Theor Appl Genet (1992) 84:286–290

Restriction fragment length polymorphism (RFLP)-based phylogenetic analysis of *Musa*

N. J. Gawel [1], R. L. Jarret [2], and A. P. Whittemore [2]

[1] USDA/ARS Western Cotton Research Laboratory, 4135 E. Broadway Rd, Phoenix, AZ 85040, USA
[2] USDA/ARS Regional Plant Introduction Station, 1109 Experiment St, Griffin, GA 30223, USA

Received August 12, 1991; Accepted November 15, 1991
Communicated by G. Wenzel

Summary. Random genomic probes were used to detect RFLPs in 19 *Musa* species and subspecies. A total of 89 phylogenetically informative alleles were scored and analyzed cladistically and phenetically. Results were in general agreement with morphology-based phylogenetic analyses, with the following exceptions: our data unambiguously places *M. boman* in section *Australimusa*, and indicates *M. beccarii* is very closely related to *M. acuminata*. Additionally, no support was found for the separation of section *Rhodochlamys* from section *Musa*. A comparison of morphology-based and RFLP-based phylogenetic analyses is presented.

Key words: Taxonomy – Banana – RFLP – *Musa* – Phylogeny

Introduction

The family Musaceae is composed of two genera, *Musa* and *Ensete*. *Musa* is composed of 25–35 species. It is found primarily in tropical regions from India to Polynesia, with maximum diversity in Indonesia (Simmonds 1962). Agronomically, *Musa* is an important genus: *M. acuminata* and hybrids of *M. acuminata* and *M. balbisiana* account for a majority of the edible bananas and plantains grown in the world. Bananas and plantains are a significant food source, especially in devloping countries; *Musa* is the principal carbohydrate source for over 100 million people world-wide (Rowe 1981). Concerns about yield declines due to the spread of the leaf-spot disease Black Sigatoka and the genetic erosion of *Musa* germplasm at its centers of diversity have generated renewed interest in *Musa* germplasm collection, identification, and taxonomy.

Despite its economic importance, *Musa* has received relatively little attention from taxonomists. The genus *Musa* was originally divided into three sections: *Physocaulis*, and *Eumusa* (Baker 1893). Cheesman (1947), using seed characters (size, shape, smoothness, markings), elevated Baker's section *Physocaulis* to generic status, to be included in *Ensete*. He then divided *Musa* on the basis of chromosome number, creating sections *Eumusa* and *Rhodochlamys* ($n=11$) and *Callimusa* (mostly $n=10$) and *Australimusa* ($n=10$). Cheesman's sectional classification is still in use, but is clearly in need of revision (Shepherd 1990). A recent phenetic analysis based on morphological and cytological characters (Simmonds and Weatherup 1990) found a very low level of consistency among the characters (in the principal component analysis, for instance, the first two components accounted for only 35.5% of the variation in the data), and suggested that section *Musa* is heterogeneous. Furthermore, crossing data do not always agree with the recognized sections (Simmonds 1954; Shepherd 1990), and a few species, notably *M. ingens* and *M. beccarii*, do not fit well into any of the recognized sections (Simmonds 1960). Recently Simmonds and Weatherup (1990) divided section *Musa* into two informal subgroups: "Eumusa-1" and "Eumusa-2".

DNA restriction fragment length polymorphisms (RLFPs) are a new and very suitable instrument for phylogenetic studies, as has demonstrated in a number of species (Song et al. 1988 a, b; Song et al. 1990; Havey and Muehlbauer 1989; Miller and Tanksley 1990). RFLPs have proven to be a valuable source of taxonomic characters with relatively low levels of homoplasy (Debner et al. 1990). In this report we demonstrate the use of restriction fragment length polymorphisms (RFLPs) in

Correspondence to: R. L. Jarret

the examination of phylogenetic relationships among representative species from the genus *Musa*.

Materials and methods

Plants used in this study are listed in Table 1; all accessions were obtained from the sources listed. DNA extraction, radioactive labelling, and autoradiography were essentially as described earlier (Gawel and Jarret 1991). Changes were made in the Southern-blotting and hybridization procedures as follows: DNA was bound to nylon membranes via UV cross-linking instead of vacuum baking. Membranes were prehybridized 3–4 h at 65 °C (6 × SSC, 5 × Denhardt's solution, 0.1% SDS, 0.001% sonicated, denatured salmon sperm DNA). Hybridizations were for 12–16 h at 65 °C (6 × SSC, 0.1% SDS, 0.001% sonicated, denatured salmon sperm DNA). Random genomic libraries of *Eco*RI-digested *M. acuminata* and *M. balbisiana* DNA were constructed in pUC18 (Maniatis et al. 1982).

Total DNA extracts of *Musa* species to be analyzed for RFLPs were digested with *Msp*I, *Eco*RI, *Hin*dIII, or *Bam*HI according to the manufacturer's directions. Data were recorded by treating each band as a separate allele, scored as present or absent. Cladistic analyses were performed using phylogenetic analysis using parsimony (PAUP) software (Swofford 1985), with the root-midpoint, swap-global, and mulpars-on options. The BOOT program (with global option) of the PHYLIP software package (Felsenstein 1985) was used to compute bootstrap analyses. Bootstrap samples were constructed by linking all alleles from a single probe/enzyme combination using the Factor option of BOOT. The multiple-state character data of Simmonds (1962) and Simmonds and Weatherup (1990) were converted to binary characters with the FACTOR program of the PHYLIP software package. Consistency analyses were conducted by eliminating all inconsistent allele-states from the data set and analyzing the remaining consistent allele-states with PAUP. Principal component analyses were performed using the PRINCOMP procedure of SAS.

Table 1. *Musa* species examined for RFLP-based phylogenetic analysis

Species	Section	*n*
M. acuminata ssp. *banksii* (F. Muell) Simmonds[a]	*Musa*	11
M. acuminata ssp. *burmannica* Simmonds[a]	*Musa*	11
M. acuminata ssp. *malaccensis* (Ridl.) Simmonds[a]	*Musa*	11
M. acuminata ssp. *microcarpa* (Becc.) Simmonds[a]	*Musa*	11
M. acuminata ssp. *siamea* Simmonds[a]	*Musa*	11
M. acuminata ssp. *truncata* Ridl.[a]	*Musa*	11
M. balbisiana Colla.[b]	*Musa*	11
M. basjoo Sieb.[a]	*Musa*	11
M. liukiuensis[b,g]	*Musa*	11
M. schizocarpa Simmonds[c]	*Musa*	11
M. beccarii Simmonds[d,f]	*Callimusa*	9[h]
M. coccinea Andr.[a]	*Callimusa*	10
M. ornata Roxb.[e]	*Rhodochlamys*	11
M. velutina Wendl. & Drude[d]	*Rhodochlamys*	11
M. boman Argent[c]	*Australimusa*	?
M. lolodensis Cheesman[c]	*Australimusa*	10
M. maclayi ssp. *ailuliai* F. Muell.[c]	*Australimusa*	10
M. peekelii ssp. *peekelii* Lauterb.[c]	*Australimusa*	10
M. textilis Nee.[b]	*Australimusa*	10

[a] Obtained from the International Network for the Improvement of Bananas and Plantains germplasm transit center at the Katholieke Universiteit, Leuven, Belgium
[b] Obtained from Fundacion Hondurena de Investigacion Agricola (FHIA), La Lima, Honduras
[c] Obtained from QDPI, Maroochy Horticultural Research Station, Nambour, Queensland, Australia
[d] Obtained from CATIE, Turrialba, Costa Rica
[e] Obtained from Dade County Fruit and Spice Park, Dade County, FL, USA
[f] Identified as *M. beccarii* by H. Tezenas Du Montcel, INIBAP, Montpellier, France
[g] Obtained as seed
[h] Also identified as $n=10$ (H. Tezenas du Montcel, personal communication)

Results

Southern blots were probed with a total of 66 genomic DNA probes; 96 alleles were revealed, 89 of which were phylogenetically informative. Of the probes used, 45 were isolated from the *M. acuminata* library and 21 from the *M. balbisiana* library. A comparison of *M. acuminata* versus *M. balbisiana* probes revealed no appreciable difference in their ability to detect polymorphisms: 68% of the probes used were from the *M. acuminata* library, and these detected 72% of the scored alleles. Of the restriction enzymes used, *Msp*I, *Hin*dIII, and *Bam*HI digests all revealed approximately equal numbers of alleles (32%, 27%, and 27%, respectively); *Eco*RI detected 14% of the alleles.

Cladistic analysis yielded a single most parsimonious tree, depicted as the unrooted cladogram in Fig. 1. Consistency analysis produced a set of 57 consistent alleles; the resulting cladogram is identical to that illustrated in Fig. 1, except that *M. basjoo* is the sister group to the *M. balbisiana*/*M. liukiuensis* branch rather than forming a separate clade (cladogram not shown). Preliminary results from cladistic analyses using *Ensete ventricosum* as the outgroup indicate the root in Fig. 1 to be in the *M. coccinea*-*M. basjoo*-*M. balbisiana* region.

The results depicted in Fig. 1 illustrate two clear groupings, one containing species from sections *Musa* and *Rhodochlamys*, the other containing species from sections *Australimusa* and *Callimusa*. The only discrepancy in these groupings is the placement of *M. beccarii* ($n=9$, also reported to be $n=10$; H. Tezenas du Montcel, personal communication), which has previously been placed in section *Callimusa* (Simmonds and Weatherup 1990). Our analysis determined the DNA of *M. beccarii* to be very similar to that of *M. acuminata* (section *Musa*).

The branches connecting the *M. acuminata*-complex with *M. schizocarpa*, *M. ornata*, and *M. velutina* are not well supported in the bootstrap analysis, and these rela-

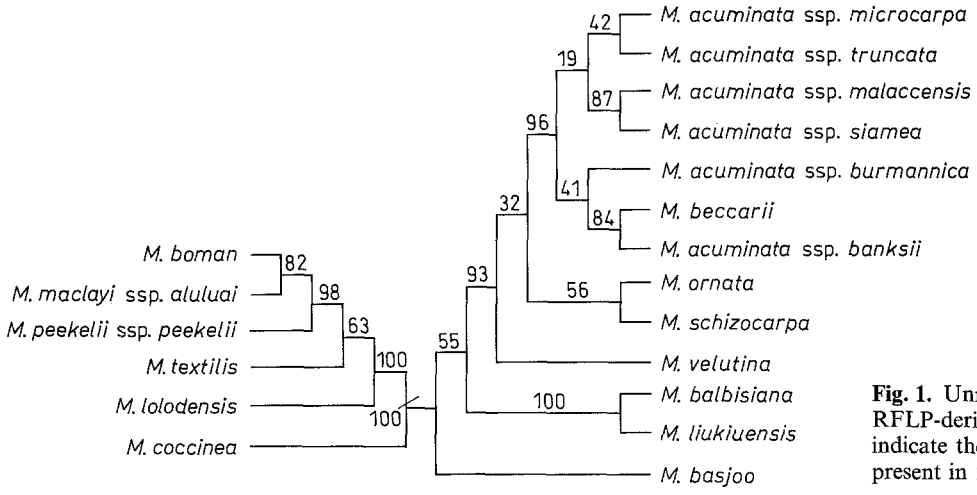

Fig. 1. Unrooted cladogram based upon RFLP-derived alleles. Numbers at nodes indicate the number of times a clade was present in 100 bootstrap analyses

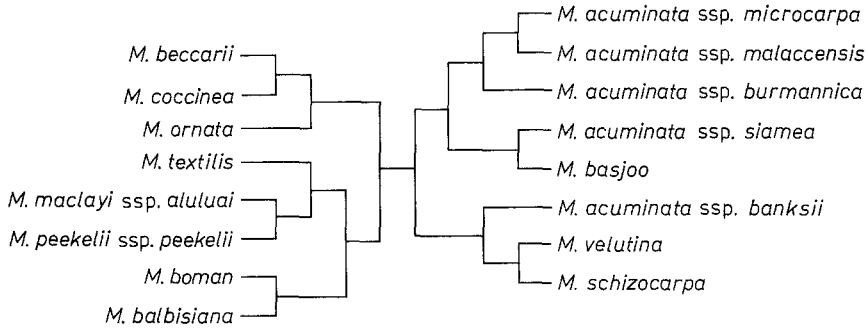

Fig. 2. Unrooted cladogram of the morphological-based descriptors of Simmonds (1962) and Simmonds and Weatherup (1990)

tionships should be considered unresolved. However, the branch leading to the unresolved group consisting of these five species is present in 93 out of 100 bootstrap samples. The two species of section *Rhodochlamys*, *M. ornata* and *M. velutina*, are not closely associated, but seem to be independently derived from section *Musa*. The close relationship among these species is also supported by crossing (Simmonds 1962) and cytological (Simmonds 1954) data. In the phenetic analysis based on morphological characters (Simmonds and Weatherup 1990), the section *Rhodochlamys* species do not group consistently. We analyzed these morphological characters (Simmonds 1962; Simmonds and Weatherup 1990) cladistically and found that section *Rhodochlamys* species remained divided between separate branches (Fig. 2).

M. liukiuensis is very similar to *M. balbisiana* and is considered to be a synonym of the latter species (Shepherd 1990). This relationship is reflected in our analysis: these two species were sister taxa in all 100 of the bootstrap analyses. *M. basjoo* consistently forms a separate branch very near the division of the $n=10$ and $n=11$ groups. The data indicate that *M. basjoo* shares alleles with many other species, including *M. acuminata* ssp. *truncata*, *M. boman*, and *M. balbisiana*, a fact reflected in the low bootstrap statistic.

Most branches on the *Australimusa*/*Callimusa* side of the cladogram in Fig. 1 are supported by high bootstrap values. The *Australimusa* species were always grouped together in the cladistic analysis, and the branch pattern within the section is well resolved. *Musa coccinea* (section *Callimusa*) is on a well-defined branch separate from the species of section *Australimusa*. *Musa boman* has been alternately placed in section *Australimusa* (Argent 1976) and section *Musa* (Simmonds and Weatherup 1990). Our data strongly support the former placement.

The analysis by Simmonds and Weatherup (1990) places *M. beccarii* close to *M. coccinea* in section *Callimusa*. We have found that *M. beccarii* remains closely associated with *M. coccinea* when the Simmonds and Weatherup data is analyzed cladistically (Fig. 2). However, cladistic analysis of our DNA RFLP data unambiguously places *M. beccarii* within the *M. acuminata* subspecies complex (Fig. 1). The close association between *M. beccarii* and the *M. acuminata* subspecies persists when the DNA data are analyzed phenetically (Fig. 3).

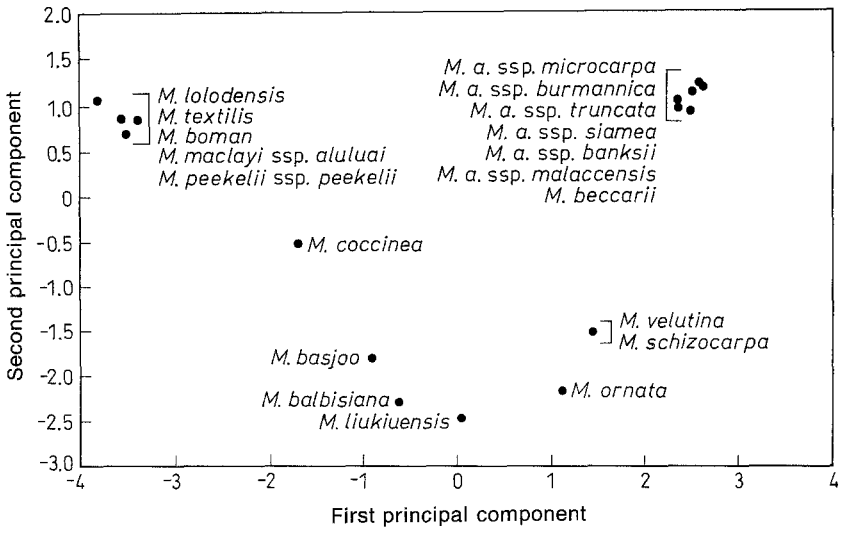

Fig. 3. Principal component analysis of RFLP-derived alleles. Principal components 1 and 2 account for 68.8% of the variability among the species examined

Principal component analyses were conducted on the RFLP-derived alleles. Results indicate that the first two principal components account for 68.8% of the variation among the species examined. As illustrated in Fig. 3, the *M. acuminata* subspecies complex (and *M. beccarii*) and species from section *Australimusa* form two widely separated clusters. Species from sections *Rhodochlamys*, *Callimusa*, and other section *Musa* species are distributed in between the clusters. As in the cladistic analysis, section *Rhodochlamys* is not well separated from section *Musa*. *M. coccinea* (section *Callimusa*) is well separated from other species in this analysis, confirming its distant relationship with members of the other sections.

Discussion

This RFLP study yielded a large number of phylogenetically useful alleles. The resulting cladogram is quite robust, with many clades shared between 95 or more of the bootstrap analyses. It is thus possible to reevaluate the taxonomic groupings currently accepted for *Musa*, and the characters on which they are based.

The separation of *Musa* into two groups with chromosome base numbers of $n=10$ and $n=11$ (plus *M. beccarii*) is strongly supported by our data. Within the $n=10$ group, our analysis generally supports previous classifications. The five species of section *Australimusa* grouped together on all of the bootstrap analyses. This strongly supports the inclusion of *M. boman*, whose original (Argent 1976) taxonomic position has been questioned (Simmonds and Weatherup 1990). The isolated position of *M. coccinea*, the only typical member of section *Callimusa* studied, is consistent with its treatment as a separate section. However, the morphologically and chromosomally aberrant *M. beccarii*, which has been placed in section *Callimusa* (Simmonds and Weatherup 1990), is unambiguously grouped with the $n=11$ species.

Within the $n=11$ group, the accepted classification is not well supported by our analysis. Currently, species with tall pseudostems and horizontal to pendent inflorescences are placed in section *Musa*, while species with short pseudostems and erect inflorescences have been treated as a separate section, *Rhodochlamys*, or (in the case of *M. beccarii*) excluded from the group altogether (Simmonds 1962). Our analysis includes two species of section *Rhodochlamys*; they arise independently from different parts of the $n=11$ clade. Because of low bootstrap statistics in this section of the cladogram, it is not possible to reject the hypothesis that they are sister taxa, but the data certainly provide no support for the recognition of section *Rhodochlamys*. Shepherd (1990) and Simmonds (1962) present cytogenetic evidence that also negates a clear distinction between *Musa* and *Rhodochlamys*. Within section *Musa*, Simmonds and Weatherup (1990) have recently emphasized the isolated position of *M. balbisiana* and its close relatives. Our data support this, but suggest an equally great isolation for *M. basjoo*.

Like the species of section *Rhodochlamys*, *M. beccarii* has a short pseudostem and erect inflorescence and has been consistently placed outside section *Musa* (Simmonds 1960). Our analysis shows *M. beccarii*, *M. ornata*, and *M. velutina* arising independently from different groups of section *Musa*. This suggests that short pseudostems and erect inflorescences have evolved repeatedly in *Musa*; thus these characters are of questionable value as indicators of relationship.

We have previously described a high level of similarity between the chloroplast DNA of *M. beccarii* and the *M. acuminata* subspecies complex, and suggested the possibility that *M. beccarii* is the product of an interspecific hybridization event between *M. acuminata* and

another unknown species (Gawel and Jarret 1991). In light of our current data, this seems unlikely. *M. beccarii* and the *M. acuminata* subspecies form a very stable group present in 96 out of 100 bootstrap analyses. The RFLP-based alleles scored for *M. beccarii* were similar to those of the *M. acuminata* subspecies complex. If *M. beccarii* were an interspecific hybrid, only a portion of the alleles would be similar to *M. acuminata*, and it would not associate consistently with only the *M. acuminata* subspecies complex. Thus, our data strongly suggest that *M. beccarii* is a chromosomally aberrant relative of *M. acuminata*, of relatively recent origin.

Comparison of the RFLP-based data in Figs. 1 and 3 with the morphology-based data in Fig. 2 (and in Simmonds and Weatherup 1990) illustrates differences between phylogenetic analyses based upon morphological versus molecular characteristics. Extreme differences in morphological characteristics are not necessarily indicative of the same degree of genetic difference (Hamrick and Godt 1989). This may be the case with *M. beccarii*: our data show it shares a recent common ancestry with *M. acuminata*. Morphological differences between *M. beccarii* and *M. acuminata* seem to be the result of rapid character evolution in the lineage leading to *M. beccarii*. Similarities between *M. beccarii* and section *Callimusa* may be due solely to convergent evolution.

The results presented in this paper demonstrate the use of RFLP-based alleles as an alternative to morphological and cytological characters in phylogenetic analyses of *Musa*. In contrast to the accepted morphology-based taxonomy of the genus, our data unambiguously place *M. boman* in section *Australimusa* and indicates that *M. beccarii* is very closely related to *M. acuminata*. Data which so strongly indicate a close relationship between *M. beccarii* and *M. acuminata* are, admittedly, unexpected. This discrepancy between morphology-based and DNA-based phylogenetic analyses merits further investigation [the specimen of *M. beccarii* used in this study was obtained from the germplasm collection of CATIE (Turrialba, Costa Rica) and identified as *M. beccarii* by H. Tezenas du Montcel (personal communication)]. In addition to these findings, we found no support for the separation of section *Rhodochlamys* from section *Musa*. Further research is needed to provide information necessary to clarify the unresolved (*M. schizocarpa*, *M. ornata*, *M. velutina*, *M. basjoo*, and the *M. acuminata* subspecies) portions of our analysis, and to classify species not examined in this study.

References

Argent GCG (1976) The wild bananas of Papua Guinea. Notes R Bot Gard Edinburgh 35:77–114

Baker JG (1893) A synopsis of the genera and species *Musaceae*. Ann Bot 7:189–229

Cheesman EE (1974) Classification of bananas. Kew Bull 97–117

Debner T, Salamini F, Gebhardt C (1990) Phylogeny of wild and cultivated *Solanum* species based on nuclear restriction fragment length polymorphisms (RFLPs). Theor Appl Genet 79:360–368

Felsenstein J (1986) PHYLIP (Phylogeny inference package) version 2.9. University of Washington, Pullman, Wash.

Gawel N, Jarret RL (1991) Cytoplasmic genetic diversity in bananas and plantains. Euphytica 52:19–23

Hamrick JL, Godt MJ (1989) Allozyme diversity in plant species. In: Brown ADH, Clegg MT, Kahler AL, Weir BS (eds) Plant population genetics, breeding, and genetic resources. Sinauer Assoc, Sunderland, Mass., pp 43–63

Havey MJ, Muehlbauer FJ (1989) Variability for restriction lengths in lentil. Theor Appl Genet 77:839–843

Maniatis T, Fritsch EF, Sambrook J (1982) Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory press, Cold Spring Harbor, N.Y.

Miller JC, Tanksley SD (1990) RFLP analysis of phylogenetic relationships and genetic variation in the genus *Lycopersicon*. Theor Appl Genet 80:437–448

Rowe PR (1981) Breeding an intractable crop. Bananas. In: Rachie KO, Lyman JM (eds) Genetic engineering for crop improvement. Working Papers, The Rockefeller Foundation, pp 66–84

Shepherd K (1990) Observations on *Musa* taxonomy. In: Jarret RL (ed) Identification of genetic diversity in the genus *Musa*. INIBAP, Parc Scientifique Agropolis-Montpellier, France

Simmonds NW (1954) Isolation in *Musa*, sections *Eumusa* and *Rhodochlamys*. Evolution 8:65–74

Simmonds NW (1960) Notes on banana taxonomy. Kew Bull 14:198–212

Simmonds NW (1962) The evolution of the bananas. Longman, London

Simmonds NW, Weatherup STC (1990) Numerical taxonomy of the wild bananas (*Musa*). New Phytol 115:567–571

Song KM, Osborn TC, Williams PH (1988a) Brassica taxonomy based on nuclear restriction framgent length polymorphisms (RFLPs). 1. Genome evolution of diploid and amphidiploid species. Theor Appl Genet 75:784–794

Song KM, Osborn TC, Williams PH (1988b) Brassica taxonomy based on nuclear restriction framgent length polymorphisms (RFLPs). 2. Preliminary analysis of subspecies within *B. rapa* (syn. *campestris*) and *B. oleracea*. Theor Appl Genet 76:593–600

Song KM, Osborn TC, Williams PH (1990) Brassica taxonomy based on nuclear restriction framgent length polymorphisms (RFLPs). 3. Genome relationships in *Brassica* and related genera and the origin of *B. oleracea* and *B. rapa*. Theor Appl Genet 79:497–506

Swofford DL (1985) PAUP-Phylogenetic analysis using parsimony, version 4.2. Illinois Natural History Survey, Champaign, Ill.