

Molecular Phylogenetics and Biogeography of the Caribbean-Centered *Croton* Subgenus *Moacroton* (Euphorbiaceae s.s.)

Benjamin W. van Ee^{1,4,5} · Paul E. Berry² ·
Ricarda Riina² · Jorge E. Gutiérrez Amaro³

¹ Department of Botany, University of Wisconsin-Madison, 430 Lincoln Dr., Madison, WI 53706, USA

² Department of Ecology and Evolutionary Biology, University of Michigan Herbarium,
3600 Varsity Drive, Ann Arbor, MI 48108, USA

³ Jardín Botánico Nacional, Carretera del Rocío, Km 32, CP. 19230, Calabazar, Boyeros, Havana, Cuba

⁴ Department of Botany, National Museum of Natural History, Smithsonian Institution,
Washington, DC 20013-7012, USA

⁵ Author for Correspondence; e-mail: bvaneec@uwalumni.com

Published online: 29 April 2008

© The New York Botanical Garden 2008

Abstract Initial molecular phylogenetic studies established the monophyly of the large genus *Croton* (Euphorbiaceae s.s.) and suggested that the group originated in the New World. A denser and more targeted sampling of *Croton* species points to a South American origin for the genus. The nuclear and chloroplast genomes indicate a different rooting for the phylogeny of *Croton*. Although we favor the rooting indicated by the chloroplast data our conclusions are also consistent with the topology inferred from the nuclear data. The satellite genera *Cubacroton* and *Moacroton* are embedded within *Croton*. These two genera are synonymized into *Croton* and a new subgenus, *Croton* subgenus *Moacroton*, is circumscribed to include them and their allied *Croton* species. *Croton* subgenus *Moacroton* is morphologically characterized by a primarily lepidote indumentum, bifid or simple styles, and pistillate flowers with sepals that are connate at the base. This newly circumscribed subgenus is found from North America to South America, and in contrast to the majority of *Croton* species most of its members are found in mesic habitats. The group is most diverse in the greater Caribbean basin. A molecular clock was calibrated to the phylogeny using the available Euphorbiaceae fossils. The timing and pattern of diversification of *Croton* is consistent with both the GAARlandia and Laurasian migration hypotheses. A single species, *Croton poecilanthus* from Puerto Rico, is placed incongruently by its nuclear and chloroplast genomes. The possibility of this species being of hybrid origin is discussed.

Resumen Estudios moleculares preliminares establecieron que el género grande *Croton* (Euphorbiaceae s.s.) es monofilético y sugirieron que el grupo origino en el Nuevo Mundo. Un muestreo mas denso y especifico de especies de *Croton* apunta a un origen Suramericano para el género. Los genomas nucleares y del cloroplasto indican una diferente raíz para la filogenia de *Croton*. Aunque favoramos la raíz

indicada por los datos del cloroplasto nuestras conclusiones también son consistentes con la topología deducida por los datos nucleares. Los géneros satélite *Cubacroton* y *Moacroton* están encajados dentro de *Croton*. Estos dos géneros son sinonimizadas dentro de *Croton* y un nuevo subgénero, *Croton* subgénero *Moacroton*, es circunscrito para incluirlos y las especies aliadas de *Croton*. *Croton* subgénero *Moacroton* esta caracterizada morfológicamente por un indumento mayormente lepidoto, estilos bifidos o simples, y flores pistiladas con sépalos unidos en la base. Este nuevamente circunscrito subgénero esta distribuido desde Norteamérica hasta Suramérica, y en contraste con la mayoría de las especies de *Croton* la mayoría de sus especies se encuentran en habitats húmedos. El grupo es mas diverso en la zona del Caribe. Un reloj molecular fue calibrado a la filogenia utilizando los fósiles de Euphorbiaceae disponibles. La sincronización y patrón de la diversificación de *Croton* son consistentes con las hipótesis de GAARlandia y la migración Laurasiana. Una especie, *Croton poecilanthus* de Puerto Rico, es puesto incongruentemente por sus genomas nucleares y del cloroplasto. La posibilidad de que esta especie tiene un origen hibrido es discutido.

Keywords Caribbean biogeography · *Croton* · Euphorbiaceae · GAARlandia · *Moacroton* · Molecular phylogenetics

Palabras claves Biogeografía del Caribe, *Croton*, Euphorbiaceae, GAARlandia, *Moacroton*, Filogenética molecular

Introduction

Croton (Euphorbiaceae s.s.) has historically been treated as one of the largest genera of flowering plants (Frodin, 2004). In the *World Checklist and Bibliography of Euphorbiaceae*, Govaerts et al. (2000) recognized 1,223 species. The genus occurs worldwide, mostly in arid and tropical areas, with greatest species-level diversity in Madagascar, Brazil, and the Caribbean. However, the group has been dogged by a lack of understanding of its phylogeny and questions about the inclusion or not of segregate genera such as *Astraea*, *Crotonopsis*, *Cubacroton*, *Eremocarpus*, *Julocroton*, and *Moacroton*. Webster (1993) proposed a fairly inclusive circumscription of *Croton*, including all of the segregate genera mentioned above except for the then poorly known Cuban endemics *Moacroton* and *Cubacroton*. He also divided the genus into 40 sections, based on suites of morphological characters, although he acknowledged that there were many remaining uncertainties.

The first molecular phylogeny of *Croton* (Berry et al., 2005) examined 88 species within the genus and related groups, and covered 29 of the 40 sections recognized by Webster (1993). Their results showed that a recently described genus from Brazil, *Brasiliacroton*, was sister to *Croton*, with the latter including all of the former generic segregates sampled except for *Astraea*. One species of *Moacroton* was included in the analysis and was tenuously placed within *Croton*, but no material of *Cubacroton* was available at the time.

In the results of Berry et al. (2005), a small clade of three American species emerged as sister to all of the remaining species of *Croton*. These species were

Croton alabamensis, a rare species distributed disjunctly between central Alabama and east-central Texas; *Croton olivaceus*, a poorly known species from montane forests in Peru and Ecuador; and *Moacroton lanceolatus*, one of five locally endemic species of *Moacroton* occurring on serpentine soils of Cuba. These species were referred to as Clade C-1 in Berry et al. (2005) and this clade is henceforth referred to here as the *Moacroton* clade. The other large sister clade of *Croton* (Clades C-2 to C-11 in Berry et al., 2005) will henceforth be referred to here as the core *Croton* clade. Because of the isolated phylogenetic position and conservation interest of *Croton alabamensis*, van Ee et al. (2006) examined the phylogenetic relationships of that species in greater detail. Their results showed that *C. alabamensis* is most closely related to *Croton* section *Corylocroton*, a group of small trees and shrubs distributed in the greater Caribbean basin and extending into South America.

The findings of Berry et al. (2005) and van Ee et al. (2006) both indicate that *Croton* likely originated in the New World, possibly in South America or an area centered on the Caribbean basin. However, the sampling was too sparse to draw more robust conclusions. These initial results led to the current study, in which our aims were to conduct a more in-depth and targeted sampling to determine what other species or groups within *Croton* belong to the *Moacroton* clade, as well as their implications for Caribbean biogeography and character evolution within the genus.

Materials and Methods

Sampling Strategy

The three species found by Berry et al. (2005) to belong in the *Moacroton* clade (*C. alabamensis*, *C. olivaceus*, and *Moacroton revolutus*) possess combinations of characters that make them difficult to place within any of Webster's (1993) 40 sections. Our sampling strategy was to sample species that exhibit some of the characters of these species, as well as to continue a broad sampling across the genus.

The remaining four of the five species of *Moacroton*, which are found exclusively in Cuba on serpentine soils (Borhidi, 1977), were targeted for sampling. These have coriaceous leaves with lepidote scales, and their staminate flowers are greatly reduced compared to *Croton* (four to six per flower with no filaments, vs. 6–130 stamens in other species of *Croton*). *Cubacroton maestrense*, a small cloud forest tree from eastern Cuba, was also sampled because it shares many vegetative and floral characters with *Moacroton*. Its main difference from *Croton* is in the male flowers, which have one or two stamens with well-developed filaments. *Cubacroton* grows on igneous substrates in cloud forests of the Sierra Maestra Mountains at some of the highest elevations in Cuba. *Croton poecilanthus*, which is endemic to cloud forests in eastern Puerto Rico, was sampled because of its remarkable similarity to *Moacroton* in all features except the staminate flowers, which have numerous stamens with well developed filaments (Fig. 1).

Additional members of *Croton* section *Corylocroton* were targeted for sampling, given the findings by van Ee et al. (2006), which placed this group closest to *C. alabamensis*.

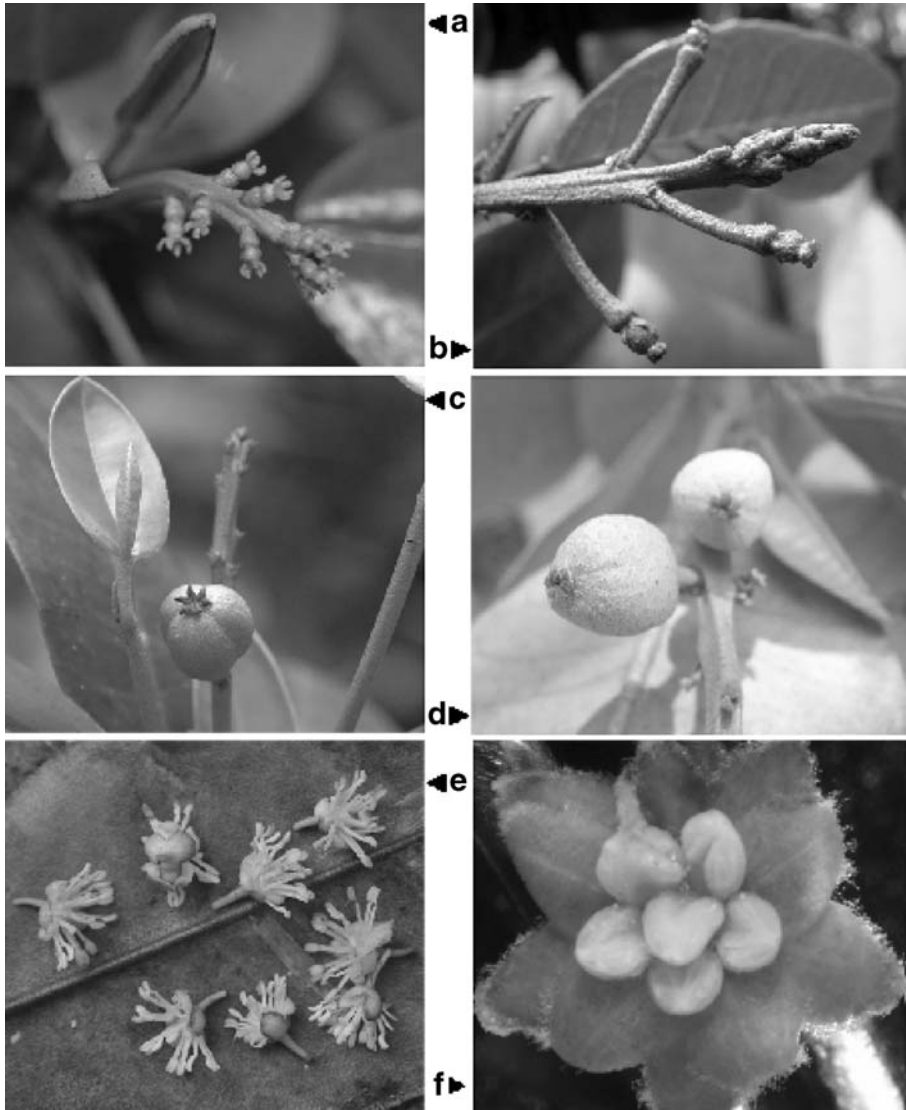


Fig. 1 *Croton poecilanthus* (left, a, c, e) and *Croton ekmanii* [*Moacrotan ekmanii*] (right, b, d, f). a and b, pistillate inflorescence. c and d, Fruits and stigma lobes. e and f, staminate flowers

Croton olivaceus is found in tropical premontane wet forests between 400 and 1,000 m in Ecuador and Peru. Its sparse covering of stellate trichomes gives it an overall glabrous appearance. It has elliptical leaves and elongated (up to 30 cm) inflorescences. A search for similar species led us to target *C. nubigenus* from Nicaragua and Honduras and *C. corinthius* from Costa Rica. Both species occur in similar habitats and share the glabrous appearance, leaf shape, and inflorescence type of *C. olivaceus*.

Data Sets

Three overlapping data sets were assembled. The first data set, a four-gene matrix (ITS, *trnL-trnF*, *ndhF*, and *rbcL*), is made up of 32 accessions consisting of three outgroup taxa and 29 *Croton* species. This data set is intended to provide a robust phylogenetic hypothesis for the major lineages of *Croton*. The second data set of 45 accessions is a three-gene matrix of ITS, *trnL-trnF*, and *ndhF* sequences. It includes more species and was assembled to show the clade membership, as well as to infer phylogenetic relationships at the species level. The third data set is a two-gene matrix of *rbcL* and *trnL-trnF* sequences of 56 species, 30 from *Croton* (including *Moacroton* and *Cubacroton*), and 26 from across Euphorbiaceae s.s. The sampling in this data set is intended for dating the divergence times of the genus and its component clades. The 150 sequences with GenBank numbers beginning with EF- were newly generated. The 68 GenBank accessions that start with AY- or DQ- were obtained from Berry et al. (2005), van Ee et al. (2006), and Wurdack et al. (2005; Appendix).

The outgroups, *Brasiliocroton*, *Astraea*, and *Acidocroton*, were selected based on the family-wide analysis of *rbcL* and *trnL-trnF* by Wurdack et al. (2005) and the tribal-level analysis of ITS and *trnL-trnF* by Berry et al. (2005). *Brasiliocroton mamoninha* is the sister taxon to all of *Croton*, and its ITS sequence is close enough to *Croton* to allow unambiguous alignment. *Astraea*, along with *Acidocroton*, is sister to *Croton* + *Brasiliocroton*.

DNA Extraction and Sequencing

Total genomic DNA was extracted from silica-dried tissue or herbarium collections of single individuals (Chase & Hillis, 1991) using QIAGEN DNeasy plant kits. The nuclear ribosomal internal transcribed spacer (ITS) and 5.8 S regions and the chloroplast *trnL* intron and *trnL-trnF* spacer and *ndhF* regions were sequenced using published primers (Baldwin et al., 1995; Urbatsch et al., 2000; White et al., 1990; Taberlet et al., 1991; Olmstead & Sweere, 1994). The plastid region *rbcL* was amplified with the same primers used by Wurdack et al. (2005; K. Wurdack, personal communication), which are 1F (5' ATGTCACCACAAACAGAAAC 3') and 1460R (5' TCCTTTTAGTAAAAGATTGGGCCGAG 3'). Amplified PCR product of *rbcL* was then cycle-sequenced using these same two primers plus the internal primers 636F (5' GCGTTGGAGAGAYCGTTTCCT 3') and 724R (5' TCACATGTACCTGCAG TAGC 3'). PCR products were cleaned using either the AMPure (Agencourt Bioscience, Beverly, Massachusetts, USA) magnetic bead method, or enzymatically with ExoSAP-IT (GE Healthcare, Chalfont St. Giles, UK), and cycle-sequenced at the University of Wisconsin Biotechnology Center with an ABI 3100 automated DNA sequencer using the manufacturers protocols (Big Dye version 3.1; Applied Biosystems, Foster City, California, USA). All molecular methods followed standard procedures and are the same as those described by Berry et al. (2005), differing only in that the *rbcL* region was amplified using 52°C for the annealing temperature in the PCR reactions. Sequences were edited and assembled in Sequencher 3.0 (GeneCodes Co., 1991–1995), and then manually aligned in MacClade 4.06 (Maddison and

Maddison, 2000). All sequences are deposited in GenBank and the phylogenetic trees and data matrices are deposited in TreeBASE.

Phylogenetic Analyses

Maximum parsimony (MP) and maximum likelihood (ML) analyses were conducted in PAUP* 4.0b10 (Swofford, 2002). *Astraea* and *Acidocroton* were designated in PAUP* as a monophyletic outgroup. All characters were weighted equally. Maximum parsimony bootstrap percentages were obtained from 1,000 bootstrap replicates with simple taxon addition (holding one tree at each step), tree-bisection-reconnection (TBR) branch swapping, and saving no more than 100 trees per bootstrap replicate.

The best fitting likelihood models for each of the data partitions, as well as for the combined data sets, were selected with Modeltest ver. 3.07 (Posada & Crandall, 1998) using both the hierarchical likelihood ratio test (hLRT) at $\alpha=0.01$ and Aikaike Information Criterion (AIC). Likelihood models were selected for the first and second, third, and all codon positions for the protein-coding regions, *rbcL* and *ndhF*, for the data sets that were calibrated for a molecular clock. Maximum likelihood searches were performed using the best-fitting models with as-is stepwise taxon addition and TBR branch swapping.

Bayesian phylogenetic analyses were performed in MrBayes ver. 3.1.2 (Huelsenbeck & Ronquist, 2001). The most suitable models of nucleotide substitution for each of the gene regions were selected in MrModeltest ver. 2.2 (Nylander, 2004). In the four- and three-gene data sets, which were the only ones to include both nuclear and chloroplast data, the nuclear ITS and chloroplast partitions were analyzed independently and in combination. In the combined analyses model parameters were fit independently to the separate data partitions. In the four- and three-gene analyses four Bayesian Markov Chain Monte Carlo (MCMC) runs of four linked chains were run for 1,000,000 generations each. A temperature setting of 0.08 was used rather than the default setting of 0.20 to increase the acceptance rates for swaps between different chains, which at the default temperature setting was lower (Altekar et al., 2004). In the MCMC chains of the four-gene data set every fiftieth generation was sampled resulting in a total of 80,004 trees, 20,001 from each run. In the MCMC chains of the three-gene data set every 100th generation was sampled resulting in a total of 40,004 trees, 10,001 from each run. The two-gene data set was run for two MCMC runs sampling every hundredth generation resulting in a total of 20,002 trees, 10,001 from each run. A likelihood-by-generation graph was produced in Excel to estimate the burn-in portion of the MCMC runs and the trees from the generations on the steeply sloping portion of the graph were discarded. The Comparetree command in MrBayes ver. 3.1.2 was used to compare the trees from different runs to determine if they had converged on the same set of trees. The Sumt command was used to calculate the consensus trees from the pooled samples. The models selected for the partitions of the data sets that were analyzed with ML and Bayesian methods are summarized in Table 1.

Incongruence between the gene regions was estimated using the incongruence length distance (ILD) test (Farris et al., 1994), implemented as the partition homogeneity test in PAUP* 4.0b10 (Swofford, 2002). For the four-gene data set, an

Table 1 Maximum Likelihood Models Selected by Modeltest ver. 3.07 for the Different Partitions of the Data Sets that were Analyzed

	AIC	hLRT
Data set one: combined	GTR+I+G	GTR+I+G
Data set one: combined chloroplast	TVM+I+G	TVM+I+G
Data set one: ITS	GTR+I+G	TrNef+G
Data set one: <i>rbcL</i>	GTR+I+G	TVM+I+G
Data set one: <i>rbcL</i> positions 1 and 2	TVM+I+G	F81+I+G
Data set one: <i>rbcL</i> position 3	TVM+G	HKY+G
Data set one: <i>ndhF</i>	GTR+I+G	GTR+G
Data set one: <i>ndhF</i> positions 1 and 2	GTR+G	TVM+G
Data set one: <i>ndhF</i> position 3	TVM+G	TVM+G
Data set one: <i>trnL-trnF</i>	GTR+I+G	HKY+G
Data set two: combined	GTR+I+G	GTR+I+G
Data set two: ITS	GTR+I+G	SYM+I+G
Data set two: <i>ndhF</i>	GTR+I+G	GTR+G
Data set two: <i>trnL-trnF</i>	GTR+I+G	HKY+G
Data set three: combined	GTR+I+G	TrN+I+G
Data set three: <i>trnL-trnF</i>	TVM+G	HKY+G
Data set three: <i>rbcL</i>	GTR+I+G	GTR+I+G
Data set three: <i>rbcL</i> positions 1 and 2	GTR+I+G	F81+I+G
Data set three: <i>rbcL</i> position 3	TVM+G	TVM+G

ILD test comparing the three chloroplast regions was first performed. Following that, an ILD test was performed comparing the nuclear ITS region with the three combined chloroplast regions. A similar strategy for testing for incongruence was used on the three-gene data set; first the two chloroplast regions, the non-coding *trnL-trnF* and the coding *ndhF*, were tested against each other, and then ITS was compared against the chloroplast sequence data. For the two-gene data set the *trnL-trnF* region was compared to the *ndhF* region. ILD tests were conducted using 100 partition homogeneity replicates of ten random sequence addition replicates each, simple taxon addition, tree bisection-reconnection (TBR) searches saving up to ten trees in each rep, and limiting the maximum number of trees saved to 1,000. The ILD test of the two-gene data set differed in that it was run for 1,000 replicates. Additional settings for the ILD test involving more replicates and more saved trees were explored, but in all cases in which this was done the *P* values were either the same or similar. Informative characters were removed before performing the ILD tests, which is advisable given that the two data partitions have different percentages of variable characters (Cunningham, 1997).

Analyses were performed with *Croton poecilanthus* both included and excluded, given the incongruence between its nuclear and chloroplast sequences (discussed below).

Character Mapping

To reconstruct ancestral character states, characters were mapped onto the molecular phylogeny in MacClade 4.06 (Maddison & Maddison, 2000) using both accelerated transformation (ACCTRAN) and delayed transformation (DELTRAN). One ecological character and seven morphological characters were mapped onto the fully resolved majority rule consensus tree of the maximum parsimony, maximum

likelihood, and Bayesian phylogenies inferred from the combined ITS, *trnL-trnF*, *ndhF*, and *rbcL* four-gene data set. This topology is most similar to that found in the combined data ML search, and the early branching pattern is the same as that indicated by the combined Bayesian phylogeny. This topology is also the same as that recovered in all analyses of the chloroplast data. Characters were scored based on examination of herbarium collections. All characters were treated as unordered. Characters scored are: habitat, presence/absence of stellate trichomes, presence/absence of lepidote trichomes, stigma type, stamen number, presence/absence of petals in pistillate flowers, sepal type in pistillate flowers, and presence/absence of bisexual cymules. These morphological characters are the main ones used by Webster (1993) in developing the current taxonomic framework of the genus. Characters and coding information is summarized in Table 2.

Divergence Time Estimates

To estimate the divergence times within *Croton*, we supplemented the Euphorbiaceae s.s.-wide sampling of *rbcL* and *trnL-trnF* by Wurdack et al. (2005) to place the main clades of *Croton* on the family-level phylogeny. Given the lack of known fossils assignable to *Croton*, or to any genera within tribe *Crotoneae*, the two-gene data set assembled in this study was necessary for incorporating the known Euphorbiaceae fossils as calibration points. This data set consists of a cross-section of Euphorbiaceae s.s. genera. The *rbcL* and *trnL-trnF* regions have been shown to be alignable across the Euphorbiaceae while providing good phylogenetic resolution (Wurdack et al., 2005). By including *Hippomane mancinella* and its sister taxon *Bonania cubana*, a node is created that can be constrained to a minimum age of 40 mya based on the oldest reliable age estimate of the fossil *Crepetocarpon perkinsii* (Dilcher & Manchester, 1988; Potter & Dilcher, 1980). By including *Acalypha californica* and its sister taxon *Argomuelleria macrophylla*, another node is created that can be constrained to a minimum of 61 mya based on the estimated age of the *Acalypha* type fossil pollen from the Early Paleocene (Sun et al., 1989). Davis et al. (2005) and Wikström et al. (2001) estimated a Cretaceous origin for Euphorbiaceae, with an origin of *Croton* as early as 65 million years ago at the Cretaceous/Tertiary boundary. The two fossil-based dates are used as minimum age constraints for the applicable nodes on the *rbcL* and *trnL-trnF* phylogeny and the maximum age constraint, the estimated age of 114 mya of Euphorbiaceae, is taken

Table 2 Ecological and Morphological Characters

Character	Codes
Habitat	0=xeric; 1=mesic
Stellate trichomes	0=absent; 1=present
Lepidote trichomes	0=absent; 1=present
Stigma type	0=simple; 1=bifid; 2=multifid
Stamen number	0=6 or less; 1=10 or more
Petals of pistillate flowers	0=absent; 1=present
Sepals of pistillate flowers	0=imbricate; 1=valvate; 2=connate at base (but not valvate); 3=obsolete
Bisexual cymules	0=absent; 1=present

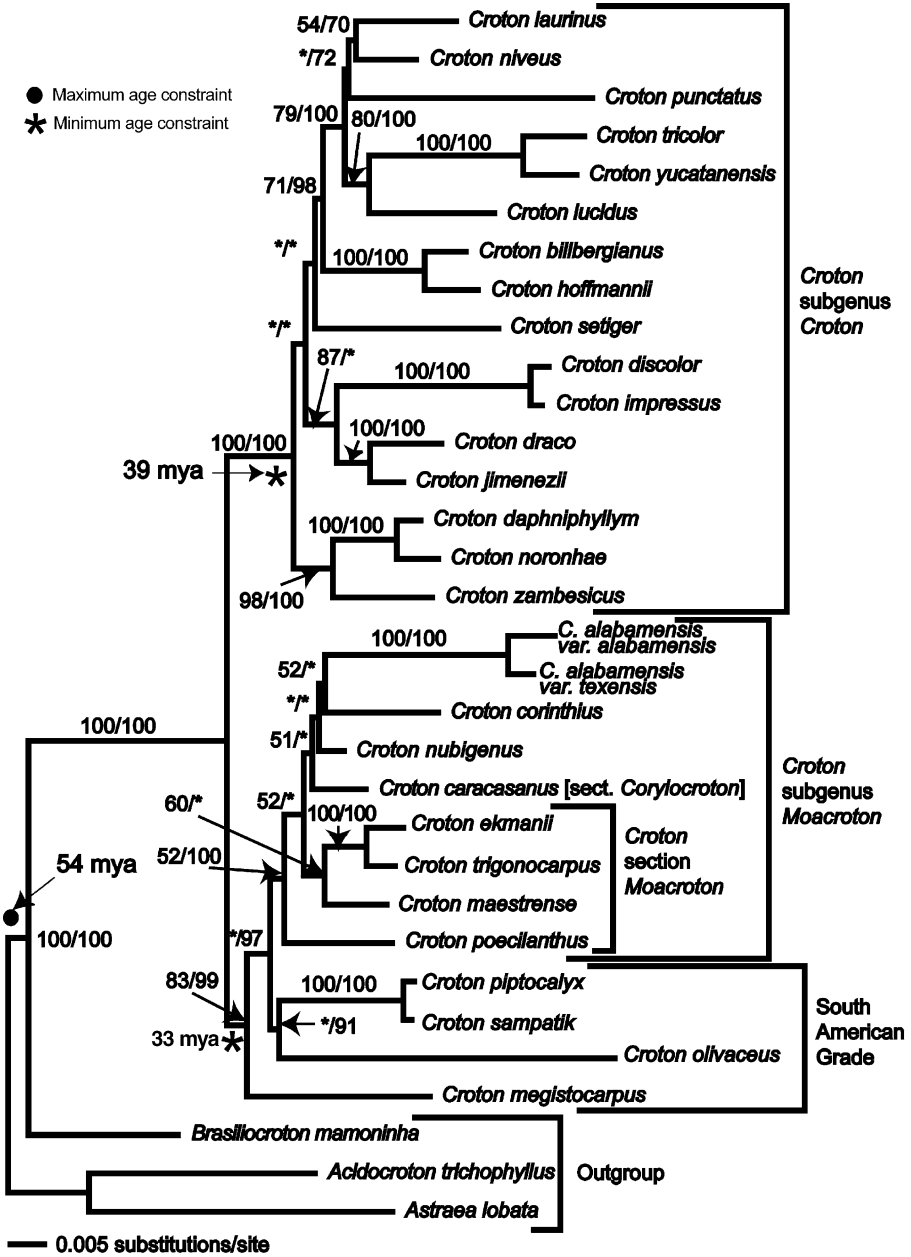


Fig. 2 Maximum likelihood phylogram of the combined *rbcL*, *ndhF*, *trnL-trnF*, and ITS data set analysis. Branches are drawn in proportion to the substitutions per site assuming the GTR+I+G model. The numbers on the nodes represent maximum parsimony bootstrap values followed by Bayesian clade credibility values. The parsimony and Bayesian topologies are not identical to the ML topology. Nodes that are either not found on the parsimony bootstrap tree or the Bayesian consensus tree, or have less than 50% bootstrap support are labeled with an asterisk (*)

from a thorough sampling of the Malpighiales that used four macrofossils and 11 palynofossils (Davis et al., 2005). The estimated ages for the divergence of *Croton*, and the major clades within *Croton*, obtained from this family-wide sampling can then be applied to the more thorough genus-level phylogeny.

The combined two-gene data set was tested for a molecular clock using the Likelihood Ratio test by comparing twice the difference of the likelihood score of a tree assuming a clock and the likelihood score of the same tree without a clock to a χ^2 distribution with $n-2$ degrees of freedom, where n =number of taxa. The first and second, third, and all three codon positions together of *rbcL* were also tested for clock-like behavior given the possibility that the first and second positions may exhibit more clock-like behavior (Magallón & Sanderson, 2005).

We are cognizant of the potential for problems with clade–age calibrations due to the few fossils available (Nixon, 2006), but currently these fossils represent the best available minimum age constraints within the family.

Divergence times were estimated independently on the maximum likelihood and Bayesian phylograms obtained from both combined and independent analyses of *rbcL* and *trnL-trnF* using penalized likelihood (PL) and non-parametric rate smoothing (NPRS) in the program r8s ver. 1.71 (Sanderson, 2003). The optimal smoothing PL parameters for the data from the different analysis methods were found using fossil-based cross-validation. To do so, *Panda oleosa* was pruned off, and the nodes directly above *Hippomane* and *Acalypha* were constrained to minimum ages of 40 and 61 mya, respectively. Despite the penalty function available in ver. 1.7 of r8s which is intended to correct for the tendency of the program to perform poorly when estimating the ages of deeper nodes from shallow calibrations, it was necessary to constrain the maximum age of the root of the tree. Not doing so would place the origin of Euphorbiaceae in the Precambrian. Although the 114 mya estimate of the origin of Euphorbiaceae used here is older than the 69–71 mya estimate of Wikström et al. (2001), it was in fact influenced by their age estimates given that Davis et al. (2005) constrained the basal node of their phylogeny with the youngest (and therefore most conservative) estimate for that age from Wikström et al. (2001).

Estimates for the divergence of *Croton* from its nearest sister group, *Brasiliocroton*, as well as estimates of the divergence of the core *Croton* and *Moacrotan* clades, were obtained from the transformed chronogram of the two-gene data set. These estimates were then used as the constraints for transforming Fig. 2 into ultrametric form using PL in r8s. The estimate of the divergence of *Croton* from *Brasiliocroton* was used as the maximum age constraint of the root. The estimates of the crown ages of the core *Croton* and the *Moacrotan* clades were used as minimum age constraints for these nodes, much as fossil constraints are used.

Results

Phylogenetic analyses

In the four-gene data set, the ITS, *trnL-trnF*, *ndhF*, and *rbcL* partitions have aligned lengths of 762, 1,504, 1,969, and 1,398, with 229 (30%), 125 (8%), 187 (9%), and 83

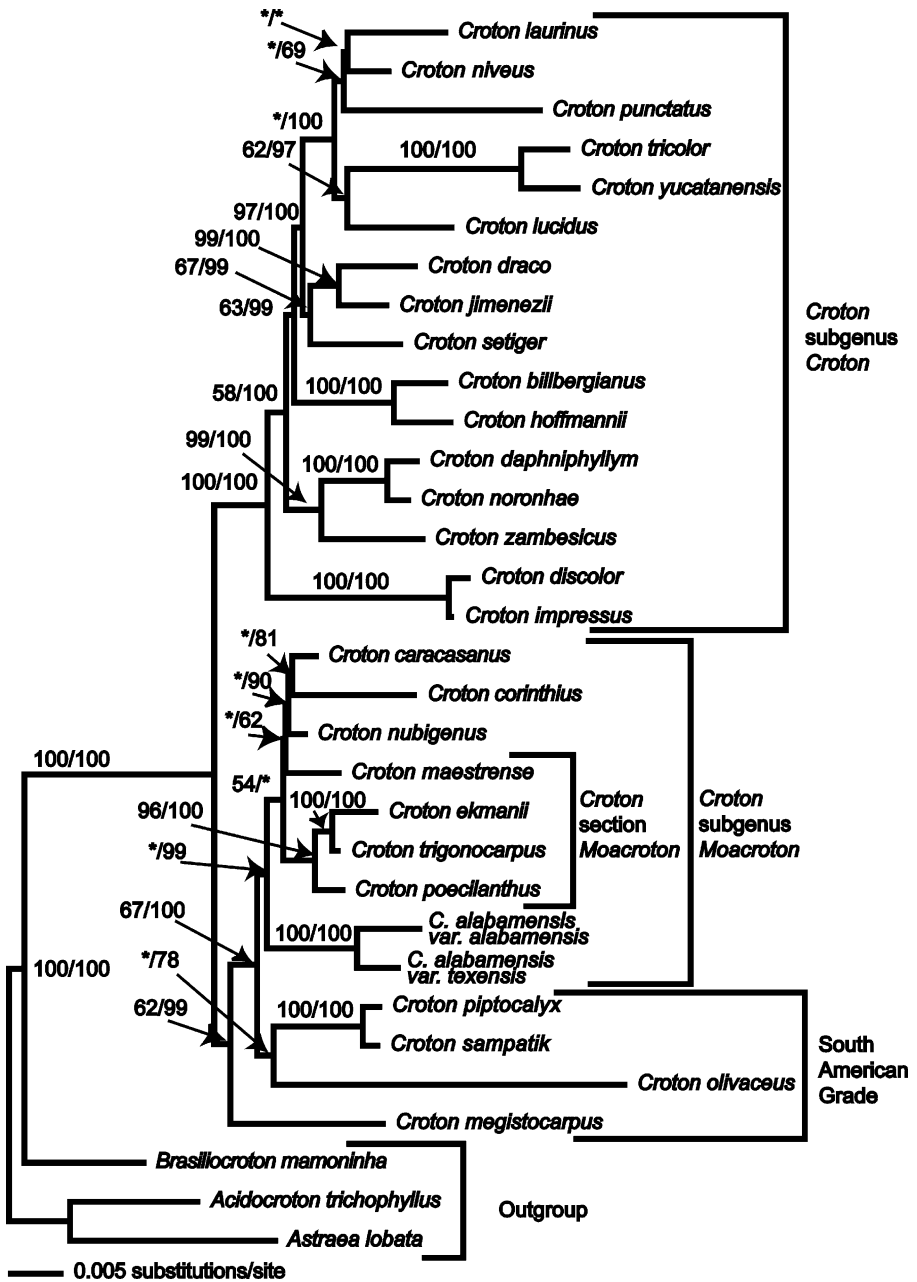


Fig. 3 Maximum likelihood phylogram of the combined chloroplast sequence (*rbcl*, *ndhF*, and *trnL-trnF*) analysis. Branches are drawn in proportion to the substitutions per site assuming the GTR+I+G model. The numbers on the nodes represent maximum parsimony bootstrap values followed by Bayesian clade credibility values. The parsimony and Bayesian topologies are not identical to the ML topology. Nodes that are either not found on the parsimony bootstrap tree or the Bayesian consensus tree, or have less than 50% bootstrap support are labeled with an asterisk (*)

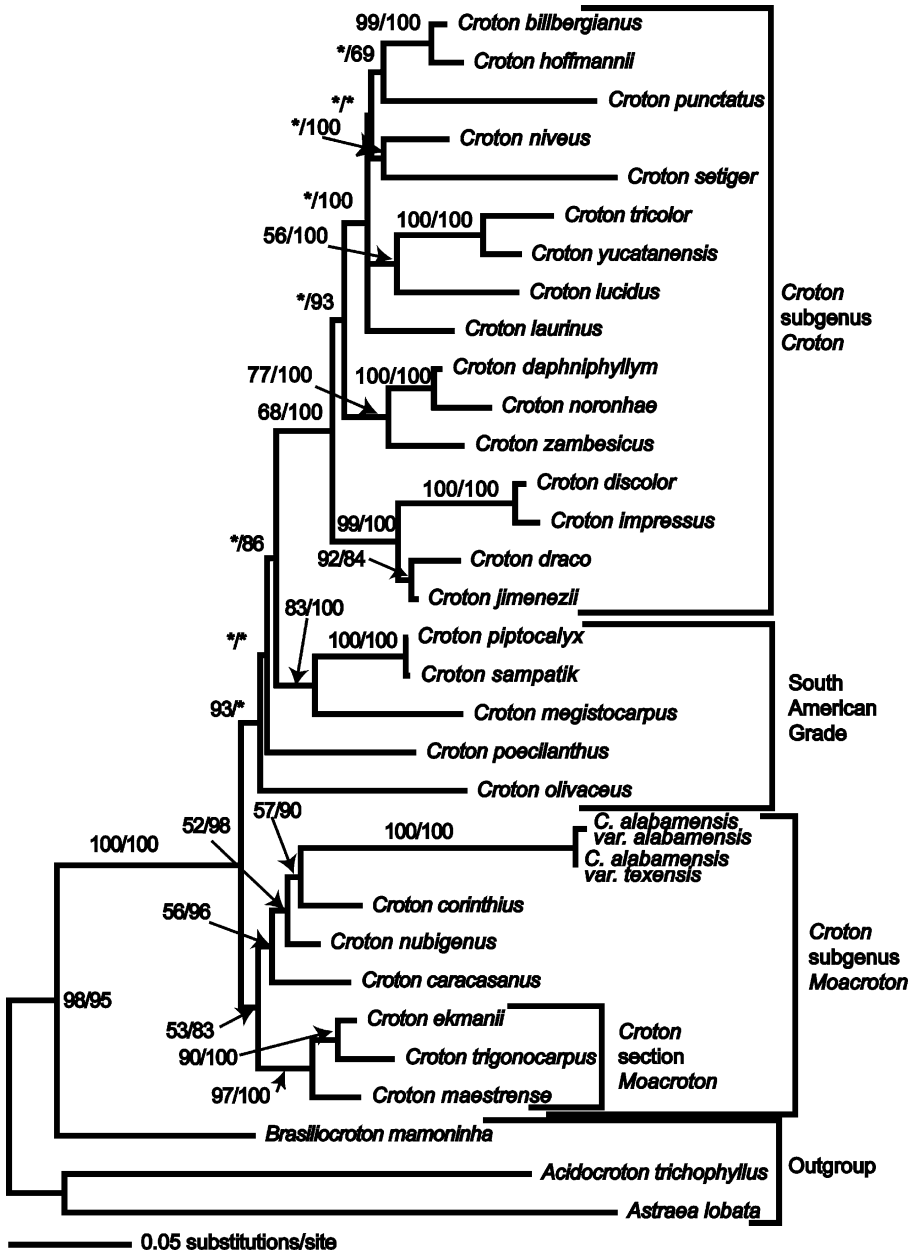


Fig. 4 Maximum likelihood phylogram of the ITS sequence analysis. Branches are drawn in proportion to the substitutions per site assuming the GTR+I+G model. The numbers on the nodes represent maximum parsimony bootstrap values followed by Bayesian clade credibility values. The parsimony and Bayesian topologies are not identical to the ML topology. Nodes that are either not found on the parsimony bootstrap tree or the Bayesian consensus tree, or have less than 50% bootstrap support are labeled with an asterisk (*)

(6%) parsimony-informative sites respectively. The first ILD test for incongruence among the three chloroplast regions (*trnL-trnF*, *ndhF*, and *rbcL*) yielded a P value of 0.14, failing to reject the hypothesis that the three regions are drawn from a common population of characters. The second ILD test, with the data partitioned between the nuclear ITS region and the three chloroplast genes rejected the hypothesis of homogeneity ($P=0.01$). Repeating the test after removing *C. setiger*, which was the taxon found to be mainly responsible for incongruence in the Berry et al. (2005) combined data matrix and *C. poecilanthus* still yielded $P=0.01$. A succession of ILD tests with an increasing number of taxa removed all yielded $P=0.01$.

The same alignment as in the four-gene data set was used in the three-gene data set for the ITS, *trnL-trnF*, and *ndhF* sequences, which are of the same aligned length in both data sets. In the three-gene analyses, the ITS, *trnL-trnF*, and *ndhF* sequences have 246 (31%), 149 (10%), and 200 (10%) parsimony-informative sites, respectively. An ILD test with the data partitioned between ITS and the two chloroplast regions (*trnL-trnF* and *ndhF*) also found conflict between the partitions ($P=0.01$). However, an ILD test of the two chloroplast regions (*trnL-trnF* and *ndhF*) found no incongruence ($P=0.52$).

Parsimony analyses of the combined four-gene data set, excluding uninformative characters, produced two most parsimonious trees of 1,619 steps each with a consistency index (CI) of 0.532 and a retention index (RI) of 0.670 (trees not shown). The two most parsimonious differed only in the placement of *C. punctatus* sister to *C. laurinus* and *C. niveus* in one, and sister to *C. tricolor*, *C. yucatanensis*, and *C. lucidus* in the other.

Both combined and separate maximum likelihood analyses implemented under the selected models of the four-gene data set returned single trees (Figs. 2, 3 and 4). The topology differs between the combined analysis (Fig. 2) and the chloroplast-only analysis (Fig. 3) in that the combined analysis places *C. poecilanthus* sister to all of the *Moacroton* clade. The chloroplast-only analysis (Fig. 3) places *C. poecilanthus* embedded within the *Moacroton* clade. The ITS-only analysis (Fig. 4) differs from the combined and chloroplast-only analyses in the attachment point of the root of *Croton* and the position of *C. poecilanthus*. The ITS-only analysis roots *Croton* with the South American taxa (*C. olivaceus*, *C. megistocarpus*, *C. piptocalyx*, and *C. sampatik*), along with *C. poecilanthus*, as a grade leading up the core *Croton* clade rather than as a grade at the base of the *Moacroton* clade. Templeton tests of the chloroplast (which is most similar to that found by the combined analysis) and nuclear ITS topologies, given both the chloroplast and ITS data sets, found them to be significantly different ($P<0.0001$).

The Bayesian MCMC runs of the four-gene data set resulted in four sets of 20,001 trees each. A visual examination of the likelihood-by-generation plot of each of these four sets of 20,001 trees revealed that by the 100,000th generation (2,000 retained trees) stability had been reached in each run. The first 2,000 trees were discarded as the burn-in. The four MCMC chains of the three-gene data set yielded four files of 10,001 trees each. Of the 40,004 trees found in the second Bayesian analysis, 36,804 were kept, with the first 800 trees from each file (40,000 generations) discarded as the burn-in. The consensus trees from each of the different runs were compared among those from the same data set. All four MCMC runs of each of the Bayesian analyses resulted in consensus trees that were topologically the same as the others of the same data set. Given this, we conclude that the different

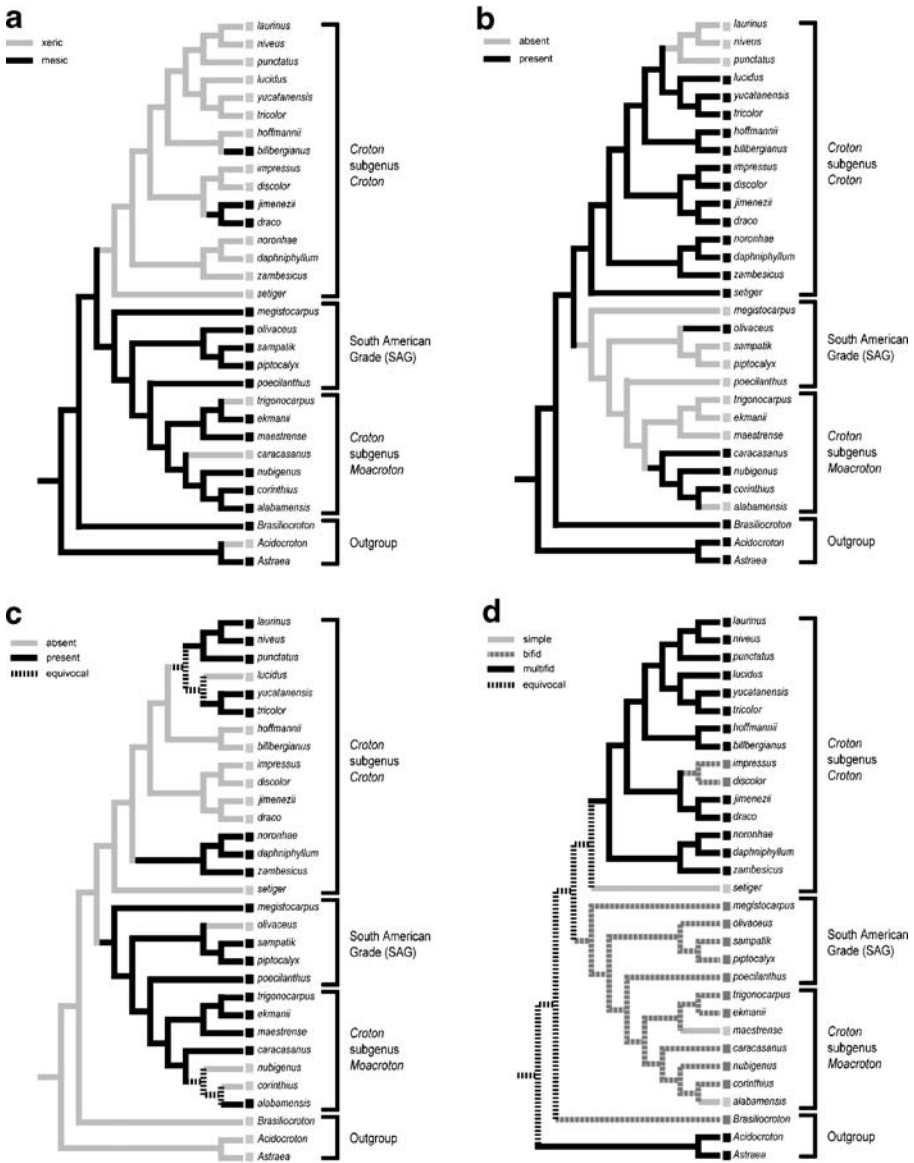


Fig. 5 a–d Patterns of ecological and morphological characters mapped under the assumption of ACCTRAN onto the topology of the majority rule consensus tree of three trees: one of two most parsimonious trees, the maximum likelihood tree, and the Bayesian consensus tree obtained from the analyses of ITS, *rbcL*, *ndhF*, and *trnL-trnF*. **a** Habitat: xeric, mesic. **b** Stellate trichomes: absent, present. **c** Lepidote trichomes: absent, present. **d** Stigma type: simple, bifid, multifid. Note: one accession of *C. alabamensis* was removed to simplify the figure. **e–h**: **e** Stamen number: six or less, ten or more. **f** Petals on pistillate flowers: absent, present. **g** Sepal type of pistillate flowers: imbricate, valvate, connate at the base (not valvate), uncertain. **h** Bisexual cymules: absent, present

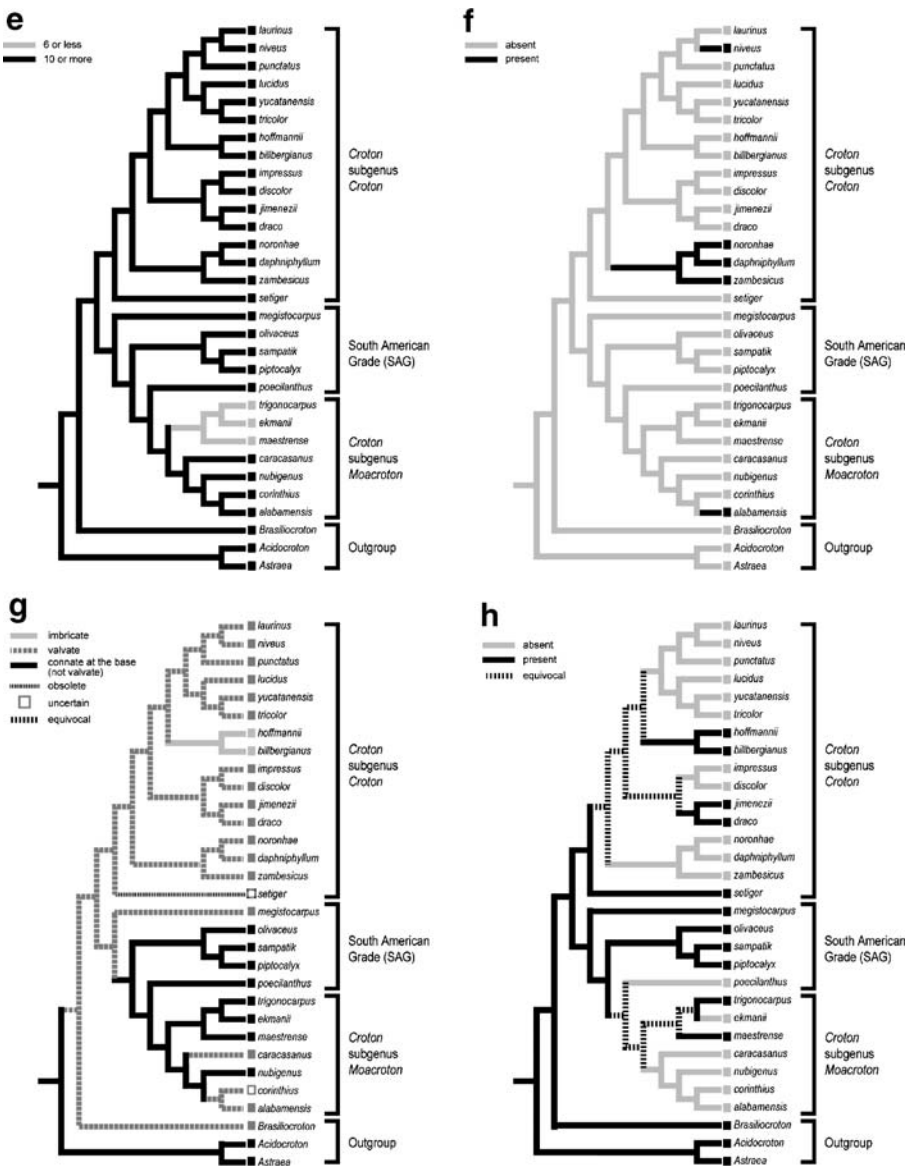


Fig. 5 Continued.

runs converged on the same set of trees and are combinable. The post burn-in trees from the four runs in each analysis were pooled to calculate the majority rule consensus tree and estimates of Bayesian clade credibility.

A fully resolved majority rule consensus tree of the ML, MP, and Bayesian phylogenies of the four-gene data set was produced to summarize the results obtained from the different analysis methods. This cladogram represents our preferred phylogenetic hypothesis and is the one on which the morphological characters are mapped (Fig. 5).

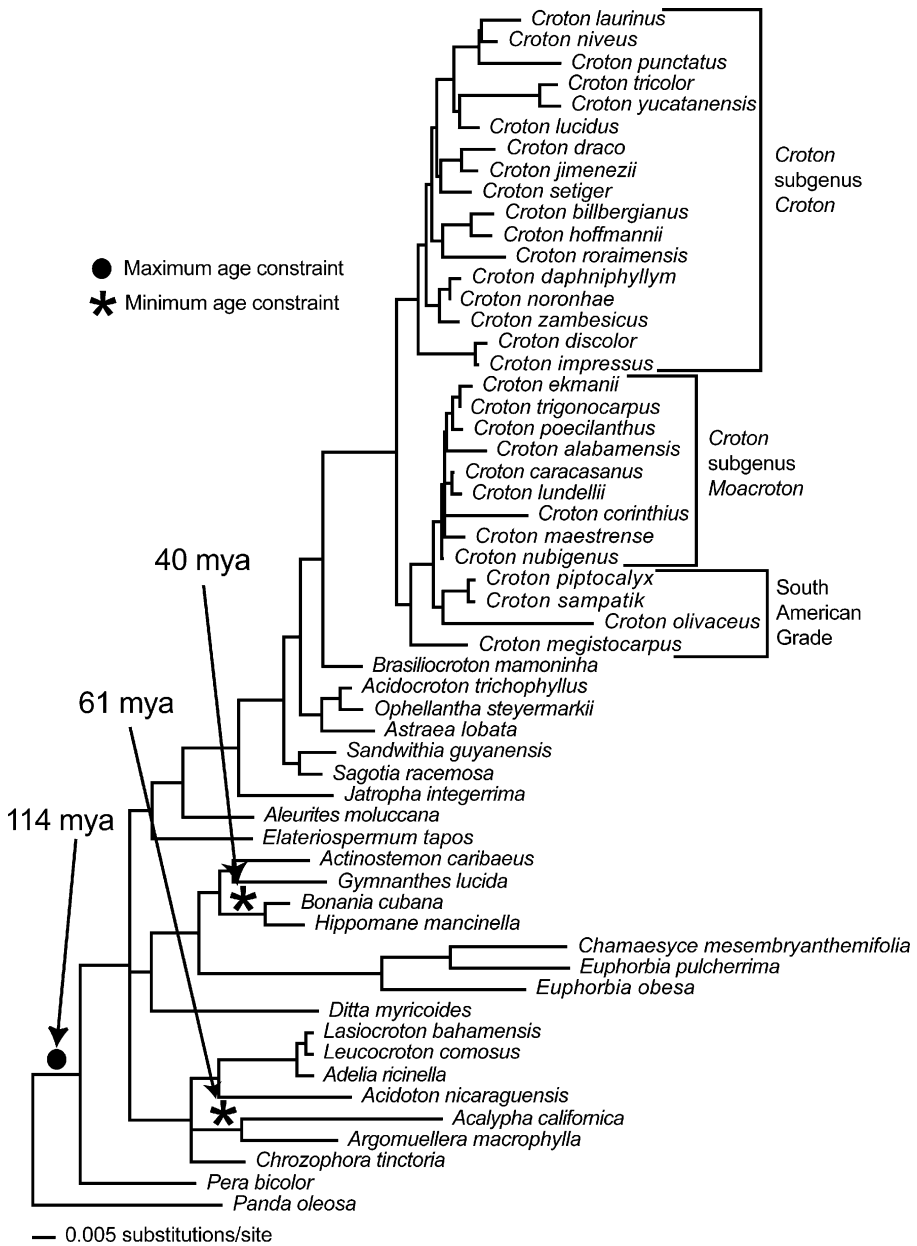


Fig. 6 Bayesian phylogram of combined *rbcL* and *trnL-trnF* analysis of Euphorbiaceae genera. Branch lengths are drawn in proportion to the estimated number of substitutions per site and represent an average of the branch length of all trees sampled in the markov chain that have that branch. The estimate of the age of Euphorbiaceae of Davis et al. (2005) is used as a maximum age constraint and fossils are used as minimum constraints

Table 3 Ages in Millions of Years Before Present Estimated by Penalized Likelihood and Nonparametric Rate Smoothing of the Bayesian Phylograms of Combined and Separate analyses

	Bayesian					
	Penalized likelihood			Nonparametric rate smoothing		
	Combined	<i>rbcL</i>	<i>trnL-F</i>	Combined	<i>rbcL</i>	<i>trnL-F</i>
<i>Croton</i> stem	<i>54.11</i>	51.13	56.72	52.70	53.91	57.90
<i>Croton</i> crown	39.29	38.60	42.23	37.86	41.95	43.64
<i>Moacrotan</i> clade crown	35.22	33.70	39.00	33.85	37.01	40.43
Core <i>Croton</i> clade crown	<i>33.26</i>	33.92	39.24	32.06	37.63	40.65
Acalyphoid fossil	61.00 ^a	94.93	61.00 ^a	61.00 ^a	94.82	70.68
<i>Crepetocarpon</i> fossil	40.00 ^a	40.00 ^a	40.00 ^a	40.01	40.00 ^a	40.02

Values with a superscripted a (^a) are those that were constrained as minimum ages. The estimates in italics are those used as constraints in the estimates of divergence times in the *Croton* data set

In the two-gene data set, the *rbcL* and *trnL-trnF* sequences have aligned lengths of 1,398 and 1,504 bases, and 183 (13%) and 397 (26%) parsimony-informative sites respectively. The ILD test conducted on the data matrix rejects ($P=0.001$) the hypothesis of homogeneity between the data partitions, despite both gene regions being from the chloroplast genome. Wurdack et al. (2005), from whom most of the non-*Croton* *rbcL* sequence data used here was obtained, analyzed the two regions separately and in combination. Our separate and combined analyses, while made up of a much smaller sampling, are consistent with the topologies obtained from their separate and combined parsimony analyses as depicted in Figs. 1, 2 and 3 in Wurdack et al. (2005). The consensus tree of all post-burn-in trees visited in the Bayesian Markov chains in the combined analysis of *rbcL* and *trnL-trnF* is shown in Fig. 6. The lengths of individual branches depicted on this phylogram represent the averages of the length of that branch on all post-burn-in trees that have that branch. The results of the clade-age estimates from the separate and combined analyses using Bayesian and maximum likelihood are summarized and compared in Tables 3 and 4. The estimates from the combined Bayesian analyses obtained by PL (italicized in Table 3) were chosen as the constraints for transforming Fig. 2.

Table 4 Ages in millions of years before present estimated by penalized likelihood and nonparametric rate smoothing of the maximum likelihood phylograms of combined and separate analyses

	Likelihood					
	Penalized likelihood			Nonparametric rate smoothing		
	Combined	<i>rbcL</i>	<i>trnL-F</i>	Combined	<i>rbcL</i>	<i>trnL-F</i>
<i>Croton</i> stem	55.05	54.13	60.90	53.66	55.28	59.11
<i>Croton</i> crown	40.10	39.03	45.79	38.78	40.62	44.06
<i>Moacrotan</i> clade crown	36.91	33.84	42.54	35.63	35.38	40.91
Core <i>Croton</i> clade crown	32.23	39.03	38.61	31.36	37.78	37.10
Acalyphoid fossil	61.00 ^a	62.16	61.00 ^a	61.02	62.26	61.00 ^a
<i>Crepetocarpon</i> fossil	40.00 ^a	40.00 ^a	40.00 ^a	40.01	10.01	40.02

Values with a superscripted a (^a) are those that were constrained as minimum ages

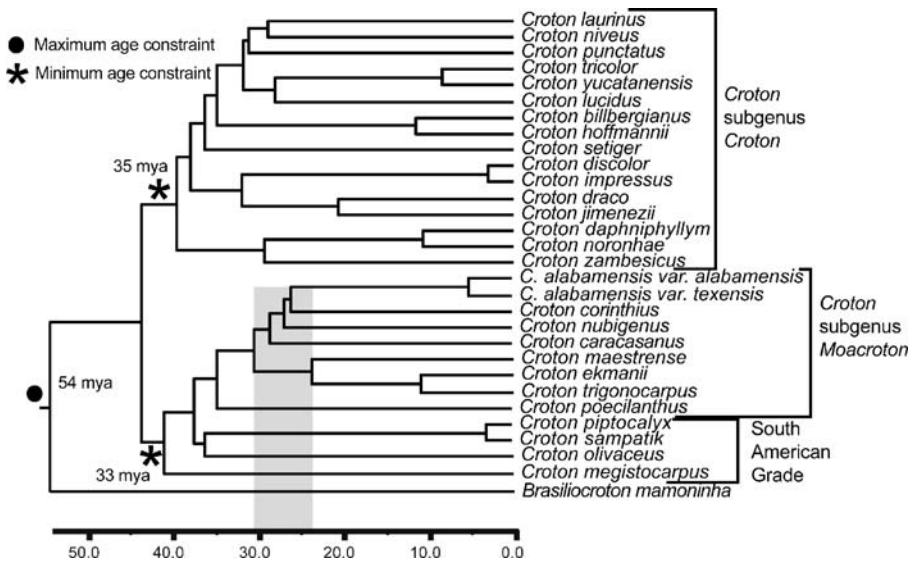


Fig. 7 Chronographic form of Fig. 2 smoothed used penalized likelihood. *Acidocroton trichophyllus* and *Astraea lobata* were removed before the phylogram was smoothed

Character Evolution

The majority of bifurcations in the phylogeny occur early-on, with the bulk of branch length leading up to the terminal taxa (Fig. 7), and six of the eight mapped characters are recovered as synapomorphies (Fig. 5). Our opinion is that this scenario may favor reversals over parallelisms in the characters in question and we therefore selected to show the reconstructions based on accelerated transformation (ACCTRAN) which maximizes reversals (Maddison & Maddison, 2000). The characters were also mapped using delayed transformation (DELTRAN) and the differences in the reconstructions were minimal (results not shown). Mapping the habitat type onto the consensus cladogram (Fig. 5) indicates that a mesic habitat is plesiomorphic in *Croton* (Fig. 5a). Even with the alternative topology indicated by the ITS-only analysis (Fig. 4), a mesic habitat would be found to be plesiomorphic. Although some groups within the core *Croton* clade possess lepidote scales and some in the *Moacrotan* clade have stellate trichomes, the preponderance of stellate trichomes is a synapomorphy for the core *Croton* clade and lepidote scales for the *Moacrotan* clade (Fig. 5b and c). The presence of multifid styles is a synapomorphy for the core *Croton* clade (with some reversals), with all of the members of the *Moacrotan* clade and the South American grade (SAG) having either bifid or simple styles (Fig. 5d). Sepals that are connate at the base in pistillate flowers is another synapomorphy, albeit with exceptions, for the *Moacrotan* clade + SAG (Fig. 5g). The remaining main characters used by Webster (1993) that were mapped on to the phylogeny are either synapomorphic for only a small group of species, such as stamen number (Fig. 5e) and the presence/absence of petals in pistillate flowers (Fig. 5f), or else homoplasious across the whole genus as in the case of the presence/absence of bisexual cymules (Fig. 5h).

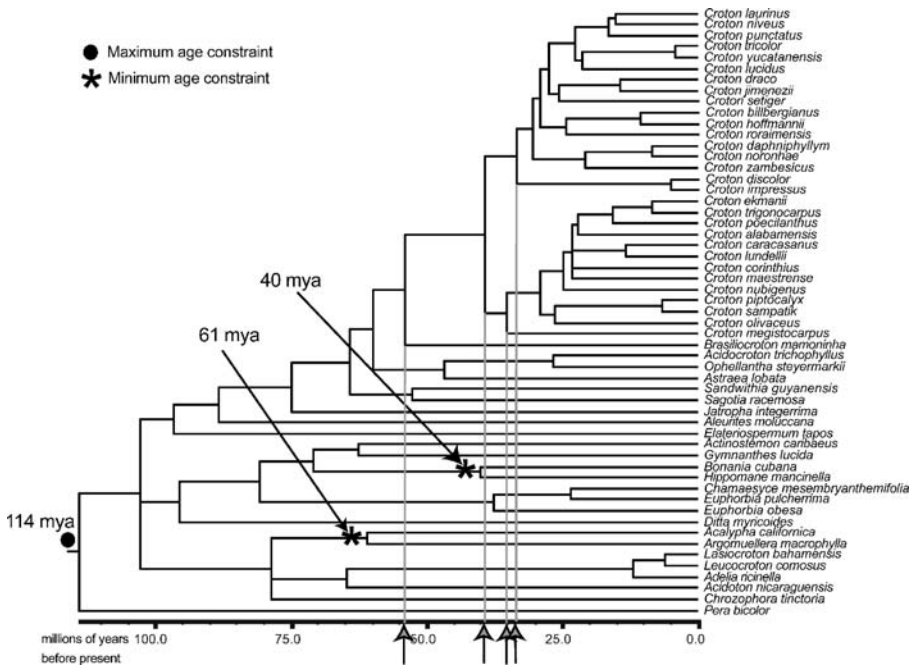


Fig. 8 Chronographic form of Fig. 6 smoothed used penalized likelihood. The nodes that were constrained are indicated along with the estimates of the stem and crown ages of *Croton* and its major clades. *Panda oleosa* was removed before the phylogram was smoothed

Divergence Time Estimates

Using the selected models, none of the combined or separate *rbcL* and *trnL-trnF* data sets, nor the first and second or third codon positions of *rbcL*, evolve in a clock-like fashion ($P < 0.05$), so a simple molecular clock cannot be used to date nodes within the phylogeny. The low smoothing parameter (100) for penalized likelihood found through cross-validation suggests that the data are far from clocklike (Sanderson, 2002).

The combined *rbcL* and *trnL-trnF* Bayesian phylogram (Fig. 6) was converted into a chronogram using penalized likelihood (Fig. 8). By constraining the maximum age of Euphorbiaceae to be no older than 114 million years ago (mya), the split between *Brasiliocroton* and *Croton* is estimated at approximately 55 mya. Davis et al. (2005) estimated the divergence of *Croton* from *Euphorbia*, which was the nearest relative to it that they included, at 65 mya. In contrast, the age of the node joining these two on Fig. 8 is estimated at 102 mya compared to the 69–71 mya estimated by Wikström et al. (2001). Considering that the reliable Euphorbiaceae fossils are 40 and 61 mya, we feel that our age estimates of the divergence of *Croton* from *Brasiliocroton* are realistic.

The divergence time estimates for *Croton* and its two major clades were estimated from both the ML and Bayesian phylograms. Results indicate that using *rbcL*, *trnL-trnF*, or both in combination, using either penalized likelihood or non-parametric rate smoothing, or smoothing a Bayesian or likelihood phylogram all recover highly similar estimates of divergence times (Tables 3 and 4). Although the ILD test found

the *rbcL* and *trnL-trnF* gene regions to be incongruent ($P=0.001$), they both recover similar topologies and comparable divergence estimates when analyzed separately. Given this, the calibrations from the combined *rbcL* and *trnL-trnF* Bayesian phylogram were chosen for use in the second step of converting the *Croton* phylogeny to ultrametric form (see the italicized figures in Table 3).

Discussion

The Major Clades Within *Croton*

Since the publication of the first molecular phylogeny of *Croton* (Berry et al., 2005), more than 500 *Croton* species have been sequenced for ITS (B. van Ee, P. E. Berry, and R. Riina, unpublished data). With the exception of the species focused on in this work, all of the species sampled to date fall into the core *Croton* clade, and the Old World species still constitute a monophyletic group within that clade. The species outside of the core *Croton* clade fall either in the well-supported *Moacrotan* clade or in the South American grade (SAG).

The first split in the phylogeny of *Croton* divides the genus into two disproportionately sized groups. Separate analyses of the nuclear ITS and chloroplast sequence data recover a different position for the South American grade (SAG) species (Figs. 3 and 4) relative to this first split. In all analyses of ITS alone, with *C. poecilanthus* included, the SAG is found at the base of the core *Croton* clade and includes *C. poecilanthus*. Excluding *C. poecilanthus* from the ITS analyses yields the same topology in the likelihood and Bayesian searches, but in the parsimony searches the SAG is attached at the base of the *Moacrotan* clade. In all of the chloroplast analyses, the SAG is likewise found at the base of the *Moacrotan* clade, and *C. poecilanthus* is placed within *Moacrotan*. With the current sampling, which covers the full phylogenetic breadth of *Croton* given our current knowledge of the genus (Berry et al. 2005), and over 5,600 base pairs of data, the relationships of the SAG species are not resolved well enough to decisively include them in either the core *Croton* or *Moacrotan* clades. Given that this is the same issue as determining the rooting of *Croton*, and the earliest branching patterns within the genus, it remains an important issue to pursue in future research. However, we feel that the balance of the evidence—sequence data, ecology, and morphology—better supports the topology recovered in the combined analysis in which the South American grade (SAG) belongs at the base of the *Moacrotan* clade (Fig. 2).

We have chosen the topology recovered by the combined analyses of the ITS and chloroplast sequence data as our preferred phylogenetic reconstruction of the genus (Fig. 2). This topology is largely concordant with that recovered by the chloroplast analysis (Fig. 3). In the analyses of Berry et al. (2005), in which three lineages from outside of the core *Croton* clade were included (*C. olivaceus*, *C. alabamensis*, and *Moacrotan revolutus*), the same topology as we favor here was recovered in all but the parsimony analysis of ITS. The ecological and morphological characters (Fig. 5a–h) map with fewer steps onto the topology in which the SAG appears at the base of the *Moacrotan* clade.

Data Set Incongruence

The incongruence length difference (ILD) tests detected incongruence between the nuclear and chloroplast partitions of the three- and four-gene data sets. Berry et al. (2005) were able to establish congruence between ITS and *trnL-trnF* in their sampling of 88 taxa, however, they included only three taxa from outside of the core *Croton* clade. Following the strategy proposed by Wiens (1998) we present both combined and separate ITS and chloroplast analyses and discuss the differences. The main differences in the topologies recovered by the nuclear and chloroplast analyses (Figs. 3 and 4) are the placement of the SAG and the position of *C. poecilanthus* (discussed above and below). The relative position of the SAG, which is also a question of the rooting of *Croton*, remains a problematic area in the phylogeny *Croton*. For this reason we are not making taxonomic changes that apply to the SAG taxa at this time. The biogeographic implications of the topology recovered in the ITS analysis (Fig. 4) does not change the implication of a South American origin for *Croton* given the South American distribution of the outgroups.

Character Evolution

The *Moacroton* clade and the South American grade (SAG) have never been recognized as groups prior to the use of molecular evidence. The reconstructions of the morphological characters indicate that the *Moacroton* clade plus the SAG is a morphologically and ecologically cohesive group generally characterized by a primarily lepidote indumentum, bifid or simple styles, pistillate flowers with sepals connate at the base, and a preference for mesic habitats, although there are some exceptions (Fig. 5a–h).

Although Fig. 9 includes an approximately equal numerical sampling of accessions from the core *Croton* clade and the *Moacroton* clade plus the South American grade (SAG), it should be noted that it is a fairly exhaustive sampling of the *Moacroton* clade and the SAG while over 1,200 species are suspected of belonging to the core *Croton* clade. The estimated age of the divergence of the core *Croton* clade from the *Moacroton* clade (Fig. 8) is shortly after the Cretaceous/Tertiary (K/T) boundary approximately 65 million years ago (Pope et al., 1998). Evidence suggests that the massive extinctions at the K/T boundary were the result of a meteorite impact causing a global cloud of dust that blocked sunlight and caused a freezing blackout (Alvarez et al., 1980), and had the long-term ecological effect of an overall drying in North and Central America (Ocampo et al., 2006). Specific morphological traits are not readily apparent as potential key innovations that contributed to the great disparity in diversification between these two main clades of *Croton*. However, the adaptation to a primarily xeric habitat in the core *Croton* clade, compared to a preference for the plesiomorphic mesic habitat in the *Moacroton* clade, appears to be a likely explanation for its much greater diversification given the period of warmer, drier climate starting about 65 mya (Zachos et al., 2001).

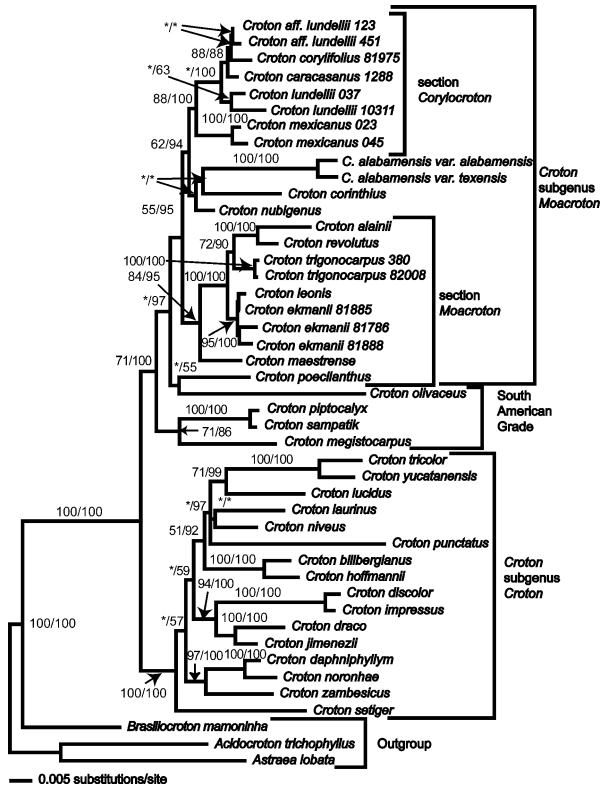


Fig. 9 Maximum likelihood phylogram of the combined *ndhF*, *trnL-trnF*, and ITS data set analysis of 45 taxa. Branches are drawn in proportion to the substitutions per site assuming the GTR+I+G model. The numbers on the nodes represent maximum parsimony bootstrap values followed by Bayesian clade credibility values. The parsimony and Bayesian topologies are not identical to the ML topology. Nodes that are either not found on the parsimony bootstrap tree or the Bayesian consensus tree, or have less than 50% bootstrap support are labeled with an asterisk (*)

Biogeographic Reconstructions

The estimated age of *Croton* is approximately 55 million years, and the estimated divergence of the Old World clade from its New World sister clade is 40 mya (Fig. 8). These dates are much younger than the latest estimated connection between South America and Africa at approximately 105 mya (Pitman et al., 1993). This argues against the current worldwide distribution of *Croton* being a Gondwanan disjunction. Instead, a possible route by which *Croton* may have reached the Old World is via the Eocene North Atlantic land bridge proposed by Tiffney (1985). Davis et al. (2002) proposed a similar migration from South America to the Old World via a Laurasian migration route for Malpighiaceae, another primarily tropical group. Although the biogeographic connection between Old and New World *Croton* is not the subject of this work it is a related issue given that a migration from South America to the Old World via a high latitude connection requires a hypothesis for how the group reached North and Central America from South America, which is the biogeographic focus of this work.

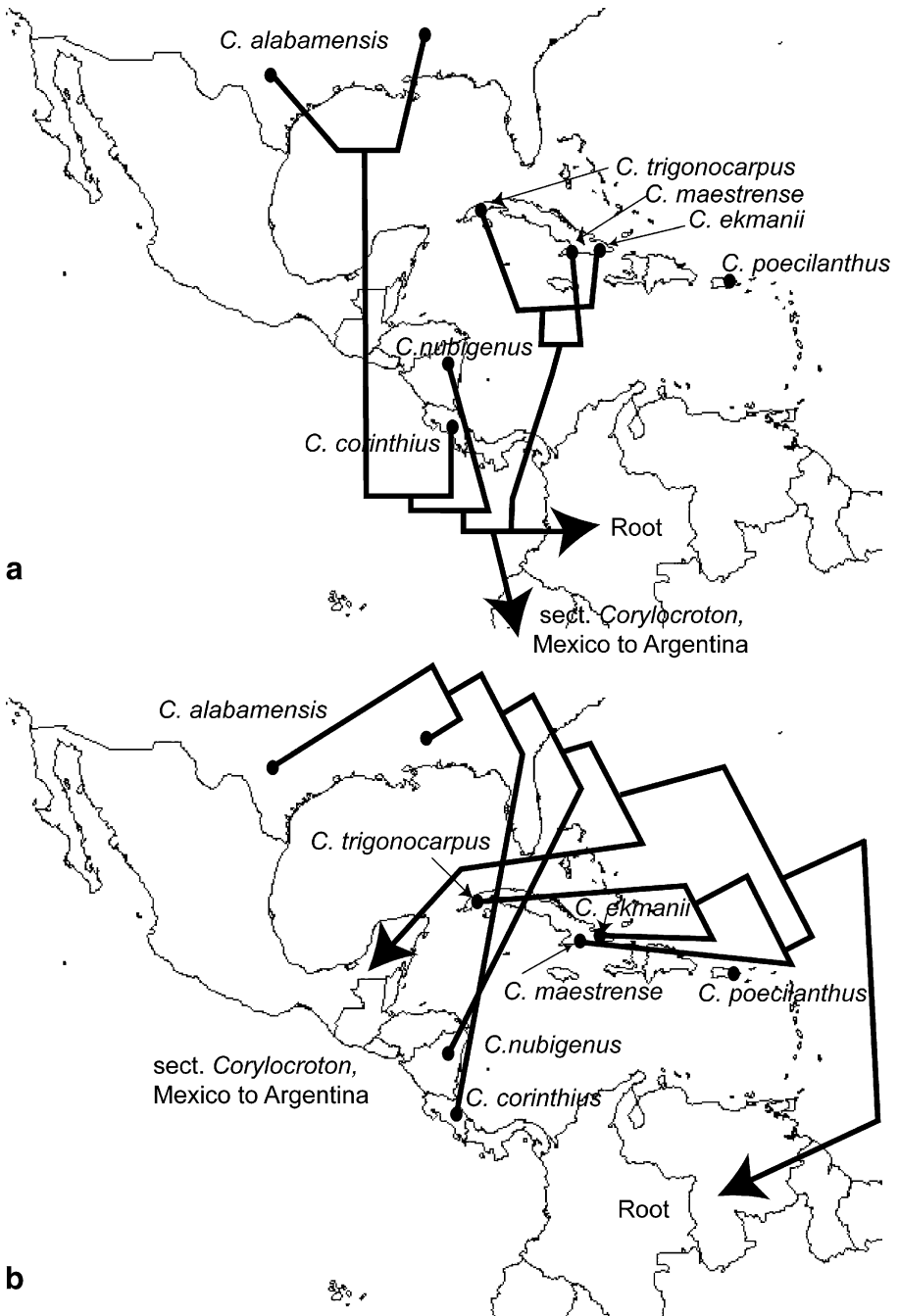


Fig. 10 **a** (Top), **b** (bottom). Phylogenetic tree for subgenus *Moacroton* approximated as biogeographic hypotheses onto a map of the greater Caribbean region. *Croton poecilanthus* was omitted given its incongruent position between the nuclear and chloroplast sequence data

The primarily Caribbean distribution of the *Moacrotan* clade provides an opportunity to examine the historical biogeography of this geologically complex region. The historical biogeography of the Caribbean region has been contentiously debated, with most of the debate concerning the relative roles of dispersal vs. vicariance, and the related issue of if, and when, there were geological connections between given land masses. *Croton* fruits consist of tripartite capsules that dehisce explosively when fully ripe; dispersing the oblong, carunculate seeds no more than a few meters. Additional dispersal by ants has been reported (Passos & Ferreira, 1996). The use of *Croton* seeds as food by birds has also been documented (Downhower & Racine, 1976), and a common name for some *Croton* species is “dove weed” given that their seeds appear to be a favorite food of doves. In a study of tropical pioneer trees in Panama, Dalling et al. (1998) found that the distribution of *Croton billbergianus* seedlings was markedly skewed towards conspecific adults, indicating limited dispersal. Although long-distance dispersal over water barriers cannot be ruled out, and undoubtedly it has played a role given the presence of some sections of the core *Croton* clade on various Caribbean islands, the limited dispersal capabilities of *Croton* suggest that there are significant barriers to gene flow between distant islands.

We have identified two main biogeographical hypotheses for the distribution of the *Moacrotan* clade in the region, although these are by no means the only alternatives (Fig. 10). The first is the Paleogene arc hypothesis in which an island arc consisting of pieces of what went on to become Cuba and Hispaniola may have been connected to South and Central America up until about 49 mya (Pitman et al., 1993; Chakrabarty, 2006). Another scenario is the Greater Antilles + Aves Ridge (GAARlandia) hypothesis, in which an exposed Aves Ridge may have connected South America to the Greater Antilles during the Eocene–Oligocene transition 35–33 mya (Iturralde-Vinent & MacPhee, 1999).

Within the *Moacrotan* clade, two main divisions can be identified. One consists of a strictly Greater Antilles lineage, which includes the members of the former genera *Moacrotan* and *Cubacrotan*, and *C. poecilanthus*, and the other of the mainland lineages, including *C. alabamensis*, *C. corinthius*, *C. nubigenus*, and *Croton* section *Corylocroton*. *Croton alabamensis* occurs in just two small areas in the southern USA (van Ee et al., 2006; Wurdack, 2006). *Croton corinthius* and *C. nubigenus* have restricted ranges in wet tropical forests in Costa Rica and Nicaragua, respectively. *Croton* section *Corylocroton* has a widespread distribution from Mexico to South America and includes at least one species in the Antilles, *C. corylifolius*, and at least two species from South America, *C. beetlei* and *C. caracasanus*. The pattern of diversification of section *Corylocroton*, in which the first split in the phylogeny divides the group into a Mexican clade (represented by *C. mexicanus*) sister to a widespread clade, appears to indicate a Central American/Mexican origin for the section rather than either a South American or Antillean origin (Fig. 9).

Iturralde-Vinent and MacPhee (1999) proposed a reconstruction of the Caribbean region with an exposed Aves Ridge creating a land bridge that connected the Guiana Shield in northern South America to the Greater Antilles during the Eocene–Oligocene transition 35–33 million years ago (see also Borhidi, 1985, MacPhee & Iturralde-Vinent, 1994, 1995 for more arguments for this land bridge). They called this the Greater Antilles + Aves Ridge (GAARlandia) hypothesis. Dávalos (2004) evaluated whether the GAARlandia hypothesis could account for the distribution

patterns of mammals in the Caribbean, and she concluded that many groups including primates and hystricognath rodents were consistent with the model, while others such as choleopodine sloths, insectivorans, and natalids cannot be reconciled with this model and thus require an alternative explanation.

Our chronogram (Fig. 7) shows that the timing (35 to 24 mya ago) and order of divergence of the lineages of the *Moacroton* clade are at least consistent with the GAARlandia hypothesis. The stem age of the clade of species endemic to Cuba is estimated at 30 mya (Fig. 7). Either the land bridge existed for slightly longer than until about 32 mya, as proposed, or else the calibration of our chronogram is off by a few million years, either of which seems possible. *Croton trigonocarpus* from far western Cuba is reconstructed as diverging from its relatives in eastern Cuba much later, approximately 12 mya (Fig. 7), which is consistent with the GAARlandia model in which the Havana-Matanzas Channel separated western Cuba from central Cuba until about 14 mya. The remaining Central and North American members of the *Moacroton* clade (*C. nubigenus*, *C. corinthius*, *C. alabamensis*, and *Croton* section *Corylocroton*), or their ancestors, may have arrived on the mainland via over water dispersal from the GAARlandia bridge, which requires invoking extinction on Cuba and/or other portions of the bridge.

In the Paleogene arc hypothesis, members of the *Moacroton* clade could have reached Central America via an arc of islands roughly in the current position of Panama. This island arc is hypothesized as later rafting eastward to form Cuba and Hispaniola (Iturralde-Vinent & MacPhee, 1999; Kerr et al., 1999). Pitman et al. (1993) suggest that the connection between this arc and Central America lasted until approximately 50 million years ago, which would have made it possible for *Croton* to use the connection shortly after diverging, which our estimates place at approximately 55 million years ago (Fig. 8). There is also evidence that an arc of volcanic islands existed in what is now central Panama as early as the late Paleocene (57 mya; Maury et al., 1995). This island arc may have become increasingly more emergent in the time leading up to the collision of Central and South America (Coates et al., 2004). Although the precision of the estimate of the early diversification of the *Moacroton* clade (approximately 30 mya on Fig. 7) may be questioned, it is clear that they originated well before the isthmus of Panama linked North and South America in the late Pliocene.

Although the Paleogene arc and GAARlandia hypotheses have been presented as alternatives to each other (Chakrabarty, 2006), both scenarios could account for the current distribution of the major clades of *Croton*. The high level of diversity and endemism of the core *Croton* clade in Central and North America, although not as high as in South America, suggests that the genus arrived there early in its evolution. The Paleogene arc could have provided the mechanism for dispersal out of South America until about 50 million years ago. However, the pattern and timing of the diversification of the *Moacroton* clade, which happened later (35–24 mya, Fig. 7), fits the GAARlandia model better.

Croton poecilanthus: A Possible Ancient Hybrid

The phylogenetic affinities inferred from the nuclear ITS and chloroplast gene regions of *Croton poecilanthus* are different (Figs. 3 and 4). For this reason, we also

analyzed the data with *C. poecilanthus* removed, which did not otherwise alter the recovered relationships.

Analyses of the chloroplast genes place *C. poecilanthus* inside *Moacrotan* sister to *C. trigonocarpus* and *C. ekmanii* with 96% bootstrap support and 100% Bayesian clade credibility (Fig. 3). Analyses of ITS place it in an unsupported position within the SAG between *C. olivaceus* and *C. megistocarpus*, two species from South and Central America, respectively (Fig. 4). Morphologically, *C. poecilanthus* is strikingly similar to *C. ekmanii*, which is the main reason why it was targeted for sampling here (Fig. 1). The leaves, trichomes, inflorescences, and pistillate flowers of *C. poecilanthus* and *C. ekmanii* are morphologically similar enough that they could easily be mistaken as being the same species. *Croton poecilanthus* differs morphologically from *C. ekmanii* in that it has a more typical number (25+) of stamens, with filaments present, in its staminate flowers. By comparison, the staminate flowers of *C. ekmanii* have six or fewer stamens, and the anthers are sessile. An explanation for the alternate placements of *C. poecilanthus* recovered in the separate chloroplast and nuclear analyses of the data could be that it is a hybrid between a species within *Moacrotan*, or its ancestor, and a South American species. The lack of a single species, or small clade, sister to *C. poecilanthus* in the ITS analyses leads us to believe that the hypothesized second hybrid parent is not included in the sampling, but that if it were it would be among the grade of South American species. The close morphological similarity between *C. poecilanthus* and *C. ekmanii* suggests that their close relationship recovered in the chloroplast analysis is not a result of horizontal gene transfer, such as chloroplast capture, but rather a genuine phylogenetic relationship.

In the GAARlandia hypothesis, the distribution of *C. ekmanii*, or its ancestor, could have overlapped along the land bridge with taxa extending up from South America. The current range of *C. poecilanthus*, which is found exclusively in cloud forests on Puerto Rico, is midway along the hypothetical bridge from Cuba (where *C. ekmanii* is endemic) to South America. The divergence of *C. poecilanthus* on the ITS phylogeny is estimated at approximately 37 to 36 mya, whereas from the chloroplast phylogeny it is estimated at 25 to 10 mya (Fig. 7). The age of the divergence of *C. poecilanthus* obtained from the combined analyses is 37 to 35 mya. Although a proper estimate of the phylogenetic position of *C. poecilanthus*, and hence its age of origin, is confounded by the incongruence between the nuclear and chloroplast data, it is clear that it is a comparatively ancient event, occurring longer than at least 10 million years ago, and quite possibly as far back as 35 to 33 mya, when the GAARlandia land bridge is hypothesized to have existed.

Future Directions

Ongoing molecular work on *Croton* is substantially increasing the number of species sampled in the genus, with over 500 species sequenced for at least one gene to date. This will allow us to develop a more comprehensive phylogenetic framework for the genus, which can serve to identify patterns of ecological and morphological characters that were not previously discernible (viz., Fig. 5). At the same time, a number of species with incongruent nuclear and chloroplast genomes have been identified, *C. poecilanthus* here, *C. draconoides* in Riina et al. (in prep.), *C.*

tenuilobus in van Ee et al. (in prep.), and *C. setiger* in Berry et al. (2005). These findings raise the possibility that evolution within the genus has not always been strictly dichotomous. To better resolve the rooting of the genus and the early branching patterns within *Croton*, future research should include more gene regions than those used to date, and possibly whole-genome studies.

New Names and Combinations

A. *Croton* subgenus *Moacroton* (Croizat) van Ee & P.E. Berry **stat. nov.**

Moacroton Croizat, J. Arnold Arbor. 26(2): 189–190 (1945). Type species: *Moacroton leonis* Croizat

The subgenus can be divided at this time into two sections and three additional species that each deserve sectional recognition, and will be treated as such in an upcoming publication. These are:

1. *Croton* section *Moacroton* (Croizat) van Ee & P.E. Berry **stat. nov.**

Moacroton Croizat, J. Arnold Arbor. 26(2): 189–190 (1945).
Cubacroton Alain, Candollea 17: 116 (1960).

The five species formerly in the Cuban endemic genus *Moacroton* Croizat are found exclusively in Cuba on serpentine soils (Borhidi, 1991). The group's distribution is disjunct between the eastern and western ends of the island (Borhidi, 1996). *Croton maestrense* (formerly *Cubacroton*) is from eastern Cuba in the Sierra Maestra mountains and is found on igneous substrates. Morphologically this group is characterized by coriaceous leaves, which typically turn a metallic hue when dried, and a reduced number and size of stamens. The staminate flowers of *C. maestrense* bear one or two stamens with 2–3 mm long filaments. The staminate flowers of the other sect. *Moacroton* species, except for *C. poecilanthus*, have between four and six stamens in which the anther is sessile on the receptacle of the flower.

Croton poecilanthus is included here as a member of sect. *Moacroton* given its close morphological similarity to *C. ekmanii* and its unambiguous placement within the group in analyses of the chloroplast data (Fig. 3). It differs morphologically from other members of the section by its numerous stamens (25+) with well-developed filaments. Like *C. maestrense*, *C. poecilanthus* is found on igneous substrates. Its alternative placements in the combined nuclear and chloroplast analyses (Fig. 2) and analyses of the nuclear ITS data (Fig. 4) are possibly due to an ancient hybridization event (see discussion above).

The only species of the section not sampled molecularly is *C. cristalensis*, which was not able to be collected.

Species Included:

– *Croton alainii* van Ee & P. E. Berry **nom. nov.**

Moacroton lanceolatus Alain, Contribuciones Ocasionales del Museo de Historia Natural del Colegio “De La Salle” 11: 4 (1952), non *Croton lanceolatus* Hornem. ex Didr., Videnskabelige Meddelelser fra Dansk Natur-

- historisk Forening i Kjøbenhavn 1857: 148 (1857), nec *Croton lanceolatus* Cav. Icon. 6: 38 (1800).
- *Croton cristalensis* Urb. [*Moacroton cristalensis* (Urb.) Croizat]
 - *Croton ekmanii* Urb. [*Moacroton ekmanii* (Urb.) Croizat, *M. gynopetalus* Borhidi, *M. tetramerus* Borhidi & O. Muñiz]
 - *Croton leonis* (Croizat) van Ee & P. E. Berry, **comb. nov.**
Moacroton leonis Croizat, Journal of the Arnold Arboretum 26(2): 190 (1945).
 - *Croton maestrense* (Alain) van Ee & P. E. Berry, **comb. nov.**
Cubacroton maestrense Alain, Candollea 17: 116 (1960). *Moacroton maestrense* (Alain) Radcl.-Sm., Gen. Euphorbiacearum 319 (2001).
 - *Croton poecilanthus* Urb.
 - *Croton revolutus* (Alain) van Ee & P. E. Berry, **comb. nov.**
Moacroton revolutus Alain, Contribuciones Ocasionales del Museo de Historia Natural del Colegio “De La Salle” 11: 3 (1952).
 - *Croton trigonocarpus* Griseb. [*Moacroton trigonocarpus* (Wright ex Griseb.) Croizat].

2. *Croton* section *Corylocroton* G.L. Webster, Taxon 42. 1993.

The circumscription of sect. *Corylocroton* given here represents more than a doubling of included species since the section was described by Webster (1993). A more thorough understanding of the group has been hampered by the rarity of some members, particularly in southern Mexico and Central America. Additional populations of sect. *Corylocroton* taxa have been discovered in South America and it is necessary to determine if these represent new species.

Species Included:

- *Croton beetlei* Croizat
- *Croton caracasanus* Pittier
- *Croton corylifolius* Lam.
- *Croton grewiiifolius* Müll. Arg. [*C. siltepecensis* Lundell]
- *Croton lundellii* Standl. [*C. petenensis* Lundell]
- *Croton mexicanus* Müll. Arg.
- *Croton oerstedianus* Müll. Arg.
- *Croton pagiveteris* Croizat
- *Croton quercetorum* Croizat
- *Croton tonduzii* Pax
- *Croton verapazensis* Donn. Sm. [*C. jalapensis* Croizat, *C. simiarum* Standl. & L.O. Williams]

Additional Species

These three members of subgenus *Moacroton* are currently not assigned to a section, although they each warrant a section of their own and will be treated as such in an upcoming publication (van Ee et al., in prep.).

- *Croton alabamensis* E. A. Smith ex Chapm.
- *Croton corinthius* Poveda & J.A. González
- *Croton nubigenus* G.L. Webster

The South American Grade

Croton pachypodus from Central America and *C. sapiifolius* from Brazil are not included in our molecular analyses here because some chloroplast sequence data are lacking. However, ITS data (not shown) place them within the SAG. The members of the South American grade are:

- *Croton megistocarpus* J.A. González & Poveda
- *Croton pachypodus* G.L. Webster
- *Croton olivaceus* Müll. Arg.
- *Croton piptocalyx* Müll. Arg.
- *Croton sampatik* Müll. Arg.
- *Croton sapiifolius* Müll. Arg.

Croton subgenus *Croton*

- *Croton* subgenus *Crotonopsis* (Michx.) Radcl.-Sm. & Govaerts, Kew Bull. 52 (1): 183 (1997).
- *Croton* subgenus *Eremocarpus* (Michx.) Radcl.-Sm. & Govaerts, Kew Bull. 52 (1): 184 (1997).
- *Croton* subgenus *Julocroton* (Michx.) Radcl.-Sm. & Govaerts, Kew Bull. 52(1): 184 (1997).

Included in the typical subgenus are all the remaining sections of *Croton* (Webster, 1993). Radcliffe-Smith and Govaerts (1997) recognized three subgenera, *Crotonopsis*, *Eremocarpus*, and *Julocroton*, in addition to the typical subgenus. We recognize these three groups as sections of *Croton* following Webster (1967, 1992, 1993).

Acknowledgments This study was supported by National Science Foundation grants DEB-0212481 (PEB) and DEB-0508725 (PEB & BVE), a grant from the Latin American Caribbean & Iberian studies (LACIS) program of UW-Madison, and a Davis foundation grant from the UW-Madison Department of Botany. Field work was made possible by generous support from the Jardín Botánico Nacional (HAJB) in Cuba, the staff of the Caribbean National Forest in Puerto Rico, R. Rueda and I. Coronado of the Herbario de la Universidad Nacional Autónoma de Nicaragua-León (HULE) in Nicaragua, N. Zamora of the Instituto Nacional de Biodiversidad (INB) in Costa Rica, B. León of El Colegio de la Frontera Sur, Unidad Chetumal (CIQR) in Mexico, Charles Sexton and Bill Reiner of Balcones Canyonlands National Wildlife Refuge, and Steve Ginzburg of the University of Alabama Herbaria. Additional material from South America was kindly provided by B. Caruzo and I. Cordeiro (SP). We are grateful to the curators of HAJB, MICH, MO, NY, SP, US, and WIS for allowing us to sample material from their collections.

Literature Cited

- Altekar G., S. Dwarkadas, J. P. Huelsenbeck & E. Ronquist.** 2004. Parallel metropolis coupled Markov chain Monte Carlo for Bayesian phylogenetic inference. *Bioinformatics* 20: 47–415.
- Alvarez, L. W., W. Alvarez, F. Asaro, & H. V. Michel.** 1980. Extraterrestrial cause for the Cretaceous-Tertiary boundary extinction. *Science*, 208: 1095–1108.
- Baldwin, B. G., M. J. Sanderson, J. M. Porter, M. F. Wojciechowski, C. S. Campbell, & M. J. Donoghue.** 1995. The ITS region of nuclear ribosomal DNA: a valuable source of evidence on angiosperm phylogeny. *Ann. Mo. Bot. Gard.* 82: 247–277.
- Berry, P. E., A. L. Hipp, K. J. Wurdack, B. van Ee & R. Riina.** 2005. Molecular phylogenetics of the giant genus *Croton* and tribe Crotonae (Euphorbiaceae sensu stricto) using ITS and *trnL-trnF* DNA sequence data. *Am. J. Bot.* 92(9): 1520–1534.
- Borhidi, A.** 1977. El Género *Moacrotan* Croiz. (Euphorbiaceae). *Acta. Bot. Hung.* 36(1–4): 7–12.
- . 1985. Phytogeographic survey of Cuba. 1. The phytogeographic characteristics and evolution of the flora of Cuba. *Acta. Bot. Hung.* 31: 3–34.
- . 1996. Phytogeography and vegetation ecology of Cuba. Akadémiai Kiadó, Budapest.
- Chakrabarty, P.** 2006. Systematics and historical biogeography of Greater Antillean Cichlidae. *Mol. Phylogenet. Evol.* 29: 619–627.
- Chase, M. W. & H. H. Hillis.** 1991. Silica gel: an ideal material for field preservation of leaf samples for DNA studies. *Taxon.* 40: 215–220.
- Cunningham, C. W.** 1997. Can three incongruence tests predict when data should be combined? *Mol. Biol. Evol.* 14(7): 733–740.
- Dalling, J. W., S. P. Hubbell & K. Silvera.** 1998. Seed dispersal, seedling establishment and gap partitioning among tropical pioneer trees. *J. Ecol.* 86(4): 674–689.
- Dávalos, L. M.** 2004. Phylogeny and biogeography of Caribbean mammals. *Bot. J. Linn. Soc.* 81: 373–394.
- Davis, C. C., C. D. Bell, S. Mathews & M. J. Donoghue.** 2002. Laurasian migration explains Gondwanan disjunctions: evidence from Malpighiaceae. *PNAS* 99(10): 6833–6837.
- , **C. O. Webb, K. J. Wurdack, C. A. Jaramillo & M. J. Donoghue.** 2005. Explosive radiation of Malpighiales supports a mid-Cretaceous origin of modern tropical rain forests. *Am. Nat.* 165(3): E36–E65.
- Dilcher, D. L. & S. R. Manchester.** 1988. Investigations of angiosperms from the Eocene of North America: a fruit belonging to the Euphorbiaceae. *Tert. Res.* 9: 45–58.
- Downhower, J. F. & C. H. Racine.** 1976. Darwin's finches and *Croton scouleri*: an analysis of the consequences of seed predation. *Biotropica* 8(1): 66–70.
- Farris, J. S., M. Källersjö, A. G. Kluge & C. Bult.** 1994. Testing significance on incongruence. *Cladistics* 10: 315–319.
- Govaerts, R., D. G. Frodin & A. Radcliffe-Smith.** 2000. World checklist and bibliography of Euphorbiaceae, 4 vols. Royal Botanic Garden, Kew, UK.
- Huelsenbeck, J. P., & F. Ronquist.** 2001. MrBayes: Bayesian inference of phylogenetic trees. *Biometrics* 17: 754–755.
- Iturralde-Vinent, M.A., & R. D. E. MacPhee.** 1999. Paleogeography of the Caribbean region: implications for Cenozoic biogeography. *Bull. Am. Mus. Nat. Hist.* 238: 1–95.
- Kerr, A. C., M. A. Iturralde-Vinent, A. D. Sanders, T. L. Babbs & J. Tarney.** 1999. A new plate tectonic model of the Caribbean: Implications from a geochemical reconnaissance of Cuban Mesozoic volcanic rocks. *Geological Society of American Bulletin* 111(11): 1581–1599.
- MacPhee, R. D. E. & M. A. Iturralde-Vinent.** 1994. First Tertiary land mammal from Greater Antilles: an Early Miocene sloth (Xenarthra, Megalonychidae) from Cuba. *Am. Mus. Novit.* 3094: 13.
- & ———. 1995. Origin of the Greater Antillean Land Mammal Fauna, 1; New Tertiary fossils from Cuba and Puerto Rico. *Am. Mus. Novit.* 3141: 31.
- Maddison, D. R. & W. P. Maddison.** 2000. *MacClade 4: analysis of phylogeny and character evolution.* Sinauer Associates, Sunderland.
- Magallón, S. A. & M. J. Sanderson.** 2005. Angiosperm divergence times: the effect of genes, codon positions, and time constraints. *Evolution* 59(8): 1653–1670.
- Nixon, K.** 2006. The land of molecular Oz: Armchair paleobotany and molecular clocks. Botanical Society of America 2006 abstract, available at: www.2006.botanyconference.org

- Nylander, J. A. A.** 2004. MrModeltest v2. Program distributed by the author. Department of Systematic Zoology, Uppsala University.
- Ocampo, A., V. Vajda & E. Buffetaut.** 2006. Unraveling the Cretaceous–Paleogene (KT) catastrophe: evidence from flora fauna and geology. In: Cockell C, Koeberl C, Gilmour I (eds) Biological processes associated with impact events, Springer-Verlag series, pp 203–227.
- Olmstead, R.G. & J. A. Sweere.** 1994. Combining data in phylogenetic systematics—an empirical approach using three molecular data sets in the Solanaceae. *Syst. Biol.* 43: 467–481.
- Passos, L. & S. O. Ferreira.** 1996. Ant dispersal of *Croton priscus* (Euphorbiaceae) in a tropical semideciduous forest in southeastern Brazil. *Biotropica* 28(4b): 697–700.
- Pitman, W. C., S. C. Cande, J. LaBreque & J. L. Pindell.** 1993. Fragmentation of Gondwana: the separation of Africa from South America. In: Goldblatt P (ed) Biological relationships between Africa and South America. Yale University Press, New Haven, pp 817–818.
- Pope, K. O., S. L. D'Hondt & C. R. Marshall.** 1998. Meteorite impact and the mass extinction of species at the Cretaceous/Tertiary boundary. *Proc. Natl. Acad. Sci.* 95: 11028–11029.
- Posada, D. & K. A. Crandall.** 1998. MODELTEST testing the model of DNA substitution. *Bioinformatics* 14: 817–818.
- Potter, F. W. & D. L. Dilcher.** 1980. Biostratigraphic analysis of Eocene clay deposits in Henry County, Tennessee. In: Dilcher DL, Taylor TN (eds) Biostratigraphy of fossil plants. Dowden, Hutchinson & Ross, Stroudsburg, PA, USA, pp 211–225.
- Radcliffe-Smith, A. & R. Govaerts.** 1997. New names and combinations in the Crotonoideae. *Kew. Bull.* 52(1): 183–189.
- Sanderson, M. J.** 2002. Estimating absolute rates of molecular evolution and divergence times: a penalized likelihood approach. *Mol. Biol. Evol.* 19: 101–109.
- . 2003. r8s: inferring absolute rates of molecular evolution and divergence times in the absence of a molecular clock. *Bioinformatics Applications Note* 19(2): 301–302.
- Sun, M., X. Sun, Y. Zhao, O. Wang, Z. Li, Z. Hu & P. Mei.** 1989. Sporo-pollen. In: Division of Comprehensive Studies on Ocean Geology of the Ministry of Geology and Minerals of P.R. China and Institute of Geology of the Chinese Academy of Geological Sciences (ed), Cenozoic paleobiota of the continental shelf of East China Sea (Donghai) (Micropaleobotanical Volume). Geological Publishing House, Beijing, pp 6–111.
- Swofford, D. L.** 2002. PAUP*: Phylogenetic analysis using parsimony (* and other methods), version 4.0 beta 10. Sinauer Associates, Sunderland, Ma, USA.
- Taberlet, P., L. Gielly, G. Pautou & J. Bouvet.** 1991. Universal primers for amplification of three non-coding regions of chloroplast DNA. *Plant. Mol. Biol.* 17: 1105–1109.
- Tiffney, B.** 1985. The Eocene North Atlantic land bridge: its importance in tertiary and modern phytogeography of the Northern hemisphere. *J. Arnold. Arb.* 66: 243–273.
- Urbatsch, L. E., B. G. Baldwin & M. J. Donoghue.** 2000. Phylogeny of the coneflowers and relatives (Heliantheae: Asteraceae) based on nuclear rDNA internal transcribed spacer (ITS) sequences and chloroplast DNA restriction site data. *Syst. Bot.* 25(3): 539–565.
- Van Ee, B., N. Jelinski, P. E. Berry & A. L. Hipp.** 2006. Phylogeny and biogeography of *Croton alabamensis* (Euphorbiaceae), a rare shrub from Texas and Alabama, using DNA sequence and AFLP data. *Mol. Ecol.* 15: 2735–2751.
- Webster, G. L.** 1967. The genera of Euphorbiaceae in the southeastern United States. *J. Arnold. Arb.* 48: 303–430.
- . 1992. Realignments in American *Croton*. *Novon.* 2: 269–273.
- . 1993. A provisional synopsis of the sections of the genus *Croton* (Euphorbiaceae). *Taxon.* 42: 793–823.
- , **M. J. del Arco Aguilar & B. A. Smith.** 1996. Systematic distribution of foliar trichome types in *Croton* (Euphorbiaceae). *Bot. J. Linn. Soc.* 121: 41–57.
- White, T. J., T. Bruns, S. Lee & J. Taylor.** 1990. Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. In: Innis MA, Gelfand DH, Sninsky JJ, White TJ (eds) PCR protocols. A guide to methods and applications, Academic Press, San Diego, pp 315–324.
- Wiens, J. J.** 1998. Combining data sets with different phylogenetic histories. *Syst. Biol.* 47(4): 568–581.
- Wikström, N., V. Savolainen & M. W. Chase.** 2001. Evolution of the angiosperms: calibrating the family tree. *Proc. R. Soc. Lond. B-Bio.* 268: 2211–2220.
- Wurdack, K. J.** 2006. The lectotypification and 19th century history of *Croton alabamensis* (Euphorbiaceae s.s.). *SIDA* 22(1): 469–483.

- , P. Hoffmann & M. W. Chase. 2005. Molecular phylogenetic analysis of uniovulate Euphorbiaceae (Euphorbiaceae sensu stricto) using plastid *rbcL* and *trnL-F* DNA sequences. *Am. J. Bot.* 92(8): 1397–1420.
- Zachos, J., M. Paganini, L. Sloan, E. Thomas & K. Billups. 2001. Trends, rhythms, and aberrations in global climate 65 Ma to present. *Science* 292: 686–693.

Appendix

Taxa, vouchers, localities, and GenBank accession numbers for all sequences analyzed. GenBank accession numbers starting with EF- were newly generated; those starting with AY- or DQ- were obtained from Berry et al. (2005), Wurdack et al. (2005), and van Ee et al. (2006).

Taxon; Origin; *Voucher*; GenBank accession numbers: *rbcL*; *trnL-trnF*, *ndhF*, ITS.

Acalypha californica Benth.; USA, California; *G. Levin 2192* (SD); AY794943; AY794776;—;—. *Acidocroton trichophyllus* Urb.; Cuba, Holguín; *HABJ 81844* (WIS); EF405827; EF408087; EF416920; EF421766. *Acidoton nicaraguensis* (Hemsl.) G.L. Webster; Nicaragua; *van Ee 590* (WIS); EF405828; EF408088;—;—. *Actinostemon caribaeus* Griseb.; Panama; *C. Galdames 3568* (NY); AY794863; AY794657;—;—. *Adelia ricinella* L.; Puerto Rico; *F. Axelrod 3874* (NY); AY794918; AY794737;—;—. *Aleurites moluccana* (L.) Willd.; Cult. USA, Missouri, MO 895208; *K. Wurdack* (US); AY794883; AY794709;—;—. *Argomuelleria macrophylla* Pax; Ghana; *H. Schmidt et al. 2240* (MO); AY794937; AY794769;—;—. *Astraea lobata* (L.) Klotzsch; Costa Rica, Guanacaste; *van Ee 296* (WIS); EF405829; EF408089; EF416921; EF421720. *Bonania cubana* A. Rich.; Bahamas, Andros; *S. Hill 3156* (NY); AY794833; AY794613;—;—. *Brasiliocroton mamoninha* P.E. Berry & I. Cordeiro; Brazil, Maranhão; *M. Lobo et al. 340* (NY); AY794907; AY794691; EF416922; AY971174. *Chrozophora tinctoria* (L.) Raf.; Greece; *B. Verdcourt 4137* (US); AY794951; AY794787;—;—. *Croton alabamensis* E.A. Sm. ex Chapm. var. *alabamensis*; USA, Alabama; *van Ee et al. 369* (WIS); EF405831; DQ227545; EF416931; DQ227513. *Croton alabamensis* E.A. Sm. ex Chapm. var. *texensis* Ginzburg; USA, Texas; *van Ee et al. 341* (WIS); EF405832; DQ227553; EF416932; DQ227521. *Croton alainii* van Ee & P.E. Berry [*Moacrotan lanceolatus* Alain]; Cuba, Holguín; *HABJ 81770* (WIS);—; EF408138; EF416956; EF421756. *Croton billbergianus* Müll. Arg.; Costa Rica, Limón; *van Ee & van Ee 595* (WIS); EF405833; EF408109; EF416933; EF421772. *Croton caracasanus* Pittier; Venezuela, Distrito Federal; *Riina 1288* (WIS); EF405834; DQ227525; EF416934; DQ227557. *Croton corinthius* Poveda & J.A. González; Costa Rica, Limón; *van Ee & van Ee 600* (WIS); EF405835; EF408110; EF416935; EF421751. *Croton corylifolius* Lam.; Cuba, Pinar del Rio; *HABJ 81975* (WIS);—; EF408094; EF416923; EF421727. *Croton daphniphyllum* Radcl.-Sm. ined.; ?Madagascar; *McPherson 18310* (MO); EF405836; DQ227531; EF416936; DQ227563. *Croton discolor* Willd.; Puerto Rico; *van Ee 547* (WIS); EF405838; EF408112; EF416938; EF421774. *Croton draco* Schltld. & Cham.; Nicaragua, Región Autónoma del Atlántico Norte, *van Ee & Coronado 592* (WIS); EF405840; EF408114; EF416940; EF421776. *Croton ekmanii* Urb. [*Moacrotan ekmanii* (Urb.) Croizat]; Cuba, Holguín; *HABJ 81786* (WIS); EF405860; EF408145; EF416961; EF421761.

Croton ekmanii Urb. [*Moacroton ekmanii* (Urb.) Croizat]; Cuba, Guantánamo; *HABJ* 81885 (WIS);-; EF408143; EF416959; EF421760. *Croton ekmanii* Urb. [*Moacroton ekmanii* (Urb.) Croizat]; Cuba, Guantánamo; *HABJ* 81888 (WIS);-; EF408147; EF416963; EF421763. *Croton hoffmannii* Müll. Arg.; Costa Rica, Cartago; *van Ee & van Ee* 598 (WIS); EF405837; EF408111; EF416937; EF421773. *Croton impressus* Urb.; Puerto Rico; *van Ee* 543 (WIS); EF405839; EF408113; EF416939; EF421775. *Croton jimenezii* Standl. & Valerio; Costa Rica, San José; *van Ee* 326 (WIS); EF405841; EF408115; EF416941; EF421777. *Croton laurinus* Sw.; Jamaica; *Cacho* 50 (WIS); EF405842; EF408116; EF416942; EF421778. *Croton leonis* (Croizat) van Ee & P.E. Berry [*Moacroton leonis* Croizat]; Cuba, Holguín; *HABJ* 81773 (WIS);-; EF408140; EF416957; EF421758. *Croton lucidus* L.; Cuba, Havana; *van Ee* 378 (WIS); EF405843; EF408117; EF416943; EF421765. *Croton lundellii* Standl.; México, Campeche; *León* 037 (WIS);-; EF408097; EF416925; EF421732. *Croton lundellii* Standl.; Belize, Cayo District; *Whitefoord* 10311 (MO);-; EF408098; EF416926; EF421764. *Croton* cf. *lundellii* Standl.; México, Yucatán; *van Ee* 123 (WIS);-; EF408099; EF416927; EF421733. *Croton* cf. *lundellii* Standl.; México, Yucatán; *van Ee* 451 (WIS);-; EF408102; EF416928; EF421736. *Croton maestrense* (Alain) van Ee & Berry [*Cubacroton maestrense* Alain]; Cuba, Granma; *HABJ* 81958 (WIS); EF405857; EF408127; EF416953; EF421753. *Croton megistocarpus* J.A. González & Poveda; Costa Rica, Puntarenas; *Gardner s.n.* (WIS); EF405845; EF408118; EF416944; EF421779. *Croton mexicanus* Müll.Arg.; México, Campeche; *León* 023 (WIS);-; EF408096; EF416924; EF421731. *Croton mexicanus* Müll.Arg.; México, Chiapas; *León* 045 (WIS);-; EF408107; EF416930; EF421746. *Croton niveus* Jacq.; Costa Rica, Heredia; *van Ee & van Ee* 593 (WIS); EF405847; EF408119; EF416945; EF421780. *Croton noronhae* Baill.; Madagascar; *McPherson* 18296 (MO); EF405848; EF408120; EF416946; EF421781. *Croton nubigenus* G.L. Webster; Nicaragua, Región Autónoma del Atlántico Norte; *van Ee & Coronado* 589 (WIS); EF405849; EF408121; EF416947; EF421716. *Croton olivaceus* Müll. Arg.; Ecuador, Napo; *Neill* 11163 (MO, QCNE); EF405830; AY794694; EF416929; AY791237. *Croton piptocalyx* Müll. Arg.; Brazil; *Caruzo* 54 (SP); EF405858; EF408132; EF416954; EF421791. *Croton poecilanthus* Urb.; Puerto Rico; *van Ee* 551 (WIS); EF405850; EF408122; EF416948; EF421782. *Croton punctatus* Jacq.; USA, North Carolina; *van Ee* 514 (WIS); EF405851; EF408123; EF416964; EF421783. *Croton revolutus* (Alain) van Ee & P.E. Berry [*Moacroton revolutus* Alain]; Cuba, Matanzas; *HABJ* 82008 (WIS);-; EF408142; EF416958; EF421759. *Croton roraimensis* Croizat; Venezuela; *Berry* 7738 (MICH); EF405852; EF408124;-;- *Croton sampatik* Müll. Arg.; Peru, Pasco; *Riina et al.* 1447 (WIS); EF405859; EF408133; EF416955; EF421792. *Croton setiger* Hook.; USA, California; *Hughey s.n.* (US); AY794818; AY794697; EF416949; AY791249. *Croton tricolor* Klotzsch ex Baill.; Brazil; *Caruzo* 87 (SP); EF405854; EF408125; EF416950; EF421752. *Croton trigonocarpus* Griseb. [*Moacroton trigonocarpus* (Griseb.) Croizat]; Cuba, Pinar del Rio; *HABJ* 81960 (WIS); EF405861; EF408146; EF416962; EF421762. *Croton trigonocarpus* Griseb. [*Moacroton trigonocarpus* (Griseb.) Croizat]; Cuba, Pinar del Rio; *van Ee* 380 (WIS);-; DQ227562; EF416960; DQ227530. *Croton yucatanensis* Lundell; Mexico, Yucatán; *van Ee* 121 (WIS); EF405855; DQ227537; EF416951; DQ227569. *Croton zambesicus* Müll. Arg.; Zambia, Songwe Gorge; *Zimba* 901 (MO); EF405856; AY971260; EF416952; AY971341. *Ditta myricoides* Griseb.; Puerto Rico; *J. Cedeño s.n.* (US); AY794871;

AY794675;—;—. *Elateriospermum tapos* Blume; Malaysia; *E. Soepadmo* & *S. Suhaimi s193* (NY); AY794873; AY794678;—;—. *Euphorbia mesembryanthemifolia* Jacq.; USA, Florida; *K. Wurdack D102* (US); AY794820; AY794601;—;—. *Euphorbia obesa* Hook. f.; Cult. USA, Maryland; *K. Wurdack D539* (US); AY794826; AY794605;—;—. *Euphorbia pulcherrima* Willd. ex Klotzsch; Cult. USA, North Carolina; *K. Wurdack D084* (US); AY794819; AY794600;—;—. *Gymnanthes lucida* Sw.; Cult. USA, Florida, Fairchild Tropical Garden X2284a; *K. Wurdack D055* (US); AY794858; AY794653;—;—. *Hippomane mancinella* L.; Cult. USA, Florida, Fairchild Tropical Garden 67260a; *K. Wurdack D053* (US); AY794835; AY794616;—;—. *Jatropha integerrima* Jacq.; Cult. USA, Fairchild Tropical Garden 63169a; *K. Wurdack D047* (US); AY794902; AY794685;—;—. *Lasiocroton ? bahamensis* Pax & K. Hoffm.; Cult. USA, Fairchild Tropical Garden 66629b; *K. Wurdack D058* (US); AY788181; AY794739;—;—. *Leucocroton comosus* Urb.; Cuba; *P. Delprete et al. 8696* (NY); AY794919; AY794738;—;—. *Ophellantha steyermarkii* Standl.; Mexico, Chiapas; *D. Breedlove 46994* (NY); AY794906; AY794690;—;—. *Panda oleosa* Pierre; Ghana; *H. Schmidt et al. 2048* (MO); AY663644; AY794815;—;—. *Pera bicolor* (Klotzsch) Müll. Arg.; Guyana; *L. Gillespie 4300* (US); AY794968; AY794808;—;—. *Sagotia racemosa* Baill.; Peru, Madre de Dios; *S. Smith 253* (US); AY794903; AY794687;—;—. *Sandwithia guyanensis* Lanj.; Guyana; *R. Ek et al. 906* (NY); AY794904; AY794688;—;—.