

類集分析과 主成分分析에 의한 韓國產 메꽃屬의 數量分類學的 研究

金 潤 植 · 崔 秉 熙
(高麗大學校 理科大學 生物學科)

A Numerical Taxonomic Study of *Calystegia* in Korea by the Cluster Analysis and Principal Component Analysis

Kim, Yun Shik and Byeung Heui Choi
(Department of Biology, Korea University, Seoul)

ABSTRACT

The relationships and character variations on 5 taxa of *Calystegia* were examined by cluster analysis and principal component analysis. Thirteen *Calystegia* population samples from the middle part of Korea were observed. Although minor differences were noted, essentially similar results were obtained from the phenograms by UPGMA, UPGMC and Ward's clustering methods, and these results were in accordance with those obtained from the ordination plots by principal component analysis.

C. soldanella is distantly connected with the other taxa mainly because of its morphologically different leaf organs. Based on the difference on the first principal component, *C. hederacae* is kept apart from the rest 3 taxa. In the relationships among *C. japonica*, *C. sepium* var. *americana* and *C. davurica*, minor differences were obtained from the 3 clustering methods. As to the character variations among different populations within a taxon, they are slight in *C. soldanella* and *C. sepium* var. *americana*, but remarkable in *C. hederacae* and *C. davurica*.

緒 論

메꽃屬 (*Calystegia*)은 南·北半球의 溫帶에서 亞熱帶 까지 分布하며 形態의으로 多形性인 것으로 알려져 있는데 (Cockayne, 1928), 韓國產에 관하여서는 Palibin (1901), Komarov (1904), Nakai (1911, 1952), 그리고 Park (1949) 등에 의하여 보고 된 바 있고, 본인들 (Kim and Choi, 1983)은 染色體 數와 外部形態 및 解剖學的 形質들을 조사하여 種 檢색표를 작성한 바 있다.

分類學에서 多變量解析法을 이용하는 목적은 객관적이고 반복적인 방법에 의하여 分類群間的 類緣關係 (relationship)나 類似度 (similarity)를 밝히려 함인데 (Radford *et al.*, 1974),

多變量解析法 중 類集分析 (cluster analysis)과 主成分分析 (principal component analysis) 등은 分類學에 자주 이용되고 있다 (Rushton, 1978). 특히 主成分分析法은 現象속에 內在해 있는 規則性 (regularity)과 秩序 (order)를 찾아 내려는 要因分析 (Rummel, 1970)에 의한 방법으로서, 類集分析에 의한 分類群間의 關係를 조사한 다음 主成分分析으로 이들간의 變異에 대한 일반적 형태를 조사하는 것이 바람직한 것으로 알려져 있다 (Sneath and Sokal, 1973).

메꽃屬에 관하여서는 Ogden(1978)이 主成分分析法을 이용하여 뉴우지일랜드産 메꽃屬의 變異에 관하여 조사 보고한 바 있다.

本 研究에서는 韓國産 메꽃屬의 다섯 分類群, 즉 *C. soldanella* Roem. et Schult. (갯메꽃), *C. hederacae* Wall. (애기메꽃), *C. japonica* (Thunb.) Chois. (메꽃), *C. sepium* var. *americana* Matsuda (큰메꽃) 및 *C. davurica* (Sims) Chois. (신메꽃)의 여러 形質을 대상으로 類集分析과 主成分分析을 이용하여 分類群間의 類緣關係와 形質變異에 대하여 考察하고 本 屬에 관한 다른 조사 결과 (Kim and Choi, 1983)와 비교 검토해 보고자 한다.

材料 및 方法

본 연구에 사용된 材料는 1983年 5月 부터 10月 사이에 中部地方의 13개 集團 (Table 1)

Table 1. *Calystegia* species and their localities observed

Scientific name	Localities
<i>Calystegia soldanella</i>	Yongjong-do (S-1), Taichon (S-2)
<i>C. hederacae</i>	Paju-gun (H-1), Suwon (H-2), Iksan-gun (H-3), Taichon (H-4)
<i>C. japonica</i>	Iksan-gun (J-1), Suwon (J-2), Seoul (J-3)
<i>C. sepium</i> var. <i>americana</i>	Suwon (A-1), Suwon (A-2)
<i>C. davurica</i>	Kangwha-do (D-1), Yongjong-do (D-2)

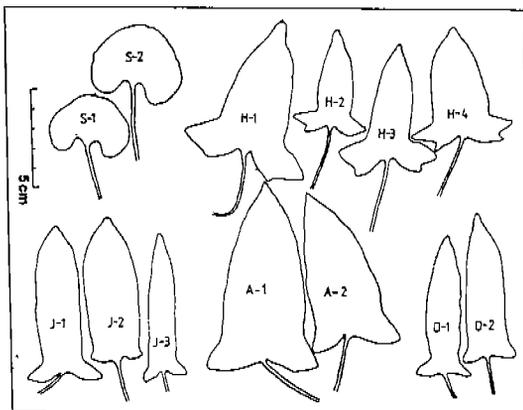


Fig. 1. Representative leaf shape in 13 *Calystegia* populations. The numbers on the leaves refer to the localities of Table 1.

을 조사하였고 各 集團을 하나의 OTU (Operational Taxonomic Unit)로 취급하였으며, 各 分類群의 同定은 본인들 (Kim and Choi, 1983)이 작성한 검색표에 의하였다.

한편 各 分類群의 集團에서 채집된 잎의 모양을 Fig. 1에 도시하였는데, 잎은 集團의 중심지역에서 자라는 개체들로부터 줄기의 중간 부위에 있는 것을 택하였다.

調査된 形質은 모두 38개로서 各 OTU당 20개체 이상 조사하여 산술평균과 표준편차를 구하였으며, 이들 중

Table 2. Morphological characters redefined as ratios of *Calystegia*

V 1 lamina length/lamina width	V 2 lamina largest width/lamina middle width
V 3 petiole length/lamina length	V 4 petiole length/peduncle length
V 5 bracteole length/bracteole width	V 6 bracteole length/calyx length
V 7 calyx length/calyx width	V 8 bracteole width/calyx width
V 9 bracteole length/corolla length	V10 corolla length/corolla width
V11 pistil length/stamen length	V12 pistil length/corolla length
V13 anther length	V14 stigma length
V15 seed length/seed width	

환경에 의한 영향이 적고 分類群間에 有意한 차이가 있다고 생각되는 것 중 (Radford *et al.*, 1974)에서 간격척도로써 19개 形質이 채택되었다. 이 19개의 形質은 다시 形質상호간의 비에 의해 15개의 변수 (Table 2)로 변환되었는데 이는 形質間的 비에 의한 방법을 사용하므로써, 측정된 形質을 직접 사용할 경우 재료의 선택방법에 의해 나타날 수 있는 形質의 크기 효과 (the effect of organism size)를 줄이고 形質의 모양을 반영할 수 있는 방법 (Corruccini, 1975)으로 생각되었기 때문이다. 한편 13개 OTU에 대한 15개의 변수에 의해서 13×15 원 자료행렬 (raw matrix)이 얻어졌다.

資料分析을 위한 계산은 우선 원자료를 표준화 (standardization) 시켰는데 (Moss, 1968; Sneath and Sokal, 1973; Radford *et al.*, 1974; Rushton, 1978), 그 방법은 각 변수에서 평균값을 빼고 다시 그 값을 표준편차로 나누어, 각 변수를 平均 0 分散 1로 만들었다.

類集分析은 OTU들간의 Euclidian distance에 의한 unweighted pair-group method using arithmetic averages (UPGMA), unweighted pair-group centroid method (UPGMC) 및 Ward's error sum of squares (Ward, 1963)의 방법으로 phenogram을 작성하였다.

主成分分析에서는 固有值 (eigenvalue) 1 이상의 3개의 主成分을 택하여 ordination plot을 작성하였고, 각 OTU의 主成分들에 대한 要因評點 (factor score)은 OTU의 표준화된 形質 값과 主成分의 固有벡터 (eigenvector)를 곱한 수치의 합으로 계산되었다.

本 研究에 이용된 統計分析법은 Jardine and Sibson (1971), Sneath and Sokal (1973), Kim (1976), Gnanadesikan (1977) 등의 이론에 따랐으며 統計分析은 Statistical Analysis System (SAS Institute, 1982) Program을 사용하여 高麗大學校에 설치되어 있는 IBM 4341 computer에 의해 수행되었다.

結 果

Cluster analysis. 形質들의 비에 의한 15개 변수를 대상으로 13개 OTU들에 대한 類集分析을 한 결과 세 가지 方法, 즉 UPGMA, UPGMC 및 Ward's method간에는 대체로 유사한 類集形態를 보였는데 (Fig. 2), UPGMC에서 선메꽃 集團인 D-1이 같은 分類群인 D-2보다 메꽃 集團인 J群과 먼저 연결된 것을 제외하면, 세 가지 類集法 모두에서 같은 種內的 集團들끼리 먼저 연결되어 이 方法이 分類群間的 식별에 유용함을 보여주었다.

큰메꽃 集團인 A-1, A-2는 세 方法 모두에서 제일 먼저 연결되어 두 집단 사이에 변이

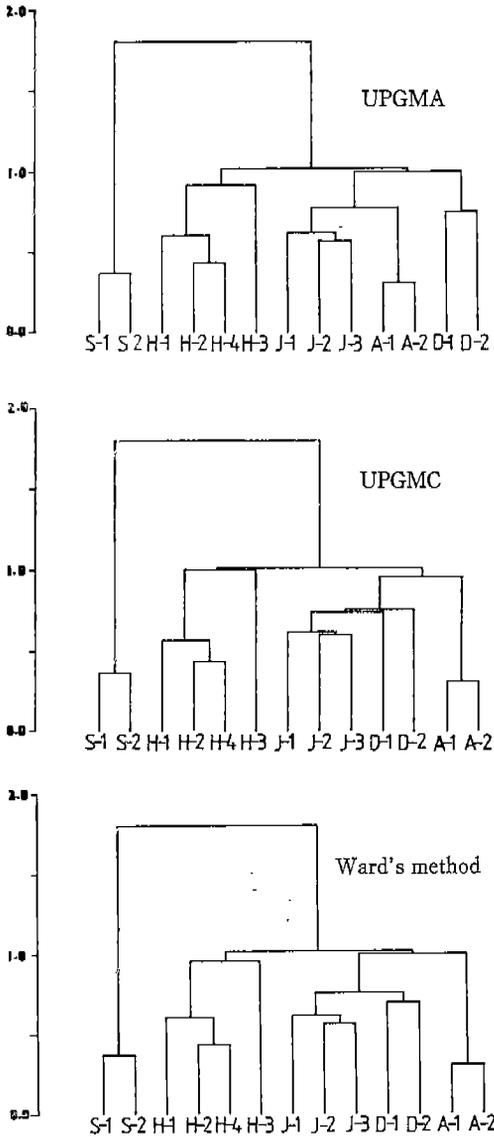


Fig. 2. Phenograms of the 13 *Calystegia* populations using 3 clustering methods-UPGMA, UPGMC, Ward's method.

이었는데 第Ⅰ主成分이 36%, 第Ⅱ主成分이 21%, 第Ⅲ主成分이 17%이었다 (Table 3).

Table 3에 rotated factor matrix를 보면 많은 변수들이 대체로 上位 主成分에 높은 因子 負荷量 (factor loading)을 나타내고 있는데, 主成分에 대한 변수들의 loading 값에 의한 ordination plot (Fig. 3)을 작성하고 그들을 主成分別로 grouping하였다.

主成分들에 대한 OTU의 要因評點 (factor score)으로 부터 작성된 ordination plot (Fig. 4)을 보면 갯메꽃은 第Ⅱ主成分에 의해서, 큰메꽃은 第Ⅲ主成分에 의해서 다른 分類群들과

가 적은 것으로 나타났는데 이 두 集團은 모두 水原에서 조사된 것이었다. S-1과 S-2는 갯메꽃 集團인데 이들도 세 가지 類集法 모두에서 낮은 수준으로 연결되어 두 集團 사이에 變異가 적은 것으로 나타났으며, 선메꽃의 두 集團인 D-1과 D-2는 대체로 높은 수준에서 연결되어 두 集團間에 變異가 심한 것으로 나타났다.

또한 세 가지 類集法에서 나타난 分類群間의 거리를 보면 갯메꽃 (S-1, S-2)은 屬內的 다른 分類群들과는 가장 높은 수준 ($d=1.80$)에서 연결되어 類緣關係가 가장 먼 種으로 해석되고, 애기메꽃 (H-1, H-2, H-3, H-4)도 나머지 세 分類群 즉 메꽃, 큰메꽃, 선메꽃들과 높은 수준에서 연결되어 分類거리가 먼 것으로 해석된다. 한편 애기메꽃 集團중에서 H-3는 다른 OTU들과 높은 수준에서 연결되어 變異가 심한 集團임을 나타냈는데, 다른 애기메꽃 集團들에 비해 花梗과 萼의 길이가 긴 것으로 全北 益山郡에서 조사된 것이었다.

메꽃 (J-1, J-2, J-3), 큰메꽃 (A-1, A-2) 및 선메꽃 (D-1, D-2)間에는, UPGMC와 Ward's method에서 메꽃과 선메꽃이 먼저 연결되었으나 UPGMA에 의하면 메꽃과 큰메꽃이 먼저 연결되어 類集法間에 약간의 차이가 있었다.

Principal component analysis. 15개 변수들을 대상으로 하여 varimax회전法에 의해서 회전한 결과 固有值 1 이상의 3개의 主成分 (principal component)이 발견되었고, 이들의 변수전체분산에 대한 설명도는 74%

Table 3. Rotated factor matrix obtained from principal component analysis

Characters	Principal component I	Principal component II	Principal component III	Communality
V 1	0.44	0.68	-0.37	0.79
V 2	0.18	-0.79	-0.26	0.73
V 3	-0.90	-0.20	0.13	0.86
V 4	-0.70	-0.28	0.32	0.67
V 5	0.25	0.56	0.28	0.46
V 6	0.85	-0.18	0.29	0.84
V 7	0.26	0.74	0.27	0.69
V 8	0.82	-0.04	0.31	0.76
V 9	0.82	0.26	-0.04	0.73
V10	0.18	0.03	0.87	0.79
V11	-0.03	-0.43	-0.75	0.75
V12	-0.05	0.05	0.91	0.84
V13	-0.13	0.79	-0.12	0.65
V14	-0.80	-0.16	-0.10	0.67
V15	0.89	-0.38	0.04	0.93
Eigenvalue	5.41	3.22	2.53	
Contribution	0.36	0.21	0.17	
Cumulative contribution	0.36	0.57	0.74	

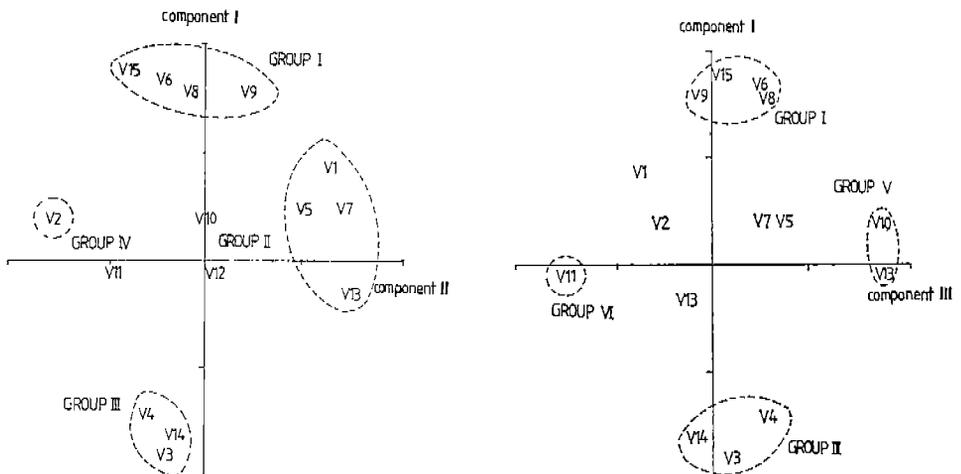


Fig. 3. Grouping of 15 characters by its factor loadings on principal component I & II and principal component I & III.

구별됨을 알 수 있으며, 또한 이 두 分類群은 세 主成分 모두에서 變異가 적게 나타남을 알 수 있었다. 한편, 애기베꽃 集團들은 第 I 主成分에 의해서 다른 分類群들과 구별 되는데, 애기베꽃 集團 內에서도 H-3 集團은 第 I 主成分에 의해서 種內의 다른 集團들과 구별

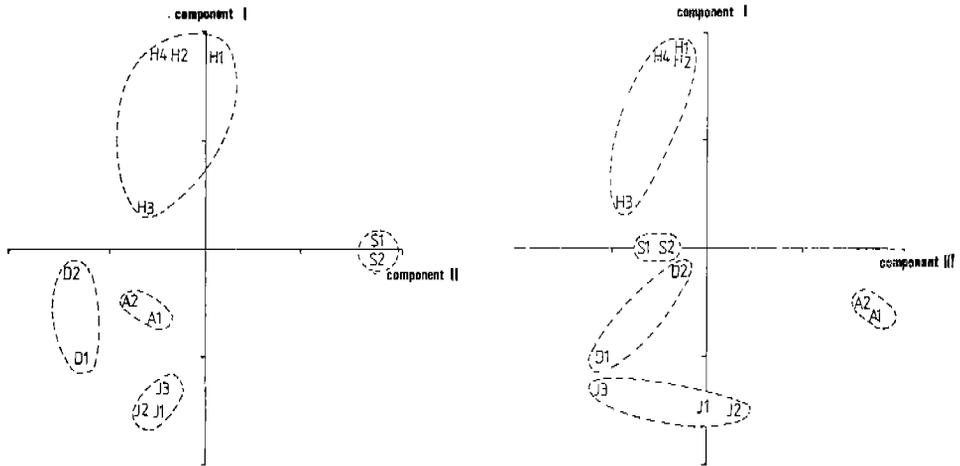


Fig. 4. Principal components analysis of the 13 *Calystegia* populations. Each population is represented by a point defined by its loading on principal component I & II and principal component I & III.

되는 것으로 보아 이는 꽃과 花梗길이의 변이임을 알 수 있었다.

메꽃의 세 集團인 J-1, J-2, J-3에 있어서 第Ⅰ主成分과 第Ⅱ主成分에 대한 loading 값들은 일정하였으나, 第Ⅲ主成分 즉 形質 group V와 VI을 보면 지역에 따라 變異가 있음을 알 수 있다.

선메꽃의 두 集團인 D-1과 D-2는 第Ⅱ主成分에 대한 값들은 비슷하였으나, 第Ⅰ과 第Ⅲ主成分에 대한 값들에 변이가 있으므로 形質 group I, III, V, VI 등이 지역에 따라 變異가 있음을 알 수 있다.

考 察

UPGMA, UPGMC 및 Ward's method에 의한 類集分析과 varimax회전법에 의한 主成分分析을 이용하여 메꽃屬의 分類거리와 形質變異 등에 관해서 조사하여 본 결과 두 多變量解析法에 의한 결과가 잘 일치되고 있으며, 集團 사이의 形質變異 등도 잘 나타내고 있다. 이는 Rushton (1978)의 *Quercus*屬에 대한 研究에서도 Ward's method에 의한 類集分析과 主成分分析의 결과가 서로 잘 일치함을 주장한 바 있어 이 두 방법은 變異集團 사이의 分類에 매우 유용하다고 생각된다.

固有值 1 이상의 세 主成分 (74%)들의 성격은 Table 4에서 보는 바와 같은데, Ogden (1978)은 뉴우지일랜드産 메꽃屬의 조사에서 主成分들 (84.5%)이 꽃, 花冠, 葉柄 및 잎의 길이 등과 相關이 깊다고 보고한 바 있으나, Ogden이 조사에 이용한 形質과 본 조사에 이용된 形質이 서로 다르기 때문에 두 조사 결과를 비교 분석하기에는 무리가 있다고 생각된다.

類集分析의 결과 나타난 phenogram (Fig. 2)에서 보면, 갯메꽃 集團 (S-1, S-2)은 다른

Table 4. Relationships between principal components and characters

Principal component	Correlation	Characters group	Corresponding characters
component I	+	Group I	bracteole length/calyx length (V6), bracteole width/calyx width (V8), bracteole length/corolla length (V9), seed length/seed width (V15)
	-	Group III	petiole length/lamina length (V3), petiole length/peduncle length (V4), stigma length (V14)
component II	+	Group II	lamina length/lamina width (V1), bracteole length/bracteole width (V5), calyx length/calyx width (V7), anther length (V13)
	-	Group IV	lamina largest width/lamina middle width (V2)
component III	+	Group V	corolla length/corolla width (V10), pistil length/corolla length (V12)
	-	Group VI	pistil length/stamen length (V11)

分類群들과의 分類거리가 가장 멀게 나타났는데, 이 조사 결과는 본인들 (Kim and Choi, 1983)이 서식지와 외부형태 및 해부학적 조사에서 갯메꽃이 屬內的 다른 分類群들과는 遠緣의 關係임을 밝힌 바와도 일치하며, 또한 Fig. 4에서 보면 이는 주로 第II主成分 즉 잎의 形態의 차이에 기인함을 알 수 있었는데 Fig. 1에서 보는 바와 같이 다른 分類群들과는 잎의 모양이 뚜렷이 구별되고 있다.

애기메꽃은 세 方法 모두에서 나머지 3分類群, 즉 메꽃, 큰메꽃, 선메꽃들과 分類거리가 멀게 나타났는데 이는 본인들 (Kim and Choi, 1983)이 줄기의 生長형태와 털의 有無 등에 의해 선메꽃을 나머지 세 分類群 즉 애기메꽃, 메꽃, 큰메꽃들로 부터 구분한 것과는 일치하지 않았으나 Fig. 4에서 보면 애기메꽃이 다른 分類群들과 구별되는 것은 第I主成分의 차이에 의한 것임을 알 수 있었다.

메꽃, 큰메꽃 및 선메꽃간에는 類集法에 따라 약간의 차이를 보이는데 이는 分類거리를 類集化시키는 方法간의 차이에서 오는 결과로 생각되며, Ward's method와 UPGMC에 의한 類集分析에서 메꽃과 선메꽃이 먼저 연결된 것은 이들이 줄기의 生長형태와 털의 유무를 제외하면 외부형태적으로 유사하다는 것을 나타내 주는 것이었다.

이상에서 보는 바와 같이 메꽃屬에 관해 類集分析을 이용하여 分類群間的 關係(거리)와 變異정도를 알아보고 主成分分析에 의한 集團間的 關係와 變異의 일반적 形態를 알아본 결과 두 방법이 서로 잘 일치하고 있음을 알 수 있었다.

分類群間的 關係(거리)에 대한 본 조사 결과는 메꽃屬에 대한 外部形態 및 解剖學的 조사에서 나타난 결과와 일치하는 면과 일치하지 않는 면이 있는데, 이는 外部形態의 量的 形質에 의한 數量分類方法에는 形質들의 환경에 대한 변화에 문제점이 있을 수 있고 또한 分類거리에 대한 類集方法間에도 차이가 있을 수 있기 때문으로 생각된다.

多變量解析法에 의한 分類群間的 거리가 반드시 그들의 類緣關係와 일치한다고는 보기 어

려우나 환경에 영향을 덜 받고 分類學的으로 의의가 있는 形質들을 반복적인 方法으로 조사함으로써 소수 形質을 대상으로 하였을 때 나타날 수 있는 오류를 보완 할 수 있을 것으로 思料된다.

摘 要

韓國 中部地方의 13개 集團에서 관찰된 메꽃屬 (*Calystegia*)의 다섯 分類群 즉 갯메꽃, 애기메꽃, 메꽃, 큰메꽃 및 선메꽃에 대하여 類集分析和 主成分分析을 이용해서 分類群間的 類緣關係와 形質變異에 관하여 조사하여 보았다.

UPGMA, UPGMC 및 Ward's method에 의한 類集分析 方法間에 類集形態가 유사하게 나타났으며, 主成分分析의 결과와도 잘 일치 하였는데 그 결과는 다음과 같다.

갯메꽃은 屬內的 다른 分類群들과의 分類거리가 가장 멀게 나타났고 이는 주로 잎의 형태적 차이에 의한 것이었다. 애기메꽃도 나머지 3分類群과 거리가 멀게 나타났는데, 이는 주로 第 I 主成分의 차이에 의한 것이었다. 메꽃, 큰메꽃 및 선메꽃간의 關係는 類集法에 따라 약간의 차이가 있었다. 集團 사이에 形質變異는 갯메꽃과 큰메꽃에서는 變異가 적었으며, 애기메꽃과 선메꽃에서는 變異가 심하였다.

메꽃屬에 대한 數量的 分類方法이 分類群間的 類緣關係와 形質變異를 조사 하는데 있어서 매우 효과적 이라고 생각된다.

參 考 文 獻

- Cockayne, L. 1928. The Vegetation of New Zealand. Reprinted 1958. Engelman, Weinheim. 456 pp.
- Corruccini, R. S. 1975. Multivariate analysis in biological anthropology: Some considerations. *J. Human Evol.* 4: 1~19.
- Gnanadesikan, R. 1977. Methods for Statistical Data Analysis of Multivariate Observations. John Wiley & Sons, New York.
- Jardine, N. and R. Sibson. 1971. Mathematical Taxonomy. John Wiley & Sons, New York.
- Kim, K. W. 1976. Methodology for Social Science. Pagyong, Seoul. pp. 456~471.
- Kim, Y. S. and B. H. Choi. 1983. Chromosome number, morphological and anatomical study on *Calystegia* in Korea. *Kor. Jour. Pl. Tax.* 13: 89~101.
- Komarov, V. L. 1904. Flora Manshuriae II.
- Moss, W. W. 1968. Experiments with various techniques of numerical taxonomy. *Syst. Zool.* 17: 31~47.
- Nakai, T. 1911. Flora Koreana II. *Jour. Coll. Sci. Univ. Tokyo* 31: 108~110.
- _____. 1952. A synoptical sketch of Korean flora. *Bull. Not. Sci. Mus.* 31: 94~95.
- Ogden, J. 1978. Variation in *Calystegia* R. Br. (Convolvulaceae) in New Zealand. *New Zealand J. Bot.* 16: 123~140.
- Palibin, J. W. 1901. Conspectus florum Koreanae. *Act. Hort. Petrop.* 14: 163~164.
- Park, M. K. 1949. An Enumeration of Korean Plants. Seoul, Korea. 197 pp.
- Radford, A. E., W. C. Dickison, J. R. Massey and C. R. Bell. 1974. Vascular Plant Systematics. Harper & Row, New York. pp. 485~500.
- Rummel, R. J. 1970. Applied Factor Analysis. Northwestern Univ. Press.

- Rushton, B. S. 1978. *Quercus robur* L. and *Quercus petraea* (Matt.) Liebl.: a multivariate approach to the hybrid problem, 1. Data acquisition, analysis and interpretation. *Watsonia* 12 : 81~101.
- SAS Institute. 1982. SAS user's guide, Statistics. SAS Institute Statistical Analysis System. pp. 309~431.
- Sneath, P. H. A. and R. R. Sokal. 1973. Numerical Taxonomy. Freeman, San Francisco.
- Ward, J. H. 1963. Hierarchical grouping to optimize an objective function. *J. Am. Statist. Ass.* 58 : 236~244.

(1984. 1. 30 接受)