‘Andean-centred’ genera in the short-branch clade of Annonaceae: testing biogeographical hypotheses using phylogeny reconstruction and molecular dating

Michael D. Pirie1*, Lars W. Chatrou1, Johan B. Mols2, Roy H. J. Erkens1 and Jessica Oosterhof1

1Nationaal Herbarium Nederland, Universiteit Utrecht Branch, Utrecht, The Netherlands; 2Nationaal Herbarium Nederland, Universiteit Leiden Branch, Leiden, The Netherlands

ABSTRACT

Aim We test biogeographical hypotheses regarding the origin of Andean-centred plant groups by reconstructing phylogeny in the short-branch clade (SBC) of Annonaceae, and estimating the timing of diversifications in four apparently Andean-centred genera: Cremastosperma R.E.Fr., Klarobelia Chatrou, Malmea R.E.Fr. and Mosannona Chatrou. The SBC includes species distributed in both the Old and New World tropics. A number of the Neotropical genera display ‘Andean-centred’ distribution patterns, with high species richness on both sides of the Andes mountain range. In particular, we test whether these groups could have originated on the South American continent during the time frame of the Andean orogeny [from c. 23 Ma (Miocene) to the present].

Methods Chloroplast DNA sequences were used to reconstruct phylogeny in related Annonaceae taxa plus outgroups, under maximum parsimony and Bayesian inference. The markers rbcL, trnL-trnF and psbA-trnH were sampled for 96 accessions to test the monophyly of each of the genera, and thus whether they might be para- or polyphyletic with respect to related groups distributed across Amazonia. To determine the sister groups of the four genera, the additional markers matK, ndhF, trnT-trnL, trnS-trnG and atpB-rbcL were sampled for 23 of the 96 accessions. Molecular dating techniques (nonparametric rate-smoothing; penalized likelihood; Bayesian inference) were then applied to estimate the age of the crown group of each genus and the age of their sister groups.

Results Monophyly was confirmed in Cremastosperma, Malmea and Mosannona. The monotypic genus Pseudephedranthus Aristeg. was found to be nested within Klarobelia, the species of which otherwise formed a monophyletic group, and a South American-centred (SAC) clade was identified. The SAC clade comprises all the SBC genera distributed in South America and generally to a limited extent into Central America, but not those endemic to Africa, Asia and Central America. Age estimations for clades within the SBC were no older than around 60 Myr; those for the crown groups of Cremastosperma, Klarobelia, Malmea and Mosannona fell largely within the last 10–20 Myr.

Main conclusions The distribution patterns of Cremastosperma, Klarobelia, Malmea and Mosannona are not the arbitrary result of the definition of para- or polyphyletic groups. We infer the presence of a common ancestor of the four genera in South America, but not by vicariance of an ancestral population on Gondwana. The age estimations, instead, may suggest that the SAC clade originated in South America by dispersal across the Boreotropics. Although the strength of this test was limited by imprecision in the molecular dating results, the
INTRODUCTION

Around a third of all flowering plants are found in the Neotropics (tropical America) (Smith et al., 2004). Two areas within the Neotropics, the tropical Andes (including forests on the eastern side of the Andes extending from Colombia through eastern Ecuador and Peru as far south as Bolivia), and the Chocó/Darién/Western Ecuador region (the narrow tropical zone on the Pacific Ocean side of the Andean mountain chain), together house 22,500 endemic plant species representing 7.5% of all species of plants worldwide (Myers et al., 2000). Understanding the origins of these biodiverse areas in the Neotropics will help us to determine why this region is so species-rich.

In his 1982 paper, Al Gentry proposed historical biogeographical scenarios that might explain the high floristic diversity of the Neotropics in general, and the areas surrounding the Andes in particular. Gentry considered the majority of South American taxa to be Gondwanan-derived, and he distinguished two groups within them on the basis of distribution patterns and growth forms. Amazon-centred taxa, sensu Gentry, are largely canopy trees and lianas. Andean-centred taxa, in contrast, are diverse in the lowlands near the base of the mountains and in middle-elevation cloud forests, areas corresponding to the tropical Andes and Chocó/Darién/western Ecuador region, with very poor representation in Amazonia. They are chiefly shrubs, epiphytes and palmettos. Gentry’s explanation for the distribution of Andean-centred taxa was that they were subject to speciation as a result of the formation of the Andean mountain chain. However, South American geological history of the past few tens of millions of years is complex, and a number of overlapping events may be critical in determining current distribution patterns – in order of decreasing age: the Andean orogeny (Miocene onwards); closure of Panama isthmus (Pliocene); and the numerous episodes of climatic changes occurring throughout the Pleistocene (Burnham & Graham, 1999). Even the relative importance of historical vs. present-day (ecological) factors in determining current plant species distributions remains a point of contention (Tuomisto & Ruokolainen, 1997).

The pantropically distributed family Annonaceae comprises c. 2500 species of trees and lianas, found predominantly in tropical rain forests. Over 900 species are recognized in the Neotropics (Chatrou et al., 2004), where they represent a significant part of plant diversity, both in terms of number of species and number of individuals (Valencia et al., 1994; Ter Steege et al., 2000). The majority of Annonaceae genera in South America would be considered Amazon-centred according to Gentry: most species are medium-to-large trees distributed across the Amazon. However, some groups, notably the genera Cremastosperma R.E.Fr., Klarobelia Chatrou, Malmea R.E.Fr. and Mosannona Chatrou, display markedly Andean-centred distributions (lowland to pre-montane forest usually only up to 1500 m) and are usually small understory trees.

Cremastosperma can be distinguished from other Neotropical Annonaceae by its raised midrib with a unique longitudinal groove, and comprises c. 35 species. Most species are found along the Andes in Peru and Ecuador, with significant diversity also extending north through Colombia into Panama and Costa Rica (Fig. 1a). Only four species have been found further east: Cremastosperma macrocarpum Maas and Cremastosperma venezuelanum Pirie in coastal Venezuela; Cremastosperma brevipes (DC.) R.E.Fr. on the Guiana shield; and Cremastosperma monospermum (Rubey) R.E.Fr. widespread from Peru across northern Bolivia and central Brazil. Of the 19 species originally placed under Malmea by Fries (1937), 12 were subsequently moved to three new genera described by Chatrou (1998) on the basis of leaf, inflorescence and seed characters: Klarobelia (now including 12 recognized species), Mosannona (14) and Pseudomalmea (three). Malmea now includes six recognized species. The distributions of species of Klarobelia, Malmea and Mosannona (Fig. 1b–d) are markedly similar to those of Cremastosperma. In all four genera, no single species is found either on both sides of the Andean mountain chain, or distributed across more than one of the further disjunct areas in the Guianas, Venezuela and tropical Andes. However, there are differences: Mosannona is distributed further into Central America, with one species found as far north-west as the Mexican states of Veracruz and Nayarit, and, as is also the case for Malmea, no species of Mosannona have been collected in coastal Venezuela. Further, the type species of Malmea, Malmea obovata R.E.Fr., is known only from one collection from the Atlantic coast of Brazil, and Malmea manausensis Maas & Mirafla is found in the heart of the Amazon basin (Fig. 1c); in neither of these two areas have species of the other three genera been collected.

Phylogeny reconstruction in Annonaceae: are Cremastosperma, Klarobelia, Malmea and Mosannona monophyletic?

The informal names long-branch clade (LBC) and short-branch clade (SBC) (corresponding to the ‘inaperturate’ and ages of crown groups of the four genera suggest that diversifications occurred within the time frame of the orogeny of the Northern Andes.

Keywords
Andes, Annonaceae, biogeography, Cremastosperma, Klarobelia, Malmea, molecular dating, Mosannona, Neotropics, phylogeny reconstruction.
The SBC phylogeny: pantropical disjunct distribution patterns and their implications for biogeographical hypotheses in the Neotropics

Recent studies of angiosperm groups, including Annonaceae, previously regarded as Gondwanan, have demonstrated using molecular dating techniques that these distributions may have originated later than the estimated timing of separation of the continents that constitute Gondwana (Renner et al., 2001; Davis et al., 2002; Doyle et al., 2004; Richardson et al., 2004). A combination of the presence of land connections and suitable climatic conditions made available a number of dispersal paths between currently isolated tropical zones during periods since the break-up of Gondwana (Morley, 2003; Pennington & Dick, 2004). The view that, since its split from Africa around 100 Ma, the geological history of South
America is that of an island continent only recently connected to North America via the Panamanian isthmus in the Pliocene (Burnham & Graham, 1999), may thus underestimate the role of dispersal in the origin of pantropical disjunctions (Morley, 2003). One proposed explanation for the pantropical distribution of clades within Annonaceae is that they may represent relics of a Palaeogene boreotropical flora (as described by Wolfe, 1975). The possible dispersal routes this represents, unlike the presumption of a common ancestor on Gondwana, would not necessarily involve an ancestral area in South America. The phylogeny of the SBC and the positions of Cremastosperma, Klarobelia, Malmea and Mosannona within it have important implications for possible reconstructions of their biogeographical history. Levels of resolution on the basis of \textit{rbcL} and \textit{trnL-F} do not exclude the possibility that most of the Neotropical SBC genera together could represent a monophyletic group (Richardson \textit{et al.}, 2004). Should, in particular, the Cremastosperma and Malmea clades be demonstrated to be sister to other South American-centred clades, this would indicate a common geographical origin (in the Neotropics) dating back at least as far as the most recent common ancestor (MRCA) of those groups. However, should such a Neotropical clade share a most recent last common ancestor with a (largely) Asian clade, then its Neotropical distribution could theoretically be as recent as the age of its crown group. The origin of its constituent species in South America would then need to be reassessed.

The importance of this question for Andean biogeography is that it is currently impossible to reject the possibility that \textit{Cremastosperma} and/or \textit{Malmea} originated in different geographical areas from \textit{Klarobelia} or \textit{Mosannona}: within or outside the South American continent, with recent dispersal to South or Central America followed by radiation. The similarity in their current distribution patterns might not reflect common biogeographical history, and the explanation of Gentry for their diversification, as a response to the Andean orogeny, might even fall outside the time frame of their presence in South America.

Aims

Gentry’s explanation for the distribution of Andean-centred taxa was that they (1) originated in South America (descending from Gondwana ancestors), and (2) were subject to a recent burst of speciation as a result of the of the Andean orogeny. These represent two hypotheses for the origin of \textit{Cremastosperma}, \textit{Klarobelia}, \textit{Malmea} and \textit{Mosannona}, which can be tested using phylogenetic reconstruction and molecular dating techniques.

![Figure 2 Phylogeny of the Annonaceae.](image-url)
The aims of this study are, therefore, firstly to determine the relationships between the four genera and other Neotropical and Palearctic SBC clades and to estimate the ages of the sister groups that span present-day tropical regions. Competing explanations for the distribution of SBC Annonaceae across the tropics can then be tested according to the time frame within which they are hypothesized to have taken place. This, in turn, may help us determine the geographical origin of the Andean-centred genera. Secondly, we aim to test the monophyly of each of these genera, and also to determine the age of the crown group of each genus. In this way we can both exclude the possibility that the distribution patterns as observed are the result of the definition of para- or polyphyletic groups, and test whether the species diversifications of Andean-centred clades took place within the time frame of the Andean orogeny.

Should the similar distribution patterns of these Andean-centred genera in Annonaceae be due to a common biogeographical history, then reconstructing that history might also offer insight into the origins of the high diversity of other taxa in north-western South America. Alternatively, the species represented by these Annonaceae genera might have originated in different ways and/or over different periods of time.

METHODS

Taxon sampling

This study largely utilized previously unpublished sequence data as well as published sequences (Sauquet et al., 2003; Mols et al., 2004; Pirie et al., 2005). Around half the total numbers of described and undescribed species for each of the four genera under study were represented by 14 samples of Cremastosperma; six of Klarobelia; four of Malmea; and seven of Mosannona. The geographical distribution of specimens sampled is indicated on the distribution maps (Fig. 1). A total of 77 SBC taxa were sampled, including increased sampling of species representing all related South American genera as presented by Richardson et al. (2004); five Asian and Central American taxa representative of the miliusoid clade, including first-branching lineage Monocarpia (Mols et al., 2004); three samples representing the Asian ‘Polyalthia hypoleuca complex’ clade (Rogstad, 1989; Mols et al., 2004); and eight representing the African genera Greenwayodendron, Piptostigma and Amniciina, first-branching lineages of the SBC (Mols et al., 2004; Richardson et al., 2004). Major lineages within the SBC and the first-branching lineages of Annonaceae, the ‘Ambavioid’ clade and Anaxagorea (Doyle & Le Thomas, 1996; Richardson et al., 2004) were represented by 16 accessions. Outgroups were selected from other families in the order Magnoliales: Magnolia and Liriodendron (Magnoliaceae) (not Eupomatiaceae, the sister group to Annonaceae, as sequences were unavailable) and Coelocaryon (Myristicaceae). GenBank accession numbers and voucher details are presented in Appendices S1 and S2 in Supplementary Material.

Character sampling

For all 96 accessions, the chloroplast DNA markers rbcL, trnL-trnF and psbA-trnH were sampled (matrix 1, see Appendix S1); and for 23 of these 96 accessions the additional markers matK, ndhF, trnT-trnL, trnS-trnG and atpB-rbcL were sampled (matrix 2, see Appendix S2). LBC accessions were excluded in matrix 2, as high sequence divergence within this clade made homology assessment in the alignment of non-coding markers ambiguous. Having confirmed the best outgroups for the SBC in matrix 1, both alignment problems and sequencing effort were minimized by the exclusion of more distant outgroups in matrix 2.

DNA extraction, PCR amplification and sequencing

Total genomic DNA was extracted using a modified cetyl trimethyl ammonium bromide (CTAB) method (Doyle & Doyle, 1987): 50 mg silica dried or herbarium leaf material was homogenized in 1300 μL CTAB and incubated for 20 min with 12 μL 2-mercaptoethanol at 65 °C, followed by 90 min ambient mixing with 1 mL 24 : 1 chloroform : isooamylalcohol. After 10 min centrifugation at 9500 g, 300 μL supernatant was purified using the Wizard DNA purification system (Promega, Leiden, The Netherlands) (without isopropanol precipitation, avoiding the co-precipitation of oxidized material; Savolainen et al., 1995).

PCR amplification conditions were modified depending on the qualities of the DNA sample available. Samples extracted from herbarium material often contain lower quantities of more fragmented DNA and higher levels of PCR-inhibiting compounds (Savoilainen et al., 1995). In most cases the rbcL gene was amplified in two pieces and sequenced using primers 1F/724R (Olmstead et al., 1992) and 636F/1460R (Fay et al., 1997, 1998). Where amplification was unsuccessful, further internal primers, 217F, 922F, 536R and 1104R (Pirie et al., 2005) and 376R: 5’-GGGTTCAAAGCTGCTAGAGCTTACG-3’ and 444F: 5’-GGTCCGCCCCATGGCAGCATTCC-3’ were amplified and sequenced using primers 1F/536R, 217F/724R or 1F/376F, 217F/536R and 444F/724R, 636F/1104R and 922F/1460R to amplify the gene in up to five overlapping pieces of between 300 and 500 bp long. Plant universal primers of Taberlet et al. (1991) were used to amplify separately and sequence the trnL intron (primers C/D) and trnL-trnF spacer (primers E/F). The psbA-trnH intergenic spacer was amplified and sequenced using primers psb A and trn H (GUG) (Hamilton, 1999). Partial matK sequences were amplified using primers 390F and 1326R (Cuénoud et al., 2002), and MintF and MintR (Pirie et al., 2005), in combination 390F/1326R or 390F/MintR and MintF/1326R and sequenced using primers 390F and 1326R. The ndhF gene was amplified and sequenced in two overlapping pieces using primers 1, 972 and 2110R (Olmstead & Sweere, 1994) and 1165R (Kim et al., 2001) in combinations 1/1165R and 972/2110R. The trnT-trnL intergenic spacer was amplified using primers A and B (Taberlet et al., 1991), AintF: 5’-CCGTTCGCCATTCGACGACC-3’ and AintR: 5’-CGTTGATGTATCCGCAATTCAAT-3’ in combination A/B or A/BintR and AintF, and sequenced
using primers A and B. The trnS-trnG intergenic spacer was amplified and sequenced using primers trn S (GCU) and trn G (UCC) (Hamilton, 1999), and the atpB-rbcL intergenic spacer using primers atpbr3c (complementary to S20 of Hoot et al., 1995) and atpbrc2 (T. Scharaschkin and J. Doyle, pers. comm.).

A standard PCR protocol was used throughout, with the addition of 1 μL 0.4% bovine serum albumin per 25-μL reaction (which was found to increase amplification in all samples), 35 cycles of 30 s, 94 °C; 1 min, 55 °C; 2 min, 72 °C, with an initial 4 min, 94 °C and final 7 min, 72 °C. PCR products were purified using QIAquick PCR purification kits (Qiagen, Venlo, The Netherlands), sequenced with the PCR primers, and analysed by electrophoresis using an automatic sequencer ABI 3730XL.

Phylogenetic analysis

DNA sequences were edited in seqman 4.0 (DNAStar Inc., Madison, WI, USA) and aligned manually, resulting in alignments of 1496 positions (rbcL, including a 34–41-aligned-positions long non-coding region on the 3' end); 1402 positions (trnL-F); 798 positions (psbA-trnH); 843 positions (matK); 2043 positions (ndhF); 1157 positions (trnT-trnL); 953 positions (trnS-trnG); and 828 positions (atpB-rbcL). Areas of the alignments where the assessment of homology was ambiguous were excluded from the analyses.

Gaps in the alignments were coded as present/absent characters where they could be coded unambiguously, following the simple gap-coding method of Simmons & Ochoterena (2000). Two excluded regions, one in psbA-trnH, the other in trnT-trnL, of 15 and 12 positions, respectively, appeared to represent inversions, with around half the accessions possessing almost exact reverse-complement sequences of the others. Under the assumption that these inversions had occurred with high frequency, the bases in one version were aligned with those of the reverse complement of the other. These characters displayed little or no homoplasy when optimized onto the bootstrap topologies, and were therefore presumed to contain phylogenetic signal and included in further analyses.

Maximum parsimony analysis

Data were analysed using the parsimony algorithm of the software package PAUP* 4.0b10 (Swofford, 2000) under the equal and unordered weights criterion (Fitch parsimony; Fitch, 1971). The length of the shortest trees was estimated for matrix 1 using the parsimony ratchet (Nixon, 1999) as implemented using PAUP* and PAUPR at (Sikes & Lewis, 2001). All shortest trees were calculated using the ‘branch and bound’ method for matrix 2, and Bremer support (Bremer, 1994) estimated using the program TREEOT (Sorenson, 1999). Support was also estimated using bootstrap analyses of 500 replicates with ‘full’ heuristic searches of 50 random addition sequences, Tree Bisection and Reconnection (TBR), saving 50 trees each time. Bootstrap percentages were interpreted following Richardson et al. (2004): 50–74% represents weak support, 75–84% moderate support, and 85–100% strong support. Bootstrap analyses were performed on matrix 2 with and without the rbcL sequence data (see Results).

Selecting the best fitting DNA substitution model

ModelTest 3.06 (Posada & Crandall, 1998) was used to select the substitution model best fitting each sequence data partition, and the combined sequence data of matrix 1, using an arbitrary most parsimonious tree topology as estimated above for matrix 1 and the most parsimonious topology for matrix 2.

Bayesian analysis

The combined data sets were also analysed using Bayesian inference, as implemented in MrBayes ver. 3.0 (Huelsenbeck, 2000). The data were partitioned according to the separate markers used, and both rates and substitution models were allowed to vary across the partitions. Prior values for the DNA substitution models were applied to each partition (as determined using ModelTest above). Prior probabilities for all topologies were equal. Coelocaryon preussii (Myristicaceae, sister group to rest of Magnoliidae; Sauquet et al., 2003) was chosen as the single outgroup taxon permitted by MrBayes in matrix 1; Cleistopholis glauca as outgroup for matrix 2. Markov chain Monte Carlo (MCMC) analyses were run for 5,000,000 generations with four simultaneous MCMC chains to calculate posterior probabilities, and one tree per 100 generations was saved. The burn-in values were determined empirically from the likelihood values and 50% majority rule consensus trees calculated together with approximations of the posterior probabilities for the observed bipartitions.

Molecular dating

Topology

Nodes present in the Bayesian consensus of matrix 1 and not contradicted by results from matrix 2 were used to constrain two further maximum parsimony (MP) searches of the sequence data of matrix 1 (‘full’ heuristic, 100 random taxon addition sequences, TBR, saving maximum of 50 shortest trees each time), from which single arbitrary most parsimonious topologies were selected. The two searches included (a) all taxa (96 in total), and (b) all taxa minus nine accessions of Malmea represented by four accessions (including those representing first-branching lineages, thus ensuring that the crown nodes remained comparable), in order to explore possible bias in age estimations according to numbers of taxa sampled.

Fossil calibration

The oldest unambiguously identifiable fossil Annonaceae remains have been found in the Maastrichtian of Nigeria.
(seeds with perichalazal ring and lamelliform ruminant endosperm; Chesters, 1955) and Colombia (reticulate monosulcate pollen; Solé de Porta, 1971). The Maastrichtian seeds were used by both Richardson et al. (2004) and Doyle et al. (2004) as one possible calibration point within Annonaceae, placed at the stem node of the LBC and SBC. Although endosperm ruminant in the basal grade of Angiosperms appears mostly irregular, taxa of the Ambavioioid clade were described as ‘lamellate, woody’ by Van Setten & Koek-Noorman (1992), and arguably lamelliform ruminant was observed by the authors in a species of Tetrameranthus. We considered the precise placement of both fossils on the Annonaceae phylogeny to be sufficiently ambiguous to prefer use of the fossil taxon Archaeanthus (Düchér & Crane, 1984), following Doyle et al. (2004) and Richardson et al. (2004), to assign a minimum age of 98 Ma to the stem node of Magnoliaceae (due to the distinctive stipules, elongate receptacle and fruits). This interpretation is not entirely uncontroversial: Archaeanthus has recently been explicitly excluded from age estimations in angiosperms by Crepet et al. (2004). They instead used two fossil flowers, Cronquistiflora and Detrusandra (Crepet & Nixon, 1998) to impose the more conservative minimum age of 90 Ma on the Magnoliaceae. However, results of Doyle et al. (2004) and Richardson et al. (2004) broadly agree with ages estimated in angiosperm-wide studies (Wikström et al., 2001; Davies et al., 2004) which suggest the calibration of Crepet et al. may represent a (greater) underestimation of the true age.

In order to test whether the sequence data of matrix 1 exhibited clock-like behaviour, a likelihood ratio test was performed on the first of the above (constrained) most parsimonious tree topologies. Likelihoods of the data with and without constraint of a molecular clock were calculated, and the likelihood ratio statistic compared with \( \chi^2 \) critical value with 94 degrees of freedom (number of taxa minus 2).

**Molecular dating using nonparametric rate smoothing and penalized likelihood**

This substitution model selected using `ModelTest` was used to calculate branch lengths for the above topologies, based on the original data using the maximum likelihood criterion as implemented in `paup*`. Confidence limits on branch lengths, reflecting stochasticity in the sampling of character changes (substitutional noise), were estimated by 100 replicates of bootstrap resampling (as also described by Wikström et al., 2001), with subsequent maximum likelihood branch-length estimation on the constrained tree topology for each bootstrap replicate. This resulted in 100 trees comprising a range of estimated lengths for each branch of the topology. Thereafter, Sanderson’s methods of nonparametric rate smoothing (NPRS) (Sanderson, 1997) and penalized likelihood (PL) (Sanderson, 2002a) were applied as implemented in the software package r8s (Sanderson, 2002b) in order to estimate divergence times. Divergence times were estimated for nodes representing the MRCAs (crown groups) of each of the four genera, and those representing the MRCAs of a number of further SBC sub-clades (see below; Table 3), with the Magnoliaceae stem lineage (the root node once the initial outgroup has been pruned out in the analyses) fixed at 98 Myr old. Analyses were performed on the trees with branch lengths derived from the original data, and for each of the 100 trees each with branch lengths derived from bootstrap resampled data. The results of the latter analyses were summarized giving mean values with SD for specified nodes using the ‘profile’ command in r8s.

**Bayesian dating using a Bayesian technique**

Bayesian molecular dating was performed following Rutschmann (2004), Renner (2004), and the `paml` (Yang, 1997) and `multidivtime` (Thorne & Kishino, 2002) manuals. Nucleotide substitutions in the combined sequence data were estimated using `paml`’s ‘baseml’ program and the F84 + G model (with five rate categories), with the single topology as above. Using the `multidivtime` package, each baseml output was converted using ‘pamlmodelinf’ for input in ‘estbranches’, to estimate branch lengths and calculate the variance–covariance structure of those estimates. These were then used as input for ‘multidivtime’ to calculate node divergence times with the following settings: 100,000 generations of the Markov chain, sampled every 10 after a burn-in of 10,000. Prior number of time units between tip and root: 98 (Myr), SD of prior: 98, prior rate at root node: 0.0003 (derived from r8s results using PL), nu: 1, SD nu: 1, and the single constraint on node times was the same as the calibration in the r8s analyses as above.

**RESULTS**

**Matrix 1 (96 taxa)**

Both 100 and 1000 iterations of the parsimony ratchet recovered numerous trees of 2137 steps, consistency index (CI) = 0.648, retention index (RI) = 0.777. Sequence and alignment lengths, numbers of variable and parsimony informative characters and tree statistics for matrix 1 are presented in Table 1. The marker `rbcL` provided the lowest total number of parsimony informative characters, despite representing higher numbers of sequenced bases, and exhibited significantly higher levels of homoplasy compared with the other two markers.

The best fitting substitution model for all sequence data of matrix 1 estimated using `ModelTest` was **GTR + I + Γ**. Best fitting models for each marker individually are presented in Table 1. Maximum parsimony bootstrap analysis and Bayesian inference of matrix 1 resulted in congruent consensus topologies, that of Bayesian inference being significantly more resolved. Both results are presented in Fig. 3, the Bayesian 50% majority rule consensus, with posterior probabilities above the nodes and MP bootstrap support (BS) values below.
Monophyly of *Cremastosperma*, *Malmea* and *Mosannona* was confirmed at 100% BS. A clade including all accessions of *Klarobelia* plus *Pseudephedranthus* received 90% BS. Further clades supported by > 50% BS correspond largely to those revealed by Richardson et al. (2004), except for a clade including all the ‘South American-centred’ SBC taxa (except a single accession, unidentified to genus, falling with 100% BS within the miliusoid/ *Monocarpia* clade) with 63% BS. This is referred to as the ‘SAC clade’. Sampling of taxa within the miliusoid clade was not sufficient to infer further the position of the unidentified accession, and it was omitted from further character sampling. One clade, that of *Malmea* as sister group to *Cremastosperma*, received 97% posterior probability, but < 50% BS (see Matrix 2 below).

**Matrix 2 (23 taxa)**

Maximum parsimony branch and bound search of matrix 2 excluding *rbcL* sequence data resulted in a single tree of 1555 steps, CI = 0.871, RI = 0.791, presented in Fig. 4 (with posterior probabilities above the nodes and BS values and Bremer support below). Inclusion of *rbcL* resulted in three trees of 1754 length, CI = 0.851, RI = 0.761. Sequence and alignment lengths, numbers of variable and parsimony informative characters and tree statistics for matrix 2 are presented in Table 2. Levels of homoplasy in *rbcL* were higher than those of all the other markers. The CI values for *rbcL*, *trnL-trnF* and *psbA-trnH* were higher in matrix 2 than in matrix 1, which is to be expected given the lower taxon sampling density; less homoplasy is revealed. Values for RI were lower in all three markers in matrix 2, but that for *rbcL* decreased markedly. Analyses were thus performed on matrix 2 with and without the *rbcL* sequence data, and results compared.

Best fitting substitution models for each marker as estimated by **MOSELT** are presented in Table 2. Bootstrap analysis of matrix 2 resulted in a fully resolved consensus tree, although one node received only weak support (Fig. 4). The SAC clade received strong support (87% BS). A sister group relationship between *Pseudeoxandra* and *Cremastosperma* was also strongly supported (100% BS), contradicting the results of Bayesian (but not of MP) analysis of matrix 1.

The bootstrap consensus when including *rbcL* included two polytomies. The two nodes with Bremer support of only one step without *rbcL* (and subject to moderate-to-weak BS) were no longer recovered when *rbcL* was included. These were the node of the *Mosannona/Klarobelia* clade as sister to the *Cremastosperma/Malmea/Pseudeoxandra* clade (58% BS), and that of the miliusoid/ *Monocarpia* clade as sister to the SAC clade (75% BS). Bootstrap support values across the topology were slightly lower when *rbcL* was included (data not shown). Bayesian analysis recovered a congruent topology, although with one polytomy: relationships between the two African clades and the rest of the ingroup were not resolved.

**Molecular dating**

The likelihoods of the sequence data of matrix 1, given the constraint topology, were significantly different according to clock-constrained and unconstrained substitution models ($P < 0.01$), indicating rate heterogeneity: the rejection of a molecular clock. Nonparametric rate smoothing, PL and Bayesian (using **MULTDIVTIME**) methods were thus applied to produce full (a) and reduced (b) taxon-sampled ultrametric trees. The cross-validation test for PL resulted in a smoothing parameter of 31.62. Maximum ages for the MRCA of the four genera were estimated at $c.20–27$ Ma using NPRS; $13–16$ Ma using PL; and $23–28$ Ma (95% probability) using **MULTDIVTIME**. The oldest estimations for the Asian/Neotropical crown group were $c.42$ Ma (NPRS, point estimation plus SD), $36$ Ma (PL, point estimation plus SD) and $51$ Ma (**MULTDIVTIME**, 95% probability). Age estimations with SD using NPRS and PL, and mean age estimations with SD and 95% confidence limits derived using **MULTDIVTIME**, are presented in Table 3 for key nodes referred to in the Discussion. A chronogram for the SBC (based on results using PL) with timing of geological events referred to in the Discussion is presented in Fig. 5.

The bounds for age estimations, according to SD or 95% confidence limits, overlapped for all nodes compared across the three different methods, and both taxon selections. For results using NPRS and **MULTDIVTIME**, all age estimations for nodes including all 96 taxa (a) were older than those

---

**Table 1** Details of matrix 1, maximum parsimony search and best fitting substitution models

| Marker     | Sequence length | Alignment length | Variable characters | Parsimony informative characters | Parsimony informative indels | CI/RI* | Model                |
|------------|-----------------|------------------|---------------------|---------------------------------|-----------------------------|--------|----------------------|
| *rbcL*†    | 1470–1480       | 1496             | 310                 | 187                             | 0                           | 0.519/0.712 | GTR + I + G       |
| *trnL-trnF*| 912–996         | [1402†]          | 454                 | 270                             | 30                          | 0.729/0.810 | K81uf + G        |
| *psbA-trnH*| (274§) 412–511  | 798              | 340                 | 215                             | 10                          | 0.689/0.812 | TVM + G          |

*CI, Consistency index; RI, retention index, optimized onto the combined topology.
†Including 3′ non-coding region.
‡Annona muricata 0525: large deletion.
§Alignment length of Chatrou et al. (unpubl. data).
produced when only 81 were included (b). The largest apparent difference was observed in the *Cremastosperma* crown node, from which clade the largest number of accessions had been excluded in (b). In contrast, ages produced for the two taxon selections using PL were almost identical, with no consistent difference apparent. More

Figure 3 Summary of results of phylogeny reconstruction using maximum parsimony and Bayesian analysis: matrix 1. Maximum parsimony bootstrap support values below nodes, Bayesian analysis posterior probabilities above.
rigorous tests are required to determine whether the results of any of these methods are sensitive to levels of taxon sampling. In this study, the widest bounds of the age estimations produced using the different methods are interpreted as the most stringent test for the biogeographical hypotheses.

**DISCUSSION**

*‘Backbone’ phylogeny of the SBC*

The resolution obtained in this study suggests that sampling a larger number of characters for a careful selection of placeholder taxa would be an efficient approach for resolving relationships between further clades identified on the basis of wider taxon sampling in Annonaceae. A number of previously unidentified clades were discovered as result of analysis of matrix 2: sister group relationships between *Cremastosperma* and *Pseudoxandra* and between the *Cremastosperma/Pseudoxandra* clade and *Malmea*, and the SAC clade. The latter result, in particular, has important implications: in the first instance, optimization of ancestral areas results in the stem nodes of each of the clades forming the SAC clade inferred to be Neotropical. Consequently, the minimum age that can be inferred for the presence of their ancestral lineages in the Neotropics is greatly constrained.

However, despite the large amount of DNA sequence data analysed, the numbers of characters supporting the ‘backbone’ nodes were low. Two nodes in particular are subject to weak or moderate support, and no longer recovered when *rbcL* was included in the analyses: the node of the *Mosannona/Klarobelia* clade as sister to the *Cremastosperma/Malmea/Pseudoxandra* clade, and that of the miliusoid/*Monocarpia* clade as sister to the South American-centred clade. The latter uncertainty effectively renders ambiguous the optimization of ancestral areas for the nodes subtending the Asian and SAC clades, increasing the maximum possible age for the presence of the SAC clade in the Neotropics (see below; Fig. 5).

**Figure 4** Summary of results of phylogeny reconstruction using maximum parsimony and Bayesian analysis: matrix 2 (excluding *rbcL* sequence data). Single most parsimonious topology presented with maximum parsimony bootstrap and Bremer support values below nodes, Bayesian analysis posterior probabilities above.

**Table 2** Details of matrix 2, maximum parsimony search and best-fitting substitution models

| Marker            | Sequence length | Alignment length | Variable characters | Parsimony informative characters | Parsimony informative indels | CI/RI* | Model          |
|-------------------|-----------------|------------------|---------------------|---------------------------------|------------------------------|--------|----------------|
| *rbcL†*           | [1470–1480]     | [1496‡]          | 127                 | 59                              | 0                            | 0.688/0.554 | TrN + I + G   |
| *matK*            | 843             | 843              | 142                 | 69                              | 0                            | 0.881/0.833 | K81uf + G    |
| *ndhF*            | 2043            | 2043             | 370                 | 170                             | 0                            | 0.840/0.763 | K81uf + G    |
| *trnT-trnL*       | 641–947         | 1157             | 178                 | 68                              | 8                            | 0.877/0.745 | TVM + G      |
| *trnL-trnF*       | [912–996‡]      | [1402§]          | 174                 | 64                              | 7                            | 0.907/0.816 | K81uf + G    |
| *psbA-trnH*       | [(274*) 412–511] | [798‡]           | 120                 | 50                              | 1                            | 0.838/0.782 | K81uf + G    |
| *trnS-trnG*       | (265–**)        | 643–851          | 953                 | 134                             | 50                           | 0.890/0.769 | TIM + G      |
| *atpB-rbcL*       | 745–770         | 828              | 135                 | 61                              | 4                            | 0.918/0.874 | K81uf + I    |

*CI, Consistency index; RI, retention index, optimized onto the combined topology.
†Including 3′ non-coding region.
‡Alignment/length derived from matrix 1.
§Alignment length of Chatrou et al. (unpubl. data).
*Annona muricata* 0525: large deletion.
**Malmea dielsiana** 0260: large deletion.
Biogeographical history in the SBC of Annonaceae: tracking the origins of Andean-centred genera on the South American continent

The oldest estimations of the age of the SBC crown group produced here (58.76 Ma), in agreement with estimations of Richardson et al. (2004) (53.1–62.5 Ma, SD = 3.6), were significantly younger than the timing of the Africa–South America break-up (c. 100 Ma; Burnham & Graham, 1999). The age of the MRCA of the SAC clade (Fig. 5, indicated by an arrow and A), estimated at between c. 45 and 16 Ma here, represents the time after which this lineage can be said with certainty to have been present in the Neotropics. According to the SBC topology presented here, the actual age of the first Neotropical ancestor of the SAC clade could date back as far as the SAC clade stem node (Fig. 5, B: 51–16 Ma). However, should the miliusoid/Monocarpia clade prove instead to be sister to the P. hypoleuca clade, this would push this age back as far as that of the Piptostigma/Greenwayodendron/Neotropics/Asia node (Fig. 5, C: 57–26 Ma).

Neither of these ages is ancient enough to allow explanation of the distribution of the major clades within the SBC through either the splitting of west Gondwana, or transatlantic dispersal from Africa to South America (which, according to Morley, 2003, may have been possible across island chains up to c. 76 Ma). However, the minimum ages are older than the first point at which North and South America were directly connected by the present-day Central American land bridge (Fig. 5) (3.5–3.1 Ma; Burnham & Graham, 1999).

If these conclusions are interpreted as support for a Boreotropical dispersal of SBC sub-clades, then South American origin of the SAC would have to imply dispersal

### Table 3: Age estimations derived using DNA sequence data of matrix 1 and the nonparametric rate smoothing (NPRS), penalized likelihood (PL) and Bayesian (multidivtime) methods:

|                         | NPRS (A) | NPRS (B) | PL (A) | PL (B) | multidivtime (A) | multidivtime (B) |
|-------------------------|----------|----------|--------|--------|------------------|-----------------|
| Cremastosperma          | 22.33    | 13.65    | 7.17   | 7.54   | 16.49            | 9.07            |
| SD                      | 4.99     | 4.99     | 8.53   | 5.83   | 5.14             | 3.84            |
| 95%                     |          |          | 8.01–27.73 | 3.45–18.20 |
| Malmea                  | 15.90    | 13.09    | 9.54   | 9.88   | 13.42            | 10.47           |
| SD                      | 4.23     | 3.29     | 5.35   | 5.02   | 4.97             | 4.04            |
| 95%                     |          |          | 5.49–24.76 | 4.28–19.94 |
| Klarobelia              | 15.96    | 10.71    | 7.69   | 7.46   | 12.51            | 9.25            |
| SD                      | 3.31     | 3.01     | 5.20   | 4.16   | 4.59             | 3.58            |
| 95%                     |          |          | 5.25–23.12 | 3.72–17.64 |
| Mosannona               | 17.38    | 10.22    | 8.62   | 7.88   | 13.58            | 10.18           |
| SD                      | 3.28     | 2.72     | 4.45   | 3.77   | 4.87             | 3.82            |
| 95%                     |          |          | 6.08–24.85 | 4.62–19.22 |
| Mosannona west of Andes (1) | 12.67  | –        | 6.04   | –      | 8.76             | –               |
| SD                      | *        | *        |        |        | 3.76             |                 |
| 95%                     |          |          | 3.26–17.78 |        |
| Mosannona east of Andes (2) | 11.98 | –        | 5.27   | –      | 5.97             | –               |
| SD                      | *        | *        |        |        | 3.19             |                 |
| 95%                     |          |          | 1.35–13.69 |        |
| Mosannona/Klarobelia clade | 31.16 | 25.57    | 19.42  | 19.60  | 26.40            | 22.15           |
| SD                      | 2.95     | 2.60     | 8.21   | 7.91   | 6.08             | 5.40            |
| 95%                     |          |          | 16.13–39.81 | 12.89–33.82 |
| SAC clade (node A)      | 37.65    | 33.09    | 24.76  | 25.43  | 31.81            | 26.66           |
| SD                      | *        | *        |        | *      | 6.23             | 5.82            |
| 95%                     |          |          | 21.23–45.44 | 16.21–39.20 |
| Neotropics/Asia (node B) | 39.23   | 34.87    | 26.29  | 27.07  | 36.72            | 31.70           |
| SD                      | 3.15     | 2.77     | 9.64   | 9.11   | 6.60             | 6.35            |
| 95%                     |          |          | 25.29–50.93 | 20.25–44.94 |
| Piptostigma/Greenwayodendron/Neotropics/Asia (node C) | 48.44 | 44.78    | 34.89  | 35.63  | 42.76            | 38.13           |
| SD                      | *        | *        |        | *      | 6.83             | 6.64            |
| 95%                     |          |          | 30.81–56.62 | 25.83–51.8  |

*No SD calculated for nodes subtended by zero-length branches in any of the bootstrap replicate trees.
between North and South America prior to the closure of the Panama isthmus. Both animal and plant fossil evidence suggest that this may have been possible (Morley, 2003). The plate-kinematic model of Pindall et al. (1988) is interpreted to suggest two windows of opportunity for dispersal associated with the formation of island chains at the leading edge of the east-drifting Caribbean plate (Morley, 2003; Pennington & Dick, 2004). Firstly, the proto-Greater Antilles formed a bridge between Yucatán and Colombia, c. 50 Ma, which was subsequently fractured as the plate drifted further east. Secondly, a land mass including the Greater Antilles and Aves Ridge (GAARlandia; Iturralde-Vinent & McPhee, 1999) formed around 35–33 Ma, which may have provided a dispersal route for around 3 Myr, before fragmenting to form the present-day Caribbean islands (Morley, 2003). The Panama isthmus itself formed in the Pliocene from the island arc associated with the trailing edge of the Caribbean plate (Pindall et al., 1988).

**Figure 5** Chronogram: branch lengths proportional to time, as estimated using penalized likelihood. Nodes A–C represent minimum ages of a common Neotropical ancestor of the SAC according to different topologies as referred to in the text. Nodes 1 and 2 represent the clades in Mosannona found west and east of the Andes, respectively. Ages are subject to SD (where estimated) of 4.45–9.64 Myr.
These two pre-Panama dispersal opportunities appear not to have been utilized by ancestors of the genera Desmopsis, Sapranthus, Stenanona and Tridimeris, which are nested with high support within the otherwise almost exclusively Asian miliusoid clade (Mols et al., 2004). Both Sapranthus and Tridimeris are endemic to Central America, as are most species of Desmopsis and Stenanona: the few South American species of these genera are found only on the Pacific coast of Colombia (P. Maas, pers. comm.), apparently limited in their dispersal into South America by the barrier presented by the Andean mountain chain (see below). Clades within the SAC clade, such as those of Cremastosperma, Klarobelia, Malmea or Mosannona, are diverse in South America. If they originated in Central America and dispersed across the Panama isthmus (within the past 3.5 Myr), it would be difficult to explain why the miliusoid clades did not. More data are needed to obtain more accurate dates for the diversifications of Central American taxa, and thus to understand better how the distributions of SBC taxa were affected by pre-Panama isthmus land bridges. However, in the absence of a fully sampled and completely resolved phylogeny of the SAC clade, we consider that these results provide convincing support for a common ancestor of Cremastosperma, Klarobelia, Malmea and Mosannona on the South American continent.

Monophyly of Cremastosperma, Klarobelia, Malmea and Mosannona

In general, monophyly of the genera sampled was confirmed. A possible exception was the genus Oxandra, the accessions of which fell into two clades plus one isolated lineage. Monophyly was confirmed for Cremastosperma, Malmea and Mosannona. Klarobelia proved broadly monophyletic, with the exception of a single accession. Pseudophedranthus fragrans, representing a monotypic genus from the upper Rio Negro on the Brazil/Venezuela/Colombia border, was nested within Klarobelia. It was sister to a clade of species not represented by Richardson et al. (2004), which together form the sister group to the rest of Klarobelia, thus explaining why its position had not previously been discovered. This result is curious, given the morphological characters (such as impressed as opposed to raised midrib, and open rather than closed bud development) that otherwise appear to represent synapomorphies for Klarobelia to the exclusion of Pseudophedranthus, and also warrants further investigation. The ‘Andean-centred’ distribution patterns thus were not contradicted, although the unique distribution of Pseudophedranthus falls between those of the disjunct localities of other species in the tropical Andes and coastal Venezuela.

Progress towards reconstructing the biogeographical history of Cremastosperma, Klarobelia, Malmea and Mosannona

One biogeographical hypothesis concerning Andean-centred groups such as Cremastosperma, Klarobelia, Malmea and Mosannona is that speciation in these groups occurred as a result of the Andean orogeny (Gentry, 1982). The inferred presence of ancestors of the four clades on the South American continent prior to the closure of the Panama isthmus, and confirmation of their distribution patterns through testing the monophyly of the accessions of these taxa, failed to reject this hypothesis. A further test was applied using molecular dating techniques to examine the possibility that diversifications in these groups occurred within the time frame of the elevation of the Andes. Burnham & Graham (1999) considered the Andean orogeny to have been an influence in the history of Neotropical vegetation since the Miocene. This would place the effective time frame for diversifications that might have been associated with the Andean orogeny at anywhere between c. 23.3 Ma and the present. However, estimations of palaeoelevation of the Central and Colombian Andes suggest that much of the Andean uplift occurred in the late Miocene and Pliocene (Gregory-Wodzicki, 2000), when elevations increased by more than 3500 m (Lundberg & Chernoff, 1992). The central Andes had reached no more than half their modern elevation by 10 Ma, and the Eastern Cordillera of the Colombian Andes was at no more than 40% of its modern elevation by 4 Ma (Gregory-Wodzicki, 2000).

Molecular dating results produced here showed that the ages of MRCA of each of the four genera fall within roughly the same wide time window (the oldest dates we estimated are those using NPRS, which give ages from 22.3 ± 4.99 to 15.96 ± 3.31 Ma). The multidivtime and PL methods suggest even younger ages for diversification in these genera, which fall even more clearly within the time frame of Andean orogeny. These dates are consistent with suggestions by Burnham & Graham (1999) that the Andean orogeny has had an effect on the history of Neotropical vegetation since the Miocene. However, imprecision in the results means that we cannot exclude the possibility that speciation within each of the four clades actually occurred during rather different time slices.

That no one species of any of these four genera occurs on both sides of the Andes provides compelling evidence that the current elevation of the northern Andes forms an effective barrier to dispersal in these groups. Furthermore, preliminary results in Mosannona presented here show that most of the species sampled fall into two clades, one west and one east of the Andes (Figs 3 & 5). If the north-Andean uplift represents a vicariance event common to this and other SAC clade genera, then this should be reflected in their species phylogenies, with congruent timing of subsequent divergences. The age of the divergence between Mosannona clades on the west and on the east of the Andes (nodes 1 and 2 in Fig. 5; Table 3) could be as early as 15 Ma, or rather more recent according to our estimates, which is consistent with the limitation in elevation below which species of the SAC clade are found (mostly 1500 m, very rarely up to c. 2000 m), which would probably place a vicariance explanation for speciation in these groups no earlier than the Pliocene. The absence of miliusoid clade taxa east of the Andes would suggest the timing was prior to the
closure of the Panama isthmus. However, alternative (or additional) causes of speciation (or indeed extinction) in these groups could have been linked to distribution shifts along the Andean elevational range during climatic changes in the Pleistocene (Hooghiemstra & van der Hammen, 1998).

CONCLUSIONS

We consider the fully resolved (although not in all nodes highly supported) topology presented here to represent a credible hypothesis of phylogenetic relationships between the major clades of the SBC – but one that should be further tested with independent data.

The identification of a clade comprising all the SBC genera distributed in South America and generally, to a limited extent, into Central America (the SAC clade), to the exclusion particularly of clades endemic to Asia and Central America, suggests a common origin of the SAC clade in South America. Origin of the SAC clade in South America as a result of dispersal across the Boreotropics is supported by the age estimations presented here, rather than Gentry’s hypothesis of origin by Gondwanan vicariance. Broad monophyly of the four genera leads us to conclude that the distribution patterns as observed are not the arbitrary result of the definition of poly- or paraphyletic groups. The ages estimated for the MRCA of each clade were not significantly different from each other, and appear to fall within the time frame of the orogeny of the Northern Andes, although the strength of this test was limited by imprecision in the molecular dating results.

Further testing of these biogeographical hypotheses requires the reconstruction of species-level phylogenies of Cremastosperma, Klarobelia, Mosannona and Malmea. Further work should concentrate both on finding further age-calibration points for the Annonaceae phylogeny, and on assessing sources of error in the techniques used to derive ultrametric trees. This approach could shed further light on the dynamic processes of invasion of Central America (Chatrou, 1997) and the origin of high species diversity in tropical America.

ACKNOWLEDGEMENTS

The authors would like to express gratitude to Jan Maas and María Paula Balcázar Vargas for help in the laboratory, to Tanya Scharaschkin and Jim Doyle for advice and access to unpublished Annonaceae primers, and to the herbaria of the New York Botanical Garden and the National Herbarium of the Netherlands, Wageningen University branch, for granting permission for the extraction of DNA from herbarium specimens. James Richardson and Jijke Koek-Noorman gave helpful criticism of the manuscript. Distribution maps in Fig. 1 were produced using ESRI data made available by the New York Botanical Garden’s ‘Digital Basemap of the Americas’. Phylogenetic trees in Figs 3–5 were produced using TreeGraph (Muller & Muller, 2004).

REFERENCES

Bremer, K. (1994) Branch support and tree stability. Cladistics, 10, 295–304.
Burnham, R.J. & Graham, A. (1999) The history of Neotropical vegetation: new developments and status. Annals of the Missouri Botanical Garden, 86, 546–589.
Chatrou, L.W. (1997) Studies in Annonaceae XXVIII. Macromorphological variation of recent invaders in northern Central America: the case of Malmea (Annonaceae). American Journal of Botany, 84, 861–869.
Chatrou, L.W. (1998) Changing genera: systematic studies in Neotropical and West African Annonaceae. PhD Thesis, Utrecht University, Utrecht, the Netherlands.
Chatrou, L.W., Rainer, H. & Maas, P.J.M. (2004) Annonaceae. Flowering plants of the Neotropics (ed. by N. Smith, S.A. Mori, A. Henderson, D.W. Stevenson and S.V. Heald), pp. 18–20. Princeton, Englewood Cliffs, NJ, USA.
Chesters, K.I.M. (1955) Some plant remains from the Upper Cretaceous and Tertiary of West Africa. Annals and Magazine of Natural History, 12, 489–504.
Crepet, W.L. & Nixon, K.C. (1998) Two new fossil flowers of magnololid affinity from the Late Cretaceous of New Jersey. American Journal of Botany, 85, 1273–1288.
Crepet, W.L., Nixon, K.C. & Gandolfo, M.A. (2004) Fossil evidence and phylogeny: the age of major angiosperm clades based on mesosossil and macrofossil evidence from Cretaceous deposits. American Journal of Botany, 91, 1666–1682.
Cuénot, P., Savolainen, V., Chatrou, L.W., Powell, M., Grayer, R.J. & Chase, M.W. (2002) Molecular phylogenetics of Caryophyllales based on nuclear rDNA and plastid rbcL, atpB, and matK DNA sequences. American Journal of Botany, 89, 132–144.
Davies, T.J., Barraclough, T.G., Chase, M.W., Soltis, P.S., Soltis, D.E. & Savolainen, V. (2004) Darwin’s abominable mystery: insights from a supertree of the angiosperms. Proceedings of the National Academy of Sciences the United States of America, 101, 1904–1909.
Davis, C.C., Bell, C.D., Mathews, S. & Donoghue, M.J. (2002) Laurasian migration explains Gondwanan disjunctions: evidence from Malpighiaceae. Proceedings of the National Academy of Sciences the United States of America, 99, 6833–6837.
Dilcher, D.L. & Crane, P.R. (1984) Archaeanthus: an early angiosperm from the Cenomanian of the Western Interior of North America. Annals of the Missouri Botanical Garden, 71, 351–383.
Doyle, J.J. & Doyle, J.L. (1987) A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochemical Bulletin, 19, 11–15.
Doyle, J.A. & Le Thomas, A. (1996) Phylogenetic analysis and character evolution in Annonaceae. Adansonia, 18, 279–334.
Doyle, J.A., Bygrave, P. & Le Thomas, A. (2000) Implications of molecular data for pollen evolution in Annonaceae. Pollen and spores: morphology and biology (ed. by M.M. Harley,
C.M. Morton and S. Blackmore), pp. 259–284. Royal Botanic Garden, Kew, UK.

Doyle, J.A., Sausquet, H., Scharaschkin, T. & Le Thomas, A. (2004) Phylogeny, molecular and fossil dating, and biogeographic history of Annonaceae and Myristicaceae (Magnoliidae). *International Journal of Plant Sciences*, 165, S53–S67.

Fay, M.F., Swensen, S.M. & Chase, M.W. (1997) Taxonomic affinities of *Medusagyne oppositifolia* (Medusagynaceae). *Kew Bulletin*, 52, 111–120.

Fay, M.F., Bayer, C., Alversson, S., de Brujin, A.Y. & Chase, M.W. (1998) Plastid *rbcL* sequence data indicate a close affinity between *Diegodendron* and *Bixa*. *Taxon*, 47, 43–50.

Fitch, W.M. (1971) Toward defining the course of evolution: minimum change for a specified tree topology. *Systematic Zoology*, 20, 406–416.

Fries, R.E. (1937) Revision der Arten einiger Annenaceen-Gattungen IV. *Acta Horti Bergiani*, 12, 221–288.

Gentry, A.H. (1982) Neotropical floristic diversity: phytogeographical connections between Central and South America, Pleistocene climatic fluctuations, or an accident of the Andean orogeny? *Annals of the Missouri Botanical Garden*, 69, 557–593.

Gregory-Wodzicki, K.M. (2000) Uplift history of the Central and Northern Andes: a review. *GSA Bulletin*, 112, 1091–1105.

Hamilton, M.B. (1999) Four primer pairs for the amplification of chloroplast intergeneric regions with intraspecific variation. *Molecular Ecology*, 8, 521–523.

Hooghiemstra, H. & van der Hammen, T. (1998) Neogene and Quaternary development of the neotropical rain forest: the forest refugia hypothesis, and a literature overview. *Earth Science Reviews*, 44, 147.

Hoot, S.B., Culham, A. & Crane, P.R. (1995) The utility of *atpB* gene sequences in resolving phylogenetic relationships: comparison with *rbcL* and 18S ribosomal DNA sequences in the Lardizabalaceae. *Annals of the Missouri Botanical Garden*, 82, 194–207.

Huelsenbeck, J.P. (2000) MrBayes: Bayesian inference of phylogeny. Department of Biology, University of Rochester, Rochester, New York, USA.

Iturralde-Vinent, M.A. & McPhee, R.D.E. (1999) Paleogeography of the Caribbean region: implications for Cenozoic biogeography. *Bulletin of the American Museum of Natural History*, 238, 1–95.

Kim, S., Park, C.-W., Kim, Y.-D. & Suh, Y. (2001) Phylogenetic relationships in family Magnoliaceae inferred from *ndhF* sequences. *American Journal of Botany*, 88, 717–728.

Lundberg, J.G. & Cheroff, B. (1992) A Miocene fossil of the Amazonian fish *Arapaima* (Teleostei, Arapaimidae) from the Magdalena River region of Colombia – biogeographic and evolutionary implications. *Biotropica*, 24, 2–14.

Mols, J.B., Gravendeel, B., Chatrou, L.W., Pirie, M.D., Bygrave, P.C., Chase, M.W. & Kessler, P.J.A. (2004) Identifying clades in Asian Annonaceae: monophyletic genera in the polyphtelic Miliuseae. *American Journal of Botany*, 91, 590–600.

Morley, R.J. (2003) Interplate dispersal paths for megathermal angiosperms. *Perspectives in Plant Ecology, Evolution and Systematics*, 6, 5–20.

 Muller, J. & Muller, K. (2004) Treegraph: automated drawing of complex tree figures using an extensible tree description format. *Molecular Ecology Notes*, 4, 786–788.

Myers, N., Mittermeier, C.G., da Fonseca, G.A.B. & Kent, J. (2000) Biodiversity hotspots for conservation priorities. *Nature*, 403, 853–858.

Nixon, K.C. (1999) The Parsimony Ratchet, a new method for rapid parsimony analysis1. *Cladistics*, 15, 407–414.

Olmstead, R.G. & Sweere, J.A. (1994) Combining data in phylogenetic systematics: an empirical approach using three molecular data sets in the Solanaceae. *Systematic Biology*, 43, 467–481.

Olmstead, R.G., Michaels, H.J., Scott, K.M. & Palmer, J.D. (1992) Monophyly of the Asteridae and identification of major lineages inferred from DNA sequences of *rbcL*. *Annals of the Missouri Botanical Garden*, 79, 249–265.

Pennington, R.T. & Dick, C.W. (2004) The role of immigrants in the assembly of the South American rainforest tree flora. *Philosophical Transactions of the Royal Society of London, Series B*, 359, 1611–1622.

Pindall, J.L., Cande, S.C., Pitman, W.C., Rowley, D.B., Dewey, J.F., LaBrecque, J. & Haxby, W. (1988) A plate-kinematic frame-work for models of Caribbean evolution. *Tectonophysics*, 155, 121–138.

Pirie, M.D., Chatrou, L.W., Erkens, R.H.J., Maas, J.W., Van der Niet, T., Mols, J.B. & Richardson, J.E. (2005) Phylogeny reconstruction and molecular dating in four Neotropical genera of Annonaceae: the effect of taxon sampling in age estimations. *Plant species-level systematics: new perspectives on pattern and process* (ed. by F.T. Bakker, L.W. Chatrou, B. Gravendeel and P.B. Pelser), pp. 149–174. Regnum Vegetabile 143. A. R. G. Gantner Verlag, Liechtenstein.

Posada, D. & Crandall, K.A. (1998) ModelTest: testing the model of DNA substitution. *Bioinformatics*, 14, 817–818.

Renner, S.S. (2004) Bayesian analysis of combined chloroplast loci, using multiple calibrations, supports the recent arrival of Melastomataceae in Africa and Madagascar. *American Journal of Botany*, 91, 1427–1435.

Renner, S.S., Clausing, G. & Meyer, K. (2001) Historical biogeography of Melastomataceae: the roles of Tertiary migration and long-distance dispersal. *American Journal of Botany*, 88, 1290–1300.

Richardson, J.E., Chatrou, L.W., Mols, J.B., Erkens, R.H.J. & Pirie, M.D. (2004) Historical biogeography of two cosmopolitan families of flowering plants: Annonaceae and Rhamnaceae. *Philosophical Transactions of the Royal Society of London, Series B*, 359, 1495–1508.

Rogstad, S.H. (1989) The biosystematics and evolution of the *Polyalthia hypoleuca* complex (Annonaceae) of Malesia. I. Systematic treatment. *Journal of the Arnold Arboretum*, 70, 153–246.
Rutschmann, F. (2004) Bayesian molecular dating using PAML/multidivtime. A step-by-step manual. University of Zürich, Switzerland, pp. 1–13.

Sanderson, M.J. (1997) A nonparametric approach to estimating divergence times in the absence of rate constancy. *Molecular Biology and Evolution*, 14, 1218–1231.

Sanderson, M.J. (2002a) Estimating absolute rates of molecular evolution and divergence times: a penalized likelihood approach. *Molecular Biology and Evolution*, 19, 101–109.

Sanderson, M.J. (2002b) r8s, version 1.50. Section of Evolution and Ecology, University of California, Davis, CA, USA.

Sauquet, H., Doyle, J.A., Scharaschkin, T., Borsch, T., Hilu, K., Chatrou, L.W. & Le Thomas, A. (2003) Phylogenetic analysis of Magnoliidae and Myristicaceae based on multiple data sets: implications for character evolution. *Botanical Journal of the Linnean Society*, 142, 125–186.

Savolainen, V., Cuénoud, P., Spichiger, R., Martinez, M.D.P., Crèvecœur, M. & Manen, J.-F. (1995) The use of herbarium specimens in DNA phylogenetics: evaluation and improvement. *Plant Systematics and Evolution*, 197, 87–98.

Sikes, D.S. & Lewis, P.O. (2001) PAUPRat. PAUP* implementation of the parsimony ratchet, version 1, beta. Department of Ecology and Evolutionary Biology, University of Connecticut, Storrs, CT, USA.

Simmons, M.P. & Ochoterena, H. (2000) Gaps as characters in sequence-based phylogenetic analysis. *Systematic Biology*, 49, 369–381.

Smith, N., Mori, S.A., Henderson, A., Stevenson, D.W. & Heald, S.V. (2004) *Flowering plants of the Neotropics*. Princeton University Press, Princeton, NJ, USA.

Solé de Porta, N. (1971) Algunos géneros nuevos de polen procedentes de la Formación Guaduas (Maastrichtiense-Paleoceno) de Colombia. *Studia Geologica Salmanticensia*, 2, 133–143.

Sørensen, M.D. (1999) TreeRot, version 2. Boston University, Boston, MA.

Swofford, D.L. (2000) PAUP*. Phylogenetic analysis using parsimony (*and other methods). Sinauer, Sunderland, MA, USA.

Taberlet, P., Gielly, L., Pautou, G. & Bouvet, J. (1991) Universal primers for amplification of the three non-coding regions of chloroplast DNA. *Plant Molecular Biology*, 17, 1105–1109.

Ter Steege, H., Sabatier, D., Castellanos, H., van Andel, T.R., Duivenvoorden, J., Adalardo de Oliveira, A., Ek, R.C., Lilwah, R., Maas, P.J.M. & Mori, S.A. (2000) An analysis of the floristic composition and diversity of Amazonian forests including those of the Guiana Shield. *Journal of Tropical Ecology*, 16, 801–828.

Thorne, J.L. & Kishino, H. (2002) Divergence time and evolutionary rate estimation with multilocus data. *Systematic Biology*, 51, 689–702.

Tuomisto, H. & Ruokolainen, K. (1997) The role of ecological knowledge in explaining biogeography and biodiversity in Amazonia. *Biodiversity and Conservation*, 6, 347–357.

Valencia, R., Balslev, H. & Paz y Miño, G. (1994) High tree alpha-diversity in Amazonian Ecuador. *Biodiversity and Conservation*, 3, 21–28.

Van Setten, A.K. & Koek-Noorman, J. (1992) Studies in Annonaceae. XVII. Fruits and seeds of Annonaceae: morphology and its significance for classification. *Bibliotheca Botanica*, 142, 1–101.

Wikström, N., Savolainen, V. & Chase, M.W. (2001) Evolution of the angiosperms: calibrating the family tree. *Proceedings of the Royal Society of London, Biological Sciences*, 1482, 2211–2220.

Wolfe, J.A. (1975) Some aspects of plant geography of the Northern Hemisphere during the late Cretaceous and Tertiary. *Annals of the Missouri Botanical Garden*, 62, 264–279.

Yang, Z. (1997) PAML: a program package for phylogenetic analysis by maximum likelihood. *Computer Applications in Biosciences*, 13, 555–556.

**BIOSKETCHES**

The authors have all specialized in the systematics of Annonaceae at the National Herbarium of the Netherlands.

**Michael D. Pirie** completed a PhD at the Utrecht University branch in 2005, working on systematics and taxonomy of Neotropical Annonaceae, in particular the genus *Cremastosperma*.

**Lars W. Chatrou** gained a PhD at Utrecht in 1998 and has continued work on the systematics of Annonaceae at Utrecht and the Royal Botanic Gardens, Kew.

**Johan B. Mols** concentrated on Asian Annonaceae of the miliusoid clade, completing a PhD in Leiden in 2004.

**Roy H.J. Erkens** is currently engaged in PhD research working on the species-rich genus *Guatteria*.

**Jessica Oosterhof** worked on Neotropical short-branch clade Annonaceae at Utrecht in 1998–2000.

Editor: Malte Ebach

**SUPPLEMENTARY MATERIAL**

The following supplementary material is available online from http://www.Blackwell-Synergy.com:

**Appendix S1** Details of accessions sampled for DNA sequence data, and *rbcL, trnL-trnF* and *psbA-trnH* GenBank accession numbers.

**Appendix S2** *ndhF, trnT-trnL, trnS-trnG*, and *atpB-rbcL* GenBank accession numbers.