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 homoelogous 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 homoelogous repeats in polyploid genomes and to test for preferential expression of homoelogous nrDNA loci. Most natural polyploids possess one predominant nrDNA homeolog in their genome. Analysis of F2 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. Throughout the majority of 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), phenomenology (Comai et al. 2000; Schranz and Osborn 2000; Ramsey and Schemske 2002), and physiology (Tai 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. 1998; 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...
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 1997a, 1997b; 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* (BB'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. 1990b; 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. 1990b; 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

---

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

**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.

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

\[
\begin{array}{|l|l|l|l|l|}
\hline
\text{Accession} & \text{Ploidy Level} & \text{Parentage (♀ × ♂)} & \text{Generation Studied} & \text{Fertility} \\
\hline
\text{Artificial polyploids} & & & & \\
A39/1 & 2n = 4× & D1 (G1316) × G. clandestina (G1253) & F2 & Fertile \\
A56/1 & 2n = 4× & D1 (G1316) × G. canescens (G1351) & F2 & Fertile \\
A57/1 & 2n = 4× & D1 (G1316) × G. canescens (G1232) & F2 & Fertile \\
A58/1 & 2n = 4× & D1 (G1316) × G. canescens (G1232) & F2 & Fertile \\
A30 & 2n = 8× & G. tomentella T2 (G1188) × G. tomentella T1 (G1133) & F2 & Fertile \\
\hline
\text{Homoploid tetraploid hybrids} & & & & \\
H204 & 2n = 4× & T1 (G1468) × T1 (G1392) & F2 & Fertile \\
\hline
\text{Homoploid diploid hybrids} & & & & \\
RP8/1 & 2n = 2× & D2 (G1413) × G. canescens (G1232) & F1 & Sterile \\
RP15/3 & 2n = 2× & D3 (G1749) × D5B (G1941) & F1 & Sterile \\
RP20/3 & 2n = 2× & D2 (G1413) × D5B (G1941) & F1 & Sterile \\
\hline
\end{array}
\]
because this could sometimes be the case for one or both homeologous nrDNA repeats, PCR amplification should be considered semi-quantitative.

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 electrophorograms (Rauscher, Doyle, and Brown 2002). When one nrDNA copy could not be detected in electrophorograms, indicating that it represented less than 5% of the total PCR 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 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 final product, we tested for its presence using primers specific for the “missing” homeolog (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 final product. Individuals corresponding to this category were, thus, classified as having a ratio >95%:5%.

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 trifoliate 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 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 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

| 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<sup>a</sup> | 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*, 18111286, 1393*, 1412, 1854 | D4 | >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 (DSB=D3) | 1747 | D3 | 68% D3 | 62; 78% DSB |
| 1469 | D3 | 68% D3 | >95% D3 |
| 2468 | D5B | 73% DSB | 95% DSB |
| 2469 | D5B | >95% DSB | >95% DSB |
| 2470 | DSB | 51% DSB | — |
| 2476 | DSB | 80% D3 | 100% DSB |
| 2557 | D3 | 66% D3 | >95% D3 |
| 2437 | DSB | 100% DSB | — |
| 1304, 1348, 1350* DSB | >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. tabacina (B′+B) | 1945 | — | >95% DSB | 100% DSB |
| 2263* | B′ | 51% B′ | >95% B′ |
| 1072*, 1080*, 1234*, 1988* | B′ | 69–77% | >95% B′ |
| 1075*, 1205*, 1255* | B′ | >95% B′ | >95% B′ |

<sup>a</sup> Chloroplast donor data for G. tomentella and G. tabacina are from Rauscher, Doyle, and Brown (unpublished data) and (Doyle et al. 1990).
allopolyploids are fixed heterozygotes, homoploid tetra-
ploid 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 allopolyplploid 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 independ-
ent 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 Allopolyplploids

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 allopolyplploids that possessed substantial
amounts of both nrDNA repeats in their genomes showed
nearly complete preferential expression of one homeolo-
gous 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),
wheras in all other cases, expression was biased toward
the major repeat class.

The same homeolog was more abundant and prefer-
entially 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. stenopitha*) copy was
always the favored repeat (table 3). In contrast, the re-
mainig 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 prefer-
ential 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 homeo-
log comprised more than 95% of the transcribed rRNA in
all four plants.
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 F2 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 F2 individuals from an artificial cross (H204) between *G. tomentella* T1 tetraploid individuals (D) 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 F2 progeny from this T1 × T1 cross were assayed for nrDNA homeolog ratio in an effort to discriminate between deletion of a homoeologous 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. 2B]); (2) G1392 has two homoeologous 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. 2C]); or (3) one plant has lost a locus and the other plant shows concerted evolution (mixed model [fig. 2D]).

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 F2 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 homoeologous 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 (D3), whereas G1392 has retained two loci, one of which has been largely converted from D3 to D1.

A subset of the F2 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 (F1) 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 homoeologs (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

Table 4

<table>
<thead>
<tr>
<th>Accession</th>
<th>Tetraploid Taxon Mimic</th>
<th>nrDNA Homeolog Proportion</th>
<th>rRNA Homeolog Proportion</th>
</tr>
</thead>
<tbody>
<tr>
<td>A39/1</td>
<td>T5 (D1 + <em>G. clandestina</em>)</td>
<td>53% D1</td>
<td>56% clandestina</td>
</tr>
<tr>
<td>A56/1</td>
<td>T5 (D1 + <em>G. canescens</em>)</td>
<td>48% D1</td>
<td>55% canescens</td>
</tr>
<tr>
<td>A57/1</td>
<td>T5 (D1 + <em>G. canescens</em>)</td>
<td>59% D1</td>
<td>59% canescens</td>
</tr>
<tr>
<td>A58/1</td>
<td>T5 (D1 + <em>G. canescens</em>)</td>
<td>63% D1</td>
<td>78% canescens</td>
</tr>
<tr>
<td>RP8/1</td>
<td>T5 (D2 + <em>G. canescens</em>)</td>
<td>73% D2</td>
<td>66% canescens</td>
</tr>
<tr>
<td>RP15/3</td>
<td>T4 (D3 × D5B)</td>
<td>58% D5B</td>
<td>&gt;95% D3</td>
</tr>
<tr>
<td>RP20/3</td>
<td>None known (D2 + D5B)</td>
<td>—</td>
<td>66% D5B</td>
</tr>
</tbody>
</table>
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 F1 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
*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 F2-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**

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

**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

<table>
<thead>
<tr>
<th>Taxon</th>
<th>Accession</th>
<th>nrDNA Proportions</th>
<th>Tissue(s)</th>
<th>rRNA Proportions</th>
</tr>
</thead>
<tbody>
<tr>
<td>G. Tomentella T4</td>
<td>2468</td>
<td>73% D5B : 27% D3</td>
<td>leaf, coty, hypo</td>
<td>95% D5B</td>
</tr>
<tr>
<td>G. Tomentella T4</td>
<td>2476</td>
<td>80% D3 : 20% D5B</td>
<td>leaf, coty, hypo</td>
<td>73% D5B</td>
</tr>
<tr>
<td>G. Tomentella T5</td>
<td>1739</td>
<td>80% clan : 20% D1</td>
<td>leaf, young leaf, roots</td>
<td>100% D5B</td>
</tr>
<tr>
<td>G. Tomentella T5</td>
<td>1969</td>
<td>74% clan : 26% D1</td>
<td>leaf, coty, hypo</td>
<td>70%–74% D5B</td>
</tr>
<tr>
<td>G. Tabacina</td>
<td>1072</td>
<td>74% B' : 26% B</td>
<td>leaf, coty</td>
<td>&gt;95% clan</td>
</tr>
<tr>
<td>G. Tabacina</td>
<td>1080</td>
<td>73% B' : 27% B</td>
<td>leaf, coty, hypo</td>
<td>&gt;95% B'</td>
</tr>
<tr>
<td>G. Tabacina</td>
<td>1234</td>
<td>71% B' : 29% B</td>
<td>leaf, coty, hypo</td>
<td>&gt;95% B'</td>
</tr>
<tr>
<td>G. Tabacina</td>
<td>1254</td>
<td>77% B' : 23% B</td>
<td>leaf, coty, hypo</td>
<td>&gt;95% B'</td>
</tr>
<tr>
<td>G. Tabacina</td>
<td>1988</td>
<td>69% B' : 31% B</td>
<td>leaf, coty</td>
<td>&gt;95% B'</td>
</tr>
<tr>
<td>G. Tabacina</td>
<td>2263</td>
<td>51% B' : 49% B</td>
<td>leaf, coty, hypo</td>
<td>&gt;95% B'</td>
</tr>
</tbody>
</table>

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 *Riber* (Keep 1962). In triticale (wheat × rye), the wheat homeolog is prelatively 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 1997a), 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 F2 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 absence 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 1997a; 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 Kridel [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).

Acknowledgments

The authors thank Jane L. Doyle and Jim P. Grace for technical assistance, as well as A. Bruneau and two anonymous reviewers for their comments on a previous version of the manuscript. The work was supported by NSERC (Canada) and FQRNT (Québec) fellowships to S.J. and by a grant to J.J.D. and A.H.D.B. from the US National Science Foundation Systematic Biology Program (DEB-0089483).

Literature Cited


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


Brandon Gaut, Associate Editor

Accepted March 18, 2004