ANALYSIS OF GENETIC DIVERSITY AND RELATIONSHIPS IN THE CHINA

ROSE GROUP

A Thesis

by

VALERIE ANN SOULES

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

December 2009

Major Subject: Plant Breeding

ANALYSIS OF GENETIC DIVERSITY AND RELATIONSHIPS IN THE CHINA

ROSE GROUP

A Thesis

by

VALERIE ANN SOULES

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

Approved by:

Chair of Committee,	David H. Byrne
Committee Members,	James R. Manhart
	Alan E. Pepper
	Hongbin Zhang
Head of Department,	Tim D. Davis

December 2009

Major Subject: Plant Breeding

ABSTRACT

Analysis of Genetic Diversity and Relationships in the China Rose Group.

(December 2009)

Valerie Ann Soules, B.S., North Dakota State University

Chair of Advisory Committee: Dr. David H. Byrne

The wild origin, early breeding history, and diversity of the China Rose group, including *R. chinensis* and its varieties, cultivars, and hybrids, are largely unknown. The aims of this study were to investigate the genetic diversity and relationships of the China Roses with related species and hybrids, including information in support of, or refuting, the hypothesis that these roses are the hybrid result of the wild *R. chinensis* var. *spontanea* and *R. odorata* var. *gigantea*. Ninety *Rosa* accessions, including China Roses, a Miscellaneous Old Garden Rose, Noisettes, early Polyanthas, Bourbons, Teas, and species from Sections Indicae and Synstylae were surveyed using 23 microsatellite primer pairs. The *trnH-psbA* chloroplast intergenic spacer was also sequenced for the China Roses, Misc. Old Garden Rose, and the species to look specifically at maternal relationships.

A total of 291 alleles were scored for the 23 microsatellites, with alleles per locus ranging from 6-22 and averaging 12.65. A dendrogram based on Dice similarity and a three-dimensional Principle Coordinate Analysis (PCoorA) graph were plotted with the

data. In the cluster analysis, the similarity coefficients ranged from ~0.15-0.99, with the cultivated roses forming well-defined groups at about 0.45 similarity. These groups generally reflected the American Rose Society horticultural classifications. A large number of sports and synonyms in the China Rose group were identified through this analysis as well. The PCoorA gave a better graphical representation of the relationships of the species and cultivars, and with the inclusion of the chloroplast sequence haplotypes, some maternal relationships could also be identified.

This study shows that the cultivated China Roses are a closely related group and identified which accessions were likely Hybrid China Roses. The results also suggest that the China Roses were maternally derived from *R. chinensis* var. *spontanea*. Based on the microsatellites and chloroplast sequence haplotypes, the identity of the *R. odorata* var. *gigantea* accessions in this study are suspect, but the China Roses may also have this species in their background as the result of natural or artificial hybridization.

ACKNOWLEDGEMENTS

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

Modern, typically tetraploid rose hybrids are considered superior to wild roses with respect to most of the economically important ornamental traits; however, these traits have their origins in a core group of rose species. Of the nearly 200 rose species that have been described worldwide, only about ten are generally considered to have contributed significantly to modern cultivars, including the diploids *Rosa chinensis* Jacq*., R. odorata* var. *gigantea* (syn. *R. gigantea*) Collet, and *R. multiflora* Thunb. (Gudin, 2000). To give us the roses grown today for gardens, cut flowers, and other rose products, extensive breeding has taken place since these species were first cultivated. Unfortunately, the circumstances of the domestication and early breeding of modern roses was poorly preserved as written records, but molecular tools are now being used to understand the genetic diversity and relationships between groups.

For the purposes of this study, China Roses will refer to the group that includes *R. chinensis* as a species in section Indicae, as well as its horticultural varieties and early Hybrid Chinas. China Roses, especially 'Old Blush' (syn. 'Parson's Pink China') and 'Slater's Crimson China', were of great importance in the background of modern roses because of the specific traits they contributed (Shepherd, 1978). Recurrent blooming, darkening from bud to aging blossom rather than fading, deep red color, and dwarf habit are all traits attributed to one or more horticultural varieties from the China Rose group

_____

This thesis follows the style of the American Society of Horticultural Science Journals.

(Shepherd, 1978). However, reports on the wild *R. chinensis* Jacq. var. *spontanea* Rehd. & Wils. do not describe it as having the most noteworthy China Rose trait: the ability to repeat bloom throughout the season (Rix, 2005; Zlesak, 2006).

Some of the complications surrounding the taxon *R. chinensis* come from the species type specimen being one of the cultivated varieties introduced to Europe from China, and not a wild specimen (Shepherd, 1978). The root of the confusion has its origin much further in the past though, when the Chinese first began cultivating roses perhaps as long as 5000 years ago (Shepherd, 1978). Guoliang (2003) lists three factors likely to have been involved in the origin of China Roses 1) large-scale cultivation of wild roses made repeatedly finding and propagating unique, desirable specimens and sports a common occurrence, 2) some degree of artificial pollination by Chinese gardeners, and 3) selection for particular traits after natural hybridization. The most often speculated candidate for natural or artificial hybridization is *R. chinensis* var. *spontanea* with *R. odorata* var. *gigantea* (Shepherd, 1978; Zlesak, 2006).

Other possible contributions to the genetic background of China Roses may have come from a species in the section Synstylae, as noted by Piola et al. (2002) based on morphological descriptions of the section Indicae and Synstylae species and the cultivated group of China Roses (Krussmann, 1981). There is also a high level of cross fertility between species in sections Indicae and Synstylae. Notable examples include the foundations for the Polyantha and Noisette horticultural classes (*R. chinensis* x *R. multiflora* and x *R. moschata* Herrm., respectively (Krussmann, 1981)), as well as the

Texas A&M Basye Rose Breeding and Genetics Program's *R. wichuriana* Crép.

'Basye's Thornless' x *R. chinensis* 'Old Blush' population (Byrne et al., 2007).

The early hybrid groups created using the China Roses as one of the parents have also played an important part as the bridge between the China Roses and today's modern roses (Figure 1). The horticultural classes of those hybrids included in this study are: Noisette, early Polyantha (pre-1900), Bourbon, and Tea Roses. The Noisette class was founded on the cross *R. moschata* x *R. chinensis* 'Old Blush' (syn. 'Parson's Pink China') made by Champney in 1802 and named 'Champney's Pink Cluster'. Historically the Noisette class has been hard to define, as they were quickly interbred with Teas and other classes (Harkness, 1978). Polyantha Roses started with the crossing of a dwarf China Rose with a repeat flowering dwarf of *R. multiflora*, and these roses tend to be short, repeat flowering plants with clusters of small blooms (Phillips and Rix, 1988). The origin of the large petaled, fragrant Bourbon Roses is less certain, but they appear to have been derived from hybrids between China Roses and Damask Roses and named for the Ile de Bourbon (Harkness, 1978). Lastly, the Tea class, which has a recorded origin that shares two important things with China Roses: both are thought to be based on hybrids between *R. chinensis* and *R. odorata* var. *gigantea*, and both had already been cultivated in China for many years before their introduction and further breeding in Europe (Harkness, 1978).
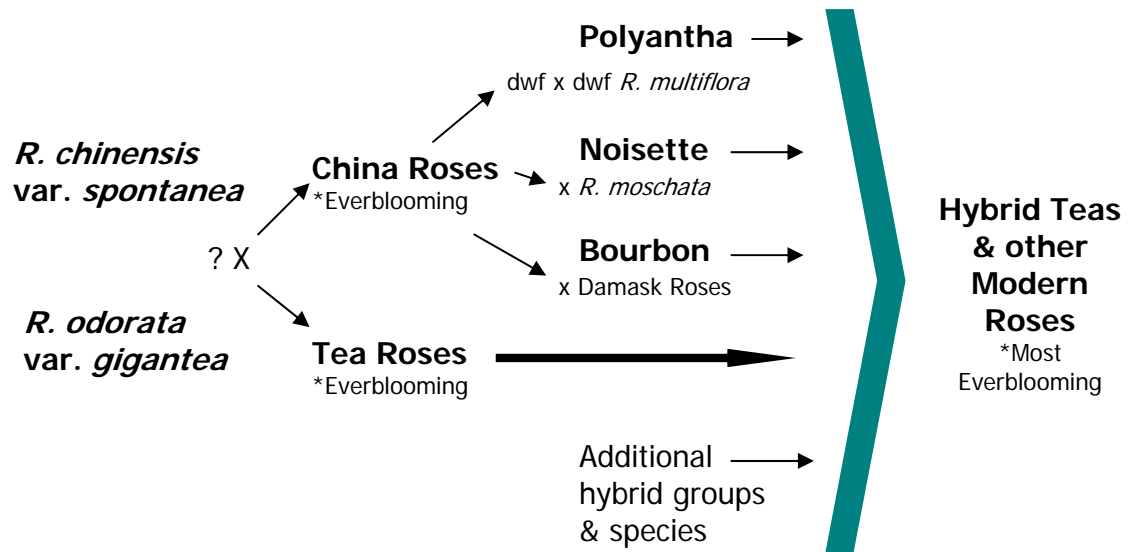
Figure 1. Outline of hybrid rose relationships. This flow chart shows some of the known and suspected hybrids created in *Rosa*, focusing on the China and Tea Roses and the everblooming gene these varieties brought to modern roses.

Because so many of the characteristics of modern roses are inherited from China Roses (Shepherd, 1978), further study is not only useful to satisfy academic curiosity, but to gain a better understanding of the diversity that has been, and may continue to be used in breeding.

The main objectives of this study are:

1) Survey the relative genetic diversity found in a large group of China Roses using microsatellite (SSR) markers.

2) Test the chloroplast intergenic spacer *trnH-psbA* for sequence divergence in hybrid rose cultivars for use in tracing maternal relationships.

3) Use SSR and chloroplast data to investigate the genetic relationships of the China Roses and early hybrid groups, as well as possible species of origin of the China Rose group.

LITERATURE REVIEW

Surveys of genetic diversity and relationships in a group of roses can be a step toward further study and breeding with those that still have the potential to be genetic resources for introgression into economically valuable types. Conventional breeding through selection of F_1 hybrids may not be enough to satisfy the demand for disease resistance, stress tolerance, production parameters, *and* flower quality in modern roses (Debener, et. al, 2004), but increased knowledge in the area of genetic diversity and relationships may be useful for reconstructing a commercial type with emphasis on a needed trait, or finding a less distantly related donor species for more efficient introgression breeding (Debener, 2002).

Previous studies have used molecular markers and other biochemical analysis approaches to study relationships and diversity within the genus *Rosa*. In molecular markers, AFLP, RFLP, RAPD, and microsatellites have all been used to investigate areas including, but not limited to, diversity (Babaei et al., 2007; Kiani, et al., 2008), mapping (Crespel et al., 2001; Debener and Mattiesch, 1999; Dugo et al., 2005), and fingerprinting for the identification of cultivars (Esselink et al., 2003; Hubbard et al., 1992). Scariot et al. (2006) showed that microsatellite markers are useful in describing genetic diversity and relationships among roses. Their choice of SSRs over other markers was based on the high repeatability of analysis, high degree of polymorphism, co-dominant inheritance, and their abundant and relatively uniform distribution in the genome. Some of the research using genetic or biochemical analysis has also included

one or more representatives of section Indicae (syn. Chinenses) (Bruneau et al., 2007; Jan, et al., 1999; Martin et al., 2001; Matsumoto et al., 1998; Millan, et al., 1996; Scariot et al., 2006; Wissemann and Ritz, 2005; Wu et al., 2001), but no diversity studies have been published focusing specifically on the group that includes the China Roses and related species and hybrids.

In order to screen on a wider scale, the above studies included smaller numbers of individuals from a greater number of species and hybrid groups. The results for the most part support current classical taxonomy, as well as previous suggestions for reclassifications in *Rosa*, but have raised additional questions in some areas. An RFLP study of chloroplast DNA revealed different cytoplasm types for accessions of *R. chinensis*, *R. chinensis* 'Mutabilis', *R. chinensis* 'Alba', and *R. gigantea*, suggesting that there are several genetically distinct species in the background of these roses (Takeuchi et al., 2000). It is also worth noting that in that study, the cultivar 'Alba' shared a cytoplasm type with *R. multiflora* accessions, and *R. gigantea* shared a cytoplasm type with *R. damascena* Mill. and *R. moschata* accessions, rather than with the other two members sampled from section Indicae, which formed individual groups. Other studies have shown *R. gigantea* to group nearer to section Synstylae than to Indicae, and Wissemann and Ritz (2005) shows evidence of consectionality of Synstylae and Indicae. The purported hybrid origin of *R. chinensis* varieties and cultivars (other than var. *spontanea*) could lead to various results when only one or two members of the group are used to represent the 'species'. If molecular evidence does indeed suggest a hybrid origin for China Roses, and especially if a species from section Synstylae is within the

complex, this knowledge would be of use in selecting accessions and analyzing the results of future phylogenetic studies.

Methods of analyzing molecular markers are an important consideration in studies including hybrid roses. Because of known hybridization, as well as the additional suspected hybridization that violates assumptions for strict phylogenetic analysis (Bruneau et al., 2007; Koopman et al., 2008), a divergent tree cannot accurately represent the evolution and breeding of the China Roses or other hybrid groups. A 2006 study (Martin et al., 2006) used Principal Component Analysis and a UPGMA (Unweighted Pair Group Method with Arithmetic Mean) dendrogram to analyze the genetic diversity and relationships of roses from some of the founder groups, including China Roses, up through the more modern horticultural classes. The data was in general agreement with current classification and showed a continuous gradient of the European/Chinese allele ratio through domestication. The varying ploidy levels across species and hybrids in roses and many other plants should also be taken into account, since polyploid accessions can not be assigned a copy number from SSR data. Kloda, et al. (2008) demonstrated that Principle Coordinate Analysis (PCoorA) with Euclidian distance is an appropriate tool for studies with related species on different ploidy levels. A UPGMA dendrogram was also used to plot a tree of the data in this study.

The inclusion of data from the chloroplast genome in addition to nuclear markers is also becoming more common in studies, and can be another way to investigate known or possible hybrid origins of plants, since most flowering plants, including *Rosa*, are thought to have maternal inheritance of chloroplasts (Corriveau and Coleman, 1988).

Different methods of surveying the chloroplast genome in *Rosa* have been attempted, but ultimately the level of sequence divergence has been found to be low (Matsumoto et al., 1998; Wissemann and Ritz, 2005). However, a recent study done using the *trnH-psbA* chloroplast intergenic spacer sequence in *Rosa* species showed a higher level of sequence divergence across the genus (Bruneau et al., 2007). This sequence has also been tested in numerous other groups of plants for possible use as part of a DNA barcoding system, and has shown higher than usual sequence divergence as compared to other chloroplast regions (Kress, et al., 2005; Kress and Erickson, 2007).

MATERIALS AND METHODS: MICROSATELLITES

Genotypes Included

Ninety rose accessions were chosen, with 42 from the China Rose group, 18 from sections Indicae and Synstylae representing eight different species, and 30 cultivars from Horticultural Classes that were bred using China Roses. The identities of the roses are listed as given by the source of the plant, tissue sample, or DNA extract, and the Horticultural Classification is as given in Modern Roses XI (Cairns, 2000). An attempt was made to obtain multiple samples of the species roses, from different sources if possible, due to the possibility of misidentification or hybridization when these species are brought into cultivation. The complete list of materials is shown in Table 1.

Table 1. Rose accessions.

Cultivar or species name[x]	Sample number	Source[y]
China Roses, Section Indicae		
Archduke Charles	C1	ARE
Arethusa	C2	ARE
Cramoisi Superieur	C3	ARE
Climbing Cramoisi Superieur	C4	ARE
Ducher	C5	ARE
Green Rose	C6	ARE
Hermosa	C7	ARE
Jean Bach Sisley	C8	ARE
Madame Laurette Messimy	C9	ARE
Martha Gonzales	C10	ARE
Papa Hemeray	C11	ARE
Rouletii	C12	ARE
Serratipetala	C13	ARE
Single Pink China	C14	ARE
Vincent Godsiff	C15	ARE
Old Gay Hill China	C16	ARE
Setina	C17	ARE
Napoleon	C18	ARE
Yue Yue Fen China	C19	TAMU Z03:4:7[z]
Yue Yue Fen 8	C20	TAMU Z03:4:24[z]
Yue Yue Fen 9	C21	TAMU Z03:5:7[z]
Louis Philippe	C22	ARE
Mutabilis	C23	ARE
Spice	C24	ARE
Old Blush	C25	ARE
Climbing Old Blush	C26	ARE
WOB26 (*R. wichuriana* 'Basye's Thornless' x 'Old Blush')	C27	TAMU W04:4:35
Slater's Crimson China	C29	Moore
Pompon de Paris	C30	Moore
Pink Pet	C31	Moore
Fellenberg	C32	ARE
Bengale d'Automne	C33	VG
Bengale Centifeuilles	C34	VG
Miss Lowe's Variety	C35	VG
Purpurea	C36	VG
Single Cerise China	C37	VG
Ferndale Red China	C38	VG
White Pearl in Red Dragon's Mouth	C39	VG
Fabvier	C40	VG
R. chinensis Jacq. var. *spontanea* Redh. & Wils.	CS1	QBG, 1988.237 Sichuan province
R. chinensis var. *spontanea*	CS2	QBG, 2001.226 Sichuan Province
R. chinensis var. *semperflorens* Koehne	CSF	China
Miscellaneous Old Garden Rose		
Fortune's Double Yellow	FDY	ARE

Table 1. Continued

Cultivar or species name[x]	Sample number	Source[y]
Other Section Indicae		
R. odorata var. *gigantea* Collet	OG1	QBG, 2002.065 cultivated form
R. odorata var. *gigantea*	OG2	QBG, 2002.218A cultivated form
R. odorata var. *gigantea*	OG3	China
Section Synstylae		
R. multiflora Thunb. var. *cathayensis* Rehd. & Wils.	MC1	QBG, 2003.382C Sichuan province
R. multiflora	M2	China
R. multiflora	M3	J&P (rootstock)
R. henryi Boulen.	H1	QBG, 1996.018 Sichuan province
R. henryi	H2	QBG, 1996.016 Sichuan province
R. longicuspis var. *longicuspis* Bertol.	LL1	QBG, 1990.206B Yunnan province
R. longicuspis var. *longicuspis*	LL2	QBG, 1992.244 Sichuan
R. longicuspis	LL3	China
R. rubus Lév. & Vaniot	R1	QBG, 1992.071 Sichuan province
R. rubus	R2	China
R. soulieana Crép.	S1	QBG, 1991.190 Sichuan province
R. soulieana	S2	QBG, 2003.438C Sichuan province
R. soulieana	S3	China
R. wichuriana Crép.'Basye's Thornless'	W1	ARE
R. brunonii Lindl.	Br	China
Noisette		
Bouquet d'Or	N1	ARE
Celine Forestier	N2	ARE
Champney's Pink Cluster	N3	ARE
Blush Noisette	N4	ARE
Jaune Desprez	N5	CRN
Jeanne D'Arc	N6	ARE
Lamarque	N7	ARE
Mme. Alfred Carriere	N8	ARE
Marechal Niel	N9	CRN
Manetti	N10	J&P (rootstock)
Polyantha		
White Pet	P1	ARE
Perle d'Or	P2	CRN
Ma Paquerette	P3	Moore
Mignonette	P4	VG
High Country Mignonette	P5	VG

Table 1. Continued

Cultivar or species name[x]	Sample number	Source[y]
Bourbon		
Coquette des Blanches	B1	ARE
Great Western	B2	ARE
La Reine Victoria	B3	ARE
Louise Odier	B4	ARE
Queen of Bourbons	B5	ARE
Zephirine Droughin	B6	ARE
Mrs. Bosanquet	B7	ARE
Souvenir de la Malmaison	B8	ARE
Tea		
Bon Silene	T1	ARE
Climbing Devoniensis	T2	ARE
Isabella Sprunt	T3	ARE
Safrano	T4	ARE
Adam	T5	ARE
Duchesse de Brabant	T6	ARE

[x]Names of accessions are listed as given by the source, and Horticultural Classes are as assigned in Modern Roses XI (Cairns (ed.), 2000). One or more synonyms may exist for many of the old rose varieties.

[y]ARE: Antique Rose Emporium
    9300 Lueckemeyer Rd., Brenham, TX 77833
    http://www.antiqueroseemporium.com
CRN: Chamblee's Rose Nursery
    10926 U.S. Highway 69 North, Tyler TX 75706
    http://www.chambleeroses.com
China: Flower Research Institute
    Yunnan Academy of Agricultural Science
    Kunming, Yunnan, China
J&P: Jackson & Perkins Roses
    2 Floral Ave, Hodges, SC 29653
    http://www.jacksonandperkins.com
Moore: Texas A&M University Material from Ralph Moore
QBG: Quarry Hill Botanical Garden
    12841 Sonoma Hwy, Glen Ellen, CA 95442
    Asian Plants Database: http://quarryhillbg.org/asianplantdatabase.html
TAMU: Basye Rose Breeding & Genetics Program
    Department of Horticultural Sciences
    Texas A&M University, College Station, TX 77843-2133
    http://aggie-horticulture.tamu.edu/rose/index.html
    Following codes indicate planting bed and location.
    [z]These plants were originally obtained as seeds, collected from various Yue Yue Fen China roses of both wild and garden origin.
VG: Vintage Gardens
    4130 Gravenstein Hwy. North
    Sebastopol, CA 95472
    http://www.vintagegardens.com

DNA Extraction

Roses belonging to the Texas A&M University Basye Rose Breeding and Genetics program, and those from locations within Texas were sampled from live plants by collecting young leaf tissue in plastic bags that were transported on ice back to the laboratory. Samples from other sources were shipped as multiple-node stem cuttings wrapped in moist paper towel and sealed in plastic bags, or as dried down, extracted DNA samples. All samples were placed in a -20°C freezer upon arrival.

The DNA extraction method used for the leaf tissue samples was a mini-preparation protocol using CTAB (cationic hexadecyl trimethyl ammonium bromide), with modifications of the method used by Doyle and Doyle (1987), Boonprakob (1996), and Jan (1996).

Approximately 50 mg/sample of frozen leaf tissue were put into 1.5 mL microtubes and cooled by adding liquid nitrogen. A handheld, battery operated power drill was used to grind the tissue with a chilled plastic pestle, sized to fit the tubes, used in place of a drill bit. Seven hundred microliters of 2X CTAB was then added and thoroughly mixed in the tubes. Samples were incubated in a 65°C water bath for 2.5 h and then cooled to room temperature. Next, 700 µL of CIA (chloroform: iso-amyl alcohol, 24:1) was mixed with the samples. Centrifugation was done at 13,200 g_n for 10 min, and repeated if the upper layer of a sample was not yet clear and colorless. Highly pigmented rose varieties did not become as light in color, but were relatively clear by the end of the extraction. The upper layer of the samples was then transferred to a new 1.5 mL tube and combined with another 700 µL of CIA. The above centrifugation steps

were repeated. The upper aqueous layer was again transferred to a new tube, and then

mixed with chilled isopropanol. After inverting the tubes several times, the samples

were placed in a -20°C freezer overnight. The following day, samples were centrifuged

at 6,000 g_n for 10 min. The supernatant was then poured off, taking care not to dislodge

the pellet at the bottom of the microtube. Samples were allowed to dry briefly, followed

by a 70% ethanol rinse, which was repeated once, with a short centrifugation in between

to ensure that the pellet would be retained in the tube. After the last ethanol rinse was

poured off, samples were allowed to dry completely at room temperature. To resuspend

the samples, 50-200 µL of TE buffer/sample was used. Samples were vortexed for about

10 min, or until the pellet was completely dissolved. DNA samples were quantified

using a DQ 300 fluorometer (Hoefer, Inc.), and stored at -20°C.

Microsatellite Primers

Thirty SSRs were selected from 40 screened for polymorphism on 3% MetaPhor

(Lonza) agarose gel in a subset of eight accessions from the project. The forward

primers for those 30 SSRs were then ordered with FAM, HEX, or NED fluorescent

labels. Twenty-three SSRs (Table 2) amplified well across the roses included in the

study and could be accurately scored following capillary electrophoresis.

Table 2. SSR primer pairs.

Primer code[y]	Reference; Linkage Group[z]	Supplier
Rw34L6-F	Hibrand-Saint Oyant et al., 2008; 1	Bioneer, Inc.
Reverse		Invitrogen Corp.
CTG21F	Hibrand-Saint Oyant et al., 2008; 1	Bioneer, Inc.
R		Invitrogen, Corp.
Rw59A12-F	Hibrand-Saint Oyant et al., 2008; 2	Applied Biosystems, Inc.
-R		Invitrogen, Corp.
Contig172-F	Hibrand-Saint Oyant et al., 2008; 2	Bioneer, Inc.
-R		Invitrogen Corp.
Rw53O21-F	Zhang et al., 2006; 3	Bioneer, Inc.
-R		Invitrogen Corp.
Rw55E12-F	Hibrand-Saint Oyant et al., 2008; 3	Applied Biosystems, Inc.
-R		Invitrogen Corp.
Rh72-F	Yan et al., 2005; 7	Bioneer, Inc.
-R		Invitrogen, Corp.
BFACT47-F	Rousseau-Gueutin et al., 2008; 4	Applied Biosystems, Inc.
-R		Bioneer, Inc.
Rw14H21-F/R	Zhang et al., 2006; 6	Eurofins MWG Operon
Rw5G14-F	Hibrand-Saint Oyant et al., 2008; 5	Bioneer, Inc.
-R		Invitrogen Corp.
H24D11-F	Hibrand-Saint Oyant et al., 2008; 6	Applied Biosystems, Inc.
-R		Invitrogen Corp.
CL2980-F	Hibrand-Saint Oyant et al., 2008; 7	Bioneer, Inc.
-R		Invitrogen Corp.
CTG623-F/R	Hibrand-Saint Oyant et al., 2008; 7	Bioneer, Inc.
H10D03-F/R	Hibrand-Saint Oyant et al., 2008; 5	Bioneer, Inc.
Rh58-F	Yan et al., 2005; 3	Applied Biosystems, Inc.
-R		Invitrogen Corp.
Rw35C24-F	Hibrand-Saint Oyant et al., 2008; 4	Applied Biosystems, Inc.
-R		Invitrogen Corp.
Rw5D11-F/R	Zhang et al., 2006	Eurofins MWG Operon
RhEO506-F	Esselink et al., 2003; 2	Bioneer, Inc.
-R		Eurofins MWG Operon
RhD221-F	Esselink et al., 2003; 4	Applied Biosystems, Inc.
-R		Eurofins MWG Operon
RhAB26-F	Esselink et al., 2003	Bioneer, Inc.
-R		Eurofins MWG Operon
RhB303-F/R	Esselink et al., 2003	Eurofins MWG Operon
RhAB13-F	Esselink et al., 2003; 4	Applied Biosystems, Inc.
-R		Eurofins MWG Operon
Rw29B1-F	Zhang et al., 2006	Bioneer, Inc.
-R		Eurofins MWG Operon

[y]-F = forward primer, -R = reverse primer
[z] Rose Linkage Group 1-7 is given if published

PCR Amplification

Amplifications of the markers were done with a 10 µL total reaction/sample that included 5.0 µL Phusion Flash Master Mix (New England BioLabs, Inc.), 3.0 µL pure water, 0.5 µL each, forward and reverse primer (10 µM stock), and 1 µL DNA (~10 ηg/µL stock). Thermal cycling was done on a Techne TC-412 (Barloworld Scientific, Ltd.) with the following program: 105°C heated lid, 98°C hot start, initial denaturation of 98°C for 10 s, and 30 cycles of 98°C for 1 s : 55°C for 5 s : 72°C for 1 min, finishing with a final extension of 72°C for 1 min and a final hold at 4°C. Presence of product was confirmed on an agarose gel, and samples were stored at -20°C. Accessions showing no amplification for a particular primer were repeated twice before being assigned as missing data.

Capillary Electrophoresis

Capillary electrophoresis was done on an ABI 3130 Genetic Analyzer (Applied Biosystems, Inc.). PCR products for the same accession were pooled into groups of three with one of each dye label (FAM, HEX, and NED), with little or no overlap in allele sizes, as determined by the screening gels. Since the concentration of the PCR products were too high to use directly, and different amounts of product for each florescent dye were needed to obtain readable peaks, a dilution series was done. Results within a useable range were obtained across all DNA and marker combinations by making a dilution plate with 1.0 µL of the FAM product and 1.2 µL each of the HEX and NED products in 130 µL of pure water. After mixing thoroughly, 1.0 µL of this

dilution was added to 4.5 µL of Rox-HDF in a 96-well semi-skirted capillary plate

(various manufacturers). The Rox-HDF contained 850 µL of Hi-Di Formamide and 50

µL Genescan 400HD [ROX] (both from Applied Biosystems, Inc.). Samples were

stored up to 24 hours in a -20°C freezer before analysis. Immediately before analysis on

the ABI 3130 Genetic Analyzer, the plates were denatured on the same TC-413 Thermal

Cycler used for PCR. The program used was: heated lid at 105°C, denature at 95°C for

5 min, and hold at 4°C. The ABI 3130's Fragment Analysis POP7 protocol was used.

Fragment Data Analysis

Software used to assign peaks from the ABI 3130 output files included

GeneMapper v. 4.0, and Peak Scanner v. 1.0 (both from Applied Biosystems, Inc.).

Ploidy level was also assigned to accessions based on the number of alleles observed per

marker for the SSRs. Once the ploidy levels were established, each diploid accession's

number of heterzygous loci was divided by the number of markers for which data was

obtained for that sample to calculate individual heterozygosity. The polyploid

individuals were not included because of the lack of knowledge of the copy number of

each allele.

NTSYS-pc v. 2.2 (Numerical Taxonomy and Multivariate Analysis System,

Exeter Software) was used for cluster analysis of the accessions to help visualize

relationships. A rectangular data set was first created from the SSR data in Microsoft

Excel (Microsoft Corp.) with the alleles as the rows, and the accessions as the columns,

using "1" for presence of an allele and "0" for absence. Dice coefficient for similarity

and UPGMA clustering were chosen to create a dendrogram of the data. In an attempt to help take into account the different ploidy levels included in this study, Principle Coordinate Analysis was also performed to look at the most significant factors in the molecular variation. Euclidian distance was chosen to de-emphasize shared absence of alleles, and the principle coordinates were graphed on a three-dimensional plot.

MATERIALS AND METHODS: CHLOROPLAST SPACER SEQUENCE

Genotypes Included

The roses used for this portion of the study included the China Roses, Miscellaneous Old Garden Rose, Section Indicae, and Section Synstylae plants (see Table 1). The DNA extracts were from the same stocks used to obtain the SSR data.

Chloroplast Intergenic Spacer

The chloroplast intergenic spacer *trnH-psbA* (forward primer aliquots provided by Dr. Alan Pepper, Department of Biology, Texas A&M University according to primer sequences referenced in Kress et al., 2005) was chosen because of significant species-level variability, short length for DNA extraction and amplification, and conserved flanking sites for universal primers in the taxa sampled to date (Kress, et al., 2005; Kress and Erickson, 2007). This sequence has also been used in a genus-wide phylogenetic study of *Rosa*, in which it was amplified and sequenced in all accessions (Bruneau et al., 2007).

PCR Amplification

Amplification of the intergenic spacer was done with a 15 µL total reaction/sample that included 7.5 µL Phusion Flash Master Mix, 4.5 µL pure water, 0.75 µL each of the primers *trnH* and *psbA* (50 ηg/µL stocks), and 1.5 µL DNA (~10 ηg /µL stock). Thermal cycling was done using the same program as for the SSRs in this study.

Gel screening was done to confirm amplification, and all samples were amplified and

showed heavy bands under ethidium bromide staining. Samples were stored at -20°C.

PCR Cleanup

The PCR products were cleaned before continuing with the sequencing process

using one of two methods, following manufacturers' instructions: ExcelaPure 96-well

UF PCR Purification Kit (Edge Biosystems) or ExoSAP-IT (USB Corp.).

Sequencing Reaction

Sequencing reactions were performed using a total reaction volume of 5.0 µL,

with 2.0 µL ABI BigDye Terminator Sequencing Kit (v. 1.1 Applied Biosystems, Inc.,

supplied from the Gene Technologies Laboratory, Texas A&M University), 1.0 µL

forward primer (*psbA* 12.5 ηg /µL, diluted from the concentrated stock), and 2.0 µL PCR

product (cleaned, 30-60 ηg /µL). The cycling parameters used on the TC-413 were

heated lid at 105°C, initial denaturation at 96°C for 1 min, 60 cycles of 95°C for 30 s :

55°C for 15 s : 60°C for 4 min, and a final hold at 4°C. The process was repeated for the

trnH primer to obtain a consensus sequence for each sample.

Dye Terminator Removal

Sequenced samples were cleaned prior to running on the ABI 3130 using a

Performa DTR V3 96-well Short Plate Kit (Edge Biosystems), according to

manufacturer's instructions. Samples were dried down by centrifuging at 1,000 g_n for 30

min, gently decanting the liquid, and spinning briefly to < 100 g_n with the 96-well plate

upside down on a paper towel. An additional 5-10 min of air drying in darkness was

required before sealing and storing the samples at -20°C.

Capillary Electrophoresis

Dried samples were resuspended in 10 µL of Hi-Di Formamide and run on the

ABI 3130 Genetic Analyzer using the Sequencing POP7 protocol for BigDye v. 1.1.

Sequence Data Analysis

Sequences were prepared for further analysis using Sequencher v. 4.8 for PC

(Gene Codes Corp.). Sequence for each sample from primers *trnH* and *psbA* were

aligned, checked for accuracy in sections with mid to low quality ratings, and trimmed

on the ends to remove the primers. The longest aligned sequence was 390-bp in length

from primer to primer. However, only cultivar accession C31 reached this length

because of insertion-deletion points (indels) present only in that sample and causing gaps

in the others, making the rest of the individual samples 374-bp or less. When aligned

pair-wise with samples from the same species, these sequences were 82bp longer (+/-

indels) than the sequence available for the accessions examined by Bruneau et al. (2007).

Sequence was obtained in the entire section from primer to primer in all samples except

accessions C2 and Br, which only had high quality sequence in one direction, and were

14- and 15-bp short on one end, respectively. In both cases the quality of the DNA stock

for the repeat attempts was suspect, and new PCR product for sequencing was not

obtained. The missing sections on the ends of the sequence showed no divergence in any of the other samples, so all samples were trimmed to the length of the shortest before further analysis.

The aligned, shortened consensus sequence was 375-bp in length, as compared to the Bruneau et al. study (2007), which had an aligned length of 503-bp in the genus *Rosa* (including an outgroup). The same study also reported being unable to reliably align nearly half of the sequence data, so over 43% of the sequence was excluded from further analysis. This issue was not observed in this study, as only two accessions had indels leaving gaps in all other accessions (a total of four indels were found), whereas nine indels were found in the 2007 study. That increased number of indels across such a large representation of the genus (70 taxa) would indeed make alignment time-consuming and unreliable, so this may affect future use of *trnH-psbA* in *Rosa* and is discussed further in the Results section of this study. The explanation for the ease of alignment in this study is likely the tighter focus of the study around China Roses and their hybrids and possible progenitor species, which represent an economically important, but relatively small portion of the genus.

The compiled sequences were analyzed using PAUP v. 4.0b10 for Macintosh (Phylogenetic Analysis Using Parsimony, Sinauer Associates). The 375-bp alignments were first manually checked, and a mononucleotide microsatellite (occurring in all samples) and two sections with indels were excluded from the PAUP analysis. The three exclusions were from base pairs 131-146, 189-191, and 195-214. Distance was calculated, and the Neighbor-Joining method was used. An unrooted phylogram of the

data was produced in PAUP to graph the main haplotypes of the accessions based on

single nucleotide changes in the sequence. To take the microsatellite into consideration

and allow for the separation of more species and groups, the region was analyzed

manually and the number of repeats were indicated on the phylogram generated by the

PAUP analysis.

RESULTS AND DISCUSSION

This study used microsatellite markers and a chloroplast intergenic spacer sequence to investigate genetic diversity and relationships in the China Rose group, their early hybrid groups, and possible contributing species.

Results and discussion will start with the general microsatellite analysis results, followed by the implications of the SSR dendrogram clustering and Principle Coordinate Analysis, and the chloroplast haplotypes found within the China Roses and species included in this study.

General SSR Analysis Results

The 23 microsatellite primers included in this study produced 291 alleles scored across the 90 accessions. The average number of alleles per primer was 12.65, and the range was from six (CTG21, Contig172, and BFACT47) to 22 (RhAB26) (Table 3). All 291 alleles were polymorphic across the plants sampled. In other rose studies, Tang et al. reported a range of 6-14 alleles per primer with an average of 8.2 across 42 accessions of species, Old Garden Roses, and Yunnan cultivars (2008). Scariot et al. reported 6-21 alleles per primer with an average of 13.7 across 65 Old Garden Roses and species (2006). In 2001, a study of 142 Malus species, hybrids, and cultivars reported a range of 6-40 alleles per primer with an average of 26.4 (Hokanson et al., 2001). Comparing these SSR studies in the Rosaceae family, it can be seen that similar ranges and averages were found for the number of alleles per primer, with differences

Table 3. SSR primers: Number of alleles and sizes.

Primer	No. of alleles	Sizes in bp
Rw34L6	13	176, 180, 182, 186, 197, 200, 202, 207, 209, 213, 217, 225, 234
CTG21	6	120, 123, 127, 129, 131, 133
Rw59A12	16	201, 204, 208, 211, 214, 216, 128, 220, 222, 224, 227, 230, 232, 235, 245, 257
Contig172	6	139, 142, 145, 148, 152, 154
Rw53O21	7	155, 158, 161, 164, 167, 171, 174
Rw55E12	16	115, 153, 155, 160, 164, 166, 168, 172, 174, 176, 178, 180, 182, 186, 192, 247
Rh72	19	240, 246, 250, 252, 254, 259, 261, 266, 270, 275, 277, 281, 284, 287, 291, 293, 296, 301, 308
BFACT47	6	144, 147, 151, 156, 163, 168
Rw14H21	15	110, 114, 117, 119, 121, 123, 125, 128, 131, 136, 138, 143, 146, 148, 153
Rw5G14	8	226, 228, 230, 233, 236, 239, 242, 249
H24D11	12	141, 148, 151, 154, 157, 160, 164, 166, 168, 172, 176, 182
CL2980	10	210, 213, 219, 221, 223, 225, 227, 229, 233, 235
CTG623	15	199, 209, 211, 214, 217, 219, 223, 227, 230, 233, 236, 241, 248, 251, 255
H10D03	14	207, 213, 215, 217, 219, 221, 223, 225, 227, 231, 233, 236, 238, 252
Rh58	16	224, 230, 239, 243, 246, 248, 250, 252, 256, 259, 262, 264, 269, 275, 279, 288
Rw35C24	14	244, 246, 248, 250, 252, 254, 256, 259, 261, 265, 268, 272, 275, 283
Rw5D11	13	214, 217, 221, 225, 228, 231, 234, 237, 240, 243, 246, 250, 253
RhEO506	12	207, 210, 213, 219, 221, 225, 227, 230, 233, 239, 251, 254
RhD221	12	202, 206, 211, 213, 215, 220, 222, 226, 230, 232, 235, 238
RhAB26	22	159, 164, 168, 170, 172, 174, 177, 180, 184, 186, 193, 198, 201, 206, 224, 230, 240, 245, 249, 254, 275, 296
RhB303	12	115, 117, 120, 122, 124, 126, 128, 130, 132, 138, 142, 145
RhAB13	18	133, 138, 139, 141, 144, 146, 148, 151, 153, 158, 162, 164, 166, 169, 171, 173, 177, 181
Rw29B1	9	340, 343, 347, 350, 352, 354, 356, 361, 365

correlating with the number of accessions included in the study, and whether the accessions represented a particular group/s in the genus or the genus as a whole.

The ploidy level of the accessions was calculated based on the SSR data (Table 4). Sixty-three of the accessions (70%) were found to be diploid, with 21 (23%) triploid, and six (6.7%) tetraploid. Among the roses classified as China Roses, 28 were diploid (67%) and 14 were triploid (33%). Some of the accessions had previously published ploidy levels, and the calculations here were in agreement with those (Cairns (ed.), 2000; Roberts (eds.) et. al, 2003).

In the diploids, where the copy number of each allele is known, the percentage of heterozygous loci of individual plants was also calculated and ranged from 30% for the Polyantha cultivar 'Ma Paquerette', to 87% in the China Rose cultivar Single Cerise China (Table 4). The average heterozygosity of the diploid accessions was 67%. Only nine of the 63 diploids (CS1, CS2, OG1, OG2, R2, S1, P3, P4, and P5) had 50% or less heterozygous loci. If the botanical species and varieties are removed from consideration, the average heterozygosity for the diploids increases to 72%. A high level of heterozygosity was expected due to the complex hybrid nature of many of the accessions, and if it could be judged accurately for the polyploid accessions as well, may have been even higher.

Molecular marker studies in roses and other plants focusing on the development of primers often calculate heterozygosity for each of the loci being tested (for example, Kimura et al., 2006), but studies have also been done that use various methods of estimating heterozygosity of individuals based on molecular markers to help describe

Table 4. Ploidy level and heterozygosity.

Ploidy level	Sample	% heterozygous loci[z]
Diploid		
	C1, Archduke Charles	74
	C2, Arethusa	70
	C5, Ducher	57
	C6, Green Rose	74
	C8, Jean Bach Sisley	83
	C9, Madame Laurette Messimy	78
	C11, Papa Hemeray	74
	C12, Rouletii	74
	C14, Single Pink China	74
	C15, Vincent Godsiff	82
	C19, Yue Yue Fen China	74
	C20, Yue Yue Fen 8	70
	C21, Yue Yue Fen 9	70
	C23, Mutabilis	74
	C24, Spice	78
	C25, Old Blush	74
	C26, Climbing Old Blush	74
	C27, WOB26	78
	C30, Pompon de Paris	74
	C31, Pink Pet	61
	C33, Bengale d'Automne	74
	C35, Miss Lowe's Variety	70
	C36, Purpurea	68
	C37, Single Cerise China	87
	C39, White Pearl in Red Dragon's Mouth	64
	CS1, *R. chinensis* var. *spontanea*	43
	CS2, *R. chinensis* var. *spontanea*	36
	CSF, *R. chinensis* var. *semperflorens*	74
	FDY, Fortune's Double Yellow	65
	OG1, *R. odorata var. gigantea*	35
	OG2, *R. odorata var. gigantea*	48
	OG3, *R. odorata var. gigantea*	78
	MC1, *R. multiflora* var. *cathayensis*	64
	M2, *R. multiflora*	57
	M3, *R. multiflora* (rootstsock)	65
	H1, *R. henryi*	78
	H2, *R. henryi*	73
	LL1, *R. longicuspis* var. *longicuspis*	59
	LL2, *R. longicuspis* var. *longicuspis*	57
	LL3, *R. longicuspis*	59
	R1, *R. rubus*	55
	R2, *R. rubus*	43
	S1, *R. soulieana*	50
	S2, *R. soulieana*	55
	S3, *R. soulieana*	77
	W1, *R. wichuriana* 'Basye's Thornless'	65
	Br, *R. brunonii*	55
	N2, Celine Forestier	77
	N3, Champney's Pink Cluster	78

Table 4. Continued

Ploidy level	Sample	% heterozygous loci[z]
Diploid		
	N4, Blush Noisette	78
	N5, Jaune Desprez	83
	N6, Jeanne D'Arc	83
	N9, Marechal Niel	74
	P1, White Pet	78
	P2, Perle d'Or	78
	P3, Ma Paquerette	30
	P4, Mignonette	39
	P5, High Country Mignonette	43
	B7, Mrs. Bosanquet	74
	T1, Bon Silene	64
	T3, Isabella Sprunt	83
	T4, Safrano	83
	T6, Duchesse de Brabant	64
Triploid		
	C3, Cramoisi Superieur	
	C4, Climbing Cramoisi Superieur	
	C7, Hermosa	
	C10, Martha Gonzales	
	C13, Serratipetala	
	C16, Old Gay Hill China	
	C17, Setina	
	C18, Napoleon	
	C22, Louis Philippe	
	C29, Slater's Crimson China	
	C32, Fellenberg	
	C34, Bengale Centifeuilles	
	C38, Ferndale Red China	
	C40, Fabvier	
	B6, Zephirine Drouhin	
	B8, Souvenir de la Malmaison	
	N1, Bouquet d'Or	
	N7, Lamarque	
	N8, Mme. Alfred Carriere	
	T2, Climbing Devoniensis	
	T5, Adam	
Tetraploid		
	B1, Coquette des Blanches	
	B2, Great Western	
	B3, La Reine Victoria	
	B4, Louise Odier	
	B5, Queen of Bourbons	
	N10, Manetti	

[z]Diploid accessions only

germplasm diversity. In *Malus* x *domestica*, heterozygosity estimates for two cultivars

were done using both AFLPs and SSRs. The apple cultivars had similar levels of

heterozygosity, but the magnitude of the heterozygosity was very different between the

AFLP method measured by the segregation in progeny (18.4 and 18.9%) and the SSR-

based estimates (82.6 and 87.0%) (Kenis and Keulemans, 2005). The authors stated that

this difference was expected with AFLPs because segregation is not detected in F_1

progeny when both parents have the same fragment, so it should be considered a

minimum estimate of heterozygosity. They also noted that SSR estimates could be

higher because they measure hyper variable non-coding regions, so this was taken into

consideration when comparing the different estimates of heterozygosity. Domesticated

apples are a hybrid group with strong self-incompatibility mechanisms that make them

highly heterozygous (Kenis and Keulemans, 2005), and comparing the SSR based results

for the *Malus* cultivars to those of the diploid non-species roses in this study shows that

they are quite heterozygous as well, with several accessions between 82% and 87% and

an average of 72% heterozygous loci for the diploid cultivars.

In *Rosa*, segregation of AFLP fragments in progeny has also been used to

estimate the heterozygosity of two species roses. An accession of *R. wichuriana* had an

estimated heterozygosity of 94%, and an *R. rugosa* had an estimated 41% heterozygosity

(Crespel et al., 2001). The authors noted the disparity between these two levels of

heterozygosity in supposedly wild species, and hypothesized that the cultivated origins

of the samples could help to explain it. The garden origin of the *R. wichuriana* could

mean that it was a hybrid, making it more heterozygous, and the select clone from

Meilland used for the *R. rugosa* accession could have been inbred and not accurately represent the wild species (Crespel et al., 2001). The level of heterozygosity for *R. wichuriana* in this study ('Basye's Thornless', 65%) was not as high as in Crespel et al., but drawing meaningful conclusions from comparisons is difficult, due to the different methods of estimating heterozygosity, and the questions on origins of the species.

SSR Dendrogram Clustering

As can be seen from the dendrogram, the species roses used in this study do not strictly cluster into groups by species, nor appear to be very closely related to the China Roses or other cultivars, but rather are more loosely grouped with Section Synstylae together at a very low coefficient of similarity as compared to the groups of cultivars in this study (Figure 2). The cultivated varieties clustered into more well-defined groups at around 0.45 similarity, and groups generally reflected the American Rose Society approved horticultural classifications listed in Modern Roses XI (Cairns (ed.), 2000). The genetic diversity of the group as a whole proved to be high, with similarity coefficients for non-identical samples ranging from ~0.15-0.99.

China Roses

Starting from the top of the dendrogram, the first group of interest in the figure, and the focus of the study, is the China Roses. They are included within the largest cluster, which also contains a sub-cluster of Tea type roses. Based on the similarity coefficients, the cultivated China Roses as a group are closely related, as are the roses in the Tea Type group, and the genetic diversity is lower for these groups with a long

Figure 2. UPGMA dendrogram of SSR data. This dendrogram was produced using the Unweighted Pair Group Method with Arithmetic Mean clustering from the Dice similarities of the SSR data. The * denotes the largest cluster containing the Chinas and Tea Types, and the main groups of interest are named near the top node of the cluster.
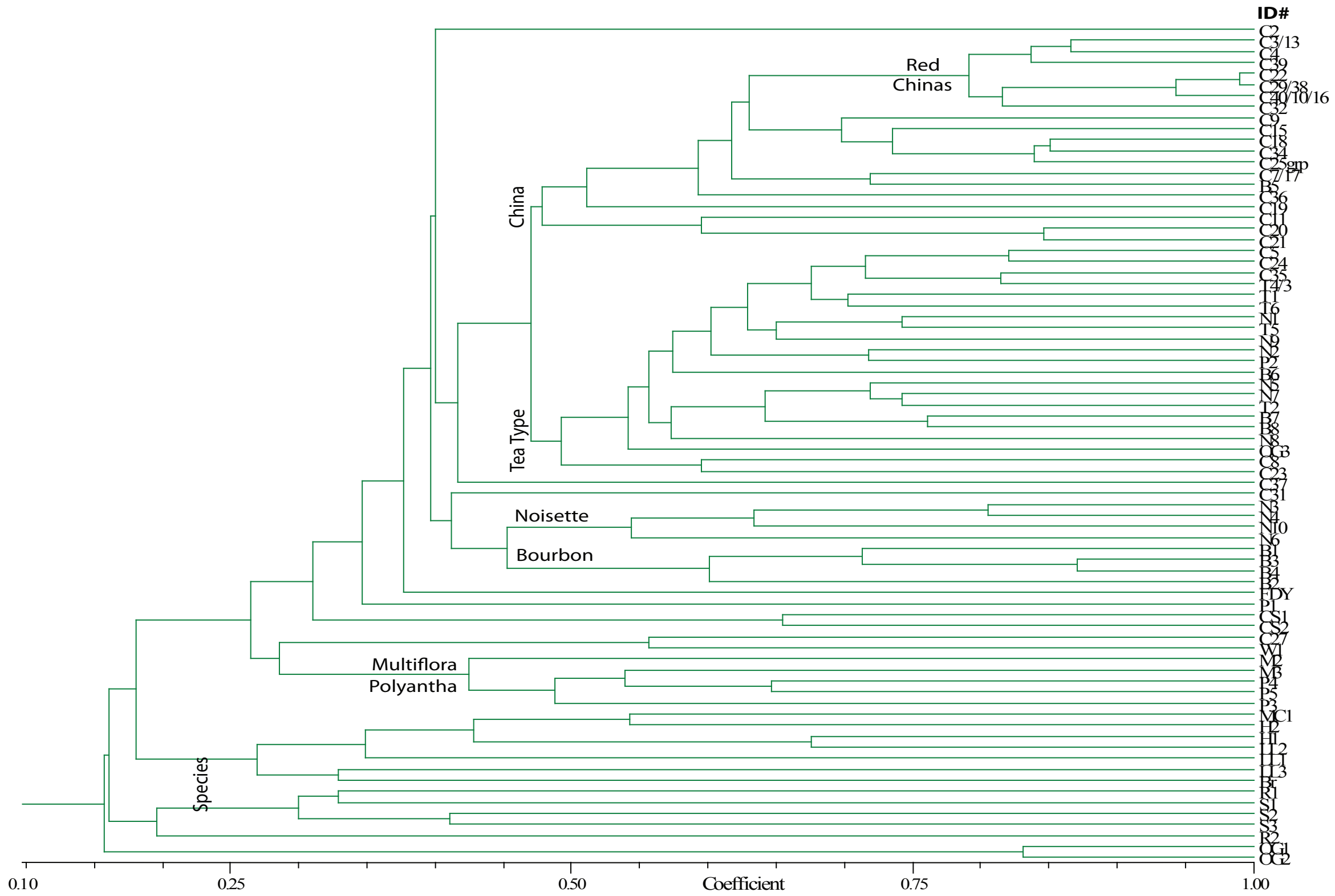
history of cultivation and selection in China and again in Europe with the limited group of cultivars introduced into western breeding.

The dendrogram clusters generally conformed to current classification, and the only accession to group in the China sub-cluster with a horticultural classification other than China or Hybrid China was B5, the Bourbon Rose 'Queen of Bourbons'. This does not necessarily imply mislabeling of the sample, or that this rose does not belong in the Bourbon class for horticultural purposes. Instead, it adds support to the historical records indicating that Bourbon Roses started from a cross speculated to be between a China Rose and a Damask (Harkness, 1978), and that this cultivar may have heavier China influence in its background. 'Queen of Bourbons' shares a 0.72 similarity coefficient with 'Hermosa' and 'Setina' (C7 and C17), the cultivars it is linked to in the dendrogram, but it also has strong affinity to the other Bourbons in the study, with 0.71 similarity to 'Louise Odier' (B4). Roses like this with complex backgrounds make classification difficult because of the different reasons for classifying roses. A classification that may be useful in the horticultural trade may not reflect the actual parentage and genetics behind the roses that is important to breeders and scientists. Since the focus of this study is genetic diversity and relationships, these kinds of divergence from current classification represent useful information for breeders concerned with not only the phenotype of the rose, but the genetics that would be passed on to progeny.

Also of interest within the China Roses sub-cluster is the group of 11 accessions that cluster at just over 0.75 similarity (see Red Chinas cluster in Figure 2). This group

with so many shared alleles is composed exclusively of China Roses that have semi-double to double red blooms, and excludes only a few other China Roses in this study that fit that same general morphological description (see Cairns (ed.), 2000 for descriptions). The presence of this group is not unexpected, since the red color exhibited in some of the China Roses had not been seen before in European roses prior to the introduction of China Roses (Shepherd, 1978), and would make these rare roses popular for both vegetative and seed propagation.

Perhaps the most easily noted detail revealed about the China Roses by the similarity matrix though, was how many accessions had identical SSR profiles. The "C25 grp" on the dendrogram represents the China Rose cultivar Old Blush, with the group's oldest recorded date of introduction into Europe of around 1752 (Cairns (ed.), 2000), and the eight synonyms or sports found in this study: 'Climbing Old Blush', 'Green Rose', 'Single Pink', 'Rouletii', 'Pompon de Paris', 'Bengale d'Automne', 'Archduke Charles', and an *R. chinensis* var. *semperflorens.* This large group of synonyms and sports still being actively propagated and sold in the trade demonstrates how important 'Old Blush' continues to be, long after being used as a parent of importance in the breeding of modern roses.

'Climbing Old Blush' was already known to be a climbing sport of 'Old Blush', and the 'Green Rose' or 'Viridiflora' had been speculated to be another sport, and was also found to be genetically identical in several previous studies (Martin, et al., 2001; Scariot, et al., 2006). The cultivar Single Pink was shown to be identical to 'Old Blush' in this study as well, supporting the hypothesis that it is a single flowered sport of 'Old

Blush' (Phillips and Rix, 1988), or visa versa. 'Rouletii', which is a miniature with double pink recurrent blooms (Cairns (ed.), 2000), appears to be a dwarf sport of 'Old Blush'. 'Pompon de Paris' has also been described as appearing to be identical to at least one source of plants identified as 'Rouletii', and for the accessions sampled in this study, these two cultivars and 'Old Blush' share the same SSR profile. Only one source of 'Rouletii' was sampled for this study though, so a future study combining field observation and genetics of multiple sources and their relationship to 'Old Blush' might be of interest to breeders, producers, and growers.

'Bengale d'Automne' may be another case of unclear identity or multiple names for the same rose. The rose sampled in this study was identical to 'Old Blush', and though propagated as a separate cultivar, the source was not convinced it could be distinguished from 'Old Blush', but was perhaps a more refined form of it (see Table 1 for information on source). 'Archduke Charles' also looks reminiscent of 'Old Blush', and shares the characteristic pale pink blushing, changing to much darker pink flowers, but with more exaggerated colors and a higher petal count (Dickerson, 1992). 'Archduke Charles' has been referenced as a possible seedling of 'Old Blush', but the accession in this study had the same SSR profile as Old Blush, and would therefore be a sport.

The accession of *R. chinensis* var. *semperflorens* in this study was in the 'Old Blush' group based on the SSR data as well, so assuming correct collection and labeling, this example of the red variety of *R. chinensis* appears to be a flower color sport of 'Old Blush' or vice versa. It is also notable that this particular specimen is not identical to the

specimen of 'Slater's Crimson China' (C29) used in this study, though they are sometimes cross-referenced because these names have been used interchangeably at times in history (Dickerson, 1992). In addition, these samples of *R. chinensis* var. *semperflorens* and 'Slater's Crimson China' were different ploidy levels: diploid and triploid, respectively. A rose that did prove to have the same profile as 'Slater's Crimson China' was the found rose 'Ferndale Red China' (C38), so it seems that rose has found its identity. However, there is more than one plant identified in the trade as 'Slater's Crimson China' (Piola, et al., 2002), so testing multiple sources could investigate the different clones in the trade, but would still not be able to say with certainty which were the original cultivar.

Another group of suspected synonyms and sports that the results of this study support is a red, semi-double flowered group based on the China Rose cultivar Fabvier (syn. 'Colonel Fabvier'). 'Fabvier', 'Martha Gonzales', and 'Old Gay Hill Red China' (C40, C10, and C16) all had identical SSR profiles. Based on catalogue descriptions of the plants in commerce (see Table 1 for information on sources), the results suggest that the found rose 'Martha Gonzales' is actually 'Fabvier', and that 'Old Gay Hill Red China', also a found rose, is a taller growth form sport of 'Fabvier'.

The dendrogram based on the SSR data also helps show the relationship of another group of red China Roses. Though both accessions C4 and C3 are triploid and rarely set hips, the double flowered rose 'Climbing Cramoisi Superieur', is said to be a seedling of 'Cramoisi Superieur' (Dickerson, 1992). The similarity of 0.87 found in this study does indeed support seedling status or other close inbred relationship, rather than

the more common situation where a climbing variety named for another cultivar is the vegetative sport of that cultivar. A rose that does appear to be a sport of 'Cramoisi Superieur' based on the SSR data is the fringe-petaled 'Serratipetala' (C13), which is also a triploid, red double-flowered rose.

The results indicate that 'Setina' (syn. 'Climbing Hermosa', C17) is the sport of 'Hermosa' (C7), and many references list this as the parentage, even in cases where these roses were grouped into the Bourbon class rather than China or Hybrid China (Cairns (ed.), 2000; Dickerson, 1992).

Tea Types

The other related group in the first cluster formed a sub-cluster consisting of plants with a variety of current horticultural classifications; however, when reviewing the cultivars included in the cluster, it can be seen that they are generally of the Tea Rose type. All of the roses in the study classified as Tea Roses were included in this cluster, as well as one accession of *R. odorata* var. *gigantea* (OG3) sent directly from China. It does not seem unlikely that OG3 might group here, since it is thought to be one of the original parents of the first Tea Roses (Harkness, 1978), but what is odd is its low similarity to the other accessions said to be of this same species in cultivation in the United States (OG1 and OG2). OG1 and OG2 group with each other at a similarity of 0.83, while OG3 only has 0.19 similarity with OG1, and 0.21 with OG2. Another recent study found that an *R. odorata* var. *gigantea* and *R. odorata* var. *erubescens* had an SSR-based similarity coefficient of ~0.64 (Tang et al., 2008). The large difference in similarity levels could be caused by several factors. It is possible that the Chinese

source used in this study may more closely represent the actual species in its wild

condition, and that the U.S. source of this species, which is of cultivated origin and far

removed from collection, may be the seed grown results of a natural or artificial hybrid

rather than representative of the wild *R. odorata* var. *gigantea*. The opposite could also

be true though, with the U.S. source (35 & 48% heterozygous loci) being cultivated

offspring of a wild representative of the species, and OG3 (78% heterozygous loci) a

cultivated variety of *R.* x *odorata* and an example of one of the first Tea Roses bred in

China, rather than the wild species native to that country. The chloroplast sequence data

gives additional information on these accessions.

The rest of the roses included in the Tea Type group are listed in the horticultural

classes of China/Hybrid China, Noisette, Bourbon, and Polyantha. All of these roses

have one or both of the following key features. The first is known parentage that

includes a rose classified as a Tea (Cairns (ed.), 2000), or unknown parentage, which

would suggest that a Tea was included, whether or not the accession's phenotype is of

another classification. The other is various past and present rose nursery catalogues

where these roses have added notes about their classification. Some may have been

listed under the current American Rose Society class, but with descriptions commenting

that they are very Tea-like, or they may simply be listed in a catalogue section for Teas

or Tea-Noisettes instead. These altered classifications and added categories are usually

implemented because the nursery feels those roses' appearances and how they grow in

the garden better fit in the altered category. Regardless of what horticultural class these

roses are considered to belong to for other purposes, knowing their affinity to the Tea

roses, despite their classification or phenotype, is useful for geneticists and breeders working with and planning crosses using these roses. The China/Hybrid China Roses grouped here will be discussed in greater detail with the results of the chloroplast spacer sequence.

Noisettes and Bourbons

The next defined cluster of interest contains the rest of the Bourbon and Noisette Roses, and the cluster splits into a Noisette group (N3, N4, N6, and N10) and a Bourbon group (B1, B2, B3, and B4). That these two groups sorted out, while the others of each classification grouped with Tea Roses, shows that these accessions are more clearly defined and genetically separated from Tea type roses. The level of similarity of these accessions is between that of the cultivated Chinas and Teas and the more diverse species. A possible reason for this is that breeding in these groups quickly progressed to more modern hybrid groups like the Hybrid Teas, in contrast to the long history of cultivation and selection the China and Tea Roses had in their country of origin before introduction to Europe.

Multiflora/Polyantha Group

The last group that shows a clear cluster contains all but two of the Polyantha class roses and both *R. multiflora* accessions, though not *R. multiflora* var. *cathayensis*. This group as a whole had a level of diversity similar to the Noisettes and Bourbons. Similarity coefficients between the Polyantha and *R. multiflora* within this group ranged from 0.39-0.56, and similarity between the group's Polyantha accessions and, as an

example China, 'Old Blush' ranged from 0.38-0.41. This supports the recorded origin of varieties of *R. multiflora* and *R. chinensis* for the Polyantha class.

Principle Coordinate Analysis

The 3-D Principle Coordinate Analysis plot that graphs the most significant factors of the SSR data is in good agreement with the results shown in the dendrogram, supporting the validity of the clusters there, and revealing additional insight into the relationships of the species and cultivars in this study that a divergent dendrogram cannot show (Figure 3).

Each pinpoint and wire graphs the accessions in three dimensions, and in the optimized view of those dimensions plotted in Figure 3, the first thing that can be readily noted is that the same groupings are seen there as in the dendrogram in Figure 2. The species, for the most part, are tightly grouped to the right and mid-way up the graph in this view, and again are not strictly separated out by type. The China Roses form a broader group spread diagonally down toward the bottom left of the view, with the Red Chinas again clustered near each other at the far edge of the group.

The Bourbons and Noisettes are not as clearly sorted here, but the ones that grouped together on the dendrogram can be found near each other on the 3-D graph, apart from those that grouped with the Teas on the dendrogram. The location of 'Queen of Bourbons' (B5) in the dendrogram raised some questions, but the Principle Coordinate Analysis shows that it is indeed set apart from the China Roses by its third dimension, which is much larger, like the main group of Bourbons and Noisettes. This
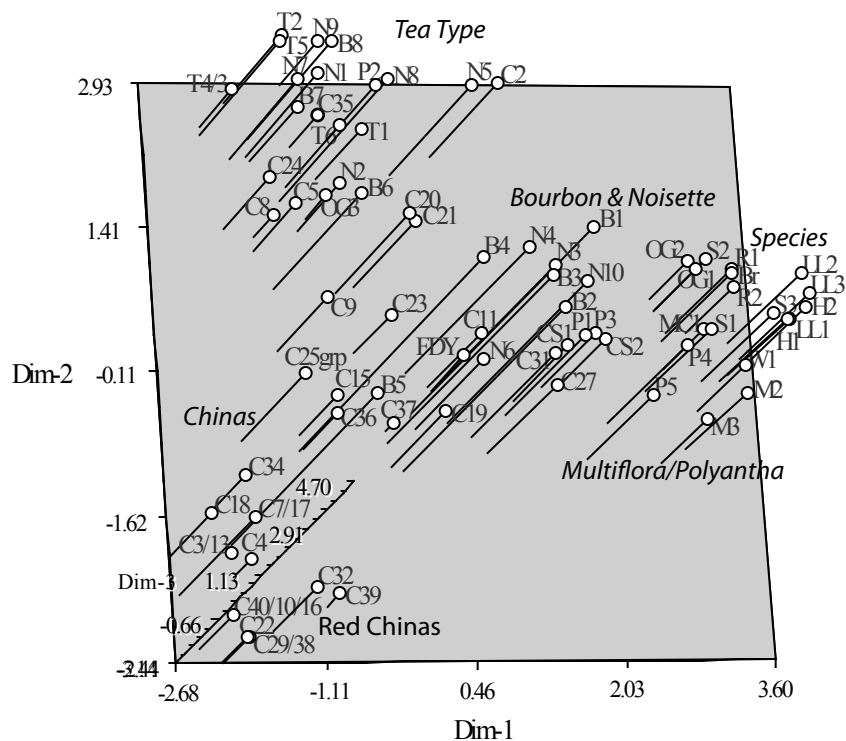
Figure 3. 3-D Principle Coordinates Analysis plot of SSR data. Accessions are labeled by sample number as listed in Table 1. Samples with identical SSR profiles are labeled with one point on the graph and separated by "/". Ex: "T4/3" is accessions T4 and T3. "C25grp" is the 'Old Blush' group and includes accessions C1, C6, C12, C14, C25, C26, C30, C33, and CSF.

supports the earlier discussion that 'Queen of Bourbons' is genetically allied with the other Bourbons, but with perhaps more China Rose influence in its background.
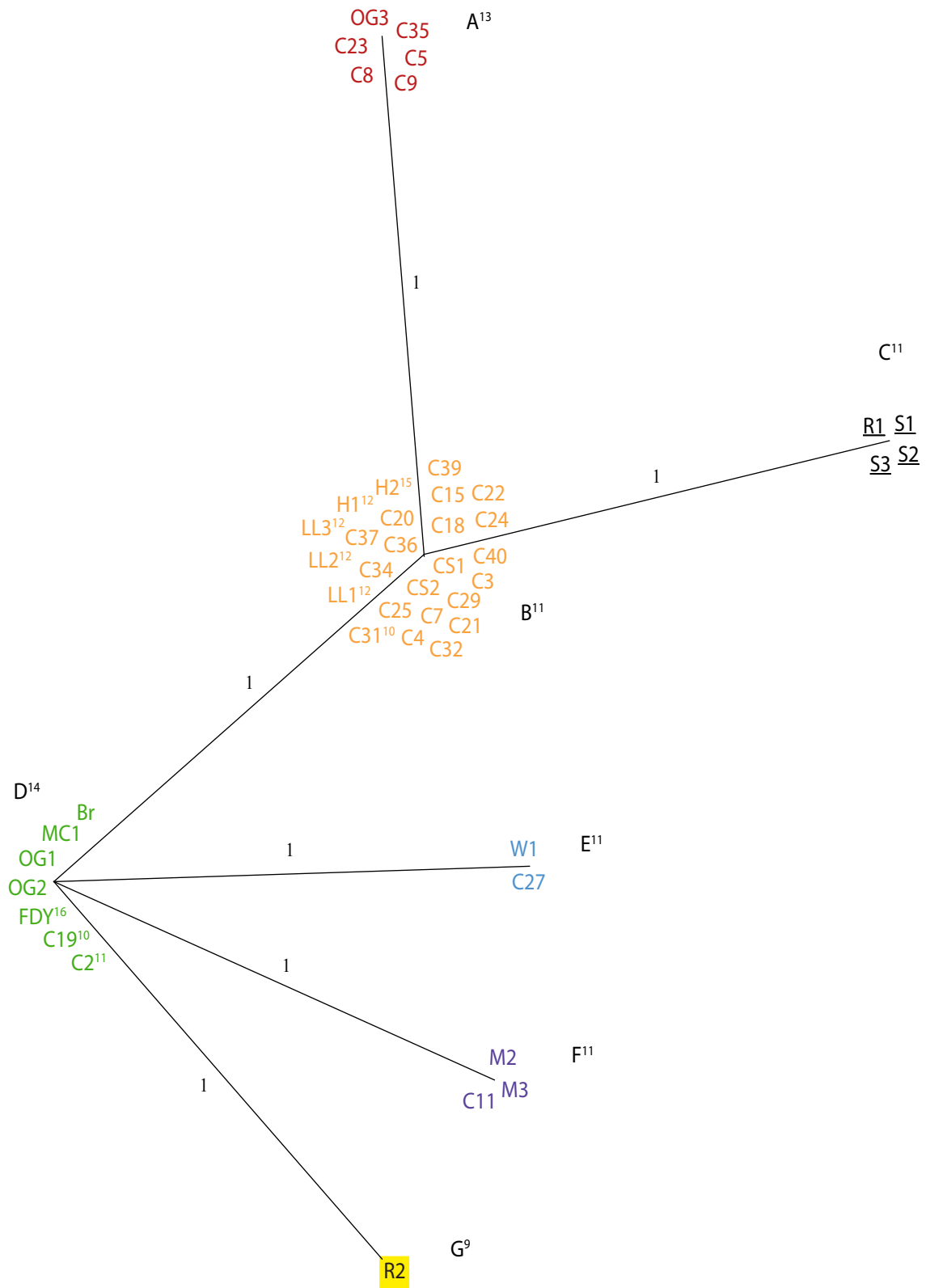
The Polyanthas are distributed between the *R. multiflora* accessions and *R. chinensis* var. *spontanea* and the cultivated Chinas. This is historically recorded to be their origin, and the three dimensions of this plot give a good representation of that. Looking at the individual similarities used to generate the dendrogram also revealed this, but the divergent cluster groups were not able to graphically display that information.

The relationships among the Tea accessions in the 3-D analysis were the same as in the dendrogram, except for one accession. 'Mutabilis' (C23) sorted into the Tea Type cluster on the dendrogram, but here is found to be outside of the tighter Tea cluster and nearer to the main China group. More insight into the placement of this cultivar and others is revealed by their haplotypes, and is further discussed in the following section with the chloroplast spacer data.

Chloroplast Intergenic Spacer Haplotypes

Figure 4 is an unrooted phylogram of the haplotypes of the China Roses, an Old Garden Rose, and the species roses included in this study. Each branch that separates the seven different main haplotypes indicates a single nucleotide change in the *trnH-psbA* chloroplast intergenic spacer sequence of those samples. The chloroplast SSR (cpSSR) also found in the sequence varied from 9-16 repeats of a single nucleotide (T), with 11 repeats being most prevalent, making it typical of SSRs found in plant chloroplasts (Weising and Gardner, 1999). The variability in the cpSSR is represented

Figure 4. *trnH-psbA* haplotypes. Accessions are labeled by sample number as listed in Table 1. Accessions with identical SSR profiles are represented with one sample number. C25 represents C1, C6, C12, C14, C25, C26, C30, C33, and CSF; C3 = C3 & C13; C40 = C10, C16, & C40; C29 = C29 & C38; C7 = C7 & C17. Haplotype groups are labeled A – G, and each branch of the phylogram indicates a single nucleotide change. Superscripts represent the number of repeats in the mononucleotide SSR. Main haplotype groups that all share the same number of repeats have that number indicated with the haplotype letter, while groups with more than one type have labels on the differing accessions.

OG3 C35   A[13]
C23   C5
C8   C9

C[11]

R1 S1
S3 S2

1

C39
H2[15] C15 C22
H1[12] C20 C18 C24
LL3[12] C37 C36
LL2[12] C34 CS1 C40
LL1[12] CS2 C3
C25 C29
C7 C21 B[11]
C31[10] C4 C32

1

D[14]
Br
MC1
OG1
OG2
FDY[16]
C19[10]
C2[11]

1 W1   E[11]
C27

1

M2   F[11]
C11 M3

1

R2   G[9]

in Figure 4 by superscripts indicating the number of repeats in each group or individual accession, and is useful in cases where it allows more species/hybrid groups to be distinguished within the larger haplotype based on single nucleotide changes.

As mentioned earlier, less variation in this sequence was found in this group of roses than was found in a survey spanning the whole genus *Rosa* (Bruneau et al., 2007). In that study, 22 of 503 characters were found to be potentially informative (for phylogenetics), and in this study only six sites of single nucleotide mutations were found across the samples included. In addition to the larger number of indels, more variation was also found in the chloroplast SSR of the other study, with 9-18 T repeats being found. Again, this is most likely due to the focus of this study where all of the roses are expected to be relatively closely related as compared to their relationship with roses from more distant sections in the genus *Rosa*.

As was the case with the nuclear SSR data, one of the first aspects noted with the chloroplast sequence data is which accessions have identical sequences. All samples that had identical SSR profiles also had identical *trnH-psbA* sequences as expected. However, unlike the SSR data, identical chloroplast sequences do not indicate that accessions are genetically identical, or nearly so in the case of vegetative sports. Since the chloroplast genome is thought to be inherited maternally in roses (Corriveau and Coleman, 1988), accessions with identical sequences indicate maternal relationship.

From the haplotypes obtained, it can be seen that some species have individuals with different haplotypes, while in other cases several species share the same haplotype. Other studies have also observed that the level of polymorphism between and within

species in a genus varies when looking at chloroplast data. One such study using chloroplast PCR-RFLP found that *R. chinensis* 'Alba' grouped with varieties of *R. multiflora*, and another *R. chinensis* grouped by itself, while the third China Rose accession in that study, *R. chinensis* 'Mutabilis', had yet another haplotype (Takeuchi et al., 2000). In this study, the results also place the cultivar Mutabilis in haplotype group A[13], which is different from the majority of the cultivated China Roses and *R. chinensis* var. *spontanea* in group B[11] (see Figure 4). Three more main haplotypes in this study also include one or more China Roses. The China Rose 'Papa Hemeray' (C11) groups with two *R. multiflora* accessions in group F[11],'Yue Yue Fen China' and 'Arethusa' (C19 and C2 respectively) have haplotype D[(10 & 11)], and 'WOB26' (C27) has haplotype E[11]. The large number of China Roses sampled, including some which are known or suspected hybrids with other groups, contributes to the high number of different haplotypes for members of this group and does not necessarily point to higher intraspecific variability. In contrast to the multiple haplotypes of the China Roses, all samples of *R. soulieana* (S1, S2, S3) and one of the two *R. rubus* accessions (R1) all share haplotype C[11], and *R. brunonii* (Br), *R. multiflora* var. *cathayensis* (MC1), and two *R. odorata* var. *gigantea* samples (OG1 & OG2) share D[14].

Comparing the sequences from this study to samples of shared species in the other rose study using *trnH-psbA* (Bruneau et al., 2007) does reveal more variability within species. Four of the five accessions had sequences varying by two or more single nucleotides from samples identified as the same species in this study. *R. henryi* had the same sequence as those in this study, but had yet another microsatellite repeat number

that fell between that of H1 (12bp) and H2 (15bp) with 13bp. Others had sequences that diverged significantly more. A sample of *R. chinensis* var. *spontanea* had two polymorphisms not found in CS1 and CS2, one of which was a mutation not found in any other rose from this study. The *R. multiflora* sample showed similar differences with two polymorphic sites not matching those found in M2 or M3, or the rest of the accessions. The *R. odorata* var. *gigantea* in Bruneau et al. (2007) also differed by having one unique single nucleotide polymorphism, a 27bp indel, and a different repeat number for the SSR. The SSR in that sample had 16bp, as opposed to OG1 and OG2's 14bp and OG3's 13bp. Two of the polymorphisms did match ones found in OG1 and OG2. The *R. odorata* var. *gigantea* sample was documented as coming from the Montreal Botanical Garden, and originally from Yunnan China, which is the same region the OG3 DNA sample was sent from. The sequence of *R. wichuriana* var. *wichuriana* showed the most variation when compared with W1 from this study. Two indels had to be identified before the samples could be aligned, the cp-SSR contained 9bp versus 11bp, and seven single nucleotide polymorphisms were found. All but one of the single nucleotide polymorphisms was unique to the study, and five were within one 14bp stretch of the sequence. The sequence did share one polymorphism with C27 and its maternal parent W1 that set them apart from any other samples in this study.

Because of the variability found when comparing studies, it appears that more intraspecific sequence divergence may be present in these species than suggested by this research or Bruneau et al. (2007) alone, where only 1-3 accessions from each species was sequenced for each study. The variability could be natural, or could mean that

misidentification occurred during collection or transfer between institutions, so further

research is needed to investigate intraspecific variability in this chloroplast region in

Rosa.

Combined Chloroplast and Nuclear Data Sets

While general information on the maternal lineage of these roses can be obtained

from the chloroplast data alone, combining those results with the nuclear SSR data

reflecting the contributions of both parents, along with the breeding history, if known,

offers more insight. Combining the different types of data available was especially

important for the focus of this study, where the accessions were not all wild-collected

species, but included the China Rose group where the origin is largely unknown and the

line between species type and hybrids is blurred from a long history of cultivation.

Figure 5 shows the same static view of the 3-D graph of the SSR data as Figure 3, but

with the sample label pins color-coded by haplotype for accessions included in the

chloroplast sequencing portion of the study. Samples with cpSSRs that differ from the

majority of the group are also indicated in Figure 5.

The rose seedling 'WOB26' and its parents *R. wichuriana* 'Basye's Thornless'

and *R. chinensis* 'Old Blush' from the Texas A&M Basye Rose program's mapping

population illustrates the clearer view of relationships that is apparent when all the data

from this study is graphed together. Breeding records for the seedling designated

'WOB26' indicate that *R. wichuriana* 'Basye's Thornless' is the maternal parent, and

'Old Blush' the pollen parent (Byrne, personal comm.). Figure 5 clearly shows

'WOB26' between the two parents where it would be expected on the graph, and that it
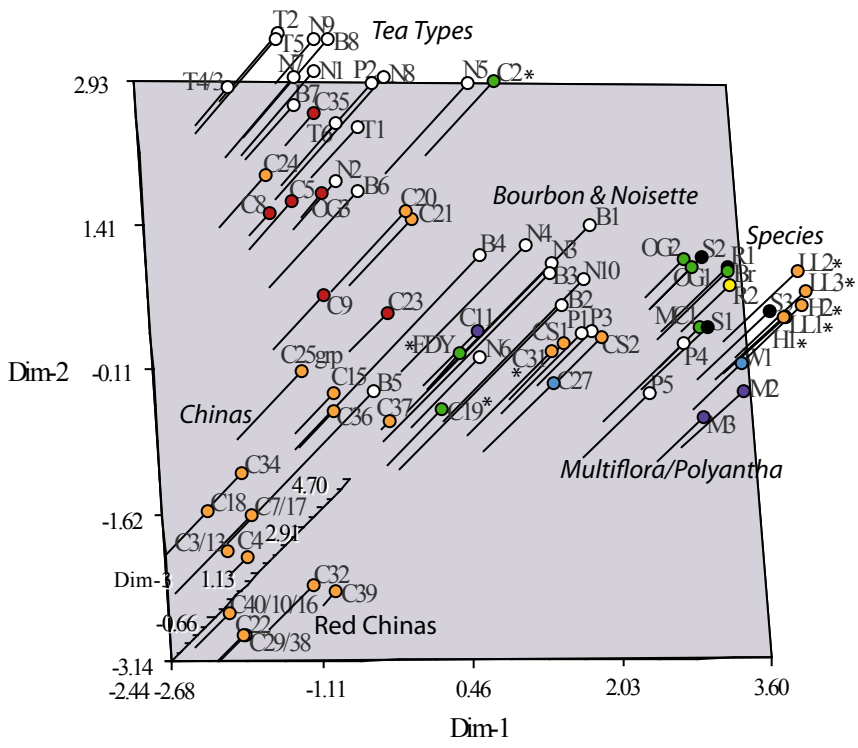
Figure 5. Combined 3-D PCoorA of SSR data and haplotypes. See Figure 3 for notes on sample name labels. Colored pins correspond to the haplytype groups in Figure 3. Red = haplotype A, Orange = B, Black = C, Green = D, Blue = E, Purple = F, and Yellow = G, and "*" signifies a sample with a cpSSR repeat number different from the majority of the group (see Figure 4 for details).

shares its haplotype with the recorded maternal parent. While not all the accessions in this study have such complete breeding history and do not likely have their parents included in the study, general relationships based on both nuclear and chloroplast sources can still be seen for species and hybrid groups.

In the China Rose group, known or suspected Hybrid Chinas become more apparent with both the nuclear and chloroplast data plotted together. The nuclear SSR data by itself revealed those China cultivars that had SSR profiles more closely matching the Teas. The maternally inherited chloroplast sequence data added to the SSR data may agree with the SSR data and support it, or help to point out China cultivars grouping with the other Chinas based on SSR that likely had a Tea Rose or other species or hybrid group on the maternal side of their ancestry. 'Jean Bach Sisley', 'Ducher', and 'Miss Lowe's Variety' (C8, C5, and C35) are three China/Hybrid China Roses that grouped in the Tea Type group based on SSR profiles, and they also share a haplotype with the *R. odorata* var. *gigantea* accession from China (OG3) rather than the haplotype of the majority of the China Roses. Of these, 'Jean Bach Sisley' was already listed as a Hybrid China in Modern Roses XI (Cairns (ed.), 2000), and 'Ducher' is listed there as a China, but can be found described in the trade as an older Tea Rose type. In these two cases the data from this study support their hybrid status. The surprising cultivar is 'Miss Lowe's Variety', which is of unknown parentage, but has been speculated to be a dwarf sport of 'Slater's Crimson China' (Cairns (ed.), 2000). The SSR data already proved this incorrect for the 'Slater's Crimson China' accession in this study, and also that it was unlikely that 'Miss Lowe's Variety' was a sport of any China Rose based on its profile.

The combination of the SSR PCoorA and the chloroplast haplotype data indicate that 'Miss Lowe's Variety' is actually a dwarf type rose with strong Tea Rose type background. The Bermuda found rose 'Spice' (C24) has the haplotype typical of the China Roses, but is within the cluster of Tea Type Roses based on the SSR data, so it appears to have a strong Tea Rose background as well. The Tea-based SSR profile and China Rose chloroplast type of this accession match with it's suspected identity of 'Hume's Blush Tea-Scented China', and if it is not that particular rose, it does appear to be an old Tea Rose type cultivar.

Several other roses included in this study that appear to be hybrids are not so clearly sorted into the Tea type group or the China Rose group. 'Arethusa' (C2) seems to have some Tea influence based on similarity and its coordinates in the Figure 5, but it is classified as a Hybrid China with unknown parentage bred in 1903 (Cairns (ed.), 2000). Because this accession also has a chloroplast sequence that is unique in this study, it does not give a clear indication of what species or group may have had a maternal contribution to this hybrid. 'Yue Yue Fen China' (C19) is another accession with a unique haplotype. The plant of this accession as well as those of 'Yue Yue Fen 8 & 9' (C20 and C21) were collected as seed from wild and garden Yue Yue Fen type roses (those like 'Old Blush') in China. They all appear to have China Rose influence based on the SSR data, but C19's haplotype indicates that it is a seedling of a rose with a different maternal background from the main group. On the other hand, C20 and C21 have the haplotype of the main China Rose group, but are located just outside the main group and closer to the Tea Type group based on similarity levels. This could indicate

that pollination was not controlled for the hips those seeds came from, and that a Tea

Rose was the likely pollen source then or in previous generations. Since the seeds were

collected and pooled together from multiple sources, it cannot be said whether either the

'Yue Yue Fen' or 'Yue Yue Fen 8 and 9' were from cultivated sources that could have

crossed with other garden roses, or whether they may be the result of natural crosses.

Because part of China is the native habitat of both *R. chinensis* var. *spontanea* and *R.

odorata* var. *gigantea*, the proposed parents of both the cultivated Chinas and Teas, as

well as the country of early cultivation of those roses, finding a range of intermediate

type roses along with those clearly belonging to one class or the other is quite likely.

Two more China class roses are also located just outside the main group,

between the China Roses and the Tea Types. 'Mme. Laurette Messimy' and 'Mutabilis'

(C9 and C23) also have the same haplotype as OG3. The parentage of 'Mme. Laurette

Messimy' includes a Tea Rose parent on one side, and a Tea Rose grandparent on the

other (Cairns (ed.), 2000), which is in keeping with its location based on the SSR data

and the haplotype. 'Mutabilis' has long been considered a China Rose by most

references, and is the best example of one of the China Rose traits – darkening from bud

through aging bloom (Dickerson, 1992). However, at least one source can be found that

puts forth the opinion that it may be derived in part from the Tea Roses (Thomas, 1980).

Both the SSR and chloroplast data support this hypothesis that 'Mutabilis' has strong

Tea influence in addition to China Rose heritage.

The background of the Hybrid China 'WOB26' (C27) is known, and has already

been discussed, but the last two roses outside the main group classified as China Roses

have backgrounds that are not as clear. 'Papa Hemeray' (C11) is located between the main group of China Roses and the Polyanthas and other hybrids, and shares a haplotype with the two *R. multiflora* accessions in this study. Based on its SSR grouping and haplotype, and its recorded introduction in 1912 after Polyantha Roses had begun to be introduced (Cairns (ed.), 2000), it is quite possible that this Hybrid China Rose has some *R. multiflora* in its background. The haplotype of 'Pink Pet' (C31) gives fewer clues to its maternal relationship. The basic haplotype of this rose is that of the main group of China Roses, however, this was the accession that had a sequence with several indels not included in any of the other roses sampled. In addition, the SSR repeat number sets it apart from the main group. This cultivar has also been classified as a Polyantha in the past (Dickerson, 1992), and it is grouped near to some of the Polyantha cultivars in this study, so it is possible that it is of Polyantha breeding.

Fortune's Double Yellow (FDY) was also included in the chloroplast spacer sequence portion of the study, since it is usually considered a Miscellaneous Old Garden Rose of uncertain ancestry (Cairns (ed.), 2000) that, like the China and Tea Roses, was bred in China before being introduced to Europe (Gault and Synge, 1971). It has been speculated that this cultivar is a Tea Rose or has *R. odorata* var. *gigantea* influence (Dickerson, 2007). The location of FDY is not within the main group of Chinas, Teas, or other hybrids, but it has similarity coefficients ranging from 0.30-0.48 with the Teas and OG3, so it does appear to have a close relationship with the Tea Type Roses. The chloroplast haplotype of FDY is another that is unique in this study based on its cpSSR number of 16, so no direct maternal connections can be presumed, though the main

haplotype matches that of the *R. brunonii*, R. *multiflora* var. *cathayensis*, and two *R. odorata* var. *gigantea* (OG1 & OG2) from this study. The SSR-based similarity levels of FDY with Br or MC1 are low though (0.17 and 0.22) and do not indicate a close relationship. The similarity between FDY and OG1 or OG2 are higher at 0.32 and 0.31, but this could simply reflect the fact that they may all be hybrids with another species in Section Synstylae if OG1 and OG2 are not true *R. odorata* var. *gigantea* specimens.

Origins of the China Roses

After examining the relationships of the individual China Rose cultivars in this study, the ultimate question is the origin of the group as a whole. The previous discussion of which China Rose accessions are likely Hybrid Chinas, and what the hybrid origin is likely to be is of great importance. At this point in history, the China Rose cultivars available are not necessarily the original China Roses in the pure species or hybrid state they existed in China; however, hybrid accessions with data that suggest breeding with other species or hybrid roses were identified. Once those hybrid accessions were set aside, examining the possible origins of the more defined main group of China Roses was possible. The other important consideration for drawing conclusions on the origins of the China Rose group based on this study was the true identity of the different *R. odorata* var. *gigantea* accessions.

As mentioned previously, the distant SSR relationship of the two different types of *R. odorata* var. *gigantea* accessions in this study make their identity questionable. The chloroplast data adds more details, but does not answer the question of which represents the wild species. In addition to being less closely related than expected based

on the SSR data, the chloroplast sequence reveals that OG3 has a different haplotype than OG1 and OG2, and that *R. odorata* var. *gigantea* from Bruneau et al. (2007) has yet another haplotype. In fact, the US source (OG1 and OG2) shares a haplotype with MC1 and Br, in addition to grouping closer to the Synstylae section species than to OG3 and the cultivated Tea Type Roses in the SSR Principle Coordinate Analysis. Based on this, the US source of *R. odorata* var. *gigantea* could be a hybrid with a rose from Section Synstylae on the maternal side. If that is the case, OG3 from *R. odorata* var. *gigantea*'s country of origin could well be the true species, but upon receiving more information from the source of this sample's DNA, it appears to have some characteristics atypical of the species. The OG3 sample is said to have the same form as other examples of *R. odorata* var. *gigantea*, but blooms about a month earlier and sometimes blooms a second time in one year (Xianqin Qiu, personal communication). Along with the bloom character disparity, OG3 also has the highest percentage of heterozygous loci of all the supposed wild species sampled in this study (78%), pointing to a probable hybrid origin of this accession. Given that OG3 otherwise looks like the common *R. odorata* var. *gigantea* though, it is likely a very primitive example of the first Tea Roses. If the hybrid seed resulted from *R. odorata* var. *gigantea* as the maternal parent, then the *trnH-psbA* sequence of OG3 should match the wild species, but this cannot be established without further testing of more typical accessions thought to be the true wild variety.

Because of this confusion, it is not possible to say whether or not this variety of the species *R. odorata* contributed to the background of the China Roses, but it is possible to comment on the relationships of these particular samples identified as *R.*

odorata var. *gigantea* to the China Roses, and to the Tea Roses. The PCoorA graph with

chloroplast haplotypes (Figure 5) shows that OG3 is more closely related to the China

and Tea Roses than OG1 or OG2, and the similarity coefficient values for selected China

and Tea Roses shows the magnitude of their similarity (Table 5). The China and Tea

Roses chosen for the condensed matrix were selected to represent the larger data set, and

OG1 and OG2 have similarity coefficients with these roses ranging from 0.13-0.24 with

an average of 0.17 (0.15 with China Roses only), while OG3's similarity coefficients

range from 0.45-0.63 with an average of 0.51 (0.50 China Roses only). This suggests

that it is indeed likely that the China Roses and the Tea Roses share a common

background with OG3 type roses.

The other half of the prospective parents thought to have contributed

significantly to the China and Tea Rose classes is *R. chinensis* var. *spontanea*. *R.*

chinensis var. *spontanea* as the parent of the China Roses is a common hypothesis in

published literature based on morphology, and the wild species was named var.

spontanea for this reason (Rix, 2005). Based on proximity to the China and Tea Roses

on the dendrogram and the three-dimensional PCoorA graph, the *R. chinensis* var.

spontanea samples in this study do not appear to be as closely related to the hybrid

groups as sample OG3. The similarity coefficients between *R. chinensis* var. *spontanea*

and the China and Tea Roses ranged from 0.27-0.51 with an average of 0.36, and 0.40

with just the China Roses. While these levels of similarity are less on average than those

of OG3 with the China Roses, it is still significant for this group when combined with

the chloroplast data that supports a maternal relationship of *R. chinensis* var. *spontanea*

Table 5. Similarity coefficients matrix of select China, Tea, and species roses.

Sample ID → ↓	CS1	CS2	C25grp	C29/38	C3/13	OG1	OG2	OG3	T1	T4/3	T5
CS1	1.00[x]	0.66	0.51	0.4	0.32	0.14	0.16	0.29	0.31	0.35	0.34
CS2	0.66	1.00	0.43	0.36	0.35	0.2	0.22	0.23	0.27	0.34	0.33
C25grp	0.51	0.43	1.00	0.68	0.56	0.14	0.16	0.49	0.53	0.66	0.62
C29/38	0.4	0.36	0.68	1.00	0.81	0.13	0.15	0.47	0.41	0.49	0.45
C3/13	0.32	0.35	0.56	0.81	1.00	0.13	0.18	0.53	0.42	0.59	0.44
OG1	0.14	0.2	0.14	0.13	0.13	1.00	0.83	0.19	0.18	0.19	0.22
OG2	0.16	0.22	0.16	0.15	0.18	0.83	1.00	0.21	0.17	0.18	0.24
OG3	0.29	0.23	0.49	0.47	0.53	0.19	0.21	1.00	0.45	0.63	0.46
T1	0.31	0.27	0.53	0.41	0.42	0.18	0.17	0.45	1.00	0.78	0.64
T4/3	0.35	0.34	0.66	0.49	0.59	0.19	0.18	0.63	0.78	1.00	0.76
T5	0.34	0.33	0.62	0.45	0.44	0.22	0.24	0.46	0.64	0.76	1.00

[x]Similarity coefficients calculated using Dice and rounded to two decimals.

with the China Roses. As discussed in the Chloroplast Intergenic Spacer Haplotype section, the main group of China Roses share the same haplotype as the *R. chinensis* var. *spontanea* accessions in this study. With the occurrence of intraspecific variation in some of the other species, the homogeneity of the China Roses and *R. chinensis* var. *spontanea* is a strong indication that *R. chinensis* var. *spontanea* is a maternal ancestor of the China Roses. Before modern breeding and hand pollination, the practice of improving roses by collecting seed from a desirable female parent may have contributed to the China Roses shared haplotype, despite the fact that they appear to have originated not only from selection, but also as a result of at least some hybridization.

The China Roses and Tea Roses also have high levels of genetic similarity with each other (average 0.51 for the samples in Table 5), and accessions in this study can be found in a gradient from one type to the next based on SSR's, which could be the result of interbreeding between China Roses and Tea Roses, and also supports the suspected common hybrid origin of the two types of roses. The history and the extent of the interbreeding between the two groups, and the natural and/or artificial hybrids that may have occurred between species to create the groups are not known. Ultimately, the hypotheses of Guoliang are still the best summary of what likely happened in the ancient breeding of these roses: The large-scale cultivation of wild roses in China would have made repeatedly finding and propagating unique, desirable specimens and sports a common occurrence in the selection of China Roses. Also, some degree of artificial pollination by Chinese gardeners could have taken place at some point in history, and

selection for particular traits may have occurred after natural hybridization (Guoliang, 2003).

Future Use of trnH-psbA in Genus Rosa

In the genus-wide phylogenetic analysis study by Bruneau et al. (2007), the chloroplast intergenic spacer *trnH-psbA* proved to be slightly more variable than previously surveyed regions, but its alignment was complicated by repetitive regions with multiple indels, leading to the exclusion of almost half of the sequence. While some useful information was still able to be obtained from the sequence, this makes *trnH-psbA* less effective for phylogenetic or diversity studies that include roses from a wide genetic background. Because the *trnH-psbA* region was easily amplified and variable sequence could be obtained from the roses sampled, it may warrant future use for cultivar identification or confirming close maternal relationship. When identifying unique samples is the goal, rather than aligning multiple sequences for comparison, the prevalent indels and highly variable areas do not pose a problem, but rather add to the distinguishing power of the region as found by Kress and Erickson (2007).

SUMMARY

Twenty-three microsatellite markers were used to investigate the genetic diversity and relationships among 90 rose accessions, including China Roses, their early hybrid groups, and possible contributing species. Further analysis was done to sequence the chloroplast intergenic spacer *trnH-psbA* in a subset of the genotypes that included the China Roses and the Section Indicae and Synstylae species.

The SSR portion of the study found 291 alleles total, with a range of 6-22 alleles per locus, and an average of 12.65, which was in line with other recent studies in *Rosa* and the Rosaceae Family (Hokanson et al., 2001; Scariot et al., 2006; Tang et al., 2008). Based on the SSR data, the ploidy levels of the accessions were calculated, revealing that approximately two-thirds of the roses classified as China and Hybrid China were diploid, and the remaining one-third were triploid. Some of the samples had previously published ploidy levels (Cairns (ed.), 2000; Roberts (eds.) et al., 2003), and the calculations here were in agreement with those. The percentage of heterozygous loci was also calculated for the diploid individuals in the study, and the values were found to range from 30-87%, with an overall average of 67%. While the small sample size and different methods of calculation used in previous *Rosa* studies makes comparison of heterozygosity estimates difficult, the most heterozygous roses in this study had the same level of heterozygosity as calculated based on SSRs in *Malus*, which is known to be highly heterozygous (Kenis and Keulemans, 2005).

To further examine the genetic diversity and relationships of the roses in this study, NTSYS-pc v. 2.2 (Numerical Taxonomy and Multivariate Analysis System, Exeter Software) was used to create a Dice similarity based dendrogram, as well as a three-dimensional graph of the Principle Coordinates of the SSR data. The dendrogram revealed clusters of cultivars that generally agreed with the American Rose Society horticultural classifications listed in Modern Roses XI (Cairns (ed.), 2000), and indicated that the cultivated China Roses have high genetic similarity within the group. A number of vegetative sports and synonyms within the China Rose group were also identified, including the largest group made up of 'Old Blush', 'Climbing Old Blush', 'Green Rose', 'Single Pink', 'Rouletii', 'Pompon de Paris', 'Bengale d'Automne', 'Archduke Charles', and an *R. chinensis* var. *semperflorens*. The three dimensions of the Principle Coordinate Analysis graph gave a clearer representation of the relationships of the species and cultivars, and showed the same groups as in the dendrogram clusters with few exceptions.

The Principle Coordinate Analysis, combined with the haplotypes generated from the chloroplast intergenic spacer sequences gave the most insight into the diversity and relationships of the roses in this study. With this information, the China Roses could be more clearly defined as a group after identifying the likely Hybrid China accessions, which aided in the process of looking at possible progenitor species for the core group. The close SSR-based genetic relationships of the main China Roses, combined with their shared chloroplast haplotype with the *R. chinensis* var. *spontanea* samples in this study suggest that the maternal ancestor of the China Roses is *R. chinensis* var. *spontanea*.

The SSR results also indicate a close relationship with one of the *R. odorata* var. *gigantea* accessions in this study (OG3), but a much more distant relationship was found with the other two samples of that species (OG1 and OG2). Due to the low similarity and different chloroplast haplotypes observed between the two sources of samples for this species, their identities are in question and it cannot be said for certain which may represent *R. odorata* var. *gigantea*. What can be said is that accession OG3 in this study had high coefficients of similarity with the cultivated China and Tea Roses, and roses of this type likely contributed to the common background between the China and Tea Rose groups.

LITERATURE CITED

Babaei, A., S.R. Tabaei-Aghdaei, M. Khosh-Khui, R. Omidbaigi, M.R. Naghavi, G.D. Esselink, and M.J.M. Smulders. 2007. Microsatellite analysis of Damask Rose (*Rosa damascena* Mill.) accessions from various regions in Iran reveals multiple genotypes. BMC Plant Biol. 7:12

Boonprakob, U. 1996. RAPD polymorphisms in diploid plums: Genetic relationships and genetic linkage maps. Doctoral Thesis, Texas A&M University, College Station.

Bruneau, A., J.R. Starr, and S. Joly. 2007. Phylogenetic relationships in the genus *Rosa*: New evidence from chloroplast DNA sequences and an appraisal of current knowledge. Syst. Bot. 32(2):366-378.

Byrne, D.H., N. Anderson, and H.B. Pemberton. 2007. The use of *Rosa wichurana* in the development of landscape roses adapted to hot humid climates. Acta Hort. 751:267-274.

Cairns, T. (ed.). 2000. Modern roses XI: The world encyclopedia of roses. Elsevier Ltd, Amsterdam.

Corriveau, J.L. and A.W. Coleman. 1988. Rapid screening method to detect potential biparental inheritance of plastid DNA and results for over 200 angiosperms species. Amer. J. Bot. 75:1443–1458.

Crespel, L., D. Zhang, J. Meynet, and S. Gudin. 2001. Estimation of heterozygosity in two botanic rose species using AFLP markers. Acta Hort. 546:187-191.

Debener, T. 2002. Molecular markers as a tool for analyses of genetic relatedness and selection in ornamentals. p.329-345. In: A. Vainstein (ed.) Breeding for ornamentals: Classical and molecular approaches. Kluwer Academic Publishers, Dordrecht, The Netherlands.

Debener, T., M. Linde, A. Dohm, and H. Kaufmann. 2004. The utilization of molecular tools for rose breeding and genetics. Acta Hort. 630:29-42.

Debener, T. and L. Mattiesch. 1999. Construction of a genetic linkage map for roses using RAPD and AFLP markers. Theor. Appl. Genet. 99:891-899.

Dickerson, B.C. 1992. The old rose advisor. Timber Press, Portland, OR.

Dickerson, B.C. 2007. Old roses: The master list: Second edition updated and expanded. iUniverse, Lincoln, NE.

Doyle, J.J., J.L. Doyle. 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochem. Bull. 19(1):11-15.

Dugo, M.L., Z. Satovic, T. Millan, J.I. Cubero, D. Rubiales, A. Cabrera, and A.M. Torres. 2005. Genetic mapping of QTLs controlling horticultural traits in diploid roses. Theor. Appl. Genet. 111:511-520.

Esselink, G.D., M.J.M. Smulders, and B. Vosman. 2003. Identification of cut rose *(Rosa hybrida*) and rootstock varieties using robust sequence tagged microsatellite site markers. Theor. Appl. Genet. 160.277-286.

Gault, S.M., P.M. Synge. 1971. The dictionary of roses in colour. Royal Horticulture Society/Royal National Rose Society/Ebury Press & Michael Joseph, London.

Gudin, S. 2000. Rose: Genetics and breeding. Plant Breeding Rev. 17:159-189.

Guoliang, W. 2003. Ancient Chinese roses, p. 387-395. In: A.V. Roberts, T. Debener, S. Gudin. (eds.). Encyclopedia of rose science. Volume One. Elsevier Ltd, Amsterdam.

Harkness, J. 1978. Roses. William Clowes & Sons Lim., London.

Hibrand-Saint Oyant, L., L. Crespel, S. Rajapakse, L. Zhang, and F. Foucher. 2008. Genetic linkage maps of rose constructed with new microsatellite markers and locating QTL controlling flowering traits. Tree Genet. & Genomes. 4:11-23.

Hokanson, S.C., W.F. Lamboy, A.K. Szewc-McFadden, & J.R. McFerson. 2001. Microsatellite (SSR) variation in a collection of *Malus* (apple) species and hybrids. Euphytica. 118:281-294.

Hubbard, M., J. Kelly, S. Rajapakse, A. Abbott, and R. Ballard. 1992. Restriction Fragment Length Polymorphisms in rose and their use for cultivar identification. HortScience. 27(2):172-173.

Jan, C.H., 1996. Rose germplasm identification and analysis with RAPD markers. Master's Thesis, Texas A&M University, College Station.

Jan, C.H., D.H. Byrne, J. Manhart, H. Wilson. 1999. Rose germplasm analysis with RAPD markers. HortScience. 34(2):341-345.

Kenis, K., J. Keulemans. 2005. Genetic linkage maps of two apple cultivars (*Malus* x *domestica* Borkh.) based on AFLP and microsatellite markers. Mol. Breed. 15:205-219.

Kiani, M., Z. Zamani, A. Khalighi, R. Fatahi, and D.H. Byrne. 2008. Wide genetic diversity of *Rosa damascena* Mill. germplasm in Iran as revealed by RAPD analysis. Scientia Hort. 115:386-392.

Kimura, T., C. Nishitani, H. Iketani, Y. Ban, T. Yamamoto. 2006. Development of microsatellite markers in rose. Mol. Ecol. Notes. 6(3):810-812.

Kloda, J.M., P.D.G. Dean, C. Maddren, D.W. MacDonald, and S. Mayes. 2008. Using principle component analysis to compare genetic diversity across polyploidy levels within plant complexes: An example from British restharrows (*Ononis spinosa* and *Ononis repens*). Heredity. 100:253-260.

Koopman, W.J.M., V. Wissemann, K. De Cock, J. Van Huylenbroeck, J. De Riek, G.J.H. Sabatino, D. Visser, B. Vossman, C.M. Ritz, B. Maes, G. Werlemark, H. Nybom, T. Debener, M. Linde, and M.J.M. Smulders. 2008. AFLP markers as a tool to reconstruct complex relationships: A case study in *Rosa* (Rosaceae). Am. J. Bot. 95:353-366.

Kress, J.W., D.L. Erickson. 2007. A two-locus global DNA barcode for land plants: The coding *rbcL* gene complements the non-coding *trnH-psbA* spacer region. PLoS ONE. 2(6):e508.

Kress, J.W., K.J. Wurdak, E.A. Zimmer, L.A. Weight, and D.H. Janzen. 2005. Use of DNA barcodes to identify flowering plants. PNAS. 102(23):8369-8374.

Krussmann, G. 1981. The complete book of roses. Timber Press, Portland, OR.

Martin, M., F. Piola, D. Chessel, M. Jay, and P. Heizmann. 2001. The domestication process of the modern rose: Genetic structure and allelic composition of the rose complex. Theor. Appl. Genet. 102:398-404.

Matsumoto, S., M. Kouchi, J. Yabuki, M. Kusunoki, Y. Ueda, and H. Fukui. 1998. Phylogenetic analysis of the genus *Rosa* using the matK sequence: Molecular evidence for the narrow genetic background of modern roses. Sciencia Hort. 77(1-2):73-82.

Millan, T., F. Osuna, S. Cobos, A.M.Torres, J.I.Cubero. 1996. Using RAPDs to study phylogenetic relationships in *Rosa*. Theor. Appl. Genet. 92:273-277.

Phillips, R., M. Rix. 1988. Roses. Random House, New York.

Piola, F., O. Raymond, J.L. Fiasson, C. Sanlaville-Boisson, P. Heizmann, and M. Jay. 2002. Modern tools in rose taxonomy: Contribution of chemical, biochemical, and molecular approaches to the study of the domestication of the modern rose. Recent Res. Devel. Plant Biol. 2:77-88.

Rix, M. 2005. *Rosa chinensis* f. *spontanea*. Curtis's Bot. Mag. 22(4):214-219.

Roberts, A.V., T. Debener, S. Gudin. (eds.). 2003. Encyclopedia of rose science. Volume One. Elsevier Ltd, Amsterdam.

Rousseau-Gueutin, M., E. Lerceleau-Kohler, L. Barrot, D.J. Sargent, A. Moufort, D. Simpson, P. Arus, G. Guerin, and B. Denoyes-Rothan. 2008. Comparative genetic mapping between octoploid and diploid Fragaria species reveals a high level of colinearity between their genomes and the essentially disomic behavior of the cultivated octoploid strawberry. Genetics. 179:2045-2060.

Scariot, V., A. Akkak, and R. Botta. 2006. Characterization and genetic relationships of wild species and old garden roses based on microsatellite analysis. J. Amer. Soc. Hort. Sci. 131(1): 66-73.

Shepherd, R.E. 1978. History of the rose. Earl M. Coleman, New York.

Takeuchi, S., K. Nomura, H. Uchiyama, K. Yoneda. 2000. Phylogenetic relationships in the genus *Rosa* based on the restriction enzyme analysis of the chloroplast DNA. J. Japan. Soc. Hort. Sci. 69(5):598-604.

Tang, K., X. Qiu, H. Zhang, S. Li, Q. Wang, H. Jian, B. Yan, and X. Huang. 2008. Study on Genetic Diversity of Some *Rosa* Germplasm in Yunnan Based on SSR markers. Acta Hort. Sinica. 35(8):1227-1232.

Thomas, G.S. 1980. Shrub roses of today. Dent & Sons, Ltd., London.

Weising, K., R.C. Gardner. 1999. A set of conserved PCR primers for the analysis of simple sequence repeat polymorphisms in chloroplast genomes of dicotyledonous angiosperms. Genome. 42:9-19.

Wissemann, V. and C.M. Ritz. 2005. The genus *Rosa* (Rosoideae, Rosaceae) revisited: Molecular analysis of nrITS-1 and *atpB-rbcL* intergenic spacer (IGS) versus conventional taxonomy. Bot. J. Linnean Soc. 147:275-290.

Wu, S., Y. Ueda, S. Nishihara, and S. Matsumoto. 2001. Phylogenetic analysis of Japanese *Rosa* species using DNA sequences of nuclear ribosomal internal transcribed spacers (ITS). J. Hort. Sci. & Biotech. 76(2):127-132.

Yan, Z., C. Denneboom, A. Hattendorf, O. Dolstra, T. Debener, P. Stam, P.B. Visser. 2005. Construction of an integrated map of rose with AFLP, SSR, PK, RGA, RFLP, SCAR and morphological markers. Theor. Appl. Genet. 110:766-777.

Zhang, L.H., D.H. Byrne, R.E. Ballard, and S. Rajapakse. 2006. Microsatellite marker development in rose and its application in tetraploid mapping. J. Amer. Soc. Hort. Sci. 131(3):380-387.

Zlesak, D.C. 2006. Rose: *Rosa x hybrida*, p. 695-738. In: N.O. Anderson (ed.). Flower breeding & genetics: Issues, challenges, and opportunities for the 21st century. Springer, Dordrecht, The Netherlands.

APPENDIX

A. 2X CTAB Buffer

2X CTAB buffer (100 ml):

2% CTAB	2.00 g
1.4 M NaCl	8.12 g
20 mM EDTA, pH 8.0	4 ml of 0.5 M
100 mM Tris HCL, pH 8.0	10 ml of 1.0 M
1% PVP-40 (polyvinylpyrollidone, M.W. 40,000)	1.00 g
β-Mercaptoethanol*	200 μL

*added immediately before use

VITA

Name:	Valerie Ann Soules
Address:	Dept. of Horticultural Sciences c/o Dr. David Byrne Texas A&M University College Station, TX 77843-2133
Email Address:	valerie.eitreim@tamu.edu
Education:	B.S., Horticultural Science, Minor: Botany, North Dakota State University, May 2007 M.S., Plant Breeding, Texas A&M University, December 2009