Ref. Ares(2024)6633096 - 19/09/2024

PROMICON

Project Collection of OMICS data in the Research Ideas and Outcomes journal

Deliverable D6.4

19 September 2024

Adriana Dechina, Nikol Yovcheva, Pavel Stoev

Pensoft Publishers

Prepared under contract from the European Commission

PROMICON

Harnessing the power of nature through PROductive MIcrobial CONsortia in biotechnology - measure, model, master

This project receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101000733.

Grant agreement No. 101000733
EU Horizon 2020 Research and Innovation Action

Project acronym:	**PROMICON**
Project full title:	**Harnessing the power of nature through PROductive MIcrobial CONsortia in biotechnology - measure, model, master**
Start of the project:	June 2021
Duration:	48 months
Project coordinator:	Dr. rer. nat. Jens O. Krömer Helmholtz Centre for Environmental Research - UFZ https://www.ufz.de/
Deliverable title:	Project collection of OMICS data in the Research Ideas and Outcomes Journal
Deliverable n°:	D6.4
Nature of the deliverable:	Report
Dissemination level:	Public
WP responsible:	WP6
Lead beneficiary:	PENSOFT
Citation:	Dechina, A., Yovcheva, N. & Stoev, P. (2024). *Project collection of OMICS data in the Research Ideas and Outcomes journal*. Deliverable D6.4 EU Horizon 2020 PROMICON Project, Grant agreement No 101000733.
Due date of deliverable:	Month 40
Actual submission date:	Month 40

Deliverable status:

Version	Status	Date	Author(s)
1.0	Draft	16 August 2024	Adriana Dechina PENSOFT
1.1	Internal review	19 August 2024	Nikol Yovcheva, Pavel Stoev, PENSOFT
1.2	Final draft	28 August 2024	Adriana Dechina PENSOFT
1.3	Coordinator review	19 September 2024	Minmin Pan, Jens Krömer UFZ
1.4	Final version	19 September 2024	Adriana Dechina PENSOFT

The content of this deliverable does not necessarily reflect the official opinions of the European Commission or other institutions of the European Union.

Table of contents

Preface

This report provides a comprehensive overview of the publication of PROMICON's OMICS data, detailing the preparation process and publication options available to partners. It outlines the steps for making datasets accessible, including guidance on how to publish data papers in scholarly journals and thus openly share datasets. This report also outlines the strategic shift in the dissemination of PROMICON's OMICS data from the Biodiversity Data Journal to the Research Ideas and Outcomes journal, aiming to enhance the visibility and accessibility of PROMICON's research outputs. It serves as a practical guide for partners, ensuring a smooth transition and effective sharing of the project's valuable data.

Summary

An Open Science collection of OMICS data will be launched in the Research Ideas and Outcomes journal with the aim of enhancing transparency and ensuring widespread dissemination of PROMICON’s research data. This collection will consolidate project results and ensure compliance with the findable, accessible, interoperable, and reusable data principles.

Initially, the project planned to publish its data in the Biodiversity Data Journal. However, after a comprehensive re-evaluation and careful consideration of the data's interdisciplinary nature, the decision was made to transition to the Research Ideas and Outcomes journal. This shift reflects PROMICON’s commitment to making its OMICS data freely available in a platform better suited for a wider range of scientific outputs, thus maximising the impact and visibility of the research.

This document details the steps taken to establish the new collection in the Research Ideas and Outcomes journal, including guidance on how to publish OMICS data papers and deposit datasets. It also outlines the strategic planning involved in this transition, ensuring that partners are well-prepared to contribute effectively. Two partners are already in the process of preparing their manuscripts for submission, while others are selecting their preferred methods for data sharing.

By ensuring that PROMICON’s data is widely disseminated and fully compliant with findable, accessible, interoperable, and reusable data principles, this approach promotes greater scientific collaboration and discovery, reinforcing the project's dedication to best practices in open data management.

List of abbreviations

BDJ	Biodiversity Data Journal
DOI	Digital Object Identifier
EU	European Union
FAIR	Findable, Accessible, Interoperable, Reusable
PO	Project Officer
RIO	Research Ideas and Outcomes

1 Introduction

To promote transparency and broad dissemination of the data generated within PROMICON, an Open Science collection of OMICS data will be established in the Research Ideas and Outcomes (RIO) journal. This collection will serve as a central location for the results produced within the project duration, ensuring that all project data adheres to the FAIR (Findable, Accessible, Interoperable, Reusable) data principles.

As PROMICON advances in its research, the project is anticipated to generate OMICS data. Initially, plans were in place to publish this data in the Biodiversity Data Journal (BDJ). However, after a thorough re-evaluation and detailed discussions – particularly as the nature and scope of the data became clearer – it was concluded that the RIO journal would better align with the project's objectives. While BDJ is highly respected for its focus on data publication, its specialisation in biodiversity posed constraints for the broader, interdisciplinary data that PROMICON will produce. In contrast, the RIO journal offers a more appropriate platform for disseminating this diverse OMICS data, thereby enhancing the reach and impact of the project's outcomes.

In alignment with its commitment, PROMICON has promised to provide open access to its OMICS data, reinforcing the project's dedication to making scientific data publicly available. Moreover, the collection will emphasise the importance of providing access to this data, underlining the critical role that open and FAIR data practices play in advancing scientific research and collaboration.

2 PROMICON OMICS data collection

The initial plan for promoting the open dissemination of PROMICON-generated data was outlined in ***Task 6.3 "Dissemination and communication activities"***. The strategy involved creating a dedicated collection of OMICS data papers in the BDJ to compile the project results. However, after evaluating the scope and publication model of the journal, several limitations were identified that could potentially hinder the effective dissemination of PROMICON's diverse research outputs. Following discussions with the project coordinators, it was determined that a deviation from the original plan was necessary. Rather than establishing a new collection in BDJ solely for OMICS data, the decision was made to utilise PROMICON's existing open-access project collection in the RIO journal.

2.1 Biodiversity Data Journal

The Biodiversity Data Journal is an open-access, peer-reviewed journal that focuses on the publication of biodiversity-related data and research. BDJ is designed to facilitate the publication of high-quality datasets, with a particular emphasis on biodiversity records, taxonomic studies, and ecological data. The journal's platform supports a broad range of data types and encourages the use of innovative formats to document and share data in a standardised, accessible manner. However, BDJ's scope is specifically aligned with biodiversity research, which can limit its applicability for projects with broader or more interdisciplinary data, such as PROMICON.

2.2 Research Ideas and Outcomes journal

Research Ideas and Outcomes is an open-access journal that offers a flexible and inclusive platform for publishing a wide variety of research outputs, ranging from traditional research articles and data papers to more unconventional formats like project reports, workshop summaries, and policy briefs. RIO is designed to support the full research lifecycle, enabling researchers to share not only their final results but also ideas, methodologies, and interim

findings. With no thematic constraints, RIO is well-suited for interdisciplinary projects like PROMICON, allowing for the comprehensive representation and open dissemination of all types of research outputs generated by the project. Moreover, the project already has a RIO collection with its own digital object identifier (DOI) (10.3897/rio.coll.239), featuring project papers and deliverables, which stakeholders will be able to access along with the OMICS data all in one place.

2.3 Deviation summary

Considering the specific focus of the BDJ on biodiversity research and its thematic constraints, alongside the flexibility and broad scope offered by RIO, it became evident that RIO would better serve the diverse and interdisciplinary nature of PROMICON's outputs. With these factors in mind, the decision to deviate from the original plan of publishing in BDJ and instead utilise RIO was reached, ensuring a more effective and comprehensive dissemination of the project's research. This alternative approach offers several significant advantages:

- **Comprehensive representation:** RIO provides a platform that supports not only traditional research outputs, such as research data and data papers, but also non-conventional outputs, including workshop reports, policy briefs, and project deliverables. This comprehensive representation of PROMICON's research is not possible within the thematic constraints of BDJ.
- **Flexibility:** Unlike BDJ, which imposes specific thematic restrictions, RIO allows for the inclusion of various types of PROMICON data within the same collection. This flexibility is particularly beneficial given the diverse range of data produced by the project.
- **Multiple publication options:** The RIO platform accommodates different methods of data publication. Researchers can publish descriptive data papers directly within the RIO collection, submit data papers to other journals and link them to the RIO collection via metadata, or upload datasets to relevant repositories and link them to the RIO collection. This approach ensures that all project data is accessible and properly documented in one place, while also allowing researchers to choose the most suitable publication format for their specific needs.

By shifting the focus to RIO, PROMICON can ensure that all generated data is disseminated effectively and in alignment with the project's broader open science objectives. This strategic adjustment not only addresses the shortcomings of the initial plan but also enhances the overall impact and accessibility of PROMICON's research outputs.

2.4 Deviation process

Following this decision, in June 2024, the project coordinator communicated with the project officer (PO) to explain the rationale behind the proposed deviation and its potential benefits for the project's dissemination strategy. The PO reviewed the justification for the change and recognised the advantages that RIO's flexible and comprehensive platform would offer for the broader dissemination of PROMICON's research outputs. Consequently, approval was granted to proceed with this deviation and again confirmed in September 2024 with plans to formally make amendments reflecting this shift. This decision ensures that all project data will be published in a manner that maximises visibility, accessibility, and impact, fully supporting PROMICON's commitment to open science and effective knowledge sharing.

3 Preparation for PROMICON OMICS data collection

To ensure a smooth and efficient publication process, PENSOFT, as the leader of WP6, undertook the responsibility of creating a comprehensive guide. This guide is designed to

provide clear instructions on how to make data available, ensuring that all project partners are well-equipped to publish their datasets effectively.

The guide covers several important aspects, including the benefits of publishing OMICS data papers. It highlights how this publishing format enhances data visibility, credibility, and potential for collaboration, while also ensuring long-term preservation and accessibility. Additionally, the guide outlines the step-by-step publication process, from data preparation and metadata documentation to submission and final publication. This process is tailored to maximise the impact of each dataset, ensuring that it meets the standards of open science and FAIR data principles.

3.1 Overview of OMICS data papers

The OMICS data paper is an innovative publishing format specifically designed to encourage the detailed documentation of metadata and to support the scholarly publication of OMICS data. This approach not only enhances the visibility and accessibility of the project's outputs but also sets a benchmark for the quality and thoroughness of data reporting in the field of biotechnology research. By utilising this novel format, the project aligns with current best practices in data management and contributes to the broader scientific community's efforts to make data openly available and reusable for future research.

OMICS data papers offer a unique opportunity to publish datasets in a manner that complements traditional research publications. While research papers focus on analysing and interpreting data to derive scientific conclusions, OMICS data papers primarily describe the methodologies used for data collection, the structure and contents of the dataset, and potential applications for other researchers. This clear division of content ensures that publishing a data paper does not impede the publication of research papers that utilise the same data, as the former generally lacks the analytical components that are central to the latter.

To maximise the benefits of this approach, it is recommended that OMICS data be published in data paper form only after the corresponding research papers have been successfully released. This strategy allows researchers to fully capitalise on the novelty and significance of their findings without concerns about prematurely exposing the underlying data.

3.1.1 Benefits of OMICS data papers

Publishing OMICS data as a data paper provides several key advantages:

- **Data sharing:** Data papers offer a comprehensive and structured way to share datasets, providing detailed descriptions that help other researchers understand the data's context, how it was collected, and its potential uses. This fosters a wider use of the data, extending its impact across multiple research projects.
- **Credibility, transparency, and reproducibility:** By citing the data paper in subsequent publications, researchers can demonstrate that their findings are based on reliable, openly accessible data. This practice bolsters the credibility of the research, supports transparency, and enhances the reproducibility of scientific results.
- **Citation and recognition:** A data paper is a recognised scientific publication that can be cited independently, contributing to the researcher's academic profile. This additional publication increases the visibility of the dataset and can generate citations, further enhancing the research's impact.
- **Facilitating meta-analyses:** Data papers are invaluable for researchers conducting meta-analyses and large-scale integrative studies. By making data readily available, these papers enable the combination of datasets from different sources, helping to uncover broader patterns and insights that individual studies might miss.

- **Long-term preservation:** Publishing a dataset as a data paper ensures its long-term preservation and accessibility. This guarantees that the data remains available for future researchers, helping to maintain continuity in scientific inquiry and supporting the long-term goals of the research community.
- **Increased collaboration opportunities:** Openly sharing OMICS data through data papers can lead to new collaborative opportunities. Other researchers with complementary expertise or additional datasets may be encouraged to work together, expanding the scope and impact of the original research.

3.2 Publication options

Partners can make their data available through several options. To begin with, OMICS data should be published via one of the following choices:

1. **Publish an OMICS data paper in the Biodiversity Data Journal**, which PENSOFT will then link to the project's collection in RIO. This option is free of charge for PROMICON partners.
2. **Publish a data paper directly in the Research Ideas and Outcomes journal**, with publication benefits and processes detailed for the partners in the guiding document. This option is also free of charge for PROMICON partners.
3. **Publish a data paper in any other journal of their choice**, such as Data in Brief. Partners will then provide PENSOFT with the link so the data can be included in the project's RIO collection. This option may incur additional charges depending on the chosen journal's policies.
4. **Deposit OMICS data in a trusted open-access repository** of choice. Partners will then have to provide PENSOFT with the link for inclusion in the RIO collection.

3.3 Partners involved

All partner organisations have been informed by PENSOFT about the opportunity to publish OMICS datasets. Out of them, UFZ, and NOVA have so far expressed interest and provided details about concrete datasets that can be potentially published in the OMICS collection.

3.4 Potential datasets

To further streamline the publication efforts, PENSOFT has prepared a table listing potential datasets for the collection, which is also included in the guide. This table includes key information such as dataset descriptions, responsible researchers, and suggested timelines for publication. By providing this organised overview, the guide helps coordinate the efforts of all involved partners, ensuring that the collection is opened in a timely and efficient manner, with high-quality data contributions from across the project.

In total, four datasets have been enlisted in the table, out of which three have already started the publishing process:

- ***Metabolomics dataset of synthetic cocultures and/or monocultures*** **-** to set up the omics platforms, the synthetic coculture of phototrophs was designed. Thus, a series of metabolomics-related data were generated. *Related to: D1.8.*
- ***Metabolomics dataset of environomics from NOVA ID*** - Analytical platform delivers quantitative data from complex WP2 samples. *Related to: M1.*
- ***Datasets of natural microbial consortium*** - NMR-environomics dataset. *Related to: D1.9.*

3.5 Publication process for publication options 1 and 2

A detailed step-by-step guide has been created by PENSOFT, ensuring partners can navigate the publication process smoothly and efficiently. This guide is also available in the document distributed to partners.

4 Next steps

With the plan for publishing PROMICON's OMICS datasets now established, partners will begin the process of making their data available. Two partners have already initiated their registrations in ARPHA and are exploring manuscript preparation. They are expected to submit their datasets, advancing the project's collection in RIO.

Meanwhile, the rest of the partners are still in the process of selecting their preferred method for data publication. They are evaluating their options to determine the most suitable approach for making their datasets available, whether through data papers or depositing in open-access repositories. As these decisions are finalised, the next steps will involve coordinating with PENSOFT to ensure seamless integration of all datasets into the RIO collection. This collaborative effort will ensure that all PROMICON data is effectively published and accessible to the broader scientific community.