*Acta Botanica Hungarica 30 (3—4), pp. 353—362 (1984)*

# A PRELIMINARY NUMERICAL TAXONOMIC STUDY OF THE *SCILLA BIFOLIA* AGGREGATE (LILIACEAE, SCILLOIDEAE) IN HUNGARY

Z. KERESZTY and J. PODANI

RESEARCH INSTITUTE FOR ECOLOGY AND BOTANY, HUNGARIAN ACADEMY OF SCIENCES, VÁCRÁTÓT

(Recevied: 15 March 1984)

Fifteen morphological and two phenological characters were recorded in twenty-six populations of the *Scilla bifolia* aggregate grown under uniform conditions. Two multivariate techniques, sum of squares clustering and principal component analysis, were used to study the variation among populations. Both methods revealed three distinct groups, suggesting the existence of three species in Hungary. These are recognized as *S. vindobonensis*, *S. kladnii* and *S. buekkensis*, all described earlier. A fourth group may also be distinguished at a lower level, its taxonomic position remains doubtful, however.

## Introduction

The geographic variation of *Scilla bifolia* L. populations in Hungary has been the subject of recent surveys applying traditional taxonomic techniques (cf. KERESZTY 1983). The authors' observations and reports from neighbouring countries (SPETA 1974, 1977) suggest that *Scilla bifolia* is a group, including a number of more or less distinct forms, rather than a single species. However, proposals to create a new taxonomy of this complex have not yet been justified by the use of numerical techniques. The present paper is therefore an attempt to meet a long need by providing cluster analysis and ordination results. Although this study is based on populations only from Hungary, it is expected that the results will be informative and useful to an overall revision of this species complex. Since *S. bifolia* is a protected species in Hungary, the findings may also have relevance to plant conservation.

## Material and methods

A collection of twenty-six population samples from thirteen different geographic regions of Hungary was assembled. The geographic localization of sampling sites is shown by the map in Fig. 1. The regions and collecting localities are listed in Table 1, together with phytosociological and phytogeographical data. The nomenclature follows SOÓ (1964) for syntaxa, and PÓCS (1968) for phytogeographical units. The code numbers refer to the grid system used by the Central European flora mapping project (cf. EHRENDORFER 1973). In order to obtain reliable data to indicate genetic differences and similarities, the sample individuals were grown under uniform environmental conditions in the garden of the Research Institute for Botany, Vácrátót, in 1979–1980. All plants were cultivated for at least two years before measurements were taken in 1982.

Fifteen morphological and two phenological characters were selected for the study. The morphological characters were measured and recorded for ten plant individuals (OTUs) from each population at the time of full flowering, i.e., when all flowers were open. If the number of flowers per inflorescence was over six, the uppermost flower was allowed to be in bud and/or the lowermost to be ceased. Eight characters are continuous, nine are determined

**Table 1**

*List of regions and collecting sites of Scilla bifolia populations used in this study*

| Geographic region | Phytogeographic region and district | Locality | Code No. | Plant community | Altitude, m | Date |
|---|---|---|---|---|---|---|
| I. Danube Basin, upper floodplain | Eupannonicum–Arrabonicum | 1. Győr, Püspökerdő | 8371 | | 100 | 8. 5. 1979 |
| | | 2. Zsejkepuszta | 8271 | | | |
| | | 3. Vámosszabadi, Határerdő | 8272 | | | |
| II. Danube Basin, middle floodplain | Eupannonicum–Praematricum | 4. Kisoroszi | 8180 | *Fraxino pannonicae–Nlmetum* | 80 | 4. 5. 1979 |
| | Eupannonicum Colocense | 5. Ráckeve, Angyali sziget | 8779 | | | 30. 4. 1979 |
| | | 6. Makád, Csepel sziget | 8979 | | | 4. 5. 1979 |
| III. The foothills of Bakony Mts | Bakonyicum–Veszprémiense | 7. Bakonypölöske, Kupi erdő | 8770 | *Aegopodio–Alnetum quercetosum* | 200 | 12. 4. 1979 |
| | | 8. Pápakovácsi, Gannai erdő | 8771 | | | |
| IV. Velencei Mts | | 9. Nadap, Templom hegy | 8777 | *Aceri tatarico–Quercetum ornetosum* | 240 | 10. 4. 1979 |
| V. Börzsöny Mts | Matricum–Neogradense | 10. Nagyhideghegy, top | 8079 | *Aconito–Fagetum* | 850 | 27. 4. 1979 |
| | | 11. Nagyhideghegy, eastern slope | | | | |
| VI. Bükk Mts | Matricum–Borsodense | 12. Cserépfalu, Hór völgye | 8089 | *Fagetum silvaticae* | 260 | 6. 5. 1979 |
| | | 13. Felsőtárkány, Várhegy | 8088 | *Querco–Fagetum* | 650 | 20. 4. 1979 |
| | | 14. Répáshuta, Kerekhegy | 8089 | *Tilio–Fraxinetum* | 650 | 27. 4. 1979 |
| VII. Zempléni Mts | Matricum–Tokajense | 15. Makkoshotyka, Völgypatak valley | 7695 | *Aegopodio–Alnetum* | 350 | 10. 5. 1979 |

| | | | | | | |
|---|---|---|---|---|---|---|
| | Carpaticum–Cassovicum | 16. Füzérradvány, Vilypuszta | 7595 | *Querco petreae–Carpinetum* | 280 | |
| VIII. Hortobágy Nat. Park, Great Hungarian Plain | Eupannonicum–Crisicum | 17. Újszentmargita, Margitai erdő | 8292 | *Galatello–Quercetum* | 100 | 25. 4. 1979 |
| IX. Western Transdanubia | Praenoricum–Castriferreicum | 18. Ikervár, erdő<br>19. Ikervár, rét | 8767 | *Querco–Ulmetum* | 150 | 12. 4. 1979 |
| X. Tolna County | Praeillyricum–Kaposense | 20. Pálfa, Simontornyai erdő | 9277 | *Helleboro dumetorum–Carpinetum* | 230 | 30. 3. 1979 |
| XI. Mecsek Mts | Praeillyricum–Sopianicum | 21. Égervölgy | 9875 | *Helleboro odoro–Fagetum* | 230 | 10. 4. 1979 |
| | | 22. Pécs, Tubes | | *Tilio argenteae–Fraxinetum* | 500 | 11. 4. 1980 |
| XII. Villányi Mts | Praeillyricum–Villányicum | 23. Szárszomlyó, at Nagyharsány | 9875 | | 300 | 10. 4. 1980 |
| | | 24. Ördögvölgy | | *Asperulo taurinae–Carpinetum* | 200 | |
| XIII. Békés County, Great Hungarian Plain | Eupannonicum–Crisicum | 25. Doboz, Fenyeresi erdő<br>26. Vargahossza erdő | 9293 | *Querco–Ulmetum* | 120<br>100 | 29. 4. 1979 |

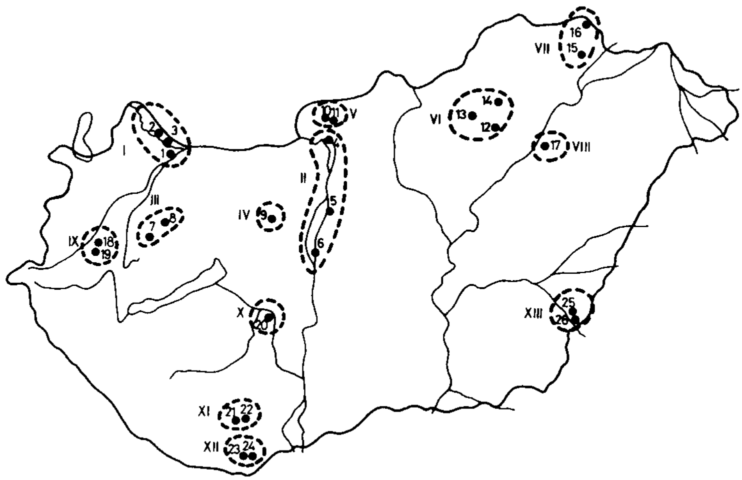


*Fig. 1.* The geographic localization of *Scilla bifolia* populations used in this study. See Table 1 for explanation of numbering

by binary codes. The continuous characters include three dimensionless ratios. The characters and character states, where applicable, are listed below together with the abbreviations used in the figures:

1. L: "leaf ratio" (length/width)
2. I: "inflorescence ratio" (stem length/length of main inflorescence axis)
3. N: number of leaves
4. P: number of pedicels
5. F: "flower ratio" (tepal length/filament length)
6. S: number of seeds per fruit
7. M: length of mature seeds
8. C: chromosome number
9. E: elaisome shape; 0: variable, surface is grained, 1: cylindrical and smooth
10. G: stem colour; 0: green, 1: reddish
11. R: root thickness; 0: thick, 1: thin
12. B: stem bud colour; 0: bluish-grey, 1: green
13. T: tepal colour; 0: white spotted, 1: blue
14. U: premature fruit colour; 0: reddish, 1: green
15. D: dry seed colour; 0: light ochre, 1: dark brown
16. X: date of complete leaf marcescence; 0: before 25 May, 1: after 25 May
17. Y: first day of flowering; 0: before 31 March, 1: after 31 March

The measurements were summarized in a 260×17 raw data matrix. The means for each character formed another input matrix to analyze relationships among populations coming from thirteen geographic regions of Hungary.

Two multivariate techniques were used to reveal group structure of populations which can be contrasted with the taxonomic considerations. First, an R-type principal component analysis (Seal 1964) was performed using program PRINCOMP (Podani 1984). The resulting ordination depicts potential groupings and/or continuous variation in the data. The component correlations obtained indicate which characters are associated with each axe of variation. The correlation matrix was scrutinized to find positively and negatively characters for future use in taxonomic descriptions. Secondly, the set of populations from thirteen regions (see Table 1) was subjected to sum of squares agglomeration clustering (cf. Orlóci 1978) based on normalized character means. The cluster analysis was performed by program NCLAS from the SYN-TAX package (PODANI 1980, 1984). The resulting classification was then compared to the groups depicted by the ordination. All computations were done on the IBM 3031 computer of the Hungarian Academy of Sciences.

## Results

### *Character correlations*

The entire correlation matrix of characters is not presented. Instead, correlation coefficients exceeding the arbitrary level of 0.6 are listed in Table 2. It is seen that three variables, namely chromosome number (C), elaisome shape (E), and the date of complete leaf marcescence (X), occur most frequently in the table. Chromosome number is strongly correlated with both phenological characters and the colour of stem buds, tepals and dry seeds. Also, as the largest negative coefficient indicates, large chromosome number is probably associated with relatively short mature seeds. The observation that the coefficients between E and other five binary variables are negative is simply the matter of coding.

**Table 2**

*Correlation coefficients exceeding the arbitrary level of* |0.6|
*See text for character symbols*

| Characters | Correlation | Characters | Correlation |
|---|---|---|---|
| N — C | 0.976 | M — C | —0.974 |
| N — P | 0.95 | E — B | —0.857 |
| X — B | 0.854 | E — U | —0.857 |
| X — T | 0.854 | E — T | —0.857 |
| X — D | 0.854 | E — D | —0.857 |
| Y — D | 0.854 | E — X | —0.731 |
| Y — B | 0.854 | E — Y | —0.731 |
| X — C | 0.843 | G — P | —0.713 |
| Y — C | 0.843 | G — L | —0.694 |
| L — M | 0.831 | S — F | —0.657 |
| S — B | 0.742 | G — N | —0.642 |
| D — S | 0.742 | E — C | —0.617 |
| C — B | 0.72 | L — F | —0.61 |
| C — T | 0.72 | | |
| C — D | 0.72 | | |
| U — B | 0.69 | | |
| U — T | 0.69 | | |
| U — D | 0.69 | | |
| Y — X | 0.675 | | |
| C — N | 0.675 | | |
| S — X | 0.653 | | |
| M — X | 0.647 | | |
| S — R | 0.614 | | |

Leaf ratio and seed length are positively correlated. The results suggest that increase in the number of leaves (i.e., three leaves instead of two) implies an increase in the number of pedicels. These changes are apparent consequences of an increased chromosome number (36 or 54, instead of the diploid 2n = 18).

## *Component correlations*

The first three components extracted from the correlation matrix explain a total of 70% of the variance. The first component accounts for 48%, which is an unusually high percentage. The second and third components explain 13% and 8%, respectively. The coefficients for the first three eigenvectors associated with the first three eigenvalues are illustrated in Figs 2–3. As seen, the following characters had strong influences on the components: on PCl, three quantitative characters (chromosome number, number of seeds per fruit, and the length of mature seeds) and six binary characters (elaisome shape, dry seed colour, stem bud colour, tepal colour, first day of flowering, date of complete leaf marcescence, and premature fruit colour); on PC2, number of leaves and the flower ratio; and on PC3, stem colour. That is, there is no character highly correlated with more than one of the first three components. Leaf ratio, inflorescence ratio, the number of pedicels, the length of mature seeds and root thickness are not strongly associated with any of these components, although these characters have some outstandingly high correlations with others (cf. Table 2).

## *Ordination and classification*

The ordination of sample individuals is shown in Figs 4–5. For clarity, the positions of OTUs are not indicated in the scattergram. Instead, each group of individuals collected within the same region is outlined in the ordination space. Note that most of these groups contain more than one population (cf. Table 1).

The bipolar first component makes a clear distinction between two groups. Large negative scores are assigned to populations coming from a variety of regions: the upper and middle floodplains of the Danube (I–II), the foothills of Bakony Mountains (III), Mecsek Mountains (XI), Villányi Mountains (XII) and Békés county, Great Hungarian Plain (XIII). These populations do not overlap completely, there is a slight separation between I and XI, and all others if the components are simultaneously examined.

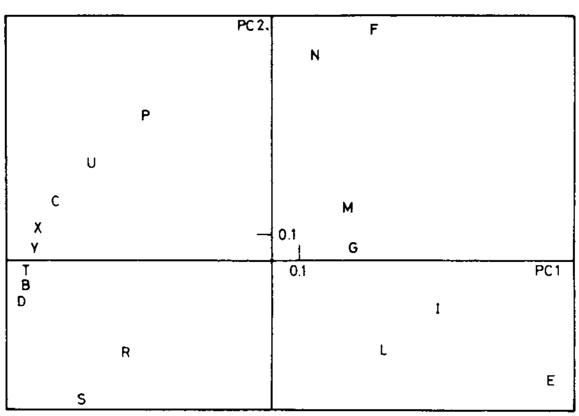


*Fig. 2.* The correlations of characters with PC1 and PC2. See text for explanations of symbols

The group on the positive end of PC1 includes populations from hilly and mountainous areas, with one exception (Hortobágy National Park). The second and third components reveal the fine structure of this group. The populations from Zemplén (VII) are separated from the others on PC2, and those from western Transdanubia (IX) on PC3. On the fourth component the populations overlap one another, therefore it is not illustrated.

The classification of populations (Fig. 6) is in a very good agreement with the ordination. At two group level, the classification equals to the partition revealed on PC1. Within

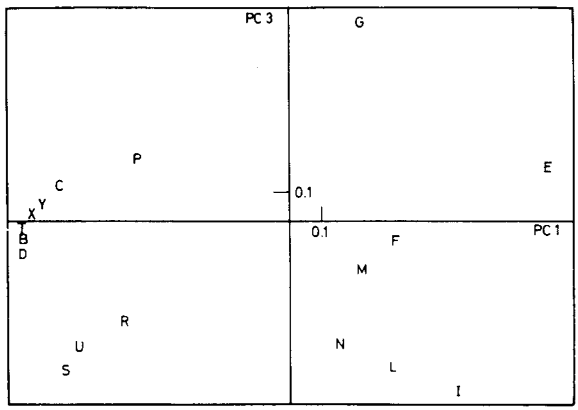


*Fig. 3.* The correlations of characters with PC1 and PC3. See text for explanations of symbols

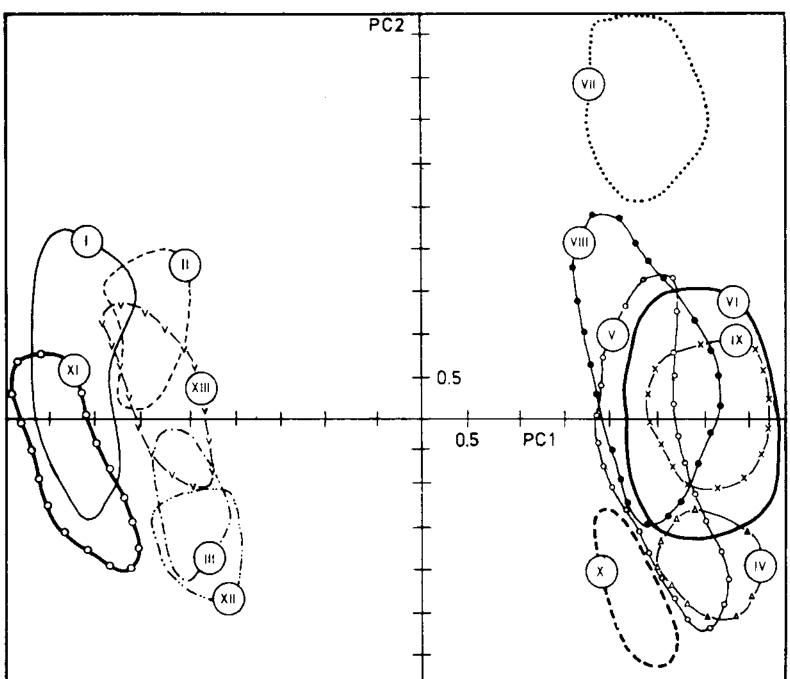


*Fig. 4.* Principal components ordination of populations on the first two components. See Table 1 for numbers

western Transdanubia are separated from the others. The close relationship between I and IX is also seen in the dendrogram. However, the hierarchy does not suggest their segregation from II, III, XII and XIII, contrary to the PCA results.
the larger group first the populations from the Zemplén Mountains, and then those from

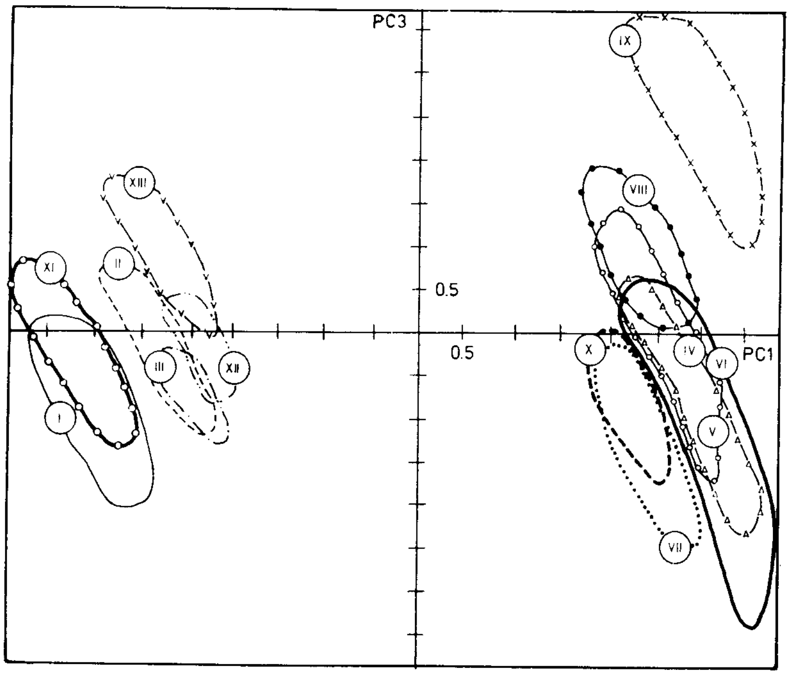


*Fig. 5.* Principal components ordination of populations on PC1 and PC3. See Table 1 for numbers

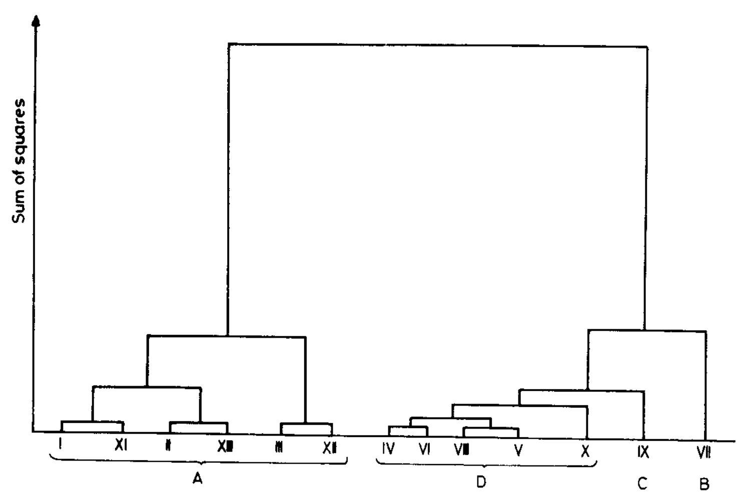


*Fig. 6.* Phenogram resulting from the sum of squares agglomeration clustering of populations coming from thirteen regions of Hungary. See Table 1 for numbers

## Discussion

The numerical analysis of *Scilla bifolia* sensu lato populations collected in Hungary proved the existence of distinct forms. The results of principal component analysis and cluster analysis suggest that at least three, possibly four, taxa can be distinguished within the aggregate. These are designated by A–D in the sequel (see also Fig. 6). The karyotype, most binary and some quantitative characters and both phenological characters appear of fundamental importance in determining group structure.

Group A contains the populations from the upper and middle floodplain of the Danube, from Mecsek, the Villányi Mountains, the foothills of Bakony, and from Békés county. All individuals in this group are diploid (2n = 18). Character states typical of these populations are: elaisomes smooth and cylindrical (sausage-shaped), tepals white spotted, dry seeds light ochre, stem buds bluish grey, premature fruits reddish. First day of flowering before 31 March, early leaf marcescence. The mature seeds are approximately 1.5 times longer than in groups B–D. The number of seeds per fruit is 16 on the average. Apparently, this group corresponds to *S. vindobonensis* described by SPETA (1974).

Groups B–D may be clearly distinguished from group A. The main differences are: elaisomes grained, mature seeds dark coloured, tepals always blue. Premature fruits usually green, later turning into brown. The stem buds are also green. Late leaf marcescence and flowering.

The two populations collected in Zemplén constitute group B. Although the chromosome number is 18, just like in group A, the morphological differences are clear-cut, as described above. This group is separated from groups C and D on the basis of a constant leaf number (2), relatively short filaments if compared to the tepals, and of the consistently low seed number (6–7) per fruit. This form seems to be identical with *S. kladnii* Schur, which has originally been described on the basis of the reduced number of seeds.

The plants in group C, collected in western Transdanubia (IX), are tetraploid (4n = 36). The morphological characteristics in which this group differs from the others are: stems green and the leaf ratio is low, i.e., the leaves are relatively short and wide. It is doubtful whether this form can be identified as *S. drunensis* Speta (cf. SPETA 1977).

The karyotype of group D, which includes the populations from the Börzsöny and Bükk Mts, Tolna County and the Hortobágy National Park, is also tetraploid. The only exception is the population collected in the Velencei Mts, which is hexaploid. Interestingly, this difference is not manifested in macromorphology. The hexaploid population cannot be distinguished from most tetraploids on the basis of characters included in this study. The number of leaves in this group is commonly three. Numerous (over 20) but small seeds are found in each fruit. This group is likely to correspond to *S. buekkensis* Speta (1977), formerly listed as *S. subtriphylla* Schur.

In summary, the following conclusions can be drawn from the present study:

1. The striking difference between the diploid and polyploid populations of the *S. bifolia* aggregate suggest their taxonomic disjunction at the species level. Moreover, two diploid species may be found in Hungary, one of them (which is probably *S. kladnii*) being more similar to the tetraploids than to the other diploids.

2. The diploid group A (*S. vindobonensis*) can not be unambiguously subdivided.

3. It seems clear from the results that the polyploid populations in Hungary do not separate from each other at species level. Nevertheless, the taxonomic status of the populations from western Transdanubia remains problematic. Further studies are needed to elucidate its relationship to *S. buekkensis* and *S. drunensis*. On the other hand, the hexaploid population exhibits no demonstrable differences from most tetraploids. Consequently, chromosome number itself is not a reliable basis for creating taxonomic categories within this aggregate.

4. Altitude has no practical influence on the distribution of groups A and D. Both types occur in lowland and mountain habitats.

5. Principal component analysis proved to be a robust technique in indicating group structure. The obvious non-linearity of many characters had no apparent influence on the results, as far as the groups recognized by cluster analysis are concerned.

## REFERENCES

Kereszty, Z. (1983): A hazai Scilla bifolia alakkör taxonómiai felülvizsgálata (Taxonomical re-examination of a Scilla bifolia aggregate in Hungary I. raising the problems, plan of research). *Bot. Közlem.* **70:** 55–60.

Kereszty, Z., Szilágyi, L. (1984): Cytological investigation of Scilla bifolia populations in Hungary I. *Acta Bot. Hung.* **30:** 53–66.

Orlóci, L. (1978): *Multivariate Analysis in Vegetation Research.* 2nd ed. Junk, The Hague.

Pócs, T. (1968): Növényföldrajz és ökológia (Plant Geography and Ecology). In: T. Hortobágyi (ed.): *Növényrendszertan és növényföldrajz* (Systematic Botany and Plant Geography). Tankönyvkiadó, Budapest.

Podani, J. (1980): SYN-TAX: Computer program package for cluster analysis in ecology, phytosociology and taxonomy. *Abstracta Bot.* **6:** 1–158.

Podani, J. (1984): SYN-TAX II: Computer programs for data analysis in ecology and systematics. *Abstracta Bot.* **8:** 73–94.

Seal, H. L. (1964): *Multivariate Statistical Analysis for Biologists.* Methuen, London.

Soó, R. (1973): *A Magyar Flóra és Vegetáció Rendszertani-Növényföldrajzi Kézikönyve V.* Synopsis systematico-geobotanica florae vegetationisque Hungariae V. Akadémiai Kiadó, Budapest; 81–82.

Speta, F. (1974): Cytotaxonomische und arealkundliche Untersuchungen an der Scilla bifolia-Gruppe in Oberösterreich, Niederösterreich und Wien. *Naturkundl. Jahrb. Stadt Linz* **19:** 9–54.

Speta, F. (1977): Neue Scilla-Arten aus dem östlichen Mittelmeerraum. *Naturkundl. Jahrb. Stadt Linz* **22:** 65–72.