Proc. Natl. Acad. Sci. USA
Vol. 87, pp. 714–717, January 1990
Botany

Multiple origins of polyploids in the *Glycine tabacina* complex inferred from chloroplast DNA polymorphism

(restriction fragment variation/plant taxonomy)

JEFF J. DOYLE*, JANE L. DOYLE*, A. H. D. BROWN†, AND J. P. GRACE†

*L. H. Bailey Hortorium, 466 Mann Library Building, Cornell University, Ithaca, NY 14853; and †Commonwealth Scientific and Industrial Research Organisation, Division of Plant Industry, GPO Box 1600, Canberra, Australian Capital Territory 2601, Australia

Communicated by Charles B. Heiser, Jr., November 13, 1989 (received for review September 8, 1989)

ABSTRACT **The *Glycine tabacina* polyploid complex has been shown to include a minimum of two morphological and crossing groups, which also differ in chloroplast DNA (cpDNA) restriction map and nuclear ribosomal gene repeat phenotype. These AAB_2B_2 and BBB_2B_2 *G. tabacina* polyploids contain plastomes referable to the A and B diploid plastome groups of subgenus *Glycine*, respectively. Eight different cpDNA variants were observed among the 65 B-type polyploids studied, six of which were identical for numerous restriction site characters to plastome types found among the highly polymorphic B genome diploid species. It is hypothesized that there have been numerous independent origins of polyploid *G. tabacina*: at least one AA × B_2B_2 event and a minimum of five BB × B_2B_2 events involving different BB types as female progenitor. Low amounts of cpDNA divergence between diploid and polyploid plastomes and among the plastomes of geographically disjunct polyploids suggest that the origin and dispersal of polyploids are relatively recent events. All hypothesized diploid progenitors are native to Australia, while both A- and B-type *G. tabacina* polyploids occur on islands of the Pacific outside the range of diploids. The presence of several different plastome types of polyploid *G. tabacina* in the Pacific islands suggests that several colonization events have occurred.**

Polyploidy has been a significant genetic process in the evolution of flowering plants. The origin of a polyploid taxon is often considered a rare event, however, despite the abundance of angiosperm species with high, presumably polyploid, chromosome numbers (1, 2), the common natural occurrence of unreduced gametes in plants (3), and the often high frequency of polyploids within plant species (4). The availability of molecular markers has recently provided a means of assessing more readily the frequency of polyploid origins, with the result that multiple origins of polyploid taxa are increasingly being noted (5–8).

Glycine subgenus *Glycine*, the perennial relatives of the cultivated soybean, comprises 15 currently recognized species. The entire genus, although presumably an ancient polyploid with $2n = 4x = 40$, is fully diploidized and only three species of subgenus *Glycine* include neopolyploids (9, 10). One of these, *Glycine tabacina*, includes diploids ($2n = 40$) that are confined to eastern Australia and polyploids ($2n = 80$) that occur both sympatrically with the diploids as well as on islands of the South Pacific and western central Pacific (11–13). Two reproductively isolated types of polyploids are known in the complex and have been designated AAB_2B_2 and BBB_2B_2 to indicate that they share one of their two diploid genomes (11–13). Consistent with this hypothesis, both polyploids are fixed hybrids for their 18S–25S nuclear rRNA gene (rDNA) repeats and share one of their repeat classes (ref. 14; J.J.D., J.L.D., and A.H.D.B., unpublished data). Studies of restriction endonuclease recognition site variation of maternally inherited (15) chloroplast DNA (cpDNA) indicate that the two polyploid classes had different chloroplast genome (plastome) donors; that of the AAB_2B_2 type belonged to the A plastome group of diploid species while the BBB_2B_2 progenitor was, as expected, a member of the B plastome group (16).

Most recently, studies of cpDNA variation within the B diploid species group of the subgenus have revealed considerable polymorphism, with 25 plastome groups identified (J.J.D. *et al.*, unpublished data). These studies, along with previous cpDNA studies of the entire subgenus, have provided markers for elucidating the origins and affinities of polyploid *G. tabacina* accessions. Here we show that the B polyploid group is polymorphic for cpDNA variants that in several cases are identical to those found among diploid accessions, suggesting that this polyploid taxon had numerous independent origins.

METHODS AND MATERIALS

Sixty-five accessions of polyploid *G. tabacina* were used in this study, most selected from the Commonwealth Scientific and Industrial Research Organisation native Australian *Glycine* collection (listed in Table 1; map coordinates given in ref. 16). These accessions were chosen to represent as fully as possible the morphological, geographical, and genetic diversity of the complex. Plants were grown in the greenhouse under uniform conditions. Voucher specimens are deposited in the herbarium of the L. H. Bailey Hortorium (BH) and in the Herbarium Australiense (CANB).

DNA was isolated from 0.01–1 g of fresh leaves from individual plants by the method of Doyle and Doyle (17). Accessions in Table 1 were screened with a subset of the more than 50 restriction site characters used in our studies of variation among taxa of the diploid B genome group of subgenus *Glycine* (J.J.D. *et al.*, unpublished data); accessions marked with asterisks in Table 1 were screened with a minimum of 45 of these characters. Digestion with restriction enzymes, gel electrophoresis, transfer to nylon membranes, hybridization with mung bean cpDNA clones (from J. Palmer, Indiana University), and autoradiography were as described elsewhere (15). Cladistic analyses were performed on diploid and polyploid accessions using algorithms available in the HENNIG86 phylogenetic analysis software package (Version 1.5; supplied by J. S. Farris, State University of New York, Stony Brook, NY) as described elsewhere (J.J.D. *et al.*, unpublished data). Character analyses and graphical depiction of cladograms used the CLADOS software package (K. C. Nixon, personal communication).

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: cpDNA, chloroplast DNA; rDNA, nuclear rRNA genes.

Table. 1. Geographic origins and plastome designations of *G. tabacina* polyploid accessions studied

Group	Origin	Accessions
1	New South Wales, Australia	1208, 1234,* 1254, 1313, 1321, 2030, 2511, 2762
	Queensland, Australia	378707, 1262, 1862, 2246, 2274, 2279, 2772
	Victoria, Australia	2709
	South Australia	1706*
	Western central Pacific	1071, 1072, 1221, 1226, 1284, 1306, 1332, 1338,* 1382, 1714, 1716
	South Pacific	1258,* 1863, 1981, 1982, 1984, 1988
2	Australian Capital Territory	1075*
	New South Wales, Australia	1080,* 1439, 2028, 2029, 2763
	Queensland, Australia	2263, 2268, 2715, 2738
	Victoria, Australia	2708, 2710
3	New South Wales, Australia	1144, 1255,* 1273,* 1503, 1860, 2765
	Queensland, Australia	1861, 2281, 2308
	Western central Pacific	2602
4	New South Wales, Australia	1141,* 1312, 1314
5	South Pacific	1205,* 1324, 1325
6	New South Wales, Australia	1329*
U	South Pacific	1986, 1989

Accessions are listed by Commonwealth Scientific and Industrial Research Organisation Native Australian *Glycine* Collection numbers (four digits) except 378707, for which the U.S. Department of Agriculture Plant Introduction number is given. Plastome groups (groups 1–6) correspond to origin numbers given in Fig. 2; group U includes the two types unlike any diploid plastome types.
*Accessions were screened for variation in a minimum of 45 restriction site characters, while other accessions were screened with a subset of diagnostic markers.

RESULTS

All BBB_2B_2 (B type) *G. tabacina* accessions were screened for several characters that had been found to be diagnostic for particular plastome types in our studies of cpDNA variation in B genome diploid taxa (J.J.D. *et al.*, unpublished data). Through this procedure, nine different plastome types were found among the 65 polyploid accessions studied. Representatives of several of these types were analyzed with additional characters used in the diploid studies; results indicated that six of the nine types were each identical to different diploid plastome types. A seventh type (G1075) was more closely related to one of these diploid-polyploid pairs than any other polyploid accession. Two accessions, each with a different plastome variant, were unlike plastomes found in any of the diploids studied (G1986, G1989). Autoradiograms showing diagnostic restriction site variants for some of these diploid-polyploid pairs are illustrated in Fig. 1; cladistic relationships of these accessions and related diploid plastome types are shown in Fig. 2. This cladogram also includes a member of diploid plastome group 2, found in B_2B_2 diploid accessions related to *G. tabacina*, illustrating its divergent B′ plastome for comparison with other B diploid and polyploid accessions. Other plastome groups found among diploid accessions belonged to major clades shown in Fig. 2 but possessed characters not found in any polyploid accession and are not shown. For example, several *Glycine microphylla* accessions possessed plastomes sharing apomorphic characters with plastome 24 (and thus with G1706) but had additional unique restriction site characters as well.

The geographic origins of polyploid accessions representing these various plastomes are given in Table 1. There were no striking correlations between plastome class and geography. The three most common plastome types (polyploid plastome groups 1, 2, and 3, following the numbering of Fig. 2 and Table 1) were scattered throughout the species' range in eastern Australia; all three types also were found on islands of the Pacific and are in some cases sympatric there. Group 4 plastomes were observed only in Australian accessions, while group 5 plastomes were found only in accessions from New Caledonia. As these types differed by only a single restriction site character, it is possible that they represent a single origin followed by a restriction site gain in accessions from Australia. However, this same site polymorphism also occurs among populations of *Glycine latifolia*, a diploid species confined to Australia; the site character showed no homoplasy in a survey of 74 diploid B genome accessions (J.J.D. *et al.*, unpublished data), and the presence of this site in polyploid accessions is interpreted as evidence of cpDNA contribution from diploids possessing it, rather than of independent origin.

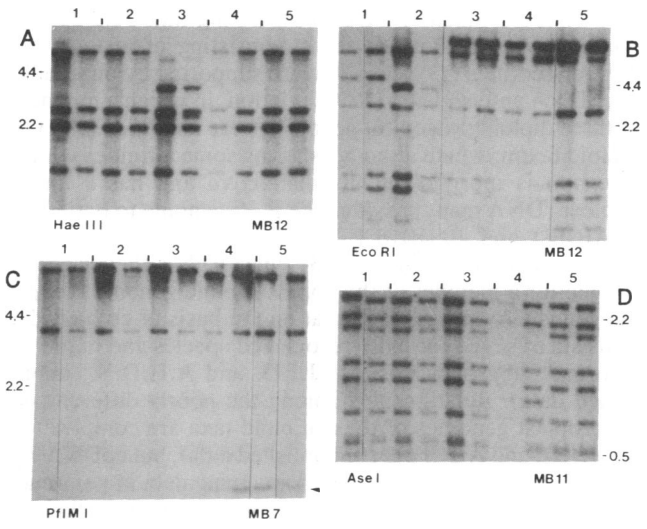

FIG. 1. Selected restriction fragment polymorphism phenotypes showing five diploid-polyploid pairs of accessions defined by unique combinations of derived restriction site gain or loss characters. Numbers above lanes indicate these pairs, in which the first accession is a diploid and the second is a polyploid, and correspond to the plastome types given in Table 1. Accessions used were the same on all four gels. Paired lanes: 1, *G. microphylla* G1143, *G. tabacina* G1338; 2, *G. microphylla* G1315, *G. tabacina* G1080; 3, *G. tabacina* G1138, *G. tabacina* G1255; 4, *G. latifolia* G1137, *G. tabacina* G1141; 5, *G. latifolia* G1160, *G. tabacina* G1205. (*A*) *Hae* III digestion probed with mung bean cpDNA clone 12 (MB12). (*B*) *Eco*RI probed with MB12 (two characters illustrated). (*C*) *Pfl*MI probed with MB7 (small fragment indicated by arrowhead). (*D*) *Ase* I probed with MB11. Size standards (in kilobases) are given for each digest.

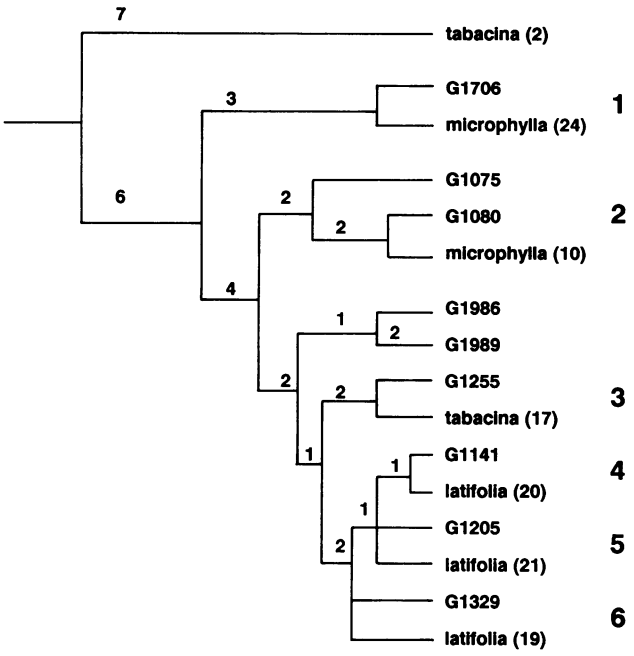

FIG. 2. Cladogram of polyploid and diploid plastome classes. Topology shown is consistent with most parsimonious trees from analysis of numerous diploid accessions with a larger character set (J.J.D. *et al.*, unpublished data). Polyploid plastomes are denoted by Commonwealth Scientific and Industrial Research Organisation accession number; diploid types are indicated by the epithet of the species most commonly having that plastome, followed by a number corresponding to the plastome type of J.J.D. *et al.* (unpublished data). The number of characters supporting each clade is given above the branch. Diploid-polyploid plastome pairs are indicated in boldface numbers to the right of the tree.

DISCUSSION

Artificial hybridization studies indicating that polyploid *G. tabacina* was a complex taxon consisting of at least two different entities (13) recently were supported by our studies of cpDNA and nuclear rDNA variation, in which a minimum of three diploid progenitor genomes were identified (16). One diploid taxon, which although sharing some affinities with *G. tabacina* is morphologically distinctive and has a unique nuclear rDNA map, provided the B_2B_2 genome to both the A (AAB_2B_2) and B (BBB_2B_2) types (ref. 16; J.J.D. *et al.*, unpublished data). All AAB_2B_2 types had as their maternal progenitor a member of the core A diploid genome group of subgenus *Glycine*, a group that on the basis of cpDNA data consists of seven formally recognized species and numerous informal subgroups (ref. 15; J.J.D. and A.H.D.B., unpublished data). Relationships among the poorly differentiated chloroplast genomes of these diploid taxa are complex (ref. 15; J.J.D. and A.H.D.B., unpublished data), but cpDNA data do not contradict nuclear rDNA mapping data in pointing to the D4 diploid isozyme race of *Glycine tomentella* as the most likely A genome parent (J.J.D. *et al.*, unpublished data).

Initial studies classifying 91 accessions of *G. tabacina* as either A or B types (16) used cpDNA characters that did not distinguish between plastomes possessed by BB or B_2B_2 diploid taxa. In more recent work, accessions of the diploid B genome species were found to comprise 25 different cpDNA variants, with a major dichotomy between the B′ plastomes of B_2B_2 plants and the numerous core B plastomes of BB diploids (J.J.D. *et al.*, unpublished data; Fig. 2). Screening for numerous characters marking this dichotomy failed to reveal any polyploids with plastomes like those of B_2B_2 diploids. Because all B-type polyploids are known to possess B_2B_2 rDNA repeats (J.J.D. *et al.*, unpublished data), it appears that the B_2B_2 diploid was the paternal genome donor of this group.

Studies of nuclear rDNA variation among BB diploids have not provided characters to discriminate among the different taxa. In contrast, the evidence of 25 cpDNA plastome types among the three recognized species (22 excluding B_2B_2 accessions) has made it possible to formulate precise hypotheses concerning the origins of polyploids derived from these diploid taxa. The existence of pairs of diploid and polyploid accessions sharing cpDNA types with unique, derived restriction sites suggests that there have been numerous independent origins of the B type of *G. tabacina*.

As has been described elsewhere, there is only partial concordance between plastome groups and taxonomic groupings based on morphology in the B diploid complex (J.J.D. *et al.*, unpublished data). Nevertheless, the end points of variation in the complex—*G. microphylla* with its narrow leaflets and small seeds, and *G. latifolia*, with rhombic leaflets, large seeds, and specialized ecology—have distinct cpDNA types that are found among polyploid *G. tabacina* accessions. Such polyploids might therefore be expected to differ from one another morphologically, although the differences would be moderated by the presence of a B_2B_2 genome in each case. Despite considerable morphologic variation among polyploids, no clear differentiation between these groups of accessions is evident. One possible explanation for this observation is that B-type polyploid accessions bearing different cpDNAs are at least partially interfertile in artificial crosses (ref. 11–13 and 16; J.J.D. *et al.*, unpublished data), suggesting that morphological homogenization through gene flow is likely. It is also possible, though unlikely, that the B-type polyploids originated prior to the morphological divergence of lineages leading to the modern diploid taxa.

A curious feature of our observations is that all *G. tabacina* allopolyploids arose by hybridizations in which B_2B_2 diploids or their derivatives served as pollen parents. There is no evidence of allopolyploidy involving hybridization between AA and BB diploids, nor have allopolyploids derived exclusively from BB genomes been observed. The unique phylogenetic position of the B_2B_2 genome inferred from cpDNA phylogenies (Fig. 2; J.J.D. *et al.*, unpublished data), basal to the core B genome, is possibly noteworthy in this context. B_2B_2 diploids are perhaps expected to be less diverged from A genome diploids than are BB taxa and thus might be better able to form viable hybrids with them that would then provide the raw material for allopolyploid formation. Artificial crosses between A and B genome diploids yield viable, but highly sterile, progeny (11, 12, 19–21); however, no natural hybridization between the two groups has been observed. Unfortunately, no laboratory crosses have been reported in which B_2B_2 diploids have been used as a parent with either A or B genome diploid plants.

"Multiple Origins" of Polyploids. The idea that polyploid formation is a rare event is implicit in much of the literature on the subject, although recent documentation of multiple allopolyploid origins in species of the fern genus *Asplenium* (5), in the weedy composite *Tragopogon* (8), and multiple autopolyploid origins in *Heuchera* (Saxifragaceae; ref. 7) indicates that this may be incorrect. The term multiple origins as it relates to polyploidy covers a host of possibilities, several of which are represented in the *G. tabacina* complex. On the one hand, it may refer to two genetically quite divergent entities constrained within the same specific epithet. The AAB_2B_2 and BBB_2B_2 polyploids described here are both classified as *G. tabacina*, presumably because their shared progenitor genome confers on them some morphological similarities. The two groups are, however, morphologically distinguishable, and although they are sympatric no natural hybrids between them have been reported. Indeed, artificial F_1 hybrids are sterile and show significant meiotic

irregularity (refs. 11–13; J.J.D. *et al.*, unpublished data). Clearly, these two groups could qualify as separate biological species.

At the other extreme, polyploidy within conspecific populations is probably a common phenomenon (2). The existence of numerous polyploid individuals within such a population, having arisen in independent events, could also be said to represent a cytotype having multiple origins. While in the majority of cases this might be considered trivial, adaptive differentiation of polyploids with multiple, presumably autopolyploid, origins has been observed in at least one genus of Saxifragaceae (7).

The B group of polyploid *G. tabacina* represents an intermediate case of multiple origins. All accessions of this polyploid are allopolyploid, as shown by fixed hybridity for nuclear rDNA (J.J.D. *et al.*, unpublished data). The two genomes involved are apparently relatively divergent; however, the source of one of the two major genomes has been from several members of a group of closely related and interfertile diploid genomes. What is recognized as a single taxon by such diverse criteria as ecology, morphology, interfertility, and shared molecular markers of the nuclear genome apparently was formed several times and in different combinations.

The Age of Formation and Colonization of *G. tabacina* Polyploids. Most of the polyploid accessions surveyed possessed plastomes that are not significantly differentiated from those encountered among diploids. If there has been no drastic slowing of cpDNA evolutionary rates in *G. tabacina*, this suggests that little time has elapsed since these polyploids originated. Given the possibility that cpDNA variation in this complex could predate the origin of the modern diploid species, the data say little about the age of polyploids relative to the B genome diploids. However, assuming even a rough molecular clock, the much greater cpDNA differentiation among the three major plastome groups of the subgenus (15) suggests that these polyploids are likely to be young relative to the age of the subgenus as a whole. The lack of cpDNA divergence among plastomes of B genome diploids and *G. tabacina* polyploids also would be consistent with continued gene flow between these two groups and with polyploid formation as an ongoing phenomenon in the complex (3).

Both the AAB_2B_2 and BBB_2B_2 types of polyploid *G. tabacina* are successful taxa, as indicated by their extensive ranges. Most strikingly, unlike any diploids in the subgenus, these polyploids have expanded their ranges into the Pacific. The A-type polyploid has reached Taiwan and the Marianas at least once, while the B type is even more widespread, occurring in the South Pacific as well. Furthermore, the existence of representatives of at least three different B-type plastome classes on different Pacific islands suggests multiple colonization events, or perhaps single events involving more than one individual. No markers have yet been found for either cpDNA or nuclear rDNA (J.J.D. *et al.*, unpublished data) that distinguish island from mainland accessions within a plastome group. These results, together with a lack of morphological or genetic differentiation, suggest that not only is polyploid *G. tabacina* a newly originated taxon, but that its colonization of Pacific islands also has been recent.

CONCLUSIONS

The A and B polyploids of *G. tabacina* are classic allopolyploids and presumably owe much of their evolutionary success to the benefits of fixed heterozygosity (22). However, allopolyploids having only a single origin are expected to include only a subset of the genetic diversity found among their diploid progenitors (23). Because *G. tabacina* plants having different progenitors are interfertile (11–13), the species has the potential to have incorporated far more of the genetic variation present in its diploid genome donors. Thus, as suggested for *Asplenium* species (5), multiple origins followed by gene exchange at the polyploid level may overcome the initially depauperate genetic condition. Combined with fixed hybridity, this may represent an important genetic strategy leading to evolutionary success of polyploids.

The ability to demonstrate multiple origins of polyploids in mosses (6), *Asplenium* (5), *Tragopogon* (8), various Saxifragaceae (7), the *G. tomentella* complex (ref. 24; J.J.D. *et al.*, unpublished data), and now in *G. tabacina* was due to the availability of polymorphic molecular markers. cpDNA has been shown to be more polymorphic than is commonly appreciated (18) and, as this study and others (6, 7) have shown, provides useful markers for investigating the origins of allopolyploid species.

The authors wish to thank Ted Hymowitz for making available to us preprints of his work. J.J.D. was supported by Grants NSF BSR-8516630, BSR-8805630, and USDA/Hatch NYC-187405. The collection of several accessions was made possible with funding from the International Board of Plant Genetic Resources to A.H.D.B.

1. Goldblatt, P. (1980) in *Polyploidy*, ed. Lewis, W. H. (Plenum, New York), pp. 219–239.
2. Lewis, W. H. (1980) in *Polyploidy*, ed. Lewis, W. H. (Plenum, New York), pp. 103–147.
3. Lewis, W. H. (1980) in *Polyploidy*, ed. Lewis, W. H. (Plenum, New York), pp. 241–268.
4. deWet, J. M. J. (1980) in *Polyploidy*, ed. Lewis, W. H. (Plenum, New York), pp. 3–15.
5. Werth, C. R., Guttman, S. I. & Eshbaugh, W. H. (1985) *Science* **228,** 731–733.
6. Wyatt, R., Odrzykoski, I. K. & Stoneburner, A. (1988) *Proc. Natl. Acad. Sci. USA* **84,** 9054–9058.
7. Soltis, D. E. & Soltis, P. S. (1989) *Am. J. Bot.* **76,** 1114–1118.
8. Soltis, D. E., Soltis, P. S. & Ness, B. D. (1989) *Evolution* **43,** 650–656.
9. Newell, C. A. & Hymowitz, T. (1978) *Am. J. Bot.* **65,** 168–179.
10. Tindale, M. D. & Craven, L. A. (1988) *Aust. Syst. Bot.* **1,** 399–410.
11. Singh, R. J. & Hymowitz, T. (1985) *Z. Pflanzenzuecht.* **95,** 289–310.
12. Singh, R. J. & Hymowitz, T. (1985) *Theor. Appl. Genet.* **71,** 221–230.
13. Singh, R. J., Kollipara, K. P. & Hymowitz, T. (1987) *Genome* **29,** 490–497.
14. Doyle, J. J. & Beachy, R. N. (1985) *Theor. Appl. Genet.* **70,** 369–376.
15. Doyle, J. J., Doyle, J. L. & Brown, A. D. H. (1990) *Evolution*, in press.
16. Doyle, J. J., Doyle, J. L., Grace, J. P. & Brown, A. D. H. (1990) *Syst. Bot.*, in press.
17. Doyle, J. J. & Doyle, J. L. (1987) *Phytochem. Bull.* **19,** 11–15.
18. Birky, C. W., Jr. (1988) in *Plant Evolutionary Biology*, eds. Gottlieb, L. D. & Jain, S. K. (Chapman & Hall, New York), pp. 23–53.
19. Putievsky, E. & Broue, P. (1979) *Aust. J. Bot.* **27,** 713–723.
20. Newell, C. A. & Hymowitz, T. (1983) *Am. J. Bot.* **70,** 334–348.
21. Grant, J. E., Grace, J. P., Brown, A. H. D. & Putievsky, E. (1984) *Aust. J. Bot.* **32,** 655–663.
22. Stebbins, G. L. (1971) *Chromosomal Evolution in Higher Plants* (Arnold, London).
23. Holsinger, K. & Gottlieb, L. D. (1988) *Syst. Bot.* **13,** 1–6.
24. Doyle, J. J., Doyle, J. L. & Brown, A. D. H. (1989) *Syst. Bot.* **14,** 398–407.