

數値分析에 의한 참나무 基本種의 類緣關係*¹馬 相 圭*²Similarity Relationship between Basic Species of the
Oak by the Numerical Method*¹Sang-Kyu Ma*²

In order to prove the similarity relationships between the basic species of oak through Electronic Data Processing System(EDPS) and numerical analysis, The analyzed species and datas were selected from the list of morphological observation in the thesis of T.B. Lee, 1961, "Phylogenetic study of the subgenus *Lepidobalanus* of the genus *Quercus* in Korea," and were coded by categories shown in Table 1. The value in the list were transformed into hundred percentage to standardize the observational value by each code into dimensionless.

The similarity index between species were computed through formula of non-metric coefficient, $N_{jk} = \sum_{i=1}^n \left(\frac{|x_{ij} - x_{ik}|}{x_{ij} + x_{ik}} \right)$, using the UNIVAC-1106, at National Computer Center.

Quercus aliena, by analysis result, is most similar to *Q. mongolica* with the similarity index, 71.6 and *Q. dentata* is most far apart from *Q. serrata* in the relationship with index, 121.4. The above thesis of Professor, T. Lee, are closely similar with the result of this research study. But, their similar relationship are proved in quantity through numerical method in our research study. In addition, The relationships among *Q. mongolica*, *Q. aliena* and *Q. serrata* are found to be very similar, but *Q. dentata* to be enough far in similarity to other species (by dendrogram shown at Fig. 1. The numerical classification through EDPS is found to be suitable method also applicable to the plant taxonomy.

電子計算組織과 數値分析에 의하여 참나무 基本種間의 類緣關係를 증명하고자 non-metric coefficient 式을 適用하며 UNIVAC-1106 으로 處理하였다.

研究結果는 李昌福教授의 論文結果와 一致하고 있으나, 筆者의 研究에서는 이들 類緣關係가 數量的으로 再證明된 것이고 樹狀型分類에 의거하면 참나무 基本種中 떡갈나무는 類緣關係가 가장 멀게 나타나고 있음을 밝혔다.

* ¹ Received for publication in September 26, 1973* ² 林業試驗場, 서울, Forest Research Institute, Seoul

資料의 分析過程은 다음과 같다.

(1) 基本調査一覽表를 表1의 code化 內容과 같이 區分하여 punch card에 옮긴다.

(2) 總觀察值가 樹種別로 差異가 있으므로 직접 類緣關係를 分析할 수 없으므로 觀察值를 標準化(dimension less)시킨다. 이 研究에 있어서는 觀察值를 標準化시키기 위하여 컴퓨터에 data를 入力시켜 百分比化시켰다.

(3) 樹種間의 類緣關係分析은 non-metric coefficient인 式①을 使用하였다.

$$N_{ik} = \sum_{j=1}^n \left(\frac{|x_{ij} - x_{ik}|}{x_{ij} + x_{ik}} \right) \dots \dots \dots \textcircled{1}$$

(4) N_{jk} 가 적으면 적은수록 類緣關係가 높다. 計算後 N_{jk} 가 가장 적은 樹種群을 찾아 낸다.

(5) N_{jk} 가 가장 적은 調査을 이룬 2樹種의 標準化된 數值를 平均하여 별도의 名稱을 부여한 後 餘他樹種과의 類緣關係 分析을 同一한 方法으로 계속 시킨다.

(6) 計算이 끝나면 類緣關係가 가장 높은 것 끼리 dendrogram化시켜 그들 간의 類緣關係를 밝힌다.

結 果

떡갈, 신갈, 갈참 및 졸참나무 간의 類緣關係를 分析한 結果는 表2와 같다. 表2에 依하면 신갈과 갈참나무 間의 類似度가 가장 높고, 졸참과 떡갈나무 間의 類緣關係가 가장 멀게 나타 났다.

表 2. 참나무 基本種間의 類似度指數
Table 2. Similarity index between basic species of the oaks

	D	M	A	S
<i>Q. dentata</i> (D)	0	103.692	95.914	121.402
<i>Q. mongolica</i> (M)		0	71.605	84.997
<i>Q. aliena</i> (A)			0	72.920
<i>Q. Serrata</i> (S)				0

既研究 結果에는 졸참과 갈참, 갈참과 신갈, 신갈과 떡갈이 보다 서로 가깝다는 結論을 주고 있으며 다음 그림과 같은 關係가 있는 듯이 보인다 하였다.

上記 그림에서 보이는 숫자는 筆者가 表2의 內容을 보충시킨 것이다.

既 研究의 結果와 數値分析에 依하여 얻어진 結果와는 거의 一致함을 알 수 있고 이들의 類緣關係가 過去와는 달리 數量的으로 나타낼 수 있음을 발견할 수 있을 것이다. 다만 떡갈과 신갈간의 類緣關係보다는 數値分析結果 갈참과 신갈간의 類緣關係가 높다는 사실이

發見된다.

類似度가 가장 높은 갈참과 신갈 group과 餘他樹種과의 類緣關係를 밝히기 위하여 갈참과 신갈에 대한 觀察階級別 出現頻度率을 合計하여 平均한 것을 MA라고 졸참과 떡갈간의 類緣關係를 分析하니 表3과 같다. 表3에 의하면 갈참과 신갈 group과의 類緣關係가 높은 것은 졸참나무이다.

表 3. 참나무 基本種間의 類似度指數
Table 3. Similarity index between basic species

	M+A	D	S
M+A	0	106.323	84.279
D		0	121.402
S			0

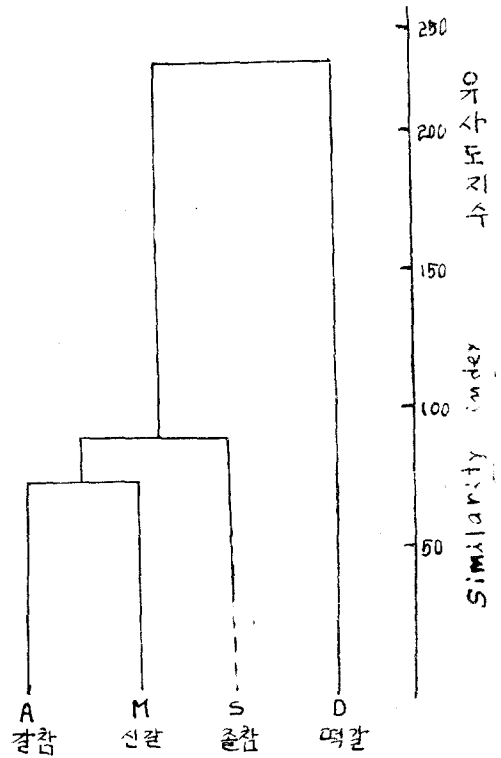


그림1. 한국산 참나무류의 基本圖간의 수상(樹狀)형분류도
Fig 1. Dendrogram between the basic species of the oaks in Korea.

동일한 方法으로 MA 와 S 의의 平均值와 D 의의 類緣關係는 226.79 이었다. 따라서 참나무 基本種間에 딱 갈나무가 가장 멀게 나타나고 있다. 이들의 關係를 dendrogram 化 시키니 그림 1 과 같다.

考 察

참나무 基本種間의 類緣關係를 數値分析한 바 갈참과 신갈間의 類似度 指數는 71.6'으로 가장 가깝고 갈참과 졸참間은 72.9 로 그 다음이며 졸참과 신갈은 85.0 으로 나타나고 있다. 갈참은 신갈과 졸참의 중간 위치에 있다고 말할 수 있다.

그리고 갈참과 딱갈間의 類似度指數는 95.9 이고 딱갈과 신갈은 103.7 로 갈참은 신갈과 딱갈의 사이에 있으며 신갈에 보다 가깝다. 또한 딱갈과 졸참間은 그 類緣關係가 가장 멀게 나타나고 있어 既研究結果와 거의 一致하고 있다. 다만 신갈과 딱갈이 보다 가깝다는 結論이 數値上으로는 오히려 신갈과 갈참이 가깝게 나타나고 있다.

同一한 觀察資料에 對한 過去의 研究結果와 電子計算組織과 數値分析에 依한 結果와는 事實 거의 類似하게 나타났지만 後者의 分析結果가 數量的으로 그들 關係를 나타내므로 우리는 確實히 그리고 客觀的으로 그들 類緣關係를 說明할 수 있다는 장점이 있다.

以上의 分析結果에 依하면 植物分類에도 電算組織에 依한 數値分析이 可能하고 實 有利한 計算結果를 주고 있다고 생각된다.

이 研究에서 實施한 數値分析은 다만 20 個의 形態의 因子의 觀測值로 이루어진 것이며 生殖器官, 榮養器官 또는 顯微鏡的 性質 등이 첨가되면 數値分析 結果는 달라질 수도 있다.

因子의 選擇은 目的에 따라 種類가 決定되겠지만 因子의 性質에는 順序尺度(alternatives)와 間隙尺度(scales)가 있으므로 因子區分에 有意하여야 한다.

毛(hair)를 예로서 說明하면;

- ① Hairs; (0) absent, (1) present
- ② Hairs, if present; (0) straight, (1) carled.
- ③ Hairs, if present; (0) 星毛, (1) 單毛.
- ④ Hair distribution; (1) all over, (2) one side (3) widely scattered.
- ⑤ Hair density; (1) Very sparse, (2) sparse (3) medium, (4) dense.
- ⑥ Hair length; (1) ≤ 1 mm, (2) $< 1 \sim \leq 2$ mm (3) $> 2 - \leq 3$ mm.

①~③은 alternatives, ④~⑥은 scales 的 性質을 나타내고 있다. 그러나 이들 數値는 서로 性質이 다르므로 직접 比較가 어렵다. 이 경우는 觀測值를 標準化시켜야 한다. 筆者는 觀測值를 100 分比로 標準化 시켰지만 다음 公式이 標準化 作業에 適合하다.

$$X_{ij} = x_{ij} - \bar{x}_i / S_i \dots \dots \dots \textcircled{1}$$

단 x_{ij} 는 j 列의 i 行의 特性值, \bar{x}_i 는 i 行의 特性值의 全標本의 平均值

S_i 는 上記의 標準偏差

X_{ij} 는 標準化된 j 列 sample 의 i 行의 特性值

이와 같이 標準化된 特性值가 準備되던 식 ① 以外에도 다음과 같은 公式을 標本の 性質에 따라 有意選定한다.

○ Sample 間 相調係數

$$r_{jk} = \frac{\sum_{i=1}^n (X_{ij} - \bar{X}_j) (X_{ik} - \bar{X}_k)}{\sqrt{\sum_{i=1}^n (X_{ij} - \bar{X}_j)^2 \sum_{i=1}^n (X_{ik} - \bar{X}_k)^2}}$$

○ Euclidean distance or taxonomic distance

$$D_{jk} = \frac{\left(\sum_{i=1}^n (X_{ij} - X_{ik})^2 \right)^{\frac{1}{2}}}{n}$$

○ Mean character difference

$$D_{jk} = \frac{\left| \sum_{i=1}^n (X_{ij} - X_{ik}) \right|}{n}$$

上記 式에서 r_{jk} 는 1 에 가까울수록, D_{jk} 는 적을수록 類似性이 높다.

○ Pattern similarity

$$S = \cos \theta = \frac{\sum X_{ij} X_{ik}}{\sqrt{\sum X_{ij}^2 \sum X_{ik}^2}}$$

Pattern 이 類似할수록 1 에 가깝다.

○ Radian distance (Pattern 間 거리)

$$R = \cos^{-1} S (S = \cos \theta)$$

Pattern 間 거리는 pattern 이 같은 境遇에는 $\pi/2 = 1.57$ 이 된다.

結 論

電子計算組織과 數値分析法에 依하여 참나무 基本種의 類緣關係를 증명하고자 “한국산 참나무류의 계통학적 연구, 이창복, 1961,” 論文의 基本調査-覽表를 供試資料로 하여 表1의 內容과 같이 code 化 시킨후 觀察值를 標準化시키기 爲하여 code 별 觀察值를 百分比化 시켰다.

樹種間的 類緣關係 分析은 non-metric coefficient,

$$N_{jk} = \sum_{i=1}^n \left(\frac{|X_{ij} - X_{ik}|}{X_{ij} + X_{jk}} \right), \text{ 式을 적용하여 UNIVAC-}$$

1106 으로 處理하였다.

分析結果는 表 2 와 같이 갈참나무와 신갈나무 間的 類似指數가 71.6으로 가장 가깝고 졸참나무와 떡갈나무 間은 類似指數 12.4 로 가장 멀게 나타나고 있어 李昌福 教授의 論文의 結果와는 一致하고 있으나 筆者의 分析으로 이들 關係가 數量的으로 再證明된 것이다. 供試樹種間的 樹狀型分類에 의거하면 갈참나무와 신갈나무 및 졸참나무 間은 類緣關係가 가깝게 나타나고 있으나 떡갈나무의 類緣關係는 상당히 적게 나타나고 있음을 밝혔다.

이 研究의 結果에 依하면 植物分類에도 電算手法에 依한 數量的 分類가 可能함을 알았다.

引 用 文 獻

1. Burt, R.L., L.A. Edye., W.T. Williams., B. Craf., and C.H.L. Nicholson. 1971. Numerical analysis of variation patterns in the genus *Stylasanthus* as an aid to plant introduction and assessment. Aust. J. Agri. Res., 22, 737-57.
2. Kyuma, K. 1972. Numerical methods in soil classification. Pedologist. 16(1) (Japan)
3. Rayer, J.H. 1966. Classification of soils by numerical methods. Jour. of Soil Sci. 17(1): 79-92.
4. Sneath, P.H.A. 1957. The application of computer to taxonomy. J. gen. Microbiol. 17. 201-26 [Rayer, J.H. 1966. J.H. 1966. J. Soil Sci. 17(1):79-90].
5. Tai-wan Kwon and Yong-hyun Lee. 1972. On the numerical analysis of Korean food supply pattern(1962-1969). Korean Jour. Food Sci. & Tech. 4(2):140-150.
6. Tchang-Bok Lee. 1961. Phytogenetic study of the subgenus *Lepicobalanus* of the genus *Quercus* in Korea. Res. Bull. Korean Agri. Soc.

1. Burt, R.L., L.A. Edye., W.T. Williams., B. Craf.,