Theor Appl Genet (1985) 70:369–376

Ribosomal gene variation in soybean (*Glycine*) and its relatives

J. J. Doyle [1] and R. N. Beachy [2]

[1] L. H. Bailey Hortorium, Cornell University, Ithaca, NY 14853, USA
[2] Department of Biology, Washington University, St. Louis, MO 63130, USA

Received September 28, 1984; Accepted November 6, 1984
Communicated by J. MacKey

Summary. The genes encoding the 18S:25S ribosomal RNA gene repeat in soybean (*Glycine max*) and its relatives in the genus *Glycine* are surveyed for variation in repeat length and restriction enzyme site locations. Within the wild species of subgenus *Glycine*, considerable differences in repeat size occur, with a maximum observed in *G. falcata*. Repeat length and site polymorphisms occur in several species, but within individual plants only single repeat types are observed. The rDNA of the cultivated soybean and its wild progenitor, *G. soja* are identical at the level of this study, and no variation is found in over 40 accessions of the two species. Data from rDNA mapping studies are congruent with those of previous biosystematic studies, and in some instances give evidence of divergences not seen with other approaches.

Key words: *Glycine* – Ribosomal genes – Restriction enzymes – Evolution

Introduction

The cultivated soybean, *Glycine max* (L.) Merr., is both a major crop and an important experimental plant for molecular biological studies.

The genus *Glycine* Willd., as currently circumscribed, consists of nine species in two subgenera (Hermann 1962; Lakey 1977; Newell and Hymowitz 1980). The Annual cultivated soybean and its wild progenitor, *Glycine soja* Sieb. & Zucc., belong to subgenus *Soja* (Moench) F. J. Herm.; both species are native to northeastern Asia and have diploid chromosome numbers of 2n = 40 (Hymowitz 1976). Although high diploid chromosome numbers (x = 20 for the genus as a whole) reflect the evolutionarily polyploid nature of the group, cytogenetic data suggest that all species in the genus are fully diploidized (Newell and Hymowitz 1983). Seven perennial species belong to subgenus *Glycine* Willd., which has its center of distribution in Australia; chromosome numbers in most of these species are 2n = 40 (Newell and Hymowitz 1978a). Both diploids and polyploids (2n = 80) occur in *G. tabacina* (Labill.) Benth., while a second species, *G. tomentella* Hayata, includes diploids, aneuploids (2n = 38, 78) and polyploids (Newell and Hymowitz 1978a, 1983). Despite several biosystematic studies on both subgenera (Broich and Palmer 1980; Broué et al. 1977; Hymowitz and Newell 1978a, b, 1980, 1983; Putievsky and Broué 1979), much remains to be learned about the evolutionary relationships of this important genus.

The tandemly repeated multigene family encoding the 18S and 25S ribosomal RNAs has proven to be a useful tool for molecular evolutionary studies in plants (Appels and Dvorak 1982a, b; Dvorak and Appels 1982; Doyle et al. 1984). We have chosen to study the evolution of this gene family in an attempt to complement biosystematic studies of the genus *Glycine*. In this paper we report our initial findings of repeat length variation and restriction endonuclease site polymorphisms within both subgenera of *Glycine*.

Materials and methods

DNA isolation

Plants grown from seeds in the Washington University greenhouse were used for all studies. Voucher specimens of these plants were made and are deposited at BH. DNA was isolated from leaves by one of two methods. The first of these (Zimmer et al. 1983) includes equilibrium density centrifugation in CsCl gradients containing ethidium bromide, while the second procedure does not involve ultracentrifugation. This latter procedure, suitable for DNA isolation from small quantities of tissue, was modified from the method of Appels and Dvorak (1982a), and differs from their procedure primarily in that the proteinase K treatment was omitted. In this method, less than 0.1 g of fresh leaf tissue was powdered in a mortar and pestle in liquid nitrogen. To this powder, 1 ml of extraction buffer (0.1 M NaCl, 0.1 M EDTA, 50 mM Tris-HCl pH 7.0) was added and grinding was continued. The slurry was extracted

with two volumes of phenol-chloroform (1 : 1), and total nucleic acids were concentrated by precipitation with ethanol. RNA was eliminated from the preparation by treatment with DNase-free pancreatic RNase, and the DNA was collected by reprecipitation with ethanol. This procedure yielded from 50 to 100 μg of DNA per 0.1 g of fresh leaf tissue.

Restriction endonuclease digests

Restriction endonucleases (New England Biolabs) were used under the conditions specified by the manufacturer. Two to five μg samples of genomic DNA were subjected to agarose gel electrophoresis in gels containing 0.89 M tris-borate, 0.089 M boric acid, 0.05 M EDTA, pH 8.0. Following electrophoresis, gels were stained with ethidium bromide and photographed. DNA was transferred from the gels to nitrocellulose following the method of Southern (1975) as modified by Wahl et al. (1979).

DNA-DNA hybridization

Recombinant DNA molecules used as hybridization probes included pGmr-1, a complete 18S–25S rDNA repeat from soybean subcloned from a Charon 4A genomic clone (Jackson 1978) into the Eco R-I site of pBR325 (Zimmer and Walbot, in preparation), and various subclones of this plasmid. Additional probes were constructed by digestion of pGmr-1 with appropriate endonucleases followed by electroelution of fragments separated on agarose gels (Maniatis et al. 1982). Hybridization probes were labelled with ^{32}P by in vitro nick translation (Maniatis et al. 1975).

DNA-DNA hybridizations were performed at 42 C in 5× SSC (1×SSC=0.15 M NaCl, 0.015 M Na-citrate), 1×Denhardts (0.1% Ficoll, 0.1% polyvinylpyrollidine 360, 0.1% bovine serum albumin), 0.02 M sodium dodecyl sulfate (SDS), 40% formamide. Filters were washed in 2×SSC, 0.4% SDS at 65 C, and exposed to Kodak XAR5 X-ray film with one intensifying screen (Dupont Cronex Lightning Plus) for 1–3 days at –70 C.

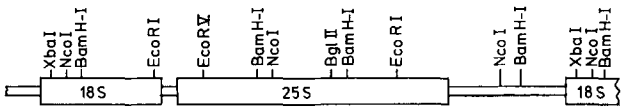

Fig. 1. Restriction endonuclease recognition site map of the *Glycine max* 18S: 25S ribosomal RNA gene repeat for six restriction endonucleases. The entire gene repeat in this species is 7.8 kb

Results

A map of the ribosomal repeat of *Glycine max* showing the recognition sites of several restriction endonucleases is shown in Fig. 1. Several of the sites on this map have previously been reported (Jackson 1978; Friedrich et al. 1979; Varsanyi-Breiner et al. 1979). The enzyme Xba-I has only a single recognition site in the ribosomal repeat of *G. max*. When Xba-I digested genomic DNA was fractionated on agarose gels, transferred to nitrocellulose and hybridized with ^{32}P pGmr-1, only a single band of hybridization was present in all *Glycine* acces-

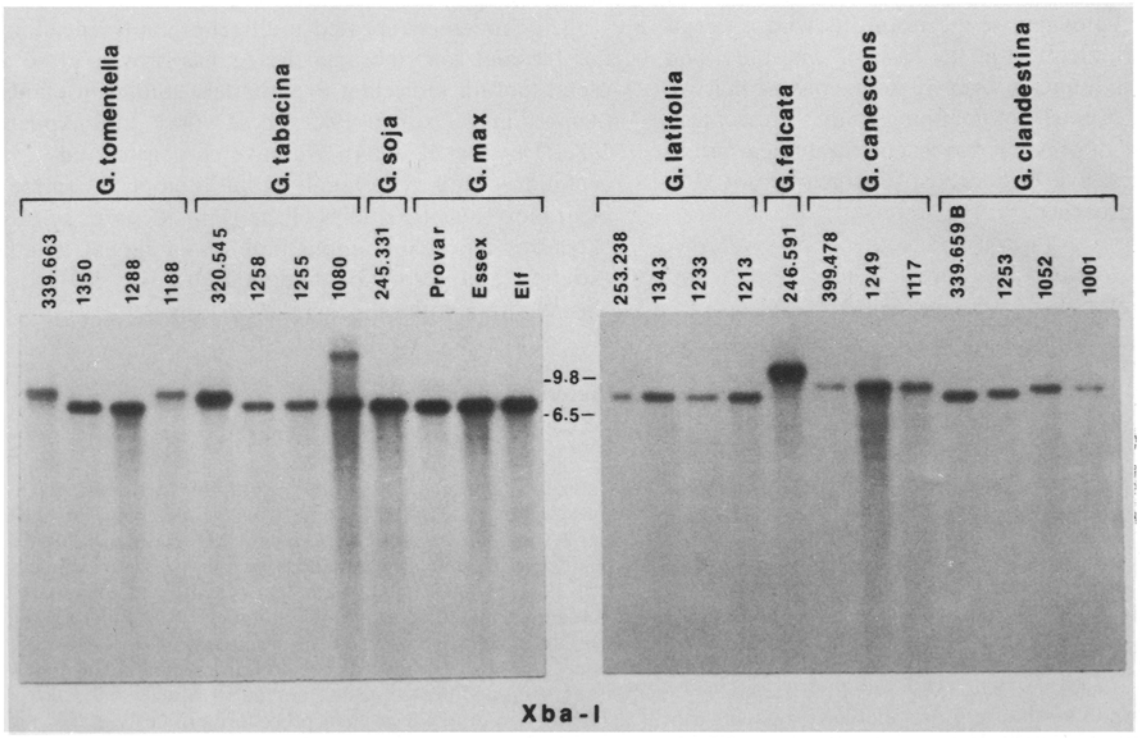

Fig. 2. Autoradiographs of Xba-I digested genomic DNA from different *Glycine* accessions separated on 0.7% agarose gels, transferred to nitrocellulose and hybridized to a cloned *G. max* rDNA repeat. Since Xba-I cuts only once per repeat, the sizes of the single band in each lane corresponds to the length of the ribosomal repeat in that accession. Numbers between the two gels are in kb, and are derived from molecular weight markers run on the same gel

Table 1. Variation among accessions of *Glycine* subgenus *Glycine* studied for rDNA variation. Accessions beginning with "G" originate with CSIRO, Australia; PI = USDA Plant Introduction. ND = not determined. NSW = New South Wales; Qld = Queensland; NT = Northern Territories. NTS = non-transcribed spacer. Polymorphic Eco R-I sites if present are denoted using abbreviations used in Fig. 6 (eg. E_c, E_d). Chromosome numbers are from a: Newell and Hymowitz 1978a; b: Newell and Hymowitz 1983; c: Broue et al. 1977; CSIRO, unpublished)

Species	Accession	Origin	Chromosome no. 2n	Repeat length (kb)	NTS Eco R-I site
canescens	G1117	NSW, Australia	ND	9.5	–
	G1249	NSW, Australia	ND	9.5	–
	G1340	NT, Australia	ND	8.2	–
	PI 339478	NT, Australia	40[a]	9.7	–
clandestina	G1001	NSW, Australia	ND	8.4	E_c
	G1052	NSW, Australia	ND	8.4	E_c
	G1253	NSW, Australia	ND	8.1	E_c
	PI 339659	NSW, Australia	40[a]	8.0	–
falcata	PI 246591	Qld, Australia	40[a]	12.0	–
latifolia	G1213	NSW, Australia	ND	8.3	E_d
	G1233	NSW, Australia	ND	8.3	–
	G1343	Qld, Australia	ND	8.3	E_d
	PI 253238	Qld, Australia	40[b]	8.3	–
	PI 319696	Qld, Australia	40[a]	8.3	E_d
	PI 321393	NSW, Australia	40[a]	8.3	E_d
latrobeana	PI 246591	Qld, Australia	ND	8.5	–
tabacina	G1143	NSW, Australia	40[d]	8.0	–
	G1199	NSW, Australia	40[d]	8.0	–
	G1209	NSW, Australia	40[d]	8.0	–
	PI 339661	NSW, Australia	40[a]	8.0	–
	PI 378703	Qld, Australia	40[a]	8.0	–
	PI 378705	unknown	40[a]	8.0	E_e
	G1080	NSW, Australia	80[d]	8.0	E_e
	G1255	NSW, Australia	80[d]	8.0	E_e
	G1258	New Caledonia	80[d]	8.0	E_e
	PI 320545	Taiwan	80[a]	9.2	–
tomentella	G1188	Qld, Australia	80[c]	9.7	–
	G1288	NSW, Australia	78[c]	7.8	–
	G1350	Taiwan	ND	7.8	–
	PI 320548	Taiwan	80[a]	8.0	E_e
	PI 339657	NSW, Australia	78[a]	7.8	–
	PI 339663	NSW, Australia	78[a]	7.8	–
	PI 441004	Qld, Australia	80[b]	9.7	–

sions tested, showing the existence of a single recognition site in their ribosomal repeats (Fig. 2). Similar results were obtained using other enzymes with one recognition site per repeat (Eco R-V, Bgl-II, data not shown). Only a single predominant repeat is visible within individual plants. The length of the ribosomal repeat varies considerably within the genus (Table 1), with several species, including *G. max,* having rather short repeats of about 8 kilobases (kb). Other species have longer repeats, with a maximum of 12 kb observed in *G. falcata* Benth. Although different repeat sizes are not observed within individuals of a given species, repeat length heterogeneity between different accessions of one species was observed.

Localization of repeat length heterogeneity in subgenus Glycine

Variation in the length of the ribosomal repeat was localized to the Eco R-I fragment bearing the non-transcribed spacer (NTS) region of the ribosomal repeat. Nitrocellulose filters bearing electrophoretically separated Eco R-I digested genomic DNAs were probed first with the entire *G. max* ribosomal repeat (plasmid pGmr-1) to establish the positions of all rDNA fragments (Fig. 3). Filters were then stripped of bound probe by washing with boiling low salt buffer and re-probed with a ^{32}P-labelled plasmid bearing the Xba-I to Bam H-I fragment of the 18S gene. Following

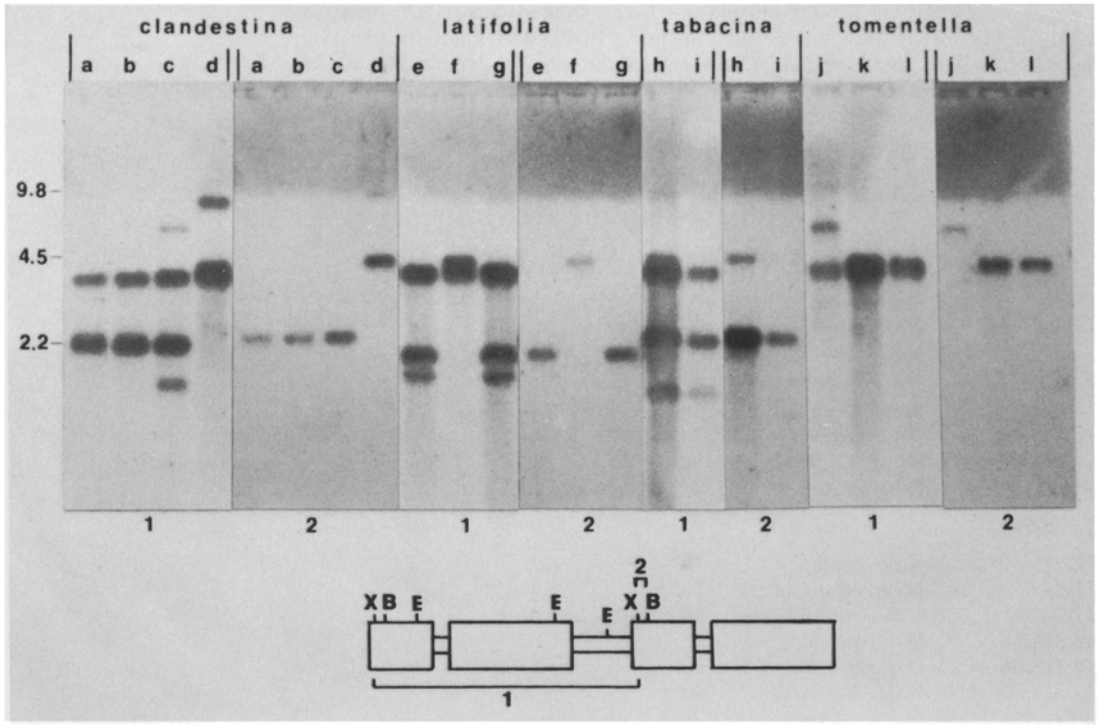

Fig. 3. Eco R-I digestions of genomic DNA from *Glycine* species showing repeat length and/or restriction site polymorphisms. *Glycine clandestina: a* G1001, *b* G1052, *c* G1258, *d* PI 339659; *G. latifolia: e* G1213, *f* G1233, *g* G1343; *G. tabacina: h* G1080, *i* G1255; *G. tomentella: j* G1188, *k* G1288. For each species two experiments are shown. Lanes labelled "1" are probed with the entire *G. max* rDNA repeat, represented on the gene map as the distance between Xba-I sites (area labelled "1" on map). Gels labelled "2" are the same filters, from which hybridized probe has been removed. These filters were subsequently re-hybridized with a probe containing the cloned Xba-I: Bam H-I fragment of the *G. max* 18S gene, the location of which is shown on the gene map (area 2). The bands hybridizing to this latter probe are a subset of the former hybridization and aid in the assignment of Eco R-I sites to the proper location on the gene map. From this experiment it is seen that the Eco R-I fragment bearing the NTS is responsible for repeat length differences, and that it is this fragment that bears the polymorphic Eco R-I sites in each species. the Eco R-I sites of one repeat are shown on the map for reference

such an experiment, the band of hybridization at 3.9 kb found in all *Glycine* species was not seen (Fig. 3). The size of the band or bands revealed by hybridization with this fragment accounted for the variation in repeat size observed in the different species. Further localization of size variation within *G. clandestina* Wendl. was also achieved by this experiment. Hybridization with this probe revealed only the larger of two NTS bands, a band common to the other two accessions bearing an NTS Eco R-I site (Fig. 3). Thus the length heterogeneity in this species occurs at the 5′ end of the NTS. Further delimitation of the principal repeat length heterogeneity among *Glycine* species was apparent by mapping repeats with the enzyme Nco-I. Although some length heterogeneity is apparent in the region of the NTS just 5′ of the 18S gene, this variation does not account for the differences observed between the shorter repeat of *G. max* and the longer repeats of several other *Glycine* species – in fact, the Nco-I fragment 5′ of the 18S gene is longer in *G. max* and *G. soja* than in any other member of the genus (Fig. 4). Repeat length variation in the genus thus occurs between the Eco R-I site in the 25S gene and the Nco-I site in the NTS.

Restriction endonuclease recognition polymorphisms for the enzyme Eco R-I were found in several species belonging to subgenus *Glycine* (Figs. 3 and 5). In each case, the polymorphic site was mapped to the NTS (Fig. 6). Although these sites map near each other in each case, those in at least two of the three species (*G. clandestina* and *G. latifolia* (Benth.) Newell and Hymowitz) are different, based on 1) their positions relative to shared sites in the NTS and 2) their occurrence in different positions within a shared Nco-I fragment of constant size.

Subgenus Soja

Over 40 different accessions of subgenus *Soja* were analyzed for length or Eco R-I site polymorphisms such as those observed elsewhere in the genus (Table 2). Sampling included accessions from throughout the native geographical ranges of *G. max* and *G. soja,* as

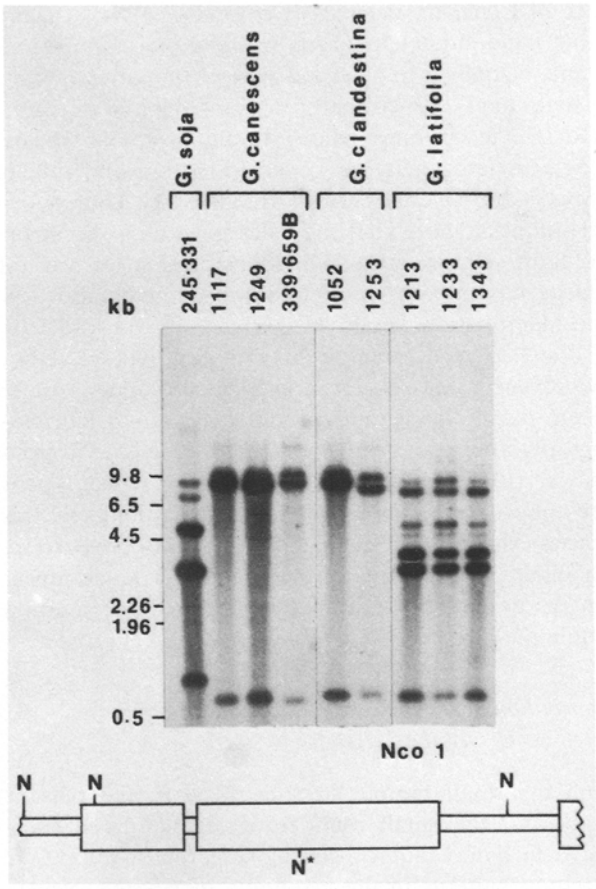

Fig. 4. Nco-I digests of *Glycine* genomic DNA probed with the *G. max* rDNA repeat. The large number of high molecular weight bands in each case is due to incomplete digestion of all repeats by the enzyme. Note that the smallest fragment seen on these gels is, in all but *G. soja*, of nearly identical size (1.1 kb)

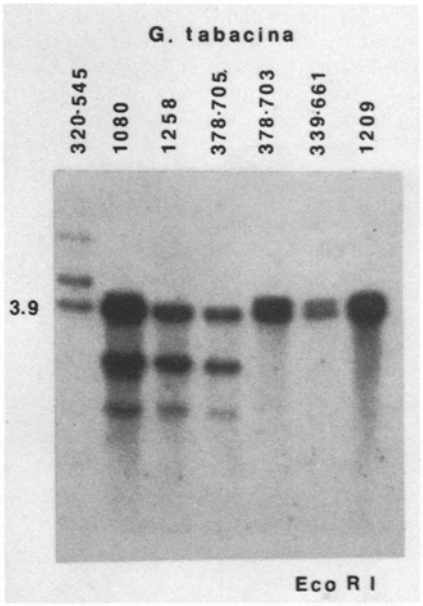

Fig. 5. Repeat length and Eco R-I site polymorphisms in *Glycine tabacina*. The probe used in this experiment is the full *G. max* rDNA repeat. Accession numbers are given above lanes

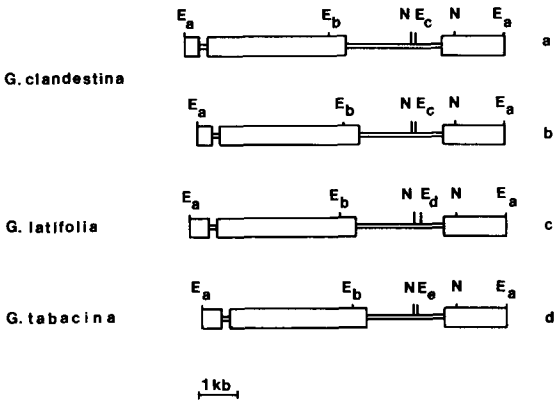

Fig. 6. Restriction maps of ribosomal repeats of subgenus *Glycine* species showing Eco R-I site polymorphisms. N = Nco-I; E = Eco R-I recognition site; subscripts are used to denote sites in different positions (E_a and E_b are common to alle *Glycine* accessions). Examples of repeat types for each species: *a* G1001; *b* PI 339659; *c* G1213; *d* G1080

well as a diversity of morphological types within each species. In addition, most of the physiological diversity within the two species, as represented by known maturity groups, was included in the sampling. Despite the diversity of the sample, no variation was found between or within these two species.

Discussion

Lack of within-individual heterogeneity

The presence of a single major ribosomal gene repeat class in each individual plant of the various *Glycine* accessions extends the findings of Friedrich et al. (1979) and Varsanyi-Breiner et al. (1979) to species of the genus other than the cultivated soybean. These findings contrast with the substantial heterogeneity reported by Jackson (1978) who found that as much as 30% of the repeats in soybean tissue culture lines were of a size class larger than the 7.8 kb type found almost exclusively in this study.

The absence of detectable quantities of repeats of either different length or restriction profile from the genomes of individual plants is of interest considering the evolutionary history of the genus *Glycine*. Although fully diploidized, at x = 20 the genus is considered an ancient polyploid whose progenitor(s) were probably x = 11 plants, like the bulk of the tribe Phaseoleae to which *Glycine* belongs (Lackey 1981). Since the nearest

Table 2. Accessions of *Glycine max* and *G. soja* studied. USDA plant introduction numbers and varietal names are given where available. All accessions have repeat lengths of 7.8 kb and lack NTS Eco R-I sites. Maturity groups of *G. max* are from Hymowitz et al. (1977)

Species	PI no.	Cultivar	Maturity group	Origin
max	81031	'Bansei'	II	Japan
	30744	'Black Eyebrow'	II	China
	54563-3	'Boone'	IV	China
	71663	'Charlee'	VII	China
	85355	'Delsoy'	VI	Korea
	–	'Elf'	–	–
	–	'Essex'	V	–
	81039	'Fuji'	III	Japan
	20405	'Habaro'	I	USSR
	54606-3	'Harbinsoy'	IV	China
	81038	'Hidatsa'	00	Japan
	22406	'Hongkong'	IV	Hongkong
	30746	'Hoosier'	I	China
	82202	'Jefferson'	IV	Korea
	–	'Keburi'	–	Japan
	40658	'Laredo'	VI	China
	30593	'Manchu'	III	China
	71664	'Missoy'	VII	China
	71608	'Monetta'	VII	China
	–	'Mo-shi-dou'	–	Japan
	80456	'Osaya'	III	Japan
	–	'Pando'	00	Korea
	128182	'Poland Yellow'	0	–
	–	'Provar'	II	–
	81021	'Sioux'	00	Japan
	19186-D	'Virginia-S'	V	China
	–	'Wayne'	III	–
	–	'Williams'	III	–
	86904-1	–	–	–
soja	65549	–	II	China
	81762	–	II	USSR
	101404-B	–	II	China
	163413	–	–	–
	245331	–	X	Taiwan
	326582-B	–	I	USSR
	339732	–	IV	Korea
	339871-A	–	V	Korea
	342619-A	–	0	USSR
	342621-A	–	00	USSR
	342622-A	–	I	USSR
	406684	–	III	Japan
	407290	–	II	China
	407294	–	II	China
	424002	–	00	USSR

extant relatives of *Glycine* are currently unknown, it is impossible to state whether the event leading to the origin of *Glycine* was primarily auto- or allopolyploid, and, hence, whether one would expect rDNA heterogeneity in individual plants due to fixed hybridity. Intra-genomic heterogeneity can be found in non-hybrid diploid plants, as in *Vicia faba,* where four different spacer lengths occur within individual plants (Yakura et al. 1984), and in polyploids, such as one race of *Claytonia virginica* (Doyle et al. 1984). On the other hand, no heterogeneity is found in autopolyploid plants belonging to a second race of the latter species, (Doyle et al. 1984), nor in autopolyploid *Tolmiea* (Doyle et al., in preparation). In allohexaploid wheats, length heterogeneity is observed within individuals (Appels and Dvorak 1982a). The diversity of these observations suggests that the rules governing the incorporation, elimination or homogenization of variant repeat types, both in diploids and polyploids, vary considerably from species to species.

Only a single example of within-individual rDNA heterogeneity was observed even in presumably more recent polyploids within *Glycine tabacina* – a morphologically anomalous polyploid (2n=80) accession from Taiwan (PI 320545) in which a secondary repeat type is present in low abundance. Other polyploids in this species exhibit only a single repeat type. In *G. tomentella,* polyploids and aneuploids contain only a single repeat length, despite the existence of repeat size variation within the species.

Correlation of rDNA variation and taxonomy in subgenus Glycine

Both the small sample sizes involved in the present study and the small number of restriction enzymes tested to date caution against overemphasis of rDNA variation patterns in the taxonomy of *Glycine.* Nevertheless, the patterns observed are congruent with previous morphological and cytogenetic data, and in some cases further elucidate taxonomic relationships.

Glycine clandestina is a morphologically complex entity, with several infraspecific groups recognized by Newell and Hymowitz (1978b, 1983) on the basis of leaf morphology, habit and seed surface morphology. Three accessions of this species, though varying in their rDNA repeat lengths, are similar in possessing a recognition site for the enzyme Eco R-I in the NTS region (Eco R-I_c in Fig. 6; Table 1). All three of these accessions represent "typical" *G. clandestina* in being twining plants with digitately trifoliolate leaves and straight legumes. A fourth accession, PI 339659, is considered atypical in the species, approaching *G. tabacina* in habit and morphology (Newell and Hymowitz 1978b). The rDNA repeat of this plant, like that of diploid *G. tabacina,* lacks the NTS Eco R-I site and is somewhat shorter than those of the other *G. clandestina* accessions. Our findings thus lend further support to the existence of definable evolutionary lineages within what is currently recognized as *G. clandestina.*

Little morphological differentiation has accompanied polyploidization in *G. tabacina* (Newell and Hymowitz 1978b). The rDNA of this species is polymorphic for an Eco R-I site (Eco R-I_e, Fig. 6) in the

NTS. Four of the five diploid accessions studied lack this site (Table 1), the single exception being a morphologically anomalous plant of unknown geographic origin (Newell and Hymowitz 1978 b). With one exception the polyploids represent a uniform group, with short repeats having three Eco R-I sites. The exception, PI 320545, is also the only representative of the narrow-leafleted Taiwanese form of the species included in our study. The major repeat in this plant is longer than that found elsewhere in the species, and lacks Eco R-I_e typical of Australian polyploids. However, a second repeat type, visible only after considerable overexposure of films, appears identical to that of Australian polyploids both in size and in the presence of Eco R-I_e. The fact that this minor repeat type exists in the genome of this plant suggests its relationship to the Australian polyploids. Ten individuals from this seed lot all possessed both repeat forms. With regard to the typical polyploids, our findings represent the first evidence of divergence between cytotypes within *G. tabacina* since Australian diploids and polyploids may be readily distinguished by the presence, in polyploids, of Eco R-I_e. This divergence is also reflected in other restriction site differences in the 18S : 25S gene repeat as well as in 5S ribosomal gene repeat length (Doyle and Beachy, in preparation).

In *Glycine canescens* F. J. Herm., no restriction site polymorphism were observed, but three different repeat lengths were found among the four accessions tested. This species and *G. clandestina* will form viable F_1 hybrids, suggesting that the two species may be more closely related than their very different morphology and ecology would suggest (Putievsky and Broué 1979; Newell and Hymowitz 1983). The rDNA of these species appears different, however, at least at the level thus far studied.

Length heterogeneity also occurs within *Glycine tomentella,* with repeat lengths of 7.8 and 9.7 kb observed. Repeat size classes were correlated with both geographic distribution and chromosome number, with the two Queensland tetraploids having the larger repeat type and the three New South Wales 78-chromosome aneuploids having a shorter repeat. These two groups are also morphologically distinguishable under greenhouse conditions. Ribosomal gene differentiation follows along the lines observed for isozymic markers (Broué et al. 1977) and crossability groups (Putievsky and Broué 1979; Newell and Hymowitz 1983). The apparent lack of differentiation between Taiwanese accession G1350, presumably a tetraploid, and the New South Wales aneuploids is also of interest. Since only a single repeat is observed in each plant, this study is unable to shed any new light on the identities of the genomes making up these polyploids and aneuploids. A single tetraploid *G. tomentella* accession, PI 320548, has a short repeat and an NTS Eco R-I site. This accession is anomalous with regard to various morphological characters and is considered by Newell and Hymowitz (1978 b) to approach *G. canescens.* The presence of an NTS Eco R-I site, however, distinguishes this accession from all representatives of *G. canescens* thus far encountered. The location of this site is identical to that of Eco R-I_e, and is, in this respect as well as in total repeat length, identical to tetraploid Australian *G. tabacina,* suggesting that its affinities may lie elsewhere than previously thought.

Recent biosystematic work has resulted in the recognition of *Glycine latifolia* as a diploid species separate from both *G. tabacina* and *G. tomentella* (Newell and Hymowitz 1980). The six accessions of *G. latifolia* studied all had identical repeat lengths, but four accessions were found to possess an NTS Eco R-I site (Eco R-I_d) not found in the other accessions. The two accessions lacking this site were in no other way remarkable. All of the numerous individual plants grown from a line lacking this site (G1133) lacked the site themselves.

The remaining two species of subgenus *Glycine* were each represented in this study by only one accession. The single accession of *G. falcata* is notable in possessing a ribosomal repeat far larger than what is common in the genus. Thus, in rDNA repeat length, as in a number of other morphological, chemical and genetic characters (Mies and Hymowitz 1973; Newell and Hymowitz 1978 b; Putievsky and Broué 1979), this species appears to stand apart from the rest of the subgenus. Based on one sample, *Glycine latrobeana* (Meissn.) Benth., has a repeat of intermediate length lacking NTS Eco R-I sites.

Subgenus Soja and its relationship to subgenus Glycine

The cultivated soybean and its presumed wild progenitor, *Glycine soja,* are indistinguishable from one another on the basis of rDNA repeat length or Eco R-I polymorphisms. Data from 5S ribosomal gene repeat lengths also indicate that these two species are nearly identical (Doyle and Beachy, in preparation), as do studies in which the degree of nucleotide homology between NTS regions of the two species' rDNA has been determined by heteroduplex formation and thermal elution (Doyle and Beachy, unpublished). These findings extend the observation of identical seed protein profiles in the two species (Mies and Hymowitz 1973) to additional molecular characters. The obvious morphological differences between the two species are nearly all associated with the transformation of a wild species into a domesticated plant (e.g. loss of vining habit and increased seed size) and mask the great similarity between the two taxa.

Whereas heterogeneity for length and/or restriction sites was observed among small numbers of accessions in species of subgenus *Glycine,* no such variation was observed in over 40 samples of subgenus *Soja.* The species of both subgenera are inbreeders, with members of the former subgenus often producing cleistogamous flowers. The wild species of subgenus *Glycine* generally occur in small, isolated populations, and this, together with their breeding system, is presumably responsible for the degree of variability observed within taxa (Newell and Hymowitz 1978a). Perhaps human interest in soybean, with concomitant wide distribution of seeds and concerted efforts at plant breeding, has served as the homogenizing force in the history of subgenus *Soja* that has been lacking for the species in subgenus *Glycine.*

Because of the great amount of variability observed within subgenus *Glycine,* little can be said at this point about the relationship between the two subgenera. Only by looking at a larger number of restriction sites can the degree of differentiation between the two groups be measured. However, preliminary results do suggest that some enzyme sites do divide the genus along subgeneric lines. Thermal elution studies indicate that differences in the most variable region of the NTS are on the order of 5–10 nucleotide percent between the two subgenera (Doyle and Beachy, unpublished). These studies should help elucidate the evolutionary relationships of *Glycine.*

Acknowledgements. The authors gratefully acknowledge the technical assistance of Jane Leclere Doyle. We also thank Dr. J. Peacock and other members of C.S.I.R.O., Australia, for providing us with numerous accessions of *Glycine* species, and Dr. Xavier Delannay for making accessible to us seed lines maintained at Monsanto Agricultural Products Company. Rob DeSalle is thanked for critical comments on the manuscript, and, with Liz Zimmer, for contributions at various stages of the project. JJD was supported by a Monsanto Postdoctoral Fellowship; RNB is supported by National Science Foundation grant PCM-8203022 and Department of Energy grant DE-AC02-81ER10888.

References

Appels R, Dvořák J (1982a) The wheat ribosomal DNA spacer region: its structure and variation in populations and among species. Theor Appl Genet 63:337–348

Appels R, Dvořák J (1982b) Relative rates of divergence of spacer and gene sequences within the rDNA region of species in the *Triticeae:* implications for the maintenance of homogeneity of a repeated gene family. Theor Appl Genet 63:361–371

Broich SL, Palmer RG (1980) A cluster analysis of wild and domesticated soybean phenotypes. Euphytica 29:23–32

Broué P, Marshall DR, Muller WJ (1977) Biosystematics of subgenus *Glycine* (Verdc.): isoenzymatic data. Aust J Bot 25:355–366

Doyle JJ, Beachy RN, Lewis WH (1984) Evolution of rDNA in *Claytonia* polyploid complexes. In: Grant WF (ed) Plant biosystematics. Academic Press, Canada Ontario, pp 321–341

Dvořák J, Appels R (1982) Chromosome and nucleotide sequence differentiation in genomes of polyploid *Triticum* species. Theor Appl Genet 63:349–360

Friedrich H, Hemleben V, Meagher RB, Key JL (1979) Purification and restriction endonucleases mapping of soybean 18S and 25S ribosomal RNA genes. Planta 146: 467–473

Hermann FJ (1962) A revision of the genus *Glycine* and its immediate allies. Techn Bull USDA 1268:1–79

Hymowitz T (1976) Soybeans. In: Simmonds NW (ed) Evolution of crop plants. Longman, London, pp 159–162

Hymowitz T, Newell CA, Carmer SG (1977) Pedigrees of soybean cultivars released in the United States and Canada. Int Agric Publ INTSOY Ser 13:1–23

Jackson PJ (1978) Characterization of the ribosomal ribonucleic acid genes of soybean cells grown in suspension culture. Doctoral Thesis, Iowa State University, Ames

Lackey JA (1977) *Neonotonia,* a new generic name to include *Glycine wightii* (Arnott) Verdcourt (Leguminosae, Papilionoideae). Phytologia 37:209–212

Lackey JA (1981) Phaseoleae. In: Polhill RM; Raven PH (eds) Advances in legume systematics. Royal Botanical Gardens, Kew, pp 301–327

Maniatis T, Fritsch EF, Sambrook J (1982) Molecular cloning: a laboratory manual. Cold Spring Harbor, New York

Maniatis T, Jeffrey A, Klein DG (1975) Nucleotide sequence of the rightward operator of phage λ. Proc Natl Acad Sci USA 72:1184–1189

Mies DW, Hymowitz T (1973) Comparative electrophoretic studies of trypsin inhibitors in seed of the genus *Glycine.* Bot Gaz 134:121–125

Newell CA, Hymowitz T (1978a) A reappraisal of the subgenus *Glycine.* Am J Bot 65:168–179

Newell CA, Hymowitz T (1978b) Seed coat variation in *Glycine* Willd. subgenus *Glycine* (Legumonisae) by SEM. Brittonia 30:76–88

Newell CA, Hymowitz T (1980) A taxonomic revision in the genus *Glycine* subgenus *Glycine* (Leguminosae). Brittonia 32:63–69

Newell CA, Hymowitz T (1983) Hybridization in the genus *Glycine* subgenus *Glycine* Willd. (Leguminosae, Papilionoideae). Am J Bot 70:334–348

Putievsky E, Broué P (1979) Cytogenetics of hybrids among perennial species of *Glycine* subgenus *Glycine.* Aust J Bot 27:713–723

Southern EM (1975) Detection of specific sequences among DNA fragments separated by gel electrophoresis. J Mol Biol 98:503–517

Varsanyi-Breiner A, Gusella JF, Keys C, Housman DE, Sullivan, Brisson N, Verma DPS (1979) The organization of a nuclear DNA sequence from a higher plant: molecular cloning and characterization of soybean ribosomal DNA. Gene 7:317–334

Wahl GM, Stern M, Stark GR (1979) Efficient transfer of large DNA fragments from agarose gels to diazobenzoxymethyl paper and rapid hybridization by using dextran sulfate. Proc Natl Acad Sci USA 79:3683–3688

Yakura K, Kato A, Tanifuji S (1984) Length heterogeneity of the large spacer of *Vicia faba* rDNA is due to the differing number of a 325 bp repetitive sequence element. Mol Gen Genet 193:400–405

Zimmer EA, Rivin C, Walbot V (1983) A DNA isolation procedure suitable for most higher plant species. Plant Mol Biol Newslett 2:93–99