Are rhododendron hybrids distinguishable on the basis of morphology and microsatellite polymorphism?

Are rhododendron hybrids distinguishable on the basis of morphology and microsatellite polymorphism?

Matteo Caser[a], Aziz Akkak[b] and Valentina Scariot[a]

[a]Department of Agronomy, Forest and Land Management, University of Turin, via L. da Vinci 44, 10095 Grugliasco (TO), Italy.

[b]Department of Arboriculture and Pomology, University of Turin, via L. da Vinci 44, 10095 Grugliasco (TO), Italy; current address: Department of Agro-Environmental Science, Chemistry and Plant Protection, via Napoli 25, 71100 Foggia, Italy.

Corresponding author to which the proofs should be sent: Matteo Caser

Phone +390116708935; Fax +390116708798; e-mail: matteo.caser@unito.it

Department of Agronomy, Forest and Land Management, University of Turin, via L. da Vinci 44, 10095 Grugliasco (TO), Italy.

Abstract

Sequence Tagged Microsatellite Sites (STMSs) and morphological traits markers were used to evaluate 33 rhododendron germplasm for genetic diversity assessment and discrimination power. The average genetic diversity estimates were 0.724 (morphological traits) and 0.174 (STMSs) markers data sets. The Shannon index was higher for morphological traits (1.797) than STMS (0.302). The correlation coefficients obtained by the Mantel matrix correspondence test, which was used to compare the cophenetic matrices for the two markers, showed that estimated values of relationships given for morphological and STMS were not significantly related ($p > 0.05$). The data set from STMS, supported by the total Probability of identity (1.13×10^{-9}) and total Paternity exclusion probability (0.9999), allowed all accessions to be uniquely identified. In summary, STMS marker proved to be an efficient tool in assessing the genetic variability among old broad leaf rhododendron genotypes. The pattern of variation appeared to be consistent, and it can be used for germplasm conservation and management for restoration of historical genetic resources.

Key words: biodiversity, *Rhododendron* spp., microsatellites, Principal Coordinate Analysis, Principal Component Analysis, cluster analysis, morphological traits

1. Introduction

Rhododendrons are one of the most popular landscape plants in the Northern Hemisphere. The genus belongs to the *Ericaceae* family and contains over 1,000 species, sorted in 8 subgenera: *Rhododendron*, *Hymenanthes*, *Pentanthera*, *Tsutsusi*, *Azaleastrum*, *Therorhodion*, *Mumeazalea* and *Candidastrum* (Chamberlain et al., 1996). The basic chromosome number is 13 and most of the species and hybrids are diploid (Väinölä, 2000) but natural polyploids (tetraploids, octoploids and dodecaploids) can be found (Janaki Ammal et al., 1950). Besides the immense number of species and cultivars derived by intra and inter subgeneric hybridization (Contreras et al., 2007), the diversity of this genus is also consequence of the plurality of its habitats. Most of the species are present in Asia, in North America and in Australia, while in Europe only six species (*R. hirsutum* L., *R. ferrugineum* L., *R. palustre* sbp. *palustre* L., *R. myrtifolium* K., *R. lapponicum* L. and *R. ponticum* L.) are indigenous.

During the end of the XVIII and the beginning of the XIX century, many species and new hybrids were introduced from Asia to Europe, especially in England and Germany, by means of the plant hunters. The hybridization activity reached remarkable levels and rhododendrons became the most popular flowering, broadleaf evergreens and ornamental woody plant, especially in northern locations with cold winters (Väinölä, 2000). As for most of the cultivated tree species, the identification of rhododendron hybrids is convoluted. The presence of numerous species and the wide geographical distribution, together with the high level of interspecific hybridization, make genetic relationships within the genus confused. The same accession name could have been accidentally given to different genotypes or one genotype can have several synonyms. In addition, the parentage of most hybrids is unknown and their relationships with horticultural groups is often not available.

The study of genetic diversity of old cultivated plants, recovered in private and public gardens, as source of desirable genes is of current interest (Khlestkina et al., 2004) and the development of

new hybrids adapt to different conditions, such as Alpine area (Nybom et al., 2004; Kjolner et al., 2004) is particularly intended. Morphological and molecular characterization could be of help for preserving and using these genetic resources and, through the study of pre-breeding and breeding germplasm diversity, for determining unique and distinct traits.

Among molecular markers, Sequenced Tagged Microsatellite Sites (STMSs) are considered to be neutral markers and more informative for characterizing germplasm collections thanks to several characteristics, including abundance in eukaryotic genomes, high levels of polymorphism, Mendelian inheritance, co-dominance, and locus specificity (Merdinoglu et al., 2005; Scariot et al., 2006; Marchese et al., 2007). However, the development of STMSs is laborious and at present only a few STMS primers were designed in *Rhododendron* (Dunemann et al., 1998; Kameyama et al., 2002; Dendauw et al., 2001). Numerous types of other molecular markers have been developed and used for phylogenetic studies and cultivar fingerprinting, such as *trn*K and *mat*K (Kron, 1997; Kurashige et al., 1998, 2001), nuclear ITS sequences (Gao et al., 2002; Tsai et al., 2003), Random Amplified Polymorphic DNA (RAPD; Scariot et al., 2007; Lanying et al., 2008), and EST derived markers (De Keyser et al., 2009). Only a few refer to subgenus *Hymenanthes* and *Rhododendron*. Jin et al. (2006) developed ISSR markers in *R. fortunei* L. and Contreras et al. (2007) and Wei et al. (2006) studied species diversity in *R. ponticum* L. and *R. catawbiense* M. using AFLP and EST derived markers.

This study evaluated (1) the usefulness of 4 STMS markers for establishing relationships in old broad leaf rhododendrons, and (2) the relatedness among cultivated hybrids (many of them never previously DNA-typed) and seven species belonging to subgenus *Hymenanthes* section *Ponticum* (*R. griffithianum* H., *R. fortunei* L., *R. ponticum* L., *R. maximum* L., *R. catawbiense* M., *R. caucasicum* P and *R. arboreum* A.) and *R. edgeworthii* H. (subgenus *Rhododendron*, section *Rhododendron*) by means of morphological and DNA markers.

2. Materials and methods

102

103 *2.1. Plant material, DNA isolation and STMS analysis*

104

105 Thirty-three rhododendron accessions (8 species and 25 cultivars) of the genus *Rhododendron*
106 subgenus *Hymenanthes* section *Ponticum* and subgenus *Rhododendron* section *Rhododendron*
107 located in the Burcina Park and Villa Taranto (Northern Italy), and in the Missouri Botanical
108 Garden (U.S.A.) were selected for this study (Table 1). Genomic DNA was extracted as described
109 by Thomas et al. (1993), from approximately 0.20 g leaf tissue. The DNA extraction buffer (CNB;
110 Crude Nuclei chromatin Buffer) contained 2.5% PVP (polyvinyl pyrrolidone K40), 0.2 M Tris-HCl
111 pH 7.6, 0.05 M Na_2EDTA pH 8, 0.25 M NaCl and 2.5% β-mercaptoethanol. This method yielded
112 up to 240 ng/µl of genomic DNA per extraction. DNA quality was examined by electrophoresis on
113 a 0.8% agarose gel and DNA concentration was quantified by means of a spectrophotometer. Four
114 STMS primer sets developed by Dunemann et al. (1998), labelled with a specific fluorochrome (6-
115 FAM or HEX), were used: GA211, RDC46, RDC45 and RDC27. Amplification reactions were
116 carried out in a final volume of 20 µl containing 50 ng template DNA, 2 µl 10X PCR reaction
117 buffer (100mM Tris-HCl, pH 8.3, 500 mM KCl), 1.5 mM $MgCl_2$, 200 µM dNTPs, 0.5 µM of each
118 primer and 0.5 U AmpliTaq Gold® DNA polymerase (Applied Biosystems, Foster City, CA). The
119 PCR amplifications were performed using the following temperature program: initial step of 9 min
120 at 95°C, followed by 28 cycles of 30 sec at 95°C, 45 sec at 50°C, 1 min 30 sec at 72°C, with a final
121 extension step of 45 min at 72°C. One µl of a mix containing amplification products was added to 3
122 µl of a mix containing 5:2:1 parts of formamide, GeneScan-500 Liz size standard and loading dye
123 (25 mM EDTA, 50 $mgmL^{-1}$ blue dextran). Fluorescent samples were denatured at 95°C for 5 min
124 and detected on a sequencing gel (5% acrylamide, 6 M Urea, 1X TBE buffer) using an ABI-
125 PRISM®377 DNA sequencer (Applied Biosystems, Foster City, CA).

126

127 *2.2. Morphological characterization*

All the plants were described by means of 11 morphological traits referring to habitus, flower and leaf (Table 2) as previously assessed by Remotti et al. (2003). Three flowers in full bloom and three mature leaves were measured on each plant.

Multistate characters were treated as follows: leaves shapes = 0 (lanceolate), 1 (oval-lanceolate), 2 (spatulate-lanceolate), 3 (oval), and 4 (spatulate); leaves surface = 0 (flat), 1 (convex), 2 (concave), 3 (corrugated), and 4 (smooth); plant habit was considered = 0 (straggly shrub), 1 (thick shrub), 2 (straggly tree), and 3 (thick tree); blooming time = 0 (early; from 1st to 30st April), 1 (semi-early; from 1st to 15st May), 2 (semi-late; from 15st to 30st May), and 3 (late; from 1st to 15st Juny); bloom density = 0 (low; distance between truss > 50 cm), 1 (medium; distance between truss ranged from 25 to 50 cm), and 2 (high; distance between truss < 25 cm); truss shapes = 0 (hemispheric), 1 (conic-hemispheric), and 2 (conic).

2.3. Data analysis

The presence or absence of fragments amplified by STMS primers was coded by 1 or 0 respectively and scored as a binary data matrix. Allele frequencies (Fig. 1), number of effective alleles, Shannon's index, diversity (h), unbiased diversity (uh) were calculated using GenAlEx 6.3 (Peakall and Smouse, 2006). Genetic distances based on STMS data were computed according to Nei (1978). Cluster analysis was performed using Neighbor-joining method, on arithmetic means (Sneath and Sokal, 1973), by means of the TREECON software (Van de Peer and De Wachter, 1994). This software, was also used to estimate the statistical stability of the branches in the tree by bootstrap analysis with 1000 replicates.

Principal Component Analysis (PCA) was performed on morphological traits. The first two axes were plotted according to the extracted Eigen vectors, using the software package NTSYS-pc version 2.1 (Applied Biostatisctics Inc., NY, USA).

With the purpose to compare morphological and genetic distance matrices, Mantel analysis (Mantel, 1967) was executed on the distance matrices (Gower 1971; Nei 1978) using Mantel Nonparametric Test Calculator for Windows, version 2.00. The significance of the statistic was evaluated by permutations (9,999 x) and expressed as probability (Smouse et al., 1986). The software IDENTITY 1.0 (Sefc et al., 1999) was used to calculate the frequency of null alleles, paternity exclusion probability, and probability of finding 2 identical genotypes.

3. Results

3.1. Levels of polymorphism

Eight species and 25 hybrids, representing diverse taxonomic sections and horticultural groups, were analyzed in duplicate with four microsatellite markers. The presence/absence of the alleles was established using a threshold value for allele assignment, according to Esselink et al. (2003).

All four STMS loci analyzed were multiallelic, fully transferable and amplified DNA fragments in all the accession analysed. As explained in Table 3 the bands ranged from approximately 106 bp to 269 bp in size. The four STMS primer pairs amplified a total number of 71 alleles, with an average of 17.75 per locus. Value of diversity among samples for each locus (h) ranged from 0.166 (GA211) to 0.192 (RDC46). Overall unbiased diversity (uh) was only slightly higher with values ranging from 0.171 (GA211) to 0.198 (RDC46). The observed genotypes did not show a single allele per locus and no sample was considered homozygous for the purpose of computing genetic distance parameters. The number of alleles per sample at each locus ranged from 5 (‘Perspicum’) to 11 (‘Fastosum Flore Pleno’). Thirteen accessions presented at least one specific allele with a total of 19 specific alleles. In order to provide information on the differentiating capacity of each marker, the number of unique banding patterns per assay unit was calculated. Overall, for each locus, the number of unique genotypes varied between 11 (RDC45) and 22 (GA211) with a total of 68 for all

loci. The probability of identity is defined as the probability with which two randomly taken genotypes display the same STMS profile. The probability of identity for each locus (Table 3) ranged from 0.0009 for RDC27 to 0.0507 for GA211 with an average of 0.0157, whereas the total probability of identity was 1.13×10^{-9}. The estimated frequency of null alleles was positive for all loci. The total paternity exclusion was 0.999. This index was the highest for GA211 (0.752) and the lowest for RDC46 (0.402), with a mean value of 0.592.

The allele frequency distribution (Fig. 1) showed an unexpected pattern that breaches the normal distribution of random phenomena. Indeed no normal distribution of allele frequency was observed. The most frequent allele was RDC46-171, which showed frequency greater than 50%. Among the others, only three alleles (4%) showed a frequency higher than 30% (RDC45-112, GA211-152 and GA211-154). On the other hand, 20 alleles (28%) were relatively infrequent (< 5%; RDC45-106, RDC45-116, RDC45-124, RDC45-130, RDC45-135, RDC45-147, RDC45-157, RDC45-163, RDC27-151, RDC45-188, RDC45-192, RDC45-194, RDC45-224, RDC45-231, GA211-142, GA211-243, GA211-269, RDC46-109, RDC46-120 and RDC46-169).

Referring to morphology, 11 characteristics were evaluated as multistate and quantitative traits. Mean, maximum, minimum and Standard Error of the quantitative traits are shown in Table 2. A large diversity was observed, indicating a high level of variation among genotypes. Overall, only leaves length and width were correlated ($R^2 = 0.66$). The mean diversity calculated from morphological traits was 0.724 (Table 2) and, range-wise, morphological data produced larger (0.415-0.955) diversity compared to STMS (0.166-0.192). Moreover, the Shannon index was higher (1.797) than the obtained from STMSs (0.302).

Relationships between hybrids both with known or unknown parentage and their supposed reference species were evaluated calculating distance matrices with both STMS and morphologic data (data not shown). The lowest genetical distance was found between *R. ponticum* and 'Madame Boyer', while the highest between 'Fastosum flore pleno' and *R. caucasicum*, and 'Fastosum flore pleno' and 'James Marshall Brooks'. Whereas, for morphological characteristics the most

correlated accessions were ‘Everestianum’ and ‘Nigrescens’. However, all plants seem to be rather diverse both genetically and morphologically. The analysis performed with IDENTITY 1.0 indicated that ‘Madame Masson’, ‘Pink pearl’, and ‘The strategist’ shared at least one allele at each locus, suggesting a possible parentage relationship.

3.2. Morphological and molecular relatedness

STMSs were at first used to estimate phylogenetic relationships among the studied species. The PCoA (Fig. 2; 43.35 % of the total variance was explained by the first two coordinates) indicated that the distribution of the *Rhododendron* species is congruent at the subgenus level with the current classification system (Chamberlain et al., 1996; Cox and Cox, 1997): six to seven species of subgenus *Hymenanthes* section *Ponticum* (*R. griffithianum, R. fortunei, R. ponticum, R. maximum, R. catawbiense*, and *R. arboreum*) clustered together, while *R. caucasisum* was apart. For low values of Coord. 2, *R. edgeworthii* (subgenus *Rhododendron*) was separated from the others studied species.

With the purpose to visualize the relationships among the *Rhododendron* hybrids, a Principal Component Analysis (PCA) was performed on morphological traits (Fig. 3) and a cluster analysis on STMS data (Fig. 4). At first sight, the two analyses did not yield similar groupings. The PCA (Fig. 3 and Table 4) showed that the first three components accounted for 70.69%, 17.47%, and 5.99% of the variance respectively, their cumulative variance being 94.15%. On the basis of the eigenvector values for traits along the first three components (Table 4), the attributes responsible for maximum separation were (with values in parentheses): the Bloom density (8.6 e^{-04}), Leaf shape (7.7 e^{-04}), and Leaf length (0.933) along the first component (PC1), while the Blooming time (5.9 e^{-04}) and the flower diameter (0.926) along the second (PC2). The first two components were used to visualize a scatter plot (Fig. 3) in which the accessions were divided in two groups. The cultivars were mainly grouped for low values of PC1 (Group A) and the species for high (Group B). In

Group B also some hybrids were included (‘Mrs. R.S. Holford’, ‘Sappho’, Princesse Hortense, ‘Jhon Walter’, and ‘White pearl’). In the same group, for low values of PC2, a small number of hybrids with *R. griffithianum, R. edgeworthii, and R. caucasicum* were grouped. In particular, these samples differed from others by an early or semi-early blooming time and larger flower diameter.

In the dendrogram based on STMS data (Fig. 4), the 33 accessions were divided, by medium-low bootstrap values, in three main groups generally in accordance with the hybrids pedigree recorded in literature. The highest values were found between ‘Onsloweanum’ and ‘Pink Pearl’ (58 %), between ‘Madame Masson’ and ‘The strategist’ (57 %), and between ‘White Pearl’ and *R. griffithianum* (50 %). In the first group were sited the species *R. ponticum*, *R. fortunei*, and *R. griffithianum* with their related hybrids ‘Fastosum Flore Pleno’ (*R. catawbiense* x *R. ponticum*), ‘Michael Waterer’ (*R. ponticum* x *R. arboreum*), ‘Madame Masson’ (*R. catawbiense* x *R. ponticum*), ‘The strategist’ (*R. griffithianum*), and ‘White Pearl’ (*R. griffithianum* x *R. maximum*). In the second group ‘Cunningham’s White’ (*R. caucasicum* x *R. ponticum*), and ‘Prince Camille de Rohan’ (*R. maximum* x *R. caucasicum*) clustered with the species *R. caucasicum* and *R. edgeworthii*. In the third, ‘John Walter’ (*R. catawbiense* x *R. arboreum*), and ‘Everestianum’ (*R. catawbiense*) grouped with *R. maximum, R. catawbiense,* and *R. arboreum.*

With the purpose to compare the morphological and molecular data sets, Mantel’s test was performed. Via a permutation procedure (9,999 x), the correlation between the morphological (Gower 1971) and molecular (Nei 1978) matrices was tested against multiple randomisations of one of them. Results showed that data were not statistically correlated ($g = 0.2329$, critical value = 1.645 for $p = 0.05$, $r = 0.016$).

4. Discussion

4.1. Polymorphisms

In agreement with Scariot et al. (2007), the four STMSs, previously developed in 'Cunningham's White' (*R. caucasicum* x *R. ponticum*), were amplified in all the *Rhododendron* species, resulting effective in detecting polymorphism in different related subgenera. This finding enables the analysis of different species using the same STMS loci set, thus reducing the cost. Moreover, the presence of common markers across species can be of great interest in genetic mapping (Fraser et al., 2005). The mean number of alleles per locus was 17.75 and comparable to the value of 16.5 in evergreen azaleas (*Rhododendron*; Scariot et al., 2007). The most informative loci were GA211 and RDC27. They were able to discriminate the 66% and 70% of genotypes, respectively.

The ability to detect multiple alleles at a single locus makes microsatellites more appropriate for many diversity studies. In this work all the loci were suitable for establishing relationships because more than one allele was found in all the analyzed accessions. Despite the importance of the marker type in analyzing variation, the choice of statistical coefficients is a rather fundamental step in studying genetic diversity. Various diversity indices may be used to summarize the genetic difference among tested genotypes (Magurran, 2004). Among statistics, the Shannon's index is one of the most correlated as described by Jabot and Chave (2009). In this study, a relatively low value of this index (0.302) was found with STMSs compared to morphological data set. Also the mean diversity (h) was the lowest (0.174), showing that STMSs were less efficient in detecting variability in the studied germplasm collection. The unbiased diversity was higher than the corresponding diversity for all loci as found in *Nelundo* by Tian et al. (2008) and in *Brassica juncea* by Khan et al. (2008). As discussed by Powell et al. (1996) diversity estimates can be affected by several factors such as, the distribution of markers in the genome and the nature of evolutionary mechanisms underlying the variation measured. Nevertheless, the number of studied samples was too small to draw conclusions on the occurrence of null alleles, as their presence can be truly ascertained only by studying their frequency in a large population (Callen et al., 1993).

Because of few STMS information available on the studied *Rhododendron* subgenus, in this work the appropriate number of loci for discriminating analyses was checked by computing the

probability of identity, estimated from the allele frequencies in the data set. According to the frequency results, the total probability of identity for all four loci was 1.13×10^{-9}. Similar results were found for example in European grapes (5.67×10^{-9}) by Najafi et al. (2006), and in *Eucalyptus leucoxylon* (6.37×10^{-9}) by Ottewell et al. (2005). In addition, the very high combined probability of paternity exclusion (0.9999) indicated that the selected STMSs allow a high level of individual identification in *Rhododendron* (Alamerew et al., 2004; Khlestkina et al., 2004; Reis et al., 2009).

4.2. Species diversity

The evaluation of genetic variability in a genus is of fundamental importance for optimal preservation of genetic resources, collection management, and plant exploitation. According to the current, and accepted, morphology based classification systems (Chamberlain et al., 1996; Cox and Cox, 1997), the genus *Rhododendron* can be divided into eight subgenera. Plants usually recognized as "true" rhododendrons belong to the subgenera: *Hymenanthes* and *Rhododendron*. In this study, results based on STMS data (Fig. 2) confirmed the apportionment of the eight studied species (except for *R. caucasicum*) to these subgenera in agreement with Wei et al. (2006), and Lanying et al. (2008). The employed STMSs were also able to discriminate within section *Ponticum*, demonstrating close relationships between *R. ponticum* and *R. maximum* than between these two and *R. catawbiense* as previously observed analysing *matK*, *trnL-F*, *RPB2*, and intron-flanking EST gene by Milne (2004), Goetsch et al. (2005) and Wei et al. (2006), respectively. Moreover, as shown by the dendrogram (Fig. 4), *R. edgeworthii* and *R. caucasicum* were grouped together, confirming a genetically differentiation from the other species.

4.3. Hybrids diversity

This study provided useful information about the phenotypic and genetic diversity of the subgenera *Hymenanthes* and *Rhododendron*, previously poorly investigated.

Based on morphological traits (Fig. 3), most accessions were sorted in a *continuum* group and no clear pattern of differentiation could be observed in relation to the known pedigree. Only 'White Pearl' grouped with its parental species *R. griffithianum* (Accati, 2001), as confirmed by the cluster analysis on STMS (Fig. 4). PCA showed that the traits, which yielded the maximum separation between the accessions studied were Blooming time, Bloom density, Leaf shape, Leaf length, and Flower diameter. The wide variation observed in these traits could be used for improving ornamental characteristics through breeding.

STMS data provided useful information for parentage determination and can be a reliable method to assess the genetic relatedness of the rhododendron accessions in less time than with morphological traits. In the three different branches of the dendrogram almost all the hybrids with known parentage were correctly grouped, except for 'Purity' (*R. edgeworthii* x *R. formosum*) and 'Lady Eleanor Cathcart' (*R. maximum* x *R. arboreum*). Concerning the hybrids with unknown pedigree, this analysis suggested relationships between accessions within the same group. These close relationships are generally in agreement with horticultural classification (Accati, 2001) and IDENTITY analysis. Therefore, on the basis of this study, some hypotheses about the origins of these hybrids can be carried out. In particular, the close relationships between 'Madame Boyer', 'Lady Rolle', and 'Sappho' with *R. ponticum*, between 'Perspicuum' and *R. caucasicum*, between 'Madame Patti' and *R. maximum*, and between 'Memoire de Dominique Vervaene' and *R. catawbiense* suggested possible parentage relationships. Then, STMSs data evaluated in this work appear to represent an useful instrument to facilitate the management of *Rhododendron* germplasm.

Because of no significant correlation between the morphological and molecular data was detected, a PCoA was also performed on the combined data (data not shown). However, this analysis did not add any further information for revealing the cultivar's pedigree and fill the lack of breeding information cumulated along the centuries.

Information on old broad leaves *Rhododendron* germplasm, provided here for the first time, can be used for breeding purposes. In the studied genotypes, as shown by Shannon index, the wide morphological variability is related to a restricted genetic pool. This result confirmed that in the last centuries only a few parental species were employed in the breeding programs. Thus, based on the known interspecific cross attitude within the genus *Rhododendron* demonstrated by Kurashige et al. (1998 and 2001) and Gao et al. (2002), new species could be used to obtain new cultivars.

Acknowledgments

We are grateful to Elena Accati and Dora Remotti for their precious advice, Roberto Botta and his staff for their support and skilful assistance, Frank Dunemann for providing STMS markers and the Missouri Botanical Garden (U.S.A.) for supplying germplasm.

References

Accati, E., 2001. Il rododendro. Gli antichi ibridi del Parco Burcina di Pollone. Supplemento al n.26 di Quaderni della Regione Piemonte-Agricoltura, pp. 9-241.

Alamerew, S., Chebotar, S., Huang, X., Roder, M., Borner, A., 2004. Genetic diversity in Ethiopian hexaploid and tetraploid wheat germplasm assessed by microsatellite markers. Gen. Res. Crop Evol. 51, 559-567.

Callen, D.F., Thompson, A.D., Shen, Y., Philips, H.A., Richards, R.I., Mulley, J.C., Sutherland, G.R., 1993. Incidence and origin of "null" alleles in the (AC)n microsatellite markers. Am. J. Hum. Genet. 52, 922-927.

Chamberlain, D.F., Hyam, R., Argent, G., Fairweather, G., Walter, K.S., 1996. The genus *Rhododendron*. Its classification and synonymy. Edinburgh Royal Botanical Garden.

Contreras, R.N., Ranney, T.G., Tallury, S.P., 2007. Reproductive behaviour of diploid and allotetraploid *Rhododendron* L. 'fragrant affinity'. HortScience 42, 31-34.

Cox, P.A., Cox, K.N.E., 1997. The encyclopedia of *Rhododendron* species. Glendoick Publishing, Perth, U.K.

De Keyser, E., Riek de, J., Bockstaele van, E., 2009. Discovery of species-wide EST-derived markers in *Rhododendron* by intron-flanking primer design. Mol. Breed. 23, 171-178.

Dendauw, J., De Riek, J., Arens, P., Van Bockstaele, E., Vosman, B., and De Loose, M., 2001. Development of sequenced tagged microsatellite site (STMS) markers in azalea. Acta Hortic. 546, 193-197.

Dunemann, F., Kahnau, R., Stange, I., 1998. Analysis of complex leaf and flower characters in *Rhododendron* using a molecular linkage map. Theor. Appl. Genet. 98, 1146-1155.

Esselink, G.D., Smulders, M.J.M., and Vosman, B. 2003. Identifiction of cut rose (*Rosa hybrida*) and rootstock varieties using robust sequence tagged microsatellite site markers. Theor. Appl. Genet. 106, 277-286.

388 Fraser, L.G., McNeilage, M.A., Tsang, G.K., Harvey, C.F., De Silva, H.N., 2005. Cross-species
389 amplification of microsatellite loci within the dioecious, polyploidy genus *Actinidia*
390 (Actinidiaceae). Theor. Appl. Genet. 112, 149-157.
391 Gao, L.M., Li, D.Z., Zhang, C.Q., Yang, L.B., 2002. Infrageneric and sectional relationships in the
392 genus *Rhododendron* (Ericaceae) inferred from ITS sequence data. Acta Bot. Sin. 44, 1351-
393 1356.
394 Goetsch, L., Eckert, A.J., Hall, B.D., 2005. The molecular systematic of *Rhododendron* (Ericaceae):
395 a phylogeny based upon RPB2 gene sequences. Syst. Bot. 30, 616-626.
396 Gower, J.C., 1971. A general coefficient of similarity and some of its properties. Biometrics 27,
397 857–874.
398 Jabot, F., Chave, J., 2009. Inferring the parameters of the neutral theory of biodiversity using
399 phylogenetic information and implications for tropical forests. Ecol. Let. 12, 239-248.
400 Janaki Ammal, E.K., Enoch, I.C., Bridgwater, M., 1950. Chromosome numbers in species of
401 *Rhododendron*. The Rhododendron Yearbook. Royal Hortic. Soc., pp. 78-91.
402 Jin, Z.X., Li, J.M., Gu, Q.P., 2006. Genetic diversity in the natural populations of *Rhododendron*
403 *fortunei* revealed by ISSR molecular markers. Acta Hortic. Sinica 33, 1263-1267.
404 Kameyama, Y., Isagi, Y., Nakagoshi, N., 2002. Related structure in *Rhododendron metternichii* var.
405 *hondoense* revealed by microsatellite analysis. Mol. Ecol. 11, 519-527.
406 Khan, M.A., Rabbani, M.A., Munir, M., Ajmal, S.K., Malik, M.A., 2008. Assessment of Genetic
407 Variation Within Indian Mustard (*Brassica juncea)* Germplasm Using Random Amplified
408 Polymorphic DNA Markers. J. Integr. Plant Biol. 50, 385-392.
409 Khlestkina, E.K., Than, M.H.M., Pestsova, E.G., Roder, M.S., Malyshev, S.V., Korzun, V., Borner
410 A., 2004. Mapping of 99 new microsatellite-derived loci in rye (*Secale cereale* L.) including 39
411 expressed sequence tags. Theor. Appl. Genet. 108, 1466-1472.

412 Kjolner, S., Sastad, S.M., Taberlet, P., Brochmann, C., 2004. Amplified fragment lenght
413 polymorphism versus random amplified polymorphic DNA markers: Clonal diversity in
414 *Saxifraga cernua*. Mol. Ecol. 13, 81-86.
415 Kron, K.A., 1997. Phylogenetic relationships of *Rhododendroideae* (Ericacae). Amer. J. Bot. 84,
416 973-980.
417 Kurashige, Y., Mine, M., Kobayashi, N., Handa, T., Takayanagi, K., Yukawa, T., 1998.
418 Investigation of sectional relationships in the genus *Rhododendroideae* (Ericaceae) based on
419 *matK* and *trnK* sequences. Plant Syst. Evol. 228, 1-14.
420 Kurashige, Y., Etoh, J.I., Handa, T., Takayanagi, K., Yukawa, T., 2001. Sectional relationships in
421 the genus *Rhododendron* (*Ericaceae*): evidence from matK and trnK intron sequences. Plant
422 Syst. Evol. 228, 1-14.
423 Lanying, Z., Yongqing, W., Li, Z., 2008. Genetic diversity and relationship of *Rhododendron*
424 species based on RAPD analysis. Am. Eur. J. Agric. Env. Sci. 3, 626-631.
425 Magurran, A.E., 2004. Measuring biological diversity. Blackwell, Oxford, UK, pp. 256
426 Mantel, N.A., 1967. The detection of disease clustering and a generalized regression approach.
427 Cancer Res. 27, 209-220.
428 Marchese, A., Tobutt, K.R., Raimondo, A., Motisi, A., Boskovic, R.I., Clarke, J., Caruso, T., 2007.
429 Morphological characteristics, microsatellite fingerprinting and determination of incompatibility
430 genotypes of Sicilian sweet cherry cultivars. J. Hortic. Sci. Biotech. 82, 41-48.
431 Merdinoglu, D., Butterlin, G., Bevilacqua, L., Chiquet, V., Adam-Blondon, A.F., Decroocq, S.
432 2005. Development and characterization of a large set of microsatellite markers in grapevine
433 (*Vitis vinifera* L.) suitable for multiplex PCR. Mol. Breed. 15, 349–366.
434 Milne, R.I., 2004. Phylogeny and biogeography of *Rhododendron* subsection *Pontica*, a group with
435 tertiary relict distribution. Mol. Phylog. Evol. 33, 389-401.

436 Najafi, J., Alipanah, L., Ghareyazie, B., Mohammadi, S.A., Nazari, A.H., This, P., 2006. Genetic
437 diversity of Iranian and some of European grapes revealed by microsatellite markers. Ir. J.
438 Biotech. 4, 36-44.
439 Nei, M., 1978. Estimation of average heterozygosity and genetic distance from a small number of
440 individuals. Genetics 89, 583-590.
441 Nybom, H., Esselink, G.D., Werlemark, G., Vosman, B., 2004. Microsatellite DNA marker
442 inheritance indicates preferential between two homologous genomes in polyploid and
443 hemisexual dogroses, *Rosa* L. sect. *Caninae*. Heredity 92, 139-150.
444 Ottewell, K.M., Donnellan, S.C., Moran, G.F., Paton, D.C., 2005. Multiplexed Microsatellite
445 Markers for the Genetic Analysis of *Eucalyptus leucoxylon* (Myrtaceae) and Their Utility for
446 Ecological and Breeding Studies in other Eucalyptus Species. J. Hered. 96, 445-451.
447 Powell, M.B., Nikles, D.G., Dieters, M.J., Matheson, A.C., Nikles, D.G., Harwood, C.E., Walker,
448 S.M., 1996. Genetic parameter estimates and predicted breeding values for diameter, height
449 and stem straightness of *Pinus elliotti*, *Pinus caribaea* var. *hondurensis* and their F_1 hybrid.
450 Tree improvement for sustainable tropical forestry. QFRI-IUFRO Conference 1, 169-172.
451 Peakall, R., and Smouse, P.E. 2006. GENALEX 6: genetic analysis in Excell. Population genetic
452 software for teaching and research. Mol. Ecol. Notes 6, 288-295.
453 Reis, A.M.M., Braga, A.C., Lemes, M.R., Gribel, R., Collevatti, R.G., 2009. Development and
454 characterization of micro satellite markers for the Brazil nut tree *Bertholletia excelsa* Hunmb. &
455 Bonpl. (Lecythidaceae). Mol. Ecol. Res. 9, 920-923.
456 Remotti, D., Duzyaman, E., Tuzel, Y., 2003. Identification and morpho-botanic characterization of
457 old *Rhododendron* hybrids of Burcina Park (Pollone-Biella) in Italy. Acta Hortic. 598, 73-77.
458 Scariot, V., Akkak, A., Botta, R., 2006. Characterization and genetic relationships of wild species
459 and old garden roses based on micro satellite analysis. J. Amer. Soc. Hortic. Sci. 131, 66-73.

460 Scariot, V., de Keyser, E., Handa, T., de Riek, J., 2007. Comparative study of the discriminating
461 capacity and effectiveness of AFLP, STMS and EST markers in assessing genetic relationships
462 among evergreen azaleas. Plant Breed. 126, 207-212.
463 Sefc, K.M., Regner, F., Turetschek, E., Glössl, J., Steinkellner, H., 1999. Identification of
464 microsatellite sequences in *Vitis riparia* and their applicability for genotyping of different *Vitis*
465 species. Genome 42, 367–373.
466 Smouse, P.E., Long, J.C., Sokal, R.R., 1986. Multiple regression and correlation extensions of the
467 Mantel test matrix correspondence. Syst. Zool. 35, 627-632.
468 Sneath, P.H.A., Sokal, R.R., 1973. Numerical Taxonomy. Freeman, San Francisco.
469 Thomas, M.R., Matsumoto, S., Cain, P., Scott, N.S., 1993. Repetitive DNA of grapevine: classes
470 present and sequences suitable for cultivation identification. Theor. Appl. Genet. 86, 985-990.
471 Tian, H.L., Xue, J.H., Wen, J., Mitchell, G., Zhou, S.L., 2008. Genetic diversity and relationships of
472 lotus (*Nelumbo*) cultivars based on allozyme and ISSR markers**.** Sc. Hortic. 116, 421-429.
473 Tsai, C.C., Huang, S.C., Chen, C.H., Tseng, Y.H., Huang, P.L., Tsai, S.H., Chou, C.H., 2003.
474 Genetic relationships of *Rhododendron* (Ericaceae) in Taiwan based on the sequence of the
475 internal transcribed spacer of ribosomal DNA. J. Hortic. Sci. Biotech. 78, 234-240.
476 Väinölä, A., 2000. Polyplodization and early screening of *Rhododendron* hybrids. Euphytica 112,
477 239-244.
478 Van de Peer, Y., De Wachter, Y., 1994. TREECON for Windows: A software package for the
479 construction and drawing of evolutionary trees for the Microsoft Windows environment. Comp.
480 Appl. Biosci. 10, 569–570.
481 Wei, H., Fu, Y., Arora, 2006. Utilization of intron-flanking EST-specific markers in the
482 phylogenetic analysis and parentage identification of Rhododendron species and hybrids. J.
483 Amer. Soc. Hortic. Sci. 131, 814-819.
484
485

Figure legends

Fig. 1. Allele frequency distribution of the four STMS markers (Dunemann et al. 1998) used to analyse 33 *Rhododendron* accessions.

Fig. 2. Scatter diagram of eight *Rhododendron* species obtained performing the Principal Coordinate Analysis (PCoA) on STMS data. The first two principal coordinates explain 24.18% and 19.17% of the total genetic variance, respectively.

Fig. 3. Scatter diagram of 33 *Rhododendron* accessions obtained performing the Principal Component Analysis (PCA) on 11 morphological characteristics (5 quantitative and 6 multistate) related to leaves, flowers, truss and habitus. The first two components explain 88.16% of the total variation.

Fig. 4. Cluster analysis of 33 *Rhododendron* accessions, based on STMS data. Values at the internodes indicate the percentage of bootstrap values from 1000 re-sampling cycles.

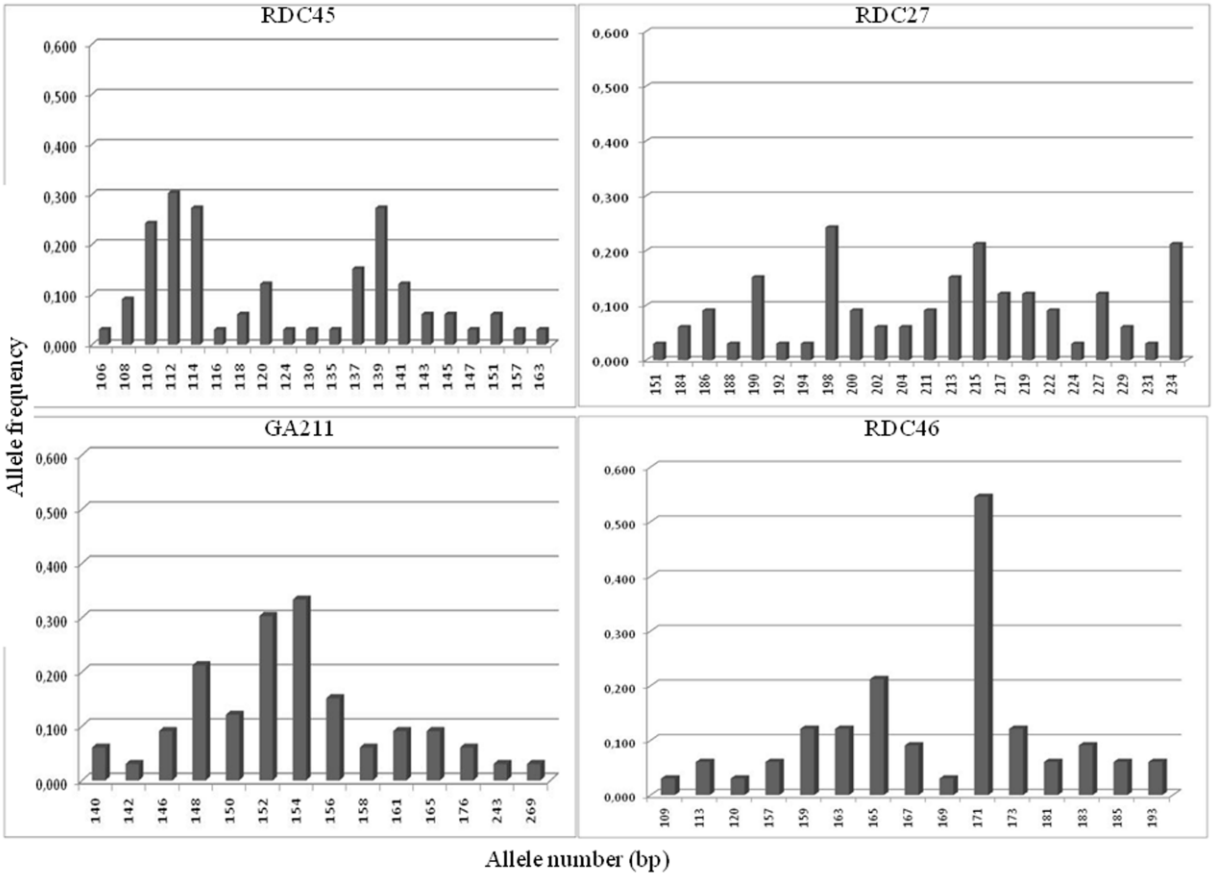


Figure 1

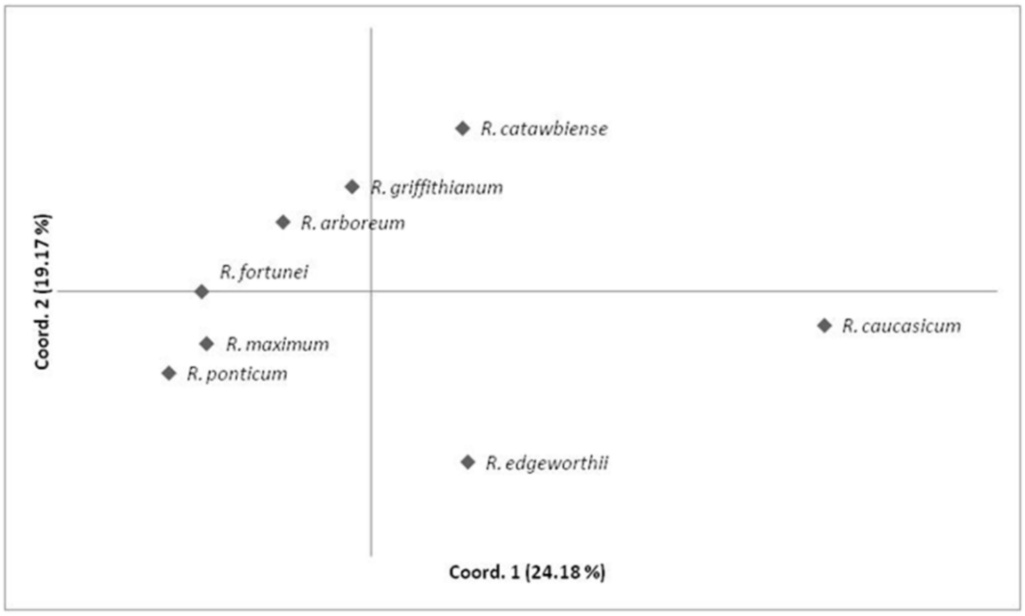
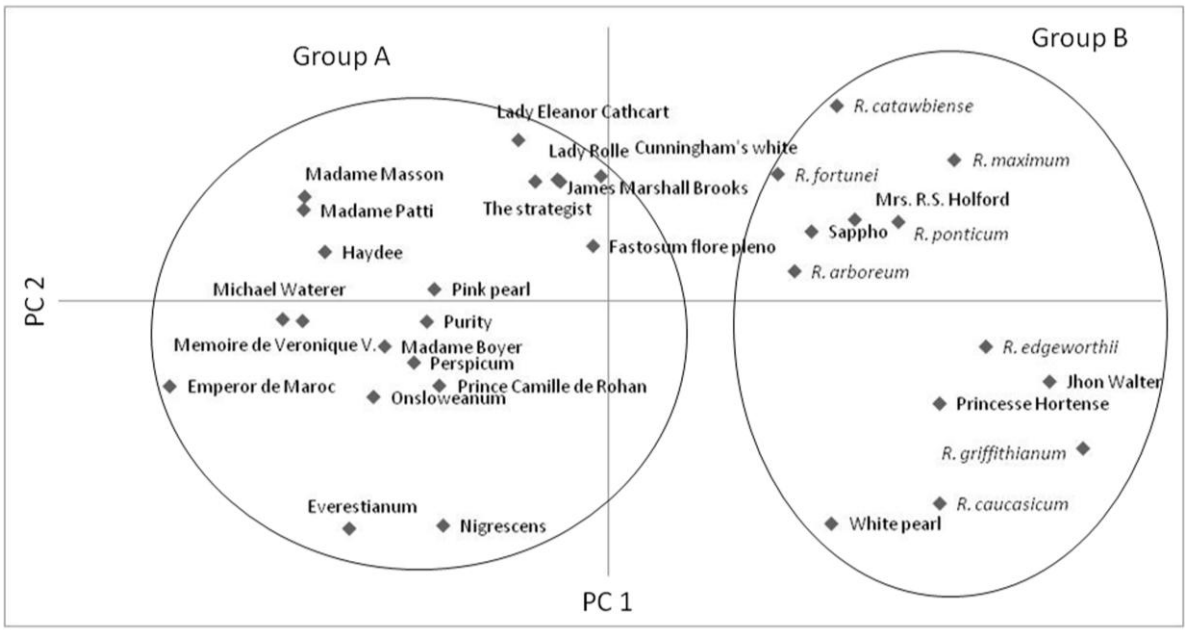


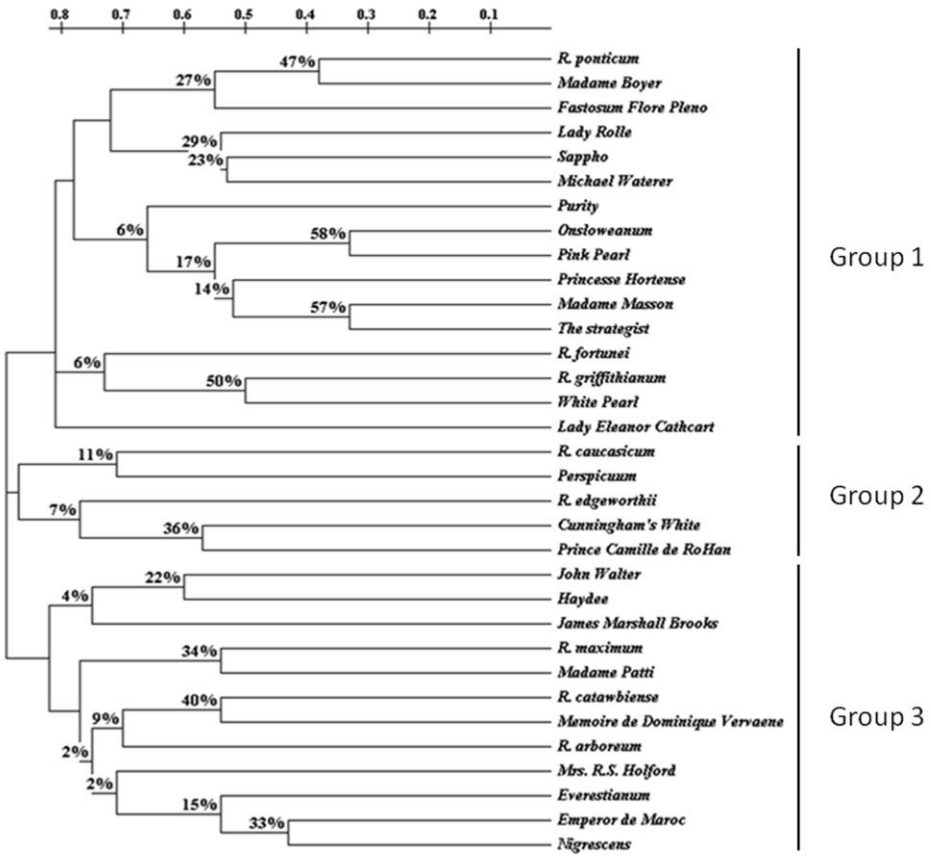
Figure 2



541

542 Figure 3

543



544

545 Figure 4

546

547