Evolution of genome size in the angiosperms[1]

Douglas E. Soltis,[2,5] Pamela S. Soltis,[3] Michael D. Bennett,[4]
and Ilia J. Leitch[4]

[2]Department of Botany and the Genetics Institute, University of Florida, Gainesville, Florida 32611-5826 USA; [3]Florida Museum of Natural History and the Genetics Institute, University of Florida, Gainesville, Florida 32611-7800 USA; and [4]Jodrell Laboratory, Royal Botanic Gardens, Kew, Richmond, Surrey TW9 3DS, UK

Genome size varies extensively across the flowering plants, which has stimulated speculation regarding the ancestral genome size of these plants and trends in genome evolution. We investigated the evolution of C-values across the angiosperms using a molecular phylogenetic framework and C-values not previously available for crucial basal angiosperms, including *Amborella*, Illiciaceae, and *Austrobaileya*. Reconstructions of genome size across the angiosperms and extant gymnosperms indicate that the ancestral genome size for angiosperms is very small (1C ≤ 1.4 pg), in agreement with an earlier analysis of Leitch et al. (1998). Furthermore, a very small genome size (1C ≤ 1.4 pg) is ancestral not only for the angiosperms in general, but also for most major clades of flowering plants, including the monocots and the eudicots. The ancestral genome of core eudicots may also have been very small given that very low 1C-values appear to be ancestral for major clades of core eudicots, such as Caryophyllales, Saxifragales, and asterids. Very large genomes occur in clades that occupy derived positions within the monocots and Santalales.

Key words: angiosperms; character evolution; C-values; genome size; phylogeny.

Both chromosome number and genome size vary tremendously across the flowering plants, having stimulated considerable speculation regarding both the original genome size and base chromosome number of the angiosperms, with further hypotheses involving trends in genome and chromosomal evolution. Many authors have proposed that the original base chromosome number for angiosperms was low, between $x = 6$ and 9 (e.g., Ehrendorfer et al., 1968; Stebbins, 1971; Raven, 1975; Grant, 1981). However, analysis of the vast range of chromosome sizes encountered in angiosperms shows that genome size, with which this paper is concerned, can vary independently of chromosome number. The amount of DNA in an unreplicated gametic nuclear genome is referred to as the 1C-value. The 1C-value is often loosely referred to as genome size, but strictly speaking, genome size is the amount of DNA in an unreplicated, basic, gametic chromosome set. Genome size equals the 2C nuclear DNA amount divided by ploidal level (Bennett et al., 1998). This formula gives an accurate estimate for individuals with constituent genomes of equal size (e.g., diploids and autopolyploids), but provides only a mean estimate for individuals with constituent genomes of different sizes (e.g., some diploid hybrids and allopolyploids). Note that for polyploids, genome sizes estimated in this way are always smaller than 1C-values. For example, in the diploid *Triticum monococcum* 2C = 12.45 pg, so 1C = 12.45/2 = 6.23 pg, which also equals the genome size. In the tetraploid *T. dicoccum*, in contrast, 2C = 24.05, so 1C = 24.05/2, which equals 12.03 pg, but the genome size is 24.05/4, or 6.01 pg.

C-values have been estimated for approximately 3500 species of angiosperms (Bennett et al., 1998; Bennett and Leitch, 2003), representing over 1% of the approximately 250 000–300 000 species of flowering plants and approximately 48% of all angiosperm families (sensu APG, 1998; APG II, 2003). The

Angiosperm DNA C-values Database (http://www.rbgkew.org.uk/cval/homepage.html) represents the largest collection of nuclear DNA amounts for any group of organisms (reviewed in Leitch et al., 1998). C-values in angiosperms span a huge range. The smallest reported values are for *Cardamine amara* (Brassicaceae; 1C = 0.05 pg; Bennett and Smith, 1991) and *Fragaria* (Rosaceae; 1C = 0.10 pg; Antonius and Ahokas, 1996); the largest value is for *Fritillaria assyriaca* (Liliaceae; 1C = 127.4 pg; Bennett and Smith, 1976).

Despite this enormous range in DNA amount, the basic complement of genes required for normal growth and development appears to be essentially the same, leading to what is referred to as the "C-value paradox" (Thomas, 1971). Several mechanisms have been proposed for this large variation in genome size in the angiosperms. One mechanism is repeated cycles of polyploidy (e.g., Leitch and Bennett, 1997; Soltis and Soltis, 1990, 2000; Otto and Whitton, 2000; Wendel, 2000). This hypothesis is supported by recent genomic evidence (Bowers et al., 2003), as well as by isozyme evidence. A number of basal lineages, as well as some eudicot families, with uniformly high chromosome numbers have numerous duplicated loci, in agreement with ancient polyploidy (Soltis and Soltis, 1990). Even *Arabidopsis* with its very small genome size and low chromosome number appears to be an ancient polyploid (*Arabidopsis* Genome Initiative, 2000; Grant et al., 2000; Walbot, 2000; Bowers et al., 2003). Transposable elements also appear to contribute to increases in genome size throughout eukaryotes (e.g., Bennetzen, 2000; Kidwell, 2002), perhaps through the large-scale accumulation of retroelements, as in Poaceae (Bennetzen, 1996, 2000; SanMiguel et al., 1996), and some nonangiosperm lineages, such as *Pinus* (Elsik and Williams, 2000).

Leitch et al. (1998) provided a histogram of C-values for 2802 species, which not only revealed a strongly skewed distribution of genome sizes, but also a very small modal size (0.7 pg). Based on these results, they designated those species with C-values of ≤1.4 pg and ≤3.5 pg (twice and five times the modal C-value for angiosperms) as having very small and small genome sizes, respectively. Species with C-values of

>14.0 pg and >35.0 pg (values 20 and 50 times the modal C-value for angiosperms) were defined as having large and very large C-values, respectively. We have reanalyzed the frequency of C-values for a larger data set of 3543 angiosperms (see Supplementary Data accompanying the online version of this article). Our findings are essentially identical to those of Leitch et al. (1998), although the modal size we calculated for angiosperms (0.6 pg) is slightly smaller than the value they obtained. We therefore followed the terminology developed initially by Leitch et al. (1998), with the addition that C-values of 3.51–13.99 are here defined as intermediate.

Leitch et al. (1998) considered genome size in angiosperms in light of a phylogeny, although they did not reconstruct character evolution. Despite the enormous range in nuclear DNA amount, most angiosperms actually had small 1C-values, between 0.1 and 3.5 pg (Leitch et al., 1998). Every higher-level group for which data were available at that time contained species with relatively small C-values (≤3.5 pg). In contrast, species with large genomes (≥14.0 pg) had a much more restricted distribution, being found only in some monocots, as well as some Ranunculales (some Papaveraceae and Ranunculaceae), Caryophyllales (one report for Droseraceae), rosids (e.g., some Brassicaceae, Rutaceae, Onagraceae), asterids (e.g., some Rubiaceae, Solanaceae, Asteraceae), and Santalales (Viscaceae). However, most members of these groups actually had small- or intermediate-sized genomes; only two groups contained members with very large genomes (≥35 pg), Santalales and monocots. They also concluded that there was a tendency for species with large genomes to be restricted to the more derived families within each of these groups.

The sample of 2802 species used by Leitch et al. (1998) included C-values for 1794 diploids, 658 polyploids, and 350 species of unknown ploidy. Importantly, Leitch et al. repeated the analysis for the 1794 species listed as diploids and on this broad scale of comparison obtained essentially the same result as for all 2802 species. Thus, species with large C-values (≥14.0 pg) occurred in only the same six groups, of which only the same two (monocots and Santalales) had very large C-values (≥35.0 pg).

Based on their analyses of C-values and a visual comparison to the *rbcL* topology for angiosperms (Chase et al., 1993), Leitch et al. (1998) concluded that the most parsimonious explanation for these observations was that the ancestral angiosperms had small genomes and that the possession of large genomes was derived. Leitch et al. (1998, 2001) further concluded that within extant seed plants the possession of a small genome was unique to the angiosperms; extant gymnosperms are generally characterized by larger C-values than angiosperms.

In the detailed analysis of Leitch et al. (1998), minimum, maximum, and mean C-values were assigned to families, or larger clades, such as Ranunculales and asterids (see Fig. 2 and Table 1 in Leitch et al., 1998). These values were not adjusted for ploidy, although an additional analysis on diploids alone showed essentially the same result (see earlier). Frequently, the highest C-values in a genus or family are for polyploids; therefore, taking ploidy into consideration when making inferences regarding evolution of genome size is important. For example, in *Magnolia*, there are three estimates of genome size (1C = 0.90, 5.98, and 7.1). *Magnolia kobus* has a 1C-value of 0.9 pg and is diploid, with 2n = 38, the lowest number for *Magnolia*. Importantly, *Liriodendron*, the sister group of *Magnolia*, also with 2n = 38, has a 1C-value of 0.80

pg, which is comparable to *M. kobus*. The higher C-values for *Magnolia* are two reports (5.98 and 7.1 pg) for *M. soulangiana*, with 2n = 76. The higher 1C-values for this species would therefore be attributed to polyploidy. (However, the genome sizes for the tetraploid *M. soulangiana* are much higher than a simple doubling of the 1C-value for the diploid *M. kobus*.) When *Magnolia kobus* and *Liriodendron* are used in MacClade reconstructions, the ancestral state for Magnoliaceae is unambiguously reconstructed as a very small genome (<1.4 pg). In contrast, when a mean value is used, the family is reconstructed as having a small (1.5–3.6 pg), rather than very small genome. We therefore reasoned that genome sizes of only diploids should be used whenever possible, so as not to obscure patterns of variation and hinder character-state reconstruction.

The study by Leitch et al. (1998) provided the first large-scale phylogenetic perspective on genome-size evolution in the angiosperms. Despite the important results of Leitch et al. (1998), there have been two significant developments that indicate that a reevaluation of the diversification of 1C-values in angiosperms is now timely. Firstly, C-value data for some crucial taxa (especially the basal-most angiosperms) are now available for the first time because of work at the Royal Botanic Gardens, Kew (e.g., Hanson et al., 2001a, b; Leitch and Hanson, 2002). The only early-branching angiosperm family represented in Leitch et al. (1998) was Nymphaeaceae. C-values for two species of *Nymphaea* (Nymphaeaceae) are both very small (1C = 0.60 and 1.10 pg). Data are now available for other early-branching families (e.g., Amborellaceae, Illiciaceae, Trimeniaceae, Austrobaileyaceae). The 1C-value of Amborellaceae was recently determined and is very small (1C = 0.89 pg; Leitch and Hanson, 2002). However, other early-branching angiosperms (e.g., members of Austrobaileyales) have larger 1C-values than do Amborellaceae and Nymphaeaceae: 1C = 3.40 pg in *Illicium anisatum* (Illiciaceae); 1C = 7.4–8.9 pg in *Kadsura* (Schisandraceae); 1C = 9.52 in *Austrobaileya* (Austrobaileyaceae); 1C = 4.08 in *Piptocalyx* (Trimeniaceae). Secondly, the tree used by Leitch et al. (1998) was based on the shortest *rbcL* trees of Chase et al. (1993); however, this *rbcL* topology does not reflect our current understanding of relationships among basal angiosperms. The Chase et al. topology placed Ceratophyllaceae as sister to all other angiosperms, rather than the basal grade of Amborellaceae, followed by Nymphaeaceae, and a clade of Illiciaceae/Schisandraceae, Trimeniaceae, and Austrobaileyaceae (= Austrobaileyales, sensu APG II, 2003), identified in a series of recent studies (e.g., Mathews and Donoghue, 1999; Qiu et al., 1999; Soltis et al., 2000; Savolainen et al., 2000; Bowe et al., 2000; Zanis et al., 2002). Recent topologies (e.g., Qiu et al., 1999; Chase et al., 2000; Soltis et al., 2000; Zanis et al., 2002) provide greater resolution and internal support of relationships than earlier single-gene analyses (e.g., Chase et al., 1993). Further, in contrast to Leitch et al. (1998), we have used character-state mapping with parsimony and MacClade (Maddison and Maddison, 1992), enabling us to reconstruct the ancestral genome sizes for clades within the angiosperms.

MATERIALS AND METHODS

Genome size estimates—We used genome-size estimates for known diploids whenever possible in our calculations and avoided using means for large clades, such as families and orders; rather we typically worked at the level of genus (we also experimented with the use of mean values for large clades;

see later). All 1C-values have been published previously (see Appendix, available as Supplementary Data accompanying the online version of this article); all are available either in the Angiosperm DNA C-values Database (Bennett and Leitch, 2003), Hanson et al. (2001a, b), or Leitch and Hanson (2002).

Reconstruction of genome size evolution—We reconstructed the evolution of 1C-values using parsimony and MacClade versions 3.04 and 4.0 (Maddison and Maddison, 1992). We placed C-values into the same categories designated by Leitch et al. (1998; see review earlier). Species with 1C-values and genome sizes of ≤1.4 pg and ≤3.5 pg were defined as having very small and small genomes, respectively. 1C-values of 3.51–13.99 pg were considered intermediate; species with C-values and genome sizes of ≥14.0 pg and ≥35.0 pg (which is 20 and 50 times the modal C-value) were defined likewise as having large and very large 1C-values, respectively. We used ACCTRAN, DELTRAN, and the "all most parsimonious states" options in our reconstructions.

Because C-value is a quantitative and continuous character, we also reconstructed the evolution of 1C-value using squared-change parsimony, which is also implemented in MacClade. Squared-change parsimony requires a fully resolved tree. Hence, we therefore explored several alternative resolutions of relationships in those instances in which polytomies are present. For example, we employed alternative resolutions among the major clades of eudicots, as well as within the asterids and monocots. Ultimately, these alternative topologies had no impact on the reconstruction of 1C-value.

Phylogenetic tree—As a framework for character reconstruction we constructed a grarfted supertree for the angiosperms, using the recent jackknife consensus topology (which shows only those clades having ≥50% support) based on three genes as a backbone (Soltis et al., 2000), with modifications among basal angiosperms (Qiu et al., 1999; Zanis et al., 2002), monocots (Chase et al., 2000), core eudicots (Soltis et al., 2003), Saxifragales (Fishbein et al., 2001), asterids (Albach et al., 2001; Bremer et al., 2002; Soltis et al., 2003), and Caryophyllales (Cuénoud et al., 2002). For Papaveraceae, which occupies a pivotal position near the base of Ranunculales, we added taxa following Hoot et al. (1997, 1999). For asterids, we added representatives for the early-diverging lineages Cornales and Ericales. The uncertainty of phylogenetic relationships among major subclades of eurosids precludes an accurate reconstruction of ancestral genome size in this clade.

Extant gymnosperms were used as an outgroup; their topology reflects the results of recent molecular phylogenetic analyses (e.g., Bowe et al., 2000; Chaw et al., 2000; Soltis et al., 2002) in which cycads followed by *Ginkgo* are sister to Gnetales and conifers. We placed Gnetales sister to the two conifer representatives employed, *Pinus* and *Larix* (Pinaceae).

Because the monocot clade comprises approximately 22% of all angiosperm species (Drinnan et al., 1994), a separate reconstruction of genome-size evolution was conducted for the monocots (Fig. 2). In the summary tree employed, *Ceratophyllum* was placed sister to the monocots (Zanis et al., 2002). The general monocot topology follows recent analyses (Chase et al., 2000; Soltis et al., 2000). *Acorus* was placed as sister to all other monocots (e.g., Qiu et al., 1999; Chase et al., 2000; Soltis et al., 2000); following *Acorus*, a clade of Alismataceae, Araceae, Tofieldiaceae, Hydrocharitaceae, and Zosteraceae was placed as sister to all remaining monocots (Chase et al., 2000). Within Araceae, the topology of Salazar et al. (unpublished data) was considered. The remainder of the topology follows Chase et al. (2000).

Taxon sampling for the trees was determined by the availability of estimates of genome size. In most cases, we provide values for genera (Fig. 1; Appendix, available as Supplementary Data accompanying the online version of this article). To simplify the topology for Fig. 1, rather than showing all genera for which values are available, we sometimes used some families as placeholders in our reconstructions. In the eudicots, family exemplars were used in the asterid, Caryophyllales, and Ranunculales clades. For the eudicot families indicated in Fig. 1 and the Appendix, variation in genome size at the diploid level was low. For example, several 1C-values have been reported for Cornaceae, Menispermaceae, Saxifragaceae, Droseraceae, Polygonaceae, and Portulacaceae, but the values for diploids are all "very small." For other families (i.e., Caryophyllaceae, Apiaceae, and the monocot family Araceae), 1C-values are provided based on reconstructions using multiple genera. In

Caryophyllaceae, 1C-values for diploids are very small with the exception of *Silene*, which has a value in the small range, but a molecular tree for the family (Smissen et al., 2002) suggests that this represents a derived condition. Similarly, genome-size estimates for most Apiaceae are very small at the diploid level, with several genera having 1C-values in the small range. Recent topologies (Plunkett and Downie, 1999) indicate, however, that these genera are in derived positions; reconstructions for the family indicate an ancestral genome that was "very small." For Papaveraceae, we reconstructed the ancestral genome size using recent topologies for the family (Hoot et al., 1997, 1999). For Ranunculaceae, we also attempted to reconstruct an ancestral genome size (as for Papaveraceae), but relationships within the family remain uncertain and a range of values occurs for some genera (e.g., *Ranunculus*; see later). Families and larger clades state (e.g., Ranunculaceae, Ericaceae, Asterales, many monocot lineages) exhibiting variation in 1C-values and for which we could not reliably reconstruct an ancestral state, were considered polymorphic; we employed all character states for each variable family or higher-level clade. In separate reconstructions we also employed mean values, but these different codings had no impact on our major conclusions and only a minor influence on the reconstructions for major clades.

In our broad angiosperm analysis, we represented the monocot clade with families (e.g., Hydrocharitaceae, Aponogetonaceae), orders (e.g., Liliales, Asparagales), and higher categories (i.e., commelinids) as placeholders. We coded these groups as polymorphic, using the entire range of values reported for all members of each clade. We also explored the impact of using mean values. However, this had no impact on the reconstruction of the ancestral state for the monocots as a whole. As noted later, we attempted to reconstruct the ancestral state for Araceae, using only those values reported for early-diverging members of the family.

Because of the small number of C-values reported for many lineages, as well as phylogenetic uncertainty, we did not attempt to reconstruct the ancestral genome sizes of most monocot families and higher-level clades. Instead, in our focused analyses of the monocots (Fig. 2), we used the range of 1C-values for families and higher-level clades (values taken from Leitch et al., 1998). In a separate analysis we also employed mean values. Again, however, the use of several values (polymorphic) vs. mean values had no impact on the reconstruction of the ancestral state for the monocots as a whole.

RESULTS AND DISCUSSION

Ancestral genome size in angiosperms and evolution in basal angiosperms—Using a gymnosperm clade as outgroup, the ancestral genome size of the angiosperms is reconstructed as very small using parsimony (Fig. 1), regardless of the transformation option used (ACCTRAN, DELTRAN, all most parsimonious states). We also obtained similar results using squared-changed parsimony and ML optimizations (not shown). Extant gymnosperms are reconstructed as having an ancestral genome in the intermediate range. It is unclear from our analyses, however, if the origin of the angiosperms was accompanied by a decrease in genome size or if this occurred earlier in land plant evolution. In their reconstruction of genome size diversification across all land plants, I. J. Leitch et al. (unpublished manuscript) found evidence for several independent decreases in genome size across the diversity of land plants. One such decrease may have occurred in the ancestor of the angiosperms, but this result depends on the trace option employed. Our data do indicate that a decrease in genome size occurred in the ancestor of *Gnetum* (Fig. 1), although the genome sizes in *Gnetum* are not as small as those in many angiosperms (see also Leitch et al., 2001).

Our results reinforce the findings of Leitch et al. (1998) that extant basal angiosperms are characterized by very small 1C-values. Not only did the ancestor of extant angiosperms have a very small genome, but our data also indicate that a very small genome was ancestral throughout basal angiosperms, the
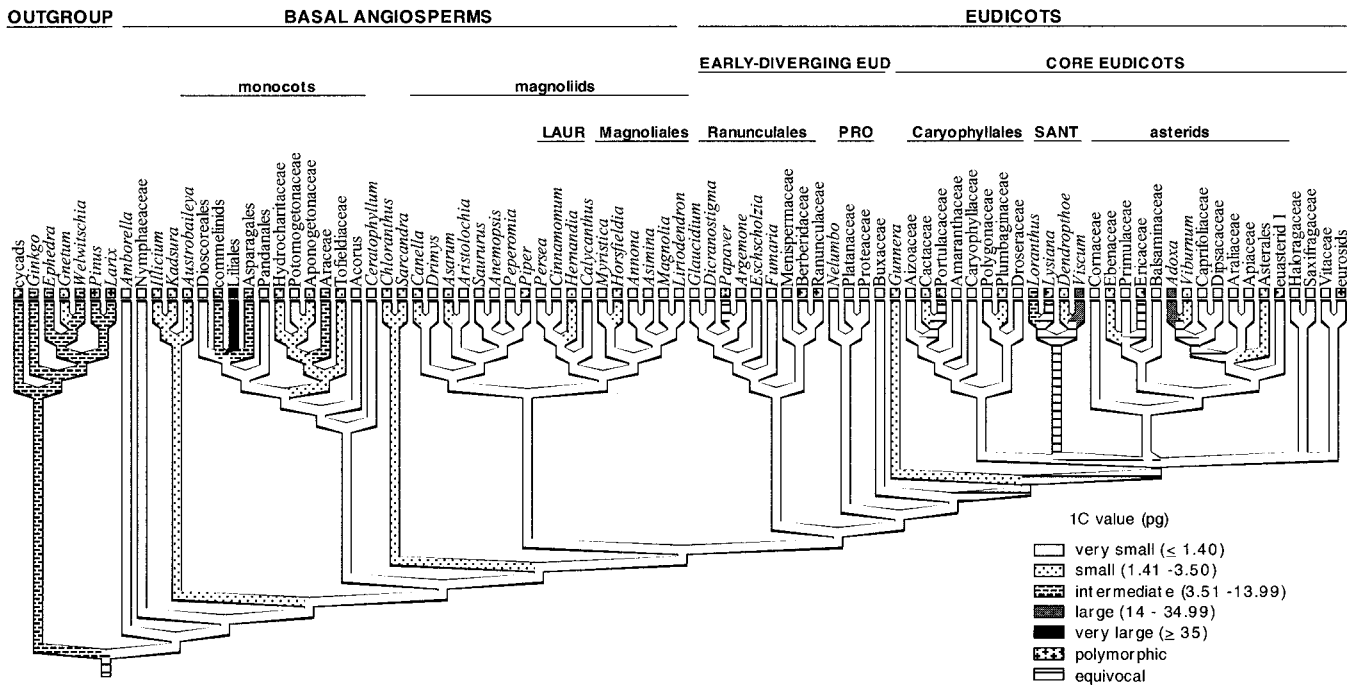
Fig. 1. Parsimony reconstruction of genome-size diversification in the angiosperms, using the ''all most parsimonious states'' resolving option of MacClade (Maddison and Maddison, 1992). General tree topology follows Soltis et al. (2000), with modifications following Chase et al. (2000), Fishbein et al. (2001), Zanis et al. (2002), and Soltis et al. (2003) (see text). 1C-values are provided in the Appendix (available as Supplementary Data accompanying the online version of this article) and are taken from Hanson et al. (2001a, b), Leitch and Hanson (2002), and Bennett and Leitch (2003). Range of values for Asparagales, commelinids, Dioscoreales, Liliales, and Pandanales are from Leitch et al. (1998); more detail for monocots is provided in Fig. 2. LAUR = Laurales; PRO = Proteales; SANT = Santalales.

monocots, the early-diverging eudicots, and some, if not all, of the major clades of core eudicots (e.g., Caryophyllales, Saxifragales, asterids). Among basal angiosperms, *Amborella* has a very small genome, as does the ancestor of Nymphaeaceae and the ancestors of the monocot and magnoliid clades (see later). Importantly, the Austrobaileyales clade is characterized by C-values in the small and intermediate ranges, rather than the very small range (Fig. 1; Appendix, available as Supplementary Data accompanying the online version of this article). Hence, this clade may represent an evolutionary lineage that long ago experienced an increase in genome size.

A very small genome size is reconstructed as ancestral through much of the monocot clade. Both *Ceratophyllum*, the apparent sister group of the monocots (Zanis et al., 2002) and *Acorus*, which is sister to all other monocots, have very small 1C-values, 0.69 pg (Leitch and Hanson, 2002) and 0.52 pg (mean value for the two *Acorus* species, *A. gramineus*, 1C = 0.4 pg; *A. calamus*, 1C = 0.65 pg; Appendix), respectively.

The ancestral state for many remaining monocots is also reconstructed as a very small 1C-value (Figs. 1, 2). Following *Acorus*, a clade of Alismataceae, Araceae, Tofieldiaceae, Hydrocharitaceae, and Zosteraceae is sister to all remaining monocots (Chase et al., 2000). Tofieldiaceae, which appear to occupy a pivotal position as sister to remaining Alismatales, have a small genome size, but the reported 1C-value is only 1.50 pg (Appendix). In contrast, genome-size estimates are high for Alismataceae (there are chromosome counts for nine species, but not all are diploids; the 1C-value for a known diploid species is 10.30 pg, which is in the intermediate range). Genome sizes reported for Araceae range from 0.33 to 15.83 pg. Some of the early-branching members of Araceae (as re-

vealed in a recent molecular phylogenetic analysis of the family; G. Salazar et al., unpublished data) have very small genome sizes (e.g., *Lemna*, 1C = 0.60 pg) or small genomes (e.g., *Xanthosoma*, 1C = 2.3 pg), whereas genome-size estimates for some early-diverging Araceae are large (e.g., *Orontium*, 1C = 15.00 pg). However, C-values for additional early-branching members of Araceae are still needed to determine whether the ancestral genome size for the family is large or small. Additional estimates are likewise needed for other early-branching monocots.

Small or very small 1C-values are also characteristic of other monocot families, such as Typhaceae, Pandanaceae, Dioscoreaceae, Bromeliaceae, and Sparganiaceae. Based on our reconstructions, very large genomes may have evolved at least three times independently in the monocots (Fig. 2): (1) once in Commelinaceae of the commelinid clade; (2) in some members of Liliaceae, Melanthiaceae, and Alstroemeriaceae of the Liliales; (3) and in some Hyacinthaceae, Alliaceae, and Amaryllidaceae of the Asparagales clade. However, the number of origins of very large genomes in the latter two orders is uncertain. Although some members of Asparagales and Liliales do have very large genomes, most have large or even smaller genomes. It remains unclear, however, if the common ancestor of these two large clades (Asparagales and Liliales) had a very large genome size or if a very large genome has originated multiple times in the two clades. Orchidaceae and Iridaceae (successive sisters to other Asparagales) contain an array of 1C-values, from very small to large. Clarifying ancestral genome size and subsequent diversification in Asparagales and Liliales will require additional genome-size estimates as well
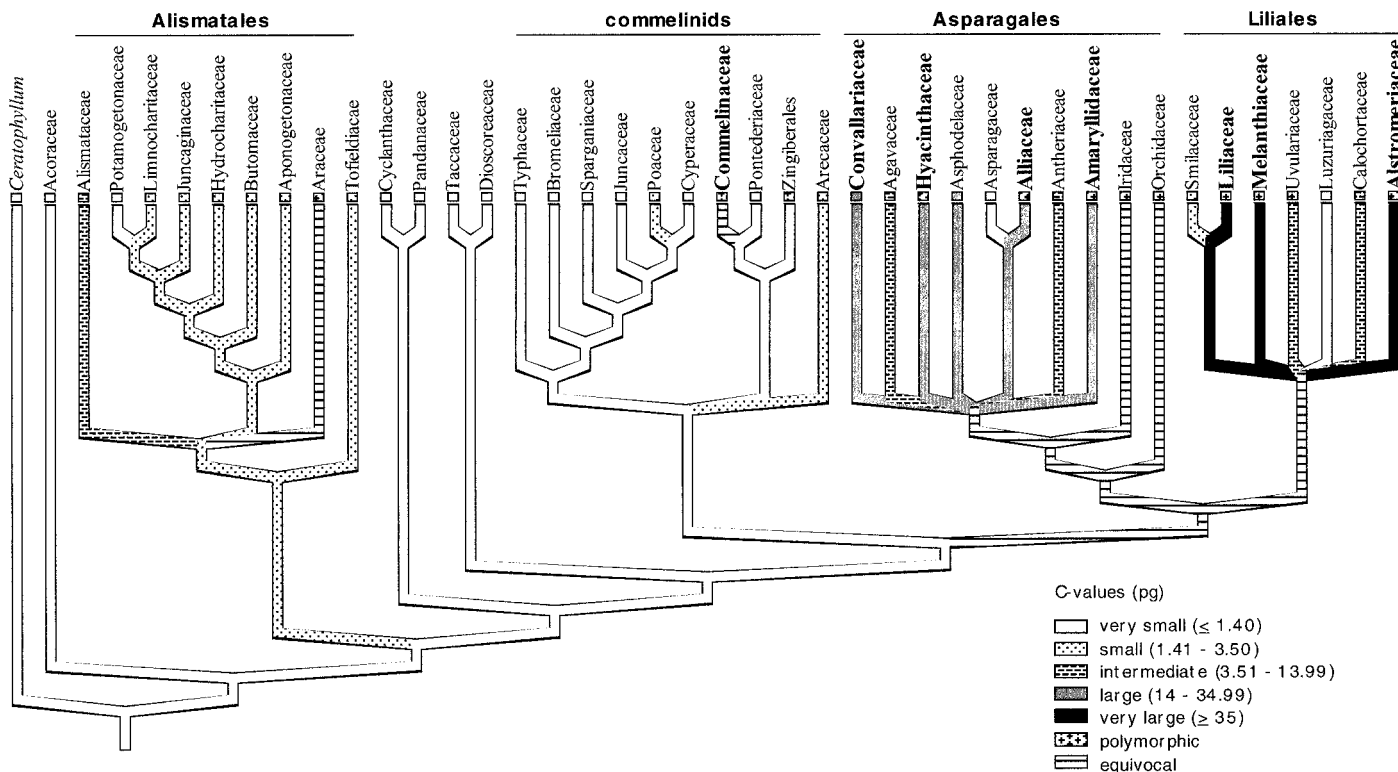
Fig. 2. Parsimony reconstruction of genome-size diversification in the monocots, using the "all most parsimonious states" resolving option of MacClade (Maddison and Maddison, 1992). Families indicated in bold have some representatives with very large genomes (see text). General tree topology follows Chase et al. (2000) (see text). 1C-values are provided for the genera shown in the Appendix (available as Supplementary Data accompanying the online version of this article); 1C-values for families are taken from Leitch et al. (1998).

as better resolution of relationships among the constituent members of these two clades.

Following the monocots, the best estimates of angiosperm phylogeny place Chloranthaceae as sister to all other angiosperms, although support for this placement is low (Zanis et al., 2002). Estimates of genome size for Chloranthaceae are in the small or intermediate range. Estimates for two species of *Chloranthus* are 1C = 2.90 and 3.59 pg, respectively; *Sarcandra* has a 1C-value of 4.35 pg. Hence, Chloranthaceae may represent another ancient lineage that, like Austrobaileyales, experienced an early increase in genome size.

New data for the large magnoliid clade (Laurales + Magnoliales and Canellales + Piperales) (sensu APG II, 2003) provide increased resolution of genome-size evolution within this clade. Genome-size estimates are now available for two members of Canellales (= Winterales of some recent investigations), *Canella* (Canellaceae; 1C = 5.83 pg), and *Drimys* (Winteraceae; 1C = 1.13 pg), for additional members of Magnoliales, such as *Myristica* and *Horsfieldia* (Myristicaceae; 1C = 1.3 and 1.7 pg, respectively), and for additional Laurales, including *Hernandia* (Hernandiaceae; 1C = 1.80 pg) and *Calycanthus* (Calycanthaceae; 1C = 0.98 pg) (Appendix). Many members of the magnoliid clade have small or very small genomes. For example, numerous estimates are available for genera of Annonaceae, and all have very small genomes (Appendix). Importantly, character-state reconstructions further indicate that the ancestral genome of the entire magnoliid clade was very small, as were the ancestral genomes of the Laurales + Magnoliales subclade and the Canellales + Piperales subclade (Fig. 1).

Consideration of ploidy indicates that the ancestral genome for the entire Piperales was probably very small. However, this pattern is obscured when mean values are used for the constituent families because the polyploid taxa have higher 1C-values than the diploids. Saururaceae have very small genomes, with 1C estimates for *Anemopsis*, *Houttuynia*, and *Saururus* of 0.8, 1.3, and 0.5 pg, respectively. Piperaceae also have small or very small genomes. The mean of 10 1C estimates for the genus *Piper* is 1C = 1.25 pg. The highest value for *Piper*, 1C = 2.4 pg, is for a high polyploid having 2n = 104. Similarly, for *Peperomia* the mean 1C-value for seven species with chromosome counts is 1C = 2.06 pg, but the three highest values (1C = 3.1, 2.0, and 3.95 pg) are for polyploids having 2n = 33, 44, and 66, respectively. When only diploids are considered (2n = 22), the mean 1C-value for the genus is only 1.30 pg and the range is 0.63–1.68 pg.

When only known diploids of Magnoliaceae (2n = 38) are considered, the genome size estimates are very small for this family. Our reconstruction of a very small genome size for extant Magnoliaceae and most other basal angiosperm lineages is intriguing in that these plants are presumed ancient polyploids whose ancestral diploids are now extinct (e.g., Stebbins, 1971; Grant, 1981), a hypothesis supported by isozyme data (Soltis and Soltis, 1990). Thus, the original genome sizes of the now extinct "paleodiploid" Magnoliaceae and other basal lineages such as Calycanthaceae, Myristicaceae, Illiciaceae, and Austrobaileyaceae were likely even smaller than those estimated for modern taxa.

Genome-size evolution in eudicots—The additional data included in the present analysis coupled with the inclusion of

only known diploids had a profound effect on the reconstruction of patterns of genome-size evolution in the eudicots. The ancestral genome for eudicots appears to have been very small (Fig. 1) based on reconstructions for key clades, such as Ranunculales and Proteales, as well as on data for Buxaceae. New data for *Platanus* (Platanaceae; 1C = 1.30 pg) and *Nelumbo* (Nelumbonaceae; 1C = 0.24 pg), combined with an earlier report for *Grevillea* (Proteaceae; 1C = 0.83 pg), suggest that the ancestral genome for Proteales was very small. Again, similar results were obtained using parsimony, squared-change parsimony, and maximum likelihood optimization methods.

Ranunculales occupy an important position as sister to all other eudicots, and our reconstructions indicate that this clade may also have had a very small ancestral genome. Our results contrast with those of Leitch et al. (1998), who employed a mean value for the entire clade (1C = 6.5 pg), which is close to the mean for all 2802 angiosperms they studied (mean 1C = 6.9 pg). However, Berberidaceae have a very small average genome size (1C = 1.26 pg for 10 known diploids) as do Menispermaceae with 1C = 0.70 pg. The large family Papaveraceae also appears to have a very small ancestral genome, but this is not evident from the mean value for the entire family. The average genome size for Papaveraceae is 1C = 2.99 pg, which is in the small range for angiosperms. However, this estimate is heavily influenced by numerous estimates for species of the large genus *Papaver*, which range from 1C = 1.75 to 8.93 pg. In addition, polyploidy contributes to the two highest values in *Papaver* (*P. orientale*, 1C = 8.93 pg, 2n = 42; *P. setigerum*, 1C = 6.68 pg, 2n = 44). The mean 1C-value for 21 known diploids in *Papaver* is 3.02 pg with the values ranging from 2.33 to 4.90 pg; the mean when all values are included regardless of ploidy is 3.41 pg. Within Papaveraceae, many early-branching members (based on the topology of Hoot et al., 1997) have very small genome sizes. *Fumaria* (1C = 0.55 pg even though it is tetraploid), *Glaucium* (1C = 0.60 pg), *Eschscholzia* (1C = 1.10 pg), and *Argemone* (1C = 0.60 pg) all have very small genomes. Using genome-size estimates for diploids in Papaveraceae and examining genome-size estimates for the family in light of recent phylogenetic trees indicates that the ancestral genome size for the family was very small (1C < 1.4 pg), which contrasts with the mean value for the family of 1C = 2.99 pg.

The ancestral genome size for Ranunculaceae, the other large family of Ranunculales, remains unclear. Although numerous values have been reported for Ranunculaceae, data are not available for crucial early-diverging members of the family, such as *Hydrastris*, *Glaucidium*, *Coptis*, and *Xanthorhiza* (e.g., Drinnan et al. 1994; Soltis et al., 2000). Values for these taxa would permit a better understanding of the ancestral genome size in this family and hence enhance character-state reconstruction in the Ranunculales as a whole.

The ancestral genome size for core eudicots is reconstructed as equivocal. The single 1C estimate for Gunnerales, the sister to all other core eudicots, is 1C = 7.44 pg (based on a single species of *Gunnera*). However, despite this relatively large 1C-value for *Gunnera*, a very small genome size may be ancestral throughout most of the core eudicots. In fact, if Saxifragales are treated as sister to the rosid clade, as found with weak support in Soltis et al. (2000), a very small genome size is reconstructed as ancestral for all core eudicots.

Because of the uncertainty of the position of Saxifragales, we have shown the core eudicots as a large polytomy (reviewed in Soltis et al., 2000, 2003). A very small genome size

is reconstructed as ancestral for Saxifragales, Caryophyllales, and asterids. Although only three C-values are available for Saxifragales, all are less than 1.4 pg. Caryophyllales also have a small ancestral genome size (Fig. 1). The mean 1C-value for the Caryophyllales clade is 1.7 pg (Leitch et al., 1998), but our MacClade reconstructions using diploid values indicate that the ancestral genome size is in the very small range. Following the topology of Albach et al. (2001), Soltis et al. (2003), and Bremer et al. (2002) for asterids, Cornales, followed by Ericales, represent the successive sisters to the euasterids. We also used the topology of Bremer et al. (2002) for relationships within Ericales (Fig. 1). Diploid Cornales and Ericales generally have 1C-values less than 1.4 pg and the ancestral condition for the asterid clade is therefore a very small genome. Within the asterids, a large genome is found only in *Adoxa* (Adoxaceae; 1C = 14.30), which occupies a derived position within the euasterid clade (Fig. 1) and is a polyploid with 2n = 36. Diploid species of *Viburnum* (also in Adoxaceae) have small genome sizes (Fig. 1). Adoxaceae are part of euasterid II, and other diploids of this clade have either small or very small genomes. Thus, character-state reconstructions for Caryophyllales and asterids, which suggest very small ancestral genomes, provide a slightly different view than the mean values of 1.7 and 3.2 pg (Leitch et al., 1998), respectively (both in the small range).

The ancestral genome size of the rosids is unclear. Vitaceae appear to represent the sister to all other rosids, and *Vitis* has a very small genome size (1C = 0.4–0.6 for 21 diploids). However, relationships within the core rosids remain poorly understood, precluding the reconstruction of an ancestral genome size for this large clade at this time. If the large eurosid clade is coded as polymorphic (diploids generally have very small or small genomes), our reconstructions suggest that the ancestral state of rosids was a very small genome (Fig. 1). However, additional phylogenetic resolution is required for the rosid clade before ancestral states can be reconstructed with any confidence.

Although Santalales are often considered to have a large genome size (see Leitch et al., 1998), very large genomes are actually confined to *Viscum* (Viscaceae) with a 1C-value (based on two similar estimates) of 76.0 and 79.3 pg. However, 1C estimates for other Viscaceae (sensu APG, 1998; APG II, 2003) are much lower. *Loranthus* has a 1C-value of 15.20 pg, and *Lysiana* has a mean 1C-value of 12.50 pg (based on values for six diploid species that range from 11.03 to 15.28 pg). The value for *Loranthus* is in our large range and the range of values for *Lysiana* represent our intermediate and large categories. *Dendrophthoe* (Santalaceae) has a mean 1C-value of only 4.3 pg based on C-values for five diploids (range 2.7–6.2 pg). Using the current best estimate of phylogeny for Santalales (Nickrent and Malécot, 2001), *Viscum* is clearly derived within Santalales. The ancestral state for the family therefore is reconstructed as equivocal. One caveat of our mapping investigation is that 1C-values are not available for early-branching Santalales. Furthermore, many genera of Loranthaceae for which 1C-values are available are not included in the Nickrent and Malécot (2001) tree; conversely, most of the genera included in the phylogenetic analyses do not have genome-size estimates. Additional 1C-values for early-diverging members of Santalales would be useful for inferring patterns of genome size evolution.

Genetic obesity hypothesis—Bennetzen and Kellogg (1997) proposed that genome-size evolution in plants would be largely unidirectional, with an overall pattern of increase as a result of the combined influence of polyploidy and the accumulation of retroelements. They suggested that plants have a ''one way ticket to genetic obesity.'' The hypothesis of unidirectional increase in genome size has only rarely been critically evaluated (Bennetzen and Kellogg, 1997; Cox et al., 1998; Wendel et al., 2002). Our results within angiosperms appear to be in general agreement with the genetic obesity hypothesis, with very large genomes confined to taxa that occupy derived positions within larger clades. However, the approach that we have taken here is coarse-grained; careful evaluation of the genetic obesity hypothesis is required within individual clades (e.g., within families and genera).

Our reconstructions indicate that a very small genome size has played a more prominent role in angiosperm evolution than previously appreciated, representing the ancestral condition for most major clades. However, a study of genome size evolution across embryophytes (land plants) provides convincing evidence for decreases in genome size, as well as increases. Genome size has decreased independently in Marsileaceae, Salviniaceae, *Gnetum*, and perhaps in the ancestor of angiosperms (I. J. Leitch et al., unpublished manuscript). On a finer scale, in *Gossypium* Wendel et al. (2002) examined genome size evolution in a phylogenetic context and found that the number of decreases in 1C-value exceeded the number of increases. Furthermore, several sources give convincing evidence that contractions in genome size have occurred throughout the angiosperms (Rabinowicz, 2000; Wendel et al., 2002). Other recent studies improve our understanding as to how such decreases can take place (e.g., Vicient et al., 1999; Kirik et al., 2000; Bennetzen, 2002; Frank et al., 2002; Hancock, 2002; Petrov, 2002; Zuckerkandl, 2002). Current data indicate that unequal recombination can slow the increase in genome size and that illegitimate recombination and other deletion processes may be the major mechanisms for decreases in genome size (Bennetzen, 2002; Petrov, 2002). The extensive Angiosperm DNA C-values Database reveals enormous variation in genome size within many genera and families, with some polyploids having genome sizes that are substantially smaller than those of diploid congeners. The evidence seems to indicate, therefore, that the evolution of genome size in the angiosperms, as well as in embryophytes in general, is dynamic with both increases and decreases.

LITERATURE CITED

ALBACH, D. C., P. S. SOLTIS, D. E. SOLTIS, AND R. G. OLMSTEAD. 2001. Phylogenetic analysis of the Asteridae s. l. based on sequences of four genes. *Annals of the Missouri Botanical Garden* 88: 163–212.

ANTONIUS, K., AND H. AHOKAS. 1996. Flow cytometric determination of polyploidy level in spontaneous clones of strawberries. *Hereditas* 124: 285.

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

APG II [ANGIOSPERM PHYLOGENY GROUP II]. 2003. An updated classification of the angiosperms. *Botanical Journal of the Linnean Society* 141: 399–436.

ARABIDOPSIS GENOME INITIATIVE. 2000. Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature* 408: 796–815.

BENNETT, M. D., AND I. J. LEITCH. 2003. Angiosperm DNA C-values Database (release 3.1, September 2001). Website: http://www.rbgkew.org.uk/cval/homepage.html.

BENNETT, M. D., I. J. LEITCH, AND L. HANSON. 1998. DNA amounts in two

samples of angiosperm weeds. *Annals of Botany* 82(Supplement A): 121–134.

BENNETT, M. D., AND J. B. SMITH. 1976. Nuclear DNA amounts in angiosperms. *Philosophical Transactions of the Royal Society of London B* 274: 227–274.

BENNETT, M. D., AND J. B. SMITH. 1991. Nuclear DNA amounts in angiosperms. *Philosophical Transactions of the Royal Society of London B* 334: 309–345.

BENNETZEN, J. L. 1996. The contributions of retroelements to plant genome organization, function, and evolution. *Trends in Microbiology* 4: 347–353.

BENNETZEN, J. L. 2000. Transposable element contributions to plant gene and genome evolution. *Plant Molecular Biology* 42: 251–269.

BENNETZEN, J. L. 2002. Mechanisms and rates of genome expansion and contraction in flowering plants. *Genetica* 115: 29–36.

BENNETZEN, J. L., AND E. A. KELLOGG. 1997. Do plants have a one-way ticket to genomic obesity? *Plant Cell* 9: 1509–1514.

BOWE, L. M., G. COAT, AND C. W. DEPAMPHILIS. 2000. Phylogeny of seed plants based on all three genomic compartments: extant gymnosperms are monophyletic and Gnetales' closest relatives are conifers. *Proceedings of the National Academy of Sciences, USA* 97: 4092–4097.

BOWERS, J. E., B. A. CHAPMAN, J. RONG, AND A. H. PATERSON. 2003. Unravelling angiosperm genome evolution by phylogenetic analysis of chromosomal duplication events. *Nature* 422: 433–438.

BREMER, B., K. BREMER, N. HEIDARI, P. ERIXON, R. G. OLMSTEAD, A. A. ANDERBERG, M. KÄLLERSJÖ, AND E. BARKHORDARIAN. 2002. Phylogenetics of asterids based on 3 coding and 3 non-coding chloroplast DNA markers and the utility of non-coding DNA at higher taxonomic levels. *Molecular Phylogenetics and Evolution* 24: 274–301.

CHASE, M. W., D. E. SOLTIS, P. S. SOLTIS, P. J. RUDALL, M. F. FAY, W. H. HAHN, S. SULLIVAN, J. JOSEPH, T. GIVNISH, K. J. SYTSMA, AND C. PIRES. 2000. Higher-level systematics of the monocotyledons: an assessment of current knowledge and a new classification. *In* K. Wilson and D. Morrison, [eds.], Proceedings of Monocots II: The Second International Symposium on the Comparative Biology of the Monocotyledons, Sydney, Australia, 3–16. CSIRO Press, Sydney, Australia.

CHASE, M. W., ET AL. 1993. Phylogenetics of seed plants: an analysis of nucleotide sequences from the plastid gene *rbcL*. *Annals of the Missouri Botanical Garden* 80: 528–580.

CHAW, S.-M., C. L. PARKINSON, Y. CHENG, T. M. VINCENT, AND J. D. PALMER. 2000. Seed plant phylogeny inferred from all three plant genomes: monophyly of extant gymnosperms and origin of Gnetales from conifers. *Proceedings of the National Academy of Sciences, USA* 97: 4086–4091.

COX, A. V., G. J. ABDELNOUR, M. D. BENNETT, AND I. J. LEITCH. 1998. Genome size and karyotype evolution in the slipper orchids (Cypripedioideae: Orchidaceae). *American Journal of Botany* 85: 681–687.

CUÉNOUD, P., V. SAVOLAINEN, L. W. CHATROU, M. POWELL, R. J. GRAYER, AND M. W. CHASE. 2002. Molecular phylogenetics of Caryophyllales based on nuclear 18S rDNA and plastid *rbcL*, *atpB*, and *matK* DNA sequences. *American Journal of Botany* 89: 132–144.

DRINNAN, A. N., P. R. CRANE, AND S. B. HOOT. 1994. Patterns of floral evolution in the early diversification of non-magnoliid dicotyledons (eudicots). *Plant Systematics and Evolution, Supplement* 8: 93–122.

EHRENDORFER, F., F. KRENDL, E. HABELER, AND W. SAUDER. 1968. Chromosome numbers and evolution in primitive angiosperms. *Taxon* 17: 337–468.

ELSIK, C. G., AND C. G. WILLIAMS. 2000. Retroelements contribute to the excess low-copy-number DNA in pine. *Molecular and General Genetics* 264: 47–55.

FISHBEIN, M., C. HIBSCH-JETTER, D. E. SOLTIS, AND L. HUFFORD. 2001. Phylogeny of Saxifragales (angiosperms, eudicots): analysis of a rapid, ancient radiation. *Systematic Biology* 50: 817–847.

FRANK, A. C., H. AMIRI, AND S. G. E. ANDERSSON. 2002. Genome deterioration: loss of repeated sequences and accumulation of junk DNA. *Genetica* 115: 1–12.

GRANT, D., P. CREGAN, AND R. C. SHOEMAKER. 2000. Genome organization in dicots: genome duplication in *Arabidopsis* and synteny between soybean and *Arabidopsis*. *Proceedings of the National Academy of Sciences, USA* 97: 4168–4173.

GRANT, V. 1981. Plant speciation. Columbia University Press, New York, New York, USA.

HANCOCK, J. M. 2002. Genome size and the accumulation of simple sequence

repeats: implications of new data from genome sequencing projects. *Genetica* 115: 93–103.

HANSON, L. K., A. MCMAHON, M. A. T. JOHNSON, AND M. D. BENNETT. 2001a. First nuclear DNA C-values for 25 angiosperm families. *Annals of Botany* 87: 251–258.

HANSON, L., K. A. MCMAHON, M. A. T. JOHNSON, AND M. D. BENNETT. 2001b. First nuclear DNA C-values for another 25 angiosperm families. *Annals of Botany* 88: 851–858.

HOOT, S. B., J. W. KADEREIT, F. R. BLATTNER, K. B. JORK, A. E. SCHWARZ-BACH, AND P. R. CRANE. 1997. Data congruency and phylogeny of the Papaveraceae s. l. based on four data sets: *atpB* and *rbcL* sequences, *trnK* restriction sites, and morphological characters. *Systematic Botany* 22: 575–590.

HOOT, S. B., S. MAGALLÓN, AND P. R. CRANE. 1999. Phylogeny of basal eudicots based on three molecular data sets: *atpB*, *rbcL*, and 18S nuclear ribosomal DNA sequences. *Annals of the Missouri Botanical Garden* 86: 1–32.

KIDWELL, M. G. 2002. Transposable elements and the evolution of genome size in eukaryotes. *Genetica* 115: 49–63.

KIRIK, A., S. SALOMON, AND H. PUCHTA. 2000. Species-specific double-strand break repair and genome evolution in plants. *EMBO Journal* 19: 5562–5566.

LEITCH, I. J., AND M. D. BENNETT. 1997. Polyploidy in angiosperms. *Trends in Plant Science* 2: 470–476.

LEITCH, I. J., M. W. CHASE, AND M. D. BENNETT. 1998. Phylogenetic analysis of DNA C-values provides evidence for a small ancestral genome size in flowering plants. *Annals of Botany* 82: 85–94.

LEITCH, I. J., AND L. HANSON. 2002. DNA C-values in seven families fill phylogenetic gaps in the basal angiosperms. *Botanical Journal of the Linnean Society* 140: 175–179.

LEITCH, I. J., L. HANSON, M. WINFIELD, J. PARKER, AND M. D. BENNETT. 2001. Nuclear DNA C-values complete familial representation in gymnosperms. *Annals of Botany* 88: 843–849.

MADDISON, W. P., AND D. R. MADDISON. 1992. MacClade: analysis of phylogeny and character evolution. Sinauer, Sunderland, Massachusetts, USA.

MATHEWS, S., AND M. J. DONOGHUE. 1999. The root of angiosperm phylogeny inferred from duplicate phytochrome genes. *Science* 286: 947–950.

NICKRENT, D. L., AND V. MALÉCOT. 2001. A molecular phylogeny of Santalales. *In* A. Fer, P. Thalouarn, D. M. Joel, L. J. Musselman, C. Parker, and J. A. C. Verkleij [eds.], Proceedings of the Seventh International Parasitic Weed Symposium, 69–74. Faculté des Sciences, Université de Nantes, Nantes, France.

OTTO, S. P., AND J. WHITTON. 2000. Polyploidy incidence and evolution. *Annual Review of Genetics* 34: 401–437.

PETROV, D. A. 2002. DNA loss and evolution of genome size in Drosophila. *Genetica* 115: 81–91.

PLUNKETT, G. M., AND S. R. DOWNIE. 1999. Major lineages within Apiaceae subfamily Apioideae: a comparison of chloroplast restriction site and DNA sequence data. *American Journal of Botany* 86: 1014–1026.

QIU, Y.-L., J. LEE, F. BERNASCONI-QUADRONI, D. E. SOLTIS, P. S. SOLTIS, M. ZANIS, E. A. ZIMMER, Z. CHEN, V. SAVOLAINEN, AND M. W. CHASE. 1999. The earliest angiosperms: evidence from mitochondrial, plastid and nuclear genomes. *Nature* 402: 404–407.

RABINOWICZ, P. D. 2000. Are obese plant genomes on a diet? *Genome Research* 10: 893–894.

RAVEN, P. H. 1975. The basis of angiosperm phylogeny: cytology. *Annals of the Missouri Botanical Garden* 62: 724–764.

SANMIGUEL, P. A., A. TIKHONOV, Y. K. JIN, N. MOTCHOULSKAIA, D. ZAKHAROV, A. MELAKE-BERHAN, P. S. SPRINGER, K. J. EDWARDS, M. LEE, Z. AVRAMOVA, AND J. L. BENNETZEN. 1996. Nested retrotransposons in the intergenic regions of the maize genome. *Science* 274: 765–768.

SMISSEN, R. D., J. C. CLEMENT, P. J. GARNOCK-JONES, AND G. K. CHAMBERS. 2002. Subfamilial relationships within Caryophyllaceae as inferred from 5′ *ndhF* sequences. *American Journal of Botany* 89: 1336–1341.

SOLTIS, D. E., A. E. SENTERS, M. J. ZANIS, S. KIM, J. D. THOMPSON, P. S. SOLTIS, L. RONSE DE CRAENE, P. K. ENDRESS, AND J. S. FARRIS. 2003. Gunnerales are sister to other core eudicots, and exhibit floral features of early-diverging eudicots. *American Journal of Botany* 90: 461–470.

SOLTIS, D. E., AND P. S. SOLTIS. 1990. Isozyme evidence for ancient polyploidy in primitive angiosperms. *Systematic Botany* 15: 328–337.

SOLTIS, D. E., P. S. SOLTIS, AND M. J. ZANIS. 2002. Phylogeny of seed plants based on evidence from eight genes. *American Journal of Botany* 89: 1670–1681.

SOLTIS, D. E., ET AL. 2000. Angiosperm phylogeny inferred from a combined data set of 18S rDNA, *rbcL* and *atpB* sequences. *Botanical Journal of the Linnean Society* 133: 381–461.

SOLTIS, P. S., AND D. E. SOLTIS. 2000. The role of genetic and genomic attributes in the success of polyploids. *Proceedings of the National Academy of Sciences, USA* 97: 7051–7057.

STEBBINS, G. L. 1971. Chromosomal evolution in higher plants. Addison-Wesley, London, UK.

THOMAS, C. A. 1971. The genetic organization of chromosomes. *Annual Review of Genetics* 5: 237–256.

VICIENT, C. M., A. SUONIEMI, K. ANAMTHAMAT-JONSSON, J. TANSKANEN, A. BEHARAV, E. NEVO, AND A. H. SCHULMAN. 1999. Retrotransposon BARE-1 and its role in genome evolution in the genus Hordeum. *Plant Cell* 11: 1769–1784.

WALBOT, V. 2000. A green chapter in the book of life. *Nature* 408: 794–795.

WENDEL, J. F. 2000. Genome evolution in polyploids. *Plant Molecular Biology* 42: 225–249.

WENDEL, J. F., R. C. CRONN, J. S. JOHNSTON, AND H. J. PRICE. 2002. Feast and famine in plant genomes. *Genetica* 115: 37–47.

ZANIS, M. J., D. E. SOLTIS, P. S. SOLTIS, S. MATHEWS, AND M. J. DONOGHUE. 2002. The root of the angiosperms revisited. *Proceedings of the National Academy of Sciences, USA* 99: 6848–6853.

ZUCKERKANDL, E. 2002. Why so many noncoding nucleotides? The eukaryote genome as an epigenetic machine. *Genetica* 115: 105–129.