

# Cluster분석에 의한 재래종 담배 품종의 분류에 관하여

安大鎮 · 金允東\*

韓國人蔘煙草研究所 陰城試驗場·研究調整部\*

## Varietal Classification on the Basis of Cluster Analysis in Local Tobacco

Dai Jin Ann and Yoon Dong Kim\*

Eumseong Tobacco Experiment Station, Division of Coordination\*  
Korea Ginseng and Tobacco Research Institute, Seoul, Korea

(Received Mar. 31, 1982)

### Abstract

Korean local and introduced varieties were classified by the cluster analysis of correlation and taxonomic distance based on nineteen growth characters.

1. Thirty six varieties can be classified into three groups (I, II, III) by WVGM (weighted variable group method)
2. Major characters for classifying cultivars were days to flowering, number of leaves, leaf length, stem diameter and width of midrib: the five characters seemed to be useful in monothetic classification.
3. Korean varieties were similar to oriental, and japanese varieties to taiwan.
4. WVGM was more accurate and meaningful than classification by WPGM (weighted paired group method) and reticulate diagram of correlation.
5. Characteristics of each group: Group I closely related to many leaves, late of maturity and broad leaf type, Group II related to medium leaves, late of maturity and narrow leaf type, Group III related to few leaves, early of maturity and medium leaf type respectively.

### 서 론

작물의 類緣關係를 규명하는 방법으로 變異학적 방법, 變異학적 방법, 교잡친화성에 의한 방법, 염색체 수와 genome 분석에 의한 방법, 생태학적 방법 및 동위효소의 조성등을 이용하는 방법이 있으며 품종 분류에 있어서 다변량해석법이 각종 분야에도 이용되어져 객관성을 띠는 품종의 분류에도 적용할 수 있게 되었고 그 중에서도 판별계수, 인자분석법, 種間相關 및 分類距離에 의한 Cluster분석 및 주성분 분석등이 종간 유전적인 유연관계를 밝히는데 수리적인 분류법으로 이용되어 왔으며 수치해석학적인

분류는 상호간의 유사성을 수치로서 평가하여 순서화한 후에 분류하는 것으로 재현성 및 객관적인 특성이 있어 많은 연구가 이루어져 밀에 있어서 genetic divergence(4), 유채품종분류(5), Oryza속의 종간 변이(7), 주성분 분석에 의한 개의 분류(8)등의 보고가 있다. 담배분야에는 Murty와 Pavate(10)의 연구결과만 있을뿐, 거의 없는 실정이다.

본 실험은 수집 보존중인 재래종 및 도입종의 質的形質을 대상으로 종합평가한 다변량해석법을 적용하여 품종간 유전적인 유연관계 즉 한국 재래종과 오리엔트종, 일본 재래종등간의 관계를 밝혀 이들 형질간 유사성을 알고자 한다.

재료 및 방법

1. 자료수집

한국인삼연구초연구소 음성시험장(충북 음성) 담배 품종보존 시험구에서 36품종(표 1)을 선택하여 1978년부터 1979년의 2년간에 35개 형질을 조사하여 그 중에서 발아일수, 전엽수, 화축장, 초장, 간장, 최대

엽의 長, 최대엽의 幅 및 위치, 엽형지수, 엽색, 개화일수, 꽃색, 엽간거리, 幹對葉의 각도, 中骨對支骨角, 中骨太細, 葉厚, 요철(凹凸)의 유무등 18개 형질의 반복에 대한 평균치를 분석에 이용하였다. 시험구는 40주 단구제(16.2m<sup>2</sup>)로 하였으며 Polyethylene 필름으로 피복한 일반말칭 재배방법으로 기타는 본 연구소 표준재배법에 준하였다.

Table 1. 36 tobacco varieties

Korean	Japanese	Taiwan	Oriental
1. Hyangtchio	13. Chulsu	27. Songyang	31. Doremisunde-
2. Mokgitchio	14. Kogi	28. Jeryenga	puridanga
3. Byultchio	15. Kokobu	29. Taewaneub	32. Priprieum
4. Kazatchio	16. Subu	30. Tongzuoe	33. Deppronggela
5. Singangtchio	17. Nanbu		34. Deciyublak
6. Suantchio	18. Daruma		35. Debratata-
7. Useultchio	19. Chudaruma		trebishoun
8. Yeatchio	20. Oudaruma		36. Municotine-
9. Osibeutchio	21. Bitchu		freemish
10. Cheongju	22. Kirigasaku		
11. Daegu-1	23. Hatano		
12. Ubangtchio	24. Giesuk		
	25. Ginhyang-1		
	26. Renka		

2. 자료분석

본 실험에 이용된 통계방법은 품종간 상관 및 분류거리에 바탕을 둔 Cluster분석이며 이들에 대한 통계학적인 방법은 Sokal(12)에 의하여 상세히 소개 되었으므로 계산과정만을 간단히 기술하고자 한다.

Cluster분석(t개의 품종, n개의 형질)

1) 각 형질의 평균:  $X_j = \frac{1}{t} \sum_{i=1}^t X_{ij}$

2) 표준편차:  $S_j = \left\{ \frac{1}{t-1} \left[ \sum_{i=1}^t X_{ij}^2 - \frac{(\sum_{i=1}^t X_{ij})^2}{t} \right] \right\}^{1/2}$

3) 표준화:  $X'_{ij} = \frac{X_{ij} - \bar{X}_j}{S_j}$

4) 상관계수(Pearson식, i번째와 k번째 품종간):

$$r_{ik} = \frac{\sum_{j=1}^n (X'_{ij} - \bar{X}'_i)(X'_{kj} - \bar{X}'_k)}{\left\{ \sum_{j=1}^n (X'_{ij} - \bar{X}'_i)^2 \sum_{j=1}^n (X'_{kj} - \bar{X}'_k)^2 \right\}^{1/2}}$$

5) 분류거리:  $D_{ik} = \left( \frac{\sum_{j=1}^n (X'_{ij} - X'_{kj})^2}{n} \right)^{1/2} = \left( \frac{\Delta_{ik}^2}{n} \right)^{1/2}$

6) 품종의 Clustering

가. 加重變數群法(Weighted Variable Group Method: WVGM) Spearman(16) 식으로 群q와 群Q 간의

상관은  $r_{qQ} = \frac{\square qQ}{\sqrt{q+2\Delta q} \sqrt{Q+2\Delta Q}}$  로서 구하였고

여기서  $\square qQ$ : 群q와 Q의 구성원간 전상관의 합.

$\Delta q$ : 群q의 구성원간 전상관의 합

$\Delta Q$ : 群Q의 구성원간 전상관의 합

q, Q: 群q와 群Q의 각각 구성원수이다.

한개의 품종 X와 群q와의 상관은  $r_{xq} = \frac{\sum r_{Xq}}{\sqrt{q+2\Delta q}}$

에 의하여 구하였다.

품종들간의 상관계수 matrix에서 상호상관계수가 가장 큰 품종끼리 하나의 품종군으로 한 다음에 이 품종군과 나머지 품종들간의 상관계수를 Spearman 식에 의하여 구하여 상관계수 matrix를 작성한다. 새로 작성된 상관계수 matrix에서 상호상관계수가 가장 큰 element끼리 합쳐서 더 큰 群으로 분류되는 과정을 반복하여서 점점 큰 群으로 분류하는 것이다.

나. 加重對群法 (Weighted Pair Group Method: WPGM)

Single link cluster 방법으로서 분류거리 (D)로 matrix를 만든 후에 최소값의  $D_{ij}$ 를 찾아  $i$ 번째와  $j$ 번째 품종을 합쳐서 群으로 하고 다시 이 群과 다른 품종  $k$ 와의 분류거리를  $\frac{D_{ik}+D_{jk}}{2}$  즉 품종  $i$ 와 품종  $j$ 를 포함한 群과 품종  $k$ 간 분류거리는 품종  $i$ 와 품종  $k$ 간 분류거리와 품종  $j$ 와 품종  $k$ 간 분류거리의 합의 평균에

의하여 구한 후에 matrix를 다시 만들어 최소값의  $D$ 를 찾아서 群으로 하며 이와같은 과정을 반복해서 Grouping하여 분류하는 방법이다.

## 결과 및 고찰

### 1. 품종간 상관

19개 형질과 36개 품종간의 612개의 상관계수의 분포를 보면 그림 1과 같으며 상관계수는  $+0.89 \sim -0.66$  사이에 넓은 분포범위를 나타내고 이중  $0.1 \sim -0.2$ 에 가장 많은 빈도를 보여주는 한편 공시품종이 모두 포함되어 있으므로 품종간의 유연관계가 적을 것으로 보이며 이러한 사실은 그림 1의 상관계수 분포에서 정규분포에 비하여 음의 상관쪽으로 기울어지는 경향으로 보아 알 수 있다.

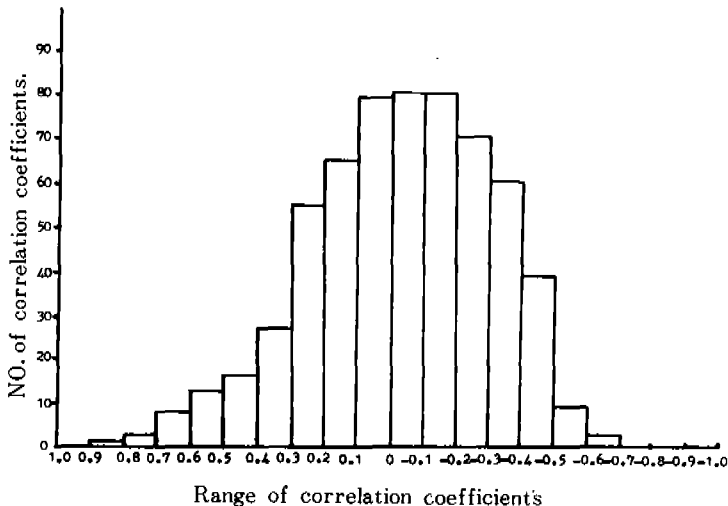


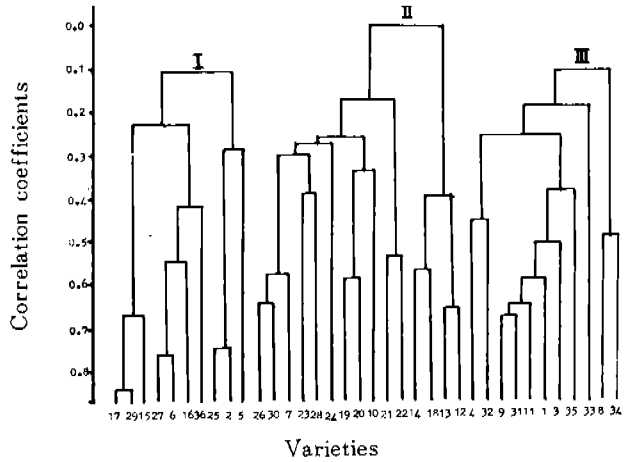
Fig. 1. Histogram of correlation coefficients among 36 varieties.

### 2. 상관계수에 의한 WVGM 분류

종간 상관계수를 기준으로 하여 WVGM에 의하여 cluster한 품종 상호간의 유연관계를 dendrogram으로 본 결과는 그림 2와 같다. 크게는 3개 群으로 분류할 수 있고 이를 다시 세분하면 각 群이 2개 亞群으로 되어 6개 亞群으로 나누어질 수 있으나 각 群에 있어서 한쪽의 亞群이 각 群의 품종의 80%를 포함하고 있으므로 3개 群으로 분류하는 것이 적당할 것 같다. 이는 Murty와 Pavate(10)가 13개 품종을 4개 群으로 분류한 것과는 양상이 다르며 품종수가

많고 지리적, 유전적 변이에 의한 환경변이가 작용한 것으로 생각되어진다(11). 한편으로 각 품종의 기원별로는 품종 群의 형성에서 한국종은 제 I, II 및 III 群에 일본종과 중국종은 제 I 및 II 群, 오리엔트종은 제 I 및 III 群에 분포되어 있고 한국종과 오리엔트종은 주로 제 III 群에 분포되며, 일본종은 주로 제 II 群에 분포되어 있고, 이들 모든 품종을 보면 그 기원에서 유연관계를 보아 한국종과 오리엔트종이 몇 품종의 차이는 있으나 유사하며, 일본종과 중국종이 유사한 것으로 나타났으며 한국종의 일부 품종은 모든 群에 포함되어 있었다.

Fig. 2. Dendrogram of the relation among 36 cultivars based on the WVGM clustering procedure using correlation coefficients.



3. 상관계수 및 WPGM에 의한 분류

상관계수에 의한 망상형 분류는 WVGM과 유사한 3개 군으로 크게 나뉘어지나 임의의 인위적인 분류로 객관성이 결여된 편이며 WPGM 분류는 4개 군으로 크게 구분되며 또 제 I군과 제 III군을 합할수 있고 제 II군과 IV군이 합쳐져서 2개 군을 형성할

수 있고 제 IV군은 개별적인 亞群이 많아 유사성이 부족한 것으로 보여져 타당성이 적다고 생각된다.

4. 분류방법간 비교

이상의 통계학적 방법에 의한 분류를 보면 대개 3개 군으로 되며 WPGM분류는 亞群을 많이 형성

Table 2. Comparison of three clustering methods

Characters	WVGM		WPGM		Reticulate-diagram	
	C. V	F	C. V	F	C. V	F
Days of germination	19.76	2.9753	17.33	1.2210	19.63	3.8532
No. of leaves	34.11	18.0643*	28.44	22.0972*	34.56	18.3999*
Length of flower axis	84.08	1.1739	58.93	0.4591	70.66	0.4217
Plant height	32.90	1.5655	9.68	8.0681	31.05	10.6824*
Stem diameter	29.21	9.9411*	9.18	36.5674*	25.63	8.9353*
Length of the largest leaf	29.64	27.0952*	26.15	30.4499*	31.71	22.4353*
Width of largest leaf	46.96	1.4722	36.55	9.6142*	46.78	2.6173
Leaf position of the largest leaf	41.49	8.8386*	40.72	12.6787*	45.77	3.4520
Leaf color	43.92	0.3692	38.13	1.6655	46.51	0.5267
Days to flowering	18.25	14.3421*	12.76	19.5789*	15.35	12.7887*
Flower color	30.08	8.3792	32.83	4.7933	36.59	2.2357
Length: leaf to leaf	75.72	6.1000	67.45	9.3388*	77.19	2.2313
Leaf angle to stem	48.27	2.2156	50.08	0.9203	45.63	1.3175
Vein angle to midrib	28.27	0.9553	27.73	0.3704	30.28	0.7284

Width of midrib	36.57	12.9220*	26.80	19.2059*	36.89	13.1602*
Leaf thickness	34.26	3.2684	26.17	1.6064	36.30	2.3062
Ripple of leaf	141.60	0.8942	178.78	1.1401	143.36	1.1860
Leaf arrangement	6.10	3.5236	2.21	1.0000	5.98	2.2268

Note ; WVGМ : Weighted variable group method.

WPGM : Weighted pair group method.

\* : Significant at 5% level.

되어 群間 변동이 일부 품종에서 엇보였고 표 2에서는 단일 형질로 분류하려면 전엽수, 최대엽의 長, 개화일수, 中骨太細 및 幹徑 등의 형질이 유용할 것으로 생각되며 특히 개화일수 및 최대엽의 長 그리고 幹長이 중요한 형질로 나타나고 있으며 이는 개화기가 분류거리에 크게 영향하여 早生群과 晩生群으로 크게 분류된다는 (7,13) 보고와 일치하는 경향이며 또한 초장과도 일치한다(11). 그리고 방법간에는 WV

GM이 WPGM보다 더욱 잘 맞는다는 보고 (7, 14)와 또한 일치하는 경향이다.

### 5. 담배 품종群의 특성비교

WVGМ에 의하여 분류된 품종群의 주요 특성은 표 3와 같으며 제 I群은 多葉, 晩生, 広葉型, 葉長이 작은 품종群이고, 제 II群은 엽수는 중간, 晩生,

Table 3. Characteristics of three groups based on WVGМ methods

Characters	Unit	Group		
		I	II	III
Days of germination	day	16.1	17.3	16.2
No. of leaves *	number	32.5	30.0	24.0
Length of flower axis	cm	53.2	51.5	58.9
Plant height	cm	191.1	193.0	179.1
Stem diameter *	cm	2.37	2.75	2.35
Length of the largest leaf *	cm	45.4	61.0	53.1
Width of the largest leaf *	cm	33.0	36.9	34.3
Leaf position of the largest leaf		6.8	6.4	5.2
Length/Width of largest leaf		1.40	1.70	1.56
Leaf color		3.0	2.9	2.9
Days to flowering *	day	51.4	51.7	45.7
Flower color *		3.0	2.5	3.1
Length : leaf to leaf	cm	61.3	58.7	83.6
Leaf angle to stem	degree	48.2	55.6	53.3
Vein angle to midrib	degree	65.3	68.1	69.1
Width of midrib *	cm	1.06	1.42	1.25
Leaf thickness	0.1 mm	38.2	37.2	41.6
Ripple of leaf		0.7	0.9	0.9
Leaf arrangement	degree	135.0	136.8	139.1
Maturity		Late	Late	Early
No. of leaves		Many	Medium	Few
Size of leaf		Small	Large	Medium

\* : Significant at 5% level.

細葉型, 잎이 큰 품종群이고, 제Ⅲ群은 早生, 少葉, 中葉型, 잎의 크기가 중간인 품종群으로 나타났으며 기타 형질에 있어서는 제Ⅰ群은 幹徑이 적고, 中骨太細가 적은 群, 제Ⅱ群은 幹徑이 크고, 葉幅이 큰 群, 제Ⅲ群은 幹徑이 작고 엽간거리가 크고 中骨太細가 중간인 群으로 나타났으며 일반적으로 개화일수에 의한 早生, 晩生의 구별(4, 13)과 일치하였고 초장에서는 Murakami 및 Nakamura(8)와 일치되지 않은 바 이는 주성분을 이용한 것과의 차이로 생각된다. 그리고 이들 형질은 분류방법 및 群間에 다소 차이가 있으나 이는 Cluster 과정중의 차이로 여겨지며, 종합적으로 보면 제Ⅱ群이 제Ⅰ群보다 다수성 품종군으로 나타나고 있다.

## 결 론

재래종 및 도입종 36개 품종을 선정하여 19개 형질을 이용하여 품종간 상관계수 및 분류거리에 의한 Cluster 분석을 하여 품종의 유연관계 및 분류에 적합한 대상형질을 찾고 품종군의 특성을 조사하여 얻은 결과는 다음과 같다.

1. WVGM (加重變數群法)에 의한 분류는 크게 3개 群으로 크게 나뉜다.
2. 단일 형질로 분류하려면 개화일수, 전엽수, 葉長, 中骨太細 및 幹徑이 중요 형질이 된다.
3. 품종의 유연관계는 주로 한국종과 오리엔트종이, 그리고 일본종과 중국종이 유사성을 보였다.
4. 분류 방법중에서는 WVGM이 가장 좋았다.
5. 群間 특성중 제Ⅰ群은 多葉, 晩生, 広葉型이고, 제Ⅱ群은 中葉, 晩生, 細葉型이며, 제Ⅲ群은 小葉, 早生, 中葉型였다.

## 참 고 문 헌

1. 崔海椿, 李正日, 育種誌. 11:179-195(1979)
2. 平野夕, 日本育雜. 27:350-358(1977)
3. 上鳥愔誌, 日本育雜. 24:261-268(1974)
4. Lee, J. and P. J. Kaltsickes. Euphytica. 22:124-131(1973)
5. 李榮萬, 育種誌. 12:61-92(1980)
6. Liang, G. H. L. and A. J. Casady. Crop. Sci. 6:76-79(1966)
7. Morishima, H. and H. Oka. Evolution 14:153-165(1959)
8. Murakami, M. and S. Nakamura. Japanese J. Breed. 26(Subscr. 2):89-90(1976)
9. Murty, B. R. and V. Arumachalam. Ind. J. Gen. and Pl. Breed. 27:60-69(1967)
10. Murty, B. R. and M. V. Pavate. Indian J. Genet. Plant Breed. 22:68-80(1962)
11. Sachan, K. S. and J. R. Sharma. Ind. Gen. Pl. Breed. 31:86-93(1961)
12. Sokal, R. R. and P. H. A. Sneath. "Principles of numerical taxonomy" P. 359-370, W. H. Freeman Company, Sanfransisco and London. (1963)
13. Somayajulu, P. I. N., A. I. Joshi and B. R. Murty. Ind. J. Gen. and Pl. Breed. 30:47-58(1970)
14. Yamada. T. and S. Suzuki. JIBP. Synthesis. 5:137-143(1975)
15. Zveritt, B. S. Biometrics 35:169-181(1979)
16. 農林水産技術會議事務局, 農林水産試驗研究のための統計的 數學的方法. Ⅲ. 特殊編:23-32(1975)