Hunt Institute for Botanical Documentation
5th Floor, Hunt Library
Carnegie Mellon University
4909 Frew Street
Pittsburgh, PA 15213-3890
Telephone: 412-268-2434
Email: huntinst@andrew.cmu.edu
Web site: www.huntbotanical.org

Department of Forestry,
Commonwealth Forestry
Institute,
University of Oxford.

Your ref: DJR/ch

16th February, 1967.

Dear Dave,

Thanks for your comments of 6th February.

I agree it would have been useful to discuss the data preparation before doing the analysis, but on the other hand I don't think the suggestions about character scoring in your latest paper would have influenced a very high proportion of 30 scored in the Meliaceae. The data is almost identical to that used in the factor analysis.

At present I am writing a fuller account of the analysis and should be grateful if you would be prepared to criticise it in due course.

Mr. White thanks you for your kind offer to send instructions for using your computer programme in Oxford, and would be pleased to receive these. He will be writing to you in the near future. There is also an agricultural botanist here interested in classifying soil types who has expressed a similar wish. So if it is not inconvenient, perhaps you could send the necessary instructions fairly soon.

Thanks for your help.

Yours sincerely,

T.D. Pennington.

Professor D.J. Rogers,
Taximetrics Laboratory,
Dept. of Botany & Plant Pathology,
Colorado State University,
Fort Collins,
Colorado 80521.

Forest Herbarium
Dept. of Forestry
University of Oxford
England.

Nov. 23rd 1966

Dear Dave,

Thank you for your letter of Oct. 21st. I have recently received the weighty package you mention, and will work through the data in the near future. When I have completed this I'll write at length on your results. Iain, who is fairly familiar with variation pattern in the <u>Meliaceae</u>, said ~~that~~ your method would be most suitable for this type of problem, so I have great hopes. I put the same data through a Principal Component Analysis, but with rather poor results.

Yours sincerely

Terry Pennington

- Taximetrics Laboratory

February 6, 1967

Dr. T. D. Pennington
Department of Forestry
Commonwealth Forestry Institute
University of Oxford
Oxford, England

Dear Terry:

We have received your interesting report on the computer analysis of the Meliaceae data. I believe there are several pertinent comments which may clarify some areas. We suffer from the fact that we can't communicate face to face. I am sure that had we been able to talk with you during the data preparation, we would have had a better chance of reflecting your knowledge and interests regarding Meliaceae. You should have had at your disposal the mimeographed paper ("A general method of taxonomic description for a computed similarity measure") we sent after we had done the computing before commencing the data preparation. That paper presented the ideas necessary to frame your data in terms more likely to produce meaningful results with our methodology.

We assume that the data that you presented to us was a reworking of the information presented to the program for factor analysis. It is too bad that your input to our program was so structured because our methodology is intended to do the least damage to taxonomic data, whereas structuring the information for factor analysis forces upon the taxonomist some unhappy data structuring.

We should definitely emphasize that we do not intend the computer program to be a dictator. We intend rather that the results of the computer's run is only a hint to you, the taxonomist. It is not intended that you take this run on the computer and consider it to be the final answer. Rather, it is quite likely that the best use of this machinery is as a tester of ideas and as a methodology to suggest other organization of the data. In your particular case, much of the information probably applies rather specifically to certain groups within the Meliaceae, and there is little connection between "private information" of one group and that of another. The computer does not know this unless it is instructed. It cannot do anything precise about it. It is also the case that unless you put information in that you would like for it to use, you cannot expect to get back the desired information. It is also true that the value of the information in classification is known by you and not by the computer. If you instruct the computer to use valueless information for the differentiation (or clustering), your output will be valueless. I mean by that statement that if you attempt to distinguish species within genera by using information only about genera, and vice versa, it is unlikely that you will succeed with the computer.

There are many different ways in which the print-out is useful. You claim that you find it impractical to describe the differences between genera with large numbers of objects using large numbers of descriptive states. If you had taken the first part of your print-out, namely, the presentation of the states describing each of the objects and listed the attributes of the objects which fell together in one cluster, you would quickly be able to discover that some of your characters were "carrying the load" as the descriptive agents holding that group together and that the different non-common information listed on the print-out were your "key characters." I suggest that had you used the machinery print-out that way, your complaint that multiple state differentiation would not be apparent.

Another note in interpretation of the relationships found is that all objects are put together first as pairs of objects. This being the case other objects joining these are likewise done in pairwise manner. This being the case, it is difficult to put together objects described by large numbers of different states. The overall similarity measure is a powerful tool. But as with any new instrument, its use takes a considerable amount of effort. We have found it useful, but only after much work to understand what is going on. If you expect to receive a monograph from the computer, you will be sadly disappointed, but if you use it as a tool, you can gain tremendous new insight into taxonomy.

Much as we would like to, we simply cannot accept any data from Mr. White or his students. We will be glad to send instructions to him to convey to his computer center (there must be one available) which can then make your data processing more convenient. As you can see, we spend much time in preparing your data before it is machine-ready. When we prepare it, we are never certain that we have done justice to your data. For these reasons, we prefer, if possible, to get our programs transferred to your computer, then you have the opportunity to experiment with the program much more readily. If Mr. White desires, we will send all the necessary programming information to him, and he can have your own programmer put it onto your computer.

Sincerely yours,

David J. Rogers
Professor of Botany

DJR:ch

TELEPHONE:
OXFORD 57891

DEPARTMENT OF FORESTRY
COMMONWEALTH FORESTRY INSTITUTE
UNIVERSITY OF OXFORD

21st January, 1967.

Dear Dave,

I have now had sufficient time to thoroughly digest your analysis of the Meliaceae subfamily Melioideae, and to discuss the results with my chief, Frank White, who is interested in your method.

Before going into the results in detail, it will help if I explain the variation patterns found in this subfamily. I made a thorough study of generic limits in this group for my thesis, which resulted in some changes, both in the circumscription and number of genera, in the latest account of the family (Harms, 1940).

In my thesis, the subfamily was divided into 35 genera placed in 7 tribes. The tribes are defined by reference to a number of overlapping morphological, anatomical and palynological characters, which in combination provide discontinuities at this level.

Below is a summary of the 7 tribes and their contents:

I.	Turraeeae.	Naregamia, Munronia, Nurmonia, Turraea,* Quivisia, Calodecaryia, Nymania.
II.	Melieae.	Melia, Azadirachta.
III.	Vavaeeae.	Vavaea.
IV.	Trichilieae.	Trichilia,* Pseudobersama, Pterorhachis, Walsura, Lepidotrichilia, Malleastrum, Ekebergia, Astrotrichilia, Owenia, Capuronianthus, Cipadessa.
V.	Aglaieae.	Aglaia,* Lansium, Aphanamixis, Sphaerosacme.
VI.	Guareeae.	Heckeldora, Cabralea, Ruagea, Turraeanthus, Guarea, Chisocheton, Megaphyllaea, Synoum, Anthocarapa, Pseudocarapa, Dysoxylum.*
VII.	Sandoriceae.	Sandoricum.

Contd./...

Within each tribe the pattern of variation is one of two basic types.

1. A tribe contains a single isolated genus or a pair of closely related genera which are isolated, and without any close relatives in the family. The tribes Melieae, Vavaeeae and Sandoriceae belong here.

2. A tribe contains a central large genus surrounded by a number of usually small "satellite" genera which are often closely related to it. The large genera in the 4 tribes concerned are asterisked. Between members of this group there are some important differences in detail.

In 3 of the tribes, the Turraeeae, Trichilieae and Guareeae, the central large genus is variable in many characters, with many structures varying independently, and because of this it cannot be satisfactorily split. They do however contain at least one unifying feature enabling them to be clearly distinguished from all closely related "satellite" genera.

The "satellite" genera in these tribes are defined by a particular combination of characters, most of which are found individually in different combinations elsewhere in the family. Such characters correlated in only small areas of the family, give rise to low correlation values when considered in relation to their distribution throughout the whole family.

In the remaining tribe, the Aglaieae, the variation pattern is somewhat different, as both the large genus Aglaia and its small "satellites" are fairly homogeneous, and the "satellites" are distinguished by a combination of unique characters unknown elsewhere in the family.

The results of the Principal Component Analysis and your Cluster Analysis are more easily interpreted in the light of this information.

The Principal Component Analysis

As the overall correlations of the majority of characters is rather low, this analysis gave generally poor results. The one exception is, as expected, the genera of the Aglaieae, which are clearly distinguished. In this group the result agrees in most details with my classification. 6 components were extracted accounting for 56% of the variation, and apart from the genera of the Aglaieae, very few genera from the other tribes are distinguished from the mass which contains representatives of many genera.

I tried computing the 4 large tribes independently with much better results. In this case most genera in both Aglaieae and Turraeeae came out, but again with disappointing results in the Trichilieae and Guareeae.

Contd./...

The Rogers Cluster Analysis

1. The isolated genera and isolated pairs.

The isolation of Sandoricum is clearly shown, and also that of Melia and Azadirachta though the limits of these two are confused. 3 of the 4 specis of Vavaea cluster at the 6th level. The 4th species would have clustered with them save for an error in theoriginal scoring.

2. The "satellite" genera.

The taxonomic position of 20 "satellites" is shown quite well. Most of these form discreet clusters at least at one similarity level; others do not receive their more isolated members until the genus has formed links with other genera.

A further 4 genera in this category would have been more clearly picked out if I had scored additional characters.

3. The "Big" genera. Turraea, Trichilia, Aglaia, Dysoxylum.

(a) The clustering of Aglaia (51 species) and its "satellites" Lansium and Aphanamixis, corresponds almost completely with my own generic limits. The points of closest relationship between these genera are also clearly indicated. In this tribe the clusters based on the degree of overall similarity are correlated with those discontinuities in the variation pattern which distinguish the genera. There is only one error in the clustering of this tribe; this is the monotyoic genus Sphaerosacme which is linked to Trichilia. A better clustering of this genus might be obtained if additional pollen characters were used.

(b) Turraea (10 species). 8 species are isolated as a cluster at Level 5, but at Level 6 when the 9th species joins, the genus links with Munronia and Quivisia. The last species to join is very isolated and does not come in until Level 9.

(c) Trichilia (36 species). At the 3rd Level with 21 single members still to join, Trichilia is divided into 3 groups: one of 4 species; one of 2 species; one of 9 species linked to 4 species of Walsura. These 3 clusters unite at the 5th Level and are also joined by a further 8 Trichilia species and Pseudobersama (monotypic). At this level there are 13 single members. These remaining Trichilias join at the 6th, 7th, 8th and 9th Levels, by which time the cluster embraces virtually the whole family and Trichilia has formed many links with members of other genera (e.g. at the 6th Level there are links with Dysoxylum, Owenia and Synoum.

(d) Dysoxylum (38 species). The limits of this genus are also confused in a similar way. At the 3rd Level the genus appears as 1 pair of species and a cluster of 10 species (26 single members). At Level 5 there are 2 isolated pairs, a group of 17 species linked to 10 species of Guarea, and 17 single members. At Level 6, 31 species are in a single cluster, but this also has links with Trichilia, Synoum Aphanamixis and Ruagea; another single species forms an isolated pair with Turraeanthus, and the remaining 6 species are single members. The last single member joins at Level 9.

Contd./...

In the case of the large genera, Turraea, Trichilia and Dysoxylum, the generic limits as defined by me do not correspond at any level of similarity to your clusters based on the degree of overall similarity. But the new arrangement of the species suggested by your analysis is not an improvement on my own because:

1. If the genera were based on clusters formed at a high level of similarity, then you maintain a small number of genera containing 2-15 species and a multitude of single species.

2. If the genera were based on clusters formed at a lower level of similarity, they would not only be inconveniently large, but they would also be extremely difficult to diagnose and almost impossible to key out, because they can be defined only by reference to a very large number of characters.

3. If the genera are based on clusters formed at a low level of similarity, then other related genera, which are at present well-defined, would become less distinct and more difficult to distinguish.

The single linkage method of clustering has certain disadvantages when applied to a large and heterogeneous group like the Trichilieae.

The criterion of a Rogers' Cluster is basically that a member of the cluster shares a certain number of characters with at least one other member, irrespective of what the characters are, and it shares less than this number of characters with any non-member. For example, 10 species are compared for 10 characters, and each may have 9 in common with one other species in the group, but the 10th differing character may be different for each species. To define such a group one must refer to all 10 characters.

Applied to the type of variation pattern found in the Trichilieae, this method causes clusters based on overall similarity to be quite cumbersome and difficult to define. For instance at Level 5, there emerges a group of 28 species, comprising 23 Trichilia, 4 Walsura and Pseudobersama: to define this one must refer to 22 characters. The ultimate test of a good classification is that it should be practicable, and these clusters do not pass the test. How would you key out such a cluster?

The alternative when attempting to classify a heterogeneous group like the Trichilieae is not to pay so much attention to overall similarity as to look for any discontinuities in the variation. These exist in the Trichilieae, based primarily on the nature of the fruit (capsule/berry/drupe). The genera Trichilia, Walsura and Ekebergia are separated by differences in the type of fruit. This grouping of the species has several advantages. The fruit characters give the genera internal unity, external discontinuity, and they are therefore easy to diagnose and can be relatively simply keyed out.

It may be that phylogenetic evidence showing that 4 species of Walsura really do belong in Trichilia will one day be discovered, but at present there is no such evidence, and until there is, the present workable classification based on the discontinuities I have mentioned must be preferred.

Contd./...

This is not the artificial weighting of a few arbitrarily chosen characters but a classification which has gradually evolved by gradual improvement of the previous classification of a group in which all members are recognised as being very similar, but which nevertheless can be conveniently split at the generic level using the discontinuous patterns of variation among a few characters.

The pattern of variation in the Guareeae is similar to that in the Trichilieae. The heterogeneous Dysoxylum is closely allied to Guarea but easily separated from it by the discontinuity in 3 characters. This distinction is blurred when the two genera are compared for overall similarity, but the new clustering suggested by your analysis is of little practical significance.

To summarise: (i) Your method picks out the isolated genera and "satellite" genera, and indicates their relationships fairly well.

(ii) In the Aglaieae, which is homogeneous except for those characters which provide the discontinuities between Aglaia and its relatives, the results of the cluster analysis also correspond closely to mine. For this case the discontinuities are correlated with the distinctions between groups based on overall similarity.

(iii) In the remaining large tribes where the central genus is variable in many characters the clusters suggested by this analysis do not correspond to my generic limits.

Although these results do not altogether coincide with the generic limits as defined by me, they have caused me to look very closely at the patterns of variation in the subfamily and this has given me a much clearer insight into the make-up of the group. I should mention of course that these comments are only my interpretation of the results and I should welcome any criticisms of them from you and your men. If you would like me to send fuller details of the stage-by-stage clustering, do let me know.

With best wishes,

yours sincerely,

Terry Pennington

T.D. Pennington.

Professor D.J. Rogers,
Taxonomy Laboratory,
Dept. of Botany & Plant Pathology,
Colorado State University,
Fort Collins, Colorado 80521.

P.S. Mr. White who is particularly interested in the application of numerical taxonomy to practical (not theoretical) problems in taxonomy has asked me to add this postcript. He suggests that your method might be of great value in groups like Leguminosae: Papilionatae where the problems chiefly arise from the large number of species involved and the large number of characters concerned, but where genera are often remarkably constant in many diagnostic characters.

In a few weeks time he will be giving a course of lectures and practical classes on modern methods in taxonomy and would like to get the students to investigate patterns of variation at the generic level in this group. If his students are sufficiently responsive and reliable would you be interested in having the data they record for a comparative analysis?

This is a copy of the species list in case you haven't your own.

J.R.

Computer No.	Your Number.
1	Munronia A 1
2	2
3	3
4	4
5	5
6	6
7	7
8	8
9	Naregamia B 1
10	Nurmonia C 1
11	Turraea D 1
12	2
13	3
14	4
15	5
16	6
17	7
18	8
19	9
20	10
21	Quivisia EE 1
22	2
23	Calodecaryia FF 1
24	Nymania GG 1
25	Azadirachta H 1
26	Melia I 1
27	2
28	3
29	Trichilia J 1
30	2
31	3
32	4
33	5
34	6
35	7
36	8
37	9
38	10
39	11
40	12
41	13
42	14
43	15
44	16
45	17
46	Trichilia J 18
47	19
48	20
49	21
50	22
51	23
52	24
53	25
54	26
55	27
56	28
57	29
58	30
59	31
60	32
61	33
62	34
63	35
64	36
65	Pseudobersama K 1
66	Pterorhachis L 1
67	Lepidotrichilia M 1
68	2
69	3
70	4
71	Heynea N 1
72	Walsura O 1
73	2
74	3
75	4
76	5
77	6
78	7
79	8
80	Malleastrum P 1
81	2
82	3
83	Ekebergia Q 1
84	2
85	3
86	4
87	Owenia R 1
88	2
89	3
90	4
91	Astrotrichilia 1

Digitized by the Hunt Institute for Botanical Documentation

Computer No.	~~Your~~ My Number.		Computer No.	~~Your~~ My Number.	2
92	Astrotrichilia	2	137	Dysoxylum	X 37
93		3	138		38
94	Cipadessa	T 1	139	Turraeanthus	Y 1
95	Sandoricum	U 1	140	Megaphyllaea	Z 1
96		2	141	Chisocheton	1
97		3	142		2
98	Pseudocarapa	V 1	143		3
99	Anthocarapa	W 1	144		4
100		2	145		5
101	Dysoxylum	X 1	146		6
102		2	147		7
103		3	148		8
104		4	149		9
105		5	150		10
106		6	151		11
107		7	152		12
108		8	153		13
109		9	154		14
110		10	155		14
111		11	156		15
112		12	157		16
113		13	158		17
114		14	159		18
115		15	160		19
116		16	161		20
117		17	162		21
118		18	163	Cabralea	1
119		19	164		2
120		20	165		3
121		21	166	Ruagea	1
122		22	167		2
123		23	168		3
124		24	169	Synoum	1
125		25	170	Hecheldora	1
126		26	171	Leplaea	1
127		27	172	Guarea	1
128		28	173		2
129		29	174		3
130		30	175		4
131		31	176		5
132		32	177		6
133		33	178		7
134		34	179		8
135		35	180		9
136		36	181		10

Computer No.		~~Your~~ M. Number.
182	Guarea	10
183		11
184		12
185		13
186		14
187		15
188	Sphaerosacme	1
189	Lansium	1
190		2
191		3
192		4
193		5
194		6
195	Aphanamixis	1
196		2
197		3
198		4
199		5
200	Aglaia	1
201		2
202		3
203		4
204		5
205		6
206		7
207		8
208		9
209		10
210		11
211		12
212		13
213		13
214		14
215		15
216		16
217		17
218		18
219		19
220		20
221		21
222		22
223		23
224		24
225		25
226		26

Computer No.		~~Your~~ M. Number.
227	Aglaia	27
228		28
229		29
230		30
231		31
232		32
233		33
234		34
235		35
236		36
237		37
238		38
239		39
240		40
241		41
242		42
243		43
244		44
245		45
246		45
247		46
248		47
249		48
250		49
251		50
252		51
253	Vavaea	1
254		2
255		3
256		4
257	Capuronianthus	1

- Taxonomy Laboratory

October 21, 1966

Dr. T. D. Pennington
Forest Herbarium
Commonwealth Forestry Institute
South Parks Rd.
Oxford, United Kingdom

Dear Dr. Pennington:

I am sorry that we have taken so long to return the results of our computer analysis of Meliaceae. I shan't bore you with the multitudinous reasons for the delay, some valid, others not so. However, today we are sending under separate cover, and by surface mail, a rather weighty package. I might mention one problem--we simply did not know your address and had to write to Prance, in Brazil, to find out where you are. His reply came only yesterday.

You will receive several documents in the package. These are:

(1) The computer print-out.
(2) A mimeographed paper on "How to Read the Printout"
(3) A piece of brown "butcher paper" with two levels of clustering done for you, as an example of the next step in the work.
(4) A reprint of a paper from Systematic Zoology that you should thoroughly digest in order to understand our method.
(5) A mimeographed preprint of a paper which will appear in next month's issue of BioScience, and should be of interest to you.

Item 1, above, has the following included information:

a. Title of the job, and the number of characters and objects.

b. Three lines of zeros, meaningless to you.

c. A line stating "No Matrices in this Study", again of no meaning to you.

d. Main input data. The listing under this heading allows you to check all characters used in the study, for each object (specimen or species). PLEASE NOTE that, for technical reasons, we had to record any attribute in any character with more than 9 attributes as letters. Thus, attribute 10 in any character is listed as "A", attribute 11 as "B", etc. This should not be confusing.

e. Immediately following the list of input data is a list of objects found to have been coded as identical to other objects. Since they are identical, they are not included in the following part. In the list of identicals you may determine which objects are identical. In the second column are listed the identical objects you will find omitted in the clustering analysis. Do not forget these objects after the study is completed.

f. Next in order is a long table of similarity measures, for each object with every other object. This will be useful to you in interpretation of the clustering results. Note that where you have two objects listed, to find a similarity measure between them, look for the higher numbered object first. Thus, for the similarity between objects 63 and 162, look in the listing where 162 appears, then follow along until you find 63 listed with it. It will appear as follows:

162 63 0.67742 31

The last two numerals, 31, indicate that specimens 162 and 63 were compared on 31 characters.

g. Following the very long table of similarity measures, about 3/4 of the way through the total printout, you will finally find the part that interests you most, namely the clustering analysis. It is at this point that document numbered (2) above will have instructions for you. Even with the computer's analysis, you have much work left to do. However, we have found that this work is extremely instructive, and something the investigator can do better than anyone else. It is unfortunate that so much distance separates us, because we could give you aid and comfort as you proceed with the job. The third document listed above will give you some of the work already, to indicate how you should proceed.

We find that your data is extremely well "structured". This means merely that there are some very clean distinctions between the taxa. Whether the differentiation is pleasing to you or not, is something which we cannot yet determine. Certainly, your work at the family level is similar to Iain's, but shows quite different structure than did that of Irwin on a section of Cassia, and quite significantly different from the intertwined, reticulate structure we find in the cultivars of Manihot esculenta.

We await your reactions to the results with considerable concern. Recall that we entered this work with the idea that we would learn from your data something of the value of the method. We may find new directions from your evaluation of the results which will aid in its improvement. Therefore, please do not consider that there is nothing more than can be done, and this method is necessarily the best, or worst. You can give us assistance in being completely candid about these results. Any comment will be considered by us, and this will be our "pay off."

Sincerely yours,

David J. Rogers
Professor of Botany

DJR:ch

Computer Number	Your Number			IDENTITY OF NUMBERS			
1	Munronia	A	1	32	Trichilia	J	4
2			2	33			5
3			3	34			6
4			4	35			7
5			5	36			8
6			6	37			9
7			7	38			10
8			8	39			11
9	Naregamia	B	1	40			12
10	Nurmonia	C	1	41			13
11	Turraea	D	1	42			14
12			2	43			15
13			3	44			16
14			4	45			17
15			5	46			18
16			6	47			19
17			7	48			20
18			8	49			21
19			9	50			22
20			10	51			23
21	Quivisia	EE	1	52			24
22			2	53			25
23	Calodecarya	FF	1	54			26
24	Nymania	GG	1	55			27
25	Azadirachta	H	1	56			28
26	Melia	I	1	57			29
27			2	58			30
28			3	59			31
29	Trichilia	J	1	60			32
30			2	61			33
31			3	62			34

63	Trichilia	J 35	94	Cipadessa	T 1
64		36	95	Sandoricum	U 1
65	Pseudobersama	K 1	96		2
66	Pterorhachis	L 1	97		3
67	Lepidotrichilia	M 1	98	Pseudocarapa	V 1
68		2	99	Anthocarapa	W 1
69		3	100		2
70		4	101	Dysoxylum	X 1
71	Heynea	N 1	102		2
72	Walsura	O 1	103		3
73		2	104		4
74		3	105		5
75		4	106		6
76		5	107		7
77		6	108		8
78		7	109		9
79		8	110		10
80	Malleastrum	P 1	111		11
81		2	112		12
82		3	113		13
83	Ekebergia	Q 1	114		14
84		2	115		15
85		3	116		16
86		4	117		17
87	Owenia	R 1	118		18
88		2	119		19
89		3	120		20
90		4	121		21
91	Astrotrichilia	1	122		22
92		2	123		23
93		3	124		24

125	Dysoxylum	X 25	156	Chisocheton	15
126		26	157		16
127		27	158		17
128		28	159		18
129		29	160		19
130		30	161		20
131		31	162		21
132		32	163	Cabralea	1
133		33	164		2
134		34	165		3
135		35	166	Ruagea	1
136		36	167		2
137		37	168		3
138		38	169	Synoum	1
139	Turraeanthus	Y 1	170	Heckeldora	1
140	Megaphyllaea	Z 1	171	Leplaea	1
141	Chisocheton	1	172	Guarea	1
142		2	173		2
143		3	174		3
144		4	175		4
145		5	176		5
146		6	177		6
147		7	178		7
148		8	179		8
149		9	180		9
150		10	181	(Accidental Repeat)	10
151		11	182		10
152		12	183		11
153		13	184		12
154	(Accidental Repeat)	14	185		13
155		14	186		14

187	Guarea	15	218	Aglaia	18
188	Sphaerosacme	1	219		19
189	Lansium	1	220		20
190		2	221		21
191		3	222		22
192		4	223		23
193		5	224		24
194		6	225		25
195	Aphanamixis	1	226		26
196		2	227		27
197		3	228		28
198		4	229		29
199		5	230		30
200	Aglaia	1	231		31
201		2	232		32
202		3	233		33
203		4	234		34
204		5	235		35
205		6	236		36
206		7	237		37
207		8	238		38
208		9	239		39
209		10	240		40
210		11	241		41
211		12	242		42
212	(Accidental Repeat)	13	243		43
213		13	244		44
214		14	245	(Repeat!)	45
215		15	246		45
216		16	247		46
217		17	248		47

249	Aglaia	48
250		49
251		50
252		51
253	Vavea	1
254		2
255		3
256		4
257	Capuronianthus	1

patelliform fused to staminal tube (8); free cyathiform (9); cyathiform fused to ovary (10); cyathiform fused to staminal tube (11); tubular (12).

85-96 Gynoecium 21. Number of locules 1(1); 1-2(2); 2(3); 2-3(4); 3(5); 3-4(6); 4(7); 4-5(8); 5(9); 5-10(10); 10(11); 10-15(12).

97-104 22. Ovules solitary (1); 1-2 collateral (2); 1-2 superposed (3); 2 collateral (4); 2 superposed (5); more than 2(6); 2 oblique (7); more than 2 with parietal placentas (8).

105-109 23. Style-head capitate (1); discoid (2); conical (3); receptaculum pollinis (4); deeply lobed (5).

110-113 Fruit 24. Capsule (1); berry (2); drupe (3); nut (4).

114-120 Seed 25. With fleshy aril (1); membraneous aril (Cabralea-type (2);
114-119+1 membraneous aril (Munronia-type (3); fleshy sarcotesta (4); mealy sarcotesta (5); basal, swollen sarcotesta (6).

121-122 26. Endosperm present (1); absent (2).

123-124 27. Cotyledons flat (1); plano-convex (2).

125-128 28. Cotyledons collateral (1); oblique (2); superposed (3); variable (4).

129-134 29. Radicle superior (1); abaxial (2); ± central & included (3); inferior (4); variable (5); adaxial (6).

135-136 30. Cotyledons free (1); fused (2).

137-138 31.

Attributes

Characters used in the Principal Component Analysis

1-4 **Leaf** 1. Simple (1); unifoliolate (2); trifoliolate or pinnate (3); bipinnate (4).

5-6 2. Indeterminate apex (1); terminal leaflet or none (2).

7-8 3. Indumentum simple (1); or stellate (2).

9-11 **Calyx** 4. Entire margin (1); or lobed (2); both entire and lobed (3).

12-14 5. Lobes imbricate (1); or not (2); both (3).

15-16 6. Calyx closed in bud and circumscissile at the base (1); or not (2).

17-22 **Corolla** 7. No. of petals 3(1); 3-4(2); 4(3); 4-5(4); 5(5); 5-~~6~~-10(6); ~~6-10(7).~~

23-25 8. Free (1); united (2); both (3).

26-28 9. Free (1); fused to staminal tube above ½ (2); fused at base (3).

29-31 10. Contorted or imbricate (1); valvate (2); both (3).

32-36 **Androecium** * 11. Ratio of length to breadth of staminal tube. ·5-1 (32); 1.05-2 (33); 2.05-5 (34); 5.05-10 (35); 10.05-25 (36).

37-42 12. Appendages 4 x as many as anthers (1); 2 x as many (2); as many as (3); absent (4); as many as to 2 x as many (5); 2 x to 4 x as many (6).

43-56 13. Number of anthers 4-5 (1); 5-6(2); 6(3); 6-7(4); 7(5); 7-8(6); 8(7); 8-9(8); 9(9); 9-10(10); 10(11); 10-15(12); 16-20(13); 21-25(14).

57-59 14. Anthers hairy (1); glabrous (2); both (3).

60-61 15. Anthers locellate (1); not locellate (2).

62-63 16. Anthers with appendages (1); without appendages (2).

64-66 17. Anthers inserted apically (1); partly exserted (2); completely included (3).

~~67~~ 67-70 18. Filaments completely fused (1); fused in upper ½ (2); fused in lower ½ (3); fused in upper or lower ½ (4).

71-72 19. Anthers in 1 whorl (1); anthers in 2 whorls (2).

73-84 **Disk** 20. Absent (1); stipitate (2); free annular (3); annular, fused to ovary (4); annular fused to staminal tube (5); free patelliform (6); patelliform fused to ovary (7);

* [illegible]

THE NEW YORK BOTANICAL GARDEN
BRONX PARK • BRONX 58 • NEW YORK

June 3, 1966

Dear Dave,

Here is the Pennington Data on your worksheets. I hope it is alright. If there are any questions I will be back in New York in 3 weeks time to answer them.

With best wishes

Jairo