Marine Genomics 61 (2022) 100922

Contents lists available at ScienceDirect

Marine Genomics

journal homepage: www.elsevier.com/locate/margen

Data

Genome analysis of a new biosurfactants source: The Antarctic bacterium *Psychrobacter* sp. TAE2020

Christopher Riccardi [a,*], Caterina D'Angelo [b], Marzia Calvanese [b], Annarita Ricciardelli [c], Maria Luisa Tutino [b], Ermenegilda Parrilli [b], Marco Fondi [a]

[a] Department of Biology, University of Florence, Via Madonna del Piano 6, Sesto Fiorentino, Italy
[b] Department of Chemical Sciences, University of Naples Federico II, Complesso Universitario Monte S. Angelo, via Cinthìa, 80125 Naples, Italy
[c] Department of Marine Biotechnology, Stazione Zoologica "Anton Dohrn", 80121 Naples, Italy

ARTICLE INFO

Keywords:
Surfactants
Cold-adapted bacteria
Psychrobacter sp.
Next generation sequencing

ABSTRACT

Biosurfactants are considered a possible green alternative to chemical surfactants for countless commercial products including detergents and cleaners, personal care products, cosmetics, pharmaceuticals and therapeutics, food additives, emulsifiers, and dispersants for bioremediation. Organisms from extreme environments are well-adapted to the harsh conditions and represent an exciting avenue of discovery of naturally occurring biosurfactants. In this study, we report the genome analysis of *Psychrobacter* sp. TAE2020, an aerobic γ-proteobacterium isolated from an Antarctic coastal seawater sample collected in the vicinity of the French Antarctic station Dumont d'Urville, Terre Adelie (66°40′ S; 140° 01′ E) which has been shown to produce biosurfactants. Biochemical assays indicate that *Psychrobacter* sp. TAE2020 can produce one or more excellent emulsifiers and a biosurfactant which is able to reduce the surface tension of a Gut medium. Next generation sequencing and genome mining allowed the identification of a plethora of biosynthetic gene clusters possibly involved in the production of emulsifying agents, just waiting to be isolated and characterized. This study paves the way for a more thorough investigation into the potential biotechnological applications of this new Antarctic strain.

1. Introduction

Chemical and microbial surfactants are amphiphilic compounds that are able to reduce both surface and interfacial tension, and to form emulsions from two immiscible liquids. These molecules have many industrial applications in pharmaceuticals, cosmetics, food additives, herbicides and pesticides fields (Tripathi et al., 2018). Compared to their chemically synthesized equivalents, biosurfactants have the advantage of being biodegradable and having low toxicity (Mukherjee et al., 2006), thus representing an environmentally friendly alternative. Moreover, their use has been extensively studied for the enhancement of oil recovery and remediation of hydrocarbon and metal-contaminated soil and water (Tripathi et al., 2018). Cold habitats are extraordinary reservoirs of molecules of biotechnological interest, such as cold-active enzymes, antibiotics, and biosurfactants (Trudgeon et al., 2020). Notably, *Pseudomonas*, *Pseudoalteromonas*, *Marinomonas*, *Halomonas*, *Rhodococcus* and *Cobetia* genera have been described as biosurfactant producers from polar oceans, deep sea, and marine sediments (Perfumo et al., 2018). Biosurfactants play key ecological roles in many cold soils and marine environments (Trudgeon et al., 2020; Dang et al., 2016), for instance, they participate in carbon-cycling processes by enhancing the bioavailability of poorly soluble compounds. Furthermore, they have the ability to interact with multiple physical phases (water, ice, hydrophobic compounds, and gases) at low and freezing temperatures, and for this reason they can be used to develop new biotechnological products and processes that have low energy demand and operate under low-temperature regimes (Trudgeon et al., 2020; Malavenda et al., 2015). In this paper we report the genome sequence and analysis for *Psychrobacter* sp. TAE2020, a previously uncharacterized Antarctic marine bacterium that is able to produce surfactants, and whose biosynthetic gene clusters could make it an eligible candidate for biotechnological applications. Minimum information for the hereby presented sequence (MIxS) is reported in Table 1.

* Corresponding author.
E-mail address: christopher.riccardi@unifi.it (C. Riccardi).

https://doi.org/10.1016/j.margen.2021.100922
Received 31 August 2021; Received in revised form 28 December 2021; Accepted 28 December 2021
Available online 3 January 2022

Table 1
MIxS Checklist for the newly sequenced strain *Psychrobacter* sp. TAE2020.

Item	Description
Investigation type	Bacteria_archaea
Project name	Transcriptional Regulatory Networks in Antarctic bacteria as a proxy for global warming effects on microbial life in the Polar Oceans
Geographic location (latitude and longitude)	66°40' S; 140° 01' E
Geographic location (sea, region)	French Antarctic station Dumont d'Urville, Terre Adélie
Collection date	1994–09
Broad-scale environmental context	Marine biome [ENVO:00000447]
Environmental medium	Terre Adelie [GAZ:00008877]
Local environment context	Coastal sea water [ENVO:00002150]
Observed biotic relationship	Free living
Encoded traits	Biosurfactants producer
Relationship to Oxygen	Aerobe
Sequencing method	Illumina HiSeq 2500
Assembly Software	SPAdes;3.12.0
Annotation	Prokka;1.12
Number of contigs	33
Feature prediction	Prodigal;2.6.3
16S recovered software	Barrnap;0.9
Number of standard tRNAs extracted	42
tRNA extraction software	Prokka;1.12

2. Data description

2.1. Bacterial strains and culture conditions

Psychrobacter sp. TAE2020 was grown in the synthetic medium Gut (L-Glutamic acid 10 g/L, NaCl 10 g/L; NH_4NO_3 1 g/L; $KH_2PO \cdot 7\ H_2O$ 1 g/L; $MgSO_4 \cdot 7\ H_2O$ 200 mg/L; $FeSO_4 \cdot 7\ H_2O$ 5 mg/L; $CaCl_2 \cdot 2\ H_2O$ 5 mg/L) in planktonic conditions at 15 °C under vigorous agitation (180 rpm) up till the stationary phase of growth (72 h). The supernatant was recovered by centrifugation at 6000 rpm for 30 min at 4 °C, sterilized by filtration through membranes with a pore diameter of 0.22 μm and stored at 4 °C until use.

2.2. Drop-collapse assay

The drop-collapse assay (Jain et al., 1991) was performed as a qualitative test for the identification of a biosurfactants producer strain. In detail, drop collapse assay was performed by the deposition of 50 μL droplets of samples (supernatant or only medium) on a hydrophobic surface. Methylene blue was added to stain the samples for photographic purposes and did not influence the shape of the droplets. The spreading of the droplet on parafilm was observed after 1 h.

2.3. Emulsification index

The emulsification index (E24), (Blesic et al., 2018) of samples was determined by adding 2 mL of dectol (decane and toluene in 65:35 volume ratio) and 1 mL of the cell-free supernatant in a test tube, vortexed at maximum speed for 3 min and allowed to stand for 24 h. The emulsification index, E24, was determined by calculating the ratio between the height of emulsifying layer and the total height, multiplied by 100.

2.4. Surface tension measurements

The surface tension, γ, was measured through De Nouy ring method using a KSV Sigma 70 digital tensiometer (Dyne Testing Ltd., Newton House, Lichfield, UK) equipped with an automatic device to set the time between two consecutive measurements and to select the rising velocity of the platinum ring (D'Errico et al., 2005). The ring rising velocity was set low enough to reach the equilibrium between the air–solution interface and the solution bulk. At least three surface tension measurements were performed on a 10 mL sample.

2.5. Psychrobacter sp. TAE2020 surfactants production

Cell-free supernatant of *Psychrobacter* sp. TAE2020 (*Psychrobacter* sp. TAE2020-SN) was analyzed by drop collapse assay (Fig. 1A), a sensitive and rapid method for the screening of surfactants producing bacteria. In general, if the liquid does not contain surfactants, polar water molecules are repelled from the hydrophobic surface and the drops remain stable, whereas if the liquid contains surfactants, the drops spread or even collapse because of the reduced interfacial tension between the liquid drop and the hydrophobic surface. As shown in Fig. 1A, the drop of *Psychrobacter* sp. TAE2020 cell-free supernatant collapsed, suggesting the presence of compound/s acting as surfactants. To evaluate if *Psychrobacter* sp. TAE2020 bioactive compound was also able to reduce the surface tension of an aqueous solution, the Du Nouy ring method was applied and the surface tension value obtained for *Psychrobacter* sp. TAE2020 cell-free supernatant resulted to be lower than that of Gut medium, 60.63 (± 0,5) and 77.49 (± 0,2) mN/m, respectively. Biosurfactants can also form emulsions that confer high long-term stability to the dispersed phase droplets in emulsions between two immiscible liquids. *Psychrobacter* sp. TAE2020-SN emulsification ability was tested in presence of 2 mL of Dectol, a "model oil", compared to the negative control (Gut medium), by measuring the emulsification index E24. *Psychrobacter* sp. TAE2020-SN showed an E24 value of 70% (Fig. 1B). Taken together, these results demonstrate that *Psychrobacter* sp. TAE2020-SN contains an excellent emulsifier and a biosurfactant that is able to reduce the surface tension of Gut medium.

2.6. Genome annotation and bioinformatics analysis

A total of 1,225,606 paired ended sequencing reads were collected and submitted to Trimmomatic (Bolger et al., 2014) for quality filtering and trimming. Reads that were not suitable for analysis due to low quality were discarded, while those having a low score towards the 3′ end were shortened, leaving a number of 1,218,486 idoneous reads, most of which measured 251 nucleotides in length (879,430, or 72.17% of the total). Genome de novo assembly was performed using SPAdes (Prjibelski et al., 2020), leading to a draft genome embedding 3,294,027 nucleotides (N50: 179582, L50: 7, contigs: 74, coverage: 89×, G + C content: 41.7%). The obtained scaffolds were then assayed for the presence of 16S rRNA gene using ContEst16S web server (Lee et al., 2017), which predicted the presence of one 1539 bp 16S rRNA gene that allowed to preliminarily assign the genome to the *Psychrobacter* genus, given the 99.22% identity at 100% query cover with *Psychrobacter urativorans* strain R10.10B. In order to further refine the ordering and orientation of the contigs obtained through the assembly procedure, all available *Psychrobacter* genomic sequences were downloaded from the NCBI and we made use of the multi draft-based scaffolder MeDuSa (Bosi et al., 2015), which remarkably improved the overall quality of the newly sequenced genome. When MeDuSa was fed with our newly assembled contigs and a set of 19 complete or draft *Psychrobacter* genomes, it was able to assemble the 74 contigs into 45 scaffolds, and increase the N50 statistic of an order of magnitude, from 179,582 to 3,206,207, thus allowing a more accurate representation of the genome. Contigs of size less than 200 bp were manually removed, leaving a final number of 3,293,789 base pairs distributed throughout 33 scaffolds. After assembly and refinement, we sought to locate the 16S rRNA gene in order to assess to which species the bacterium under investigation is more closely related to. Using Barrnap software for ribosomal RNA prediction (github.com/tseemann/barrnap) on the newly scaffolded genome we were able to retrieve six complete ribosomal RNA sequences (four 5S, one 16S and one 23S) which all displayed between 93 and 98% identity to *P. urativorans* strain R10.10B and *Psychrobacter cryohalolentis*

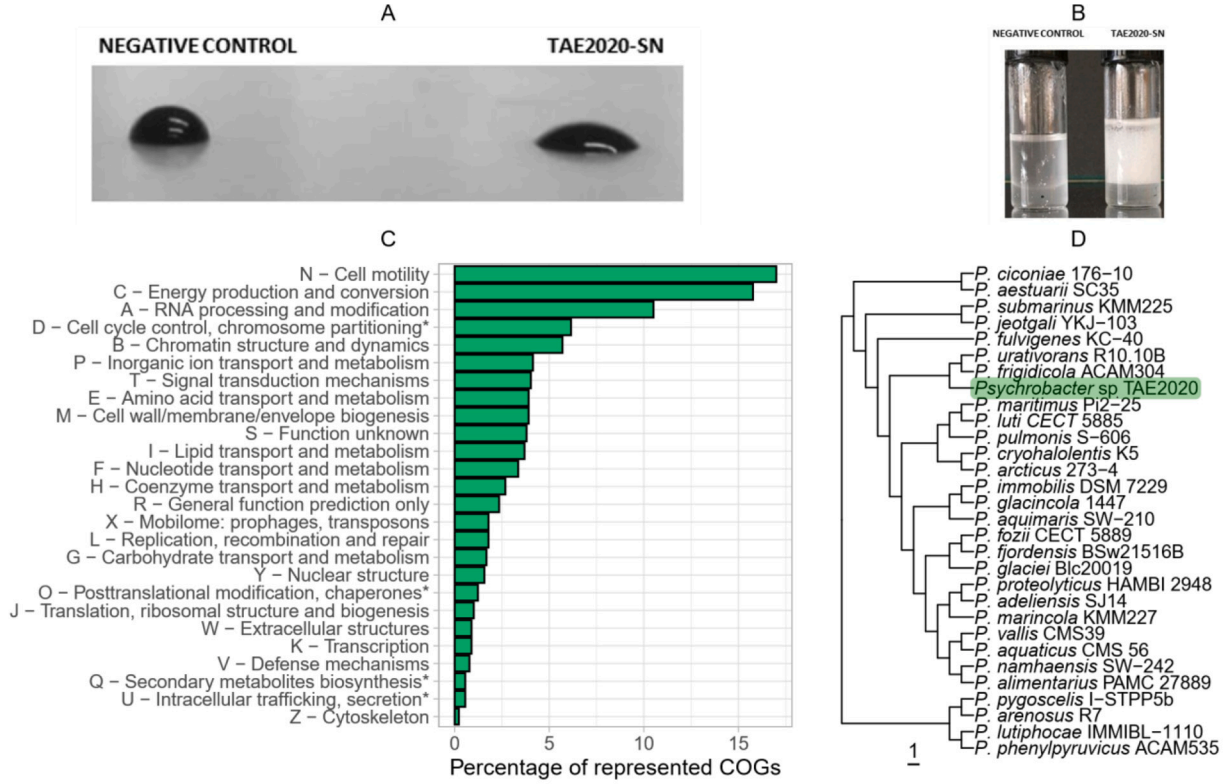

Fig. 1. (A) Drop collapse assay: drop of *Psychrobacter* sp. TAE2020 cell-free supernatant (TAE2020-SN) and negative control (L-Glutamate culture medium) on a hydrophobic surface; (B) Emulsification index: Emulsion of 1 mL of *Psychrobacter* sp. TAE2020 cell-free supernatant (TAE2020-SN) mixed to 2 mL of Dectol after 24 h, in comparison to the negative control (mixture of L-Glutamate culture medium and Dectol, left). (C) COG categories distribution in *Psychrobacter* sp. TAE2020. Descriptions marked with an asterisk were slightly simplified for better readability. Full descriptions available at https://www.ncbi.nlm.nih.gov/research/cog#. (D) Phylogenetic analysis showing the taxonomic relatedness among the newly sequenced strain and the other *Psychrobacter* representatives used within the context of this work. All nodes displayed a 100% bootstrap support therefore bootstrap values were omitted from the visualization. Isolate identifier was added to those species to which a strain has not been assigned yet.

strain FDAARGOS_308, respectively. When limiting the analysis to the 16S rRNA gene sequence, we obtained a match with the corresponding 16S sequence of *P. urativorans* strain DSM 14009 with a query coverage of 99% and sequence identity of 99.93%, and with that of *P. urativorans* strain R10.10B with a query coverage of 100% and sequence identity of 99.22%. To further investigate the genomic background that is behind *Psychrobacter* sp. TAE2020's biosurfactants production, a whole-genome annotation was first conducted using prokka (Seemann, 2014), which allowed to identify 2695 protein products, 931 of which were labelled as "hypothetical protein", and 42 tRNAs. The proteome file in FASTA format that was generated by prokka was then uploaded to eggNOG mapper genome functional annotation webserver (Huerta-Cepas et al., 2019), while the genomic FASTA file was submitted to antiSMASH 6.0 webserver (Blin et al., 2021) for genome mining analysis. The most represented cluster of orthologous groups (COG) category is Cell Motility with 152 genes (17%), followed by Energy Production and conversion (15.7%) and RNA Processing and Modification (10.51%), (Fig. 1C). The predicted, biologically relevant gene cluster (BGC) that shares the highest similarity to a known cluster is involved in the production of a siderophore that is likely to resemble vibrioferrin (36% identity), whereas the remaining ones share a somewhat mild similarity to clusters producing emulsan, surfactin, TP-1161 (Thiopeptide), K53 capsular polysaccharide and, more in general, saccharides (Table 2). It did not escape our notice that the similarity values shared with other

Table 2
antiSMASH-predicted biosynthetic gene clusters. Each column respectively specifies: the region number, the product types as detected by antiSMASH, the nucleotide range within the region, the closest compound from the MiBIG database (Kautsar et al., 2020) and the fraction of genes within the closest known compound that have a significant BLAST hit to the genes within each region.

Region	Type	From	To	Most similar known cluster	Similarity	
Region 1.1	fatty_acid	126,562	145,897			
Region 1.2	siderophore	325,289	340,438	vibrioferrin	Other	36%
Region 1.3	saccharide	350,792	372,207			
Region 1.4	saccharide	491,914	519,018			
Region 1.5	terpene	1,136,338	1,158,478			
Region 1.6	saccharide	1,218,766	1,242,740	K53 capsular polysaccharide	Saccharide	15%
Region 1.7	saccharide	1,913,419	1,930,991	emulsan	Saccharide	13%
Region 1.8	fatty_acid	2,367,607	2,386,641			
Region 1.9	fatty_acid	2,460,581	2,481,819	TP-1161	RiPP:Thiopeptide	16%
Region 1.10	saccharide	2,652,530	2,680,125			
Region 1.11	saccharide	2,857,505	2,882,929	surfactin	NRP:Lipopeptide	8%
Region 4.1	saccharide	1	8863	emulsan	Saccharide	9%

(known) gene clusters involved in biosurfactant/emulsan production is quite low (ranging from 8% to 15%). This raises questions on the actual structure of the biosurfactant gene cluster in *Psychrobacter* sp. TAE2020 and on the exact metabolic steps leading to the synthesis of this valuable compound in this strain. Further work will be necessary to elucidate this point in the future.

2.7. Phylogenetic analysis

To further investigate the evolutionary relatedness between *Psychrobacter* sp. TAE2020 and other members of the same genus, all *Psychrobacter* representative genomes were downloaded from the NCBI to build a phylogenetic tree on a core genome composed of 97 conserved markers (Supplementary Material). Despite being labelled as "Representative" on the RefSeq (O'Leary et al., 2016) database, some of those genomes were heavily fragmented. As Roary developers recommend against using highly fragmented scaffolds, the N50 statistic was used as a filter to select a final number of 30 genomes for which at least half of the nucleotides in the assembly were represented in contigs with length equal to or greater than the 1st Quartile (142680), (Supplementary Material). The phylogenetic tree was inferred through the Neighbor-Joining (Saitou and Nei, 1987) method on a multiple sequence alignment of 79,363 nucleotide positions from a core set of markers conserved across *Psychrobacter*, which was computed using Roary (Page et al., 2015). The percentage of (1000) replicate trees in which the associated taxa clustered together in the bootstrap test resulted in a 100% accordance throughout all the nodes. The obtained midpoint-rooted tree (Fig. 1D) shows that *Psychrobacter* sp. TAE2020 is a close relative of *Psychrobacter frigidicola* ACAM304 and *P. urativorans* R10.10B. Parallel to the phylogenetic analysis, we also sought to assess the Average Nucleotide Identity (ANI) between the newly sequenced strain and the 30 other selected genomes using FastANI (https://github.com/ParBLiSS/FastANI). The highest average genomic similarity 78.9% between *Psychrobacter* sp. TAE2020 and *P. frigidicola* ACAM304 (Supplementary Material). Mega 7 (Kumar et al., 2016) was used for inferring the evolutionary history through NJ while the processing of raw data was performed through the awk scripting language and the images rendered using the ggtree (Yu, 2020) package in R (https://www.r-project.org/).

Data accessibility

Biological source materials used in the herein presented genome announcement are available under the BioSample database accession JAHLMU000000000.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

Research supported by MIUR (Programma Nazionale di Ricerche in Antartide 2018, grant PNRA18_00335) and by Regione Campania (POR Campania FESR 2014/2020, Azione 1.5, grant 'GENOMAeSALUTE').

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.margen.2021.100922.

References

Blesic, M., Dichiarante, V., Milani, R., Linder, M., Metrangolo, P., 2018. Conference paper evaluating the potential of natural surfactants in the petroleum industry: the case of hydrophobins. Pure Appl. Chem. 90, 305–314.

Blin, K., Shaw, S., Kloosterman, A.M., Charlop-Powers, Z., van Wezel, G.P., Medema, M. H., Weber, T., 2021 May 12. antiSMASH 6.0: improving cluster detection and comparison capabilities. Nucleic Acids Res., gkab335 https://doi.org/10.1093/nar/gkab335. Epub ahead of print. PMID: 33978755.

Bolger, A.M., Lohse, M., Usadel, B., 2014 Aug 1. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics. 30 (15), 2114–2120. https://doi.org/10.1093/bioinformatics/btu170. Epub 2014 Apr 1. PMID: 24695404; PMCID: PMC4103590.

Bosi, E., Donati, B., Galardini, M., Brunetti, S., Sagot, M.F., Lió, P., Crescenzi, P., Fani, R., Fondi, M., 2015 Aug 1. MeDuSa: a multi-draft based scaffolder. Bioinformatics. 31 (15), 2443–2451. https://doi.org/10.1093/bioinformatics/btv171. Epub 2015 Mar 25. PMID: 25810435.

Dang, N.P., Landfald, B., Willassen, N.P., 2016. Biological surface-active compounds from marine bacteria. Environ. Technol. (United Kingdom) 37, 1151–1158. https://doi.org/10.1080/09593330.2015.1103784.

D'Errico, G., Ciccarelli, D., Ortona, O., 2005 Jun 15. Effect of glycerol on micelle formation by ionic and nonionic surfactants at 25 degrees C. J. Colloid Interface Sci. 286 (2), 747–754. https://doi.org/10.1016/j.jcis.2005.01.030 (PMID: 15897093).

Huerta-Cepas, J., Szklarczyk, D., Heller, D., Hernández-Plaza, A., Forslund, S.K., Cook, H., Mende, D.R., Letunic, I., Rattei, T., Jensen, L.J., von Mering, C., Bork, P., 2019 Jan 8. eggNOG 5.0: a hierarchical, functionally and phylogenetically annotated orthology resource based on 5090 organisms and 2502 viruses. Nucleic Acids Res. 47 (D1), D309–D314. https://doi.org/10.1093/nar/gky1085. PMID: 30418610; PMCID: PMC6324079.

Jain, D.K., Collins-Thompson, D.L., Lee, H., Trevors, J.T., 1991. A drop-collapsing test for screening biosurfactant-producing microorganisms. J. Microbiol. Methods 13 (4), 271–279. https://doi.org/10.1016/0167-7012(91)90064-W.

Kautsar, S.A., Blin, K., Shaw, S., Navarro-Muñoz, J.C., Terlouw, B.R., van der Hooft, J.J. J., van Santen, J.A., Tracanna, V., Suarez Duran, H.G., Pascal Andreu, V., Selem-Mojica, N., Alanjary, M., Robinson, S.L., Lund, G., Epstein, S.C., Sisto, A.C., Charkoudian, L.K., Collemare, J., Linington, R.G., Weber, T., Medema, M.H., 2020 Jan 8. MIBiG 2.0: a repository for biosynthetic gene clusters of known function. Nucleic Acids Res. 48 (D1), D454–D458. https://doi.org/10.1093/nar/gkz882. PMID: 31612915; PMCID: PMC7145714.

Kumar, S., Stecher, G., Tamura, K., 2016 Jul. MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. Mol. Biol. Evol. 33 (7), 1870–1874. https://doi.org/10.1093/molbev/msw054 (Epub 2016 Mar 22. PMID: 27004904).

Lee, I., Chalita, M., Ha, S.M., Na, S.I., Yoon, S.H., Chun, J., 2017 Jun. ContEst16S: an algorithm that identifies contaminated prokaryotic genomes using 16S RNA gene sequences. Int. J. Syst. Evol. Microbiol. 67 (6), 2053–2057. https://doi.org/10.1099/ijsem.0.001872. Epub 2017 Jun 22. PMID: 28639931.

Malavenda, R., Rizzo, C., Michaud, L., Gerçe, B., Bruni, V., Syldatk, C., Hausmann, R., Lo Giudice, A., 2015. Biosurfactant production by Arctic and Antarctic bacteria growing on hydrocarbons. Polar Biol. 38, 1565–1574. https://doi.org/10.1007/s00300-015-1717-9.

Mukherjee, S., Das, P., Sen, R., 2006. Towards commercial production of microbial surfactants. Trends Biotechnol. 24, 509–515. https://doi.org/10.1016/j.tibtech.2006.09.005.

O'Leary, N.A., Wright, M.W., Brister, J.R., Ciufo, S., Haddad, D., McVeigh, R., Rajput, B., Robbertse, B., Smith-White, B., Ako-Adjei, D., Astashyn, A., Badretdin, A., Bao, Y., Blinkova, O., Brover, V., Chetvernin, V., Choi, J., Cox, E., Ermolaeva, O., Farrell, C. M., Goldfarb, T., Gupta, T., Haft, D., Hatcher, E., Hlavina, W., Joardar, V.S., Kodali, V.K., Li, W., Maglott, D., Masterson, P., McGarvey, K.M., Murphy, M.R., O'Neill, K., Pujar, S., Rangwala, S.H., Rausch, D., Riddick, L.D., Schoch, C., Shkeda, A., Storz, S.S., Sun, H., Thibaud-Nissen, F., Tolstoy, I., Tully, R.E., Vatsan, A. R., Wallin, C., Webb, D., Wu, W., Landrum, M.J., Kimchi, A., Tatusova, T., DiCuccio, M., Kitts, P., Murphy, T.D., Pruitt, K.D., 2016 Jan 4. Reference sequence (RefSeq) database at NCBI: current status, taxonomic expansion, and functional annotation. Nucleic Acids Res. 44 (D1), D733–D745. https://doi.org/10.1093/nar/gkv1189. Epub 2015 Nov 8. PMID: 26553804; PMCID: PMC4702849.

Page, A.J., Cummins, C.A., Hunt, M., Wong, V.K., Reuter, S., Holden, M.T., Fookes, M., Falush, D., Keane, J.A., Parkhill, J., 2015 Nov 15. Roary: rapid large-scale prokaryote pan genome analysis. Bioinformatics. 31 (22), 3691–3693. https://doi.org/10.1093/bioinformatics/btv421. Epub 2015 Jul 20. PMID: 26198102; PMCID: PMC4817141.

Perfumo, A., Banat, I.M., Marchant, R., 2018. Going green and cold: biosurfactants from low-temperature environments to biotechnology applications. Trends Biotechnol. 36, 277–289. https://doi.org/10.1016/j.tibtech.2017.10.016.

Prjibelski, A., Antipov, D., Meleshko, D., Lapidus, A., Korobeynikov, A., 2020 Jun. Using SPAdes De Novo Assembler. Curr. Protoc. Bioinformatics 70 (1), e102. https://doi.org/10.1002/cpbi.102 (PMID: 32559359).

Saitou, N., Nei, M., 1987 Jul. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol. Biol. Evol. 4 (4), 406–425. https://doi.org/10.1093/oxfordjournals.molbev.a040454 (PMID: 3447015).

Seemann, T., 2014 Jul 15. Prokka: rapid prokaryotic genome annotation. Bioinformatics. 30 (14), 2068–2069. https://doi.org/10.1093/bioinformatics/btu153. Epub 2014 Mar 18. PMID: 24642063.

Tripathi, L., Irorere, V.U., Marchant, R., Banat, I.M., 2018. Marine derived biosurfactants: a vast potential future resource. Biotechnol. Lett. 40, 1441–1457. https://doi.org/10.1007/s10529-018-2602-8.

Trudgeon, B., Dieser, M., Balasubramanian, N., Messmer, M., Foreman, C.M., 2020. Low-temperature biosurfactants from polar microbes. Microorganisms 8, 1–14. https://doi.org/10.3390/microorganisms8081183.

Yu, G., 2020 Mar. Using ggtree to visualize data on tree-like structures. Curr. Protoc. Bioinformatics 69 (1), e96. https://doi.org/10.1002/cpbi.96 (PMID: 32162851).