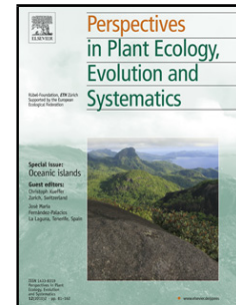


Accepted Manuscript

Title: An integrative and dynamic approach for monographing species-rich plant groups – building the global synthesis of the angiosperm order Caryophyllales

Author: Thomas Borsch Patricia Hernández-Ledesma Walter
G. Berendsohn Hilda Flores-Olvera Helga Ochoterena
Fernando O. Zuloaga Sabine von Mering Norbert Kilian



PII: S1433-8319(15)00045-1
DOI: <http://dx.doi.org/doi:10.1016/j.ppees.2015.05.003>
Reference: PPEES 25273

To appear in:

Received date: 16-1-2015
Revised date: 17-4-2015
Accepted date: 3-5-2015

Please cite this article as: Borsch, T., Hernández-Ledesma, P., Berendsohn, W.G., Flores-Olvera, H., Ochoterena, H., Zuloaga, F.O., von Mering, S., Kilian, N., An integrative and dynamic approach for monographing species-rich plant groups ndash building the global synthesis of the angiosperm order Caryophyllales, *Perspectives in Plant Ecology, Evolution and Systematics* (2015), <http://dx.doi.org/10.1016/j.ppees.2015.05.003>

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

An integrative and dynamic approach for monographing species-rich plant groups – building the global synthesis of the angiosperm order Caryophyllales

Thomas Borsch^{1,2*}, Patricia Hernández-Ledesma¹, Walter G. Berendsohn^{1,2}, Hilda Flores-Olvera³, Helga Ochoterena³, Fernando O. Zuloaga⁴, Sabine von Mering¹, Norbert Kilian¹

¹ Botanischer Garten und Botanisches Museum Berlin-Dahlem, Dahlem Centre of Plant Sciences, Freie Universität Berlin, Königin-Luise Str. 6-8, 14195 Berlin, Germany; t.borsch@bgbm.org; phl@st.ib.unam.mx; w.berendsohn@bgbm.org; s.vonmering@bgbm.org

² Berlin-Brandenburg Institute of Advanced Biodiversity Research (BBIB), Altensteinstr. 6, 14195 Berlin, Germany

³ Departamento de Botánica, Instituto de Biología, Universidad Nacional Autónoma de México (UNAM), Apartado Postal 70-233, C.P. 04510 México, D.F. México; mahilda@ib.unam.mx; helga@ib.unam.mx

⁴ Instituto de Botánica, Darwinion, Argentina; fzuloaga@darwin.edu.ar

*Corresponding author:

Professor Dr. Thomas Borsch
Botanischer Garten und Botanisches Museum Berlin-Dahlem
Dahlem Centre of Plant Sciences
Königin Luise-Str. 6-8
14195 Berlin
Germany

+49-30-838-50132 (Phone)

+49-30-838-4-50133 (Fax)

t.borsch@bgbm.org

Abstract

One of the major goals of systematics is to provide a synthesis of knowledge on the diversity of a group of organisms, such as flowering plants. Biodiversity conservation and management call for rapid and accurate global assessments at the species level. At the same time the rapid development of evolutionary biology with a spectrum of approaches to test species relationships and species limits, has revolutionized and is still revolutionizing the science of plant systematics including taxonomy. We explore the relevant scientific and technological developments, ~~but also organizational and logistic requirements,~~ with the aim to suggest a conceptual framework for an integrated monographic synthesis ~~of species diversity~~ which can reach global coverage. ~~Our exemplar group are the Caryophyllales, as an example of a globally distributed~~ which are a plant lineage of worldwide distribution, comprising approx. 5% of flowering plant species diversity. The current situation of ~~organism~~-classification is marked by a transition from pre-phylogenetic treatments to taxonomic treatments increasingly evaluated in an evolutionary context. Structured data (both molecular and morphological), linked to well-documented specimens will be important as fundamental entities of information that can be subjected to evolutionary analysis. As a result, taxon concepts are established as hypotheses which then can be used as basis for a classification system in a second step. ~~The process of generating knowledge and subsequent classification of the recognized entities is step by step, including processes of reciprocal illumination.~~ Global syntheses need to provide information and use a classification system that reflects the current state of knowledge. ~~they need to be dynamic.~~ In order to accommodate the constantly improved understanding of the organisms, eventually also resulting in the change of taxon concepts, the treatments need to be dynamic. The workflow for a global monographic synthesis as outlined here is supported by currently available biodiversity informatics tools such as the EDIT Platform for Cybertaxonomy. ~~Based on these conceptual findings we outline the practical implementation steps towards a synthesis of the Caryophyllales, as an example of a globally distributed plant lineage, comprising approx. 5% of flowering plant species diversity. While the actual~~ The availability of electronic sources (names, protologues, type images, literature) greatly facilitates the access to information, but as our case shows, considerable efforts for data curation and research are still needed. ~~T~~he implementation of a global monographic synthesis such as the Caryophyllales requires the involvement of the global scientific community.

Keywords: taxonomic workflow, biodiversity informatics, species concept, botanical monography, monographing, systematics, classification

Introduction

Synthesizing the current knowledge on lineages of organisms is still a major goal of systematic biology. This is normally achieved through monographs, which usually focus on the genus level and provide access to the known diversity of species. Monographs make explicit statements on taxonomic concepts and summarize the history of classification of lineages. In this way they deliver correct nomenclature including accepted names and synonyms and guide us through the previous literature where often differing taxon concepts were applied. Monographs serve us with comparative data on the phenotype, increasingly also on the genotype (molecular data), distribution, ecology, and provide means to identify species and infraspecific taxa. Identification keys combined with thorough descriptions and illustrations not only enable the user to decide from a selection of characters and character states to which species an individual belongs but also help to recognize hitherto undescribed taxa.

Studies of species diversity have been revolutionized by phylogenetics. Technological developments in DNA sequencing yield the data sets that allow reconstructing trees and networks of evolutionary relationships. Still, in most studies the coverage is limited for practical reasons to a selection of individuals, constituting a sample of species, to infer a representative phylogeny of the group. But increasingly high coverage of species is being achieved due to methodological advances and decreasing costs per base sequenced (see, e.g. Mansion et al., 2012).

It is evident, however, that the establishment of molecular methods in systematic biology significantly affects monographing. ~~In the long run,~~ factual integration of molecular (genetic) and morphological (phenotypic) characters for an individual or taxon will be an essential task of any species diagnosis (González et al., 2013) and any monograph. Also other kinds of data in the “omics” age will need to be assigned to well-documented individual specimens representing species. Therefore, efficient approaches to synthesize this knowledge in truly integrative monographs are needed. It is needless to say that unambiguous names for clearly and transparently described biological entities (species) are fundamental to all fields of biology as well as biodiversity conservation and use (Patterson et al., 2010; Hardisty and Roberts, 2013). On the other hand, the question of which species concept would apply to a group of organisms has been intensely discussed (~~De Queiroz, 2005;~~ De Queiroz and Donoghue, 2007; Goldstein and DeSalle, 2010; Mallet, 2007; McDade, 1995), and some evolutionary biologists even demurred from producing classifications because they would have had to make a compromise. We argue that the approach can be more pragmatic without ~~loose~~inig scientific ~~scrutiny~~precision.

In flowering plants, for example, species may either be monophyletic (in line with a species concept that proposes to only recognize monophyletic entities as species; e.g., Donoghue, 1985) or at least represent lineages with common origin (in line with a phylogenetic species concept that proposes to recognize entities that are the result of phylogenetic history; Nixon and Wheeler, 2008). There may be cases of incomplete lineage sorting (e.g. Gurushidze et al.,

2010; Flores-Rentería et al. 2013), resulting in paraphyletic assemblages of morphologically very similar individuals while single populations out of the common gene pool may have adapted and phenotypically diversified with new traits (consistent with the phylogenetic species concept, but not with the monophyletic species concept). In any case, (almost) all species concepts can be evaluated with phylogenetic methods (as it is the case also for the biological or typological species concepts, Mayr 1969), but in the case of plants, non-hierarchical speciation has also been documented. For example, hybrid and allopolyploid speciation is in fact frequent (e.g. Kim et al., 2008; Rieseberg et al., 1990; Soltis and Soltis, 2009), even among [more](#) distantly related genera. Therefore the biological species concept is not always applicable, but thanks to the application of molecular evolutionary methods, this [type of](#) species origin can now be inferred with high levels of certainty. Our approach to describe species is to take the best available knowledge on the evolutionary history of a species, including continuities and discontinuities in the distribution of characters and character states among individuals and populations, as a basis to formulate a taxonomic concept. In this sense, biologically different mechanisms that have led to speciation are reflected in the taxonomic concept, and the purpose of species classification is to make this transparent. However, we do not argue that the criterion of monophyly for classifying higher level taxa (Hennig, 1966) should be questioned. We rather want to stress that a consistency of concepts of classification that spans from above to below the species level may not be appropriate or even pragmatic. Thus, non-dichotomous speciation histories related to complex patterns of gene flow can also be formulated in respective taxon concepts, which then will receive a name.

A further challenge for monographing is the constant accumulation of new knowledge. Reconstructing the tree of life has successfully illuminated the overall relationships of major groups such as flowering plants, providing the base for classifying monophyletic entities at the level of orders and families (APG III, 2009). Considering the genus level, the coming years and even decades will yield a wealth of new insights. [Like in other large groups such as the \(e.g. for grasses; \(Vorontsova and Simon, 2012\) this is also the case in Caryophyllales ; eacti; \(e.g., Nyffeler and Eggli, 2010a-b; Fuentes-Bazan et al., 2012; Dillenberger and Kadereit, 2014; -Hernández-Ledesma et al., in prep.\).](#) And the same will happen at the species level, but probably much slower in terms of total coverage. [We are therefore in a transition-phase from pre-phylogenetic to evolution-based taxon concepts and classifications.](#) Apart from clear and stable rules in nomenclature, dynamic approaches are needed that can efficiently handle changing views on the concepts for taxa and which easily allow the incorporation of new knowledge into existing treatment frameworks. Traditional hard copy monographs by mostly individual authors, which take years or even decades to complete but become quickly outdated, are no longer an adequate means of synthesizing current knowledge on lineages of organisms.

And then there is the enormous task of comparing organisms from different geographical parts of the world in order to unravel their evolutionary history, for which a representative set of characters needs to be compared in order to understand what species are and to evaluate alpha-taxonomic concepts. As a basis for this task, international collaboration has to be established. Piles of literature, often in a variety of languages, need to be evaluated. Collections, built upon centuries of exploration, have to be worked through and supplemented by new collecting activities and fieldwork to answer specific questions. And laboratory work has to be organised that will efficiently include representative sets of samples and will keep up with methodological advances. Often it takes years before the necessary material and a team of workers to analyse different types of characters are brought together. Nevertheless, the individual specialist is comparable to a historian who has to be specialized on an epoch, a

country or a social group in order to connect the knowledge into a profound synthetic treatment.

Under these circumstances, the question arises how a monographic project can be efficiently organized? How can the data, gathered with considerable investment of resources, be collected, organized and stored in a dynamic, highly effective as well as sustainable way, which optimises their future use in updated synthesizing approaches, as well as making them available for practical purposes? While the work is becoming increasingly facilitated by electronic access to type collections (e.g. JSTOR Global Plants, JSTOR 2013-), historical literature (e.g. Biodiversity Heritage Library, BHL 2005-) and specimens and species occurrence observations in general (e.g. Global Biodiversity Information Facility, GBIF 2014-) the complexity of the task of preparing an integrative biological monograph still remains.

Despite this enormous complexity, the products of this research are urgently needed. Societies demand reliable and accessible knowledge about organisms in order to efficiently deal with challenges in conservation and sustainable use of biodiversity. This makes a need for efficient approaches to cover as much species diversity as possible in modern monographs. The need for taxonomic research as a basis has been repeatedly expressed (e.g. the Global Taxonomy Initiative, CBD 1998-). It is the only source of data for assessing the extinction risk of plants (Von Staden et al., 2013) and other organisms. It is of utmost importance to consider the discovery, collection, storing in collections, study and description of specimens and tissues of the still unknown species of the biosphere, before they turn extinct, as an absolute priority for biology in our century of extinctions (Dubois, 2010). As a consequence there are several proposals to revolutionize and streamline taxonomic research, using new technologies (Godfray, 2002; Mallet and Wilmot, 2003; Agnarsson and Kutner, 2007; Godfray et al., 2007; Mayo et al., 2008).

Both a lack of knowledge or confusion about species limits can pose problems for conservation (Mace, 2004). A case study on Juncaceae and Potamogetonaceae indicated, for example, that accurate monographing revealed up to 25% of taxa categorized in the IUCN red lists representing synonyms or otherwise doubtful entities (Kirschner and Kaplan, 2002), a percentage that poses serious challenges to the efficiency of conservation measures and investments. More recently, the Global Strategy of Plant Conservation (GSPC) as a programme of the United Nations Convention on Biological Diversity (CBD, 2006) with 16 practical targets, manifested the need for an “online flora of all known plants by 2020” (CBD-SBSTTA, 2012). While all these facts, needs, and arguments are well appreciated, precise conceptual work to develop the implementation of larger taxonomic treatments and monographic syntheses is still lacking.

A global treatment of plant species diversity is essentially not a new call – there have been several initiatives such as the Species Plantarum project (Brummitt et al., 2001) aiming at monographic coverage, Encyclopedia of Life (EOL, 2014) using an aggregator approach with existing online resources, and attempts to create global checklists, as the result of database merging (The Plant List, TPL, 2013) or in a federated approach (~~Species 2000~~, Roskov et al., 2014). However, neither of these initiatives has progressed to monographically treat a significant amount of global plant diversity. In parallel, however, several network initiatives became quite successfully established in their respective communities, providing global up-to-date treatments for all species of those groups through electronic permanently curated portals (e.g., Brassicaceae – Koch et al., 2012; Cichoreae Portal - ICN, 2009-; Arecaceae – PalmWeb, 2010-; Solanaceae - Solanaceae Source, 2011-; *Campanula* - Campanula Portal 2013-). Also, there are project websites on certain groups of plants that aim to organise

taxonomic information and coordinate research (e.g. Euphorbiaceae - Rina and Berry, 2014, Sileneae – Oxelman et al., 2013).

Most of the current technical reports and opinion papers focus on either the development of information standards, data base and portal systems (Berendsohn et al., 2011; Holetschek et al., 2012) or the information needs for conservation (Paton, 2009; Nic Lughadha and Miller, 2009; Hardisty and Roberts, 2013). Only few studies address the taxonomic or monographic workflow as such (e.g., Stuessy and Lack, 2011; Marhold and Stuessy, 2013). Our motivation is to bridge between these issues in light of the recent, enormous advances in the study of phylogeny and speciation of plants that is fuelled by methodological advances in the genomics age. And we feel the importance to also consider developments in the scientific community that govern the generation of knowledge that then will be synthesized in monographic endeavours. In this paper we aim at (1) reviewing the current perspective for doing large scale monographs in an interdisciplinary and international environment, and (2) developing a perspectivework-flow for implementing a global synthesis of the angiosperm order Caryophyllales as a model of the angiosperms.

~~More precisely, this paper aims at (1) presenting a conceptual framework for an integrative monograph, (2) discussing the logistics of doing large scale monographs in an interdisciplinary and international environment, and (3) developing a perspective for a global synthesis of a group of angiosperms. We will do this through outlining more practical implementation steps using the Caryophyllales as a model in the third part of this paper.~~

Logistics of Rationale for large scale syntheses: Setting the base for the approach in Caryophyllales and integrative monographing

The step-wise approximation in classifying and naming biological entities

~~Since there~~In an ideal world, evolutionary relationships and species limits will have been investigated including all types of data (morphology, anatomy, embryology, palynology, cytology, secondary compounds, DNA sequences; see Stuessy, 1990), using a comprehensive sampling of all taxa so far identified and of populations representing the whole range of putative species, and using a suite of inference methods ranging from tree and network reconstruction to the population-level modelling of shifts in allele frequencies and gene flow. However, in the real world this is not the case. Data sets and taxon sampling will be complemented over time by usually different workers, leading to~~is~~ a step-wise gain in knowledge on the evolutionary history and diversity of organisms. As a consequence, their classification is also improving in a step-wise manner.

~~Depending on the state of work in different a groups of plants~~In more practical terms~~the real world including the Caryophyllales this implieswe find~~ that (a) there are different concepts of species and genera readily available, (b) ~~they~~that concepts are differently used by different user groups, (c) ~~and that~~ some of them are affected by rather rapid knowledge turnover. Scientific names need to be applied to such recognized entities even at preliminary stages where our knowledge is far from complete. This is necessary because taxonomy/systematics does not operate in a vacuum – concrete uses and other traits (e.g. conservation status,

invasiveness, role in landscape and ecosystem) are linked to these entities and are widely applied.

The ~~issue-need~~ of managing (different) taxonomic concepts has ~~already~~ been pointed out (Berendsohn, 1995; Franz and Peet, 2009 and Franz and Cardona-Duque, 2013). The rank-issue (“mandatory categories”, De Queiroz and Gauthier, 1992) advocated by phylocodists is only a small part of this problem [e.g. Moore, 2003; but see the entire special issue of *The Botanical Review* 69(1)]. In most cases where new evidence implies changes, these have to be either in the concept (content) or in the name. There is no benefit in maintaining a name stable if its content is dramatically changed. On the contrary, names have the purpose of communication and should therefore be as stable as possible with respect to their content.

It is evident that there have to be clear rules on how to translate evolutionary insights, and as a consequence taxonomic concepts into a formal system of nomenclature (e.g. McNeill et al., 2012). The objective of this paper is not to discuss the philosophy behind differing approaches such as a “Linnaean”, hierarchical, rank-based nomenclature on the one hand (see Turland, 2013) and a phylogenetic nomenclature (sensu PhyloCode, Cantino and Queiroz, 2010). We do believe, though, that names of organisms have to facilitate unambiguous communication of taxonomic concepts which will not be achieved through applying the PhyloCode. Rules for describing these concepts should enable us to simplify and categorise the detailed hypothesis on the evolutionary history of a group of organisms that led to the formulation of the concept.

For our synthesis of the Caryophyllales, we agree with the view that monophyly is an important criterion to make higher-level taxon concepts (i.e. above the level of species such as genera and families) as predictive as possible. Using current phylogenetic techniques, monophyly can be established with high degrees of confidence in a macroevolutionary context. At the level of species and below, where evolutionary histories are much more complex (microevolution) and cannot be necessarily described by dichotomous patterns, we should explicitly state what the concept for a species as classified includes: (a) an assemblage of populations that all go back to a common ancestor; (b) an assemblage of populations that represents a partial spectrum of ancestral genotypes so that the species is paraphyletic (it is often very difficult to distinguish between incomplete lineage sorting or hybrid origin); (c) populations that go back to one or several polyploidization events (inferred from molecular data); (d) an assemblage of populations that go back to one or several hybridization events without genome duplication. There is no theoretical justification to apply the criterion of monophyly as universal base for species-level and higher rank classification (Rieppel, 2009) because the distinctness of a species as a biological entity can be explicitly formulated as a hypothesis to entail a defined group of individuals. As Rieppel (2009) put it: “Monophyletic groups (clades) stand in phylogenetic relations to each other, constituents of species stand in tokogenetic relations to each other”. Biologically meaningful assessments of species relationships and species limits therefore require a comprehensive approach (see e.g. Naciri and Linder 2015, for a recent review). This also means that in order to be reproducible, the recognized entities and the respective taxon concepts have to be based on explicit sets of-character data (both molecular and morphological) linked to individual specimens.

Considering ~~that in that~~ a stepwise process there will be discrepancies in circumscriptions of species envisaged by different authors, a responsible taxonomist should procure the least number possible of taxonomic changes in taxon concepts in light of new evidence. Therefore, an important issue is the economy of change in classification systems procuring nomenclatural stability (Nixon et al., 2003). For instance, different philosophical criteria, such as reflecting monophyly in a classification, have been the cause of profound supraspecific changes in classifications. Only accepting monophyletic taxa could still result in contrasting

opinions on supraspecific taxa. Therefore, changes should also be based on an expert view on clade stability (in light of the reliability of existing phylogenetic hypotheses to try to depict what it is considered “the true organismal phylogeny” or at least the best supported hypothesis) and of phenotypic diagnosability (e.g. Vences et al., 2013).

In spite of this, many of the current changes in classification do not consider economy of nomenclatural changes. Conservation of names is a currently seldom-used resource to promote name stability in this context. An analysis of stability in both concepts (species circumscription) and names of vascular plants and mosses in Germany (Berendsohn and Geoffroy, 2007) concluded that there is a considerable instability in a substantial proportion of plant taxa, even among works in current use.

~~A significant proportion of publications presenting the results of phylogenetic analyses does not translate their results into any classification system, a phenomenon that has been addressed as the “phylogeny classification gap” (Franz, 2005). While some authors think that phylogenetic trees are best communicated directly, it was convincingly argued how important classification systems are that attach easily remembered names to organisms for the vast majority of users (Franz 2005). Another quality of a “phylogeny classification gap”—practical limitation, is that most current phylogenetic studies do not provide the needed data to support a species level classification, mostly due to limitations in taxon sampling. Thus, so far only the first rounds of evolutionary knowledge gain have been implemented.~~

Comment [A1]: The idea is still presented but in the context of the Caryophyllales work program (see p. 23).

Standards and tools for managing integrative concept-based classification systems

Classifications and species circumscriptions evolve with the growth of knowledge ~~about the natural world~~. In taxonomy and elsewhere in science, there is a fundamental difference between information provision by information systems and information provision directly mediated by human beings (normally specialists). The practical value of being able to name and identify a biological entity lies in the provision of an indexing system for knowledge about these entities, which allows integrating information from different sources on different topics over large periods of time. In traditional printed publications, the integration of information has been mediated by specialists who are able to qualify the relationships between the named and circumscribed taxonomic concepts. In contrast, in information managed by computer systems and published on the Internet, the possibility to automatically link information on taxa from various sources using the scientific name is increasingly used. Given the uncertainties stated above, without expert validation this poses a high risk of misinformation.

~~As recognized already two decades ago,~~ In order to come to terms with this problem public information systems need integrated explicit knowledge to reliably transmit information linked to taxonomic concepts (in this case: named biological entities; Beach et al., 1993; Berendsohn, 1995). Information models were developed ~~subsequently~~, providing the theoretical base for handling taxonomic concepts (e.g. Zhong et al., 1996; Berendsohn, 1997; Pullan et al., 2000), followed by software development that demonstrated the practicability of such models for taxonomic data (e. g. Pullan et al., 2000; Gradstein, 2001; Berendsohn et al., 2003; Berendsohn and Geoffroy, 2007). In parallel, a number of projects drove standardisation of data items, especially in the context of the organisation for Biodiversity Information Standards (TDWG, formerly the Taxonomic Databases Working Group). Examples include the Taxonomic Concept Transfer Schema (TCS; Kennedy et al., 2006), Structured Descriptive Data (SDD; Hagedorn et al., 2005), Access to Biological Collection

Data (ABCD; Berendsohn, 2007-2010, Holetschek et al. 2012) and Darwin Core (DwC; Wiczorek et al., 2009). This development culminated in the definition of an object-oriented Common Data Model for taxonomic information (CDM) that was devised for and forms the basis of the EDIT Platform for Cybertaxonomy (Berendsohn, 2010; Berendsohn et al., 2011).

The EDIT Platform is a collection of more or less tightly coupled tools especially devised for taxonomists, incorporating the entire scope of nomenclatural, taxonomic, descriptive and geographic information contained in the products of taxonomic work (monographs, treatments for faunas and floras, and checklists; Berendsohn, 2010; Venin et al., 2010). The EDIT Platform is natively concept-oriented. Therefore, the system allows applying a name to deviating taxonomic entities, all of which include the type specimen but are more or less inclusive. Concept relationships can be made explicit, stating that the circumscriptions according to the first concept is congruent to, overlapping, included in or includes the second one. Alternative classification systems can be shown (e.g. the alternative classifications for *Hieracium* and *Pilosella* in the *Cichorieae* portal; ICN, 2009-). It is easy to adapt the classification to a new concept once there are new research results, without having to laboriously re-enter information (protologues, images, etc.) used previously. [This will also be of considerable importance in the synthesis of Caryophyllales.](#) Available Platform tools include those for local and on-line data input/editing and on-line and print data output (Ciardelli & al., 2009). Import and export interfaces handle files that adhere to community data standards such as ABCD, DwC, SDD, and TCS. The Xper2 software (Ung et al., 2010), which can also be used as a stand-alone tool, handles descriptive information, i.e. character/character state lists that can be augmented by illustrations and used for keys and species descriptions as well as exported as a matrix into NEXUS (Maddison et al., 1997) for phylogenetic analysis. Descriptive information may refer to taxa or to specimens, or to taxa synthesising information from taxa of a lower rank or from specimens (Kilian ~~et~~ al., submitted). Nomenclatural details are covered in full for botany and zoology, taking into account the often very specific requirements of taxonomists and taxonomic publications. On the technical side, the EDIT Platform is independent of the operating system and database management system used, and a full range of machine-accessible web services (Booth et al., 2004) provide the base for full interoperability e.g. in workflow environments, as well as connectivity with the global biodiversity informatics infrastructures. This is exemplified in the Biodiversity Virtual e-Laboratory (BioVeL) project using the taxonomic data of the Catalogue of Life (Mathew et al., 2014).

We posit that for a truly integrative approach to taxonomic information a system like this is necessary, combining the existing rigorous traditions in handling data from taxonomic research and nomenclature with the possibility of concept handling while remaining open for the advances that are currently being made in Internet technologies.

Modern monographs will be produced by teams rather than individuals

~~It is evident that especially large genera are in need of revision (e.g. von Staden et al., 2013), in particular those distributed across wider geographic areas and spanning several to many countries. Teamwork and international collaboration in connection with current approaches in phylogenetics offer a great potential to tackle this problem, with an initial step to break up taxa with hundreds of species into more workable smaller monophyletic study groups (Mansion et al., 2012). At the same time the inclusion of multiple individuals from geographically different parts of a species' range becomes feasible.~~

Nevertheless, while an increase in new species descriptions on an international scale can be observed (Joppa et al., 2011), there is a decline in revisionary work exemplified by the number of taxa treated per researcher (Von Staden et al., 2013; Joppa et al., 2011). This may be explained by several factors: it is quicker to describe new species than to work on taxonomic concepts for all species in a group, and it is easier to just generate and selectively compare the data needed to justify the description of a new species than to comprehensively study a whole clade. Revisions require access to a wide spectrum of materials and data which is much more difficult to achieve and is best done through a concerted effort.

Current evaluation criteria for researchers' performance are mostly focused on the impact factor (IF) of the journals chosen for publication, leaving in a second level considerations about the time required to produce results. Monographic work is always laborious and time consuming and rarely publishable in journals with high impact factors. Therefore, in order to promote revisionary work, it is essential that institutions adopt local and inter-institutional policies committed to contributing research on basic taxonomy. Interdisciplinary collaboration among researchers who are capable of analysing data with modern tools, but who support and value the knowledge provided by alpha taxonomists can lead to integration of the taxonomic information in such a way that its analyses becomes publishable in journals of high impact (e.g. Alberts, 2013). A vision of leading researchers, on the one hand recognizing the synergies created by teamwork of specialists in different fields as well as the explicit willingness to value the contributions to the research results at all levels is essential for such an integration. The recent trend of some journals to describe the individual contribution of each author is a good advance in promoting team spirit. Within the taxonomic community itself, teamwork of specialists in different subdisciplines such as molecular phylogenetics, anatomy etc. as well as regional experts knowledgeable of the local flora is necessary. Common appreciation in the taxonomic community is needed to recognize the value of multi-authored studies, but consequently also the need for proper inputs by each partner.

If different disciplines and geographic regions are to be united in the common goal of providing an integrated treatment of a taxonomic group, heterogeneity of knowledge will be an issue, not only in terms of the degree of knowledge of a region or a group, but also of experience of the participants. For instance, phylogenetic hypotheses do not exist for all genera, nor are there molecular data for all species. Also, not all local taxonomists have access to molecular tools while others are not aware of the presence of samples in regional collections or in specific localities. Overarching interdisciplinary monographic work should therefore pragmatically aim at the compilation of available information as a first step, thus establishing the degree of knowledge and priorities for future research on a group or a region. Because available information will vary among regions and groups, three options are available for monographic work: wait until complete information is available for all groups and the entire distributional range, discard valuable information if it is not available for all, or to accept monographs with heterogeneous content. Given the need for taxonomic information, the first case is not acceptable. In our opinion, the third case is rendered acceptable by explicitly stating the lack of information, thus serving also as a primary source for detecting research priorities.

Gathering, integrating and analysing with the same depth different types of information needs teamwork. Therefore, a team built to integrate available monographic knowledge should ideally include a coordinator with a broad experience willing to facilitate communication; geographically restricted (floristic) alpha taxonomists, taxonomic specialist(s) in a group who are also specimen curators, phylogenetic (molecular) systematists, and also experts in

~~biodiversity informatics to connect existing data to the workflow. The roles of each team member must be defined, understood, agreed and communicated in the team. Among the team members, the role of the taxonomic specialist is key (i.e. the researcher who specializes in maintaining an overview on all species concepts in a study group such as a genus) to the success of these goals because these persons are likely to provide conceptual and practical leadership to orient the research towards proper sampling, key questions, gaps in knowledge, etc.~~

Continuous actualization of knowledge and classification through involvement of the scientific community and the use of modern information technology

The beginning of the 21st century has seen important steps towards realising teams working on integrative monographs, often facilitated or even incentivized by the technical advances in information sharing, networking and new forms of publishing (Marhold and Stuessy, 2013). ~~Consortia and networks have formed to actually progress with research on large genera or families in a collaborative way. These may or may not include smaller integrative projects for PhD students as advocated by Marhold and Stuessy (2013) and they may (and should) include links to ongoing regional (Flora) work. In the following we will have a more detailed look at those initiatives already existing on some plant families. This should help to arrive at the most~~ consolidate a sensitive approach for large scale syntheses of diverse plant groups such as the Caryophyllales.

~~During recent years several teams started out to summarize knowledge on flowering plant families using information systems and As one of the first projects using internet technologies. For example,~~ the Solanaceae Source (2011-) began ~~as a project~~ in 2004 to work towards a species-level treatment for *Solanum*, relying on existing modern monographs and revisions and coordinating work on the remaining groups using existing international taxonomic expertise. The continuous assembly of information on names, types, and specimens as well as bibliographic sources, and combining this with field work and phylogenetic research (e.g. Särkinen et al., 2013) resulted in an increasingly complete coverage of the information on the family. The network is coordinated by experts in the family, and a sustainably supported software platform (Scratchpads, Smith et al., 2011) provides the means for information storage, display and communication among the experts. The Lecythidaceae Pages (Mory and Prance, 2006-; Mory et al., 2010-) similarly attempt to provide a continuously amended information source on this Neotropical family, based on a widely distributed and institutionally supported network. The information system is now based on the commercial KE-EMU software (KEsoftware, 2014). The site developed for Melastomataceae (MELnet, 2007-14), though not yet going much beyond a species checklist and providing distribution data, is based on the Diversity Workbench Platform (Triebel et al., 1999-). The Cichorieae Network (ICN, 2009-) and Palmweb (2010-) served as exemplar implementations for the above described EDIT Platform for Cybertaxonomy, a suite of software tools aimed at comprehensively supporting the entire taxonomic workflow. The EDIT platform thereby supports alternative classifications, which are not found in other online syntheses.

For the monocots, eMonocot (2012-) provides a portal integrating several information sources, among those Palmweb (2010), GrassBase (Clayton et al., 2006), ~~and~~ several Scratchpad-based systems (e.g. eMonocot, 2012-a,b) ~~and GrassBase (2006-)~~. For Brassicaceae, the more recently created BrassiBase ~~site~~ (Koch et al., 2012; Kiefer et al., 2013) aims at “providing an online-accessible knowledge and database system of cross-referenced information and resources on Brassicaceae” including full taxonomic coverage of the family as well as

character and trait studies. ~~The network includes international specialists taking responsibility for different aspects and sub-disciplines of the biology of the family.~~ The software was developed specifically for the project.

Apart from sites with truly monographic aims there are checklist sites that cover plant groups, either covering certain taxa (e.g. WCSP, 2014) or geographic areas (see Crouch et al., 2013 for examples from southern Africa). A pioneer for collaborative organisation of such sites is the ILDIS network for Legumes, which started in the beginning of the 1990ies using the ALICE software (Bisby, 1993) and continues to be available on the Internet (ILDIS, 1996-).

~~Several phylogeny working groups~~ ~~Other initiatives merely~~ came together to coordinate sampling and phylogenetic analysis of diverse families ~~in order to obtain representative trees~~ and usually include the effort of revised classification above the species level based on the criterion of monophyly (e.g. Poaceae, Grass Phylogeny Working Group, 2001; Fabaceae, LPWG, 2013).

~~However, their goals usually do not explicitly aim at a species level monographic treatment.~~

We can conclude that these initiatives have been quite successful for a number of plant groups to collaboratively achieve the collection, merging and electronic publishing of existing information sources (published and unpublished) and thus can serve as proof of concept for a de-centralised approach ~~towards treating the entire world flora~~. As exemplified for ILDIS, however, institutional collaboration ~~and commitment~~ will be needed to ensure long-term sustainability (Crouch et al., 2013; see also Costello et al., 2014 ~~for a general consideration of the sustainability of biodiversity databases~~). ~~Also, the While existing portals supported by existing initiatives more or less present taxonomic information (treatments), as products online rather than they hardly making taxon concepts transparent with links between existing names, literature (incl. protologues), specimens (incl. types) and character data (from the phenotype and genotype). We also therefore also conclude that a synthesis of Caryophyllales that builds upon an organised and electronically supported taxonomic workflow has to still pioneer in several ways.~~

Information resources and standards supporting de-centralized taxonomic workflows

~~Apart from the taxonomic platform system, which holds the critically reviewed data, the taxonomic workflow increasingly relies on a number of external resources that allow access to previous work or to ongoing work in unrelated communities. This not only includes the information published in the literature (libraries) but also unpublished data (e.g. morphological descriptions of specimens, character matrices, lists of annotated specimens or annotated herbarium specimens). These resources become increasingly digitised and openly available: publications (e.g. BHL, 2005-; JSTOR, 2000-), digital specimens from individual herbaria or from aggregators (e.g. JSTOR Global Plants, JSTOR, 2013), general specimen and species observation data from GBIF (2004-), nomenclators (e.g. IPNI, 2004-) or entire checklists available for download or as a machine-accessible webservice (e.g. PESI, 2014).~~

~~For global synthesis projects, a broad level of coverage is essential and will still require substantial work. For the revision process, access to type specimens and other original material that has been used to name new taxa is essential. JSTOR Global Plants (GP, JSTOR, 2013) is a prime electronic information source aggregating images of type specimens and particularly important historical specimen collections from more than 300 herbaria worldwide. Preliminary gap analyses presented during the meeting of the Global Plants Initiative in~~

September 2014 show that basionyms which are represented by at least one specimen designated as some sort of a type in JSTOR Global Plants are: more than 70% for the comprehensive African Plants Database (2014L, Gautier, pers. comm.); 91% for the accepted taxa recorded for Mexican Lamiaceae (M. R. García Peña and H. Ochoterena, unpublished data); 74% for Lecythidoideae (M. Tulig and S. Mori, pers. comm.). The conclusion is that JSTOR Global Plants is becoming the unique resource for botanical type specimens. It is noted, however, that these studies are clearly biased towards Africa and the Americas, where the digitisation initiative funded by the Andrew W. Mellon Foundation was particularly active. As a consequence, this may explain, for example, why only 16 holotype or lectotype specimens are available for the names (species and infraspecific) in the largely Asian genus *Nepenthes*, which counts 259 species names in TPL. Moreover, the cited gap analyses confirmed the results of Smith and Figueiredo (2013) on the African species of *Polygala*, which clearly showed that the quality of the textual metadata in JSTOR Global Plants is not yet satisfactory.

The advances in standardising these data achieved by bodies such as TDWG make it increasingly possible to seamlessly interchange information. Standardised information is essential to enable and ensure the information flow between different teams that treat overlapping taxonomic subjects, e.g. between Flora initiatives and global synthesis networks. Berendsohn et al. (2011) provide an overview of the standards relevant in this context. Awareness of standard data definitions and adherence to standards in the creation of datasets is essential for the synthetic approach here described. Modern taxonomic tools largely respect this, but it is also possible for the individual to create and maintain data in a form that makes them reusable in a collaborative context. Almost equally important is the assignation of globally unique identifier for the physical objects (e.g. specimens) and data items (e.g. names) used in the taxonomic workflow (Güntsch et al., 2013).

Global synthetic monographs of organismic lineages taxonomic slicing and regional floras will mutually benefit

Global syntheses like the one envisaged for *Caryophyllales* have the advantage that they can consider information from study biological entities independent of political borders or regionally defined project settings. Hence, the importance of clade or taxon-wide monographs that are delimited only by scientific criteria, based on the analysis are essential for the understanding of the evolutionary history. In fact, very few clades exclusively occur within project regions such as countries: on the genus level such taxa are mostly restricted to islands (e.g., *Podonephelium* of the Sapindaceae in New Caledonia [Munzinger et al., 2013]). Nevertheless, a geographically much wider sampling had to be used to detect their monophyly. Considering further that such taxa constitute only a fraction of the respective floras, the knowledge obtained through large scale taxonomic syntheses will generally be beneficial to more regional flora projects. In fact, there are many cases of species or groups of closely allied species, for which strikingly different taxon concepts have been published. A *Caryophyllales* example is *Alternanthera halimifolia* from the (Amaranthaceae) that may be an endemic from the lomas of Peru or a widespread Neotropical species (TROPICOS alone lists 48 synonyms incl. infraspecific taxa). It is obvious that a regional flora project will neither have the resources nor the mandate to clarify species concepts in such cases. Global syntheses will therefore be essential to provide the knowledge on species concepts that can then also be used for more regional treatments.

On the other hand, there are also clear advantages of regional studies such as good specimen coverage, ~~and~~ the availability of a local/regional network that allows representative field sampling and knowledge about relevant specimens in regional herbaria. Floristic research often requires specific knowledge of localities, and thus involvement of local people and institutions. The question of whether a species occurs in certain geographic area can often only be clarified at that level. The same applies to the assessment of data which cannot be derived from a specimen in a collection, such as on habitat, phenology, pollinators etc. In many cases botanists acquire enormous expertise on the plants of their region but do not study plants beyond this region.

Furthermore, regional treatments offer obvious benefits for the user. A selection of species in a particular area will lead to more usable, smaller keys, and in fact may be the only feasible way to allow plant identification by non-specialists of a group. Therefore, there are good reasons to produce both regional floras and worldwide treatments. ~~Floras will better serve the practical needs of their region, e.g. by taking into account regionally restricted character variation.~~ Some major flora projects that cover very large and well delimited biogeographic regions may go beyond regional syntheses and may include the provision of more large scale original taxonomic research (e.g. Flora Malesiana, Roos et al., 2011). But even here, the overlap at genus and species level between the area covered by Flora Malesiana and the Flora of Nepal was recently estimated in the range of 30-40% (Pendry and Watson, 2009). The authors therefore suggested that coordinated efforts on the respective taxonomic groups could be a way to facilitate the production of treatments in both flora projects. Along the same line, Funk (2006) nicely pointed out that Floras are actually part of a continuum between revisions and monographs on one hand, and field work and data bases on the other.

~~Ideally~~ We conclude that, the authors who prepare treatments of specific taxonomic groups for a Flora project, ~~are should also ideally be~~ involved in the respective taxonomic network. However, the interaction between taxonomic networks (global slicing) and regional flora projects has to be actively developed by the respective scientific communities and the communication has to be intense. We believe that certain mechanisms and work-flows need to be developed to support this interaction, which should also be an objective of the Caryophyllales synthesis.

~~This interaction is straightforward when Flora treatments are currently prepared but require additional efforts when treatments are older (printed Floras have often no means to update the published treatments). Funk (2006) nicely pointed out that Floras are actually part of a continuum between revisions and monographs on one hand, and field work and data bases on the other. While generally agreeing, we argue for an even increased interaction between various approaches in the generation and dissemination of knowledge on the diversity of organisms. However, consider it as a prerequisite is that the interaction between taxonomic networks (looking at a taxon such as the Caryophyllales globally) and regional flora projects and taxonomic networks are is actively developed by the respective scientific communities and that communication is intense.~~

User needs versus needs of the scientific community and the individual researcher -- the example of the conservation community

During our preparations for ~~of~~ a global synthesis of Caryophyllales it became evident that there is quite a gap between the interests of users of monographs and taxonomic information and the needs of the actual scientists generating this knowledge. We believe that large scale

~~syntheses need to take that into account in order to be successful. The fundamental importance of a solid taxonomic knowledge base is well known in all communities dealing in one way or another with the diversity of organisms.~~ The Global Taxonomy Initiative, established by the conference of the parties of the Convention on Biological Diversity (~~CBD~~) in ~~(CBD 1998; 2006)~~ highlighted the importance of taxonomy at the political level. User needs have then been more clearly ~~and prominently~~ recognized through the Global Strategy for Plant Conservation (GSPC) that as a programme of the CBD called upon the provision of a widely accessible working list of known plant species (Target 1 for 2011) as a step towards a complete world flora (adopted by the COP in 2011 as Target 1 in the updated version of the GSPC for 2020). However, the underlying discussions were largely conducted outside the academic community, and without much regard to issues such as scientific methodology and workflows, research funding, good scientific practice, ~~or~~ academic curricula and career development. This may partly be the result of distributed responsibilities between ministries of the environment or natural resources (with responsibility for the CBD) on the one hand, and ministries for education and research (with responsibility for science) on the other. The result is not only a huge gap in funding for taxonomic research but also a decrease in quality. As a consequence, user needs are often only superficially met. Trying to fulfil such targets without responsible involvement of the scientific community, i.e. without the standard mechanisms of science-funding and science evaluation in effect, even increases the taxonomic impediment by cutting off many scientists and institutions from a positively evaluated contribution to solving global challenges. We hope that the facts and thoughts presented in this paper will not only stimulate a solution-oriented discussion on the example of the Caryophyllales in the scientific community but also a wider stakeholder dialog.

User needs appear to be very pragmatic in the conservation community, and in other communities using taxonomic information: a stable reference system for their factual data that refer to organisms. A user may feel satisfied once a name is attached to a plant, regardless if a natural species concept is reflected by the obtained name. User needs may contradict the actual research needs and are on short term not necessarily correlated with the quality of the science ~~(but see introduction for the problems that can result from building applied work on unclear species concepts)~~. It is thus most important to convey beyond the taxonomic community that the provision of reliable information is more than a mere mobilization of existing knowledge. In addition to the successful development of a biodiversity informatics infrastructure during the past two decades (Hobern et al., 2013), ~~the~~ aspects of data curation and synthesis as well as the actual research processes need to be supported. The research community can and will contribute, provided that appropriate incentives exist i.e. that their contributions further careers of individuals or positive evaluation and funding of institutions. Synergies can be generated ~~and user needs be met with quality products~~ if research projects that ~~normally often~~ span three to four years (e.g. PhD theses) ~~or in phases of two to three years each in most funding schemes~~ are embedded in a sustainable structure that supports the management and synthesis of primary data. We believe that such synergies will be very important for global initiatives of practical relevance for conservation such as the World Flora Online (WFO; CBD-SBSTTA, 2012) to be successful.

There are first positive developments (e.g. the German Federation for the Curation of Biological Data financed by the German research council, DFG; Diepenbroek et al., 2014), ~~but~~ Nevertheless, at the same time, there needs to be sufficient long term institutional funding to support data curation, large scale synthesis and result/data dissemination. It should be noted that in spite of the lack of adequate funding, much knowledge is generated by numerous smaller research projects worldwide. However, there so far has been no conceptual framework for integrating results into a global synthesis. Offering

this together with the respective infrastructure, as proposed in this article, could lead to benefits for users and the scientific community at large.

~~On the level of the individual scientist, comparable problems exist. There is an increase in the pressure on researchers to publish more papers in impact factor journals. As a consequence it is better for researchers to show quick results in shorter publications, than to devote many years of work on a monograph that in the end won't be published in an ISI journal due to its length. A frequent but striking case in this sense is that a new species can be published in an ISI journal while the monograph in a monograph series that provides a comprehensive set of data and hypotheses on species concepts will not have ISI recognition and counts as zero. It is the latter, however, which provides the sound basis for, e.g., conservation work, also on the local level. At a global level, there is a patchwork of knowledge. In initiatives of practical relevance for conservation such as WFO (CBD SBSTTA, 2012) there will be the need to integrate treatments from different regional floras; this requires the study of taxonomic concepts and this again can only be achieved by a monographic approach. There are ongoing attempts to overcome this dilemma. For example, both IAPT (the International Association of Plant Taxonomists; K. Marhold, pers. comm.) and CETAF (the Consortium of European Taxonomic Facilities, M. Price, pers. comm.) are urging indexing services such as the Web of Science to provide a category for taxonomy instead of subsuming taxonomic work under other biological sciences". Several recent publications analyse and provide further ideas to overcome the problem (e.g. Ebach et al., 2011; Wägele et al., 2011; Payne et al., 2012), including the citation of all taxonomic literature that was used to conduct the research or the authors who published the respective species concepts used. We believe that the approach of a global network like the one laid out here for the Caryophyllales and synthesis helps to increase the productivity of the individual researcher, by offering access to infrastructure, to a structured knowledge base, as well as to methodological background knowledge that directly furthers the research effort. Therefore, a positive effect on research quality can also be expected, with benefits are also likely in the sense of a social network and by increasing the visibility and impact of publications. Therefore, a positive effect on research quality can also be expected. Furthermore, the network effecting the Synthesis will actively discuss means to increase the impact of publications that form output from the online contributions.~~

The angiosperm order Caryophyllales as a model

Function of the model group

On a small scale, prototypic treatments exist that largely fulfil the criteria for an integrative monographic approach ~~that as have been~~ outlined above. However, this has not yet been attempted on a large scale, ~~with groups~~ involving thousands of species and representing a major clade of ~~the~~ angiosperms. Simply upscaling existing methods will not be sufficient. Rather a new, much more collaborative approach is needed that focuses on the knowledge generation process in biological systematics ~~as outlined in the previous chapters~~ and also builds upon the existing data and information resources.

The approach outlined here needs to emphasize short and long-term goals. Taxonomists are challenged to provide as soon as possible a treatment which globally covers the species diversity in order to provide a base for decision making, e.g. in conservation. On the other hand there is the goal of building a scientifically comprehensive information base and synthesis on the group that is fully based on documented evidence and falsifiable in its postulated taxonomy. Nevertheless, several information components (e.g. database of names,

citations, protologues, type specimen images) are essential in both short and long term perspective. An electronic information system will therefore have to structure and manage available published knowledge, evidence gathered in formulating new knowledge (including primary data linked to specimens), and provide the means to gather and re-integrate new data and knowledge (e.g. altered taxon concepts) and thus actively support the development towards a truly integrated scientific monograph of the entire group.

Evolution, diversity and importance of the Caryophyllales

The Caryophyllales were chosen as a group to serve as an exemplar for this new approach because the order is ~~not overly large but at the same time large enough~~ of significant size to have an ~~significant~~ impact in advancing our knowledge on flowering plants. Caryophyllales comprise about 5% of flowering plant species diversity, and are classified into 37 families. Currently, 732 genera (Hernández-Ledesma et al., in prep.) and an estimated 12,500 species are recognized (~~Table 1~~). Caryophyllaceae, Aizoaceae and Cactaceae are the most species-rich families, while ten families comprise only one to three species (~~Table 1~~). The Caryophyllales are an ecologically diverse lineage with traits such as C₄ photosynthesis, salt and/or gypsum tolerance (~~Moore et al., 2014~~), succulence and carnivory. Centres of diversity are temperate and subtropical regions, especially in Mediterranean climates, deciduous vegetation with seasonal distribution of rainfall, semi-deserts and deserts, and montane vegetation. Due to the high number of endangered plants recognized in the order, prominently in Cactaceae, Droseraceae and Nepenthaceae, the group is also relevant in the context of the Global Strategy for Plant Conservation and the Convention on International Trade in Endangered Species of Wild Fauna and Flora (CITES). Furthermore, Caryophyllales comprise many species of economic importance such as ornamentals (e.g. Cactaceae, Caryophyllaceae, carnivorous groups), cereals and green vegetables (e.g. spinach, amaranth, quinoa, and sugar beet), but also as noxious weeds (e.g. *Alternanthera philoxeroides* (Mart.) Griseb., *Amaranthus spinosus* L., *Opuntia* spp., *Mirabilis* spp.) and allergens (e.g. *Amaranthus retroflexus* L., *Atriplex* spp., *Salsola kali* L.).

Several studies have shown that Caryophyllales form a monophyletic group (e.g. Cuénoud et al., 2002; Hilu et al., 2003; Schäferhoff et al., 2009; Brockington et al., 2009; ~~Qiu et al., 2010; Soltis et al., 2011~~). The concept of the order has changed from the pre-phylogenetic Centrospermae (Eichler, 1878) to the expanded Caryophyllales that is largely based on DNA sequence data. A summary of the taxonomic history will be presented in Hernández-Ledesma et al. (in prep.) along with a more detailed discussion on the current understanding of phylogenetic relationships.

Existing data and information sources for Caryophyllales

The Caryophyllales are typical for most major groups of flowering plants, in that their family-level classification is beginning to stabilise while generic realignments based on phylogenetic insights are well underway (see section on establishing a taxonomic backbone below). Nevertheless, the big work ahead relates to the circumscription and treatment of species and infraspecific taxa, which will therefore be discussed here. Since integration of information from existing electronic sources is fundamental to streamline the work of the taxonomist, we focus on publicly available electronic data sources for Caryophyllales.

The International Plant Names Index (IPNI 2004–) provides a starting point for names and their original publication. Having been produced from three datasets (Index Kewensis, Gray Index and Australian Plant Name Index), there is considerable duplication of names. Tropicos (2014) similarly provides nomenclatural data including the original publication but also indications of the nomenclatural status of names and the usage of names in the form of being accepted or a synonym according to specific bibliographic references. It is especially this feature that renders Tropicos one of the important sources for Caryophyllales information, albeit it is somewhat geographically biased. A total of 36,835 species names are registered in Tropicos for the Caryophyllales (as of Oct. 22, 2014). IPNI and Tropicos were used to complement the so-far existing global family checklists to form The Plant List (TPL, 2013). TPL is a checklist indicating a status (accepted, synonym or unresolved) for most species-level plant names that have been published. TPL lists a total of 50,192 names, about 37% of which are unresolved as to their status as correct names or synonyms (see [Table-Fig. 1](#) for details [in four families, and Appendix 1](#)). It is the only comprehensive species-level taxonomic checklist that covers the entire Caryophyllales on a global scale.

We carried out a review to assess the species numbers for the different families in Caryophyllales comparing TPL and expert views taken from the literature ([Table 1 Appendix 1](#)). While species number estimates match in some medium sized or small families (e.g. Didiereaceae, Frankeniaceae), numbers differ by a quarter and more for families such as Cactaceae, Caryophyllaceae, Droseraceae or Polygonaceae. This indicates the amount of work that has to be done in order to clarify species limits and arrive at a stable classification. In addition, the number of unresolved names in TPL points to the task of clarifying existing usage of species names. In order to tackle these problems, easy access to literature and specimen data is instrumental.

Fortunately, more and more of the [taxonomic literature necessary information](#) becomes accessible on-line [such as: publications \(e.g. BHL, 2005-; JSTOR, 2000-\), digital type specimens from individual herbaria or from aggregators \(e.g. JSTOR Global Plants, JSTOR, 2013-\), general specimen and species observation data from GBIF \(2004-\), nomenclators \(e.g. IPNI, 2004-\) or entire checklists available for download or as a machine-accessible webservice \(e.g. PESI, 2014\)](#). No reliable estimate can be given at this time of the degree of on-line availability of taxonomic literature relevant for the Global Synthesis of Caryophyllales but an indication may be the number of names with digitized and accessible protologues. For example, for the genus *Anabasis* (Basellaceae) about 30% of the names in Tropicos offer a link to a protologue openly accessible in the Biodiversity Heritage Library; for *Nepenthes* (Nepenthaceae) this number reaches 45%. However, with these numbers it has to be considered that there is a big amount of digitized literature available on-line from other, non-specialized sources, e.g. from JSTOR (2000-) and from the many commercially published journals with taxonomic content, albeit these sources are often behind a paywall.

Examples of electronic sources for Caryophyllales that specifically treat a defined geographic region are presented in [the appendix \(Appendix-endix 24\)](#). These include taxonomic checklists and Flora treatments in a wide variety of formats, from scanned documents that are not immediately usable for structured data access in an information system, to highly structured databases that can produce output which can be reused almost instantly. Although the list given in the appendix is by no means fully comprehensive, it becomes clear that still many treatments are not openly accessible (in spite of calls to the contrary, e.g. by Brach and Bouffort, 2011). Looking at the over-all available literature it also becomes clear that quality treatments and monographs on the species level are far from fully covering both the taxonomic diversity and the geographic spread of taxa in the Caryophyllales.

For the revision process, access to type specimens and other original material that has been used to name ~~new~~ taxa is essential. [JSTOR Global Plants \(GP, JSTOR, 2013-\)](#) is a [prime electronic information source aggregating images of type specimens and particularly important historical specimen collections from more than 300 herbaria worldwide. A dataset for Caryophyllales families provided by JSTOR \(As of 8 December 2014\) to the authors included records representing](#) a total of close to 75,000 specimen images [pertaining to Caryophyllales that](#) were accessible through GP (JSTOR 2013-), of which more than 52,000 carried some indication by the data provider that the specimen was considered to be a type or original material. More than 12,000 were marked as holotypes, lectotypes or neotypes, 4,500 as syntypes. Nearly 20,500 specimens were designated as isotypes, isolectotypes, isoneotypes and isosyntypes. Around 1000 records each refer to paratypes and original material. The remaining more than 13,000 records were just called Type, with more than 3,000 explicitly marked as questionable. Because we lack a complete checklist of Caryophyllales species (and infraspecific) names and because the names given in GP are not fully standardized, we cannot be sure how many different names are represented in this dataset. To understand the situation in more detail, we looked at examples such as the seven genera of the tribe Pisonieae (Nyctaginaceae; Hernández-Ledesma unpublished data), with an estimated 260 species. Global Plants provides images of the holotypes, lectotypes or neotypes for 89 of these names, i.e. for about one third. At least one type specimen (this then also includes isotypes and isolectotypes) is available for 167 names, i.e. for about two thirds, while for one third no type is [present/available](#). This level of coverage is in rough agreement with the findings of preliminary gap analyses on other plant groups as cited above under Information Resources. However, it may also be noted that the coverage is better for Africa and the Americas, where the digitisation initiative funded by the Andrew W. Mellon Foundation was particularly active. As a consequence, this may explain why only 16 holotype or lectotype specimens are available for the names (species and infraspecific) in the largely Asian genus *Nepenthes*, which counts 259 largely unresolved species names in TPL.

[-Further work is therefore necessary to localise and digitize the remaining type specimens, the degree of which will vary from genus to genus, and may be substantial in several genera. Moreover, when the number of accepted species names and the overall total of names at species level \(Table 1\) is compared to the number of type images available \(Fig. 14\), another important issue becomes evident: in many cases a relatively large amount of specimens was actually digitized including types, specimens that are probably types or just historical material without scrutinizing their correct status. The most important task will therefore be to verify the type status of the material available. Global Plants is therefore a source that facilitates taxonomic work but not yet a ready-to-use taxonomic product. the cited gap analyses](#) This confirms [the results of Smith and Figueiredo \(2013\) on the African species of *Polygala*, which clearly showed that the quality of the textual data \(e.g. the category of types, collector information, etc.\) in GP is not yet satisfactory. The conclusion is that JSTOR Global Plants is becoming the unique global resource for botanical type specimens, but that specialist input is needed for quality control of the type specimen data. The amount of work needed is difficult to assess and will not only depend on the respective genus or family but also on how Global Plant data curation interacts with the taxonomic work flow maintained by the teams working on these taxa. Within the Caryophyllales synthesis, this work should take place as one of the first steps in the context of verifying protologues, citations and type images as an activity of the scientific community. The results can be incorporated into commented global checklists which can also be published, at least as data papers. However, JSTOR also needs to provide resources to manage the information backflow and to implement curation standards for Global Plants.](#)

With regard to specimens in general, herbaria have made much progress in providing ~~digital catalogues, now increasingly seconded by~~ digital images. Some major herbaria have been completely digitised, e.g. those in Paris and Leiden, others provide varying degrees of mobilisation of their holdings. Access is still not satisfactorily organised, although the Global Biodiversity Information Facility (GBIF, 2004-) provides the technical infrastructure for a global, single point of entry. Currently (October 2014), GBIF offers access to 43.5 million records that have been marked as plant specimens, of which close to 2.4 million are classified as Caryophyllales. However, the number of already digitised specimens exceeds this by far. For example, the large herbarium in Paris, although fully digitised, is not included in this number, neither are large already digitised specimen holdings, e.g. in China, Brazil, and India. The conclusion is that, similar to the view on the literature presented above, incorporation of specimen information will be eased by the electronic resources available, and that this will make the results more comprehensive and reliable, but that this will in no way replace taxonomic scrutiny and synthesis.

As for DNA electronic data, the INSDC (International Nucleotide Sequence Database Collaboration; Karsch-Misrachi, 2011) for example retrieves 1,623,294 records for Caryophyllales (<http://www.ncbi.nlm.nih.gov/nucleotide/?term=caryophyllales>: accessed on December 22 2014). This amount of information would be paramount for the understanding of the phylogeny and hence naturalness of many taxa, but it is difficult to assess how many of those sequences are effectively linked to well-documented and accessible voucher specimens (even fewer of them would be linked to specimen images). Also, the sequenced loci are not standardized across taxa. It will therefore be necessary to access the original source in order to include sequence data into a global synthesis. Actual research will then have to evaluate if the taxonomic identification can be reassessed and the data placed into the respective taxon circumscription, in particular at the level of species.

In conclusion, a formidable amount of information is accessible and provides a broad base to accelerate the step-wise research process towards the Global Synthesis of Caryophyllales. Further in-depths analysis of the existing information resources is part of the work in progress including the identification of priorities for further data mobilisation and digitisation.

Main goals and application needs to be fulfilled

In the following paragraphs we will consider the organisational prerequisites and present a strategy towards a global Synthesis of Caryophyllales ([Fig. 2](#)). There are two major lines of activity:

- (1) A global **Caryophyllales Network** needs to be established in a way that
 - integrates research activities to organize a quality and first hand outlet of knowledge, building on effective communication within the scientific community and formal agreements among the participating institutions
 - facilitates interaction between established (taxonomic) specialists and young researchers, thereby promoting postgraduate education and career development
 - promotes scientific exchange by organising workshops or conferences.
- (2) A scholarly **information system** has to be established on-line that

- provides the state of knowledge on species diversity
- allows to present the currently accepted concepts of species (and supraspecific as well as infraspecific taxa)
- promotes the identification and discussion of conflicting taxon circumscriptions
- integrates information on names, synonyms, protologues, descriptions, and keys
- integrates character data both phenotypic (“morphological”) and DNA (“molecular”)
- facilitates the development of a standardized terminology used in Caryophyllales (in the sense of an ontology e.g. for morphological characters)
- provides synthetic distribution data and specimen-based distribution records for accepted taxa
- credits all sources of information and thereby increases the visibility of research
- informs original sources about data corrections, enhancements, and annotations.
- references source objects and provides stable references for newly created data objects
- represents an exemplar high quality data set through linking of specimen information (including digitized images) with their meta-data (e.g. geographical coordinates) and character data
 - to advance the recognition and classification of species on the basis of well-studied biological entities
 - to assess and model species distributions including their future trends
 - to reconstruct the evolution of lineages in an approach illuminating molecular and morphological diversification
- provides a dynamic and easy to update platform to identify areas of future research and to promote collaboration
- enhances visibility of the network, its partner institutions and contributing scientists.

Strengthening the Caryophyllales Network

In the taxonomic community, many examples of collaborations of individual researchers exist. Forces are often joined in the context of projects with defined kick off and completion dates or to review the state of knowledge about particular taxa or topics. Some collaborations also provide an organisational framework including resources for coordination, data management and curation, apart from the actual research activities. But such networks are often grounded in project financing, and consequently, collaborative activities tend to dwindle once the project period runs out. For long-term sustainability, we therefore believe that institutional commitments are imperative. Institutions need to sustain coordination and information management tasks in addition to allocating the resources for the dedicated individuals forming an expert network.

For the Caryophyllales network, an initial institutional partnership consists of the Freie Universität Berlin’s Botanic Garden and Botanical Museum Berlin-Dahlem (BGBM, herbarium B), the Universidad Nacional Autónoma de México’s Instituto de Biología (incl. herbarium MEXU) and the Instituto de Botánica Darwinion in Argentina ([herbarium SI](#)). All three are committed to contribute their existing expertise in Caryophyllales research into a larger scale initiative. The BGBM has defined Caryophyllales as one of the focus areas on their long-term research agenda with a structural backup of a full position dedicated to

sustainably support the network. A memorandum of understanding (MoU) among these core partners will formalise the agreement on the project's main aims, structure and governance, in order to coordinate the elaboration of products and the execution (and further development) of the work plan (including methods and tools as well as milestones and time frame). Efforts are under way to extend the geographic coverage to include further institutions with Caryophyllales on their research agenda.

The actual network, however, consists of the individual scientists. There are several benefits arising from participation, and the intention is to develop the network as to create mutual benefits. Apart from easier communication about research questions, the information infrastructure (incl. linked resources of structured data on Caryophyllales) will facilitate the generation of a variety of publications, including scholarly papers, flora treatments, and data papers. The network will be instrumental in increasing the visibility and recognition of the participating scientists and institutions in the scientific community and for users, which in turn can have positive effects for evaluation and funding purposes.

A distinct advantage of a jointly curated repository of the information on Caryophyllales is the possibility of linking into other initiatives by means of information interfaces. For example, the World Flora Online (WFO) project (CBD-SBSTTA, 2012) has synergies with the Caryophyllales monographic synthesis. WFO aims for a compilation of current floristic knowledge across the entire world while the Caryophyllales synthesis aims for a quality treatment maintained for an important group of flowering plants. It is not the intention of WFO to replace current regional floristic or revisionary efforts. The interaction between the Caryophyllales synthesis data and the (presumably much less detailed) WFO can be streamlined using existing data interfaces (Fig. 3) that, on the Caryophyllales side, need only minimal adjustments to adapt to the WFO data ingestion standard (a variant of the Darwin Core standard, in Darwin Core Archive format). In the future, we think that WFO data ingestion will be achieved by using the Web Services of the EDIT Platform, which already can provide Darwin Core standard data.

In building the human Caryophyllales network, a detailed assessment of relevant publications, information sources and experts was the starting point. In order to organize efforts towards the understanding of Caryophyllales promoting communication and international collaboration, we have contacted specialists, who are currently working on the systematics and/or taxonomy of Caryophyllales or who have contributed to the knowledge of the group at least in the last decades. Of the 72-110 specialists contacted, 50-more than 80 have responded positively to the initiative of a global synthesis. These botanists are studying most diverse families of the order and are based in universities and other research institutions in 26 different countries. Since further integration and participation of the community (established senior researchers and students are invited to participate in the network) will be essential for the success of the initiative the overall positive responses from these specialists to our initiative are encouraging. An important step in further constituting the network community ~~will be~~ the ongoing organisation of an international Caryophyllales conference in the autumn of 2015.

Implementing the work-flow and information system for the Synthesis

An overview of the work-flow towards to synthesis of Caryophyllales is given in Fig. 2. In order to achieve dissemination of synthesis results as quickly as possible and to provide structured information sources for the community involved, the workflow generally follows two paths:

- (1) Synthesis of existing information
 - a. Compilation of species names (data ingestion)
 - b. Integration of resources after expert review and verification (data ingestion: taxonomic treatments, literature, online resources, specimen information)
 - c. Data publication and presentation (output products)
- (2) Phylogenetic evaluation of taxon concepts based on structured data
 - a. Assessment of molecular and morphological character data and ontology
 - b. Testing hypotheses with phylogenetic methods
 - c. Translating the results to species nomenclature and classification

The two main paths are strongly interrelated and will take place in parallel and with different speed in different taxonomic groups. Global coverage will therefore be reached independently for different sub-groups, and for the majority of these at an alpha-taxonomic level first. A fundamental principle is to streamline the replacement ~~of~~ with more thoroughly evaluated treatments once they become available.

In the first step, existing resources such as taxonomic information already online as well as printed taxonomic treatments, other literature and specimens are being made technically available to be integrated (~~see chapter Existing resources for details~~). Most of the necessary interface specifications to access and import existing taxonomic information resources already exist, but the greater part of the electronic information available is insufficiently standardised to be directly used. Exceptions are nomenclatural and checklist information as well as specimen data, where many sources produce reasonably standardised output. Part of the expert review and verification will be to link these resources, e.g. the names from IPNI with the publication information present in JSTOR (2000-) and BHL (2005-) as well as with the type specimen information present in JSTOR Global Plants (JSTOR 2013-), the unresolved names in TPL (2013) with ~~there~~ accepted names ~~names~~ (where possible), and unambiguously circumscribed species from eFloras with expertly verified specimens that serve as vouchers (for ~~, e.g.,~~ information on sequences, pollen, morphological characters, and geographic occurrence etc.). New data generated during that process include the references that were used in establishing these linkages and the information needed to give proper credit to the person who carried out the review and verification. In contrast to specimens and names, the available data from flora or monographic taxonomic treatments (i.e. descriptions, keys and supplemental information) are largely unstandardized. Thus, one of the first steps here will be to identify priority data sources, electronic or in print, in order to be able to mobilise that information and make it available for the Synthesis. These sources have of course the advantage to come with an authorship that can directly be used for proper credit at this stage. However, the process of transforming a taxonomic treatment from print to database involves not only substantial technical but also a considerable amount of taxonomic expertise (Hamann et al., 2013), which both need to be properly credited.

In the following step, specialists for the different Caryophyllales groups will be using these existing resources as a basis for the revision and verification of information, and also for deeper evolutionary analysis. This step is crucial in terms of quality control and gap filling and highlights the necessity for the involvement of the respective specialist community. A literature database containing relevant literature on taxonomic, systematic, nomenclatural, phylogenetic, biogeographic and ecological aspects of Caryophyllales groups is currently under development. Several specialists have contributed reference libraries that are integrated and de-duplicated to be made available online for the Caryophyllales community. The bibliography will be updated continuously to provide a useful tool for research.

[Translating the results of phylogenetic and evolutionary analysis into a formal classification system as part of the work flow; will thus help to overcome the “phylogeny-classification gap” \(Franz, 2005\). But this does not only hold true for the importance of classification systems that attach easily remembered names to organisms considering the vast majority of users \(Franz 2005\). It also addresses another quality of a “phylogeny-classification gap” – which is that most phylogenetic studies still do not provide the needed data to support a species level classification, mostly due to limitations in taxon sampling. Our goal therefore is to more comprehensively integrate evolutionary analysis and classification work.](#)

Establishing a dynamic taxonomic backbone as an integral part of the Synthesis

As a first step the current understanding of family circumscriptions was evaluated with the aim of establishing well defined monophyletic groups. This resulted in the recognition of 37 families in the order, largely in congruence with the last classification of the Angiosperm Phylogeny Group (APG III, 2009), but incorporating some recently published research results. Along the same line and aiming at full taxonomic coverage an annotated generic checklist of Caryophyllales is being created in an extensive collaborative effort (Hernández-Ledesma et al., in prep.). It will provide the first version of the genus level taxonomic backbone made available both as a dynamic electronic source and as a scientific paper. It builds upon the comprehensive but mostly pre-phylogenetic genus-level treatment for the entire order published [by various family specialists](#) in the “Families and genera of vascular plants” series; volumes two and five (Kubitzki et al. [eds.], 1993; Kubitzki and Bayer [eds.], 2003). In order to reach a comprehensive cover of relevant genus names, the electronic version of Names in Current Use 3 (Greuter et al., 1997) was used as a starting point. More recent comprehensive family treatments including all genera were considered (e.g. for Aizoaceae --several authors in Hartmann, 2001a–b; Cactaceae -- Hunt, 2006; Basellaceae -- Eriksson, 2007) as well as many specific papers that reflect the enormous progress in understanding the evolution of Caryophyllales in the last two decades. While the generic checklist provides a complete genus-level treatment of the order that represents current knowledge and identifies gaps, it must be emphasized that the individual quality of generic circumscriptions depends of the stage reached in the step-wise progress in research (see above). Ideally, genera represent well-defined monophyletic entities, but the treatment can also be provisional (phylogenetic trees lack resolution and support in respective nodes) or still at alpha-taxonomic level (no phylogenetic data available). Through explicitly stating the stage of knowledge, the generic synopsis also provides the framework for facilitating future knowledge generating activities, which will again lead to constant updates of the genus level treatment. The actual editorial procedures supporting and documenting this process are on the agenda for discussions in the Network.

To meet user needs, a full-coverage species-level taxonomic backbone will be built initially using available electronic sources (e.g. The Plant List) if there are no expert-reviewed resources (see below). Access will be name-based and conflicting classifications can be depicted in parallel (see the historical Provisional Global Plant Checklist, IOPI, 1997-2003). At the species level (including infraspecific taxa), the percentage of taxon concepts which have already been evaluated by truly evolutionary methods at this point cover only a small fraction of the estimated 12,500 entities. In case there are recent monographs (at least providing the character data in support of species concepts accepted) the species-level taxonomic backbone can reflect those monographs, easily giving credit to the respective author(s). This may extend to Flora treatments if they include whole genera or monophyletic subgeneric entities. For all Caryophyllales species names, the citations, protologues and type

images will be included in or linked to the EDIT platform (Fig. 32) and presented on-line. This will be particularly important for those genera for which just an automatically compiled species list is available. The Caryophyllales network will help to identify colleagues who can more easily work out “expert validated species checklists” which are essentially alpha-taxonomic but progressed in quality over a simple names compilation and which will replace the initial species list. The resulting product will be credited to its author(s) and can be citable using formats such as data-papers (Chavan and Penev, 2011) or even a journal publication. It is estimated that a larger proportion of the total species diversity can be treated at least to this level until 2020.

The expert-validated species checklists will then also inform sampling plans for integrated morpho-molecular sampling of documented specimens in the course of evolutionary analyses. The hypotheses on biological entities identified in this research can then lead to an evaluation and eventually update of the species level treatment of the taxonomic backbone, then replacing the expert checklists. Since this kind of evolutionary research is more comprehensive, its publication as journal paper is likely, making it easier to credit authors. Developments in science evaluation (impact factors etc.) will be reflected in the publication policies of the network. Print (or print-like) publications generated from the online resource will also be possible at appropriate times for significant taxonomic groups (e.g. families) or following specific stakeholder demands, also as part of journal articles dealing with new research results in evolutionary analysis.

Working towards a standardized ontology for characters, structured descriptions and linking specimens to character data

With regard to descriptive data, a forward-looking approach will turn to account for the plethora of knowledge on Caryophyllales accumulated in the literature and the research collections, while at the same time it will be oriented towards a more sustainable way of organising systematics than the common state of the art. This is a decisive part of the vision of the integrative on-line monographic approach.

Sustainability in systematics presupposes maximum data build-up. The corresponding requirements start with (1) unequivocal links of the gathered character data to the specimen from which they originate, continue with (2) standardised basic entities of information (i.e. terms and concepts of morphological characters and states) and (3) lead to iterative data aggregation procedures using congruent and interoperable character and state matrices from specimens to primary and higher taxa (see Kilian et al., submitted). The first and last are basically information technological requirements. Standardised terms and concepts for character data, in contrast, are core issues of systematics: they span from term ontologies to homology issues, they bridge between literature data and new data to be gathered, and they are fundamental to the entire process from data gathering to data aggregation and synthesis of information.

Playing such a central role, predefined standardised terms and concepts for character data receive a high priority in the timing and structure of the work process in the Caryophyllales Synthesis. Developing terms and concepts for a standard Caryophyllales character data matrix should thus be high on the agenda of the Network, starting out with an analysis of the characterisations (“descriptions”) of the genera in the taxonomic backbone, taken from literature and unpublished own sources. In a first resulting work flow, contributing specialists extract terms and concepts of character data with references into a terminology wiki (Karam

and Fichtmüller, 2014-), forming the referenced elements of a growing illustrated standard term glossary with ontologies, which also builds on existing glossaries (Beentje, 2012; Egli, 1993) and other activities regarding work flows and ontologies (Franz and Thau, 2010, Deans et al., 2012, Mathew and Güntsch, 2013). In a second resulting work flow, the unstructured textual genus descriptions are transformed into structured characterisations in a character data matrix (using the open source Xper2 software, see Ung et al., 2010, 2010a) within the EDIT Platform (Fig. 32), with the support of the contributing specialists and employing the predefined standard entries of the Caryophyllales glossary. Working up all Caryophyllales genera in this way ensures that a vast majority of the actual range of characters and states in this order will be covered and compared. Of course, there will be a number of specialised characters that apply only for a specific group within the order, but we believe that also for these their descriptive value will be greatly enhanced with the inclusion into a joined semantic framework.

In parallel, three important tasks can be achieved in this way: (1) a standard Caryophyllales glossary and referenced ontology; (2) a standard Caryophyllales character data matrix; (3) a multi-access key to the genera of the Caryophyllales by using the functionalities of the Xper software on the basis of the Caryophyllales character data matrix. The structured way of data storage in both the standard glossary and the standard matrix enables their efficient editing, improving and supplementing, which is inevitable even in the long run. The standardisation of matrix and glossary ensures congruence and build up of the character data across the entire order.

In addition to this top-down approach, we envisage a complementary bottom-up approach, wherever appropriate, for specimen-based taxonomic work. This approach is based on the proposed solution to extent (a) the use of character data matrices from taxon characterisation to specimen data gathering, and (b) the data aggregation from lower to higher taxon ranks to the iterative aggregation of specimen-based data at species (or any other primary taxon) level (see Kilian et al., submitted). Employing the Caryophyllales standard matrix for specimen data guarantees the data congruence. It brings the further advantage that subsequently revised taxon delimitations and specimen identifications can be easily brought into account by triggering a re-aggregation of the datasets.

Finally, since for the time being few structured data sets for taxon descriptions below generic rank will be available compared to the numerous unstructured textual descriptions in the literature (mostly of limited regional scope, as from floras), we envisage a mixed approach for taxon descriptions at lower taxon ranks: unstructured textual descriptions from the literature can be analysed in the same way as will be done for generic descriptions, thus adding both to the standard glossary and standard matrix. A literature item can then be transformed and incorporated as a separate item into an otherwise specimen-based application of the Caryophyllales standard character data matrix along with the specimen items. The mixed matrix data set can subsequently be aggregated into a taxon level matrix data set. Moreover, with growing numbers of specimen-based items, comparative evaluations of the specimen-based and literature-based items become possible. The data of the literature-based items are then either increasingly backed by the specimen-based data, or the discrepancies become apparent. In this way, our approach supplies a qualified assessment of the concept relationships (Berendsohn, 1995; Geoffroy and Berendsohn, 2003) between literature-based items and a preferred taxonomic concept.

Data publication and presentation

The synthesis of information will be carried out in the EDIT Platform for Cybertaxonomy (see above). Different interfaces are used to integrate information from all components, specialists (or editors) review and re-organise the content on-line, with support from the Network coordinators.

One key product of the global synthesis will be the Caryophyllales Portal, a scholarly information system that will serve as a single entry point for expert-reviewed quality information about the order. The Portal will be a product of the EDIT Platform (Fig. 32), web services and data exchange via standardised formats will be used to secure connectivity to international biodiversity informatics initiatives such as IPNI and GBIF, wherever considered appropriate by the Network and technically feasible. The platform web services are also the primary interface to advanced computational workflow environments. In terms of products, scholarly content and scientific publications are at the core of activities, but products for the interested general public (education purposes; updated quality material for teaching) should not be neglected. Illustrations (photographs, botanical illustrations) must be included.

The portal is not seen as a means to publish primary research results; it should provide the interface and synthesis of such publications and make the results available in standard formats for open data exchange. However, authors may wish to publish expertly revised alpha-taxonomic species catalogues as starting points for treatments through the portal. Using standard interfaces, quality data with respect to the taxonomic backbone as well as with respect to detailed species level treatments can also be made available to ongoing flora projects, the WFO initiative, and to other initiatives like the Catalogue of Life and the Encyclopedia of Life.

The editorial processes and responsibilities involved in establishing the taxonomic backbone (and thus the baseline for the Synthesis) will be discussed and developed as a collaborative effort of the Caryophyllales Network. Basically, the principle is that the scientists or teams of scientists who work on a family or genus are also updating the backbone. The exact editorial process is being discussed. Clearly, this will be and for some time stay a highly iterative process, with new research results being constantly fed back into the information system. The resulting instability will represent a problem to users outside the taxonomic community. As in other systems (e.g. the Catalogue of Life, Roskov et al., 2014), regular stabilised and citable versions of the backbone and the Synthesis should be part of the output.

What is a realistic timeline for the Synthesis?

Of course, this is a question that is intimately related to the availability of resources, monetary and intellectual, that can be dedicated to the initiative. Commitment of resources thus will be one of the principal discussion items within the network. The online bibliography is expected to be available by the end of 2015, and the genus-level backbone is already implemented in the EDIT Platform; its first edition will be available and published in 2015, which will also be the time for the first launch of the Data Portal. With this available, the import of species names from online sources can commence and the EDIT Platform will have feedback mechanisms to these sources for corrections generated in the integration process. Where feedback interfacing cannot be automated, an annotation system based on the AnnoSys approach (Tschöpe et al., 2013) can be used. Specialist' review and updating of the species level backbone will start with a few exemplar groups, and first publication of comprehensive species checklists can be expected for 2016, including links to on-line protologues, type specimen images and (partial) distribution information (as soon as expertly identified

specimens become available, point distribution maps can be generated by the portal). Work on the integration of descriptions and other character data from published works will be an ongoing accompanying process, in many cases seconded by negotiations with publishers on open access to such resources. Species and generic boundaries are still controversial in many groups and detailed studies at the species-level are often lacking. Once smaller monophyletic groups are identified in broader studies, these clades can be the focus of more detailed studies. During this step-wise process knowledge gaps will be identified and result in research providing evolutionary hypotheses for the generic and species limits in clades that are currently understudied. Information on the geographic distribution based (initially) on TDWG units (Brummitt, 2001) will considerably increase the utility of the portal over the usual checklist formats, as will the explicit stating of caveats with respect to taxonomic classification, distribution, nomenclature etc., in order to specify potential problems for proper usage.

We do think that a minimum coverage of all species in a taxonomically consistent and comprehensive taxonomic backbone should be achievable by 2020, which incidentally is the deadline formulated by the CBD for a World Flora Online to become available for the conservation community – (CBD-SBSTTA, 2012). How far the comprehensive Synthesis including the species level sampling for phylogenetic analysis and the character matrix and ontology will have progressed by that year cannot be forecast at this stage, it is too dependent on the resources made available for research as well as on the degree of community integration achieved by the Network.

Conclusion

The current international initiatives to integrate and synthesize the actual taxonomic and systematic knowledge of flowering plants [such as WFO or Catalogue of Life](#) provide information standards for the compilation and accessibility of the available data using modern electronic tools. Nevertheless, [theirese](#) strategies neither launch international collaboration of experts nor the guidance of new generation of scientists focused on the study of biodiversity. They also do not actively promote more in-depth research, although [only](#) a rather small proportion of the species concepts currently used has hitherto been evaluated in an evolutionary context. We propose mechanisms to achieve these objectives using Caryophyllales as a model group. We aim to implement a system that allows not only the compilation of existing data, but also the generation of new information that can replace it in a stepwise process. This iterative process intends to shed light on the currently controversial taxon circumscriptions [including at](#) different hierarchical levels. The basis of this progressive workflow is the integration of phylogenetic methodology to test current hypotheses of both taxa and characters based on specimens using structured and re-usable character data that are clearly linked to specimens. Our approach will also provide a perspective to the further development of the science of systematics (and taxonomy) in terms of improving the additivity and re-usability of its results.

Considering the methodological developments in systematics and the current stage of knowledge on the diversity of flowering plants, as exemplified by the order Caryophyllales, we conclude that users and the scientific community need to get into a closer dialogue. This will be essential in order to maintain the long-term quality of information for the users but also to promote application-oriented research in systematics, and thus highlight the relevance of this work. In this sense, the integrated global synthesis of the Caryophyllales can contribute to global initiatives such as the World Flora Online and can be a model how to sustainably organize this contribution through involving the scientific community.

Acknowledgements

We thank A. Güntsch and N. Turland for very helpful comments on a previous version of the manuscript. The Deutsche Forschungsgemeinschaft, the Bundesministerium für Bildung und Forschung (BMBF), the German Federal Agency for Nature Conservation and the European Commission are acknowledged for supporting a range of projects that contributed to the development of ideas towards this paper.

References

- African Plants Database, 2014. Version 3.4.0. Conservatoire et Jardin botaniques de la Ville de Genève and South African National Biodiversity Institute, Pretoria. Published at <http://www.ville-ge.ch/musinfo/bd/cjb/africa/> [last accessed 19 Oct. 2014].
- Agnarsson, I., Kutner, M., 2007. Taxonomy in a changing world: seeking solutions for a science in crisis. *Syst. Biol.* 56, 531–539. <http://dx.doi.org/10.1080/10635150701424546>
- Alberts, B., 2013. Impact factor distortions. *Science* 340, 787.
- Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J., 1990. Basic local alignment search tool. *J. Mol. Biol.* 215, 403–410. [http://dx.doi.org/10.1016/S0022-2836\(05\)80360-2](http://dx.doi.org/10.1016/S0022-2836(05)80360-2)
- APG III, 2009. An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG III. *Bot. J. Linn. Soc.* 161, 105–121. <http://dx.doi.org/10.1111%2Fj.1095-8339.2009.00996.x>
- Beach, J.H., Pramanik, S., Beaman, J.H., 1993. Hierarchic taxonomic databases, in: Fortuner, R. (Ed.), *Advances in computer methods for systematic biology: artificial intelligence, databases, computer vision*. John Hopkins University Press, Baltimore, pp. 241–256.
- Beentje, H.J., 2012. *The Kew Plant Glossary: an illustrated dictionary of plant terms*. Kew Publishing, Richmond.
- 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 <http://wiki.tdwg.org/twiki/bin/view/ABCD/> [last accessed 7 Jan. 2015].
- Berendsohn, W.G., Döring, M., Geoffroy, M., Glück, K., Güntsch, A., Hahn, A., Kusber, W.-H., Li, J.-J., Röpert, D., Specht, F., 2003. The Berlin Taxonomic Information Model. *Schr.reihe Veg.kd.* 39, 15–42.
- Berendsohn, W.G., Geoffroy, M., 2007. Networking taxonomic concepts – uniting without “unitaryism”, in: Curry, G., Humphries, C. (Eds.), *Biodiversity databases - techniques, politics, and applications*. CRC Taylor & Francis, Baton Rouge, pp. 13–22.
- Berendsohn, W.G., 2010. Devising the EDIT Platform for Cybertaxonomy, in: Nimis, P.L., Vignes-Lebbe, R. (Eds.), *Tools for identifying biodiversity: progress and problems*. EUT Edizioni Università di Trieste, Trieste, pp. 1–6. Also published at <http://hdl.handle.net/10077/3737> [last accessed 7 Jan. 2015].
- Berendsohn, W.G., Güntsch, A., Hoffmann, N., Kohlbecker, A., Luther, K., Müller, A., 2011. Biodiversity information platforms: From standards to interoperability. *Zookeys* 150, 71–78. <http://dx.doi.org/10.3897/zookeys.150.2166>
- BHL, 2005- [continuously updated]. Biodiversity Heritage Library. BHL Project. Published at <http://www.biodiversitylibrary.org/> [last accessed 19 Oct. 2014].

- Bisby, F.A., 1993. The ILDIS prototype for Legumes. *Ann. New York Acad. Sci.* 700: 159–164. <http://dx.doi.org/10.1111/j.1749-6632.1993.tb26316.x>
- [Bittrich, V., Kühn, U., 1993. Nyctaginaceae, in: Kubitzki, K., Rohwer, J.G., Bittrich, V. \(Eds.\), The families and genera of vascular plants 2. Springer, Berlin, pp. 473–485.](#)
- Booth, D., Haas, H., McCabe, F., Newcomer, E., Champion, M., Ferris, C., Orchard, D. (Ed.), 2004. W3C Web Services Architecture. W3C Working Group Note 11 February 2004. Published at <http://www.w3.org/TR/ws-arch/> [last accessed 7 Jan. 2015].
- Brach, A.R., Bouffort, D.E., 2011. Why are we still producing paper floras? *Ann. Missouri Bot. Gard.* 98, 297–300. <http://dx.doi.org/10.3417/2010035>
- [Brockington S.F., Alexandre, R., Ramdial, J., Moore, M.J., Crawley, S., Dhingra, A., Hilu, K.W., Soltis, P.S., Soltis, D.E., 2009. Phylogeny of the Caryophyllales sensu lato: Revisiting hypotheses on pollination biology and perianth differentiation in the core Caryophyllales. *Int. J. Pl. Sci.* 170, 627–643.](#)
- Brummitt, R.K., 2001. World geographical scheme for recording plant distributions. Edition 2. Plant Taxonomic Database Standards 2. TDWG, Hunt Institute, Pittsburg. Published at http://www.nhm.ac.uk/hosted_sites/tdwg/TDWG_geo2.pdf [last accessed 19 Oct. 2014].
- Brummitt, R.K., Castrovieja, S., Chikuni, A.C., Orchard, A.E., Smith G.F., Wagner, W.L., 2001. The Species Plantarum Project, an international collaborative initiative for higher plant taxonomy. *Taxon* 50, 1217–1230. <http://www.jstor.org/stable/1224752>
- [Calonje, M., Martín Bravo, S., Dobeš, C., Gong, W., Jordon Thaden, I., Kiefer, C., Kiefer, M., Paule, J., Schmickl, R., Koch, M.A., 2009. Non coding nuclear DNA markers in phylogenetic reconstruction. *Pl. Syst. Evol.* 282, 257–280.](#)
- Campanula Portal 2013-: The Campanula data portal. – Available at <http://dev.e-taxonomy.eu/dataportal/campanulaceae> [last accessed 16 Jan. 2015]
- Cantino, P.D., Queiroz, K. de, 2010. International Code of Phylogenetic Nomenclature, version 4c. Published at <http://www.ohio.edu/phylocode/> [last accessed 7 Jan. 2015].
- CBD, 1998-. Global Taxonomy Initiative. Published at <https://www.cbd.int/gti/> [last accessed 19 Oct. 2014].
- CBD, 2006. Convention on Biological Diversity. Guide to the Global Taxonomy Initiative. CBD Technical Series No. 30. Published at <http://www.cbd.int/doc/publications/cbd-ts-30.pdf> [last accessed 20 Oct. 2014].
- 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> [last accessed 7 Jan 2015].
- [CBOL Plant Working Group, 2009. A DNA barcode for land plants. *Proceedings of the National Academy of Sciences, USA*, 106, 12794–12797.](#)
- Chavan, V., Penev, L., 2011. The data paper - a mechanism to incentivize data publishing in biodiversity science. *BMC Bioinformatics* 12(Suppl. 15): S2. <http://dx.doi.org/10.1186/1471-2105-12-S15-S2>
- 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.
- Clayton, W.D., Vorontsova, M.S., Harman, K.T., Williamson, H., 2006-. GrassBase - The Online World Grass Flora. Published at [last accessed 20 Oct. 2014].
- Costello, M.J., Appeltans, W., Bailly, N., Berendsohn, W.G., de Jong, Y., Edwards, M., Froese, R., Huettmann, F., Los, W., Mees, J., Segers, H., Bisby, F.A., 2014. Strategies for the sustainability of online open-access biodiversity databases. *Biol. Conserv.* 173, 155–165. <http://dx.doi.org/10.1016/j.biocon.2013.07.042>
- [Cox, C.J., Goffinet, B., Wickett, N.J., Boles, S.B., Shaw, A.J., 2010. Moss diversity: A molecular phylogenetic analysis of genera. *Phytotaxa* 9, 175–195.](#)

- Criste, C., Kozłowski, G., Frey, D., Betrisey, S., Maharramova, E., Garfi, G., Pirintsos, S.A., Naeiri, Y., 2014. Footprints of past intensive diversification and structuring in the genus *Zelkova* (Ulmaceae) in south-western Eurasia. *J. Biogeogr.* 41, 1081–1093. <http://dx.doi.org/10.1111/jbi.12276>
- Crouch, N., Smith, G.F., Figueiredo, E., 2013. From checklists to an e-flora for Southern Africa: Past experiences and future prospects for meeting target 1 of the 2020 global strategy for plant conservation. *Ann. Missouri Bot. Gard.* 99, 153–160. <http://dx.doi.org/10.3417/20111107>
- Crowl, A.A., Mavrodiev, E., Mansion, G., Haberle, R., Pistarino, A., Kamari, G., Phitos, D., Borsch, T., Cellinese, N., 2014. Phylogeny of Campanuloideae (Campanulaceae) with emphasis on the utility of nuclear pentatricopeptide repeat (PPR) genes. *PLoS ONE* 9(4):e94199. <http://dx.doi.org/10.1371/journal.pone.0094199>
- Cuénoud P., Savolainen, V., Chatrou, L.W., Powell, M., Grayer, R.J., Chase, M.W., 2002. Molecular phylogenetics of Caryophyllales based on nuclear 18S rDNA and plastid *rbcL*, *atpB*, and *matK* sequences. *Amer. J. Bot.* 89, 132–144.
- Deans, A.R., Yoder, M.J., Balhoff, J.P., 2012. Time to change how we describe biodiversity. *Trends Ecol. Evol.* 27, 78–84. <http://dx.doi.org/10.1016/j.tree.2011.11.007>
- ~~De Queiroz, K., 2005.~~
- De Queiroz, K., Donoghue, M. 2007. Species concepts and species delimitation. *Syst. Biol.* 56, 879–886. <http://dx.doi.org/10.1080/10635150701701083>
- De Queiroz, K., Donoghue, M., 1988. Phylogenetic systematics and the species problem. *Cladistics* 4, 317–338. <http://dx.doi.org/10.1111/j.1096-0031.1988.tb00518.x>
- De Queiroz, K., Gauthier, J., 1992. Phylogenetic taxonomy. *Ann. Rev. Ecol. Syst.* 23, 449–480.
- Dick, C.W. & Kress, W.J., 2009. Dissecting tropical plant diversity with forest plots and a molecular toolkit. *Bioscience* 59, 745–755.
- Dickson, W.C., 2003. Physenaceae, in: Kubitzki, K., Bayer, C. (Eds.), *The families and genera of vascular plants* 5. Springer, Berlin, pp. 332–333.
- Diepenbroek, M., Glöckner, F., Grobe, P., Güntsch, A., Huber, R., König-Ries, B., Kostadinov, I., Nieschulze, J., Seeger, B.; Tolksdorf, R., Triebel, D. 2014. Towards an Integrated Biodiversity and Ecological Research Data Management and Archiving Platform: The German Federation for the Curation of Biological Data (GFBio). – In: Plöedereder, E., Grunke, L., Schneider, E. & Ull, D. (eds): *Informatik 2014 – Big Data Komplexität meistern*. GI-Edition: Lecture Notes in Informatics (LNI) – Proceedings 232: 1711–1724. Köllen Verlag, Bonn.
- Dillenberger M.S., Kadereit J.W. 2014. Maximum polyphyly: Multiple origins and delimitation with plesiomorphic characters require a new circumscription of *Minuartia* (Caryophyllaceae). *Taxon* 63, 64–88.
- Donoghue, M.J., 1985. A critique of the biological species concept and recommendations for a phylogenetic alternative. *The Bryologist* 88, 172–181. <http://www.jstor.org/stable/3243026>
- Droege, G., Barker, K., Astrin, J.J., Bartels, P., Butler, C., Cantrill, D., Coddington, J., Forest, F., Gemeinholzer, B., Hobern, D., Mackenzie-Dodds, J., Tuama, É.Ó., Petersen, G., Sanjur, O., Schindel, D., Seberg, O., 2013. The global genome biodiversity network (GGBN) data portal. *Nucleic Acids Res.* 2013, 1–6. <http://dx.doi.org/10.1093/nar/gkt928>
- Dubois, A., 2010. Taxonomy in the century of extinctions: taxonomic gap, taxonomic impediment, taxonomic urgency. *Taprobanica* 2, 1–5. <http://dx.doi.org/10.4038/tapro.v2i1.2702>

- [Ebach, M.C., Valdecasas, A.G., Wheeler, Q.D., 2011. Impediments to taxonomy and users of taxonomy: accessibility and impact evaluation. *Cladistics* 27, 550–557. <http://dx.doi.org/10.1111/j.1096-0031.2011.00348.x>](#)
- Eggl, U., 1993. Glossary of botanical terms with special reference to succulent plants. British Cactus and Succulent Society, Richmond.
- Eichler, A.W., 1878. Blüthendiagramme, Zweiter Teil. W. Engelmann, Leipzig.
- eMonocot, 2012-. An online resource for monocot plants. Published at <http://e-monocot.org/> [last accessed 20 Oct. 2014].
- eMonocot Team, 2012a. eMonocot Cyperaceae. Published at <http://cyperaceae.e-monocot.org/> [last accessed 20 Oct. 2014].
- eMonocot Team, 2012b. CATE Araceae. Published at <http://araceae.e-monocot.org/> [last accessed 20 Oct. 2014].
- [Endress, M.E., Bittrich, V., 1993. Molluginaceae, in: Kubitzki, K., Rohwer, J.G., Bittrich, V. \(Eds.\), *The families and genera of vascular plants 2*. Springer, Berlin, pp. 419–425.](#)
- EOL, 2014. Encyclopedia of Life. Published at <http://www.eol.org> [last accessed 19 Oct. 2014].
- Ericksson, R., 2007. A synopsis of Basellaceae. *Kew Bull.* 62, 297–320.
- [Farr, E.R., Zijlstra, G. \(Eds.\), 1996. *Index Nominum Genericorum \(Plantarum\)*. Published at <http://botany.si.edu/ing/> \[last accessed 19 Oct. 2014\].](#)
- Flores-Rentería, L., Wegier, A., Ortega Del Vecchyo, D., Ortíz-Medrano, A., Piñero, D., Whipple, A.V., Molina-Freaner, F., Domínguez, C.A., 2013. Genetic, morphological, geographical and ecological approaches reveal phylogenetic relationships in complex groups, an example of recently diverged pinyon pine species (Subsection *Cembroides*). *Mol. Phylogenet. Evol.* 69, 940–949. <http://dx.doi.org/10.1016/j.ympev.2013.06.010>
- [Forrest, L.L., Davis, E.C., Long, D.G., Crandall-Stotler, B.J., Clark, A., Hollingsworth, M.L., 2006. Unraveling the evolutionary history of the liverworts \(Marchantiophyta\): multiple taxa, genomes and analyses. *Bryologist* 109, 303–334. \[http://dx.doi.org/10.1639/0007-2745\\(2006\\)109\\[303:UTEHOT\\]2.0.CO;2\]\(http://dx.doi.org/10.1639/0007-2745\(2006\)109\[303:UTEHOT\]2.0.CO;2\)](#)
- Franz, N.M., 2005. On the lack of good scientific reasons for the growing phylogeny/classification gap. *Cladistics* 21, 495–500. <http://dx.doi.org/10.1111/j.1096-0031.2005.00080.x>
- Franz, N.M., Peet, R.K., 2009. Towards a language for mapping relationships among taxonomic concepts. *Syst. Biodiv.* 7, 5–20. <http://dx.doi.org/10.1017/S147720000800282X>
- Franz, N.M., Cardona-Duque J., 2013. Description of two new species and phylogenetic reassessment of *Perelleschus* O'Brien & Wibmer, 1986 (Coleoptera: Curculionidae), with a complete taxonomic concept history of *Perelleschus* sec. Franz & Cardona-Duque, 2013. *Syst. Biodiv.* 11, 209–236. <http://dx.doi.org/10.1080/14772000.2013.806371>
- Franz, N.M., Thau, D. 2010. Biological taxonomy and ontology development: scope and limitations. *Biodiversity Informatics* 7, 45–66. Also published at <https://journals.ku.edu/index.php/jbi/article/view/3927/3852> [last accessed 7 Jan 2015]
- [Fuentes-Bazan, S., Uotila, P., Borsch, T., 2012. A novel phylogeny-based generic classification for *Chenopodium sensu lato*, and a tribal rearrangement of *Chenopodioideae* \(Chenopodiaceae\). *Willdenowia* 42, 5–24. <http://dx.doi.org/10.3372/wi.42.42101et al. 2012>](#)
- Funk, V.A., 2006. Floras: a model for biodiversity studies or a thing of the past? *Taxon* 55, 581–588. Also published at <http://www.ingentaconnect.com/content/iapt/tax/2006/00000055/00000003/art00006> [last accessed 7 Jan. 2015].

- Gaskin, J.F., 2003. Tamaricaceae, in: Kubitzki, K., Bayer, C. (Eds.), *The families and genera of vascular plants 5*. Springer, Berlin, pp. 363–368.
- GBIF, 2004-. Global Biodiversity Information Facility. Published at <http://www.gbif.org> [last accessed 20 Oct. 2014].
- Gemeinholzer, B., Dröge, G., Zetzsche, H., Haszprunar, G., Klenk, H.-P., Güntsch, A., Berendsohn, W.G., Wägele, J.-W., 2011. The DNA Bank Network: The start from a German initiative. *Biopreserv. Biobank* 9, 51–55. <http://dx.doi.org/10.1089/bio.2010.0029>
- Gilbert, M.G., 1993. A review of *Gisekia* (Gisekiaceae). *Kew Bull.* 48, 343–356. <http://dx.doi.org/10.2307/4117942>
- Godfray, H.C.J., 2002. Challenges for taxonomy. *Nature* 417, 17–19. <http://dx.doi.org/10.1038/417017a>
- Godfray, H.C.J., Clark, B.R., Kitching, I.J., Mayo, S.J., Scoble, M.J., 2007. The web and the structure of taxonomy. *Syst. Biol.* 56, 943–955. <http://dx.doi.org/10.1080/10635150701777521>
- Goldstein, P.Z., DeSalle, R., 2010. Phylogenetic species, nested hierarchies, and character fixation. *Cladistics* 16, 364–384. <http://dx.doi.org/10.1006/clad.2000.0141>
- González Gutiérrez, P.A., Köhler, E., Borsch, T., 2013. New species of *Buxus* (Buxaceae) from northeastern Cuba based on morphological and molecular characters, including some comments on molecular diagnosis. *Willdenowia* 43, 125–137. <http://dx.doi.org/10.3372/wi.43.43115>
- 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. <http://dx.doi.org/10.2307/1224722>
- ~~GrassBase 2006 bzw. siehe Clayton et al. 2006~~
- Grass Phylogeny Working Group, 2001. Phylogeny and subfamilial classification of the grasses (Poaceae). *Ann. Missouri Bot. Gard.* 88, 373–457.
- ~~Greuter, W., Burdet, H.M., Long, G. (Eds.), 2006. Med Checklist—A critical inventory of vascular plants of the circum-mediterranean countries (electronic version). Published at <http://www2.bgbm.org/mcl/home.asp> [last accessed 19 Oct. 2014].~~
- Greuter, W., Zimmer, B., Berendsohn, W.G. (Eds.), 1997. NCU-3e. Names in current use for extant plant genera. Electronic version 1.0. Published at <http://www.bgbm.org/iapt/ncu/genera/> [last accessed 19 Oct. 2014].
- ~~Güntsch, A., Kelbert, P., Eckert, S., Patterson, D., Penev, L., Hagedorn, G. 2013. Strategies for improved cooperation and interoperability between infrastructures. Pro iBiosphere Report D4.1 Published at http://wiki.pro-ibiosphere.eu/w/media/1/13/Pro-ibiosphere_WP4_FUB_BGBM_VFF_20122013.pdf [last accessed 8 Jan. 2015].~~
- Gurushidze, M., Fritsch, R.M., Blattner, F.R., 2010. Species-level phylogeny of *Allium* subgenus *Melanocrommyum*: Incomplete lineage sorting, hybridization and trnF gene duplication. *Taxon* 59, 829–840. <http://www.jstor.org/stable/25677671>
- Hagedorn, G., Thiele, K., Morris, R., Heidorn, P.B., 2005. The Structured Descriptive Data (SDD) w3c-xml-schema, version 1.0. Published at <http://www.tdwg.org/standards/116/> [last accessed 19 Oct. 2014].
- 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. <http://dx.doi.org/10.12705/632.11>
- Hardisty A., Roberts, D., 2013. A decadal view of biodiversity informatics: challenges and priorities. *BMC Ecology* 13, e16. <http://dx.doi.org/10.1186/1472-6785-13-16>
- Hartmann, H.E.K. (Ed.), 2001a. Aizoaceae A-E, in: Egli, U., Hartmann, H.E.K. (Eds.): *Illustrated handbook of succulent plants*. Springer, Heidelberg.

- Hartmann, H.E.K. (Ed.), 2001b. Aizoaceae F-Z, in: Egli, U., Hartmann, H.E.K. (Eds.): Illustrated handbook of succulent plants. Springer, Heidelberg.
- Hebert, P.D.N., Cywinska, A., Ball, S.L., deWaard, J.R., 2003. Biological identifications through DNA barcodes. *Proc. R. Soc. Lond. B* 270, 313–321. <http://dx.doi.org/10.1098/rspb.2002.2218>
- Hennig, W., 1966. Phylogenetic systematics. Univ. Illinois Press, Urbana.
- Hillis, M.H., Thieret J.W., Morefield J.D., 2003. *Sarcobatus*, in: Flora of North America Editorial Committee (Eds.), *Flora of North America 4*. Oxford University Press, New York and Oxford, pp. 387–389. Published at http://www.efloras.org/florataxon.aspx?flora_id=1&taxon_id=129190 [last accessed 7 Jan. 2015].
- Hilu K.W., Borsch, T., Müller, K., Soltis, D.E., Soltis, P.S., Savolainen, V., Chase, M.W., Powell, M.P., Alice, L.A., Evans, R., Sauquet, H., Neinhuis, C., Slotta, T.A.B., Rohwer, J.G., Campbell, C.S., Chatrou, L.W., 2003. Angiosperm phylogeny based on *matK* sequence information. *Amer. J. Bot.* 90, 1758–1776.
- Hoborn, D., Apostolico, A., Arnaud, E., Bello, J.C., Canhos, D., Dubois, G., Field, D., García, E.A., Hardisty, A., Harrison, J., Heidorn, B., Krishtalka, L., Mata, E., Page, R., Parr, C., Price, J., Willoughby, S. 2013. Global Biodiversity Informatics Outlook – Delivering Biodiversity Knowledge in the Information age. Copenhagen. Published at www.biodiversityinformatics.org/ [last accessed 28 dec. 2014].
- Hunt, D., 2006. The new cactus lexicon. Dh Books, Milborne Port.
- Holetsek, J., Dröge, G., Güntsch, A., Berendsohn, W., 2012. The ABCD of primary biodiversity data access. *Plant Biosyst.* 146, 771–779. <http://dx.doi.org/10.1080/11263504.2012.740085>
- ICN (Hand, R., Kilian, N., Raab-Straube, E. von; general eds.), 2009- [continuously updated]. International Cichorieae Network: Cichorieae Portal. Published at <http://wp6-cichorieae.e-taxonomy.eu/portal/> [last accessed 17 Oct. 2014].
- ILDIS, 1996- International Legume Database & Information Service. Published at <http://www.ildis.org/> [last accessed 19 Oct. 2014].
- IOPI 1997-2003. Provisional Global Plant Checklist. Published at <http://www.bgbm.org/iopi/gpc/default.asp> [last accessed 2 Jan 2015].
- IPNI, 2004- [continuously updated]. The International Plant Names Index. The plant names project. Published at <http://www.ipni.org> [last accessed 2 Jan. 2014].
- Jakob, S., Blatter, F.R., 2006. A chloroplast genealogy of *Hordeum* (Poaceae): Long term persisting haplotypes, incomplete lineage sorting, regional extinction, and the consequences for phylogenetic inference. *Mol. Biol. Evol.* 23, 1602–1612. <http://dx.doi.org/10.1093/molbev/msl018>
- Joppa, L.N., Roberts, D.L., Myers, N., Pimm, S.L., 2011. Biodiversity hotspots house most undiscovered plant species. *Proc. Natl. Acad. Sci. U.S.A.* 108, 13171–13176. <http://dx.doi.org/10.1073/pnas.1109389108>
- JSTOR, 2000- [continuously updated]. JSTOR. ITHACA. Published at <http://www.jstor.org/> [last accessed 19 Oct. 2014].
- JSTOR, 2013- [continuously updated]. JSTOR Global Plants. New York. Published at <http://about.jstor.org/download/file/fid/22128> [last accessed 19 Oct. 2014]
- Karam, N., Fichtmüller, D. 2014. GFBio Terminology Server. Published at <http://terminologies.gfbio.org/> [last accessed 8 Jan 2015].
- Karsch-Mizrachi, I., Nakamura, Y., Cochrane, G. 2011. The International Nucleotide Sequence Database Collaboration". *Nucleic Acids Res.* 40 (Database issue), D33–D37. <http://dx.doi.org/10.1093/nar/gkr1006>
- KEsoftware, 2014. About KE Software's Electronic Museum management system, EMu. Published at <http://emu.kesoftware.com/about-emu/> [last accessed 19 Oct. 2014].

- Kennedy, J., Hyam, R., Kukla, R., Paterson, T., 2006. Standard data model representation for taxonomic information. *OMICS: A Journal of Integrative Biology* 10, 220–230. <http://dx.doi.org/10.1089/omi.2006.10.220>
- [Kiefer et al., 2013](#)
- Kim, S-T., Sultan, S., Donoghue, M.J., 2008. Allopolyploid speciation in *Persicaria* (Polygonaceae): insights from a low-copy nuclear region. *Proc. Natl. Acad. Sci. USA* 105, 12370–12375. <http://dx.doi.org/10.1073/pnas.0805141105>
- Kirschner, J., Kaplan, Ž., 2002. Taxonomic monographs in relation to global red lists. *Taxon* 51, 155–158. <http://dx.doi.org/10.2307/1554973>
- Koch, M.A., Kiefer, M., German, D.A., Al-Shehbaz, I.A., Franzke, A., Mummenhoff, K., Schmickl, R., 2012. BrassiBase: Tools and biological resources to study characters and traits in the Brassicaceae – version 1.1. *Taxon* 61, 1001–1009. Also published at <http://www.ingentaconnect.com/content/iapt/tax/2012/00000061/00000005/art00007> [last accessed 7 Jan 2015].
- [Korotkova, N., Borsch, T., Quandt, D., Taylor, N.P., Müller, K.F., Barthlott, W., 2011. What does it take to resolve relationships and to identify species with molecular markers? An example from the epiphytic *Rhipsalideae* \(Cactaceae\). *Am. J. Bot.* 98, 1549–1572. <http://dx.doi.org/10.3732/ajb.1000502>](#)
- Kubitzki, K., Bayer, C. (Eds.), *The families and genera of vascular plants 5*. Springer, Berlin, pp. 28–29.
- [Kubitzki, K., 2003b. Droseraceae, in: Kubitzki, K., Bayer, C. \(Eds.\), *The families and genera of vascular plants 5*. Springer, Berlin, pp. 198–202.](#)
- [Kubitzki, K., 2003c. Drosophyllaceae, in: Kubitzki, K., Bayer, C. \(Eds.\), *The families and genera of vascular plants 5*. Springer, Berlin, pp. 203–205.](#)
- [Kubitzki, K., 2003d. Frankeniaceae, in: Kubitzki, K., Bayer, C. \(Eds.\), *The families and genera of vascular plants 5*. Springer, Berlin, pp. 206–212.](#)
- [Kubitzki, K., 2003e. Nepenthaceae, in: Kubitzki, K., Bayer, C. \(Eds.\), *The families and genera of vascular plants 5*. Springer, Berlin, pp. 320–324.](#)
- [Kubitzki, K., 2003f. Simmondsiaceae, in: Kubitzki, K., Bayer, C. \(Eds.\), *The families and genera of vascular plants 5*. Springer, Berlin, pp. 355–358.](#)
- Kubitzki, K., Rohwer, J.G., Bittrich, V. (Eds.), *The families and genera of vascular plants 2*. Springer, Berlin, pp. 253–280.
- [Legume Phylogeny Working Group, 2013. Legume phylogeny and classification in the 21st century: progress, prospects and lessons for other species-rich clades. *Taxon* 62, 217–248.](#)
- [Lipscomb, B.L., 2003. Achatocarpaceae, in: *Flora of North America Editorial Committee \(Eds.\), Flora of North America 4*. Oxford University Press, New York and Oxford, pp. 12–13.](#)
- [LPWG, 2013. The Legume Phylogeny Working Group. Legume phylogeny and classification in the 21st century: Progress, prospects and lessons for other species-rich clades. *Taxon* 62, 217–248. <http://dx.doi.org/10.12705/622.8>](#)
- Mace, G.M., 2004. The role of taxonomy in species conservation. *Phil. Trans. R. Soc. B* 359, 711–719. <http://dx.doi.org/10.1098/rstb.2003.1454>
- Maddison, D.R., Swofford D.L., Maddison, W.P., 1997. NEXUS: An extensible file format for systematic information. *Syst. Biol.* 46, 590–621. <http://dx.doi.org/10.1093/sysbio/46.4.590>
- Mallet, J., 2007. Species, concepts of, in: Levin, S. et al. (Eds.), *Encyclopedia of Biodiversity 5*. Academic Press, pp. 427–440. Online update 1, pp. 1–15, Elsevier, Oxford.
- Mallet, J., Willmott, K., 2003. Taxonomy: renaissance or Tower of Babel? *Trends Ecol. Evol.* 18, 57–59. [http://dx.doi.org/10.1016/S0169-5347\(02\)00061-7](http://dx.doi.org/10.1016/S0169-5347(02)00061-7)

- Mansion, G., Parolly, G., Crowl, A.A., Mavrodiev, E., Cellinese, N., Oganessian, M., Fraunhofer, K., Kamari, G., Phitos, D., Haberle, R., Akaydin, G., Ikinci, N., Raus, T., Borsch, T., 2012. How to handle speciose clades? Massive taxon sampling as a strategy towards illuminating the natural history of *Campanula* (Campanuloideae). *PLoS One* 7(11), e50076. <http://dx.doi.org/10.1371/journal.pone.0050076>.
- Marhold, K., Stuessy, T. (Eds.) in collaboration with Agababian, M., Agosti, D., Alford, M.H., Crespo, A., Crisci, J.V., Dorr, L.J., Ferencová, Z., Frodin, D., Geltman, D.V., Kilian, N., Linder, H.P., Lohmann, L.G., Oberprieler, C., Penev, L., Smith, G.F., Thomas, W., Tulig, M., Turland, N., Zhang, X.-C., 2013. The future of botanical monography. Report from an international workshop, 12-16 March 2012, Smolenice, Slovak Republic. *Taxon* 62, 4–20. Also published at <http://www.ingentaconnect.com/content/iapt/tax/2013/00000062/00000001/art00003> [last accessed 7 Jan 2015].
- Mathew C. & Güntsch, A. 2013. A Workflow for Data Refinement. *BioVeL Newsletter* No. 2, Autumn 2012: 3-5. Published at <https://www.biovel.eu/images/publications/newsno2autumn12.pdf> [last accessed 7 Jan 2015].
- Mathew, C., Güntsch, A., Obst, M., Vicario, S., Haines, R., Williams, A.R., de Jong, Y., Goble, C., 2014. A semi-automated workflow for biodiversity data retrieval, cleaning, and quality control. *Biodiversity Data Journal* 2, e4221. <http://dx.doi.org/10.3897/BDJ.2.e4221>
- Mayo, S.J., Allkin, R., Baker, W., Blagoderov, V., Brake, I., Clark, B., Govaerts, R., Godfray, C., Haigh, A., Hand, R., Harman, K., Jackson, M., Kilian, N., Kirkup, D.W., Kitching, I., Knapp, S., Lewis, G.P., Malcolm, P., Raab-Straube, E. von, Roberts, D.M., Scoble, M., Simpson, D.A., Smith, C., Smith, V., Villalba, S., Walley, L., Wilkin, P., 2008. Alpha e-taxonomy: responses from the systematics community to the biodiversity crisis. *Kew Bull.* 63, 1–16. <http://dx.doi.org/10.1007/s12225-008-9014-1>
- Mayr, E., 1969. *Principles of Systematic Zoology*. McGraw-Hill, New York.
- McDade, L.A., 1995. Species concepts and problems in practice: Insight from botanical monographs. *Syst. Bot.* 20, 606–622. <http://dx.doi.org/10.2307/2419813>
- McNeill, J., Barrie, F.R., Buck, W.R., Demoulin, V., Greuter, W., Hawksworth, D.L., Herendeen, P.S., Knapp, S., Marhold, K., Prado, J., Prud'homme van Reine, W.F., Smith, G.F., Wiersema, J.H., Turland, N.J. (Eds.), 2012. International code of nomenclature for algae, fungi, and plants (Melbourne Code) adopted by the Eighteenth International Botanical Congress Melbourne, Australia, July 2011. *Gantner, Ruggell [Regnum Veg. 154]*.
- MELNet, 2007-14. *Melastomataceae.Net*. A Site with information on the biodiversity of Melastomataceae. Published at www.melastomataceae.net [last accessed 19 Oct. 2014].
- [Mison, N.R., 2005. Plumbaginaceae, in: Flora of North America Editorial Committee \(Eds.\), Flora of North America 5. Oxford University Press, New York and Oxford, pp. 602–611.](#)
- [Monaghan, M.T., Wild, R., Elliot, M., Fujisawa, T., Balke, M., Inward, D.J.G., Lees, D.C., Ranivosolo, R., Eggleton, P., Barraclough, T.G., Vogler, A.P., 2009. Accelerated species inventory on Madagascar using coalescent-based models of species delineation. Syst. Biol. 58, 298–311. http://dx.doi.org/10.1093/sysbio/syp027](#)
- [Moore, G. 2003. Should taxon names be explicitly defined? Bot. Rev. 69, 2-21.](#)
- [Moore et al., 2014](#)
- Morýš, S. A., Prance, G.T., 2006-2010. *The Lecythidaceae Pages*. The New York Botanical Garden, Bronx, New York.

- Morya, S. A., Smith, N.P., Cornejo, X., Prance, G.T., 2010-. The Lecythidaceae Pages (<http://sweetgum.nybg.org/lp/index.php>). The New York Botanical Garden, Bronx, New York. Published at <http://sweetgum.nybg.org/lp/index.html> [last accessed 7 Jan. 2015].
- Munzinger, J., Lowry II, P.P., Callmänder, M.W., Buerki, S., 2013. A taxonomic revision of the endemic New Caledonian genus *Podonephelium* Baill. (Sapindaceae). *Syst. Bot.* 38, 1105–1124. <http://dx.doi.org/10.1600/036364413X674814>
- [Naciri, Y., & Linder, P.L., 2015. 2015. Species delimitation and relationships: The dance of the seven veils. *Taxon* 64, 3–16. <http://dx.doi.org/10.12705/641.24>](#)
- [Naciri & Linder 2015](#)
- Nic Lughadha, E., Miller, C., 2009. Accelerating global access to plant diversity information. *Trends Pl. Sci.* 14, 622–628. <http://dx.doi.org/10.1016/j.tplants.2009.08.014>
- [Nikiforova, S.V., Cavalieri, D., Velasco, R., Goremykin, V., 2013. Phylogenetic analysis of 47 chloroplast genomes clarifies the contribution of wild species to the domesticated apple maternal line. *Mol. Biol. Evol.* 30, 1751–1760. <http://dx.doi.org/10.1093/molbev/mst092>](#)
- [Nixon, K.C., Carpenter, J.M., Stevenson, D.W., 2003. The PhyloCode is fatally flawed, and the “Linnaean system” can easily be fixed. *Bot. Rev.* 69, 111–120.](#)
- Nixon, K.C., Wheeler, Q.D., 2008. An amplification of the phylogenetic species concept. *Cladistics* 6, 211–223. <http://dx.doi.org/10.1111/j.1096-0031.1990.tb00541.x>
- [Njuguna, W., Liston, A., Cronn, R., Ashman, T.L., Bassil, N., 2013. Insights into phylogeny, sex function and age of *Fragaria* based on whole chloroplast genome sequencing. *Mol. Phylogenet. Evol.* 66, 17–29. <http://dx.doi.org/10.1016/j.ympev.2012.08.026>](#)
- Nyffeler, R., Eggli, U., 2010a. Disintegrating Portulacaceae: A new familial classification of the suborder Portulacinae (Caryophyllales) based on molecular and morphological data. *Taxon* 59, 227–240. <http://www.jstor.org/stable/27757065>
- Nyffeler, R., Eggli U., 2010b: A farewell to dated ~~ideas~~ and concepts - molecular phylogenetics and a revised suprageneric classification of the family Cactaceae. *Schumannia* 6 = *Biodiv. Ecol.* 3, 109–149.
- Oxelman, B., Rautenberg, A., Thollesson, M., Larsson, A., Frajman, B., Eggens, F., Petri, A., Aydin, Z., Töpel, M., Brandtberg-Falkman, A. 2013. Sileneae taxonomy and systematics. Published at <http://www.sileneae.info> [last accessed 5 Jan. 2015].
- [Payne, J.L., Smith, F.A., Kowalewski, M., Krause Jr., R.A., Boyer, A.G., McClain, C.R., Finnegan, S., Novack-Gottshall, P.M., Sheble, L., 2012. A lack of attribution: closing the citation gap through a reform of citation indexing practices. *Taxon* 61, 1349–1354. Also published at <http://www.ingentaconnect.com/content/iapt/tax/2012/00000061/00000006/art00030> \[last accessed 7 Jan. 2015\].](#)
- Palmweb, 2010-. Palmweb: Palms of the World Online. Published at <http://www.palmweb.org/> [last accessed 19 Oct. 2014].
- Patterson, D.J., Cooper, J., Kirk, P.M., Pyle R.L., Remsen, D.P., 2010. Names are key to the big new biology. *Trends Ecol. Evol.* 25, 686–691. <http://dx.doi.org/10.1016/j.tree.2010.09.004>
- Paton, A., 2009. Biodiversity informatics and the plant conservation baseline. *Trends Pl. Sci.* 14, 629–637. <http://dx.doi.org/10.1016/j.tplants.2009.08.007>
- Pendry, C.A., Watson, M.F., 2009. Flora Malesiana and the Flora of Nepal: floristic links and the potential for collaboration. *Blumea* 54, 18–22. <http://dx.doi.org/10.3767/000651909X474050>
- PESI, 2014. Pan-European Species directories Infrastructure. Published at <http://www.eunomen.eu/portal>. [last accessed 19 Oct. 2014]

- Petersen, G., Aagesen, L., Seberg, O., Larsen, I.H., 2011. When is enough, enough in phylogenetics? A case in point from *Hordeum* (Poaceae). *Cladistics* 27, 428–446. <http://dx.doi.org/10.1111/j.1096-0031.2011.00347.x>
- Pullan, M.R., Watson, M.F., Kennedy, J.B., Raguenaud, C., Hyam, R., 2000. The Prometheus Taxonomic Model: a practical approach to representing multiple classifications. *Taxon* 49, 55–75. <http://dx.doi.org/10.2307/1223932>
- Porembski, S., 2003. *Ancistrocladaceae*, in: Kubitzki, K., Bayer, C. (Eds.), *The families and genera of vascular plants 5*. Springer, Berlin, pp. 25–27.
- Porembski, S., Barthlott, W., 2003. *Dioneophyllaceae*, in: Kubitzki, K., Bayer, C. (Eds.), *The families and genera of vascular plants 5*. Springer, Berlin, pp. 178–181.
- Prance, G.T., 2003. *Rhabdodendraceae*, in: Kubitzki, K., Bayer, C. (Eds.), *The families and genera of vascular plants 5*. Springer, Berlin, pp. 339–341.
- Rabeler, R.K., Hartman, R.L., 2005. *Caryophyllaceae*, in: *Flora of North America Editorial Committee (Eds.), Flora of North America 5*. Oxford University Press, New York and Oxford, pp. 3–8.
- Rohwer, J.G., 1993a. *Phytolaccaceae*, in: Kubitzki, K., Rohwer, J.G., Bittrich, V. (Eds.), *The families and genera of vascular plants 2*. Springer, Berlin, pp. 506–515.
- Rohwer, J.G., 1993b. *Stegnospermaceae*, in: Kubitzki, K., Rohwer, J.G., Bittrich, V. (Eds.), *The families and genera of vascular plants 2*. Springer, Berlin, pp. 592–594.
- Rieppel, O., 2009. Species monophyly. *J. Zoolog. Syst. Evol. Res.* 48, 1–8. <http://dx.doi.org/10.1111/j.1439-0469.2009.00545.x>
- Rieseberg L.H., Carter I.R., Zona, S., 1990. Molecular tests of the hypothesized hybrid origin of two diploid *Helianthus* species (Asteraceae). *Evolution* 44, 1498–1511. <http://dx.doi.org/10.2307/2409332>
- Rina R. and Berry, P.E. (coordinators). 2014. *Euphorbia plantetary biodiversity inventory database*. <http://app.tolkin.org/projects/72/taxa> (updated continuously).
- Roos, M.C., Berendsohn, W.G., Desein, 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. *Gardens' Bulletin Singapore* 63, 189–195. <http://biostor.org/reference/140281>
- Roskov, Y., Kunze, T., Orrell, T., Abucay, L., Paglinawan, L., Culham, A., Bailly, N., Kirk, P., Bourgoin, T., Baillargeon, G., Decock, W., De Wever, A., Didžiulis, V. (Eds.), 2014. *Species 2000 & ITIS Catalogue of Life, 2014 Annual Checklist*. DVD. Species 2000, Naturalis, Leiden, the Netherlands.
- Rothfels, C.J., Larsson, A., Kuo, L. Y., Korall, P., Chiou, W. L., Pryer, K.M., 2012. Overcoming deep roots, fast rates, and short internodes to resolve the ancient rapid radiation of eupolypod II ferns. *Syst. Biol.* 16, 490–509. <http://dx.doi.org/10.1093/sysbio/sys001>
- Särkinen, T., Olmstead, R.G., Bohs, L., Knapp, S., 2013. A phylogenetic framework for evolutionary study of the nightshades (Solanaceae): a dated 1000-tip tree. *BMC Evol. Biol.* 13, 212. <http://dx.doi.org/10.1186/1471-2148-13-214>
- Sanchez, A., Schuster, T.M., Burke, J.M., Kron, K., 2011. Taxonomy of Polygonoideae (Polygonaceae): a new tribal classification. *Taxon* 60, 151–160. Also published at <http://www.ingentaconnect.com/content/iapt/tax/2011/00000060/00000001/art00013> [last accessed 7 Jan. 2015].
- Sánchez-del Pino, I., Borsch, T., Motley, T.J., 2009. trnL-F and rpl16 sequence data and dense taxon sampling reveal monophyly of unilocular anthered Gomphrenoideae (Amaranthaceae) and an improved picture of their internal relationships. *Syst. Bot.* 34, 57–67. <http://dx.doi.org/10.1600/036364409787602401>

- Schäferhoff, B., Müller, K. F., Borsch, T., 2009. Caryophyllales phylogenetics: disentangling Phytolaccaceae and Molluginaceae and description of Microteaceae as a new isolated family. *Willdenowia* 39, 209–228. <http://dx.doi.org/10.3372/wi.39.39201>
- Schuettpeitz, E., Pryer, K.M., 2007. Fern phylogeny inferred from 400 leptosporangiate species and three plastid genes. *Taxon* 56, 1037–1050. <http://dx.doi.org/10.2307/25065903>.
- Seberg, O., Petersen, G., 2009. How many loci does it take to DNA barcode a *Crocus*? *PLoS One* 4: e4598. <http://dx.doi.org/10.1371/journal.pone.0004598>
- Sites, J.W., Marshall, J.C., 2004. Operational criteria for delimiting species. *Ann. Rev. Ecol. Evol. Syst.* 35, 199–227. <http://dx.doi.org/10.1146/annurev.ecolsys.35.112202.130128>.
- Škondrić, S., Aleksić, J.M., Lukušić, D., 2014. *Campanula cichoracea* (Campanulaceae), a neglected species from the Balkan-Carpathian *C. lingulata* complex as inferred from molecular and morphological characters. *Willdenowia* 44, 77–96. <http://dx.doi.org/10.3372/wi.44.44111>
- Small R.L., Ryburn J.A., Cronn R.C., Seelanan T., Wendel J.F., 1998. The tortoise and the hare: Choosing between noncoding plastome and nuclear ADH sequences for phylogeny reconstruction in a recently diverged plant group. *Amer. J. Bot.* 85, 1301–1315. <http://dx.doi.org/10.2307/2446640>.
- Smith, G.F., Figueiredo E., 2013. Type specimens online: what is available, what is not, and how to proceed. Reflections based on an analysis of the type specimens of southern African *Polygala* L. (Polygalaceae) accessible on the worldwide web. *Taxon*. 62, 801–806. <http://dx.doi.org/10.12705/624.5>
- Smith, V.S., Rycroft, S., Brake, I., Scott, B., Baker, E., Livermore, L., Blagoderov, V., Roberts, D., 2011. Scratchpads 2.0: a Virtual Research Environment supporting scholarly collaboration, communication and data publication in biodiversity science. *Zookeys* 150: 53–70, <http://dx.doi.org/10.3897/zookeys.150.2193>
- Solanaceae Source 2011-. A global taxonomic resource for the nightshade family. Published at <http://www.solanaceaesource.org/> [last accessed on 28 Dec. 2014].
- Soltis, P.S., Soltis, D.E., 2009. The role of hybridization in plant speciation. *Ann. Rev. Plant Biol.* 60, 561–588. <http://dx.doi.org/10.1146/annurev.arplant.043008.092039>
- Spellenberg, R.W., 2003. *Nyctaginaceae in Flora of North America* Editorial Committee (Eds.), *Flora of North America* 4. Oxford University Press, New York and Oxford, pp.14–74.
- Staats, M., Erkens, R.H.J., Vossenbergh, B., van de Wieringa, J.J., Kraaijeveld, K., Stielow, B., Geml, J., Richardson, J.E., Bakker, F.T., 2013. Genomic treasure troves: Complete genome sequencing of herbarium and insect. *PLoS ONE* 8, e69189. <http://dx.doi.org/10.1371/journal.pone.0069189>
- Stuessy, T.F., 1990. *Plant Taxonomy. The systematic evaluation of comparative data.* Columbia University Press, New York, 514 S.
- Stuessy, T. F., Lack, H. W. (Eds.), 2011. *Monographic plant systematics. Fundamental assessment of plant biodiversity.* Koeltz, Königstein [Regnum Veg. 153].
- TPL, 2013. *The Plant List. Version 1.1.* Published at <http://www.theplantlist.org/> [last accessed 19 Oct. 2014].
- Triebel, D., Hagedorn, G., Jablonski, S., Rambold, G. (Eds.), 1999-. *Diversity Workbench – A virtual research environment for building and accessing biodiversity and environmental data.* Published at <http://www.diversityworkbench.net> [last accessed 19 Oct. 2014].
- Tropicos, 2014. *Tropicos.org.* Missouri Botanical Garden. Published at <http://www.tropicos.org> [last accessed 19 Oct. 2014].

- Tschöpe, O., Macklin, J.A., Morris, R.A., Suhrbier, L., Berendsohn, W.G., 2013. Annotating biodiversity data via the internet. *Taxon* 62, 1248–1258.
<http://dx.doi.org/10.12705/626.4>
- Turland, N.J., 2013. The code decoded. A user's guide to the International Code of Nomenclature for algae, fungi, and plants. Koeltz, Königstein [Regnum Veg. 155].
- Ung, V., Dubus, G., Zaragüeta-Bagils, R., Vignes-Lebbe, R., 2010. Xper²: introducing e-taxonomy. *Bioinformatics* 26, 703–704.
<http://dx.doi.org/10.1093/bioinformatics/btp715>
- Ung, V., Causse, F., Vignes Lebbe, R., 2010a. Xper2: managing descriptive data from their collection to e-monographs, in: Nimis, P.L., Vignes-Lebbe, R. (Eds.), *Tools for identifying Biodiversity: Progress and Problems*. EUT Edizioni Università di Trieste, Trieste, pp. 113–120. Also published at
http://dbiodbs1.units.it/bioidentify/files/volume_bioidentify_low.pdf [last accessed 7 Jan. 2015].
- Vences, M., Guayasamin, J.M., Miralles, A., de la Riva, I., 2013. To name or not to name. Criteria to promote economy of change in Linnaean classification schemes. *Zootaxa* 3636, 201–244. <http://dx.doi.org/10.11646/zootaxa.3636.2.1>
- Venin, M., Kirchoff, 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, in: Nimis, P.L., Vignes-Lebbe, R. (Eds.), *Tools for identifying Biodiversity: Progress and Problems*. EUT Edizioni Università di Trieste, Trieste, pp. 7–11. <http://hdl.handle.net/10077/3738>
- Vivrette, N.J., Bleck, J.E., Ferren Jr., W.R., 2003. *Aizoaceae*, in: *Flora of North America Editorial Committee (Eds.), Flora of North America 4*. Oxford University Press, New York and, pp. 75–91.
- Von Staden, L., Raimondo, D., Dayaram, A., 2013. Taxonomic research priorities for the conservation of the South African flora. *S. Afr. J. Sci.* 109, 1–10.
<http://dx.doi.org/10.1590/sajs.2013/1182>
- Vorontsova, M.S., Simon, B.K., 2012. Updating classifications to reflect monophyly: 10 to 20 percent of species names change in Poaceae. *Taxon* 61, 735–746. Also published at <http://www.ingentaconnect.com/content/iapt/tax/2012/00000061/00000004/art00003> [last accessed 7 Jan. 2015].
- Wägele, H., Klussmann Kolb, A., Kuhlmann, M., Haszprunar, G., Lindberg, D., Koch, A., Wägele, J.W., 2011. The taxonomist—an endangered race. A practical proposal for its survival. *Front. Zool.* 8, 25. <http://dx.doi.org/10.1186/1742-9994-8-25>
- WCSP, 2014. World Checklist of Selected Plant Families. Facilitated by the Royal Botanic Gardens, Kew. Published at <http://apps.kew.org/wcsp/> [last accessed 20 Oct. 2014].
- Wieczorek, J., Döring, M., De Giovanni, R., Robertson, T., Vieglais, D., 2009. Darwin Core terms: a quick reference guide. Published at <http://rs.tdwg.org/dwc/terms/index.htm> [last accessed 20 Oct. 2014].
- Zhong, Y., Jung, S., Pramanik, S., Beaman, J.H., 1996. Data model and comparison and query methods for interacting classifications in a taxonomic database. *Taxon* 45, 223–241.
<http://www.jstor.org/stable/1224663>

Fig. 1. Comparison of the estimate of species diversity in four Caryophyllales families based current consensus in literature by family experts and the numbers given in The Plant List. In all cases, there are significant numbers of unassessed names in TPL. From the digitized

specimens made available by the Global Plant Initiative (GBI) all families only a small percentage has so far been verified as to their accurate type status.

Fig. 2. Workflow for the global synthesis of a group of organisms such as the Caryophyllales that aims at complete species-level coverage. The compilation of names with the corresponding original descriptions (protologues) is the starting point. Specimens are a central information source. An alpha-taxonomic approach leads to proposing species concepts based on a series of criteria put forward by individual authors, along with descriptions, keys and maps (process shown in the area with white background). Because all species recognized have been treated using an alpha-taxonomic approach, this process can lead to full coverage relatively quickly. Nevertheless, only the assessment of a representative set of character data (both phenotype and genotype) from a representatively sampled number of individuals (specimens), and the subsequent analysis with an evolutionary methodology will allow the postulation of biological entities as reproducible and testable hypotheses. These entities can be named according to nomenclatural rules, leading to evolutionary-based treatments. At the same time, the process of testing hypotheses allows refining the character and state lists, which directly affects the assessment of phenotypic and genotypic character data. Since clear sets of individuals with their character data can be assigned to a species concept, the respective descriptions and keys can be generated in a structured way. This approximation process is shown in the area with gray background. While ongoing research will continuously improve the knowledge base and thus improve the treatments it will take considerable time to cover all species at this level. Alpha-taxonomic descriptions and keys will therefore be replaced in a step-by-step process as more knowledge becomes available. Circles represent actions, square boxes products. Green portions of the boxes reflect the status quo in Caryophyllales; blue corresponds to the estimated amount of work still needed. Different symbols at the specimen boxes reflect different quality of the specimens. Solid lines indicate direction of the actions in the workflow; dotted lines indicate options for iterative analyses.

Fig. 2. Components of the EDIT Platform for Cybertaxonomy used in the synthesis workflow. Data input and editing can be effected by software clients (*Taxonomic EDITor*) over the internet or in a local area network; for descriptive data the Xper² software is used (*Descriptive Editor*). *Standard File Import* interfaces handle input in the form of community-standard data files. The *Data Store* holds the entire information used (names, character data, etc.) in highly standardized form (Common Data Model, CDM). The *Programming Library* is the part of the software that manages all operations including input, output, and complex processing of requests (such as putting a correct name into synonymy). Programmers use the Library to create application software for the users, such as the editors or the *Web Portals* and *Print Publisher* (for web and document-type output, including maps, descriptions and keys). *Standard File Export* interfaces allow to generate community standard files for import elsewhere (e.g. for the World Flora Online initiative) while *Web Services* allow machine-to-machine access to the Platform functionality in order to incorporate it into software elsewhere (e.g. specialized access portals with “life” access to the Platform, data quality control tools, research workflows, alternative in- and output tools, etc.).

Table Appendix 1. A general approximation of the Caryophyllales diversity compared with the number of accepted and unresolved names (unplaced/unassessed) in The Plant List (TPL). 1= Lipscomb, 2003; 2= Rohwer, 1993a; 2*= Rohwer, 1993a (exclud. *Agdestis*, *Barbeuia*, *Gisekia*, *Lophiocarpus* and *Microthea*); 3= Vivrette et al., 2003; 4= Sánchez del Pino et al., 2009; 5= Kühn, 1993; 6= Nyffeler and Egli, 2010; 7= Porembski, 2003; 8= Kubitzki, 2003a; 9= Eriksson, 2007; 10= Hunt 2006; 11= Rabeler and Hartman, 2005; 12= Porembski and Barthlott, 2003; 13= Kubitzki, 2003b; 14= Kubitzki, 2003c; 15= Kubitzki, 2003d; 16= Gilbert, 1993; 17= Endress and Bittrich, 1993; 17*= Endress and Bittrich, 1993 (exclud. *Corbichonia*, *Hypertelis*, *Limeum* and *Macarthuria*); 18= Kubitzki, 2003e; 19= Spellenberg, 2003; 20= Bittrich & Kühn 1993; 21= Dickison, 2003; 22= Morin, 2005; 23= Sánchez et al., 2011; 24= Prance, 2003; 25= Hillis et al., 2003; 26= Kubitzki, 2003f; 27= Rohwer, 1993b; 28= Gaskin 2003; 29= Nyffeler and Egli, 2010b. The Plant List does not distinguish Amaranthaceae and Chenopodiaceae, so the number refers to the clade containing both.

Family	Num. Genera (Hernández- Ledesma <i>et</i> <i>al.</i> in prep.)	Num. Species	Accepted names in TPL	Unresolved names in TPL
<u>Achatocarpaceae</u>	<u>2</u>	<u>16¹</u>	<u>11</u>	<u>--</u>
<u>Agdestidaceae</u>	<u>1</u>	<u>1²</u>	<u>1</u>	<u>--</u>
<u>Aizoaceae</u>	<u>125</u>	<u>2500³</u>	<u>2271</u>	<u>2083</u>
<u>Amaranthaceae/Chenopodiaceae</u>	<u>82+104 = 186</u>	<u>840⁴ + 1400⁵ =2240</u>	<u>2052</u>	<u>2385</u>
<u>Anacamptserotaceae</u>	<u>3</u>	<u>39⁶</u>	<u>58</u>	<u>33</u>
<u>Ancistrocladaceae</u>	<u>1</u>	<u>20⁷</u>	<u>21</u>	<u>1</u>
<u>Asteropeiaceae</u>	<u>1</u>	<u>8⁸</u>	<u>8</u>	<u>--</u>
<u>Barbeuiaceae</u>	<u>1</u>	<u>1²</u>	<u>1</u>	<u>--</u>
<u>Basellaceae</u>	<u>4</u>	<u>19⁹</u>	<u>19</u>	<u>3</u>
<u>Cactaceae</u>	<u>132</u>	<u>1450-1870^{10, 29}</u>	<u>2233</u>	<u>5505</u>
<u>Caryophyllaceae</u>	<u>94</u>	<u>3000¹¹</u>	<u>2456</u>	<u>4255</u>
<u>Didiereaceae</u>	<u>7</u>	<u>19-20⁶</u>	<u>22</u>	<u>2</u>
<u>Dioncophyllaceae</u>	<u>3</u>	<u>3¹²</u>	<u>3</u>	<u>2</u>
<u>Droseraceae</u>	<u>3</u>	<u>100+¹³</u>	<u>189</u>	<u>10</u>
<u>Drosophyllaceae</u>	<u>1</u>	<u>1¹⁴</u>	<u>1</u>	<u>3</u>
<u>Frankeniaceae</u>	<u>1</u>	<u>70¹⁵</u>	<u>73</u>	<u>16</u>
<u>Gisekiaceae</u>	<u>1</u>	<u>7¹⁶</u>	<u>7</u>	<u>2</u>
<u>Halophytaceae</u>	<u>1</u>	<u>1⁶</u>	<u>1</u>	<u>--</u>
<u>Limeaceae</u>	<u>1</u>	<u>20¹⁷</u>	<u>20</u>	<u>25</u>
<u>Lophiocarpaceae</u>	<u>2</u>	<u>6^{2, 17}</u>	<u>5</u>	<u>4</u>
<u>Microtaceae</u>	<u>1</u>	<u>9²</u>	<u>9</u>	<u>6</u>
<u>Molluginaceae</u>	<u>9</u>	<u>90^{17*}</u>	<u>85</u>	<u>103</u>
<u>Montiaceae</u>	<u>15</u>	<u>226⁶</u>	<u>245</u>	<u>164</u>
<u>Nepenthaceae</u>	<u>1</u>	<u>85¹⁸</u>	<u>7</u>	<u>247</u>
<u>Nyctaginaceae</u>	<u>31</u>	<u>300¹⁹-400²⁰</u>	<u>450</u>	<u>430</u>
<u>Physenaceae</u>	<u>1</u>	<u>2²¹</u>	<u>2</u>	<u>--</u>
<u>Phytolaccaceae</u>	<u>13</u>	<u>50- 52^{2*}</u>	<u>52</u>	<u>81</u>
<u>Plumbaginaceae</u>	<u>29</u>	<u>775²²</u>	<u>635</u>	<u>1075</u>
<u>Polygonaceae</u>	<u>47</u>	<u>1200²³</u>	<u>1384</u>	<u>2120</u>
<u>Portulacaceae</u>	<u>1</u>	<u>116⁶</u>	<u>109</u>	<u>147</u>
<u>Rhabdodendraceae</u>	<u>1</u>	<u>3²⁴</u>	<u>3</u>	<u>--</u>
<u>Sarcobataceae</u>	<u>1</u>	<u>2²⁵</u>	<u>2</u>	<u>3</u>
<u>Simmondsiaceae</u>	<u>1</u>	<u>1²⁶</u>	<u>2</u>	<u>3</u>
<u>StegnospERMataceae</u>	<u>1</u>	<u>3²⁷</u>	<u>5</u>	<u>--</u>
<u>Talinaceae</u>	<u>3</u>	<u>28⁶</u>	<u>29</u>	<u>34</u>
<u>Tamaricaceae</u>	<u>4</u>	<u>78²⁸</u>	<u>79</u>	<u>134</u>
<u>Incertae sedis</u>	<u>3</u>	<u>12</u>		<u>--</u>
<u> <i>Hypertelis</i></u>		<u>9¹⁷</u>	<u>7</u>	<u>4</u>
<u> <i>Macarthuria</i></u>		<u>3¹⁷</u>	<u>0</u>	<u>10</u>
<u> <i>Jorena?</i></u>		<u>?</u>		<u>--</u>

Total	<u>732</u>	<u>12489–12592</u>	<u>12557</u>	<u>18890</u>
Bittrich, V., Kühn, U. 1993. Nyctaginaceae. In Kubitzki K., Rohwer J. G. & Bittrich V. (Eds.), The families and genera of vascular plants 2. Springer, Berlin, pp. 473–485.				
Dickison, W.C. 2003. Physenaceae in: Kubitzki, K., Bayer, C. (Eds.), The families and genera of vascular plants 5. Springer, Berlin, pp. 332–333.				
Endress, M.E., Bittrich V. 1993. Molluginaceae in: Kubitzki, K., Rohwer, J.G. and Bittrich, V. (Eds.), The families and Genera of vascular plants 2. Springer, Berlin, pp. 419–425.				
Ericksson, R. 2007. A synopsis of Basellaceae. Kew Bull. 62, 297–320.				
Gaskin, J.F. 2003. Tamaricaceae in: Kubitzki, K., Bayer, C. (Eds.), The families and genera of vascular plants 5. Springer, Berlin, pp. 363–368.				
Gilbert, M.G. 1993. A review of <i>Gisekia</i> (Gisekiaceae). Kew Bull. 48, 343–356.				
Hillis, M.H., Thieret J.W., Morefield J.D. 2003. <i>Sarcobatus</i> in: Flora of North America Editorial Committee (Eds.), Flora of North America 4. Oxford University Press, New York and Oxford, pp. 387–389.				
Hunt, D. 2006. The new cactus lexicon. Dh Books, Milborne Port.				
Kubitzki, K. 2003a. Asteropeiaceae in: Kubitzki, K., Bayer, C. (Eds.), The families and genera of vascular plants 5. Springer, Berlin, pp. 28–29.				
Kubitzki, K. 2003b. Droseraceae in: Kubitzki, K., Bayer, C. (Eds.), The families and genera of vascular plants 5. Springer, Berlin, pp. 198–202.				
Kubitzki, K. 2003c. Drosophyllaceae in: Kubitzki, K., Bayer, C. (Eds.), The families and genera of vascular plants 5. Springer, Berlin, pp. 203–205.				
Kubitzki, K. 2003d. Frankeniaceae in: Kubitzki, K., Bayer, C. (Eds.), The families and genera of vascular plants 5. Springer, Berlin, pp. 206–212.				
Kubitzki, K. 2003e. Nepenthaceae in: Kubitzki, K., Bayer, C. (Eds.), The families and genera of vascular plants 5. Springer, Berlin, pp. 320–324.				
Kubitzki, K. 2003f. Simmondsiaceae in: Kubitzki, K., Bayer, C. (Eds.), The families and genera of vascular plants 5. Springer, Berlin, pp. 355–358.				
Kühn, U. 1993. Chenopodiaceae in: Kubitzki, K., Rohwer, J.G. and Bittrich, V. (Eds.), The families and genera of vascular plants 2. Springer, Berlin, pp. 253–280.				
Lipscomb B.L. 2003. Achatocarpaceae in: Flora of North America Editorial Committee (Eds.), Flora of North America 4. Oxford University Press, New York and Oxford, pp. 12–13.				
Mison, N.R. 2005. Plumbaginaceae in: Flora of North America Editorial Committee (Eds.), Flora of North America 5. Oxford University Press, New York and Oxford, pp. 602–611.				
Nyffeler, R., Eggli, U. 2010a. Disintegrating Portulacaceae: A new familial classification of the suborder Portulacinae (Caryophyllales) based on molecular and morphological data. <i>Taxon</i> 59, 227–240.				
Nyffeler, R. & Eggli U. 2010b: A farewell to dated ideas and concepts - molecular phylogenetics and a revised suprageneric classification of the family Cactaceae. <i>Ein Abgesang auf veraltete Ideen und Konzepte</i> \ (mi Phylogenetik und eine revidierte supragenerische Klassifikation der Familie Cactaceae. – <i>Schumannia</i> 6 = Biodiv. Ecol. 3: 109–149.				
Porembski, S. 2003. Ancistrocladaceae in: Kubitzki, K., Bayer, C. (Eds.), The families and genera of vascular plants 5. Springer, Berlin, pp. 25–27.				
Porembski, S., Barthlott, W. 2003. Dioncophyllaceae in Kubitzki, K., Bayer, C. (Eds.), The families and genera of vascular plants 5. Springer, Berlin, pp. 178–181.				
Prance, G.T. 2003. Rhabdodendraceae in: Kubitzki, K., Bayer, C. (Eds.), The families and genera of vascular plants 5. Springer, Berlin, pp. 339–341.				
Rabeller, R.K., Hartman, R.L. 2005. Caryophyllaceae in Flora of North America Editorial Committee (Eds.), Flora of North America 5. Oxford University Press, New York and Oxford, pp. 3–8.				

- Rohwer, J. G. 1993a. Phytolaccaceae in: Kubitzki, K., Rohwer, J.G. and Bittrich, V. (Eds.), *The families and genera of vascular plants 2*. Springer, Berlin, pp. 506–515.
- Rohwer, J. G. 1993b. Stegnospermaceae in: Kubitzki, K., Rohwer, J.G. and Bittrich, V. (Eds.), *The families and genera of vascular plants 2*. Springer, Berlin, pp. 592–594.
- Sanchez, A., Schuster, T.M., Burke, J.M., Kron K. 2011. Taxonomy of Polygonoideae (Polygonaceae): a new tribal classification. *Taxon* 60, 151–160.
- Sánchez-del Pino, I., Borsch, T. & Motley, T.J. 2009. trnL-F and rpl16 sequence data and dense taxon sampling reveal monophyly of unilocular anthered Gomphrenoideae (Amaranthaceae) and an improved picture of their internal relationships. *Syst. Bot.* 34, 57–67.
- Spellenberg, R.W. 2003. Nyctaginaceae in *Flora of North America* Editorial Committee (Eds.), *Flora of North America 4*. Oxford University Press, New York and Oxford, pp.14–74.
- The Plant List. 2013. Version 1.1. Published on the Internet; <http://www.theplantlist.org/> (accessed 14–24 April–2014).
- Vivrette, N.J., Bleck, J.E. Ferren Jr., W.R. 2003. Aizoaceae in *Flora of North America* Editorial Committee (Eds.), *Flora of North America 4*. Oxford University Press, New York and Oxford, pp.75–91.
- ¹ African Plants Database (version 3.4.0). Conservatoire et Jardin botaniques de la Ville de Genève and South African National Biodiversity Institute, Pretoria, <http://www.ville-ge.ch/musinfo/bd/cjb/africa/> (last accessed 5 Jan. 2015).
- ¹ SANBI, 2005-2014. Plants of southern Africa: an online checklist. POSA ver. 3.0 June 2009. <http://posa.sanbi.org/> (last accessed 5 Jan. 2015).
- ¹ The Digitised Flora of Central Africa (Democratic Republic of the Congo, Rwanda & Burundi). Botanic Garden Meise, published at <http://www.br.fgov.be/RESEARCH/DATABASES/FOCA/> (last accessed 5 Jan. 2015).
- ¹ Marloth, R., 1913. The flora of South Africa with synoptical tables of the genera of the higher plants, Vol. 1. Capetown, Darter Bros. & Co. <http://dx.doi.org/10.5962/bhl.title.65674>
- ¹ Euro+Med, 2006-. Euro+Med PlantBase - the information resource for Euro-Mediterranean plant diversity. Published online at <http://ww2.bgbm.org/EuroPlusMed/> (last accessed 5 Jan. 2015).
- ¹ Greuter, W., Burdet, H.M., Long, G. (Eds.), 1984, 1986, 1989. Med-Checklist - A critical inventory of vascular plants of the circum-mediterranean countries. Vols 1, 2 & 4. Geneva, Conservatoire et Jardin botanique de la Ville Genève/ Berlin, Botanischer Garten & Botanisches Museum Berlin-Dahlem. Published at <http://ww2.bgbm.org/mcl/> (last accessed 5 Jan. 2015).
- ¹ Flora Zambesiaca. Published at <http://apps.kew.org/efloras/search.do> (last accessed 5 Jan. 2015).
- ¹ Madagascar Catalogue, 2015. Catalogue of the Vascular Plants of Madagascar. Missouri Botanical Garden, St. Louis, U.S.A. & Antananarivo, Madagascar. Published at <http://www.tropicos.org/project/mada> (last accessed 7 Jan. 2015).
- ¹ Hooker, W.J., 1849. Niger flora; or. An enumeration of the plants of western tropical Africa. London, H. Bailliere. <http://dx.doi.org/10.5962/bhl.title.594>
- ¹ Swaziland National Trust Commission, 2014. Swaziland' Flora Database. Published at <http://www.sntc.org.sz/flora/> (last accessed 5 Jan. 2015).
- ¹ van Steenis, C.G.G.J. (ed.), 1948-. Flora Malesiana. Series I, Spermatophyta. Djakarta, Noordhoff-Kolff. <http://dx.doi.org/10.5962/bhl.title.40744>
- ¹ Flora Malesiana DataPortal. Published at <http://dev.e-taxonomy.eu/dataportal/flora-malesiana/> (last accessed 5 Jan. 2015).
- ¹ Elven, R. (Ed.), 2015. Annotated Checklist of the Panarctic Flora (PAF), Vascular plants. Published at <http://nhm2.uio.no/paf/> (last accessed 5 Jan. 2015).
- ¹ Podlech, D., 2012. Checklist of flowering plants of Afghanistan. Published at http://www.sysbot.biologie.uni-muenchen.de/de/personen/podlech/flowering_plants_afghanistan.pdf (last accessed 5 Jan. 2015).
- ¹ Zhengyi, W., Raven, P.H., Deyuan, H., 1994-. Flora of China. Published at http://efloras.org/flora_page.aspx?flora_id=2 (last accessed 5 Jan. 2015).
- ¹ Flora of Israel Online. Published at <http://flora.huji.ac.il/browse.asp?lang=en> (last accessed 1 june 2014).
- ¹ Flora of Japan. Published at <http://foj.c.u-tokyo.ac.jp/gbif/> (last accessed 5 Jan. 2015).
- ¹ Ohwi, J., 1965. Flora of Japan. Washington, Smithsonian Institution. <http://dx.doi.org/10.5962/bhl.title.43786>
- ¹ University of Greifswald, Institute of Botany and Landscape Ecology, Institute of Geography and Geology, Computer Centre, 2010- (continuously updated). FloraGREIF - Virtual Flora of Mongolia. Computer Centre of

- University of Greifswald, D-17487 Greifswald, Germany. Published at <http://greif.uni-greifswald.de/floragreif/> (last accessed 5 Jan. 2015).
- ¹ Kress, W.J., DeFilipps, R.A., Farr, E., Kyi, D.Y.Y., 2003. A Checklist of the Trees, Shrubs, Herbs, and Climbers of Myanmar. Contributions from the United States National Herbarium 45: 1-590. Published online at <http://botany.si.edu/myanmar/checklistNames.cfm> (last accessed 5 Jan. 2015).
- ¹ Updated online version of Press, J.R., Shrestha, K.K. & Sutton, D.A., 2000. Annotated Checklist of the Flowering Plants of Nepal. London, The Natural History Museum. Published at http://www.efloras.org/flora_page.aspx?flora_id=110 (last accessed 5 Jan. 2015).
- ¹ Missouri Botanical Garden, 1995-. Pakistan Plant Database (PPD). Published at <http://www.tropicos.org/Project/Pakistan> (last accessed 5 Jan. 2015).
- ¹ Zohary, M., N. Feinbrun-Dotan, N., 2012. Flora of Palaestina, 2nd ed. Revised and edited by the Editorial Board (Danin, A., Heller, D., Fragman-Sapir, O., Kislev, M., Vered-Leschner, H., Plitmann, U.). Published at <http://www.tropicos.org/Project/Palaestina> (last accessed 5 Jan. 2015).
- ¹ Post, G.E., 1896. Flora of Syria, Palestine and Sinai; from the Taurus to Ras Muhammad, and from the Mediterranean Sea to the Syrian desert. Beirut, Syria, Syrian Protestant College. <http://dx.doi.org/10.5962/bhl.title.20319>
- ¹ Pelsner, P.B., Barcelona, J.F., Nickrent, D.L. (eds.), 2011-. Co's Digital Flora of the Philippines. Published at <http://www.philippineplants.org/> (last accessed 7 Jan. 2015).
- ¹ Norton, J.A., Abdul Majid, S., Allan, D.R., Al Safran, M., Böer, B., Richer, R., 2009. An Illustrated Checklist of the Flora of Qatar. Gosport, UK, Browndown Publications. Published at http://www.unesco.org/new/uploads/media/Plant_Species_Checklist_of_Qatar_03.pdf (last accessed 7 Jan. 2015).
- ¹ Published at <http://www.biodiversitylibrary.org/bibliography/43751#/summary>
- ¹ Instituto de Botánica Darwinion – IBODA, 2015. Catálogo de las Plantas Vasculares del Conosur. Published at <http://www.darwin.edu.ar/Proyectos/FloraArgentina/fa.htm> (last accessed 5 Jan. 2015).
- ¹ Milliken, W., Klitgård, B., Baracat, A. (eds.), 2009-. Neotropikey - Interactive key and information resources for flowering plants of the Neotropics. Published at www.kew.org/neotropikey (last accessed 5 Jan. 2015).
- ¹ Flora Argentina. Plantas vasculares de la República Argentina. Published at <http://www.floraargentina.edu.ar/> (last accessed 5 Jan. 2015).
- ¹ Maycock, J.D., 1830. Flora barbadensis: a catalogue of plants, indigenous, naturalized, and cultivated, in Barbados : to which is prefixed, a geological description of the island. London, J. Ridgeway. <http://dx.doi.org/10.5962/bhl.title.40099>
- ¹ Catálogo de las Plantas Vasculares de Bolivia. Published at <http://www.tropicos.org/Project/BC> (last accessed 7 Jan. 2015).
- ¹ Lista de Espécies da Flora do Brasil. Jardim Botânico do Rio de Janeiro. Published at <http://floradobrasil.jbrj.gov.br/> (last accessed 7 Jan. 2015).
- ¹ Martius, C.F.P. von, Eichler, A.W, Urban, I. (eds.), 1840-1906. Flora brasiliensis. Munich & Leipzig: R. Oldenbourg. Published at <http://florabrasiliensis.cria.org.br/> (last accessed 7 Jan. 2015).
- ¹ Reiche, C., 1896-1911. Flora de Chile. Published at http://efloras.org/flora_page.aspx?flora_id=60 (last accessed 7 Jan. 2015).
- ¹ Instituto Nacional de Biodiversidad, Missouri Botanical Garden, 1995. Manual de la Flora de Costa Rica. Published at <http://www.mobot.org/manual.plantas/lista.html> (last accessed 7 Jan. 2015).
- ¹ Standley, P.C. 1937. Flora of Costa Rica. Field Museum of Natural History, Botanical series 18, pt. 1-4. <http://dx.doi.org/10.5962/bhl.title.2251>
- ¹ Burger, W.C. (ed.), 1983. Flora Costaricensis. Fieldiana, Botany, new ser. 13. <http://dx.doi.org/10.5962/bhl.title.2593>
- ¹ Jørgensen, P.M., León-Yáñez, S. (eds.), 1999. Catalogue of the vascular plants of Ecuador. Monogr. Syst. Bot. Missouri Bot. Gard. 75, i-viii, 1-1182. Published online at <http://www.tropicos.org/Project/CE> (last accessed 7 Jan. 2015).
- ¹ Standley, P.C., Steyermark, J.A., 1946. Flora of Guatemala. Fieldiana, Botany 24, pt. 4. <http://dx.doi.org/10.5962/bhl.title.2233>
- ¹ Boggan, J., Funk, V., Kelloff, C., Hoff, M., Cremers, G., Feuillet, C., 1997. Checklist of the plants of the Guianas (Guyana, Surinam, French Guiana), 2nd ed. Published at <http://botany.si.edu/bdg/pdf/guulist2nd.pdf> (last accessed 7 Jan. 2015).
- ¹ Davidse, G., Sousa Sánchez, M., Knapp, S., Chiang Cabrera, F. (eds.). Flora Mesoamericana. Universidad Nacional Autónoma de México, México. Published at <http://www.tropicos.org/Project/FM> (last accessed 7 Jan. 2015).

- ¹ Macfadyen, J., 1837. The flora of Jamaica; a description of the plants of that island, arranged according to the natural orders. London, Longman, Orme, Brown, Green, & Longman. <http://dx.doi.org/10.5962/bhl.title.4970>
- ¹ Fawcett, W., Rendle, A.B. 1914, 1926. Flora of Jamaica, containing descriptions of the flowering plants known from the island, Vols. 3, 5. London, Printed by order of the Trustees of the British Museum. <http://dx.doi.org/10.5962/bhl.title.2760>
- ¹ Stevens, W.D. et al. (eds.), 2001. Flora de Nicaragua. St. Louis, Missouri Botanical Garden Press. Published online at <http://www.tropicos.org/projectwebportal.aspx?pagename=Home&projectid=7> (last accessed 7 Jan. 2015).
- ¹ D'Arcy, W.G., 1987. Flora of Panama: Checklist and Index. Part I: The Introduction and Checklist. Monographs in Systematic Botany 17. Published online at <http://www.tropicos.org/Project/PAC> (last accessed 7 Jan. 2015).
- ¹ Inventario biológico de Paraguay. Published at <http://www.tropicos.org/Project/Paraguay> (last accessed 5 Jan. 2015).
- ¹ Brako, L., Zarucchi, J.L., 1993. Catalogue of the Flowering Plants and Gymnosperms of Peru. Monographs in Systematic Botany 45. Published online at <http://www.tropicos.org/Project/PEC> (last accessed 7 Jan. 2015).
- ¹ Flora of Peru. Fieldiana series, Chicago. Published online at <http://www.biodiversitylibrary.org/browse/collection/floraofperu> (last accessed 7 Jan. 2015).
- ¹ Flora of Puerto Rico and the Virgin Islands. Published at <http://botany.si.edu/antilles/prflora/index.html> (last accessed 5 Jan. 2015).
- ¹ Pulle, A., 1906. An enumeration of the vascular plants known from Surinam: together with their distribution and synonymy. Leiden, E. J. Brill. <http://dx.doi.org/10.5962/bhl.title.44906>
- ¹ Flora of North America Editorial Committee (eds.), 1993–. Flora of North America North of Mexico. 18+ vols. New York and Oxford. Published online at http://www.efloras.org/flora_page.aspx?flora_id=1 (last accessed 7 Jan. 2015).
- ¹ Acevedo-Rodríguez, P., Strong, M.T. 2007. Catalogue of the seed plants of the West Indies Website. Published at <http://botany.si.edu/antilles/WestIndies/> (last accessed 7 Jan. 2015).
- ¹ Australian Plant Name Index, IBIS database, Centre for Australian National Biodiversity Research, Australian Government, Canberra. Published at <http://www.cpbr.gov.au/cgi-bin/apni> (last accessed 5 Jan. 2015).
- ¹ Australian Plant Census, IBIS database, Centre for Australian National Biodiversity Research, Council of Heads of Australasian Herbaria, <http://www.chah.gov.au/apc/index.html> (last accessed 5 Jan. 2015).
- ¹ Flora of Australia Online. Published at <http://www.environment.gov.au/biodiversity/abrs/online-resources/flora/main/> (last accessed 7 Jan. 2015).
- ¹ Bentham, G. 1863–1878. Flora australiensis: a description of the plants of the Australian territory. London, L. Reeve and Co. <http://dx.doi.org/10.5962/bhl.title.16515>
- ¹ Smith, A.C., 1979–1991. Flora Vitiensis nova: a new Flora of Fiji. Lawaii, Hawaii; Pacific Tropical Botanical Garden. <http://dx.doi.org/10.5962/bhl.title.44033>
- ¹ Wagner, W.L., Herbst, D.R., Lorence, D.H., 2005–. Flora of the Hawaiian Islands website. Published at <http://botany.si.edu/pacificislandbiodiversity/hawaiianflora/query2.cfm> (last accessed 5 Jan. 2015).
- ¹ Wagner, W.L., Lorence, D.H., 2002–. Flora of the Marquesas Islands website. <http://botany.si.edu/pacificislandbiodiversity/marquesasflora/index.htm> (last accessed 5 Jan. 2015).
- ¹ Wagner, W.L., Herbst, D.R., Tornabene, M.W., Weitzman, A., Lorence, D.H., 2012–. Flora of Micronesia website. <http://botany.si.edu/pacificislandbiodiversity/micronesia/index.htm> (last accessed 5 Jan. 2015).
- ¹ Breitwieser, I., Brownsey, P.J.; Heenan, P.B., Nelson, W.A., Wilton, A.D. (eds.), 2010–. Flora of New Zealand Online. Published at <http://www.nzflora.info/> (last accessed 7 Jan. 2015).
- ¹ Duretto, M.F. (Ed.), 2009–. Flora of Tasmania Online. Hobart, Tasmanian Herbarium, Tasmanian Museum & Art Gallery. Published at www.tmag.tas.gov.au/floratasmania (last accessed 7 Jan. 2015).
- ¹ Bennett, A.W., 1896. The flora of the Alps: being a description of all the species of flowering plants indigenous to Switzerland; and of the Alpine species of the adjacent mountain districts of France, Italy, & Austria including the Pyrenees, Vol. 1. London, J.C. Nimmo. <http://dx.doi.org/10.5962/bhl.title.33124>
- ¹ Castroviejo, S. (coord. gen.), 1986–2012. Flora iberica 1-8, 10-15, 17-18, 21. Real Jardín Botánico, CSIC, Madrid. Published at <http://www.floraiberica.es> (last accessed 5 Jan. 2015).
- ¹ Nikolić, T. (Ed.), 2014. Flora Croatica Database. Faculty of Science, University of Zagreb. Published at <http://hirc.botanic.hr/fcd> (last accessed 5 Jan. 2015).
- ¹ Hand, R., Hadjikyriakou, G.N., Christodoulou, C.S. (Eds.), 2011– (continuously updated). Flora of Cyprus – a dynamic checklist. Published at <http://www.flora-of-cyprus.eu/> (last accessed 5 Jan. 2015).
- ¹ Deutschlandflora, 2014. Published at <http://www.deutschlandflora.de/map.phtml> (last accessed 5 Jan. 2015).
- ¹ Buttler, K.P., Thieme, M. & Mitarbeiter, 2014. Florenliste von Deutschland – Gefäßpflanzen, Version 6. Frankfurt am Main, August 2014. Published online at <http://www.kp-buttler.de> (last accessed 7 Jan. 2015).

¹ [Sociedade Portuguesa de Botânica, 2012-2014. Portal Flora-On, published at http://www.flora-on.pt/ \(last accessed 5 Jan. 2015\).](http://www.flora-on.pt/)

Accepted Manuscript