

## **Reconciliation commission meeting**

**Time - Date: 21:00 (CEST) - 06.12.2023**

**Attendees:** Luis Miguel Roridguez Rojas, Marike Palmer, Maria Chuvochina, Taylor Priest

### **Agenda points**

1. Official decision on “Omnitrophus fodinae” request ([seqco.de/r:d841ck02](https://seqco.de/r:d841ck02))

### **Outcome of discussion**

The Reconciliation Commission deliberated on the request (Appendix A) to grant an exception under Rule 18a (SeqCode Statutes 1.0) and accept the proposed type material to validate the name “Omnitrophus fodinae”. Taking into consideration the public discussion (<https://github.com/seq-code/seqcode/discussions/2>) as well as the associated response by the original authors of the submission (Appendix B), the Reconciliation Commission came to a majority vote to accept the request.

## Appendix A

### Request for Decision and Community Discussion

Marike Palmer, Maria Chuvochina, Taylor Priest, Luis M. Rodriguez-R

**Validation request:** proposal to grant an exception under Rule 18a in order to validate the name ‘*Omnitrophus fodinae*’

**Author/s of the original request:** Chris Rinke<sup>1</sup>

<sup>1</sup> Australian Centre for Ecogenomics, School of Chemistry and Molecular Biosciences, University of Queensland, Queensland, Australia

**Reason for request:** type material does not meet the quality standards, but abundant evidence of the species existence and novelty has been accumulated in the literature.

**Effective publication:** Rinke, C. et al. Insights into the phylogeny and coding potential of microbial dark matter. *Nature* 499, 431–437 (2013). <https://doi.org/10.1038/nature12352>

**Proposed type genome:**

Single-Cell Amplified Genome GCA\_000405945.1 [65.01% Completeness, 0.56% Contamination (CheckM2); 70.8% Completeness, 0.9% Contamination (Automated Registry Estimate by MiGA); Complete 16S rRNA and 23S rRNA genes phylogenetically concordant with other marker genes]

**Evidence for taxon and its novelty:**

In the original publication (Rinke et al., 2013), four SAGs were identified based on a Small Sub-Unit (SSU) ribosomal RNA (rRNA) gene phylogeny as belonging to the candidate grouping OP3, previously identified from rRNA gene sequences (Hugenholtz et al., 1998; Glöckner et al., 2010). Phylogenetic reconstruction based on 38 marker sequences also supported delineation of these four genomes as distinct from other phyla (Rinke et al., 2013), and the highest estimated quality SAG (GCA\_000405945.1 = SCGC AAA011-A17) was selected to serve as type for a novel phylum. Since the initial proposal, the novelty of this taxon has further been supported by phylum-level phylogenomics and comparative genomics study (Seymour et al., 2023), where phylogenies inferred from 4 different conserved marker sets, comprising 56 to 120 marker sequences, along with overall genome relatedness indices, Average Nucleotide and Amino Acid Identity (ANI and AAI), support its novelty and its exclusivity of other taxa represented by medium- or high-quality genomes belonging to the phylum (grouping into >200 putative species clusters).

(i) The species can be unambiguously identified using genomic metrics and phylogenetic analyses; (ii) the proposed type genome is a representative genome of the species cluster bearing this name in GTDB; (iii) the publication that introduce this species and its associated phylum name has been cited over thousand times and these names have been used in literature for a decade; (iv) it is the first genomic representative of the phylum.

**Classification:** ‘*Omnitrophus*’, ‘*Omnitrophaceae*’, ‘*Omnitrophales*’, ‘*Omnitrophia*’, ‘*Omnitrophota*’ (syn. ‘*Omnitrophica*’, OP3)

**Submitter comments:** “This genome has lower completeness than suggested by standards, but we request lowering of the requirement here, because this is a historical name and allows us to keep the nomenclature for the entire phylum. Despite several attempts at recovering a higher completeness genome from metagenomic data, there has been no other genomes recovered for the genus, thus in order to retain the historical genus through phylum names with the least amount of confusion, we request special consideration for validation with this genome as type.”

**Taxon description:** “This species is the type of the genus *Omnitrophus*. No data on cell size in the species is available. This species is represented by only one medium-quality genome, however, for historical significance, this genome serves as the nomenclatural type for the species. The %GC content of the genome is 49.56%. The estimated genome size is 2.9 Mb, with the available genome estimated to be 65 % complete. The genome available for this species encodes 2 very large ORFs with domains typically implicated in eukaryotic cell adhesion. This organism originated from groundwater sampled at Homestake Mine (Sanford Lab Homestake), South Dakota. The nomenclatural type for the species is the genome GCA\_000405945.1.”

**Historical significance:**

The effective publication (Rinke et al., 2013) has accrued 1,365 citations at the time of writing, and the derived phylum name “*Omnitrophota*” (= “*Omnitrophica*”) is found in the title or the abstract of 33 journal articles (Clarivate Web of Science).

**Relevant SeqCode Rules and Recommendations:**

**Rule 18a**

“The type of a species or subspecies is a designated DNA sequence that is compliant with the minimum standards designated by the SeqCode Committee for genome, metagenome-assembled genome, or single-amplified genome sequences.... Upon recommendations of the SeqCode Committee or subcommittees on the taxonomy of specific groups, the SeqCode Committee may approve other minimal standards as suitable types for specific groups.”

Appendix I, Data quality necessary for completion of SeqCode Registry\*:

Required

- Type genome assembly quality for MAGs and SAGs: >90% complete and <5% contaminated (modified from Bowers et al., 2017). For isolates, read coverage  $\geq 10x$  (Field et al., 2008).
- Agreement between genome and 16S rRNA taxonomic assignments

Recommended

- 16S rRNA genes >75% complete and passes chimera checks >80% of tRNAs present (modified from Bowers et al., 2017).

- High genome integrity (contig # <100; N50 >25 kb; max. contig >100 kb).
- MAG/SAG read coverage  $\geq 10x$ .

**\* Data quality will be assessed by automated pipelines or other approaches. Exceptions for lower data quality should be justified by authors in the effective publication.**

**Reconciliation Commission comments:**

As the effective publication predates any recommended community quality standards, and any recommendations in the SeqCode, no justification for lowering data quality standards was provided in the effective publication. However, justification in the SeqCode Registry was provided. We request opinions from the SeqCode community on whether this taxon name can be validated based on the proposed genome sequence so that the genus name can serve as a nomenclatural type for higher taxonomic ranks including the rank of phylum, thus preserving the historical name for this taxon. Any additional attempts to recover a better quality genome for this species failed (personal communication).

Potential outcomes:

A. Reject the proposal to validate the name “*Omnitrophus fodinae*” because there is not enough evidence to support the taxon description.

B. Accept the proposal to validate the name “*Omnitrophus fodinae*” based on the medium-quality draft genome.

## Appendix B

### Response to Community Discussion #2

Marika Palmer ([@MarikaPalmer](#)), Maria Chuvochina ([@uqmchuvo](#)), Taylor Priest ([@tpriest0](#)), Luis M. Rodriguez-R ([@lmrodriguezr](#))

**Community Discussion on validation request for ‘*Omnitrophus fodinae*’:** proposal to grant an exception under Rule 18a in order to validate the name ‘*Omnitrophus fodinae*’ (<https://github.com/seq-code/seqcode/discussions/2>)

### Summary of Public Discussion

Broad attempts at garnering community input were made, including a banner on the SeqCode Registry, email correspondence with registered community members, announcements on Twitter, and a post of public discussion on GitHub. This resulted in input from nine SeqCode Community members that included two of the four authors of the original request. Five contributors (excluding the authors of the request) expressed their support for validation, while two contributors were against validation of the name ‘*Omnitrophus fodinae*’ with the single-cell amplified genome GCA\_000405945.1 as nomenclatural type. The main arguments against validating the name ‘*Omnitrophus fodinae*’ were focused on whether i) the genome is of sufficient quality to allow recognition of distinct and conspecific taxa in the future, and ii) lowering of the data standards for this taxon could lead to habitual propagation of cases with lower quality nomenclatural types. A potential solution proposed by two participants was to validate the name and replace the type in the future with an alternative genome, but the current rules of the SeqCode do not allow this.

In response to these concerns, it was made clear that the genome has an estimated contamination that is within the recommended guidelines, and only falls outside of the guidelines with respect to estimated completeness. Due to the low contamination, employing this genome as the nomenclatural type will have low probability of introducing ambiguously defined species into databases, and will allow for the classification of conspecific and distinct species. Additionally, if validated, this name would constitute a single exception for the data quality standards and would not influence data quality standards across the board going forward.

Furthermore, if the name ‘*Omnitrophus fodinae*’ is not validated at this time, other members of the phylum may be validated in the near future (currently underway), resulting in a replacement of the phylum and class names, although proposed revisions under consideration by the Legislative Commission (Whitman et al., 2023, *ISME Commun.*) may provide time until 1 January, 2027 to recover a higher quality genome for this taxon. Given the extensive use of the associated phylum and class names in literature over the past decade, we believe that replacement of these names will be unfavorable, as it may cause confusion and a loss of knowledge being transferred over time.

## Conclusion

Based on viewpoints expressed in the public discussion, the majority of contributors agreed that validation of the name '*Omnitrophus fodinae*' would pose limited or no threat to the quality standards of the SeqCode and has limited potential for ambiguous species interpretations. The majority of the contributing community members considered the benefits of retaining the historical phylum name '*Omnitrophota*' and species name '*Omnitrophus fodinae*', to outweigh the potential threat of designating the lower completeness genome as nomenclatural type.

## Ballot

- A.** Reject the proposal to validate the name "*Omnitrophus fodinae*" with the medium-quality low-completeness draft genome as nomenclatural type.
- B.** Accept the proposal to validate the name "*Omnitrophus fodinae*" with the medium-quality low-completeness draft genome as nomenclatural type.