Systematic Botany (2007), 32(4): pp. 819–836

Delimiting Species Boundaries in *Rosa* Sect. *Cinnamomeae* (Rosaceae) in Eastern North America

SIMON JOLY[1] and ANNE BRUNEAU

Institut de recherche en biologie végétale, Université de Montréal, 4101 Sherbrooke Est, Montréal, Québec, Canada H1X 2B2

[1]Author for correspondence, present address: Allan Wilson Centre for Molecular Ecology and Evolution, Massey University, Private Bag 11 222, Palmerston North, New Zealand (s.joly@massey.ac.nz)

Communicating Editor: Gregory M. Plunkett

ABSTRACT. This study investigates species boundaries in the polyploid complex of *Rosa* sect. *Cinnamomeae* east of the Rocky Mountains. This complex is characterized by extensive intra-specific polymorphism that is the consequence, in part, of hybridization and polyploidy. An objective multivariate approach is employed to delimit species in the complex, which involved cluster and ordination analyses of 25 quantitative morphological characters and of amplified fragment length polymorphisms (AFLPs). Because polyploid individuals blurred species boundaries in the complex, they were discriminated a priori using stomata guard cell lengths in order to investigate species boundaries at each ploidy level separately. Four distinct species were found at the diploid level: *R. blanda* – *R. woodsii*, *R. foliolosa*, *R. nitida* and *R. palustris*. According to the morphological and molecular data, *R. blanda* and *R. woodsii* are indistinguishable and should be considered as a single species. Three species were identified at the polyploid level, *R. arkansana*, *R. carolina*, and *R. virginiana*, albeit with evidence of hybridization between them. The genetic and morphological similarity between individuals of the polyploid species and those of the different diploid species allowed us to identify possible parents for the polyploid species. *Rosa arkansana* likely originated from *R. blanda* (incl. *R. woodsii*), *R. carolina* from a hybrid between *R. blanda* and *R. palustris*, and *R. virginiana* from *R. palustris*. Although the multivariate approach was not able to differentiate species when all individuals were considered together, a classification tree showed that it is indeed feasible to identify species in the complex without prior knowledge of the ploidy level of individuals.

KEYWORDS: Amplified fragment length polymorphisms (AFLP), multivariate analyses, polyploidy, *Rosa*, species delimitation, taxonomy.

Species delimitation in the genus *Rosa* has always been challenging. François Crépin accurately summarized the historical taxonomic work on this genus by remarking that ''species [of the genus *Rosa*] become more obscure and less recognizable as the work upon them has multiplied'' (Crépin 1896). This taxonomic complexity is for the most part the consequence of the important polymorphism in *Rosa* species and is reflected in complexes such as among the native species of section *Cinnamomeae* in eastern North America. But where some people saw polymorphism, others saw distinct species. This explains why Rydberg (1920) recognized 23 native species whereas Erlanson (1966) and Lewis (1957) recognized only nine species in this complex (the total number of species excludes hybrid taxa).

The native species of *Rosa* section *Cinnamomeae* east of the Rocky Mountains form a polyploid species complex that comprises five diploid species, *R. blanda* Ait., *R. foliolosa* Nutt., *R. nitida* Willd., *R. palustris* Marsh., and *R. woodsii* Lindl., and three tetraploid species, *R. arkansana* Porter, *R. carolina* L., and *R. virginiana* Mill. The ninth species of *Rosa* sect. *Cinnamomeae* in eastern North America recognized by Erlanson (1966), *R. acicularis* Lindl. ($2n = 6x, 8x$), is morphologically distinct from the other species (Lewis 1957) and will not be included here. Moreover, its circumboreal distribution (Lewis 1959) implies that investigating its species status would require a much broader sampling than that used in this study. Previous studies of this complex have focussed on morphology (Erlanson 1930, 1934; Lewis 1957, 1958, 1959, 1962), cytology (Erlanson 1929; Lewis 1957, 1966), and experimental crosses (Erlanson 1934; Ratsek et al. 1939, 1940; Lewis and Basye 1961). Although this important biosystematic work has greatly stabilized the number of species accepted, and although some species recognized by Rydberg were shown to be indistinguishable from one of the nine species recognized by Erlanson in 1966 (Bruneau et al. 2005), species delimitation in this complex remains problematic. Taxonomic problems are known at the diploid level where some species hybridize and are morphologically difficult to distinguish (e.g., *R. blanda* and *R. woodsii*; Lewis 1962), but the problem is particularly acute at the polyploid level. *Rosa carolina*, a species that is widespread east of the Mississippi River, is known to hybridize with *R. arkansana* in the western portion of its distribution (Lewis 1957; Erlanson MacFarlane 1966), but also in the east with *R. virginiana* (Fernald 1922; Lewis 1957). Morphological similarity cuts across ploidy levels and no single morphological character can be used to distinguish one species from another.

Moreover, although the polyploids were suggested to have independent origins (Joly et al. 2006), it is still not clear whether these represent distinct evolutionary entities.

To delimit species, it is important to first determine what a species is and how it is to be recognized in nature. In this paper, species are considered unique and distinct from other hierarchic levels of classification (Ghiselin 1975; Hull 1976), making them ''the real units of evolution'' (Mayr 1969). It is at the species level that adaptations, the end product of natural selection, are fixed and allowed to be passed to sibling species. Characteristics that are important for considering species the unit of evolution are their spatiotemporal continuity with potential for evolution and their cohesiveness (Hull 1976). In this sense, species can be viewed as the most inclusive group of organisms that has the potential of maintaining cohesion and that evolve independently from other such groups. This definition allows predictions to be made to identify species in nature. Indeed, if a species is a cohesive group of organisms that evolves independently from other such groups, then it should eventually become morphologically or genetically distinct with time. Intentionally, the nature of the mechanisms that lead to speciation and that are responsible for maintaining cohesion within species is not mentioned. It is principally on this topic that most species concepts differ, in part because the relevant mechanisms vary among groups of organisms. However, whether the cohesive mechanisms involve gene flow, reproductive systems, competition, ecology, or other factors, any of these forces will eventually create genetically and morphologically distinct species with time.

Species delimitation in any group showing extensive morphological polymorphisms is never simple but investigation of species boundaries in a polyploid species complex is even more problematic (Rieseberg et al. 2006). Although polyploid individuals are generally reproductively isolated from their diploid parents (i.e., are distinct species), they may not be morphologically differentiated, thereby challenging traditional approaches for determining species boundaries (Diamond 1992). In situations where polyploids cause problems for species delimitation, one approach (as used herein) is to discriminate diploids and polyploids a priori and to analyse them independently.

Other than polyploidy, hybridization is also commonly reported as blurring species boundaries in several species complexes (Anderson 1949; Diamond 1992). This occurs in part because hybrid individuals commonly show intermediate morphological (and genetic) characteristics relative to their parents (Schilling and Heiser 1976; Neff and Smith 1978; McDade 1997), and because recombination further complexifies the picture by creating all degrees of intermediacy between the parents (Anderson and Hubricht 1938; Jensen and Eshbaugh 1976; Jensen et al. 1993). For the same reasons, allopolyploidy is also expected to cause problems with species delimitation.

In this study, we use an objective multivariate approach for delimiting species boundaries that does not rely upon a priori taxonomic identifications, the null hypothesis being that a single species is present in the complex of *Rosa* sect. *Cinnamomeae* east of the Rocky Mountains. In order to define species in this complex, quantitative morphological characters and amplified fragment length polymorphisms (AFLP) were investigated. Species boundaries were investigated by looking for gaps in the morphological and molecular variation of organisms using both clustering and ordination methods. The goals of the present study were fourfold: (1) Define the species boundaries in the complex, (2) identify the morphological characters that best differentiate the species, (3) investigate potential parental relationships for the polyploid species, (4) and evaluate to what extent polyploidy and hybridization affect our ability to detect species boundaries in this complex.

Materials and Methods

Morphological Analyses. A total of 186 individuals were investigated for the morphological analyses (Appendix 1). Individuals sampled covered the entire geographic distribution of each species in order to represent the full extent of the morphological variation. For most species, one individual per population was selected, even when more than one individual was available. However, in some populations of *R. palustris, R. nitida*, *R. virginiana*, and *R. woodsii*, two or more individuals were sampled (see Appendix 1). Henceforth, individuals will be referred to by the a priori species name followed by the collection number (e.g., *nitida*1010-1).

Morphological characters were selected to be applicable to as many specimens as possible. Because roses flower only 2–3 weeks each year, the reproductive characters chosen were those that could be measured either from flower buds, flowers, or fruits. This decision should not bias the results because most characters considered important in identifying rose species are of this type (Erlanson 1934; Lewis 1957).

Twenty-five quantitative morphological characters were examined and measured either from herbarium specimens (Appendix 2). The character values included in the analyses for each individual were the mean of four/five measurements per specimen, except for the number of leaflets per leaf, which was estimated from ten leaves. The measurements per specimen were sometimes fewer if material was insufficient. All lengths were measured using electronic callipers with a precision of 0.01 mm and a dissection microscope when necessary. Length measures were log transformed (using the natural logarithm) before the analyses. The distribution of all

characters (after transformations) was verified to ascertain that none included extreme outliers.

ESTIMATION OF MISSING DATA. Of the 186 individuals sampled, six had missing data for the four floral characters. Because these individuals were also surveyed for the molecular analyses, these missing characters were estimated to avoid having to remove these individuals from the combined analyses. For four of them (*palustris*417, *nitida*1016-1, *virginiana*520, *virginiana*444), other individuals were available from the same population and missing data were estimated by taking the population mean for each character. Because it was not possible to use a population mean for the two remaining individuals (*carolina*502, *carolina*15783), the missing data were estimated by multiple linear regression using all non-floral characters as independent variables. Removing these individuals from the analyses did not alter our conclusions.

PLOIDY LEVEL DETERMINATION. Stomata guard cell lengths, known to be strongly correlated with ploidy level in roses (Lewis 1957), were measured to estimate the ploidy level of the individuals studied. The average length of guard cells for each individual was estimated from 20 stomata from a terminal leaflet. A K-means analysis of two clusters, performed using the ''kmeans'' function in R (Ihaka and Gentleman 1996; R Development Core Team 2005), was used to objectively discriminate diploids and polyploids. The results were compared to the ploidy level expected from the a priori taxonomic identifications. When these disagreed, pollen size was used as second ploidy level estimator. Pollen size has also been shown to be a robust indicator of the ploidy level in roses (Lewis 1957). Pollen was measured using a microscope with a 63 × objective. Because the length of the stomata guard cells was used a priori to classify individuals as diploid or polyploid, this character was not used in subsequent analyses, except when explicitly mentioned.

Molecular Analyses. Of the 186 individuals included in the morphological analyses, 115 were also investigated at the molecular level using AFLPs (Appendix 1). This represents the specimens collected by the authors and colleagues for which leaves were preserved in silica gel. The sampling for the molecular analyses is therefore not as complete as for the morphological analyses. Species that were most affected by this reduced sampling are *R. foliolosa*, for which only three populations were sampled, and *R. carolina* and *R. palustris*, for which southern populations were lacking. The impact of this limited sampling on the results is discussed later.

DNA was extracted using the cetyltrimethylammonium bromide (CTAB) method of Doyle and Doyle (1987) as modified in Joly et al. (2006). Amplified fragment length polymorphisms (AFLPs; Vos et al. 1995) were used to characterize the genetic constitution of individuals at the genomic level. We followed the protocol of Applied Biosystems (Foster City, California) for large genomes with the modifications described by Bruneau et al. (2005). From a preliminary screening of 18 different primer combinations, four were selected that showed the greatest number of polymorphic bands: EcoRI-AAG + MseI-CAC, EcoRI-ACA + MseI-CAC, EcoRI-AAG + MseI-CTG, and EcoRI-ACA + MseI-CTG. Two primer combinations (differentiated by distinctive chromophores) were run simultaneously on an ABI3100-*avant* sequencer (Applied Biosystems). Unambiguous bands were scored using Genographer (Benham 2001).

Delimitation of Species. To identify distinct groups of individuals, two multivariate approaches were used simultaneously. For each group of individuals analysed, both ordination and cluster analyses were performed, as they give complementary information for determining the number of distinct groups present in the data. The cluster analysis defines groups of individuals from the complete amount of information contained in the data (if the similarity coefficient does not exclude information), whereas ordination analyses illustrate the dispersion of individuals in a few dimensions that explain the greatest amount of variance. In the following analyses, groups of individuals were considered distinct only if the results of the clustering and ordination analyses were congruent. A stepwise approach was used to objectively delimit distinct groups of individuals. First, all individuals were analysed together and distinct groups of individuals were identified. Then, these groups were divided in different datasets and were analysed independently to identify further groups within them. This was important because the presence of outliers, which are responsible for most of the variance present in the dataset, reduces the capacity of these analyses to differentiate the more similar individuals. This is particularly true of ordination methods. Groups identified in this second round of analyses were further analysed, and so on until no further clearly distinct groups of individuals were found. This was done separately for the molecular and morphological datasets: groups found with both datasets were considered to be strongly supported as distinct species whereas those found with only one of the datasets were considered to require more careful interpretation. When the assignment of certain individuals to species differed between the morphological and the molecular analyses, a combined analysis was used to assign these individuals to one species.

Cluster analyses were performed on the morphological dataset, on the molecular dataset, and on a combined dataset. Ward's (1963) minimum variance method was used because it uses an objective function that minimizes the within-group sum of squares, and thus should result in groups that correspond to species according to our definition. The Euclidean distance from standardized variables was used for the morphological matrix, whereas the Jaccard (1900) distance was used for the molecular dataset to avoid considering the shared absence of a band as a similarity and therefore maximizing homology for the data included in the analysis. The distance matrix used in the combined analyses was the mean of the morphological and the molecular matrices, recalculated to include only the individuals that had both morphological and molecular information. To give the datasets approximately the same weight, morphological and molecular matrices were scaled so that the maximum distance in each matrix was 1. Ward's phenograms were obtained from the ''agnes'' function of the ''cluster'' package in R (Maechler 2005).

Ordination analyses involved Principal Component Analysis (PCA) for the morphological data and Principal Coordinate Analysis (PCoA) for the molecular and the combined datasets. The PCA was performed using the ''prcomp'' function in R from the correlation matrix, scaling the character vector lengths to 1. The PCoA analysis was performed using the ''cmdscale'' function in R from the Jaccard distance matrix for the molecular dataset. For the combined analysis, the PCoA was performed on the same matrix that was used for the combined cluster analysis.

Origin of the Polyploids. To investigate potential origins for the polyploid species, we compared the overall morphological and molecular similarity of the individuals of each polyploid with the individuals of each diploid species. Pairwise distances were based on the standardized matrix of Euclidean distances for the morphological data and on the Jaccard distance matrix for the molecular dataset. Both matrices were scaled so that the maximum distance equalled 1. Only polyploid individuals from allopatric populations were used for these calculations. For each polyploid, the mean distance obtained with the different diploid species were tested using Tukey's HSD test (5% level - Tukey 1953;

Kramer 1956) to determine whether the differences obtained were significant.

Classification Tree. Although distinct groups of individuals were found in the analyses, it is impossible to differentiate any species from the others by using a single character (see Results). Yet, it may be still possible to properly identify species if a hierarchic method is used, such as classification trees. Classification trees aim at recovering pure groups of species a priori identified by dividing the individuals into groups in a tree-like fashion using the input characters until the groups obtained are pure or until a dividing threshold is attained. They therefore result in a hierarchic key that can be helpful for differentiating species and identifying characters that are most useful in delimiting species or groups of species. Classification trees were constructed using the species identified in the present study. For polyploids, only individuals from allopatric populations were included. Two trees were constructed: both included all characters, but one also included the length of the stomata guard cells. The minimum number of individuals at a node and in a newly formed group was set to ten and four, respectively. The length characters were not log-transformed here because this does not affect the analysis. The analysis was performed using the ''tree'' package in R (Ripley 2005).

RESULTS

Ploidy Level Determination. The distribution of the lengths of the stomata guard cells showed two modes that also correspond to the K-means clusters (Fig. 1). The K-means cluster with the shortest length consisted of 119 individuals and had a mean length of 17.43 μm, whereas the other cluster comprised 67 individuals and had a mean length of 22.86 μm. These two clusters were assumed to comprise diploid and polyploid individuals, respectively. According to the a priori taxonomic identifications, only 6 individuals (3.2%) were misclassified: *foliolosa*4184, *woodsii*2008, *carolina*967, *virginiana*454, *virginiana*03-57-1, and *virginiana*6. Pollen grains, available and measured for two of these individuals (*virginiana03-57-1*: 36.3 μm ± 4.1, *virginiana454*: 35.1 μm ± 4.6; mean pollen size for diploid and polyploid is 27.9 ± 0.04 and 33.4 ± 0.05, Lewis 1957), disagreed with the results of the guard cell analysis which had considered them to be diploids. For both these individuals, the pollen results were used because they also agreed with the a priori taxonomic identifications. For the specimens that lacked pollen information, the ploidy level obtained from the K-means analysis of the guard cell length was used.

Morphological Differentiation of Diploids. Because no distinct groups of individuals were evident when all individuals were analysed together (data not shown), diploid and polyploid individuals were investigated separately. When all diploid individuals were analysed for the morphological data, the PCA and the cluster analysis

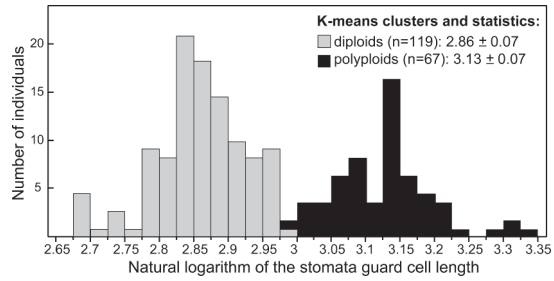

FIG. 1. Distribution of the natural logarithm of the stomata guard cell lengths in *Rosa* sect. *Cinnamomeae*. The k-means clusters and statistics that correspond to the diploid and polyploid individuals are given.

delimited two distinct groups of individuals (Fig. 2). If we consider the a priori species delimitation, one group consisted of all *R. blanda* and *R. woodsii* individuals whereas the other consisted of *R. foliolosa*, *R. nitida* and *R. palustris* individuals. Characters most important to differentiate these groups according to the PCA were the glands on the pedicels and hypanthia, pubescence of leaflets, and the leaflet tooth length (Table 1).

When these two groups were reconsidered in separate analyses, no distinct groups of individuals were evident in the *R. blanda* – *R. woodsii* group (not shown). In contrast, two distinct groups were found in the other: one consisted of all but one *R. palustris* individuals and of three *R. nitida* individuals and the other consisted of all *R. foliolosa* individuals and the remaining *R. nitida* (20) and *R. palustris* (1) individuals (Fig. 3). Characters that mostly differentiated these groups are the number of hypanthium glands, leaflet pubescence, and the number of leaflets (Table 2).

When analysed alone, the *R. palustris* group showed evidence of two distinct groups of individuals (Fig. 4). The *R. foliolosa* – *R. nitida* group also showed evidence for two distinct groups of individuals (Fig. 5), one of which comprised all *R. foliolosa* individuals and one *R. nitida* individual and the other the remaining *R. nitida* individuals. The most important characters for differentiating these last two groups were pedicel length, number of bristles, and hair number on the primary veins (Table 3). No further distinct groups of individuals were found in these two groups (data not shown).

Molecular Differentiation of Diploids. A total of 107 AFLP unambiguous bands were scored. Among diploids, eight bands were constant in all individuals and three were unique. The analysis of all diploid individuals identified four distinct groups of individuals with the first PCoA and Ward's clusters (Fig. 6). Considering the a priori species identification, these groups consisted of *R. nitida*, *R. palustris*, *R. foliolosa*, and of both *R. blanda*

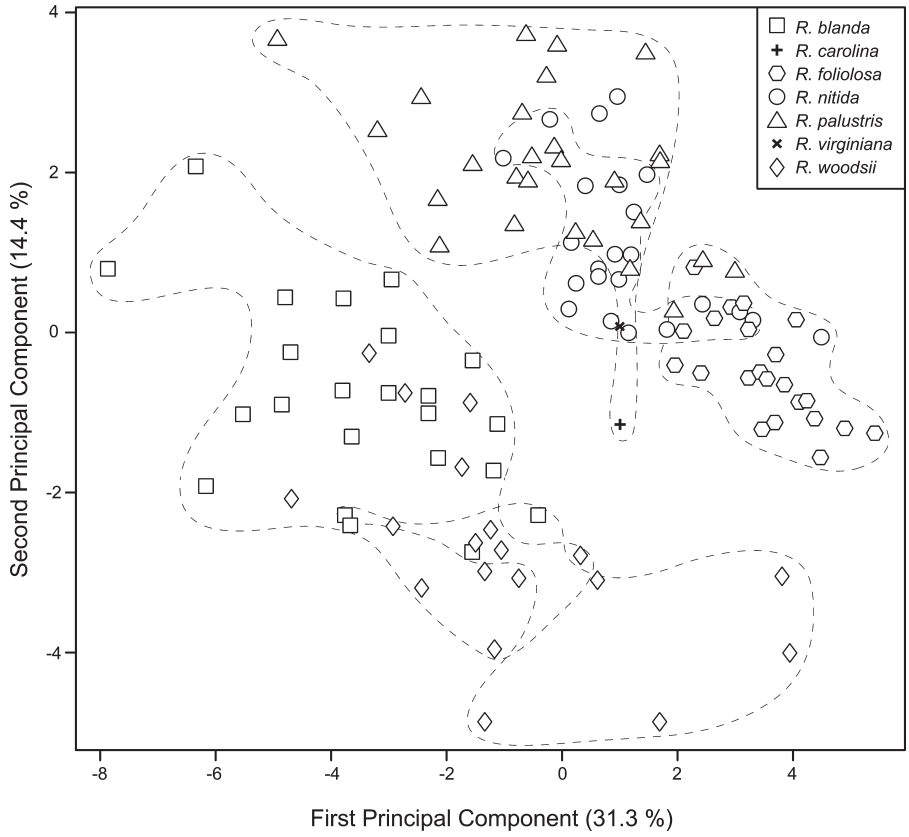

FIG. 2. Principal component analysis of diploid individuals for the morphological characters (character loadings in Table 1). The a priori species identifications are indicated by symbols. The outlines represent the major groups found in a cluster analysis.

and *R. woodsii* together. With further analyses, the *R. palustris* group showed evidence of distinct groups of individuals (Fig. 7), as did the *R. nitida* group (the groups can be observed in Fig. 6 for the 2nd and 3rd components). No further distinct groups were observed (not shown).

Differentiation of Polyploids. Because hybridization potentially occurs among polyploid species, only samples from allopatric populations were included in cluster analyses even though all samples were included in the ordinations. For the morphological dataset, the cluster analysis identified two distinct groups of allopatric individuals that also were separated on the ordination (filled symbols, Fig. 8; Table 4). One group consisted of *R. carolina* individuals and the other of *R. arkansana* and *R. virginiana* individuals, according to a priori taxonomic identifications. Even when this second group was analysed alone, *R. arkansana* and *R. virginiana* individuals did not fall into distinct groups nor was there evidence of other distinct groups of individuals (analyses not shown). When individuals from sympatric populations were also considered, it was impossible to distinguish distinct groups of individuals (see open symbols in Fig. 8; cluster analyses not shown).

In the analysis of the molecular data, the cluster and ordination analyses identified three distinct groups of allopatric individuals (filled symbols, Fig. 9). According to a priori identifications, these three groups consisted exclusively of *R. arkansana*, *R. carolina* and *R. virginiana* individuals. Again, it was more difficult to identify distinct groups of individuals when sympatric individuals were included in the analysis (open symbols, Fig. 9).

Origin of Polyploid Species. The origin of polyploid species was evaluated for the three groups that were recovered in the molecular analyses. Although the group that corresponded to *R. arkansana* and *R. virginiana* could not be distinguished in the morphological cluster analysis, they were slightly distinct in the ordination suggesting a slight morphological differentiation (see discussion). *Rosa arkansana* was equally close to both *R. blanda* and *R. nitida* at the morphological level, but it was closest to *R. blanda* according to

TABLE 1. Character loadings for the different principal components (PC) in the PCA of all diploid individuals (Fig. 2). Because the length of each of the 25 character vectors was scaled to 1, a variable is considered to contribute significantly to the ordination in reduced space if its projection is greater than 0.2.

	PC 1	PC 2	PC 3
NLFT	0.16	−0.13	0.17
N1SER	−0.08	0.39	−0.16
N2SER	0.02	−0.04	0.35
LLFT	−0.27	0.28	−0.08
L1SER	−0.26	−0.07	−0.17
WLFT	−0.32	0.11	−0.01
LWLFT	−0.30	0.20	−0.08
LTEET	−0.25	−0.18	0.13
PULFT	−0.17	−0.24	−0.15
P1VEIN	−0.27	−0.11	−0.22
P2VEIN	−0.26	−0.17	−0.16
PBLFT	−0.19	−0.26	−0.15
G1VEIN	0.04	−0.23	0.04
GBLFT	0.00	−0.21	−0.07
B1YW	0.08	0.09	0.42
P1YW	0.15	0.16	−0.05
LSTP	−0.27	0.25	0.06
LAUR	−0.25	0.08	0.35
WAUR	−0.26	0.06	0.28
WSTP	−0.25	−0.04	0.31
GSTP	−0.02	−0.20	0.17
GHYP	0.11	0.38	−0.20
GPED	0.12	0.30	0.15
LPED	−0.18	0.02	0.24
NFLW	−0.14	0.08	−0.10

TABLE 2. Character loadings for the different principal components (PC) in the PCA of diploid individuals from the *R. foliolosa* – *R. nitida* – *R. palustris* group (Fig. 3). Because the length of each of the 24 character vectors was scaled to 1, a variable is considered to contribute significantly to the ordination in reduced space if its projection is greater than 0.2.

	PC1	PC2	PC3
NLFT	−0.24	−0.13	0.23
N1SER	0.31	0.03	0.00
N2SER	−0.02	−0.24	−0.27
LLFT	0.33	−0.02	0.22
L1SER	0.19	0.18	0.15
WLFT	0.33	−0.08	0.09
LWLFT	0.33	0.00	0.21
LTEET	0.04	−0.20	0.47
PULFT	0.17	0.15	−0.16
P1VEIN	0.29	0.14	−0.21
P2VEIN	0.26	0.13	−0.22
PBLFT	0.23	0.14	−0.28
G1VEIN	−0.09	0.07	0.07
B1YW	−0.07	−0.30	−0.37
P1YW	0.06	−0.02	0.11
LSTP	0.31	−0.14	0.06
LAUR	0.14	−0.39	0.01
WAUR	0.15	−0.33	0.11
WSTP	0.10	−0.39	−0.06
GSTP	−0.02	−0.34	0.10
GHYP	0.20	0.15	−0.01
GPED	0.06	−0.20	−0.23
LPED	0.12	−0.24	−0.24
NFLW	0.16	0.04	0.19

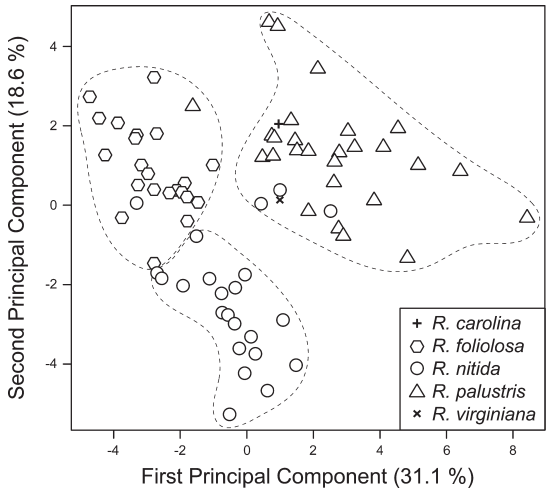

FIG. 3. Principal component analysis of morphological characters for the diploid group that consists of *R. foliolosa*, *R. nitida*, and *R. palustris* individuals identified in the analysis of all diploids (Fig. 2; character loadings in Table 2). The a priori species identifications are indicated by symbols and the outlines represent the major groups found in a cluster analysis.

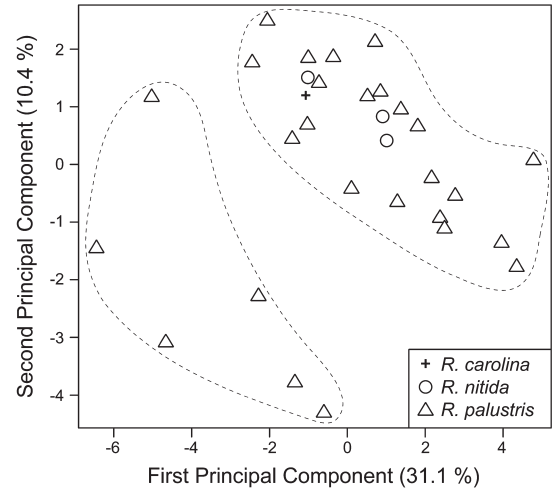

FIG. 4. Principal component analysis of morphological characters for the group that consists principally of *R. palustris* individuals, identified in the analysis of the group of *R. foliolosa* and *R. nitida*, and *R. palustris* (Fig. 3). The a priori species identifications are indicated by symbols and the outlines represent the major groups found in a cluster-analysis.

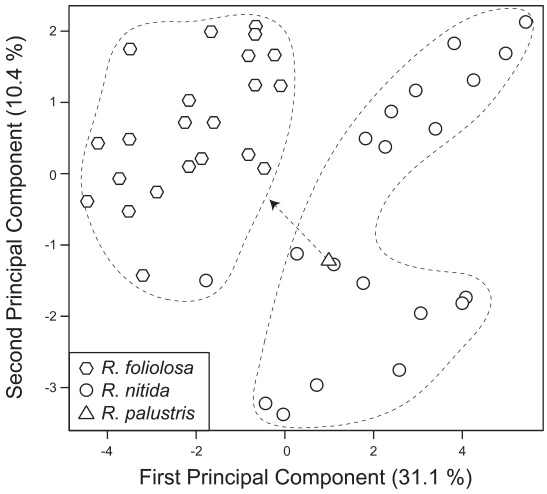

FIG. 5. Principal component analysis of morphological characters for the group that consists principally of *R. foliolosa* and *R. nitida* individuals, identified in the analysis of the group of *R. foliolosa*, *R. nitida*, and *R. palustris* (Fig. 3; character loadings in Table 3). The a priori species identifications are indicated by symbols and the outlines represent the major groups found in a cluster analysis. The grouping of the single *R. palustris* individual was incongruent among the ordination and clustering analyses and the arrow indicates its affiliation in the cluster analysis.

molecular distances (Fig. 10). *Rosa carolina* was morphologically closest to both *R. foliolosa* and *R. nitida*, but the AFLP dataset suggested it was closest to both *R. blanda* and *R. palustris* (Fig. 10). Finally, *R. virginiana* was closest to *R. nitida* according to the morphological characters, whereas the molecular data suggested it was closest to *R. palustris* (Fig. 10).

TABLE 3. Character loadings for the different principal components (PC) in the PCA of diploid individuals from the *R. foliolosa* – *R. nitida* group (Fig. 5). Because the length of each of the 24 character vectors was scaled to 1, a variable is considered to contribute significantly to the ordination in reduced space if its projection is greater than 0.2.

	PC1	PC2	PC3
NLFT	−0.13	0.26	−0.07
N1SER	0.24	−0.09	0.10
N2SER	0.14	−0.21	−0.07
LLFT	0.29	0.16	0.04
L1SER	−0.01	0.03	−0.11
WLFT	0.33	−0.04	−0.01
LWLFT	0.26	0.21	0.01
LTEET	0.13	0.36	−0.05
PULFT	−0.04	0.16	0.09
P1VEIN	0.08	−0.38	−0.16
P2VEIN	−0.07	0.11	−0.63
PBLFT	−0.07	0.11	−0.63
G1VEIN	−0.10	−0.01	0.07
B1YW	0.19	−0.35	−0.04
P1YW	0.02	0.27	0.09
LSTP	0.29	0.09	−0.14
LAUR	0.32	0.01	−0.07
WAUR	0.28	0.13	−0.12
WSTP	0.30	−0.03	−0.10
GSTP	0.27	0.21	−0.03
GHYP	0.06	0.30	0.18
GPED	0.22	−0.02	0.11
LPED	0.28	−0.27	0.00
NFLW	0.04	0.24	0.15

Classification Tree. To construct the classification trees, the four distinct groups of diploid individuals and the individuals from allopatric populations for the three groups of polyploids found in the previous analyses were used. The individuals *foliolosa*4184 and *woodsii*2008 that were

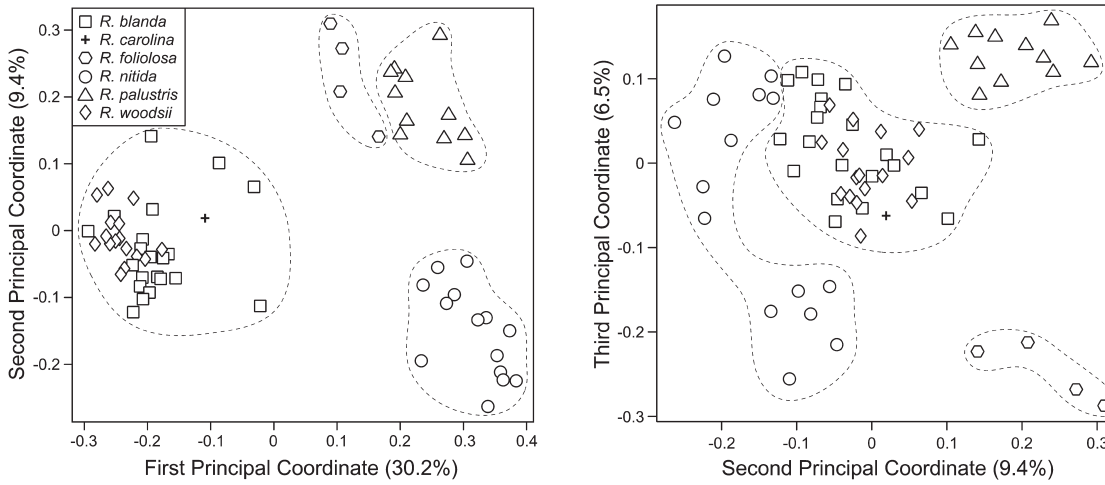

FIG. 6. Principal coordinate analysis of the molecular data for the diploid individuals. The a priori species identifications are shown by symbols. The outlines represent the major groups found in a Ward cluster analysis.

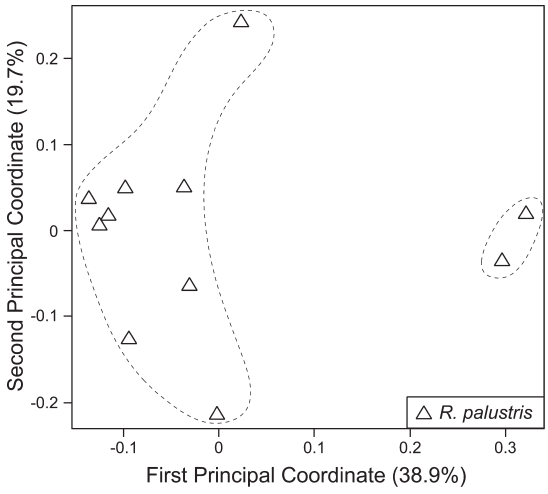

FIG. 7. Principal coordinate analysis of the molecular data for the group that consists of *R. palustris* individuals, as identified in the analysis of all diploid individuals (Fig. 6). The outlines represent the major groups found in a Ward cluster analysis.

considered to be polyploids (stomata size) were removed from the analysis because of their ambiguous positions. Individuals *carolina*967 and *virginiana*6 were also removed because of their unstable position in the analyses. The tree constructed including the stomata guard cell length had only 8 misclassifications out of the 147 individuals included (5.4%; Fig. 11A). When the stomata character was omitted from the analysis, the number of misclassifications obtained was slightly higher (14 of 147, for 9.5%; Fig. 11B).

TABLE 4. Character loadings for the different principal components (PC) in the PCA of polyploid individuals (Fig. 8). Because the length of each of the 25 character vectors was scaled to 1, a variable is considered to contribute significantly to the ordination in reduced space if its projection is greater than 0.2.

	PC1	PC2	PC3
NLFT	−0.28	0.12	−0.20
N1SER	0.00	−0.37	−0.06
N2SER	0.12	−0.21	0.08
LLFT	−0.01	−0.45	0.14
L1SER	−0.12	−0.08	0.09
WLFT	−0.13	−0.37	0.17
LWLFT	−0.06	−0.39	0.20
LTEET	−0.16	−0.28	0.16
PULFT	−0.13	0.08	0.02
P1VEIN	−0.27	0.19	0.24
P2VEIN	−0.26	0.19	0.28
PBLFT	−0.24	0.19	0.31
G1VEIN	−0.07	0.04	0.28
GBLFT	−0.06	−0.01	0.32
B1YW	−0.24	0.07	0.06
P1YW	0.28	−0.11	−0.14
LSTP	−0.27	−0.16	−0.25
LAUR	−0.26	−0.18	−0.27
WAUR	−0.30	−0.13	−0.26
WSTP	−0.28	−0.05	−0.29
GSTP	−0.05	−0.10	0.29
GHYP	0.18	−0.02	0.01
GPED	0.21	−0.02	0.10
LPED	−0.17	−0.06	0.07
NFLW	−0.28	0.12	−0.20

Figure 12 shows the variation in the morphological characters for all species to illustrate the overlap present for every character and to help in identifying species.

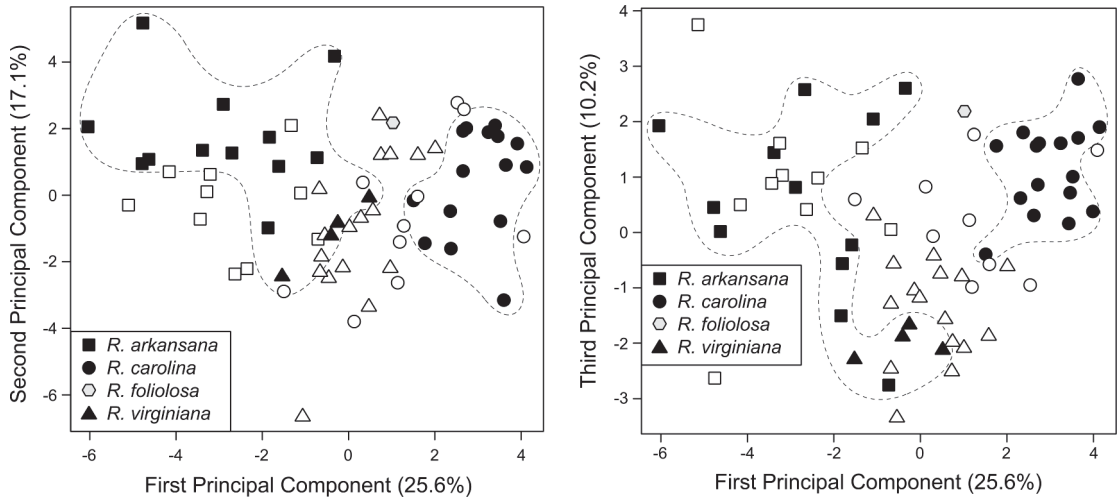

FIG. 8. Principal component analysis of polyploid individuals for the morphological characters (character loadings in Table 4). The a priori species identifications are indicated by symbols; filled forms indicate allopatric polyploid populations and empty forms sympatric populations. The individual identified a priori as *R. foliolosa* but that was classified as a polyploid (stomata size) is represented, but not the *R. woodsii* individual that showed extreme character values. The outlines represent the major groups found in a cluster analysis of allopatric individuals.

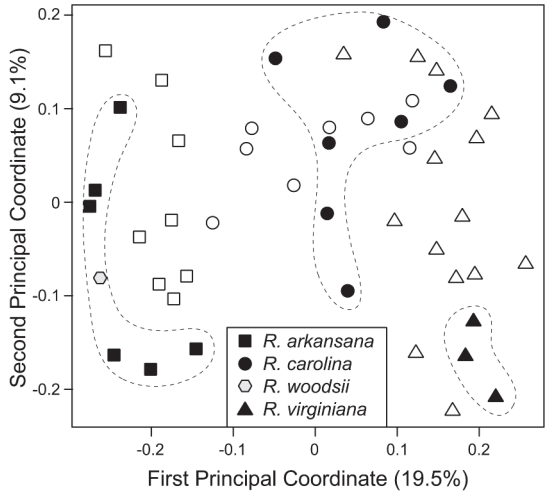

FIG. 9. Principal coordinate analysis of the AFLP data for polyploid individuals. The a priori species identifications are shown by symbols. The outlines represent the major groups found in a Ward cluster analysis of the allopatric individuals; filled and empty forms indicate allopatric and sympatric individuals, respectively. The individual that was identified a priori as *R. woodsii* but that was classified as a polyploid (stomata size) is represented.

DISCUSSION

When all individuals of the complex were analysed in a single analysis, it was impossible to find evidence for distinct groups of individuals (data not shown). This was mostly due to the presence of polyploid individuals because species boundaries became evident when diploids and polyploids were analysed independently. This study also shows that the analysis of individual characters is not sufficient to identify species in the complex (Fig. 12). None of the characters studied can completely discriminate one species from others: the range of variation always overlaps. Only with a multivariate approach can species boundaries be identified by finding gaps in the morphological and molecular variation. Such analyses were used here to delimit species at the diploid and polyploid levels.

How Many Diploid Species? The morphological and the molecular analyses generally revealed the same groups of individuals, although the gaps among these were more pronounced in the molecular analyses. Four distinct groups of diploid individuals were identified by both morphological and molecular analyses in *Rosa* sect. *Cinnamomeae* east of the Rocky Mountains. When these were compared to a priori taxonomic identifications, one group comprised *R. blanda* and *R. woodsii* individuals, whereas the other groups corresponded to *R. foliolosa*, *R. nitida*, and *R. palustris*.

The group consisting of *R. blanda* and *R. woodsii* is clearly distinct morphologically from the three other species of the complex. These two distinct groups of diploids were once placed in separate sections: *R. foliolosa*, *R. nitida* and *R. palustris* in sect. *Carolinae* and *R. blanda* and *R. woodsii* in sect. *Cinnamomeae* (Crépin 1889). Although sect. *Carolinae* is not monophyletic (Wissemann and Ritz 2005; Joly et al. 2006; Bruneau et al. 2007), it nonetheless reflects a morphological differentiation between these species groups. The principal characteristics that differentiate these two groups are the glands on hypanthia and pedicels (absent in *R. blanda* – *R. woodsii*) and the length of the leaflet teeth (larger for *R. blanda* – *R. woodsii*).

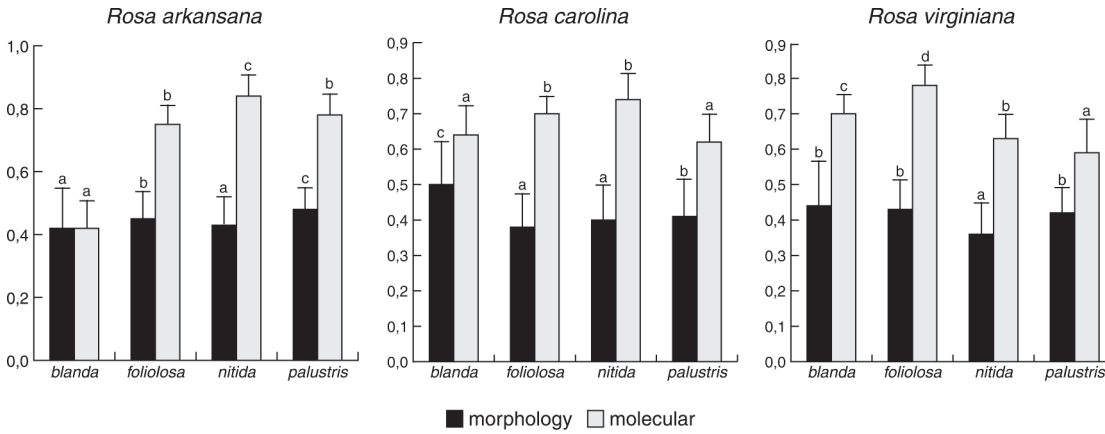

FIG. 10. Histograms showing the mean pair-wise morphological and molecular distances between individuals of each polyploid species (only those from allopatric populations) and those of each diploid species identified in the analyses. For each dataset, different letters above the bars indicate that these means are significantly distinct according to Tukey's HSD test (5% level; Tukey 1953; Kramer 1956).

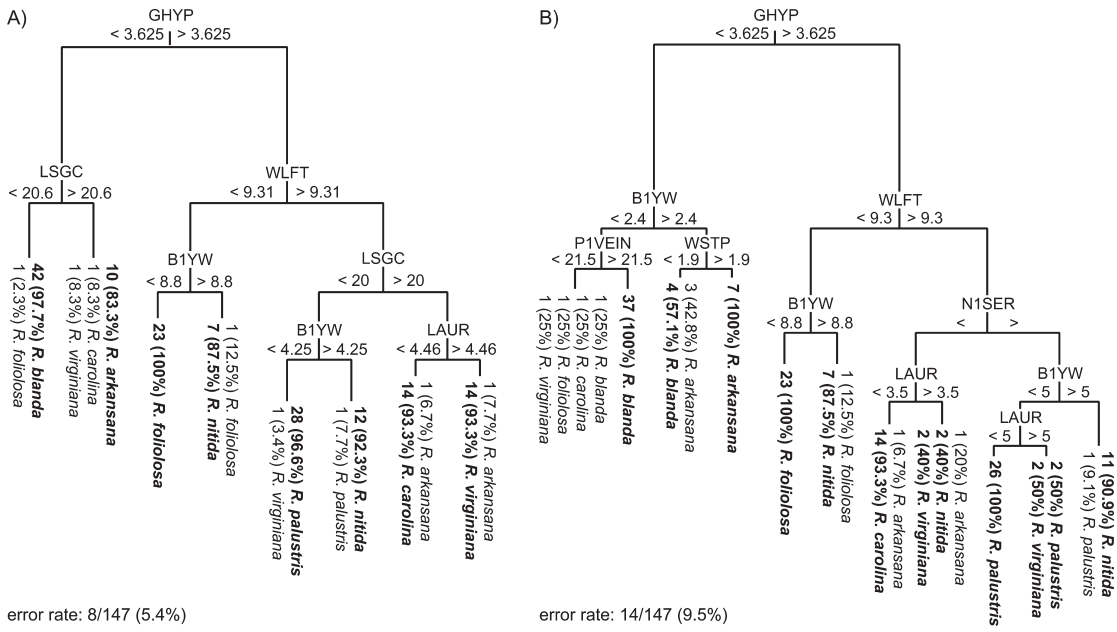

FIG. 11. Classification trees obtained from the species identified in this paper. The two trees represent the best solution when the length of the stomata guard cell length (LSGC) is (A) included and (B) excluded from the analysis. For each terminal group, the number of individuals of each species (and its percentage in the group) is indicated. The lengths of the edges are proportional to the amount of deviance (impurity) resolved by the split at the node above.

The *R. blanda* – *R. woodsii* group is very polymorphic, but there is no evidence of genetic or morphological structure within it. These two species have been differentiated by presence (*R. woodsii*) *vs.* absence (*R. blanda*) of prickles on the stems, although other characters such as leaflet form and glands have also been used (Lewis 1962). Particular importance seems to have been given to the prickle character because it represents a trait of interest to rose breeders and because *R. blanda* is generally considered to be the only unarmed rose species. In addition to prickle characteristics, the geographic distribution of these species also has played a role in their delimitation. *Rosa woodsii* is found from the Pacific coast to Manitoba/Minnesota, whereas *R. blanda* occurs from Manitoba/Minnesota to the Atlantic coast. Thus, other than in the region where the two species overlap, there was not much concern in discerning these species as only one of them was expected. The close relationship of these two species was also noted by Erlanson (1934), who obtained fertile hybrids from interspecific crosses, and in an investigation of three single-copy nuclear genes that failed to reveal a distinction between them (Joly and Bruneau 2006). A hybrid zone in the region of sympatry of these two species has also been proposed (Lewis 1962). However, the present results do not support a hybrid zone because the clines observed in the morphological characters do not all occur in the same geographical region. Instead they either show no fixed differences between *R. blanda* and *R. woodsii* or show a continuous gradient from west to east (Joly 2006). Moreover, even if only individuals from allopatric populations are analysed, the two species still cannot be differentiated (data not shown). Together, these results suggest that *R. blanda* and *R. woodsii* are a single species, which would require placing *R. woodsii* Lind. in synonymy with *R. blanda* Ait. based on priority. But this group is complex and warrants further study. For example, varieties have been described within *R. woodsii* that were not sampled here and these are sometimes considered distinct species (e.g., *R. woodsii* var. *glabrata* (Parish) Cole, *R. woodsii* var. *fendleri* (Crépin) Cole). Denser sampling of *R. woodsii* is needed to completely resolve the taxonomic problems in the *R. blanda* – *R. woodsii* complex.

Rosa foliolosa is the species in this complex with the most restricted geographic distribution: it grows in mesic prairies of northeastern Texas, western Arkansas, and Oklahoma (Lewis 1958). It is also the species with the smallest individuals in this group, and indeed the characters that best differentiate *R. foliolosa* from the other species in these analyses are the narrow terminal leaflets and the short pedicels. Although only four individuals

(three populations) were sampled for the molecular study, it is also clearly distinct at the molecular level. These conclusions are unlikely to change with additional sampling given the distinct genetic nature of the individuals included here and given that this species was clearly distinct in the morphological analyses where sampling covers the entire range of the species.

Rosa nitida and *R. palustris* are the only two species studied here that grow in bogs and poorly drained soils. The morphological analyses showed that *R. nitida* has more bristles, longer auricles, wider stipules, and more double serrations than its close allies *R. foliolosa* and *R. palustris*. Although molecular analyses revealed evidence of sub-structure in *R. nitida*, this differentiation is not reflected at the morphological level nor is it related to the geographic distance among individuals: one group consists of one Newfoundland and one New Brunswick population, whereas the other consists of two Newfoundland and one Québec population (data not shown). Because only five populations were sampled for the molecular study, it is possible that the observed intra-specific genetic differentiation within *R. nitida* is an artefact of the limited sampling and that it would disappear if more populations were analysed. Alternatively, the sub-structure observed may also represent two evolutionary lineages. Further investigations are needed to clarify this.

Rosa palustris individuals are taller, have larger leaflets, and more leaflet hairs than individuals of *R. foliolosa* and *R. nitida*, and they have a greater number of serrations on the leaflets and more hypanthium glands than individuals of any other species in the complex. Sub-structure within *R. palustris* was found in both the molecular and the morphological analyses, although there is no concordance between the groups found. The two morphological groups identified mainly differ in degree of pubescence and there is no correspondence with geographic distance between these groups. The distinction is thus considered to represent intraspecific variation. At the molecular level, only two individuals differ from the others and both are from the same population, which happens to be the eastern-most population sampled. However, because this distinction is based on only two individuals, the *R. palustris* group is considered to represent a single species. Although southern populations of *R. palustris* were not sampled for the molecular dataset, the clear distinctiveness of this species in the molecular analyses and in the morphological analyses suggests that it likely would remain genetically distinct with increased sampling.

How Many Polyploid Species? If no a priori information regarding taxonomic identification of species is used at the polyploid level, the results suggest the presence of a single species: there are no gaps observable in the variation of these individuals (Figs. 8, 9). Yet, if hybridization were occurring among two or more species, this might also be the expected result. To limit the impact of hybridization on species delimitation, we decided to rely partly on current taxonomic species to identify allopatric populations that could be analysed independently. When only individuals from allopatric populations of the three currently recognized species were included in cluster analyses, two groups of individuals were clearly distinct (i.e. in both morphological and molecular analyses): one of these groups is equivalent to *R. carolina*, whereas the other comprises individuals that belong to *R. arkansana* and *R. virginiana*. The molecular dataset suggests that this second group is further divided into distinct groups that perfectly match the a priori species *R. arkansana* and *R. virginiana*. Although these latter groups were not recovered in the cluster analyses of the morphological data, they were slightly differentiated in the PCA. Because these two groups clearly seem to have distinct evolutionary histories according both to the present study (see Origin of Polyploids below) and to a previous genealogical investigation of the single copy *GAPDH* nuclear gene (Joly et al. 2006), and because they occupy distinct geographic areas (see below), they are considered to represent two distinct species. Indeed, it is surprising that *R. arkansana* and *R. virginiana* are difficult to differentiate here because, traditionally, taxonomic problems have mainly involved *R. carolina* with either *R. arkansana* or *R. virginiana* (Lewis 1957). The molecular data also suggest that *R. arkansana* and *R. virginiana* are the two most distant among the polyploid species (Fig. 9). As such, the superficial confusion between *R. arkansana* and *R. virginiana* in the present morphological study may be a consequence of the morphological characters used.

The observation that no morphological or molecular gaps exist among these species when all individuals are considered strongly suggests that hybridization may be common. Lewis (1957) suggested that hybridization contributed to the taxonomic confusion among these polyploid species, but this has never been clearly demonstrated. Even if confusion is expected when differentiating these species in nature, the classification tree shows that it might still be possible to differentiate them with reasonable accuracy. Following is a description of the characters that were

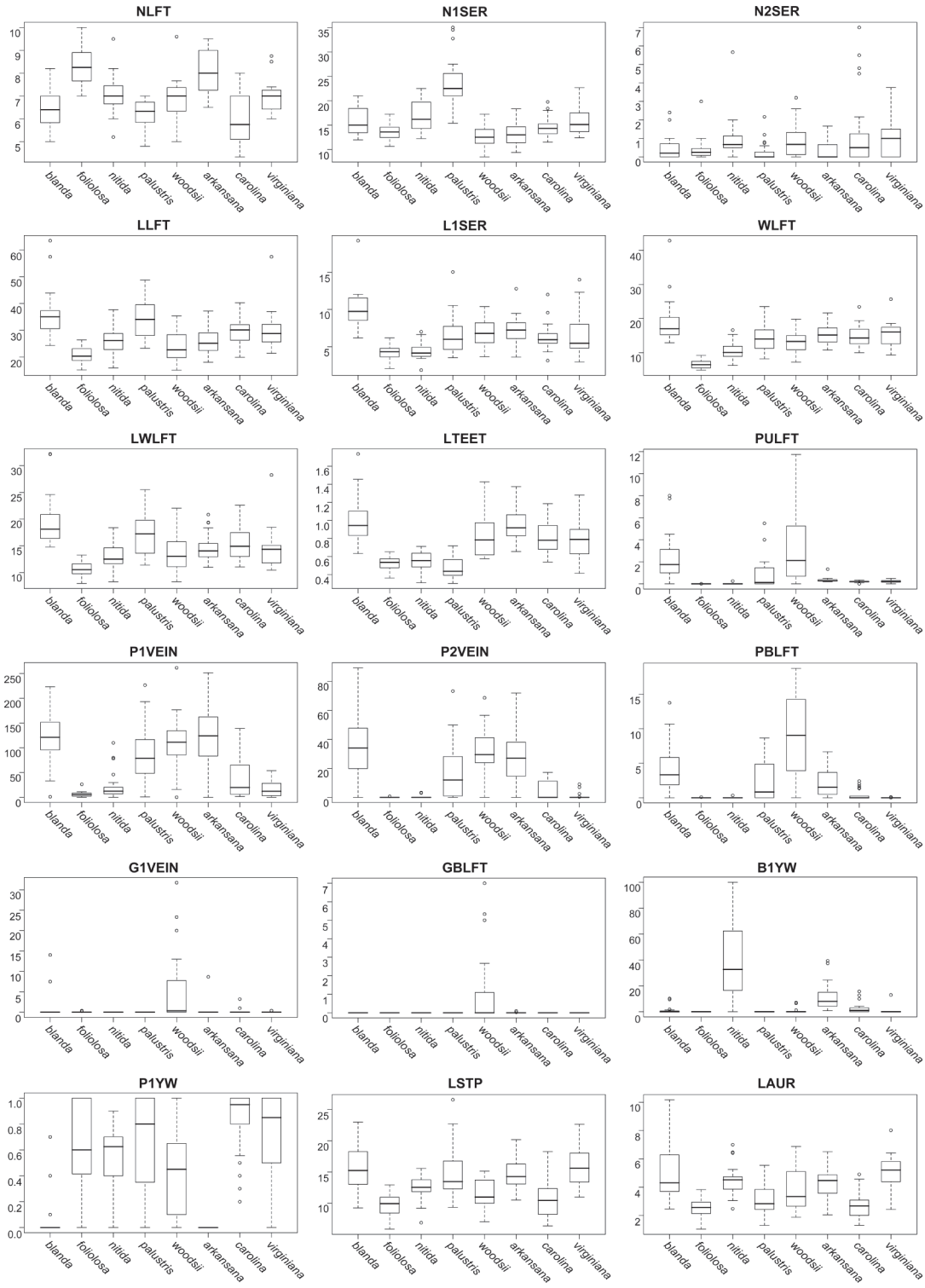

FIG. 12. Boxplots illustrating the variation of the morphological characters investigated for each species identified in this study, with the exception of individuals identified a priori as *Rosa woodsii* that are distinguished from *R. blanda*. Character abbreviations are given in Appendix 2.

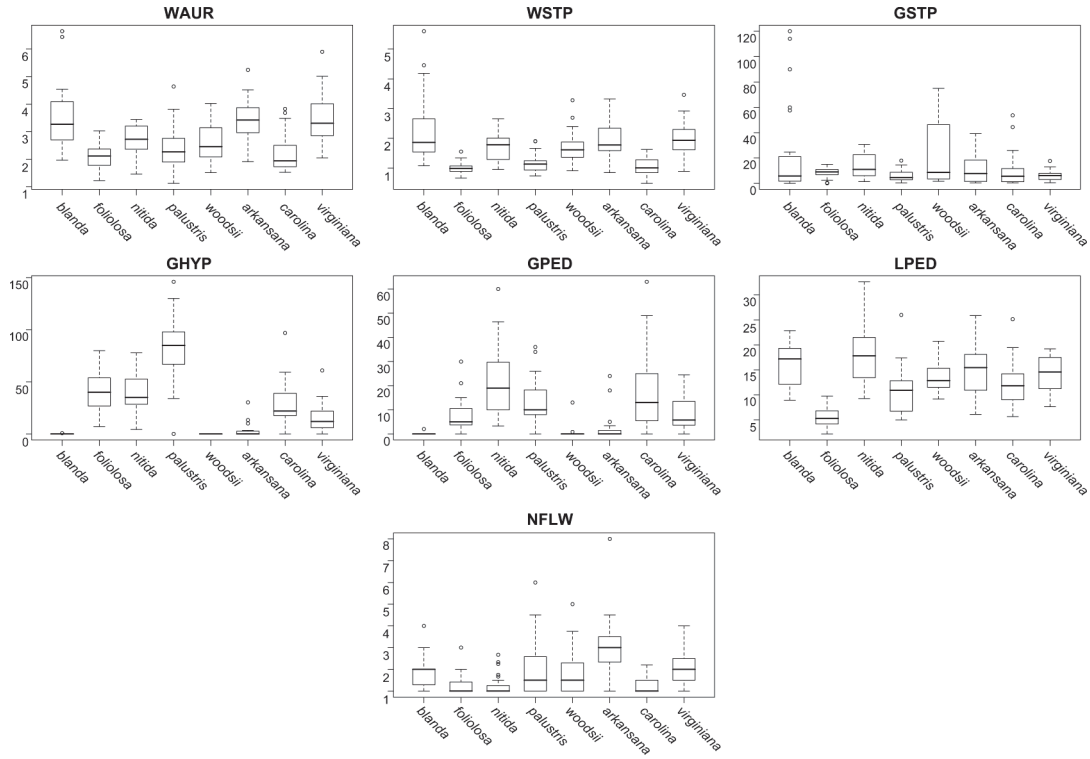

FIG. 12. Continued.

found to be most useful for differentiating each species.

Rosa arkansana grows in the prairies, approximately from the Rocky Mountains to the Mississippi River. It is differentiated from the other polyploid species by a greater number of leaflets per leaf, by the presence of more hairs on the veins and on the limb of the terminal leaflet, by more bristles, by an absence of infrastipular prickles, by very few (or no) glands on hypanthia and pedicels, and by more flowers per inflorescence.

Rosa carolina and *R. virginiana* are difficult to differentiate. Geographically, *R. carolina* is widespread east of the Mississippi River whereas *R. virginiana* is only found along the Atlantic coast from Newfoundland to Virginia. *Rosa carolina* grows in dry soils whereas *R. virginiana* is common on dunes and along the edges of salt marshes. Morphologically, *R. carolina* is often slightly more pubescent, particularly on the secondary veins of the leaflet, has more bristles, and generally has fewer flowers per inflorescence than *R. virginiana*. However, the length and width of the stipules and the length of the auricules seem to be the most useful characters for differentiating these species, with *R. virginiana* having larger values for these characteristics. Another character that is useful in differentiating these species is infrastipular prickle morphology: *R. carolina* usually has straight, not especially broad-based infrastipular prickles, whereas *R. virginiana* has stout prickles that often are broad-based (Lewis 1957).

Origins of Polyploid Species. The overall similarity of polyploids and diploid species (e.g., Heiser et al. 1965; Schilling and Heiser 1976) and their position in ordinations (e.g., Perný et al. 2005) are often used to investigate the origins of polyploid species based on the assumption that hybrids are intermediate to their parental species (Neff and Smith 1978; McDade 1997). Of these approaches, the comparison of similarities is expected to be more accurate than the position of individuals in ordinations because the latter do not use the full extent of information contained in the data, but only the portion represented in the reduced space of the ordination. The comparison of similarities also has the advantage, when pooled by species, that it allows one to statistically test if a species is closer to one putative parent than another.

For all three polyploid species, the molecular and the morphological datasets disagreed as to which diploid species should be regarded as potential ancestors. In these kinds of analyses, interpretation is more straightforward with the

molecular data because the similarity measures are based on the shared presence of bands between diploids and polyploids. In contrast, morphological characters are more likely to be influenced by the polyploid event (Levin 2002; Ramsey and Schemske 2002) and by environmental conditions, although further research is needed to clearly evaluate the effect of this latter factor on character variation in the complex. Moreover, even though hybrid individuals are often thought to be morphologically intermediate relative to their parents (Neff and Smith 1978; McDade 1997), they can either be more extreme than their parents for some morphological characteristics (transgressive hybridization - Rosenthal et al. 2002) or they can be closer to one of their parents because of introgressive hybridization (Whiffin 1973), dominance of characters (Ramon 1968), or recombination in backcross generations (Knops and Jensen 1980). As such, a hybrid may not even fall on the straight line between the two parents in an ordination (Ornduff and Crovello 1968). In addition, a hybrid between two distant individuals may be most similar to a non-parental species that is morphologically intermediate relative to the two parental species (McDade 1997). Consequently, the molecular results may be better suited for hypothesizing diploid ancestors for the polyploid species. This would suggest that *R. arkansana* has evolved from *R. blanda* (incl. *R. woodsii*), that *R. carolina* originated from a cross between *R. blanda* (incl. *R. woodsii*) and *R. palustris*, and that *R. palustris* is the ancestor of *R. virginiana*. Similar conclusions were reached by genealogical analysis of the *GAPDH* nuclear gene (Joly et al. 2006), although it was impossible to distinguish *R. palustris* from *R. nitida* and *R. foliolosa* with this marker.

These results suggest that the three polyploid species have distinct evolutionary origins. This implies that the hybrid zones between these species are secondary hybrid zones, i.e., the hybridizing populations evolved separately prior to contact (Endler 1977; Barton and Hewitt 1985). This is important for species delimitation because it suggests that these species evolved separately and adds further support to their distinct species status.

Delimiting Species in a Polyploid Complex. Ideally, one would like to be as objective as possible when delimiting species in a particular group. Multivariate numerical methods help by adding objectivity in species delimitation, yet the threshold at which groups are given species status remains a subjective decision for the taxonomist (Sneath and Sokal 1973). In order to avoid this subjectivity, we followed an approach that only recognizes groups found jointly in clustering and ordination analyses. At each stage, groups identified are analysed separately to identify smaller groups until no clearly distinct groups emerge. The groups that are assigned species status are those found in both the morphological and the molecular analyses, whereas groups found in only one analysis need more careful interpretation.

POLYPLOIDY. The approach described above was not sufficient to differentiate species in this complex because of the presence of polyploid individuals. Indeed, a recent literature survey showed that polyploidy is one of the most important factors for taxonomic confusion in species complexes (Rieseberg et al. 2006). Although polyploids can be easily distinguished from diploid species in some complexes (e.g., Suda and Lysák 2001; Perný et al. 2005), this is not possible in other groups (e.g., Vanderhoeven et al. 2002). The solution advocated here is to analyse diploid and polyploid individuals separately, which is justified given that polyploids and diploids are reproductively isolated in roses (see Erlanson 1929).

Given the problems polyploids cause for species delimitation, it is worth questioning whether species can be identified in this complex if the ploidy level of individuals is not known. This is relevant because measuring stomata guard cell length or pollen size is not always straightforward. Interestingly, the classification tree from which the length of the stomata guard cells was removed suggests that it is still possible to identify species with limited misclassifications when diploids and polyploids are not differentiated a priori. This suggests that a hierarchic classification can differentiate species even when it is not possible to distinguish species when all individuals are examined simultaneously. This result is even more encouraging because the classification tree discriminates groups using a single character per node, suggesting a possible increase in accuracy if more than one character is used at each node.

HYBRIDIZATION. As with polyploidy, hybridization also has been thought to be responsible for taxonomic confusion in many species complexes (Diamond 1992). Although no clear evidence of hybridization was evident at the diploid level, hybridization clearly causes problems for delimiting species boundaries at the polyploid level. Therefore, if only the objective criterion described above had been used for delimiting polyploid species, the number of species identified would have been underestimated. Hybrids bridge gaps between species, but they can also distort the true

species relationships in phenograms (Heiser et al. 1965; Jensen and Eshbaugh 1976; McDade 1997). This shows that objectivity in delimiting species, although an attractive property, should not be applied blindly because biological events such as hybridization and polyploidy, which are prevalent in plants, may well hinder the detection of real species boundaries.

Identification Key to Species of the Complex. A logical result of the analyses presented here is an identification key. The key below was inspired by the classification trees, but other characters useful in distinguishing these species were also included. *Rosa acicularis*, although not studied here, was included in the key based on Lewis' monograph (1962).

1. Hypanthium glabrous . . . 2
   2. Long and straight prickles present throughout the stems . . . *R. acicularis*
   2. Prickles absent from the stems or, if present, either short (bristles) or curved . . . 3
      3. Infrastipular prickles stout and broad-based . . . *R. virginiana*
      3. Infrastipular prickles absent or not especially stout or broad-based . . . 4
         4. Generally fewer than 2 hairs per mm^2 on the abaxial leaf surface; infrastipular prickles always absent on new stems; bristles always present on new stems; leaflets from 7 to 9 per leaf . . . *R. arkansana*
         4. Generally more than 2 hairs per mm^2 on the abaxial leaf surface; infrastipular prickles either present or absent on new stems; bristles generally absent on new stems; leaflets from 5 to 7 per leaf . . . *R. blanda*
1. Hypanthium with glands . . . 5
   5. Bristles present on new branches . . . 6
      6. Hairs present on the abaxial surface of the terminal leaflet . . . 7
         7. Infrastipular prickles present . . . *R. carolina*
         7. Infrastipular prickles absent . . . *R. arkansana*
      6. Hairs absent on the abaxial surface of the terminal leaflet . . . *R. nitida*
   5. Bristles absent on new branches . . . 8
      8. Width of the terminal leaflet less than 9 mm . . . *R. foliolosa*
      8. Width of the terminal leaflet more than 9 mm . . . 9
         9. Hypanthium typically with more than 86 glands; terminal leaflet oblong, generally with more than 20 small teeth per margin . . . *R. palustris*
         9. Hypanthium typically with fewer than 86 glands; terminal leaflet ovate, elliptic or obovate, generally with fewer than 20 teeth per margin . . . 10
            10. Bristles absent on new stems; auricules more than 3.8 mm long; stipules more than 1.1 mm wide; infrastipular prickles stout, broad based, and often curved . . . *R. virginiana*
            10. Bristles present or absent on new stems; auricules less than 3.8 mm long; stipules less than 1.1 mm wide; infrastipular prickles slender and not especially broad based or curved . . . *R. carolina*

ACKNOWLEDGEMENTS. The authors first thank Julian Starr without whose help this study would not be what it is. The authors also acknowledge the help of Walter Lewis, Barbara Ertter, Luc Brouillet, Beth Dickson, and Félix Hébert for plant collections. Finally, Richard Jensen, François-Joseph Lapointe, Chris Eckert, Luc Brouillet, and Volker Wissemann helped to improve this study through their critical comments and Walter Lewis kindly tested our identification key. Financial assistance for this study was obtained from a National Science and Engineering Research Council (NSERC) scholarship to S. J. and NSERC and Fonds Québécois de la Recherche sur la Nature et les Technologies (FQRNT) grants to A. B.

LITERATURE CITED

ANDERSON, E. 1949. *Introgressive hybridization*. New York: John Wiley.

——— and L. HUBRICHT. 1938. Hybridization in *Tradescantia* III. The evidence for introgressive hybridization. *American Journal of Botany* 25: 396–402.

BARTON, N. H. and G. M. HEWITT. 1985. Analysis of hybrid zones. *Annual Reviews in Ecology and Systematics* 16: 113–148.

BENHAM, J. J. 2001. Genographer, version 1.6.0. Bozeman, Montana: Montana State University.

BRUNEAU, A., S. JOLY, J. R. STARR, and J.-N. DROUIN. 2005. Molecular markers indicate that the narrow Québec endemics *Rosa rousseauiorum* and *Rosa williamsii* are synonymous with the widespread *Rosa blanda*. *Canadian Journal of Botany* 83: 386–398.

———, J. R. STARR, and S. JOLY. 2007. Phylogenetic relationships in the genus *Rosa*: new evidence from chloroplast DNA sequences and appraisal of current knowledge. *Systematic Botany* 32: 366–378.

CRÉPIN, F. 1889. Sketch of a new classification of roses. *Journal of the Royal Horticultural Society* 11: 217–228.

———. 1896. *Rosae Americanae* I. Observations upon the genus *Rosa* in North America. *Botanical Gazette* 22: 1–34.

DIAMOND, J. M. 1992. Horrible plant species. *Nature* 360: 627–628.

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

ENDLER, J. A. 1977. *Geographic variation, speciation, and clines*. Princeton, New Jersey: Princeton University Press.

ERLANSON, E. W. 1929. Cytological conditions and evidences for hybridity in North America wild roses. *Botanical Gazette* 87: 443–506.

———. 1930. The phenological procession in North American wild roses in relation to the polyploid series. *Papers of the Michigan Academy of Sciences, Arts and Letters* 11: 137–150.

———. 1934. Experimental data for a revision of the North American wild roses. *The Botanical Gazette* 96: 197–259.

ERLANSON MACFARLANE, E. W. 1966. The old problem of species in *Rosa* with special reference to North America. *American Rose Annual* 51: 150–160.

FERNALD, M. L. 1922. Notes on the Flora of western Nova Scotia 1921. *Rhodora* 24: 157–180.

GHISELIN, M. T. 1975. A radical solution to the species problem. *Systematic Zoology* 23: 536–544.

HEISER, JR., C. B., J. SORIA, and D. L. BURTON. 1965. A numerical taxonomic study of *Solanum* species and hybrids. *The American Naturalist* 99: 471–488.

HULL, D. L. 1976. Are species really individuals? *Systematic Zoology* 25: 174–191.

IHAKA, R. and R. GENTLEMAN. 1996. R: a language for data analysis and graphics. *Journal of Computational and Graphical Statistics* 5: 299–314.

JACCARD, P. 1900. Contribution au problème de l'immigration post-glaciaire de la flore alpine. *Bull. Soc. Vaudoise Sci. Nat.* 36: 87–130.

JENSEN, R. J. and W. H. ESHBAUGH. 1976. Numerical taxonomic studies of hybridization in *Quercus* I. Populations of restricted areal distribution and low taxonomic diversity. *Systematic Botany* 1: 1–10.

———, S. C. HOKANSON, J. G. ISEBRANDS, and J. F. HANCOCK. 1993. Morphometric variation in oaks of the Apostle Islands in Wisconsin: evidence of hybridization between *Quercus rubra* and *Q. ellipsoidalis* (Fagaceae). *American Journal of Botany* 80: 1358–1366.

JOLY, S. 2006. *Évolution des roses (*Rosa*: Rosaceae) indigènes de la section Cinnamomeae à l'est des Montagnes Rocheuses*. Ph.D. dissertation. Montréal, Canada: Université de Montréal.

——— and A. BRUNEAU. 2006. Incorporating allelic variation for reconstructing the evolutionary history of organisms from multiple genes: an example from *Rosa* in North America. *Systematic Biology* 55: 623–636.

———, J. R. STARR, W. H. LEWIS, and A. BRUNEAU. 2006. Polyploid and hybrid evolution in roses east of the Rocky Mountains. *American Journal of Botany* 93: 412–425.

KNOPS, J. F. and R. J. JENSEN. 1980. Morphological and phenolic variation in a three species community of red oaks. *Bulletin of the Torrey Botanical Club* 107: 418–428.

KRAMER, C. Y. 1956. Extension of multiple range tests to group means with unequal numbers of replications. *Biometrics* 12: 309–310.

LEVIN, D. A. 2002. *The role of chromosomal change in plant evolution*. Oxford: Oxford University Press.

LEWIS, W. H. 1957. *A monograph of the genus* Rosa *in North America east of the Rocky Mountains*. Ph.D. thesis University of Virginia.

———. 1958. A monograph of the genus *Rosa* in North America. II. *R. foliolosa*. *The Southwestern Naturalist* 3: 145–153.

———. 1959. A monograph of the genus *Rosa* in North America. I. *R. acicularis*. *Brittonia* 11: 1–24.

———. 1962. Monograph of the genus *Rosa* in North America. IV. *R.* x *dulcissima*. *Brittonia* 14: 65–71.

———. 1966. Chromosome numbers of phanerogams. 1. *Annals of the Missouri Botanical Garden* 53: 100–101.

——— and R. E. BASYE. 1961. Analysis of nine crosses between diploid *Rosa* species. *Proceedings of the American Society for Horticultural Science* 78: 573–579.

MAECHLER, M. 2005. Cluster: Cluster analysis, version 1.10.2.

MAYR, E. 1969. *Principles of systematic zoology*. New York: McGraw-Hill.

MCDADE, L. A. 1997. Hybrids and phylogenetic systematics. III. Comparison with distance methods. *Systematic Botany* 22: 669–683.

NEFF, N. A. and G. R. SMITH. 1978. Multivariate analysis of hybrid fishes. *Systematic Zoology* 28: 176–196.

ORNDUFF, R. and T. J. CROVELLO. 1968. Numerical taxonomy of Limanthaceae. *American Journal of Botany* 55: 173–182.

PERNÝ, M., A. TRIBSCH, T. F. STUESSY, and K. MARHOLD. 2005. Allopolyploid origin of *Cardamine silana* (Brassicaceae) from Calabria (southern Italy): karyological, morphological and molecular evidence. *Botanical Journal of the Linnean Society* 148: 101–116.

R DEVELOPMENT CORE TEAM. 2005. R: A language and environment for statistical computing, version 2.2.1. Vienna, Austria: R Foundation for Statistical Computing.

RAMON, S. 1968. A numerical taxonomic study of certain taxa of *Haplopappus*, section *Blepharodon*. *University of Kansas Science Bulletin* 47: 863–900.

RAMSEY, J. and D. W. SCHEMSKE. 2002. Neopolyploidy in flowering plants. *Annual Reviews in Ecology and Systematics* 33: 589–639.

RATSEK, J. C., W. S. FLORY JR., and S. H. YARNELL. 1940. Crossing relations of some diploid and polyploid species of roses. *Proceedings of the American Society for Horticultural Science* 38: 637–654.

———, S. H. YARNELL, and W. S. FLORY JR. 1939. Crossing relations of some diploid species of roses. *Proceedings of the American Society for Horticultural Science* 37: 983–992.

RIESEBERG, L. H., T. E. WOOD, and E. J. BAACK. 2006. The nature of plant species. *Nature* 440: 524–527.

RIPLEY, B. 2005. Tree: classification and regression trees, version 1.0–20.

ROSENTHAL, D. M., A. E. SCHWARZBACH, L. A. DONOVAN, O. RAYMOND, and L. H. RIESEBERG. 2002. Phenotypic differentiation between three ancient hybrid taxa and their parental species. *International Journal of Plant Science* 163: 387–398.

RYDBERG, A. 1920. Roses of northeastern America. *Bulletin of the Torrey Botanical Club* 47: 45–66.

SCHILLING, E. E. and C. B. HEISER JR. 1976. Re-examination of a numerical taxonomic study of *Solanum* species and hybrids. *Taxon* 25: 451–462.

SNEATH, P. H. A. and R. R. SOKAL. 1973. *Numerical Taxonomy*. San Francisco: W. H. Freeman and Company.

SUDA, J. and M. A. LYSÁK. 2001. A taxonomic study of the *Vaccinium* sect. *Oxycoccus* (Hill) W.D.J. Koch (Ericaceae) in the Czech Republic and adjacent territories. *Folia Geobotanica* 36: 303–320.

TUKEY, J. 1953. A problem of multiple comparisons. Mimeographed manuscript. Reprinted as pp.1-300 in *The collected works of John W Tukey Vol. VIII. Multiple comparisons : 1948–1983,* ed. H. I. Braun. New York: Chapman and Hall.

VANDERHOEVEN, S., O. HARDY, X. VEKEMENS, C. LEFÈVRE, M. DE LOOSE, J. LAMBINON, and P. MEERTS. 2002. A morphometric study of populations of the *Centaurea jacea* complex (Asteraceae) in Belgium. *Plant Biology* 4: 403–412.

VOS, P., R. HOGERS, M. BLEEKER, M. REIJANS, T. VAN DE LEE, M. HORNES, A. FRIJTERS, J. POT, J. PELEMAN, M. KUIPER, and M. ZABEAU. 1995. AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Research* 23: 4407–4414.

WARD, J. H. 1963. Hierarchical grouping to optmize an objective function. *Journal of the American Statistical Association* 58: 236–244.

WHIFFIN, T. 1973. Analysis of a hybrid swarm between *Heterocentron elegans* and *H. glandulosum*. *Taxon* 22: 413–423.

WISSEMANN, V. and C. M. RITZ. 2005. The genus *Rosa* (Rosoideae, Rosaceae) revisited: molecular analysis of nrITS-1 and *atp*B-*rbc*L intergenic spacer (IGS) versus conventional taxonomy. *Botanical Journal of the Linnean Society* 147: 275–290.

APPENDIX 1. Specimens included in the multivariate analyses of morphological and molecular data. Specimens are ordered by their a priori taxonomic identifications then by country and state. Only the principal collector is given, followed by the collection number, herbarium, and the analysis in which the individual was included (M = included in morphological analyses; MM = included in morphological and molecular studies). When no collection number was indicated on a specimen, the collection date is given in parentheses. Voucher specimens not separated by semicolons are from the same population.

Rosa arkansana. **CANADA: Alberta:** *Moss 82* (US) M; **Manitoba:** *Joly 730* (MT) MM; *Joly 738* (MT) MM; **Saskatchewan:** *Ryan 3* (MT) MM; *Ryan 7* (MT) MM; *Ryan 8* (MT) MM; **USA: Colorado:** *Cary 66* (US) M; **Idaho:** *Allen (1873)* (US) M; **Iowa:** *Hayden 11581* (US) M; *Joly 601* (MT) MM; **Kansas:** *Lewis 15792-2* (MO) MM; *Lewis 15837-1* (MT) MM; *Lewis 15840-1* (MT) MM; **Minnesota:** *Joly 655* (MT) MM; *Joly 663* (MT) MM; *Joly 673* (MT) MM; **Montana:** *Standley 17604* (US) M; **New Mexico:** *Arsène 17732* (US) M; **North Dakota:** *Joly 763* (MT) MM; **Texas:** *Ruth 726* (US) M; **Wisconsin:** *Joly 605* (MT) MM; **Wyoming:** *Tweedy 3224* (US) M.

Rosa blanda. **CANADA: Manitoba:** *Joly 699* (MT) MM; *Joly 722* (MT) MM; **New Brunswick:** *Joly 409* (MT) MM; *Joly 962* (MT) MM; *Joly 988* (MT) MM; *Joly 993* (MT) MM; **Ontario:** *Joly 582* (MT) M; *Joly 788* (MT) MM; **Québec:** *Bruneau 1214* (MT) MM; *Bruneau 1219* (MT) MM; *Bruneau 1236* (MT) MM; *Bruneau 1239* (MT) M; *Drouin 98016* (MT) MM; *Joly 1011-1* (MT) MM; **USA: Michigan:** *Joly 784* (MT) MM; **Minnesota:** *Joly 657* (MT) MM; *Joly 678* (MT) MM; *Joly 692* (MT) MM; *Joly 770* (MT) MM; **Missouri:** *Lewis 15887* (MT) MM; **Wisconsin:** *Joly 622* (MT) MM; *Joly 636* (MT) MM; *Joly 780* (MT) MM.

Rosa carolina. **CANADA: New Brunswick:** *Joly 967* (MT) MM; **Ontario:** *Joly 576* (MT) MM; *Joly 580* (MT) MM; **USA: Alabama:** *Pollard 89* (US) M; **Arkansas:** *Scully 1326* (US) M; **Florida:** *Palmer 35234* (US) M; **Georgia:** *Allard 101* (US) M; **Kansas:** *Lewis 15843-1* (MO) MM; **Kentucky:** *Braun 3117* (US) M; **Louisiana:** *Thieret 22907* (US) M; **Massachusetts:** *Joly 460* (MT) MM; **Minnesota:** *Joly 651* (MT) MM; **Mississippi:** *McDougall 1620* (US) M; **Missouri:** *Lewis 15779* (MO) MM; *Lewis 15783* (MO) M; *Lewis 15844* (MO) MM; **New Jersey:** *Joly 502* (MT) MM; **Pennsylvania:** *Joly 491* (MT) MM; **Tennessee:** *Lewis 15879-1* (MT) MM; *McDougall 1635* (US) M; **Texas:** *Lewis 2064* (US) M; **Virginia:** *Joly 523* (MT) MM, *Joly 524* (MT) MM; **West Virginia:** *Joly 545* (MT) MM; **Wisconsin:** *Joly 620* (MT) MM; *Joly 775* (MT) MM.

Rosa foliolosa. **USA: Arkansas:** *Erlanson 9529* (MO) M; **Oklahoma:** *Emig 614* (MO) M; *Emig 758* (MO) M; *Engelmann (1897)* (MO) M; *Griffith 3484-1* (MO) M; *Hill 11782* (MO) M; *Houghton 3968* (MO) M; *Lewis 15846-1* (MO) MM, *Lewis 15846-2* (MO) MM; *Lewis 15979* (MO) MM; *Merrill 783* (MO) M; *Palmer 13079* (MO) M; *Palmer 42016* (MO) M; *Palmer 8306* (MO) M; *Waugh 125* (MO) M; **Texas:** *Butler 11074* (MO) M; *Eggert (1899)* (MO) M; *Erlanson 9526* (MO) M; *Heller 4184* (MO) M; *Lindheimer 608* (MO) M; *Lundell 13902* (MO) M; *O'Kennon 19069A* (MT) MM.

Rosa nitida. **CANADA: New Brunswick:** *Joly 941* (MT) MM, *Joly 943* (MT) MM, *Joly 944* (MT) MM; *M.-Victorin 46572* (MT) M; **Newfoundland:** *Brouillet 03-55-1* (MT) MM, *Brouillet 03-55-2* (MT) MM, *Brouillet 03-55-3* (MT) MM; *Joly 1016-1* (MT) MM, *Joly 1016-2* (MT) MM, *Joly 1016-3* (MT) MM; *Joly 1018-1* (MT) MM, *Joly 1018-4* (MT) MM, *Joly 1018-5* (MT) MM; **Nova Scotia:** *Smith 8288* (MT) M; **Prince Edward Island:** *Fernald 7664* (MT) M; **Québec:** *Bergeron 81-39* (MT) M; *Cinq-Mars 66-226* (MT) M; *Hamel 12486* (MT) M; *Joly 1010-1* (MT) MM, *Joly 1010-2* (MT) MM, *Joly 1010-3* (MT) MM; *M.-Victorin 49425* (MT) M; **FRANCE: St.-Pierre et Miquelon:** *LeGallo 460* (MT) M.

Rosa palustris. **CANADA: New Brunswick:** *Joly 417* (MT) MM, *Joly 418* (MT) MM; **Ontario:** *Joly 573* (MT) MM; *Taylor 2141* (US) M; **Québec:** *Bowers 2182* (MT) M; *M.-Victorin 2362* (MT) M; *R.-Germain 7114* (MT) M; *R.-Germain 7115* (MT) M; *Raymond (1947)* (MT) M; **USA:** *Lewis 15980-1* (MT) MM; **Connecticut:** *Joly 476* (MT) MM; **Florida:** *Small 8652* (US) M; **Georgia:** *Duncan 6222* (US) M; **Illinois:** *Lewis 2406* (US) M; **Michigan:** *Joly 587* (MT) MM; *Joly 588* (MT) MM; **Missouri:** *Palmer 6159* (US) M; **Ohio:** *Lewis 2305* (US) M; **New Hampshire:** *Lewis 2156* (US) M; **Pennsylvania:** *Joly 548* (MT) M, *Joly 549* (MT) MM; *Joly 560* (MT) MM, *Joly 561* (MT) MM; **South Carolina:** *Godfrey 734* (US) M; **Tennessee:** *McDougall 1362* (US) M; **West Virginia:** *Allard 11491* (US) M; **Wisconsin:** *Joly 644* (MT) MM.

Rosa virginiana. **CANADA: New Brunswick:** *Hébert 6* (MT) M; *Joly 431* (MT) MM; *Joly 946* (MT) MM; *Joly 973* (MT) MM; **Newfoundland:** *Brouillet 03-57-1* (MT) MM; *Brouillet 03-60-1* (MT) MM; *Joly 1017-1* (MT) MM; *Joly 1019-1* (MT) M; **Nova Scotia:** *Joly 1015-1* (MT) MM; *Joly 924* (MT) MM; *Joly 928* (MT) MM; *Lewis 15898* (MT) MM; **Prince Edward Island:** *Fernald 7667* (MT) M; **Québec:** *Joly 997* (MT) MM; **USA: Connecticut:** *Joly 474* (MT) MM; **Maine:** *Joly 444* (MT) MM; **Maryland:** *Joly 517* (MT) MM, *Joly 520* (MT) MM; **Massachusetts:** *Joly 454* (MT) MM; **New Jersey:** *Bartram 3668* (MT) M; *Joly 496* (MT) MM; **Rhode Island:** Collins (1920) (MT) M.

Rosa woodsii. **CANADA: Alberta:** *Dickson 2008* (MT) MM; **British Columbia:** *Lewis 15848-1* (MO) MM; *Lewis 15850-2* (MO) MM; **Manitoba:** *Joly 741* (MT) MM; **Northwest Territories:** *Porsild 16664* (MT) M; **Saskatchewan:** *Joly 750* (MT) MM; *Joly 754* (MT) MM; *Ryan 1* (MT) M; **USA: California:** *Ertter 17989* (JEPS, MT) MM; *Ertter 18307* (JEPS, MT) MM; **Colorado:** *Joly 1005-1* (MT) MM, *Joly 1005-2* (MT) MM; *Joly 1008-1* (MT) MM; **Idaho:** *Ertter 18005* (JEPS, MT) MM; **Montana:** *Hitchcock 13164* (MT) M; **Nevada:** *Ertter 17525* (JEPS, MT) MM; **New Mexico:** *Spellenberg 12555* (MT) MM; **North Dakota:** *Joly 758* (MT) MM; **Oregon:** *Ertter 17990* (JEPS, MT) MM; **Utah:** *Ertter 18289c* (JEPS, MT) MM.

APPENDIX 2. Description of the morphological characters used for delimiting species boundaries in *Rosa* sect. *Cinnamomeae* east of the Rocky Mountains.

Leaf Characters. **NLFT:** Number of leaflets per leaf. **N1SER:** Number of primary leaflet serrations on one side of the terminal leaflet, including the terminal serration. **N2SER:** Number of double serrations on one side of the terminal leaflet. **LLFT:** Length of the terminal leaflet, from the base to the extremity of the limb of the leaflet, in mm. **L1SER:** Length of the terminal leaflet along the primary vein from the base of the limb to the point perpendicular to the first serration on either side of the leaflet, in mm. **WLFT:** Width of the leaflet at the widest point of the leaflet, in mm. **LWLFT:** Length along the primary vein of the terminal leaflet from the base of the limb to the point perpendicular to the greatest width of the leaflet, in mm. **LTEET:** Depth of serration sinuses near the middle of the leaflet, in mm. **PULFT:** Number of hairs on the adaxial surface of the terminal leaflet in a 1.71 × 1.71 mm area. **P1VEIN:** Number of hairs along 1 cm of the primary vein on the abaxial side of the terminal leaflet. **P2VEIN:** Number of hairs along 1 cm of a secondary vein on the abaxial side of the terminal leaflet. **PBLFT:** Number of hairs on the abaxial surface of the terminal leaflet in a 1.71 × 1.71 mm area. **G1VEIN:** Number of glands along 1 cm of the primary vein on the abaxial side of the terminal leaflet. **GBLFT:** Number of glands on the abaxial surface of the terminal leaflet in a 1.71 × 1.71 mm area.

Stipule Characters. **LSTP:** Length of the part of the stipule that is adnate to the petiole, in mm. **LAUR:** Length from the point where the stipule diverges from the petiole to the

extremity of the auricle, in mm. **WAUR:** Width of the auricle, from the middle of the petiole to the portion of the stipule that is furthest from the rachis, in mm. **WSTP:** Width of the stipule at half length, in mm, from the middle of the petiole to the blade of the stipule, at the middle of the portion of the stipule attached to the rachis. **GSTP:** Number of glands along the border of the stipule on one side of the petiole.

Prickle and Bristle Characters. **B1YW:** Average number of bristles along 1 cm of current-year stems. **P1YW:** Proportion of leaves below which there are infrastipular prickles on branches of the current year. Infrastipular prickles are always in pairs on either side of the leaf and immediately below its point of attachment.

Floral characters. **NFLW:** Number of flowers per inflorescence. **GHYP:** Number of glands found on one side of the hypanthium. **GPED:** Number of glands on one side of the pedicel over its whole length. **LPED:** Length of the pedicel, in mm.