PhycoBank: Repository for algal novelties

Andreas Müller[‡], Anton Güntsch[‡], Regine Jahn[‡], Andreas Kohlbecker[‡,§], Wolf-Henning Kusber[‡], Jonas Zimmermann[‡]

‡ Freie Universität, Botanic Garden Berlin, Berlin, Germany
§ ZKM | Center for Art and Media Karlsruhe, Karlsruhe, Germany

Corresponding author: Andreas Müller (a.mueller@bo.berlin)

Abstract

The International Code of Nomenclature (ICN) for algae, fungi, and plants calls for indexing of names in nomenclatural repositories (Turland 2018, Art. 42). Scientific names, new combinations, validations, and typifications of algae are novelties tracked by PhycoBank, the registration system for algae.

PhycoBank was established and institutionalized at the Botanic Garden Berlin as the repository for nomenclatural acts concerning algae. Since June 2018, PhycoBank staff have been operating the registration system permanently. All data entered into the system undergo a curatorial process to assure a high level of data quality.

PhycoBank's three main components comprise a user-friendly data entry web application available for all registered submitters (self-registration allowed) and curators, a public data access portal, and a search engine that integrates most of the larger online repositories for algae names. The latter is not only a prerequisite for reliable data curation during the registration process but also a valuable, publicly available online tool for algae names. A fourth component still under development is handling the tight integration of the registration system with the workflow of digital publishers.

The resulting data are expected to be of high quality and to have been intensively checked against existing nomenclatural acts worldwide. A crucial requirement for the data entry application is thus an intensive support for data validation and curation. Besides the search index giving access to a huge number of existing external names, the data entry application offers a set of tools for quality assurance. It supports and requires the creation of mostly fully atomized data that strictly follow the rules of the ICN. Additionally, completeness for core data is mandatory. Both, atomization and completeness are a precondition to achieve the required uniqueness in the dataset. Completeness also implies that new combinations can only be registered after or together with their original name. The same applies to bi- or trinomials and the respective uni- and binomials they are built on.

©

The publicly available data portal allows accessing published registrations by searching for scientific names, nomenclatural authors, higher ranks and for each part of a bibliographic reference (such as bibliographic author, journal, title, year) stored in PhycoBank.

Being a name centric application PhycoBank is neutral with respect to taxonomic opinions. Therefore, only nomenclatural synonymies (basionym or replaced synonym relationships) are stored in the system and names are not attached to a unique classification. Instead, to facilitate search via higher ranks as required by users, classification information is stored as a directed graph of higher taxa with registered names linking to this graph. By this, searches can be performed on multiple classifications simultaneously and can return all possible matches.

PhycoBank uses http-based persistent identifiers (e.g. http://phycobank.org/102170), which makes them resolvable and actionable. These identifiers link to nomenclatural acts only and not to the information in the act. That is, a PhycoBank identifier should not be used to refer to a scientific name even if this name was established in the nomenclatural act.

PhycoBank makes use of the application stack offered by the EDIT Platform for Cybertaxonomy (Kohlbecker 2017). However, as registration workflows essentially differ from other taxonomic workflows, a completely new user interface for editing has been developed, which provides the users with an intuitive and fluent user experience despite the high complexity of the data.

The Common Data Model (Müller 2017) is the core data model for the Platform and already covered the vast majority of data types and fields required by the registration in terms of completeness and degree of atomization. For PhycoBank, it only needed to be extended by a single data type representing the registration/nomenclatural act itself.

As of July 2022, PhycoBank includes 4,332 registrations, of which 4,202 are name novelties (1,523 in preparation or ready, 239 under curation, 2,407 published, and 33 rejected); 130 registrations refer to lectotypes or epitypes of existing names.

PhycoBank will apply for recognition as a repository in 2022. This is a prerequisite for a proposal to make registration of nomenclatural acts for algae mandatory. This is possible before, at or after the 20th International Botanical Congress 2024.

Keywords

nomenclature, phycology, registration, names backbone, EDIT Platform

Presenting author

Andreas Müller

Presented at

TDWG 2022

Funding program

PhycoBank was initially funded by the German Research Foundation (DFG, JA 874/8-1).

Conflicts of interest

References

- Kohlbecker A, et al. (2017) The EDIT Platform for Cybertaxonomy, a Brief Overview. Proceedings of TDWG 1 https://doi.org/10.3897/tdwgproceedings.1.20368
- Müller A, et al. (2017) A Comprehensive and Standards-Aware Common Data Model (CDM) for Taxonomic Research. Proceedings of TDWG 1 https://doi.org/10.3897/tdwgproceedings.1.20367
- Turland N, et al. (Ed.) (2018) International Code of Nomenclature for algae, fungi, and plants. Regnum Vegetabile https://doi.org/10.12705/code.2018