CHAPTER 4

NUMERICAL STUDY

Introduction

Biological classification is confined to a grouping of organisms base mainly on their structures. From major categories, such as Division, Class down to minor categories, i.e. genus, species and infraspecific taxa (Clifford and Stephenson, 1975). In contrast, numerical taxonomy is the grouping of operation taxonomic unit (OTU) into taxa by numerical methods on the basis of their quantitative and qualitative characters (Sneath & Sokal, 1973). It has been used to classify various groups of organisms and to elucidate the problems in classification. Sneath and Sokal (1973) have pointed out the advantages of numerical taxonomy over conventional taxonomy as followed:-

"it has the power to integrate data from various sources, such as morphology, physiology, chemistry, amino acid sequences of protein and more. In numerical taxonomy, greater efficiency are promoted by automation of taxonomic process. Therefore, much taxonomic work can be done by less highly skilled workers. It can be used for re-examine the principles of taxonomy and of the purposes of classification" (Sneath & Sokal, 1973).

The general problem of variations in natural populations is that the taxonomic unit was difficultly classified by conventional taxonomy. This delimitation may be improved by methods of numerical taxonomy (Sneath and Sokal 1973). Previously, numerical taxonomy have been used in classification and in recognizing species and infraspecific taxa of many species complexes, examples included Thompson and Lammers (1997); Speer and Hilu (1998); Kephart et al. (1999); Boonkerd, Saengmanee and Baum (2002). Furthermore, it has been used in recognizing a new taxonomic unit or a new combination in species complex (Forster and Liddle, 1991; Chatrou, 1997; Ohta and Takamiya, 1999; Mikkelsen and Seberg, 2001).

In this thesis, Cluster Analysis (CA) and Canonical Discriminant Analysis

(CDA) were used to solve classification problems in "*H. parasitica* complex" in Thailand. Cluster analysis used to place individual specimens into groups. A subset of characters that maximized differences among the groups determined by cluster analysis and grouping based on morphological and anatomical data (see Chapter 3) was selected by stepwise discriminant analysis. The objectives of this study thus were: 1) to determine the important of morphological and anatomical characters that contribute to the discrimination among them, and 2) to clarify the taxonomic status of the nine forms previously recognized based on classical classification in the "*H. parasitica* complex" using morphometric multivariate analyses.

Materials and Methods

1. Specimen collections

Specimens of the nine forms of "*H. parasitica* complex", previously recognized by classical taxonomy were used for morphometric analyses. All specimens were gathered from 50 localities in their natural habitats throughout Thailand (Figure 3.1 and Table 3.1) during 2003 to 2004. Specimens of each form including leaves and flowers were preserved in 70% ethanol. All measurements of macroscopic characters were carried out using electronic digital caliper (Keiba, model 3) and those of microscopic characters were carried out using micrometer under light microscope.

2. Data analysis

Five hundreds and thirty-four specimens were used for all analyses. In general, thirty-five quantitative characters of both vegetative and reproductive characters (Figure 4.1 and Table 4.1) were subjected to discriminant and cluster analyses.

2.1 Cluster analysis, agglomerative, hierarchical and nested (SAHN) clustering (Sneath and Sokal, 1973) were performed using average taxonomic distance and the unweighted pair-group method with arithmetic averages (UPGMA) implemented in NTSYS-pc package version 2.10m (Rohlf, 2000) to place individual specimens into groups. To reduce the effects of different scales of measurement for different characters, the values for each character were standardized using procedure STAND.

2.2 A subset of characters that maximized differences among the groups determined by cluster analysis was selected by stepwise discriminant analysis. To characterize mean differences among the species, canonical discriminant analysis was used to acquire insight into group differences and to estimate character weights from correlations between canonical variables and original variables. Procedure CLASSIFY in SPSS/PC for Windows, release 10.0 (Anonymous, 1999) was used to analyze a set of discriminant analysis.

In this analysis, three grouping criteria were used. Firstly, the 534 specimens were divided into 9 categories (Forms I-IX) according to morphological and anatomical diagnostic in chapter 3. Secondly, Forms I-IV were omitted. Only Forms VI-IX were considered whose members were very closely related, the specimens gathered from these four forms were numerous, thus encompassed most kind of variation provided a data set for robust statistical analysis. In this criterion, these categories according to the morphological and anatomical studies in chapter 3 were defined. Thirdly, grouping criterion was restricted to a priori result from the cluster analysis.

Table 4.1 Thirty five quantitative characters with their methods of scoring used in the study of "*H. parasitica* complex".

Abbreviation	Characters
LL	leaf length in cm
LW	leaf width in cm
DBL	distance from base to the widest point of leaf in cm
LS	leaf shape (calculated by DBL/LL)
PETL	petiole length in cm
PETW	petiole width in cm
PECL	pedicel length in mm
PECW	pedicel width in mm
SPL	sepal length in mm
SPW	sepal width in mm
DCO	diameter of corolla in mm
COL	corolla length in mm
COLL	corolla lobe length in mm
DCOT	diameter of corolla tube in mm
DBCL	distance from base to the widest point of corolla lobe in mm
COAL	corolla lobe apex length in mm
COTL	corolla tube length in mm
COLW	corolla lobe width in mm
COBW	corolla lobe base width in mm
DCN	diameter of corona in mm
DCNR	diameter of coronal receptacle in mm
DCOR	distance from corpuscle to the outer point of receptacle in mm
DCNL	distance between corona lobes in mm
CNLL	corona lobe length in mm
DBCN	distance from base to the widest point of corona lobe in mm
CNLS	corona lobe shape (calculated by DBCN/CNLL)
CNLW	corona lobe width in mm
RCRD	ratio of corona diameter and receptacle diameter
RCCD	ratio of corona diameter and corolla tube diameter
POLL	pollinium length in micron
POW	pollinium width in micron
COPL	corpusculum length in micron
UCNL	upper apex of corpusculum length in micron
LCNL	lower apex of corpusculum length in micron
COPW	corpusculum width in micron
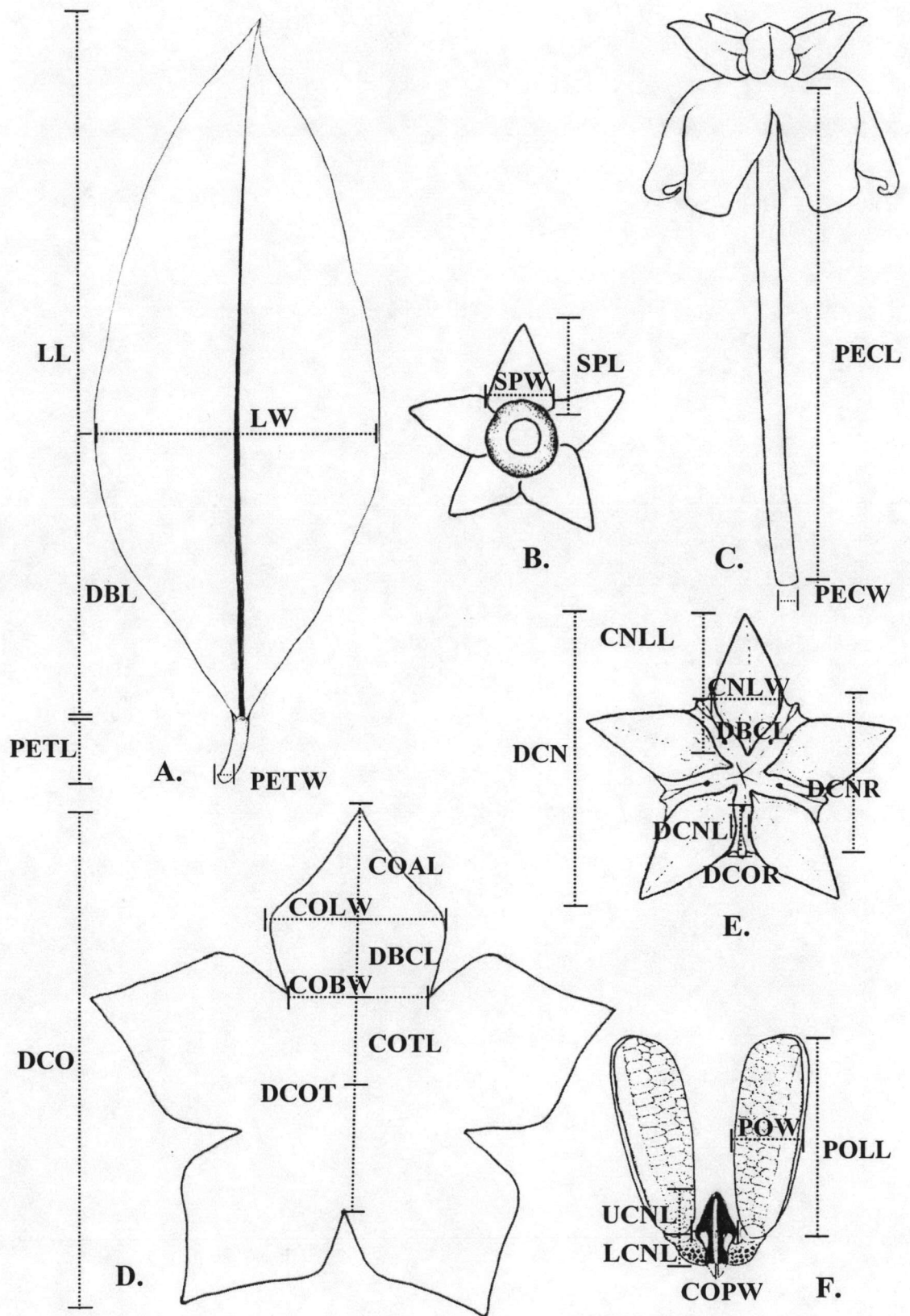
COL=COTL+DBCL+COAL, COLL= DBCL+COAL, COPL= UCPL+LCPL

Figure 4.1 Measurements of leaf and floral parts. A. leaf, B. calyx, C. flower, D. corolla, E. corona, F. pollinariun.

Results and Discussion

1. Cluster analysis

The result of cluster analysis is showed in Figure 4.2. The dendrogram split the 534 specimens into either two or three groups at the 1.80 and 2.80 of average taxonomic distance, respectively. In the two-clusters grouping, specimens classified as group 1 consisted of all members of Form I. Specimens classified as group 2 encompassed two subgroups. Subgroup 1 consisted of all members of Form II. Subgroup 2 included members from Forms III-IX. In the three-clusters grouping, group 1 is the same as in the two-clusters grouping, consisted of all Form I. All members of Form II were placed in group 2. Group 3 consisted of Form III-IX.

Both the two and three-clusters groupings demonstrate a clear separation of Form I. On the other way, the dendrogram shows close relationships between Form II and Forms III-IX. However, Form II is clearly separated from the others on the dendrogram. While Form III, IV, VI, VII, VIII and IX are shown to be very closely related groups and their status are not distinct.

2. Canonical discriminant analysis

2.1 Discriminant analysis of nine categories

Thirty-five characters were used in this analysis. The linear discriminant function classification results showed 84.6% correctly classified. For this reason, the linear discriminant function (Table 4.2) could be used for identification of specimens of the *Hoya parasitica* complex in Thailand. The nature of the entries differences was showed by the pooled within canonical structure (Table 4.3). Canonical variable 1 was 87.4 % correlated with the thirty-five characters and explained 46.6% of the total variance (Table 4.4). It was highly associated with character LW, PETW, PECL, COL, DCO, COAL, and COBW. Canonical variable 2 explained 21.1% of the total variance. This axis is mostly associated with SPL. Canonical variable 3 explained 16.7 % of the total variance. The twelve variables CNLW, POLL, COLL, DCOT, DBCN, COTL, DBCL, RCCD, RCRD, DCOR, UCNL, and CNLS were not used in the analysis according to the result of stepwise discriminant analysis (Table 4.3).
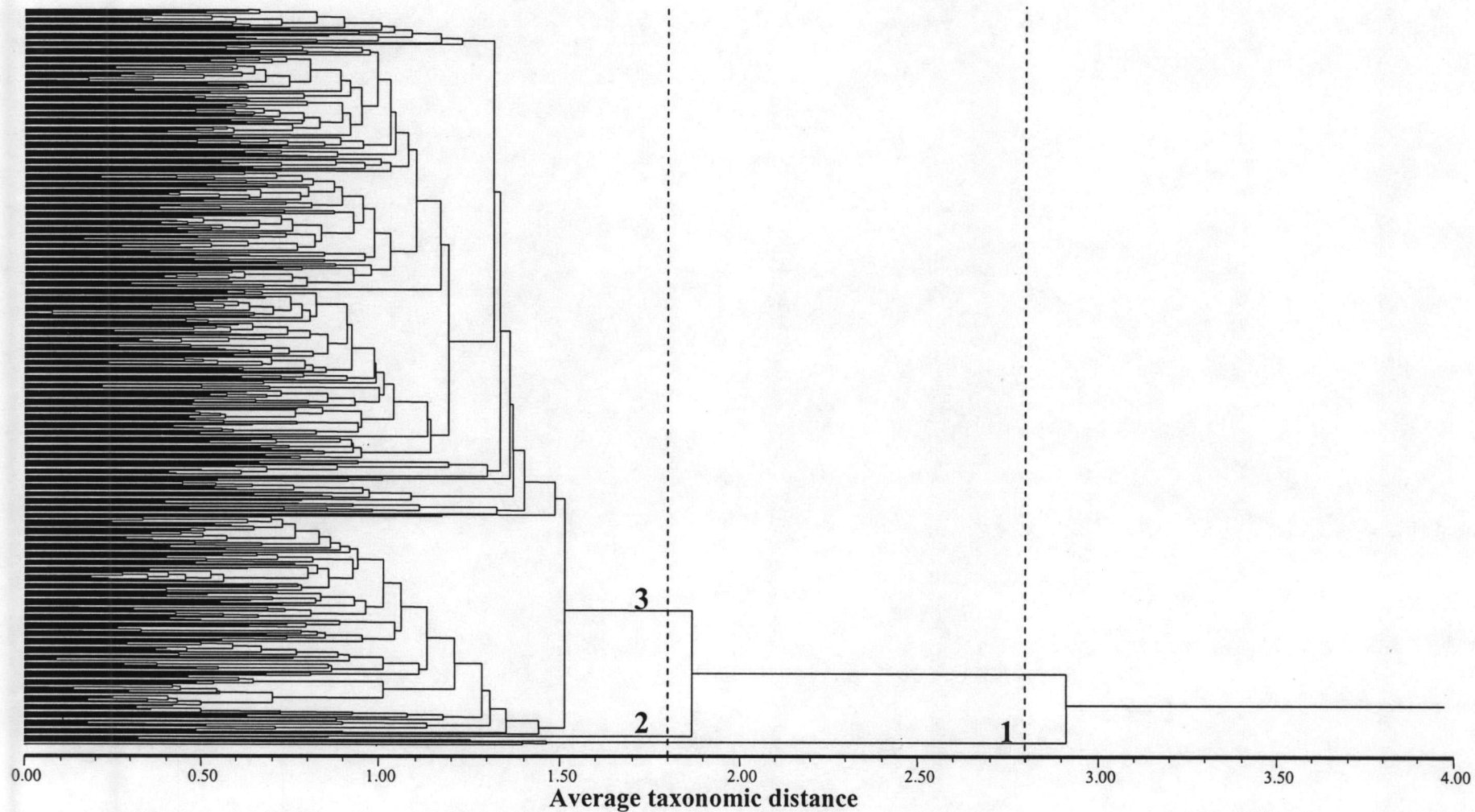
Figure 4.2 UPGMA clustering of 534 OTUs based on 35 quantitative characters of *Hoya parasitica* complex in Thailand (1-Form I, 2-Form II, 3-Form III-IX).

The ordination plot on the two canonical axes shows that the nine categories are not clearly distinct (Figure 4.3). The 9 categories separated the 534 specimens into 2 groups in canonical axis 1. Group 1 consists of Form I and II of *Hoya parasitica* complex. Group 2, the largest group, composed of Form III-IX of *Hoya parasitica* complex. In canonical axis 2, 9 categories can be divided into 2 groups. This axis separated Form I to group 1 and the other Forms to group 2.

2.2 Discriminant analysis of four categories according to 4 forms (VI, VII, VIII, and IX) based on morphological and anatomical data

Thirty-five characters were used in this analysis. The linear discriminant function classification results showed 84.7% correctly classified. For this reason, the linear discriminant function (Table 4.5) could be used for identification of specimens of the *Hoya parasitica* complex in Thailand. The nature of the entries differences was showed by the pooled within canonical structure (Table 4.6). Canonical variable 1 was 85.5% correlated with the thirty-five characters and explained 75.3% of the total variance (Table 4.7). It was highly associated with character LW, CNLW, PECL, POW, POLL, LL, CNLS, and COLW. Canonical variable 2 explained 15.1% of the total variance. This axis was highly associated with COPL, DCNL, COPW, DBL, SPW, and RCCD. Canonical variable 3 explained 9.5% of the total variance. This axis was highly associated with PETW, COL, LCNL, SPL, and RCRD. The sixteen variables LS, DCO, COAL, DBCN, DCNR, DCOT, COLL, COBW, UCNL, DCN, DCOR, DBCL, PECW, COTL, PETL, and CNLL were not used in the analysis according to the result of stepwise discriminant analysis (Table 4.6).

The ordination plot on the three canonical axes showed that the four categories were not apparent (Figure 4.4). The four categories, i.e. forms VI, VII, VIII, and IX could not be separated clearly into distinct group in canonical axis 1, 2 and 3.

2.3 Discriminant analysis of three categories according to the result from cluster analysis

Thirty-five characters were used in this analysis. The linear discriminant function classification results showed 100% correctly classified. For this reason, the linear discriminant function (Table 4.8) could be used for identification of specimens of the *Hoya parasitica* complex in Thailand. The nature of the entries differences was

showed by the pooled within canonical structure (Table 4.9). Canonical variable 1 was 78.4% correlated with the thirty-five characters and explained 67.3% of the total variance (Table 4.10). It was highly associated with character SPL, DCN, DCNR, DCO, PETL, COL, LS, COLW, SPW, COLL, PETW, and PECL. Canonical variable 2 explained 32.3 % of the total variance. This axis was highly associated with COPW, LW, POW, COBW, COAL, COPL, and DCNL. Canonical variable 3 explained 16.7% of the total variance. The sixteen variables DCOR, CNLL, PECW, COTL, DCOT, CNLW, LL, DBCN, DBCL, DBL, RCRD, RCCD, POLL, LCNL, UCNL, and CNLS were not used in the analysis according to the result of stepwise discriminant analysis (Table 4.9).

The ordination plot on the two canonical axes showed that the three categories were distinct (Figure 4.5). The three categories were separated into three groups on canonical axis 1. Group 1 consisted of Form I of *Hoya parasitica* complex, while Group 2 included Form II. Group 3, the largest group, was composed of Form III-IX of *Hoya parasitica* complex. In canonical axis 2, the three categories were divided into two groups. This axis separated Form II to group 1 and the other forms to group 2.
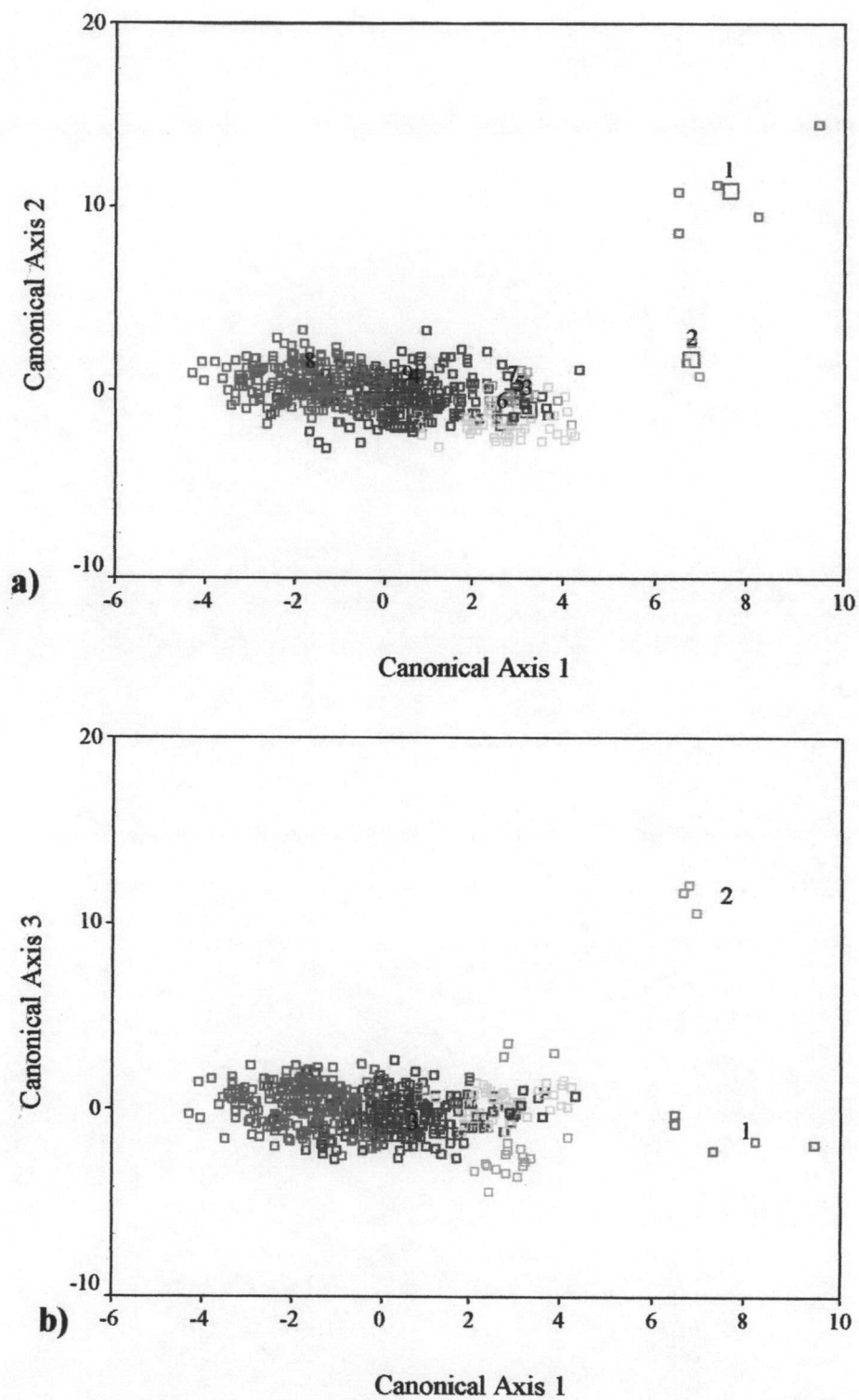
Figure 4.3 The ordination plot of 9 forms (vegetative and reproductive character) using 9 categories as priori groups, a)- the ordination plot on the canonical axes 1 and 2, b)- the ordination plot on the canonical axes 1 and 3.
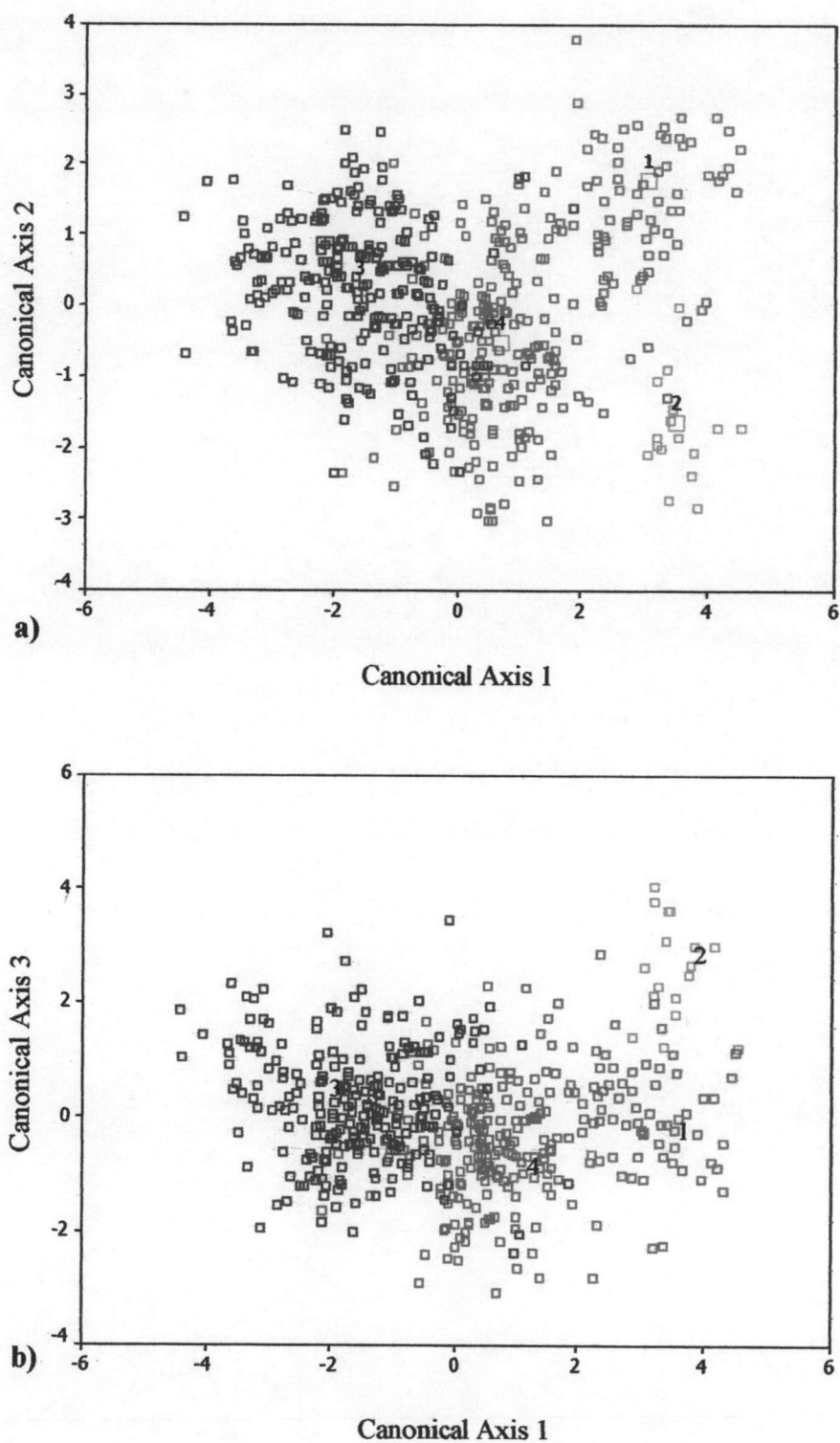
Figure 4.4 The ordination plot of 4 forms (vegetative and reproductive character) using 4 categories as priori groups, a)- the ordination plot on the canonical axes 1 and 2, b)- the ordination plot on the canonical axes 1 and 3. (1=VI, 2=VII, 3=VIII, and 4=IX).
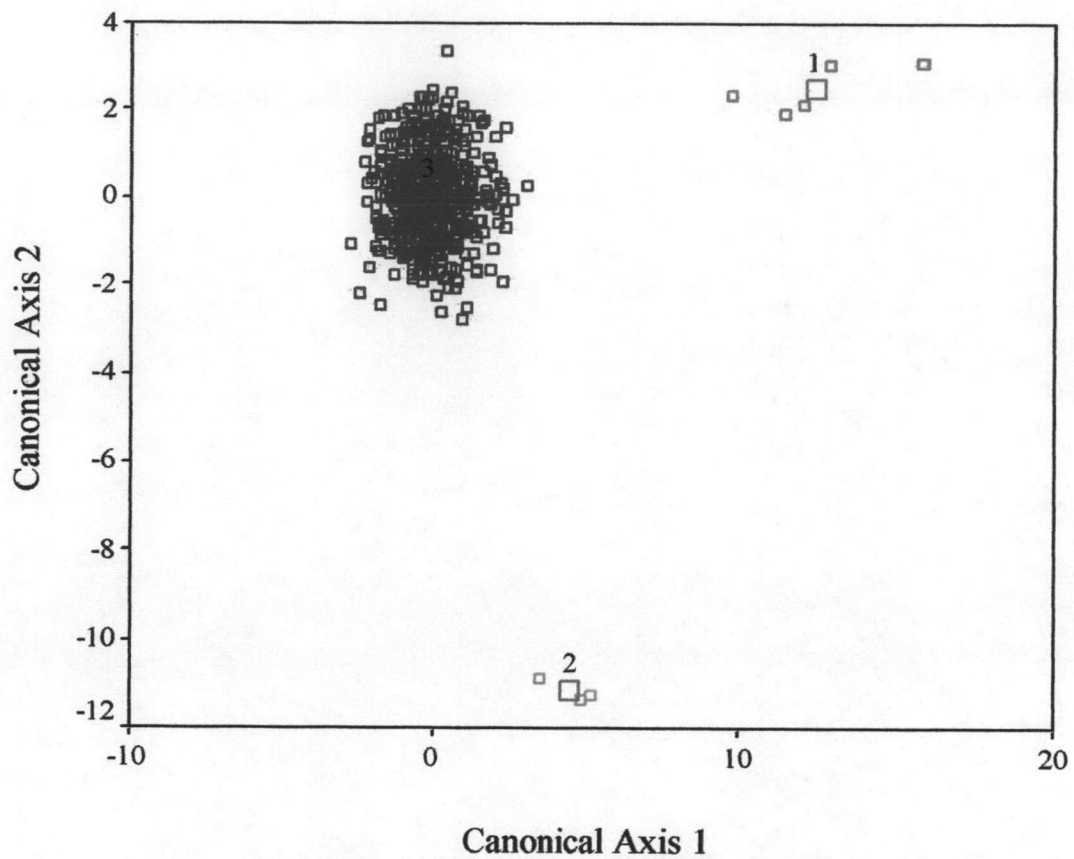
Figure 4.5 The ordination plot of 3 clustering groups (vegetative and reproductive character) using 3 categories as priori groups.

Conclusion

Two techniques of numerical taxonomy were used to investigate the taxonomic status of the nine forms (delimited by morphological and anatomical data) in the *"Hoya parasitica* complex" in Thailand. It should be concluded that the results from numerical taxonomic study as well as the comparison of qualitative morphological and anatomical characters of leaf and flower provide justification for recognition of the segregation of the nine groups of the complex (Form I-Form IX). The results from this study agreed with the treatment of the *"H. parasitica* complex" in Thailand into three groups based on morphological and anatomical studies in Chapter 3. However, it should be noted that there were some continuous variations in quantitative characters of leaf and flower.

The cluster analysis demonstrated a separation of the Form I and Form II. The clear cut separation of these forms from the remainder are probably due to their extreme difference in leaf and flower characters from the others. Furthermore, there are close relationships on the dendrogram among the Form III, IV, V, VI, VII, VIII, and IX and their taxonomic status are not distinct.

In the discriminant analyses, it can be concluded that when nine categories or nine forms was used as priori grouping, the dendrogram (Fig. 4.2) shows the separation of the nine forms based on the morphological and anatomical data in chapter 3. In addition, forms: VI, VII, VIII and IX were not distinct which is in accordance with the morphological and anatomical data in chapter 3. Finally, when the three clustering grouping from a result of the cluster analysis was used as a priori in the discriminant analyses, the three groups were recognized (Fig. 4-Fig. 6), i.e. group 1 (Form I), group 2 (Form II) and group III (Form III-IX). The most 8 important characters for separation of the three groups were sepal length (SPL), diameter of corona (DCN), diameter of coronal receptacle (DCNR), diameter of corolla (DCO), petiole length (PETL), corolla length (COL), corpusculum width (COPW) and leaf width (LW).

In all, the results from multivariate analyses of morphological and anatomical data guide to the conclusion that the nine forms should not be infraspecific variations of a

single species, *Hoya parasitica* (Roxb.) Wall. ex Wight. The "*H. parasitica* complex" in Thailand should be consisted at least of three species.
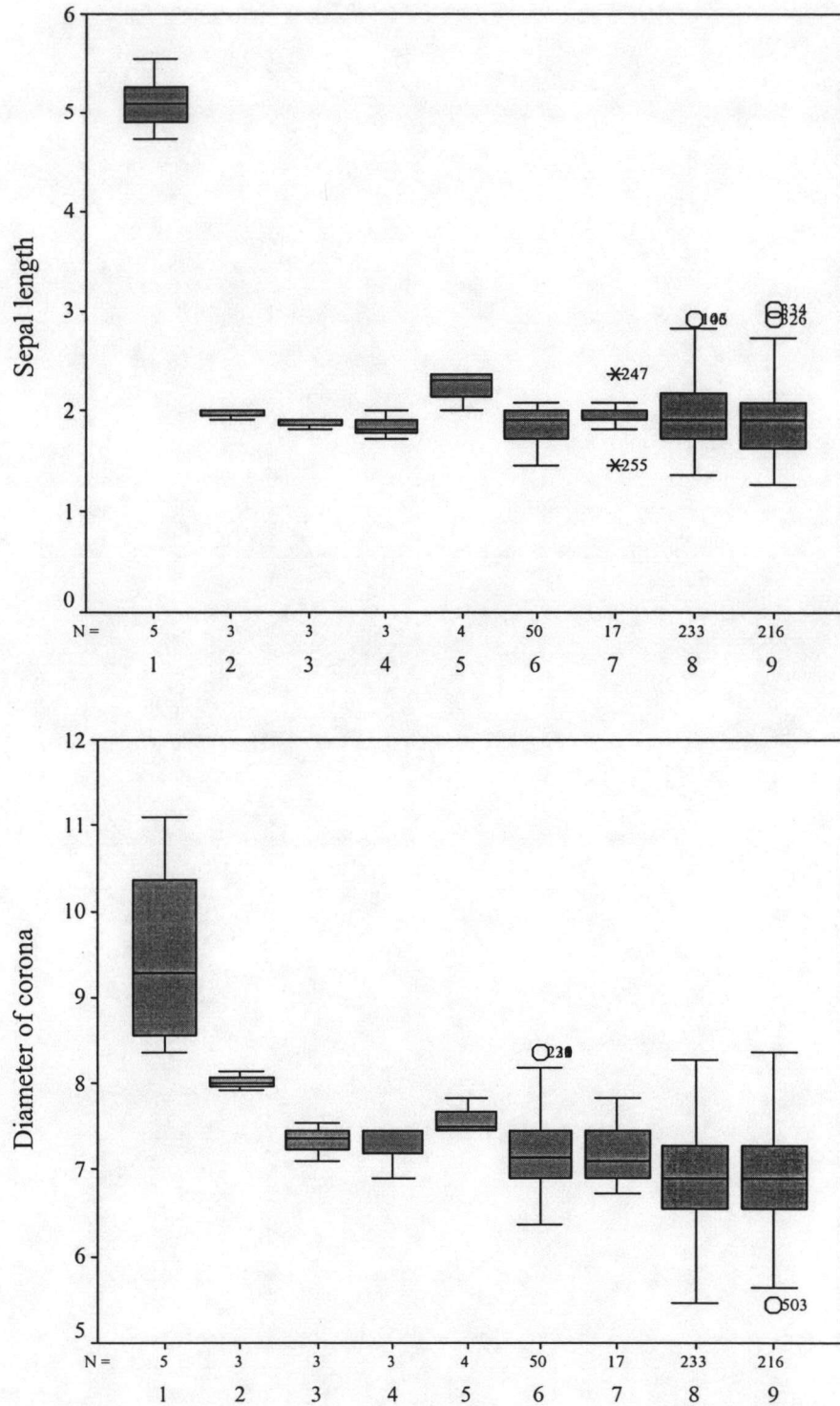
Box plots of eight most important characters are demonstrated in Figure 4.6 and Table 4.11. It is noted that sepal length, corpusculum width and leaf width are useful quantitative characters for discrimination of the three groups. An identification key to the species of the "*H. parasitica* complex" in Thailand based on quantitative characters is provided below.

1. Sepals longer than the corolla tubes, more than 4.5 mm long ..... *Hoya rigida*
    (Form I)

1. Sepals equal in length or shorter than corolla tubes, less than

    2 mm long

    2. Corpusculum less than 70 μm wide and leaf more than 11 cm

    wide .... *Hoya* sp. nov.
    (Form II)

    2. Corpusculum more than 100 μm wide and leaf less than 9 cm

    wide.. *Hoya parasitica*
    (Form III-IX)

Table 4.11 Means and standard deviation of 35 quantitative characters of the 3 taxa of *H. parasitica* complex in Thailand.

Characters	Taxa							
	H. rigida (form I)		*H. sp* (form II)		*H. parasitica* (form III-IX)		Total	
	Mean	±SD	Mean	±SD	Mean	±SD	Mean	±SD
LL	15.850	1.419	16.000	1.127	11.549	2.721	11.614	2.755
LW	7.710	0.829	11.967	1.069	5.103	1.506	5.166	1.603
DBL	5.170	0.952	5.233	0.448	4.987	1.269	4.990	1.263
LS	0.326	0.036	0.328	0.024	0.433	0.051	0.432	0.052
PETL	3.080	0.228	2.133	0.058	1.314	0.512	1.335	0.540
PETW	0.550	0.053	0.530	0.029	0.444	0.080	0.446	0.080
PECL	23.145	2.173	18.317	1.458	18.975	3.350	19.010	3.356
PECW	1.177	0.086	0.939	0.105	0.952	0.089	0.954	0.091
SPL	5.109	0.316	1.970	0.053	1.931	0.303	1.961	0.430
SPW	2.147	0.160	1.655	0.146	1.630	0.205	1.635	0.210
DCO	19.618	1.345	13.713	0.258	13.902	1.330	13.954	1.436
COL	10.226	0.487	7.576	0.140	7.715	0.774	7.738	0.806
COLL	6.344	0.393	4.654	0.386	4.985	0.550	4.996	0.564
DCOT	7.854	1.268	5.669	0.136	6.137	0.642	6.151	0.668
DBCL	2.546	0.392	1.829	0.018	1.841	0.329	1.847	0.335
COAL	3.799	0.318	2.824	0.393	3.145	0.369	3.149	0.375
COTL	3.882	0.127	2.909	0.240	2.730	0.378	2.742	0.392
COLW	5.976	0.437	5.038	0.428	4.891	0.427	4.902	0.439
COBW	5.017	0.624	3.529	0.096	3.990	0.399	3.997	0.413
DCN	9.527	1.174	8.013	0.111	6.982	0.532	7.012	0.596
DCNR	4.983	0.404	3.736	0.013	3.786	0.242	3.797	0.269
DCOR	1.417	0.139	1.060	0.052	1.024	0.090	1.028	0.098
DCNL	0.695	0.041	0.790	0.103	0.577	0.136	0.579	0.136
CNLL	4.548	0.268	3.917	0.074	3.509	0.296	3.521	0.312
DBCN	2.109	0.284	1.459	0.155	1.506	0.207	1.511	0.216
CNLS	0.462	0.043	0.372	0.033	0.429	0.043	0.429	0.043
CNLW	2.382	0.176	1.742	0.013	1.880	0.207	1.884	0.212
RCRD	1.909	0.107	2.153	0.039	1.846	0.107	1.848	0.109
RCCD	1.232	0.193	1.414	0.053	1.146	0.108	1.148	0.111
POLL	560.394	25.024	456.881	1.369	503.223	46.250	503.498	46.415
POW	219.037	21.482	156.748	0.374	183.532	16.615	183.713	17.065
COPL	189.709	7.557	203.395	9.433	168.626	18.429	169.019	18.604
UCNL	95.131	2.528	96.267	3.235	81.153	14.805	81.369	14.800
LCNL	94.577	6.500	107.129	6.199	87.473	10.044	87.650	10.122
COPW	130.794	5.859	69.810	1.402	116.920	10.566	116.785	11.158

Note: The characters in **bold** letter represent the important variables for separation of the three taxa

Figure 4.6 Boxplots of eigth most important characters of *Hoya parasitica* complex. (1-I, 2-II, 3-III, 4-IV, 5-V, 6-VI, 7-VII, 8-VIII, 9-IX)
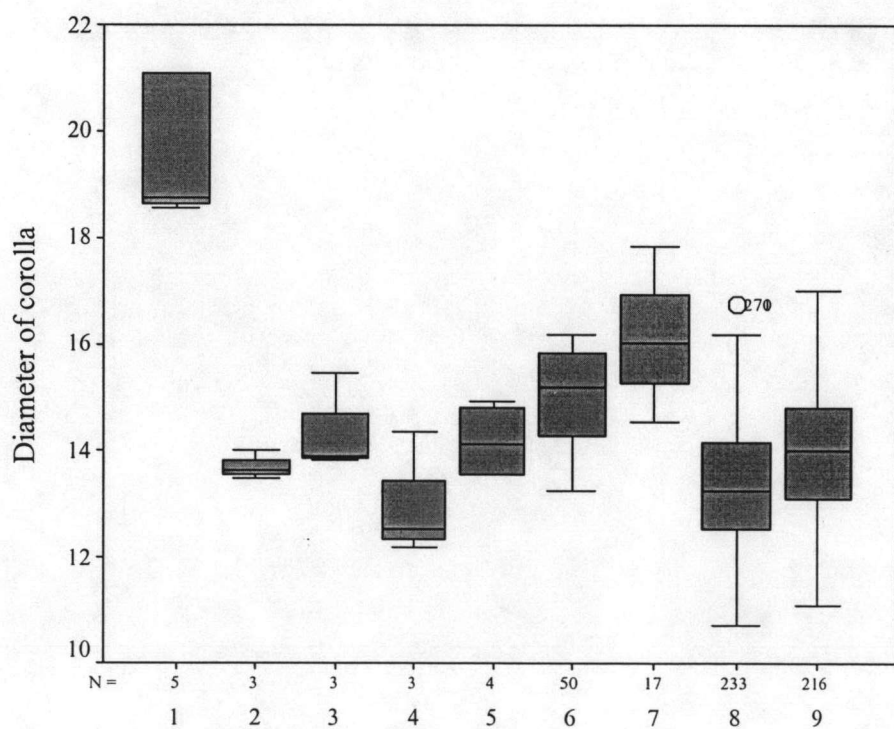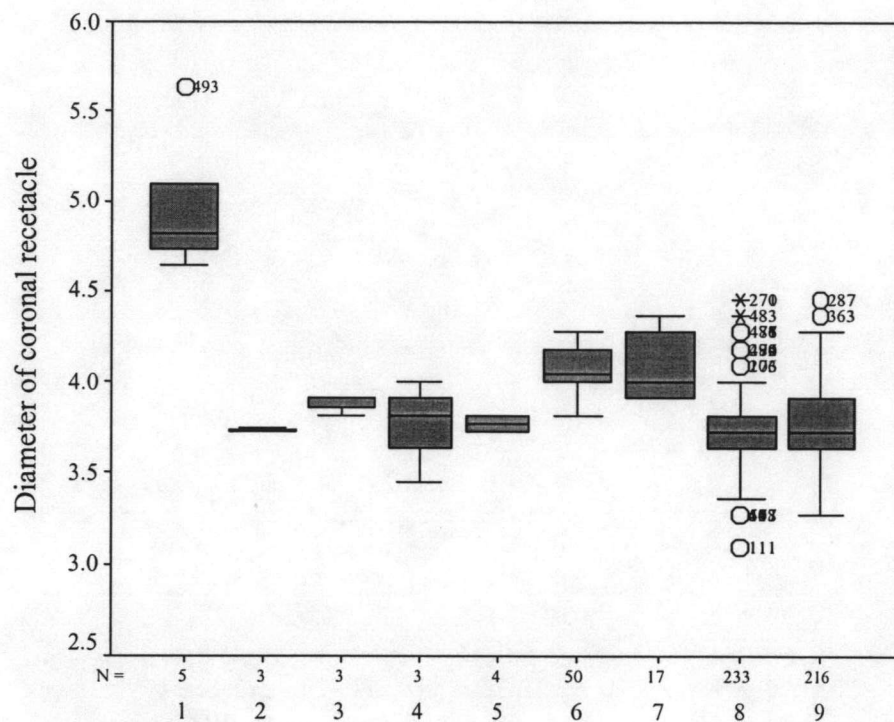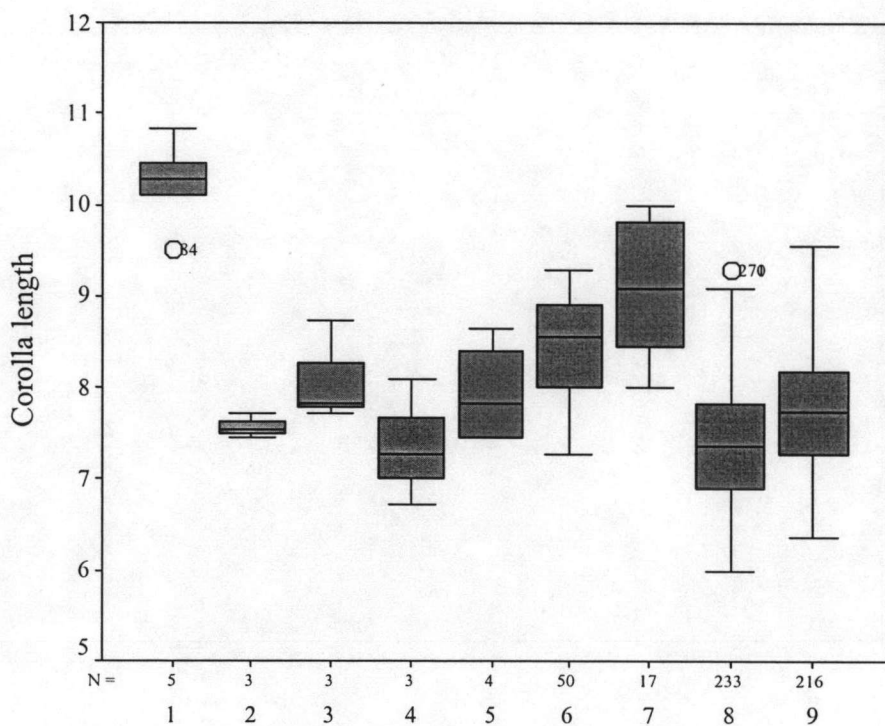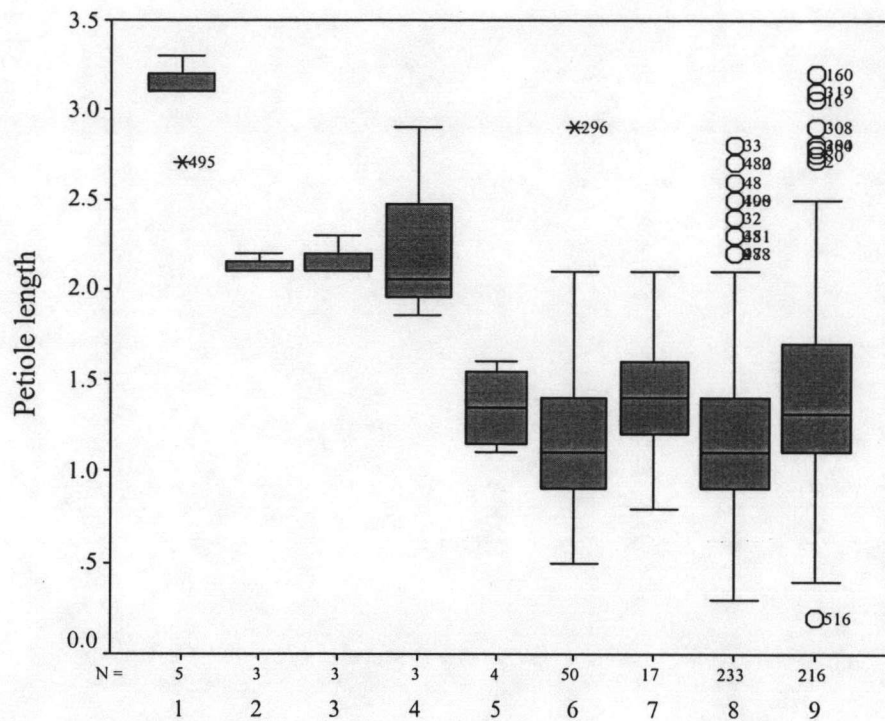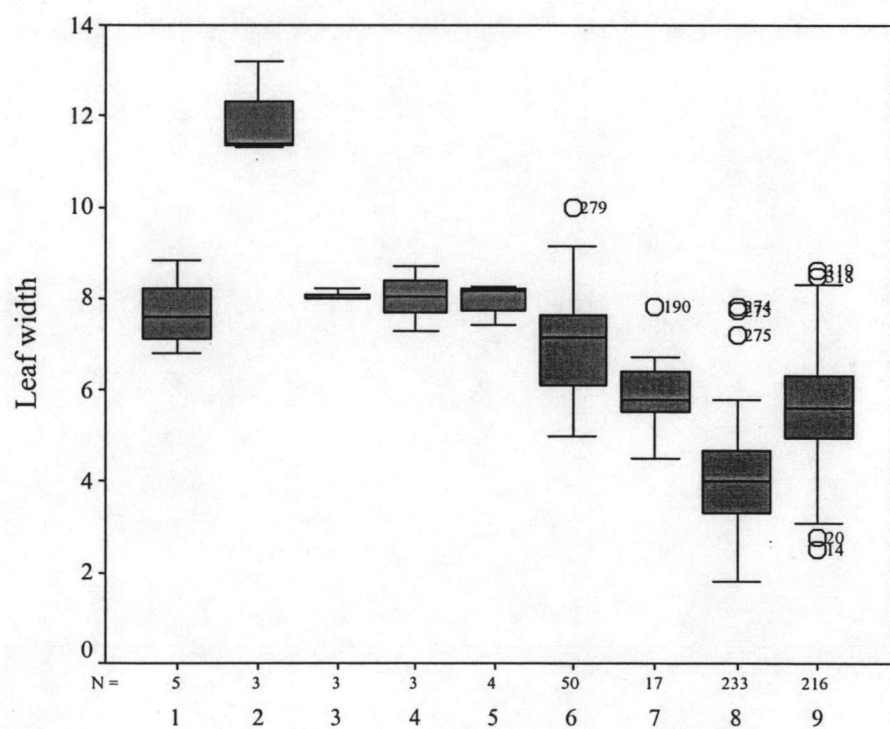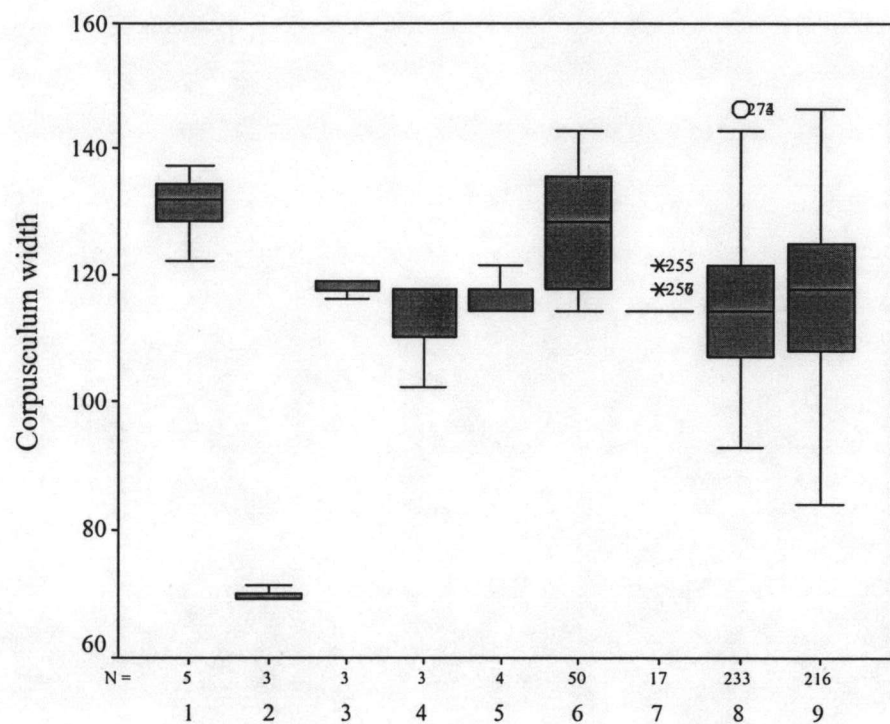
Figure 4.6 (continued)

Figure 4.6 (continued)

Figure 4.6 (continued)