Theor Appl Genet (2000) 101:554–563

I.V. Bartish · L.P. Garkava · K. Rumpunen
H. Nybom

Phylogenetic relationships and differentiation among and within populations of *Chaenomeles* Lindl. (Rosaceae) estimated with RAPDs and isozymes

Received: 23 December 1999 / Accepted: 17 January 2000

Abstract RAPD and isozyme analyses based on numerous markers have been used for the first time to investigate patterns of phenetic and genetic differentiation among and within nine wild populations of the genus *Chaenomeles* represented by the species *C. japonica*, *C. speciosa*, *C. cathayensis* and *C. thibetica*. Highly significant correlations were found between the two different marker systems for both phenetic distances and gene diversity estimates. In agreement with previous studies on cultivated *Chaenomeles* material, *C. japonica* was clearly differentiated from *C. speciosa* and *C. cathayensis*. The recently recognised species *C. thibetica* appeared to be rather closely related to *C. cathayensis*. Populations of *C. japonica* and *C. speciosa* were considerably more diverse than populations of *C. cathayensis* and *C. thibetica*. Correspondingly, most of the total variability could be attributed to the within-population differentiation in the case of *C. japonica* and *C. speciosa*, and to the between-population differentiation in the case of *C. cathayensis*. Differences in mating systems among the species can be suggested as a possible explanation of the results. A discordant pattern was found between RAPDs and isozymes in the analyses of population structure within *C. japonica*. This may be explained by a higher proportion of non-neutral markers for isozymes than for RAPDs. This finding also shows the importance of using multiple molecular marker systems in studies of population structure.

Key words Gene diversity · Isozyme · Non-neutrality · RAPD · Rosaceae

Communicated by P.M.A. Tigerstedt

I.V. Bartish · L.P. Garkava · K. Rumpunen (✉) · H. Nybom
Balsgård–Department of Horticultural Plant Breeding,
Swedish University of Agricultural Sciences,
Fjälkestadsvägen 123-1, S-291 94 Kristianstad, Sweden
e-mail: kimmo.rumpunen@hvf.slu.se
Fax: +46 44 755 30

Introduction

Chaenomeles is a genus within the subfamily Maloideae comprising four diploid (2n=34) species (Phipps et al. 1990). One species is endemic to Japan, and three originate in central Asia. Presently they are being investigated for possible domestication and improvement through breeding in northern Europe. The plants within the genus *Chaenomeles* are all moderate-sized shrubs adaptable to cultivation in temperate areas. *C. japonica* (Thunb.) Lindl. ex Spach (Japanese quince) is outcrossing and strongly self-incompatible (Rumpunen et al. 1999). The mating system has not yet been properly investigated for the other species in the genus. *C. speciosa* (Sweet) Nakai (flowering quince) is a traditional medicinal plant (Anonymous 1989) and grows in Central and southern China, Tibet and Burma. Recently it has also become cultivated in eastern China (Wang et al. 1998). In the province of Yunnan, *C. speciosa* grows sympatrically with *C. cathayensis* (Hemsl.) Schneider (Chinese quince), the largest shrub in the genus, which can also be found in Bhutan and Burma. The fourth species, *C. thibetica* Yü (Tibetan quince), first described in 1963 (Yü and Kuan) and included in the most recent check list of the subfamily Maloideae (Phipps et al. 1990), grows in Tibet and western Sichuan. *C. japonica*, *C. cathayensis* and *C. speciosa* have been used to create several interspecific hybrids for 400 years, resulting in more than 500 cultivars (Weber 1963), which today are widely appreciated as ornamentals.

Efficient methods to clarify the taxonomic status of both the wild and the cultivated material are much needed. Insights into the relative gene diversity among and within wild populations of *Chaenomeles* would be useful in plant breeding and also for the development of strategies for *ex situ* conservation of plant genetic resources. Molecular markers [random amplified polymorphic DNAs (RAPDs) and isozymes] have recently been used to evaluate gene diversity in plant samples of three species of *Chaenomeles*: *C. cathayensis*, *C. japonica* and *C. speciosa*. The accessions were obtained from several botanical gardens, and included samples of domesticated

Table 1 Accession identity (ID), number of individuals and origin of the nine *Chaenomeles* populations investigated

ID	Individuals	Species	Origin (co-ordinates[a,b] and altitude)
P9701	21	*C. japonica*	Shionomuro, Imaichi, Tochigi, Japan (36°45'N, 139°48'E[a]; 275 m)
P9702	19	*C. japonica*	Noguchi, Nikko, Tochigi, Japan (36°45'N, 139°43'E[a]; 450 m)
P9724	21	*C. japonica*	Ibuki, Shiga, Japan (35°15'N, 136°20'E[b]; 900 m)
P9801	19	*C. speciosa*	Zhenyuan, Yunnan, China (23°50'N, 101°10'E[b]; 1800–2300 m)
P9802	21	*C. speciosa*	Hutiaoxia, Zhongdian, Yunnan, China (27°50'N, 99°30'E[b]; 1800–2100 m)
P9803	21	*C. cathayensis*	Hutiaoxia, Zhongdian, Yunnan, China (27°50'N, 99°30'E[b]; 1800–2100 m)
P9804	21	*C. cathayensis*	Caojian, Yunlong, Yunnan, China (25°35'N, 98°05'E[b] , 2500–2600 m)
P9805	21	*C. speciosa*	Dali, Yunnan, China (25°35'N, 100°10'E[b]; 1900–2400 m)
P9806	21	*C. thibetica*	Yi'ong, Bomi, Tibet (30°00'N, 95°00'E[b]; 2500 m)

[a] Precise
[b] Estimated

populations (Bartish et al. 1999b, Garkava et al. 2000). Both marker systems appeared to be useful in the discrimination of accessions. Cluster and multidimensional scaling analyses yielded similar results for the two molecular marker methods (Garkava et al. 2000). The two methods also showed a concordant pattern of relative diversity within species. *C. japonica* and *C. speciosa* were found to be considerably more diverse than *C. cathayensis*. However, because of possible interspecific hybridisation among accessions from botanical gardens, it was not clear to what extent the obtained results would be representative for indigenous populations of *Chaenomeles*.

Isozymes and RAPDs are among the most frequently used molecular markers for taxonomic and systematic analyses of plants. Correspondence between results with nuclear DNA and isozyme markers using cluster and ordination analyses of the same plant sample is generally moderate (Heun et al. 1994), but highly correlated estimates of genetic distances have been reported between RAPD and isozyme datasets (Lifante and Aguinagalde 1996). Some authors have considered RAPD-based analysis to be more definitive in its separation of clusters than isozyme-based analysis (Heun et al. 1994), mostly because RAPDs yield lower coefficients of variation than isozymes and thus allow for a higher level of discrimination (Staub et al. 1997). A different view was expressed by Chan and Sun (1997), who suggested that the reliability of RAPD and isozyme data is comparable at the intraspecific level but that at the interspecific level RAPD markers are less suitable for studying phylogenetic relationships than both isozyme and restriction fragment length polymorphism (RFLP) markers.

When population structure is being investigated, different molecular marker systems and the interpretation of the scored fingerprints as phenetic or genetic data may produce different results. Gene diversity estimates obtained from RAPDs may be higher than those obtained from isozymes (Aagaard et al. 1998) because "an inherently higher rate of detectable mutations and weaker degree of selective constraint at RAPD compared to allozyme loci". Furthermore, estimates of among-population differentiation could be inflated when based on RAPD phenotypes compared to RAPD and isozyme genotypes (Isabel et al. 1999). Several studies have also questioned the neutrality of isozyme loci, suggesting that both balancing (Karl and Avise 1992; Raybould et al. 1996) and diversifying (Berry and Kreitman 1993; Lönn 1993; Le Corre et al. 1997) selection may operate on some of the enzymes. The simultaneous application of both techniques thus seems to be a more reliable approach when estimating population parameters and phylogenetic relationships, as indeed has been advocated by, for example, Chan and Sun (1997) and Ayres and Ryan (1999).

The purpose of our investigation was: (1) to determine population structure and gene diversity for species and populations within *Chaenomeles*, and (2) to investigate the correspondence between RAPD- and isozyme-based datasets in analyses of phylogenetic relationships, population structure and relative gene diversity estimates.

Materials and methods

Plant material

All of the accessions were collected in the wild in China and Japan during the autumn of 1997, except for P9801, P9802 and P9803, which were purchased from local markets (the fruits had been collected in the nearby mountains). Only one fruit, with numerous seeds, was picked from each shrub. The seeds from one fruit were presumed to constitute a family. Following cold stratification, the seeds were germinated in a greenhouse. When the seedlings were 3 months old, 7 plants from each of three randomly selected families per population were sampled for molecular analyses. However, one of the populations (P9724), was represented by only one family (of 21 seedlings to achieve a more balanced plant sampling). Altogether, 189 plants representing nine populations of *C. japonica*, *C. speciosa*, *C. cathayensis* and *C. thibetica* were initially sampled (Table 1). One-year old plants were used for isozyme extractions. Since 4 plants died before being sampled for isozyme analysis (1 plant per family in two families of population P9702 and in two families of population P9801), the total number of plants used in both analyses was 185.

Isozyme analysis

Enzymes were extracted from leaflets and phloem tissue, and polyacrylamide gel electrophoresis was carried out as previously described (Garkava et al. 2000).

The enzyme assay procedures for acid phosphatase (ACP, EC 3.1.3.2), and esterases (EST, EC 3.1.1.-) were as described by Wendel and Weeden (1989). Glutamate oxaloacetate transaminase (GOT, EC 2.6.1.1), shikimate dehydrogenase (SKDH, EC 1.1.1.25), and phosphoglucomutase (PGM, EC 5.4.2.2) were stained according to Vallejos (1983). Peroxidase (PRX, EC 1.11.1.7) was stained with 0.09% benzidine in 0.32 *M* sodium acetate buffer adjusted to pH 5.4, and bands were revealed with 0.01% hydrogen peroxide. All isozyme systems were stained in the dark at 37°–38°C. Bands were scored visually as presence and absence of a marker.

DNA extraction and polymerase chain reaction (PCR) amplification

Young unexpanded leaves (5–15 mg per individual) were collected, placed into an Eppendorf tube (1.5 ml) and stored at –80°C. Leaves were ground with a minipestle to a homogenate in 200–250 µl of 2×CTAB extraction buffer and incubated for 45 min at 65°C. From this point on all procedures for DNA extraction and PCR amplification followed the previously described method of Bartish et al. (1999b). The only exception was the use of a different thermocycler, PTC–100 TM (MJ Research).

Ten primers were selected out of 100 (Operon Technologies sets A, B, D, E, F) that had been tested previously (Bartish et al. 1999b). The main criterion for selection was reproducibility of amplification. However, some bias in the direction of higher polymorphism among RAPD markers generated by the primers used in the present study (compared to random sampling of primers) cannot be excluded.

Products of amplification from DNA samples of families from different populations were run on the same gel. In all cases, except population P9803, separate families from the same population were analysed on separate gels. For each primer, three samples of each population were replicated in different PCR experiments and analysed on separate gels. Products of PCR amplification were scored visually as presence or absence of a marker. Bands of identical size, amplified with the same primer, were considered to be homologous. Bands, which were very faint and rare (frequency less than 3%), were not included into the analysis.

Data analysis

Binary data matrices were produced from the scored isozyme and RAPD markers for individual plants. Phenotype frequencies of molecular markers within families of *C. japonica* and within the total sample of populations and species were calculated and used to create datasets of marker frequencies. The datasets were used to compute phenetic (Euclidean) distances among populations (or families in the case of *C. japonica*) according to the method of Harrison et al. (1997) and Hancock and Bringhurst (1979) for RAPD-based and isozyme-based data, respectively. This approach allows evaluation of phylogenetic relationships between populations when precise genetic data for molecular markers are absent.

In addition, a RAPD-based dataset of null allele frequencies within populations was produced for the whole material. The markers were considered to represent separate loci. Null allele frequencies were calculated for these markers in accordance with Lynch and Milligan (1994) and Bartish et al. (1999a). The dataset of RAPD null allele frequencies for populations was then used to compute Nei's genetic distances (Nei 1972) among populations.

All datasets were analysed with NTSYS-pc software (version 1.8, Rohlf 1997). The SAHN programme from this package was used to produce UPGMA dendrograms from each of two phenetic distance matrices and one genetic distance matrix to show the amount of relatedness among populations and species. Principal co-ordinate analysis (DCENTER and EIGENVECTOR programmes) was run for each of the two matrices of phenetic distances among populations, and for two more matrices of phenetic distances among families within *C. japonica*. We used this analysis to obtain additional information about interpopulation relations, which was revealed by different methods. For graphical representation of principal co-ordinate analyses, two-dimensional plots for families and three-dimensional plots for populations were produced by NTSYS.

We used the matrix of null allele frequencies within populations from the RAPD dataset for cluster analysis in PHYLIP 3.573c software (Felsenstein 1993). The matrix was bootstrapped to produce 100 random matrices, and Nei's genetic distances (1972) were calculated by GENDIST from PHYLIP for each of the 100 bootstraps. These matrices were subjected to cluster analysis in the NEIGHBOR programme from PHYLIP with the neighbour-joining algorithm. A strict consensus tree, summarising the relationships described by all 100 resampled datasets (with bootstrap support for each node), was finally produced.

Two of the most frequently used approaches in RAPD analysis were employed to evaluate gene diversities on both RAPD and isozyme datasets: estimates of unbiased values of expected heterozygosity, as in Lynch and Milligan (1994), and Shannon's index, as in Bussell (1999). Monomorphic markers were included into the Shannon's index calculations as suggested by Bussell (1999), in order to standardise values and enable comparisons between studies in which a different number of markers were recorded. We used the modified standardisation procedure suggested by Liu and Furnier (1993) because the diversity of the total plant sample may have a substantial influence on within-population diversity estimations (Bartish et al. 1999c). Since portions of the genome that are monomorphic for the recessive allele cannot be observed, doubling the number of markers monomorphic for the dominant allele (presence of a band) should yield a good estimate of the number of monomorphic markers and hopefully improve the total estimate of within-species polymorphism. We calculated the total number of monomorphic and polymorphic molecular markers separately for each species; thus, the combined number of markers included into all gene diversity estimates was different for each of the species.

The Lynch and Milligan index was calculated for each RAPD marker, *i*, separately for each population, $H_{pop}(i)$, species, $H_{sp}(i)$, and the genus, $H_{ge}(i)$. The mean Lynch and Milligan index for a population, H_{pop}, was then calculated by averaging $H_{pop}(i)$ over all markers. H_{sp} and H_{ge} were calculated similarly.

Shannon's index was calculated for each RAPD or isozyme marker, *i*, separately for each population, $H'_{pop}(i)$, species, $H'_{sp}(i)$, and the genus, $H'_{ge}(i)$. The average Shannon's index for a population H'_{pop} was then calculated by averaging $H'_{pop}(i)$ over all markers as in Monaghan and Halloran (1996). H'_{sp} and H'_{ge} were calculated similarly. The mean Shannon's index was then calculated for each marker within a species, $H'_{sp}(i)$, by averaging $H'_{pop}(i)$ over all populations within each species (3 populations for *C. japonica* and *C. speciosa*, 2 populations for *C. cathayensis*) . Similarly $\underline{H}'_{ge/sp}(i)$ was calculated by averaging $H'_{sp}(i)$ over all species, and $H'_{ge/pop}(i)$ by averaging $H'_{pop}(i)$ over all populations within the genus.

We calculated the components of diversity between groups for different levels of taxonomic hierarchy analogous to Bussel (1999). For each locus the component of diversity between populations within species, $G'_{pop/sp}(i)$, was $[H'_{sp}(i)–H'_{sp}(i)]/H'_{sp}(i)$, and the component between populations within the genus, $G'_{pop/ge}(i)$, was $[H'_{ge}(i)–H'_{ge/pop}(i)]/H'_{ge}(i)$. The component of diversity between species within genus, $G'_{sp/ge}(i)$, was correspondingly $[H'_{ge}(i)–H'_{ge/sp}(i)]/H'_{ge}(i)$. Mean estimates of $G'_{pop/sp}$, $G'_{pop/ge}$ and $G'_{sp/ge}$ were calculated by averaging $G'_{pop/sp}(i)$, $G'_{pop/ge}(i)$ and $G'_{sp/ge}(i)$ over all markers. We then calculated 99% confidence intervals for $G'_{pop/sp}$ of each species to reveal markers that deviated significantly from the parametric mean of all markers. The same procedure was also applied to only two populations (P9701 and P9702) of *C. japonica*.

Two matrices of phenetic distances among all individual plants (RAPD and isozyme datasets of polymorphic markers) were used as input distance matrices in AMOVA (Huff et al. 1993), and Φ-statistics was calculated from the variance components. In this paper we only report the proportions of variance differentiating among populations within species ($\Phi_{pop/sp}$), among populations

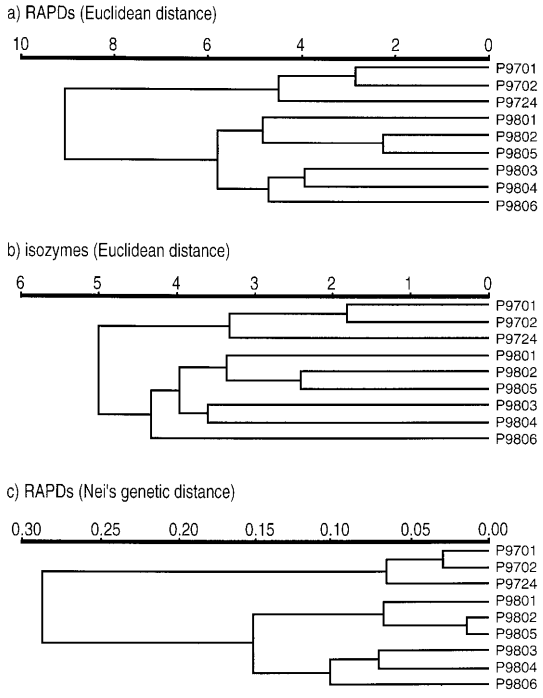

Fig. 1a–c UPGMA dendrograms of RAPD-based (Euclidean distances) (**a**), isozyme-based (Euclidean distances) (**b**) and RAPD-based (**c**) Nei's genetic distances between populations of *Chaenomeles*

within the genus ($\Phi_{pop/ge}$) and among species within genus ($\Phi_{sp/ge}$), which correspond to the *G'*-statistics.

We used Pearson product-moment correlation analysis and the standard *t*-test (paired comparisons), respectively, to evaluate associations between estimates of within-population gene diversity obtained with different molecular datasets. This analysis was also used to estimate associations between some of the molecular markers.

The product-moment correlation and the Mantel test statistic were computed with NTSYS (MXCOMP programme) to measure the degree of relationship between the distance matrices.

Results

Phenetic relationships among populations

For the total plant sample, 271 (including 17 monomorphic) RAPD and 103 (9 monomorphic) isozyme markers were scored. Phenetic (Euclidean) distances between populations varied in the range 2.87–10.07 for RAPD and 1.84–5.77 for isozyme datasets, respectively. Genetic (Nei's) distances based on RAPDs varied from 0.024 to 0.332. The complete matrices of phenetic and genetic distances for all datasets are available on request. All matrices were significantly (P<0.001) correlated with each other according to Mantel's tests. Coefficients of correlation varied from r=0.889 (between matrices of RAPD and isozyme phenetic distances) to r=0.921 (between the matrix of RAPD genetic distances and the matrix of RAPD phenetic distances).

In accordance with the high values of correlation between the matrices of phenetic and genetic distances,

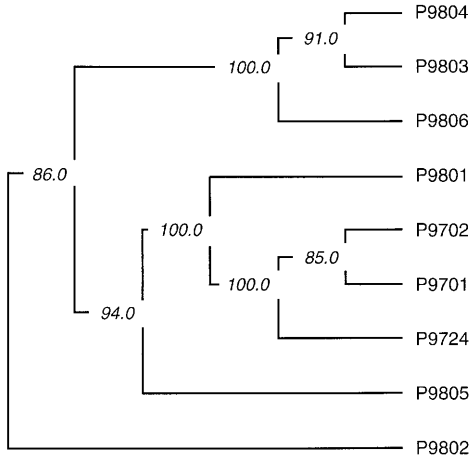

Fig. 2 Dendrogram of phylogenetic relationships between populations of *Chaenomeles* calculated on RAPD null allele frequencies by the neighbour-joining method with bootstrap support (%) for each node of the tree

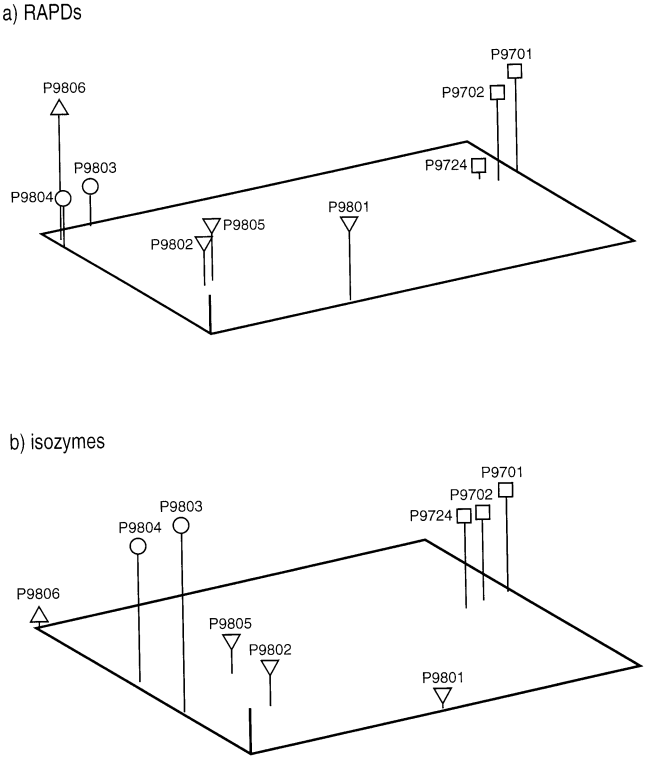

Fig. 3a, b Three-dimensional representation of principal co-ordinate analysis of phenetic relationships between populations of *Chaenomeles*. Percentage explained variability was for **a** the phenetic RAPD-based dataset: PC1 62.4%, PC2 16.0% and PC3 5.9%; **b** the isozyme-based dataset: PC1 44.2%, PC2 17.2% and PC3 12.9%

UPGMA analysis of phylogenetic relationships between populations of *Chaenomeles* produced congruent dendrograms (Fig. 1). Analysis of molecular markers grouped populations from the same species (as inferred from their morphological characters) in the same cluster. This result confirmed a monophyletic origin for all

Table 2 Gene diversity estimated by Shannon's index (*H'*) and by the Lynch and Milligan index (*H*) with standard error, as well as the percentage polymorphic markers (*P*) within populations and species of *Chaenomeles*, based on RAPD and isozyme data

Group	H'_{pop} RAPD	H'_{pop} isozyme	H_{pop} RAPD	*P*(%) RAPD	*P*(%) isozyme
C. cathayensis					
P9803	0.140 (0.024)[a]	0.069 (0.025)	0.066 (0.012)	15.8	7.1
P9804	0.282 (0.032)	0.081 (0.026)	0.139 (0.017)	30.3	8.9
C. japonica					
P9701	0.487 (0.032)	0.422 (0.047)	0.229 (0.017)	54.6	50.9
P9702	0.467 (0.032)	0.375 (0.050)	0.212 (0.016)	50.9	44.6
P9724	0.424 (0.031)	0.413 (0.050)	0.196 (0.016)	49.1	46.4
C. speciosa					
P9801	0.547 (0.030)	0.503 (0.049)	0.253 (0.015)	63.5	54.5
P9802	0.520 (0.030)	0.419 (0.044)	0.239 (0.014)	62.4	53.6
P9805	0.589 (0.030)	0.460 (0.044)	0.258 (0.014)	66.1	60.7
C. thibetica					
P9806	0.286 (0.029)	0.189 (0.052)	0.141 (0.015)	34.3	14.3
Group	H'_{sp} RAPD	H'_{sp} isozyme	H_{sp} RAPD	*P*(%) RAPD	*P*(%) isozyme
C. cathayensis	0.337 (0.032)	0.284 (0.056)	0.160 (0.015)	39.9	22.3
C. japonica	0.620 (0.032)	0.612 (0.050)	0.278 (0.016)	69.0	67.9
C. speciosa	0.706 (0.029)	0.655 (0.043)	0.306 (0.014)	78.2	75.9
Group	H'_{ge} RAPD	H'_{ge} isozyme	H_{ge} RAPD	*P*(%) RAPD	*P*(%) isozyme
Chaenomeles	0.746 (0.013)	0.665 (0.024)	0.282 (0.008)	93.7	92.0

[a] Standard error in parenthesis

Table 3 Correlation (Pearson's) between within-population gene diversity values estimated by Shannon's index (H'_{pop}) and by the Lynch and Milligan index (H_{pop}) and percentage polymorphic markers (*P*), based on RAPD and isozyme datasets

	H'_{pop} RAPD	H'_{pop} isozymeH_{pop}	RAPD	*P* RAPD
H'_{pop} isozyme	0.946***			
H_{pop} RAPD	0.996***	0.942**		
P RAPD	0.995***	0.952***	0.993***	
P isozyme	0.963***	0.979***	0.950***	0.963*

* $P<0.05$, ** $P<0.01$, *** $P<0.001$

conspecific populations. Populations of *C. japonica* always clustered separately from the other populations. Results of cluster analysis based on RAPD genetic distances between populations (Fig. 1c) were in good agreement with cluster analyses based on phenetic distances obtained from RAPD and isozyme datasets, respectively (Fig. 1a, b).

High values of bootstrap support were obtained for each node of the phylogenetic tree based on the dataset of genetic distances and the neighbour-joining algorithm (Fig 2). However, populations of *C. speciosa* were widely scattered in this dendrogram.

The pattern of phenetic relationships, revealed by UPGMA cluster analyses, was further supported by principal co-ordinate analysis (Fig. 3). RAPD- and isozyme-based phenetic distances produced results which were concordant in: (1) grouping populations P9701, P9702 and P9724 (*C. japonica)* separately from all other populations; (2) joining populations P9802 and P9805 between population P9801 (*C. speciosa*) and populations P9803 and P9804 (*C. cathayensis*), respectively; (3) showing the relative closeness of population P9806 (*C. thibetica*) to populations P9803 and P9804.

Gene diversity

Analysis of gene diversity by the Lynch and Milligan index (*H*) and Shannon's index (*H'*) within populations and species of *Chaenomeles* showed that *C. japonica* and *C. speciosa* were substantially more diverse than *C. cathayensis* and *C. thibetica*, independent of the method used for evaluation (Table 2). Percentages of polymorphic markers (*P*) and values of Shannon's index were similar when RAPD- and isozyme-based datasets were compared, although RAPD-based estimates were generally somewhat higher. The only clear difference among estimates of within-population gene diversity was found in population P9804. According to RAPD-based estimates, P9804 was approximately twice as diverse as P9803,

Table 4 Components of diversity among populations within species, among populations within genus and among species within genus, estimated from phenetic RAPD and isozyme datasets by AMOVA (Φ-statistics with *P*-value) or derived from Shannon's index (*G'*-statistics with standard error)

Group variable	$G'_{pop/sp}$ RAPD	$G'_{pop/sp}$ isozyme	$\Phi_{pop/sp}$ RAPD	$\Phi_{pop/sp}$ isozyme
Populations within				
C. cathayensis	0.606 (0.036)	0.830 (0.026)	0.229 (0.001)	0.615 (0.001)
C. japonica	0.278 (0.031)	0.323 (0.036)	0.126 (0.367)	0.178 (0.001)
C. speciosa	0.176 (0.027)	0.268 (0.032)	0.100 (0.257)	0.181 (0.001)
Group variable	$G'_{pop/ge}$ RAPD	$G'_{pop/ge}$ isozyme	$\Phi_{pop/ge}$ RAPD	$\Phi_{pop/ge}$ isozyme
Populations	0.535 (0.013)	0.491 (0.019)	0.325 (0.001)	0.404 (0.001)
Group variable	$G'_{sp/ge}$ RAPD	$G'_{sp/ge}$ isozyme	$\Phi_{sp/ge}$ RAPD	$\Phi_{sp/ge}$ isozyme
Species	0.446 (0.015)	0.435 (0.018)	0.303 (0.001)	0.312 (0.001)

[a] Standard error in parenthesis

whereas values of gene diversity for these two populations were similar in the isozyme dataset (Table 2).

In the absence of precise genetic information on individual loci and the lack of random sampling of molecular markers since selected primers were used for RAPD analysis, we refrained from trying to compare absolute values of gene diversity estimates between the two molecular methods. Instead, the Shannon's index values for population diversity estimated by RAPDs and isozymes, respectively, were tested by a *t*-test (paired comparisons). Our results clearly showed (Table 3) that the null hypothesis of the two sets of molecular markers being independent of each other could be rejected with a high level of significance ($P<0.001$). This was also true for the comparison of the RAPD-based Lynch and Milligan (H_{pop}) and Shannon's index (H'_{pop}) diversity estimates ($P<0.001$, Table 3).

Partitioning of components of molecular variance

G'-statistics or Φ-statistics were used to evaluate the variability within and between taxa for RAPDs and isozymes, respectively. The analysis was carried out for three species separately (*C. thibetica* was not included since it was represented by only one population) and for the whole genus. Between-population variability was always lower when estimated by Φ-statistics than when estimated with *G'*-statistics (Table 4). *G'*-statistics obtained with the RAPD data resulted in higher estimates of differentiation among species and among populations across the entire genus, whereas isozyme-based values resulted in higher differentiation within species. Estimates from Φ-statistics for isozymes provided a better discrimination between groups of plants at all levels of taxonomic hierarchy than RAPDs.

The major part of the molecular variability could be attributed to the between-population component for *C. cathayensis*, as estimated from *G'*-statistics and from Φ-statistics for isozyme markers (Table 4). In contrast, much less of the total molecular variability was attributed to this component for *C. japonica* and *C. speciosa*. Within- and between-species components of total vari-

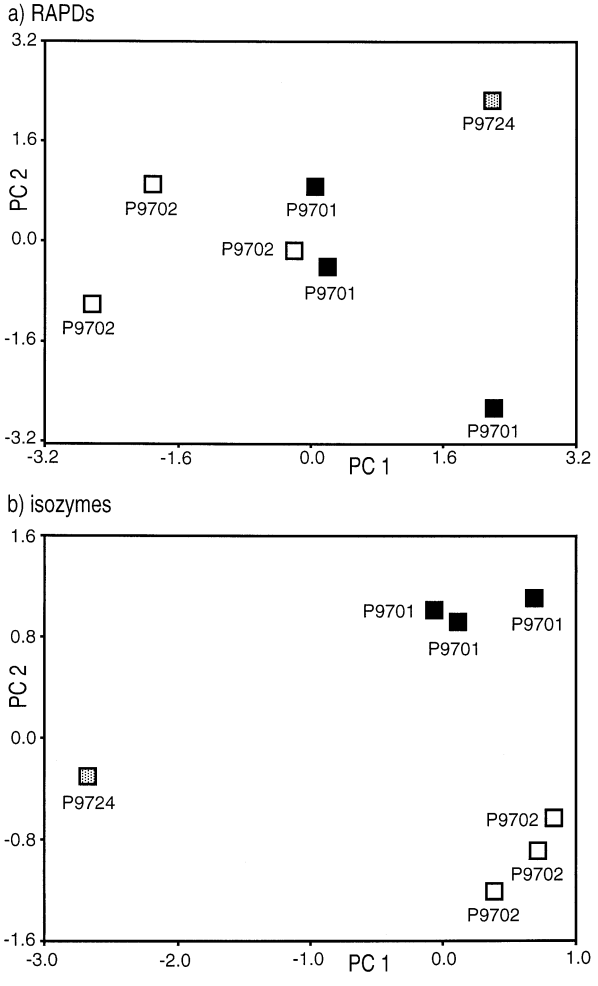

Fig. 4a, b Plot of principal co-ordinates of phenetic distances between families of *C. japonica.* **a** RAPD-based dataset: PC1 27.3%, PC2 20.1%; **b** isozyme-based dataset: PC1 39.3%, PC2 25.9%

Table 5 Differentiation between conspecific populations derived from Shannon's index [$G'_{pop/sp}(i)$] for individual markers, which deviate significantly ($P<0.01$) from the parametric means for the total set of markers. Markers beginning with OP are RAPD markers, the reminder are isozyme markers

Species	Marker	$G'_{pop/sp}(i)$
C. japonica	PEst–2	0.726
	PPrx–7	0.842
	PAcp–8	0.741
	OPA18.220	0.682
	OPF08.800	0.837
C. speciosa	PAat–9	0.689
	PAat–10	0.689
	OPE07.380	0.726
C. japonica (P9701 and P9702)	LEst–2	0.753
	LEst–3	0.676
	OPE07.1050	0.473
	OPD08.1070	0.508

ability in *Chaenomeles* were partitioned more evenly when derived from Shannon's index compared to AMOVA, with slightly more of the total variance distributed within species. The within-species component was considerably higher in the *Φ*-statistics (Table 4).

Correlations between estimates of the among-population (species) components were relatively high and significant when different methods of calculations were compared (RAPDs: *G'* vs. *Φ*, $r=0.805$, $P<0.05$; isozymes : *G'* vs. *Φ*, $r=0.983$, $P<0.01$).

Differentiation between families of *C. japonica*

RAPD- and isozyme-based matrices of phenetic distances between families of *C. japonica* were not significantly correlated ($r=0.479$, $P=0.061$). We also found substantial discrepancies between results of principal co-ordinate analyses applied to different matrices of phenetic distances. A two-dimensional plot of principal co-ordinates for families of *C. japonica* revealed only a weak population structure when RAPDs were analysed (Fig. 4a). In contrast, a clear subdivision of families into separate groups was observed for isozymes, with families from the same population grouping together (Fig. 4b).

Deviating markers

G'-statistics for individual molecular markers were used to reveal those markers, which deviated significantly ($P<0.01$) from the parametric mean for the total set of markers (Table 5). Three such isozyme markers were found for *C. japonica* (PEst–2, PPrx–7, and PAcp–8; 3.7% of the total set of markers) and 2 for *C. speciosa* (PAat–9, PAat–10; 2.3% of the total set of markers). Only 2 such RAPD markers were found for *C. japonica* (OPF08.800, OPA18.220) and 1 for *C. speciosa* (OPE07.380), comprising about 1% and 0.5% of the total set of markers, respectively. No isozyme or RAPD marker deviated significantly from the parametric mean of the *G'*-statistics for *C. cathayensis*.

When only populations P9701 and P9702 were analysed, 2 esterase markers (LEst–2 and LEst–3) and 2 RAPD markers (OPE07.1050 and OPD08.1070) deviated significantly from the parametric means for the total set of markers (Table 5).

Markers that deviated significantly from the parametric mean were found to co-occur in the same plants. Within *C. japonica* 3 isozyme (PEst–2, PPrx–7, and PAcp–8) and 1 RAPD marker (OPA18.220) were strongly associated (correlation among markers ranged from 0.74 to 0.83). Associations of deviating markers within *C. speciosa* were even stronger. Both isozyme markers (PAat–9, PAat–10) were absolutely negatively correlated ($r=-1.0$) and the correlation between OPE07.380 and the isozyme markers was high ($r=0.81$).

Discussion

Phylogenetic relationships among populations

From previous publications it is still not clear to what extent results from the different molecular marker systems are congruent. In direct comparative analyses, correlations between estimates of genetic distances, obtained from isozyme- and RAPD-based data, have proven to be significant, with estimates of correlation coefficients from moderate (Heun et al. 1994, $r=0.34$) to high (Lifante and Aguinagalde 1996, $r=0.83$). In general, isozyme- and RAPD-based estimates of gene diversity have also proven to be in good correspondence, but some discrepancies have been found in virtually all studies.

In the present study we found that correlation between distance estimates, obtained from isozyme and RAPD datasets, varied according to the geographic distances between the plant samples analysed. It could be very high, as in the case of the total sample of nine populations from four species ($r=0.891$, $P<0.001$), and it could be lower and insignificant ($r=0.479$, $P=0.061$), as in the case of families from the three populations of *C. japonica*.

The phenetic estimates of phylogenetic relationships are independent of Hardy-Weinberg equilibrium (Bussel 1999). Conversely, RAPD-based genetic estimates are dependent on this assumption. Therefore, the high correspondence between phenetic and genetic estimates obtained for all populations in the RAPD-based dataset may indicate that (1) the underlying assumption of Hardy-Weinberg equilibrium for calculation of genetic estimates do not influence the phylogenetic relationships revealed in this study or (2) the studied populations are at an equilibrium. We believe that the first alternative is the most likely since *C. cathayensis* and *C. thibetica* are supposed to have a mixed mating system and, furthermore, we have found some indications of interspecific hybridisation between *C. speciosa* and *C. cathayensis*. Both of these phenomena are known to cause deviations from the Hardy-Weinberg equilibrium. Violation of the equilibrium can, however, not be confirmed directly in this study

since we did not have access to precise genetic information about the isozyme markers that would enable the calculation of fixation indices.

The most important discrepancy between the dendrograms of phylogenetic relationships between populations was observed for *C. speciosa*. Two populations of this species (P9802 and P9805) clustered together with a third population (P9801) in the UPGMA dendrograms, but in the neighbour-joining based dendrogram they were intermediate between the *C. cathayensis* (P9803, P9804) – *C. thibetica* (P9806) cluster on one hand and the *C. japonica* (P9701, P9702, P9724) – *C. speciosa* (P9801) cluster on the other hand (Fig. 2). However, based on a more detailed analysis of genetic relationships between families in these two populations and in populations P9801, P9803 and P9804, spontaneous interspecific hybridisation between *C. cathayensis* and *C. speciosa* can be suspected (manuscript in preparation). This is probably the reason for the discrepancies observed between the two different representations.

The phylogenetic relationships between species of *Chaenomeles*, inferred from previous studies based on plant accessions from botanical gardens (Bartish et al. 1999b, Garkava et al. 2000), were completely congruent with the present results based on samples of naturally growing plants. This finding makes it possible to infer that the possible spontaneous hybridisations among plants cultivated in botanical gardens have not changed the taxonomic status of the species and that the taxonomic identities of the previously analysed accessions from several botanical gardens (Bartish et al. 1999b, Garkava et al. 2000) were correct.

Gene diversity

In accordance with previous reports (Liu and Furnier 1993; Peakall et al. 1995; Aagaard et al. 1998), our results showed that, in general, RAPDs and isozymes reveal similar patterns of gene diversity and that these markers probably experience similar evolutionary forces in *Chaenomeles*. Estimates of polymorphism of both RAPD and isozymes markers were very high and similar for the whole genus (93.7% and 92.0%, respectively, Table 2). Thus, relatively few molecular markers were fixed in the sampled populations.

Contrasting gene diversity estimates were obtained for *C. japonica* and *C. speciosa* on one hand and for *C. cathayensis* and *C. thibetica* on the other hand (Table 2). This is rather common in plants when congeneric species have contrasting mating systems (Barrett 1989; Gottlieb 1973). Our findings thus imply that outcrossing may be the prevailing mating system for *C. japonica* and *C. speciosa*, whereas *C. cathayensis* and *C. thibetica* may have a mixed mating system. Both selfing and biparental inbreeding may be suggested for the latter species to explain the observed distribution of variance, but self-incompatibility has been suspected for *C. cathayensis* (Weber 1964). Therefore, biparental inbreeding seems to be a more probable mating system for this species. Direct investigation of the mating systems within *Chaenomeles* would enable us to draw more convincing conclusions. Since our *C. cathayensis* and *C. thibetica* plant material is still juvenile, we have not yet been able to perform the necessary study. Using the more conventional method to calculate H_{pop}-values (i.e. including only markers which are polymorphic within species) yielded less differentiation between *C. japonica* (0.254) and *C. speciosa* (0.275) on the one hand, and *C. cathayensis* (0.214) and *C. thibetica* (0.238) on the other hand (results not shown in further detail). Obviously, the inclusion of monomorphic markers into our main set of calculations, according to Liu and Furnier (1993), instead emphasizes the homogeneity of the latter two species.

Shannon's index gene diversity estimates obtained with isozymes and RAPDs separately, for populations of *C. japonica* and *C. speciosa* were relatively similar within each species (Table 2). By contrast, diversity estimates based on RAPD markers showed considerably more differentiation between populations of *C. cathayensis* than did isozymes. Furthermore, a relative difference in gene diversity estimates between populations of *C. japonica* and *C. speciosa* on one hand and *C. cathayensis* and *C. thibetica* on the other was much more pronounced for isozyme markers, with two- to threefold higher relative differences (Table 2). This may indicate balancing selection on some of the isozyme markers in the presumably partly inbreeding species *C. cathayensis* and *C. thibetica*. A similar phenomenon has been suggested for oysters by Karl and Avise (1992) based on gene diversity estimates with isozymes and nuclear RFLP.

Relative gene diversity estimates derived by the Lynch and Milligan formula and by Shannon's index, respectively, were similar. The assumption of Hardy-Weinberg equilibrium therefore did not seem to influence the relative estimates when RAPD null-allele frequencies were used in the diversity analyses for the total sample of populations. This is in accordance with our findings in the phylogenetic analysis.

Partitioning of components of molecular variance

G'-statistics and *Φ*-statistics are analogous to *F*-statistics and can be used to analyse haplotypic diversity of molecular markers at different levels of hierarchial subdivisions. We found that the between-population estimates from *Φ*-statistics were always lower than corresponding estimates from *G'*-statistics (Table 4). In one case we obtained a particularly low estimate of the between-population component of total variance (*C. cathayensis*, RAPD dataset, Table 4). However the *G'*-statistics and the *Φ*-statistics were highly correlated (RAPDs: r=0.805, P<0.05; isozymes r=0.983, P<0.01).

Isozyme analysis (reviewed by Hamrick and Godt 1989) and RAPD analysis (Bartish et al. 1999a) have shown that outcrossing species, in general, retain most of their genetic variability within populations. We found that within-population components of total variability for *C.*

japonica and *C. speciosa* were always higher than between-population components, independent of kind of molecular markers or method for statistical evaluation. This is in agreement with the general result for outcrossing species. In contrast, *C. cathayensis* almost always yielded comparatively higher values for between-population components. This may indicate a mixed mating system and partial inbreeding in this species, as has also been suspected from the overall estimates of gene diversity.

Deviation of some isozyme markers from the neutral model

Contrasting patterns of genetic differentiation between conspecific populations have been previously revealed in plants as a result of comparison between isozyme and DNA markers (Raybould et al. 1996; Latta and Mitton 1997). In both these studies, DNA markers (RFLP or RAPD) suggested a higher rate of differentiation between populations than did isozymes. Balancing selection on some of the isozyme markers was brought forward as a possible explanation.

In our study, a comparative analysis of population structure within *C. japonica* by RAPD and isozyme markers also revealed substantial differences between the methods. However, isozyme markers were, in general, more differentiated than RAPDs. If we assume neutrality of RAPDs, then diversifying selection acting on some of the isozyme markers could be suspected. A similar explanation has been suggested by Le Corre et al. (1997) who compared RAPD and isozyme diversity estimates among populations of sessile oak. This is also in accordance with the general prediction made by Lewontin and Krakauer (1973) and the diversifying selection found to be acting on some plant isozymes (Lönn 1993).

The component of diversity between populations derived from Shannon's index for individual isozyme and RAPD markers was useful to find markers that deviated significantly from the parametric mean for the total set of markers (Table 5). Such markers, especially isozymes, may be involved in the adaptation to environmental stress factors (Lönn 1993). Alternatively, none of these markers represent genes directly involved in adaptation to local environment, but they are "hitchhiking" with such genes through selection due to linkage disequilibrium (Aquadro et al. 1994). The association between most of the deviating markers to such a high extent that 4 of 5 markers within *C. japonica* and all 3 markers within *C. speciosa* may be associated with the same linkage groups supports such a scenario. A note of precaution should be made however. Since our confidence interval for distribution of $G'_{pop/sp}(i)$ values accounts for sampling variation only and not for stochastic variation (Slatkin and Arter 1991), it may be incorrect to assume that selection is the only possible source of significant deviations. Difference in mutation rates between markers should also be taken into account.

Acknowledgements We are most grateful to Dr. Munetaka Sugiyama, Botanical Gardens, University of Tokyo, Japan, and Prof. Roland von Bothmer, Department of Plant Breeding, Swedish University of Agricultural Sciences, Svalöv, Sweden, who collected seeds of *C. japonica*. We are likewise very grateful to Prof. Hong De Yuan and Dr. Hong Yu, Institute of Botany, Chinese Academy of Sciences, Beijing, China, for providing seeds of the Chinese species. IVB and LPG obtained grants from the Swedish Institute and the Swedish Royal Academy of Sciences. This study was also carried out with financial support from the Commission of the European Communities, Agriculture and Fisheries (FAIR) specific RTD programme, CT 97–3894, 'Japanese quince (*Chaenomeles japonica*) – a new European fruit crop for production of juice, flavour and fibre'. It does not necessarily reflect its views and in no way anticipates the Commission's future policy in this area. The experiments performed comply with current Swedish laws.

References

Aagaard JE, Krutovskii KV, Strauss SH (1998) RAPDs and allozymes exhibit similar levels of diversity and differentiation among populations and races of Douglas-fir. Heredity 81: 69–78

Anonymous (1989) Medicinal plants in China – a selection of 150 commonly used species, WHO Regional Publ, Western Pacific Ser No. 2, Manila, pp 72–73

Aquadro CF, Begun DJ, Kindahl EC (1994) Selection, recombination, and DNA polymorphism in *Drosophila*. In: Golding B (ed) Non-neutral evolution: theories and molecular data. Chapman and Hall, New York, pp 46–56

Ayres DR, Ryan FJ (1999) Genetic diversity and structure of the narrow endemic *Wyethia reticulata* and its congener *W. bolanderi* (Asteraceae) using RAPD and allozyme techniques. Am J Bot 86:344–353

Barrett SCH (1989) Mating system evolution and speciation in heterostylous plants. In: Otte D, Endler JA (eds) Speciation and its consequences. Sinauer, Sunderland, Mass., pp 257–283

Bartish IV, Jeppsson N, Nybom H (1999a) Population genetic structure in the dioecious pioneer plant species *Hippophae rhamnoides* investigated by random amplified polymorphic DNA (RAPD) markers. Mol Ecol 8:791–802

Bartish IV, Rumpunen K, Nybom H (1999b) Genetic diversity in *Chaenomeles* (Rosaceae) revealed by RAPD analysis. Plant Syst Evol 214:131–145

Bartish IV, Jeppsson N, Bartish GI, Lu R, Nybom H (1999c) Inter- and intraspecific genetic variation in *Hippophae* (Elaeagnaceae) investigated by RAPD markers In: Jeppson N (ed) Genetic variation and fruit quality in sea buckthorn and black chokeberry. PhD thesis, Swedish University of Agricultural Sciences, Agraria 199, Alnarp

Berry A, Kreitman M (1993) Molecular analysis of an allozyme cline: alcohol dehydrogenase in *Drosophila melanogaster* on the East Coast of North America. Genetics 134:869–893

Bussell JD (1999) The distribution of random amplified polymorphic DNA (RAPD) diversity amongst populations of *Isotoma petraea* (Lobeliaceae). Mol Ecol 8:775–789

Chan KF, Sun M (1997) Genetic diversity and relationships detected by isozyme and RAPD analysis of crop and wild species of *Amaranthus*. Theor Appl Genet 95:865–873

Felsenstein J (1993) PHYLIP (Phylogeny interference package) version 3.5c. Distributed by the author. Department of Genetics, University of Washington, Seattle, Wash.

Garkava LP, Rumpunen K, Bartish IV (2000) Genetic relationships in *Chaenomeles* (Rosaceae) revealed by isozyme analysis. Sci Hortic 85:21–35

Gottlieb LD (1973) Genetic differentiation, sympatric speciation and the origin of a diploid species of *Stephanomeria*. Am J Bot 60:545–553

Hamrick JL, Godt MJW (1989) Allozyme diversity in plant species. In: Brown AHD, Clegg MT, Kahler AL, Weir BS (eds) Plant population genetics, breeding, and genetic resources. Sinauer, Sunderland, Mass., pp 43–63

Hancock JF Jr, Bringhurst RS (1979) Ecological differentiation in perennial octoploid species of *Fragaria*. Am J Bot 66: 367–375

Harrison RE, Luby JJ, Furnier GR, Hancock JF (1997) Morphological and molecular variation among populations of octoploid *Fragaria virginiana* and *F. chiloensis* (Rosaceae) from North America. Am J Bot 84:612–620

Heun M, Murphy JP, Phillips TD (1994) A comparison of RAPD and isozyme analyses for determining the genetic relationships among *Avena sterilis* L. accessions. Theor Appl Genet 87: 689–696

Huff DR, Peakall R, Smouse PE (1993) RAPD variation within and among natural populations of outcrossing buffalograss [*Buchloe dactyloides* (Nutt.) Engelm.]. Theor Appl Genet 86: 927–934

Isabel N, Beaulieu J, Theriault P, Bousquet J (1999) Direct evidence for biased gene diversity estimates from dominant random amplified polymorphic DNA (RAPD) fingerprints. Mol Ecol 8:477–483

Karl SA, Avise JC (1992) Balancing selection at allozyme loci in oysters: implications from nuclear RFLPs. Science 256:100–102

Latta RC, Mitton JB (1997) A comparison of population differentiation across four classes of gene marker in Limber pine (*Pinus flexilis* James). Genetics 146:1153–1163

Le Corre V, Dumolin-Lapegue S, Kremer A (1997) Genetic variation at allozyme and RAPD loci in sessile oak *Quercus petraea* (Matt.) Liebl.: the role of history and geography. Mol Ecol 6:519–529

Lewontin RC, Krakauer J (1973) Distribution of gene frequency as a test of the theory of the selective neutrality of polymorphisms. Genetics 74:175–195

Lifante ZD, Aguinagalde I (1996) The use of random amplified polymorphic DNA (RAPD) markers from the study of taxonomical relationships among species of *Asphodelus* sect. *Verinea* (Asphodelaceae). Am J Bot 83:949–953

Liu Z, Furnier GR (1993) Comparison of allozyme, RFLP, and RAPD markers for revealing genetic variation within and between trembling aspen and bigtooth aspen. Theor Appl Genet 87:97–105

Lynch M, Milligan BG (1994) Analysis of population genetic structure with RAPD markers. Mol Ecol 3:91–99

Lönn M (1993) Genetic structure and allozyme-microhabitat associations in *Bromus hordeaceus*. Oikos 68:99–106

Monaghan BG, Halloran GM (1996) RAPD variation within and between natural populations of morama [*Tylosema esculentum* (Burchell) Schrieber] in Southern Africa. S Afr J Bot 62: 287–291

Nei M (1972) Genetic distance between populations. Am Nat 106:283–292

Peakall R, Smouse PE, Huff DR (1995) Evolutionary implications of allozyme and RAPD variation in diploid populations of dioecious buffalograss *Buchloe dactyloides*. Mol Ecol 4: 135–147

Phipps JB, Smith PG, Rohrer JR (1990) A checklist of the subfamily Maloideae (Rosaceae). Can J Bot 68:2209–2269

Raybould AF, Mogg RJ, Clarke RT (1996) The genetic structure of *Beta vulgaris* ssp. *maritima* (sea beet) populations: RFLPs and isozymes show different patterns of gene flow. Heredity 77:245–250

Rohlf FJ (1997) NTSYS/pc. Numerical taxonomy and multivariate analysis system. Version 1.80. Exeter Software, Setauket, N.Y.

Rumpunen K, Kviklys D, Kaufmane E, Garkava L (1999) Breeding *Chaenomeles* – a new aromatic fruit crop. In: Tobutt KR, Alston FH (eds) Proc Eucarpia Symp Fruit Breed Genet. Acta Hortic 484:211–216

Slatkin M, Arter HE (1991) Spatial autocorrelation methods in population genetics. Am Nat 138:499–517

Staub JE, Box J, Meglic V, Horejsi T, McCreight JD (1997) Comparison of isozyme and random amplified polymorphic DNA data for determining intraspecific variation in *Cucumis*. Genet Resour Crop Evol 44:257–269

Vallejos CE (1983) Enzyme activity staining. In: Tanksley SD, Orton TJ (eds) Isoenzymes in plant genetics and breeding, Part A. Elsevier Science Publ, Amsterdam, pp 469–516

Wang, JX, Wang XL, Guan ZG, Teng ZQ (1998) Preliminary investigation and classification of *Chaenomeles* cultivars (in Chinese). J Beijing Forestry Univ 20:123–125

Weber C (1963) Cultivars in the genus *Chaenomeles*. Arnoldia 23:17–75

Weber C (1964) The genus *Chaenomeles* (Rosaceae). J Arnold Arbor 45:161–205, 302–345

Wendel JF, Weeden NF (1989) Visualization and interpretation of plant isozymes. In: Soltis DE, Soltis PS (eds) Isozymes in plant biology. Dioscorides Press, London, UK, pp 5–45

Yü TT, Kuan KC (1963) Taxa nova Rosacearum sinicarum I. Act Phytotaxon Sin 8:214–220