Portugaliae Acta Biologica **19**: 49-59. Lisboa, 2000

THE FLORA HELLENICA DATABASE

A. Strid
Botanic Garden, Carl Skottsbergs gata 22, S-41319 Göteborg, Sweden
E-Mail: arne.strid@gotbot.se

Strid, A. (2000). The Flora Hellenica database. *Portugaliae Acta Biologica* **19**: 49-59.

Flora Hellenica will be the first complete Flora of Greece to appear in about 100 years. Volume one was published in 1997 and a further eight volumes are planned. A floristic database registering all herbarium collections, literature reports and field notes, and currently holding c. 506,000 records, is an important feature of the project. The structure and contents of the database are presented, and examples are given of retrieval and use of stored data in quantitative phytogeography, phenology, etc.

Key words: Greece, flora, database.

Strid, A. (2000). Base de dados da Flora Helenica. *Portugaliae Acta Biologica* **19**: 49-59.

A *Flora Hellenica* será a primeira Flora de Grécia completa a aparecer nos últimos 100 anos. O primeiro volume foi publicado em 1997 e estão planeados mais oito volumes adicionais. Uma base de dados florística que regista todas as colecções de herbário, citações bibliográficas e registos de campo, e que contem actualmente cerca de 506.000 registos, é uma das características importantes do projecto. Apresenta-se a estrutura e conteúdo da base de dados, assim como exemplos da recuperação e uso dos dados armazenados em estudos quantitativos de fitogeografia, fenologia, etc.

Palavras chave: Grécia, flora, base de dados.

INTRODUCTION

Greece is a country with great floristic traditions, being the subject of the splendid *Flora Graeca* of SIBTHORP & SMITH, which appeared in the years 1806-40. The story of this great Flora with nearly 1,000 hand-coloured copper engravings based on the drawings by Ferdinand Bauer has recently been told in a book by LACK & MABBERLEY (1999). The most recent complete Flora of

Greece is that of HALÁCSY which appeared in 1900-04 with supplements in 1908 and 1912; this is a careful and critical work which was excellent for its time but is now long since outdated.

Vast amounts of published and unpublished data relevant to the vascular plants of Greece have accumulated since the time of Linnaeus. A recent bibliography (STRID 1996a) listed 10,241 titles from 1753 to 1994, and the number has subsequently increased to 11,436 (11.10. 1999). Major floristic works have appeared for certain regions, especially the Aegean islands and the mountains. Greek islands and mountains have also proved to be suitable for the study of evolution and speciation, and many papers in the field of cyto-taxonomy and experimental taxonomy have appeared from the nineteen sixties onwards. A brief account of the botanical exploration of Greece is given in the introduction to volume one of *Flora Hellenica* (STRID & KIT TAN 1997: xxx-xxxiv).

The flora of Greece is currently estimated to comprise c. 5,800 species of native and naturalised vascular plants, including some 740 endemics. Modern national Floras now exist for most other countries in the Mediterranean area, and it has been evident for some time that a critical survey of the floristic richness of Greece is urgently needed.

FLORA HELLENICA AND ITS DATABASE

After several years of preparation the first tome of a planned nine-volume work, *Flora Hellenica*, appeared in September 1997. Families and genera are presented in the same sequence as in *Flora Europaea*, the first volume covering Gymnospermae to Caryophyllaceae. *Flora Hellenica* includes the elements normally associated with a critical Flora, i.e. diagnostic keys, typification and synonymy, descriptions, and data on ecology, distribution, flowering time, chromosome numbers, variation and affinities, etc. Dot maps are given for all species and subspecies; they have been automatically generated from an electronic database where each record includes geographical coordinates (degrees and minutes of latitude and longitude).

Work on the *Flora Hellenica Database* started around 1988 and has mainly been carried out in Copenhagen and Lund (cf. STRID & KIT TAN 1997: xxxvi). By 11.10. 1999 the main database (the *Specimen Database*) comprised 391,209 records; to this should be added 115,116 records, mainly from Crete and the other south Aegean islands, which have been kindly made available in a compatible format by Ralf Jahn and Bernhard Egli, but are kept separate for the time being. It is believed that the specimen database now comprises at least half of all the existing collections, literature records and field notes of Greek vascular plants; it is thus comprehensive enough to become a powerful tool not only in the preparation of the Flora but for a wide range of purposes.

STRUCTURE AND CONTENTS OF THE DATABASE

The *Flora Hellenica Database* is a relational database created in the well-known commercial programme Paradox (DOS version). Data is organised in 27 fields and includes all the information normally found on a good herbarium label, e.g. name of plant, collector, collecting number, date of collecting, locality, altitude, and herbaria where the specimens are kept. Most importantly, there are separate fields for latitude and longitude (degrees and minutes); values in these fields can be linked to a graphics programme and are used for generating dot maps of individual species (or maps based on any other combination of search criteria).

To minimise errors, data validation in the form of spelling checks is employed for names of taxon, collector, nomos (county or province), eparchia (a smaller administrative unit), mountain and island. Collecting date must be in the form yyyy-mm-dd. The altitude field is divided into two subfields since altitude is often given as a range (e.g., 650-800 m); altitude is in meters and must be an integer between 0 and 2917 (the altitude of Mt. Olympus, the highest summit in Greece). Attempt to type a name or value not included in the list of accepted values for the particular field will generate an error message and an alternative will sometimes be suggested. The spelling mistake *"Ranunculus gvacilis"*, for instance, will generate the suggestion *"Ranunculus gracilis"*.

Typing mistakes are more easily made with numerals than with text. An important item of data validation is crosschecking between the Eparchia field and the Latitude and Longitude fields. Each eparchia has a given north-south and east-west extension, eparchia Githiou in nomos Lakonias (S Peloponnisos), for instance, is between lat. 36°23′-36°49′N and long. 22°22′-22°35′E. Once the name Githiou has been typed in the Eparchia field, an attempt to type a value outside the correct range for latitude and longitude in the respective fields will result in an error message.

Data entry is very time-consuming, and attempts have been made to rationalise the process in different ways, primarily to avoid the typing of repetitive information as far as possible. Obvious examples are names of taxa: when, for instance, the operator types *"Johrenia distans"* for the name of a species, the values "Apiaceae" for family and "(Griseb.) Halácsy" for authority are added automatically to their respective fields, having been copied from a linked Taxa Database.

Most modern collectors have consecutively numbered series of specimens. For instance, Runemark & Snogerup nos 20739-20760 were all collected in the same place on the same day ("Nom. Lakonias, ep. Epidavrou Limiras: Monemvasia, cliffs around the castle, 20-50 m. 23.5. 1964. Lat. 36°41′N, long. 23°03′E"). Such repetitive data, whenever known, has been typed into a Locality Database, which is linked to the Specimen Database and now contains more than 24.000 records. With an actual specimen in front of him, the operator will only need to type, for instance, "Runemark & Snogerup 20746", and data will be

automatically copied from the Locality Database. Similarly, a simple repeat function makes it possible to register multiple records from the same locality, e.g. from a published list.

The database has the normal functions for sorting, querying, producing reports, exporting data, etc. Special scripts are used for printing herbarium labels and for producing lists of specimens (e.g., all records of *Anemone coronaria* from the East Aegean islands). A contributor commencing work on a particular genus for *Flora Hellenica* is provided with a complete list of specimens and a set of dot maps for the individual species and subspecies as well as a list of relevant literature (the latter extracted from the separate bibliographical database). It is up to the individual contributor to make a critical assessment of the information provided (weeding out, for instance, dubious literature records and correcting misidentifications after having studied the appropriate herbarium material) and to add new data. All records are coded into one of four categories: ”V” for verified herbarium material (i.e., specimens seen and confirmed by the person revising this particular genus for *Flora Hellenica*; ”H” for other herbarium material; ”L” for published literature record (reference to author, year and page being given in a special box); and ”F” for field note (taken from notebooks, typewritten lists, etc.). In critical groups (e.g., *Alchemilla*) it may be decided to map only records in the ”V” category, whereas for unmistakable species (e.g., *Euphorbia dendroides*) all records will presumably be accepted.

USE OF THE DATABASE

Floristics and Phytogeography

An obvious use of databased floristic information is in the field of quantitative biogeography. Several examples are given in STRID (1996b, c) and KIT TAN & STRID (1999). Floristic similarities and differences, e.g. between individual mountains or islands, can be calculated with relative ease, and generalised distribution patterns and phytogeographical borders can be revealed and visualised.

Volume one of *Flora Hellenica* comprises 622 species, or c. 11 per cent of the total. In the database there are 55,625 records of these, or an average of 89.4 records per species. For the purpose of *Flora Hellenica* Greece has been divided into 13 regions (Fig. 1); this subdivision is largely a practical one, but also reflects phytogeographical borders to some extent, at least in the Aegean area.

The number of records, the number of species and the average number of records per species for each of the 13 regions are indicated in Fig. 1. The North Aegean Islands (NAe) can be disregarded as a somewhat atypical region, since it comprises only three islands (Thasos, Samothraki and Limnos). The remaining 12 regions are more or less comparable, and the number of records per species can be taken as a rough indication of the level of floristic exploration.

Fig. 1. Map of Greece indicating the 13 regions adopted for the *Flora Hellenica* project. The number of database records, the number of species, and the number of records per species are shown for each of the 13 regions, based on taxa included in volume one of the Flora (c. 11 per cent of the total). Further explanation in the text.

Floristic activity has obviously been most intense in the South Aegean area (Kiklades with 35.8 records per species and Kriti/Karpathos with 26.1). In the Kiklades most of the records have been provided by Runemark, Snogerup and co-workers in Lund, and in the Kriti/Karpathos area mainly by Rechinger, Greuter, Raus and several other German collectors. The vast amount of field notes by Runemark & al. for the Kiklades and by Jahn, Egli and others for Kriti/Karpathos, which have not yet been included in the main database, would have underlined further the fact that these regions are the best explored in Greece from a floristic point of view. Also the East Aegean islands (EAe) and Peloponnisos (Pe) are high on the list. For the East Aegean islands large numbers

of records have been provided by Carlström (Rodos), Snogerup (Chios) and Nielsen (Lesvos); for the Peloponnese by many different collectors including several from the University of Patras.

When disregarding the atypical North Aegean region, the lowest values for records per species are found in the Southern Pindos (SPi) with 8.5 and the East Central (EC) with 9.8. In the Southern Pindos, about one third of the records have been provided by a single collector (Aldén) who, however, concentrated heavily on the upper montane and alpine levels; the lowlands and the lower montane level remain rather poorly explored. In the East Central region, Mt Ossa, Mt. Pilio and the Magnisias peninsula SE of Volos have been well explored by Raus, but the rest of the region is largely made up of cultivated lowland which has proved unattractive to collectors.

Data such as these can obviously be useful when planning the further floristic exploration of Greece. Many field botanists have a tendency to come back to well-known, classical localities and to disregard ”trivial” areas in between. In some cases this may be justified (for instance, the heavily cultivated Thessalian plain is never likely to yield much of botanical interest) whereas other areas, such as low altitudes in the Pindos and the extreme north east of Thraki, are clearly underexplored in relation to their floristic importance.

Rare and Common Species

The largest genus in volume one of *Flora Hellenica* is *Silene* s.l. with 119 species (account by W. Greuter); for comparison, there are 83 species in *Flora Iberica* and the *Flora of Turkey* (incl. Supplement) lists 133. The total number of records for this genus in the *Flora Hellenica Database* is 10,046, and the average number of records per species is thus 84.4, or very close to the average for all species in volume one which is 89.4 (cf. above).

The number of records per species varies within wide limits; at the lower end of the scale are some very rare species with less than five records, at the upper end are *S. vulgaris* with 930, *S. sedoides* with 752 and *S. italica* with 699. Plotting the number of species against the number of records per species (Fig. 2) it becomes obvious that most species are comparatively rare; 29 of the 119 have less than 10 records each in the database, and no less than 68, or more than half, have less than 40 records; 42 species have between 40 and 259 records each, and at the upper end of the scale are nine species with more than 259.

These figures are probably representative for the Mediterranean region where habitat fragmentation - natural and man-made - provides for much local and regional differentiation. It can be assumed that a similar study for an area north of the Alps would show relatively fewer rare species.

Species Richness as a Function of Altitude

Of the 10,046 records of *Silene* in the *Flora Hellenica Database*, the altitude has been indicated for 6,840; *Silene* species are found from sea level up to the

highest summits (c. 2,900 m on Mt. Olympus). On herbarium labels, and consequently also in the database, altitude is often given as a range (e.g., 800-1,000 m), and in such cases the average for each record has been calculated.

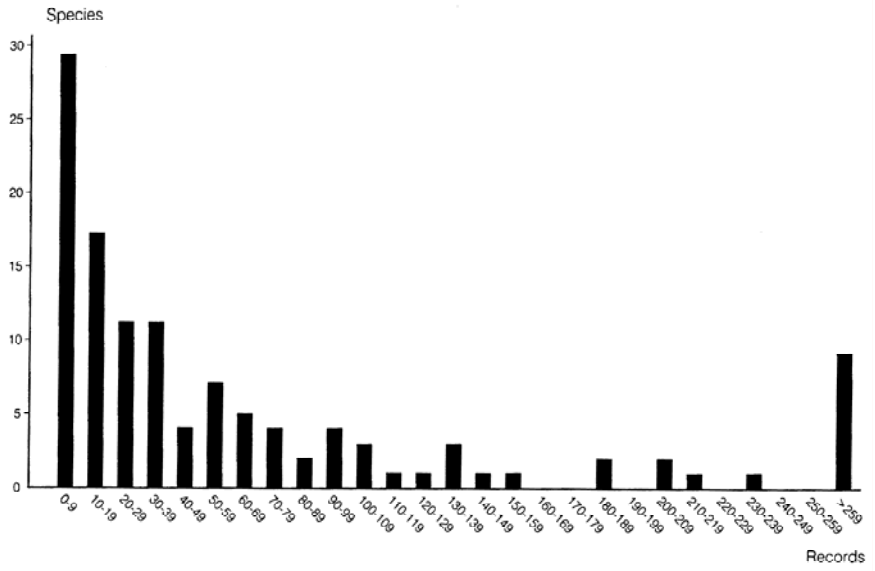

Fig. 2. The genus *Silene* in Greece (119 species). Most of the species are rare or relatively rare, having less than 40 records each in the *Flora Hellenica Database*; a few are very common with more than 259.

The 119 species of Greek *Silene* comprise 70 perennial and 49 annual or biennial (information on life form is not included in the database, but has been added for the purpose of this study). It is a well-known fact that annuals constitute a high percentage of the Mediterranean flora as compared to most other regions of the world, and that they are most prevalent in disturbed habitats. By and large, the perennials are more frequent in natural and semi-natural habitats.

Plotting the number of species as a function of altitude for 100-m intervals (Fig. 3) reveals significant differences between the two groups. The largest number of annual species (35, or 71.4 %) is found at the lowest altitude (0-99 m), but the numbers remain high (between 55.1 % and 44.9 %) up to c. 800 m and then decline gradually so that virtually no annual species are found above 1,900 m. The perennials on the other hand show a slight gradual increase in frequency from 0 to 900 m, remain at a high level up to 1,700 m and then decline gradually in frequency, but occur up to high altitudes with 17 species recorded between 2,200 and 2,299 m (versus not a single annual in this interval).

Adding other life form parameters, climatic and edaphic data, chromosome numbers, etc., to the database would make it possible to perform more sophisticated calculations and correlations of environmental and evolutionary significance.

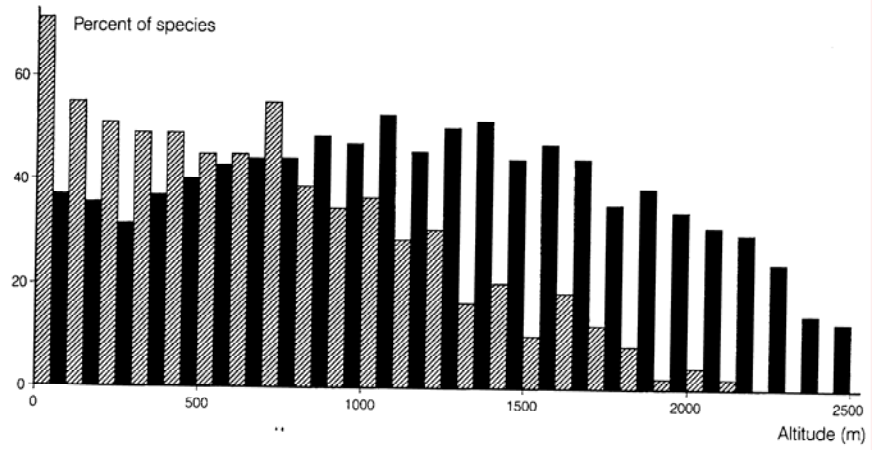

Fig. 3. The genus *Silene* in Greece. Altitudinal distribution of 70 perennial species (black columns) and 49 annual or biennial species (hatched columns). The latter are most frequent at low altitudes (up to 800 m) and scarcely extend above 1900 m, whereas the perennials are more evenly distributed and extend to higher altitudes.

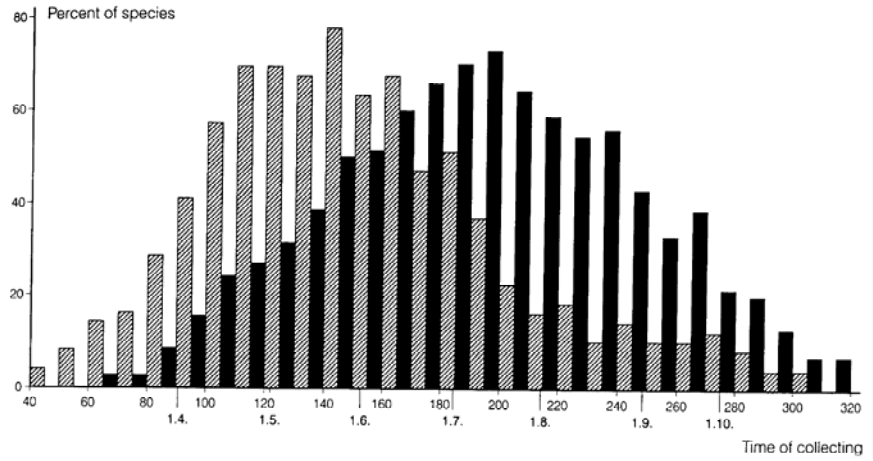

Fig. 4. The genus *Silene* in Greece. Time of collecting (more or less corresponding to time of flowering) for the 70 perennial and 49 annual or biennial species (black and hatched columns, respectively). For the perennials peak flowering is in the first half of July, for the annuals almost two months earlier.

Flowering Time

Of the 10.046 *Silene* records, precise date of collecting or observation has been recorded for 8.952; the vast majority of records can be assumed to refer to flowering material. In Fig. 4, the date of collecting or observation for annual/biennial and perennial species, respectively, has been scored for ten-day intervals throughout the vegetation period. For the annual/biennial species there is a distinct peak from end of April to beginning of June, and the ”tail” of the diagram from c. 1.8. to 1.11. probably refers to freak late-germinating and late-flowering plants or to material collected in fruit. For the perennial species the peak is in the months of June and July, and there is a less pronounced ”tail” at the upper end of the diagram. The later flowering of the perennial species may to some extent reflect the fact that they tend to occur at higher altitudes, but also comparisons restricted to a particular altitudinal ínterval show significant differences in flowering period between the two life forms.

The data can be broken down in various ways. Fig 5 shows the dates of collecting/observation for four common species of *Silene* in parts of central and southern Greece (Sterea Ellas, Peloponnisos, West Aegean area and Kiklades). This can be compared to information from other sources. In an article on the flora of Attica (Attiki), HELDREICH (1877) published a ”Calendarium Florae Atticae” with flowering times for the individual species, based on his long experience from that area. The bars at the lower part of the diagrams in Fig. 5 show flowering times in Attica as indicated by Heldreich. For the annual *S. sedoides* (Fig. 5 A) Heldreich appears to have been quite accurate. For the two perennial species, *S. vulgaris* (*S. inflata*) and *S. italica* (Fig. 5 C-D), records later than mid-June are likely to refer to material collected at high altitudes in Sterea Ellas and the Peloponnese, so also in these cases Heldreich seems to have been accurate with respect to Attica. For *S. gallica* (Fig. 5 B), however – a circum-Mediterranean weed of fields, olive groves and fallow land – the flowering time indicated by Heldreich is clearly too early. GREUTER (in *Flora Hellenica* 1: 318, 1997) more accurately states the flowering time as ”March to June”, and COODE & CULLEN (in *Flora of Turkey* 2: 238, 1966) give ”April to June” as the flowering time in Turkey.

Conclusions

The structure and contents of a major floristic database must be carefully planned with two points in mind: (1) to make the structure simple and robust and to include only what is strictly necessary, and (2) to rationalise the process of data entry, use data validation to minimise errors and to avoid repetitive work whenever possible. For the *Flora Hellenica Database* with around 500,000 records, an extra 15 seconds spent on each record would translate into more than one man-year of extra labour.

Planning is the easy part. The hard, slow and painstaking part is the actual entry of data. It requires perseverance and accuracy as well as broad knowledge

of the flora and geography of the area; it is not trivial work than can be easily left to undergraduates or technicians. Any researcher planning a database of this magnitude should be prepared to spend – personally – 2-3 hours per day for the next ten years.

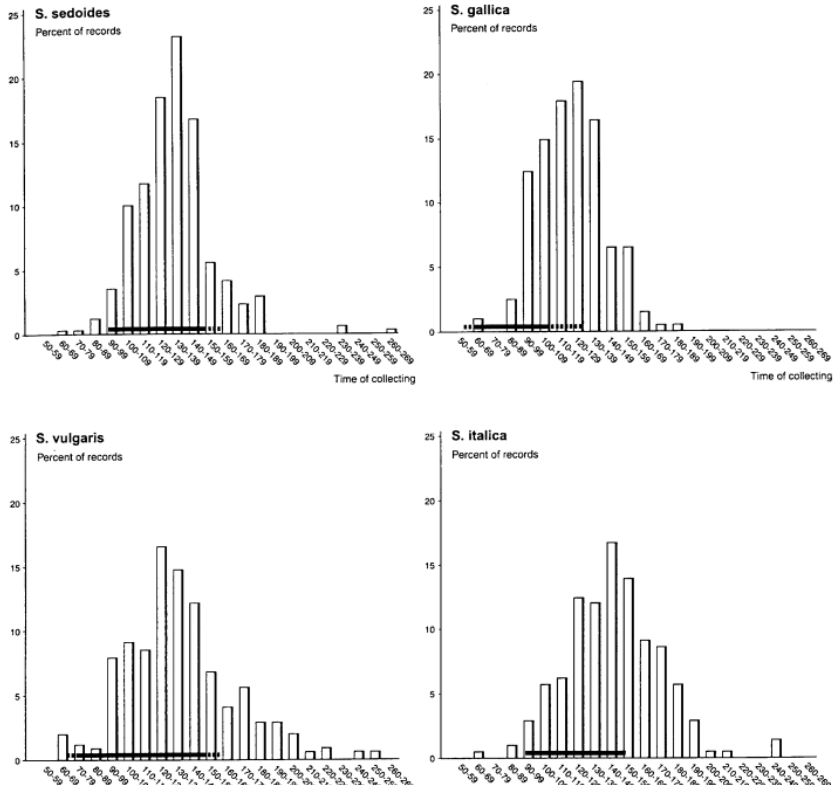

Fig. 5. Time of collecting for four common species of *Silene* (including only records for StE, Pe, WAe and Kik). The bars indicate flowering time in Attica according to Heldreich (1877), who appears to have been rather accurate *for S. sedoides, S. vulgaris* and *S. italica*, but off the mark for *S. gallica*

The *Flora Hellenica Database* has now reached a size where it becomes a powerful tool in any study related to the plant life of Greece. Reasonably accurate distribution maps for any species (or maps based on any other combination of search criteria) can be produced and printed within minutes. Adding new data, e.g. life form parameters or general distribution data for the individual species, or linking the database to other databases containing climatic or edaphic information, will make it even more useful in taxonomy, biogeography, ecology and conservation.

ACKNOWLEDGEMENTS

Financial support for the *Flora Hellenica* project has been received from several sources, notably the EU (contract no. ERBFMBICT961489), the Danish Natural Science Research Council (grant no. 9801450), the Danish Agricultural Research Council (grant no. 9601228) and the Carlsberg Foundation

REFERENCES

COODE, M.J.E. & CULLEN, J. 1966: *Silene* L. In P.H. DAVIS (ed.): Flora of Turkey and the East Aegean Islands, 2: 179-242. – Edinburgh: Edinburgh University Press.

HELDREICH, TH. v. 1877: *Die Pflanzen der attischen Ebene. In A. Mommsen (Hrsg.): Griechische Jahreszeiten* 5: pp. 471-597. – Schleswig: J. Bergas.

HALÁCSY, E. v. 1900-08: *Conspectus Florae Graecae*, vol. 1-3 + supplement. – Lipsiae: Guilelmi Engelmann (second supplement *in Magyar Bot. Lapok* 11: 114-202, 1912).

LACK, H.W. & MABBERLEY, D.J. 1999: *The Flora Graeca story. Sibthorp, Bauer, and Hawkins in the Levant.* – Oxford, etc.: Oxford University Press, xxi + 327 pp.

SIBTHORP, J. & SMITH, J.E. 1806-40: *Flora Graeca*, vol. 1-10. – Londini: Richardi Taylor et socii.

STRID, A. 1996a: Flora Hellenica bibliography. A critical survey of taxonomic, floristic and phytogeographical literature relevant to the vacsular plants of Greece, 1753-1994. – *Fragm. Florist. Geobot. Suppl.* 4: i-x, 1-508.

STRID, A. 1996b: The Greek mountain flora, with special reference to the Central European element. – *Bocconea* 5(1): 99-112.

STRID, A. 1996c: Phytogeographia Aegaea and the Flora Hellenica Database. – *Ann. Naturhist. Mus. Wien* 98B Suppl.: 99-112.

STRID, A. & TAN, KIT (eds) 1997: *Flora Hellenica*, vol. 1. – Koenigstein: Koeltz Scientific Books, xxxvi + 547 pp.

TAN, KIT & STRID, A. 1999: Phytogeographical notes on the Parnonas range in Peloponnisos, southern Greece. – *Preslia* 71: xxx-xxx (in press).