

Using the EDIT Platform for Cybertaxonomy to prepare and publish a treatment for the Caryophyllales Network: an online synthesis of the Nepenthaceae

Authors: Berendsohn, Walter G., Borsch, Thomas, Güntsch, Anton, Kohlbecker, Andreas, Korotkova, Nadja, et al.

Source: Willdenowia, 48(3) : 335-344

Published By: Botanic Garden and Botanical Museum Berlin (BGBM)

URL: <https://doi.org/10.3372/wi.48.48301>

BioOne Complete (complete.BioOne.org) is a full-text database of 200 subscribed and open-access titles in the biological, ecological, and environmental sciences published by nonprofit societies, associations, museums, institutions, and presses.

Your use of this PDF, the BioOne Complete website, and all posted and associated content indicates your acceptance of BioOne's Terms of Use, available at www.bioone.org/terms-of-use.

Usage of BioOne Complete content is strictly limited to personal, educational, and non - commercial use. Commercial inquiries or rights and permissions requests should be directed to the individual publisher as copyright holder.

BioOne sees sustainable scholarly publishing as an inherently collaborative enterprise connecting authors, nonprofit publishers, academic institutions, research libraries, and research funders in the common goal of maximizing access to critical research.

WALTER G. BERENDSOHN¹*, THOMAS BORSCH¹, ANTON GÜNTSCH¹, ANDREAS KOHLBECKER¹, NADJA KOROTKOVA¹, KATJA LUTHER¹, ANDREAS MÜLLER¹, PATRICK PLITZNER¹ & SABINE VON MERING¹

Using the EDIT Platform for Cybertaxonomy to prepare and publish a treatment for the *Caryophyllales* Network: an online synthesis of the *Nepenthaceae*

Version of record first published online on 17 September 2018 ahead of inclusion in December 2018 issue.

Abstract: The *Caryophyllales* Network strives to assemble an online dynamic synthesis of the order *Caryophyllales*, uniting the current knowledge about the phylogeny of the order with up-to-date information on the individual taxa contained. Capturing taxonomic data and the decision processes involved in the definition and circumscription of the taxa requires highly complex specialized software. The *Caryophyllales* Network uses the EDIT Platform for Cybertaxonomy for that purpose. In the context of the online treatment of the family *Nepenthaceae*, we describe the steps taken to assemble the database, the interaction with other electronic sources, the links with the World Flora Online initiative, and the prospects for the maintenance and further development of the *Nepenthaceae* segment of the *Caryophyllales* database. *Nepenthaceae* constitute an example of a family with a relatively recent flora treatment (*Flora Malesiana*, published in 2001), which to a large extent covers its total range of distribution, but with further species subsequently described as new to science in mostly regional treatments, and with an analysis of relationships and species limits on the basis of evolutionary methods just emerging. A snapshot of the current state of the database is provided as an annotated checklist in PDF format in the Supplementary Material online, which includes 176 species and nine naturally occurring named hybrids and treats 435 species and infraspecific names.

Key words: *Caryophyllales*, checklist, cybertaxonomy, database, EDIT Platform, monographing, *Nepenthaceae*, *Nepenthes*, taxonomy, WFO, World Flora Online

Article history: Received 3 August 2018; peer-review completed 24 August 2018; received in revised form 27 August 2018; accepted for publication 28 August 2018.

Citation: Berendsohn W. G., Borsch T., Güntsch A., Kohlbecker A., Korotkova N., Luther K., Müller A., Plitzner P. & Mering S. von 2018: Using the EDIT Platform for Cybertaxonomy to prepare and publish a treatment for the *Caryophyllales* Network: an online synthesis of the *Nepenthaceae*. – Willdenowia 48: 335–344. doi: <https://doi.org/10.3372/wi.48.48301>

Introduction

The *Caryophyllales* Network — This specialists' network was initiated in 2011 (Arias & al. 2016) with the aim to mobilize and coordinate the international endeavour to assemble a “Global synthesis of species diversity in the angiosperm order *Caryophyllales*” (Borsch & al. 2015). A networked and informatics-supported approach is taken to build up baseline information on species in a single open online resource. Providing open access to that information will facilitate and inspire research, avoid duplication of effort, constantly improve quality

and consistency of the underlying data, stimulate contributions and directly contribute to other global initiatives. Concurrently, open research questions in biodiversity informatics but also evolution and taxonomy become obvious and can be tackled. This already goes beyond the call for online floras issued by Brach & Boufford (2011). The first result of this collaborative effort was the publication of a family and genus-level checklist based on the latest (published) phylogenetic knowledge (Hernández-Ledesma & al. 2015), which has since been updated frequently on the *Caryophyllales* website (*Caryophyllales* Network 2015+).

¹ Botanic Garden and Botanical Museum Berlin, Freie Universität Berlin, Königin-Luise-Straße 6–8, 14195 Berlin, Germany; *e-mail: w.berendsohn@bgbm.org (author for correspondence).

The database, online portal and publication pipelines of the *Caryophyllales* Network are using the EDIT Platform for Cybertaxonomy (Ciardelli & al. 2009; Berendsohn 2010) to capture, process, maintain and publish the data. Short-term aims in the assembly of information are to document the state of the art with respect to:

- Nomenclatural data including references, links to publication of related nomenclatural acts (e.g. protologues in the Biodiversity Heritage Library, BHL 2005+), nomenclatural status, types, and monotypic groups of names
- A consistent taxonomic backbone, as far as possible based on results of phylogenetic analyses
- Distribution data increasingly directly or indirectly based on (expert revised!) specimen locations

Medium and long term aims:

- Building an ontology of descriptive terms
- Integrating (monographic) descriptive data (morphological and molecular)
- Identification keys based on structured descriptive data (character / character state data)
- Additively aggregated, structured descriptive data for taxa from specimens
- Integrating local, regional and global taxonomic treatments
- Output of localized treatments
- Handling of differing taxonomic concepts
- Inclusion of quality assessments with regard to the state of current knowledge of the taxa

The *Caryophyllales* initiative explicitly aims at integrating published results and data, ongoing research, and usage of the information, not merely at aggregating data. Its main target audience is researchers, although it will also serve as a quality source for a broad range of user communities, either directly or by contributing to aggregators such as the Global Biodiversity Information Facility (GBIF 2004+), Catalogue of Life (CoL; Bisby 1997; Roskov & al. 2018), Encyclopedia of Life (EoL; Parr & al. 2014) and the World Flora Online (WFO; Wyse Jackson & Miller 2015) initiative.

The World Flora Online initiative — (Wyse Jackson & Miller 2015; Loizeau & Wyse Jackson 2017). The World Flora Online will be an open-access, web-based compendium of the world's plant species. This is answering Target 1 of the updated Global Strategy for Plant Conservation (GSPC) 2011–2020 of the U.N. Convention on Biological Diversity (CBD; COP 2010; Sharrock 2012). In response to this target, an international consortium of institutions and organizations was formed to collaborate on providing the content. The 11th Conference of the Parties to the Convention on Biological Diversity adopted a decision welcoming the World Flora Online initiative

(CBD-SBSTTA 2012). Since January 2013, up to June 2018, 41 institutions have signed a Memorandum of Understanding, agreeing to their participation in the World Flora Online initiative.

WFO will assemble this compendium by aggregating information from monographs, taxonomic networks and floras based on a consensus “taxonomic backbone”, i.e. a list of taxa with their synonyms. The initial backbone is formed by the checklist assembled by a consortium of institutions under the lead of Missouri Botanical Garden and Royal Botanic Gardens, Kew as the outcome of Target 1 of the GSPC 2001–2010 (COP 2002), namely “The Plant List” (TPL 2013). A unique WFO Identifier has been assigned to each of the names in the backbone and will be maintained permanently. Presently, WFO's pragmatic approach to aggregation will be name-based, i.e. the incoming information will be linked to the backbone taxon to which the incoming name belongs, thus for the time being ignoring possible concept conflicts (see below). The *Caryophyllales* Network was recognized by the WFO Council as one of WFO's TENs (Taxonomic Expert Networks), whose responsibility it is to take on a specific “taxonomic slice” of the WFO and revise the taxonomy. In this context, one of the first aims for the Network is to provide an updated WFO backbone for *Caryophyllales*, which implies that all names in the WFO backbone derived from The Plant List have to be resolved in the *Caryophyllales* database, and are therefore to be placed within the best currently available taxonomic treatment for the respective species, genus or family.

Why Nepenthes? — The assembly of the taxonomic backbone is a tedious task, given the number of unresolved names that have been introduced into The Plant List and therefore into the original WFO backbone. In 2016, we were looking for a taxonomic group to try out the necessary procedures and, if possible, go beyond mere backbone data. We were therefore in need of a relatively small family with clear monophyly also at the generic level, and with existing, high-quality, electronic, open, taxonomic data that we could use as a starting point. The genus *Nepenthes* is clearly monophyletic (Meimberg & Heubl 2006; Meimberg & al. 2006). With most of the taxa in the monogeneric *Nepenthaceae* occurring in the *Flora Malesiana* area and therefore already present in the EDIT Platform, this family appeared to be an ideal candidate. Thus commenced the first author's dive into the exciting world of pitcher plants and their protagonists around the world.

Material and methods

Sources — The most important single source of data for the treatment presented here was the *Nepenthaceae* treatment in *Flora Malesiana* (Cheek & Jebb 2001), which was parsed and imported into an EDIT Platform database between 2010 and 2015 by Thomas Hamann and Andreas

Müller (Roos & al. 2012; Hamann & al. 2014). Unfortunately, the project supporting this process ended prematurely and no further resources could be committed up to now to iron out the remaining problems in the online version of *Flora Malesiana* (<http://portal.cybertaxonomy.org/flora-malesiana/>). The *Nepenthes* treatment was extracted from the *Flora Malesiana* database and imported into the *Caryophyllales* database. Further data-cleaning was effected to correct problems remaining from the import, and the extra-Malesian taxa were added, initially using the “skeletal revision” of the genus by Jebb & Cheek (1997). New taxa were added after reviewing many taxonomic treatments, which cannot be listed in full here (they appear as sources in the online version); the online Wikipedia pages for *Nepenthes* proved to be important pointers to new publications. Further sources that should be mentioned here include the entertaining books by McPherson & al. (McPherson 2009, 2011) and the treatments by Clarke (1997, 2001), unfortunately with a number of newly described species behind a print-only paywall and thus raising an unnecessary obstacle to reusing the data. The travel diary of Catalano (2010) from Thailand should also be cited; he gracefully published the protologues and articles with further species descriptions from Indo-China online on his website (<http://www.marcellocatalano.com>). The ongoing publication of contributions to a full revision of the family by Cheek & Jebb (see Cheek & al. 2018 for a list of references) are an important source, also providing some data on the subgeneric classification for the species, mostly formulated as informal species groups. With regard to species-level phylogenetic analyses, Brittnacher (2011) gave an overview of developments up to that year, citing 78 species sampled, and Schwallier & al. 2016 (five species) and Bunawan & al. 2017 (11 species) are apparently the most recently published analyses. It is noteworthy that the available molecular phylogenetic investigations on *Nepenthes* (e.g. Meimberg & Heubl 2006; Meimberg & al. 2006; Bunawan & al. 2017) employ few genomic regions and sample mostly only a single individual per species. The evolutionary testing of species limits and species relationships is therefore just at the beginning. Clarke & al. (2018) summarized the results and also provided a checklist of species (stating the caveat that “even the four authors of this article cannot agree upon the status of approximately 10% of *Nepenthes* species”), reviewed the taxonomic groups established by, and since the revision of, Danser (1928), and critically reviewed the “taxonomic inflation” caused by a recent flurry of new species publications.

The Name Matching Tool of the Tropicos website (Tropicos 1991+) was valuable for detecting discrepancies in nomenclatural reference citations. The slight taxonomic complication of having so many present-day carnivorous plant enthusiasts discussing species limits etc. was insignificant compared to the problem of sorting out names coined by horticulturists in the second half of the 19th century. Periodicals like the *Gardeners’*

Chronicle contain numerous names, often of artificial hybrids, which have to be considered in a nomenclatural context. Many of these are contained in the International Plant Names Index (IPNI 2004+), sometimes including names attributed to later (nomenclaturally superfluous) publications. From there, they also found their way into The Plant List (TPL 2013) and consequently into the taxonomic backbone of the World Flora Online database. This meant that they had to be treated without exception, and in processing those names additional ones emerged from that type of horticultural literature. That these periodicals are almost completely available online, mostly in the Biodiversity Heritage Library (BHL 2005+), greatly facilitated the investigation. Protologues were also located using JSTOR (2000+), Google Scholar and results obtained using search engines.

Tools: the EDIT Platform — The EDIT Platform for Cybertaxonomy (Ciardelli & al. 2009; Berendsohn 2010; Kohlbecker 2017) provides a comprehensive set of tools for editing and management of taxonomic data (individually or collaboratively), fully customizable online access to that data, and other means of data publication and data exchange (BGBM 2011+). The development of the Platform predates the EU-funded EDIT (European Distributed Institute of Taxonomy) project, which provided major funding; the Platform is based on information models and software developed starting in the early 1990s (Berendsohn & al. 2003). At the core of the platform is the EDIT Common Data Model (CDM), a comprehensive information model covering the entire information domain: names and classifications, descriptive data (morphological, molecular), multimedia, geographic information, literature, specimens, people and user-defined content items (Anonymous 2008+; Müller & al. 2017). The open-source software provides solid and open interfaces including a Java programmers’ library and a Web Service Layer. The most important applications are the highly configurable Taxonomic Editor (TaxEditor) for data entry and manipulation, the CDM Data Portal (Kohlbecker & al. 2008) including geo-services (Roca & al. 2009) for online display, and the Platform Web Services (Güntsch & al. 2017a). The TaxEditor (available for Windows, Linux and Macintosh operating systems) is the local client that allows users to enter, edit, export and otherwise manipulate the data in the CDM database, which resides on a CDM Server connected to the Internet. Multiple authors can work simultaneously on the same database. The *Caryophyllales* database is hosted by the BGBM. Data in Platform databases can be exported into various formats. In the context of the work presented here, we used the CDM-light export to generate tables for further processing. The EDIT Platform is used by a number of initiatives; see BGBM (2016+) for a list of publicly accessible Data Portals. The role of the EDIT Platform in the context of the *Caryophyllales* synthesis is more fully detailed in Borsch & al. (2015).

Taxonomic concepts and source citations in the EDIT Platform — To understand the Platform's handling of taxonomic concepts, a brief clarification of terminology is required.

The **Classification** is the uppermost hierarchical element in the Platform's handling of taxa. Several classifications can reside in a single database. This is useful, for example, when there are alternative views on taxon circumscriptions (e.g. in the treatment of the genera *Hieracium* and *Pilosella* in the *Cichorieae*, see Kilian & al. 2009+). An accepted name ("correct name" according to the rules of nomenclature) nested within a classification designates a **Taxon Node**, representing a taxon in a given classification. If a taxon node is assigned to a taxon of a higher rank, the latter is referred to as the **parent taxon**, the former as the latter's **child taxon**.

A **Taxon** is a taxonomic group with the data that define its circumscription and describe its properties. The circumscription of the taxon is indicated by means of a circumscription or concept reference ("**sec.-**" or "**secundum-**" reference, Berendsohn 1995, 1997; Berendsohn & Geoffroy 2007), normally a bibliographic reference clarifying the distinction of this taxon from other taxa. One and the same taxon may occur in several classifications, but it is also possible that two different taxa (taxon concepts) carry the same name in separate classifications. In themselves, classifications should be taxonomically consistent, i.e. every name should only occur once (as a taxon name or a synonym) in a given classification (except when cited as a misapplied name or pro-parte synonym).

Scientific names (as well as the names given to pseudotaxa such as unnamed clades) are assigned to taxa, synonyms or misapplied names. **Ranks of names** follow the hierarchy defined in the nomenclatural codes. However, the user is free to insert further ranks at any place in the tree thus forming new or mixed hierarchical levels.

The EDIT Platform offers various ways to indicate the source of a taxonomic concept. In addition to the "sec. reference" that indicates the circumscription of the taxon, the assignment of synonyms or of misapplied names may carry an individual "sec. reference", too, here indicating who has linked this name to the concept of the taxon. This allows addition of such information without having to create a new taxon concept. In case of misapplied names, this attribution is distinguished from the "sensu" reference that indicates the actual misapplication. Moreover, factual data (as virtually all data items in the Platform) may be individually referenced (including the name used in the source), e.g. geographic distribution, ecological data, or uses of the taxon. It is also possible to assign a block of sourced factual data (e.g. the treatment in a Flora) to the taxon. Notes (or otherwise designated free-text fields) can be used to indicate the state of revision of the concept. These features were all used in the *Nepenthaceae* database presented here.

The Platform also allows linkage of different taxonomic concepts. An example: *Nepenthes dubia* Danser was treated by Jebb & Cheek (1997) and Cheek & Jebb (2001) as a taxon including *N. tenuis* Nerz & Wistuba; a hypothesis that was rejected by Clarke (2001), McPherson (2009) and Clarke & al. 2018, who treat *N. tenuis* as a separate taxon. In the primary classification in the database (visible in the portal), we presently follow Clarke's opinion, so we do have a taxon *N. dubia* sec. Clarke (2001), but we maintain the *Flora Malesiana* treatment as a taxon *N. dubia* sec. Cheek & Jebb (2001) in a parallel classification (for technical reasons not yet shown in the portal). The EDIT Platform allows us to establish a concept relationship (Gradstein & al. 2001) between the two Taxon Nodes, in this case "*N. dubia* sec. Cheek & Jebb (2001) includes, and is not congruent with, *N. dubia* sec. Clarke (2001)".

Specimens in the EDIT Platform — The Platform covers specimens in a comprehensive way. The data items available in the ABCD (Access to Biological Collection Data; Berendsohn 2007–2010; Holetschek & al. 2012) and DwC (Darwin Core; Wieczorek & al. 2009) standards are included in the Platform, such as those relating to the collection site, collection event, processing, herbarium location, and identification history of the specimen. Moreover, the hierarchical modelling of such data in the CDM allows linkage of all derivatives of materials collected, i.e. specimen duplicates, images taken at the collection site, images of specimens, tissue samples, DNA samples, etc. (Kilian & al. 2015; Plitzner & al. 2017, 2017a).

Descriptive information in the EDIT Platform — The hierarchical structure for specimens is also important in the context of descriptive information. Aggregation of atomized character data from specimens is a current research topic in biodiversity informatics (Plitzner & al. 2017, 2017a; Henning & al. 2018). The CDM already fully supports the comprehensive TDWG Standard SDD (Structured Descriptive Data; Hagedorn & al. 2005), and linkage to ontologies or terminology services (Karam & al. 2016) are features under development, as well as integrating such data and services into the Platform's user interfaces. Structured descriptive data may be generated using the stand-alone Xper² software (Ung & al. 2010; Venin & al. 2010). In parallel, several structures support the entry and storage of categorized descriptive data items. Content under categories like "Notes", "Etymology", "General Distribution", etc.), as often found in Floras and monographs, can be captured as unstandardized text for each category. However, standard prescribed vocabularies may be defined for a category (e.g. IUCN status). Categories may be structured hierarchically (e.g. geographic areas), and some structures can even relate to multiple levels of a hierarchical category (e.g. occurrence status related to an individual distribution area).

Current state of the database and editorial rules

The *Caryophyllales* Network is still in the process of defining the editorial rules to be applied to the entire treatment of the order, so the specifications given below may change in the future (this will be indicated on the web-site). It is clear, however, that the full citation of sources of information as outlined above should be followed wherever possible. As the database normally provides highly atomized content, the output can be adapted to different formats with ease. The format used in the data portal or in the PDF edition (supplemental materials) in many respects follows the rules formulated by Kilian & al. (2009+) for the *Cichorieae* Network, another initiative using the EDIT Platform. Although the Platform offers a way to enter information items such as references, author teams, person names, specimens and even scientific names as unstructured text, the eventual aim is to enter these in a fully atomized and thus re-usable form. Updated versions of the PDF-edition will be provided as new data become available, while the data on the portal always represent the actual state of the database.

References — All references cited in the treatment should eventually be entered into the EDIT Platform database in fully atomized form (i.e. with authors forming an author team assembled from persons, with articles including the title also for nomenclatural references although this does not figure in the output of these, etc.). The “in” references for parts of a work in nomenclatural citations are also a bibliographic reference. In the database, two reference records are created, one for the “in” citation of the full work with its respective Reference Type (usually a book), and one for the part of the work (type “Book section” or “Section”). For example, the nomenclatural citation “*Nepenthes epiphytica* A. S. Rob., Nerz & Wistuba in S. McPherson, New *Nepenthes* 1: 36. 2011” creates two reference records in the database. To make the publication citable, care has to be taken to enter the full last name(s) of the author(s) and the full title as well as their abbreviated standard forms.

Nomenclatural citation of taxon names (adapted from Kilian & al. 2009+) — Titles of monographs are abbreviated in conformity with Taxonomic literature, ed. 2 (TL-2; Stafleu & Cowan 1976–1988; Stafleu & Menega 1992–2000; Dorr & Nicolson 2008–2009), but all components start with capital letters. Titles of monographs not listed in TL-2 follow the version used in IPNI, with the exception of the recently introduced inclusion of the editor in the title. Titles not mentioned in any of these sources are abbreviated as in serials. Titles of serials are abbreviated in conformity with BPH-2 (Bridson & al. 2004; online edition); titles of serials not listed in BPH-2 are abbreviated according to the standards therein defined.

Volume and fascicle numbers: volume numbers of monographs are as a rule given only if volumes have separate pagination. The same applies to fascicles of serial volumes. The publication year always refers to the actual year of publication of the page cited. If a work has been published in parts and its publication dates have been established, the publication year is that of the corresponding part and can differ from the year(s) cited on the (main) title page of that work; the latter may also be simply incorrect. In the case of known differences between the year cited in the publication and the real publication date, the former is set in quotation marks and the real value in square brackets. Series that form part of the title are normally included in the standard abbreviation for nomenclatural citations.

Authors, Persons and Teams — At present, the following rules are applied in the database: Single persons (authors, collectors, etc.) are represented by a Person record. Multiple persons (e.g. author teams) are represented by a Team record, which consists of several Team Members, which in turn are Person records. For nomenclatural authors, the standard abbreviation has to be entered in the Person record. The abbreviation must be in accordance with the authors of plant names appearing in IPNI (2004+). Full person names and initials should be entered for Persons acting as authors of bibliographic references, as collectors of specimens, and for users (i.e. those with a password for the respective database).

Specimen data — Currently, specimen data in the *Nepenthaceae* database are mostly restricted to nomenclatural types (see following section) and the input is not normally atomized. It is planned to link the cited specimens to digitized specimen images (e.g. JSTOR 2013+) as far as available and to atomize the data according to the aforementioned specimen data standards. Care is taken to cite resolvable and unique stable identifiers (Güntsch & al. 2017; Groom & al. 2017) of the specimens, if available.

Nomenclatural types of names — The original citation of type material from the protologue should be entered. Data on individual type specimens are gathered from different sources, such as directly from the protologues, from secondary literature, and in the revisionary process from the type specimen labels. However, currently type information is mostly added cursorily as a string in a field reserved for preliminary information. The EDIT Platform allows categorization of type specimens as holo-, syn-, neo-, lectotypes, etc., and their duplicates as iso-, isosyn-, isoneo-, isolectotypes, etc., or as original material. If no such specification is possible, they are simply indicated as “type”.

Geographic distribution and presence status of the taxon — Geographical distributions of *Nepenthes* species are currently entered as a text string in the field for general dis-

tribution. However, the Platform allows for the definition of geographic vocabularies and hierarchies, so the aim is to list the individual areas where the taxon is present, and to give details on the “distribution status” in that area (e.g. “native”, “introduced”, etc.). A first step could be the use of TDWG areas (Brummitt 2001) and ISO countries (ISO 2013), which are already available as standard vocabularies in the Platform. In future, ideally the distribution and geographical status will be derived from geo-referenced, individual specimen records, where present. The Platform Data Portal allows generation of area-based maps as well as dot-maps for specimen records.

Descriptive data — The descriptive morphological data obtained as a result of the parsing process of the *Flora Malesiana* treatments were preliminarily stored in the structure used for general descriptive data items (“factual data” in Platform terminology), i.e. as character-character state data with non-standardized states as extracted from the text. Several dozens of characters were thus defined, among them a number highly specific to the *Nepenthaceae*, for example 17 characters refer to aspects of the pitchers, such as, character: “Spur”, state: “7–9 mm long, filamentous, forked 4–6.5 mm from the apex”.

The full text description is stored in parallel (and shown in the portal). At present, no molecular data are included in the database.

Interaction with the World Flora Online — A close understanding of WFO necessities could be incorporated in the workflows through a close link with the Technical Working Group of the World Flora Online (<http://worldfloraonline.org>) connecting the two initiatives. In a first step, the WFO technical team provided the name data (originating from TPL) including the respective WFO name IDs. WFO-IDs were then added to the names already existing in the *Caryophyllales* database, and additional TPL names were imported, together with those found to be missing (TPL lacks names published after 2010, most infraspecific names, and some names found in the horticultural literature). Those additional names were assigned a WFO ID from an allocation of WFO IDs given to the *Caryophyllales* Network. A WFO-conformant data output of the *Nepenthaceae* data from our database was created as a DwC-A file, which will be ingested into WFO and replace the *Nepenthaceae* segment of the WFO taxonomic backbone (see WFO 2017+). The next step will be the provision of the descriptive data to WFO. Acting as a WFO Taxonomic Expert Network (TEN), the *Caryophyllales* Network will provide regular updates to the WFO taxonomic backbone also in the future. Where a consensus of experts exists, this feedback to WFO may also be extended to provide a preferred view to descriptive data or new versions of descriptive data matching up-to-date taxon concepts. The exact procedures are still being discussed by the WFO Council and working groups.

Discussion

As stated in the introduction, the prime purpose of the *Nepenthaceae* treatment presented here is to demonstrate the technical and editorial possibilities provided by the EDIT Platform in the context of the *Caryophyllales* Network and its aims, including the publication of a “data paper” (Chavan & al. 2011). An institutional commitment for the long-term sustainability of the database and the EDIT Platform itself is provided by the BGBM. Further development of its content strongly depends on the developments in the *Caryophyllales* Network. Tasks that should be tackled to maintain the content include the addition of new taxa and of taxonomic changes into the database to maintain its role as the taxonomic backbone. Expert members of the *Caryophyllales* Network are invited to take editorship of taxonomic groups and can rely on technical assistance by the Platform team. A structured geographical scheme for the known (past or present) distribution can be added with comparatively little effort. Standardization of the terminology for descriptive characters, especially of the *Nepenthes*-specific pitcher characters (see Cheek & Jebb 2013; for some ecological traits, Clarke & Moran 2011) and complete coverage (where possible) of the taxa seems to be attainable when considering the published data. Transferring this into a standardized data structure and thus making it available for the generation of localized or general identification keys is a next step. Given that there is a strong community of enthusiasts for this group, assembly of images of the taxa should also be possible.

To identify gaps in our current knowledge of the taxa, as well as to inform users about possible problems in the application of the data (e.g. in the context of conservation), some kind of standardized assessments of data quality are needed. With respect to the scientific base and therefore stability of the accepted taxon concept at species level, the classification proposed by Borsch & al. (2018) may be applied by the specialists in the group.

The Plant List (TPL 2013) listed 262 plant names for *Nepenthes* (vs the 435 specific and infraspecific names we now treat in the database), almost all of which were considered “unresolved”. The taxonomic knowledge existing in the published literature was much further advanced, but not present in the databases that provided the base for TPL. The revision by Jebb & Cheek (1997) recognized 82 species, with 74 in the area of the *Flora Malesiana*. In their treatment for that Flora, Cheek & Jebb (2001) already recognized 80 species in the area, plus three naturally occurring hybrids. Of the species recognized in the Supplementary Material of this article, 32 were published between 2001 and 2010, and an additional 47 from 2011 onward. In their checklist, Clarke & al. (2018) accepted 170 species in eight sections, probably not considering four species that were published after preparation of their manuscript. The *Nepenthaceae* section in the *Caryophyllales* database includes those

newly described species and accepts *N. extincta* and *N. kurata*, thus recognizing 176 species and nine naturally occurring, named hybrids.

Roles of the authors and Acknowledgements

Co-authors are cited alphabetically. Based on the *Flora Malesiana* treatment and other accessible literature, Berendsohn effected most of the taxonomic and nomenclatural investigation and data entry as well as the transformation of the Platform's generic output to the document provided in Supplementary Material online. Müller prepared the electronic version of Cheek & Jebb's *Flora Malesiana* treatment and effected the importation of WFO-IDs and of additional names gathered from IPNI and the WFO backbone. Luther designed the Platform's generic "CDM-light" export that served as the base for the DwC-Archive for WFO and for the creation of the text provided as Supplementary Material. Kohlbecker assisted with the online portal. Plitzner helped to resolve issues with the editing software. Güntsch, Borsch, von Mering and Korotkova assisted in the formulation of editorial rules and (the latter two) contributed their experience of Platform usage in other taxa within *Caryophyllales*. All authors contributed to the text and approved its final version.

Thanks go to Nick Turland and Werner Greuter for advice on nomenclatural issues and Frank Specht for his cooperation in the programming of the WFO DarwinCore Archive file. Sophia Wutke and Johanna Bellack copied the descriptions of species published in print-only publications.

The reviewers Mark Watson, Quentin Groom and Richard Rabeler contributed considerably to improve the text. In addition to that, Andreas Fleischmann pointed out some recent research that needed to be included in the treatment.

References

- Anonymous 2008+: Common Data Model. – Published at <http://dev.e-taxonomy.eu/trac/wiki/CommonDataModel> [accessed 31 May 2018].
- Arias S., Berendsohn W. G., Borsch T., Flores-Olvera H., Ochoterena H., von Mering S. & Zuloaga F. O. 2016: *Caryophyllales* 2015 in: Berlin and the Global *Caryophyllales* Initiative. – *Taxon* **65**: 427–428.
- Berendsohn W. G. 1995: The concept of "potential taxa" in databases. – *Taxon* **44**: 207–212.
- Berendsohn W. G. 1997: A taxonomic information model for botanical databases: the IOPI model. – *Taxon* **46**: 283–309.
- Berendsohn W. G. (ed.) 2007–2010: ABCD – Access to Biological Collection Data. – Published at <https://archive.bgbm.org/tdwg/codata/schema/> [accessed 26 Aug 2018].
- Berendsohn W. G. 2010: Devising the EDIT Platform for Cybertaxonomy. – Pp. 1–6 in: Nimis P. L. & Vignese-Lebbe R. (ed.), Tools for identifying biodiversity: progress and problems. – Trieste: EUT Edizioni Università di Trieste. – Also published at <http://hdl.handle.net/10077/3737> [accessed 26 Aug 2018].
- Berendsohn W. G., Döring M., Geoffroy M., Glück K., Güntsch A., Hahn A., Kusber W.-H., Li J.-J., Röper D. & Specht F. 2003: The Berlin taxonomic information model. – *Schriftenreihe Vegetationsk.* **39**: 15–42.
- Berendsohn W. G. & Geoffroy M. 2007: Networking taxonomic concepts – uniting without "unitaryism". – Pp. 13–22 in: Curry G. & Humphries C. (ed.), Biodiversity databases – techniques, politics, and applications. – Baton Rouge: CRC Taylor & Francis.
- BGBM 2011+: EDIT Platform for Cybertaxonomy. – Published at <http://www.cybertaxonomy.org> [accessed 26 Aug 2018].
- BGBM 2016+: EDIT Platform for Cybertaxonomy, reference projects. – Published at <https://cybertaxonomy.eu/dataportal/reference> [accessed 26 Aug 2018].
- BHL 2005+ [continuously updated]: Biodiversity Heritage Library. BHL project. – Published at <http://www.biodiversitylibrary.org/> [accessed 26 Aug 2018].
- Bisby F.A. 1997: Putting names to things and keeping track: the Species 2000 programme for a co-ordinated catalogue of life. – Pp. 59–68 in: Bridge P., Morse D. & Scott P. R. (ed.), Information technology, plant pathology and biodiversity. – Wallingford: CAB International.
- Borsch T., Flores-Olvera H., Zumaya S. & Müller K. 2018 [in press]: Pollen characters and DNA sequence data converge on a monophyletic genus *Iresine* P. Browne (*Amaranthaceae*, *Caryophyllales*) and help to elucidate its species diversity. – *Taxon*.
- Borsch T., Hernández-Ledesma P., Berendsohn W. G., Flores-Olvera H., Ochoterena H., Zuloaga F. O., von Mering S. & Kilian N. 2015: An integrative and dynamic approach for monographing species-rich plant groups – Building the global synthesis of the angiosperm order *Caryophyllales*. – *Perspect. Pl. Ecol. Evol. Syst.* **17**: 284–300.
- Brach A. R. & Boufford D. E. 2011: Why are we still producing paper floras? – *Ann. Missouri Bot. Gard.* **98**: 297–300.
- Bridson G. D. R., Townsend S. T., Polen E. A. & Smith E. R. 2004: BPH-2. Periodicals with botanical content. Updated online edition. – Pittsburgh: Hunt Institute. – Published at <http://fmhibd.library.cmu.edu/HIBD-DB/bpho/findrecords.php> [accessed 26 Aug 2018].
- Brittnacher J. 2011: *Nepenthes* phylogeny. International Carnivorous Plant Society. – Published at <http://www.carnivorousplants.org/cp/evolution/Nepenthes> [accessed 26 Aug 2018].
- Brummitt R. K. 2001: World geographical scheme for recording plant distributions. Edition 2. Plant taxonomic database standards 2. – Pittsburg: TDWG,

- Hunt Institute. – Also published at <https://github.com/tdwg/wgsrpd> [accessed 26 Aug 2018].
- Bunawan H., Choong C. Y., Yaakop S. & Noor N. M. 2017: Phylogenetic inferences of *Nepenthes* species in Peninsular Malaysia revealed by chloroplast (*trnL* intron) and nuclear (ITS) DNA sequences. – *BMC Res. Notes* **10**: 67.
- Catalano M. 2010: *Nepenthes* della Thailandia – diario di viaggio. – Prague: WOW s.r.o.
- Caryophyllales Network 2015+ [continuously updated]: A global synthesis of species diversity in the angiosperm order *Caryophyllales*. – Published at <http://caryophyllales.org> [accessed 26 Aug 2018].
- CBD-SBSTTA 2012: Global Strategy for Plant Conservation: World Flora Online by 2020. Note by the Executive Secretary. – Published at <http://www.cbd.int/doc/meetings/sbstta/sbstta-16/information/sbstta-16-inf-38-en.pdf> [accessed 26 Aug 2018].
- Chavan, V. & Penev L. 2011: The data paper – a mechanism to incentivize data publishing in biodiversity science. – *BMC Bioinf.* **12**(Suppl. 15): S2.
- Cheek M. & Jebb M. H. P. 2001: *Nepenthaceae*. Flora Malesiana, Ser. **I**, **15**. – Leiden: Nationaal Herbarium Nederland.
- Cheek M. & Jebb M. H. P. 2013: Recircumscription of the *Nepenthes alata* group (*Caryophyllales*–*Nepenthaceae*), in the Philippines, with four new species. – *Eur. J. Taxon.* **69**: 1–23.
- Cheek M. & Jebb M. H. P., Murphy B. & Mambor F. 2018: *Nepenthes* section *Insignes* in Indonesia, with two new species. – *Blumea* **62**: 174–178.
- Ciardelli P., Kelbert P., Kohlbecker A., Hoffmann N., Güntsch A. & Berendsohn W. G. 2009: The EDIT Platform for Cybertaxonomy and the taxonomic workflow: selected components. – *Lecture Notes in Informatics (LNI)* **154**: 625–638.
- Clarke C. 1997: *Nepenthes* of Borneo. – Kota Kinabalu: Natural History Publications.
- Clarke C. 2001: *Nepenthes* of Sumatra and peninsular Malaysia. – Kota Kinabalu: Natural History Publications.
- Clarke C. & Moran J. A. 2011: Incorporating ecological context: a revised protocol for the preservation of *Nepenthes* pitcher plant specimens (*Nepenthaceae*). – *Blumea* **56**: 225–228.
- Clarke C., Schlauer J., Moran J. & Robinson A. 2018: Systematics and evolution of *Nepenthes*. – Pp. 58–69 in: Ellison A. & Adamec L. (ed.), *Carnivorous plants: physiology, ecology, and evolution*. – Oxford: Oxford University Press.
- COP 2002: COP 6 Decision VI/9 – Global Strategy for Plant Conservation. Conference of the Parties, Convention on Biological Diversity. – Published at <https://www.cbd.int/decision/cop/default.shtml?id=7183> [last accessed 26 Aug 2018].
- COP 2010: COP 10 Decision X/17 – Consolidated update of the Global Strategy for Plant Conservation 2011–2020. – Published at <https://www.cbd.int/decision/cop/?id=12283> [last accessed 26 Aug 2018].
- Danser B. H. 1928: The *Nepenthaceae* of the Netherlands Indies. – *Bull. Jard. Bot. Buitenzorg* **3**: 249–438.
- Dorr L. J. & Nicolson D. A. 2008–2009: Taxonomic literature: a selective guide to botanical publications and collections with dates, commentaries and types, ed. 2. Supplements **7**, **8**. – Ruggell: Gantner [= Regnum Veg. **149**, **150**].
- GBIF 2004+: Global Biodiversity Information Facility. – Published at <http://www.gbif.org> [accessed 26 Aug 2018].
- Gradstein S. R., Sauer M., Braun M., Koperski M. & Ludwig G. 2001: TaxLink, a program for computer-assisted documentation of different circumscriptions of biological taxa. – *Taxon* **50**: 1075–1084.
- Groom Q., Hyam R. & Güntsch A. 2017: Data management: stable identifiers for collection specimens. – *Nature* **546**: 33.
- Güntsch A., Hyam R., Hagedorn G., Chagnoux S., Röpert D., Casino A., Droege G., Glöckler F., Gödderz K., Groom Q., Hoffmann J., Holleman A., Kempa M., Koivula H., Marhold K., Nicolson N., Smith V. S. & Triebel D. 2017: Actionable, long-term stable and semantic web compatible identifiers for access to biological collection objects. – *Database (Oxford)* 2017: bax003.
- Güntsch A., Kohlbecker A., Müller A., Berendsohn W. G., Luther K. & Plitzner P. 2017a: EDIT Platform web services in the biodiversity infrastructure landscape. – *Biodivers. Inform. Sci. & Standards* **1**: e20363.
- Hagedorn G., Thiele K., Morris R. & Heidorn P. B. 2005: The Structured Descriptive Data (SDD) w3c-xml-schema, version 1.0. – Published at <https://github.com/tdwg/sdd> [accessed 25 Aug 2018].
- Hamann T. D., Müller A., Roos M. C., Sosef M. & Smets E. 2014: Detailed mark-up of semi-monographic legacy taxonomic works using FlorML. – *Taxon* **63**: 377–393.
- Henning T., Plitzner P., Güntsch A., Berendsohn W. G., Müller A. & Kilian N. 2018: Building compatible and dynamic character matrices – current and future use of specimen based character data. – *Botany Letters*, DOI: 10.1080/23818107.2018.1452791.
- Hernández-Ledesma P., Berendsohn W. G., Borsch T., Mering S. V., Akhani H., Arias S., Castañeda-Noa I., Eggli U., Eriksson R., Flores-Olvera H., Fuentes-Bazán S., Kadereit G., Klak C., Korotkova N., Nyffeler R., Ocampo G., Ochoterena H., Oxelman B., Rabeler R. K., Sanchez A., Schlumpberger B. O. & Uotila P. 2015: A taxonomic backbone for the global synthesis of species diversity in the angiosperm order *Caryophyllales*. – *Willdenowia* **45**: 281–383.
- Holetschek J., Dröge G., Güntsch A. & Berendsohn W. G. 2012: The ABCD of primary biodiversity data access. – *Pl. Biosyst.* **146**: 771–779.

- IPNI 2004+ [continuously updated]: The International Plant Names Index. The plant names project. – Published at <http://www.ipni.org> [accessed 26 Aug 2018].
- ISO 13166-1:2013 Codes for the representation of names of countries and their subdivisions – Part 1: country codes. – Geneva: International Organization for Standardization. – Published at <https://www.iso.org/standard/63545.html> [accessed 28 Aug 2018].
- Jebb M. H. P. & Cheek M. 1997: A skeletal revision of *Nepenthes*. – *Blumea* **42**: 1–106.
- JSTOR 2000+ [continuously updated]: JSTOR. – New York: ITHACA. – Published at <http://www.jstor.org/> [last accessed 26 Aug 2018].
- JSTOR 2013+ [continuously updated]: JSTOR Global Plants. – New York: ITHACA. – Published at <https://plants.jstor.org/> [accessed 26 Aug 2018].
- Karam N., Müller-Birn C., Gleisberg M., Fichtmüller D., Tolksdorf R. & Güntsch A. 2016: A terminology service supporting semantic annotation, integration, discovery and analysis of interdisciplinary research data. – *Datenbank-Spektrum* **16**: 195–205.
- Kilian N., Hand R. & Raab-Straube E. von (general ed.) 2009+ [continuously updated]: *Cichorieae* Systematics Portal. – Published at <http://cichorieae.e-taxonomy.net/portal/> [accessed 26 Aug 2018].
- Kilian N., Henning T., Plitzner P., Müller A., Güntsch A., Stöver B. C., Müller K. F., Berendsohn W. G. & Borsch T. 2015: Sample data processing in an additive and reproducible taxonomic workflow by using character data persistently linked to preserved individual specimens. – *Database* **2015**: bav094.
- Kohlbecker A., Müller A., Berendsohn W. G., Güntsch A., Luther K. & Plitzner P. 2017: The EDIT Platform for Cybertaxonomy, a brief overview. – *Biodivers. Inform. Sci. & Standards* **1**: e20368.
- Kohlbecker A., Müller A. & Hoffmann, N. 2008: EDIT Data Portal – A flexible publishing tool for biodiversity information networks. – *Proceedings TDWG* **2008**: 69–70. – Published at https://static.tdwg.org/conferences/2008/tdwg_2008_proceedings.pdf [accessed 26 Aug 2018].
- Loizeau P.-A. & Wyse Jackson P. 2017: World Flora Online mid-term update. – *Ann. Missouri Bot. Gard.* **102**: 341–346.
- McPherson S. 2009: Pitcher plants of the Old World. – Poole: Redfern Natural History Publications.
- McPherson S. 2011: New *Nepenthes* 1. – Poole: Redfern Natural History Publications.
- Meimberg H. & Heubl G. 2006: Introduction of a nuclear marker for phylogenetic analysis of *Nepenthaceae*. – *Pl. Biol.* **8**: 831–840.
- Meimberg H., Thälhammer S., Brachmann A. & Heubl G. 2006: Comparative analysis of a translocated copy of the *trnK* intron in the carnivorous genus *Nepenthes* (*Nepenthaceae*). – *Molec. Phylogen. Evol.* **39**: 478–490.
- Müller A., Berendsohn W. G., Kohlbecker A., Güntsch A., Plitzner P. & Luther 2017: A comprehensive and standards-aware Common Data Model (CDM) for taxonomic research. – *Biodivers. Inform. Sci. & Standards* **1**: e20367.
- Parr C., Wilson N., Leary P., Schulz K., Lans K., Walley L., Hammock J., Goddard A., Rice J., Studer M., Holmes J. & Corrigan Jr. R. 2014: The Encyclopedia of Life v2: providing global access to knowledge about life on earth. – *Biodivers. Data J.* **2**: e1079.
- Plitzner P., Henning T., Müller A., Güntsch A., Karam N. & Kilian N. 2017: Bottom-up taxon characterisations with shared knowledge: describing specimens in a semantic context. – *CEUR Workshop Proceedings* 1933. – Published at <http://ceur-ws.org/Vol-1933/paper-4.pdf> [accessed 28 Aug 2018].
- Plitzner P., Müller A., Güntsch A., Berendsohn W. G., Kohlbecker A., Kilian N., Henning T. & Stöver B. 2017a: The CDM applied: unit-derivation, from field observations to DNA sequences. – *Biodivers. Inform. Sci. & Standards* **1**: e20366.
- Roca P., Sastre P., Lobo J., Meganck B., Theeten F., Mergen P., Müller A., Kohlbecker A., Dusan S. & Mikiewicz D. 2009: Enhancing the visualization of biological data: EDIT geographic tools. – *Proceedings TDWG* **2009**: 12–13.
- Roos M. C., Berendsohn W. G., Dessein S., Hamann T., Hoffmann N., Hovenkamp P., Janssen T., Kirkup D., de Kok R., Sierra S. E. C., Smets E., Webb C. & van Welzen P. C. 2011: e-Flora Malesiana: state of the art and perspectives. – *Gard. Bull. Singapore* **63**: 189–195; <http://biostor.org/reference/140281> [last accessed 26 Aug 2018].
- Roskov Y., Abucay L., Orrell T., Nicolson D., Bailly N., Kirk P. M., Bourgoin T., DeWalt R. E., Decock W., De Wever A., Nieukerken E. van, Zarucchi J. & Penev L. (ed.) 2018: Species 2000 & ITIS Catalogue of Life, 2018 annual checklist. DVD. – Leiden: Species 2000, Naturalis.
- Schwallier R., Raes N., de Boer H. J., Vos R. A., van Vugt R. R. & Gravendeel B. 2015: Phylogenetic analysis of niche divergence reveals distinct evolutionary histories and climate change implications for tropical carnivorous pitcher plants. – *Diversity & Distrib.* **22**: 97–110.
- Sharrock S. 2012: GSPC – Global Strategy for Plant Conservation – A Guide to the GSPC. – Richmond: Botanic Gardens Conservation International. – Published at http://www.plants2020.net/files/Plants2020/popular_guide/englishguide.pdf [accessed 26 Aug 2018].
- Stafleu F. A. & Cowan R. S. 1976–1988: Taxonomic literature: a selective guide to botanical publications and collections with dates, commentaries and types, ed. 2. Volumes 1–7. – Utrecht/Antwerpen: Bohn, Scheltema & Holkema; The Hague/Boston: dr. W. Junk b.v. [= *Regnum Veg.* **94**, **98**, **105**, **110**, **112**, **115**, **116**].

- Stafleu F. A. & Mennega E. A. 1992–2000: Taxonomic literature: a selective guide to botanical publications and collections with dates, commentaries and types, ed. 2. Supplements 1–6. – Königstein: Koeltz Scientific Books [= *Regnum Veg.* **125**, **130**, **132**, **134**, **135**, **137**].
- TPL 2013: The Plant List. Version 1.1. – Published at <http://www.theplantlist.org/> [accessed 26 Aug 2018].
- Tropicos, 1991+ [continuously updated]: Missouri Botanical Garden Tropicos database portal. – Published at <http://www.tropicos.org/> [accessed 26 Aug 2018].
- Ung V., Dubus G., Zaragüeta-Bagils R. & Vignes-Lebbe R. 2010: Xper²: introducing *e*-taxonomy. – *Bioinformatics* **26**: 703–704.
- Venin M., Kirchhoff A., Fradin H., Güntsch A., Hoffmann N., Kohlbecker A., Kuntzelmann E., Maiocco O., Müller A., Vignes Lebbe R. & Berendsohn W. G. 2010: Descriptive data in the EDIT Platform for Cybertaxonomy. – Pp. 7–11 in: Nimis P. L. & Vignes-Lebbe R. (ed.), *Tools for identifying biodiversity: progress and problems*. – Trieste: EUT Edizioni Università di Trieste. – Also published at <http://hdl.handle.net/10077/3738> [accessed 26 Aug. 2018].
- WFO 2017+: World Flora Online guidelines for contributors. – Published at <http://about.worldfloraonline.org/wfodocuments.shtml> [accessed 26 Aug 2018].
- Wieczorek J., Döring M., De Giovanni R., Robertson T. & Vieglaes D. 2009: Darwin Core terms: a quick reference guide. – Published at <http://rs.tdwg.org/dwc/terms/index.htm> [last accessed 26 Aug 2018].
- Wyse Jackson P. & Miller J. S. 2015: Developing a World Flora Online: A 2020 challenge to the world's botanists from the international community. – *Rodriguésia* **66**: 939–946.

Willdenowia

Open-access online edition www.bioone.org/loi/will



Online ISSN 1868-6397 · Print ISSN 0511-9618 · Impact factor 1.500

Published by the Botanic Garden and Botanical Museum Berlin, Freie Universität Berlin

© 2018 The Authors · This open-access article is distributed under the CC BY 4.0 licence