NUMERICAL TAXONOMY OF WEED AND CULTIVATED SORGHUMS

By

MARGARET HSUAN LU

Bachelor of Science

National Taiwan University

Taipe, Taiwan

The Republic of China

1964

Submitted to the faculty of the Graduate College
of the Oklahoma State University
in partial fulfillment of the requirements
for the degree of
MASTER OF SCIENCE
July, 1967

NUMERICAL TAXONOMY OF WEED AND CULTIVATED SORGHUMS

Thesis Approved:

J. m. J. de wet
Thesis Adviser

Margaret H. Brooks

John E. Thomas

N Durham
Dean of the Graduate College

688470

ACKNOWLEDGEMENTS

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

The genus Sorghum was studied cytogenetically by Huskins and

Smith (1932, 1934), and by Garber (1950) who recognized six sub-

genera. This major subdivision of the genus was also recognized by

Celarier (1959), except that he treated Sorghastrum as a distinct

genus. The usually accepted classification is by Snowden (1936, 1954),

who recognized fifty-two species belonging to section Eu-sorghum. He

further subdivided Eu-sorghum into two subsections, Arundinacea,

without rhizomes, and Halepensia, with well-developed rhizomes.

Snowden subdivided subsection Arundinacea into two series, Spontanea,

with more or less fragile racemes and series Sativa with tough racemes.

He also indicated that the wild species of series Spontanea were

closely related to the cultivated species (series Sativa) and that

they readily intercrossed when grown together. Among the fifty-two

species of Snowden, thirty-one are cultivated, seventeen are wild

fodder species and four are naturally occuring rhizomatous species.

Clayton (1961) transferred all the Snowdenian species into Sorghum

bicolor (Linn.) Moench. Since there are no real barriers to prevent

gene exchange between the Snowdenian species, they should belong to

one heteromorphic species, de Wet and Huckabay (1967). Recently, two

more rhizomatous species, Sorghum randolphianum Parodi and Sorghum

almum Parodi were described to include an artificially produced and a

1

natural hybrid respectively between rhizomatous and grain sorghums.

Snowden's classification was based strictly on subjective morpho-
logical characters. In the current study, fifty-three quantitative
and qualitative morphological characters were employed in an effort
to determine, by means of computer analysis, the extent of variability
among twenty-five species recognized by Snowden. This should also
provide a basis for classifying the collections made for this study
and to determine the phylogenetic affinities between the species as
recognized by Snowden.

CHAPTER II

REVIEW OF LITERATURE

As early as 1898 Heincke used a measure of phenetic distance to distinguish between races of the herring. It was early realized that biometrics could be applied to systematics. Since then people have paid attention to biosystematics and its usefulness. Nuttall (1901, 1904) used numerical taxonomy in comparative serology. Proom and Woiwod (1949) and Micks and Ellis (1952) used it in conjunction with two-way paper chromatography. Randal et al. (1951) were pioneers in applying the technique of infrared spectroscopy to the study of the taxonomy of microorganisms. The numerical taxonomy technique is constantly being improved and becoming more useful in many fields.

Similarity coefficients are the basic tools of numerical analysis. There are coefficients of association, coefficient of correlation, and coefficients of distance.

In numerical studies, the fundamental taxonomic units are referred to as operational taxonomic units. This is because, according to Sokal and Sneath (1963) the hierarchic level of the taxonomic unit employed in numerical studies will differ and it will not be correct to refer to such a unit by a specific fundamental taxonomic unit.

In the association coefficients, Sneath (1957) used the coefficient of Jaccard. The formula is $S_J = n_{JK} / (n_{JK} + U)$, $U = n_{Jk} + n_{jK}$, n_{JK} = positive matches of characters between two operational

taxonomic units (OTU's), n_{Jk} and n_{jK} are unmatched characters between

two OTU's. $s_J \longrightarrow o$ when $n_{JK}/u \longrightarrow o$, $s_J \longrightarrow 1$ when $u \longrightarrow o$. This

formula has been used in R-type (in which the similarity of pairs of

characters be examined over all OTU's) and Q-type (in which the

similarity of pairs of OTU's be examined over all characters) studies

in ecology. It omits the negative matches. Sokal and Sneath (1963)

used the simple matching coefficient. The formula is $S_{sm} = (n_{JK} +$

$n_{jk})/n$, where n_{JK} is the positive matches and n_{jk} is the negative

matches between two OTU's. n is the total character numbers. This

coefficient includes negative matches and is more stable for the more

constant denominator. When first suggested by Sokal and Michener (1958),

it was not restricted to characters with only two states. Sokal and

Sneath (1963) restricted it to two states to obtain less bias in

consequent statistical analysis. Rogers and Tanimoto (1960) developed

the coefficient of Rogers and Tanimoto. The formula is $S_{RT} = (n_{JK} +$

$n_{jk})/(n + n_{Jk} + n_{jK})$, where the symbols are the same as those described

above. In this formula, the unmatched pairs carry twice the weight of

the matched pairs in the denominator. It is more elaborate than the

simple matching coefficient. Yule and Kendall (1950) used the phi

coefficient. The formula is $S\phi = (n_{JK}n_{jk} - n_{jK}n_{Jk})/(n_j n_J n_K n_k)^{\frac{1}{2}}$, $S\phi$ is

the coefficient. n_J and n_K are the marginal totals of positive

characters of OTU j and k; n_j and n_k are the marginal totals of

negative character. The phi coefficient is frequently used in

statistics and is important for its relation to χ^2, that is $\chi^2 = nS\phi^2$

with degree of freedom equal to one. This permits a test of signifi-

cance, but it is doubtful whether any meaning can be applied to such a

test when more new characters are added which influence the homology

of the column vectors. Smirnov (1960) proposed the Smirnov coefficient of similarity. The formula is $t_{fg} = \sum_{i=1}^{m} (\Sigma\, W_{Ei})/n$, where m is the number of characters, n is the total scored number of all the characters, f, g are two OTU's, W_{Ei} represents different weight of score of different character. Smirnov's coefficient can not be supported because it is not agreed that the similarity in rare structures is more important than the similarity in commonly occuring structures. Besides, the similarity does not result in unity when comparing OTU's with themselves.

Under two states characters condition, the simple matching coefficient is the simplest and the easiest one to interpret in numerical analyses.

Pearson's product-moment had been first used in numerical taxonomy by Michener and Sokal (1957) as correlation coefficient. The formula is $r_{jk} = \sum_{i=1}^{n} (X_{ij} - \overline{X}_j)(X_{ik} - \overline{X}_k) \Big/ \sqrt{\sum_{i=1}^{n} (X_{ij} - \overline{X}_j)^2 \sum_{i=1}^{n} (X_{ik} - \overline{X}_k)^2}$, where r_{jk} is the coefficient between taxon j and k. X_{ij} is the character value of character i in OTU j, $\overline{X}_j$ is the mean of all states values for OTU j, and n is the number of characters sampled. Correlation coefficients have been repeatedly used in Q-type studies in both psychology and ecology.

In measuring distance, Cain and Harrison (1958) proposed the mean character difference, (M.C.D.). The formula is $\sum_{i=1}^{n} |X_{ij} - X_{ik}| \Big/ n$, which is the distance between OTU's j and k, where n is the character sampled. The M.C.D. is always less than the true distance, and thus will not permit the participating which is possible in mean squares. Sokal (1961) employed $d_{jk}^2 = \sum_{i=1}^{n} (X_{ij} - X_{ik})^2 \Big/ n$, as the taxonomic distance. Where d_{jk}^2 is the distance between two points in an

n-dimensional space. It can be used to measure unequal characters.

When the taxonomic distance is large, the degree of association is

small. Clark (1952) used $CD_{jk} = \sqrt{\sum_{i=1}^{n} \lfloor (X_{ij} - X_{ik})/(X_{ij} + X_{ik}) \rfloor} /n$

as the coefficient of divergence in comparison of snake populations.

The symbols are the same as those described in Sokal's taxonomic

distance. Karl Pearson (1926) developed the coefficient of racial

likeness (C.R.L.). The formula is C.R.L. $= \sqrt{1/n \sum_{i=1}^{n} \lfloor (\overline{X}_{ij} - \overline{X}_{ik})^2 /}$

$(S_{ij}^2/n_j + S_{ik}^2/ n_k) \rfloor} - 2/n$, where $\overline{X}_{ij}$ stands for the sample mean of

the ith character for entity j, S_{ij}^2 for variance of the same, n_j for

the sample size of entity j. The C.R.L. can measure the resemblance

between samples of various origins. The standard error of C.R.L.

approaches $1/\sqrt{2n}$ when n is sizable. The coefficient was used to

measure the anthropological material, mostly skulls.

The simple matching coefficient S_{Sm} among the association coeffi-

cients, the Pearson's product-moment r of correlation coefficient and

the taxonomic distance d based on standard characters among the distance

measures, are the most preferable ones. All three can be used in

cluster analyses while only the correlation coefficient can be applied

to factor analyses.

Between the correlation coefficient and distance coefficients, the

former is preferred because the variance analysis can be made and the

coefficient of determination between each two OTU's can be determined.

Rohlf and Sokal (1965) suggested that a correlation coefficient should

be used rather than a distance, whenever most of the characters used

in a study are measurements of various parts of an organism and the

OTU's differ much in overall size. When characters are independent of

size, distance coefficients seem more meaningful. The size is more

important than shape in most materials.

Association and distance are easier to interpret conceptually than correlation, because they are found in unit taxonomic characters between two organisms. The conventional interpretation of a correlation coefficient, perhaps the most applicable for numerical analyses, is through its square which identifies the proportion of the common variance of the two OTU's. When the characters are in multistates, correlation and distance are preferred. If there are many characters that have small ranges, the coefficients are not preferred either. Under this condition, if the two state characters with not too many matches, the association coefficients are recommended. It was decided to use the simple matching coefficient in this paper.

There are three methods generally used in grouping: differential shading of similarity matrix, cluster analyses, and factor analyses.

Differential shading of similarity matrix had been used by Robinson (1951). When many taxa are studied, the diagrams become quite unwieldy.

In cluster analyses, elementary cluster analysis was used by Boeke (1942). When the criterion for admission of similarity coefficients is lowered, more OTU's join the established clusters. Sooner or later clusters will overlap by this method. Sneath discussed the single linkage group in 1957. When two clusters are linked by the technique on the basis of a single bond, many of the members of two clusters may be quite far removed from each other. The average linkage grouping methods were proposed by Sokal and Michener (1958). They used correlation coefficients in this technique. But, according to Sokal and Sneath (1963), other types of similarity coefficients can also be applied to these methods. Among them, the weighted variable

group method is better than the weighted pair group method, because the
latter permits only the two most highly correlated stems to join at each
clustering cycle and gives more troubles in calculations. As soon
as all the prospective members have joined their clusters, a new
similarity matrix of all clusters with each other and with single stems
is recalculated. The weighted variable group method is better than the
unweighted variable group method because the former reduces the weight
of the members admitted earlier and increases the weight of those OTU's
admitted later. This technique also has more phylogenetic or evolution-
ary meaning. The unweighted pair group method is also good, even
though it has some bias of early joined members. But the calculation
is too complicated relative to the value of the improved efficiency.
Central or nodal clustering was proposed by Rogers and Tanimoto in
1960. Association coefficients are used in this method. It is too
complicated in calculation because the homogeneity must be tested
after each clustering. A graph theory model for systematic grouping
was suggested by Wirth, Estabrook and Rogers (1966). The advantage
of this method is that the graph figures show how each cluster is
formed. But the disadvantage is that the moat which is the numerical
value of the isolation of a C-cluster, whould be larger. The small
similarity error still cannot be corrected. The hierarchical
agglomerative centroid method was used by Lance and Williams (1966).
The method is to synthesize the populations or to group the progressive
subdivisions in accordance with the predetermined measure of "likeness"
of members of the populations. It starts from the similarities
"information statistic" and "non-metric statistic" coefficients. Both
are somewhat confusing in their components. The information statistic

can be used in two-states characters while the non-metric statistic

always takes the unknown values as no value. In this method, some

degree of sacrifice of the homogeneity of groups still is found.

Factor analysis was first employed in the grouping method by

Sokal (1958). It applies the smallest number of factors to interpret

the complex interrelationships among taxa. The multiple factor analysis

is a branch of multivariate statistics. The isolation of OTU's is

pregrouping but not exgrouping. It purifies the elements of species

when investigating and classifying them. The factor analysis indicates

the cluster to which an OTU belongs and the degree to which each of the

OTU's resembles an "average" representative of the cluster. This

might be an advantage in that it prevents one from attempting to

interpret differences which are probably not reliable. Usually the

results are similar to the weighted pair group method.

Although the analysis is more precise in factor analysis, the

calculating procedures are too complicated. It is necessary to use a

computer in every step. The weighted variable grouping method is

more convenient and also gives reasonable answers. For these reasons,

it was decided to use Sokal and Michener's weighted variable group

method of average linkage clustering technique in the numerical

analysis of <u>Sorghum</u> <u>bicolor</u>.

CHAPTER III

MATERIALS AND METHODS

The collections studied include cultivated, wild, and semiwild
species from Africa, Asia, India, Australia, United States and other
parts of the world where sorghums are grown. The collections were
grown in a uniform nursery at the Oklahoma Agricultural Experimental
Station at Stillwater, Oklahoma. Herbarium specimens preserved for
the morphological studies are filed with the biosystematic herbarium at
the Oklahoma State University. Seventy-two collections were studied
(Table I). Five mature plants of each collection were studied in de-
tail. Bud materials were also collected and studied for cytological
analysis.

The first step of the numerical analysis is to calculate the
similarity matrix. Each similarity coefficient is based upon differ-
ences and similarities between the fifty-three morphological characters
(Table II) of the two collections being compared. The data of charac-
ters of each collection were punched on IBM cards. The comparison was
made with a 7040 computer. This gave a matrix with 2,556 indices
(Table IV). The seventy-two collections were grouped by Sokal and
Michener's (1958) weighted variable group method of average linkage
clustering. Finally a dendogram of relationships of the seventy-two
collections was constructed. These numerical observations were
correlated with data on hybridization ranges and the cytological
analysis to draw our final conclusions.

10

TABLE I

CLASSIFICATION AND ORIGIN OF THE SORGHUM COLLECTIONS STUDIED
(Following de Wet, 1967)

NUMBER	COLLECTION	SNOWDENIAN SPECIES	ORIGIN
1	2784	S. controversum	India
2	2784a	S. controversum	India
3	5866	S. halepense	India
4	6358	S. controversum	India
5	10602	S. controversum	India
6	10766a	S. halepense	Australia*
7	10766d	S. halepense	Australia*
8	10766f	S. halepense	Australia*
9	10755	S. halepense	Australia*
10	10954	S. halepense	Australia*
11	4027	S. miliaceum	India
12	5311	S. miliaceum	India
13	6345	S. miliaceum	India
14	9373	S. miliaceum	India
15	9380	S. miliaceum	India
16	7172	S. propinquum	Philipines
17	8022	S. propinquum	Philipines
18	10836	S. propinquum	Singapore
19	10837	S. propinquum	Singapore
20	10838	S. propinquum	Singapore
21	5667a	S. halepense	Cultivated-U.S.A.
22	5867	S. almum	Cultivated-U.S.A.
23	5867a	S. almum	Cultivated-U.S.A.
24	5830	S. aethiopicum	Sudan
25	5830a	S. aethiopicum	Sudan
26	2511	S. virgatum	Sudan
27	5655	S. virgatum	Ethiopia

TABLE I (Continued)

NUMBER	COLLECTION	SNOWDENIAN SPECIES	ORIGIN
28	5677	S. virgatum	Ethiopia
29	5873	S. virgatum	Rhodesia
30	7171	S. virgatum	Egypt
32	5736	S. arundinaceum	Guinea
33	7276	S. arundinaceum	Ghana
34	8315c	S. arundinaceum	India
35	10012	S. arundinaceum	India
36	2792	S. arundinaceum	Rhodesia
37	2792a	S. arundinaceum	Rhodesia
38	2792b	S. arundinaceum	Rhodesia
39	2792c	S. arundinaceum	Rhodesia
40	2792h	S. arundinaceum	Rhodesia
41	4853	S. pugionifolium	Burma
42	8590	S. pugionifolium	Burma
43	4841	S. verticilliflorum	Rhodesia
44	5657	S. verticilliflorum	Rhodesia
45	5748	S. verticilliflorum	Malawi
46	5881	S. verticilliflorum	S. Africa
48	4832	S. drummondii	Portugal
49	5648	S. drummondii	U.S.A.
50	4832	S. niloticum	Ethiopa
51	4832a	S. hewisonii	Portugal
52	5735	S. hewisonii	Guinea
53	10602e	S. hewisonii	India
54	10868a	S. hewisonii	Australia*
55	10868d	S. hewisonii	Australia*
56	4851	S. roxburghii	Burma
57	5654	S. roxburghii	Burma

TABLE I (Continued)

NUMBER	COLLECTION	SNOWDENIAN SPECIES	ORIGIN
58	5687	S. caffrorum	Cultivated-U.S.A.
59	4838	S. nigricans	Portugal
60	4872	S. cernuum	Near East
61	4863	S. durra	Burma
62	5684	S. durra	India
63	4857	S. subglabrescens	Burma
64	5835	S. ankolib	Ethiopa
65	4840	S. bicolor	Portugal
66	5679	S. bicolor	Cultivated-U.S.A.
67	5738	S. bicolor	Portugal
68	4839	S. bicolor	Portugal
69	10868f	S. dochna	Australia*
70	5667	S. elegans	Cultivated-U.S.A.
71	4874	S. milliforme	Kenya
73	6117	S. simulans	Cultivated-U.S.A.
76	7167a	S. halepense	U.S.S.R.
77	7167e	S. halepense	U.S.S.R.

*Introduced roadside weeds.

TABLE II

CHARACTERS STUDIED

CHARACTER	STATE*	
1. Somatic number of chromosomes	10(1)	15(2)
2. Diploid chromosome number $2\underline{n}$ = 20	yes(1)	no(2)
3. Rhizomes present	yes(1)	no(2)
4. Culm nodes bearded	yes(1)	no(2)
5. Leaf $\leqslant$3 cm. wide	yes(1)	no(2)
6. Leaf $\geqslant$5 cm. wide	yes(1)	no(2)
7. Inflorescence compact	yes(1)	no(2)
8. Inflorescence pyramidal	yes(1)	no(2)
9. Inflorescence branches whorled	yes(1)	no(2)
10. Inflorescence branches divided	yes(1)	no(2)
11. Racemes fragile	yes(1)	no(2)
12. Pedicellate spikelet absent	yes(1)	no(2)
13. Pedicellate spikelet rudimentary	yes(1)	no(2)
14. Pedicellate spikelet persistent	yes(1)	no(2)
15. Sessile spikelet glabrous	yes(1)	no(2)
16. Sessile spikelet shortly hairy	yes(1)	no(2)
17. Awn of sessile spikelet small or absent	yes(1)	no(2)
18. Awn of sessile spikelet$>$50 mm. long	yes(1)	no(2)
19. Lower glume of sessile spikelet $>$10-nerved	yes(1)	no(2)
20. Lower glume of sessile spikelet winged	yes(1)	no(2)
21. Lower glume of sessile spikelet nerves obscured	yes(1)	no(2)
22. Lower glume of sessile spikelet obtuse	yes(1)	no(2)
23. Lower glume of sessile spikelet length/width ratio$>$2	yes(1)	no(2)
24. Sessile spikelet lanceolate	yes(1)	no(2)
25. Lower glume of sessile spikelet wrinkled-depressed	yes(1)	no(2)

TABLE II (Continued)

CHARACTER	STATE*	
26. Sessile spikelets>6 mm.	yes(1)	no(2)
27. Callus of sessile spikelet acute	yes(1)	no(2)
28. Lemma of sessile spikelet copiously ciliate	yes(1)	no(2)
29. Lemma of sessile spikelet bilobed	yes(1)	no(2)
30. Lemma of sessile spikelet mucronate	yes(1)	no(2)
31. Grain exposed by gaping glumes	yes(1)	no(2)
32. Grain longer than glumes	yes(1)	no(2)
33. Annual	yes(1)	no(2)
34. Domesticated	yes(1)	no(2)
35. Weedy	yes(1)	no(2)
36. Confined to Africa	yes(1)	no(2)
37. Confined to Austral-polynesia	yes(1)	no(2)
38. Confined to Asia	yes(1)	no(2)
39. First leaf length/width ratio	ranges from 10.9 to 50	
40. Leaf pubscent	yes(1)	no(2)
41. Leaf strongly pubescent	yes(1)	no(2)
42. Leaf sheath pubescent	yes(1)	no(2)
43. Leaf sheath strongly pubescent	yes(1)	no(2)
44. Inflorescence axis length	ranges from 200 to 699 mm.	
45. Number of nodes per primary axis of inflorescence	ranges from 10 to 20	
46. Total number of primary inflorescence branches	ranges from 30 to 90	
47. Number of branches/number of nodes ratio	ranges from 2 to 5.9	
48. Length of lower primary branch of the inflorescence	ranges from 130 to 209 mm.	
49. Inflorescence branch pilose	yes(1)	no(2)
50. Number of nodes per primary branch of the inflorescence	ranges from 0 to 11	
51. Number of secondary branches per primary branch	ranges from 0 to 20	
52. Number of sessile spikelets per secondary branch	ranges from 0 to 6	
53. Meiosis regular	yes(1)	no(2)

*Numbers inside parentheses are the codes used.

TABLE II (Continued)

CHARACTER	STATE*	
26. Sessile spikelets>6 mm.	yes(1)	no(2)
27. Callus of sessile spikelet acute	yes(1)	no(2)
28. Lemma of sessile spikelet copiously ciliate	yes(1)	no(2)
29. Lemma of sessile spikelet bilobed	yes(1)	no(2)
30. Lemma of sessile spikelet mucronate	yes(1)	no(2)
31. Grain exposed by gaping glumes	yes(1)	no(2)
32. Grain longer than glumes	yes(1)	no(2)
33. Annual	yes(1)	no(2)
34. Domesticated	yes(1)	no(2)
35. Weedy	yes(1)	no(2)
36. Confined to Africa	yes(1)	no(2)
37. Confined to Austral-polynesia	yes(1)	no(2)
38. Confined to Asia	yes(1)	no(2)
39. First leaf length/width ratio	ranges from 10.9 to 50	
40. Leaf pubescent	yes(1)	no(2)
41. Leaf strongly pubescent	yes(1)	no(2)
42. Leaf sheath pubescent	yes(1)	no(2)
43. Leaf sheath strongly pubescent	yes(1)	no(2)
44. Inflorescence axis length	ranges from 200 to 699 mm.	
45. Number of nodes per primary axis of inflorescence	ranges from 10 to 20	
46. Total number of primary inflorescence branches	ranges from 30 to 90	
47. Number of branches/number of nodes ratio	ranges from 2 to 5.9	
48. Length of lower primary branch of the inflorescence	ranges from 130 to 209 mm.	
49. Inflorescence branch pilose	yes(1)	no(2)
50. Number of nodes per primary branch of the inflorescence	ranges from 0 to 11	
51. Number of secondary branches per primary branch	ranges from 0 to 20	
52. Number of sessile spikelets per secondary branch	ranges from 0 to 6	
53. Meiosis regular	yes(1)	no(2)

*Numbers inside parentheses are the codes used.

The fifty-three characters studied include qualitative and quantitative ones. The former were coded as yes and no states. The ranges of quantitative characters varied and were coded from one to seven. Both the qualitative and quantitative characters were recorded and treated into 2 x 2 tables which consisted of n characters scored for each two OTU pairs. Each 2 x 2 table had positive matches, negative matches and unmatched characters (the characters were positive or negative for either one of the OTU's respectively). Following Sokal and Sneath's (1963) simple matching coefficient formula which is the total number of positive and negative matches, total number of characters studied, the first similarity matrix with 2,556 indices (Table IV) was completed. Then Sokal and Michener's (1958) weighted variable group method of average linkage clustering technique was employed. At first, the highest similarity coefficient in the matrix between two OTU's was determined. This was the first nucleus of the clustering cycle. The next taxon considered in this nucleus has the highest average similarity with it. It should not be lower than the necessary average similarity $\overline{S}_n$ which varies depending on number of characters, different material, different similarity and stage of clustering. If the joining OTU's average similarity was lower than $\overline{S}_n$, the next high similarity coefficient among the matrix indices was taken to be the next nucleus. The grouping method continued, basing admission of any individual into a cluster on the average of the similarities of that individual with the members of the cluster. As the clusters grew, more remote relatives were considered as prospective members. When any one prospective member's average similarity was lower than the determined $\overline{S}_n$, it was excluded from the cluster.

These procedures continued until there were only two clusters left. Single OTU's with too low an average similarity to associate with other clusters were included in the next correlation coefficient calculations.

After the first clustering cycle, the new correlation coefficients among all the clusters and OTU's were calculated and a new matrix by Spearman's sums of variables method was obtained. This method correlates the sums of the variables making up any one cluster with sums of variables in any other cluster. The formula is $r_{qQ} = \Box qQ / (q + 2 \Delta q)^{\frac{1}{2}} (Q + 2 \Delta Q)^{\frac{1}{2}}$, where $\Box qQ$ is the sum of all correlations (associations in first cycle) between members of one group with the other group, Δq is the sum of all correlations (associations in first cycle) between members of the first group, ΔQ is the sum between members of the second group, q is the number of OTU's in group 1, and Q is the number of OTU's in group 2. The formula of the correlation of the sums of the variable making up one cluster (q) with a single OTU(x) is $r_{xq} = \Sigma r_{xq} / (q + 2 \Delta q)^{\frac{1}{2}}$, where Σr_{xq} is the sum of all correlations (associations in first cycle) between members of one group (q) with the single OTU(x). After having arrived at a new matrix which consisted of correlation coefficients, the clustering described above was repeated. The clustering cycles and recomputation of new matrixes continued until finally a single cluster was obtained. Nine clustering cycles were necessary in this study. From these clusters the dendogram was constructed and the relationships between OTU's graphically presented (Figure 1).

CHAPTER IV

RESULTS

The seventy-two collections studied are listed in Table I. Except

for S. propinquum with 2n = 20, members of S. bicolor subsp. halepense

as well as the hybrid S. almum have a 2n = 40 somatic chromosomes with

more or less irregular chromosome behavior during meiosis of micro-

sporogenesis. All the members of S. bicolor subsp. bicolor have 2n= 2 0

somatic chromosomes with complete pairing during meiosis of micro-

sporogenesis. The members of S. bicolor subsp. halepense are rhizoma-

tous, while the members of S. bicolor subsp. bicolor are non-rhizomatous.

The weed varieties of subsp. bicolor, and members of subsp. halepense,

have loose and open inflorescences and grains which are completely

enclosed by the glumes, while the cultivars usually have more compact

inflorescences. Race bicolor has the inflorescence more loose than

kafir, durra or guinea races. Except for race bicolor, cultivated

sorghums have large grains which are often extruded from the glumes.

All the members of halepense and weedy members of bicolor have fragile

racemes. The cultivars are characterized by tough racemes.

All the members of subspecies halepense and the semiwild represent-

atives of subsp. bicolor are perennial. Cultivated sorghums are annual

or weakly biannual. Cultivated sorghum is now widely distributed

across the Old and New worlds. They probably were first cultivated

in Africa and were transported from there by man first to the Near-East

and India, and later to the Far-East. Subspecies _halepense_ extends from the Mediterranean region, through India to the islands of South-East Asia.

Members of race _durra_ can be recognized by their distinctly wrinkled glumes. The semiwild complexes differ from the cultivars, in usually having large lanceolate sessile spikelets. All the species studied have inflorescences with divided branches that are not whorled and rarely pilose. The awns of the sessile spikelets of all the species are small or often absent. The lower glumes of the sessile spikelets are always winged and have more than ten nerves, the callus is acute, the lemmas are bilobed and ciliate, and all the species have nodes which are essentially glabrous.

The above mentioned morphological, cytological, and distribution traits are useful to identify the large complexes. The morphology of $\underline{S}$. _bicolor_ is extremely variable, and the ranges of variation for some of the quantitative characters studied are listed in Table III.

The similarity coefficients among them, based on fifty-three quantitative and qualitative characters, were compared, and put in a matrix (Table IV). Clustering the OTU's with a drop-off value (necessary average similarity) of $\overline{S}_n$ .02, as described in the materials and methods section, resulted in thirteen groups of two units each and forty-six unattached OTU's. The correlation coefficients among the groups and OTU's were then calculated and a new matrix was obtained. Clustering from the secondary matrix, with a drop-off value of .02, one group of four units, elven groups of two units and thirty-three unattached OTU's were formed. Repeating the procedures described above, the third clustering from a third new matrix gave a group of

TABLE III

THE QUANTITATIVE CHARACTERS OF THE SNOWDENIAN SPECIES STUDIED

Snowdenian Species	First Leaf Length/ Width Ratio	Average Inflorescence Axis Length (mm.)	Average No. of Nodes / Pri. Axis of Inflorescence	Average No. of Pri. Infl. Branches
S. controversum	12.5	327.5	14.5	50.5
S. halepense	20.9	445.0	17.8	60.4
S. miliaceum	15.6	395.0	16.0	67.0
S. propinquum	22.1	270.0	16.8	35.3
S. almum	21.4	440.4	15.2	48.6
S. aethiopicum	12.4	292.0	13.2	31.0
S. virgatum	17.3	200.0	13.0	29.7
S. arundinaceum	24.2	315.0	16.3	51.0
S. pugionifolium	16.5	521.0	14.4	71.6
S. verticilliflorum	13.9	373.0	15.8	61.0
S. drummondii	10.3	382.0	13.0	83.2
S. niloticum	20.6	474.4	15.0	55.4
S. hewisonii	14.8	389.4	14.2	61.8
S. roxburghii	7.0	274.8	13.4	62.6
S. caffrorum	6.9	233.0	16.8	57.0
S. nigricans	7.3	147.0	13.8	82.2
S. cernuum	7.4	205.0	13.4	104.4
S. durra	6.9	126.5	13.5	110.0
S. subglabrescens	6.5	220.6	15.6	99.3
S. ankolib	9.9	345.0	16.8	71.0
S. bicolor	8.9	269.0	11.4	69.4
S. dochna	5.5	170.0	14.3	57.8
S. elegans	14.3	328.0	13.8	47.8
S. miliiforme	8.0	248.0	14.8	108.0
S. simulans	6.9	260.8	14.0	66.0

TABLE III (Continued)

Snodenian Species	No. of Branches/ No. of Nodes	Length of Lower Pri. Branch of Inflores- cence (mm.)	No. of Nodes/Pri. Branch of Inflores- cence	No. of Sec. Branch/ Pri. Branch	No. of Sessile Spikelets/ Sec. Branch
S. controversum	3.5	170.0	9.0	10.5	4.5
S. halepense	2.9	209.4	11.8	12.8	5.2
S. miliaceum	4.2	173.6	9.8	10.8	4.0
S. propinquum	2.1	157.5	12.3	12.5	6.5
S. almum	3.2	195.4	10.0	11.6	4.8
S. aethiopicum	2.4	131.0	9.0	10.2	3.8
S. virgatum	2.3	82.5	5.6	6.6	3.6
S. arundinaceum	3.1	144.8	10.0	11.3	4.0
S. pugionifolium	5.1	223.0	11.6	12.8	5.0
S. verticillifrorum	3.9	165.0	11.8	12.8	4.2
S. drummondii	6.4	227.8	13.2	14.0	4.6
S. niloticum	3.9	234.6	13.6	15.0	3.6
S. hewisonii	4.4	168.6	10.6	11.6	4.8
S. roxburghii	4.7	118.0	10.8	12.2	4.2
S. caffrorum	3.4	78.6	10.6	11.6	3.2
S. nigricans	6.0	54.8	9.4	10.8	4.8
S. cernuum	7.7	65.0	7.8	9.6	4.8
S. durra	8.2	48.5	12.5	14.0	3.0
S. subglabrescens	6.3	90.0	11.6	13.2	4.0
S. ankolib	4.5	156.0	13.6	15.2	3.2
S. bicolor	6.3	171.0	9.4	10.6	3.8
S. dochna	4.4	75.4	8.2	9.2	4.2
S. elegans	3.5	166.2	10.8	12.0	4.4
S. milliforme	7.0	82.8	9.6	10.6	4.0
S. simulans	4.8	89.8	13.0	15.3	3.0

TABLE IV

MATRIX I OF SIMILARITY INDICES AMONG SEVENTY-TWO

COLLECTIONS OF SORGHUM

CODE 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77

(large lower-triangular numeric similarity matrix, rows labelled 1 through 76)

five units, four groups of two units and thirty-two unattached OTU's.
In the fourth clustering with drop-off value .02, one group of two
units and eighteen unattached OTU's were formed. From the fifth
clustering with .03 drop value, there were one group of four units,
two groups of three units and fifteen unattached OTU's formed. From
the sixth clustering with a drop value of .35, one group of four units,
one group of three units, two groups of two units and seven unattached
OTU's were obtained. From the seventh clustering with .02 drop value,
one group of four units, one group of three units and two groups of
two units were formed. From the eighth cycle with drop value .035, the
two groups A and B (Figure 1) were joined together to form a large
complex A'. At the ninth cycle with drop value .02, the last two
groups, which were unattached to any group at cycle eight, were
connected. This cluster formed the complex B'. Finally, these two
large complexes clustered together as illustrated in Figure 1.

As shown in the dendogram, the large complex groups resemble in
content the subdivision presented by de Wet (1967). This classifica-
tion is followed in Table V. The species as recognized by Snowden are
listed for convenience in descriptions.

Figure 1.

DENDOGRAM OF RELATIONSHIPS AMONG SEVENTY-TWO

COLLECTIONS OF SORGHUM OBTAINED BY

WEIGHTED VARIABLE GROUP METHOD

The horizontal number 1 to 77 represent operational
taxonomic units, and the vertical numbers .83 to
.99 represent the similarity coefficients.
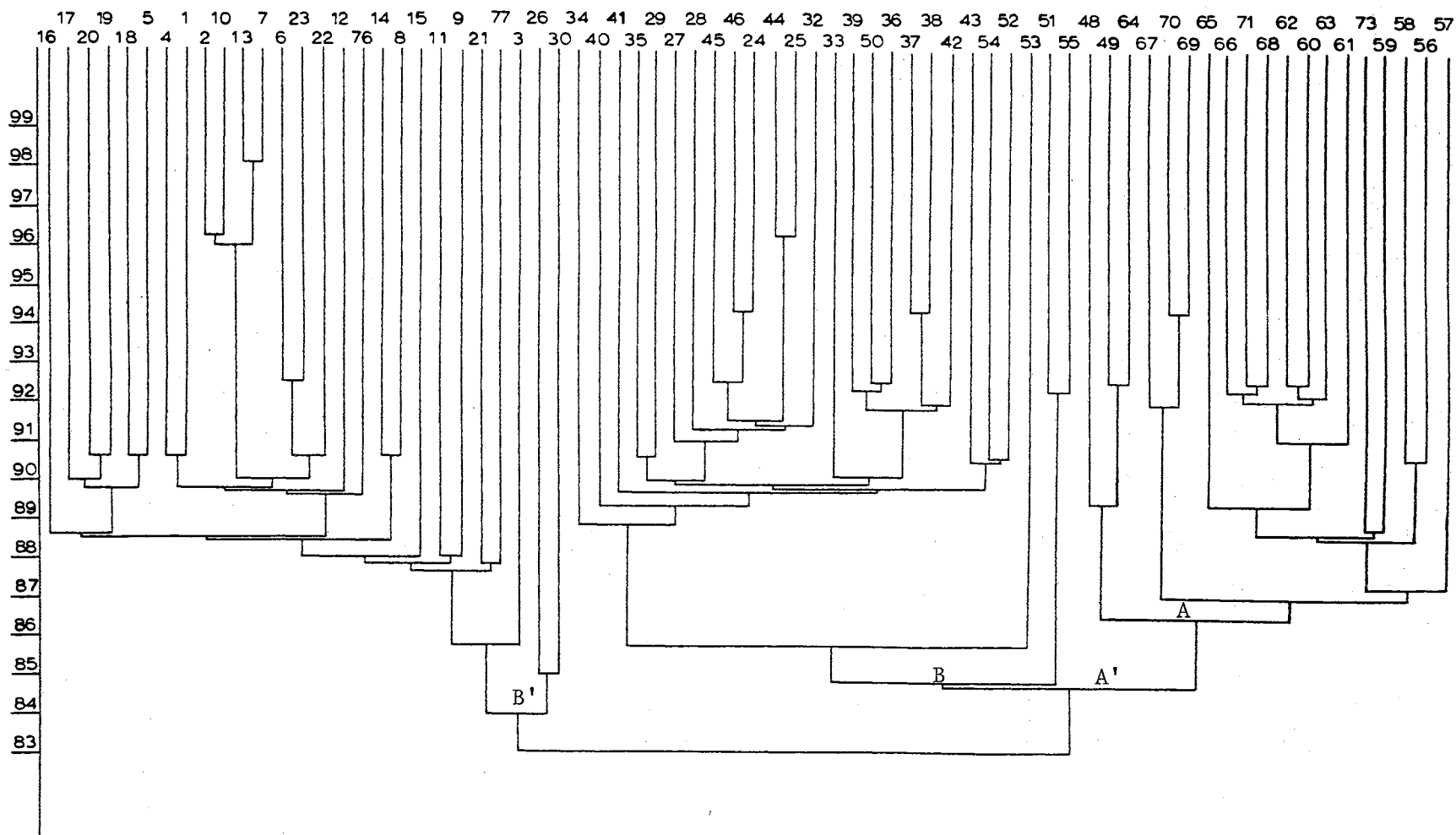
TABLE V

TAXONOMY OF THE SNOWDENIAN SPECIES STUDIED

(Following de Wet, 1967)

Sorghum bicolor

    subspecies halepense

                S. controversum

                S. halepense

                S. miliaceum

                S. propinquum

    subspecies halepense -x- subspecies bicolor

                S. almum

    subspecies bicolor

        variety aethiopicum

                S. aethiopicum

                S. virgatum

        variety arundinaceum

                S. arundinaceum

        variety verticilliflorum

                S. pugionifolium

                S. verticilliflorum

        variety verticilliflorum -x- variety bicolor

                S. drummondii

                S. niloticum

        variety aethiopicum -x- variety bicolor

                S. hewisonii

        variety bicolor

TABLE V (Continued)

race _bicolor_

      _S_. _ankolib_

      _S_. _bicolor_

      _S_. _dochna_

      _S_. _elegans_

      _S_. _miliiforme_

      _S_. _simulans_

race _durra_

      _S_. _cernuum_

      _S_. _durra_

      _S_. _subglabrescens_

race _guinea_

      _S_. _roxburghii_

race _kafir_

      _S_. _caffrorum_

      _S_. _nigricans_

CHAPTER V

DISCUSSION

According to Snowden's (1936, 1954) classification, the genus
Sorghum is divided into section Eu-sorghum with two subsections,
Arundinacea and Halepensia. There are two series in Arundinacea,
Spontanea and Sativa. The former series contains seventeen species.
The latter series comprises thirty-one species which are subdivided
into the subseries Drummondii, Guineensia, Nervosa, Bicoloria, Caffra,
and Durra. Four species are included in subsection Halepensia.
Recently de Wet and Huckabay (1967) included all these species in
Sorghum bicolor (Clayton, 1961) and subdivided the species into two
subspecies, halepense and bicolor. The former subspecies concides
with subsection Halepensia of Snowden which contains S. halepense,
S. miliaceum, S. controversum and S. propinquum. The latter includes
four varieties, var. verticilliflorum, var. aethiopicum, var. arundina-
ceum and var. bicolor. The first mentioned three include the weedy
species of series Spontanea as recognized by Snowden, while the
cultivated variety bicolor coincides with series Sativa of Snowden.
The variety bicolor is further subdivided into four races, kafir,
durra, bicolor and guinea. The major morphological differences
between subsp. halepense and subsp. bicolor are that the former has
forty somatic chromosomes and rhizomes (except for S. propinquum with
$2n = 20$ chromosomes), while the latter has twenty somatic chromosomes

29

and no rhizomes. Almost complete absence of a genetic barrier to prevent gene exchange between the $2n$ = 20 chromosome cultivated and weedy species indicate that they should belong to one heteromorphic complex species (de Wet, 1967). The results from the numerical analysis of this study more or less confirm de Wet's classification.

In the dendogram, the complex B' contains twenty-seven collections from different areas. These belong to S. propinquum, S. controversum, S. halepense and S. miliaceum which de Wet (1967) included in S. bicolor subsp. halepense. It also includes two collections of S. almum which is a hybrid between subsp. halepense and subsp. bicolor, and two of S. virgatum. The last mentioned species more typically belongs with subsp. bicolor (group B).

All the collections of each of the species S. propinquum, S. controversum, S. halepense, and S. miliaceum do not cluster together. Especially S. halepense (7) is closely related to S. miliaceum (13); S. halepense (10) is related to S. controversum (2); S. halepense (6) is allied to S. almum (23); S. halepense (8) is allied to S. miliaceum (14); and S. halepense (9) is connected to S. miliaceum (11). This probably is due to a continuous gene exchange between S. miliaceum, S. controversum and S. halepense wherever they are sympatric. Furthermore, cultivated sorghum introgresses with them wherever they grow together. This gave rise to a continuously variable group as shown in the dendogram.

The variability within species as recognized by Snowden is often greater than that between species. This is particularly obvious in the dendogram when the various collections of S. halepense, S. miliceum and S. controversum are compared. This probably is due to the fact

that these species cross with local sorghums in different ecological
and geographical regions.

The different collections of S. propinquum clustered closely
together, and with one collection of S. controversum from India,
which is somewhat close to a S. propinquum collection at Singapore.
Gene exchange between S. propinquum (2n = 20) and S. controversum
(2n = 20) probably is taking place in this area.

Two collections of S. virgatum, one from the Sudan and the other
from Egypt, form a group distinct from the rest of S. virgatum. These
two collections are relatively close to subsp. halepense and probably
introgressed with S. halepense with which they are sympatric in North
Africa. The major differences between them are that subsp. halepense
has rhizomes and 2n = 40 somatic chromosomes, while S. virgatum has
only twenty somatic chromosomes and no rhizomes.

The complex A' comprises essentially S. bicolor subsp. bicolor as
described by de Wet (1967). There are two groups, A and B, in this
large complex.

The complex B contains the weeds of S. bicolor subsp. bicolor.
There are twenty-six collections from different areas included in
this group. These belong to S. arundinaceum, S. pugionifolium, S.
verticilliflorum, S. aethiopicum, S. virgatum and S. niloticum.
Morphological studies indicated that the last mentioned species
represents a hybrid between var. verticilliflorum and a cultivated
sorghum, however, this hybrid looks like var. arundinaceum. Five
collections of S. hewisonii, a species which seemingly originated as
a cross between var. aethiopicum and var. bicolor, are morphologically
similar to members of var. verticilliflorum.

The species _S. arundinaceum_, as recognized by Snowden, is included in var. arundinaceum; _S. pugionifolium_ and _S. verticilliflorum_ belong to var. verticilliflorum; and _S. virgatum_, as well as _S. aethiopicum_, belong to var. aethiopicum (de Wet and Huckabay, 1967). The usually recognized species form a continuously variable group in complex B. All the collections of each species do not cluster together. This is probably because they are widely scattered in different areas, and crossed with local weedy or cultivated sorghums. This is typically represented in the collections from different places of _S. arundinaceum_, _S. hewisonii_, _S. verticilliflorum_ and _S. aethiopicum_. The species within each variety also do not cluster together. One collection of _S. virgatum_ (29) seems allied to _S. arundinaceum_ (35); _S. aethiopicum_ (24) is closely related to _S. verticilliflorum_ (46); _S. aethiopicum_ (25) is correlated with _S. verticilliflorum_ (44); _S. niloticum_ (50) connects with _S. arundinaceum_ (36);and _S. hewisonii_ (52, 54) is closely related to _S. verticilliflorum_ (43). The varieties do not cluster together either. All the species of Snowden in subspecies bicolor are closely related.

Actually all the weeds of subsp. bicolor are somewhat genetically homologous. They are restricted to Africa where sorghums are widely cultivated, and hybridization among them as well as with cultivars gave rise to a complex that is extremely difficult to classify. This is also true of cultivated sorghums. The weedy sorghums of subsp. bicolor in other areas of the world, probably were introduced from Africa. The concept of three varieties are acceptable only for practical convenience. They are distributed in different areas of Africa, (var. arundinaceum in western Africa, var. aethiopicum in

eastern and central Africa, var. verticilliflorum in southern and southeastern Africa), they are morphologically quite different, and probably have given rise to different races of cultivated sorghums (de Wet and Huckabay, 1967).

The complex A contains the cultivated species of Snowden. These are S. drummondii, S. ankolib, S. bicolor, S. elegans, S. dochna, S. miliiforme, S. durra, S. cernuum, S. subglabrescens, S. simulans, S. nigricans, S. caffrorum, and S. roxburghii. They all belong to S. bicolor subsp. bicolor variety bicolor of de Wet (1967). Among them, S. drummondii is weedy, and only sometimes cultivated. The species S. ankolib, S. bicolor, S. elegans, S. dochna, S. miliiforme, S. simulans belong to race bicolor; S. durra, S. cernuum and S. subglabrescens belong to race durra; S. nigricans and S. caffrorum belong to race kafir; and S. roxburghii belongs to race guinea. Race bicolor probably originated from var. aethiopicum and is now widely distributed in the Near East, India and some areas of Africa; race guinea originated from var. arundinaceum and is distributed in western and southeastern Africa; race kafir originated from var. verticilliflorum and is distributed in most areas of Africa. Members of race durra probably originated as selections from race kafir, with infiltration of some genes from race bicolor. It is now distributed in India, northeastern Africa and the Near-East.

Different species of each race do not always cluster together. Neither do collections of the same species necessarily cluster together. This is probably because different selections were made in different areas by man. Different species often seem closer allied than different collections of a species. This is shown in the connection of S. drummondii (49) with S. ankolib (64), and S. nigricans (59)

with S. simulans (73). The dendogram clearly shows the complex
variation patterns among the members of variety bicolor, these
must have come about as a result of man's continuous selection and
hybridization of cultivated sorghums with weedy sorghums of S. bicolor
subsp. bicolor since the earliest time of sorghum cultivation in east
Africa. Cultivated sorghums are not directly connected with subsp.
halepense.

In total, there are two large complexes, S. bicolor subsp. bicolor
and S. bicolor subsp. halepense in the dedogram. Subspecies bicolor
is connected with subspecies halepense through some collections of S.
virgatum. Subspecies bicolor further contains two complexes, the
cultivated variety bicolor and weedy varieties aethiopicum, verticilli-
florum and arundinaceum. Variety bicolor is connected to the weed
varieties through S. drummondii which is a semi-domesticated sorghum.
These observations indicate that the affinities between the members
of subsp. halepense and that of subsp. bicolor are less close than the
affinities between the members of the weed varieties and cultivated
sorghums. This is in accord with the results of de Wet and Huckabay
(1967). Since there is no real genetic barrier among the species of
Snowden (1936, 1954), they are recognized as one single heteromorphic
species.

On the basis of gross morphology, distribution, and affinity to
cultivated sorghums, the concept that S. bicolor (Linn.) Moench can
be divided into two subspecies halepense and bicolor is accepted.
The former members are rhizomatous with forty somatic chromosomes. The
latter contains two groups with twenty somatic chromosomes, the weeds
and the cultivars. There are four morphorlogical races, kafir, bicolor,

guinea and durra in the cultivated variety bicolor. The three weedy varieties are var. aethiopicum, var. verticilliflorum and var. arundinaceum.

CHAPTER VI

SUMMARY

An attempt was made to determine the accuracy of the various classification systems presented for Sorghum, by using techniques of numerical taxonomy. The fifty-three characters studied include both quantitative and qualitative ones. These were used to calculate similarity coefficients of the seventy-two collections belonging to twenty-five species as recognized by Snowden. The Sokal and Sneath simple matching coefficient, and Sokal and Michener weighted variable group method of average linkage clustering were applied to this numerical analysis.

The relationships of the seventy-two collections are presented as a dendogram. The collective species Sorghum bicolor (Linn.) Moench, recognized to include all the weed and cultivated sorghums, is divided into two subspecies, Sorghum bicolor subsp. bicolor, and Sorghum bicolor subsp. halepense. Subspecies bicolor was found to contain two groups. One comprises the weed varieties aethiopicum, verticilliflorum and arundinaceum. The other group includes the cultivated variety bicolor. This cultivated complex in turn contains four races, bicolor, durra, kafir and guinea.

LITERATURE CITED

Boeke, J. E. 1942. On quantitative statistical methods in taxonomy; subdivision of a polymorphous species: _Planchonella Sandwicensis_ (Gray) Pierre. Blumea. 5: 47-65.

Cain, A. J. and G. A. Harrison. 1958. An analysis of the taxonomist's judgement of affinity. Proc. Zool. Soc. London 131: 85-98.

Celarier, R. P. 1959. Cytotaxonomy of the Andropogoneas III. Subtribe Sorghae, Genus _Sorghum_. Cytologia 23(4): 395-418.

Clark, P. J. 1952. An extension of the coefficient of divergence for use with multiple characters. Copeia 2: 61-64.

Clayton, W. D. 1961. Proposal to conserve the generic name _Sorghum_ Moench. (Gramineae) versus _Sorghum_ Adans. (Gramineae). Taxon 10(8): 242-243.

de Wet, J. M. J. 1967. The origin of _Sorghum bicolor_. I. Morphological variation. Bot. Gaz. (In Press).

de Wet, J. M. J. and J. P. Huckabay. 1967. The origin of _Sorghum bicolor_. II. Distribution and domestication. Evolution (In Press).

Garber, E. D. 1950. Cytotaxonomic studies in the genus _Sorghum_. Univ. Calif. Publ. Bot. 23: 283-361.

Heincke, F. 1898. Naturgeschichte des Herings. I. Die Lokalformen und die Wanderungen des Herings in den europaischen Meerer. Abh. Deutsch. Seefischerei-Vereins 2: 1-223.

Huskins, C. L. and S. G. Smith. 1932. A cytological study of the genus _Sorghum_ Pers. I. The somatic chromosomes. J. Genet 25: 241-249.

Huskins, C. L. and S. G. Smith. 1932. A cytological study of the genus _Sorghum_ Pers. II. The meiotic chromosomes. J. Genet. 28: 387-395.

Lance, G. N. and W. T. Williams. 1966. Computer programs for hierarchical polythetic classification ("Similarity analyses"). Computer Jour. 9: 60-64.

Michener, C. D. and R. R. Sokal. 1957. A quantitative approach to a problem in classification. Evolution 11: 130-162.

Micks, D. W. and J. P. Ellis. 1952. Free amino acids in the developmental stages of the mosquito. Proc. Soc. Exp. Biol. Med. 79: 191-193.

Nuttall, G. H. F. 1901. The new biological test for blood in relation to zoological classification. Proc. Roy. Soc. Lond. Ser. B 69: 150-153.

Nuttall, G. H. F. 1904. Blood Immunity and Blood Relationship: A Demonstration of Certain Blood-relationships Amongst Animals by Means of the Precipitin Test for Blood. Cambridge University Press, Cambridge.

Pearson, K. 1926. On the coefficient of racial likeness. Biometrika 18: 105-117.

Proom, H. and A. J. Woiwod. 1949. The examination, by partition paper chromatography, of the nitrogen metabolism of bacteria. J. Gen. Microbiol. 3: 319-327.

Randall, H. M., D. W. Smith, A. C. Colm and W. J. Nungester. 1951. Correlation of biologic properties of strains of Mycobacterium with infra-red spectrums. I. Reproducibility of extracts of M. tuberculosis as determined by infra-red spectroscopy. Am. Rev. Tuberc. 63: 373-380.

Robinson, W. S. 1951. A method for chronologically ordering archaeological deposits. Am. Antiquity 16: 293-301.

Rogers, D. J. and T. T. Tanimoto. 1960. A computer program for classifying plants. Science 132: 1115-1118.

Rohlf, F. J. and R. R. Sokal. 1965. Coefficients of correlation and distance in numerical taxonomy. Univ. Kansas Sci. Bull. 45: 3-27.

Smirnov, E. S. 1960. Taxonomic analysis of a genus. (In Russian, English summary.) Zhurnal Obshenye Biologii 21: 89-103.

Sneath, P. H. A. 1957. Some thoughts on bacterial classification. J. Gen. Microbiol. 17: 184-200.

Snowden, J. D. 1936. The Cultivated Races of Sorghum. Adlard and Son London.

Snowden, J. D. 1954. The wild fodder Sorghums of the section Eusorghum. J. Linn. Soc. Lond. 55: 191-260.

Sokal, R. R. 1958. Quantification of systematic relationships and of phylogenetic trends. Proc. Xth Intern. Cong. Entomol. 1: 409-415.

Sokal, R. R. 1961. Distance as a measure of taxonomic similarity. Syst. Zool. 10: 70-79.

Sokal, R. R. and C. D. Michener. 1958. A statistical method for evaluating systematic relationships. Univ. Kansas Sci. Bull. 38: 1409-1438.

Sokal, R. R. and P. H. A. Sneath. 1963. _Principles of Numerical Taxonomy_. W. H. Freeman, San Francisco.

Wirth, M., G. F. Estabrook and D. J. Rogers. 1966. A graph theory model for systematic biology, with an example for the _Oncidiinae_ (Orchidaceae). Syst. Zool. 15(1): 59-69.

Yule, G. V. and M. G. Kendall. 1950. _An Introduction to the Theory of Statistics_. 14th ed., Hafner, New York.

VITA

Margaret Hsuan Lu

Candidate for the Degree of

Master of Science

Thesis: NUMERICAL TAXONOMY OF WEED AND CULTIVATED SORGHUMS

Major Field: Botany and Plant Pathology

Biographical:

Personal Data: Born in Chung-ching, China, May 1, 1942, the
    daughter of Chang-Yin and Mu-Yuin Wang Lu.

Education: Obtained the degree of Bachelor of Science in
    Agricultural College from the National Taiwan University
    with a major in Agronomy, in June, 1964; completed the
    requirement for the Master of Science degree in July,
    1967.

Professional experience: Graduate Research Assistant, August,
    1964 to May, 1965; a Replacement Graduate Teaching
    Assistant, March-May, 1965, in the Agronomy Department,
    National Taiwan University; Graduate Research Assistant,
    June, 1965-August, 1967, in Botany and Plant Pathology
    Department, Oklahoma State University.

Member of: American Genetic Association, American Institute of
    Biological Sciences.