Molecular Phylogenetics and Evolution 48 (2008) 23–33

Contents lists available at ScienceDirect

Molecular Phylogenetics and Evolution

journal homepage: www.elsevier.com/locate/ympev

A species-level phylogenetic study of the *Verbena* complex (Verbenaceae) indicates two independent intergeneric chloroplast transfers

Yao-Wu Yuan *, Richard G. Olmstead

Department of Biology, University of Washington, P.O. Box 355325, Seattle, WA 98195 5325, USA

ARTICLE INFO

Article history:
Received 8 August 2007
Revised 17 March 2008
Accepted 5 April 2008
Available online 14 April 2008

Keywords:
Species-level phylogenetics
Chloroplast transfer
cpDNA
Waxy
Verbena
Glandularia
Junellia

ABSTRACT

Two major impediments to infer plant phylogenies at inter- or intra- species level include the lack of appropriate molecular markers and the gene tree/species tree discordance. Both of these problems require more extensive investigations. One of the foci of this study is examining the phylogenetic utility of a combined chloroplast DNA dataset (> 5.0 kb) of seven non-coding regions, in comparison with that of a large fragment (ca. 3.0 kb) of a low-copy nuclear gene (*waxy*), in a recent, rapidly diversifying group, the *Verbena* complex. The complex includes three very closely related genera, *Verbena* (base chromosome number $x = 7$), *Glandularia* ($x = 5$), and *Junellia* ($x = 10$), comprising some 150 species distributed predominantly in South and North America. Our results confirm the inadequacy of non-coding cpDNA in resolving relationships among closely related species due to lack of variation, and the great potential of low-copy nuclear gene as source of variation. However, this study suggests that when both cpDNA and nuclear DNA are employed in low-level phylogenetic studies, cpDNA might be very useful to infer organelle evolutionary history (e.g., chloroplast transfer) and more comprehensively understand the evolutionary history of organisms. The phylogenetic framework of the *Verbena* complex resulted from this study suggests that *Junellia* is paraphyletic and most ancestral among the three genera; both *Glandularia* and *Verbena* are monophyletic and have been derived from within *Junellia*. Implications of this phylogenetic framework to understand chromosome number evolution and biogeography are discussed. Most interestingly, the comparison of the cpDNA and nuclear DNA phylogenies indicates two independent intergeneric chloroplast transfers, both from *Verbena* to *Glandularia*. One is from a diploid North American *Verbena* species to a polyploid North American *Glandularia* species. The other is more ancient, from the South American *Verbena* group to the common ancestor of a major *Glandularia* lineage, which has radiated subsequently in both South and North America. The commonly assumed introgressive hybridization may not explain the chloroplast transfers reported here. The underlying mechanism remains uncertain.

1. Introduction

It is a common observation that plant phylogenetics at the species or population level has lagged behind deeper-level phylogenetic studies, largely due to the lack of appropriate molecular markers (Schaal et al., 1998; Small et al., 1998; Bailey et al., 2004; Shaw and Small, 2004; Shaw et al., 2005, 2007; Hughes et al., 2006). Chloroplast and mitochondrial DNA sequences have too little variation for many plant groups at species or population level, and so do the conventional nuclear ITS sequences (Hughes et al., 2006). Single or low-copy nuclear DNA sequences may have sufficient variation, but inferences from these data are often confounded by their four times longer coalescent time in comparison with organelle DNA (Moore, 1995), which is related to the other major impediment—the potential gene tree/species tree discordance (Pamilo and Nei, 1988; Doyle, 1992; Maddison, 1997). Random lineage sorting accounts for much of this gene tree/species tree discordance as well as the incongruence among gene trees inferred from different loci (Pamilo and Nei, 1988; Maddison, 1997). It is particularly problematic for recent and rapidly diversifying species, because times between species divergences are too short for the loss of ancestral polymorphism and thus lineage sorting is largely incomplete, leading to the gene tree/species tree discordance. Despite these difficulties, the importance of a well-resolved species level phylogeny in studying organismal evolution, including hybridization, introgression, polyploidization, adaptive speciation, character evolution, etc., can never be overemphasized.

Methods for extracting information from incongruent gene trees to infer a species tree recently have been developed (Maddison and Knowles, 2006; Liu and Pearl, 2006; Ané et al., 2007; Carstens and Knowles, 2007), and these methods are likely to become more sophisticated by future improvements. Nonetheless, choice of appropriate molecular markers for inferring plant phylogenies

* Corresponding author.
E-mail address: colreeze@u.washington.edu (Y. Yuan).

1055-7903/$ - see front matter © 2008 Elsevier Inc. All rights reserved.
doi:10.1016/j.ympev.2008.04.004

at inter- or intra-species level requires more extensive investigations. To address this issue, we are carrying out a series of studies that examine the utility of various molecular markers, including non-coding chloroplast DNA (cpDNA), large fragments of low-copy nuclear genes, and transposable element insertions, in resolving interspecies relationships in a recent and rapidly diversifying group, the *Verbena* complex (Verbenaceae). This paper represents the first part of these studies and is focused on the application of combined cpDNA sequences of relatively fast-evolving non-coding regions, in comparison with that of a large fragment of a low-copy nuclear gene.

Seven non-coding chloroplast regions were chosen for sequencing based on a previous study that examined relative utility of 21 non-coding cpDNA regions for general phylogenetic analysis (Shaw et al., 2005). These include intergenic spacers and/or introns in *trn*D–*trn*T, *trn*S–*trn*G, *trn*S–*trn*fM, *trn*T–*trn*L, *trn*G, *trn*L, and *trn*L–*trn*F, totaling ca. 5.3 kb in the *Verbena* complex. The nuclear DNA sequence data are from the granule-bound starch synthase I (GBSSI or *waxy*) gene. The intron sequences of this gene have been used to resolve relationships among closely related species in *Castilleja* (Orobanchaceae), and optimized primers to amplify this locus in Lamiales are available (Tank, 2006). The region we sequenced is from the end of exon 7 to the beginning of exon 13, which is about 3.0 kb long in the *Verbena* complex due to the relatively large introns.

The *Verbena* complex includes *Verbena*, *Glandularia*, and *Junellia*, with each genus containing 40–50 species (Botta, 1989; Sanders, 2001). The relationships among these genera are complex, and they were often combined in the genus *Verbena* in early treatments (Briquet, 1895; Perry, 1933; Troncoso, 1974). The most consistent character to separate these three groups is base chromosome number. *Verbena* has a base chromosome number of $x = 7$ and *Glandularia* has $x = 5$ (Lewis and Oliver, 1961; Umber, 1979). *Junellia*, with $x = 10$, was suggested to be more closely related to *Glandularia* than to *Verbena*, and the base number $x = 10$ was thought to be derived from $x = 5$ through polyploidy (Botta and Brandham, 1993). The underlying assumption of these hypotheses is that the smaller base number, $x = 5$, is ancestral to $x = 10$. A phylogenetic framework is highly desirable to evaluate such prior assumptions. *Junellia* is restricted to South America and contains both diploids and polyploids, whereas both *Glandularia* and *Verbena* have a disjunct distribution between the arid regions of temperate North and South America. A curious reciprocal cytogeographic pattern was pointed out by Lewis and Oliver (1961). *Verbena* is predominantly diploid in North America and mostly polyploid in South America, whereas *Glandularia* is predominantly diploid in South America and entirely polyploid in North America (Table 1). A robust phylogeny is essential to understand this intriguing cytogeographic pattern.

Chloroplast transfer (or "chloroplast capture"), the introgression of a chloroplast from one species into another, has been reported many times, including the well-known examples of *Gossypium*, *Helianthus*, *Quercus* (reviewed by Rieseberg and Soltis, 1991). Since this review, phylogenetic studies have suggested widespread chloroplast introgression in various taxa, mainly on the basis of incongruence between cpDNA and nuclear ribosomal DNA (ITS) phylogenies. The majority of these previously reported examples are interspecies introgressions, with only one case of intergeneric introgression (Soltis et al., 1991). Our study of the *Verbena* complex indicates two independent intergeneric chloroplast transfers, both from the genus *Verbena* to *Glandularia*. While one of these transfers is recent, the other is more ancient, from a source in South American *Verbena* to the common ancestor of a major *Glandularia* lineage that has radiated subsequently in both South and North America.

2. Materials and methods

2.1. Taxon sampling

Fifty species of the *Verbena* complex and two outgroup species have been included in this study. Our sampling represents both geographic distribution and ploidy level (Table 1). Outgroup species (*Lippia salsa* and *Aloysia virgata*) were chosen based on a preliminary phylogenetic study of Verbenaceae (Olmstead, unpublished data), which identified a sister group relationship between the *Verbena* complex and the *Lippia*/*Lantana*/*Aloysia* complex.

2.2. Molecular methods

Total DNA was extracted from either field-collected, silica-gel dried tissue (45 species) or herbarium specimens (7 species, Table 1) using the modified CTAB method (Doyle and Doyle, 1987). PCR primers with corresponding references are listed in Table 2. The seven cpDNA regions were amplified in four fragments. The *trn*S–*trn*G spacer and *trn*G intron were amplified together as the *trn*S–G fragment, and the *trn*T–*trn*L spacer, *trn*L intron, and *trn*L–*trn*F spacer were amplified together as the *trn*T–F fragment. However, for some of the herbarium specimens, these regions were amplified in overlapped smaller pieces. PCR procedures for these cpDNA regions were as follows: one cycle of 3 min at 94 °C, 4 cycles of 15 s at 94 °C, 15 s at 55 °C, and 90 s at 72 °C, followed by 30 cycles of 15 s at 94 °C, 15 s at 50 °C, and 60 s at 72 °C, with a final extension step for 10 min at 72 °C. Amplified PCR products were purified by precipitation from a 20% polyethylene glycol (PEG) solution and washed in 70% ethanol prior to sequencing. To ensue accuracy, both strands of the cleaned PCR products were sequenced using the DYEnamic ET Terminator Cycle Sequencing Kit (Amersham Biosciences, Piscataway, New Jersey, USA).

For the nuclear DNA data, PCR reactions were performed using PfuUltra™ II Fusion HS DNA Polymerase (Stratagene, La Jolla, CA, USA) to reduce erroneous nucleotide incorporation during PCR. Original PCR and sequencing primers used to amplify and sequence the *waxy* locus (exon 7–13) were designed for Lamiales (Tank, 2006). Once a number of sequences were obtained from species of the *Verbena* complex, specific primers were designed for generating the *waxy* data (Table 2). PCR reactions were performed under the following conditions: one cycle of 3 min at 94 °C, 4 cycles of 15 s at 94 °C, 15 s at 57 °C, and 3.5 min at 72 °C, followed by 30 cycles of 15 s at 94 °C, 15 s at 55 °C, and 3 min at 72 °C, with a final extension step for 10 min at 72 °C. The combination of PfuUltra™ II Fusion HS DNA Polymerase (Stratagene, La Jolla, CA, USA), specific primers, and our PCR conditions is such that a strong single band was produced for all samples that have good quality total DNA. PCR products were cleaned by gel excision using the Wizard SV Gel and PCR Clean-Up System (Promega, Madison, Wisconsin, USA). After A-tailing and subsequent clean up, the PCR products were cloned using the TOPO TA Cloning kit (Invitrogen, Carlsbad, CA, USA). Depending on the ploidy level, 8–40 positive clones were screened by sequencing with one primer. Distinct clones were sequenced for the entire region in both strands. *Waxy* sequences were successfully collected from 35 out of the 52 taxa sampled in this study. There was difficulty in amplifying the entire target region (ca. 3 kb) in the remaining taxa, particularly those for which DNA was extracted from herbarium specimens. Sequences generated in this study have been deposited in GenBank (*trn*T–F, EF571518–EF571570; *trn*S–fM, EF576665–EF576717; *trn*S–G, EF583067–EF583119; *trn*D–T, EF583120–EF583172; *waxy*, EF584666–EF584737).