Research on Wild Relatives of Fruit and Nut Crops at the Davis Repository

M.K. Aradhya[a] and J.E. Preece
Geneticist and Research Leader
National Clonal Germplasm Repository, USDA-ARS
University of California, Davis
Davis, CA 95616
USA

Keywords: crop wild relatives, genetic characterization, genetic conservation, *Juglans*, phylogeny, *Vitis*

Abstract

The USDA germplasm repository in Davis is responsible for acquiring, conserving and distributing a broad spectrum of diversity of subtropical and temperate fruit and nut species germplasm to stakeholders around the world. Currently the repository holds over 7000 germplasm accessions including *Actinidia* (kiwi fruit), *Diospyros* (persimmon), *Ficus* (fig), *Juglans* (walnuts), *Morus* (mulberry), *Olea* (olive), *Pistacia* (pistachio), *Prunus* (stonefruits and almond), *Punica* (pomegranate), *Vitis* (grape), and other minor genera. Wild relatives are widely represented in the collections with nearly 50% of the accessions and >90% of the taxa (196 out a total of 215 taxa) representing the wild gene pools. Research at the repository is mainly focused on genetic characterization of germplasm using molecular markers and morphological traits to quantify and describe genetic structure and differentiation within and among species and gene pools. Various population genetic, multivariate and phylogenetic approaches are utilized to classify and elucidate genetic and evolutionary relationships within and among taxa and gene pools. Most of our crop genera are tertiary disjuncts with modern distributions showing disjunction between Eurasia and the Americas. There is rich fossil history and excellent opportunities for analyzing the phylogeny and historical biogeography to understand the paleobotanical and evolutionary events that led to the modern disjunctions. In addition to traditional uses of germplasm for genetic improvement of crops, the collections are increasingly being used in association genetic analyses for gene discovery and to dissect complex phenotypes by exploiting historical genetic recombination. We will review several studies we have conducted to illustrate different methods and approaches to characterize germplasm collections, and discuss results and implications for effective conservation, management, and utilization of germplasm collections.

INTRODUCTION

The National Clonal Germplasm Repository (NCGR) in Davis is one of more than 25 federally funded repositories in the National Plant Germplasm System (NPGS) of the United States Department of Agriculture, Agricultural Research Service (USDA-ARS) and operates in cooperation with the University of California, Davis. The primary goals of the repository are to acquire and conserve a broad spectrum of genetic diversity in the genetic resource collections assigned to this repository, and to characterize these collections in the interest of facilitating their use in plant research and crop development. The collections in the Davis repository include fruit and nut crop species that are important to California agriculture in particular and national and global agriculture in general. The current holdings at the repository stand at ~7,200 accessions representing a 12 crop genera comprising ~220 species, of which ~200 are wild relatives (Table 1). Although, most of the crop genera are fairly well represented by a number of wild

[a] aradhya@ucdavis.edu

relatives, the genetic diversity within the wild gene pools is still under represented. In recent years we are focusing on enriching the wild gene pools of these crop genera through exchange and plant explorations in the centers of diversity. All our crops have centers of diversity in the Mediterranean regions, Asia Minor, southern and Eastern Europe, and Central Asia.

Current Status of Conservation of Crop Wild Relatives

Crop Wild Relatives (CWRs) include all those taxa representing spontaneous wild and weedy forms and progenitors that are closely linked to domesticated/cultivated species. CWRs are the reservoirs of genes for tolerance to biotic and abiotic stresses and pests and diseases, early vigor and establishment, and other traits generally lacking in domesticates. From the gene pool perspective, CWRs represent a spectrum of taxonomic variability within and between the primary and secondary gene pools. Gene flow between crops and their wild relatives produces novel genetic variability for continued evolution, domestication and genetic improvement of crop species. Crop wild relatives and their hybrids with domesticates have been utilized in the genetic improvement of scion and rootstock cultivars in tree crops. As major gene donors, CWRs continue to play a significant part in the improvement and sustainable production of domesticated crops. Surprisingly, little is known about the genetic diversity and conservation status of most CWRs of fruit and nut crops.

Over exploitation, fragmentation, and drastic modification of species habitats has contributed to systematic erosion of taxonomic and genetic diversity. This loss of genetic diversity from the reduction of wild populations has reduced our ability to develop new crop varieties in response to current and future agricultural crises likely to be exacerbated by global climate change.

Tremendous efforts by gene banks around the world to assemble crop genetic resources during the past four decades have resulted in the assemblage of huge collections of CWRs in most economically important crops. Nevertheless, the wild relatives of crop species are still seriously under-represented in many collections. Traditionally, germplasm collection efforts and research on wild relatives were focused on annual agronomic crops and little attention and resources have been devoted to research and conservation of wild relatives of woody perennial crops.

Many woody perennial plant collections have representatives of CWRs, however, the genetic variability across the natural range of taxa is usually under represented. This makes the collections less useful for the extraction of functional variation needed for genetic improvement of crops. Although vast scientific knowledge on population and ecological genetics of woody perennials has accumulated, to this day, it has not been utilized effectively to plan and design in situ and ex situ conservation strategies. It is important for food, fodder, and energy security and production to focus on conservation of biological diversity of CWRs.

Genetic Conservation of Woody Perennial Crops

Genetic conservation and management of woody perennial crop species present many unique challenges and opportunities: (1) Because of the heterozygosity of woody perennials, genetic diversity is typically preserved as discrete clones representing local cultivars, indigenous varieties, mutants and bud sports of economic importance, genetic stocks, interspecific hybrids, rootstocks, and clones of wild species; (2) Germplasm is generally collected and propagated through vegetative means and may often be an evolutionary dead-end unlike seed propagation where dynamic mutation-recombination systems permit for continued evolution; (3) Genetic diversity in the clonal collection represents historical genetic structure of the species as they were part of the breeding population in their natural range before they were collected; (4) Conservation of clonal crop germplasm is expensive because they are preserved as live plants/trees under field plantings with backup collections maintained in screen or shade house conditions; (5) Germplasm accessions are exchanged or supplied mostly as vegetative propagules, often

with severe restrictions on plant material movement due to phytosanitary restrictions such as potential virus infection; and (6) In contrast to conservation of genetic diversity in cultivated species as distinct clones/cultivars, genes of undomesticated and wild relatives can be preserved in the form of seeds, seedling progenies, and/or pollen to maximize variability within and among species.

The contribution of genes from wild relatives has not often been fully realized due to the difficulty in making interspecific crosses between wild and domesticated species. Tree crop breeders generally use genetic variability in the primary gene pools, especially highly domesticated elite germplasm to avoid drastically altering the horticultural base while improving one or two traits in an otherwise good cultivar. It is time we look beyond traditional primary gene pools for useful genes to develop new scion and rootstock cultivars that can help address the ever changing needs of growers, consumers, and the markets.

Climate change will continue to offer many challenges to woody species including droughts, insufficient chilling hours for some crops, changes in growing seasons, and effects on insect and pathogen populations. Therefore, breeders will have to move useful functional variability from outside traditional sources.

Developments in molecular biology permit for comparative genomics of cultivated and wild taxa to understand better the genomic composition and organization, and should facilitate transfer of novel traits across wide gene pools into cultivated crop species to expand adaptation, disease-pest resistance and to further break the yield barriers. Identification of genes for resistance or adaptation in wild relatives may be especially valuable since the regulatory hurdles are reduced when gene transfers occur among related plant species rather than non-plant sources.

Genetic Characterization of Germplasm

Population and evolutionary genetics contribute to conservation biology by providing: (1) the basic theory to understand the role of evolutionary forces in shaping genetic diversity, structure and differentiation within and between species populations; (2) the methods to quantify and describe the genetic diversity within and among populations at different hierarchical levels of biological organization; and (3) the conceptual basis for defining and identifying the goals, methods, and priorities for developing genetically sound conservation programs. These studies facilitate achieving one of the major goals of gene banks to conserve the broadest possible genetic diversity found in the target species with the expectation that it will serve as a reservoir of potentially economically useful genes for sustainable improvement of crop species.

Genetic and phenotypic characterization is key for efficient utilization of germplasm collections. Although traditional methods of assessing diversity using morphological and quantitative traits, are expensive and are often influenced by the genotype-environment interactions, they are becoming increasingly important to gain insights into genotype-phenotype correlations. Advances in molecular marker and genomic technologies have led to novel, precise and efficient screening methods to evaluate genetic diversity and structure in germplasm collections. With the advent of molecular techniques, a range of marker systems have become available to detect genetic polymorphisms at the DNA level in germplasm collections (Karp and Edwards, 1997). This knowledge can be used to develop collection, conservation and management strategies, and efficient utilization of genetic diversity in different gene pools. Molecular markers offer direct estimation of genetic diversity at the level of the basic genetic material itself and in combination with classical field characterization provide the most comprehensive evaluation of genetic diversity and structure. Among them, PCR based markers such as SSRs, ISSRs and AFLPs, offer quick and accurate means for genotyping and are preferred over hybridization based markers like RFLPs. SSRs or microsatellite markers are widely available for several woody perennial crops due to the development of methods to produce and enrich genomic libraries for screening and isolation of microsatellites (Ostrander et al., 1992; Edwards et al., 1996; Fisher et al., 1996). Most

importantly, the co-dominant nature of SSRs makes them the markers of choice for population genetic analysis to assess genetic structure and differentiation in germplasm collections and natural populations. Targeted-PCR and sequencing of many *cp*DNA regions, especially non-coding spacers, offer a rapid evaluation of systematic and evolutionary relationships within and between cultivated and wild relatives. Many statistical methods are available to analyze marker data to describe the genetic diversity within and between populations, within and between species, and to understand the dynamics of spatial and temporal patterns (see review by Gonzalez-Candelas and Palacios, 1997).

At the Davis repository, we are evaluating the wild gene pools of all our crops. Comprehensive germplasm characterization is a prerequisite to address the following key issues related to conservation, management, and utilization of germplasm: (1) Authentication of germplasm accessions – establishment of morphological and genetic identity of germplasm accessions by genotyping and phenotyping, and to identify potential duplicates in collections; (2) Lineage based conservation – perform marker and/or sequence based phylogenetic analysis to assess evolutionary relationships, rate of evolution within among different evolutionary lineages and incorporate results into developing conservation and management strategies; (3) Germplasm evaluation – assessment and description of pattern of distribution genetic diversity, structure, and differentiation within and among gene pools; (4) Development of strategies to enrich diversity in the collection and identify deficiencies; (5) Evolve strategies for enriching diversity based on genetic characterization; and (6) Documentation of evaluation data – effective presentation of evaluation data, including morphological, metric, and molecular data in easily understandable tabular and graphic formats on the database, websites, and publication of results in peer reviewed journals.

The following are examples of our gene pool evaluation of two major crops in our germplasm collection, i.e. grape and walnut. We extracted and summarized the key data from earlier publications: Aradhya et al. (2007, 2008).

GENETIC STRUCTURE AND DIFFERENTIATION WITHIN THE GENUS *VITIS* (Extracted and Summarized from Aradhya et al., 2008)

The genus *Vitis* (*Vitaceae*) consists of two subgenera, *Vitis* (2*n*=6*x*=38) comprising all but ~60 taxa described in the genus, and *Muscadinia* (2*n*=6*x*=40), which includes three taxa, *V. rotundifolia*, *V. munsoniana*, and *V. popenoei*. The genus is a Tertiary disjunct with distributions in North America and Eurasia. The biogeographic history of *Vitis* suggests it occurred widely throughout the Northern Hemisphere during much of the Tertiary (Axelrod, 1966) including the Miocene floras western United States (Chaney and Axelrod, 1959), eastern Asia (Chaney and Hu, 1940), and central Europe (Szafer, 1952). *Vitis* is phylogenetically a complex group and much of its taxonomy is solely based on morphological criteria and is riddled with difficulties in establishing the identities of different taxa due to widespread introgression and clonal variation masking taxonomic boundaries. The existing classification schemes (Planchon, 1887; Munson, 1909; Bailey, 1934; Comeaux, 1984) do not agree with each other in the circumscription and delineation of species and their relationships (Barrett et al., 1969).

A population genetic approach was taken to elucidate the evolutionary relationships within the genus *Vitis* using 273 accessions representing 52 taxa with both Old and New World distributions, including the cultivated *V. vinifera* ssp. *vinifera* and its putative progenitor, *V. v*. ssp. *sylvestris*, and three taxa of *Muscadinia* genotyped for 18 SSR loci and 24 *Eco*RI/*Mse*I AFLP primer-pairs. The combined SSR and AFLP binary data were used to compute a standard distance matrix based on the proportion of shared bands between two accessions for all possible pair-wise combinations. Optimum Minimum Evolution (ME) tree was constructed using the Close-Neighbor-Interchange algorithm (Nei and Kumar, 2000) to elucidate the evolutionary history the genus. Bootstrap interior branch test (Dopazo, 1994) was used to test the reliability of each interior branch on the tree. The multilocus SSR data was pooled into series within *Vitis* as

per Comeaux (1984) classification and analyzed to partition the total variance into variation within and among taxa and series the using the analysis of molecular variance (AMOVA) implemented in ARLEQUIN version 3.01 (Excoffier et al., 2005). The variance components from AMOVA were used to estimate the population subdivisions (Φ statistics) within and among taxa and series.

The evolutionary history of the genus *Vitis* inferred from the ME tree (Fig. 1) revealed the identity of many taxa, but the taxonomic affinities at the series level roughly matched the taxonomic series based on morphological criterion within the New World group. The two subgenera showed a major divergence with *Muscadinia* forming a well-supported distinct basal sister group (C-1) to the subg. *Vitis*, which formed four major clusters C-2, C-3, and C-4 representing the New World taxa and the clusters C-5 and C-6 exclusively made up the Old World *Vitis* from the Eurasian and East Asian distributions, respectively.

Within the New World *Vitis*, the ser. *Labruscae* showed differentiation between *V. labrusca* closely allied with *V. aestivalis* var. *aestivalis*, both of which have overlapping distributions in the northeastern U.S. and its sister taxa, *V. mutangensis* and *V. shuttleworthii*, showing affinity with *V. cinerea* var. *cinerea*, all of which are sympatric in the southeastern U.S. Members of the ser. *Precoces*, *V. riparia* and *V. rupestris* which are somewhat parapatric in distribution occupied cluster C-2, while their associate taxon *V. acerifolia* with more westerly distribution clustered within cluster C3. Except for *V. cinerea* var. *cinerea* and *V. tilifoia*, which occupied cluster C-2, all other members of the ser. *Cinerascentes*, *V. c*. var. *berlandieri* syn. *V. c*. var. *helleri*, *V. c*. var. *floridana*, *V. biformis*, and *V. peninsularis* are scattered within the cluster C-3. Member of the ser. *Aestivales* were diverse and found in all three New World clusters. The members of the ser. *Occidentales* mostly from the southwestern U.S., *V. californica, V. treleasei*, and *V. monticola* were found in the cluster C-4 while *V. bloodworthiana* from Baja California and *V. girdiana* from Arizona occupied the cluster C-3. Members of the ser. *Vulpinae*, *V. vulpina* and *V. palmata* with sympatric distribution in the southeastern U.S. occupied the cluster 2a. Among the East Asian taxa, *V. ficifolia* showed differentiation and the identity of the remaining taxa, although evident did not show definite affinities within and between them.

The significant amount of total variation is accounted for within taxa component (~74%) as compared to between taxa within series (~17%) and between series (~9%) suggesting considerable gene flow within and between taxa and series (Table 2). This data is further substantiated by the lack of breeding barriers among taxa with overlapping ecological distributions in the eastern North America, perhaps in East Asia leading to extensive gene flow overriding the factors responsible for genetic differentiation in their natural ranges.

Overall, the East Asian species from China exhibited significant divergence from the North American group, but some aligned with American species/series. Partitioning of molecular variation suggested a significant amount of total variation is accounted for by differences among genotypes within species as compared to among species within series and among series within the genus *Vitis*. Although significant gene flow was evident at all levels of classification there was reasonable differentiation among species and series. The genetic relationships within East Asian taxa need re-examination on a larger sampling basis. Ecological and phenological isolation mechanisms may play a role in reproductive isolation among *Vitis* taxa, but the results suggest significant gene flow within and between taxa and series. Natural selection, drift due to generally skewed sex ratios in natural populations, combined with gene flow within and between taxa and series seem to have played a significant role in shaping the phylogenetic structure within the genus *Vitis*. Genetic structure and differentiation within the subg. *Vitis* suggest a complex interaction of evolutionary forces combined with the persistence of ancestral polymorphisms and reticulate evolution within and between the Eurasian and North American taxa has resulted in incomplete lineage sorting within the subg. *Vitis*.

PHYLOGENY AND BIOGEOGRAPHY OF THE GENUS *JUGLANS* (Extracted and Summarized from Aradhya et al., 2010)

The genus *Juglans* contains ~21 extant taxa divided into four sections mainly based on fruit morphology, wood anatomy, and foliage architecture (Manning, 1978) and has disjunct distribution in East Asia and eastern North America (Manchester, 1987). The genus is divided into four sections based mainly on the morphology (Manning, 1978). Sect. *Rhysocaryon* (black walnuts), which is endemic to the New World, comprises five North American temperate taxa: *J. californica*, *J. hindsii*, *J. nigra*, *J. major*, and *J. microcarpa*; three Central American subtropical taxa: *J. mollis*, *J. olanchana*, and *J. guatemalensis*; and two South American tropical taxa, *J. neotropica* and *J. australis* mainly occurring in the highlands. They typically bear four-chambered nuts with thick nutshells and septa. Sect. *Cardiocaryon* (Asian butternuts) contains four taxa: *J. hopeiensis*, *J. ailantifolia*, *J. mandshurica*, and *J. cathayensis*, all native to East Asia, while sect. *Trachycaryon* consists of the only North American butternut taxon, J. cinerea. Butternuts possess two-chambered nuts with thick nutshells and septa. Sect. *Juglans* includes two taxa: the cultivated Persian or English walnut, *J. regia*, occurring naturally in the Balkan, north Iran, Turkey, the south Caspian region, central Asia, the Himalayas, and China, and bears four-chambered nuts with thin nutshells and papery septa, and cultivated throughout the subtropical regions of the World. The second one is called iron walnut, *J. sigillata*, which is restricted to southern China and Tibetan regions bears thick rough-shelled nuts and the characteristic dark-colored kernels (Dode, 1909), and often considered as an ecotype of *J. regia*, but some botanists have treated it as a separate species (Kuang et al., 1979). Complete descriptions of the morphological variation, ecological distribution, and taxonomic treatment of the genus *Juglans* are found in Manning (1957, 1960, 1978).

Earlier molecular systematic studies based on nuclear RFLPs (Fjellstrom and Parfitt, 1995) and matK and ITS sequences (Stanford et al., 2000) supported the traditional taxonomic classification of *Juglans* and are consistent with what is known about the geological history of the genus (Manning, 1978; Manchester, 1987). In the present study, we examine the utility of some of these *cp*DNA non-coding spacer (NCS) sequences for phylogenetic reconstruction, and for assessing the level of evolutionary divergence within and among sections of *Juglans*. We further explore the biogeography of the genus *Juglans* based on the phylogenetic inferences.

Seventeen taxa representing the four sections of *Juglans* and two outgroup taxa, *Pterocarya stenoptera* and *Carya illinoiensis* were used to reconstruct the phylogeny using sequences from the five *cp*DNA non-coding spacer (NCS) regions: *trnT-trnF* (Hodges and Arnold, 1994), *psb*A-*trn*H (Sang et al., 1997), *atp*B-*rbc*L (Taberlet et al., 1991), *trn*V-16S *r*RNA (Al-Janabi et al., 1994), and *trn*S-*trn*fM (Demesure et al., 1995). The sequences (over 3.8 kb) were aligned region-wise using the software SequencherTM (GeneCodes Corp. Ann Arbor, Michigan) and later adjusted manually. Congruence of IGS sequences was examined with the incongruence length difference (ILD; Farris et al., 1994, 1995) test as implemented in PAUP* (partition homogeneity test).

Phylogenetic analyses were performed with PAUP* 4.0b10 (Swofford, 2002) using the maximum parsimony (MP) and maximum likelihood (ML) methods. MODELTEST version 3.06 (Posada and Crandall, 1998) was used to select appropriate model of evolution for ML analysis. Bootstrap and Decay analyses were performed to estimate the support for different nodes. The molecular clock hypothesis (Zuckerkandl and Pauling, 1965) was tested by computing the difference in the log likelihood scores between ML trees with and without a molecular clock assumption (2Δ = log Lno clock – Log Lclock), which follows the chi-square distribution with (n-2) degrees of freedom where n is the number of sequences or taxa. The age of nodes representing the divergence between the sects. *Juglans* and *Rhysocaryon* (JR) and between *Rhysocaryon* and *Cardiocaryon* (RC) was estimated with the molecular clock assumption as well as using the nonparametric rate smoothing (NPRS; Sanderson, 1997) and the penalized likelihood (PL; Sanderson, 2002) methods.

Phylogeny

The cladograms from both the MP and ML analyses were concordant and contained three well-supported, monophyletic clades corresponding to the sects. *Juglans*, *Cardiocaryon*, and *Rhysocaryon-Trachycaryon* described within the genus *Juglans* (Fig. 2). The ML analysis was performed using the TVM+I+G model identified by the Model Test as the best-fit for the NCS data. The clades exhibit a high degree of differentiation and differ significantly in leaf architecture, wood anatomy, and pollen and fruit morphology (Manchester, 1987). However, monophyly of the genus was not evident, probably due to past extinctions obscuring the evolutionary history. The low consistency index apparently indicates that the spacer regions contain a moderate level of homoplasy across the lineages during the evolution and diversification of *Juglans*.

The single North American butternut species, *J. cinerea* with nut characteristics (two-chambered nuts) resembling the members of sect. *Cardiocaryon*, is placed within the *Rhysocaryon* clade, members of which are characterized by four-chambered nuts with indehiscent hulls. The placement of *J. cinerea* within *Rhysocaryon* was supported in a recent phylogenetic study based on chloroplast *mat*K sequences, whereas the phylogeny based on nuclear internal transcribed spacer (ITS) sequences, nuclear genome RFLPs, and the combined data set placed *J. cinera* sister to *Cardiocaryon* (Fjellstrom and Parfitt, 1995; Stanford et al., 2000). This controversial placement of butternut into the black walnut clade by *cp*DNA, with a strong bootstrap support (90%) and decay index >5, suggests historical introgression of *Rhysocaryon* chloroplast into an ancestral member of sect. *Cardiocaryon*, which may later have given rise to the North American butternut, *Trachycaryon*. The introgression may have occurred during range reduction and selective extinction of juglandaceous taxa in general and of *Juglans* in particular in northern latitudes including some of the ancestral butternuts in North America sometime in the early Neogene. Fossil records indicate that butternuts were widely distributed throughout the northern latitudes during the late Eocene and Oligocene. Chloroplast capturing has been reported in several plant groups, perhaps the best studied is in cotton (Wendel et al., 1991). The present day *Trachycaryon* is represented by a single taxon, J. cinerea, found only in eastern North America and sympatric with members of *Rhysocaryon*.

Members of the sect. *Rhysocaryon* are not well resolved; however, in the MP and ML analyses, they are somewhat segregated into three biogeographic groups reflecting specific adaptations to the temperate, sub-tropical and tropical highland environments in which they are found (Fig. 2). The clade as a whole is well supported with a bootstrap value and decay index of 90% and >5, respectively. Many of these taxa have accumulated a number of autopomorphic mutations along with some homoplasious ones shared mostly within and to a less extant between different clades. The basal placement of southern California black walnut, *J. californica* within the RT clade, well separated from its putative close relatives, *J. hindsii* and *J. major*, is surprising, because *J. hindsii* has often been treated as a conspecific variant within *J. californica* (Wilken, 1993), and a sister relationship between these two taxa has been reported in other studies (Fjellstrom and Parfitt, 1995; Stanford et al., 2000). The basal placement of *J. californica* is probably due to two substitutions that it shares with the sect. *Cardiocaryon*, which may represent convergence. Lower resolution within the black walnut section is probably indicative of: (1) relatively recent diversification, possibly in the Pliocene; (2) reticulate evolution within the section; and (3) persistence of ancestral polymorphisms through speciation. On the contrary the fossil evidence suggests that the earliest evolutionary split within *Juglans* during the middle Eocene involved the origin of black walnut and butternut sections and thus these two sections would have had enough time for inter- and intrasectional diversification.

Sect. *Cardiocaryon* is well supported and resolved as a monophyletic lineage. Within *Cardiocaryon*, *J. hopeiensis* is moderately supported as sister to the remaining three Asian butternuts, *J. ailantifolia*, *J. cathayensis*, and *J. mandshurica*, which are well supported as a clade in all three analyses. In overall tree morphology, *J. hopeiensis* closely resembles the Persian walnut, *J. regia*, but the nut characters are similar to *J.*

mandshurica, and it has been considered as either an interspecific hybrid between *J. regia* and *J. mandshurica* (Rehder, 1940) or as a subspecies of *J. mandshurica* (Kuang et al., 1979). In contrast to earlier studies which placed *J. mandshurica* as sister to *J. ailantifolia* and *J. cathayensis* (Stanford et al., 2000; Fjellstrom and Parfitt, 1995), in our study *J. cathayensis* and *J. mandshurica* are closely united with five unique synapomorphies.

The Persian walnut, *J. regia*, and its sister taxon *J. sigillata* (sect. *Juglans*), form a distinct clade sister to both *Cardiocaryon* and *Rhysocaryon-Trachycaryon* in both MP and ML analyses. This is in contrast to earlier studies, which placed the cultivated walnut, *J. regia* within either *Cardiocaryon* (Fjellstrom and Parfitt, 1995; Stanford et al., 2000) or *Rhysocaryon* (Manos and Stone, 2001). The early evolutionary split of this clade within the genus *Juglans* contradicts the traditional taxonomic treatments and fossil evidence, both of which supported the almost simultaneous ancient divergence of sects. *Cardiocaryon* and *Rhysocaryon*, and the origin of the genus sometime in the middle Eocene (Manchester, 1987). Within the sect. *Juglans*, the cultivated species, *J. regia* with thin-shelled four chambered nuts has differentiated from its sister taxon *J. sigillata*, which retains many primitive nut characteristics such as thick rough-shelled nuts with dark colored kernels (Dode, 1909a) and may represent a semi-domesticated form within the section.

Biogeography

The extant species of *Juglans* show an intercontinental disjunction with the modern distributions of sects. *Juglans* and *Cardiocaryon* limited to Eurasia and sect. *Rhysocaryon* endemic to the Americas. A single butternut species, *J. cinerea*, with modern distribution in eastern North America, is generally considered to be a disjunct of *Cardiocaryon* (Asian butternuts) (Manchester, 1987). These disjunctions could have arisen as a result of either a vicariance event disrupting the geographic continuity of ancestral populations that once spanned from Eurasia to North America, or a long-distance dispersal from one region to the other. The vicariance hypothesis is favored over the long-distance dispersal theory because of the large fruit size in *Juglans*, which does not appear to have great dispersal ability.

Based on fossil evidence, Manchester (1987) proposed that the divergence of *Petrocarya* and *Juglans* may have occurred sometime during the late Paleocene or early Eocene (~54 Mya), and that the initial split of sects. *Rhysocaryon* and *Cardiocaryon* probably occurred during the middle Eocene (45 Mya) in North America, but the two sections were clearly resolved only in the early Oligocene (38 Mya). However, Hills et al. (1974) based on extensive analysis of nut specimens of a fossil walnut, *J. eocinerea* from the Beaufort Formation (Tertiary), southwestern Banks Island, arctic Canada, concluded that it is closely related and probably ancestral to fossil *J. tephrodes* from the Early Pliocene Germany and the extant *J. cinerea* from the eastern United States. Further, they argued that butternuts may have evolved independently in the Arctic attaining a broad distribution in the upper latitudes of the Northern Hemisphere by the Miocene and that subsequent geoclimatic changes (Axelrod and Bailey, 1969; Wolfe, 1971) resulted in the southward movement of the floras across the Bering Strait. However, the early Pleistocene glaciations have completely eliminated butternuts from Europe and northwestern parts of North America leaving small disjunct populations in eastern Asia to evolve into three major present day taxa, *J. cathayensis*, *J. mandshurica*, and *J. ailantifolia*, and one south of the glacial limit in North America to evolve to its present form, *J. cinerea*. The geographic and stratigraphic fossil distribution strongly supports the above hypothesis that butternuts may have originated and radiated from high northern latitudes. At about the same time, black walnuts spanned throughout North America and extended into the Southern Hemisphere reaching Ecuador by the late Neogene, and remained endemic to the Americas throughout their evolutionary history.

One can argue, if butternuts and black walnuts diverged from a common ancestor in North America during the middle Eocene, as suggested by Manchester (1987), there would have been ample opportunity for both groups to become established in both Asia

and North America, because both the Bering and North Atlantic land bridges were in continuous existence throughout the Eocene. In addition, there was still a favorable climate in these upper latitudes for the establishment and dispersal of broad-leaved deciduous taxa (Wolfe, 1972, 1978; Tiffney, 1985). However, the distributional range of the Tertiary fossils of butternuts and black walnuts does not overlap except in the northwestern parts of the United States around 40^{o} N latitude, strongly suggesting they may have evolved independently as suggested by Hills et al. (1974). The weak support for the sister relationship between these two groups observed in our phylogenetic analysis further substantiates this point and also suggests they may not share an immediate common ancestor or it may not be represented among the extant taxa.

An analysis of the comparative rates of molecular evolution along the branches of the cladogram indicated the rates did not conform to the expectation of the molecular clock hypothesis (Zuckerkandl and Pauling, 1965). Nevertheless, it has been shown that estimations of nucleotide substitution rates, or time since divergence using the molecular clock hypothesis, although based on uncertain assumptions and approximate values, are still helpful to understand the tempos of evolution and plant historical geographies (Parks and Wendel, 1990; Crawford et al., 1992). The cladogenesis and estimates of time since divergence for JR and RC nodes suggest that the diversification within the genus must have occurred sometime in middle of Eocene and ancestors of the sect. *Juglans* were the first to split (47-50 mya) following the divergence between *Rhysocaryon* and *Cardiocaryon* (41-45 mya) (Fig. 3). On the contrary, the fossil evidence suggests the split between *Rhysocaryon* and *Cardiocaryon* almost simultaneously with the origin of the genus itself or sometime soon after. However, the low resolution within *Rhysocaryon* suggests either relatively recent diversification sometime during the late Pliocene and early Pleistocene or extinction of many ancestral taxa at the base of the clade. The extant taxa within the clade exhibit some level of differentiation into temperate, subtropical and tropical groups. Finally, the evolutionary history of *Juglans* is riddled with range reduction, geographic isolation, local and regional extinctions within and between clades, and as consequence the extant taxa may not adequately represent the entire evolutionary history of the genus.

CONCLUSIONS

Crop wild relatives are generally under-represented in most plant genetic resource collections around the world. This is especially true with woody perennial crops. While sincere efforts are being made to collect and establish ex situ collections, very little or no systematic attempts have been made to develop in situ conservation strategies for woody perennial crops. Wild relatives are the major sources of useful genes that are selectively maintained as co-adapted gene complexes through delicate balance of evolutionary forces over millions of years. They offer an abundant supply of genetic variability for crop genetic improvement to aid in the development of new cultivars resistant to ever-changing biotic and abiotic stresses.

Phenomic and genomic characterization of CWRs is the key to; (1) efficiently utilizing the functional variability within and among populations and taxa for crop improvement; (2) identifing deficiencies in collections and planning for tactically enriching ex situ collections and developing effective in situ conservation strategies for long term conservation and sustainable utilization; and (3) facilitating gene discovery by establishing genotype-phenotype correlations to exploit the hidden functional genetic variability.

Literature Cited

Al-Janabi, M., McClelland, M., Petersen, C. and Sobral, W.S. 1994. Phylogenetic analysis of organellar DNA sequences in the Andropogoneae: Saccharinae. Theor. Appl. Genet. 88:933-944.

Aradhya, M.K., Potter, D., Gao, F. and Simon, C.J. 2007. Molecular phylogeny of *Juglans* (*Juglandaceae*): a biogeographic perspective. Tree Genetics and Genomics

3:363-378.
Aradhya, M., Koehmstedt, A., Prins, B.H., Dangl, G.S. and Stover, E. 2008. Genetic structure, differentiation, and phylogeny of the genus *Vitis*: Implications for genetic conservation. Acta Hort. 799:43-49.
Axelrod, D.I. 1966. Origin of deciduous and evergreen habits in temperate forests. Evolution 20:1-15.
Axelrod, D.I. and Bailey, H.P. 1969. Paleotemperature analysis of Tertiary floras. Palaeogeography, Palaeoclimatology, Palaeoecology 6:163-195.
Bailey, L.H. 1934. The species of grapes peculiar to North America. Gentes Herbarum 3:151-244.
Barrett, H.C., Carmer, S.G. and Rhodes, A.M. 1969. A taximetric study of interspecific variation in *Vitis*. Vitis 8:177-187.
Chaney, R.W. and Axelrod, D.I. 1959. Miocene floras of the Columbia Plateau. Publication of Carnegie Institute, Washington, 617, Pt. II.
Chaney, R.W. and Hu, H.H. 1940. A Miocene flora from Shantung province, China. Publication of Carnegie Institute, Washington, 507.
Comeaux, B.L. 1984. Taxonomic studies on certain native grapes of eastern North Carolina. Ph.D. dissertation, North Carolina State University, Raleigh. 178 p.
Crawford D.J., Lee, M.S. and Stuessy, T.F. 1992. Plant species disjunctions: Perspectives from molecular data. Aliso 13:395-409.
Demesure, B., Sodzi, N. and Petit, R.J. 1995. A set of universal primers for amplification of polymorphic non-coding regions of mitochondrial and chloroplast DNA in plants. Mol. Ecol. 4:129-131.
Dode, L.A. 1909. Contribution to the study of the genus *Juglans* (English translation by R.E. Cuendett). Bull. Soc. Dendrology, France 11:22-90.
Dopazo, J. 1994. Estimating errors and confidence intervals for branch lengths in phylogenetic trees by bootstrap approach. J. Mol. Evol. 38:300-304.
Edwards, K.J., Barker, J.H.A., Daly, A., Jones, C. and Karp, A. 1996. Microsatellite libraries enriched for several microsatellite sequences in plants. BioTechniques 20:758-760.
Excoffier, L., Laval, G. and Schneider, S. 2005. Arlequin ver. 3.0: An integrated software package for population genetics data analysis. Evol. Bioinformatics Online 1:47-50.
Farris, J.S., Kallersjo, M., Kluge, A.G. and Bult, C. 1995. Constructing a significance test for incongruence. Systematic Biol. 44:570-572.
Farris, J.S., Kallersjo, M., Kluge, A.G. and Bult, C. 1994. Testing significance of incongruence. Cladistics 10:315-319.
Fisher, P.J., Gardner, R.C. and Richardson, T.E. 1996. Single locus microsatellites isolated using 5' anchored PCR. Nucl. Acids Res. 24:4369-4371.
Fjellstrom, R.G. and Parfitt, D.E. 1995. Phylogenetic analysis and evolution of the genus *Juglans* (*Juglandaceae*) as determined from nuclear genome RFLPs. Plant Systematics Evol. 197:19-32.
Gonzalez-Candelas, F. and Palacios, C. 1997. Analyzing molecular data for studies of genetic diversity. In: W.G. Ayad, T. Hodgkin, A. Jaradat and V.R. Rao (eds.), Molecular genetic techniques for plant genetic resources, Report of an IPGRI workshop. Rome, Italy 9-11 October 1995.
Hills, L.V., Klovan, J.E. and Sweet, A.R. 1974. *Juglans eocinerea* n. sp., Beaufort Formation (Tertiary), southwestern Banks Inland, Arctic Canada. Canadian J. Bot. 52:65-90.
Hodges, S.A. and Arnold, M.L. 1994. Columbines: a geographic widespread species flock. Proc. Natl. Acad. Sci., USA 91:5129-5132.
Karp, A. and Edwards, K.J. 1997. Molecular techniques in the analysis of the extent and distribution of genetic diversity. In: W.G. Ayad, T. Hodgkin, A. Jaradat, and V.R. Rao (eds.), Molecular genetic techniques for plant genetic resources, Report of an IPGRI workshop. Rome, Italy 9-11 October 1995.
Kuang, K., Cheng, S., Li, P. and Lu, P. 1979. Juglandaceae (In Chinese, unpublished

translation provided by W.E. Manning). p.8-42. In: K.Z. Kuang and C.P. Li (eds.), Flora Reipublicae Popularis Sinicae, Vol. 21. Institutum Botanicum Academiae Sinicae, Peking.
Manchester, S.R. 1987. The fossil history of *Juglandaceae*. Missouri Botanical Garden Monograph 21:1-137.
Manning, W.E. 1957. The genus *Juglans* in Mexico and Central America. J. Arnold Arboretum 38:121-150.
Manning, W.E. 1960. The genus *Juglans* in South America and West Indies. Brittonia 12:1-26.
Manning, W.E. 1978. The classification within the Juglandaceae. Annals of Missouri Botanical Garden 65:1058-1087.
Manos, P.S. and Stone, D.E. 2001. Evolution, phylogeny, and systematics of the Juglandaceae. Annals of Missouri Botanical Garden 88:231-269.
Munson, T.V. 1909. Foundations of American grape culture. T.V. Munson and Son, Denison, Texas. 252p.
Nei, M. and Kumar, S. 2000. Molecular Evolution and Phylogenetics. Oxford University Press. 333p.
Ostrander, E.A., Jong, P.M., Rine, J. and Duyk, G. 1992. Construction of small-insert genomic DNA libraries highly enriched for microsatellite repeat sequences. Proc. Natl. Acad. Sci., USA 89:3419-3423.
Parks, C.R. and Wendel, J.F. 1990. Molecular divergence between Asian and North American species of Liriodendron (*Magnoliaceae*) with implications for interpretation of fossil floras. Amer. J. Bot. 77:1243-1256
Planchon, J.E. 1887. Monographie des Ampelideae vraies. Monographia Phanerogamerum 5:305-368.
Posada, D. and Crandall, K.A. 1998. MODELTEST: testing the model of DNA substitution. Bioinformatics 14:817-818.
Rehder, A. 1940. Manual of cultivated trees and shrubs in North America. MacMillan Co., New York, USA.
Sanderson, M.J. 1997. A nonparametric approach to estimating divergence times in the absence of rate consistency. Mol. Biol. Evol. 14:1218-1231.
Sanderson, M.J. 2002. Estimating absolute rates of molecular evolution and divergence times: A penalized likelihood approach. Mol. Biol. Evol. 19:101-109.
Sang, T., Crawford, D. and Stuessy, T. 1997. Chloroplast DNA phylogeny, reticulate evolution, and biogeography of Paeonia (*Paeoniaceae*). Amer. J. Bot. 84:1120-1136.
Stanford, A.M., Harden, R. and Parks, C.R. 2000. Phylogeny and biogeography of *Juglans* (*Juglandaceae*) based on matK and ITS sequence data. Amer. J. Bot. 87:872-882.
Szafer, W. 1952. An outline of general plant geography. State Scientific Publishers, Warsaw. 428 p.
Taberlet, P., Gielly, L., Pautou, G. and Bouvet, J. 1991. Universal primers for amplification of three non-coding regions of chloroplast DNA. Plant Mol. Biol. 17:1105-1109.
Tiffney, B.H. 1985. Perspectives on the origin of the floristic similarity between eastern Asia and eastern North America. J. Arnold Arboretum 66:73-94.
Wendel, J.F., Stewart, J. McD. and Rettig, J.H. 1991. Molecular evidence of homoploid reticulate evolution among Australian species of Gossypium. Evolution 45:694-711.
Wilken, D.H. 1993. Juglandaceae. p.709. In J.C. Hickman (ed.), The Jepson manual: higher plants of California. Univ. of California Press, Berkeley, CA, USA.
Wolfe, J.A. 1971. Tertiary climatic fluctuations and methods of analysis of Tertiary floras. Palaeogeography, Palaeoclimatology, Palaeoecology 9:27-57.
Wolfe, J.A. 1972. An interpretation of Alaskan Tertiary floras. p.201-233. In: A. Graham (ed.), Floristics and paleofloristics of Asia and estern North America. Elsevier, Amsterdam, The Netherlands.
Wolfe, J.A. 1978. A paleobotanical interpretation of Tertiary climates in the Northern

Hemisphere. Amer. Scientist 66:694-703.
Zuckerkandl, E. and Pauling, L. 1965. Evolutionary divergence and convergence in proteins. p.97-106. In: V. Bryson (ed.), Evolving genes and proteins. Academic Press, New York, NY, USA.

Tables

Table 1. Current status of germplasm collection at the National Clonal Germplasm Repository, USDA-ARS, Davis, California, USA.

Genus	Cultivated spp.	Wild spp.	Wild (% total)	Cultivated accessions	Wild accessions	Wild (% total)
Actinidia	2	25	93	56	220	80
Diospyros	1	4	80	138	24	15
Eriobotrya	1	1	50	35	1	3
Ficus	1	3	75	332	20	6
Juglans	1	22	96	435	218	33
Morus	1	5	83	32	44	58
Olea	1	2	67	144	11	7
Pistacia	1	12	92	201	133	40
Prunus	7	83	92	1070	599	36
Punica	1	0	0	233	0	0
Vitis	2	44	96	1408	1918	58
Total	19	201	91	4084	3188	44

Table 2. Hierarchical partitioning of molecular variation in the genus *Vitis*.

Source of variation	d.f.	Sum of squares	Variance components	Percentage of variation
Among series	13	258.933	0.22922**	8.76
Among species within series	46	294.104	0.44116**	16.85
Within species	626	1219.224	1.94764**	74.39
Total	685	1772.261	2.61803	

Fixation Indices
ΦST = 0.25607** Nm = 0.7262
ΦSC = 0.18468** Nm = 1.1032
ΦCT = 0.08756** Nm = 2.6200
Nm = Number of migrants per generation

Figures

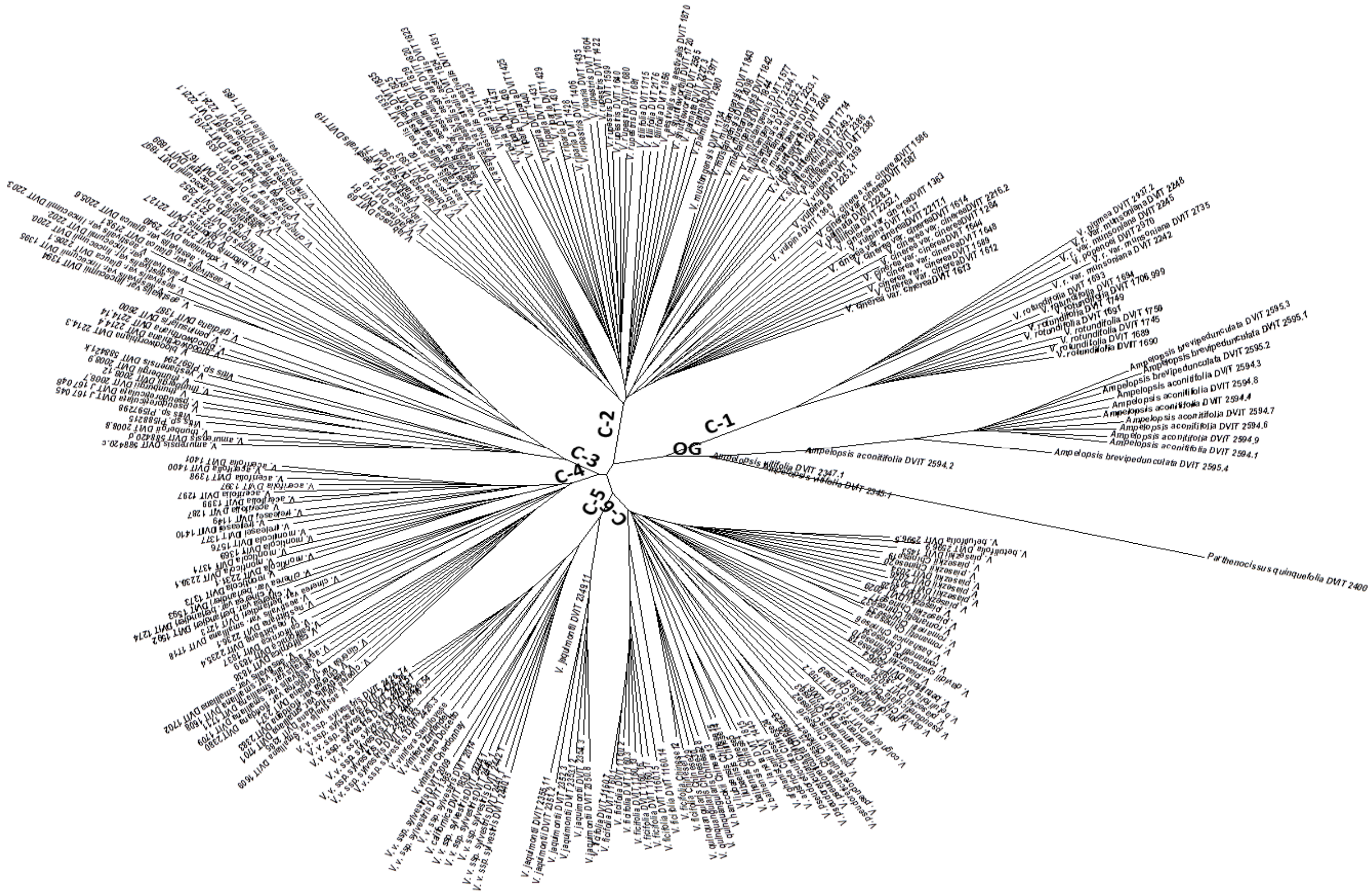

Fig. 1. Minimum evolution tree depicting the evolutionary relationships within the genus *Vitis*. All the clusters (C-1 through C-6) are well-supported at the cutoff confidence probability value ($P \geq 0.95$) as per bootstrap interior branch test.

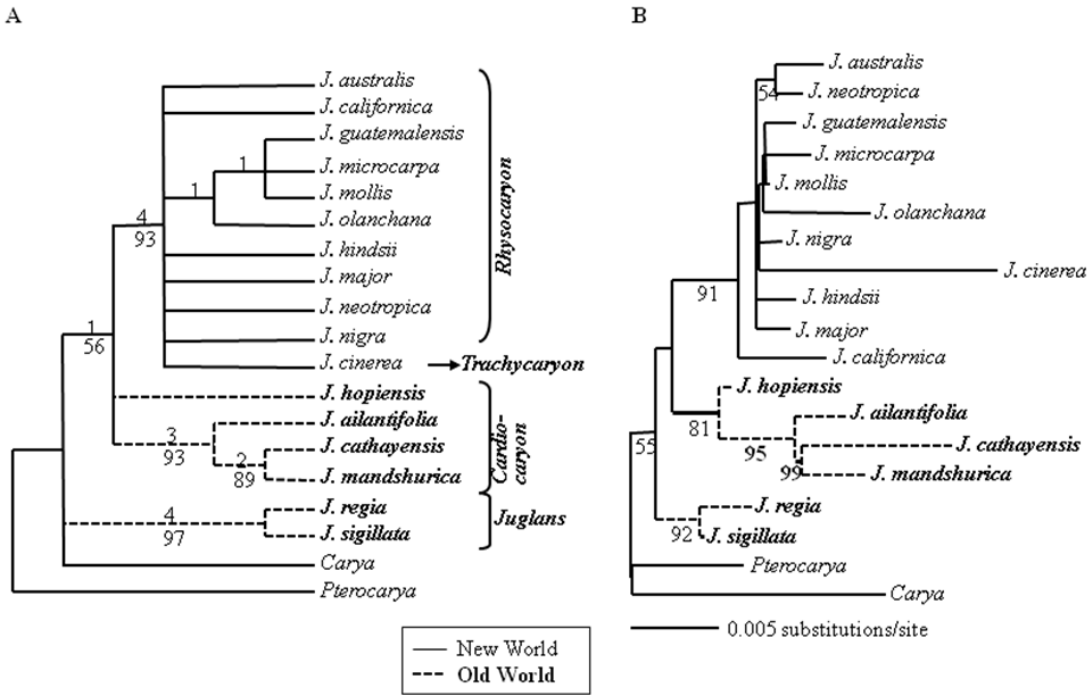

Fig. 2. Phylogenetic analyses of *Juglans* using the *cp*DNA non-coding spacer sequence data. A) Strict consensus tree of 197 equally most parsimonious trees, length = 168 steps, CI = 0.635 (excluding autapomorphies), RI = 0.810. Numbers above branches are decay values, numbers below branches are bootstrap values. B) Maximum likelihood tree based on TVM+I+G model of sequence evolution with the best log-likelihood score (-6251.1768).

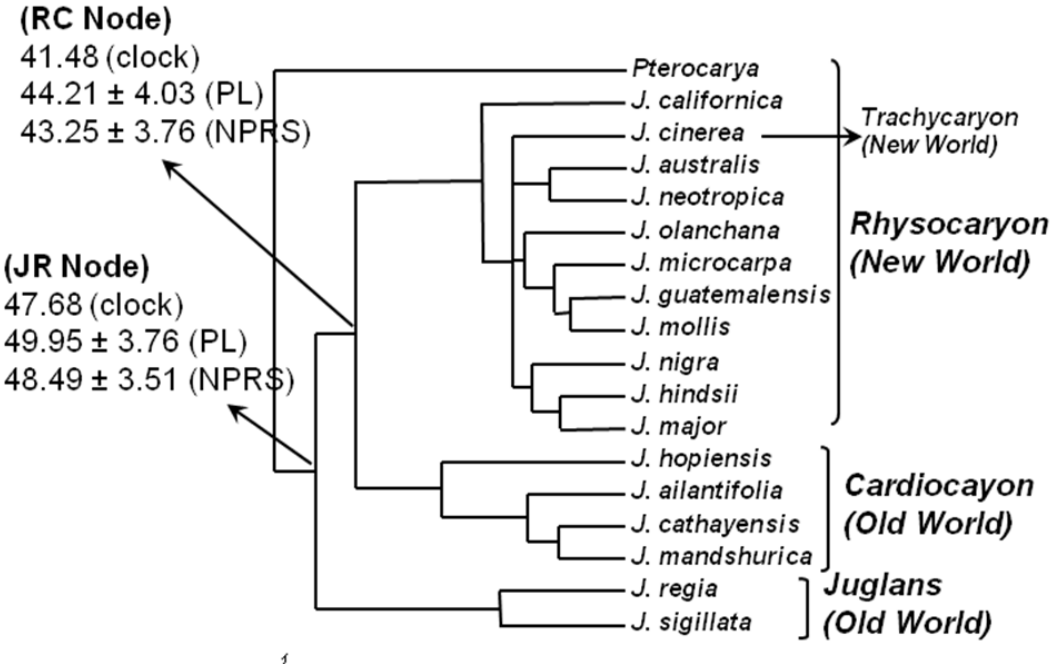

Fig. 3. Linearized tree of *cp*DNA non-coding spacer sequence data showing the time of divergence between the sects. *Juglans-Rhysocaryon* (JR Node) and *Rhysocaryon-Cardiocaryon* (RC Node) in the genus *Juglans*.