The need for and the requirements of EuroSL, an electronic taxonomic reference list of all European plants


Abstract: Biodiversity informatics has experienced tremendous developments in the last 15 years. There are now comprehensive online checklists for plant taxa as well as many large plant-taxon related databases, including the vegetation-plot databases registered in the Global Index of Vegetation-Plot Databases (GIVD; http://www.givd.info). However, efficient maintenance, analysis, and integration of these databases are still much impeded by the failure of presently available electronic taxonomic reference lists of plants to fully meet the requirements of such applications. Here we outline the principal specifications of an electronic taxonomic reference list for Europe ("EuroSL" = European standard list of plant taxa) and identify features not met in current practice. EuroSL should cover all macroscopic taxa of vascular plants, bryophytes, lichens, and algae that occur in European vegetation in a uniform database, irrespective of their floristic status (e.g. native, archaeophyte, neophyte, casual). The adoption of informal aggregates is essential to cover deviating species concepts and to capture legacy data. EuroSL should not only assign names but also match taxonomic concepts. This task cannot be fully automated, as the same correctly applied taxon name can have different meanings depending on the taxonomic concept applied. In order to be a useful tool, EuroSL would need to be better documented than most existing electronic checklists and be released in fixed versions. Every subsequent version should contain an unambiguous connection linking each taxon to the corresponding unit in the previous version. We identify possible components of EuroSL, of which Euro+Med PlantBase, the recent European checklists of bryophytes, and the taxonomic crosswalks between various national Turboveg checklists collected for SynBioSys Europe, are the major ones. Concepts developed for GermanSL might be adopted for EuroSL, but implemented in a software framework that is yet to be developed from existing tools. Such a framework would allow documented editing of the content by specialists distributed across Europe. To become successful, EuroSL would require intensive collaboration between taxonomists, ecologists and biodiversity informaticians, as well as appropriate funding. Establishing EuroSL would dramatically enhance the usability and reliability of plant-taxon related databases in Europe for the purposes of pure and applied research and conservation legislation. Its development should therefore be of highest priority.

Keywords: alga; biodiversity informatics; bryophyte; checklist; ecoinformatics; GBIF; GIVD; lichen; vascular plant; vegetation-plot database.

Abbreviations: GBIF = Global Biodiversity Information Facility; EuroSL = European standard list of plant taxa; GermanSL = German standard list of plant taxa; GIVD = Global Index of Vegetation-Plot Databases; ICN = International Code of Nomenclature for algae, fungi, and plants (see McNeill & Turland 2011).

Received: 3 May 2012 – Accepted: 14 May 2012 – Co-ordinating Editor: Gabriela Lopez-Gonzalez.

Introduction

Biodiversity informatics (Bisby 2000, Canhos et al. 2004, Guralnick & Hill 2009) and ecoinformatics (Bekker et al. 2007, Dengler et al. 2011a) are two young and closely related disciplines that have developed tremendously during the last 15 years. For example, the Global Biodiversity Information Facility (GBIF; http://www.gbif.org; see Wheeler 2004) is currently the most important data source of biodiversity information, containing 324 million indexed records of species occurrences (as of 16 March 2012). For plants and vegetation, a large and steadily growing amount of vegetation-plot data is becoming available in electronic format (Ewald 2001, Schaminée et al. 2009, Dengler et al. 2011b). These databases go beyond the single-species occurrence data in GBIF by providing fine-resolution spatial and temporal co-occurrence information on all plant species found in a particular species assemblage (plant community), often combined with plot-level structural and environmental data. In the summer of 2010, the Global Index of Vegetation-Plot Databases (GIVD; http://www.givd.info; Dengler et al. 2011b) was launched as a global metadata base (registry) of such databases. Currently, GIVD indexes 182 databases that...
together contain more than 2.8 million independent vegetation plots, corresponding to an estimated 50 million occurrence records of individual plant species (Jansen et al. 2012).

In addition to species occurrence databases and vegetation-plot databases, there are many other plant-taxon related databases, such as taxonomic and phylogenetic databases (e.g. Stevens 2008), plant trait databases (e.g. LEDA: http://www.leda-traitbase.org, Kleyer et al. 2008 or TRY: http://www.try-db.org/, Katte et al. 2011), or red list and other conservation-related databases (e.g. IUCN 2011). The vegetation-plot databases, combined with some of the other taxon-related databases, or connected to available environmental data (e.g. WorldClim: Hijmans et al. 2005) offers unprecedented opportunities in both pure and applied ecological research. Examples of promising applications of large vegetation-plot databases are (i) developing consistent supra-national vegetation classifications, (ii) testing ecological and evolutionary theories, and (iii) analysing and forecasting the effects of global change (see examples reviewed by Schaminée et al. 2009, Dengler et al. 2011b).

However, despite the existence of the databases and many ideas on how to analyse them, their integrated use at the European level is seriously impeded by the lack of appropriate tools for handling plant taxonomy. This situation prompted a team of vegetation ecologists, plant taxonomists, and biodiversity informaticians to launch the idea of EuroSL (SL = standard list), an electronic reference list of all plant taxa in Europe. On 23 December 2011, a group of 10 researchers met for the first EuroSL Workshop in Göttingen. This meeting brought together specialists from the European Vegetation Survey (EVS, a Working Group of the IAVS: Rodwell et al. 1995), the Working Group on Vegetation Databases (within the German Network for Phytodiversity, NwPhyD; see http://www.hswt.de/fh/fakultaet/w/ldozenten/ewald/vegetationsdaten-banken/vegetation-databases.html), and the Biodiversity Informatics Group of the Botanical Garden and Botanical Museum in Berlin-Dahlem (BGBM). Based on the results of this workshop, our paper aims to describe in more detail the remaining requirements and specifications for a EuroSL, and propose potential ways forward. While this article is focused on Europe and thus uses examples from this region, its analysis of problems and potential solutions are applicable to other parts of the world.

What are the problems?

Plant nomenclature is not consistently applied across space and time. This is a particularly big problem in Europe, whose many small- to medium-sized countries typically have one or more floras with partly incompatible taxonomic concepts. The problem has even been recognised in the political arena, with the European Union administration launching the Pan-European Species directory Infrastructure (PESI; http://www.eu-nomen.eu/nesi/; de Jong et al. 2010) for the purposes of standardising taxonomic names relevant to legal documents and monitoring programmes. There are now taxonomic databases available that are able to assign most synonyms correctly (e.g. The Plant List – TPL: http://www.theplantlist.org/, see Kalwij in press; Euro+Med PlantBase: http://www.emplantbase.org/; see Euro+Med 2006). However, these databases still leave the following major issues unresolved:

- **Concept taxonomy**: While focusing on the correct application of plant names, the databases neglect the fact that different delimitations of the same correctly applied species name, or, more generally, taxon name (taxonyn), are in use. This adds another level of complexity (Bendersohn 1995, Jansen & Dengler 2010), which has hardly ever been addressed even in taxonomic checklists (but see Koperski et al. 2000) or online sources. However, it is highly relevant particularly for legacy data such as those in vegetation-plot databases. Therefore, relying on names alone to link different information resources will often produce unexpected results (Berendsohn & Geoffroy 2007).

- **Comprehensiveness**: There is no single resource that combines all species found in European vegetation. This is because existing databases are mostly restricted to one of the major plant groups (vascular plants, bryophytes, lianas, algae). Similarly, non-naturalised taxa (casual aliens), hybrids and informal taxa (e.g. species groups, aggregates) are not usually included in taxonomic checklists because they do not conform to certain formal criteria. However, the non-naturalised taxa and hybrids are an integral part of the flora and vegetation of the territory (they occur in vegetation plots), while informal taxa are necessarily recorded in vegetation plots because not all plants can be determined to the species level at every stage of their development.

**Versioning**: While the Internet and online databases brought incredible benefits to biodiversity informatics, lack of clear versioning in many of the electronic information systems causes serious drawbacks compared to traditional paper publications. One such drawback is that no publication with defined content exists that can be cited as an unambiguous reference. Consequently, comparing the content and definitions at a later point is difficult or impossible.

**Requirements and specifications of EuroSL**

It is thus evident that there is an urgent need for a uniform, up-to-date, versioned, well-documented and easily applicable electronic reference list of all taxa of European vegetation. Specific requirements for such a list will be outlined in more detail in this section. Most of the aspects have already been addressed during the development of GermanSL (SL means standard list) by Jansen & Dengler (2008), which serves as a template for “EuroSL”, a comparable electronic reference list at a continental scale. However, amplification of the GermanSL model to the European level is much more complicated for the reasons listed below.

(a) There are many more species and different taxonomic views.

(b) In contrast to Germany and its well-studied flora, in some other parts of Europe, particularly the Mediterranean region, many new taxa are still being described.

(c) While GermanSL could be based on recent national taxonomic checklists of vascular plants (Wisskirchen & Haeupler 1998), bryophytes (Koperski et al. 2000), and lichens (Scholz 2000) recent checklists at the European scale exist only for the bryological divisions (Grolle & Long 2000, Hill et al. 2006) and (at different quality levels for different families) for the angiosperms in Euro+Med Plantbase.

(d) GermanSL does not provide a prototype for a sustained solution, being launched as a one-off product, implemented mainly in TURBOVEG (a major database program for vegetation-plot data; Hennekens & Schaminée 2001), but for budgetary and other reasons lacking the originally-intended dynamic updates.
Thus, while EuroSL could build, *inter alia*, upon the ideas and concepts used in GermanSL, it would need to go much further still. Details of EuroSL requirements and specifications are outlined in the following subsections.

**Completeness**

In order for EuroSL to be considered complete, it should comprise, in one uniform database, all species, certain infraspecific taxa, and the above-mentioned informal groups found in the vegetation of the continent, according to the following criteria:

- **All groups of macroscopically visible photoautotrophic taxa of terrestrial, freshwater and marine habitats** should be treated because they are considered part of the vegetation (Dengler 2003). This includes the following major groups: vascular plants (*Spermatophyta, Pteridophyta*), bryophytes (*Bryophyta, Marchantiophyta, Anthocerotophyta, lichens* [lichenized groups of *Ascomycota* and *Basidiomycota*]), and the polyphyletic group of macroscopic algae (*e.g.* *Chlorophyta, Chrysophyta* p.p., *Heterokontophyta* p.p., *Rhodophyta*). Unicellular algae and procaryotic “bluegreen algae” (*Cyanobacteria*) should be included if they form macroscopically visible aggregations.

- **There should be complete geographic coverage of Europe** as delimited in *Flora Europaea* (Tutin et al. 1968–1993: “Europe s.s.”). However, where data from adjacent geographic regions are easily available, because their flora are encompassed in existing European databases (“Europe s.l.”), these data should be kept in EuroSL even when this additional geographic range information is not included for all main groups of taxa. Examples of such extensions are the Canary Islands (Grolle & Long 2000, Euro+Med 2006, Hill et al. 2006) and all circum-Mediterranean and partially Caucasian countries (Euro+Med 2006). While the removal of taxa occurring in Europe s.l. but absent in Europe s.s. would be time consuming, keeping them (as separate concepts) could be useful in two ways: (a) native species in adjacent areas might sooner or later also occur as non-native taxa in Europe s.s.; and (b) it could be the basis for the possible future extension of EuroSL to cover larger areas.

- **As well as native and naturalised non-native taxa** (archaeophytes and neophytes), EuroSL should include non-naturalised neophytes (casual aliens) and cultivated plants that are regularly planted outside gardens. The first two groups are normally found in published checklists, but the latter two are often excluded (Pyšek 2003, see also definitions in Wisskirchen & Haeupler 1998). However, to serve as a taxonomic reference for vegetation-plot data, all taxa that can occur in relevés need to be covered, irrespective of their floristic status (Jansen & Dengler 2008).

- **While most checklists include only those hybrids that occur at least partly independently of their parents (e.g. Wisskirchen & Haeupler 1998), EuroSL would need to cover all hybrids known to occur in the wild.** These are far more than those usually listed; for example, Jansen & Dengler (2008) included several hundred hybrids in GermanSL, mostly based on Jäger & Werner (2005).

- **Floras and checklists aim at comprehensive coverage of species, but the extent of inclusion of infraspecific taxa varies.** For European liverworts and hornworts, Grolle & Long (2000) restricted themselves to species, while perhaps the majority of recent floras try to be comprehensive down to subspecies level (e.g. Tutin et al. 1968–1993, Euro+Med 2006, Buttler & Hand 2008). Several floras/checklists of vascular plants (e.g. Wisskirchen & Haeupler 1998, Jonsell 2000–2001, Jonsell & Karlsson 2010) and many for bryophytes and lichens (Koperski et al. 2006, Buettel & Hand 2008) also include varieties. It should be noted that the ranks of subspecies and variety, and partially even species and subspecies, are sometimes used interchangeably when comparing floras of different regions or main taxa (Stuessy 2009). Only rarely is a clearly different meaning assigned to these ranks (but see Jonsell 2004). In conclusion, this means that EuroSL should cover all species, subspecies, and varieties as far as they are biosystematically meaningful units (see also recommendation of Stuessy 2009: p. 161). By contrast, subvarieties, forms, and subforms, which are hardly ever treated in modern floras/checklists, because their biosystematic meaning is often doubtful, should be excluded from EuroSL (compare Stuessy loc. cit.) except when they are synonyms of what is now considered a variety, subspecies, or species.

- **Regarding supraspecific taxa between the level of species and genus, there are both formal (e.g. section, series) and informal (aggregate, species group) examples in use, often corresponding to each other (e.g. in Central Europe *Centaurea sect. Acerocentron = C. scabiosa agg.*). Originally, the concept of species aggregates (also called aggregate species, species groups or species complexes) was introduced by plant taxonomists (e.g. Manton 1958) to deal with groups of presumably closely related and morphologically similar species, whose taxonomic relationship is not yet fully resolved or where individual species cannot always be distinguished. Such cases are frequent in apomictic groups (e.g. *Rubus, Hieracium, Taraxacum*) and polyplody complexes (e.g. *Festuca ovina agg., Leucanthemum vulgare agg.*). In the past, species aggregates (or similar concepts) were widely used as informal taxonomic units in some major floras/checklists, both of vascular plants (Tutin et al. 1968–1993, Ehrendorfer 1973, Wisskirchen & Haeupler 1998) and bryophytes (Frahm & Frey 2004). While “species aggregates” (and equivalents) are still widely used by plant taxonomists in their publications (1040 hits with this topic in the *Web of Science* categories “plant sciences” and “mycology” since 1985 as of 2012-04-29), there is a strong tendency to “avoid” aggregates in recent checklists. For example, Grolle & Long (2000), Kopperski et al. (2000), Scholz (2000), Santesson et al. (2004), Hill et al. (2006) also include varieties. It should be noted that the ranks of subspecies and variety, and partially even species and subspecies, are sometimes used interchangeably when comparing floras of different regions or main taxa. Without aggregations it would be impossible to assign the vast amount of legacy data to recent taxonomic views (without considerable information loss at the genus level). Moreover, identification at species level is not always possible even for specialists (e.g. “microspecies” of *Taraxacum* can only be determined, if at all, based on plants collected in at the optimum time). For practical applications, it is therefore crucial that checklists like EuroSL adopt species aggregates (and/or similar taxonomic
Handling of taxa

Beyond the question of which taxa are included, how they are represented and "handled" is crucial for the functionality and applicability of EuroSL:

- The main taxonomic ranks from species upwards (species – genus – family – order – class – division) should be populated throughout. Subspecies and varieties should be added where accepted as meaningful taxonomic entities, as should formal and informal units between species and genus where this is appropriate for practical reasons (see above).
- Each accepted taxon should be placed in the taxonomic hierarchy, with connection to both the included taxa of the next lower level and the superior taxon of the next higher level. These connections allow for easy aggregation when a partly incomplete determination, or a particular research question, calls for analyses at a higher level than the terminal taxonomic units available. It is important that the units of the subsequent levels are fully nested, and this is also the case with informal taxa between species and genus level. Such informal taxa will be defined by listing the included members in the next lower taxonomic level.
- To reflect this full nestedness and the basic logic of the classification system, if both varieties and subspecies are accepted within one species, all varieties need to be placed within one of the subspecies, and not directly under the species. This should also be reflected in the naming, even though this is not explicitly required by the International Code of Nomenclature for algae, fungi, and plants (ICN, see McNeill & Turland 2011; formerly ICBN, McNeill et al. 2006). When, for example, var. serpentinicola is accepted within Rumex acetosa subsp. acetosa, the output name should be R. acetosa subsp. acetosa var. serpentinicola and not just R. acetosa var. serpentinicola.
- Hybrids should be included with both their formula and, if existing, their binomen. In the latter case, the formula should be treated like a synonym that refers to the binomen as the accepted name.
- Concept synonymsies (Berendsohn 1995, Koperski et al. 2000, Jansen & Dengler 2008) should be supported by the system and included wherever it is useful in practice. One main purpose of EuroSL, like that of any printed or electronic taxonomic checklist, is to provide one accepted/preferred/standard taxonomic view to be applied. However, another is to assign the many synonyms that have been used as accepted names in relevant datasets or publications correctly to units of the standard view. The number of such names which have to be treated here typically exceeds that of the accepted names (Jansen & Dengler 2008, Kalwij in press). EuroSL would clearly need to go beyond usual taxonomic checklists by focusing on the concepts behind the names used, irrespective of their correct use in the past.

Major considerations for applying the EuroSL to vegetation-pilot or other survey databases are:
- Generally, the names of the original sources should not be replaced with the accepted name of the EuroSL. Instead, both the original name and the “interpreted” name should be stored in the system, together with information on who made this assignment, preferably connected with a reasoning (e.g. by referencing to a taxonomic publication).
- A taxon name from a source should be assigned automatically to an accepted name of the EuroSL only if there is no ambiguity. In all other cases, the assignments should be made manually, but the system should preselect possible assignments and their likelihood under the given circumstances. The following cases preclude automatic assignment of name concepts: (i) Taxonyms are the most important problem regarding the number of affected names. Here, the same name is applied in the literature to different taxonomic units. Such a case is exemplified by Jansen & Dengler (2010: their Fig. 1), who show that the correctly applied name “Festuca ovina L.” means different things in three different floras, corresponding to either F. ovina L., F. ovina L. + F. guestfalica Boenn. ex Rchb., or F. ovina agg. in Wisskirchen & Haecpul (1998). (ii) Misapplied names (pseudonyms) are names used for content that does not match its nomenclatural type. This problem can only be solved by knowing for which content a name has been used in a certain regional flora (e.g. based on the determination keys used there). (iii) Homonyms occur when two different taxa have been described under the same name, one being named illegitimately and needing to be given another name according to ICN. It is important to be aware that adding authorities to a name would only solve the problem of homonyms, but not of taxonyms (when the content is largely different despite the same correctly applied authority) nor of pseudonyms (when the users are not normally aware of their misapplication, so the authority given in the source just applies to the nomenclatural publication event, not to the concept of the original author).
- To handle cases where a former species has been split into two or more species due to new taxonomic information, EuroSL would need a way to handle information (such as relevés) connected to the older, wider species concept that is now superseded by a narrower concept. One example is the moss species Hedwigia ciliata, which was split into two species in 1994, a concept now followed by nearly all recent checklists. According to ICN, one of the narrowly delimited species takes the name of the formerly more widely delimited species without any modification, while the other gets a new name (H. stellata). In such cases of species splitting, we propose the establishment of “automatic aggregates” (Dengler 2006, Jansen & Dengler 2010). In this case, “Hedwigia ciliata agg.” corresponds to H. ciliata records before 1994 or H. ciliata + H. stellata (after 1994). Care must be taken that legacy data based on the old, broader concept are not connected to H. ciliata but to H. ciliata agg. in the new taxonomic view.
- If a recent taxonomic study proposes the splitting of a species, subspecies or variety into two or more new taxa, but this new view is not yet accepted as the standard view in EuroSL (but might well be adopted in the future), we suggest “informal segregates” to store and handle this information. Jansen & Dengler (2008) proposed that taxa that are not yet generally accepted should be indicated in the standard view of German with “*” before the epithet instead of “subsp.” or “var.”. This might also be an option for EuroSL. In the Hedwigia example, this would mean that before the concept of two species was officially adopted in EuroSL, for a
while there would be an intermediate stage with one species *H. ciliata* containing two informal segregates: *H. ciliata* *l* *l* *i* *t* *i* *a* and *H. ciliata* *s* *t* *e* *l* *l* *a* *t* *a*.

If a database system is used that handles concept synonyms, this can be depicted as concept relationship between different treatments.

Finally, for the best handling of taxa with the least possible information loss when joining data from different sources, some further aspects need to be considered:

- **EuroSL should contain, wherever possible, distribution information for each taxon, in terms of the European countries or comparable widely-used entities** such as large islands and archipelagos (e.g. Tutin et al. 1968–1993, Söderström et al. 2002, Euro+Med 2006). In the simplest case, this content could be just presence/absence information, but it could also be any available information on the floristic status (native, archaeophyte, neophyte, non-naturalised neophyte, only occurring in cultivation) in each of the geographic units.

- **Beyond being valuable information for analyses, such presence/absence data at all hierarchical levels would help improve data quality in the case of “regionally monotypic taxa”**. This term refers to the widespread practice of researchers assigning or using only the species name when they actually mean the typical, or only, subspecies (or variety) occurring in a certain territory. This practice causes undue information loss once the data are combined with data from other territories where other subspecies of the species occur (Söderström et al. 2002, Jansen & Dengler 2010). With appropriate distribution information at the infraspecific level, this information loss could be avoided. For example, if the system “knows” that, of the four subspecies of *Silene latifolia* accepted by Tutin et al. (1993), only subsp. *alba* is present in Germany, any record of “Silene latifolia” from Germany could automatically be interpreted as *S. latifolia* subsp. *alba* before joining the German data with data from other countries (Jansen & Dengler 2010).

**Documentation and versioning**

While electronic checklists have many advantages over their printed predecessors, inappropriate documentation and lack of versioning might lead to lower data quality and usability in other aspects.

Here we highlight some key issues that we consider crucial for a future EuroSL, but which might also be seen as good practice in the preparation of electronic reference lists in general.

- **Documenting the meaning:** Just providing accepted names and synonyms is insufficient for an easily applicable and meaningful electronic reference list. Instead, such a list would also need to “explain” to its users the meaning of the taxonomic units and the justifications for the specific taxonomic treatment adopted. The optimal solution would be a complete European flora, with all the relevant information in one place and including the accepted name, synonyms, justification for the adopted treatment, determination key, description of the taxon, and distribution information. While name, synonyms and country-based distribution information should be the core content of EuroSL, the justification, determination key and description would not need to be included directly and in a uniform manner across all taxa, but instead could be presented via referencing to external sources.

- **Versioning:** Besides having a dynamic checklist where the content can be changed continuously to adopt new knowledge and eliminate mistakes, it is also indispensable to have accessible fixed versions. Only fixed versions can readily be used as a reference for, *inter alia*, vegetation-plot databases. Up to now, only few online information systems on species have recognised this need, among them *Species2000/Catalogue of Life* (Bisby et al. 2012; [http://www.catalogueoflife.org/annual-checklist/]), which provides both a “dynamic checklist” and fixed editions every year, which can be retrieved at a later point in time. For easy handling and flexible application, major fixed versions should, in their entirety, be available not only as an interactive database, but also as a publication. The latter should be of structured format, such as a book, which could be printed on demand or downloaded as a pdf file.

- **Connecting concepts:** Finally, clear connections between fixed versions are crucial. They would allow automatic transfer of all data connected to the name concepts between successive versions in an automatic manner without undue information loss. This necessitates that each major new version provides explicit matches for any taxon accepted in the previous version. While this fundamental principle is also implemented for GermanSL (Jansen & Dengler 2008), most authors of other (printed or electronic) checklists seem to believe that once the new version is published, the last one is immediately outdated. In their mind it would therefore be wasted effort to document the relationships of the taxonomic concepts between versions. For example, the most recent checklists of German vascular plant (*Buttler & Hand 2008*) and *lichen* (*Wirth et al. 2011*) floras come without any documented connection to their immediate predecessors (*Wisskirchen & Haeupler 1998, Scholz 2000*). To ensure usability of existing databases, each new edition of a checklist should explicitly document differences compared to the preceding edition, including an unambiguous “translation” of taxon concepts that have changed. It would be even better if such connections were also provided to some of the major floras (see the excellent example of Koperski et al. 2000). However, as this example clearly shows, documenting the relationships of the taxonomic concepts in such a way represents a considerable additional effort, and it should be discussed to what extent this can be applied for all taxa.

**Steps towards establishment of EuroSL**

During the first EuroSL workshop in Göttingen, and subsequently, the authors of this contribution have discussed how the ideas outlined in the previous section could be put into practice. First, data sources and software tools that already fulfil some of the requirements, and could thus form part of a future EuroSL, have to be identified. Then we need to find further collaborators, supporting institutions, and funding.

**Possible ingredients of EuroSL**

For vascular plants, *Flora Europaea* (Tutin et al. 1968–1993), though taxonomically outdated in many respects, is still the most recent and best documented flora for the whole of Europe. While no new edition or successor of *Flora Europaea* is in sight, *Euro+Med PlantBase* (Euro+Med 2006) aims to develop an online checklist with distribution information, which is mainly based on *Flora Europaea* (Tutin et al. 1968–1993). *MedChecklist* (Greuter et al. 1984 et seq.) and
Flora of Macaronesia (Hansen & Sunding 1993), updated by subject editors for individual plant families. Presently Euro+Med PlantBase covers about 90% of vascular plant species of Europe, the circum-Mediterranean countries, the Madeira Archipelago and Canary Islands, and is already in use as a thesaurus for BioCASE portals (Biological Collection Access Services, http://www.biocase.org) of the GBIF network. However, it currently exists only as a non-versioned, continuously updated online source. As Euro+Med PlantBase is the only modern continent-wide checklist of vascular plants, it is the natural basis of EuroSL for vascular plants. However, for the approximately 10% of species that are missing, EuroSL would have to rely on other sources, chiefly the classical treatments in Flora Europaea, until they are replaced by a new treatment from Euro+Med PlantBase in future versions of EuroSL. Other shortcomings of Euro+Med PlantBase that will have to be “worked around” in EuroSL are: (i) the current lack of fixed versions (but planned for the future); (ii) the lack of explicit and consistent links to the taxonomic concepts in the Flora Europaea (on which the Euro+Med PlantBase is based) or the more recent Atlas Flora Europaea (Jalas & Suominen 1994, 1996, 1999, Kurtto et al. 2004–2010); (iii) the low coverage of informal aggregates, varieties, hybrids, non-naturalised neophytes and cultivated plants; and (iv) the lack of links between new taxonomic concepts and references or justifications.

The bryophyte portion of EuroSL can be based on up-to-date European and Macaronesian checklists for liverworts and hornworts (Grolle & Long 2000, Sõderström et al. 2002) and mosses (Hill et al. 2006). There is even country-based distribution information for liverworts and hornworts (Sõderström et al. 2002, 2007) and peat mosses (Sõnens & Sõderström 2009). This information could enter EuroSL directly or via Euro+Med PlantBase, which plans to incorporate the bryological divisions from the same sources in the future.

For lichens, in the complete absence of a comprehensive continental checklist or flora, the EuroSL treatment can only be based on the most comprehensive national to subcontinental floras and checklists: Italy (Nimis & Mantellus 2008: 2,345 taxa), Fennoscandia (Santesson et al. 2004: 2,538 taxa), Iberian Peninsula (Hladun & Llimona 2007: 2,767 taxa), Great Britain and Ireland (Smith et al. 2009: 1,873 taxa), and Germany (Wirth et al. 2011: 2,380 taxa). As the integration of the different taxonomic concepts of these sources is beyond the scope of EuroSL, the best solution would be if the lichenologists themselves prepare a continental checklist with uniform taxonomy.

For the “algal groups”, the situation is even more complicated, as the knowledge on individual classes is split among many specialists. AlgaeBase (Guiry & Guiry 2012) might be a major contribution, but it is unclear how to extract the European alga from this global database.

Unfortunately, all the cited checklists largely exclude non-naturalised neophytes, cultivated plants, hybrids, and informal aggregates. Therefore, EuroSL would need to add this content from other sources, of which SynBioSys Europe (Schaminée et al. 2007) and GermanSL (Jansen & Dengler 2008) probably form the best starting point. However, it would require a huge specialist effort to match these sources with the taxonomic views adopted from a certain reference list. For aggregates, Flora Europaea (Tutin et al. 1968–1993), Atlas Flora Europaea (Jalas & Suominen 1994, Jalas et al. 1996, 1999, Kurtto et al. 2004–2010), Ehrendorfer (1973) and the “automatic aggregates” (see subsection Handling of concepts) and is in use as a starting point. Even more complicated would be the “algal groups”, the situation is even more complicated, as the knowledge on individual classes is split among many specialists. AlgaeBase (Guiry & Guiry 2012) might be a major contribution, but it is unclear how to extract the European alga from this global database.

For the “algal groups”, the situation is even more complicated, as the knowledge on individual classes is split among many specialists. AlgaeBase (Guiry & Guiry 2012) might be a major contribution, but it is unclear how to extract the European alga from this global database.

Unfortunately, all the cited checklists largely exclude non-naturalised neophytes, cultivated plants, hybrids, and informal aggregates. Therefore, EuroSL would need to add this content from other sources, of which SynBioSys Europe (Schaminée et al. 2007) and GermanSL (Jansen & Dengler 2008) probably form the best starting point. However, it would require a huge specialist effort to match these sources with the taxonomic views adopted from a certain reference list. For aggregates, Flora Europaea (Tutin et al. 1968–1993), Atlas Flora Europaea (Jalas & Suominen 1994, Jalas et al. 1996, 1999, Kurtto et al. 2004–2010), Ehrendorfer (1973) and the “automatic aggregates” (see subsection Handling of concepts) and is in use as a starting point. Even more complicated would be the “algal groups”, the situation is even more complicated, as the knowledge on individual classes is split among many specialists. AlgaeBase (Guiry & Guiry 2012) might be a major contribution, but it is unclear how to extract the European alga from this global database.

Medium- and long-term perspective

The participants of the first EuroSL Workshop concluded that EuroSL can be initiated based on Euro+Med PlantBase (providing a standard view on plant taxonomy in Europe) and the aggregated checklist of approximately 30 national and regional TURBOVEG species lists in SynBioSys Europe (connecting this view to many different views used across Europe). Once a first fixed version of Euro+Med PlantBase is released, this should be augmented by including the missing taxonomic groups (remaining vascular families, bryophytes, lichens, algae) to achieve a standard list. This released version should then be linked to the adopted (but not documented) standard view of SynBioSys Europe, taxon by taxon, which in turn would provide a reasonable first connection to a variety of regional taxonomic views. When this integration of the major sources is achieved, the resulting product could be launched as EuroSL 1.0.

After this launch, continuous updates would be needed to: (i) add missing taxa (rare casual aliens, aggregates, small algal groups, lichens from not yet covered countries); (ii) add and update distribution information on the taxa; (iii) edit and correct the concept synonymy relationships; and (iv) incorporate new taxonomic views resulting from recent research, firstly as alternative views and, when they are widely accepted as the standard views of EuroSL, to supersede the previous taxonomic views.

Such complex processes will need appropriate workflows and functioning software tools to implement them. These tools should allow the distribution of work among various, scattered specialists, assist comprehensive documentation of decisions, and allow for a reviewing process before new pieces go live. As discussed above, the workflow should combine continuous updates (after review) and the release of fixed major versions at intervals of one to several years. The software developed by the European Network of Excellence EDIT (European Distributed Institute of Taxonomy), the “EDIT Platform for Cybertaxonomy” (e.g. Berendsohn 2010), already covers many of the prerequisites named here (e.g. handling of concepts) and is in use for PESI including Euro+Med PlantBase. It is based on the EDIT Common Data Model (CDM), which in turn attempts to comprehensively cover existing commu-
inity data standards and thus afford the greatest possible interoperability with the multitude of available data sources. However, the effort needed to fully adapt such a comprehensive information system according to the ambitious plan outlined here has to be carefully analysed and will certainly need considerable additional resources.

Evidently, a high-quality EuroSL will only be achievable when a broad consensus among specialists is reached to accept EuroSL as a common goal for the different disciplines involved. This will also largely depend on the ability to attract appropriate funding to support the necessary new research and enormous initial editorial effort. In the long term, institutional commitment from organisations executing taxonomic research as well as those providing biodiversityinformatics infrastructures in Europe, will be crucial to ensure sustainability.

Conclusions and outlook

In this contribution, we have emphasised the urgent need for a novel electronic reference list of all plants in Europe, described its specifications, and briefly outlined our strategy to get there. The need for EuroSL has been mainly articulated by vegetation scientists who see it as an essential tool when combining large vegetation-plot databases across national borders. However, it is clear that the proposed features of EuroSL would also be highly beneficial for any other type of plant-taxon related database, such as plant-trait databases, distribution databases, mapping and collection databases, phylogenetic databases or conservation-related databases. EuroSL could also add, at least for plants, currently absent functionalities to the PESI database of the European Union.

While the establishment of EuroSL seems highly desirable and important, reaching this goal will not be easy. We need to motivate specialists from many different fields to contribute their competences, to attract substantial funding and to find institutions that are willing to commit themselves to serve as reliable, permanent hosts of a continuously updated EuroSL. Putting EuroSL into practice will only be possible when taxonomists, ecologists and other users, as well as biodiversity data specialists, efficiently work together and are willing to understand the perspectives and needs of the other groups. The authors of this article will continue to develop the ideas outlined here and work towards their implementation. We are inviting individuals and institutions that share our vision to contribute their knowledge or infrastructure to join the planning process.

Acknowledgements

We thank Gabriela Lopez-Gonzalez, Jörg Ewald, and one anonymous referee for a very fast peer review and useful suggestions on an earlier version of the manuscript. Will Simonson improved the English language usage. W.-H.K. acknowledges the GBIF-D project, supported by the Federal Ministry of Education and Research (01 LI 1001 A).

References


22

Biodiversity & Ecology 4 2012


Jürgen Dengler* (juergen.dengler@uni-hamburg.de) Biodiversity, Evolution and Ecology of Plants (BEE), Biocentre Klein Flottbek and Botanical Garden, University of Hamburg Ohnhorstr. 18 22609 Hamburg, GERMANY

Walter G. Berendsohn (berendsohn@bgbm.org), Wolf-Henning Kusber (w.h.kusber@bgbm.org), Andreas Müller (a.mueller@bgbm.org) & Eckhard von Raab-Straube (e.raab-straube@bgbm.org) Botanic Garden and Botanical Museum Berlin-Dahlem, Freie Universität Berlin Königstr.-Luisenstr. 6–8 14195 Berlin, GERMANY

Erwin Bergmeier (erwin.bergmeier@bio.uni-goettingen.de) Dept. Vegetation and Phytodiversity Analysis, Albrecht von Haller Institute of Plant Sciences, University of Göttingen Untere Karpspüle 2 37073 Göttingen, GERMANY

Milan Chytrý (chytry@sci.muni.cz) & Jiří Danhelka (danhelk@sci.muni.cz) Department of Botany and Zoology, Masaryk University Kotlářská 2 611 37 Brno, CZECH REPUBLIC

Jiří Danhelka Department of Vegetation Ecology, Institute of Botany, Academy of Sciences of the Czech Republic Lidická 25/27 602 00 Brno, CZECH REPUBLIC

Florian Jansen (jansen@uni-greifswald.de) Landscape Ecology and Ecosystem Dynamics (LEED), Institute of Botany and Landscape Ecology, University of Greifswald Grimmer Str. 88 17487 Greifswald, GERMANY