Metagenome-Assembled Genome (MAG) of Oceanicaulis alexandrii NP7 isolated from Mediterranean Sea polluted marine sediments and its bioremediation potential

Filippo Dell’Anno¹, Leonardo Joaquim van Zyl², Marla Trindade², Christophe Brunet¹, Antonio Dell’Anno³, Adrianna Ianora¹, Clementina Sansone¹

¹Stazione Zoologica Anton Dohrn, Istituto Nazionale di Biologia, Ecologia e Biotecnologie marine, 80121 Naples, Italy
²Institute for Microbial Biotechnology and Metagenomics (IMBM), Department of Biotechnology, University of the Western Cape, 7535 Bellville, Cape Town, South Africa
³Department of Life and Environmental Science, Università Politecnica delle Marche, Via Brecce Bianche, 6031 Ancona, Italy

Corresponding author: filippo.dellanno@szn.it

Abstract

Oceanicaulis alexandrii strain NP7 is a marine bacterium which belongs to the Hyphomonadaceae family and was isolated from sediments highly contaminated with metals and polycyclic aromatic hydrocarbons released for decades by industrial activities in the Gulf of Naples (Mediterranean Sea). Here, we report the partial genome sequence and annotation of Oceanicaulis alexandrii strain NP7 that contains a chromosome of 2,954,327 bp and encodes for 2914 predicted coding sequences and 44 RNA-encoding genes. Although the presence of some coding sequences for genes involved in hydrocarbon degradation processes (e.g., alkB) have already been described in the literature associated with the Oceanicaulis, this is the first time that more than 100 genes involved in metal detoxification processes and hydrocarbon degradation are reported belonging to this genus. The presence of a heterogeneous set of genes involved in stress response, hydrocarbon degradation, heavy metal resistance and detoxification suggests a possible role for Oceanicaulis alexandrii NP7 in the bioremediation of these highly contaminated marine sediments.

© The Author(s) (2021). Published by Oxford University Press on behalf of the Genetics Society of America. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.
Keywords: Oceanicaulis alexandrii, Metagenome Assembled Genome, bioremediation, Mediterranean Sea

1. Introduction

Coastal sediments subjected to high anthropogenic pressure can accumulate large amounts of chemical contaminants (Tamburini et al. 2020) which can have detrimental consequences on ecosystem and human health (Ali et al. 2019; Rasheed et al. 2019). Thus, there is an urgent need to find sustainable and eco-compatible solutions able to efficiently recover and remediate contaminated marine sediments. Among these, eco-friendly strategies based on bioremediation are gaining increasing attention for their potential in the clean-up of contaminated marine coastal ecosystems (Catania et al. 2015; Brar et al. 2017).

Several microbes involved in the removal of pollutants have been identified and include bacteria belonging to the genera Alcaligens, Bacillus, Enterobacter, Flavobacterium, Pseudomonas (Ojuederie and Babalola 2017; Wahhab et al. 2021), Achromobacter, Acinetobacter, Alteromonas, Arthrobacter, Burkholderia (Xu et al. 2018) and obligate hydrocarbonoclastic bacteria such as Alcanivorax, Thallassolutus, Cycloclasticus, Oleispira (Yakimov et al. 2007). The development of high throughput sequencing technologies has provided new insights and information on new microorganisms with potential for the bioremediation of contaminated marine matrices (Czaplicki and Gunsch 2016).

Here we describe the metagenome-assembled genome (MAG) of Oceanicaulis alexandrii strain NP7 isolated from highly contaminated sediments of the Bagnoli-Coroglio bay in the Gulf of Naples (Southern Tyrrhenian Sea, Mediterranean Sea).

This site is a typical example of a coastal area chronically contaminated by different pollutants (i.e., metals, aliphatic and aromatic hydrocarbons and polychlorinated biphenyl) which have been
released for decades by industrial activities and stopped at the beginning of nineties (Romano et al. 2009, 2018). As annotation of the *Oceanicaulis alexandrii* strain NP 7 genome has highlighted the presence of numerous genes involved in hydrocarbon degradation and heavy metal detoxification, results are discussed in the framework of its bioremediation potential.

2. Material and Methods

2.1 Sample Collection and bacteria isolation

The superficial sediment samples (0-20 cm) used for the isolation of bacteria were collected in November 2017 with a Van Veen grab in the Bagnoli-Coroglio area (Gulf of Naples): 40°48’29.0”N 14°09’54.2”E. Duplicate samples were immediately placed into sterile bags (Whirl-Pak, Nasco) and stored at 4 °C in the dark, until their processing in the laboratory.

The sediment was plated on Marine Agar (MA) (Bacto-Agar, Difco) and incubated at 28°C for 48 hours.

2.2 DNA extraction, Illumina sequencing and gene annotation

For genome sequencing, DNA was extracted from bacterial consortia, following sequential dilution on agar plates, with the DNeasy Blood & Tissue kit (Qiagen), according to the manufacturer's instructions. DNA concentration was determined using a Qubit fluorometer (Thermo-Fisher).

Sequence library preparation was performed using the Nextera DNA Flex kit (Illumina, Hayward, USA) with 1 ng input DNA according to the manufacturer's instructions. The resulting libraries were sequenced on the Illumina MiSeq platform at the University of Western Cape (South Africa) sequencing facility using a MiSeq Reagent kit V2 (500 cycle) with a 10% phiX v3 spike generating 2 × 250 bp reads per sample. Metagenome assembly was performed using CLC Genomics Workbench version 6.5. The raw reads were trimmed and demultiplexed, and ≤500 bp contigs were removed from the final assembly. Binning of metagenomic contigs was performed using
MyCC (Lin and Liao 2016), while completeness and contamination of metagenome-assembled genome (MAG) as well as genome quality were determined using CheckM using the lineage-specific workflow and default parameters (Parks et al. 2015). The CGView (Circular Genome Viewer) software was used to obtain a circularized map of the chromosome, using a default server setting (Stothard and Wishart 2005). Gene prediction and annotation were performed by Rapid Annotation using the Subsystem Technology (RAST) (http://rast.nmpdr.org/) (Overbeek et al. 2014) and MicroScope pipelines (Vallenet et al. 2017).

2.3 Taxonomic analysis

The Genome Taxonomy Database (https://gtdb.ecogenomic.org/) implemented through K-Base (www.kbase.us) was used to perform the classification.

3) Results and discussion

Metagenome sequencing of the bacterial consortium cultured on marine agar, followed by binning of the metagenome contigs highlighted the dominance of MAGs representing three different bacterial species: Halomonas sp. SZN1, Epibacterium sp. SZN4 and Oceanicaulis alexandrii NP7. Since the phylogenetic classification and the ability to degrade hydrocarbons and reduce the toxicity of metals for Halomonas sp. SZN1, and Epibacterium sp. SZN4 has already been described in Dell’ Anno et al. (2020), the focus of this work was on the description of the Oceanicaulis alexandrii NP7 MAG (Table 1) and its bioremediation potential.

Furthermore, we have chosen to thoroughly analyze the genome of Oceanicaulis alexandrii NP7 as little information is available in the literature on the biotechnological potential of members belonging to the Hyphomonadaceae family, and to our knowledge, no member of Oceanicaulis alexandrii has ever been used for bioremediation purposes.

The draft genome of Oceanicaulis alexandrii NP7, an Alphaproteoacteria belonging to the Order Caulobacterales, contains 2,954,327 bases (Figure 1).
The closest reference strain on the genomic taxa database, based on an Average Nucleotide Identity (ANI) score of 97.5% is Oceanicaulis alexandrii DSM 11625, and on this basis we have assigned the name Oceanicaulis alexandrii NP7. CheckM analysis showed a completeness of 96.5% (27 markers missing) and a contamination of 0.32% (1 marker duplicated) (Table 2).

The draft genome sequence for Oceanicaulis alexandrii NP7 has a GC content of 62.71%, and is composed of 40 contigs and 2914 predicted coding sequences with an average length of 937.58 bp having a protein coding density of 91.51%. Of the total predicted CDSs, 1995 (69.4%) were assigned to a function, 842 (29.3%) were classified as hypothetical and 44 (1.5%) as coding for RNAs (Table 2). Seven Genomic Islands (GIs) were predicted using the integration of IslandPath-DIMOB, SIGI-HMM and Island Pick provided IslandViewer4 (Bertelli et al. 2017), with a total of 129,287 bp (4.39% of the genome) and 151 predicted CDSs of which 42 were classified as proteins of unknown function (Table S1). Predicted CDSs encode for regulators, transmembrane proteins, multi-drug resistance genes, exopolysaccharide production proteins and oxidoreductases. Among these GIs, 12 mobile genetic elements, such as integrase/recombinase, and transposase genes were found, suggesting that these GIs can self-mobilize and promote horizontal gene transfer. To further explore the ability of Oceanicaulis alexandrii NP7 to survive in highly contaminated environments, we annotated and analyzed the gene functional categories (Table S2). Thirty-seven genes related to “Virulence, Disease and Defense” functions, thus likely involved in playing a role in the resistance to antibiotics and toxic compounds. Thirty-one genes belonged to the subcategory of resistance to antibiotics and toxic compounds. The RAST annotation system highlighted the presence of 5 genes involved in antibiotic resistance such as Beta-lactamase (3 genes) and resistance to fluoroquinolones (2 genes). A prevalence of genes involved in heavy metal detoxification were identified, such as copper homeostasis and tolerance (8 genes); cobalt-, zinc-, and cadmium- resistance (7 genes); mercuric reductase (1 gene); a mercuric resistance protein (1
gene) and multidrug resistance efflux pumps (6 genes). The presence of genes for heavy metal resistance suggests that Oceanicaulis alexandrii NP7 may have evolved to cope with highly metal contaminated environments. This observation is corroborated by the presence of 33 genes coding for superoxide dismutase and glutathione related pathways, known for their antioxidant and detoxification capacity (Espinosa-Diez et al. 2015). Since the production of exopolysaccharides favours the immobilization of metals through the formation of metal-exopolysaccharide complexes (Gupta and Diwan 2017), the presence of 10 genes identified in the genome of Oceanicaulis alexandrii NP7 - 5 coding for dTDP rhamnose synthesis and 5 coding for rhamnose containing glycans – suggests that this species may have a potential role in the bioremediation of a heavy metal-contaminated matrix.

The annotations generated by the RAST and MicroScope pipelines revealed the presence of numerous genes involved in the degradation of hydrocarbons and polycyclic aromatic compounds, e.g., genes involved in the degradation of the following hydrocarbons: chlorocyclohexane (2 genes), benzoate (7 genes), fluorobenzoate (1 gene), toluene (2 genes), chloroalkane and chloroalkene (2 genes), naphtalene (2 genes), aminobenzoate (4 genes) and ethylbenzene (1 gene). Furthermore, 14 genes coding for pathways involved in the metabolism of central aromatic intermediates such as the catechol branch of the beta ketoadipate pathway and salicylate, gentisate and homogentisate pathways, were identified.

The possibility of using Oceanicaulis alexandrii in bioremediation processes has previously been suggested by Oh et al. (2011), on the basis of the identification of the epoxide hydrolase, cyclohexanone monooxygenase, 6-hexanolactone hydrolase, and phytase genes. Here we expand the identification to include 100 genes in the Oceanicaulis alexandrii NP7 genome proposed to be involved in the degradation of hydrocarbons, and in mechanisms involved in the detoxification of metals. This further emphasizes the potential ability of this bacterial strain to efficiently degrade
organic contaminants and/or to be highly tolerant/resistant to metal contamination, and thus useful for bioremediation purposes.

Data availability

The complete genome sequence of *Oceanicaulis alexandrii* NP7 has been deposited in the GenBank database under number JADWML000000000.1 and bioproject number PRJNA669418. Sequence read archives are available at SRR13577792.

Author contributions

C.S. and F.D. carried out the in situ samplings and the laboratory experiments. F.D, L.J.v.Z., M.T., performed the molecular/bioinformatics analysis and analysed the results. F.D., C.B., L.J.v.Z., M.T., A.I., C.S. and A.D., interpreted the results. F.D. drafted the manuscript, all authors revised and approved the final manuscript.

Funding and Competing Interests

Dell’Anno F. was funded by a PhD grant from the SZN and UNIVPM. This study was supported by the projects ABBaCo funded by the Italian Ministry for Education, University and Research (grant number C62F16000170001), Ocean Medicines (H2020-MSCA-RISE-2015) and MERCES (H2020-SC5-2015, grant number 689518).

Conflict of interest statement

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.

References

Ali H, Khan E, Ilahi I (2019) Environmental chemistry and ecotoxicology of hazardous heavy metals:
Environmental persistence, toxicity, and bioaccumulation. J Chem 2019:.

https://doi.org/10.1155/2019/6730305

Bertelli C, Laird MR, Williams KP, et al (2017) IslandViewer 4: Expanded prediction of genomic islands for larger-scale datasets. Nucleic Acids Res 45:W30–W35.

https://doi.org/10.1093/nar/gkx343

Brar A, Kumar M, Vivekanand V, Pareek N (2017) Photoautotrophic microorganisms and bioremediation of industrial effluents: current status and future prospects. 3 Biotech 7:1–8.

https://doi.org/10.1007/s13205-017-0600-5

Catania V, Santisi S, Signa G, et al (2015) Intrinsic bioremediation potential of a chronically polluted marine coastal area. Mar Pollut Bull 99:138–149.

https://doi.org/10.1016/j.marpolbul.2015.07.042

Czaplicki LM, Gunsch CK (2016) Reflection on Molecular Approaches Influencing State-of-the-Art Bioremediation Design: Culturing to Microbial Community Fingerprinting to Omics. J Environ Eng (United States) 142:1–13. https://doi.org/10.1061/(ASCE)EE.1943-7870.0001141

Espinosa-Diez C, Miguel V, Mennerich D, et al (2015) Antioxidant responses and cellular adjustments to oxidative stress. Redox Biol 6:183–197.

https://doi.org/10.1016/j.redox.2015.07.008

Gupta P, Diwan B (2017) Bacterial Exopolysaccharide mediated heavy metal removal: A Review on biosynthesis, mechanism and remediation strategies. Biotechnol Reports 13:58–71.

https://doi.org/10.1016/j.btre.2016.12.006

Lin HH, Liao YC (2016) Accurate binning of metagenomic contigs via automated clustering sequences using information of genomic signatures and marker genes. Sci Rep 6:12–19.
Oh HM, Kang I, Vergin KL, et al (2011) Genome sequence of Oceanicaulis sp. strain HTCC2633, isolated from the western Sargasso Sea. J Bacteriol 193:317–318. https://doi.org/10.1128/JB.01267-10

Ojuederie OB, Babalola OO (2017) Microbial and plant-assisted bioremediation of heavy metal polluted environments: A review. Int. J. Environ. Res. Public Health 14

Overbeek R, Olson R, Pusch GD, et al (2014) The SEED and the Rapid Annotation of microbial genomes using Subsystems Technology (RAST). Nucleic Acids Res 42:206–214. https://doi.org/10.1093/nar/gkt1226

Parks DH, Imelfort M, Skennerton CT, et al (2015) CheckM: Assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. Genome Res 25:1043–1055. https://doi.org/10.1101/gr.186072.114

Rasheed T, Bilal M, Nabeel F, et al (2019) Environmentally-related contaminants of high concern: Potential sources and analytical modalities for detection, quantification, and treatment. Environ Int 122:52–66. https://doi.org/10.1016/j.envint.2018.11.038

Romano E, Bergamin L, Ausili A, et al (2009) The impact of the Bagnoli industrial site (Naples, Italy) on sea-bottom environment. Chemical and textural features of sediments and the related response of benthic foraminifera. Mar Pollut Bull 59:245–256. https://doi.org/10.1016/j.marpolbul.2009.09.017

Romano E, Bergamin L, Celia Magno M, et al (2018) Temporal changes of metal and trace element contamination in marine sediments due to a steel plant: The case study of Bagnoli (Naples, Italy). Appl Geochemistry 88:85–94. https://doi.org/10.1016/j.apgeochem.2017.05.012
Stothard P, Wishart D (2005) Circular genome visualization and exploration using CGView. Bioinformatics 21:537–539.

Tamburini E, Doni L, Lussu R, et al (2020) Impacts of Anthropogenic Pollutants on Benthic Prokaryotic Communities in Mediterranean Touristic Ports. Front Microbiol 11:1–16. https://doi.org/10.3389/fmicb.2020.01234

Vallenet D, Calteau A, Cruveiller S, et al (2017) MicroScope in 2017: An expanding and evolving integrated resource for community expertise of microbial genomes. Nucleic Acids Res 45:D517–D528. https://doi.org/10.1093/nar/gkw1101

Wahhab BHA, Samsulrizal NH, Edbeib MF, et al (2021) Genomic analysis of a functional haloacid-degrading gene of Bacillus megaterium strain BHS1 isolated from Blue Lake (Mavi Gölü, Turkey). Ann Microbiol 71:12. https://doi.org/10.1186/s13213-021-01625-9

Xu X, Liu W, Tian S, et al (2018) Petroleum Hydrocarbon-Degrading Bacteria for the Remediation of Oil Pollution Under Aerobic Conditions: A Perspective Analysis. Front Microbiol 9: https://doi.org/10.3389/fmicb.2018.02885

Yakimov MM, Timmis KN, Golyshin PN (2007) Obligate oil-degrading marine bacteria. Curr Opin Biotechnol 18:257–266. https://doi.org/10.1016/j.copbio.2007.04.006

Figure Legend:

Figure 1. Circular representation of the Oceanicaulis alexandrii NP7 genome. The different rings represent (from outer to inner) predicted protein-coding sequences (CDS) on the forward (outer wheel) and the reverse (inner wheel) strands (circle 2 and 3) colored according to the assigned COG classes (circle 1, 4), G+C content (circle 5), GC skew (circle 6), genomic position (circle 7). The COG colors represent functional groups (A, RNA processing and modification; B, chromatin structure and dynamics; J, Translation, ribosomal structure and biogenesis; K, Transcription; L, Replication, recombination and repair; D, Cell cycle control,
cell division, chromosome partitioning; O, Posttranslational modification, protein turnover, chaperones; M, Cell wall/membrane/envelope biogenesis; N, Cell motility; P, Inorganic ion transport and metabolism; T, Signal transduction mechanisms; U, Intracellular trafficking, secretion, and vesicular transport; V, Defense mechanisms; W, Extracellular structures; Y, Nuclear structure; Z, Cytoskeleton; C, Energy production and conversion; G, Carbohydrate transport and metabolism; E, Amino acid transport and metabolism; F, Nucleotide transport and metabolism; H, Coenzyme transport and metabolism; I, Lipid transport and metabolism; Q, Secondary metabolites biosynthesis, transport and catabolism; R, General function prediction only; S, Function unknown).

**Tables:**

| Item                        | Description                                                          |
|-----------------------------|----------------------------------------------------------------------|
| Classification              | Domain Bacteria                                                     |
|                             | Phylum Proteobacteria                                               |
|                             | Class Alphaproteobacteria                                           |
|                             | Order Caulobacterales                                               |
|                             | Family Maricaulaceae                                                |
|                             | Genus Oceanicaulis                                                  |
|                             | Species Oceanicaulis alexandri                                      |
| Investigation type          | Bacteria                                                            |
| Project name                | Oceanicaulis alexandri strain:NP7                                   |
| Collection date             | 10-05-2018                                                          |
| Geographic location         | 40°48'28.6"N 14°10'03.0"E                                          |
| (latitude and location)     | Gulf of Naples, Mediterranean Sea, Italy                            |
| Geographic location         | Sediment environmental package                                      |
| (country and/or sea, region)| Sediment                                                            |
| Environment (biome)         | Sea                                                                 |
| Environment (feature)       | sediments                                                           |
| Environment (material)      | Sea sediments                                                        |
| Depth                       | -4 m                                                                |
Table 1. Oceanicaulis alexandrii NP7: general features and MIGS mandatory information.

| Genome features                  | %     |
|----------------------------------|-------|
| CheckM completeness              | 96.5% |
| checkM Contamination             | 0.3%  |
| Size, bp                         | 2,954,327 |
| G+C content, %                   | 62.71% |
| N50                              | 209292 |
| L50                              | 2     |
| Number of contigs                | 40    |
| Number of Subsystems             | 383   |
| Number of coding sequences       | 2837  |
| Function assigned                | 1995  |
| Hypothetical                     | 842   |
| Number of RNAs                   | 44    |

Table 2. Genome statistics of Oceanicaulis alexandrii NP7
