks in the figure against the Dictionary of Natural Products Database (DNP) tentatively identified some compounds and other molecules present in the extract with no matches in the database. **Figure 7** represents the structures of the compounds tentatively identified in the lipidic extract. Not surprisingly, these compounds present in their structure units of C$_{17:0}$ iso 3-OH, C$_{15:0}$ iso, and C$_{15:1}$ iso fatty acids are identified as major fatty acids in Section "Description of *Leeuwenhoekiella parthenopeia* sp. nov." NMR analysis corroborated the presence of iso-fatty acids as the major components in the lipid extract.

4. Discussion

4.1. Isolation and ecology

The isolation of *L. parthenopeia* strain Mr9^T from a night seawater sample in autumn supports previous studies on the dynamic and seasonal effect on the microbial population. This

FIGURE 5

(A) Effect of total lipid extract from *Leeuwenhoekiella parthenopeia* Mr9[T] cell extract in the cell viability of DU-145, U-87 MG, and HaCaT after 24, 48, and 72 h of stimulation with gradual doses. Untreated cells (0 μg/mL) were used as control and indicate cells treated with dimethyl sulfoxide (DMSO) without *L. parthenopeia* Mr9[T] extract. Control corresponds to 100% of viability. The histograms represent the mean ± SD of three independent experiments performed in triplicate (*n* = 9), normalized to untreated control cells. Statistical analysis: *p*-value: 0.002 (**), 0.0002 (***), and <0.0001 (****) compared to the control. **(B)** IC_{50} inhibitory concentrations of the total lipid extract (TLE) required to inhibit DU-145, U-87 MG, and HaCaT after 24, 48, and 72 h of stimulation.

isolation could be associated with the natural phenomenon of the DVM and the mixing of the water itself during the cold months since these favor the ascent toward the surface of planktonic particles and solid matter, acting as vectors of the hidden microbiota. The particulate was ascertained by the visual observation and the measurement of TDS that showed high detection in night samples. Genomic and phenotypic features of this strain showed that it constitutes a new species of the genus *Leeuwenhoekiella*, for which we propose the new name *L. parthenopeia* sp. nov.

It is probable that *L. parthenopeia*, as a member of the *Flavobacteriia*, is trapped in small particles such as marine picoplankton, where this class is dominant and typically found at depth (Kirchman et al., 2003; Alonso et al., 2007; Gómez-Pereira et al., 2010; Mestre et al., 2017). In cold waters, a high abundance of some members of *Flavobacteriia* has also been reported (Ghiglione and Murray, 2012; Grzymski et al., 2012), and, in fact, two species of *Leeuwenhoekiella*, *L. aequorea* (Nedashkovskaya et al., 2005) and *L. polynyae* (Si et al., 2015),

were isolated from Antarctic Sea water, which also explains that cold seasons would favor the presence of microbial population minorities as a rare biosphere. The lack of isolates of microorganisms belonging to this genus or any member of the phylum *Bacteroidota* in day samples, besides the reduction of particles, demonstrates that the cultivation of *L. parthenopeia* as well as other members within this phylum is not causality and indicates a close association with planktonic particulate, organic matter, algae or neustons colonizing their surface acting as ectosymbiont (Kirchman, 2002; Fernández-Gómez et al., 2013; Mestre et al., 2017; Baumas et al., 2021). Thus, the probability of isolation of similar bacteria may depend on the composition of phytoplankton and temporal dynamics (Rooney-Varga et al., 2005; Teeling et al., 2016; Camarena-Gómez et al., 2021).

Physiological characteristics showed that *L. parthenopeia* has a heterotrophic and versatile metabolism, assimilating various substrates as a carbon and energy source, and suggesting that this bacterium can thrive in a wide range of nutritional circumstances in its habitat. This versatility highly contributed
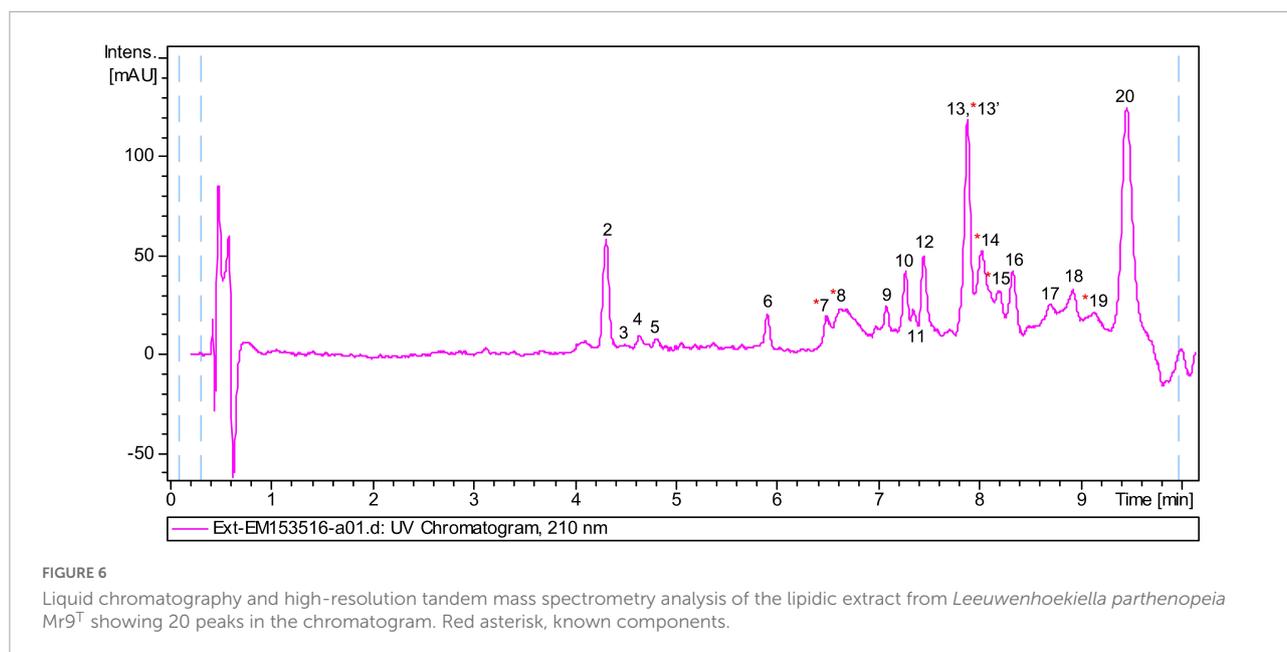
FIGURE 6
Liquid chromatography and high-resolution tandem mass spectrometry analysis of the lipidic extract from *Leeuwenhoekiella parthenopeia* Mr9[T] showing 20 peaks in the chromatogram. Red asterisk, known components.

TABLE 3 Peaks with molecule formula and putative components.

Peak	Molecular formula	Putative dereplicated component
1, 2	$C_{22}H_{26}O_6$	Fermentation medium component
3	$C_{15}H_{31}NO_2$	No coincidences in the DNP
4	$C_{21}H_{33}N_3O_3$	Fermentation medium component
5	$C_{14}H_{18}O_5$	Fermentation medium component
6	$C_{21}H_{37}N$	Fermentation medium component
7	$C_{33}H_{67}NO_6S$	Sulfobacin like component
8	$C_{34}H_{69}NO_6S$	Sulfobacin A
9	$C_{33}H_{50}O_8P_2$	Fermentation medium component
10	$C_{35}H_{68}O_6S$	No coincidences in the DNP
11	$C_{36}H_{70}N_2O_6$	No coincidences in the DNP
12	$C_{37}H_{72}N_2O_6$	No coincidences in the DNP
13	$C_{33}H_{50}O_7P_2$	No coincidences in the DNP
13′	$C_{36}H_{66}N_2O_7$	WB 3559A
14	$C_{37}H_{68}N_2O_7$	WB 3559B
15	$C_{22}H_{43}NO$	Docosenamide
16	$C_{33}H_{60}O_3S_2$	No coincidences in the DNP
17	$C_{35}H_{62}O_3$	No coincidences in the DNP
18	$C_{53}H_{58}O_7P_2$	No coincidences in the DNP
19	$C_{34}H_{65}NO_5$	Topostin B-567
20	$C_{33}H_{52}O_6P_2$	No coincidences in the DNP

to its fast laboratory domestication, exhibiting the ability to grow in synthetic marine media without requiring natural seawater in culture media as the main component of primary cultures. Its role in nature might fulfill its function as a heterotroph, that is, degrading organic matter (Alonso et al., 2007; Gómez-Pereira et al., 2010; Xue et al., 2020).

The almost undetectable genomic recruitment patterns of the new species *L. parthenopeia* when mapping the 41 metagenomes from the Mediterranean Sea revealed that this species is poorly represented with a mean relative abundance as low as 0.0029%, considering that members of the class *Flavobacteriia* (phylum *Bacteroidota*) are prominent in marine environments across different seasons and depth, accounting from 4.7 to 13.9% (Cram et al., 2015) and from 3 to 20% in planktonic particles (Mestre et al., 2017; Baumas et al., 2021). This infimum percentage reveals the scarce occurrence of this novel bacterium in the total microbial communities analyzed. Based on the arbitrary threshold of 0.1 or 0.01% of relative abundance for defining rare biosphere (Pedrós-Alió, 2012), *L. parthenopeia*, similar to *L. blandensis* (Pedrós-Alió, 2013), is part of the low-abundant fraction of microbial population and therefore belongs to the rare biosphere.

According to all the categories of metagenomes analyzed, we could not observe a pattern of distribution that allows us to establish the depth or the specific season at which *L. parthenopeia* is relatively more abundant. However, we can affirm that despite its scarcity, *L. parthenopeia* is widely distributed in the Mediterranean Sea. A remarkable fact is the highest abundance of *L. parthenopeia* found in the Ionian Sea metagenome, at a depth of 3,000 m, which might suggest that *L. parthenopeia* could be more abundant at abyssal depths. In fact, its closest relative, the species *L. nanhaiensis*, was isolated at a depth of 2,000 m (Liu et al., 2016), while a culture-independent study about the depth and seasonal distribution of *Bacteroidota*, showed that the members of the genus *Leeuwenhoekiella*
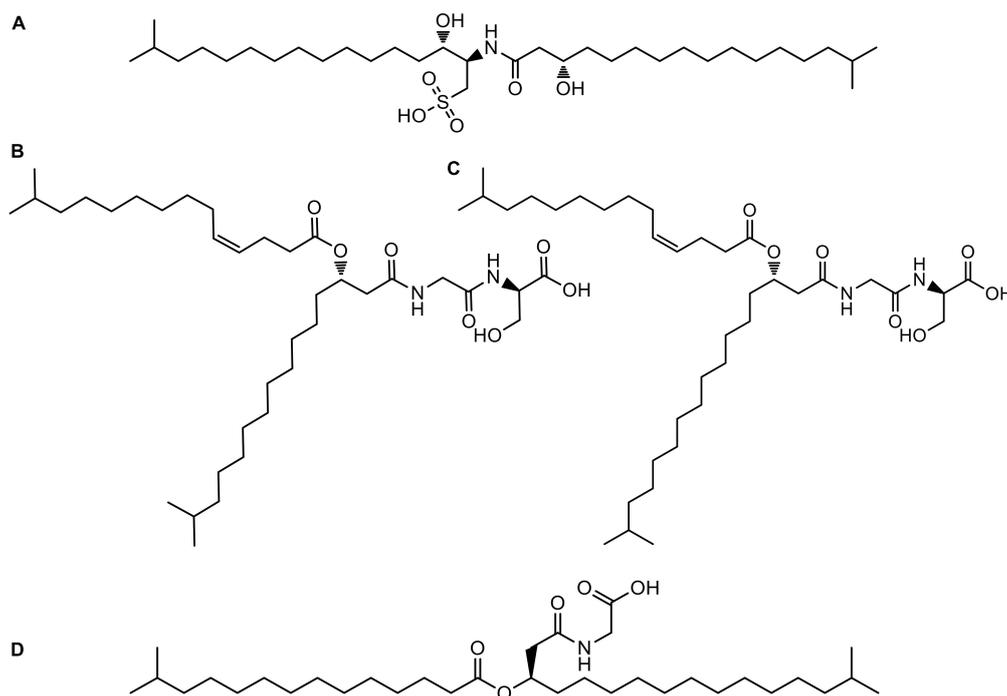
FIGURE 7
Structures of compounds tentatively identified in the liquid chromatography hyphenated to UV detection and high-resolution tandem mass spectrometry (HPLC-UV-HRMS) analysis of the lipidic extract of *Leeuwenhoekiella parthenopeia* Mr9[T]: **(A)** sulfobacin A; **(B)** WB 3559A; **(C)** WB 3559B; **(D)** topostin B-567.

were detected in deep marine waters (Díez-Vives et al., 2019), therefore, supporting that vertical migration besides the mixing waters are determinant factors to reach the rare microbiota. Our study shows that the sample pre-incubation increases the abundance of *L. parthenopeia*, as observed in the metagenomes from incubated samples collected in October, where the presence of *L. parthenopeia* is doubled, especially after 14 h with respect to the other metagenomes studied. This result indicates that the pre-incubation of the sample before the culturing increases the possibility of isolation of poorly represented species. However, this procedure would contribute equally to the overgrowth of representatives of the most abundant phyla of the marine microbial community, such as *Pseudomonadota* (formerly *Proteobacteria*) and *Cyanobacteria* (Fernández-Gómez et al., 2013).

4.2. Biosynthetic potential

The biosynthetic gene profiling of *L. parthenopeia* Mr9[T] and all species of the genus *Leeuwenhoekiella* showed that members of this genus have in common three BGCs: two terpenes and one T3-PKS. Of these, only one BGC terpene showed a low homology of 28% with a carotenoid molecule, while the rest of the BGCs do not exhibit similarity with any experimentally characterized gene encoding the synthesis

of known chemical metabolites. Even though most of the gene clusters did not show a match with known molecules, the nucleotide translated sequences of those BGCs revealed that this genus, including *L. parthenopeia* Mr9[T], possesses multifunctional enzymes related to the synthesis of squalene, lycopene, fatty acids, flavonoids, and resveratrol; thus, they are probably synthesizing these types of molecules. In contrast to the rest of the species, only *L. parthenopeia* Mr9[T] possesses NRPS and the significance of this type of BGC relies upon its well-known synthesis of antibiotic and antitumor compounds, thus the importance of *L. parthenopeia* Mr9[T] with respect to the other species.

Additionally, the low homology of the core peptide with patented proteins confirms that *L. parthenopeia* Mr9[T] does not synthesize the same proteins that those patented, suggesting its high metabolic capability to produce novel compounds with entirely different biotechnological applications.

Overall, the biosynthetic profile of species of *Leeuwenhoekiella* is very streamlined in comparison with the profile of the species of *Salinispora*, the abyssal marine bacteria producer of antitumor compounds, that harbor 176 distinct BGCs, of which only 24 have been linked to their products (Letzel et al., 2017). Despite this simplicity and given the novelty of these BGCs, they are probably synthesizing molecules with potentially different functionalities.

The similarity networking and evolutive clustering of the BGCs linked to enzyme phylogenies revealed that all species of *Leeuwenhoekiella* form four prominent gene cluster families (GCF) groups and according to this classification, species of this genus synthesize molecules such as flavine, lycopene, and squalene/phytoene. Moreover, the presence of uncharacterized conserved proteins and beta-ketoacyl synthase also suggests the synthesis of molecules of unknown function by species of this genus. Beyond these GCFs, also in evolutionary terms, *L. parthenopeia* Mr9[T] remains again potentially more interesting than its peers, since the presence of NRPS, which feature domains AMP-binding enzyme, not shared with the other species, confirms that this bacterium may synthesize other compounds of biotechnological interest.

4.3. Biological activity: Lipidomic bioassay

Regarding the antiproliferative effect that exhibits the total lipid extract (TLE) from *L. parthenopeia* Mr9[T] against prostate adenocarcinoma cell line, DU-145, and glioblastoma cell line, U-87 MG, it is important to take into account the nature of these cell lines, considered as moderate and high malignancy, respectively. In fact, the total lipid extract of *L. parthenopeia* proved to be more effective on prostate cells DU-145 than on glioblastoma cells U-87 MG. This result could be expected due to the intrinsic characteristic of the U-87 MG cell line, whose heterogeneous nature is associated with clonal plasticity that makes glioblastoma extremely resistant to current treatments (Osuka and Van Meir, 2017; Oliver et al., 2020). The antecedent is that glioblastoma is one of the most aggressive types of brain cancer, characterized by its rapid multiplication and invasion (Oraiopoulou et al., 2017), which could explain the low effect of the lipid at doses lower than 60 μg/mL at 24 and 48 h, as shown in the line graph of MTT results (**Figure 5A**). The primary cell line U-87 MG (ATCC HTB-14) has a double time of generation ranging from 24 to 29 h (Kato et al., 2011) and this faster multiplication was also observed under laboratory conditions, for which higher doses of extract and a long exposure time are required. The lipid extract was more effective on DU-145 causing a homogeneous decrease of cell viability from the first 24 h of treatment. In both tumor cell lines, DU-145 and U-87 MG, the cell viability gradually decreased in a time- and dose-dependent manner reaching a maximum efficiency at 72 h. The lower effect of the extract in immortalized human keratinocyte HaCaT, used as no tumor cell line, suggests a selective effect on tumor cells. Individual screenings with purified fractions will allow us to determine if the cell death is caused by a single compound or if it is a synergistic effect of individual compounds, then it will be necessary to identify the molecular mechanism that induces death. In any case, the lipid extract of the cell membrane of *L. parthenopeia* Mr9[T] possesses an effective antiproliferative activity on tumor cell lines DU-145 and U-87 MG, not previously reported in rare bacteria.

4.4. Lipidome analysis

The chemical composition of the lipid extract from *L. parthenopeia* Mr9[T] consisted of four putative compounds with known chemical structures, besides nine molecules without coincidences in the dictionary of natural products (DNP) identified after the dereplication. Among the known molecules are: sulfobacin A (Kamiyama et al., 1995), WB 3559A, WB-3559B (Yoshida et al., 1985), and topostin B-567 (Suzuki et al., 1990), all are derived from species of the class *Flavobacteriia*, except topostin B-567 isolated from a culture broth of *Flexibacter topostinus* belonging to the class *Cytophagia*. Regarding their biological activities, sulfobacin A is a sulfonolipid that was described as an antagonist of von Willebrand factor receptor (vWF), a blood glycoprotein involved in hemostasis and thus proposed as an antithrombosis agent (Kamiyama et al., 1995). Furthermore, sulfonolipid analogous (flavocristamide A and B) were found to have inhibitory activity against DNA polymerase α in eukaryotic cells (Kohayashi et al., 1995). Subsequently, the cytotoxic properties of this molecule were reported against four cancer cell lines with maximum activity against human mammary adenocarcinoma (Chaudhari et al., 2009). Concerning the compounds WB-3559 A and B, they were reported as potent fibrinolytic agents produced by *Flavobacterium* sp. No. 3559, which stimulates the euglobulin clot lysis time of rabbit plasma (Uchida et al., 1985; Yoshida et al., 1985). Docosenamide, also known as erucamide is a fatty acid amide belonging to a family of brain lipids that induce sleep (Cravatt et al., 1995). This molecule acts as a human metabolite, and it is more used in the neuroscience area. Finally, Topostin B-567 is an inhibitor of mammalian topoisomerase I and was isolated from a broth culture of *F. topostinus* B-572 (Ikegami et al., 1990; Suzuki et al., 1990; Brinkmann et al., 2022). Topoisomerase I is the molecular target for anticancer drugs since the activity of this enzyme is increased in tumor cells (Buzun et al., 2020). Topoisomerase suppressors constitute a family of antitumor agents with cytostatic effect, which mechanism of action is the inhibition of enzyme activity leading to the interruption of DNA strands and causing cell death (Bailly, 2000; Galluzzi et al., 2015). Considering the biological activities of the known compounds, we can hypothesize that the antiproliferative effect of DU-145 and U-87 MG observed *in vitro* in this study could be attributed to sulfobacin A and topostin B-567. The reduction of cell viability could be due to the action of both compounds since both inhibit DNA replication. However, the only one that does so selectively with tumor cells is topostin B-567, which being a specific suppressor of the enzyme topoisomerase I would explain the death of tumor cells and the low effect on epithelial tissue cells. Beyond the interesting

known molecules found in the lipid extract, the great potential of *L. parthenopeia* Mr9^T remains in the nine molecules without coincidences in the dictionary of natural products and still to be chemically elucidated.

Identifying the active fractions and their chemical structure will allow us to decipher the possible genes involved in the antiproliferative activity shown *in vitro*. Thus, based only on the *in silico* study, we cannot accurately demonstrate which of the BGCs present in the genome of *L. parthenopeia* Mr9^T can be responsible for this activity. Nonetheless, the simplicity of the biosynthetic profile of *L. parthenopeia* would make feasible the discrimination of the genes involved in the antitumor activity circumscribed to the four categories of BGC of its profile.

In conclusion, our study demonstrates that the new bacterium *L. parthenopeia* and the other species of the genus *Leeuwenhoekiella* contain biosynthetic genes with potential antiproliferative activities and encourage to carry out similar approaches to explore the marine rare biosphere as a promising source of bioactive compounds.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/**Supplementary material**.

Author contributions

GG, RH, JM, FR, and PC: formal analysis. GG and PC: investigation. GG, RH, JM, FR, CS-P, AF, CZ, SG-F, NS-M, and PC: methodology. GG, RH, JM, SG-F, NS-M, and PC: software and visualization. GG: data curation and writing original draft preparation. RH, FR, CS-P, MV, ER, NS-M, AV, and PC: validation. RH, AV, and PC: writing – review and editing. CS-P, AV, and PC: funding acquisition, project administration, and resources. PC: conceptualization and supervision. All authors contributed to the manuscript revision, read, and approved the submitted version.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2022.1090197/full#supplementary-material

References

Agrawal, S., Acharya, D., Adholeya, A., Barrow, C. J., and Deshmukh, S. K. (2017). Nonribosomal peptides from marine microbes and their antimicrobial and anticancer potential. *Front. Pharmacol.* 8:828. doi: 10.3389/fphar.2017.00828

Alonso, C., Warnecke, F., Amann, R., and Pernthaler, J. (2007). High local and global diversity of flavobacteria in marine plankton. *Environ. Microbiol.* 9, 1253–1266. doi: 10.1111/j.1462-2920.2007.01244.x

Angelini, R., Corral, P., Lopalco, P., Ventosa, A., and Corcelli, A. (2012). Novel ether lipid cardiolipins in archaeal membranes of extreme haloalkaliphiles. *Biochim. Biophys. Acta* 1818, 1365–1373. doi: 10.1016/j.bbamem.2012.02.014

Auch, A. F., von Jan, M., Klenk, H.-P., and Göker, M. (2010). Digital DNA-DNA hybridization for microbial species delineation by means of genome-to-genome sequence comparison. *Stand. Genomic Sci.* 2, 117–134. doi: 10.4056/sigs.53 1120

Bailly, C. (2000). Topoisomerase I poisons and suppressors as anticancer drugs. *CMC* 7, 39–58. doi: 10.2174/0929867003375489

Baker, B. J., Appler, K. E., and Gong, X. (2021). New microbial biodiversity in marine sediments. *Annu. Rev. Mar. Sci.* 13, 161–175. doi: 10.1146/annurev-marine-032020-014552

Bankevich, A., Nurk, S., Antipov, D., Gurevich, A. A., Dvorkin, M., Kulikov, A. S., et al. (2012). SPAdes: A new genome assembly algorithm and its applications to single-cell sequencing. *J. Computat. Biol.* 19, 455–477. doi: 10.1089/cmb.2012.0021

Barrow, G. I., and Feltham, R. K. A. (1993). "Characterization tests," in *Cowan and steel's manual for the identification of medical bacteria*, eds G. I. Barrow and R. K. A. Feltham (Cambridge: Cambridge University Press), 219–238. doi: 10.1017/CBO9780511527104.020

Baumas, C. M. J., Le Moigne, F. A. C., Garel, M., Bhairy, N., Guasco, S., Riou, V., et al. (2021). Mesopelagic microbial carbon production correlates with diversity across different marine particle fractions. *ISME J.* 15, 1695–1708. doi: 10.1038/s41396-020-00880-z

Belknap, K. C., Park, C. J., Barth, B. M., and Andam, C. P. (2020). Genome mining of biosynthetic and chemotherapeutic gene clusters in *Streptomyces* bacteria. *Sci. Rep.* 10:2003. doi: 10.1038/s41598-020-58904-9

Bernardet, J. (2015). "Flavobacteriales ord. nov," in *Bergey's manual of systematics of archaea and bacteria*, eds W. B. Whitman, F. Rainey, P. Kämpfer, M. Trujillo, J. Chun, P. DeVos, et al. (New York, NY: Wiley). doi: 10.1002/9781118960608.obm00033

Bligh, E. G., and Dyer, W. J. (1959). A rapid method of total lipid extraction and purification. *Can. J. Biochem. Physiol.* 37, 911–917. doi: 10.1139/o59-099

Blin, K., Shaw, S., Kloosterman, A. M., Charlop-Powers, Z., van Wezel, G. P., Medema, M. H., et al. (2021). antiSMASH 6.0: Improving cluster detection and comparison capabilities. *Nucleic Acids Res.* 49, W29–W35. doi: 10.1093/nar/gkab335

Boccellato, C., Kolbe, E., Peters, N., Juric, V., Fullstone, G., Verreault, M., et al. (2021). Marizomib sensitizes primary glioma cells to apoptosis induced by a latest-generation TRAIL receptor agonist. *Cell Death Dis.* 12:647. doi: 10.1038/s41419-021-03927-x

Bolger, A. M., Lohse, M., and Usadel, B. (2014). Trimmomatic: A flexible trimmer for illumina sequence data. *Bioinformatics* 30, 2114–2120. doi: 10.1093/bioinformatics/btu170

Bosch, H. F., and Rowland Taylor, R. (1973). Diurnal vertical migration of an estuarine cladoceran, *Podon polyphemoides*, in the Chesapeake Bay. *Mar. Biol.* 19, 172–181. doi: 10.1007/BF00353589

Brinkmann, S., Spohn, M. S., and Schäberle, T. F. (2022). Bioactive natural products from bacteroidetes. *Nat. Prod. Rep.* 39, 1045–1065. doi: 10.1039/D1NP00072A

Buzun, K, Bielawska, A., Bielawski, K., and Gornowicz, A. (2020). DNA topoisomerases as molecular targets for anticancer drugs. *J. Enzyme Inhib. Med. Chem.* 35, 1781–1799. doi: 10.1080/14756366.2020.1821676

Camarena-Gómez, M. T., Ruiz-González, C., Piiparinen, J., Lipsewers, T., Sobrino, C., Logares, R., et al. (2021). Bacterioplankton dynamics driven by interannual and spatial variation in diatom and dinoflagellate spring bloom communities in the Baltic Sea. *Limnol. Oceanogr.* 66, 255–271. doi: 10.1002/lno.11601

Chaudhari, P. N., Wani, K. S., Chaudhari, B. L., and Chincholkar, S. B. (2009). Characteristics of sulfobacin A from a soil isolate *Chryseobacterium gleum*. *Appl. Biochem. Biotechnol.* 158, 231–241. doi: 10.1007/s12010-008-8417-7

Chun, J., and Rainey, F. A. (2014). Integrating genomics into the taxonomy and systematics of the bacteria and archaea. *Int. J. Syst. Evol. Microbiol.* 64, 316–324. doi: 10.1099/ijs.0.054171-0

Chun, J., Oren, A., Ventosa, A., Christensen, H., Arahal, D. R., da Costa, M. S., et al. (2018). Proposed minimal standards for the use of genome data for the taxonomy of prokaryotes. *Int. J. Syst. Evol. Microbiol.* 68, 461–466. doi: 10.1099/ijsem.0.002516

Chwastek, G., Surma, M. A., Rizk, S., Grosser, D., Lavrynenko, O., Rucińska, M., et al. (2020). Principles of membrane adaptation revealed through environmentally induced bacterial lipidome remodeling. *Cell Rep.* 32:108165. doi: 10.1016/j.celrep.2020.108165

Corral, P., de la Haba, R. R., Infante-Domínguez, C., Sánchez-Porro, C., Amoozegar, M. A., Papke, R. T., et al. (2018). *Halorubrum chaoviator* Mancinelli et al. 2009 is a later, heterotypic synonym of *Halorubrum ezzemoulense* Kharroub et al. 2006. Emended description of *Halorubrum ezzemoulense* Kharroub et al. 2006. *Int. J. Syst. Evol. Microbiol.* 68, 3657–3665. doi: 10.1099/ijsem.0.003005

Cram, J. A., Chow, C.-E. T., Sachdeva, R., Needham, D. M., Parada, A. E., Steele, J. A., et al. (2015). Seasonal and interannual variability of the marine bacterioplankton community throughout the water column over ten years. *ISME J.* 9, 563–580. doi: 10.1038/ismej.2014.153

Cravatt, B. F., Prospero-Garcia, O., Siuzdak, G., Gilula, N. B., Henriksen, S. J., Boger, D. L., et al. (1995). Chemical characterization of a family of brain lipids that induce sleep. *Science* 268, 1506–1509. doi: 10.1126/science.7770779

de la Haba, R. R., Corral, P., Sánchez-Porro, C., Infante-Domínguez, C., Makkay, A. M., Amoozegar, M. A., et al. (2018). Genotypic and lipid analyses of strains from the archaeal genus *Halorubrum* reveal insights into their taxonomy, divergence, and population structure. *Front. Microbiol.* 9:512. doi: 10.3389/fmicb.2018.00512

de la Haba, R. R., López-Hermoso, C., Sánchez-Porro, C., Konstantinidis, K. T., and Ventosa, A. (2019). Comparative genomics and phylogenomic analysis of the genus *Salinivibrio*. *Front. Microbiol.* 10:2104. doi: 10.3389/fmicb.2019.02104

Díez-Vives, C., Nielsen, S., Sánchez, P., Palenzuela, O., Ferrera, I., Sebastián, M., et al. (2019). Delineation of ecologically distinct units of marine Bacteroidetes in the Northwestern Mediterranean Sea. *Mol. Ecol.* 28, 2846–2859. doi: 10.1111/mec.15068

Edgar, R. C. (2004). MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 32, 1792–1797. doi: 10.1093/nar/gkh340

Felsenstein, J. (1981). Evolutionary trees from DNA sequences: A maximum likelihood approach. *J. Mol. Evol.* 17, 368–376. doi: 10.1007/BF01734359

Felsenstein, J. (1985). Confidence limits on phylogenies: An approach using the bootstrap. *Evolution* 39, 783–791. doi: 10.1111/j.1558-5646.1985.tb00420.x

Fenical, W., Jensen, P. R., Palladino, M. A., Lam, K. S., Lloyd, G. K., and Potts, B. C. (2009). Discovery and development of the anticancer agent salinosporamide A (NPI-0052). *Bioorg. Med. Chem.* 17, 2175–2180. doi: 10.1016/j.bmc.2008.10.075

Fernández-Gómez, B., Richter, M., Schüler, M., Pinhassi, J., Acinas, S. G., González, J. M., et al. (2013). Ecology of marine bacteroidetes: A comparative genomics approach. *ISME J.* 7, 1026–1037. doi: 10.1038/ismej.2012.169

Fitch, W. M. (1971). Toward defining the course of evolution: Minimum change for a specific tree topology. *Syst. Biol.* 20, 406–416. doi: 10.1093/sysbio/20.4.406

Fuhrman, J. A., Cram, J. A., and Needham, D. M. (2015). Marine microbial community dynamics and their ecological interpretation. *Nat. Rev. Microbiol.* 13, 133–146. doi: 10.1038/nrmicro3417

Galluzzi, L., Buqué, A., Kepp, O., Zitvogel, L., and Kroemer, G. (2015). Immunological effects of conventional chemotherapy and targeted anticancer agents. *Cancer Cell* 28, 690–714. doi: 10.1016/j.ccell.2015.10.012

García-Davis, S., Reyes, C. P., Lagunes, I., Padrón, J. M., Fraile-Nuez, E., Fernández, J. J., et al. (2021). Bioprospecting antiproliferative marine microbiota from submarine volcano Tagoro. *Front. Mar. Sci.* 8:687701. doi: 10.3389/fmars.2021.687701

Ghiglione, J. F., and Murray, A. E. (2012). Pronounced summer to winter differences and higher wintertime richness in coastal Antarctic marine bacterioplankton: Temporal variation in Southern Ocean coastal bacterioplankton. *Environ. Microbiol.* 14, 617–629. doi: 10.1111/j.1462-2920.2011.02601.x

Gilbert, J. A., Steele, J. A., Caporaso, J. G., Steinbrück, L., Reeder, J., Temperton, B., et al. (2012). Defining seasonal marine microbial community dynamics. *ISME J.* 6, 298–308. doi: 10.1038/ismej.2011.107

Giovannoni, S. J., and Vergin, K. L. (2012). Seasonality in ocean microbial communities. *Science* 335, 671–676. doi: 10.1126/science.1198078

Gómez-Pereira, P. R., Fuchs, B. M., Alonso, C., Oliver, M. J., van Beusekom, J. E. E., and Amann, R. (2010). Distinct flavobacterial communities in contrasting water masses of the North Atlantic Ocean. *ISME J.* 4, 472–487. doi: 10.1038/ismej.2009.142

Goris, J., Konstantinidis, K. T., Klappenbach, J. A., Coenye, T., Vandamme, P., and Tiedje, J. M. (2007). DNA–DNA hybridization values and their relationship to whole-genome sequence similarities. *Int. J. Syst. Evol. Microbiol.* 57, 81–91. doi: 10.1099/ijs.0.64483-0

Grzymski, J. J., Riesenfeld, C. S., Williams, T. J., Dussaq, A. M., Ducklow, H., Erickson, M., et al. (2012). A metagenomic assessment of winter and summer bacterioplankton from Antarctica Peninsula coastal surface waters. *ISME J.* 6, 1901–1915. doi: 10.1038/ismej.2012.31

Gurevich, A., Saveliev, V., Vyahhi, N., and Tesler, G. (2013). QUAST: Quality assessment tool for genome assemblies. *Bioinformatics* 29, 1072–1075. doi: 10.1093/bioinformatics/btt086

Haro-Moreno, J. M., López-Pérez, M., and Rodriguez-Valera, F. (2021). Enhanced recovery of microbial genes and genomes from a marine water column using long-read metagenomics. *Front. Microbiol.* 12:708782. doi: 10.3389/fmicb.2021.708782

Haro-Moreno, J. M., López-Pérez, M., de la Torre, J. R., Picazo, A., Camacho, A., and Rodriguez-Valera, F. (2018). Fine metagenomic profile of the Mediterranean stratified and mixed water columns revealed by assembly and recruitment. *Microbiome* 6:128. doi: 10.1186/s40168-018-0513-5

Haro-Moreno, J. M., Rodriguez-Valera, F., López-García, P., Moreira, D., and Martin-Cuadrado, A.-B. (2017). New insights into marine group III Euryarchaeota, from dark to light. *ISME J.* 11, 1102–1117. doi: 10.1038/ismej.2016.188

Ikegami, Y., Takeuchi, N., Hanada, M., Hasegawa, Y., Ishii, K., Andoh, T., et al. (1990). Topostin, a novel inhibitor of mammalian DNA topoisomerase I from *Flexibacter topostinus* sp. nov. II. Purification and some properties of topostin. *J. Antibiot.* 43, 158–162. doi: 10.7164/antibiotics.43.158

Infante-Domínguez, C., de la Haba, R. R., Corral, P., Sanchez-Porro, C., Arahal, D. R., and Ventosa, A. (2020). Genome-based analyses reveal a synonymy among *Halorubrum distributum* Zvyagintseva and Tarasov 1989; Oren and Ventosa 1996, *Halorubrum terrestre* Ventosa et al. 2004, *Halorubrum arcis* Xu et al. 2007 and *Halorubrum litoreum* Cui et al. 2007. Emended description of *Halorubrum distributum* Zvyagintseva and Tarasov 1989; Oren and Ventosa 1996. *Int. J. Syst. Evol. Microbiol.* 70, 1698–1705. doi: 10.1099/ijsem.0.003956

Jones, D. T., Taylor, W. R., and Thornton, J. M. (1992). The rapid generation of mutation data matrices from protein sequences. *Bioinformatics* 8, 275–282. doi: 10.1093/bioinformatics/8.3.275

Jukes, T. H., and Cantor, C. R. (1969). "Evolution of protein molecules," in *Mammalian protein metabolism*, ed. H. N. Munro (Amsterdam: Elsevier), 21–132. doi: 10.1016/B978-1-4832-3211-9.50009-7

Kamiyama, T., Umino, T., Satoh, T., Sawairi, S., Shirane, M., Onshima, S., et al. (1995). Sulfobacins A and B, Novel von Willebrand factor receptor antagonists. I. Production, isolation, characterization and biological activities. *J. Antibiot.* 48, 924–928. doi: 10.7164/antibiotics.48.924

Kato, T. A., Tsuda, A., Uesaka, M., Fujimori, A., Kamada, T., Tsujii, H., et al. (2011). In vitro characterization of cells derived from chordoma cell line U-CH1 following treatment with X-rays, heavy ions and chemotherapeutic drugs. *Radiat. Oncol.* 6:116. doi: 10.1186/1748-717X-6-116

Kautsar, S. A., Blin, K., Shaw, S., Navarro-Muñoz, J. C., Terlouw, B. R., van der Hooft, J. J. J., et al. (2019). MIBiG 2.0: A repository for biosynthetic gene clusters of known function. *Nucleic Acids Res.* 48, D454–D458. doi: 10.1093/nar/gkz882

Kelly, L. W., Nelson, C. E., Haas, A. F., Naliboff, D. S., Calhoun, S., Carlson, C. A., et al. (2019). Diel population and functional synchrony of microbial communities on coral reefs. *Nat. Commun.* 10:1691. doi: 10.1038/s41467-019-09419-z

Kirchman, D. L. (2002). The ecology of *Cytophaga*–flavobacteria in aquatic environments. *FEMS Microbiol. Ecol.* 39, 91–100. doi: 10.1111/j.1574-6941.2002.tb00910.x

Kirchman, D. L., Yu, L., and Cottrell, M. T. (2003). Diversity and abundance of uncultured *Cytophaga* -like bacteria in the Delaware Estuary. *Appl. Environ. Microbiol.* 69, 6587–6596. doi: 10.1128/AEM.69.11.6587-6596.2003

Kohayashi, J., Mikami, S., Shigemori, H., Takao, T., Shimonishi, Y., Izuta, S., et al. (1995). Flavocristamides A and B, new DNA polymerase α inhibitors from a marine bacterium *Flavobacterium* sp. *Tetrahedron* 51, 10487–10490. doi: 10.1016/0040-4020(95)00631-H

Konstantinidis, K. T., and Tiedje, J. M. (2005). Genomic insights that advance the species definition for prokaryotes. *Proc. Natl. Acad. Sci. U.S.A.* 102, 2567–2572. doi: 10.1073/pnas.0409727102

Krzywinski, M., Schein, J., Birol, İ, Connors, J., Gascoyne, R., Horsman, D., et al. (2009). Circos: An information aesthetic for comparative genomics. *Genome Res.* 19, 1639–1645. doi: 10.1101/gr.092759.109

Letunic, I., and Bork, P. (2021). Interactive Tree Of Life (iTOL) v5: An online tool for phylogenetic tree display and annotation. *Nucleic Acids Res.* 49, W293–W296. doi: 10.1093/nar/gkab301

Letzel, A.-C., Li, J., Amos, G. C. A., Millán-Aguiñaga, N., Ginigini, J., Abdelmohsen, U. R., et al. (2017). Genomic insights into specialized metabolism in the marine actinomycete *Salinispora*: Genomic insights into specialized metabolism. *Environ. Microbiol.* 19, 3660–3673. doi: 10.1111/1462-2920.13867

Liu, C.-M., Wong, T., Wu, E., Luo, R., Yiu, S.-M., Li, Y., et al. (2012). SOAP3: Ultra-fast GPU-based parallel alignment tool for short reads. *Bioinformatics* 28, 878–879. doi: 10.1093/bioinformatics/bts061

Liu, Q., Li, J., Wei, B., Zhang, X., Zhang, L., Zhang, Y., et al. (2016). *Leeuwenhoekiella nanhaiensis* sp. nov., isolated from deep-sea water. *Int. J. Syst. Evol. Microbiol.* 66, 1352–1357. doi: 10.1099/ijsem.0.000883

López-Pérez, M., Haro-Moreno, J. M., Gonzalez-Serrano, R., Parras-Moltó, M., and Rodriguez-Valera, F. (2017). Genome diversity of marine phages recovered from Mediterranean metagenomes: Size matters. *PLoS Genet.* 13:e1007018. doi: 10.1371/journal.pgen.1007018

Maeda, H., and Khatami, M. (2018). Analyses of repeated failures in cancer therapy for solid tumors: Poor tumor-selective drug delivery, low therapeutic efficacy and unsustainable costs. *Clin. Transl. Med.* 7:11. doi: 10.1186/s40169-018-0185-6

Martín, J., Crespo, G., González-Menéndez, V., Pérez-Moreno, G., Sánchez-Carrasco, P., Pérez-Victoria, I., et al. (2014). MDN-0104, an antiplasmodial betaine lipid from *Heterospora chenopodii*. *J. Nat. Prod.* 77, 2118–2123. doi: 10.1021/np500577v

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

Mena, C., Reglero, P., Balbín, R., Martín, M., Santiago, R., and Sintes, E. (2020). Seasonal niche partitioning of surface temperate open ocean prokaryotic communities. *Front. Microbiol.* 11:1749. doi: 10.3389/fmicb.2020.01749

Mestre, M., Borrull, E., Sala, M. M., and Gasol, J. M. (2017). Patterns of bacterial diversity in the marine planktonic particulate matter continuum. *ISME J.* 11, 999–1010. doi: 10.1038/ismej.2016.166

Mestre, M., Höfer, J., Sala, M. M., and Gasol, J. M. (2020). Seasonal variation of bacterial diversity along the marine particulate matter continuum. *Front. Microbiol.* 11:1590. doi: 10.3389/fmicb.2020.01590

Mestre, M., Ruiz-González, C., Logares, R., Duarte, C. M., Gasol, J. M., and Sala, M. M. (2018). Sinking particles promote vertical connectivity in the ocean microbiome. *Proc. Natl. Acad. Sci. U.S.A.* 115, E6799–E6807. doi: 10.1073/pnas.1802470115

MIDI (2012). *Sherlock microbial identification system operating manual*. Newark, DE: MIDI.

Mistry, J., Chuguransky, S., Williams, L., Qureshi, M., Salazar, G. A., Sonnhammer, E. L. L., et al. (2021). Pfam: The protein families database in 2021. *Nucleic Acids Res.* 49, D412–D419. doi: 10.1093/nar/gkaa913

Navarro-Muñoz, J. C., Selem-Mojica, N., Mullowney, M. W., Kautsar, S. A., Tryon, J. H., Parkinson, E. I., et al. (2020). A computational framework to explore large-scale biosynthetic diversity. *Nat. Chem. Biol.* 16, 60–68. doi: 10.1038/s41589-019-0400-9

Nedashkovskaya, O. I., Vancanneyt, M., Dawyndt, P., Engelbeen, K., Vandemeulebroecke, K., Cleenwerck, I., et al. (2005). Reclassification of [*Cytophaga*] *marinoflava* Reichenbach 1989 as *Leeuwenhoekiella marinoflava* gen. nov., comb. nov. and description of *Leeuwenhoekiella aequorea* sp. nov. *Int. J. Syst. Evol. Microbiol.* 55, 1033–1038. doi: 10.1099/ijs.0.63410-0

Oliver, L., Lalier, L., Salaud, C., Heymann, D., Cartron, P. F., and Vallette, F. M. (2020). Drug resistance in glioblastoma: Are persisters the key to therapy? *Cancer Drug Resist.* 3, 287–301. doi: 10.20517/cdr.2020.29

Omand, M. M., Steinberg, D. K., and Stamieszkin, K. (2021). Cloud shadows drive vertical migrations of deep-dwelling marine life. *Proc. Natl. Acad. Sci. U.S.A.* 118:e2022977118. doi: 10.1073/pnas.2022977118

Oraiopoulou, M.-E., Tzamali, E., Tzedakis, G., Vakis, A., Papamatheakis, J., and Sakkalis, V. (2017). in vitro/in silico study on the role of doubling time heterogeneity among primary glioblastoma cell lines. *Biomed. Res. Int.* 2017, 1–12. doi: 10.1155/2017/8569328

Oren, A., and Garrity, G. M. (2021). Valid publication of the names of forty-two phyla of prokaryotes. *Int. J. Syst. Evol. Microbiol.* 71:005056. doi: 10.1099/ijsem.0.005056

Osuka, S., and Van Meir, E. G. (2017). Overcoming therapeutic resistance in glioblastoma: The way forward. *J. Clin. Invest.* 127, 415–426. doi: 10.1172/JCI89587

Paoli, L., Ruscheweyh, H.-J., Forneris, C. C., Hubrich, F., Kautsar, S., Bhushan, A., et al. (2022). Biosynthetic potential of the global ocean microbiome. *Nature* 607, 111–118. doi: 10.1038/s41586-022-04862-3

Parks, D. H., Imelfort, M., Skennerton, C. T., Hugenholtz, P., and Tyson, G. W. (2015). CheckM: Assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res.* 25, 1043–1055. doi: 10.1101/gr.186072.114

Parte, A. C., Sardà Carbasse, J., Meier-Kolthoff, J. P., Reimer, L. C., and Göker, M. (2020). List of prokaryotic names with standing in nomenclature (LPSN) moves to the DSMZ. *Int. J. Syst. Evol. Microbiol.* 70, 5607–5612. doi: 10.1099/ijsem.0.004332

Pascoal, F., Magalhães, C., and Costa, R. (2020). The link between the ecology of the prokaryotic rare biosphere and its biotechnological potential. *Front. Microbiol.* 11:231. doi: 10.3389/fmicb.2020.00231

Pedrós-Alió, C. (2007). Dipping into the rare biosphere. *Science* 315, 192–193. doi: 10.1126/science.1135933

Pedrós-Alió, C. (2012). The rare bacterial biosphere. *Annu. Rev. Mar. Sci.* 4, 449–466. doi: 10.1146/annurev-marine-120710-100948

Pedrós-Alió, C. (2013). "Rare biosphere," in *Encyclopedia of biodiversity*, ed. S. Levi (Amsterdam: Elsevier), 345–352. doi: 10.1016/B978-0-12-384719-5.00407-X

Pesant, S., Not, F., Picheral, M., Kandels-Lewis, S., Le Bescot, N., Gorsky, G., et al. (2015). Open science resources for the discovery and analysis of Tara Oceans data. *Sci. Data* 2:150023. doi: 10.1038/sdata.2015.23

Price, M. N., Dehal, P. S., and Arkin, A. P. (2010). FastTree 2 – Approximately maximum-likelihood trees for large alignments. *PLoS One* 5:e9490. doi: 10.1371/journal.pone.0009490

Reichenbach, H. (1992). Validation of the publication of new names and new combinations previously effectively published outside the IJSB: List no. 41. *Int. J. Syst. Bacteriol.* 42, 327–328. doi: 10.1099/00207713-42-2-327

Richter, M., and Rosselló-Móra, R. (2009). Shifting the genomic gold standard for the prokaryotic species definition. *Proc. Natl. Acad. Sci. U.S.A.* 106, 19126–19131. doi: 10.1073/pnas.0906412106

Rooney-Varga, J. N., Giewat, M. W., Savin, M. C., Sood, S., LeGresley, M., and Martin, J. L. (2005). Links between phytoplankton and bacterial community dynamics in a coastal marine environment. *Microb. Ecol.* 49, 163–175. doi: 10. 1007/s00248-003-1057-0

Roth, P., Mason, W. P., Richardson, P. G., and Weller, M. (2020). Proteasome inhibition for the treatment of glioblastoma. *Exp. Opin. Invest. Drugs* 29, 1133–1141. doi: 10.1080/13543784.2020.1803827

Saitou, N., and Nei, M. (1987). The neighbor-joining method: A new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* 4, 406–425. doi: 10.1093/oxfordjournals.molbev.a040454

Sanz-Sáez, I., Salazar, G., Sánchez, P., Lara, E., Royo-Llonch, M., Sà, E. L., et al. (2020). Diversity and distribution of marine heterotrophic bacteria from a large culture collection. *BMC Microbiol.* 20:207. doi: 10.1186/s12866-020-01884-7

Sasser, M. (1990). *Bacterial identification by gas chromatographic analysis of fatty acid methyl esters (GC-FAME)*. Newark, DE: MIDI.

Scherlach, K., and Hertweck, C. (2021). Mining and unearthing hidden biosynthetic potential. *Nat. Commun.* 12:3864. doi: 10.1038/s41467-021-24133-5

Shade, A., Jones, S. E., Caporaso, J. G., Handelsman, J., Knight, R., Fierer, N., et al. (2014). Conditionally rare taxa disproportionately contribute to temporal changes in microbial diversity. *mbio* 5:e01371–14. doi: 10.1128/mBio.01371-14

Shimodaira, H., and Hasegawa, M. (1999). Multiple comparisons of log-likelihoods with applications to phylogenetic inference. *Mol. Biol. Evol.* 16, 1114–1116. doi: 10.1093/oxfordjournals.molbev.a026201

Si, O.-J., Kim, S.-J., Jung, M.-Y., Choi, S.-B., Kim, J.-G., Kim, S.-G., et al. (2015). *Leeuwenhoekiella polynyae* sp. nov., isolated from a polynya in Western Antarctica. *Int. J. Syst. Evol. Microbiol.* 65, 1694–1699. doi: 10.1099/ijs.0.000160

Siegel, R. L., Miller, K. D., Fuchs, H. E., and Jemal, A. (2021). Cancer statistics, 2021. *CA Cancer J. Clin.* 71, 7–33. doi: 10.3322/caac.21654

Sogin, M. L., Morrison, H. G., Huber, J. A., Welch, D. M., Huse, S. M., Neal, P. R., et al. (2006). Microbial diversity in the deep sea and the underexplored "rare biosphere.". *Proc. Natl. Acad. Sci. U.S.A.* 103, 12115–12120. doi: 10.1073/pnas.0605127103

Stackebrandt, E., and Goebel, B. M. (1994). Taxonomic note: A place for DNA-DNA reassociation and 16S rRNA sequence analysis in the present species definition in bacteriology. *Int. J. Syst. Evol. Microbiol.* 44, 846–849. doi: 10.1099/00207713-44-4-846

Sunagawa, S., Coelho, L. P., Chaffron, S., Kultima, J. R., Labadie, K., Salazar, G., et al. (2015). Structure and function of the global ocean microbiome. *Science* 348:1261359. doi: 10.1126/science.1261359

Suzuki, K., Yamaguchi, H., Miyazaki, S., Nagai, K., Watanabe, S.-I., Saito, T., et al. (1990). Topostin, a novel inhibitor of mammalian DNA topoisomerase I from

Flexibacter topostinus sp. nov. I. Taxonomy, and fermentation of producing strain. *J. Antibiot.* 43, 154–157. doi: 10.7164/antibiotics.43.154

Tahon, G., Lebbe, L., De Troch, M., Sabbe, K., and Willems, A. (2020). *Leeuwenhoekiella aestuarii* sp. nov., isolated from salt-water sediment and first insights in the genomes of *Leeuwenhoekiella* species. *Int. J. Syst. Evol. Microbiol.* 70, 1706–1719. doi: 10.1099/ijsem.0.003959

Tatusova, T., DiCuccio, M., Badretdin, A., Chetvernin, V., Nawrocki, E. P., Zaslavsky, L., et al. (2016). NCBI prokaryotic genome annotation pipeline. *Nucleic Acids Res.* 44, 6614–6624. doi: 10.1093/nar/gkw569

Tavaré, S., and Miura, R. M. (1986). "Some probabilistic and statistical problems in the analysis of DNA sequences," in *Lectures on mathematics in the life sciences*, ed. R. M. Miura (Providence, RI: American Mathematical Society), 57–86.

Teeling, H., Fuchs, B. M., Bennke, C. M., Krüger, K., Chafee, M., Kappelmann, L., et al. (2016). Recurring patterns in bacterioplankton dynamics during coastal spring algae blooms. *Elife* 5:e11888. doi: 10.7554/eLife.11888

Uchida, I., Yoshida, K., Kawai, Y., Takase, S., Itoh, Y., Tanaka, H., et al. (1985). Structure and synthesis of WB-3559 A, B, C and D, new fibrinolytic agents isolated from *Flavobacterium* sp. *Chem. Pharm. Bull.* 33, 424–427. doi: 10.1248/cpb.33.424

van Heel, A. J., de Jong, A., Song, C., Viel, J. H., Kok, J., and Kuipers, O. P. (2018). BAGEL4: A user-friendly web server to thoroughly mine RiPPs and bacteriocins. *Nucleic Acids Res.* 46, W278–W281. doi: 10.1093/nar/gky383

Wenley, J., Currie, K., Lockwood, S., Thomson, B., Baltar, F., and Morales, S. E. (2021). Seasonal prokaryotic community linkages between surface and deep ocean water. *Front. Mar. Sci.* 8:659641. doi: 10.3389/fmars.2021.65 9641

Westram, R., Bader, K., Prüsse, E., Kumar, Y., Meier, H., Glöckner, F. O., et al. (2011). "ARB: A software environment for sequence data," in *Handbook of molecular microbial ecology I*, ed. F. J. de Bruijn (Hoboken, NJ: John Wiley & Sons, Inc.), 399–406. doi: 10.1002/9781118010518.ch46

Xue, C., Zhang, H., Lin, H., Sun, Y., Luo, D., Huang, Y., et al. (2020). Ancestral niche separation and evolutionary rate differentiation between sister marine flavobacteria lineages. *Environ. Microbiol.* 22, 3234–3247. doi: 10.1111/1462-2920. 15065

Yoon, S.-H., Ha, S.-M., Kwon, S., Lim, J., Kim, Y., Seo, H., et al. (2017). Introducing EzBioCloud: A taxonomically united database of 16S rRNA gene sequences and whole-genome assemblies. *Int. J. Syst. Evol. Microbiol.* 67, 1613–1617. doi: 10.1099/ijsem.0.001755

Yoshida, K., Iwami, M., Umehara, Y., Nishikawa, M., Uchida, I., Kohsaka, M., et al. (1985). Studies on WB-3559 A,B,C and D, new potent fibrinolytic agents. I. Discovery, identification, isolation and characterization. *J. Antibiot.* 38, 1469–1475. doi: 10.7164/antibiotics.38.1469

Zhou, Q., Hotta, K., Deng, Y., Yuan, R., Quan, S., and Chen, X. (2021). Advances in biosynthesis of natural products from marine microorganisms. *Microorganisms* 9:2551. doi: 10.3390/microorganisms9122551

Zuchegna, C., Di Zazzo, E., Moncharmont, B., and Messina, S. (2020). Dual-specificity phosphatase (DUSP6) in human glioblastoma: Epithelial-to-mesenchymal transition (EMT) involvement. *BMC Res. Notes* 13:374. doi: 10.1186/s13104-020-05214-y