Organization of American and Asian Ginseng Germplasm Using Randomly Amplified Polymorphic DNA (RAPD) Markers

C.L. Boehm, H.C. Harrison, G. Jung, and J. Nienhuis
Department of Horticulture, University of Wisconsin, Madison, WI 53706

ABSTRACT. Genetic differences among eleven cultivated and eight wild-type populations of North American ginseng (Panax quinquefolium L.) and four cultivated populations of South Korean ginseng (P. ginseng C.A. Meyer) were estimated using RAPD markers. Cultivated P. ginseng population samples were collected from four regions of S. Korea. Cultivated P. quinquefolium population samples were collected from three regions in North America: Wisconsin, the Southeastern Appalachian region of the United States, and Canada. Wild-type P. quinquefolium was collected from three states in the United States: Pennsylvania, Tennessee, and Wisconsin. Evaluation of germplasm with 10 decamer primers resulted in 100 polymorphic bands. Genetic differences among populations indicate heterogeneity. The genetic distance among individuals was estimated using the ratio of discordant bands to total bands scored. Multidimensional scaling of the relationship matrix showed independent clusters corresponding to the distinction of species, geographical region, and wild versus cultivated types. The integrity of the clusters was confirmed using pooled chi-square tests for fragment homogeneity.

American ginseng (Panax quinquefolium) and Asian ginseng (P. ginseng) are perennial herbs of the Araliaceae family that are cultivated for their edible roots. Ginseng roots contain saponins that are thought to provide medicinal effects. Differences among ginseng roots in levels of saponins have been observed (Smith et al., 1996); however no information is currently available regarding the relative importance of genetic and environmental effects in the levels of saponins contained in roots.

American and Asian ginseng are commercially valuable crops with no registered cultivars or improved populations of known genetic or phenotypic distinction. Both species are self-pollinated allotetraploids (2n = 4x = 48) (Bai et al., 1997) and are propagated by seed from assumed unimproved land races.

Cultivated P. quinquefolium plants of the United States and Canada were first selected from local wild-type populations by farmers >100 years ago. Wild P. quinquefolium is indigenous throughout eastern North America from Northern Georgia to Canada and extending westward to the Mississippi river. Wild P. quinquefolium has been, and continues to be, collected throughout the Eastern United States (Duke, 1984). Genetic relationships among regional populations of the cultivated and wild P. quinquefolium are unknown.

Cultivation of P. ginseng in Korea began more than 300 years ago and now the wild P. ginseng of China and Korea is exceptionally rare (Duke, 1984). Like P. quinquefolium the genetic structure of P. ginseng populations are unknown. However, two studies have reported related information for P. quinquefolium. Blair (1975) found no differences in karyotype among two geographically separated, wild-type populations of North Carolina P. quinquefolium while Bai et al. (1997) found a cultivated population of P. quinquefolium from Ontario to be highly heterogeneous with respect to RAPD markers.

Molecular markers have been used in numerous crop species to estimate genetic relationships. Patterns of molecular marker polymorphism have been shown to be consistent with expectations based on pedigrees and breeding behavior in several crop species (dos Santos et al., 1994; Nienhuis et al., 1993; Tivang et al., 1996). RAPD markers have been shown to cost effectively produce polymorphic marker numbers sufficient for estimating genetic relationships (Tivang et al., 1996).

Plant geneticists found that RAPD data can be reliably generated and scored (Williams et al., 1990). The analysis of genetic relationships based on RAPD data is consistent with analysis based on other molecular marker data (dos Santos et al., 1994). Also, Skroch and Nienhuis (1995b) and Spooner et al. (1996) showed that the frequency of random errors (scoring error and experimental error) was low (<2%) among replicated DNA samples scored for RAPD markers. These findings justified the application of RAPD–PCR to a Panax sp. genome sample for the purpose of germplasm organization.

The objectives of this study were to 1) measure genetic distance within and among populations of cultivated P. ginseng and cultivated and wild-type P. quinquefolium populations and 2) contrast the observed genetic integrity of the two species, cultural types, and regional populations.

Materials and Methods

EXPERIMENTAL DESIGN. Representative samples of cultivated and wild-type P. quinquefolium of North America and cultivated P. ginseng C.A. Meyer from South Korea were obtained for DNA extraction and analysis of polymorphism. Leaf or root tissues were collected from 191 individual plants. The P. quinquefolium population was composed of two ecotype populations (cultivated
and wild type) and each ecotype population was composed of three regional populations (Table 1 and Fig. 1). DNA extraction and RAPD–PCR were used to produce and identify polymorphic RAPD bands for use in statistical analysis of genetic relationships among populations.

**Panax quinquefolium.** Local populations were researched and identified through public and personal references. Individual plants were randomly sampled in on-site visits to 14 population locales. Nine cultivated populations were randomly sampled by their field managers who then submitted the fresh or dry root tissues for analysis. The 23 groups were distributed across 7 U.S. states and 3 Canadian provinces (Fig. 1). Populations were arranged by ecotypes (wild type vs. cultivated) and within the two ecotypes divided by three regions. There were 159 *P. quinquefolium* plants, with 67 being wild type and 92 cultivated (Table 1).

**Panax ginseng.** Thirty two fresh root samples from four cultivated populations were provided by Yong-Pho Lim, Department of Horticulture, Chungnam National University, Daejeon, S. Korea. The four populations were located at Kumsan, Jinan, Muju, and Punggi of S. Korea (Table 1, Fig. 1). These samples were sent to the University of Wisconsin–Madison, Dept. of Horticulture, by next-day-air delivery.

**DNA ISOLATION.** Fresh leaves or roots were collected during July and August of 1997 and stored at 5 °C. DNA was extracted within 5 d. One gram of leaf or root tissue from each plant was frozen in liquid nitrogen and then ground with a mortar and pestle. Ground tissue was combined with CTAB (hexadecyltrimethylammonium bromide) extraction buffer, and DNA was extracted as described in Rogers and Bendich (1988). RAPD reactions were performed using the methods of Skroch and Nienhuis (1995a).

**PRIMER SELECTION.** A randomly selected set of 40 plants was evaluated for polymorphism using 80 ten-base RAPD primers (Operon Technologies, Alameda Calif.). Ten primers were selected based on the clarity and number of polymorphic bands: A4, BC 98, BC 203, D3, D5, H4, H5, O15, U10, and U15. Primers BC 98 and BC 203 were first selected based on the results of Bai et al. (1997) and custom synthesized by Operon Technologies.

**GENETIC RELATIONSHIP.** The genetic relationships among ginseng populations were estimated using the methods of Tivang et al. (1996). Comigrating polymorphic fragments ranging from 0.2 to 2.0 kb were scored for presence (1) or absence (0). Genetic distance (GD) among the 191 individual plants was estimated as the complement to the simple matching coefficient (Gower, 1972). The resulting 191 × 191 genetic distance matrix was fitted in two dimensions using the Kruskal monotonic dimensional scaling (MDS) procedure in SAS 6.10 (SAS Institute, 1995; Young, 1987).

**POPULATION INTEGRITY.** Species, ecotype, and regional populations were compared pairwise for differences in individual marker frequencies. The genetic differences as indicated by the MDS plot, were verified by comparing fragment frequencies on a band-by-band basis. The significance of these observed frequency differences was tested using the chi-square test of goodness of fit (Snedecor and Cochran, 1967). Populations were evaluated pairwise for each band. Under the null hypothesis, the mean fragment frequency became the expected fragment frequency. Bands were defined as polymorphic if the mean fragment frequency was not fixed (i.e., 1 or 0). Chi-square values and

---

### Table 1. The species, region of origin and sample size for *Panax quinquefolium* and *P. ginseng* germplasm.

<table>
<thead>
<tr>
<th>Species</th>
<th>Type</th>
<th>Region</th>
<th>Sample size</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>P. quinquefolium</em></td>
<td>Cultivated</td>
<td>Wisconsin</td>
<td>43</td>
</tr>
<tr>
<td></td>
<td>Cultivated</td>
<td>Canada</td>
<td>24</td>
</tr>
<tr>
<td></td>
<td>Cultivated</td>
<td>Kentucky, Missouri, and North Carolina</td>
<td>25</td>
</tr>
<tr>
<td>Total cultivated</td>
<td></td>
<td></td>
<td>92</td>
</tr>
<tr>
<td>Wild type</td>
<td></td>
<td>Pennsylvania</td>
<td>15</td>
</tr>
<tr>
<td>Wild type</td>
<td></td>
<td>Tennessee</td>
<td>26</td>
</tr>
<tr>
<td>Wild type</td>
<td></td>
<td>Wisconsin</td>
<td>26</td>
</tr>
<tr>
<td>Total wild type</td>
<td></td>
<td></td>
<td>67</td>
</tr>
<tr>
<td><em>P. ginseng</em></td>
<td>Cultivated</td>
<td>S. Korea</td>
<td>32</td>
</tr>
</tbody>
</table>
degrees of freedom were pooled over all polymorphic bands. Degrees of freedom for a paired contrast were lost if the expected value (frequency \times number of genotypes) fell below one. All such observations were grouped into a single class with one degree of freedom (Snedecor and Cochran, 1967).

**Sample variance.** When estimates of genetic distance are calculated using a random subset of marker bands, sampling variance can occur if that subset does not equal the value obtained from all possible bands. As larger numbers of random markers are generated, an increasingly uniform distribution is produced. Consequently, the sample variance, due to under or over-sampling regions of the genome, is reduced while the precision of estimating genetic relationships improves (Nienhuis et al., 1995). Genetic distance is the mean over loci sampled; therefore, the sample variance for a given random subset of markers can be calculated using the formula given by Skroch and Nienhuis (1995b).

**Experimental error.** The frequency of random errors in scoring the data were estimated by replicating 7 randomly selected plants and blindly scoring the replications as part of the set of 191 genotypes.

### Results and Conclusions

**RAPD polymorphism.** There were 10 primers scored for 4 to 15 polymorphic bands each, resulting in 100 bands scored between *P. quinquefolium* and *P. ginseng*. The 3 primers—BC 203, D5, and A4—accounted for 15, 14, and 12 polymorphic bands among 191 plants, respectively. Among the *P. quinquefolium* regional populations, the number of polymorphic bands observed for pairwise comparisons ranged from 59 to 82 with a mean of 68. Examples of RAPD amplification products are displayed in Fig. 2, and some unique polymorphic bands between the two species and among the cultivated and wild-type *P. quinquefolium* are indicated.

**Experimental error.** A comparison of 100 RAPD bands scored among the 7 replications found that the frequency of random scoring errors was <3% (data not shown).

**Sample variance.** The relationship between the coefficient of variation for the number of polymorphic RAPD bands of four *Panax* populations was found to be <25% for 100 RAPD bands (Fig. 3). All four populations had a reduced ratio of mean GD to the standard deviation for GD as the number of RAPD bands increased. This value of 25% (\( \mu \) GD/s GD) is consistent with the findings of other germplasm organization studies using RAPD markers (Nienhuis et al., 1995; Tivang et al., 1996).

**Genetic relationships.** The genetic distance within a population gives an estimate of the degree of genetic similarity among that population’s members while the genetic distance among populations gives an estimate of the degree of genetic similarity among entire populations. A value of 0 means maximum similarity, and a value of 1 means no similarity. The high GD values among cultivated *P. ginseng*, cultivated *P. quinquefolium*, and wild-type *P. quinquefolium* indicates there are genetic differences among species, ecotypes and regional populations (Table 2). Genetic distances among species and regional populations are displayed with an MDS analysis of the 191 × 191 matrix of genetic distance values. An inspection of the MDS plot shows five discrete clusters (Fig. 4). The first dimension separates two major groups corresponding to the two *Panax* species. The second dimension separates the *P. quinquefolium* group into three distinct clusters corresponding to the cultivated population and two wild-type, regional populations (Fig. 5). These results confirm that *P. ginseng* and *P. quinquefolium* may be distinct species and that *P. quinquefolium* can be subclassified as either cultitpes or wild ecotypes.

**Population integrity.** The population comparisons, based on chi-square goodness-of-fit tests, were also significantly different and were consistent with the separation of populations observed in the MDS plots. The chi-square values (Table 3) and the MDS analysis of the genetic distance matrix (Fig. 4) show *P. ginseng* and *P. quinquefolium* to be genetically distinct. They also indicate that there are no distinct genetic populations of cultivated *P. ginseng* or *P. quinquefolium* and that the Pennsylvania wild-type *P. quinquefolium* may be of cultivated origin. However, Wisconsin and Smoky Mountain National Park (Tennessee) wild-type *P. quinquefolium* may be distinct species and that *P. quinquefolium* can be subclassified as either cultitpes or wild ecotypes.
Table 2. The mean genetic distance values of the *Panax* species and regional populations, and mean genetic distance values between the *Panax* populations which formed distinct clusters in the multidimensional scaling (MDS) plot.

<table>
<thead>
<tr>
<th>Species</th>
<th>Type</th>
<th>Regions</th>
<th>Within GD</th>
<th>Among GD</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>P. quinquefolium</em></td>
<td>All</td>
<td>All</td>
<td>0.20</td>
<td>---</td>
</tr>
<tr>
<td></td>
<td>Cultivated</td>
<td>N. America</td>
<td>0.15</td>
<td>---</td>
</tr>
<tr>
<td></td>
<td>Cultivated</td>
<td>Wisconsin</td>
<td>0.14</td>
<td>---</td>
</tr>
<tr>
<td></td>
<td>Cultivated</td>
<td>Canada</td>
<td>0.14</td>
<td>---</td>
</tr>
<tr>
<td></td>
<td>Cultivated</td>
<td>Southern U.S.</td>
<td>0.16</td>
<td>---</td>
</tr>
<tr>
<td></td>
<td>Wild type</td>
<td>U.S.</td>
<td>0.24</td>
<td>---</td>
</tr>
<tr>
<td></td>
<td>Wild type</td>
<td>Pennsylvania</td>
<td>0.13</td>
<td>---</td>
</tr>
<tr>
<td></td>
<td>Wild type</td>
<td>Smoky Mt. Natl. Park</td>
<td>0.19</td>
<td>---</td>
</tr>
<tr>
<td></td>
<td>Wild type</td>
<td>Wisconsin</td>
<td>0.19</td>
<td>---</td>
</tr>
</tbody>
</table>

*P. quinquefolium* vs. *P. ginseng*  
*P. quinquefolium*^1^ vs. TN^2^  
*P. quinquefolium* cultivated^3^ vs. WI^4^  
*P. quinquefolium* cultivated vs. PA^5^  

^1^All *P. quinquefolium* populations excluding Tennessee.  
^2^Wild *P. quinquefolium* populations from Smoky Mountain National Park, Tennessee.  
^3^All cultivated *P. quinquefolium* populations.  
^4^Wild *P. quinquefolium* populations from Wisconsin.  
^5^Wild *P. quinquefolium* from Pennsylvania.

*Panax quinquefolium* populations are genetically unique as displayed with the coordination of genetic distance values calculated with RAPD bands (Fig. 5).

The genetic relationships of the *P. quinquefolium* samples are consistent with the history of cultivation and selection within populations. In Wisconsin, cultivated *P. quinquefolium* are said to have been collected from wild stands in Marathon county in the 1920s (Persons, 1994). Artificial selection, by Wisconsin ginseng farmers for >75 years may have led to the genetic distance observed between the wild-type and cultivated populations of Wisconsin. Alternatively these differences may be due to the founding population effects caused by the increase and distribution of seed from wild stands in Marathon county only. This conclusion is supported by the fact that there were no significant differences between the fragment frequencies of regional, cultivated populations.

In Pennsylvania, *P. quinquefolium* has been gathered for >150 years. Contemporary gatherers may select wild-type seed and roots for parents of future garden populations (J. Epler, Epler Fur Co., personal communication). The Pennsylvania populations in this study were collected from wild plants that had been collected from the wild and transplanted into private woodland gardens. However, their collection from the wild does not guarantee their origin to the wild type. Pennsylvania ginseng retail companies also sell *P. quinquefolium* seed to wild ginseng harvesters. This seed is often imported from cultivated populations in Wisconsin and Canada. The Pennsylvania populations were collected in areas that had an extensive history of ginseng harvest (Epler, personal communication). The plants may have been from cultivated seed (Wisconsin or Canada) that was planted in that area in previous years by ginseng harvesters. This may explain their relatively low GD and similarity to cultivated populations. Alternatively, it is possible that Pennsylvania wild-type populations served as the source of imported seed for the cultivated populations of Wisconsin and Canada. However, there are no records that provide any evidence for this hypothesis.

The first explanation for the similarity between Pennsylvania samples and other cultivated *P. quinquefolium* samples is likely considering that the populations which were genetically distant from the cultivated populations were growing in protected parks and reserves. This evidence suggests that, in areas with a history of ginseng gathering, wild ecotypes may be mixed with cultivated populations. This is supported by the fact that wild populations are genetically unique as displayed with the coordination of genetic distance values calculated with RAPD bands (Fig. 5).

Fig. 4. Multidimensional scaling plot of the genetic distance values for individual Panax ginseng and *P. quinquefolium* plants showing distance between species and within *P. quinquefolium*. Cultivated and wild *P. quinquefolium* are indicated by the cluster above and the Tennessee wild type is found below.
landraces. Smoky Mountain National Park is home to many protected populations of *Panax quinquefolium*. These populations appear to have an unique genetic integrity and may represent a distinct center of genetic diversity. These results demonstrate the utility of RAPD markers for characterizing *Panax* sp. germplasm. They indicate that there are no genetically distinct cultivated populations of either species. However, they suggest that wild-type populations may present *P. quinquefolium* breeders with a source of genetic diversity for improving populations. Specifically, populations from Wisconsin and Tennessee might provide sources of unique genetic material for the establishment of a core collection of *P. quinquefolium*. The core collection could be evaluated for horticulturally important traits, such as disease resistance, which might be introgressed into cultivated genotypes thusly increasing the value of American ginseng crops.

### Literature Cited


