Evolutionary Dynamics and Preferential Expression of Homeologous 18S-5.8S-26S Nuclear Ribosomal Genes in Natural and Artificial *Glycine* Allopolyploids

Simon Joly,[*][1] *Jason T. Rauscher,*[*][2] *Susan L. Sherman-Broyles,* *A. H. D. Brown,†*
*and Jeff J. Doyle**

*L. H. Bailey Hortorium, Cornell University, Ithaca, New York; and †Centre for Plant Biodiversity Research, CSIRO Plant Industry, Canberra, Australia

Polyploidy is an important evolutionary process in plants, but much remains to be learned about the evolution of gene expression in polyploids. Evolution and expression of the 18S-5.8S-26S ribosomal gene family was investigated at homeologous loci in the *Glycine* subgenus *Glycine* perennial soybean polyploid complex, which consists of several diploid genomes that have formed allopolyploids in various combinations, often recurrently. A semiquantitative PCR method targeting the internal transcribed spacer (ITS) of the 18S-5.8S-26S nuclear ribosomal DNA (nrDNA) was used to survey the ratio between homeologous repeats in polyploid genomes and to test for preferential expression of homeologous nrDNA loci. Most natural polyploids possess one predominant nrDNA homeolog in their genome. Analysis of F_2 segregation in an artificial cross suggested that in some plants, most or all repeats at one homeologous locus have been lost, whereas in other plants two loci remain, but both have been homogenized by concerted evolution. In most natural allopolyploids harboring a relatively balanced ratio of homeologs, one homeolog was expressed preferentially, but in the majority of plants, low levels of transcription could be detected from the other homeolog. Individuals within some tetraploid taxa varied as to which homeolog was expressed preferentially. In some plants, the degree of preferential expression also varied among tissues. Preferential expression was absent in synthetic polyploids and in some artificial diploid hybrids, suggesting that nucleolar dominance is not necessarily a direct result of hybridization or polyploidization. The establishment of preferential expression in *Glycine* allopolyploids appears to be either stochastic within lineages or genotype specific.

Introduction

Polyploidy is a prevalent feature in the plant kingdom (Stebbins 1950; Otto and Whitton 2000), being found abundantly in algae (Nichols 1980), mosses (Crosby 1980), ferns (Wagner and Wagner 1980), and flowering plants (Stebbins 1950; Grant 1981). For example, between 50% and 70% of flowering plants are considered to have experienced a polyploid event in their past, depending on whether the threshold haploid chromosome number for a plant to be considered a polyploid is set to $n = 11$ (50% [Stebbins 1950]) or to $n = 10$ (70% [Grant 1981]). Interestingly, these numbers could underestimate the proportion of polyploid taxa, given that even *Arabidopsis thaliana*, with a haploid chromosome number of $n = 5$, likely has undergone two ancient polyploid events (Simillion et al. 2003; Ziolkowski, Blanc, and Sadowski 2003).

Besides its prevalence, polyploidy is also important for its role in plant speciation (Otto and Whitton 2000), presumably because of repercussions of polyploidy on the ecology (Favarger 1967; Ehrendorfer 1980; Lewis 1980; Petit and Thompson 1999), phenology (Comai et al. 2000; Schranz and Osborn 2000; Ramsey and Schemske 2002), and physiology (Tal 1980; Levin 1983; Thompson et al. 1997; Levin 2002) of polyploid species. If these character-

istics and the preponderance of polyploid taxa are indicative of potential advantages relative to their diploid progenitors, it is important to understand how such advantages may arise. Genome evolution has often been invoked to explain the evolutionary success of polyploids (Soltis and Soltis 1999, 2000; Wendel 2000). Along with structural changes in polyploid genomes, some of which are known to occur rapidly (Song et al. 1995; Liu, Vega, and Feldman 1998; Liu et al. 1998; Ozkan, Levy, and Feldman 2001; Shaked et al. 2001), gene regulation is known to be affected by polyploidy (Galitski et al. 1999; Comai et al. 2000; Adams et al. 2003; Osborn et al. 2003) and could also contribute to the adaptive potential of polyploids.

The nuclear 18S-5.8S-26S ribosomal gene family (nrDNA) is a classic example of a locus for which studies have been conducted at both the structural and the expression levels in polyploids. At the structural level, it is a useful marker for systematic studies (White et al. 1990; Baldwin et al. 1995), and it also has been used to study genome evolution (e.g., concerted evolution [Hamby and Zimmer 1992; Baldwin et al. 1995]). Expression studies of nrDNA are embedded in the more general subject of nucleolar dominance (Reeder 1985; Flavell 1986; Pikaard and Chen 1998; Comai 2000; Pikaard 2000), an epigenetic phenomenon in which the nrDNA of only one parent is transcribed in a diploid hybrid or allopolyploid (reviewed in Pikaard and Chen [1998]). Although the mechanisms for establishment and maintenance of nucleolar dominance are beginning to be understood (Pikaard and Lawrence 2002), there is no single hypothesis that can explain the range of expression patterns found in the different organisms investigated to date (Pikaard and Chen 1998). Moreover, the few species that have been investigated for nucleolar dominance at the molecular level are almost
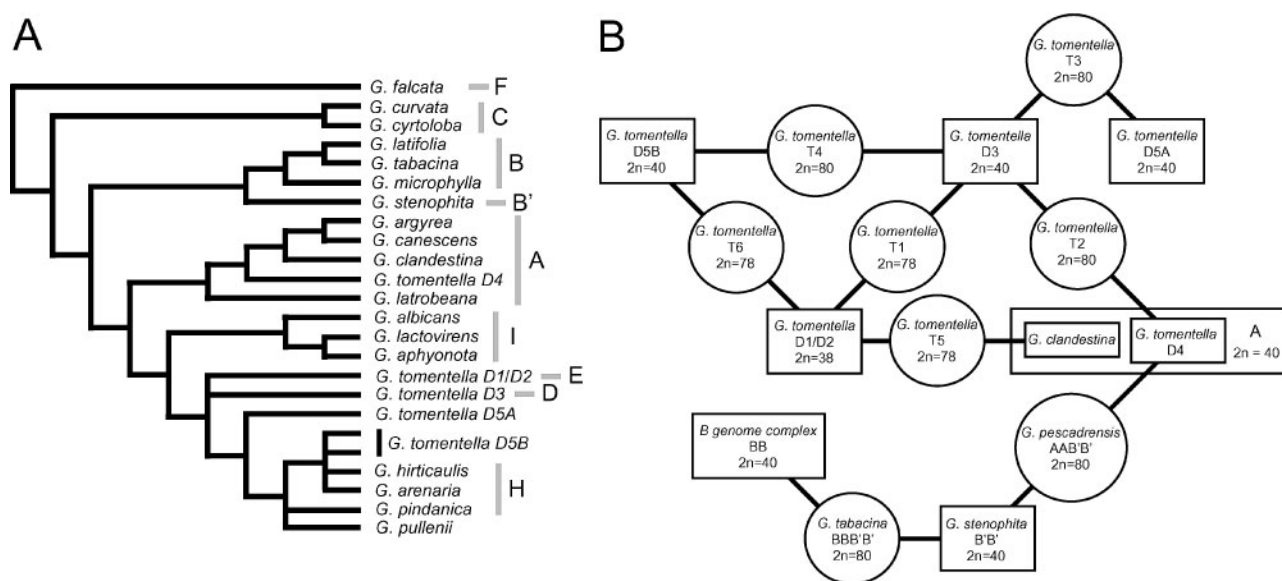
Fig. 1.—Relationships of diploid taxa (*A*) of *Glycine* subgenus *Glycine* based on histone H3-D sequences (modified from Brown et al. [2002]). The genome groups of species are indicated to the right of the cladogram. (*B*) Polyploid taxa of the perennial *Glycine* polyploid complex and their relationships to their diploid progenitors as inferred from histone H3-D data (modified from Doyle et al. [2002]). Diploid taxa are in boxes and tetraploids are circled. Taxa and diploid chromosome numbers are indicated.

exclusively "model" taxa such as *Xenopus* (Honjo and Reeder 1973; Caudy and Pikaard 2002), *Arabidopsis* (Chen, Comai, and Pikaard 1998; Lewis and Pikaard 2001; Pontes et al. 2003), *Brassica* (Chen and Pikaard 1997*a*, 1997*b*; Frieman et al. 1999; Hasterok and Maluszynska 2000), and wheat (Houchins et al. 1997; Neves et al. 1997), and, to date, most studies have mainly focused on artificial hybrids. Few wild plant taxa have been studied, and there has been little attention paid to geographical sampling or to investigating variation between individuals in natural populations.

The genus *Glycine*, which includes the soybean (*G. max*), offers a good model for studying nrDNA evolution in natural allopolyploids. The primarily Australian perennial subgenus *Glycine* includes a large, recently formed allopolyploid complex comprising several diploid genomes that have formed polyploids in various combinations, often recurrently (Doyle et al. [2004] and figure 1). This complex has long been recognized as polytypic, with polyploids classified under three distinct species epithets: *G. tabacina* (Labill.) Benth, *G. pescadrensis* Hayata (formerly known as *G. tabacina* AAB′B′) and *G. tomentella* Hayata (reviewed in Doyle et al. [2004] and figure 1). The diploid progenitors of *G. pescadrensis* are reproductively isolated from one another (Hymowitz, Singh, and Kollipara 1998; A.H.D. Brown and R. Palmer, unpublished data), as are the diploid taxa that gave rise to the *G. tomentella* polyploids (*G. tomentella* D1-D5B taxa and related species), with the exception of the *G. tomentella* D1 and D2 taxa (Doyle, Grant, and Brown 1986). In contrast, the diploid progenitors of the allotetraploid *G. tabacina* (BBB′B′), *G. stenophita* (B′B′), and species of the B genome complex (fig. 1) are partially fertile at the diploid level (A.H.D. Brown and R. Palmer, unpublished data). Whereas diploid taxa of the complex are almost totally restricted to

Australia, almost all polyploids (except for the restricted *G. tomentella* T5 and T6 taxa) have also colonized islands of the Pacific Ocean, some reaching Taiwan and the Ryukyu Islands (Doyle et al. 1990*b*; Brown et al. 2002; Doyle et al. 2002).

Previous studies of the ITS region of the 18S-5.8S-26S nrDNA gene family in *Glycine* have shown that almost all of these allopolyploids are additive for parental loci, although often in unequal amounts (Doyle et al. 1990*b*; Rauscher, Doyle, and Brown 2002, 2004). Here, we use semiquantitative PCR methods (Rauscher, Doyle, and Brown 2002) to survey for variation in the ratio of homeologous repeats in polyploid genomes and to test for preferential expression of these repeats. We do so at the level of the whole complex, among plants within a polyploid taxon (genome combination) that sometimes evolved recurrently, within plants from a single polyploid origin, between siblings of a single plant, and between tissues of a single plant. We compare these natural polyploids with artificial crosses in an effort to distinguish among patterns due to the hybridization event, to polyploidy itself, to factors occurring soon after polyploidy (e.g., epigenetic changes), or to events occurring during the evolution of the polyploid lineages.

Materials and Methods
Plant Material

Accessions were selected from the CSIRO Perennial Soybean Germplasm Collection (Canberra, Australia) to represent all six *G. tomentella* polyploid taxa (genome combinations) and *G. tabacina* polyploids. Multiple accessions were chosen to represent known cases of recurrent origins within each polyploid taxon and to cover the geographic ranges of each (table 1). Accessions generally

Table 1
Tetraploid Accessions of *Glycine Tomentella* Taxa and *G. Tabacina* Used in This Study

Taxon	Accession Number (G)	Provenance[a]	Tissues Investigated for rRNA Proportions
Glycine tomentella Hayata			
T1	1133	QLD	—
	1136	NSW	—
	1288	NSW	Leaves
	1361	Papua New Guinea	—
	1367	Papua New Guinea	Leaves
	1392	QLD	Leaves
	1427	QLD	—
	1468	QLD	—
	1763	QLD	—
T2	1134	QLD	—
	1188	QLD	Leaves
	1286	QLD	—
	1393	QLD	Leaves
	1412	QLD	—
	1811	QLD	—
	1854	Taiwan	—
T3	1359	Papua New Guinea	—
	1397	QLD	—
	1930	NT	Leaves
	2098	W. Timor, Indonesia	Leaves
	2099	W. Timor, Indonesia	Leaves
	2100	W. Timor, Indonesia	—
	2539	W. Timor, Indonesia	—
T4	1304	NT	—
	1348	Bataan Island, Philippines	—
	1350	Maopitou, Taiwan	Leaves
	1469	QLD	Leaves
	1747	QLD	Leaves
	2437	WA	—
	2468	QLD	Leaves, cotyledons, hypocotyl
	2469	QLD	Leaves
	2470	QLD	—
	2476	QLD	Leaves, cotyledons, hypocotyl
	2557	QLD	—
T5	1487	NSW	Leaves
	1739	NSW	Leaves, developing leaf, roots
	1969	NSW	Leaves, cotyledons, hypocotyl
T6	1945	WA	Leaves
Glycine tabacina (Labill.) Benth			
	1072	Mariana Islands	Leaves, cotyledons, hypocotyl
	1075	ACT	Leaves
	1080	NSW	Leaves, cotyledons, hypocotyl
	1205	New Caledonia	Leaves
	1234	NSW	Leaves, cotyledons, hypocotyl
	1254	NSW	Leaves, cotyledons, hypocotyl
	1255	NSW	Leaves
	1988	New Caledonia	Leaves, cotyledons
	2263	QLD	Leaves, cotyledons, hypocotyl

NOTE.—The table lists the G number (CSIRO Perennial *Glycine* Collection) and the provenance. The USDA PI (Plant Introduction) accession number, the latitude and longitude, and further details for these collections are given on the AusPGRIS Web site. For *G. tomentella*, the different taxa (species; T1–T6) are also indicated. The tissue(s) investigated for rRNA proportions are indicated for each accession, when applicable.

[a] Australian state or Pacific island. Australian state abbreviations: NSW = New South Wales, NT = Northern Territory, QLD = Queensland, WA = Western Australia.

consist of seed multiplied by selfing a single initial wild-collected individual. Because *Glycine* species are predominantly selfers with low levels of heterozygosity, individuals within an accession are expected to exhibit very little genetic polymorphism.

We also studied hybrid plants resulting from artificial crosses that had been made primarily to investigate cytogenetic relationships among species (e.g. Grant et al. [1984]); Doyle, Grant, and Brown [1986]; and A. H. D.

Brown and R. Palmer, unpublished). Because the hybrids were constructed for this purpose, only a small number of genome combinations found in natural polyploids were available, and some represented combinations not found in nature. Three main types of crosses were studied. The first group included artificial polyploids derived from colchicine treatment of hybrids between diploid or tetraploid accessions (table 2). The second type consisted of an artificial homoploid hybrid (*sensu* Rieseberg [1997] and Ferguson

Table 2
Characteristics of the Artificial Hybrids Included in the Study

Accession	Ploidy Level	Parentage (♀ × ♂)	Generation Studied	Fertility
Artificial polyploids				
A39/1	2n = 4×	D1 (G1316) × *G. clandestina* (G1253)	F$_2$	Fertile
A56/1	2n = 4×	D1 (G1316) × *G. canescens* (G1351)	F$_2$	Fertile
A57/1	2n = 4×	D1 (G1316) × *G. canescens* (G1232)	F$_2$	Fertile
A58/1	2n = 4×	D1 (G1316) × *G. canescens* (G1232)	F$_2$	Fertile
A30	2n = 8×	*G. tomentella* T2 (G1188) × *G. tomentella* T1 (G1133)	F$_2$	Fertile
Homoploid tetraploid hybrids				
H204	2n = 4×	T1 (G1468) × T1 (G1392)	F$_2$	Fertile
Homoploid diploid hybrids				
RP8/1	2n = 2×	D2 (G1413) × *G. canescens* (G1232)	F$_1$	Sterile
RP15/3	2n = 2×	D3 (G1749) × D5B (G1941)	F$_1$	Sterile
RP20/3	2n = 2×	D2 (G1413) × D5B (G1941)	F$_1$	Sterile

and Sang [2001]) between accessions of natural tetraploid taxa to give a fertile tetraploid hybrid (table 2). The third type of cross involved diploid accessions from different taxa to produce three sterile diploid hybrids (table 2).

All the plants evaluated for rRNA expression were grown from seed in greenhouses at Cornell University at 12 hours of daylight, except for sterile diploid hybrids that were generated, grown from seeds, and validated at CSIRO (Canberra, Australia). Leaf samples from these latter plants were fixed in RNA*later* (Ambion Inc.) and shipped Cornell University for DNA and RNA extraction.

Amplification, Sequencing, and Quantification of nrDNA Repeat Ratios

Genomic DNA was isolated using the method of Doyle and Doyle (1987). DNA of most accessions used for this investigation had been extracted for previous studies (Doyle, Doyle, and Brown 1999; Brown et al. 2002; Doyle et al. 2002). However, DNA was re-extracted from several of these accessions to test the reproducibility of nrDNA homeolog ratios in different individuals of the same accession. RNA extractions were performed from living material harvested between 12:30 and 14:30 to limit potential variation in expression during the day. For RNA extractions made from leaf material, only mature, fully expanded leaves were used. The RNA extractions were performed with Trizol Reagent (Gibco), and the DNA was extracted from the phenol phase after the RNA extraction. Both extractions were conducted according to the manufacturer's recommendations. After RNA extraction, potential contaminant DNA was removed using the DNA-*free* kit (Ambion Inc.), treating 5 μl of the RNA extraction with 1 to 4 U of DNase in a total volume of 10 μl. This step was repeated twice, including a heat denaturation step (95°C for 3 min, followed by quenching on ice) between the treatments to remove possible RNA-DNA heteroduplexes.

Reverse transcription of RNA to DNA was conducted using 5 pMol of either the "universal" primer ITS-4 (White et al. 1990) or the "angiosperm-specific" primer AB101 (Sun et al. 1994) in 1X reverse transcriptase (RT) buffer (New England Biolabs), 0.8 μM dNTP, 0.2 mg/ml BSA,

1.2% RNA secure, and 0.3 U of reverse transcriptase in 25 μl reactions; the reaction was incubated at 37°C for 30 min. Amplifications by polymerase chain reaction (PCR) were performed on either DNA or RT product (RT-PCR) using the angiosperm-specific primer ITS-5ang (Vasquez 2001) as a forward primer and either the ITS-4 or the ITS-MR2 (Rauscher, Doyle, and Brown 2004) primer as a reverse primer. PCR was also conducted on the RNA extractions not treated with reverse transcriptase (RT$^-$) to ensure that there was no contaminant DNA left in the RNA extracts. PCR reactions were conducted using standard reaction and cycling conditions for *Glycine* ITS (Rauscher, Doyle, and Brown 2002). Each PCR reaction was split into three tubes for amplification and then recombined before purification of the PCR product (QIAquick column, QIAGEN). This was done to reduce PCR drift (Wagner et al. 1994) and the relative effect of potential polymerase induced mutations.

Sequencing was conducted with the appropriate reverse primer on an ABI 3700 (Applied Biosystems) at the Cornell BioResource Center. The ITS sequence of a polyploid accession was first compared with ITS sequences from its known diploid progenitors (figure 1 and Rauscher, Doyle and Brown [2004]) to categorize polymorphic sites. Relative peak heights for each polymorphic peak were then quantified using the EditView program version 1.0.1 (Applied Biosystems). The average relative peak heights at polymorphic sites have been shown to represent accurately the proportion of products in a mixture obtained by PCR amplification (Rauscher, Doyle, and Brown 2002). Because PCR reactions consist of an exponential geometric amplification, the nrDNA proportions at the end of the reaction should also indicate their proportions before the reaction. The proportions obtained should, therefore, indicate either the relative amounts of the homeologous nrDNA repeats in the genome in the case of DNA or their actual proportional transcription levels in the targeted tissues in the case of RNA because we are amplifying a transcribed region, albeit a spacer that is not present in the functional ribosome. However, because DNA quantification could be biased when the PCR reaction is outside the linear range of amplification (when the PCR signal is proportional to the input copy number), and

because this could sometimes be the case for one or both homeologous nrDNA repeats, PCR amplification should be considered semiquantitative.

The sensitivity threshold of the method has been estimated to be 5%, below which the minor product peak cannot be distinguished from background noise in electropherograms (Rauscher, Doyle, and Brown 2002). When one nrDNA copy could not be detected in electropherograms, indicating that it represented less than 5% of the final product, we tested for its presence using primers specific for the "missing" homeolog (Rauscher, Doyle, and Brown 2002, 2004). PCR reactions for each specific primer were as previously described (Rauscher, Doyle, and Brown 2002). Absence of amplification with the specific primers suggests that this minor repeat is absent from the genome or is not expressed (Rauscher, Doyle, and Brown 2002). When the PCR reaction resulted in amplification of the minor nrDNA or rRNA class, that template could represent up to 5% of the total PCR product. Individuals corresponding to this category were, thus, classified as having a ratio >95:<5. Because specific primers were designed only for the *G. tomentella* diploid progenitors, we did not test for the presence of the minor homeolog in *G. tabacina* polyploid accessions.

Reproducibility

Rauscher, Doyle, and Brown (2002) documented the reproducibility of these methods for assessing the ratio of homeologous nrDNA repeats in a polyploid genome, but no comparable data are available for rRNA expression levels. Here, a single plant of the synthetic polyploid hybrid A30 (table 2) was used to evaluate the method for quantifying the amounts of products because it possesses two principal nrDNA homeologs in its genome and because preliminary results showed that it expresses both (see *Results*). Six independent RNA extractions were performed simultaneously from a single trifoliolate leaf, two from the terminal and two from each lateral leaflet. rRNA homeolog ratios from each extraction were evaluated as described above to test the reproducibility of the quantification method.

Natural Polyploids

nrDNA proportions were evaluated for all natural polyploid accessions sampled (table 1). Expression in leaf tissues was evaluated for most natural polyploid accessions for which living material was available, except in two cases: (1) Individuals that were found to be fixed for one nrDNA homeolog in their genome ("100%" in table 3) were not tested for rRNA expression. (2) When several individuals of a single polyploid taxon were known to possess the same nrDNA repeat at levels higher than 95% and have the same maternal origin (J. T. Rauscher, J. J. Doyle, and A. H. D. Brown, unpublished data), typically only one individual was evaluated.

Variation in rRNA Expression Ratios Among Individuals of an Accession

For two natural allopolyploid accessions that possess both nrDNA homeologs in substantial amounts in their

Table 3
18S-5.8S-26S Ribosomal Homeolog Proportions in the Genome (nrDNA) and in Mature Leaf Transcriptome (rRNA) of Natural Allopolyploid *Glycine tomentella* and *G. tabacina*

Polyploid Taxon and Accession	Maternal Taxon[a]	nrDNA Percentage	rRNA Percentage
G. tomentella T1 (D1+D3)			
1392	D3	>95% D1	>95% D1
1133, 1136, 1361, 1367*, 1427	D3	>95% D3	>95% D3
1288*, 1763	D1	>95% D3	100% D3
G. tomentella T2 (D3+D4)			
1134, 1188*, 1811	D4	>95% D4	>95% D4
1286, 1393*, 1412, 1854	D3	>95% D4	>95% D4
G. tomentella T3 (D5A+D3)			
1397	D3	100% D5a	—
1359	D3	100% D3	—
1930	D3	>95% D3	>95% D3
2098*, 2100, 2539	D3	>95% D3	>95% D3
2099	D3	>95% D3	100% D3
G. tomentella T4 (D5B+D3)			
1747	D3	68% D3	62; 78% D5B
1469	D3	68% D3	>95% D3
2468	D5B	73% D5B	95% D5B
2469	D5B	>95% D5B	>95% D5B
2470	D5B	51% D5B	—
2476	D5B	80% D3	100% D5B
2557	D3	66% D3	>95% D3
2437	D5B	100% D5B	—
1304, 1348, 1350*	D5B	>95% D3	>95% D3
G. tomentella T5 (D1+*G. clandestina*)			
1739	Clan	80% Clan	>95% Clan
1487	Clan	63% Clan	>95% Clan
1969	Clan	74% Clan	>95% Clan
G. tomentella T6 (D1+D5B)			
1945	—	>95% D5B	100% D5B
G. tabacina (B′+B)			
2263*	B′	51% B′	>95% B′
1072*, 1080*, 1234*, 1988*	B′	69–77%	>95% B′
1075*, 1205*, 1255*	B′	>95% B′	>95% B′

NOTE.—When multiple accessions had the same nrDNA proportions and the same maternal origin (cpDNA), only individuals identified by an asterisk (*) were investigated for rRNA expression. Dashes indicate individuals that were not evaluated either because it was not relevant (e.g., major nrDNA = 100%) or because of lack of living material. Accessions are classified by taxon and according to their maternal diploid progenitor.

[a] Chloroplast donor data for *G. tomentella* and *G. tabacina* are from Rauscher, Doyle, and Brown (unpublished data) and (Doyle et al. 1990*a*), respectively.

genomes (G2468, G1969), four sibling plants were evaluated for variation in rRNA proportions. This experiment was conducted to investigate the potential for variation in rRNA expression in the progeny of a single plant.

Artificial Hybrids

One plant per accession was evaluated for the diploid hybrids and for artificial polyploids, but 29 F₂ individuals were evaluated for the homoploid tetraploid cross. This is because although diploid hybrids are sterile and artificial

allopolyploids are fixed heterozygotes, homoploid tetraploid hybrids are expected to segregate for any genetic differences between the two tetraploids involved in the cross. All artificial hybrids were evaluated for both nrDNA and nrRNA, with the exception of the homoploid tetraploid cross for which 15 out of 29 individuals were evaluated only for their genomic nrDNA content.

Tissue Specificity

For some natural allopolyploid accessions that were found to possess readily detectable amounts of both nrDNA homeologs in their genomes, rRNA expression was investigated in tissues other than mature leaves. RNA was extracted from cotyledons and hypocotyls of several accessions of both *G. tomentella* and *G. tabacina* (table 1). RNA from both tissues was extracted when the seedling expanded out from the seed coat, about 2 to 3 days after imbibition. For one individual (G1739), rRNA proportions from 1 cm root tips and young leaves (unexpanded, average 1 cm length) were also evaluated.

Results

Reproducibility of the Semiquantitative PCR Method

rRNA proportions were estimated from six independent extractions of the same trifoliolate leaf for artificial polyploid accession A30. The D3 rRNA component in each isolation comprised between 81% and 85% (mean = 83.2, SD = 1.5), indicating that the method is reproducible. The much higher expression of the D3 homeolog, some six times more abundant than the D4 component, was in contrast to the nearly equal nrDNA proportion (46% D3) estimated from this same plant.

Natural Allopolyploids

Natural polyploids fell into two main groups based on their nrDNA homeolog ratio. One class was found to possess both nrDNA homeologs in their genomes in amounts detectable by direct sequencing of PCR products (i.e., the minor nrDNA type represented more than 5% of the total product); three polyploid taxa (*G. tomentella* T4 and T5 and *G. tabacina*) possess individuals of this type (table 3). In the other class, polyploids were nearly or completely fixed for one nrDNA homeolog, with the principal nrDNA type representing more than 95% of the product and the minor repeat either undetectable or only detectable using homeolog-specific primers (table 3). Most polyploids investigated were of this second type, and such individuals were found in all polyploid taxa but one (*G. tomentella* T5), with some taxa containing only such individuals (*G. tomentella* T1, T2, T3, and T6). However, the same copy did not always predominate in all individuals within a taxon (Rauscher, Doyle, and Brown 2004), and there was no relationship between chloroplast donor (J. T. Rauscher, J. J. Doyle, and A. H. D. Brown, unpublished data) and the nrDNA copy that predominated in the genome of these taxa (table 3).

At the expression level, all accessions that were nearly fixed for one nrDNA type (>95%) also expressed the same homeolog in mature leaves at proportions greater than 95% (table 3). In most cases in which the minor repeat at the genomic level could be amplified with homeolog-specific primers, transcripts of that repeat were also detected using the same primers. However, in four accessions, including three of this type, minor transcripts were undetectable (table 3).

All natural allopolyploids that possessed substantial amounts of both nrDNA repeats in their genomes showed nearly complete preferential expression of one homeologous repeat type except for one individual (*G. tomentella* T4 G1747) for which transcripts of both homeologous loci were detected by direct sequencing of RT-PCR products (table 3). Even for this individual, however, one locus was expressed at a level greater than its estimated proportion in the genome. In two individuals surveyed, the minor repeat was preferentially expressed (G1747 and G2476; both T4), whereas in all other cases, expression was biased toward the major repeat class.

The same homeolog was more abundant and preferentially expressed in all accessions surveyed from three of the six polyploid taxa for which multiple accessions were available. In *G. tomentella* T5, it was the *clandestina* homeolog and not the D1 copy, and in *G. tomentella* T2, it was the D4 homeolog rather than the D3 (table 3). In *G. tabacina* polyploids, the B' (*G. stenophita*) copy was always the favored repeat (table 3). In contrast, the remaining four polyploid taxa were polymorphic for the abundance and expression of their nrDNA loci (T1, T2, T3, and T4 [table 3]). For example, accessions of *G. tomentella* T4 varied as to which homeolog was more abundant, and the preferentially expressed homeolog was either the D5B type or the D3 type (table 3).

The *G. tomentella* T4 taxon is hypothesized to have arisen recurrently (Doyle et al. 2002). Individuals G2468, G2469, and G2476 show the same multilocus genotype when data from the histone H3-D gene (Doyle et al. 2002), ITS sequences (Rauscher, Doyle, and Brown 2004), and cpDNA data (J. T. Rauscher, J. J. Doyle, and A. H. D. Brown, unpublished data) are considered and are, thus, thought to represent a single origin of this taxon. Genomic proportions of nrDNA vary among these individuals, with the D5B repeat being almost fixed in accession G2469, predominant in G2468 (73%), and in the minority in G2476 (20%). Nevertheless, all these individuals showed preferential expression of the D5B homeolog, irrespective of their genomic composition (table 3).

Variation of rRNA Expression in Individual Progeny

Variation in transcribed rRNA proportions between sibling plants was tested in four plants of two *G. tomentella* accessions (G2468 [T4] and G1969 [T5], respectively). For G2468, the D5B component of rRNA was greater than or equal to 95% for all four plants investigated, distinctly biased from its mean value for nrDNA of 75% D5B (range, 73% to 76%). The result for the four G1969 plants was the same numerically. The major nrDNA component was the *clandestina* homeolog at about 75%, whereas this homeolog comprised more than 95% of the transcribed rRNA in all four plants.

Artificial Crosses
Artificial Polyploids

Four artificial allotetraploid plants were available that mimicked the *G. tomentella* T5 taxon in combining a *G. tomentella* D1 genome with an A-genome (in one case, *G. clandestina*, and in the others, two different accessions of the closely related *G. canescens* [fig. 1]). One F_2 individual each from two of these artificial allopolyploids (A39/1 and A56/1) possessed very similar amounts of their two nrDNA homeologs (48% to 53% D1 [table 4]) and showed no evidence of preferential expression. The other two plants (A57/1 and A58/1) had somewhat more skewed nrDNA ratios (59% to 63% D1 [table 4]). In both of these plants, expression favored the minority (*G. canescens*) homeolog.

Fertile Homoploid Tetraploid Hybrids

Twenty-nine F_2 individuals from an artificial cross (H204) between *G. tomentella* T1 polyploid individuals that differed in the proportions of nrDNA homeologs (G1392: >95% D1; G1468: >95% D3 [table 3]) were studied for nrDNA content and a subset was also studied for expression. Segregating F_2 progeny from this T1 × T1 cross were assayed for nrDNA homeolog ratio in an effort to discriminate between deletion of a homeologous locus or concerted evolutionary conversion of one locus to the other in the polyploids involved in the cross. There are three hypotheses (fig. 2): (1) G1392 has lost most or all of its D3 locus, and G1468 has lost most or all of its D1 locus (dual locus deletion [fig. 2*B*]); (2) G1392 has two homeologous loci populated predominantly with D1 repeats caused by concerted evolution of its D3 locus in favor of D1 repeats, and G1468 has the opposite condition, also caused by concerted evolution (dual concerted evolution [fig. 2*C*]); or (3) one plant has lost a locus and the other plant shows concerted evolution (mixed model [fig. 2*D*]).

Figure 3 presents the segregation results as a histogram of D1 homeolog percentages in the 29 individuals scored. The two largest classes are the class of individuals with more than 95% D1 and that of individuals having approximately equal ratios of nrDNA homeologs. No individuals with from 60% to 95% D1 were observed. Grouping phenotypes into three classes, greater than 95% D1, a mixture of D1 and D3, and less than 5% D1, the expected F_2 ratios are 1:3:1 for the dual locus deletion hypothesis, 1:14:1 for the dual concerted evolution hypothesis, and 4:11:1 for the mixed model hypothesis (table 5 and fig. 2). The observed frequencies departed significantly from expectations under the dual locus deletion model ($P = 0.009$ [table 5]) and the dual concerted evolution model ($P = 0$). In contrast, the data do not reject the mixed model ($P = 0.11$). The fit is greatly improved if the assumption of equal numbers of repeats at the two homeologous loci is relaxed. For example, if the D1 locus, whether initially populated with D1 or D3 repeats, contains only 75% of the nrDNA copies as the D3 locus, then the expected ratio is 6:9:1, and the corresponding probability of the observed segregation is $P = 0.93$. The fit to the mixed model suggests that G1498 has only a single nrDNA locus

Table 4
18S-5.8S-26S Ribosomal RNA Homeolog Proportions in the Genome (nrDNA) and Transcriptome (rRNA) of Artificial Polyploids and of Homoploid Diploid Hybrids

Accession	Tetraploid Taxon Mimic	nrDNA Homeolog Proportion	rRNA Homeolog Proportion
Artificial polyploids			
A39/1	T5 (D1 + *G. clandestina*)	53% D1	56% *clandestina*
A56/1	T5 (D1 + *G. canescens*)	48% D1	55% *canescens*
A57/1	T5 (D1 + *G. canescens*)	59% D1	59% *canescens*
A58/1	T5 (D1 + *G. canescens*)	63% D1	78% *canescens*
Homoploid diploid hybrids			
RP8/1	T5 (D2 + *G. canescens*)	73% D2	66% *canescens*
RP15/3	T4 (D3 + D5B)	58% D5B	>95% D3
RP20/3	None known (D2 + D5B)	—	66% D5B

(D3), whereas G1392 has retained two loci, one of which has been largely converted from D3 to D1.

A subset of the F_2 individuals was tested for rRNA expression (fig. 3). All individuals possessing homeolog genomic ratios greater than 95% expressed the predominant copy preferentially. Of the five plants that possessed both nrDNA types in significant proportions, two showed greater than 95% expression of the D3 homeolog. The three remaining individuals expressed both homeologous types of rRNA, but showed strong preferential expression of the D3 copy (c. 85% [fig. 3]).

Sterile Interspecific Diploid Hybrids

Three different first generation (F_1) diploid hybrids were tested for nrDNA content and expression. Two showed additive nrDNA contents, and the third (RP20/3), for which DNA data were unavailable, presumably also is additive, given that it expressed both D2 and D5B homeologs (table 4). One hybrid, RP15/3 (D3 × D5B = T4-like), showed strong preferential expression of the D3 locus. In RP8/1, the D2 repeat was more prevalent in the genome, but the *G. canescens* repeat was responsible for 66% of the rRNA transcribed in this plant. A similar expression ratio was observed in RP20/3 (table 4).

rRNA Expression in Different Tissues

rRNA expression was evaluated in cotyledons, hypocotyls, young leaves, and roots of *G. tomentella* T4 and T5 and *G. tabacina* accessions that possessed substantial amounts of both homeologous nrDNA loci in their genomes. All showed preferential expression of one homeolog in their mature leaves (tables 3 and 7). T5 accession G1739, which was investigated for rRNA expression in its roots and in developing leaves, showed no tissue-specific variation in expression, always expressing preferentially the *G. clandestina* nrDNA homeolog (table 6). The rRNA proportions of the cotyledons and the hypocotyls of natural *G. tomentella* accessions were sometimes similar to those found in the mature leaves, but some accessions showed a much more balanced ratio of homeolog expression in both of these seedling tissues (table 6). *Glycine tabacina* accessions generally showed strong preferential expression of the B′ homeolog in all
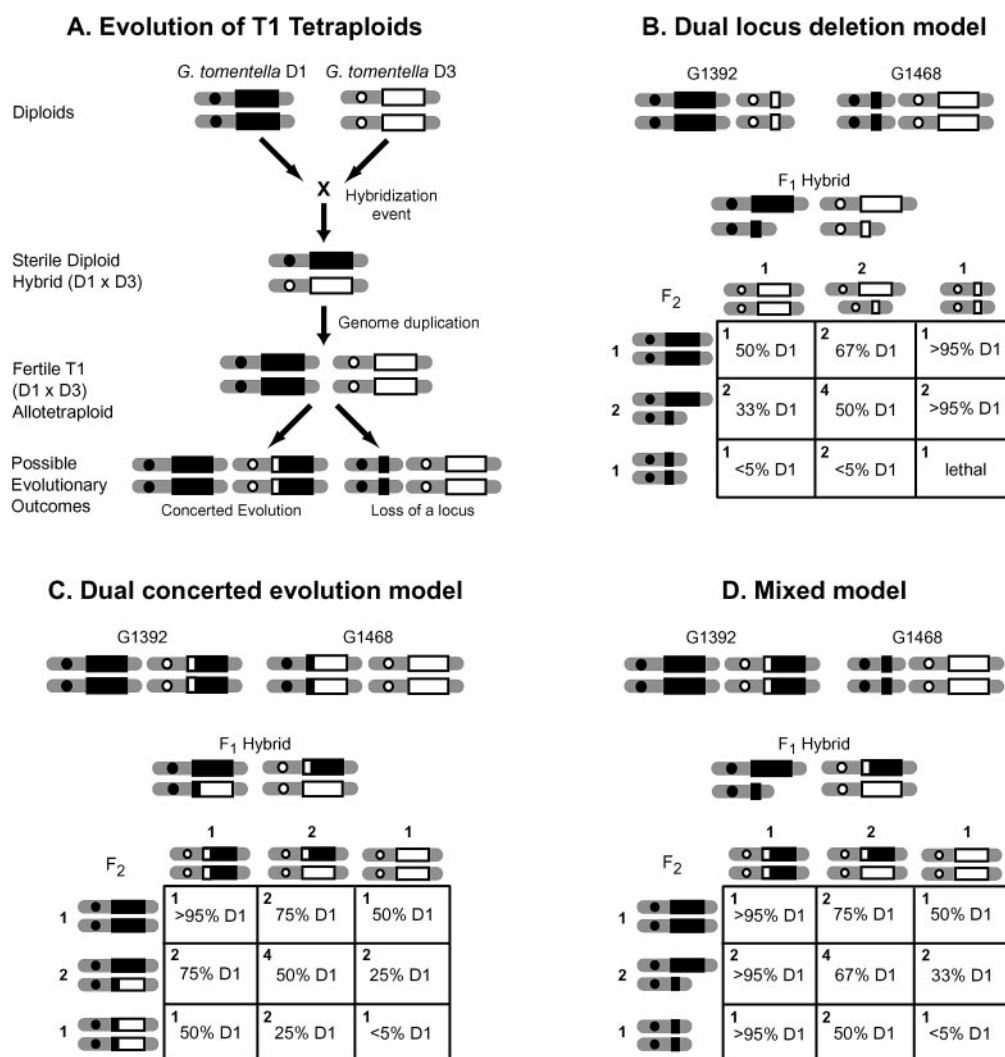
FIG. 2.—Possible models of evolution of homeologous nrDNA loci in two *G. tomentella* T1 accessions (G1392 and G1468) to explain the segregation pattern observed in second generation T1 × T1 hybrids (fig. 3). nrDNA loci are represented by rectangles and centromeres are represented by circles. D1 and D3 features are in black and white, respectively. (*A*) Standard evolution of a T1 allotetraploid from diploid *G. tomentella* D1 and D3 and possible evolutionary outcomes of the nrDNA loci: concerted evolution or loss of a locus, both of which may or may not be complete (shown as incomplete in the figure). Three different models are shown: a dual locus deletion model (*B*), a dual concerted evolution model (*C*), and a mixed model (*D*). For each model, the constitution of the natural tetraploids and F_1 hybrids is shown, as is the segregation pattern in second-generation progeny. The expected frequency of each segregation class is indicated in the upper left corner of each cell in the segregation tables.

tissues, but in three accessions, there was appreciable expression of the other (B) homeolog in either cotyledon or hypocotyl, although never in both (table 6).

Discussion
Evolution of nrDNA Loci in *Glycine* Allopolyploids

For any given nuclear gene, a recently formed allotetraploid is expected to possess two homeologous loci, one from each of its diploid progenitors. Our results from artificial diploid F_1 hybrids and artificial polyploids (tables 4 and 6) suggest that *Glycine* allopolyploids had roughly equal numbers of homeologous nrDNA copies at their formation. In contrast, we find that most natural *Glycine* allopolyploids possess one predominant nrDNA homeolog comprising more than 95% but less than 100% of the total nrDNA.

Two different processes could result in this large imbalance. First, copy number increase at one locus or decrease at the other, with no interchange between the two loci, could be responsible. Alternatively, interhomeologous locus concerted evolution could lead to the replacement of repeats at one locus with repeats from the other (Wendel, Schnabel, and Seelanan 1995). One source of relevant information for discriminating between these alternatives is the number and location of nrDNA loci in the genomes of polyploids and their diploid progenitors, such as is provided by fluorescent in situ hybridization (FISH).

In *Glycine*, FISH studies indicate that most diploids possess a single major nrDNA locus per haploid genome and that most polyploids possess two putatively homeologous loci (Krishnan et al. 2001; Singh, Kim, and Hymowitz 2001). Examples include accessions of *G. tabacina* and *G. tomentella* T2 and T3. Both of the
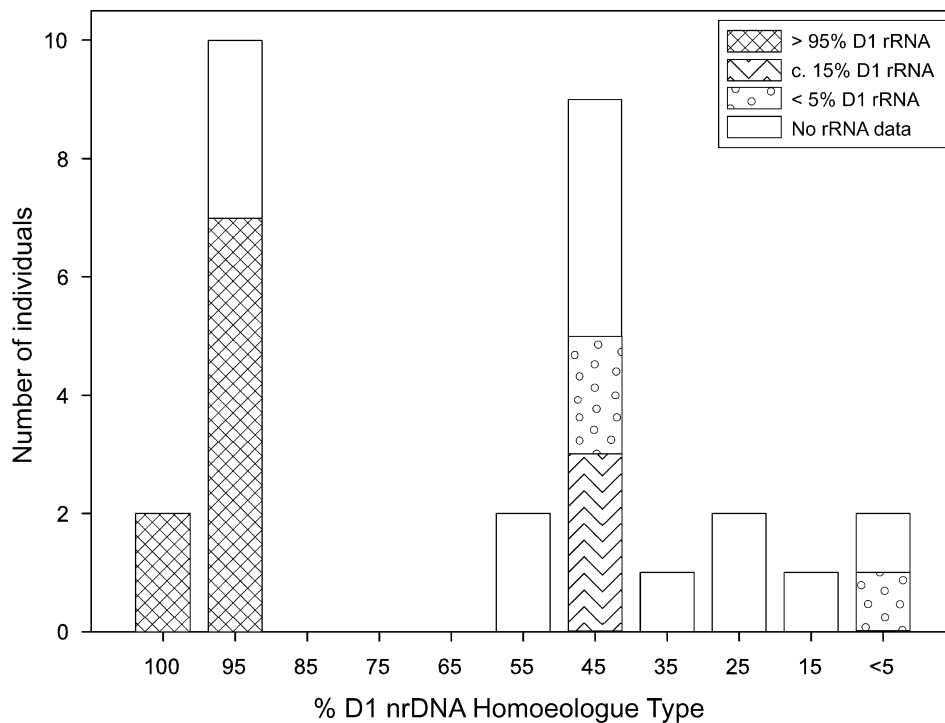
Fig. 3.—nrDNA proportions in the progeny of the T1×T1 homoploid tetraploid hybrid (H204) and their rRNA expression, classified according to the proportion of their D1 nrDNA homeolog.

G. tomentella accessions used in FISH studies (T2: G1188 = USDA PI 441005; T3: G1359 = PI 446988) also were included in our study, and both were found by us to possess only one nrDNA type exclusively (G1359) or nearly so (G1188). The observation of two loci but only one repeat type supports the hypothesis of interlocus concerted evolution for these two polyploid accessions.

The F_2-segregation data from a cross between two T1 polyploids, each nearly fixed for alternative homeologous nrDNA loci, point to roles for both concerted evolution and copy number loss in different individuals of this taxon. FISH data are available only for a T1 accession not included in our study (G1978 = PI483218 [Singh, Kim, and Hymowitz 2001]), and this plant had only a single

locus. Therefore, our hypothesis that G1468 has lost its D1 locus is consistent with FISH data for at least one T1 polyploid. In contrast, we infer that G1392 retains two loci but that both have D1 repeats caused by concerted evolutionary conversion of the D3 locus. Thus, it cannot be assumed that the presence of only a single homeologous repeat in an allopolyploid is always caused by concerted evolution. Moreover, homogenization of nrDNA can occur by more than one mechanism, even within a single species.

In *Glycine* polyploids, the direction of fixation was not determined by maternal progenitor. In the *G. tomentella* T1 taxon, accessions with *G. tomentella* D3 as a maternal parent were fixed for either the D1 or the D3 homeolog, whereas in *G. tomentella* T2, polyploids with

Table 5
Observed Frequencies, Expected Proportions, and Expected Frequencies of Each Homeolog Proportion Class Given the Model, Along with the Chi-Square Value and Probability of Each Model Given the Data

	Homoeolog Proportion Class				
	D1 ≥ 95%	95% > D1 > 5%	5% ≥ D1	Chi-Square	Probability (*P*)
Observed frequencies	12	15	2		
Expected frequencies					
Model 1: Dual locus deletions (1:3:1)	5.8	17.4	5.8	9.4	0.009
Model 2: Dual locus concerted evolution (1:14:1)	1.8	25.4	1.8	61.5	0
Model 3: Mixed model (4:11:1)	7.3	19.9	1.8	4.4	0.11

NOTE.—Expected frequencies are shown in parentheses. Model is as shown in figure 2.

Table 6
rRNA Expression in Mature Leaves, Cotyledons, Hypocotyls, Young Leaves, and Roots of *G. Tomentella* and *G. Tabacina* Allopolyploid Accessions That Possess Both nrDNA Homeologs in Substantial Amounts

Taxon	Accession	nrDNA Proportions	Tissue (s)	rRNA Proportions
G. tomentella T4	2468	73% D5B : 27% D3	leaf	95% D5B
			coty, hypo	77% D5B
G. tomentella T4	2476	80% D3 : 20% D5B	leaf	100% D5B
			coty, hypo	70%–74% D5B
G. tomentella T5	1739	80% clan : 20% D1	leaf, young leaf, roots	>95% clan
G. tomentella T5	1969	74% clan : 26% D1	leaf, coty, hypo	>95% clan
G. tabacina	1072	74% B′ : 26% B	leaf, hypo	>95% B′
			coty	83% B′
G. tabacina	1080	73% B′ : 27% B	leaf, coty	>95% B′
			hypo	89% B′
G. tabacina	1234	71% B′ : 29% B	leaf, coty	>95% B′
			hypo	87% B′
G. tabacina	1254	77% B′ : 23% B	leaf, coty, hypo	>95% B′
G. tabacina	1988	69% B′ : 31% B	leaf, coty	>95% B′
G. tabacina	2263	51% B′ : 49% B	leaf, coty, hypo	>95% B′

NOTE.—Leaf indicates mature leaves, coty indicates cotyledons, and hypo indicates hypocotyls.

different maternal origins (D3 or D4) were all fixed for the D4 repeat (table 3). Cytoplasmic effects are thought to be important in plant evolution (Levin 2002, 2003), and work with synthetic polyploids in *Brassica* has suggested that there is evolution toward the maternal progenitor in some crosses (Song et al. 1995). That such trends may not be universal, however, is shown by our findings as well as by studies in wheat that failed to find an evolutionary bias toward the genome of the cytoplasmic donor (Liu, Vega, and Feldman 1998; Liu et al. 1998).

Although most *Glycine* polyploids investigated possess a preponderant nrDNA repeat in their genomes, there are some natural polyploid accessions of *G. tabacina* and the *G. tomentella* T4 and T5 taxa that still possess both parental nrDNA homeologs in substantial amounts. It is difficult to explain why some polyploids should retain both nrDNA copies and yet others become fixed (or almost so) for one homeolog. Because synthetic *Glycine* polyploids have roughly equal numbers of homeologous repeats, imbalance in homeolog copy number must occur over time, perhaps gradually. If so, then it is possible that polyploids that still possess both nrDNA homeologs have formed more recently than those with one predominant homeolog. However, there is no correlation between imbalance of repeat ratio and the age of a given polyploid, as estimated from the divergence between nrDNA ITS or histone H3-D alleles of polyploids and their diploid progenitors (data not shown). Similarly, there is no correlation between the genetic distance between homeologous genomes and the degree to which repeat ratios are skewed. Of the three taxa in which at least some individuals had relatively balanced homeolog ratios, one (*G. tomentella* T5) is the product of a wide hybridization, whereas the other two (*G. tomentella* T4 and *G. tabacina*) were formed from crosses involving relatively closely related genomes (see figure 1).

rRNA Expression in *Glycine* Allopolyploids
General Pattern of Expression

Preferential expression of one nrDNA homeolog was observed in all but one natural *Glycine* polyploid, but in most cases, dominance was incomplete. The predominant genomic repeat was expressed preferentially in all but two natural polyploids and in some artificial hybrids. In *G. tomentella* T5 and in *G. tabacina*, all individuals sampled showed consistent preferential expression of the same homeolog, suggesting a dominance hierarchy (e.g., *clandestina* dominant to D1 in T5 plants), as has been reported in other genera (Keep 1962; Wallace and Langridge 1971). However, the existence of an intrinsic dominance hierarchy is not supported for the T5 polyploid, because synthetic hybrids and polyploids that mimic its genomic composition express both homeologs.

Variation of Expression Between Accessions

Within some *Glycine* allopolyploid taxa, individuals differ in which homeolog is expressed preferentially. This parallels observations in other taxa (Keep 1962; Neves et al. 1977; Flavell and O'Dell 1979; Chen, Comai, and Pikaard 1998; Pontes et al. 2003). In *Glycine*, however, unlike in some of these cases (e.g., Keep et al. [1962]), preferential expression is not strictly correlated with gene dosage. In some cases, the less abundant nrDNA copy in the genome is preferentially expressed. Pontes et al. (2003) suggested that the specific genotypes involved in the original cross might be responsible for the observed variation in the strength of nucleolar dominance between individuals of allopolyploid *Arabidopsis suecica*. The variation in preferential expression among individuals within *Glycine* allopolyploid taxa suggests that the direction of preferential expression could be either mostly stochastic or genotype specific.

Incomplete Preferential Expression

In *Glycine*, dominance of one homeologous nrDNA locus is rarely complete, and expression is frequently observed from the other locus. Similar findings were described by Lewis and Pikaard (2001) for *Arabidopsis suecica*, and are consistent with cytological observations in *Ribes* (Keep 1962). In triticale (wheat × rye), the wheat homeolog is preferentially expressed, but a small proportion

of cells express the rye homeolog as well (Neves et al. 1997). Thus, low levels of expression of the minor homeolog in *Glycine* individuals could be caused either by low levels of expression in all cells or by higher levels of expression in a minority of cells, as in triticale.

Variation of Expression Between Tissues

Variation in expression was sometimes detected between tissues of a single plant. Such variation of transcribed nrDNA homeolog proportions suggests that the mechanisms responsible for the maintenance of preferential expression and nucleolar dominance can be relaxed in some tissues. Decreased dominance in hypocotyl and cotyledons was not observed in all *Glycine* polyploids tested, however, again suggesting stochastic or genotype-specific effects. It is possible that in *Glycine*, there is also a species-specific effect. Both accessions from the *G. tomentella* T4 taxon that were tested for expression in tissues other than mature leaves lacked nucleolar dominance in cotyledons and hypocotyls, whereas both individuals of the *G. tomentella* T5 taxon did not show any variation in expression among tissues. It appears, however, that genotype-specific effects dominate at least in some *Glycine* species, as illustrated by *G. tabacina* polyploids (table 6). Variation in patterns of ribosomal gene expression between tissues is also known in *Brassica*, in which nucleolar dominance was absent from root meristems (Hasterok and Maluszynska 2000) and from floral tissues (Chen and Pikaard 1997*a*), although present in leaves.

Establishment of Nucleolar Dominance in *Glycine*

Nucleolar dominance in natural allopolyploids could be caused by the hybridization event itself, by factors associated with polyploidization, or by factors occurring later in the evolution of the polyploid lineages. For the *G. tomentella* T5 taxon, information was available from artificial crosses about the relative effects of hybridization, allopolyploidy, and subsequent evolution. A diploid hybrid that mimics the presumed origin of the T5 taxon showed considerable expression of both homeologs (RP8/1 [table 4]), suggesting that hybridization alone is not sufficient to induce nucleolar dominance. That this is not necessarily the case in all *Glycine* polyploids, however, is illustrated by a synthetic hybrid that mimics the T4 taxon, in which strong preferential expression is observed (RP15/3).

Like the diploid hybrid, resynthesized T5 allopolyploids lacked nucleolar dominance, indicating that polyploidization of a hybrid also was not sufficient to induce preferential expression of the *G. clandestina* homeolog. In contrast, in all natural polyploids of the *G. tomentella* T5 taxon, the *G. clandestina* homeolog represents more than 95% of the rRNA amplified by RT-PCR. Thus, for this *Glycine* species, nucleolar dominance does not seem to be induced by the hybridization event nor by the polyploid event, but presumably only develops later in the evolution of the polyploid lineage. Preferential expression could evolve rapidly, in the early generations after the polyploid event, or more gradually over time. A lag time for establishment of nucleolar dominance was suggested by studies in *Arabidopsis* (Chen, Comai, and Pikaard 1998),

but there it lasted only a single generation, with nucleolar dominance being present in all F_2 individuals. In *Glycine tomentella* T5, more than two generations seem to be necessary for preferential expression to become established. Thus, in this *Glycine* polyploid, there is no evidence for the kind of epigenetic events known to occur in some early-generation polyploids (Comai et al. 2000; Osborn et al. 2003) and which in other species are thought to be responsible for nucleolar dominance (Pikaard and Chen 1998; Comai 2000; Pikaard 2001). The combination of diverged regulatory networks in allopolyploids has been suggested to be responsible for altering gene expression in the absense of epigenetic interactions or major genomic rearrangements (Riddle and Birchler 2003). Such mechanisms would be expected to vary among genotypes and could explain some of the variation observed in *Glycine*.

Several models exist to explain the establishment of nucleolar dominance (reviewed in Pikaard and Chen [1998]). Models that rely on intrinsic structural or sequence differences between homeologous repeats have been rejected on several grounds (Chen and Pikaard 1997*a*; Neves et al. 1997; Chen, Comai, and Pikaard 1998; Pikaard and Chen 1998; Frieman et al. 1999). The finding of variation in the direction of preferential expression in several *Glycine* polyploids with similar genome combinations and variation in the strength of nucleolar dominance in *Arabidopsis suecica* (Pontes et al. 2003) also suggest that intrinsic differences in the rRNA loci are not likely to be responsible for differential expression. Mechanisms responsible for the establishment of nucleolar dominance in *Glycine* appear to be complex. Regulation of gene expression may involve independent chromosomal loci (see Durica and Krider [1977, 1978]) or may be influenced by the chromosomal context of the nrDNA loci (see Schubert and Künzel [1990]). This is likely to be a factor in allopolyploids, given that chromosomal changes often occur rapidly after their origin (Song et al. 1995; Liu et al. 1998; Ozkan, Levy, and Feldman 2001; Shaked et al. 2001).

Literature Cited

Adams K. L., R. Cronn, R. Percifield, and J. F. Wendel. 2003. Genes duplicated by polyploidy show unequal contributions to the transcriptome and organ-specific reciprocal silencing. Proc. Natl. Acad. Sci. USA **100**:4649–4654.

Baldwin, B. G., M. J. Sanderson, J. M. Porter, M. F. Wojciechowski, C. S. Campbell, and M. J. Donoghue. 1995. The ITS region of nuclear ribosomal DNA: a valuable source of evidence on angiosperm phylogeny. Ann. MO Bot. Gard. **82**:247–277.

Brown, A. H. D., J. J. Doyle, J. P. Grace, and J. L. Doyle. 2002. Molecular phylogenetic relationships within and among diploid races of *Glycine tomentella* (Leguminosae). Aust. Syst. Bot. **15**:37–47.

Caudy, A. A., and C. S. Pikaard. 2002. *Xenopus* ribosomal RNA gene intergenic spacer elements conferring transcriptional enhancement and nucleolar dominance-like competition in oocytes. J. Biol. Chem. **277**:31577–31584.

Chen, Z. J., L. Comai, and C. S. Pikaard. 1998. Gene dosage and stochastic effects determine the severity and direction of uniparental ribosomal RNA gene silencing (nucleolar dominance) in *Arabidopsis* allopolyploids. Proc. Natl. Acad. Sci. USA **95**:14891–14896.

Chen, Z. J., and C. S. Pikaard. 1997*a*. Transcriptional analysis of nucleolar dominance in polyploid plants: biased expression/silencing of progenitor rRNA genes is developmentally regulated in *Brassica*. Proc. Natl. Acad. Sci. USA **94**:3442–3447.

———. 1997*b*. Epigenetic silencing of RNA polymerase I transcription: a role for DNA methylation and histone modification in nucleolar dominance. Genes Dev. **11**:2124–2136.

Comai, L. 2000. Genetic and epigenetic interactions in allopolyploid plants. Plant Mol. Biol. **43**:387–399.

Comai, L., A. P. Tyagi, K. Winter, R. Holmes-Davis, S. H. Reynolds, Y. Stevens, and B. Byers. 2000. Phenotypic instability and rapid gene silencing in newly formed *Arabidopsis* allotetraploids. Plant Cell **12**:1551–1567.

Crosby, M. R. 1980. Polyploidy in bryophytes with special emphasis on mosses. Pp. 193–198 *in* W. H. Lewis, ed. Polyploidy: biological relevance. Plenum Press, New York.

Doyle, J. J., and J. L. Doyle. 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochem. Bull. **19**:11–15.

Doyle, J. J., J. L. Doyle, and A. H. D. Brown. 1999. Origins, colonization, and lineage recombination in a widespread perennial soybean polyploid complex. Proc. Natl. Acad. Sci. USA **96**:10741–10745.

Doyle, J. J., J. L. Doyle, A. H. D. Brown, and R. G. Palmer. 2002. Genomes, multiple origins, and lineage recombination in the *Glycine tomentella* (Leguminosae) polyploid complex: histone H3-D gene sequences. Evolution **56**:1388–1402.

Doyle, J. J., J. L. Doyle, A. H. D. Brown, and J. P. Grace. 1990*a*. Multiple origins of polyploids in the *Glycine tabacina* complex inferred from chloropast DNA polymorphism. Proc. Natl. Acad. Sci. USA **87**:714–717.

Doyle, J. J., J. L. Doyle, J. P. Grace, and A. H. D. Brown. 1990*b*. Reproductively isolated polyploid races of *Glycine tabacina* (Leguminosae) had different chloroplast genome donors. Syst. Bot. **15**:173–181.

Doyle, J. J., J. L. Doyle, J. T. Rauscher, and A. H. D. Brown. 2004. Evolution of the perennial soybean polyploid complex (*Glycine* subgenus *Glycine*): a study of contrasts. Biol. J. Linn. Soc.(in press).

Doyle, M. J., J. E. Grant, and A. H. D. Brown. 1986. Reproductive isolation between isozyme groups of *Glycine tomentella* (Leguminosae) and spontaneous doubling in their hybrids. Austral. J. Bot. **34**:523–535.

Durica, D. S., and H. M. Krider. 1977. Studies on the ribosomal RNA cistrons in interspecific *Drosophila* hybrids. I. Nucleolar dominance. Dev. Biol. **59**:62–74.

———. 1978. Studies on the ribosomal RNA cistron in interspecific *Drosophila* hybrids. II. Heterochromatic regions mediating nucleolar dominance. Genetics **89**:37–64.

Ehrendorfer, F. 1980. Polyploidy and distribution. Pp. 45–60 *in* W. H. Lewis, ed. Polyploidy: biological relevance. Plenum Press, New York.

Favarger, C. 1967. Cytologie et distribution des plantes. Biol. Rev. **42**:163–206.

Ferguson, D., and T. Sang. 2001. Speciation through homoploid hybridization between allotetraploids in peonies (*Paeonia*). Proc. Natl. Acad. Sci. USA **98**:3915–3919.

Flavell, R. B. 1986. The structure and control of expression of ribosomal RNA genes. Oxf. Surv. Plant Mol. Cell Biol. **3**:252–274.

Flavell, R. B., and M. O'Dell. 1979. The genetic comtrol of nucleolus in wheat. Chromosoma **71**:135–152.

Frieman, M., Z. J. Chen, J. Saez-Vasquez, A. Shen, and C. S. Pikaard. 1999. RNA polymerase I transcription in a *Brassica* interspecific hybrid and its progenitors: tests of transcription factor involvement in nucleolar dominance. Genetics **152**:451–460.

Galitski, T., A. J. Saldanha, C. A. Styles, E. S. Lander, and G. R. Fink. 1999. Ploidy regulation of gene expression. Science **285**:251–254.

Grant, J. E., J. P. Grace, A. H. D. Brown, and E. Putievsky. 1984. Interspecies hybridization in *Glycine* Willd. Subgenus *Gylcine* (Leguminosau). Aust. J. Bot. **32**:655–663.

Grant, V. 1981. Plant speciation. 2nd edition. Columbia University Press, New York.

Hamby, R. K., and E. A. Zimmer. 1992. Ribosomal RNA as a phylogenetic tool in plant systematics. Pp. 50–91 *in* D. E. Soltis, P. S. Soltis, and J. J. Doyle, eds. Molecular systematics of plants. Chapman and Hall, New York.

Hasterok, R., and J. Maluszynska. 2000. Nucleolar dominance does not occur in root tip cells of allotetraploid *Brassica* species. Genome **43**:574–579.

Honjo, T., and R. H. Reeder. 1973. Preferential transcription of *Xenopus laevis* ribosomal RNA in interspecies hybrids between *Xenopus laevis* and *Xenopus mulleri*. J. Mol. Biol. **80**:217–228.

Houchins, K., M. O'Dell, R. B. Flavell, and J. P. Gustafson. 1997. Cytosine methylation and nucleolar dominance in cereals hybrids. Mol. Gen. Genet. **255**:294–301.

Hymowitz, T., R. J. Singh, and K. P. Kollipara. 1998. The genomes of the *Glycine*. Pp. 289–317 *in* J. Janick, ed. Plant breeding reviews. John Wiley, New York.

Keep, E. 1962. Satellite and nucleolar number in hybrids between *Ribes nigrum* and *R. grossularia* and in their backcrosses. Can. J. Genet. Cytol. **4**:206–218.

Krishnan, P., V. T. Sapra, K. M. Soliman, and A. Zipf. 2001. FISH mapping of the 5S and 18S-28S rDNA loci in different species of *Glycine*. J. Hered. **92**:295–300.

Levin, D. A. 1983. Polyploidy and novelty in flowering plants. Am. Nat. **122**:1–25.

———. 2002. The role of chromosomal change in plant evolution. Oxford University Press, Oxford, UK.

———. 2003. The cytoplasmic factor in plant speciation. Syst. Bot. **28**:5–11.

Lewis, M. S., and C. S. Pikaard. 2001. Restricted chromosomal silencing in nucleolar dominance. Proc. Natl. Acad. Sci. USA **98**:14536–14540.

Lewis, W. H. 1980. Polyploidy in species populations. Pp. 103–144 *in* W. H. Lewis, ed. Polyploidy: biological relevance. Plenum Press, New York.

Liu, B., J. M. Vega, and M. Feldman. 1998. Rapid genomic changes in newly synthesized amphiploids of *Triticum* and *Aegilops*. II. Changes in low-copy coding DNA sequences. Genome **41**:535–542.

Liu, B., J. M. Vega, G. Segal, S. Abbo, M. Rodova, and M. Feldman. 1998. Rapid genomic changes in newly synthesized amphiploids of *Triticum* and *Aegilops*. I. Changes in low-copy noncoding DNA sequences. Genome **41**:272–277.

Neves, N., M. Silva, J. S. Heslop-Harrison, and W. Viegas. 1997. Nucleolar dominance in triticales: control by unlinked genes. Chromosome Res. **5**:125–131.

Nichols, H. W. 1980. Polyploidy in algae. Pp. 151–161 *in* W. H. Lewis, ed. Polyploidy: biological relevance. Plenum Press, New York.

Osborn, T. C., J. C. Pires, J. A. Birchler et al. (11 co-authors). 2003. Understanding mechanisms of novel gene expression in polyploids. Trends Genet. **19**:141–147.

Otto, S. P., and J. Whitton. 2000. Polyploid incidence and evolution. Annu. Rev. Genet. **34**:401–437.

Ozkan, H., A. A. Levy, and M. Feldman. 2001. Allopolyploidy-induced rapid genome evolution in the wheat (*Aegilops-triticum*) group. Plant Cell **13**:1735–1747.

Petit, C., and J. D. Thompson. 1999. Species diversity and ecological range in relation to ploidy level in the flora of the Pyrenees. Evol. Ecol. **13**:45–66.

Pikaard, C. S. 2000. Nucleolar dominance: uniparental gene silencing on a multi-megabase scale in genetic hybrids. Plant Mol. Biol. **43**:163–177.

———. 2001. Genomic change and gene silencing in polyploids. Trends Genet. **17**:675–677.

Pikaard, C. S., and Z. J. Chen. 1998. Nucleolar dominance. Pp. 277–294 *in* M. R. Paule, ed. Transcription of ribosomal RNA genes by eukaryotic RNA polymerase I. Springer-Verlag, Berlin.

Pikaard, C. S., and R. J. Lawrence. 2002. Uniting the path to gene silencing. Nat. Genet. **32**:1–2.

Pontes, O., R. J. Lawrence, N. Neves, M. Silva, J.-H. Lee, Z. J. Chen, W. Viegas, and C. S. Pikaard. 2003. Natural variation in nucleolar dominance reveals the relationship between nucleolus organizer chromatin topology and rRNA gene transcription in *Arabidopsis*. Proc. Natl. Acad. Sci. USA **100**:11418–11423.

Ramsey, J., and D. W. Schemske. 2002. Neopolyploidy in flowering plants. Annu. Rev. Ecol. Syst. **33**:589–639.

Rauscher, J. T., J. J. Doyle, and A. H. D. Brown. 2002. Internal transcribed spacer repeat-specific primers and the analysis of hybridization in the *Glycine tomentella* (Leguminosae) polyploid complex. Mol. Ecol. **11**:2691–2702.

———. 2004. Multiple origins and nrDNA ITS homoeologue evolution in the *Glycine tomentella* (Leguminosae) allopolyploid complex. Genetics **166.**

Reeder, R. H. 1985. Mechanisms of nucleolar dominance in animals and plants. J. Cell Biol. **101**:2013–2016.

Riddle, N. C., and J. A. Birchler. 2003. Effects of reunited diverged regulatory hierarchies in allopolyploids and species hybrids. Trends Ecol. Evol. **19**:597–600.

Rieseberg, L. H. 1997. Hybrid origins of plant species. Annu. Rev. Ecol. Syst. **28**:359–389.

Schranz, M. E., and T. C. Osborn. 2000. Novel flowering time variation in the resynthesized polyploid *Brassica napus*. J. Hered. **91**:242–246.

Schubert, I., and G. Künzel. 1990. Position-dependent NOR activity in barley. Chromosoma **99**:352–359.

Shaked, H., K. Kashkush, H. Ozkan, M. Feldman, and A. A. Levy. 2001. Sequence elimination and cytosine methylation are rapid and reproducible responses of the genome to wide hybridization and allopolyploidy in wheat. Plant Cell **13**:1749–1759.

Simillion, C., K. Vandepoele, M. C. E. Van Montagu, M. Zabeau, and Y. Van de Peer. 2003. The hidden duplication past of *Arabidopsis thaliana*. Proc. Natl. Acad. Sci. USA **99**:13627–13632.

Singh, R. J., H. H. Kim, and T. Hymowitz. 2001. Distribution of rDNA loci in the genus *Glycine*. Theor. Appl. Genet. **103**:212–218.

Soltis, D. E., and P. S. Soltis. 1999. Polyploidy: recurrent formation and genome evolution. Trends Ecol. Evol. **14**:348–352.

Soltis, P. S., and D. E. Soltis. 2000. The role of genetic and genomic attributes in the success of polyploids. Proc. Natl. Acad. Sci. USA **97**:7051–7057.

Song, K., P. Lu, K. Tang, and T. C. Osborn. 1995. Rapid genome change in synthetic polyploids of *Brassica* and its implications for polyploid evolution. Proc. Natl. Acad. Sci. USA **92**:7719–7723.

Stebbins, G. L. 1950. Variation and evolution in plants. Columbia University Press, New York.

Sun, Y., D. Z. Skinner, G. H. Liang, and S. H. Hulbert. 1994. Phylogenetic analysis of *Sorghum* and related taxa using internal transcribed spacers of nuclear ribosomal DNA. Theor. Appl. Genet. **89**:26–32.

Tal, M. 1980. Physiology of polyploids. Pp. 61–75 *in* W. H. Lewis, ed. Polyploidy: biological relevance. Plenum Press, New York.

Thompson, J. N., B. M. Cunningham, K. A. Segraves, D. M. Althoff, and D. Wagner. 1997. Plant polyploidy and insect/plant interactions. Am. Nat. **150**:730–743.

Vasquez, M. L. 2001. Molecular and morphological studies on Mexican red oaks (*Quercus* sect. Lobatae). Ph.D. Thesis, Cornell University, Ithaca, NY.

Wagner, A., N. Blackstone, P. Cartwright, M. Dick, B. Misof, P. Snow, G. P. Wagner, J. Bartels, M. Murtha, and J. Pendleton. 1994. Surveys of gene families using polymerase chain reaction: PCR selection and PCR drift. Syst. Biol. **43**:250–261.

Wagner, W. H., Jr., and F. S. Wagner. 1980. Polyploidy in pteridophytes. Pp. 199–214 *in* W. H. Lewis, ed. Polyploidy: biological relevance. Plenum Press, New York.

Wallace, H., and W. H. R. Langridge. 1971. Differential amphiplasty and the control of ribosomal RNA synthesis. Heredity. **27**:1–13.

Wendel, J. F. 2000. Genome evolution in polyploids. Plant Mol. Biol. **42**:225–249.

Wendel, J. F., A. Schnabel, and T. Seelanan. 1995. Bidirectional interlocus concerted evolution following allopolyploid speciation in cotton (*Gossypium*). Proc. Natl. Acad. Sci. USA **92**:280–284.

White, T. J., T. Bruns, S. Lee, and J. Taylor. 1990. Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenies. Pp. 315–322 *in* M. A. Innis, D. H. Gelfand, J. J. Sninsky and T. J. White, eds. PCR protocols. Academic Press, San Diego, Calif.

Ziolkowski, P. A., G. Blanc, and J. Sadowski. 2003. Structural divergence of chromosomal segments that arose from successive duplication events in the *Arabidopsis* genome. Nucleic Acids Res. **31**:1339–1350.