

**ESSAI DE CLASSIFICATION
DES POPULATIONS DE PANICUM
MAXIMUM JACQ. D'AFRIQUE
DE L'EST SUR DES CARACTERES
MORPHOLOGIQUES QUALITATIFS**

R. et J. RENÉ CHAUME

J. PERNÈS

D. COMBES

Biologie

Amélioration

Plantes

Utiles



OFFICE DE LA RECHERCHE SCIENTIFIQUE ET TECHNIQUE OUTRE-MER

CENTRE D'ADIOPODOUME — COTE D'IVOIRE

Novembre 1969

B.P.20 — ABIDJAN

OFFICE DE LA RECHERCHE SCIENTIFIQUE ET TECHNIQUE OUTRE-MER

Centre d'Adiopodoumé

Laboratoire de Génétique

ESSAI DE CLASSIFICATION DES POPULATIONS DE PANICUM
MAXIMUM Jacq. D'AFRIQUE DE L'EST SUR DES CARACTERES
MORPHOLOGIQUES QUALITATIFS.

par

R. et J. RÉNÉ-CHAUME - J. PERNÈS - D. COMBES

Novembre 1969

P L A N

A - Introduction	1
B - Collecte des données	3
C - Coefficients d'association	5
1. Généralités	5
2. Indice de similarité et probabilité	6
3. Indice de SMIRNOV	8
a. description	8
b. Essai d'étude des propriétés statistiques	9
4. Indice de Rogers et Tanimoto	12
D - Regroupement	13
1. Dendrogrammes	13
2. Méthode nodale de Rogers et Tanimoto	14
E - Résultats	16
1. Matériel	16
2. Graphiques	16
3. Conclusions	17
F - Bibliographie	21

Résumé

En tenant compte des informations apportées par GOODALL sur les relations entre indice de similarité et probabilité le choix d'un coefficient d'association pour un essai de classification des populations de Panicum maximum sur des caractères morphologiques qualitatifs s'est porté sur l'indice de SMIRNOV qui donne plus d'importance aux états rares. Quelques propriétés statistiques de l'indice sont abordées. La comparaison avec l'indice de ROGERS et TANIMOTO est faite sur les données de 11 populations d'Afrique de l'Est. Les regroupements par construction de dendrogrammes et par la méthode nodale lorsque des groupes fermés existent montrent que les populations de Panicum maximum d'Afrique de l'Est sont en cours de triage de phénotypes propres à chaque région.

A - Introduction

La mise au point de méthodes de taxonomie numérique s'inscrit dans le programme général du laboratoire de Génétique, dont le but est l'étude des possibilités d'amélioration de l'espèce fourragère Panicum maximum Jacq. Elle doit permettre d'apporter des renseignements dans plusieurs parties du programme.

.. L'étude des descendance par graines de plantes mises en condition d'autopollinisation et l'analyse cytologique ont montré que le mode de reproduction est une apomixie facultative. Le taux de sexualité de 2 % permet l'obtention de hors-types. L'analyse des caractères qualitatifs par les méthodes de taxonomie numérique permettra d'établir et d'apprécier les relations phénotypiques entre les hors-types et leurs parents. La filiation étant connue, elle permettra de mettre à l'épreuve la validité biologique des techniques de classification. Le regroupement par familles des hors-types de 2ème génération d'autopollinisation devrait être un critère important.

.. Des premiers essais ont montré qu'il est peut être possible d'obtenir sans castration, par ensachage double, des hybrides ; par les méthodes de taxonomie numérique on pourra faire l'étude d'un ensemble réunissant parents, hybrides et descendants par auto-fécondations.

.. Le but de la mission de prospection de Panicum maximum en Afrique de l'Est effectuée du 1-7 au 23-7-67 par J. PERNES et D. COMBES était l'étude de la variabilité de l'espèce, liée aux structures de l'environnement (milieu géographiquement hétérogène), aux modes de reproduction (recherches d'éventuelles populations à taux de reproduction sexuée élevé). Ces plantes prospectées en Afrique de l'Est présentent des phénotypes qualitativement extrêmement variés. Cet ensemble de populations doit être analysé qualitativement afin d'obtenir une description permettant en particulier de constituer en groupes les divers phénotypes. Plusieurs types de population ont été repérés : des populations monomorphes entre lesquelles il existe des points de discontinuité ; des populations polymorphes contenant des Panicum non maximum ; des po-

pulations contenant des *Panicum maximum* diploïdes ; des populations polymorphes voisines des précédentes.

Parallèlement une analyse quantitative est en cours (J. PERNES) selon les modèles décrits dans "les populations naturelles ivoiriennes de l'espèce *Panicum maximum* et les types analogues introduits". Chaque essai correspond à un groupe de populations appartenant à une unité géographique.

.. Une seconde prospection en Afrique de l'Est (J. PERNES et D. COMBES février 1969) a permis de rapporter plusieurs phénotypes diploïdes distincts issus : de la population trouvée lors de la précédente prospection ; de populations géographiquement peu éloignées, mais de milieux différents ; d'une nouvelle population sans relation avec les précédentes peu distante de la côte, où coexistaient diploïde et tetraploïde de phénotype très original. L'analyse qualitative par les méthodes de taxonomie numérique de l'ensemble des diploïdes est envisagée.

La comparaison des résultats quantitatifs et qualitatifs sera une bonne source de renseignements. La mise au point de méthodes de taxonomie numérique sur les données des populations d'Afrique de l'Est de la prospection 1967 est le but de ce travail.

B - Collecte des données.

Les variables choisies pour établir une classification des différents clones de *Panicum maximum* d'Afrique de l'Est sont les valeurs des caractères morphologiques décrivant le mieux possible les divers phénotypes. Un phénotype sera d'autant mieux décrit que le nombre de caractères sera grand.

Les caractères choisis doivent être facilement repérables et appréciables sans trop d'ambiguïté ; ils doivent être suffisamment stables dans le temps et au sein de chaque clone. La stabilité dans le temps est assez relative car il est évident que les saisons (saison des pluies et saison sèche) influencent la nature des caractères morphologiques. La notation devra donc se faire dans un temps assez court à l'intérieur d'une saison. Elle sera de plus faite par le même (ou les mêmes) observateur (s) pour éviter les variations extérieures. Les conditions de luminosité (par exemple pour les couleurs) ou d'humidité (pour les pilosités) devront être les plus homogènes possibles tout au long de l'observation.

L'homogénéité du terrain sur lequel se trouve la collection à observer doit être grande et le niveau de fertilité suffisamment élevé pour que les phénotypes s'expriment pleinement. L'âge des plantes, ou des différentes parties des plantes doit être constant.

Les différents caractères ont été réunis dans les tableaux de I à V. Ils se répartissent en cinq groupes. Un groupe pour l'aspect général de la plante, un groupe pour la description de la gaine et de la base du limbe (feuilles I), un groupe pour les inflorescences et un pour les épillets. Les états des caractères sont codés de 0 à 5. 57 caractères au total sont définis, ce qui semble d'après la littérature un nombre suffisant pour obtenir une bonne séparation des clones.

Les caractères purement qualitatifs, et les caractères semi-quantitatifs (c'est à dire ordonnés, mais non métriques) sont traités ensemble. Ces derniers posent en général moins de problèmes que les autres. Les caractères hauteur, grosseur densité sont faciles à définir ; mais des difficultés apparaissent pour

ASPECT GENERAL

N°	DESIGNATION	PRECISION	ETATS							
			0	1	2	3	4	5		
1	TOUFFE	HAUTEUR (en. Cm)	A FLORAISON	$h < 70$ Ex : 280	$70 \leq h \leq 100$	$100 \leq h < 140$ Ex : 267	$140 \leq h < 170$ Ex: Type II	$h > 170$ Ex : K 187		
		REMONTAISON	^{5 semaines} APRES FAUCHE	AUCUN PIED N' EST FLEURI	MOINS DE LA MOITIE EST FLEURIE	PLUS DE LA MOITIE	TOUS LES PIEDS SONT FLEURIS			
		PORT DU PIED		ETALE	DEMI ETALE	DRESSE				
		PORT FEUILLE	3 ^{ème} FEUILLE	DRESSEE	CASSEE	RETOMBANTE				
		SECTION	APRES FAUCHE	FAIBLE	MOYENNE	GRANDE				
6	TAILLES	GROSSEUR		Ex: 280	Ex : 267	Type II	Ex: K 187			
		DENSITE		FAIBLE	MOYENNE	FORTE				
		ANGULATION	DES ENTRE NOEUDS	ABSENCE	PRESENCE					
		PILOSITE TIGE		ABSENCE	VERS LE HAUT	REPARTITION UNIFORME				
		PILOSITE NOEUD		ABSENCE	PEU	BEAUCOUP				
		COLORATION ROUGE SOUS LE NOEUD		ABSENCE	EN ANNEAU DE $\frac{1}{2}$ cm -	SUR PLUSIEURS cms				

FEUILLES SERIE I

°	DESIGNATION	PRECISION	E T A T S						
			0	1	2	3	4	5	
1 2 3 4 5 GAINÉ	DENSITE		ABSENCE OU INSTABLE		FORTE DENSITE				
	LOCALISATION	LONGITUDINALE	ABSENCE DE PILOSITE	VERS LE BAS	VERS LE HAUT	REPARTITION UNIFORME			
		TRANSVERSALE	ABSENCE DE PILOSITE	AU DOS DE LA GAINÉ	REPARTITION UNIFORME				
	ASPECT		ABSENCE DE PILOSITE	POILS DURS ET COURTS	POILS DURS ET LONGS	DUVET	POILS MOUS ET LONGS		
	CILIATURE	BORD DE LA GAINÉ	ABSENCE	PRESENCE					
6 7 8 9 0 BASE DU LIMBE	LIGULE	HAUTEUR	BASSE	HAUTE					
		PILOSITE	ABSENCE	POILS COURTS	POILS LONGS				
	NERVURE CENTRALE PILOSITE	FACE SUP.		ABSENCE	PRESENCE				
		FACE INF.	DENSITE	ABSENCE	FAIBLE	FORTE			
			ASPECT		ABSENCE	POILS DURS	DUVET	POILS MOUS ET LONGS	POILS TRES LONGS
BASE DU LIMBE	PILOSITE	ABSENCE		TRES PILEUX Ex: 267					
ANGLE GAINÉ LIMBE	PILOSITE	ABSENCE	PRESENCE	FORTE PILOSITE					

TABLEAU III

FEUILLES

SERIES II

N°	DESIGNATION	PRECISION	ETATS					
			0	1	2	3	4	5
23	PILOSITE FACE SUPERIEURE	DENSITE	ABSENCE OU INSTABLE					
24		LOCALISA- TION	ABSENCE	VERS LE BAS	VERS LE HAUT	REPARTITION UNIFORME		
25		ASPECT	ABSENCE	POILS DURS	DUVET	POILS MOUS ET LONGS	POILS TRÈS LONGS	POILS DURS ET DUVET MELANGES
26	PILOSITE FACE INFERIEURE	DENSITE	ABSENCE OU INSTABLE					
27		LOCALISA- TION	ABSENCE	VERS LE BAS	VERS LE HAUT	REPARTITION UNIFORME		
28		ASPECT	ABSENCE	POILS DURS	DUVET	POILS MOUS ET LONGS	POILS TRÈS LONGS	POILS DURS ET DUVET MELANGES
29	COULEUR		VERT JAUNE	VERT	VERT BLEU			
30	LONGUEUR	DE LA 3 ^{eme} FEUILLE	15 à 20 cm	40 cm	60 cm			
31	LARGEUR	DE LA 3 ^{eme} FEUILLE		Ex: 280	Ex : 267	Ex: Type II	Ex : K 187	
32	DENTELURE	BORD DU LIMBE	ABSENCE	PRESENCE				
33	MALADIE	CERCOSPORIUM	ABSENCE		TRES REPANDUE			
34		COLETOTRICUM	ABSENCE		TRES REPANDUE			
35		ROUILLE	ABSENCE		TRES REPANDUE			

TABLEAU IV

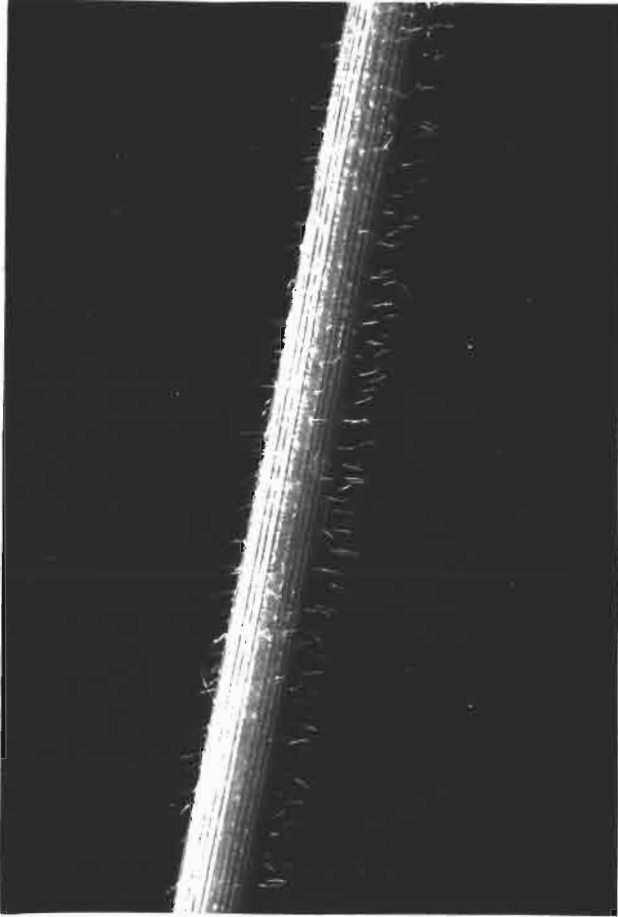
INFLORESCENCES

N°	DESIGNATION	PR- PRECISION	E. T. A. T. S.					
			0	1	2	3	4	5
36	FORME		TRIANGULAIRE	BOULOIDE Ex : 65				
37	LONGUEUR	DES RAMIFI- CATIONS DU VERTICILLE	$< \frac{2}{3}$ DE LA LONGUEUR TOTALE	$= \frac{2}{3}$	$> \frac{2}{3}$			
38	POSITION DES RAMIFICATIONS	SECONDAIRES PAR RAPPORT AUX PRIMAIRES	PARTOUT	INCERTAIN	PLUTOT VERS LE HAUT			
39	PORT DES RA- MIFICATIONS	SECONDAIRES	SERREES SUR LES PRI- MAIRES	ECARTEES DES PRI- MAIRES				
40	A X E	PORT	MOU Ex : Sotuba	RAIDE				
41		PILOSITE	DENSITE	ABSENCE	PRESENCE			
42			POSITION	ABSENCE	VERS LE BAS	VERS LE HAUT	UNIFORME	
43	VERTICILLE	PORT	MOU	RAIDE	TRES RAIDE			
44		PILOSITE	ABSENCE		FORTE Ex : 267			
45	PRUINE		ABSENCE	PRESENCE				

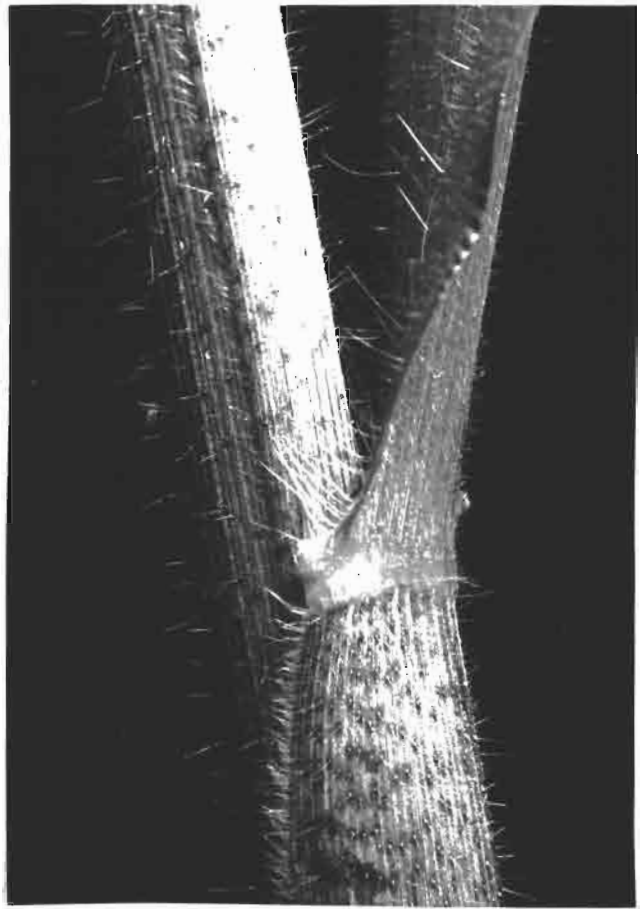
N°	DESIGNATION	PRECISION	ETATS					
			0	1	2	3	4	5
46	DENSITE		FAIBLE	MOYENNE	FORTE			
47	DISTRIBUTION	LE LONG DES RAMIFICA- TIONS	DISPERSE	UNIFORME	PAR GROUPES	RACEME SPICIFORME		
48	FORME		TOSSE	NORMAL	ALLONGE (P.B)			
49	GROSSEUR		PETIT	NORMAL	GROS			
50	COULEUR	PRINCIPALE	JAUNE FONCE	VERT PALE	VERT			
51		TACHES	ABSENCE	BRUN	ROUGE	VIOLET		
52	PILOSITE		ABSENCE	Ex: 280	POILS LONGS Ex: H1			
53	SOIE		ABSENCE	RARE	PRESENCE SUR PRESQUE TOUS	PRESENCE SUR TOUS	PLUS DE 2 SOIES	SOIES COURTES
54	1ère FLEUR	SEXE	STERILE	MALE VARIABLE	TOUJOURS MALE			
55	SOROSPORIUM		ABSENCE		TRES REPANDUE			
56	FUSARIUM		ABSENCE		TRES REPANDUE			
57	CEREBELLA		ABSENCE		TRES REPANDUE			

certains caractères qualitatifs. L'aspect de la pilosité de la nervure centrale concernant la face inférieure de la base du limbe n'a pu être défini qu'à l'aide de 6 états (0 à 5) ; le 0 correspondant à l'absence de pilosité, ce qui est une entrave à la rigueur puisque 0 pour le caractère 19 (densité) entraîne automatiquement 0 pour le caractère 20 (aspect). 5 caractères sur les 57 présentent cette anomalie qu'il n'a pas été possible de supprimer. Lorsqu'un caractère est uniforme sur toute la population étudiée il est évidemment à éliminer ce qui donne à la description un aspect non absolu. Par exemple le caractère 3⁶ du groupe inflorescences qui décrit la forme de l'inflorescence est dans la plupart des cas à supprimer ; l'état bouloïde ne se rencontrant que très rarement (seuls les clones 65 et 88 provenant du Congo et du Brésil respectivement présentent ce caractère, tous les autres clones rencontrés ont des inflorescences triangulaires).

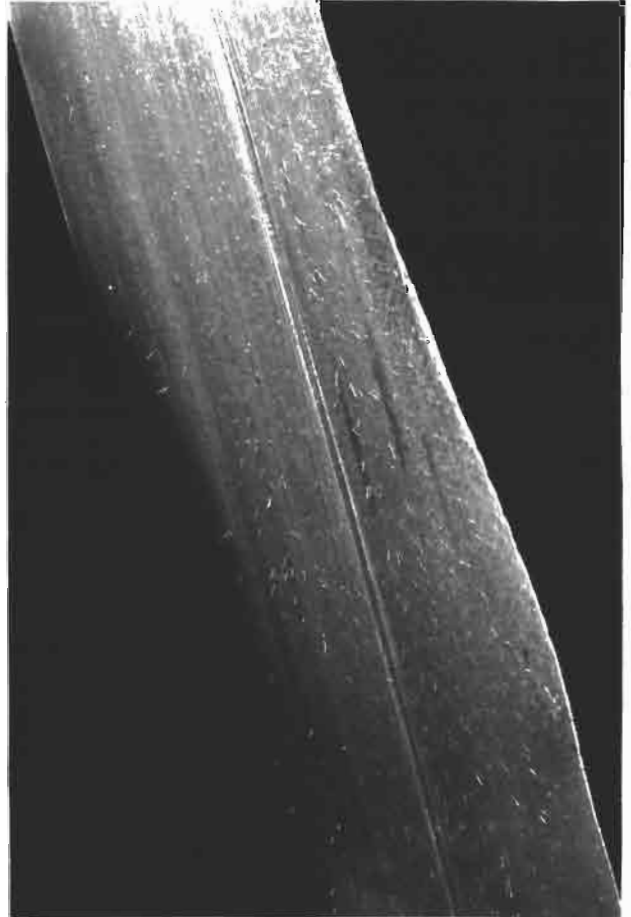
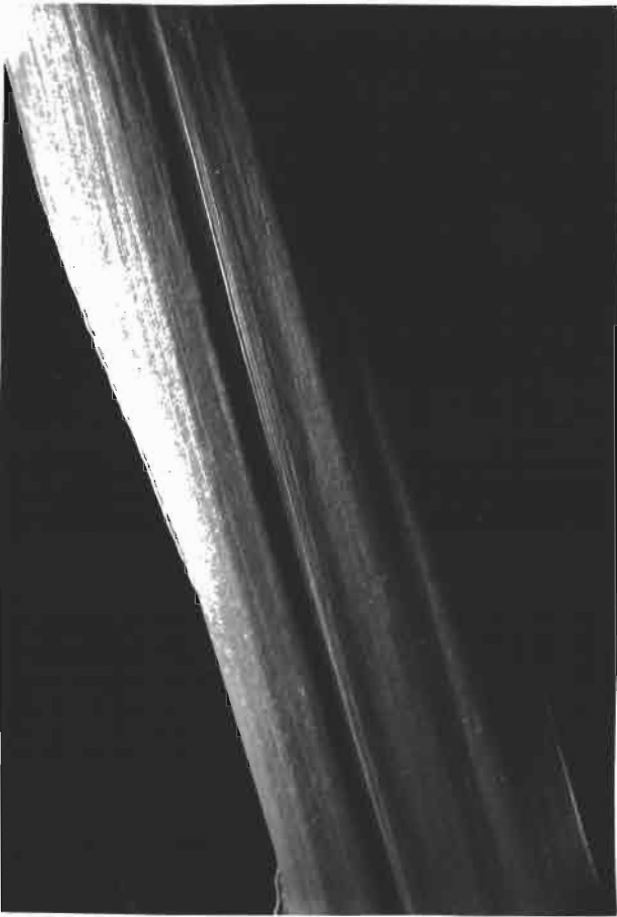
Les observations sont faites par groupe de caractères dans un temps suffisamment court pour avoir le moins de variation possible due aux conditions extérieures. Pour les deux derniers groupes (inflorescences et épillets) les observations sont faites lorsque la moitié, au moins, des plantes est à floraison.



PILOSITE DE LA TIGE

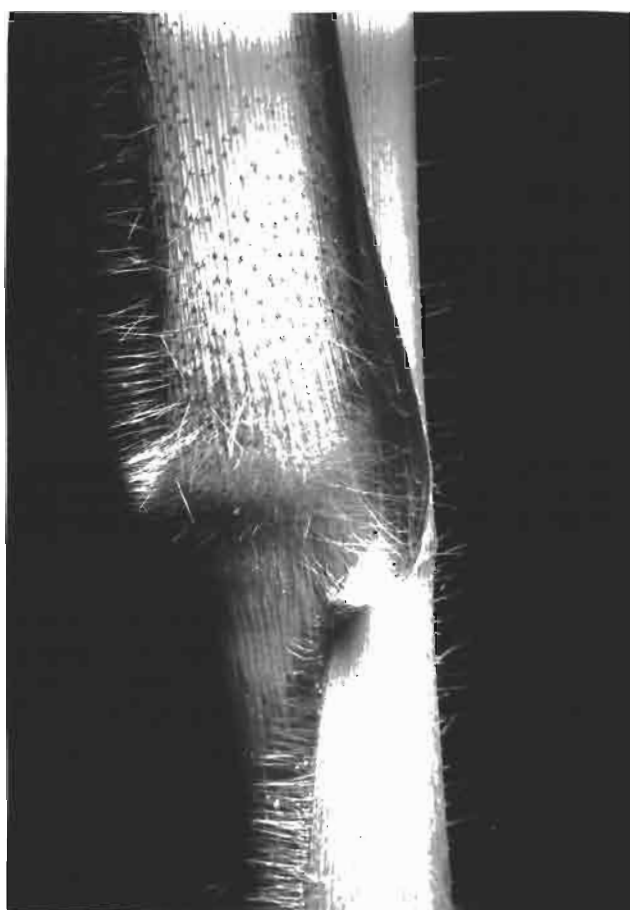
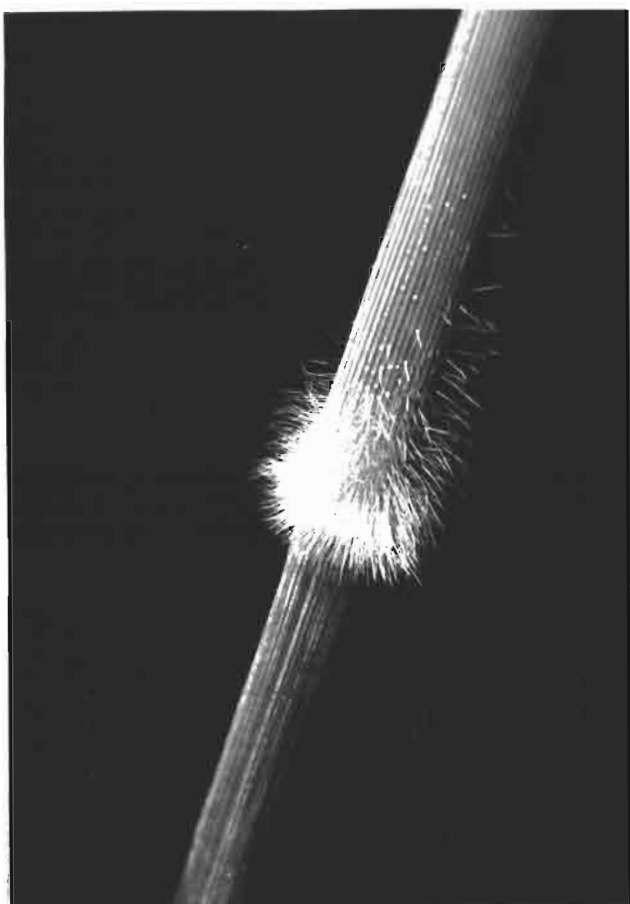


PILOSITE TIGE GAINÉ ET BASE DU
LIMBE



PILOSITE DU LIMBE

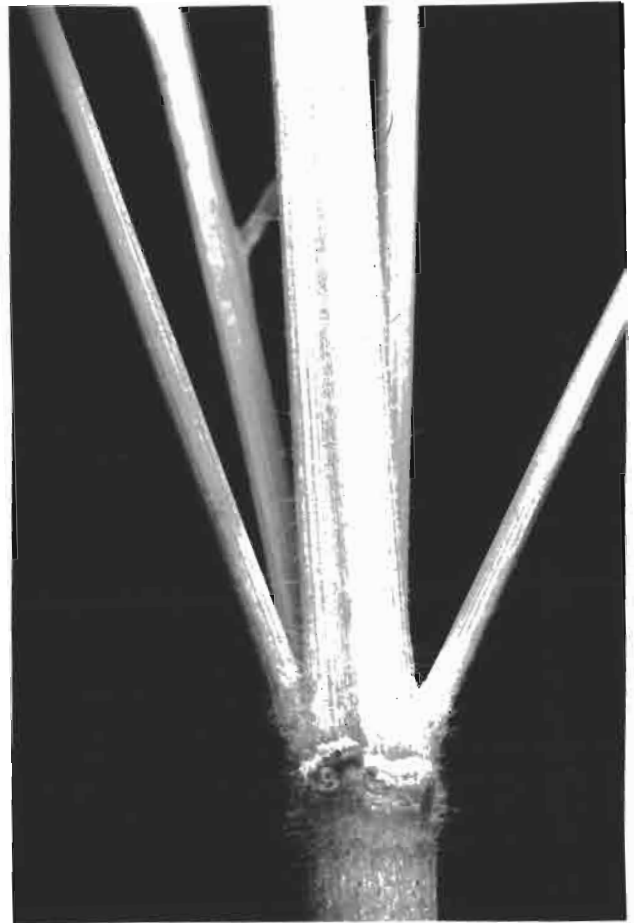
Photos Y. GILLON



PILOSITE DU NOEUD



PILOSITE DE LA LIGULE



PILOSITE DU VERTICILLE



SOIES SUR LES EPILLETS

Photos Y. GILLON

C - Coefficients d'association.

1 - Généralités.

Le problème qui se pose est un problème de classification selon la définition de P. DAGNELIE. "On considère un échantillon (ou une population) de n individus, pour chacun desquels on dispose des valeurs de p variables, et on souhaite subdiviser l'ensemble des n individus en un nombre limité, souvent mal défini de groupes d'individus semblables ; on souhaite définir ou "faire" des classes, c'est à dire établir une classification".

Les valeurs des p variables permettent de donner une mesure à la ressemblance entre un couple d'individus, qu'on appelle "indice de similarité".

Le problème de la définition d'un indice de similarité biologiquement valable est complexe, et les solutions données par les auteurs consultés sont nombreuses.

Il s'est avéré qu'un nombre de plus en plus grand de taxonomistes font appel aux ordinateurs pour le traitement de leurs données. L'importance des calculs nécessités par les différentes méthodes montre que la machine electro-mécanique devient insuffisante.

SOKAL et SNEATH (1963) donnent leurs principes de base de la taxonomie numérique.

"1. La taxonomie idéale est celle dans laquelle le taxon a le plus grand contenu d'information et qui est basée sur le nombre de caractères le plus grand possible.

2. A priori chaque caractère est de poids égal dans la création naturelle des taxa.

3. Toute similarité (ou affinité) entre deux entités quelconques est une fonction de la similarité de beaucoup de caractères pour lesquels on les compare.

4. Des taxa distincts peuvent être construits en raison des corrélations des divers caractères qu'on étudie.

5. La taxonomie numérique comme nous la concevons est donc strictement une science empirique.

6. L'affinité est estimée indépendamment des considérations phylogénétiques.

Sur le point 5 tous les auteurs ne sont pas d'accord. GOODALL (1964-1966 a - 1966 b - 1966 c - 1966 d - 1966 e - 1968) ne considère pas la taxonomie numérique comme une science empirique mais raisonne en termes probabilistes.

2 - Indice de similarité et probabilité.

Dans le problème qui nous préoccupe les variables décrivant les individus sont des caractères qualitatifs à deux ou plusieurs états et des caractères non métriques mais ordonnés à deux ou plusieurs classes. Ce sont des caractères morphologiques sujets à des processus aléatoires puisqu'ils dépendent du génotype de l'organisme lui même fonction des effets aléatoires de mutation et de recombinaison. Partant de ce fait GOODALL a construit un indice de similarité basé directement sur la théorie des probabilités.

Pour chaque caractère i et pour chaque couple d'états (i_k, i_l) , la similarité est déterminée de telle manière qu'elle engendre une relation d'ordre permettant de ranger les couples d'états. Etant donné deux couples (i_k, i_l) et (i_m, i_n) ou bien $(i_k, i_l) < (i_m, i_n)$ ou bien $(i_k, i_l) \geq (i_m, i_n)$ pour la similarité.

Soit (i_m, i_n) un couple pris au hasard on définit Prob $\left\{ (i_m, i_n) \geq (i_k, i_l) \right\} = P(i_k, i_l)$

La similarité entre i_k et i_l est alors

$$S(i_k, i_l) = 1 - P(i_k, i_l).$$

L'indice de similarité pour chaque couple est donc défini comme le

complément de la probabilité qu'un couple pris au hasard ait une similarité supérieure ou égale à celle du couple en question. Ces probabilités sont combinées en faisant l'hypothèse que les valeurs prises par les différents caractères dans le même individu sont indépendants. Le calcul imposé par la méthode exacte devient très rapidement impraticable même pour un calculateur électronique lorsque ce nombre de caractères et le nombre d'états par caractère augmentent. La solution de FISHER pour la combinaison des probabilités continues de tests indépendants peut être utilisée comme approximation. Pour les probabilités discrètes, la méthode donne un biais positif qui peut être très grand. Une amélioration a été proposée par LANCASTER (1949). Malgré toutes ces simplifications le calcul ne peut être fait que sur ordinateur.

Il est évident que cet indice est original par rapport à ceux couramment utilisés en taxonomie numérique. Il s'applique à toutes les catégories de caractères, qualitatifs, ordonnés ou quantitatifs. Les concordances qui sont les moins probables donnent une meilleure preuve de ressemblance donc une contribution plus forte dans la similarité entre deux individus. Cet indice donne donc aux caractéristiques rares le poids accru que les taxonomistes tendent à préférer. De plus un indice de déviation défini par GOODALL 1966, ouvre la voie à un processus complet de taxonomie numérique contrôlé par des niveaux de signification. A moins qu'on ait une information extérieure sur les fréquences des caractères dans la population, elles doivent être estimées à partir des individus considérés pris comme un échantillon de la population. Cet indice dépend donc du contexte dans lequel le couple d'individus est considéré ce n'est pas un indice absolu.

D'autres auteurs tout en établissant des indices empiriquement se rapprochent de l'idée de GOODALL par le fait qu'ils pensent qu'il est souhaitable de donner plus d'importance aux états rares des caractères. C'est à dire un état d'un caractère représenté dans 10 % de la population aura plus de poids dans la similarité entre deux individus qu'un état représenté dans 90 % de la population. Pour SMIRNOV les poids accordés aux caractères sont fonction de la fréquence de ceux-ci dans la population. HALL (1965)

pour les caractères binaires accorde aux états des poids fonction de leur abondance dans la population. WILLIAMS et DALE (1964) calculent un χ^2_{jk} entre chaque couple (jk) de caractères à deux états (présence-absence) et $\sum_{k \neq j} \chi^2_{jk}$ est calculée sur tous les χ^2 se rapportant au caractère j. WILLIAMS et DALE s'adressent spécialement aux études écologiques, le coefficient d'association est pour eux une distance euclidienne pondérée par les χ^2 définis ci-dessus.

Pour le problème qui nous préoccupe l'indice de SMIRNOV est apparu comme plus conforme aux données. Pour se rendre compte de l'information supplémentaire apportée par la pondération des caractères un indice simple de ROGERS et TANIMOTO est essayé sur les mêmes données.

3 - Indice de SMIRNOV

a) Description

L'indice de SMIRNOV se formule de la façon suivante :

s = nombre de clones

E_i = nombre de clones présentant le caractère E à l'état i.

e_i = nombre de clones ne possédant pas le caractère E à l'état i.

On a : $E_i + e_i = s$ et $\sum_i E_i = s$ $i \in [1, e]$ pour l'état E.

La pondération se fait de la façon suivante :

pour un caractère E et un état i considérés le poids accordé W_{Ei} sera :

* $\frac{e_i}{E_i}$ si les deux clones sont à l'état i pour le caractère E.

* * $\frac{E_i}{e_i}$ si les deux clones ne sont pas à l'état i pour le caractère E.

* * * -1 si l'un des clones est à l'état i, l'autre ne l'étant pas pour le caractère E.

On fait ensuite la moyenne des poids pour tous les états pour le caractère E.

$$\bar{W}_E = \frac{1}{e} \sum_{i=1}^e W_{Ei}$$

Si on considère enfin tous les caractères $E^1 E^2 E^3 \dots E^m$;

$$\bar{W}_{E^\alpha} = \sum_{i=1}^{e^\alpha} W_{Ei}^\alpha$$

L'indice de similarité entre les deux clones f et g sera :

$$t_{f,g} = \frac{1}{n} \sum_{\alpha=1}^m \bar{W}_{E^\alpha} ; \quad n = \sum_{\alpha=1}^m e^\alpha$$

$$t_{f,g} = \frac{1}{n} \sum_{\alpha=1}^m \left[\sum_{i=1}^{e^\alpha} W_{Ei}^\alpha \right]$$

Cet indice peut prendre des valeurs positives ou négatives supérieures ou inférieures à 1 en valeur absolue.

b) Essai d'étude des propriétés statistiques de l'indice de SMIRNOV.

Hypothèses.

On considère la population théorique dans laquelle pour le caractère α ($\alpha = 1, 2, \dots, m$) l'état i est représenté avec une probabilité $\frac{E_i^\alpha}{s}$

on a les relations suivantes :

$$\sum_i \frac{E_i^\alpha}{s} = 1 \quad (i = 1, 2, \dots, e^\alpha)$$

$$\sum_{\alpha=1}^m e^\alpha = n$$

$$\frac{E_i^\alpha}{s} + \frac{e_i^\alpha}{s} = 1$$

Les m caractères sont mutuellement indépendants. Cette population contient toutes les associations possibles des états des caractères. Soient deux individus X et Y appartenant à cette population. On peut les représenter par les vecteurs ayant pour composantes les codes des différents caractères :

$$X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_\alpha \\ \vdots \\ x_m \end{bmatrix} \quad Y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_\alpha \\ \vdots \\ y_m \end{bmatrix}$$

On considère la fonction $f_i^\alpha(x_\alpha, y_\alpha)$, contribution à la similarité entre X et Y pour l'état i du caractère ; cette fonction est définie comme suit :

$$f_i^\alpha(x_\alpha, y_\alpha) = \frac{e_i^\alpha}{E_i^\alpha} \text{ si } x_\alpha = y_\alpha = i ; \text{ Prob} \left\{ x_\alpha = y_\alpha = i \right\} = \left(\frac{E_i^\alpha}{s} \right)^2$$

$$f_i^\alpha(x_\alpha, y_\alpha) = \frac{E_i^\alpha}{e_i^\alpha} \text{ si } x_\alpha \neq i, y_\alpha \neq i ; \text{ Prob} \left\{ x_\alpha \neq i, y_\alpha \neq i \right\} = \left(\frac{e_i^\alpha}{s} \right)^2$$

$$f_i^\alpha(x_\alpha, y_\alpha) = -1 \text{ si } x_\alpha = i, y_\alpha \neq i \text{ ou } x_\alpha \neq i, y_\alpha = i ;$$

$$\text{Prob} \left\{ (x_\alpha = i, y_\alpha \neq i) \cup (x_\alpha \neq i, y_\alpha = i) \right\} = \frac{2 E_i^\alpha e_i^\alpha}{s^2}$$

$$\text{On vérifie que } \left(\frac{E_i^\alpha}{s} \right)^2 + \left(\frac{e_i^\alpha}{s} \right)^2 + \frac{2 E_i^\alpha e_i^\alpha}{s^2} = \left(\frac{E_i^\alpha + e_i^\alpha}{s} \right)^2 = 1.$$

Moyenne

L'espérance mathématique de la variable aléatoire f_i^α se calcule simplement :

$$E \left\{ f_i^\alpha \right\} = \frac{e_i^\alpha}{E_i^\alpha} \cdot \left(\frac{E_i^\alpha}{s} \right)^2 + \frac{E_i^\alpha}{e_i^\alpha} \left(\frac{e_i^\alpha}{s} \right)^2 - \frac{2 E_i^\alpha e_i^\alpha}{s^2}$$

$$= \frac{E_i^\alpha e_i^\alpha}{s^2} + \frac{E_i^\alpha e_i^\alpha}{s^2} - \frac{2 E_i^\alpha e_i^\alpha}{s^2} = 0.$$

On considère la variable aléatoire $s^\alpha = \sum_i f_i^\alpha$

$$E \left\{ s^\alpha \right\} = \sum_i E \left\{ f_i^\alpha \right\} = 0.$$

L'espérance mathématique de s^α est donc nulle ; si on considère $S(x, y) = \frac{1}{n} \sum_\alpha s^\alpha$ l'espérance mathématique de $S(x, y)$, qui est l'indice de similarité de Smirnov défini ci-dessus, est nulle.

Variance

La variance de $S(x,y)$ se calcule moins simplement.

$$S(x,y) = \frac{1}{n} \sum_{\alpha} s^{\alpha}$$

$$\sigma_s^2 = E \left\{ \left[(S(x,y) - E \{ S(x,y) \})^2 \right] \right\} = E \left\{ [S(x,y)]^2 \right\}$$

$$\sigma_s^2 = \frac{1}{n^2} E \left\{ \left(\sum_{\alpha} s^{\alpha} \right)^2 \right\} = \frac{1}{n^2} \left[\sum_{\alpha} E \left\{ (s^{\alpha})^2 \right\} + \sum_{\alpha \neq \beta} E \left\{ s^{\alpha} s^{\beta} \right\} \right]$$

mais les caractères α et β sont indépendants donc

$$\sum_{\alpha \neq \beta} E \left\{ s^{\alpha} s^{\beta} \right\} = \sum_{\alpha \neq \beta} E \left\{ s^{\alpha} \right\} \cdot E \left\{ s^{\beta} \right\} = 0.$$

reste à calculer $E \left\{ (s^{\alpha})^2 \right\}$

$$(s^{\alpha})^2 = \left(\sum_i f_i^{\alpha} \right)^2 = \sum_i (f_i^{\alpha})^2 + \sum_{i \neq j} f_i^{\alpha} f_j^{\alpha}$$

mais les f_i^{α} ne sont pas indépendants.

Le tableau VI résume les différentes valeurs des f_i, f_j lorsqu'on considère 2 états i et j .

$$\begin{aligned} \sum \text{Prob} &= \frac{1}{s^2} (E_i^{\alpha 2} + E_j^{\alpha 2} + 2E_i^{\alpha} u_{ij}^{\alpha} + 2E_i^{\alpha} E_j^{\alpha} + 2E_j^{\alpha} u_{ij}^{\alpha} + u_{ij}^{\alpha 2}) \\ &= \frac{1}{s^2} (u_{ij}^{\alpha} + E_i^{\alpha} + E_j^{\alpha})^2 = \frac{s^2}{s^2} = 1 \end{aligned}$$

$E \left\{ f_i^{\alpha} f_j^{\alpha} \right\}$ somme de la sixième colonne du tableau VI.

$$\begin{aligned} E \left\{ f_i^{\alpha} f_j^{\alpha} \right\} &= \frac{E_i^{\alpha} E_j^{\alpha}}{s^2} \left[\frac{e_i^{\alpha}}{e_j^{\alpha}} + 2 - 2 \frac{u_{ij}^{\alpha}}{e_j^{\alpha}} + \frac{e_j^{\alpha}}{e_i^{\alpha}} - 2 \frac{u_{ij}^{\alpha}}{e_i^{\alpha}} + \frac{u_{ij}^{\alpha}}{e_i^{\alpha} e_j^{\alpha}} \right] \\ &= \frac{E_i^{\alpha} E_j^{\alpha}}{e_i^{\alpha} e_j^{\alpha}} \cdot \frac{1}{s^2} \left[e_i^{\alpha 2} + 2e_i^{\alpha} e_j^{\alpha} - 2u_{ij}^{\alpha} e_i^{\alpha} + e_j^{\alpha 2} - 2u_{ij}^{\alpha} e_j^{\alpha} + u_{ij}^{\alpha 2} \right] \\ &= \frac{E_i^{\alpha} E_j^{\alpha}}{e_i^{\alpha} e_j^{\alpha}} \cdot \frac{1}{s^2} \left[e_i^{\alpha} + e_j^{\alpha} - u_{ij}^{\alpha} \right]^2 \\ &= \frac{E_i^{\alpha} E_j^{\alpha}}{e_i^{\alpha} e_j^{\alpha}} \end{aligned}$$

$$E \left\{ (f_i^{\alpha})^2 \right\} = \left(\frac{E_i^{\alpha}}{s} \right)^2 \left(\frac{e_i^{\alpha}}{E_i^{\alpha}} \right)^2 + \left(\frac{e_i^{\alpha}}{s} \right)^2 \left(\frac{E_i^{\alpha}}{e_i^{\alpha}} \right)^2 + \frac{2E_i^{\alpha} e_i^{\alpha}}{s^2}$$

Tableau .N° VI

x_d	y_d	Prob.	f_i^α	f_j^α	Prob ($f_i^\alpha f_j^\alpha$)
i	i	$\frac{E_i^{\alpha 2}}{\Delta 2}$	$\frac{E_i^\alpha}{E_i^\alpha}$	$\frac{E_j^\alpha}{f_j^\alpha}$	$\frac{E_i^\alpha E_j^\alpha}{\Delta 2} \frac{E_i^\alpha}{f_j^\alpha}$
i	j	$\frac{E_i^\alpha E_j^\alpha}{\Delta 2}$	-1	-1	$\frac{E_i^\alpha E_j^\alpha}{\Delta 2}$
i	$\neq i, j$	$\frac{E_i^\alpha u_{ij}^\alpha}{\Delta 2}$	-1	$\frac{E_j^\alpha}{f_j^\alpha}$	$-\frac{E_i^\alpha E_j^\alpha u_{ij}^\alpha}{\Delta 2 f_j^\alpha}$
j	i	$\frac{E_i^\alpha E_j^\alpha}{\Delta 2}$	-1	-1	$\frac{E_i^\alpha E_j^\alpha}{\Delta 2}$
j	j	$\frac{E_j^{\alpha 2}}{\Delta 2}$	$\frac{E_i^\alpha}{E_i^\alpha}$	$\frac{f_j^\alpha}{E_j^\alpha}$	$\frac{E_i^\alpha E_j^\alpha}{\Delta 2} \frac{f_j^\alpha}{E_i^\alpha}$
j	$\neq i, j$	$\frac{E_j^\alpha u_{ij}^\alpha}{\Delta 2}$	$\frac{E_i^\alpha}{E_i^\alpha}$	-1	$-\frac{E_i^\alpha E_j^\alpha u_{ij}^\alpha}{\Delta 2 E_i^\alpha}$
$\neq i, j$	i	$\frac{E_i^\alpha u_{ij}^\alpha}{\Delta 2}$	-1	$\frac{E_j^\alpha}{f_j^\alpha}$	$-\frac{E_i^\alpha E_j^\alpha u_{ij}^\alpha}{\Delta 2 f_j^\alpha}$
$\neq i, j$	j	$\frac{E_j^\alpha u_{ij}^\alpha}{\Delta 2}$	$\frac{E_i^\alpha}{E_i^\alpha}$	-1	$-\frac{E_i^\alpha E_j^\alpha u_{ij}^\alpha}{\Delta 2 E_i^\alpha}$
$\neq i, j$	$\neq i, j$	$\frac{u_{ij}^{\alpha 2}}{\Delta 2}$	$\frac{E_i^\alpha}{E_i^\alpha}$	$\frac{E_j^\alpha}{f_j^\alpha}$	$\frac{E_i^\alpha E_j^\alpha}{\Delta 2} \frac{u_{ij}^{\alpha 2}}{E_i^\alpha f_j^\alpha}$
Total .		1			$\frac{E_i^\alpha E_j^\alpha}{E_i^\alpha f_j^\alpha}$

$$u_{ij}^\alpha = \Delta - E_i^\alpha - E_j^\alpha$$

$$= \frac{1}{s^2} (e_i^\alpha + E_i^\alpha)^2 = 1.$$

On en déduit que

$$E \left\{ (s^\alpha)^2 \right\} = e^\alpha + \sum_{i \neq j} \frac{E_i^\alpha E_j^\alpha}{e_i^\alpha e_j^\alpha}$$

$$\text{et } \sigma_s^2 = \frac{1}{n^2} \sum_{\alpha} \left(e^\alpha + \sum_{i \neq j} \frac{E_i^\alpha E_j^\alpha}{e_i^\alpha e_j^\alpha} \right)$$

$$\sigma_s^2 = \frac{1}{n} + \frac{1}{n^2} \sum_{\alpha} \sum_{i \neq j} \frac{E_i^\alpha E_j^\alpha}{e_i^\alpha e_j^\alpha}$$

Cas particulier des caractères binaires.

$$i = 1, 2. \quad E_1^\alpha + e_1^\alpha = s \quad \text{et} \quad E_1^\alpha + E_2^\alpha = s$$

$$E_2^\alpha + e_2^\alpha = s$$

$$\text{donc } E_2^\alpha = e_1^\alpha \quad e_2^\alpha = E_1^\alpha$$

$$s^2 = \frac{1}{n} + \frac{2}{n^2} \sum_{\alpha} 1 = \frac{1}{n} + \frac{2m}{n^2}$$

$$\text{mais } n = 2m. \quad \text{donc } \sigma_s^2 = \frac{1}{2m} + \frac{2m}{4m^2} = \frac{1}{m}$$

m = nombre de caractères

$$\sigma_s^2 = \frac{1}{m}$$

4 - Indice de ROGERS et TANIMOTO.

Il se rapporte aux caractères à plusieurs états, pouvant ou non être ordonnés. La formule de ROGERS et TANIMOTO (1960) généralise le cas des caractères à deux états l'indice de similarité est égal au rapport du nombre de caractères ayant même état pour les deux variétés au nombre d'états représentés par ses deux variétés pour tous les caractères.

D - Regroupement.1 - Dendrogrammes.

La matrice de similarité étant calculée, soit avec l'indice de SMIRNOV soit avec l'indice de ROGERS et TANIMOTO le problème qui reste à résoudre est celui de la formation des groupes. La construction du dendrogramme ne conduit pas effectivement à la formation des groupes mais à l'établissement d'une structure hiérarchisée des différents clones. Des groupes pourraient être formés si on pouvait déterminer le niveau de signification de l'indice utilisé.

L'établissement du dendrogramme peut être fait par deux méthodes qui peuvent l'une et l'autre être pondérées Weighted, (unweighted) variable-groupe method et weighted (unweighted) pair group method. Dans la première méthode tous les groupements possibles deux à deux sont faits à chaque cycle. Dans la seconde seul le groupement des deux individus les plus ressemblants est effectué à chaque cycle. La matrice de similarité entre les groupes ainsi formés peut se faire avec ou sans pondération. Dans le premier cas on utilise la méthode de SPEARMAN pour le calcul du coefficient de corrélation entre groupe

$$r_{qQ} = \frac{\sum_{qQ} r_{ij}}{\sqrt{q + 2\Delta q} \sqrt{Q + 2\Delta Q}}$$

où r_{qQ} est le coefficient de corrélation entre les groupes q et Q

\sum_{qQ} est la somme de tous les coefficients de corrélation entre les membres d'un groupe avec l'autre.

Δq est la somme de tous les coefficients de corrélation du premier groupe, ΔQ du deuxième groupe.

q est le nombre d'unités du premier groupe Q celui du second.

Dans le second cas on utilise la moyenne arithmétique des coefficients entre les membres d'un groupe avec l'autre. La méthode avec pondération présente l'inconvénient dans certains cas de donner des "renversements". C'est à dire que l'on peut trouver des coefficients entre groupe supérieurs à tous les coefficients à l'intérieur des deux groupes, ce qui rend difficile la construction du dendrogramme.

2 - Méthode nodale de ROGERS et TANIMOTO.

Cette méthode mise au point pour l'indice de ROGERS et TANIMOTO peut s'appliquer à n'importe quel coefficient d'association ou indice de similarité.

Soit S_{ij} l'indice de similarité entre les clones i et j . On appelle indice de "typicalité" le nombre R_i d'unités taxonomiques ayant au moins un état commun avec la variété i . (Nombre de $S_{ij} \neq 0$ pour i donné). Soit t le nombre de taxa ; $R_i \leq t - 1$. Le taxon dont l'indice de typicalité R_i est le plus élevé est le plus "typique".

$$\text{On calcule } H_i = \prod_{j=1}^t S_{ij} \text{ pour } j \neq i \text{ et } S_{ij} > 0.$$

Le taxon i ayant la plus grande valeur R_i et la plus grande valeur H_i est considéré comme le plus "typique".

C'est le centroïde du système lorsque les coefficients sont en logarithme. Le taxon le plus typique est appelé premier noeud de l'étude et autour on forme un groupe avec tous les taxa de forts coefficients de similarité avec lui. Le second noeud choisi est le taxon suivant dans l'ordre de typicalité. Le rayon de groupe entourant le premier noeud doit être tel qu'il ne doit pas englober le second noeud.

On introduit maintenant la distance $d_{ij} = -\log_2 S_{ij}$ comprise entre 0 et l'infini (de $S_{ij} = 0$ à $S_{ij} = 1$).

$$-\log_2 H_i = \sum_{j=1}^{j=t} d_{ij} = \sum_{j=1}^{j=t} (-\log_2 S_{ij}) \quad j \neq i \quad S_{ij} > 0.$$

$d_{1,2}$ est la distance entre le premier et le second noeud. Toute unité dont la distance au premier noeud est inférieure à $d_{1,2}$ est lié au premier noeud. On forme ainsi un groupe dont on teste l'homogénéité par le calcul de la quantité $u_n [(dij)]$

$u_n [(dij)]$ est une mesure de la non-homogénéité du groupe

$$u_n [(dij)] = 1 - \frac{E_n(dij)}{E_n(g,h)}$$

n nombre d'unités dans le groupe en plus du noeud primaire.

dij matrice des distances entre clones dans le groupe

g nombre de zéros dans (dij) au-dessus de la diagonale principale.

h nombre d'éléments infinis dans (dij) au-dessus de la diagonale principale.

$$E_n(g,h) = \log_2 \left[\frac{n-g}{2} (n-g-1) - h \right]$$

$$E_n(dij) = - \frac{1}{2} \sum'_{ij} \frac{dij}{T_n [(dij)]} \log_2 \frac{dij}{T_n [(dij)]}$$

\sum' sommation sur les termes finis.

$$T_n [(dij)] = \frac{1}{2} \left(\sum'_{ij} dij \right).$$

Le calcul est effectué chaque fois qu'on ajoute un élément au groupe. S'il y a un saut dans cette valeur on supprime cette dernière variété. Les "limites naturelles" du groupe étant franchies.

Après la détermination de ce premier groupe on enlève tous ses éléments de l'étude et on recommence pour les groupes secondaires.

E - Résultats

1 - Matériel.

La figure 1 représente la carte de la prospection 1967 du Kenya. Les unités géographiques sont matérialisées en rouge.

K1. population polymorphe I contenant le *Panicum infestum* K83. A : population "Meru", précède la population polymorphe proprement dite B. C : population "Embu", suit la population B trois clones K70 K72 K74 appartiennent à A, B et C.

K2. population polymorphe II contenant le *Panicum trichochledum* K138 - population TENGERU

K3. populations de KOROGWE contenant le diploïde K189.

K4. populations de MARALAL populations dans une zone marginale de l'extension de l'espèce en Afrique de l'Est.

K5. population polymorphe III population de MOMBASA à VOI

K6. population de NAIROBI à NAMANGA - zone homogène qui fait la jonction entre le nord et le sud de la zone prospectée

K7. population de la basse côte humide de MOMBASA à LUNGA LUNGA:

K8. population de la savane à *Panicum* - MONTS PARE NORD

K9. population de BARON'S FALLS à MOMBO.

2 - Graphiques.

a) Représentation des matrices de similarité sur l'indice de SMIRNOV.

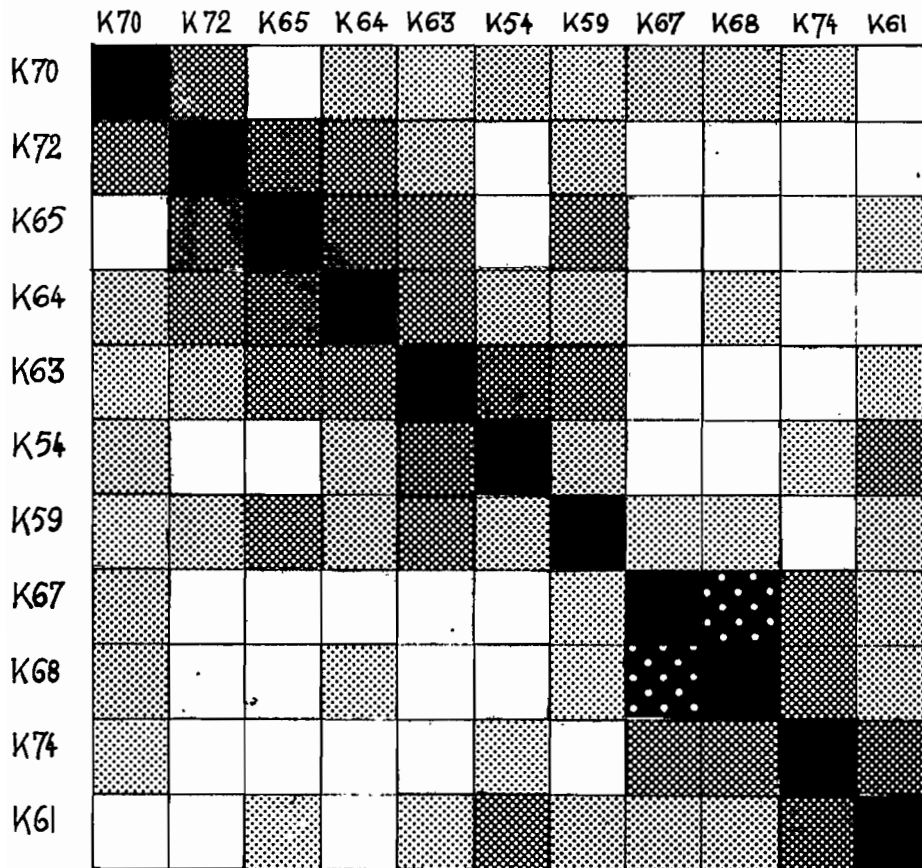
b) Dendrogrammes

Les dendrogrammes marqués S sont construits sur l'indice de SMIRNOV, ceux marqués RT sont construits sur l'indice de ROGERS et TANIMOTO.

c) Méthode nodale de ROGERS et TANIMOTO

Trois populations seulement sont représentables par cette méthode.

K1^A MERU.EMBU



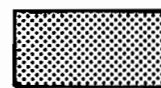
> + 0,40



> + 0,20



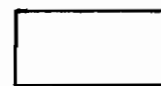
> 0



> - 0,10



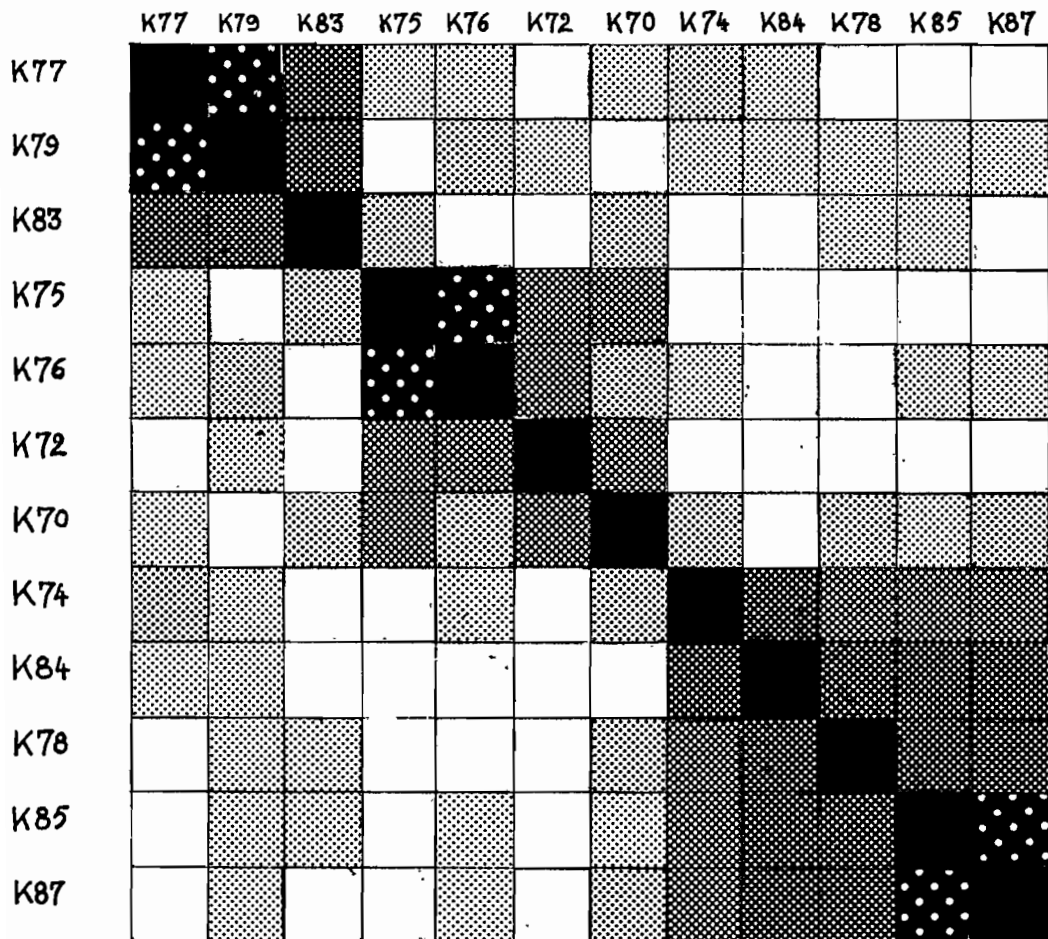
> - 0,20



> - 0,30

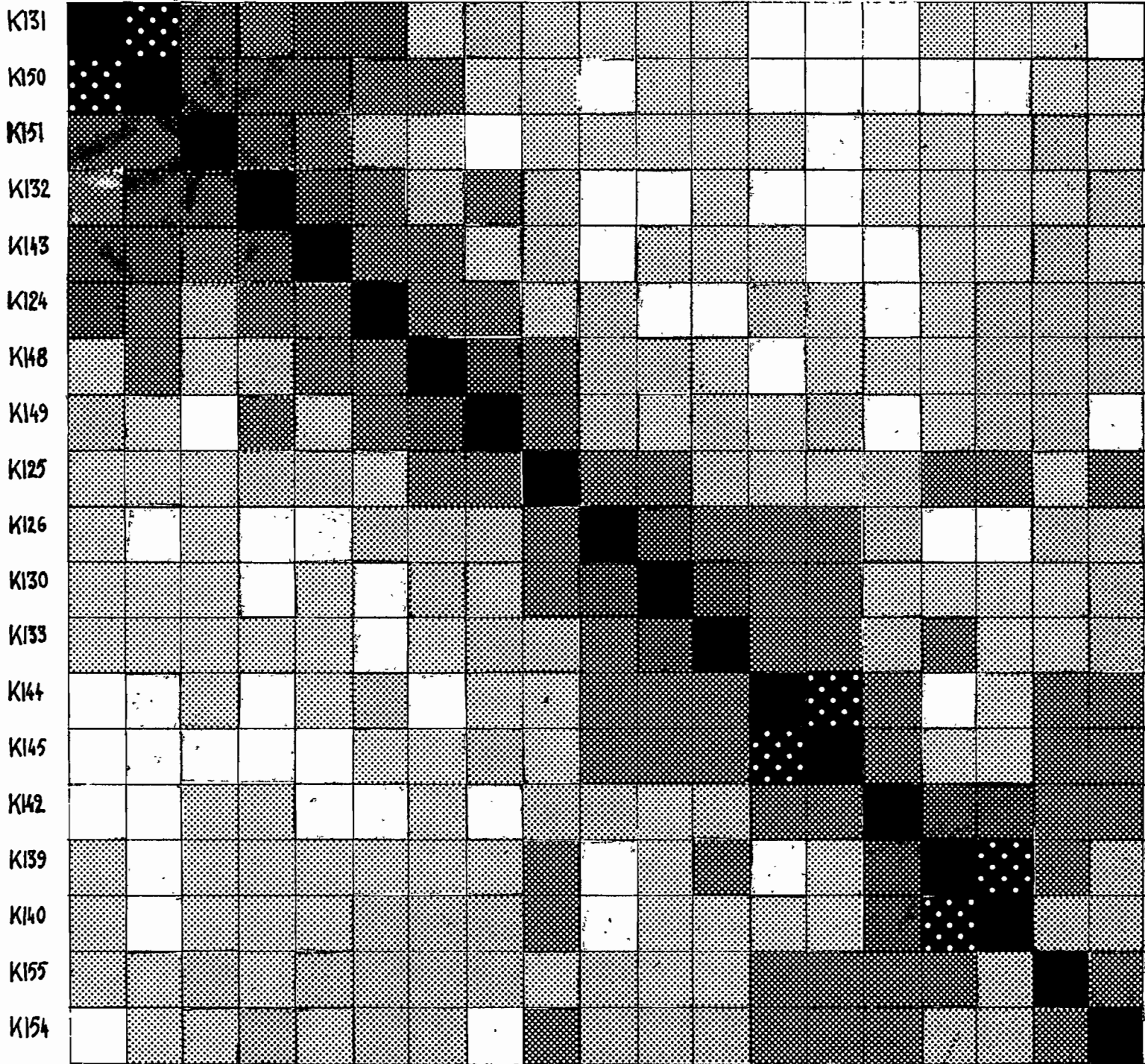
VALEURS DES INDICES DE SMIRNOV

K1^B MERU EMBU

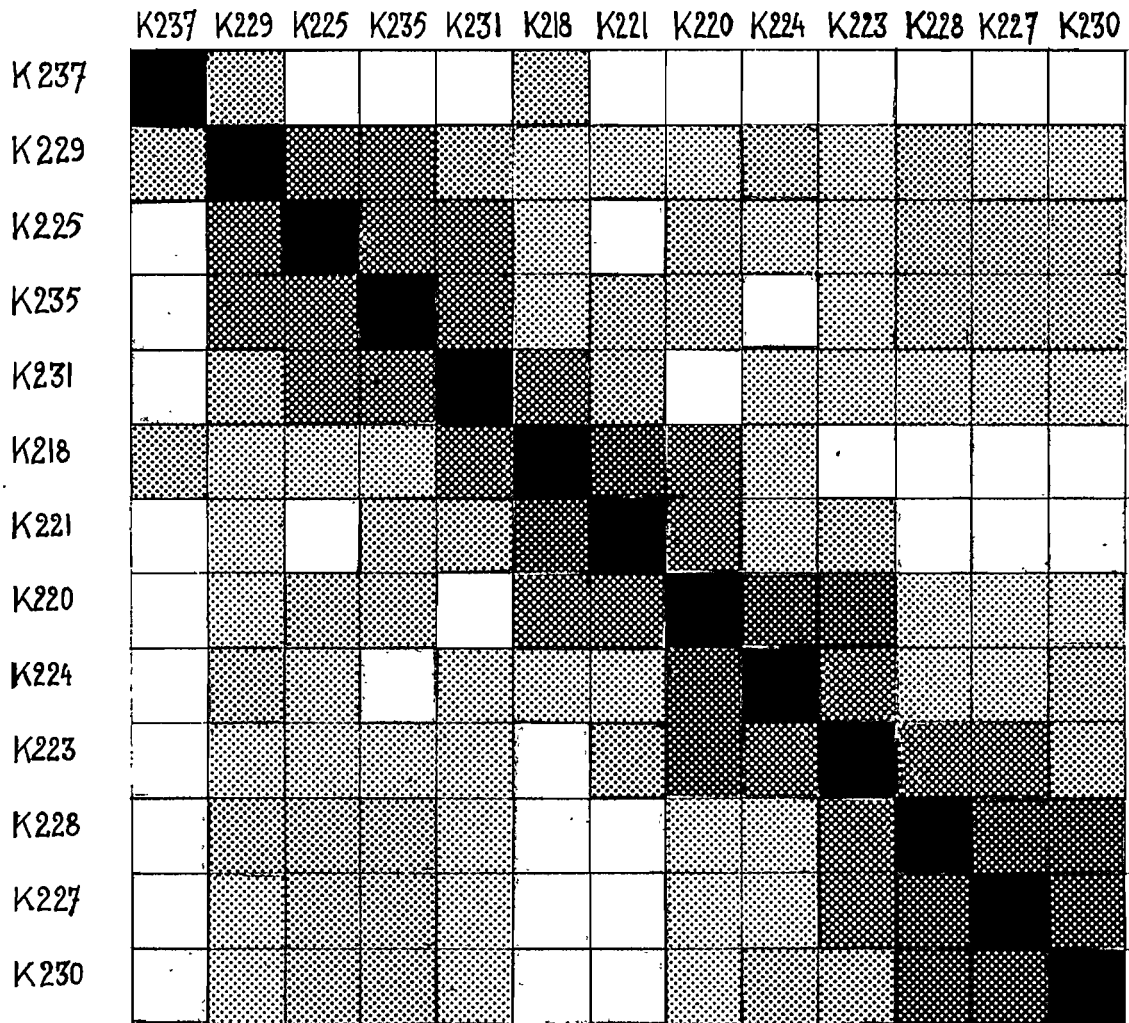


K2 Tengeru

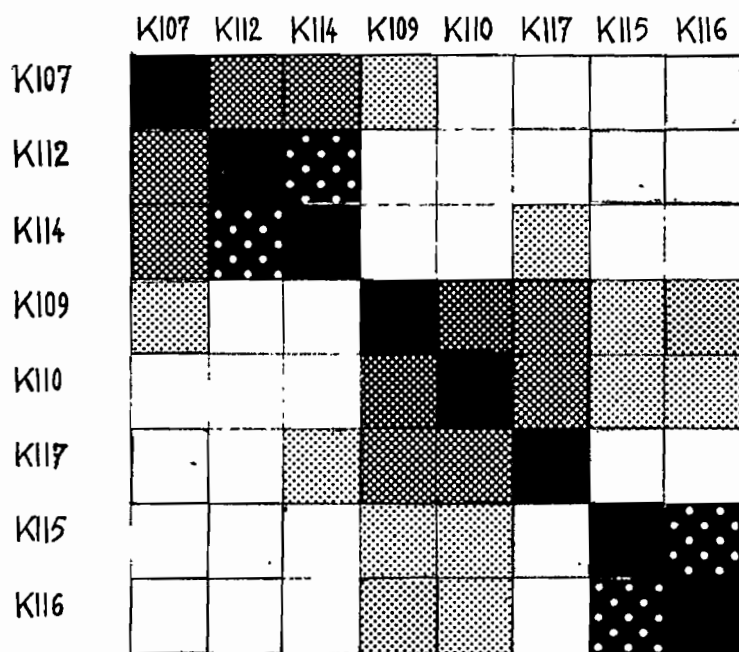
K131 K150 K151 K132 K143 K124 K148 K149 K125 K126 K130 K133 K144 K145 K142 K139 K140 K155 K154



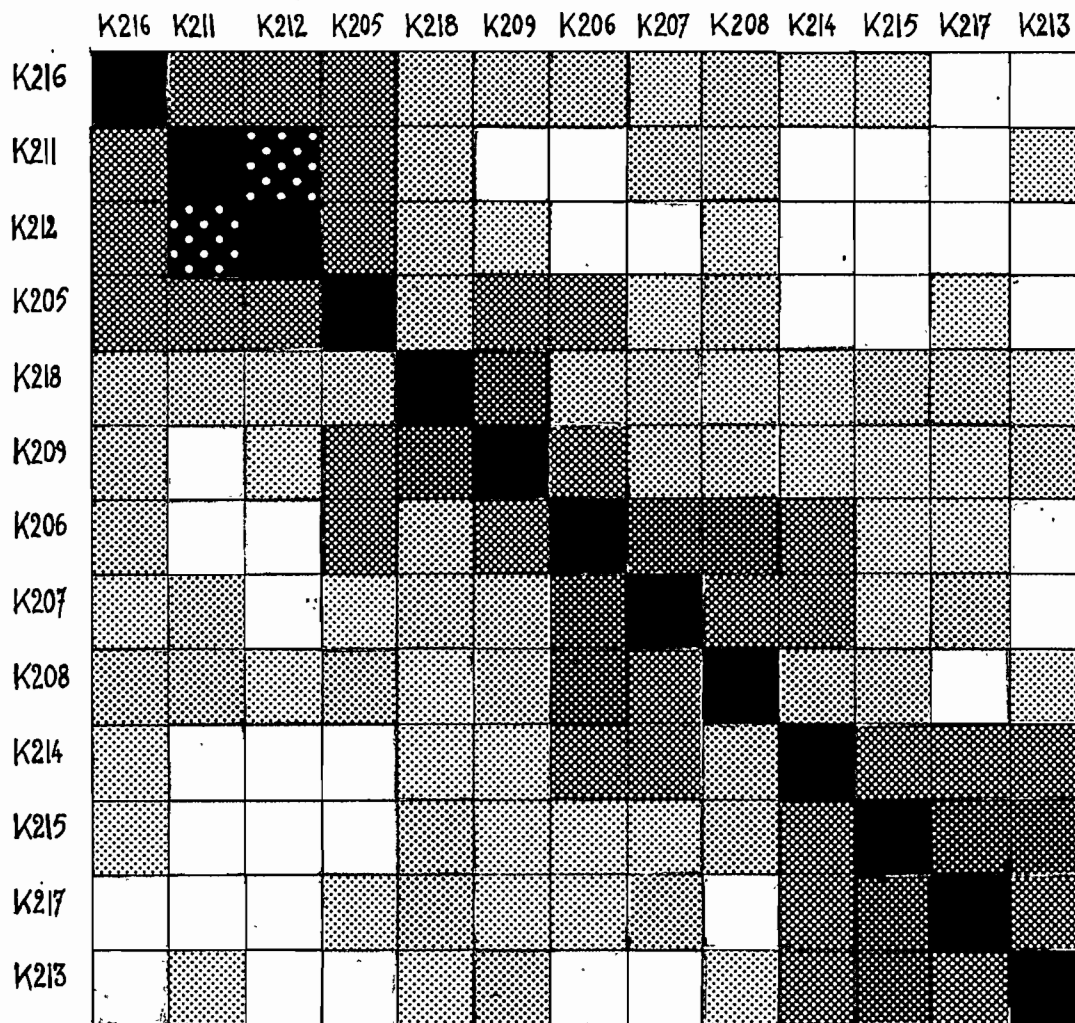
K5 - MOMBASA VOI



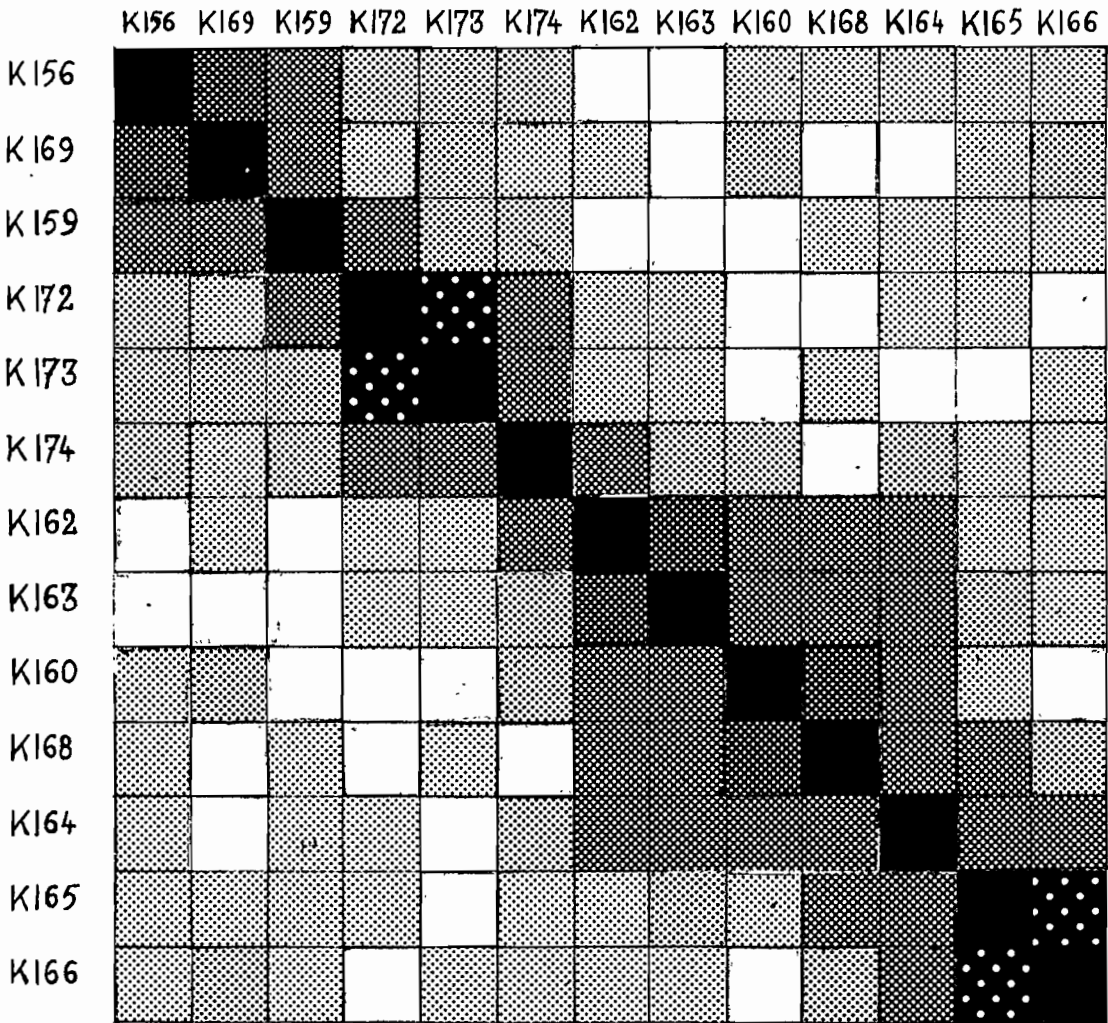
_ K6 _ DE NAIROBI A NAMANGA



- K7 MOMBASA LUNGA_LUNGA

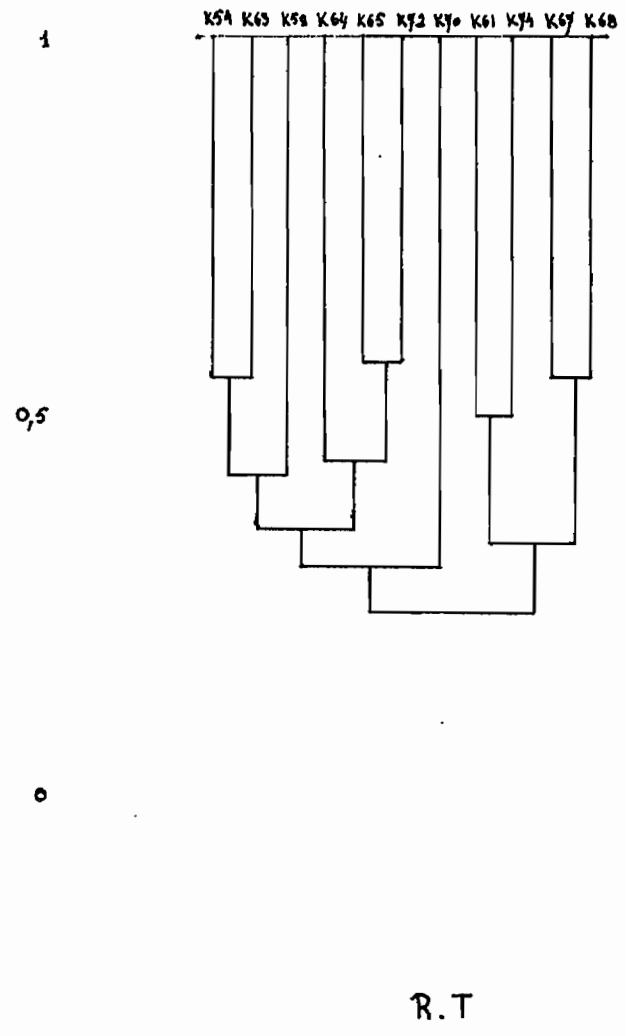
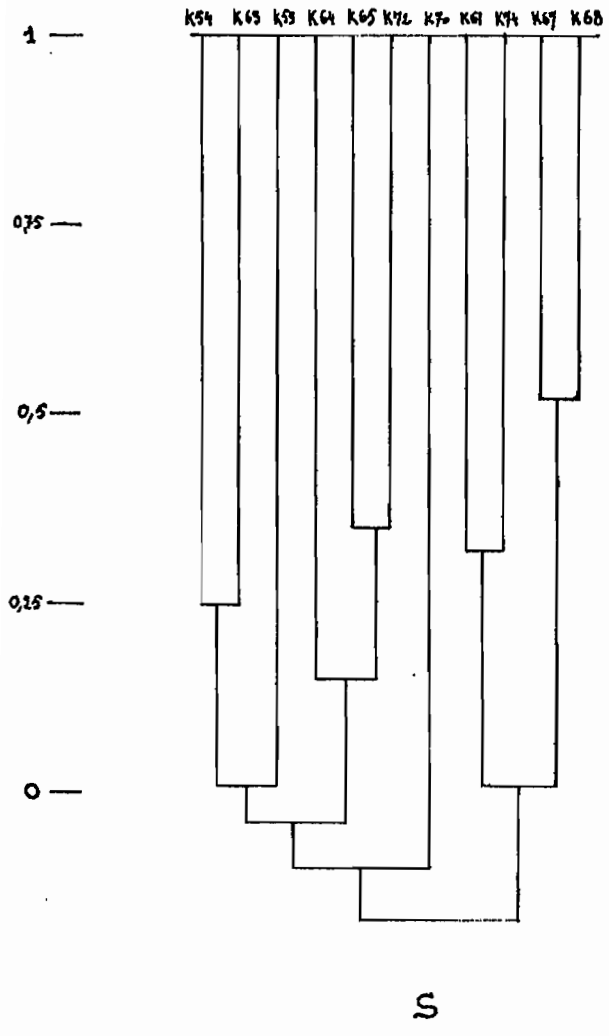


K8. MONTS PARE NORD

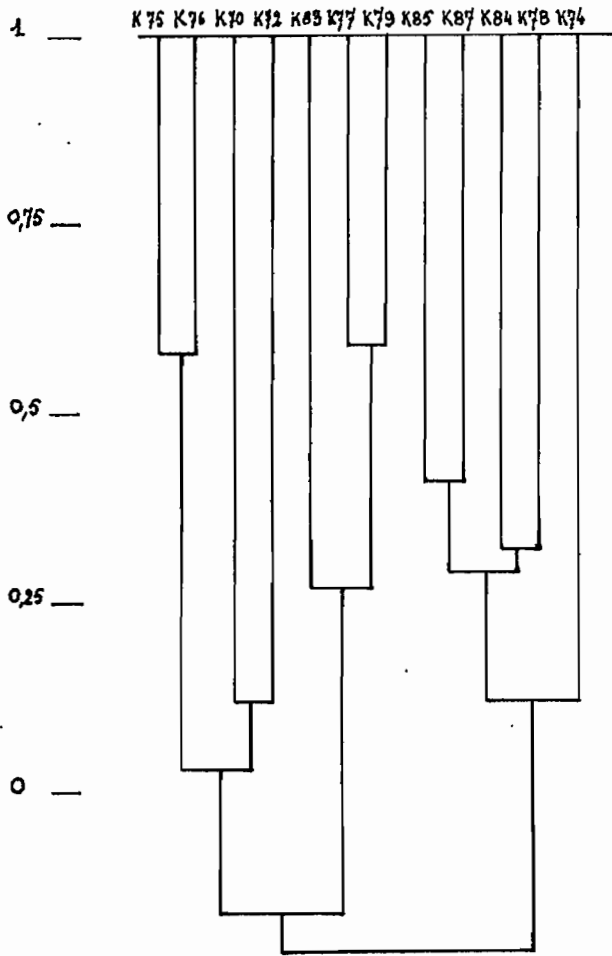


K9. DE BARON'S FALLS A MOMBO

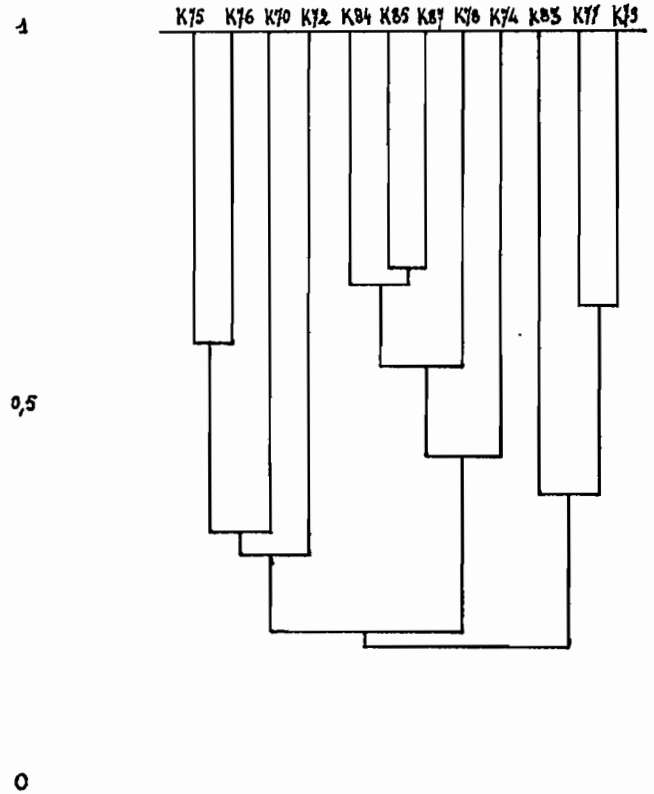
	K181	K182	K178	K175	K176	K177	K180	K183	K184	K185	K187
K181	Black	Dark	Dark	Dark	Dark	Dark	Dark	Dark	Dark	White	Dark
K182	Dark	Black	Dark	Dark	Dark	Dark	Dark	Dark	White	White	Dark
K178	Dark	Dark	Black	Dark	Dark	Dark	White	White	White	White	Dark
K175	Dark	Dark	Dark	Black	Dark	Dark	Dark	White	White	Dark	Dark
K176	Dark	Dark	Dark	Dark	Black	Dark	Dark	Dark	White	White	Dark
K177	Dark	Dark	Dark	Dark	Dark	Black	Dark	Dark	Dark	Dark	Dark
K180	Dark	Dark	White	Dark	Dark	Dark	Black	Dark	Dark	Dark	Dark
K183	Dark	Dark	White	White	Dark	Dark	Dark	Black	White	Dark	Dark
K184	Dark	White	White	White	Dark	Dark	Dark	White	Black	Dark	Dark
K185	White	White	White	Dark	White	Dark	Dark	Dark	Dark	Black	Dark
K187	Dark	Dark	Dark	Dark	White	Dark	Dark	Dark	Dark	Dark	Black



K1A - MERU-EMBU

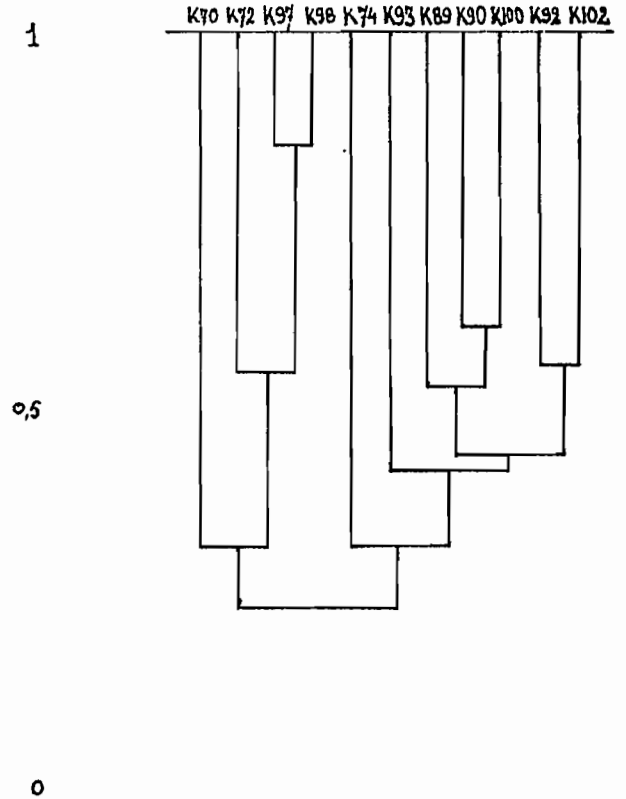
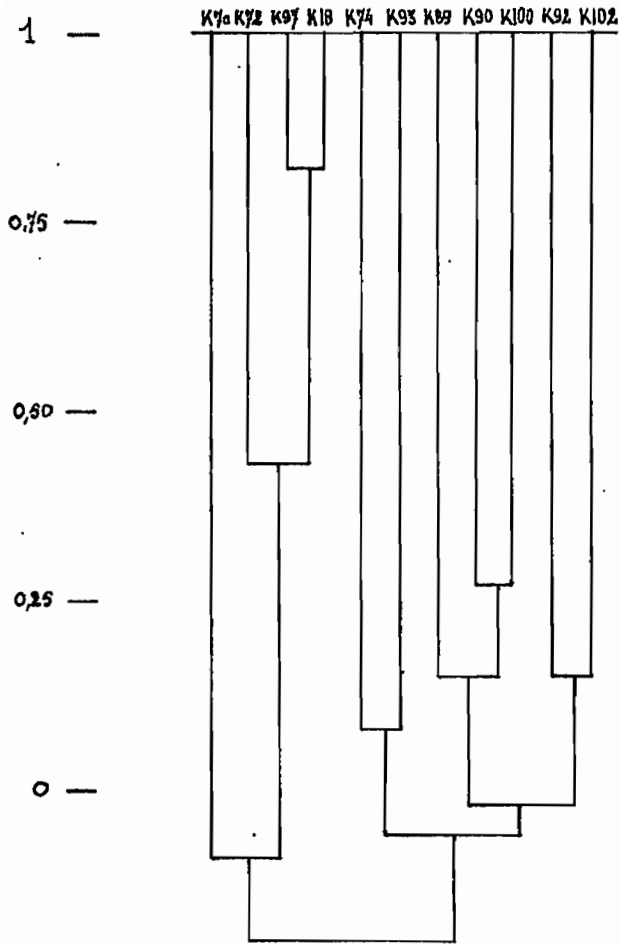


S



R.T

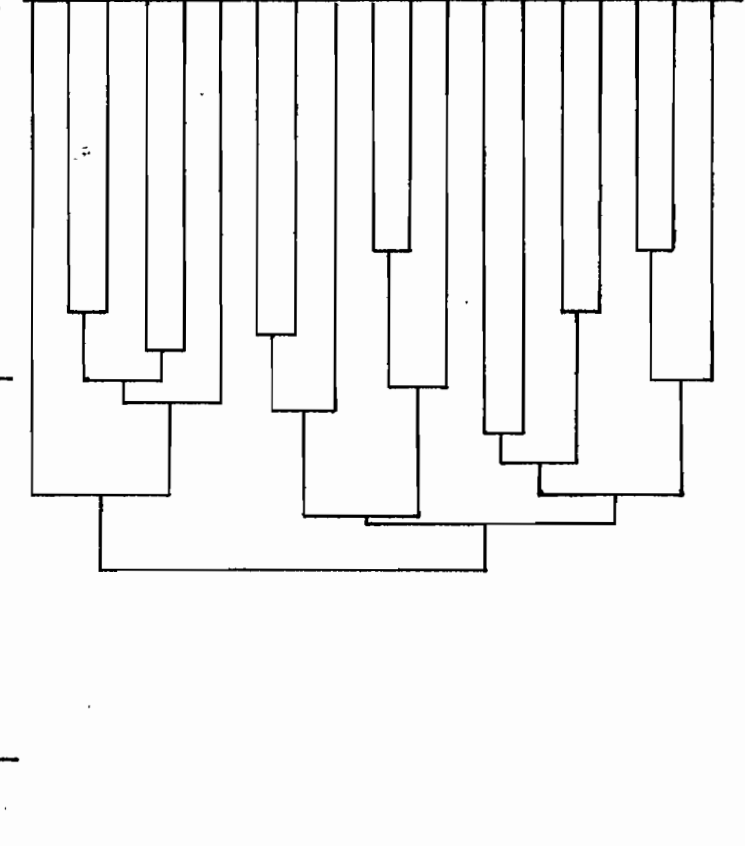
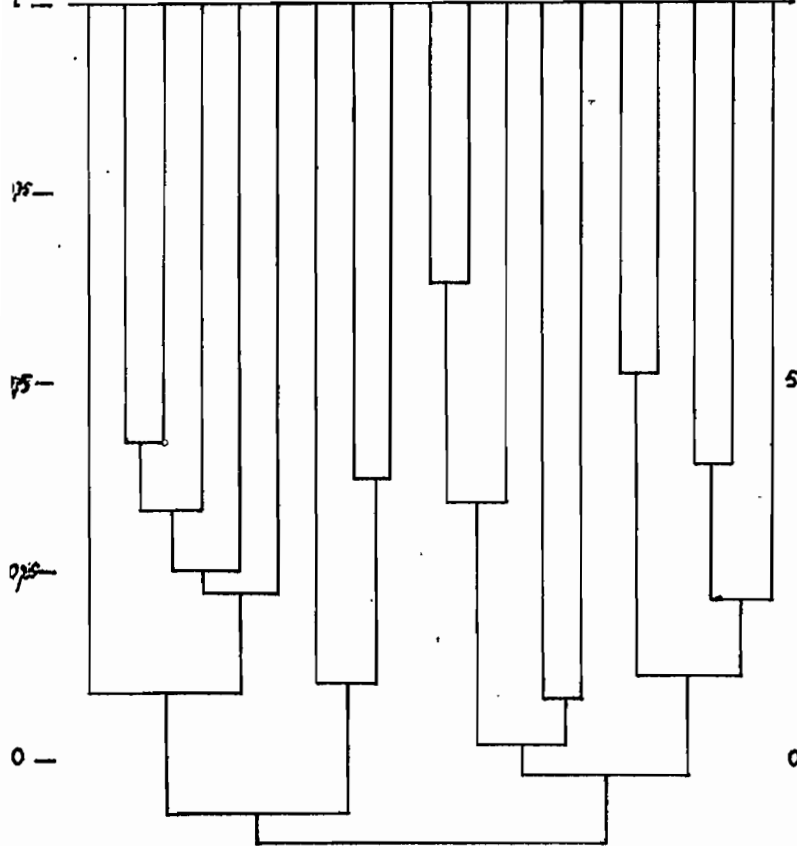
K1B . MERU-EMBU .



K.1C MERU-EMBU

K124 K131 K150 K151 K152 K143 K125 K148 K149 K139 K140 K142 K154 K155 K144 K145 K126 K130 K135

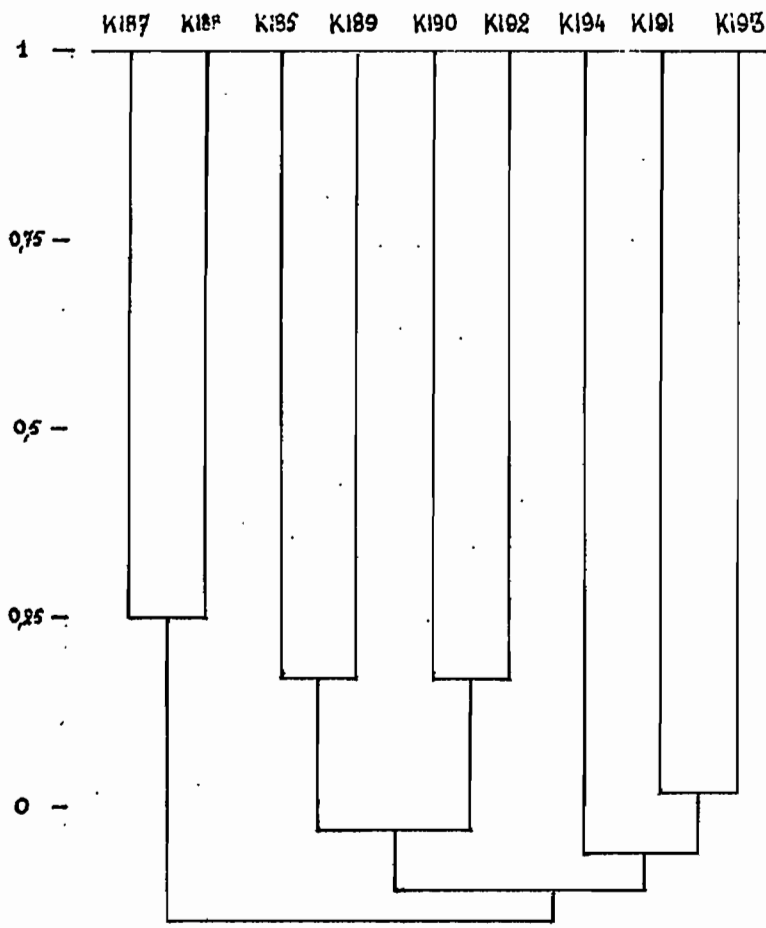
K124 K131 K150 K151 K142 K143 K125 K148 K149 K139 K140 K142 K154 K155 K144 K145 K126 K130 K135



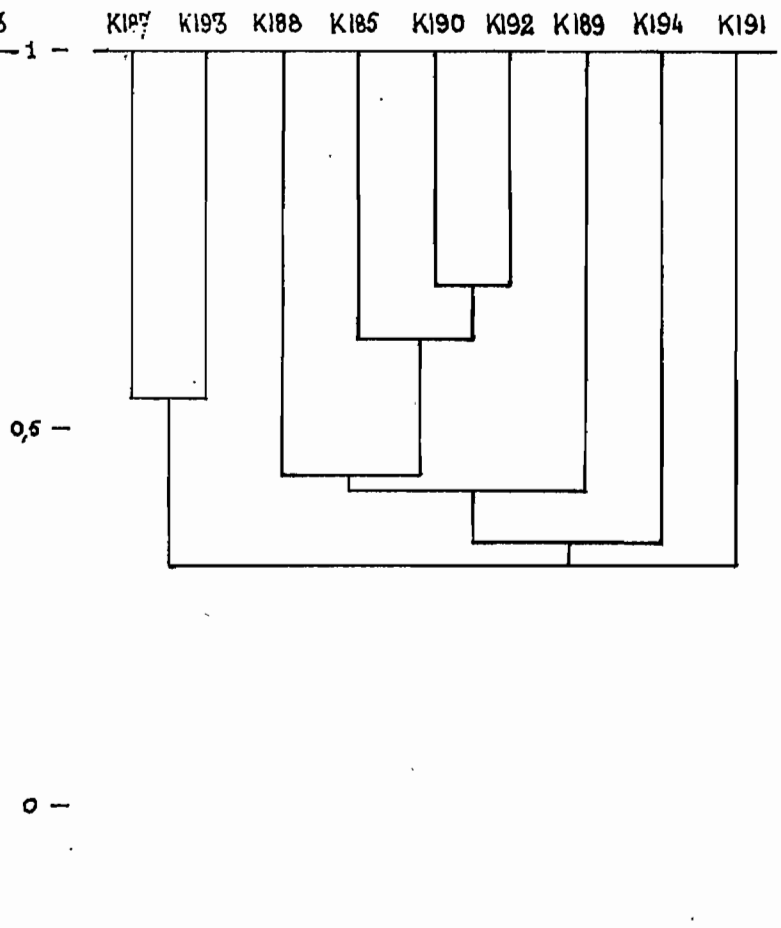
S

R.T

K2 . TENERU

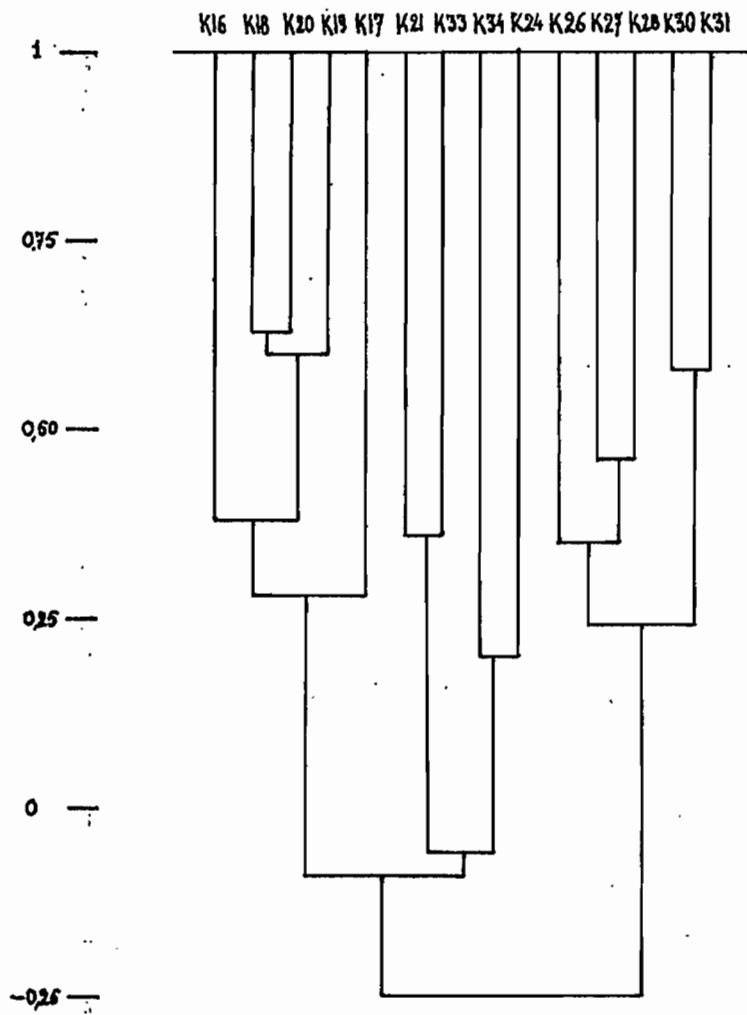


S

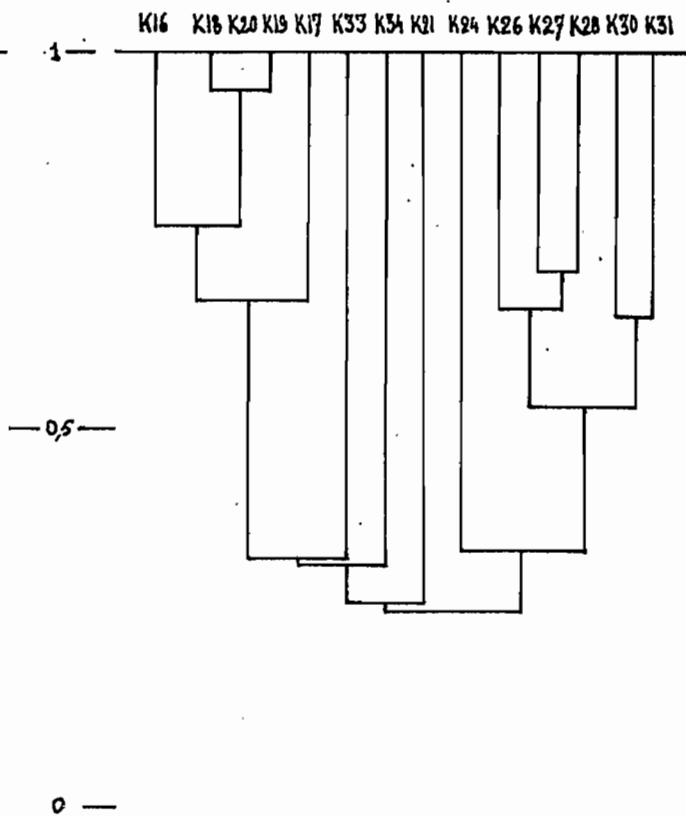


R.T

K3. KOROGWE

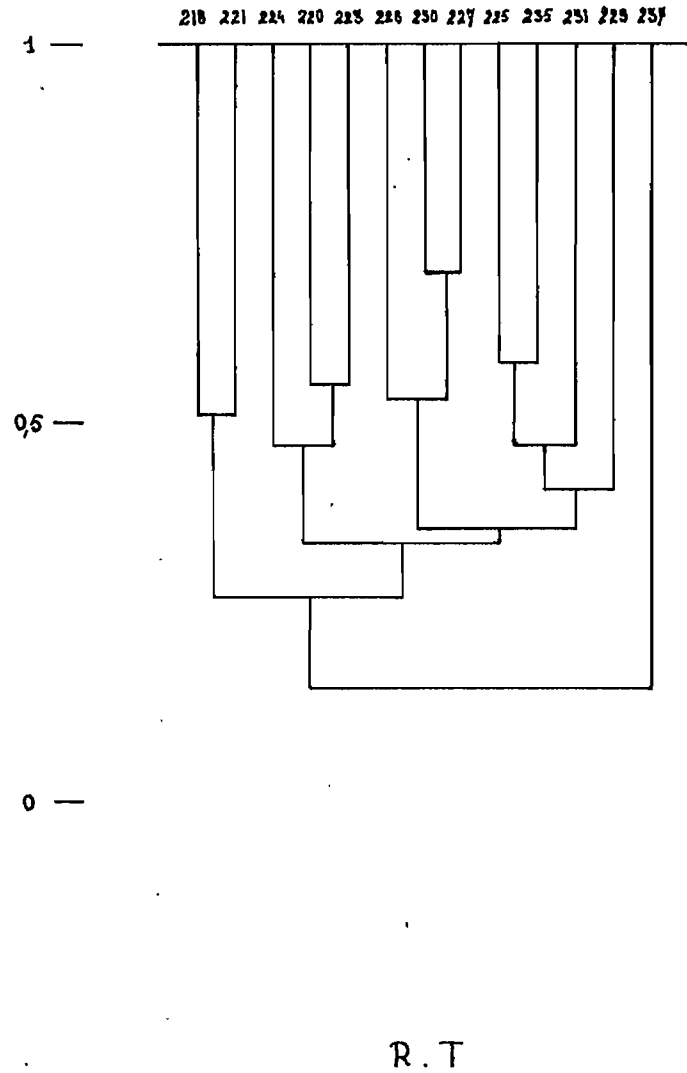
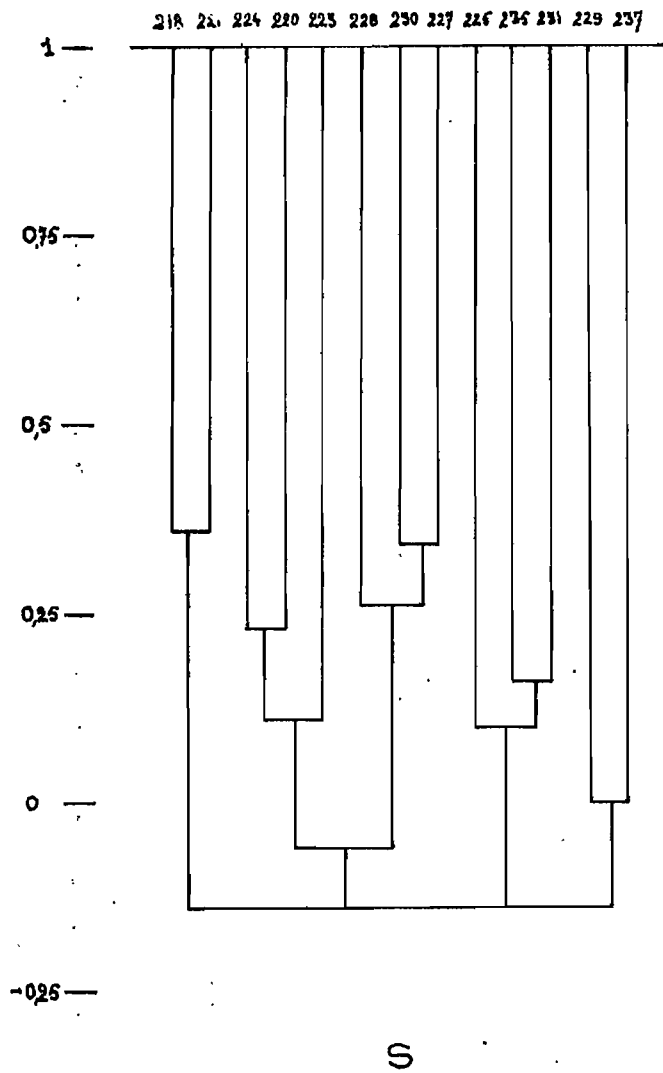


S

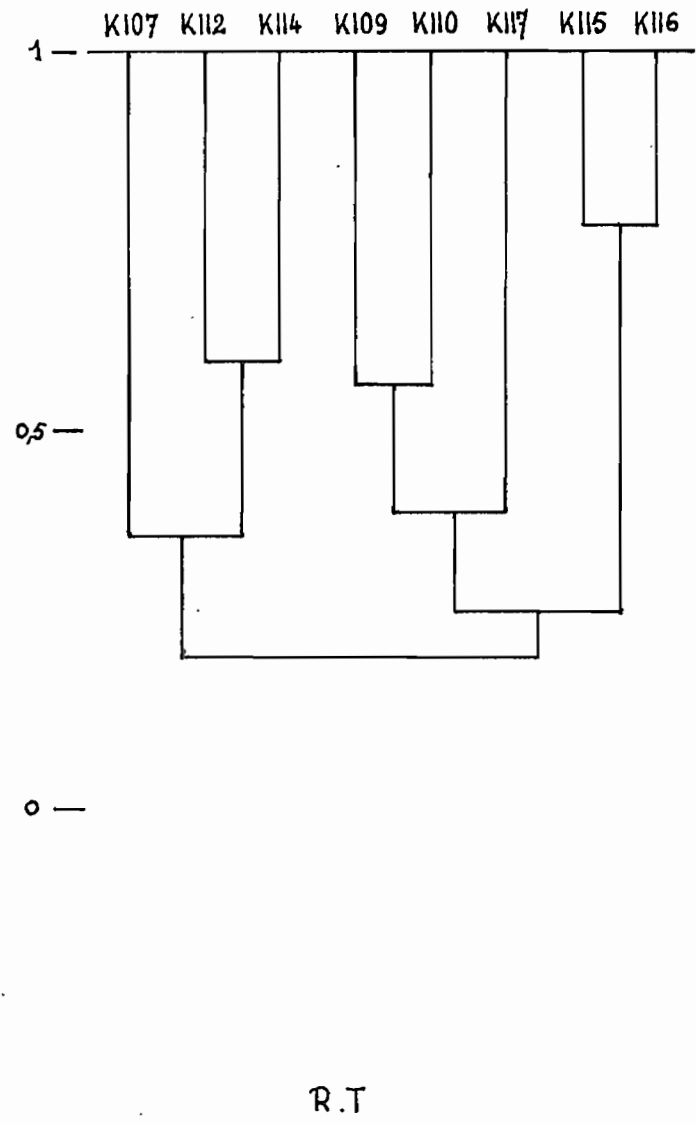
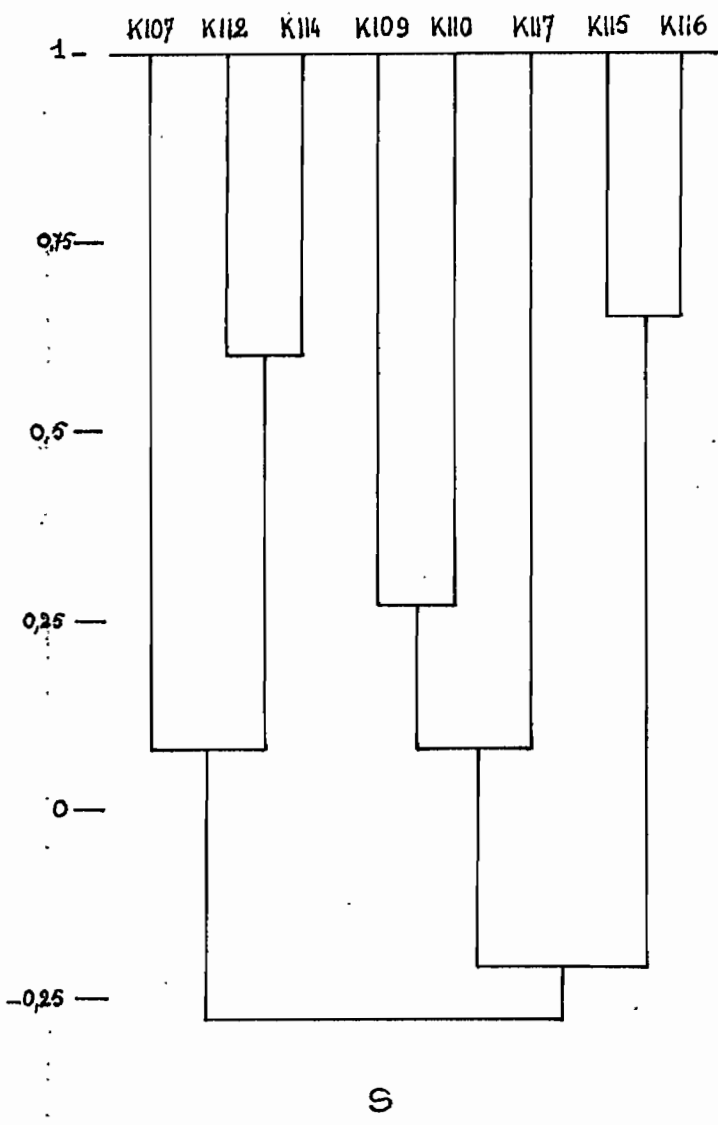


R.T

K4 - MARALAL -

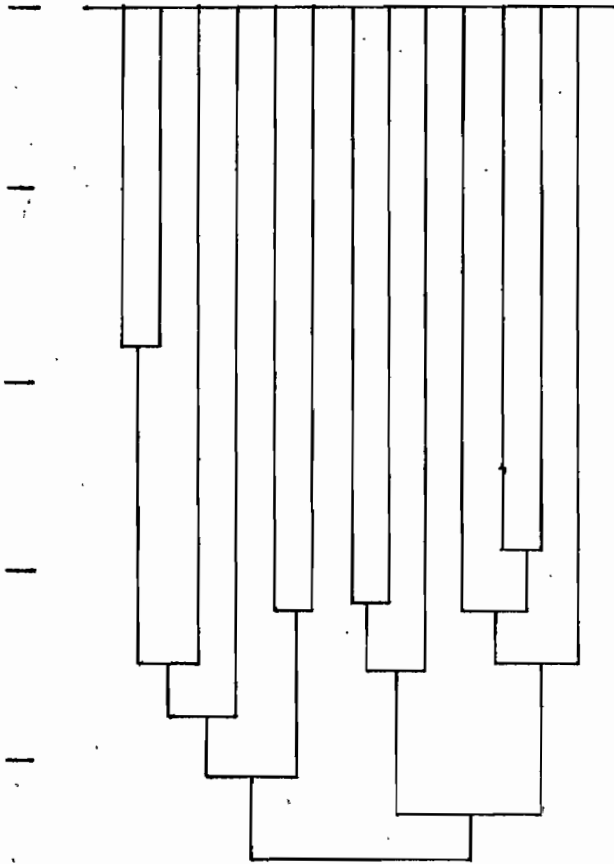


K5 - MOMBASA-VOI



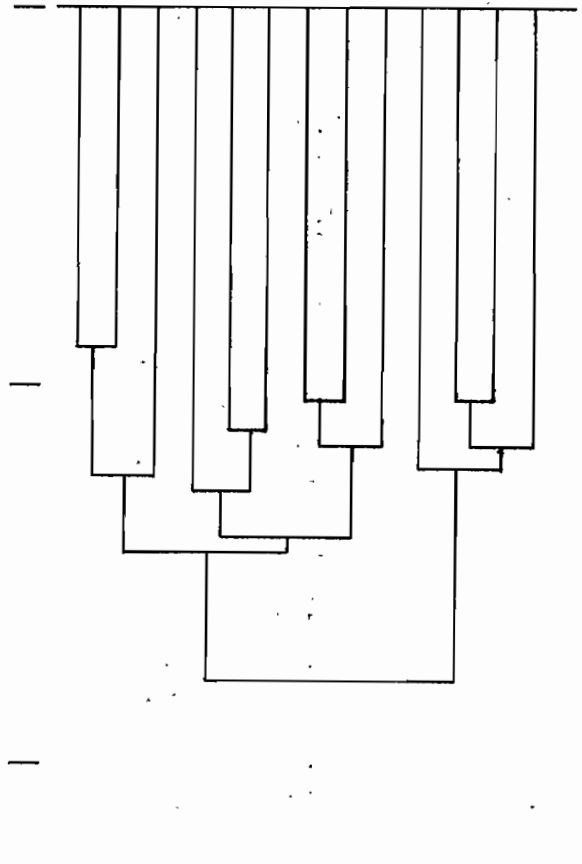
K6 .NAIROBI . NAMANGA

K201 K212 K216 K205 K205 K216 K206 K207 K208 K213 K215 K217 K214



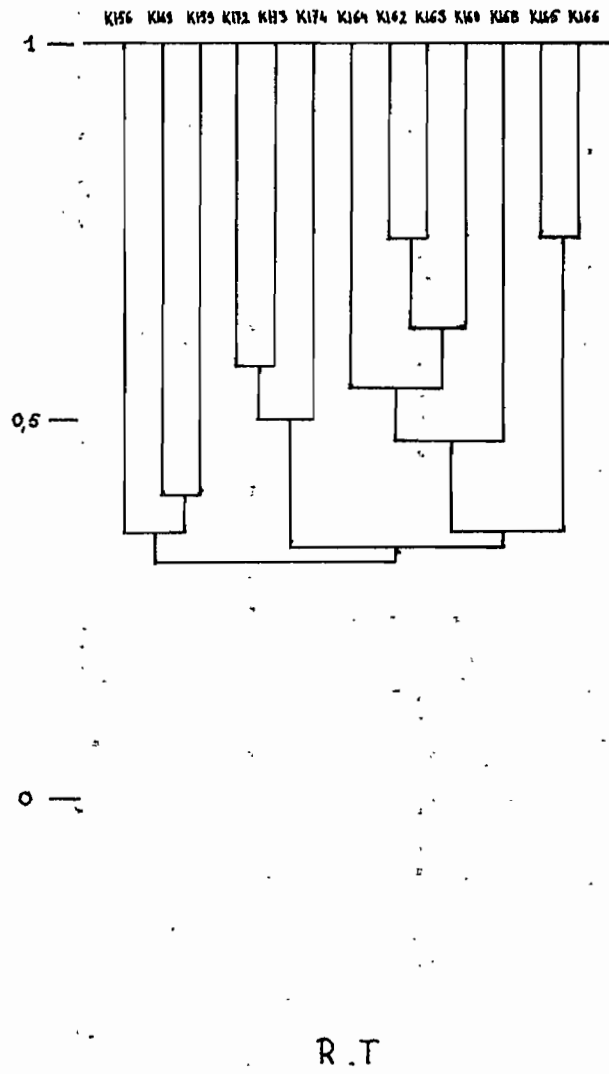
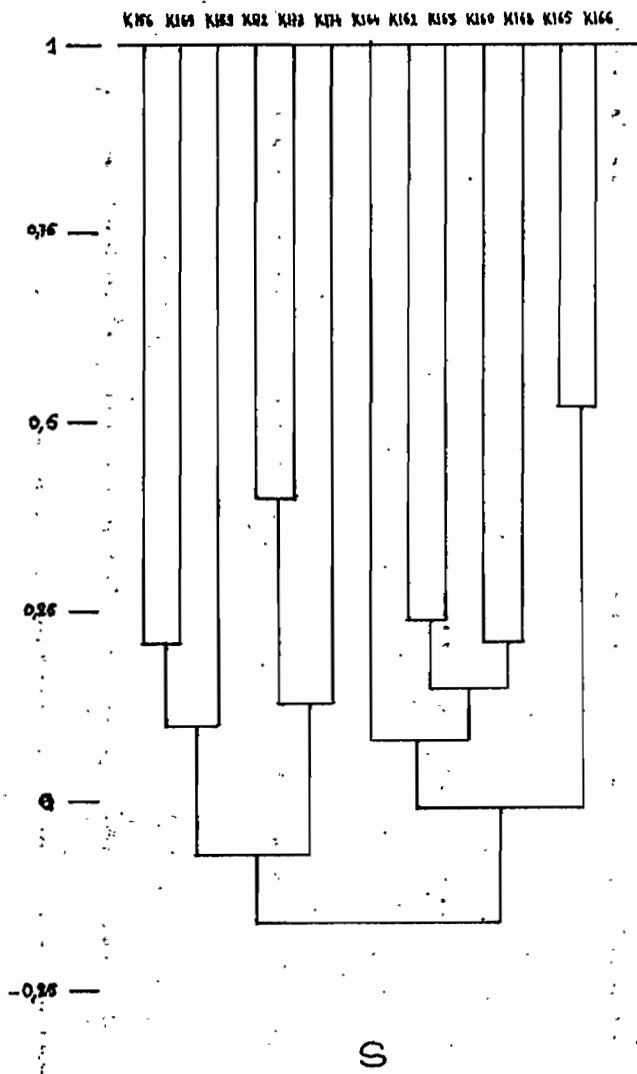
S

K201 K212 K216 K205 K205 K216 K206 K207 K208 K213 K215 K217 K214

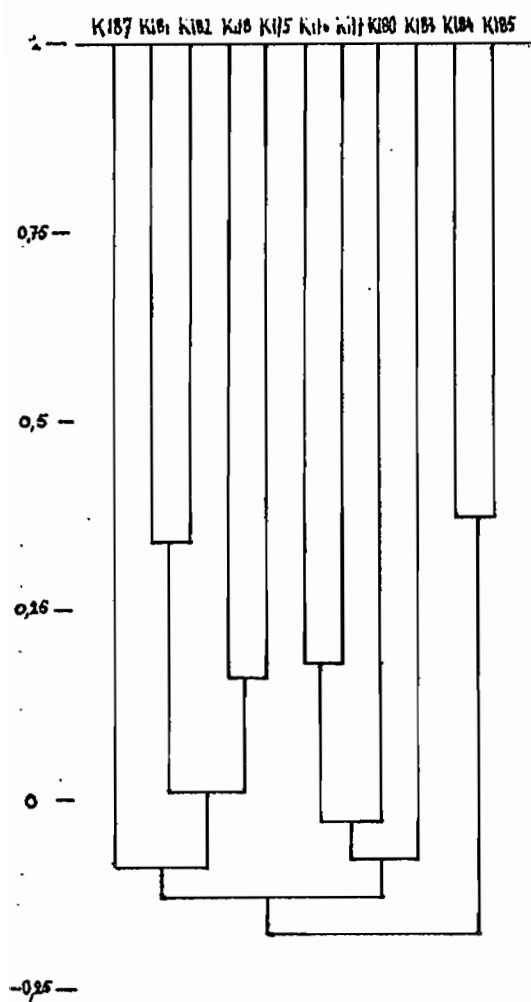


R.T

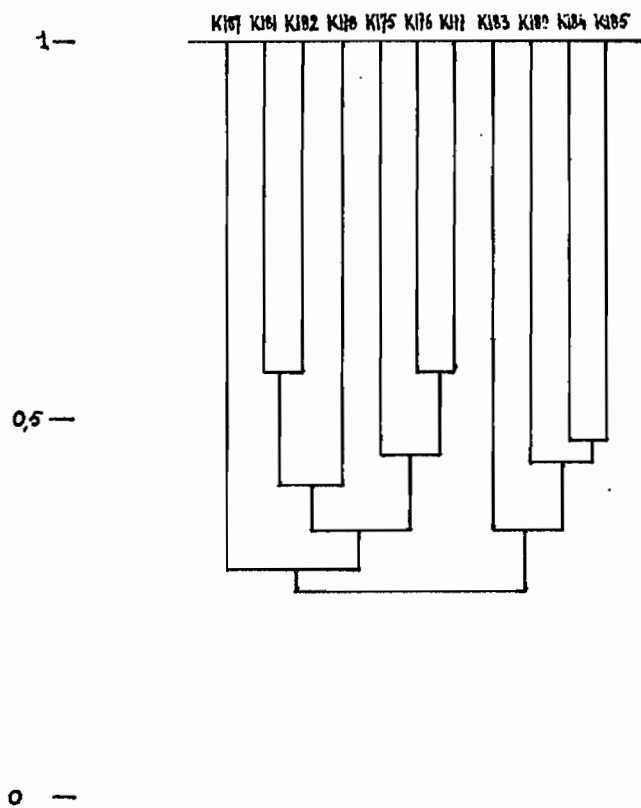
K7. MOMBASA LUNGA-LUNGA.



K8. MONTS PARE - NORD



S

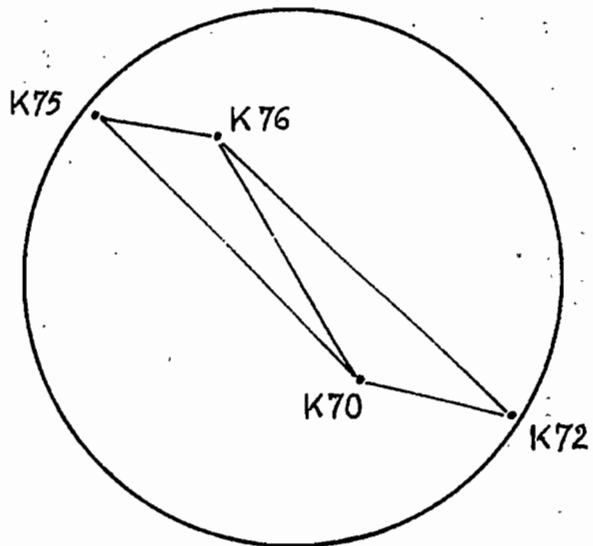
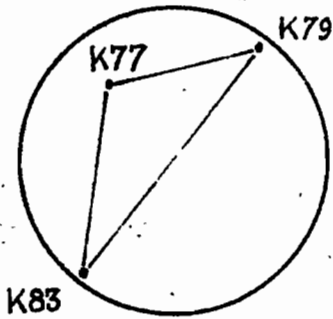
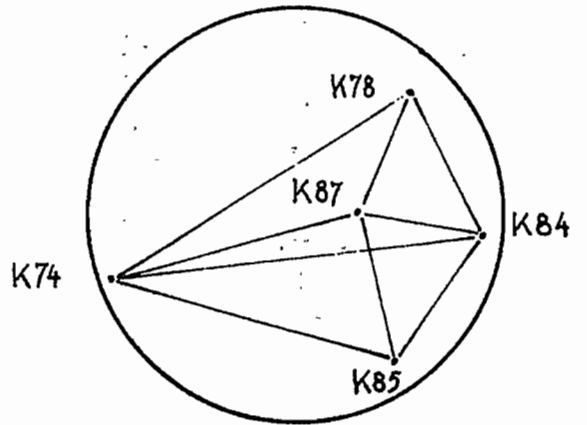


R.T

K9 De BARONS'S FALLS A MOMBO

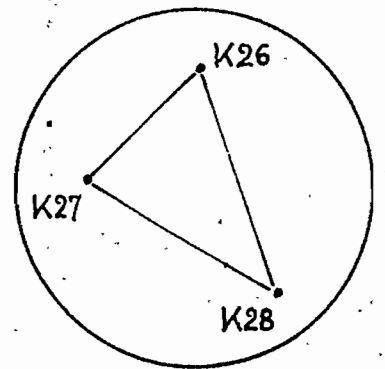
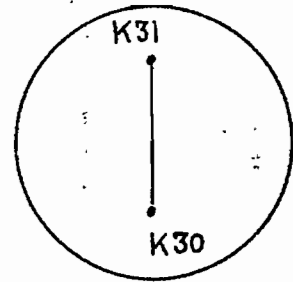
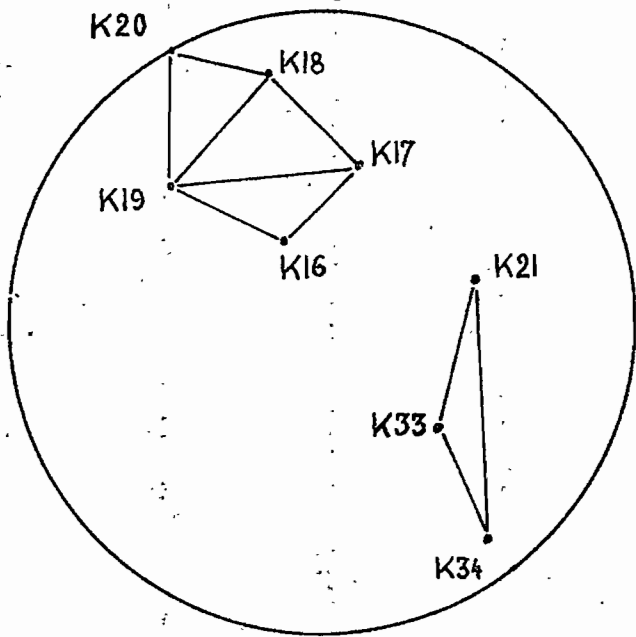
K1^B MERU-EMBU

METHODE NODALE DE ROGERS ET TANIMOTO



K4 MARALAL

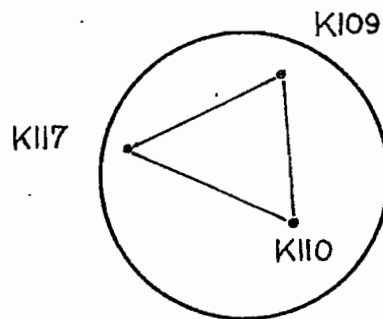
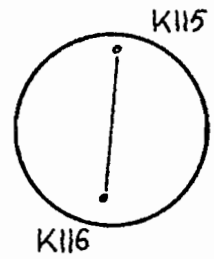
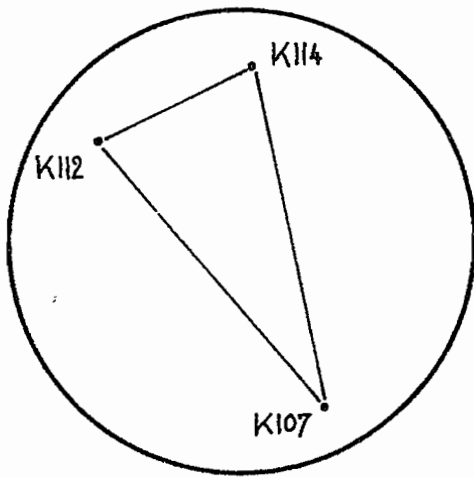
METHODE NODALE DE ROGERS ET TANIMOTO



K24

K6 DE NAIROBI A NAMANGA

METHODE NODALE DE ROGERS ET TANIMOTO



3 - Conclusions.

Les dendrogrammes basés sur les 2 indices ont été construits pour les 11 populations. Pour 2 population K1 A et K6 les deux indices donnent les mêmes liaisons et les mêmes embranchements. Pour trois populations seulement la méthode nodale est applicable et donne des groupes qui se referment :

a) K1B population polymorphe I - MERU EMBU

Trois groupes sont formés par la méthode nodale et correspondent à 3 embranchements du dendrogramme de SMIRNCV. Ils peuvent être caractérisés par la pilosité des feuilles.

I - K75 - K76 à feuilles non pileuses

II - K78-K84-K85-K87 à feuilles duveteuses

III - K77-K79-K83 - qui contient K83 *Panicum infestum* et deux hybrides.

Les trois clones voisins de la population polymorphe I se rattachent soit au groupe I, K70, K72 soit au groupe II, K74 ; K70 est voisin de K75 et K76 par la pilosité de la gaine (même densité et même aspect) ; K72 est voisin de K75 et K76 par l'absence de pilosité des feuilles.

K74 ressemble à K84 pour les caractères de gaines, à K85 pour les caractères inflorescences à K78 pour les épillets ce qui fait que dans l'ensemble K74 est assez proche de ce groupe II Le groupe III semble plus proche du groupe I que du groupe II ce qui pourrait faire penser que K77 et K79 sont hybrides entre *Panicum infestum* et un *Panicum maximum* de phénotype du groupe I.

b) K4. population autour de MARALAL, dans une zone marginale de l'extension de l'espèce en Afrique de l'Est.

On distingue très nettement 2 groupes.

I. K16 - K17 - K18 - K19 - K20 population proche de RUMBRUTI caractérisés par des gaines très pileuses des feuilles à poils longs des épillets avec soie.

II. K26 - K27 - K28 - K30 - K31. population proche de MARALAL gaines peu pileuses, feuilles à poils très longs, épillets sans soie. entre ces deux groupes se trouvent K21 - K33 - K34 qui ne peuvent se séparer du groupe I par la méthode nodale.

K24 est intermédiaire entre le groupe (K21 - K33 - K34) et le groupe II.

Le dendrogramme sur l'indice de SMIRNOV rattache K24 à K34 alors que le dendrogramme sur l'indice de ROGERS et TANIMOTO rattache K24 au groupe II.

c) K6. population de NAIROBI à NAMANGA zone homogène (AMBOSELI).

On trouve trois groupes très nettement séparés, aussi bien dans le dendrogramme sur l'indice de SMIRNOV, que sur le dendrogramme sur l'indice de ROGERS et TANIMOTO (qui est équivalent), que par la méthode nodale. Cette dernière renseigne sur le fait que les 3 groupes sont équidistants et de même homogénéité. Ces trois groupes se sont formés sur tout un ensemble de caractères sans qu'on puisse extraire de caractéristiques pour chacun des groupes.

I K107 - K112 - K114, II K109 - K110 - K117, III K115-K116.

Pour la population K2 Tengeru, population polymorphe II la méthode nodale ne donne pas de groupes fermés cependant le dendrogramme permet de caractériser 4 branches sur les caractères de pilosité. Il semble que cette population soit en plein cours d'évolution et que la différenciation ne soit pas encore faite.

- I. K131 - K132 - K143 - K150 - K151. gaines et limbes non pileux
- II. K148 - K149 feuilles à poils durs
- III. K139 - K140 - K142 - K154 - K155 - feuilles à poils longs densité de pilosité moyenne

K124. se rattache au groupe I car il a les limbes non pileux mais les gaines pileuses.

K125. ne se rattache spécifiquement pour aucune catégorie de caractères à l'un des 3 groupes mais se rattache au groupe II pour l'ensemble des caractères.

Population K3 KOROGWE, population contenant le diploïde K189,

K194, (K187-K188), (K191-K193) se différencient d'un noyau (K190 - K192 - K185) le diploïde K189 fortement lié à K185 pour les caractères feuilles II (limbes proprement dit) se rattache à ce noyau ce qui dans le dendrogramme sur l'ensemble des caractères donne la priorité au rattachement K185 - K189 avant le rattachement (K195-K190-K192) l'examen de la matrice de similarité montre que K185 est lié avec K190 et K192 positivement alors que le raccordement de (K185 - K189) à (K190 - K192) est à un niveau-3.

K7, K8, K9. Populations voisines de la population contenant les diploïdes.

Ces populations sont composés de clones ayant des phénotypes liés dans leur ensemble les uns aux autres.

Pour K7 de MOMBASA à LUNGA LUNGA population de la basse côte humide on a 4 groupes enchevêtrés.

K216 K211 K212 K205, K218 K209 K206 K207 K208 K214 K213 K215 K217

pour K8 MONTS PARE NORD autour de la savane à Panicum.

On a 4 groupes reliés entre eux par 2 ou 3 phénotypes

K156 K169 K159, K172 K173 K174, K162 K163 K160 K168 K164, K165 K166,

pour K9 de BARON'S FALLS à MOMBO.

K181 K182 K178 K175 K176 K177 K180 K183 K184 K185 K187

Toutes ces liaisons qui se font à un niveau de ressemblance toujours faible sont le reflet d'un ensemble de caractères sans particularisation proprement dite.

On a vu que la moyenne théorique des indices de similarité entre les individus d'une population entièrement aléatoire est 0. Le faible niveau de ressemblance entre les groupes montre que les populations K7, K8, K9 sont peu structurées.

Ces populations sont aussi en cours de différenciation les phénotypes en apparence homogènes commencent à se particulariser, cependant aucun groupe n'est vraiment formé.

F - BIBLIOGRAPHIE

- BONHAN-CARTER (G.F.) - 1965 - A numerical method of classification using qualitative and semi-quantitative data as applied to the facies analysis of Pimestons. Bull. Canadian Petroleum Geol. Vol. 13 : p. 482-502.
- BONNER (R.E.) - 1964 - On some clustering techniques IBM jour. Vol. 8 : p. 22-32.
- CATTELL (R.B.) COULTER (M.A.) - 1966 - Principles of behavioural taxonomy and the mathematical basis of the taxonome computer. Brit. J. Math. Statist. Psychol. Vol. 19 : p. 237-269.
- CLIFFORD (H.T.) GOODALL (D.W.) - 1967 - A numerical contribution to the classification of the Poaceae. Aust. J. Bot. Vol. 15, p. 499-519.
- CONSTANTINESCU (P.) - 1967 - A method of cluster analysis. Brit. J. Math. Statist. Psychol. Vol. 20 : p. 93-106.
- CROVELLO (T.J.) - 1968 - Different concepts of relevance in a numerical taxonomic study. Nature Lond. Vol. 218 : p. 492.
- DACNELIE (P.) - 1963 - L'étude des communautés végétales par l'analyse à plusieurs variables : un exemple 5th International Biometric conference. Cambridge pp. 2-13.
- DAGNELIE (P.) - 1966 - A propos de différentes méthodes de classification numérique. Rev. Statist. Appl. Vol. 14, n° III, p. 55-75.
- EDWARDS (A.W.F.) CAVALLI-SFORZA (L.L.) - 1965 - A method for cluster analysis. Biometrics, Vol. 21, p. 362-375.
- FLAKE (R.H.) TERNER (B.L.) - 1968 - Numerical classification for taxonomic problems. J. Theoret. Biol. Vol. 20 : p. 260-270.

- GHENT (A.W.) - 1963 - Kendall's tau coefficient as an index of similarity in comparisons of plant or animal communities. *Canad. Entom.* Vol. 95 : p. 568-575.
- GOODALL (D.W.) - 1964 - A probabilistic similarity index *Nature*, Vol. 203 : p. 1098.
- GOODALL (D.W.) - 1966a - A deviant index : a new tool for numerical taxonomy. *Nature*, Vol. 210 : p. 216.
- GOODALL (D.W.) - 1966b - A new similarity index based on probability. *Biometrics*, Vol. 22, n° 4 : p. 882-907.
- GOODALL (D.W.) - 1966c - Numerical taxonomy of bacteria some published data re-examined. *J. gen. Microbiol.* Vol. 42 : p. 25-37.
- GOODALL (D.W.) - 1966d - Classification probability and utility. *Nature*, Vol. 211 : p. 53-54.
- GOODALL (D.W.) - 1966e - Hypothesis testing in classification. *Nature*, Vol. 211 : p. 329-330.
- GOODALL (D.W.) - 1968 - Affinity between an individual and a cluster in numerical taxonomy. *Biometrie - Praximetrie*, Vol. IX, n° 1 : p. 52-55.
- GRANT (W.F.) ZANDSTRA (I.I.) - 1968 - The biosystematics of the genus *Lotus* (leguminosae) in Canada II. Numerical chemo-taxonomy. *Canad. J. Bot.* Vol. 46, p. 585-89.
- GRIFFITH (J.G.) - 1967 - Numerical taxonomy and textual criticism of a classical latin author. *Nature*, Vol. 215 : p. 326.
- HALL (A.V.) - 1965 - The peculiarity index, a new function for use in numerical taxonomy. *Nature*, Vol. 206 : p. 952.
- HALL (A.V.) - 1967 - Studies in recently developed group - forming procedures in taxonomy and ecology. *J.S. Afr. Bot.*, Vol. 33 : p. 185-196.

- HARRISON (P.J.) - 1968 - A method of eluster analysis and some applications applied Statistics, Vol. 17 n° 9.
- HEISER (C.B.) SORIA (J.) BURTON (D.L.) - 1965 - A numerical taxonomic study of Solanum species an hybrids. Amer. nat. Vol. 99, p. 471-488.
- HIERNAUX - 1965 - Une nouvelle mesure de distance anthropologique entre population utilisant simultanément des fréquences géniques, des pourcentages de trait descriptifs et des moyennes métriques. Compte rendu de l'Académie des Sciences de Paris Tome 260 : p. 1748-1750.
- HOWD (F.H.) - 1964 - The taxonomy program. A computer technique for classifying geologie data. Quart. Colo. Sch. Mines, Vol. 59, n° 4 : p. 207-222.
- JIZBA (Z.V.) - 1964 - A contribution to a statistical theory of classification. Stanford Univ. Publ., Géol. Sci., Vol. 9, n° 2 : p. 729-756.
- MACNAUGHTON-SMITH (P.) - 1965 - Some statistical and other numerical techniques for classifying individuals. London Her Olajesty's stationery office.
- PARKS (J.M.) - 1966 - Cluster analysis applied to multivariate geologie problems. The journal of Geology, Vol. 74, n° 5, part. 2 : p. 703-715.
- PERNES (J.), COMBES (D.) - 1967 - Rapport de mission de prospection Panicum maximum en Afrique de l'Est du 1/7 au 23/7/1967.
- PERNES (J.) - Les populations naturelles ivoiriennes de l'espèce Panicum maximum et les types analogues introduits.
- ROSE (M.J.) - 1964 - Classification of a set of elements. Computer 5., Vol. 7 : p. 208-211.

- ROUX (G.) ROUX (M.) - 1967 - A propos de quelques méthodes de classification en phytosociologie. Rev. Statist. Appl. Vol. 15, n° II : p. 59-72.
- SIMON (J.P.) GOODALL (D.W.) - 1968 - Relationship in annual species of Medicago. VI. Two-dimensional chromatography of the phenolics and analysis of the results by probabilistic similarity methods. Aust. J. Bot. Vol. 16, n° 1 : p. 89-100.
- SOKAL (R.R.) - 1961 - Distance as a measure of taxonomy. Syst. Zool. Vol. 10 : p. 70-79.
- SOKAL (R.R.) SNEATH (P.H.A.) - 1963 - Principles of numerical taxonomy. W.H. Freeman and Company. San Francisco and London.
- SNEATH (P.H.A.) 1957 - The application of computers to taxonomy. Jour. Gen. Microbiol., Vol. 17 : p. 201-226.
- STRINGER (P.) - 1967 - Cluster analysis of non-verbal judgements of facial expressions. Brit. J. Math. Statist. Psychol., Vol. 20 : p. 71-79.
- WILLIAM (G.F.) - 1965 - Methods in taxonomy. Nature, Vol. 207, p. 245-247.
- WILLIAMS (W.T.) DALE (M.B.) - 1964 - An objective method of weighting in similarity analysis. Nature, Vol. 201 : p. 426.
- WILLIAMS (W.T.) LANCE (G.N.) - 1965 - Logic of Computer-based intrinsic classifications. Nature, Vol. 207 : p. 159-161.
-