

# Whole genome sequence of *Paracoccus* sp. NFXS7, a carotenoid-producing bacterium isolated from a marine saltern

Constança D.F. Bertrand<sup>1,2</sup>, Rodrigo Martins<sup>1,2</sup>, Francisco Quintas-Nunes<sup>1,2</sup>, Pedro Reynolds-Brandão<sup>1,2</sup>, Maria T. Barreto Crespo<sup>1,2</sup> and Francisco X. Nascimento<sup>1,2,\*</sup>

## Abstract

The study presents the whole genome sequence of the carotenoid-producing *Paracoccus* sp. NFXS7, isolated from a marine saltern in Setúbal, Portugal. The carotenoid-producing strain NFXS7 contains homologs of the *crt* genes involved in astaxanthin biosynthesis, making it a promising candidate for biotechnological applications.

## DATA SUMMARY

This Whole Genome Shotgun project has been deposited in the National Center for Biotechnology Information (NCBI) under the accession number JAKEKV000000000.1. The whole genome sequence reads used can be accessed in the NCBI's Sequence Read Archive under the accession number PRJNA795090 (<https://www.ncbi.nlm.nih.gov/sra/PRJNA795090>).

The authors confirm all supporting data, code and protocols have been provided within the article or through supplementary data files.

## INTRODUCTION

Marine environments are rich in biodiversity and serve as a valuable resource for the discovery of novel microbial strains of industrial relevance. Among these, *Paracoccus* has gained significant attention due to its ability to produce carotenoids [1, 2], a class of pigments with potential applications in various industries, including food, pharmaceuticals and cosmetics [3]. Several species of *Paracoccus* (*Paracoccus carotinifaciens*, the most studied) produce high yields of carotenoids such as astaxanthin, which makes them suitable for industrial applications [4–6]. Still, despite the known biotechnological potential of *Paracoccus*, not much is understood about the genomic properties of these strains, the genetic mechanisms regulating carotenoid biosynthesis, and the synthesis of other relevant secondary metabolites.

In this study, we report the genome sequence of a novel carotenoid-producing *Paracoccus* strain (NFXS7) isolated from marine saltern water in Setúbal, Portugal (38°29'35.0" N 8°46'26.4" W), bringing new insights into the genomic properties of carotenoid-producing *Paracoccus*.

## METHODS

The strain was isolated using marine agar (Condalab, Spain) and was selected due to its deep orange coloration. Strain NFXS7 was routinely cultivated in marine agar and marine broth at 26 °C (at 180 r.p.m. when cultivated in liquid broth). Total DNA was isolated from an overnight grown culture using the PureLink Genomic DNA Kit (Invitrogen, USA) according to the manufacturer's instructions. The bacterial DNA was sent to MicrobesNG (<https://microbesng.com/>) (UK) and processed under the company-defined pipeline (<https://microbesng.com/documents/methods/>). The genomic DNA libraries were prepared using the Nextera XT Library

*Access Microbiology* is an open research platform. Pre-prints, peer review reports, and editorial decisions can be found with the online version of this article. Received 13 June 2024; Accepted 11 July 2024; Published 23 July 2024

**Author affiliations:** <sup>1</sup>IBET, Instituto de Biologia Experimental e Tecnológica, Apartado 12, 2781-901 Oeiras, Portugal; <sup>2</sup>ITQB-NOVA, Instituto de Tecnologia Química e Biológica António Xavier, Universidade Nova de Lisboa, Av. da República, 2780-157 Oeiras, Portugal.

**\*Correspondence:** Francisco X. Nascimento, [francisco.nascimento@ibet.pt](mailto:francisco.nascimento@ibet.pt)

**Keywords:** carotenoid; genome; marine; *Paracoccus*.

**Abbreviation:** NCBI, National Center for Biotechnology Information.

SRA: <https://www.ncbi.nlm.nih.gov/sra/PRJNA795090>. WGS: <https://www.ncbi.nlm.nih.gov/nucleotide/JAKEKV000000000.1>

000862.v2 © 2024 The Authors



This is an open-access article distributed under the terms of the Creative Commons Attribution License.

Prep Kit (Illumina, San Diego, USA) following the manufacturer's protocol with the following modifications: input DNA was increased twofold, and PCR elongation time was increased to 45 s. The DNA quantification and library preparation were conducted using a Hamilton Microlab STAR automated liquid handling system (Hamilton Bonaduz AG, Switzerland). The libraries were sequenced using Illumina NovaSeq 6000 (Illumina, San Diego, USA) using a 250-bp paired-end protocol. The obtained reads were trimmed using Trimmomatic version 0.30 [7] with a sliding window quality cutoff of Q15. A total of 1 682 720 reads with a median insert size of 642 bp were obtained. The genome sequence of strain NFXS7 was assembled using the SPAdes v 3.15.5 [8] and annotated using the National Center for Biotechnology Information Prokaryotic Annotation Pipeline [9]. The functional annotation of the genome was performed using BlastKOALA [10]. Secondary metabolite clusters were predicted using the antiSMASH version 7.1.0 [11].

## GENOME DESCRIPTION AND FUTURE OUTLOOK

The whole genome sequence assembly of *Paracoccus* sp. NFXS7 comprises 35 contigs (N50 = 296 487) totaling a genome size of 3 845 849 bp (3.845 Mbp) and a GC of 67%. The total genome coverage was 103X. A total of 3767 genes were identified, of which 3626 are complete coding sequences and 50 are encoding RNA-related genes. The BlastKOALA functional annotation resulted in 2097 entries (57.8%), mostly involved in genetic information processing (408), environmental information processing (278), carbohydrate metabolism (232) and signalling and cellular processes (201), followed by the metabolism of amino acids (127), cofactors and vitamins (113) and nucleotides (84). The antiSMASH annotation revealed the presence of gene clusters involved in the biosynthesis of ectoine, a siderophore, a carotenoid, an unknown Non-ribosomal peptide synthase and *N*-acyl-*L*-homoserine lactones. Additionally, analysis of the NFXS7 carotenoid biosynthetic cluster (*di*, *crtWZYIBEX*) showed that it presented a high similarity (97.39%) to the functional carotenoid biosynthetic cluster of *Paracoccus* sp. N81106 (GenBank accession: AB206672.1) is involved in the biosynthesis of astaxanthin and its intermediates [12].

The whole genome sequence assembly of *Paracoccus* sp. NFXS7 provides a valuable resource for future research aiming to understand *Paracoccus* physiology, secondary metabolite biosynthesis and stress resistance mechanisms, which are the key factors for the development of novel potential biotechnological applications.

### Funding information

The research was conducted in the scope of the project 'PhycouBiome: Understanding and harnessing the power of the microalgae microbiome aiming the maximization of marine microalgae productivity' funded by Fundação para a Ciência e Tecnologia/Ministério da Ciência, Tecnologia e Ensino Superior (FCT/MCTES, Portugal), grant number PTDC/BAA-BIO/1262/2020 (<http://doi.org/10.54499/PTDC/BAA-BIO/1262/2020>). The research was performed with the support of the R and D Unit 'GREEN-IT-Bioresources for Sustainability' (UIDB/04551/2020, DOI: 10.54499/UIDB/04551/2020 and UIDP/04551/2020, DOI: 10.54499/UIDP/04551/2020), iNOVA4Health (UIDB/04462/2020, and UIDP/04462/2020) and LS4FUTURE Associated Laboratory (LA/P/0087/2020, DOI: 10.54499/LA/P/0087/2020) funded by FCT/MCTES. C.B., R.M., F.Q.N. and P.R.B. acknowledge receiving a Ph.D. fellowship (2023.02799.BD, 2023.02007.BD, 2022.10633.BD and 2021.07927.BD514, respectively) funded by FCT/MCTES.

### Author contributions

C.D.F.B.: investigation, formal analysis and writing - original draft; R.M.: investigation, formal analysis and writing - original draft; F.Q.N.: investigation and formal analysis; P.R.B.: investigation and formal analysis; M.T.B.C.: funding acquisition and writing - review and editing; F.X.N.: conceptualization, resources, supervision, project administration, funding acquisition, and writing - review and editing.

### Conflicts of interest

The authors declare that there are no conflicts of interest.

### References

- Hayashi M, Ishibashi T, Kuwahara D, Hirasawa K. Commercial production of astaxanthin with *Paracoccus carotinifaciens*. *Adv Exp Med Biol* 2021;1261:11–20.
- Mussagy CU, Dufossé L. A review of natural astaxanthin production in a circular bioeconomy context using *Paracoccus carotinifaciens*. *Bioresour Technol* 2023;369:128499. Epub ahead of print 1 February 2023.
- Barreiro C, Barredo JL. Carotenoids production: a healthy and profitable industry. *Methods Mol Biol* 2018;1852:45–55.
- Mussagy CU, Pereira JFB, Dufossé L. Astaxanthin production using *Paracoccus carotinifaciens*: a way forward? *Trends Biotechnol* 2023;41:996–999.
- Khomlaem C, Aloui H, Oh WG, Kim BS. High cell density culture of *Paracoccus* sp. LL1 in membrane bioreactor for enhanced co-production of polyhydroxyalkanoates and astaxanthin. *Int J Biol Macromol* 2021;192:289–297.
- Lee JH, Kim YT. Cloning and characterization of the astaxanthin biosynthesis gene cluster from the marine bacterium *Paracoccus haundaensis*. *Gene* 2006;370:86–95.
- Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 2014;30:2114–2120.
- Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, et al. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol* 2012;19:455–477.
- Angiuoli SV, Gussman A, Klimke W, Cochrane G, Field D, et al. Toward an online repository of Standard Operating Procedures (SOPs) for (Meta)genomic annotation. *Omi A J Integr Biol* 2008;12:137–141.
- Kanehisa M, Sato Y, Morishima K. BlastKOALA and GhostKOALA: KEGG tools for functional characterization of genome and metagenome sequences. *J Mol Biol* 2016;428:726–731.
- Blin K, Shaw S, Augustijn HE, Reitz ZL, Biermann F, et al. antiSMASH 7.0: new and improved predictions for detection, regulation, chemical structures and visualisation. *Nucleic Acids Res* 2023;51:W46–W50.
- Misawa N, Satomi Y, Kondo K, Yokoyama A, Kajiwara S, et al. Structure and functional analysis of a marine bacterial carotenoid biosynthesis gene cluster and astaxanthin biosynthetic pathway proposed at the gene level. *J Bacteriol* 1995;177:6575–6584.