Systematic Botany (2002), 27(1): pp. 55–73

Combining Data from DNA Sequences and Morphology for a Phylogeny of Moringaceae (Brassicales)

MARK E. OLSON

Missouri Botanical Garden, PO Box 299, St Louis, Missouri 63166-0299;
Current address: Instituto de Biología, Universidad Nacional Autónoma de México, Circuito Exterior s/n, Ciudad Universitaria, Copilco, Coyoacán A. P. 70-367, C. P. 04510, México, D. F. México

Communicating Editor: Thomas Lammers

ABSTRACT. The Old World dry tropical family Moringaceae is remarkable for the great diversity of habit and floral morphology found within its only genus, *Moringa*. To infer the phylogenetic relationships of all 13 species, parsimony analyses of morphological data and DNA sequences from a low-copy nuclear region (*PEPC*), a chloroplast region (*trnG*), and a tandemly-repeated nuclear region present in high copy number (ITS) were conducted of each data set separately and combined. Characters from studies of ontogeny substantially enhanced the resolution of the morphological data set. The Incongruence Length Difference test indicated the congruence of all data sets, as did Templeton tests comparing the single tree resulting from the combined analysis in the context of the individual data sets. This tree is presented as the preferred topology, in which the four bottle trees appear in a basal paraphyletic assemblage, with the three species of slender trees (including the economically important *M. oleifera*) forming a clade that is sister to a clade of the six species of tuberous shrubs and trees of northeast Africa. *Moringa* is currently divided into three sections, but because of the basal grade, it cannot be divided into useful monophyletic infrageneric taxa. The phylogeny-based informal terms ''bottle tree grade'', ''slender tree clade'', and ''tuberous clade'' are suggested as alternatives. Relationships within *Moringa* were found to be largely congruent with a previous study of wood anatomy.

With just 13 species, *Moringa* Adans. (the only genus of the family Moringaceae) is for its size one of the most phenotypically varied groups of angiosperms. Ranging from huge ''bottle trees'' to tiny tuberous shrubs, and spanning the range from radial to bilateral floral symmetry (Fig. 1), the small number of species in the genus makes it useful for investigating the diversification of plant form. One species, *M. oleifera* (species authors are listed in Appendix 1), is cultivated throughout the tropics as the source of nutritious leaf and fruit vegetables, high-quality seed oil, pharmacologically active compounds, and water clarification agents (Oliveiera et al. 1999; Ghasi et al. 2000; Kalogo and Verstraete 2000; Saleem and Meinwald 2000; Jahn 2001). The other twelve species all have local uses, but only *M. oleifera* has been the subject of applied research or breeding. To provide a framework for basic and applied studies, I use DNA sequence data from one chloroplast and two nuclear loci, in addition to morphological data, to construct a phylogenetic hypothesis for the Moringaceae.

The Moringaceae and sister family Caricaceae are part of the ''mustard-oil plants'' clade (Rodman et al. 1998; the Brassicales of APG 1998), along with such families as Brassicaceae, Capparaceae, and Tropaeolaceae. *Moringa* is found in the seasonally dry tropics of Africa, Asia, and Madagascar (map, Fig. 1K). The pachycaul species with massive, water-storing trunks and fleshy roots and actinomorphic flowers (Fig. 1A-C; see also Table 1 for life form categories) occur in Africa and Madagascar (one species is apparently extinct in the wild, Olson and Razafimandimbison 2000). The remainder of the family have bilaterally symmetrical flowers, including the three species of slender trees (Fig. 1D-F). This habit class, which includes the economically important *M. oleifera*, is characterized by a more conventional trunk and tough, fibrous roots, and is principally south and southwest Asian. The dry tropical habitats in Kenya, Ethiopia, and Somalia support the highest number of *Moringa* species, and the entirety of what I call the ''tuberous'' species (Fig. 1E-J). These species are small trees or shrubs of varying habits, but always with fleshy, water-storing roots. As is true of many taxa of the dry tropics, several species of this group are very poorly-known. *Moringa arborea* has been seen twice by scientists and is known only from a single remote canyon on the Kenya-Ethiopia border. *Moringa pygmaea* is known only with certainty from the type collection in northern Somalia, though a sterile collection from the northern coast of Somalia (listed in Appendix 1) may represent this species. Verdcourt (1985) cites a probable undescribed species on the Kenya-Somalia border; subsequent fieldwork has shown this to be *M. longituba* (Olson 712 in Appendix 1). Despite the great morphological diversity in the family, the monophyly of *Moringa* is supported by numerous distinctive synapomorphies such as gum ducts in the pith and monothecal, bisporangiate anthers.

Differing species groups within *Moringa* have been proposed depending on the characters studied by the author, including leaf and floral morphology (Engler 1902; Verdcourt 1985), palynology (Ferguson 1985), and wood anatomy (Olson and Carlquist 2001). Current sectional classification (Verdcourt 1985) is based on floral morphology (Table 1; Fig. 1) and divides the

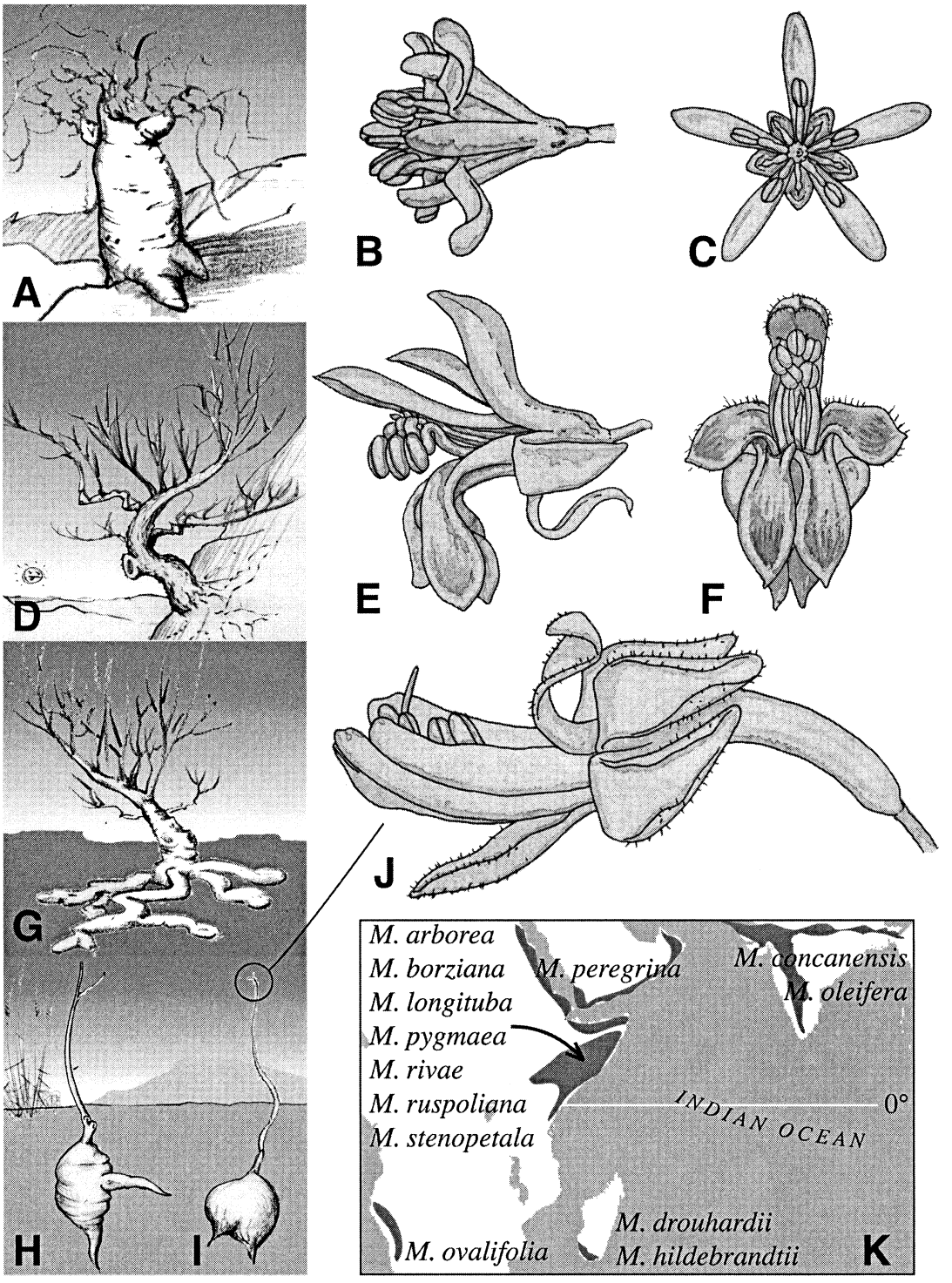

FIG. 1. Morphological diversity and range of *Moringa*. A. Bottle tree habit. B-C. Bottle tree flowers: radially symmetrical, all petals and sepals with an equal degree of flexion. D. Slender tree habit. E-F. Slender tree flowers: bilaterally symmetrical. G. Sarcorhizal trees (*M. arborea, M. ruspoliana*) have slender trunks and fleshy, brittle, tuberous roots; flowers similar to those in Figs. 1E, 1F. H. Tuberous shrubs of northeast Africa (*M. borziana, M. longituba, M. pygmaea, M. rivae*) have slender stems that are often shed in times of severe drought, with massive, very soft tubers underground. Flowers similar to Figs. 1E, 1F, with the exception of *M. longituba* (habit shown in Fig. 1I), which has bilaterally symmetrical flowers with a long tubular hypanthium (Fig. 1J). K. Range of *Moringa* showing India at upper right, part of Arabia at upper left, Madagascar at bottom center, and part of Africa at left.

TABLE 1. Species Groups in *Moringa*.

Species categorized by section (Verdcourt, 1985)	Species categorized by habit (Olson and Carlquist, 2001)	Species categorized by wood anatomy (Olson and Carlquist, 2001)	Groups shown in phylogenetic trees
Donaldsonia: flowers radially symmetrical or nearly so. *M. drouhardii, M. hildebrandtii, M. ovalifolia, M. stenopetala*	Bottle trees (Called bottle or tank trees because of their bloated trunks that are filled with water). Equivalent in species to Section Donaldsonia	Stem and root xylem have bands of confluent paratracheal parenchyma alternating with bands of libriform fibers, some of which may be wide and parenchyma-like (such fibers rare in *M. drouhardii*). Equivalent in species to Section Donaldsonia	Bottle trees: paraphyletic or polyphyletic
Moringa: flowers bilaterally symmetrical with short hypanthium. *M. arborea, M. borziana, M. concanensis, M. oleifera, M. peregrina, M. pygmaea, M. rivae, M. ruspoliana*	Slender trees: Trunks slender, with thick, tough bark and tough roots. *M. concanensis, M. oleifera, M. peregrina*	Stem xylem with libriform fibers of varying shapes and sizes, but not forming seasonal bands. Root xylem with alternating bands of libriform fibers and axial parenchyma; these parenchyma bands are not wider than bands of fibbers. Equivalent in species to the slender trees	Slender trees: monophyletic in all analyses
Dysmoringa: flowers bilaterally symmetrical with long hypanthium. *M. longituba*	Sarcorhizal trees: Slender trunks with a network of very thick, fleshy, brittle roots. *M. arborea, M. ruspoliana*	Stem xylem with narrow libriform fibers sometimes varying in size with season; sometimes with bands of paratracheal axial parenchyma. Root secondary xylem almost entirely axial parenchyma, with or without libriform fibers. *M. arborea, M. borziana, M. pygmaea, M. rivae*	"*rivae* clade": equivalent to this wood anatomy category
	Tuberous shrubs of NE Africa: Slender, often deciduous shoots, with huge tuberous roots. *M. borziana, M. longituba, M. pygmaea, M. rivae*	Stem wood with narrow libriform fibers sometimes varying in size with season without bands of paratracheal axial parenchyma. Root secondary xylem almost entirely axial parenchyma, with or without libriform fibers *M. longituba, M. ruspoliana*	"tuberous" clade: equivalent to these wood anatomy categories red-flowered clade: equivalent to this wood anatomy category

genus into three sections: *Donaldsonia* (radial symmetry), *Moringa* (bilateral symmetry with a short hypanthium), and *Dysmoringa* (bilateral symmetry with a long hypanthium). Species groups based on habit and wood anatomy are contrasted in Table 1. Verdcourt (1985) provides the only phylogenetic hypothesis for the family to date (Fig. 2), based on his study of herbarium specimens. His branching diagram lends itself to interpretation in cladistic terms, and can be read to hypothesize a monophyletic bottle tree clade (corresponding to Section *Donaldsonia*) as the sister group to the rest of the family, with a clade of slender trees forming the sister group to a clade of the tuberous species. Section *Moringa* (all of the species but *Donaldsonia* and *M. longituba*) is thus hypothesized to be paraphyletic because of the position of *M. longituba*, the sole member of Section *Dysmoringa*. How these various, sometimes conflicting groupings compare to a cladistic reconstruction based on data from various sources is examined here.

To construct a phylogenetic hypothesis of *Moringa*, three molecular loci were selected showing interspecific variation in *Moringa*: one low-copy nuclear locus (*PEPC*), one chloroplast locus (*trnG*), and a tandemly-repeated nuclear region present in high copy number (ITS). The enzyme phosphoenolpyruvate carboxylase (*PEPC*; Enzyme Comission code 4.1.1.31) is of major importance in CO_2 fixation in C_4 and CAM plants and is implicated in anaplerotic carbon metabolism in C3 plants (Latzko and Kelly 1983). The genes coding for this enzyme appear to be present in low copy numbers in small multigene families (e.g., Panstruga et al. 1995 found ~3 copies in *Solanum tuberosum*). Most plant PEPC genes that have been examined are characterized by nine introns; sequences of the ca. 450 bp-long fourth intron have been used in our laboratory for phylogeny reconstruction in Solanaceae, Chenopodiaceae, and Tamaricaceae. The tRNA gene for the amino acid glycine is located in the large single-copy region of the chloroplast genome. Hamilton et al. (1999) report population-level variation in the locus amplified by the primer pair *trnG-trnS* in *Corythophora alta* (Lecythidaceae). The tandemly-repeated internal transcribed spacer region (ITS) of the 18s–26s nuclear ribosomal DNA has been used to reconstruct phylogenies at the interspecific level in numerous groups of plants, e.g., Asteraceae (Baldwin 1992; Bayer et al. 1996), Loasaceae (Moody et al. 2001), and Rosaceae (Lee and Wen 2001).

In addition, a data set based on morphological characters was assembled to examine the effectiveness of characters drawn from the study of ontogeny. Examining immature ontogenetic stages can reveal additional phylogenetic characters not identifiable in mature organs (e.g., characteristics of cotyledons). Such characters can be considered ''instantaneous'' in that only one ontogenetic stage is required for their observation. A second category of characters are ontogenetic transformations (e.g., leaves lobed -> leaves entire, which require knowledge of more than one ontogenetic stage). Such characters are expected to improve phylogenetic resolution (de Queiroz 1985).

METHODS

Taxon Sampling for Molecular Data Sets. Taxa selected and gene regions sequenced are summarized in Appendix 1, along with voucher, locality, and species author information. At least one sample of each species of *Moringa* was available. In the case of wide-ranging species, multiple samples from distributional extremes were collected (e.g., *M. longituba*, *M. rivae*, and *M. peregrina*; a sample of ''wild type'' *M. oleifera* and a cultivar, PKM, were also included). Members of all four genera of Caricaceae (*Carica* L., *Cylicomorpha* Urban, *Jacaratia* A. DC., and *Jarilla* Rusby) were selected as an outgroup.

Tissue Collection and DNA Extraction. Most tissue was collected in the field or from cultivated specimens and immediately dried in silica gel. Voucher specimens for these collections are listed in Appendix 1. Leaves were usually used, but in some cases the plants were leafless and stem tissue was prepared by separating the bark from the xylem cylinder, removing the phellem, and drying the remaining bark layers in silica gel. Tissue was ground very finely in a mortar with a small amount of sterilized silica sand. DNA was extracted from ground tissue using the protocol of Edwards et al. (1994), followed by two 700μl 24:1 chloroform: isoamyl alcohol extractions (suggestion of Scott Hodges, pers. comm.). Material was available from the two known specimens of *Moringa pygmaea*, but extraction using the Edwards et al. (1994) protocol, a CTAB protocol (Doyle and Doyle 1987), and Qiagen DNeasy Plant Mini Kit recovered only degraded DNA.

PCR Amplification and Sequencing. The fourth intron of *PEPC* was amplified using the primers PPCX4F and PPCX5R. PPCX4F (sequence 5′ ACTCCACAGGATGAGATGAG) binds to the 4th exon and promotes extension across 4th intron; PPCX5R (sequence: 5′ GCGCCATCATTCTAGCCAA) binds to the 5th exon and promotes extension back across 4th intron. These primers were designed and kindly provided by John Gaskin. The *trnG* and ITS regions were amplified with the primers of Hamilton 1999 (*trnG*) and Bayer et al. 1996 (ITS−1,−2,−3,−4). The PCR thermal cycling profile consisted of a 90 second denaturation at 94°C followed by 30 cycles of 94°C for 50 seconds, 55°C for 70 seconds, and 72°C for 90 seconds. After these cycles, the samples were subjected to a final extension at 72°C for 3 minutes, and 30°C for 1 minute. Each reaction contained a final concentration of 2.5 mM $MgCl_2$, 10 mM Tris HCl (pH 9.0), 50 mM KCl, 0.2 mM of each dNTP, 0.2 mM of each primer and 0.5 U/μl taq polymerase. Each PCR consisted of five separate 22.5 μl reactions which were combined for purification. PCR products were separated on agarose gels, purified with a Qiaquick Gel Extraction Kit (Qiagen), and quantified using GibCo Low DNA Mass Ladder. Most sequencing reactions used Applied Biosystems Incorporated Big Dye terminators and were run on an Applied Biosystems Incorporated Model 373 or Model 377 Prism DNA Automated Sequencer. GenBank accession numbers are listed for all sequences in Appendix 1. Sequences were aligned by eye using the Se-Al Sequence Alignment Editor v1.0 alpha 1 (Rambaut 1996–1998). Alignments are available from the TreeBASE database.

Morphological Characters. A data matrix of 28 morphological characters was constructed from herbarium, pickled, and living material of all 13 species of *Moringa*. *Cylicomorpha* was selected as an outgroup based on morphological and anatomical similarity to *Moringa* (Carlquist 1998; Olson and Carlquist 2001), and Badillo's (1971) comments suggesting that, of all Caricaceae, *Cylicomorpha* bears the largest proportion of plesiomorphic character states. The characters and their coding are given in Appendix 2; vouchers are listed in Appendix 1. For microscopy, most samples were collected from living plants and preserved in 50–70% aqueous ethanol. Methods used to study wood and roots are detailed in Olson and

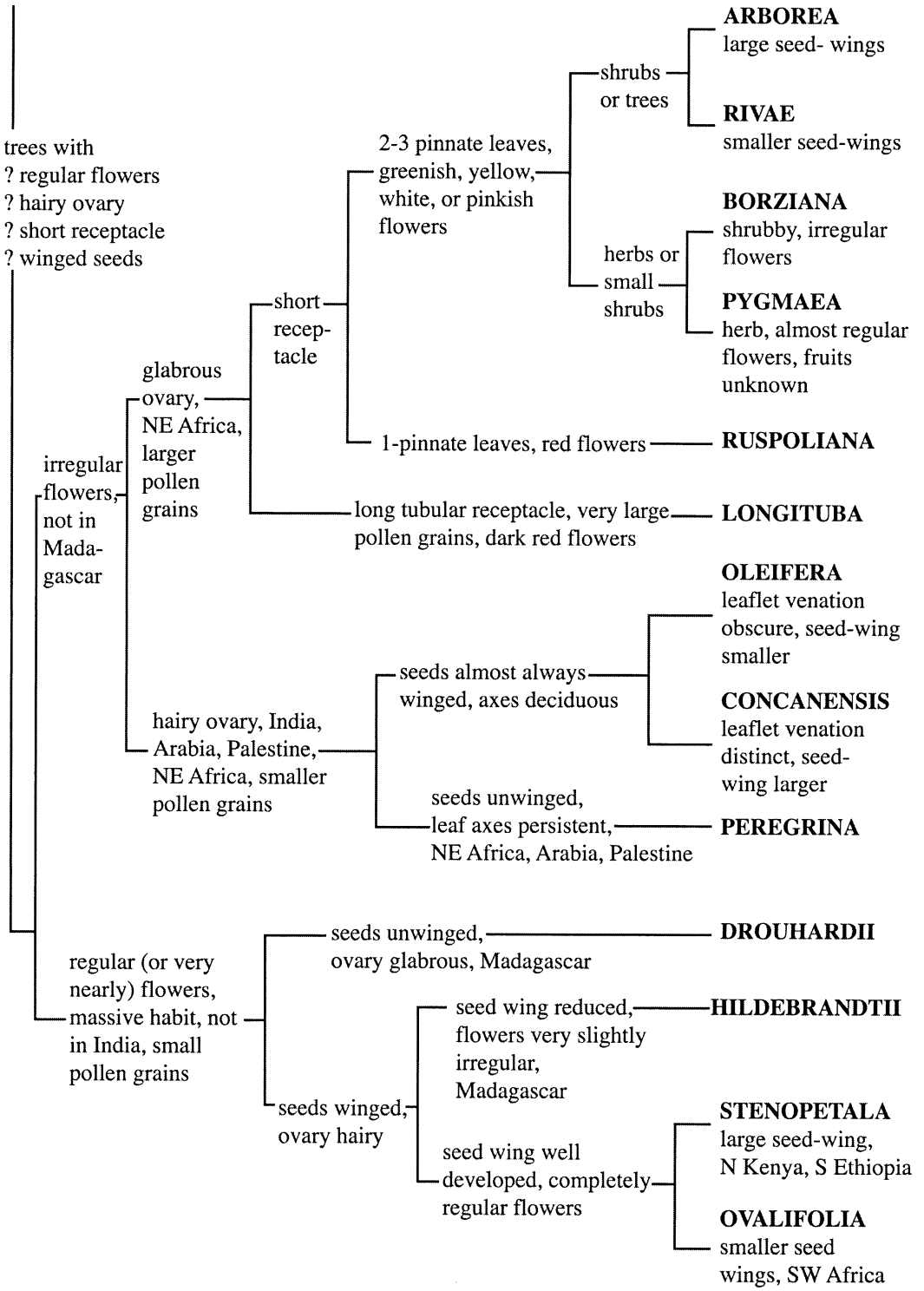

FIG. 2. Phylogeny based on Verdcourt's (1985) "possible phylogeny based on the guess that the original ancestral species were trees with regular hypogynous flowers having hairy ovaries and winged seeds", constructed from similarities observed between herbarium specimens. Verdcourt envisioned three major groups (top to bottom): tuberous trees and shrubs of northeast Africa (defined by the node marked "glabrous ovary, NE Africa, larger pollen grains"), that is sister to a slender tree clade, and a bottle tree clade that is sister to the rest of the family.

TABLE 2. ILD Test pairwise comparisons.

Pairwise comparison	p
PEPC + *trnG*	0.428
PEPC + ITS	0.108
PEPC + morphology	1
trnG + ITS	0.084
trnG + morphology	0.331
ITS + morphology	0.016

Carlquist 2001. For sectioning, leaves and flowers were passed through a dehydration series from 70% aqueous ethanol to 95%, through three changes of absolute ethanol, ending with three changes of tertiary butyl alcohol, with the sample being allowed to remain in each solution at least overnight. Samples were embedded in paraffin, sectioned on a rotary microtome at 13 μm and stained in a series corresponding to Northen's modification of Foster's ferric chloride-tannic acid staining series (Johansen 1940), with ferric aluminum sulfate substituted for ferric chloride. For scanning electron microscope (SEM) observations, leaf and floral dissections were dehydrated to absolute ethanol, critical-point dried, and mounted on aluminum stubs. The samples were sputter-coated on a Polaron E-5000 and observed with a Hitachi S-450 SEM at 20 kilovolts.

Analysis. Phylogenetic analyses employed PAUP* 4.0b8 (Swofford 2001). Searches were heuristic with the parsimony optimality criterion, unweighted and unordered characters, starting trees found via 1000 random additions, TBR branch swapping, the COLLAPSE and STEEPEST DESCENT options off, MULTREES and ACCTRAN options in effect. Gaps were interpreted as missing data and ignored. Bootstrap values were derived from 1000 replicates of a full heuristic search. AutoDecay (Eriksson 1998) was used to calculate decay indices (Bremer support).

"Total evidence" Analysis. A combined data set was constructed to include all three molecular data sets and the morphological data for each species. Three species are represented by more than one sample: *M. oleifera* (*PEPC, trnG* = Stanley, s.n.; ITS = Olson, s.n.) *M. ovalifolia* (*PEPC* and *trnG* = Olson 718, ITS = Olson, s.n.); *M. rivae* (*PEPC* and *trnG* = Olson 677, ITS = Olson 701). The other species were represented by the following samples: *Cylicomorpha parviflora* C. Kayombo 1296; *M. drouhardii* Olson 679; *M. hildebrandtii* Olson "2"; *M. peregrina* Danin, s.n.; *M. borziana* Olson 678; *M. longituba* Olson 708; *M. ruspoliana* Olson 702. While not ideal, it is sometimes justifiable to allow different individuals to contribute sequence data from different regions, especially if the species in question represent monophyletic lineages. *Moringa hildebrandtii* was absent from the *trnG* data set and these characters were coded as missing for this species.

Congruence testing. There is ample evidence that different topologies may be recovered from phylogenetic data sets of different origin (e.g., chloroplast and nuclear regions, Dumolin-Lapègue, Kremer, and Petit 1999; morphological and molecular data sets, Larson 1994). The Incongruence Length Difference test (ILD; Farris et al. 1994; the Partition Homogeneity Test option of PAUP*) and the Templeton test (Templeton 1983; Larson 1994; Mason-Gamer and Kellogg 1996; Johnson and Soltis 1998) were used to assess the level of congruence between the data sets.

ILD TEST. For this test, each possible pair of individual data sets are combined to form a single data set. Thus the four data sets used in this study were combined to form six data sets composed of two partitions each. These pairs are listed in Table 2. One thousand Partition Homogeneity Test replicates were run, using the same heuristic search settings as those of the parsimony analyses of the individual data sets. The null hypothesis that the sum of the treelengths from the random partitions should not be statistically significantly different from the sum of the treelengths produced by the original partitions are rejected in cases that displayed a P value <0.01 (as suggested by Johnson and Soltis 1998).

TEMPLETON TEST. The Templeton test compares the partitioning of a data set (the "test data" ; terminology of Mason-Gamer and Kellogg 1996) onto two trees, one resulting from an analysis of that data set (the "test tree") and a tree resulting from an analysis of another data set (the "rival tree"). The null hypothesis for this test is that both the test tree and the rival tree represent statistically equivalent, if not equally parsimonious, explanations of a given data set, as assessed by an application of a Wilcoxon signed-ranks test (Templeton 1983; Felsenstein 1985; Larson 1994).

The "total-evidence" data set was used for Templeton tests. *Moringa pygmaea* was absent from these comparisons because it was present only in the morphology data set. Similarly, Caricaceae sequences were not included in these tests because they were not used in the *PEPC* and *trnG* analyses. Topologies resulting from individual data set analyses were compared to each other, and the topology resulting from the total-evidence analysis was used as a rival tree with respect to the individual data sets. Polytomies in rival trees were resolved by constraining a parsimony analysis of the test data set with the rival tree to find the resolution(s) of polytomies most compatible with the test data. This approach avoids inflating the rival tree length with polytomies, which are here assumed to represent a lack of resolving power of the data set rather than multiple simultaneous divergences (choice of rival trees is discussed by Mason-Gamer and Kellogg 1996; Cunningham 1997; Graham et al. 1998). The two-tailed critical values calculated using PAUP* 4.0b8 (Swofford 2000) were halved to compare them to the one-tailed values in Table 30 of Rohlf and Sokal (1981). In addition, the null hypothesis was not rejected in comparisons had four characters or less that differed in length on the two trees (i.e., in cases where $N<5$).

RESULTS

Abbreviations and Figure Notes. Bootstrap values (generally those greater than 50%) and decay indices are shown in the Figures below branches, and branch lengths are shown above. The following abbreviations are used in the text and figure legends; TL = tree length in number of steps; CI = consistency index (Kluge and Farris 1969); RI = retention index (Farris 1989); RC = rescaled consistency index (Farris 1989). The species groups referred to as the bottle trees (*M. drouhardii*, *M. hildebrandtii*, *M. ovalifolia*, *M. stenopetala*), slender trees (*M. concanensis*, *M. oleifera*, *M. peregrina*), and the tuberous shrubs and trees of northeast Africa (the "tuberous clade": *M. arborea*, *M. borziana*, *M. longituba*, *M. pygmaea*, *M. rivae*, *M. ruspoliana*) are labeled on all trees. Within this latter group, the "*rivae group*" (consisting of *M. rivae* and the three morphologically similar species *M. arborea*, *M. borziana*, and *M. pygmaea*) and Section Dys*moringa* (consisting only of *M. longituba*) are also identified on the total-evidence tree (Fig. 6). Indel typology follows Golenberg et al. (1993) as modified by Hoot and Douglas (1998), where Type 1a indels are simple repeats or deletions of the same nucleotide, Type 1b indels are repeated motifs of two or more bases that include more than one nucleotide, and all other indels are referred to the Type II category.

PEPC. The sequence used corresponds approximately to bases 799–1297 of the *Arabidopsis thaliana* phosphoenolpyruvate carboxylase sequence (GenBank accession AF071788, Paterson, K. M. and H. G. Nimmo, unpubl.). Bases before position 798 and past position 1297 were excluded from phylogenetic analyses because they were missing for some species. Because of

alignment ambiguity between the ingroup and outgroup sequences, the outgroup sequences were excluded and the trees were rooted with *Moringa drouhardii*. This species is strongly supported as the sister taxon to the rest of the family in the ITS and morphological analyses, and shares many morphological features with the Caricaceae (Olson and Carlquist 2001). Three percent of the cells were coded as missing, mostly due to missing sequence for *M. hildebrandtii* Olson ''W'' and *M. peregrina* Danin s.n. Four Type 1a indels of 1bp each also contribute to the cells coded as missing (alignment deposited as TreeBASE M1027). Of the 494 bp used in the analysis, 39 were phylogenetically informative (8% of the total characters). Two most-parsimonious trees were recovered of 69 steps (CI = 0.95; RI = 0.97; RC = 0.92). The strict consensus of these trees is shown in Fig. 3A. The trees differed in their arrangement within the *rivae* group (a *M. arborea—M. rivae* clade sister to *M. borziana* vs. *M. arborea* as the sister group to a *M. borziana—M. rivae* clade).

trnG. The first ca. 144 bases align to the tRNA (Gly) gene of *Sinapis alba* (Liere and Link 1994) before extending into the adjacent noncoding spacer. Beyond approximately the 350th position, the sequences were characterized by alternating poly-A and poly-T tracts and a region of more than 100bp in length that distinguished the *rivae* group species but was unalignable to the rest of the family or the outgroup. These and subsequent bases were excluded from the analysis. As in the *PEPC* analysis, *M. drouhardii* was used to root the tree. Of the 4.5% of cells coded as missing, most of these are accounted for by three Type 1a indels and three Type 1b indels 1–7 bp in length. One region of Type 1a indels varied from 6–16 A residues and was bounded by Gs on each side. This area was aligned such that the Gs aligned and gaps were introduced to create contiguous A blocks adjacent to the 5′ G (alignment deposited as TreeBASE M1026). Of the 335 characters used in the analysis, 20 were phylogenetically informative (6% of the total characters). Four most-parsimonious trees were recovered of 30 steps (CI = 0.87; RI = 0.96; RC = 0.83). The strict consensus of these trees is shown in Fig. 3B. The trees differed in their arrangements within the slender tree clade and *rivae* group species + *M. ruspoliana*.

ITS. Several tracts of outgroup sequences were excluded from the analysis because they were so diverged as to preclude unambiguous alignment (at positions 68–170; 243–259; 450–469; 478–490; 517–530; 603–669). Of the cells in this alignment, 15.63% were scored as missing, partly because of missing sequence within the 5.8s gene for five species. Indels were more common in the spacer regions than in the coding 5.8s gene: the ITS-1 spacer had eight indels of Type 1a, and two Type 1b indels; the ITS-2 spacer had seven Type 1a indels and two Type 1b indels; the 5.8s gene had just one indel, of Type 1a (alignment deposited as TreeBASE M1028). Of the 705 characters used in the analysis, 228 were phylogenetically informative (32% of the total characters). Eighteen most-parsimonious trees were recovered of 431 steps (CI = 0.81; RI = 0.90; RC = 0.73). The strict consensus of these trees is shown in Fig. 4. Most of the differences between the most-parsimonious trees were different arrangements of the *rivae* group species and the multiple samples of *M. longituba*. The bottle tree *M. drouhardii* is the sister taxon to the rest of the family, making the bottle trees poly-, or perhaps para-, phyletic. The slender tree and tuberous clades are recovered with high support from bootstrap and decay indices.

Morphological Characters. The 28 characters in this data set and their coding are enumerated in Appendix 2. In this data matrix, 2.04% of the cells were scored as missing. Ten most-parsimonious trees were recovered of 54 steps (CI = 0.61; RI = 0.73; RC = 0.45). These trees differ in their arrangments of species within the slender trees and the tuberous clade. The resolution recovered from the bootstrap analysis is shown in Fig. 5. The bottle trees occur in this tree in a paraphyletic grade at the base of the tree, with the slender tree and tuberous clades sister to each other with moderate support. Nonmolecular characters provide support at all levels of relationship, e.g., *M. arborea* and *M. rivae* are paired by the presence of crystalliferous tyloses, a unique situation in the family (character 7; see lower tree in Fig. 5). Likewise, *M. ruspoliana* and *M. longituba* are united by entirely lacking paratracheal axial parenchyma in their shoots (character 6).

To examine the effectiveness of characters based solely on mature morphologies, the nine characters that derive from ontogenetic studies (marked by an asterisk in Appendix 2) were removed. The 4 most-parsimonious trees from this analysis (40 steps; CI = 0.75; RI = 0.86; RC = 0.65) showed less resolution, particularly at the base of the tree. This tree is also shown in Fig. 5.

Comparison of Individual Data Set Analyses. The groupings based on life form classes show differing patterns of phylogenetic status: 1) The slender trees form a clade in all analyses with strong support except for the *trnG* analysis. Although the monophyly of the slender trees is clear, no pattern of relationship among the three species that form the clade emerges in these analyses. 2) The tuberous clade appears strongly supported in all analyses. In the *PEPC* analysis, two major divisions appear within the tuberous clade: the *rivae* group, and a clade consisting of the two red-flowered species *M. longituba*, and *M. ruspoliana*. In the ITS and *trnG* analyses, *M. ruspoliana* pairs with the *rivae* group. In all analyses, little resolution was found within the *rivae* group. 3) The bottle trees are para- or poly- phyletic in all analyses, with some members of this class

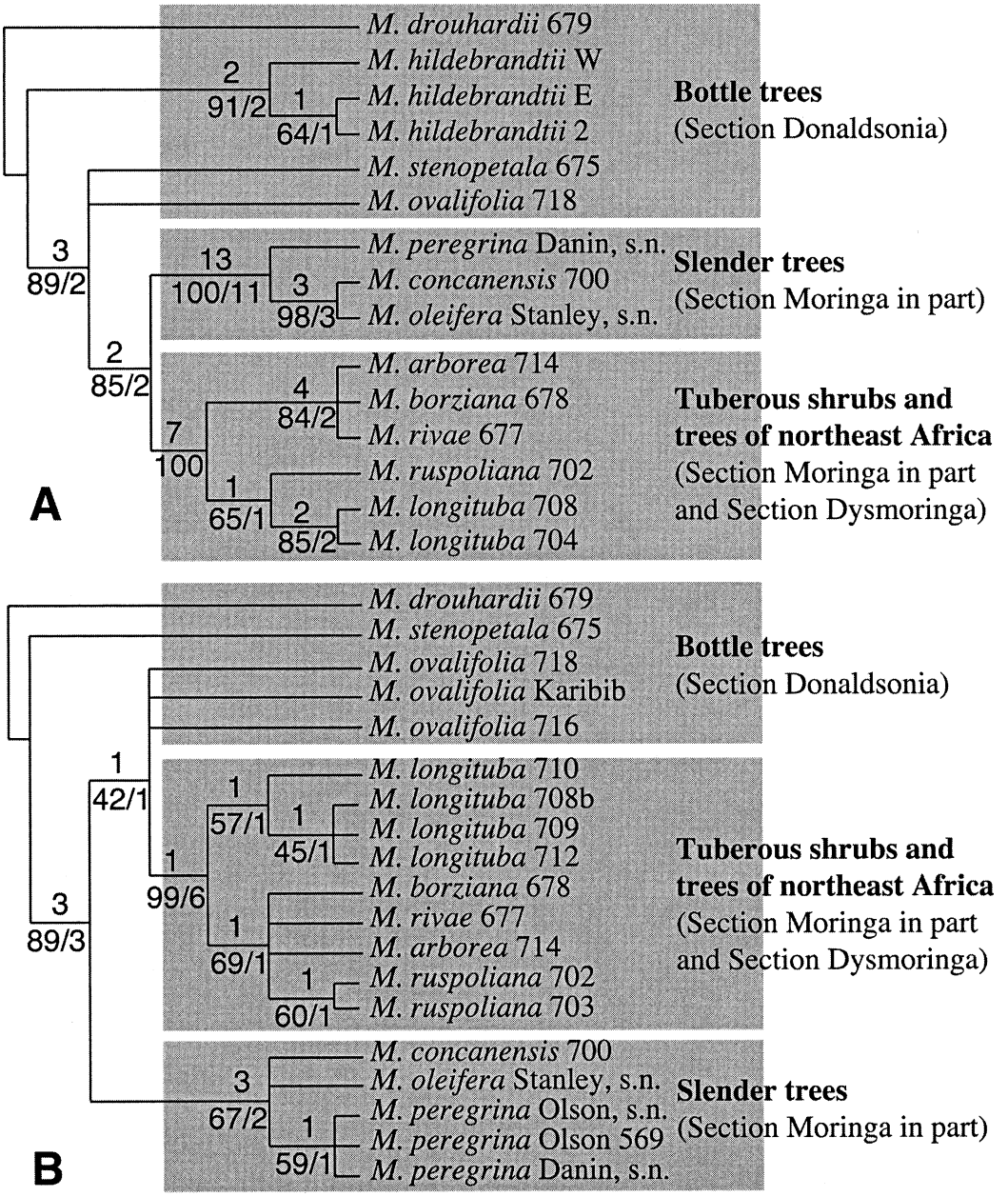

FIG. 3A-B. *PEPC* and *trnG* analyses. A. Strict consensus of the 2 most-parsimonious trees recovered in the analysis of PEPC sequence data (TL 69, CI 0.95, RI 0.97, RC 0.92). B. Strict consensus of the 4 most-parsimonious trees recovered in the analysis of trnG data (TL 30, CI 0.87, RI 0.96, RC 0.83).

grouping weakly with the tuberous clade. In the ITS and morphological analyses, *M. drouhardii* is well-supported as the sister taxon to the rest of the family.

Total Evidence Analysis. Of the 1563 characters used in the analysis, 155 were phylogenetically informative (11% of the total characters). In the analysis of the ITS data set described above, 32% of the characters were phylogenetically informative. However, of the characters in the ITS partition of the total-evidence data set, the proportion of phylogenetically-informative characters was reduced to 11%. This reduction in variation is due to the removal of multiple samples of species such as *M. longituba*, of which six were included in the ITS analysis compared to the single sample included in the total evidence analysis. The percentages of phylogenetically informative characters contributed by the PEPC, trnG, and morphological data sets were comparable to those found in the individual

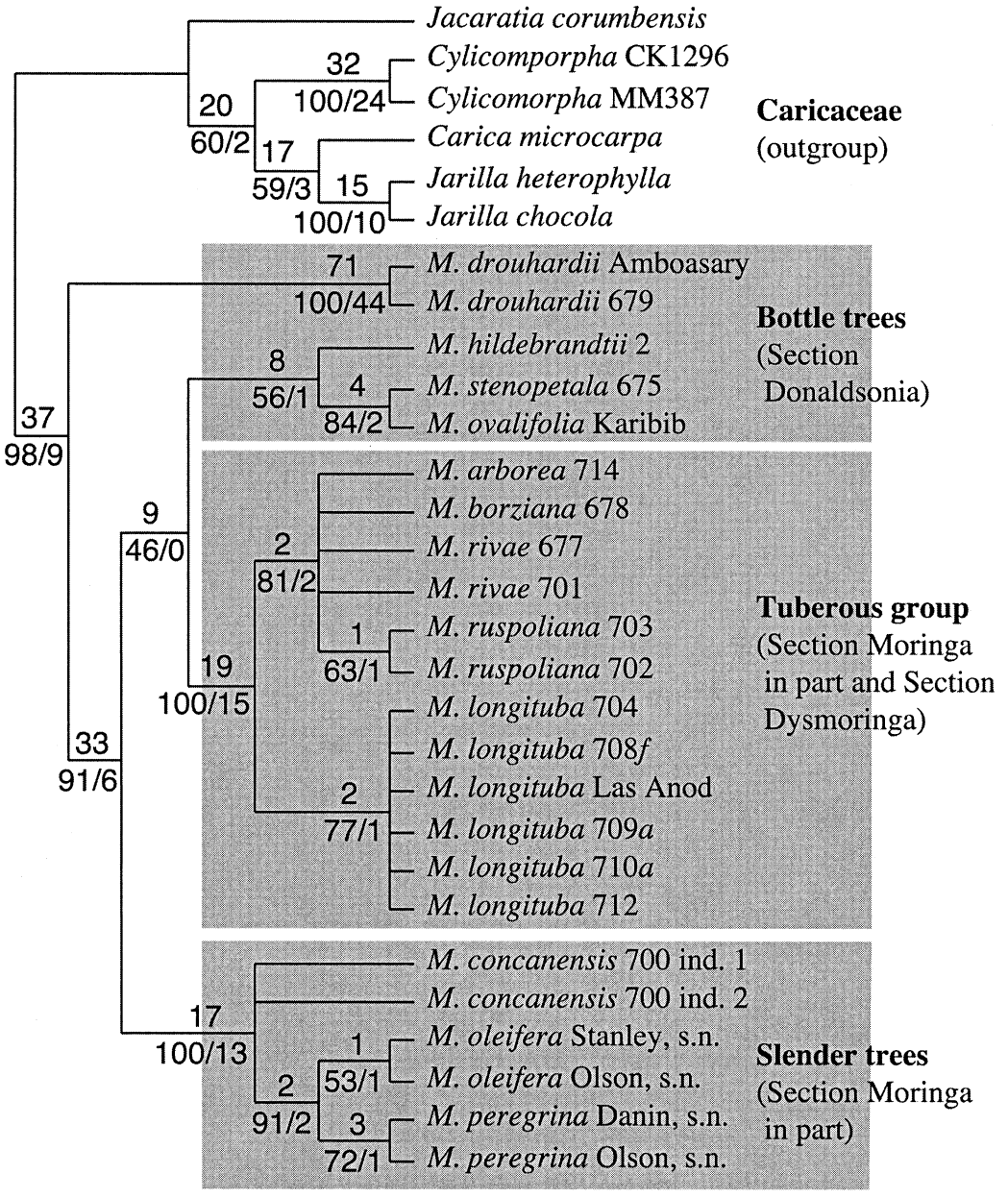

FIG. 4. Strict consensus of the 18 most-parsimonious trees recovered in the analysis of ITS data (TL 431, CI 0.81, RI 0.90, RC 0.73).

data set analyses. A single most-parsimonious tree of 432 steps was recovered (CI = 0.86; RI = 0.85; RC = 0.73; this tree is shown in Fig. 6). The relationship between the major groups in *Moringa* is resolved with better support in the total evidence analysis relative to the individual analyses. The bottle trees are paraphyletic in this analysis, with robustly-supported slender tree and tuberous clades. *Moringa ruspoliana* pairs with 81% bootstrap and a decay index of 3 with the other red-flowered species, *M. longituba*.

Congruence Testing. ILD TEST. Pairwise comparisons indicate a generally high degree of congruence between the data sets, with $P > 0.01$ in all cases; all P values were greater than 0.05 except for the ITS + morphology comparison. Test statistics are summarized in Tab 2.

TEMPLETON TEST. In just less than half (46%) of the comparisons between trees derived from analyses of the individual data sets, the null hypothesis was rejected, indicating that the test trees had a signifi-

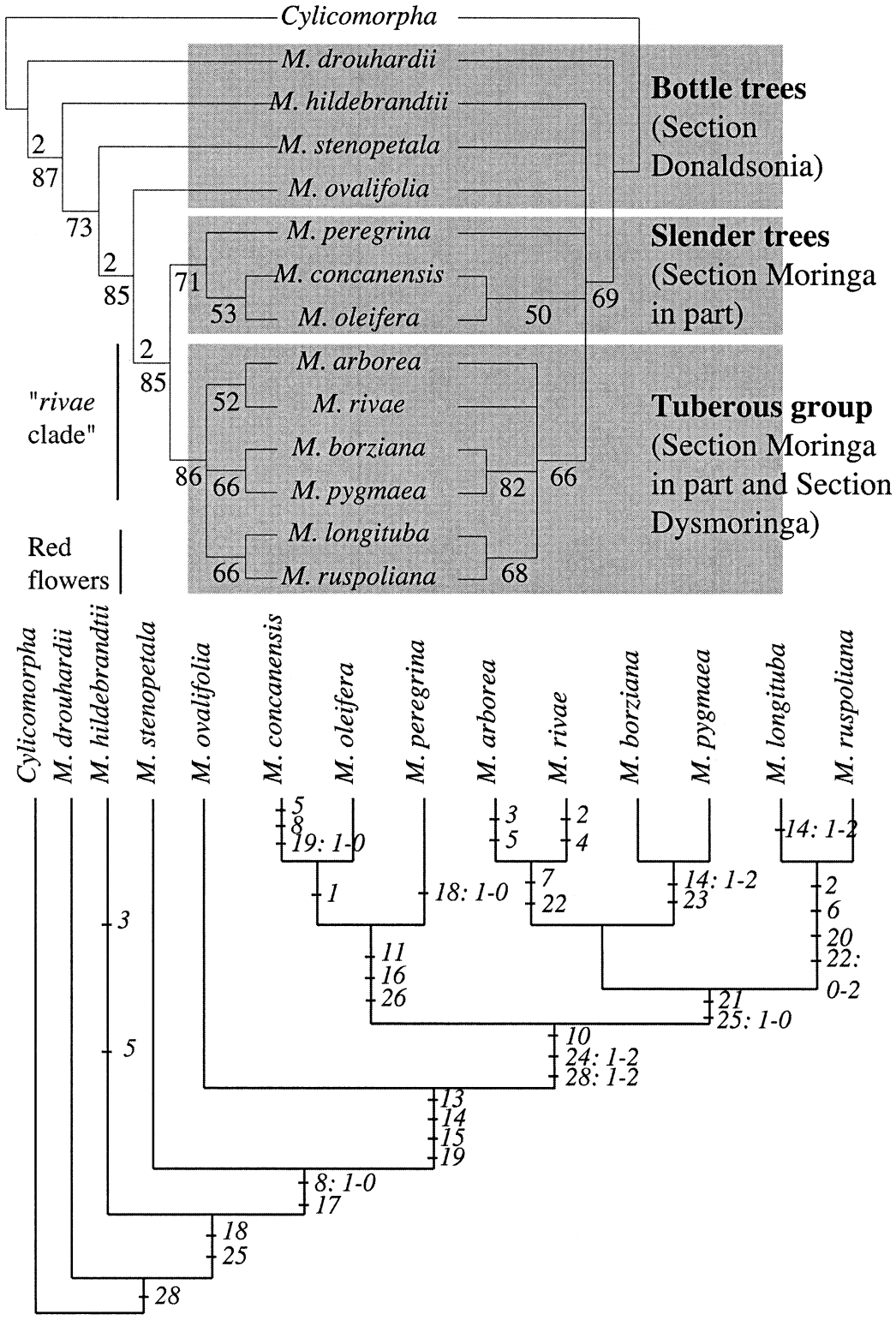

FIG. 5. Top left tree: Consensus tree from bootstrap analysis of morphological data (TL 54, CI 0.61, RI 0.73, RC 0.45). Top right tree: Strict consensus of the 4 most-parsimonious trees recovered in the analysis of the morphological data with the 9 ontogenetic characters removed (TL 40, CI 0.75, RI 0.86, RC 0.65). Lower tree: same as top left tree with unambiguous changes indicated on branches. The numbers correspond to the characters as numbered in App. 2. Unless indicated, change is to the derived state (0 to 1).

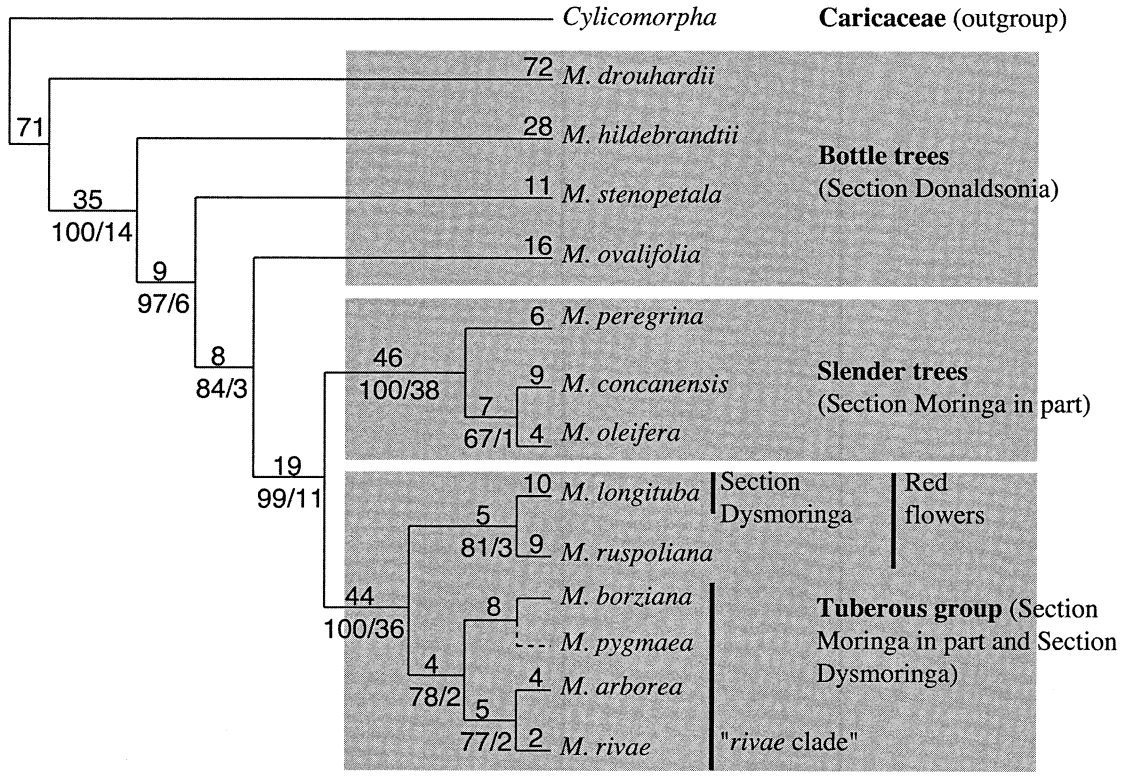

FIG. 6. Single most-parsimonious tree recovered in the ''total evidence'' analysis of all molecular, morphological, anatomical, and ontogenetic data, with *M. hildebrandtii trnG* characters (not sequenced for this species) coded as missing. (TL 432, CI 0.86, RI 0.85, RC 0.73). The position of *M. pygmaea* is shown in dashed lines based on the morphological analysis.

cantly more parsimonious partitioning of the test data than did the rival trees. For example, the *PEPC* test data set was mutually congruent with both the *trnG* and morphology data sets, but the null hypothesis was rejected in all comparisons with ITS rival trees. In the reciprocal test, the *PEPC* rival trees were compatible with the ITS test data set. The results for all comparisons, with number of characters differing in length between the two topologies and critical values, are given in Table 3. In contrast to the variable results of Templeton tests involving trees from individual data set analyses, none of the comparisons of the total-evidence rival tree in the context of individual data sets resulted in a rejection of the null hypothesis. In other words, the tree derived from a pooling of all of the data represents a topology that is compatible with each individual data set.

Population Sampling. Nine of the 13 described *Moringa* species were represented by more than one sample in at least one data set (data set and support for the node subtending all the samples of the species >50% indicated in parentheses): *M. concanensis* (ITS), *M. drouhardii* (ITS, 100%, 44 decay), *M. hildebrandtii* (*PEPC*, 91%, 2 decay), *M. longituba*, (*PEPC*, 85%, 2 decay; *trnG*, 57%, 1 decay; ITS, 77%, 1 decay), *M. oleifera* (ITS, 53%, 1 decay), *M. ovalifolia* (*trnG*), *M. peregrina* (*trnG*, 59%, 1 decay; ITS, 72%, 1 decay), *M. rivae* (ITS), and *M. ruspoliana* (*trnG*, 60%, 1 decay; ITS, 63%, 1 decay). None of the analyses rejected the monophyly of any of these species, though support was particularly weak for the pairings of the multiple samples of *M. concanensis*, *M. oleifera*, *M. ovalifolia* and *M. rivae*.

DISCUSSION

Congruence Testing and a Phylogeny of Moringaceae. Congruence of the data sets is indicated by the ILD test and that none of the individual data sets was able to reject the total evidence topology in the Templeton test comparisons. I assume that this is because the total-evidence rival topology faithfully represents the phylogenetic signal present in each of the data sets. Therefore, I choose the total evidence tree (Fig. 6) as the best skeleton for a phylogenetic hypothesis of the Moringaceae. *Moringa pygmaea* was included only in the morphological analysis where it appears as the sister species to *M. borziana*. The best estimate of its position in the *Moringa* phylogeny is therefore sister to *M. borziana*, and it is shown in this position with dashed lines in Fig. 6.

TABLE 3. Templeton test values. N = number of characters of different lengths between test tree and rival tree.

Test data and tree	Rival tree	N	P	Test data and tree	Rival tree	N	P
PPC1	trnG 1	5	0.1797	morph 4	trnG 1	10	0.1655
PPC1	trnG 2	3	0.0833	morph 5, 7–8	trnG 1	10	0.0578*
PPC2	trnG 1	3	0.0833	morph 6	trnG 1	11	0.2362
PPC2	trnG 2	5	0.1797	morph 9	trnG 1	12	0.1967
PPC1	ITS 1	7	0.0082*	morph 10	trnG 1	13	0.2597
PPC1	ITS 2	9	0.0196*	morph 1–3	trnG 2	10	0.0578*
PPC2	ITS 1	9	0.0196*	morph 4	trnG 2	12	0.1967
PPC2	ITS 2	7	0.0082*	morph 5, 7–8	trnG 2	8	0.0339*
PPC 1–2	morph 1–2	2	1	morph 6	trnG 2	13	0.2597
trnG 1–4	PPC	2	0.1573	morph 9	trnG 2	10	0.1655
trnG 1–4	ITS 1–2	4	0.0455*	morph 10	trnG 2	11	0.2362
trnG 1–4	morph 1–4	2	0.1573	morph 1–3	ITS	11	0.0018*
ITS 1	PPC	5	0.1797	morph 4	ITS	16	0.0221*
ITS 2	PPC	11	0.1317	morph 5, 7–8	ITS	13	0.0030*
ITS 1	trnG	9	0.0209*	morph 6, 9	ITS	15	0.0144*
ITS 2	trnG	11	0.0075*	morph 10	ITS	17	0.0184*
ITS 1	morph 1–3	7	0.0588*	trnG 1–4	total	2	0.1573
ITS 2	morph 1–3	13	0.0522*	ITS 1–2	total	11	0.1317
morph 1–3	PPC	2	1	morph 6	total	7	0.7055
morph 4	PPC	3	0.5637	morph 1–3	total	2	1
morph 5	PPC	7	0.7055	morph 4, 10	total	5	0.6547
morph 6–7	PPC	5	0.6547	morph 5, 7–8	total		same topology
morph 8–10	PPC		same topology	morph 9	total	3	0.5637
morph 1–3	trnG 1	8	0.0339*				

Ontogenetic Characters. Ontogenetic studies provided characters that greatly improved the resolution of the morphological analysis (cf. trees resulting from analyses with and without ontogenetic characters in Fig. 5). Many of these are simply characters drawn from immature stages, e.g., germination type (character 15). The three basal species have cotyledons that emerge from the seed, whereas in the rest of the family the cotyledons remain within the seed coat. Character 14, seasonal persistence of the shoot at different ontogenetic stages, can be considered a non-instantaneous character in that it is necessary to observe more than one ontogenetic stage to determine the state present in each species. Another such character, the pattern of anther orientation in ontogeny (character 24), unites the slender trees and tuberous clades. Examination of this character substantially altered homology interpretations that were based only on mature flowers. The radially symmetrical flowers of the pachycaul species seemed clearly to be symplesiomorphic with the actinomorphic flowers of Caricaceae. However, examination of their ontogeny showed bilateral symmetry to be a property of all species of *Moringa* (Olson, unpubl. data; see also Appendix 2).

Phylogeny and the Distribution of Moringa. Two monophyletic groups in *Moringa* show clear geographical associations (highlighted on the total evidence tree in Fig. 7). Within these groups, there are pairs of sister species that share morphological similarities and appear to have largely allopatric geographical distributions. One of these clades, the slender trees, is nearly restricted to Asia, and its species occur in a broken band from Arabia to Bangladesh. *Moringa peregrina* occurs from the Dead Sea to southern Arabia and northern Somalia. The pair *M. concanensis* and *M. oleifera* are very similar in habit, leaves, and flowers but differ most conspicuously in bark morphology. *Moringa concanensis* is widely distributed from Pakistan to Bangladesh and along the length of peninsular India, but *M. oleifera* is apparently native to dry lowlands of northern India and has never been recorded co-occurring with *M. concanensis* (though recent documentation that *M. oleifera* still occurs in the wild is scant).

The tuberous group forms a monophyletic clade with three species pairs, all of which are restricted to the Horn of Africa. The members of the red-flowered species pair, *M. longituba* and *M. ruspoliana* co-occur in the same general region only in extreme northeast Kenya and central northern Somalia. Otherwise, their distributions are exclusive, with *M. longituba* having a more southerly range extending nearly 300 km farther south into Kenya, from Moyale to Wajir, and reaching northern Somalia via an arc roughly south of the Somalia-Ethiopia border. In contrast, the northern and southern extremes of the distribution of *M. ruspoliana*, where it co-occurs with *M. longituba*, are connected by an arc extending through southeastern Ethiopia along the line roughly delimited by the northwestern edge of the hot Ogaden lowlands. A member of another species pair, *M. borziana*, is well-documented from southeastern Kenya and southern Somalia, always from within 200 km of the coast. Its sister species, *M. pyg-*

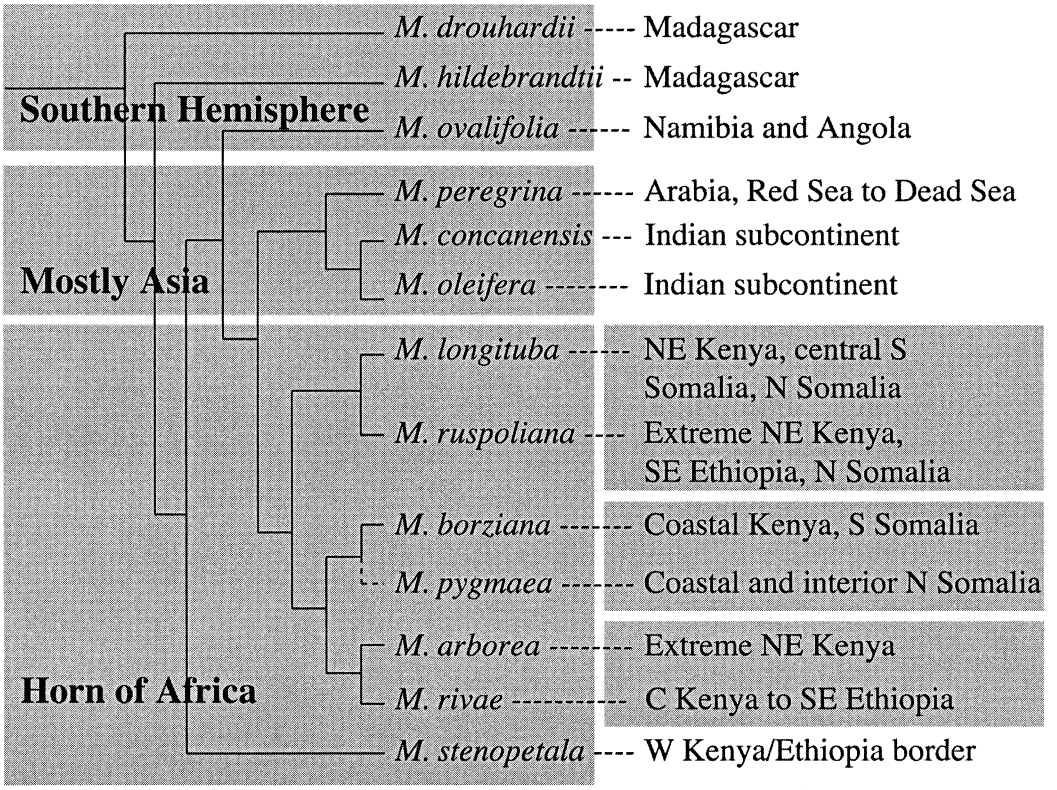

FIG. 7. Geographic ranges of *Moringa* species and phylogeny. Broad regions are shaded in gray and labeled at left. More specific ranges are cited at right. The small gray boxes at right highlight largely allopatric species pairs within the tuberous clade.

maea, may replace *M. borziana* in similar situations in northern Somalia, but exploration of this area is still too limited for definitive conclusions. Likewise, very little is known about the geographical distributions of both members of the species pair *M. rivae—M. arborea*, but the species do appear to have differing habitat preferences, with *M. arborea* exploiting the dynamic habitat of canyon bottoms and *M. rivae* preferring drier, more exposed sites.

In addition to forming a non-monophyletic group, the bottle trees show a less coherent pattern of distribution than the slender tree and tuberous clades. Three of the four species making up the basal grade in the family have austral distributions with *M. drouhardii* and *M. hildebrandtii* being restricted to Madagascar and *M. ovalifolia* reaching from central southern Namibia to southwestern Angola. *Moringa stenopetala* is found well to the north in the western Horn of Africa just to the west of the area occupied by the tuberous group.

Previous Classifications of Moringa. That a genus of just 13 species should be divided into infrageneric taxa is a reflection of its remarkable morphological diversity. In 1902, Engler placed seven of the eight species then known into Section (*Eu*)*Moringa*, and created Section *Dysmoringa* to emphasize the unique flowers of *M. longituba* (Fig. 1J). Verdcourt (1958, 1985) noted the morphological cohesiveness of the four phylogenetically basal species and transferred them to Section *Donaldsonia*. However, both *Donaldsonia* and *Moringa* emerge in phylogenetic reconstructions as paraphyletic, and the single species of Section *Dysmoringa* is embedded within the Section *Moringa* clade. Because of the paraphyletic assemblage at the base of the family, there is little gain to communication in dividing *Moringa* into monophyletic groupings (e.g., creating a section for each of the four species of the basal grade). I therefore recommend that the sections should be dismantled. In keeping with the intent of previous authors to provide a means of designating distinctive groups within the genus, the informal terms ''bottle tree grade,'' ''slender tree clade,'' and ''tuberous'' clade, which is divided into the ''*rivae* clade'' and the ''red-flowered clade'' seem suitable for this purpose (these groups are highlighted on the tree in Fig. 6).

When groupings based on life form (Table 1) are superimposed on the phylogeny, the bottle trees appear paraphyletic, and slender trees monophyletic; the sarcorhizal trees (*M. arborea* and *M. ruspoliana*) are polyphyletic within the tuberous shrubs. In constrast,

the species groups delimited by wood anatomy (Table 1) are entirely congruent with the major groupings recovered in phylogenetic analyses. Particularly noteworthy is the confirmation of the pairing of *M. longituba* and *M. ruspoliana*, a relationship not proposed before wood anatomy studies. Likewise, the grouping of the tuberous species into a clade mirrors the conclusions of Ferguson (1985) based on pollen size variation. Verdcourt's (1985) phylogeny (Fig. 2), developed from often fragmentary herbarium specimens and without benefit of an outgroup, is very similar to the one favored here. It differs from the tree in Fig. 6 only in depicting a monophyletic bottle tree clade and *M. ruspoliana* sister to the tuberous clade rather than to *M. longituba*.

Verdcourt (1985) noted that ''adequate materials for a really thorough monograph were no more available today than they were 25 years ago when the idea was first conceived by Mr. J. B. Gillett.'' The statement is equally true now. As they were 40 years ago, collections are most seriously lacking from the tuberous clade of the Horn of Africa. In particular, there have been almost no collections from the key region of southeastern Ethiopia during this time. Human conflict in the Horn, the heart of *Moringa* species diversity, has proven a significant deterrent to exploration. The remoteness of the localities of most of these species and the difficulty of access to large areas of this dry tropical region are further challenges to exploration. Better sampling within the tuberous clade should clarify phylogenetic relationships within the *rivae* clade, provide additional samples of *M. pygmaea*, and will almost certainly reveal undescribed species.

ACKNOWLEDGEMENTS. Grassie, Barbara; 's'avuto dir? My continued gratitude for the collaboration of David Odee in Nairobi and Joseph Machua in the field. Gilfrid Powys was an inspiration in getting to the field and sky. Field work would have been impossible without Ambia A. Osman and Mohammed, Abdiaziz ''Jack'' Bashir, Halima Abdi Mohammed and Ahmad Salat Omar, Geoffrey Muluvi, Hassan A. Sheikh, Shahina Ghazanfar, Martin Fisher, Sylvain Razafimandimbison, V. Amalan Stanley, Fr. K. M. Mathew, and Herta Kolberg. DNA lab work at Washington University was possible thanks to the graciousness of Barbara Schaal. Thank you to Peter Raven, Mick Richardson, and Allan Larson, for their time and endless support. James Rodman and Bernard Verdcourt provided much encouragement and ideas. Joe and Mirella Olson have been amazingly patient. Sherwin Carlquist is a wealth of assistance, encouragement, and inspiration. Peter Stevens, Elizbeth Kellogg, Mike Dyer, Jessica Ingram, and Mike Veith gave help and suggestions. Chuck Hanson, Avinoam Danin, Tom VanDevender, Burl Mostul, David Orr, Nathan Wong, and Winnie Singeo generously provided material. Jason Bradford, Ana Lucía Caicedo Samper, John Gaskin, Paula Kover, Simon Malcomber, Allison J. Miller, Ken Olsen, and Jason Rauscher made lab work a pleasure. Sara Hoot and an anonymous reviewer generously improved the manuscript. Field and lab work were supported by grant # 6141–98 from the Committee for Research and Exploration of the National Geographic Society, United States National Science Foundation Doctoral Dissertation Improvement Award DEB-9801128, and the Andrew Mellon Foundation.

LITERATURE CITED

APG (ANGIOSPERM PHYLOGENY GROUP). 1998. An ordinal classification for the families of flowering plants. Annals of the Missouri Botanical Garden 85: 531–553.

BADILLO, V. M. 1971. *Monografia de la familia Caricaceae.* Maracay: Asociación de Profesores.

BAYER, R. J., D. E. SOLTIS, and P. S. SOLTIS. 1996. Phylogenetic inferences in *Antennaria* (Asteraceae: Gnaphalieae: Cassiinae) based on sequences from nuclear ribosomal DNA internal transcribed spacers (ITS). American Journal of Botany 83: 516–527.

CARLQUIST, S. 1998. Wood and bark anatomy of Caricaceae: correlations with systematics and habit. International Association of Wood Anatomists Journal 19: 191–206.

CUNNINGHAM, C. W. 1997. Can three incongruence tests predict when data should be combined? Molecular Biology and Evolution 14: 733–740.

DE QUEIROZ, K. 1985. The ontogentic method for determining character polarity and its relevance to phylogenetic systematics. Systematic Zoology 34: 280–299.

DOYLE, J. J. and J. L. DOYLE. 1987. A rapid isolation procedure for small quantities of fresh leaf tissue. Phytochemistry Bulletin 19: 11–15.

DUMOLIN-LAPÈGUE, S., A. KREMER, and R. J. PETIT. 1999. Are chloroplast and mitochondrial DNA variation species independent in oaks? Evolution 53: 1406–1413.

EDWARDS, K., C. JOHNSTONE, and C. THOMSON. 1994. A simple method of extraction. Nucleic Acid Research 19: 1349.

ENGLER, A. 1902. Contribuzioni alla conoscenza della flora dell'Africa orientale. Annuario Reale dell'Istituto Botanico di Roma 9: 241–256.

ERIKSSON, T. 1998. AutoDecay ver.4.0 (program distributed by the author). Department of Botany, Stockholm University. Stockholm.

FARRIS, J. S. 1989. The retention index and the rescaled consistency index. Cladistics 5: 417–419.

———, M. KÄELLERSJÖ, A. G. KLUGE, and C. BULT. 1994. Testing significance of incongruence. Cladistics 10: 315–319.

FELSENSTEIN, J. 1985. Confidence limits on phylogenies with a molecular clock. Systematic Zoology 34: 152–161.

FERGUSON, I. K. 1985. Pollen morphology of the Moringaceae. Kew Bulletin 40: 25–34.

GHASI, S., E. NWODOBO, and J. O. OFILI. 2000. Hypocholesterolemic effects of crude extract of leaf of *Moringa oleifera* Lam in high-fat diet fed wistar rats. Journal of Ethnopharmacology 69: 21–25.

GOLENBERG, E. M., M. T. CLEGG, M. L. DURBIN, J. DOEBLY, and D. P. MA. 1993. Evolution of a noncoding region of the cholorplast genome. Molecular Phylogenetics and Evolution 2: 52–64.

GRAHAM, S. W., J. R. KOHN, B. R. MORTON, J. E. ECKENWALDER, and S. C. H. BARRETT. 1998. Phylogenetic congruence and discordance among one morhpological and three moelcular data sets from Pontederiaceae. Systematic Biology 47: 545–567.

HAMILTON, M. B. 1999. Four primer pairs for the amplification of chloroplast intergenic regions with intraspecific variation. Molecular Ecology 8: 513–525.

HOOT, S. B. and A. W. DOUGLAS. 1998. Phylogeny of the Proteaceae based on *atpB* and *atpB-rbcL* intergenic spacer region sequences. Australian Systematic Botany 11: 301–320.

JAHN, S. A. 2001. Drinking water from Chinese rivers: challenges of clarification. Journal of Water Supply Research and Technology-Aqua 50: 15–27.

JOHANSEN, D. A. 1940. *Plant microtechnique.* New York: McGraw Hill.

JOHNSON, L. A. and D. E. SOLTIS. 1998. Assessing incongruence: empirical examples from molecular data. Pp 297–343 in *Mo-*

lecular systematics of plants II. DNA sequencing, eds. D. E. Soltis, P. S. Soltis, and J. J. Doyle. Boston: Kluwer.

JUMELLE, M. H. 1930. Les *Moringa* de Madagascar. Annales du Musée Colonial deMarseille sér. 4. 8: 1–20.

KALOGO, Y. and W. VERSTRAETE. 2000. Technical feasibility of the treatment of domestic wastewater by a CEPS-UASB system. Environmental Technology 21: 55–65.

KLUGE, A. G. and J. S. FARRIS. 1969. Quantitative phyletics and the evolution of anurans. Systematic Zoology 18: 1–32.

LARSON, A. 1994. The comparison of morphological and molecular data in phylogenetic systematics. Pp. 371–390 in *Molecular ecology and evolution: Approaches and applications*, eds. B. Schierwater, B. Streit, G. P. Wagner, and R. DeSalle. Basel, Switzerland: Birkhäuser Verlag.

LATZKO, E. and G. J. KELLY. 1983. The many-faceted function of phosphoenolpyruvate carboxylase in C3 plants. Physiologie Végétale 21: 805–815.

LEE, S. and J. WEN. 2001. A phylogenetic analysis of Prunus and the Amygdaloideae (Rosaceae) using ITS sequences of nuclear ribosomal DNA. American Journal of Botany 88: 150–160.

LIERE, K. and G. LINK. 1994. Structure and expression characteristics of the chloroplast DNA region containing the split gene for tRNA(Gly) (UCC) from mustard (*Sinapis alba* L.) Current Genetics 26: 557–563.

MASON-GAMER, R. J. and E. KELLOGG. 1996. Testing for phylogenetic conflict among molecular data sets in the Triticeae (Gramineae). Systematic Biology 45: 524–545.

MOODY, M. L., L. HUFFORD, D. E. SOLTIS, and P. E. SOLTIS. 2001. Phylogenetic relationships of Loasaceae subfamily Gronovioideae inferred from matK and ITS sequence data. American Journal of Botany 88: 326–336.

OLIVEIRA, J. T. A., S. B. SILVEIRA, I. M. VASCONCELOS, B. S. CAVADA, and R. A. MOREIRA. 1999. Compositional and nutritional attributes of seeds from the multiple purpose tree *Moringa oleifera* Lamarck. Journal of the Science of Food and Agriculture 79: 815–820.

OLSON, M. E. and S. CARLQUIST. 2001. Stem and root anatomical correlations with life form diversity, ecology, and systematics in *Moringa* (Moringaceae). Botanical Journal of the Linnean Society 135: 315–348.

——— and S. G. RAZAFIMANDIMBISON. 2000. *Moringa hildebrandtii*: a tree extinct in the wild but preserved by indigenous horticultural practices in Madagascar. Adansonia sér. 3 22: 217–221.

PANSTRUGA, R., A. SEILER, H.-J. HIRSCH, and F. KREUZALER. 1995. Genomic structure of a phosphoenolpyruvate carboxylase gene from potato (*Solanum tuberosum*). Plant Physiology 109: 1126.

RAMBAUT, A. 1996–1998. Se-Al Sequence Alignment Editor v1.0 alpha 1. Obtained at http://evolve.zps.ox.ac.uk/software/Se-Al/Se-Al10a1.hqx.

RODMAN, J. E., P. A. SOLTIS, D. E. SOLTIS, K. J. SYTSMA, and K. G. KAROL. 1998. Parallel evolution of glucosinolate biosynthesis inferred from congruent nuclear and plastid gene phylogenies. American Journal of Botany 85: 997–1006.

ROHLF, F. J. and R. R. SOKAL. 1981. *Statistical tables.* W. H. Freeman, New York.

SALEEM, R. and J. MEINWALD. 2000. Synthesis of novel hypotensive aromatic thiocarbamate glycosides. Journal of the Chemical Society-Perkin Transactions 1: 391–394.

SANG, T. and Y. ZHONG. 2000. Testing hybridization hypotheses based on incongruent gene trees. Systematic Biology 49: 422–434.

SIIKAMÄKI, P. 1999. Developmental instability in hybrids bewtween *Lychnis viscaria* and *Lychnis alpina* (Caryophyllaceae). American Journal of Botany 86: 1683–1686.

SWOFFORD, D. L. 2000. PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods). Version 4. Sinauer Associates, Sunderland, Massachusetts.

TABERLET, P. L., F. GIELLY, F. PAUTOUT, and J. BOUVET. 1991. Universal primers for amplification of noncoding regions of cpDNA. Plant Molecular Biology 17: 1105–1109.

TEMPLETON, A. R. 1983. Phylogenetic inference from restriction endonuclease cleavage site maps with particular reference to the evolution of humans and the apes. Evolution 37: 221–244.

VERDCOURT, B. 1958. Moringaceae: a correction. Kew Bulletin 13: 385.

———. 1985. A synopsis of the Moringaceae. Kew Bulletin 40: 1–23.

APPENDIX 1. Species sampling, loci sequenced, and voucher information. Caricaceae distribution notes from Badillo, 1971.

Species	Collector and #	Locality	Regions sequenced and GenBank accession number	Herbaria with vouchers
Carica microcarpa Jacquin	Waimea Arboretum 90p260	Cultivated plant from Waimea Arboretum, Haleiwa, Hawaii: originally from Venezuela: Aragua: Parque Nacional H. Pittier	PEPC AF378604; trnG AF378623; ITS AF378578	MO
Cylicomorpha parviflora Urban	Kayombo 1296	Tanzania: Iringa: Mufindi District: Lulanda Forest Reserve	PEPC AF378605; trnG AF378627; ITS AF378579	MO
	Mwangoka 387	Tanzania: Tanga: 05°04′44′S 38°25′34″E	trnG AF378626; ITS AF378580	MO
Jacaratia corumbensis Kuntze	Waimea Arboretum 84p423	Cultivated plant from Waimea Arboretum, Haleiwa, Hawaii: reported from Bolivia, Paraguay, Argentina	trnG AF378622	MO
	Olson, s.n.	Cultivated plant from Arid Lands Greenhouses, Tucson, AZ	ITS AF378575	MO
Jarilla chocola Standley	Reina 99-962A	Mexico: Sonora: Municipio de Yecora: Curea	trnG AF378624; ITS AF378577	MO
Jarilla heterophylla (Cerv. ex La Llave) Rusby	Olson, s.n.	Cultivated plant kindly provided by Sherwin Carlquist; native to central western Mexico	trnG AF378625; ITS AF378576	MO
Moringa arborea Verdc.	Olson and Machua 714	Kenya: Northeastern Province: Mandera District: ca. 35 km NW of Rhamu on road to Malka Mari, in small rocky canyon	PEPC AF378616; trnG AF378638; ITS AF378592	MO, EA, K
Moringa borziana Mattei	Olson and Muluvi 678	Kenya: Coast Province: Taita District: SW of Tsavo East National Park Voi gate	PEPC AF378617; trnG AF378639; ITS AF378593	MO, RHT, FT, K
Moringa concanensis Nimmo	Olson and Stanley 700 (individual 2 unless specified)	India: Tamil Nadu: Udimelpet Dist. Coimbatore Rd. to Parapaddur Dam in Palni Hills ca. 200 km WSW of Tiruchirapalli	PEPC AF378613; trnG AF378633; ITS AF378586, AF378587	MO, EA, FT, K
Moringa drouhardii Jum.	Olson 679 and ''Amboasary''	Madagascar: Tulear: near Amboasary	PEPC AF378606, AF378607, trnG AF378628; ITS AF378581, AF378582	MO, K, TAN
Moringa hildebrandtii Engl.	Olson ''E'' and ''W''	Madagascar: Tulear: cultivated in village of Ambohimahavelona	PEPC AF378608; AF378609	MO, EA, FT, K, TAN
	Olson ''2''	Cultivated plant provided by Rare Plant Research, Portland, Oregon	PEPC AF378610; ITS AF378584	MO, K, TAN
Moringa longituba Engl.	Olson and Machua 704	Kenya: Northeastern Province: Mandera District: ca. 20 km WNW of Mandera near locality of Filqo	PEPC AF378620; ITS AF378598	MO, K
	Olson and Machua 708	Kenya: Northeastern Province: Wajir District: ca. 20 km N of Wajir	PEPC AF378619; trnG AF378644; ITS AF378599	MO, EA, FT, K
	Olson and Machua 709	Kenya: Northeastern Province: Wajir District: Lafaley village, ca. 8 km N of Wajir	trnG AF378646; ITS AF378601	MO, EA, FT, K
	Olson and Machua 710	Kenya: Northeastern Province: Wajir District: ca. 12 km E of Wajir	trnG AF378643; ITS AF378602	MO, EA, FT, K
	Olson and Machua 712	Kenya: Northeastern Province: Wajir District: ca. 30 km E of Wajir (the ''*Moringa* sp. Kenya'' of Verdcourt, 1985)	trnG AF378645; ITS AF378603	MO, EA, FT, K
	Horwood, s.n.	Somalia: Nugal: Las Anod: near Las Anod	ITS AF378600	MO
Moringa oleifera Lam.	Stanley s.n.	India: Tamil Nadu: Chrompet (a neighborhood of Chennai)	PEPC AF378614; trnG AF378634; ITS AF378588	MO
	Olson, s.n.	Commercially available annual cultivar called ''PKM'', provided by V. Amalan Stanley, Chennai, India	ITS AF378589	MO

APPENDIX 1. Continued.

Species	Collector and #	Locality	Regions sequenced and GenBank accession number	Herbaria with vouchers
Moringa ovalifolia Dinter and A. Berger	Olson 716 and 718	Namibia: Namib-Naukluft Park: W of Kuiseb Pass	PEPC AF378612; trnG AF378629, AF378630	MO, WIND
	Olson, s.n.	Namibia: Karibib: seed purchased from B&T World Seeds	trnG AF378631; ITS AF378583	MO
Moringa peregrina (Forssk.) Fiori	Olson 569	Oman: Dhofar: in broad Wadi between Raysut and Mughsayl	trnG AF378637	MO, EA, FT, K
	Olson, s.n.	Seeds purchased in Cairo market, provided by Ben Stern, University of Bradford, UK	trnG AF378636; ITS AF378591	MO
	Danin, s.n.	Israel: Southern Region: En-Gedi	PEPC AF378615; trnG AF378635; ITS AF378590	MO
Moringa pygmaea Verdc.	Nugent 25	Somalia: Northern region: 13 km E of Qardho airstrip	none successful	EA
	Glover and Gilliland 1194	Somalia: North coast near Berbera	none successful	EA, K
Moringa rivae Chiov.	Olson and Powys 677	Kenya: Eastern Province: Marsabit District: E slope of Baio Mountain	PEPC AF378618; trnG AF378640; ITS AF378594	MO, EA, FT, K
	Olson and Machua 701	Kenya: Northeastern Province: Mandera District: ca. 4 km N of Rhamu	ITS AF378595	MO, EA, FT, K
Moringa ruspoliana Engl.	Olson and Machua 702	Kenya: Northeastern Province: Mandera District: Rhamu-Dimtu Division: around village of Yabicho	PEPC AF378621; trnG AF378641; ITS AF378597	MO, EA, FT, K
	Olson and Machua 703	Kenya: Northeastern Province: Mandera District: on Kenya-Somalia border S of Dawa River	trnG AF378642; ITS AF378596	MO, EA, FT, K
Moringa stenopetala (Baker f.) Cufodontis	Olson 675	Kenya: Rift Valley Province: Baringo District: Parmalok Island, Lake Baringo	PEPC AF378611; trnG AF378632; ITS AF378585	MO, EA, FT, K

APPENDIX 2. Morphological characters and data set.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28
Cylicomorpha	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	?	0	0
M. drouhardii	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1
M. hildebrandtii	0	0	1	0	1	0	0	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1	1	?	0	1
M. ovalifolia	0	0	1	1	1	0	0	0	1	0	0	0	1	1	0	0	1	1	1	?	0	0	0	1	1	0	0	1
M. stenopetala	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1	0	1	1	0	0	0	0	0	1	1	?	0	1
M. concanensis	1	0	0	0	1	0	0	1	1	1	1	0	1	1	1	1	1	1	0	0	0	0	0	2	1	1	1	2
M. oleifera	1	0	0	1	0	0	0	0	1	1	1	0	1	1	1	1	1	1	1	0	0	0	0	2	1	1	1	2
M. peregrina	0	0	0	1	0	0	0	0	0	1	1	0	1	1	1	1	1	0	1	0	0	0	0	2	1	1	1	2
M. arborea	0	0	1	0	1	0	1	1	0	1	0	0	1	1	1	?	1	1	1	0	1	1	0	2	0	0	2	2
M. rivae	0	1	0	1	0	0	1	1	0	1	0	1	1	1	1	0	1	1	1	0	1	1	0/1	2	0	0	2	2
M. borziana	0	0	0	0	0	0	0	0	0	1	0	1	1	2	1	0	1	1	1	0	1	0	1	2	0	0	2	2
M. pygmaea	0	0	0	0	0	?	0	1	0	1	0	1	1	2	1	?	1	1	1	0	?	0	1	2	0	?	2	2
M. longituba	0	1	0	0	0	1	0	0	0	1	0	1	1	2	1	0	1	1	1	1	1	2	0	2	0	0	2	2
M. ruspoliana	0	1	0	0	?	1	0	1	0	1	0	0	1	1	1	?	1	1	1	1	1	2	0	2	0	0	2	2

Assignment of morphological character states. Characters marked with an asterisk are derived from ontogenetic studies.

A. Wood, bark, and root anatomical characters. The following characters are for the most part depicted and discussed in Carlquist (1998) and Olson and Carlquist (2001).

1. Callose plugs on sieve plates: Both in transections and longisections stained with safranin, large pink-staining plugs, presumably callose, are conspicuous on the sieve plates of most species. 0 = present, 1 = absent.
2. Rhomboidal crystals in phloem rays. 0 = absent, 1 = present.
3. Phloem ray sclereids. 0 = absent, 1 = present.
4. Druses in cortical sclerenchyma. The thick-walled cells present in the outer bark of some species contain druses that are completely immobilized within their massive walls. 0 = absent, 1 = present.
5. Phelloderm sclereids. 0 = absent, 1 = present.
6. Paratracheal axial parenchyma in shoots: 0 = present, 1 = absent.
7. Druses in tyloses. 0 = absent, 1 = present.
8. Druses in xylem rays. 0 = present, 1 = absent.
9. Uniserate wings on multiseriate rays. 0 = present, 1 = absent.
10. Principal xylem cell type in stems. Some species are characterized by large amounts of paratracheal axial parenchyma and few libriform fibers, other species by the opposite condition. 0 = paratracheal axial parenchyma, 1 = libriform fibers.
11. Growth rings. Most species are characterized by seasonal change in xylem cell size and shape and often cell type. A few species show little if any seasonal fluctuation in cell shape. 0 = present, 1 = absent.
12. Libriform fibers in root secondary xylem. 0 = present, 1 = absent.

B. Seed, seedling, and leaf characters.

13. *Ground tissue proliferation in epicotyl. All *Moringa* and all Caricaceae examined form swollen, often tuberous hypocotyls early in ontogeny. In many species, this swelling of the lower part of the stem extends above the insertion of the cotyledons into the epicotyl. In other species, there is a strong differentiation into a bloated underground tuberous epicotyl/root and a slender aboveground stem that usually survives only one season before dying back to the tuber. 0 = epicotyl swollen, 1 = epicotyl slender.
14. *Seasonal persistence of the shoot at different ontogenetic stages. In several species the shoot formed upon germination grows very rapidly in height and, barring injury, eventually forms the main bole of the tree. These shoots may be considered permanent in that they are not shed naturally by the plant. In contrast, the juvenile plants of other species often persist through many seasons as tubers that send up seasonal shoots when conditions are favorable and die back to the tuber during drought. A subset of the *Moringa* species with ephemeral juvenile shoots eventually form root systems that are sufficiently large to support permanent shoots, and the plant begins to grow into a tree with a permanent trunk. Still other species maintain the characteristic of regularly dying back to the tuber in times of drought throughout the life of the plant. 0 = juvenile and adult shoots permanent, 1 = juvenile shoots ephemeral, adult shoots permanent, 2 = juvenile and adult shoots ephemeral.
15. *Germination phanerocotylar (cotyledons emerge from the seed coat) = 0, 1 = germination cryptocotylar (cotyledons remain in the seed coat).
16. *Leaves palmate = 0, leaves pinnate = 1. In the Caricaceae, most leaves are palmate, especially in *Cylicomorpha* and *Jacaratia*. In *Moringa*, adult leaves are pinnate. However, the juvenile leaves of many species are distinctly palmate, and the transition to pinnate leaves occurs over the first seven leaves produced by the seedling.
17. *Leaf margin of 1st leaves. The first leaves of some species are characterized by irregular margins with occasional lobes and indentations (here denoted ''irregular margins''). Other species have entire leaf margins. 0 = irregular margins, 1 = entire margins.
18. Seed wings. Most *Moringa* species are characterized by wide, hyaline wings that run the length of the seeds. Two species, and the Caricaceae, are characterized by wingless seeds with variously irregular surfaces. 0 = wings absent, 1 = wings present.
19. Leaftip epidermal layers. Some species have just one epidermal layer at the leaf margin, whereas others have several. Surprisingly, this character does not correlate with leaflet size. 0 = one epidermal layer, 1 = multiple epidermal layers.
20. Leaf trichomes. Whereas the young leaves of many species are covered with sometimes dense indumentum, the adult leaves of most of these species are apparently glabrous. Character states were assessed both from examination of intact leaves and of leaf serial sections. 0 = absent, 1 = present.

21. Epidermal cell dimples. In some species, the cells surrounding the stomata show dimples or folds along the surfaces contacting the guard cells and perpendicular to them. 0 = dimples absent, 1 = dimples present.

C. Floral characters.

22. Flower color 1. Distribution of red pigmentation. Some species have localized maroon or pink patches, whereas others have dense red pigmentation distributed throughout the perianth. 0 = absent or along midvein, 1 = distinctly aggregated at petal tips, 2 = throughout perianth.
23. Flower color 2. Yellow. Flowers cream, with or without pink or brown lines or blotches, or red = 0, 1 = bright yellow.
24. *Anther orientation in ontogeny. The five anthers of all species studied initially point to the center of the flower. As anthesis approaches, the filaments twist. In some species, three of the anthers come to face one way while the others come to face the opposing direction (here denoted 3/2 orientation). In other species, four anthers point the same direction and only one faces the opposing direction (here denoted 4/1 orientation). The Caricaceae appear to maintain centrally-pointing anthers throughout ontogeny. 0 = no change in orientation; 1 = 3/2, 2 = 4/1.
25. Ovary pubescence. The basal half of the ovary in some species is clothed in a dense coat of long unicellular trichomes that seem to form a barrier to the nectariferous lower part of the hypanthium. 0 = ovary glabrous, 1 = ovary pubescent.
26. *Carpel emergence. During floral ontogeny, the ovary emerges at the same time as the anthers in some species, but only well after the anther primordia are distinct in others. 0 = ovary emergence coincident with anthers, 1 = ovary emergence after anthers.
27. *Filament and staminode postgenital adhesion. The filaments and staminodes of all species are free upon differentiation from the primordia. In some species, they remain so throughout all of ontogeny. In other species, the filaments and staminodes bring themselves into contact in a semicircle and adhere to one another via an unknown substance, presumably a sticky secretion, visible in sections.
28. *Flora symmetry in ontogeny. All species of Caricaceae have radially symmetrical flowers, and the species studied show this arrangement from the earliest stages of development. All *Moringa* species studied exhibit bilateral symmetry very early in ontogeny. Some species have adult floral morphologies that are radially symmetrical or nearly so, whereas others have very clearly bilaterally symmetrical flowers. 0 = symmetrical throughout development, 1 = bilateral to radial, 2 = bilateral throughout development.