Systematics and biogeography of the rattan Calamus javensis complex (Arecaceae, Calamoideae) in Malesia

Atria, M.

Citation

Atria, M. (2023, February 21). *Systematics and biogeography of the rattan Calamus javensis complex (Arecaceae, Calamoideae) in Malesia*. Retrieved from https://hdl.handle.net/1887/3563955

Chapter 2

Morphometric Analysis of the rattan *Calamus javensis* Complex (Arecaceae, Calamoidae)

Mega Atria, Harald van Mil, William J. Baker, John Dransfield, Peter C. van Welzen
Systematic Botany (2017), 42(3): pp. 494–506.

Abstract -- *Calamus javensis* (Calamoideae) is a slender rattan common in tropical rainforests. The species is very polymorphic and forms a species complex together with some taxonomically still non-recognized forms and the species *C. acuminatus*, *C. amplijugus*, *C. congestiflorus*, *C. corrugatus*, *C. elopurensis*, *C. hypertrichosus*, *C. impar* and *C. tenompokensis*. Within the complex the morphological variation and similarities among the entities are too difficult to solve with traditional morphological observation. Therefore, two multivariate analyses, MDS (Multidimensional scaling) and Hierarchical Cluster Analysis (HCA) with UPGMA (Unweighted Pair Group Method with Arithmetic Mean) were used for two datasets. The inclusion of categorical characters did not make significant differences in the cluster results. Of all 9 taxa included in the complex, only 2 clusters resulted from the analysis next to two forms based on few specimens. Only *C. tenompokensis* was recognized as a separate cluster, while all other taxa were combined into one typical *C. javensis* cluster. More material may result in the recognition of the two forms as distinct species.

2.1. Introduction

Calamus javensis Blume is an understorey rattan belonging to the subtribe Calaminae, tribe Calameae of the palm subfamily Calamoideae (Dransfield et al.

2008). This slender, generally sparsely spiny, climbing palm is widely distributed in South-East Asia, ranging from Southern Thailand and Peninsular Malaysia, to Sumatra, Java, Borneo and Palawan. Within the genus *Calamus*, *C. javensis* is known as an extremely polymorphic species (Dransfield, 1992). Characters are (most typical ones in italics): stem without leaf sheath 2–6 mm in diameter; with sheaths to 10 mm in diameter; internodes up to 30 cm long but usually shorter; *ocrea conspicuously deep crimson when young*; leaves pinnate, ecirrate, up to 40 cm long; *terminal leaflets flabellate, lowermost pair often swept back across the stem*; flagellum up to 75 cm long; *inflorescence long, with red crimson rachilla*; ripe fruit ovoid (Dransfield, 1992; Dransfield and Manokaran 1993).

People make bindings, rope and baskets from the cane. The fruits are edible for small mammals and a few birds. The raw cabbage can be eaten and is sometimes used to cure coughs. In Sarawak the cane is considered excellent because of its length and strength, but it is only sold locally for baskets, cordage and handicrafts. In Palawan, the cane is deemed only second in quality to *Calamus caesius* Blume (a well-known species in the rattan trade) (Mogea, 1994; Shim & Tan 1994). According to the FAO, *C. javensis* is included in the major commercial species of rattan as identified for Asia by Dransfield and Manokaran (1993). Then *C. javensis* may become a new alternative rattan source for industrial needs. However, the fact that *C. javensis* is a polymorphic species complex with variable quality of the cane makes it less valuable unless the complex is taxonomically resolved.

Calamus javensis was first described by Blume (1847). He transferred his own *Calamus equestris* into the synonymy of *C. javensis*. He also noted morphological variation within *C. javensis* and distinguished 3 varieties based mainly on the variation in the leaves and spines on the leaf sheaths (Table 2-1; var. α is considered to be the typical variety and should according to modern nomenclature be called var. *javensis*). Beccari (1884, 1908) recognized Blume's varieties and made modifications, creating his own infraspecific taxonomy. Since then (Table 2-1), many changes in the infraspecific taxa occurred, forms were united, separated, etc. Also, on the species level many exceptional forms were described as new species and synonimized again, the latter especially by Dransfield (1992, 1999), who mainly recognized one species without infraspecific taxa. As a result, *C._javensis* (Dransfield, 1999) is taxonomically the most difficult rattan species complex in the Southeast Asian region. Table 2-1 shows that no subdivision of the complex provided a satisfactory solution, forms were described at various levels or moved to different levels, which also shows that the characters used for recognition were doubtful too. Madulid (1981) treated *C. javensis* as a species complex because of the extreme variation within the species and the relatively subtle differences among morphologically similar groups of specimens. To solve the problem, Dransfield (1999) suggested 2 options, either maintaining several varieties or blend them all into one species. Alternatively, add comprehensive studies in the field next to phenetic and molecular-phylogenetic techniques to examine the circumscription of *C. javensis*.

In species complexes variation is great often showing complex patterns of disjunctions in the ranges of variation (often recognised by separate names), some of which can be linked to ecological factors or geography, others being apparently random. Subsuming all variation under one name results in loss of information. In Java, where the type was originally collected, the species is not very variable (Dransfield, 1999). In West Java, *C. javensis* exhibits a more homogeneous form, though forms with smaller leaflets and shorter rachillae are among the specimens collection from Ujung Kulon (Atria, 2008, unpublished). Variation in Sumatra is similar to that in Peninsular Malaysia (Dransfield, 1999). In the Malay Peninsula, many intermediates are named as varieties. Dransfield (1979) included all names in *C. javensis*. In Sabah (N Borneo) variation is more discontinuous, certain populations are rather distinct, but there appear to be intermediates forms. One particular distinct form is a non-climbing plant on ridge tops above Tenom (Sabah), it has narrow, lanceolate leaflets, like *C. acuminatus* Becc., but the fruit is ovoid just like the other populations of *C. javensis* (Dransfield, 1994). In Sarawak (NW Borneo), on Mount Mulu, a plant was found to resemble *C. tenompokensis* Furtado, but the leaflets are much smaller and have a different texture and the inflorescence is much smaller (Dransfield, 1992).

Problems of intraspecific and infraspecific variation in plants have been commonly addressed by multivariate statistical techniques (Chandler & Crisp, 1998; Loo et al., 2001; Gengler-Nowak, 2002; Henderson & Martins, 2002; Knudsen, 2002; Bacon & Bailey, 2006; Henderson, 2006). In the *C. javensis* complex, Madulid (1981) analysed the leaf variation using Principal Components Analysis (PCA), which resulted in a few discernible groupings. However, Madulid also showed a continuous range in leaf variation. A re-appraisal of *C. javensis* is necessary to clear and put into order its present chaotic taxonomic status. “Morphological characters of plant parts other than the leaves need to be thoroughly examined for more taxonomic evidence that could lead to a better understanding of the true nature of the *C. javensis* complex” (Madulid, 1981). A preliminary study of *C. javensis*’ morphological variation over its distribution was done by using PCA; the results showed that geographic variation in morphology exists, which probably reflects phenotypic responses to environmental gradients (Atria, 2008, unpublished), but this study did not cover all forms in the distribution and all named taxa within the *C. javensis* complex.

In the present study, the selected morphological characters will be analysed using multivariate analyses to see the character distribution in the *C. javensis* complex and to investigate the distinctiveness of some of the character combinations to recognize (some) forms as distinct taxa.

The objectives of this study are: (1) to analyse the character distribution within the species complex using overall morphological characters; (2) to characterize probable characters that significantly influence the species complex. With the right selection of characters and the availability of adequate material throughout the distribution area, it is expected that the distribution of morphological variation in the *C. javensis* complex in West Malesia becomes clearer. The result of this study will be used as a preliminary for the molecular examination of the species complex in order to construct a comprehensive classification of the *C. javensis* complex.

2.2. Material and Methods

Herbarium Specimens

From a total of 226 herbarium specimens, we selected 177 as a representative subset for multivariate statistical analyses; 86 of these were from northern Borneo, where the greatest variation occurs. The specimens of *C. javensis* and related taxa were selected from three herbaria, which store the best and most variable collections: Herbarium Bogoriense (BO), Leiden (L) and Kew (K). Photo material of types of varieties was received from Herbarium Firenze (FI), and particular forms were examined in two Bornean herbaria, Sandakan (SAN) and Kuching (SAR). Type

Table 2-1. The chronological history of names and taxa involved in the Calamus javensis Blume species complex (adapted from Madulid, 1981).

Blume (1847)	**Beccari & Hook.f. (1892)**	**Ridley (1907)**	**Beccari (1908, 1913)**	**Ridley (1925)**	**Furtado (1956)**	**Dransfield (1984, 1992)**
C. javensis	*C. javensis*	*C. penicillatus*	*C. javensis*	*C. javensis*	*C. javensis*	*C. javensis*
var. α	'forma *typica*'					(Dransfield & Manokaran, 1993; Govaerts & Dransfield, 2005)
var. β (*firmus*)						
		var. inermis		var. *inermis*	var. *inermis*	
					var. *laevis*	
	var. *peninsularis* subvar *purpurascens*	var. *purpurascens*	var. *peninsularis* subvar *purpurascens*	var. *purpurascens*	var. *peninsularis*	
					var. *purpurascens*	
	subvar. *tenuissimus*		var. *tenuissimus*		var. *tenuissimus*	
	subvar. *pinangianus*	var. *pinangianus*	var. peninsularis	var. *pinangianus*	var. *pinangianus*	
			var. pinangianus			
	subvar. polyphyllus		var. *polyphyllus*		var. *polyphyllus*	
			var. *intermedius*			
var. γ (*tetrastichus*)			var. *tetrastichus*			
			subvar. *mollispinus*			
			var. *sublaevis*			
			var. *acicularis*			
			var. *exilis*			
					C. kemamanensis	
			C. corrugatus			*C. corrugatus*
			C. filiformis			

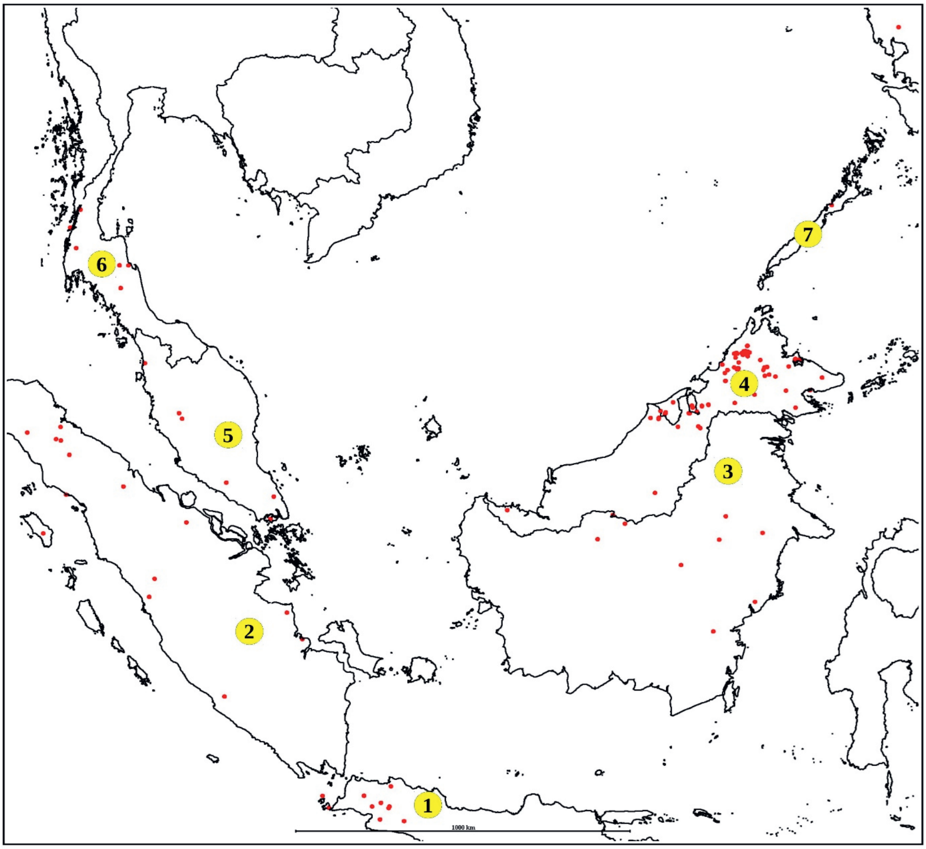

Figure 2-1. Samples locality and collection sites. 1=West Java, 2=Sumatra, 3=Kalimantan, 4=Northern Borneo, 5=Malay Peninsula, 6=Thailand, 7=the Philippines.

specimens were included except for Beccari's varieties as the photo material was unsuitable for measurements. Representative herbarium specimens were also collected during fieldwork in West Java and several parts of Sabah and Brunei. Collection and preparation followed the standard procedure by Dransfield (1986). Fertile and sterile specimens were both collected to represent the distributions of all forms. Sample localities and collection sites are shown in Figure 2-1.

Representatives of as many as 9 taxa that are generally accepted (Blume, 1847; Beccari, 1908; Furtado, 1956; Dransfield, 1984; Dransfield, 1992; Govaerts & Dransfield, 2005) were included in this study: *Calamus acuminatus* Becc., *C. amplijugus* J.Dransf., *C. congestiflorus* J.Dransf., *C. corrugatus* Becc., *C. elopurensis* J.Dransf., C. *hypertrichorus* Becc., *C. impar* Becc., *C. javensis* Blume and *C. tenompokensis* Furtado (see also Table 2-1).

Morphometry was conducted according to Rohlf (1990), Chandler & Crisp (1997), Loo et al. (2001) and Henderson (2006). Specimens were not assigned to groups a priori and each herbarium collection was considered as an individual Operational Taxonomic Unit (OTU) in tabulating the data matrix (Loo et al., 2001). Multiple herbarium sheets from a single collection were grouped into a single OTU. The characters were chosen based on the differences among taxa. Thirty-two quantitative and qualitative characters, like sizes, shapes and textures of vegetative and floral parts, were measured using a ruler, measuring tape and calliper (Table 2-2) (Gengler-Nowak, 2002; Henderson & Martins, 2002; Sreekumar et al., 2006). Missing values were left blank in the matrix. The measurements derived from observations on herbarium specimens were assigned to a matrix. The 32 characters by 177 Operational Taxonomic Units (OTUs) data matrix was then analysed using multivariate analyses.

Multivariate Analyses

All statistical analyses were performed in the statistical computing environment R (R Core Team, 2016). Missing values were imputed using Multiple Imputation by Chained Equation (MICE) algorithms provided by the MICE package (van Buuren & Groothuis-Oudshoorn, 2011). The default in the MICE package for R is k=5, "k" is how many cases should be in each match set. That is, each case with missing data on x is matched to 5 cases (with data present) that have the closest predicted values (Allison, 2015). To investigate the effect of categorical variables on the classification problem, we made a second dataset without categorical data (qualitative characters).

We used two types of statistical learning algorithms, unsupervised and supervised to allow for two different perspectives on the problem. Multidimensional Scaling (MDS) was applied to overview the distribution of the variation within the *C. javensis* complex. MDS has become a popular technique for both multivariate and exploratory data analysis (Krauss, 1996; Chandler & Crisps, 1998; Streiber et al., 1999; Wicklermaier, 2003). Additionally, we used Hierarchical Cluster Analysis (HCA) using an unweighted cluster method (UPGMA = Unweighted

Pair Group Method with Arithmetic Mean)).

The results presented in this paper were obtained using an Euclidean similarity measure for the numerical dataset and the Gower measure for the dataset with categorical data (Gower, 1971; Struyf et al., 1997). In the Hierarchical Cluster Analysis (HCA) this was combined with the average agglomeration measure for UPGMA. Other distance and agglomeration measures were applied to check their effect on the cluster analysis. To assess the uncertainty of the HCA we used a bootstrap method as implemented in the pvclust R package (Suzuki & Shimodaira, 2006).

To get another perspective on the data supervised methods of Classification Trees (CT; Breiman et al., 1984; De'ath, 2002; Hothorn et al., 2006) and Random Forest learning algorithm (RF; Breiman, 2001; Liaw & Wiener, 2002; Cutler et al., 2007) were used to see how well these algorithms predicted the classifications done by experts. Here we draw on a much used metaphor used in the context of tree methods like CT and RF; the resulting decision trees and community of decision trees are seen as a virtual expert (CT) or community of experts (RF). Confusion matrices and cross validation error were used to evaluate quality of the classification procedure where confusion is equated as the lack of consensus between experts.

Table 2-2. List of morphological characters, states and units of measurement used in the study.

No	Character (names and abbreviation)	States / Measurem/ents / Type of data	Vegetative/ Generative	remarks
1	Stem diameter with sheath (LS diam)	Cm / numeric / Continuous	Vegetative	
2	Stem diameter without sheath	Cm / numeric / Continuous	Vegetative	
3	Leaf sheath spine density (spinedens)	(0) no spine; (1) sparse; (2) moderate; (3) dense; (4) very dense / discrete / semi-continuous	Vegetative	used in the analysis
4	Spine shape (Spinshap)	(0) tape; (1) swollen-based; (2) broad swollen-based short; (3) swollen-based pointing; (4) clawed; (5) no spines / discrete	Vegetative	used in the analysis
5	Ocrea (Ocrea)	(0) none, (1) obvious (<2 cm), (2) conspicuous (2 cm) / categorical	Vegetative	used in the analysis
6	Knee (Knee)	(0) none, (1) obvious, (2) conspicuous	Vegetative	used in the analysis
7	Leaf length	Cm / numeric / Continuous	Vegetative	used in the analysis

No	Character (names and abbreviation)	States / Measurem/ents / Type of data	Vegetative/ Generative	remarks
8	Petiole length (Petiolength)	Cm / numeric / Continuous	Vegetative	used in the analysis
9	Leaflet pair number including flabellate pair (Leafpair)	Count / numeric / Discrete	Vegetative	used in the analysis
10	Leaflets uppermost length (Upperlength)	Cm / numeric / Continuous	Vegetative	used in the analysis
11	Leaflets uppermost width (Upperwidth)	Cm / numeric / Continuous	Vegetative	used in the analysis
12	Leaflets penultimate length (Penlength)	Cm / numeric / Continuous	Vegetative	used in the analysis
13	Leaflets penultimate width (Penwidth)	Cm /numeric / Continuous	Vegetative	used in the analysis
14	Leaflets median length (Medlength)	Cm / numeric / Continuous	Vegetative	used in the analysis
15	Leaflets median width (Medwidth)	Cm / numeric / Continuous	Vegetative	used in the analysis
16	Leaflets lowermost length (Lowlength)	Cm / numeric / Continuous	Vegetative	used in the analysis
17	Leaflets lowermost width (Lowwidth)	Cm / numeric / Continuous	Vegetative	used in the analysis
18	Prophyll armature (Propharm)	(0) smooth; (1) prickly; (2) spiny; (3) with indument / Categorical	Generative	used in the analysis
19	Peduncular bracts shape (Pedbrashape)	(0) loose; (1) tightly sheathing; (2) tightly sheathing with limb; (3) tightly sheathing with very long limb/ Categorical	Generative	used in the analysis
20	Rachis bracts armature (Rachbracharm)	(0) smooth; (1) prickly; (2) spiny; (3) with indument / Categorical	Generative	used in the analysis
21	Number of partial inflorescences (cluster)	Count / numeric / Discrete	Generative	
22	Cluster length (the length of the partial inflorescence)	Cm / numeric / Continuous	Generative	
23	Number of rachillae	Count / numeric / Discrete	Generative	used in the analysis
24	Rachilla length (Rachilleng)	Cm/ numeric/ Continuous	Generative	
25	Inflorescence length	(0) short (<1 m); (1) long (>1 m) / Categorical	Generative	
26	Prophyll length	(0) short (<10 cm); (1) long (>10 cm) / Categorical	Generative	
27	Leaflet blade texture	(0) glabrous; (1) hairy / Categorical	Vegetative	
28	Stem habit (Stem)	(0) climbing; (1) short erect / Categorical	Vegetative	
29	Pistillate flower bract texture	(0) smooth; (1) striate; (2) prickly / Categorical	Generative	
30	Pistillate flower bract shape	(1) loose; (2) tight / Categorical	Generative	
31	Staminate flower bract texture	(0) smooth; (1) striate; (2) prickly / Categorical	Generative	
32	Staminate flower bract shape	(1) loose; (2) tight / Categorical	Generative	used in the analysis

2.3. Results

The results of different types of analysis are shown, starting with the results of (alpha-) taxonomic comparisons of the specimens, field observations and finally continued by multivariate results.

Morphological Observations

The observations and examination of the *C. javensis* complex resulted in several tentative form groups, form 1–4 next to various forms recognized as species (Table 2-3), whereby several groups are represented by a very limited number of specimens. Two of Beccari's species, *C. corrugatus* and *C. hypertrichosus*, were grouped with the typical *C. javensis*. The two species have a close resemblance with typical *C. javensis*. Of *C. hypertrichosus* only a single specimen was observed, which had one very distinctive character, white hairy leaflets and leaf sheaths. In *C. corrugatus* intermediate specimens showed that the ring-like corrugated/ wrinkled leaf sheath is variable and not distinctive, it ranges between a half circle to a full ring-like, corrugated sheath.

Detailed observations from field and herbarium examinations showed that members of the *C. javensis* complex exhibit a wide range of morphological variation. Intermediates were identified, especially among the very diverse specimens from Northern Borneo. Specimens were collected in an altitude range of 20–2,000 m above sea level. Fieldwork showed the overlapping occurrence of two taxa, *C. javensis* var. *polyphyllus* and *C. acuminatus*, and other growth forms together in the same place with typical *C. javensis*.

Variable characters observed in the *C. javensis* complex are:

- ◊ Stem and leaf sheath. All members of the *C. javensis* complex are clustering rattans, which climb with the flagella except for *Form 3* and *C. tenompokensis*, which are short-stemmed, erect rattans (acaulescent). The stem diameter, including the sheaths, varies from 2–30(–35) mm in diam. with 5–10 mm as common diameter found. The smallest stem diameter belongs to *C. javensis* and sturdier stems belong to *C. tenompokensis.*
- ◊ The spines on the leaf sheaths differ in size, shape and density. The shape varied from small, clawed to triangular flat, short to long pointed, with mostly a swollen base. Almost all specimens are spiny, with scattered to many, slender

Table 2-3. Groups distinguished in the Calamus javensis complex based on morphological observations.

No	Taxa	No of Specimens	Remarks	Distribution
1	*Calamus amplijugus*	10	Broadly elliptic leaflets, rachilla very long up to 25 cm	Brunei, Malaysia: Sabah
2	*Calamus congestiflorus*	6	Short and congested rachilla	Indonesia: East Kalimantan, West Java; Malaysia: Sabah
3	*Calamus elopurensis*	4	Leaflets in 2 or 3 pair rosette to the terminal, Ocrea conspicuous, tubular triangle limb, reddish, up to 6 cm, petiole long, c. 9-15 cm, floccose, sometimes abundant.	Malaysia: Sabah
4	*Calamus impar*	9	Leaflets 2 pairs, deeply bilobed flabellate pair, ocrea short-truncated	Indonesia: East Kalimantan, West Java; Malaysia: Sabah; Thailand: Phang Nga
5	*Calamus javensis* (typical)	89	ocrea conspicuously deep crimson when young; leaflets 4-7 pairs, terminal leaflets flabellate, lowermost pair often swept back across the stem; inflorescence long, with red crimson rachilla; ripe fruit ovoid	Indonesia: West Java, Sumatra, Kalimantan; Southern Thailand, Malaysia: Malay Peninsula, Sabah, Sarawak.
	Calamus corrugatus	6	Leaf sheath glabrous with corrugated horizontal lines.	Indonesia: Central Kalimantan Sarawak
	Calamus hypertrichosus	1	Leaf sheath glabrous, leaflets with abundant hairs on both surfaces.	Sarawak
6	*Calamus javensis* var. *polyphyllus* + *Calamus acuminatus*	15 + 13	Leaflets 8-12 pairs, apetiolate to short petiolate	Brunei, Indonesia: Sumatra, Malaysia: Peninsular Malaysia, Sabah, Sarawak. Sabah: Tenom, Keningau
7	*Calamus tenompokensis*	10	Stem short, moderate stem with robust spines, flagellum absent, inflorescence short,	Malaysia: Sabah, Sarawak
8	*Form 1*	2	Lack of flagellum and cirrus, short staminate inflorescence, warty rachilla bracts	Sarawak (Mt. Mulu), 1200-1500 m alt.
9	*Form 2*	1	Short setae along the veins and one long black setae abaxial main veins	The Philippines, Pampanga, Mount Arayat
10	*Form 3*	2	Leaflets only 2 or 3 pairs, tomentose	Malaysia: Sabah, Sarawak
11	*Form 4*	9	Ocrea rough and hirsute, pistillate inflorescence short	Indonesia: Sumatra, Malaysia: Sabah

to robust spines, but *C. acuminatus* specimens have an almost unarmed leaf sheath with only very few, scattered spines.

◊ Flagella are always present except in *Form 3* and *C. tenompokensis*, where, if present the flagellum is short, less than 1 m long. The flagella are armed with clawed spines, 2–5 arranged in a group. The tip of the spines is almost always black (black-tipped spine).

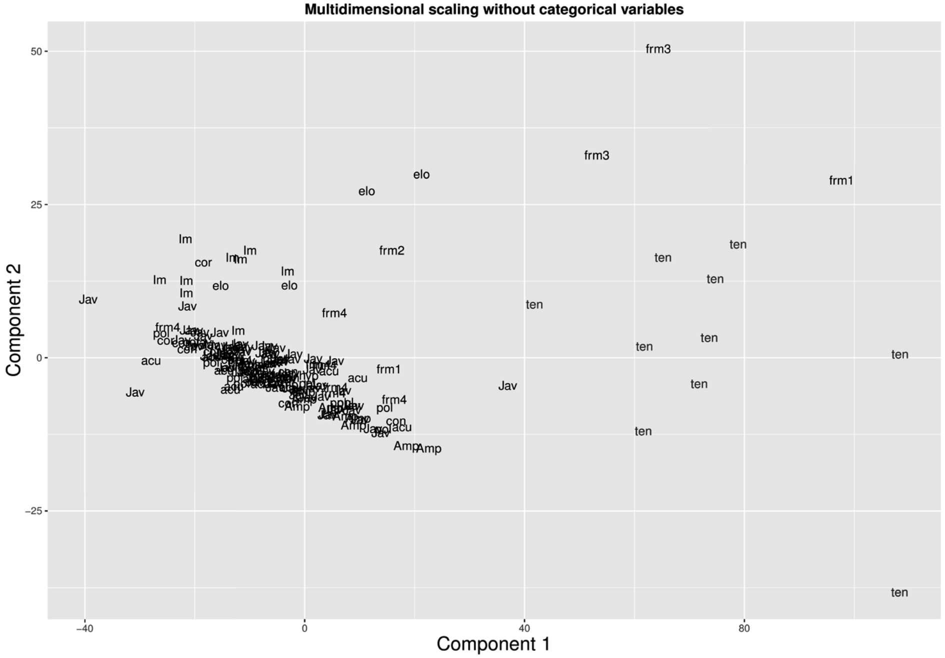

Figure 2-2. Clustering result of unsupervised method (MDS without categorical variables). acu = C. acuminatus, amp = C. amplijugus, con = C. congestiflorus, cor = C. corrugatus, elo = C. elopurensis, frm1 = Form 1, frm2 = Form 2, frm3 = Form 3, frm4 = Form 4, hyp = C. hypertrichosus, im = C. impar, jav = C. javensis, pol = C. javensis var. polyphyllus, ten = C. tenompokensis.

◊ The ocrea is tubular, dark red (maroon) and sometimes torn apart in older plant parts. Ocrea length varies from 0.5 to more than 2 cm. The ocrea is usually glabrous, but in *Form 4* it is hirsute with stiff hairs. In *C. elopurensis* it is prolonged on the far side to the petiole whereas in other forms the ocrea is prolonged next to the petiole.

◊ The leaves and leaflets show most variation. The **leaf** is pinnately compound and ecirrate with a length of 20 cm to more than 140 cm (*C. tenompokensis*). Mature plants usually have a very short petiole, but *C. javensis* var. *polyphyllus* and *C. acuminatus* have sessile leaves. Juvenile or young plants usually have a longer petiole than the older plants. In *C. elopurensis*, the petiole is quite distinctive from c. 9-15 cm long. *Calamus tenompokensis* has the longest petioles, 22–27(–60) cm long.

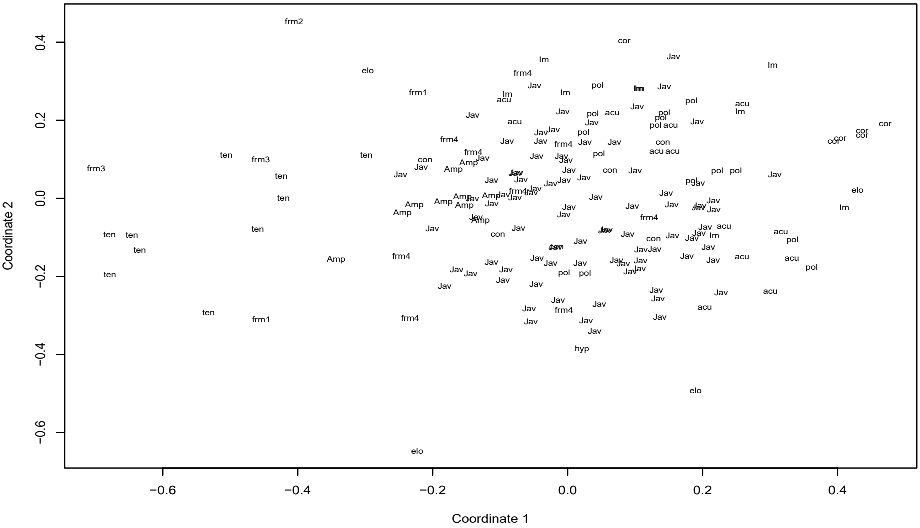

Figure 2-3. Clustering result of unsupervised method (MDS with categorical variables). acu = C. acuminatus, amp = C. amplijugus, con = C. congestiflorus, cor = C. corrugatus, elo = C. elopurensis, frm1 = Form 1, frm2 = Form 2, frm3 = Form 3, frm4 = Form 4, hyp = C. hypertrichosus, im = C. impar, jav = C. javensis, pol = C. javensis var. polyphyllus, ten = C. tenompokensis.

◊ Four distinct leaflet shapes are observed, elliptic, oblong, oblanceolate and lanceolate. The apices are usually acute, but can also be acuminate. The leaflets have black bristle apically and along the margins.

◊ Leaflets are in 2 (*C. impar*)–8(–12: *C. tenompokensis*, *C. acuminatus*, *C. javensis* var. *polyphyllus* and *Form 1*) pairs; most *C. javensis* specimens have 5–7 pairs.

◊ The uppermost leaflets are partially bifid and joined for 1/3–2/3 of their length. The basal or the lowest pair usually are the smallest leaflets, which vary from swept back across the stem (enclosing the stem) to spreading.

◊ The staminate and pistillate inflorescences are quite similar. The inflorescences are long and mainly pendulous, whereby the first part is erect followed by a slightly curved part and the longest, terminal part is hanging. The inflorescence

Table 2-4. Tree model confusion table. horizontal = actual taxon; vertical = predicted taxon. Abbreviation of taxa: acu = C. acuminatus, amp = C. amplijugus, con = C. congestiflorus, cor = C. corrugatus, elo = C. elopurensis, frm1 = Form 1, frm2 = Form 2, frm3 = Form 3, frm4 = Form 4, hyp = C. hypertrichosus, im = C.impar, jav = C. javensis, pol = C. javensis var. polyphyllus, ten = C. tenompokensis. Number indicates the number of OTUs (or collection specimens).

PREDICTED

ACTUAL	acu	amp	con	cor	elo	frm1	frm2	frm3	frm4	hyp	im	jav	pol	ten
Acu	4	0	0	0	0	0	0	0	0	0	0	1	8	0
amp	0	5	0	0	0	0	0	0	0	0	0	5	0	0
Con	0	0	0	0	0	0	0	0	0	0	0	6	0	0
Cor	1	0	0	0	0	0	0	0	0	0	1	4	0	0
Elo	0	0	0	0	0	0	0	0	0	0	2	0	0	2
frm1	0	0	0	0	0	0	0	0	0	0	0	0	1	1
frm2	0	0	0	0	0	0	0	0	0	0	0	0	0	1
frm3	0	0	0	0	0	0	0	0	0	0	0	0	0	2
frm4	0	0	0	0	0	0	0	0	6	0	0	3	0	0
Hyp	0	0	0	0	0	0	0	0	0	0	0	1	0	0
Im	0	0	0	0	0	0	0	0	0	0	9	0	0	0
Jav	1	4	0	0	0	0	0	0	1	0	2	81	0	0
Pol	3	1	0	0	0	0	0	0	0	0	0	2	9	0
Ten	0	0	0	0	0	0	0	0	0	0	0	1	0	9

length varies from 50 cm to more than 1m. Short and erect inflorescences are present in *C. tenompokensis* (less than 1 m long), *Form 1* and *Form 3*. Typical *C. javensis* has (2–5)–8 partial inflorescences; inflorescences bearing > 8 partial inflorescences can be found in *C. javensis* var. *polyphyllus*, *C. elopurensis*, and *C. amplijugus*.

◊ The prophyll is 7–30 cm long, closely sheathing, with spines, but some forms have an unarmed prophyll, or the prophyll is glabrous except for a spiny tip. Rachis bracts are mostly spiny or prickly, some are glabrous, others have hairs.

◊ The pistillate rachillae are borne on the 2nd order branches and staminate rachillae are on the 3rd order branches. Staminate and pistillate rachillae

Table 2-5. Confusion matrix of Random Forest analysis. acu = C. acuminatus, amp = C. amplijugus, con = C. congestiflorus, cor = C. corrugatus, elo = C. elopurensis, frm1 = Form 1, frm2 = Form 2, frm3 = Form 3, frm4 = Form 4, hyp = C. hypertrichosus, im = C. impar, jav = C. javensis, pol = C. javensis var. polyphyllus, ten = C. tenompokensis. Number indicates the number of OTUs (or collection specimens).

	PREDICTED													
ACTUAL	acu	amp	con	cor	elo	frm1	frm2	frm3	frm4	hyp	im	jav	pol	ten
Acu	4	0	0	0	0	0	0	0	0	0	0	6	3	0
amp	0	4	0	0	0	0	0	0	0	0	0	6	0	0
con	0	0	0	0	0	0	0	0	0	0	0	6	0	0
cor	0	0	0	4	0	0	0	0	0	0	0	1	1	0
Elo	0	0	0	0	3	0	0	0	0	0	0	1	0	0
frm1	0	0	0	0	0	0	0	0	0	0	0	1	0	1
frm2	1	0	0	0	0	0	0	0	0	0	0	0	0	0
frm3	0	0	0	0	0	0	0	2	0	0	0	0	0	0
frm4	0	0	0	0	0	0	0	0	5	0	0	4	0	0
hyp	0	0	0	0	0	0	0	0	0	0	0	1	0	0
Im	0	0	0	0	0	0	0	0	0	0	7	2	0	0
Jav	1	3	0	0	0	0	0	0	1	0	0	80	4	0
Pol	4	0	0	0	0	0	0	0	0	0	0	9	2	0
Ten	0	0	0	0	0	1	0	0	0	0	0	0	0	9

have the same color, red when mature and green in younger stages. Staminate rachillae have a more simple structure, one bracteole subtending one staminate flower along each side of the rachilla; while pistillate rachillae are more complex, with an arrangement in ’dyads’, with one rachilla bract subtending 2 flowers, each with its own bracteole, one fertile pistillate flower and one sterile staminate (which is usually already caducous when collected and leaving a scar next to the pistillate flower). The rachilla length in *C. javensis* varies from 4–8 cm in staminate inflorescences and is up to 12 cm long in pistillate inflorescences. The average length of the staminate rachilla is 5.5 cm and that of the pistillate rachillae 6.5 cm. In other forms the rachillae are 2–12 cm long; short and congested in *C. congestiflorus* with the rachillae only 2–4

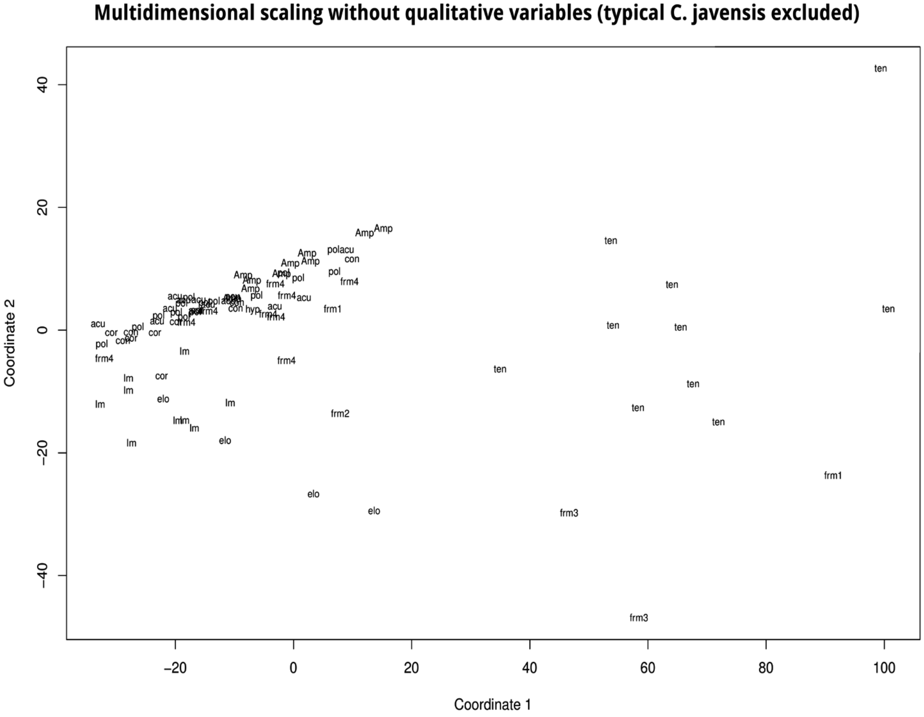

Figure 2-4. MDS clustering analysis without categorical variables and typical C. javensis excluded from the data matrix. Abbreviation of taxa: acu = C. acuminatus, amp = C. amplijugus, con = C. congestiflorus, cor = C. corrugatus, elo = C. elopurensis, frm1 = Form 1, frm2 = Form 2, frm3 = Form 3, frm4 = Form 4, hyp = C. hypertrichosus, im = C. impar, pol = C. javensis var. polyphyllus, ten = C. tenompokensis.

cm long, but up to 25 cm long in *C. amplijugus*.

Multivariate analysis

The dataset consisted of 177 OTUs with 32 characters. The dataset suffered from missing data, mostly in the inflorescence characters. Missing data were due to incomplete parts of the specimens, as can be expected, fruiting specimens have no floral characters, pistillate specimens no fruits or staminate floral characters, etc. To minimize the missing values, there was a pre-selection of variables based on the availability of observed characters. Twenty characters were selected, and the

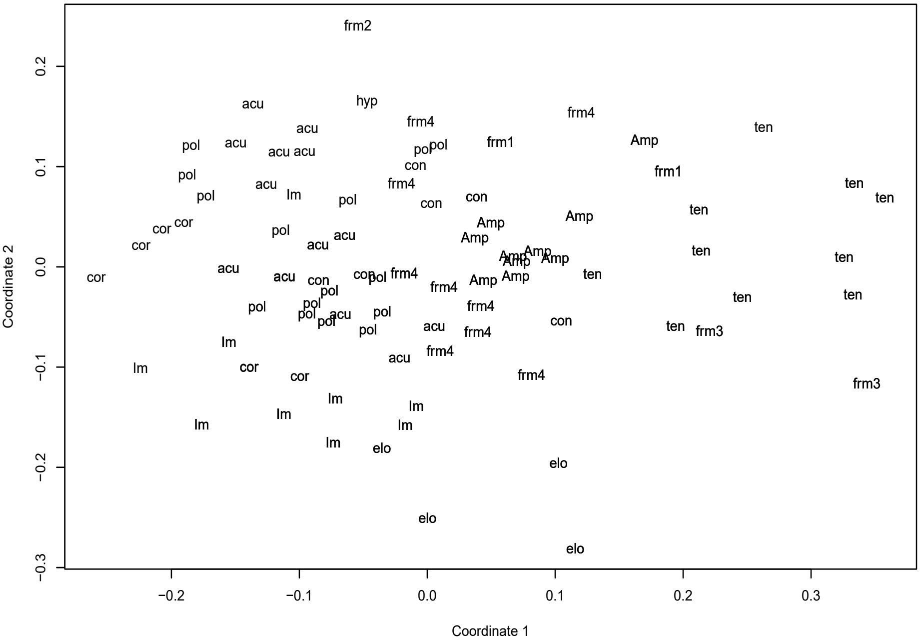

Figure 2-5. MDS clustering analysis with categorical variables, with typical C. javensis excluded from the data matrix. Abbreviation of taxa: acu = C. acuminatus, amp = C. amplijugus, con = C. congestiflorus, cor = C. corrugatus, elo = C. elopurensis, frm1 = Form 1, frm2 = Form 2, frm3 = Form 3, frm4 = Form 4, hyp = C. hypertrichosus, im = C. impar, pol = C. javensis var. polyphyllus, ten = C. tenompokensis.

remaining missing data were imputed using MICE (Table 2-2).

We made a comparison between two datasets: a dataset consisting only of numerical data (MICE + numerical data) and a dataset with numerical and categorical data (MICE + Numerical and Categorical data). Categorical data were included in order to show their influence on ordination and clustering.

Unsupervised methods were applied to both datasets, with and without categorical variables. Distance matrices were computed, using a Euclidean measure, and visualized using a heat map representation. Only one small cluster, consisting of *C. tenompokensis*, is observed in an otherwise homogenous heat map; using different distance measures did not affect this result. MDS produces

Table 2-6. Tree model confusion table for matrix without typical C. javensis. Abbreviation of taxon: acu = C. acuminatus, amp = C. amplijugus, con = C. congestiflorus, cor = C. corrugatus, elo = C. elopurensis, frm1 = Form 1, frm2 = Form 2, frm3 = Form 3, frm4 = Form 4, hyp = C. hypertrichosus, im = C. impar, jav = typical C. javensis, pol = C. Javensis var. polyphyllus, ten = C. tenompokensis. Number indicates the number of OTUs (or collection specimens).

PREDICTED

ACTUAL	acu	amp	con	cor	elo	frm1	frm2	frm3	frm4	hyp	im	jav	pol	Ten
Acu	4	0	0	1	0	0	0	0	0	0	0	0	8	0
Amp	0	10	0	0	0	0	0	0	0	0	0	0	0	0
Con	0	0	0	0	0	0	0	0	6	0	0	0	0	0
Cor	1	0	0	5	0	0	0	0	0	0	0	0	0	0
Elo	0	0	0	0	0	0	0	0	0	0	2	0	0	2
frm1	0	0	0	0	0	0	0	0	0	0	0	0	1	1
frm2	0	0	0	0	0	0	0	0	0	0	0	0	0	1
frm3	0	0	0	0	0	0	0	0	0	0	0	0	0	2
frm4	0	0	0	0	0	0	0	0	8	0	0	0	0	1
Hyp	0	0	0	1	0	0	0	0	0	0	0	0	0	0
Im	0	0	0	0	0	0	0	0	2	0	7	0	0	0
Jav	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Pol	3	1	0	0	0	0	0	0	0	0	0	0	11	0
Ten	0	0	0	0	0	0	0	0	0	0	0	0	0	10

analogous results with a compact cluster for all OTUs except *C. tenompokensis* with goodness of fit measures of 0.876 (Figure 2-2).

When categorical variables are included, members of *C. tenompokensis* remain separate away from the main cluster; the main cluster included all other species/specimens but is less condensed; the goodness of fit of the non-metric fit is again high (0.928). We note, however, that the main cluster is partitioned into two regions that are related to the character 'bract shape' but OTUs representing the same species were present in both partitions (Figure 2-3).

Hierarchical clustering of both datasets gave similar results independent of the agglomeration method used: *C. tenompokensis* forming a separate cluster in the presence of *Form 1* and *3*. Bootstrapping the cluster tree indicated the *C. tenompokensis* cluster to be very robust with approximately unbiased p-values

Table 2-7. Confusion matrix of Random forest analysis for matrix data without typical C. javensis. Abreviation of taxon: acu = C. acuminatus, amp = C. amplijugus, con = C. congestiflorus, cor = C. corrugatus, elo = C. elopurensis, frm1 = Form 1, frm2 = Form 2, frm3 = Form 3, frm4 = Form 4, hyp = C. hypertrichosus, im = C. impar, pol = C. J*avensis var. polyphyllus, ten = C. tenompokensis. Number indicates the number of OTUs (or collection specimens).*

PREDICTED

ACTUAL	acu	amp	con	cor	elo	frm1	frm2	frm3	frm4	hyp	im	pol	ten
Acu	8	0	0	0	0	0	0	0	0	0	0	5	0
Amp	0	10	0	0	0	0	0	0	0	0	0	0	0
Con	0	3	2	0	0	0	0	0	0	0	0	1	0
cor	0	0	0	4	0	0	0	0	0	0	1	1	0
elo	0	0	0	0	3	0	0	0	0	0	1	0	0
frm1	0	0	1	0	0	0	0	0	0	0	0	0	1
frm2	1	0	0	0	0	0	0	0	0	0	0	0	0
frm3	0	0	0	0	0	0	0	1	0	0	0	0	1
frm4	0	0	1	0	0	0	0	0	8	0	0	0	0
hyp	0	0	0	0	0	0	0	0	0	0	0	1	0
im	1	0	0	1	0	0	0	0	0	0	7	0	0
pol	2	2	0	1	0	0	0	0	0	0	0	10	0
ten	0	0	0	0	0	1	0	0	0	0	0	0	9

greater than 0.95 (Suzuki & Shimodaira, 2006). The other clusters consisted of a mixture of all OTUs representing a mix of the species but without a particular distribution in distinct clusters as observed earlier in the MDS with categorical data for the bract shape.

When using the above unsupervised methods, we did not include the information of the species. When using supervised methods, species are used as a supervisory signal for classification problems. From the confusion matrix of the Confusion Table (CT) analysis (Table 2-4), we infer that most species are classified as *C. javensis*, including one member of *C. tenompokensis*. Nine *C. tenompokensis* are classified correctly as is the complete sample of *C. impar*. Some *C. elopurensis*, *Form 1* and *Form 2* are misclassified as *C. tenompokensis* and some *C. javensis, C. elopurensis and C. corrugatus* were misclassified as *C. impar*.

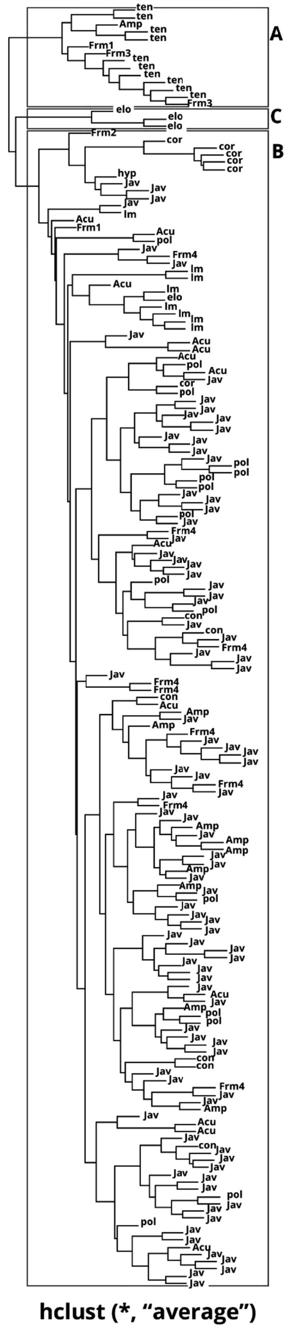

Figure 2-6. Dendogram based on all characters for all specimens using Hclust (‘mcquitty’) analysis.

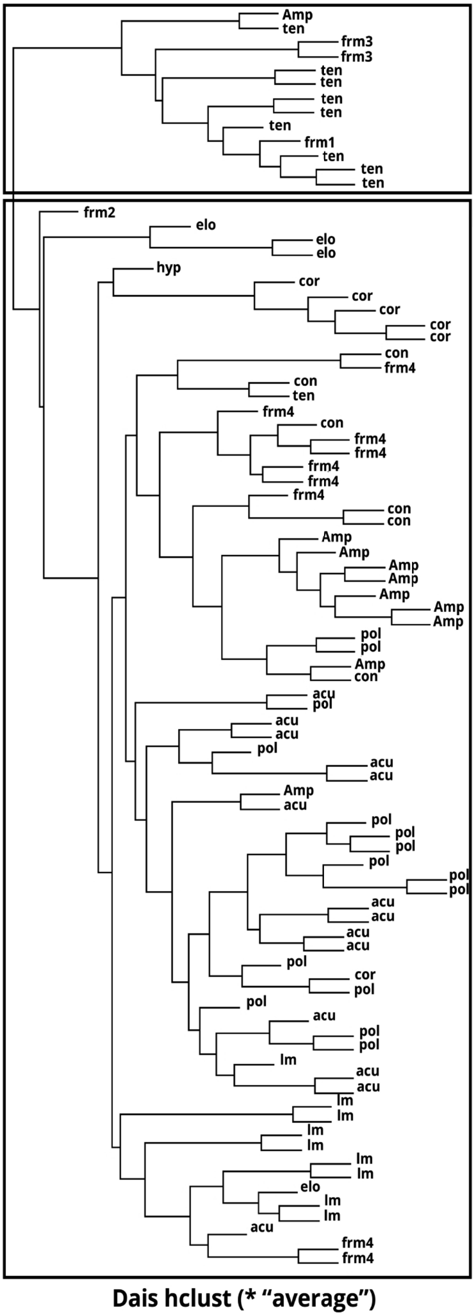

Figure 2-7. Dendogram based on all characters for all specimens excluding typical C. javensis using Hclust ('average') analysis.

In Table 2-5 we tabulate the confusion matrix computed using the Random Forest algorithm. Here the confusion matrix is based on a consensus tree, the decision tree with the highest occurrence in the forest.

The evaluation over the correct or incorrect prediction made by the model can be shown by the error rate value. Here the error rate of 28.25% is misleading due to the number of correctly classified members of *C. javensis*. If we disregard *C. javensis* from this table, 57 % species are misclassified and thus in disagreement with the experts. *Calamus tenompokensis* is, however, for 90 % correctly classified. These results point again to the problematic classification of the *C. javensis* complex based on the morphological features.

The Exclusion of typical C. javensis from the data matrix

Most extreme forms (often recognized as separate species) are only represented by a low number of OTUs, the majority of specimens represent typical *C. javensis*. It may well be that the bulk of typical *C. javensis* OTUs obscures groupings of the extreme forms, therefore we excluded all typical *C. javensis* OTUs from the data matrix. The unsupervised MDS analysis without categorial variables showed no significant differences (Figure 2-4) with when we included typical *C. javensis* in the analysis (Figure 2-2). Two clusters were observed: one big cluster consisting of all forms except members of *C. tenompokensis* (ten in Figure 2-4) and one member of *Form 1* and 2 members of *Form 3*. Members of *C. impar* (im) and *C. elopurensis* (elo) are at the periphery of the big cluster. Form 1 and Form 4 were mingled- with the big cluster (Figure 2-4). The same result was obtained when the categorical characters were included (Figure 2-5), but the clustering is less compact, but the group of *C. tenompokensis*, *Form 1* and *Form 3* still show a tendency to form a separate cluster. Typical here is the grouping of *C. acuminatus* (acu) and *C. javensis* var. *polyphyllus* (pol) (Figure 2-5).

The exclusion of typical *C. javensis* from the data matrix, gave a slightly different result for the data prediction in the supervised methods as shown in Table 2-6 and Table 2-7. All Forms (frm1–frm4) were misclassified as *C. tenompokensis*. Members of *C. acuminatus* and *C. javensis* var. *polyphyllus* were alternately misclassified. Species *C. corrugatus* and *C. impar* were almost correctly classified and the only *C. hypertrichosus* was misclassified as *C. corrugatus* or *C. javensis* var. *polyphyllus*. Almost all members of *C. amplijugus* were perfectly classified.

Hierarchical Custer Analysis produced dendograms for 2 sets of data, one including *C. javensis* (Figure 2-6) and one excluding *C. javensis* (Figure 2-7). The two dendograms support the distribution of the characters from the MDS analysis and the prediction made by the TREE and rF analysis. There are two groups, one small group (group A) consisting of *C. tenompokensis*, Form 1 and Form 3 (Figure 2-6 and Figure 2-7); and the rest formed a large group (group B) consisting of all typical *C. javensis* and the rest of the species. Noted here a small group within group B, indicated as C, contains members of *C. elopurensis* only (Figure 2-6). The exclusion of typical *C. javensis* yielded a similar result, there are also 2 main groups but here, in group B, all the species with peculiar characters are grouped together (Figure 2-7).

2.4. Discussion

Herbarium studies and characters

The characters used in the analyses did not lead to a clear distinction between species except for the recognition of *C. tenompokensis* separate from *C. javensis*. There are two reasons why *C. javensis* appears to be a good, but variable species. 1. The variation in characters is well spread among all OTUs, thus extreme forms do not show a constant set of different character states. 2. Specimens with intermediate forms between the distinguished taxa were observed during the fieldwork and in the herbarium material.

Not all characters and characters states were used in the multivariate analyses (often present in a single taxonomic unit or too many missing data). These and the characters used are discussed here.

Leaves and Leaf sheath are the two most variable characters. Beccari (1913) described *C. hypertrichosus* based on a sterile specimen from Borneo; the type has leaflets in 4 pairs and a very long petiole (c. 30 cm long). Seemingly the type was a young plant. Other specimens examined were full-grown (flowers, fruits) and showed leaves with 7 pairs of leaflets in an arrangement that resembles typical *C. javensis*. The type specimen of *C. hypertrichosus* only differs from typical *C. javensis* in having the leaflets covered with whitish hairs. It was the only specimen with this type of indumentum, and, therefore, the character state was not used in the multivariate analyses. On the other hand, because the white hairs are only shown by a single, immature specimen, no taxonomic value is given to this character state.

Smooth or almost smooth leaf sheaths were found among the OTUs called *C. corrugatus*. Instead of having spines, members of *C. corrugatus* have corrugated ring-like lines on the leaf sheath. However, the arrangement of the leaflets is similar to that of typical *C. javensis*. The corrugate lines are not completely circular in all specimens and a specimen from Central Kalimantan (*Mogea 3615*) has the typical *Calamus javensis* feature of an incompletely corrugated leaf sheath.

Leaflet shape is another character that is variable and apparently distributed over all samples. Already in typical *C. javensis* there is a broad variety in shapes.

Calamus amplijugus can easily be recognized by its many broadly elliptic leaflets, a long pistillate rachilla and narrow fruit scales. In the confusion matrix (Table 2-6 and Table 2-7) *C. amplijugus* was well recognized and correctly classified. Intermediate specimens with *C. javensis* were found among the herbarium specimens and several were observed during field work. All these collections are from Sabah (N. Borneo) only.

The number of leaflet pairs varies with age (less when young, more when older), but several forms were distinguished based on the number of leaflet pairs. *Calamus javensis* var. *polyphyllus* and *C. acuminatus* are synonymous. Beccari (1908) described *C. javensis* var. *polyphyllus* based on the presence of 9–12 pairs of leaflets; and later he (1913) described *C. acuminatus* as a species with 10–11 pairs, but with an acuminate leaf tip (see also below).

A low number (2 or 3) of leaflet pairs is observed among the members of *C. elopurensis*, *C. impar* and *Form 4*. The arrangement of these leaflets (not used as character in the multivariate analyses) can differ from typical *C. javensis*. The penultimate pair in *C. javensis* is almost always opposite and close to the flabellum; whereas *Form 4* has subopposite or alternate penultimate leaflets of which one is very close to the terminal pair. *Form 4* resembles *C. elopurensis* (3 pairs of leaflets) and *C. impar* (2 pairs of leaflets). However, the size of the leaflets of *C. impar* is smaller than those of *C. elopurensis* (up to 20 cm and up to 35 cm, respectively). Moreover, in *C. elopurensis* the staminate inflorescence has conspicuous bracts along the peduncle with a long limb of up to 10 cm long; in *C. impar* the bract of peduncle has tubular, c. 13 cm long, closely sheathing, distally with an extended limb to 2.5 cm long; whereas the peduncle bracts of *Form 4* are tubular, closely sheathing, later with extended elliptic limb to 5 mm long, c. 35 cm long.

Imputation (MICE) eliminated several generative characters, because these were not present in all OTUs and did not contribute significantly to the results of the multivariate analyses. One character remaining was the rachilla length. Rachilla length was expected to differentiate *C. congestiflorus* from the other taxa in the complex. Dransfield (1982) described this taxon because of a congested, short rachilla (2.5–5 cm long) and a lanceolate rachis limb. However, specimens were found with a variable short rachilla length and various degrees of congestion.

Multivariate analyses and taxon recognition

The conclusion from all multivariate analyses is that the *C. javensis* complex with its 9 described species, several infraspecific taxa and 4 tentative forms can be reduced to only 2 species. The majority of specimens constitutes *C. javensis* and a small group can be recognized as *C. tenompokensis*, retained as a different species. The tentative forms still need further attention, but extra material is needed to correctly establish their status. This division is supported by all analyses used.

The unsupervised analysis showed generally that two clusters can be distinguished: All OTUs forming one group (*C. javensis*), except those forming *C. tenompokensis*. The big cluster of typical *C. javensis* indicates an even distribution of all variation among the OTUs. There were no significant groups within this big cluster, only an indication that several forms could be distinguished. In fact, the analyses without typical *C. javensis* gave a similar, but somewhat less compact clustering, which means that all extreme forms, recognized as separate species, do not differ in distinct character states from each other, intermediates are always present. The two confusion tables of the data set including typical *C. javensis* (Table 2-4 and Table 2-5) support the conclusion of a less restricted definition of *C. javensis*, because many to most to all members of special forms appeared to be part of typical *C. javensis*.

Members of *C. tenompokensis* formed a separate group. The clustering made with MDS, when tested, showed in the confusion tables (Table 2-4 and Table 2-5) that only one OTU of *C. tenompokensis* was misclassified. The members of the *C. tenompokensis* group share the presence of flabellate leaflets with *C.*

javensis. However, *C. tenompokensis* can easily be distinguished from other taxa within the *C. javensis* complex by the vegetative parts: the angular petiole and rachis, a very different leaf sheath appearance, the sheaths being massive and robust, and the number and arrangement of the leaflets, *C. tenompokensis* has 9 pairs of large, lanceolate leaflets, which are almost always regularly arranged. The staminate inflorescence resembles those of *C. javensis*, but the base of the calyx is swollen. The pistillate inflorescences have rachilla bracts that are different in the broadly cupuliform limb (specimen *Chew & Corner RSNB 1892*).

Several peculiar forms, *Form 1–4* cannot be treated well for now, because of a lack of sufficient samples. *Form 4* is already discussed above.

Form 2 is only represented by a single sample collected from the Philippines, and the leaflets have scattered setae along veins and one black long seta on the main vein on the abaxial surfaces. The leaflets resemble the arrangement of typical *C. javensis*, but with quite long petioles and peculiar rachis spines.

Morphological examination of *Form 1*, consisting of two specimens, shows that one is quite distinct, grouping with *C. tenompokensis* (Figure 2-2, Figure 2-3, Figure 2-4 and Figure 2-5), the other is part of the typical *C. javensis* group, a result confirmed by the confusion tables (Table 2-4, Table 2-5, Table 2-6, and Table 2-7). This second specimen appears to be intermediate. *Form 1* is, like *C. tenompokensis* and *Form 3*, only known from higher altitudes (1200–1500 m). One member of *Form 1* is clearly different from *C. javensis* in the leaves and arrangement of leaflets, leaf sheath features, the short staminate inflorescences, tubular, warty rachilla bracts and the lack of a flagellum and a cirrus. *Calamus tenompokensis* resembles this taxon, also lacks the flagellum, but *Form 1* has smaller leaflets (c. 15 cm long, *C. tenompokensis* up to 35 cm long), the tubular and warty rachilla bracts (*C. tenompokensis*: explanate and smooth) and the staminate inflorescence borne on the 2nd branch (3rd branch in *C. tenompokensis*). *Form 3*, collected from Sarawak (*S 44664*) and Sabah (*SAN 144483*), remains separate from the typical *C. javensis* cluster and is closer to *C. tenompokensis* in all analyses (Figure 2-2, Figure 2-3, group A in Figure 2-6). Specimen *SAN 144483* has leaflets without hairs, but there are remnants of hairs. This entity differs from other specimens in the *C. javensis* complex, because it is stemless, has a peculiar arrangement of the tomentose leaflets, and the stigmas show a curled tip. Tomentose leaflets are also present in *C. hypertrichosus*, but the latter is climbing, has 7 leaflets pair (compare to 3 pairs in this taxon) with the leaflets arranged as in typical *C. javensis*.

The resemblance between *C. acuminatus* and *C. javensis* var. *polyphyllus* was shown to be high. The confusion tables (Table 2-4, Table 2-5, Table 2-6, and Table 2-7) also clearly indicate this as both species are interchangeable as they alternately show up as misclassified. Some members of these two taxa are also misclassified as typical *C. javensis*, which also indicates that they are part of this cluster, because they share many character states. This is also clearly shown in Figure 2-7, where members of *C. acuminatus* and *C. javensis* var. *polyphyllus* are in the same cluster. Beccari (1908) stated that *C. acuminatus* is indistinguishable from some *C. javensis* varieties. Based on the observations of the first author, *C. acuminatus* is similar to *C. javensis* var. *polyphyllus* (Beccari, 1908), but *C. acuminatus* has an almost smooth leaf sheath, smaller flowers and fruits and the bracts of the peduncle are more cupuliform. Beccari (1913) also suspected that *C. acuminatus* may be considered as subspecies. However, there are specimens observed with intermediate character states between *C. acuminatus* and *C. javensis*, e.g., specimens with leaflets varying between lanceolate and narrowly elliptic and leaf sheaths ranging from smooth to moderately spiny; the inflorescence can range from fine to the size of typical *C. javensis*. In conclusion, both *C. acuminatus* and *C. javensis* var. *polyphyllus* are part of the typical *C. javensis* complex and no separate entities.

Calamus impar and *C. elopurensis* also show a tendency to group together, but always in close proximity of the *C. javensis* cluster. When the categorical data were added the clustering became less distinct. The differences with *C. javensis* are discussed above. The confusion matrix of both analysis TREE and Random Forest (Table 2-4, Table 2-5, Table 2-6, and Table 2-7) resulted in a clear indication that each form is distinct, but they are not significantly different enough to separate them from typical *C. javensis*.

All other species are part of the typical *C. javensis* complex as indicated by Figure 2-2, Figure 2-3, Figure 2-4 and Figure 2-5.

The dendogram of all specimens (Figure 2-6) shows that *C. javensis* specimens are among almost all members of the complex, except group A of *C. tenompokensis*, *Form 1* and *Form 3*. *Calamus javensis* is a polymorphic species whereby many intermediate characters were found among the members of the complex (Group B, Figure 2-6). A number of *C. corrugatus* and *C. impar* specimens form subgroups but other members of both species are mingled with the rest of the specimens. This can best be seen in Figure 2-7, where the distinctive variation among extreme forms is shown.

Calamus javensis was recently discussed by Syam et al. (2016) in their work on the *Calamus flabellatus* complex in the Malesian region. They included *C. javensis* and other species within the complex, like *C. acuminatus, C. amplijugus, C. congestiflorus, C. corrugatus, C. hypertrichosus* and *C. ruvidus*. Their use of *C. flabellatus* to indicate the complex is unfortunate as *C. javensi*s is the oldest name and the most widespread species. The results of Syam et al. (2016) differ from our study. Two species, *C. corrugatus* and *C. hypertrichosus*, were separate from the *C. javensis* cluster in their cluster analysis. Our analysis showed that *C. corrugatus* was among typical *C. javensis*, which is corroborated by the presence of intermediate specimens; whereas *C. hypertrichosus*, with only one observed specimen, was also placed among typical *C. javensis*, though slightly separated in Figure 2-5 due to the distinctive hairs. Like with us, Syam et al. (2016) found a major cluster consisting of *C. javensis* and satellite species. Strangely enough, they named all groups of *C. flabellatus*, but did not do the same with those of *C. javensis*. We do not consider their study to be very representive, as the samples are few (seemingly 20 OTUs in the cluster analysis made with an unknown program) and not well-spread over the variation of the species, mainly N. Borneo is lacking. Unfortunately, as no data matrices are provided their work cannot be reproduced.

2.5. Conclusion

The supervised classification results in a set of rules that can be represented as a decision tree that experts would use to classify a species. The resampling methods of Random Forest produces many decision trees and thus creates, as its name indicates, even a decision tree forest. In the current study we interpret the miss-classifications observed, based on the morphological variables, as disagreement between the virtual experts and the real experts. The observation that most of the samples are classified as *C. javensis* points in the direction of a variable species or cryptic species rather than distinct species with the exception of *C. tenompokensis*. The multivariate analysis clearly indicates that *C. javensis* is a distinct species in its very broad definition, with most other entities, ever separated from it as species, as synonyms. The only exception is *C. tenompokensis*, which can still be recognized as distinct. If more material becomes available, some of the four forms might also be recognized as species.

A future molecular analysis will be used to test the two species hypothesis as put forward by the multivariate analyses. Once agreement exists a new classification and description of *C. javensis* and *C. tenompokensis* will be provided.