Molecular Phylogenetics and Evolution 50 (2009) 608–617

Contents lists available at ScienceDirect

Molecular Phylogenetics and Evolution

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

Phylogenetic utility of nuclear nitrate reductase: A multi-locus comparison of nuclear and chloroplast sequence data for inference of relationships among American Lycieae (Solanaceae)

Rachel A. Levin *, Jessica Blanton, Jill S. Miller

Department of Biology, McGuire Life Sciences Building, Amherst College, Amherst, MA 01002, USA

ARTICLE INFO

Article history:
Received 28 August 2008
Revised 13 November 2008
Accepted 4 December 2008
Available online 13 December 2008

Keywords:
GBSSI
ITS
Lycium
Nitrate reductase
Phylogeny
Solanaceae
trnD–trnT
trnT–trnF

ABSTRACT

There has been considerable interest and research into the benefits of multiple low-copy nuclear regions for phylogenetic studies at low taxonomic levels. In this study, the phylogenetic utility of DNA sequence data from two low-copy nuclear genes, nitrate reductase (NIA) and granule-bound starch synthase I (GBSSI), was compared with data from nuclear ribosomal ITS and the cpDNA spacers *trnT–trnF* and *trnD–trnT* across 33 closely related taxa in tribe Lycieae (Solanaceae). The NIA data (introns 1–2) had the most parsimony-informative characters, with over twice the number provided by GBSSI, but NIA also had greater homoplasy. Although gene trees were generally concordant across the four regions, there were some notable areas of incongruence, suggesting both incomplete lineage sorting as well as possible reticulate origins.

1. Introduction

Efforts at unraveling phylogenetic relationships among angiosperms using molecular sequence data have largely relied on data from either the chloroplast genome (cp; reviewed in Clegg and Zurawski, 1992; Shaw et al., 2005, 2007) or the nuclear ribosomal internal transcribed spacer regions (ITS; Baldwin, 1992; Baldwin et al., 1995; Álvarez and Wendel, 2003). Likely this is due to the availability of universal primers (e.g., Taberlet et al., 1991; Baldwin et al., 1995) and to the relative ease of PCR amplification given high copy number. Further, the combination of these two types of data often provides robust phylogenetic support, given a mixture of the generally more slowly evolving cp regions and the faster evolving ITS region (e.g., see review in Small et al., 2004; Mort et al., 2007). Inclusion of data from both the maternally evolving plastid genome and the biparentally inherited nuclear genome can also be advantageous in the detection of hybridization. For example, disagreement in the placement of species in cp and ITS trees has been used to suggest hybrid origins for several angiosperm lineages (Sang et al., 1997; Mansion et al., 2005; Lihová et al., 2006).

However, for understanding relationships among closely related species within genera, further molecular data are often needed, and low-copy nuclear genes are increasingly being used to address these phylogenetic questions. One such region is the granule-bound starch synthase I gene (GBSSI, *waxy*), which has been shown to be useful in a variety of plant families including Poaceae (Mason-Gamer et al., 1998; Fortune et al., 2007, 2008), Malvaceae (Small, 2004), Boraginaceae (Moore and Jansen, 2006), and Proteaceae (Mast et al., 2005). In Solanaceae GBSSI appears to be single copy and phylogenetically informative among species within genera (Peralta and Spooner, 2001; Levin and Miller, 2005; Levin et al., 2005, 2006, 2007; Dillon et al., 2007).

The nuclear nitrate reductase gene (NIA) catalyzes the reduction of nitrate to nitrite, facilitating in plants the uptake of nitrogen from the soil. The gene is usually comprised of four exons and three introns (Fig. 1). Among Solanaceae NIA appears to be single copy (*Solanum lycopersicon*, Daniel-Vedele et al., 1989; *Solanum tuberosum*, Harris et al., 2000; *Petunia hybrida*, Salanoubat and Ha, 1993). However, there are two loci in *Nicotiana tabacum* due to its allopolyploid origin (Vaucheret et al., 1989). NIA has also been found to be at a single locus in Goodeniaceae (*Scaevola*, Howarth and Baum, 2002), Anacardiaceae (Yi et al., 2007, 2008) and Chenopodiaceae (*Spinacia oleracea*, Tamura et al., 1997), whereas there may be two loci in certain *Olea* (Oleaceae) species (Hamman-Khalifa et al., 2007), and at least two loci in Brassicaceae (*Arabidopsis thaliana*, Cheng et al., 1988) and Poaceae (*Oryza sativa*, Hamat et al., 1989). Zhou and Kleinhofs (1996) examined the

* Corresponding author. Fax: +1 413 542 7955.
E-mail address: rlevin@amherst.edu (R.A. Levin).

1055-7903/$ - see front matter © 2008 Elsevier Inc. All rights reserved.
doi:10.1016/j.ympev.2008.12.005

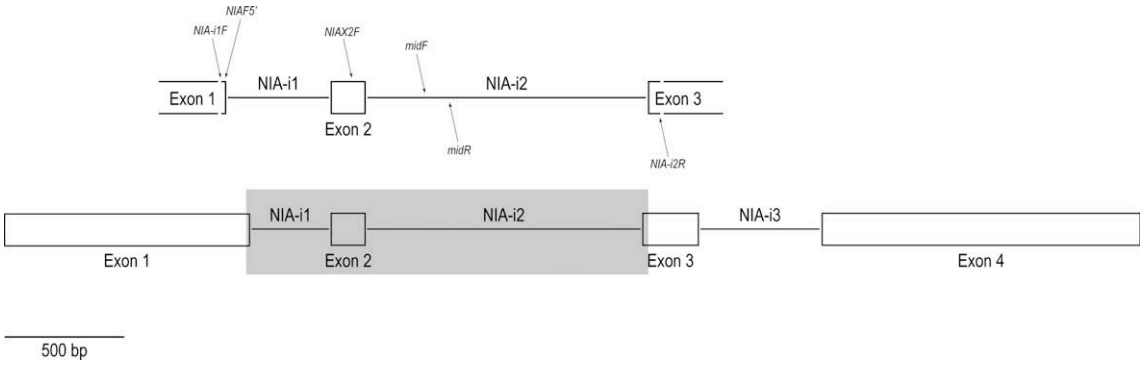

Fig. 1. Diagram of the region of the nitrate reductase gene that was sequenced and the primers that were used for PCR and/or sequencing (above). The structure of the NIA gene as determined for *Petunia hybrida* (Salanoubat and Ha, 1993) is shown below; shading indicates the part of the gene that was sequenced in the present study. Locations of the primers are as follows (reference sequence *L. andersonii* FJ444895): NIAF5′ = 1–20 bp (NIA-i1F was not sequenced, but this primer ends 29 bp before NIAF5′), NIAX2F = 535–552 bp, midF = 791–810 bp, midR = 873–893 bp, NIA-i2R = 1665–1684 bp.

molecular evolution of nitrate reductase sequences and suggested that NIA may be a useful region for evolutionary studies of relationships across distantly related plant lineages. Further, Howarth and Baum (2002) suggested that introns of this gene are useful for evolutionary studies of recently diverged species, and showed a high rate of evolution in comparison to three other nuclear gene regions in *Scaevola* (Howarth and Baum, 2005). Sequence data for nitrate reductase has also been shown to be informative for examining relationships within *Rhus* and *Pistacia* (Anacardiaceae; Yi et al., 2007, 2008) and among closely related wild potatoes (*Solanum* sect. *Petota*; Rodriguez and Spooner, 2004, 2005).

Lycium is a genus of ca. 85 species distributed mainly in southwestern North America, South America, and southern Africa (tribe Lycieae; Solanaceae). Although *Lycium* is by far the largest genus in tribe Lycieae, the tribe also includes the monotypic *Phrodus microphyllus* and small (3 spp.) genus *Grabowskia* (Hunziker, 2001). However, previous studies (Levin and Miller, 2005; Levin et al., 2007; Olmstead et al., 2008) have indicated that these genera, which are endemic to the Americas, may be nested within *Lycium*. Results suggest that *Lycium* species have diversified relatively recently, particularly in southern Africa and within certain groups of American taxa (Shak et al., 2006, Miller and Levin, unpublished data). Here we focus on the paraphyletic American *Lycium*, a large group of species within which the Old World *Lycium* species are nested (Levin and Miller, 2005; Levin et al., 2007). Specifically, our goals are to: (1) characterize the nitrate reductase region between exons 1 and 3 for *Lycium*, (2) determine whether this region is useful for phylogenetic inference, and (3) compare topologies across three nuclear (NIA, ITS, and GBSSI) and two chloroplast (*trnD–trnT* and *trnT–trnF*) regions.

2. Materials and methods

2.1. Taxon sampling

Thirty-three species were involved in this study, including 28 *Lycium* species, two *Grabowskia* species, and *P. microphyllus*. As Old World *Lycium* are strongly supported as monophyletic (Levin and Miller, 2005; Levin et al., 2007), and our focus here was on relationships among American species, we included only two representative Old World species (*L. arenicola* and *L. villosum*). Thus, we included 26 species of American *Lycium* (15 species from North America and 11 from South America), which is equivalent to approximately half of all New World taxa. In addition, two Chilean *Nolana* species (*N. coelestis* and *N. werdermannii*) were included as a monophyletic outgroup. The GBSSI data and all but three sequences for the chloroplast *trnT–trnF* data are from Levin and Miller (2005) and Levin et al. (2007). The majority of ITS sequences are from Miller (2002) and Yeung et al. (2005), with over one third (12 sequences) generated for the present study. All taxa, GenBank Accession Nos., and voucher information are in Table 1. Total genomic DNA was extracted from fresh or silica gel-dried leaf material following Miller (2002) and Levin et al. (2004).

2.2. Nitrate reductase (NIA)

Amplification of the 3′-end of exon 1 through the 5′-end of exon 3 (Fig. 1) was conducted using the primers NIA-i1F (5′-TCG GAA AGC TAT TAT CAT TAC AAG G-3′) and NIA-i2R (5′-CCA TGT CTC TCC TCC ATC CA-3′), designed by D. Spooner and F. Rodriguez (University of Wisconsin, Madison, Wisconsin USA). Alternatively, the forward primer NIAF5′ (5′-GCT GAA CTT GCT AAC GCT GA-3′) was used. Twenty-five microliter reactions contained 1× buffer, 3.0 mM $MgCl_2$, 0.20 mM dNTPs, 0.40 μM of each primer, 1× Qiagen Q-solution (Qiagen, Inc., Valencia, CA), 0.625 U of Taq polymerase, and 1 μL DNA. The thermal cycler program had an initial denaturation at 95 °C for 3 min; 40 cycles at 95 °C for 45 s, 56.5 °C for 2 min, 72 °C for 2 min; ending with an extension at 72 °C for 7 min. Occasionally, this region was PCR amplified in two separate pieces, using primer pairs NIAF5′ and midR (5′-CGT CTR AAA AGC AAT ACA TGA-3′) and midF [5′-TGG TTG AGA CAC CCC TCT GA-3′; occasionally NIAX2F (5′-ACT CAG CGA CCT TAC ACG-3′) was used instead] and NIA-i2R. Thermal cycler conditions for these amplifications had a 4 min denaturation at 94 °C; 40 cycles at 94 °C for 30 s, 52–54 °C for 1 min, 72 °C for 1 min; with an extension at 72 °C for 7 min. PCR products were cleaned using the QIAquick PCR purification Kit (Qiagen, Inc., Valencia, CA). Products were directly sequenced with the same primers as for amplification and internal primers midF (or occasionally NIAX2F) and midR (Fig. 1) using an ABI automated sequencer at the DNA Sequencing Facility of the Biotechnology Resource Center, Cornell University (Ithaca, NY, USA). Several directly sequenced PCR products yielded incomplete sequences; thus, 60% of all PCR amplicons (only those where the entire region from exon 1 to exon 3 was amplified in a single PCR) were cloned prior to sequencing. Cloning reduced sequencing problems given allelic length polymorphisms and also facilitated determination as to whether NIA is at a single locus in *Lycium*.

Cleaned amplification products were cloned using the pT7Blue Perfectly Blunt Cloning Kit (Novagen, EMD Chemicals, Inc., Madison, WI). Colonies were PCR amplified in 50 μL reactions using 10 μL of each colony prep and the vector primers R20 and U19 at final concentrations of 0.01 mM. Reactions contained 1.25 U of Taq polymerase, 1× buffer, 0.20 mM dNTPs, and 3.0 mM $MgCl_2$/$MgSO_4$. The thermal cycler program had an initial denaturation

Table 1
Taxa, collection localities, voucher information, and GenBank Accession Nos. for all sequences included in this study. GenBank Accession Nos. are listed in the following order: NIA, GBSSI, ITS, *trnT–trnL*, *trnL–trnF*, *trnD–trnT*. Voucher specimens are deposited in the following herbaria: ARIZ, University of Arizona; BLFU, University of the Free State; CORD, Universidad Nacional de Córdoba; MASS, University of Massachusetts; TAIC, Texas A&M University, Kingsville; UT, University of Utah.

Tribe Lycieae Hunz.

Grabowskia Schltdl.
G. boerhaviaefolia Schltdl.—Argentina, *Bernardello 894* (CORD); FJ444896, EF137751, FJ439754, DQ124431, DQ124554, FJ444876
G. obtusa Arnott—Argentina, *Bernardello 891* (CORD); FJ444897, DQ124498, FJ439755, DQ124433, DQ124556, FJ444877

Lycium L.
L. americanum Jacq.—Argentina, *Barboza 525* (CORD); FJ444898, DQ124502, DQ124619, DQ124438, DQ124561, FJ189630
L. andersonii A. Gray—Mexico, *Miller 97-12* (ARIZ); FJ444895, DQ124503, DQ124620, DQ124439, DQ124562, FJ189631
L. arenicola Miers—South Africa, *Venter 647* (BLFU); FJ444899, DQ124504, FJ439756, DQ124440, DQ124563, FJ444878
L. athium Bernardello—Argentina, *Miller* et al. *05-25* (MASS); FJ444900, EF137756, FJ439757, FJ439766, N/A, FJ444879
L. berlandieri Dunal—Arizona, USA, *Miller 01-1* (ARIZ); FJ444901, DQ124506, AF238989, DQ124444, DQ124567, FJ444880
L. brevipes Benth.—Mexico, *Miller 97-19* (ARIZ); FJ444902, DQ124508, DQ124621, DQ124446, DQ124569, FJ444881
L. californicum Nutt. ex A. Gray—Arizona, USA, *Miller & Levin 04-12* (MASS); FJ444903, DQ124510, DQ124648, DQ124448, DQ124571, FJ444882
L. carolinianum Walt.—Texas, USA, *Hempel 843* (TAIC); FJ444904, DQ124512, DQ124622, DQ124450, DQ124573, FJ444883
L. cestroides Schltdl.—Argentina, *Bernardello 878* (CORD); FJ444905, DQ124513, DQ124623, DQ124451, DQ124574, FJ189633
L. cooperi A. Gray—Arizona, USA, *Miller 97-1* (ARIZ); FJ444906, DQ124518, AF238984, DQ124456, DQ124579, FJ444884
L. cuneatum Dammer—Argentina, *Bernardello & Vesprini 897* (CORD); FJ444907, DQ124519, FJ439758, DQ124457, DQ124580, FJ189638
L. elongatum Miers—Argentina, *Bohs 2940* (UT); FJ444909, DQ124520, DQ124624, DQ124458, DQ124581, FJ444885
L. exsertum A. Gray—Arizona, USA, *Miller 95-1* (ARIZ); FJ444910, DQ124522, AF238994, DQ124460, DQ124583, N/A
L. fremontii A. Gray—Mexico, *Miller 01-4* (ARIZ); FJ444911, DQ124524, DQ124626, DQ124462, DQ124585, FJ444886
L. infaustum Miers—Argentina, *Bernardello 893* (CORD); FJ444912, DQ124529, DQ124627, DQ124467, DQ124590, FJ444887
L. leiospermum I. M. Johnst.—Mexico, *Miller & Levin 05-53* (MASS); FJ444913, EF137779, FJ439759, FJ439767, N/A, FJ444888
L. macrodon A. Gray—Arizona, USA, *Miller 97-21* (ARIZ); FJ444914, DQ124530, DQ124628, DQ124469, DQ124592, FJ444889
L. nodosum Miers—Argentina, *Barboza 515* (CORD); FJ444915, EF137783, DQ124630, DQ124471, DQ124594, FJ189641
L. pallidum Miers—Arizona, USA, *Miller 97-20* (ARIZ); FJ444916, DQ124534, DQ124631, DQ124473, DQ124596, FJ444890
L. parishii A. Gray—Arizona, USA, *Miller 97-22* (ARIZ); FJ444917, DQ124535, DQ124632, DQ124474, DQ124597, FJ189643
L. puberulum A. Gray—Texas, USA, *Levin 97-6* (ARIZ); FJ444918, DQ124537, AF238985, DQ124476, DQ124599, FJ189644
L. rachidocladum Dunal—Chile, *Miller* et al. *04-82* (MASS); FJ444919, EF137787, FJ439760, FJ189731, N/A, FJ189646
L. shockleyi A. Gray—Nevada, USA, *Miller 98-1* (ARIZ); FJ444920, DQ124540, AF238987, DQ124481, DQ124604, FJ189648
L. tenuispinosum Miers—Argentina, *Bernardello 892* (CORD); FJ444921, EF137795, DQ124633, DQ124485, DQ124608, FJ444891
L. texanum Correll—Texas, USA, no voucher; FJ444922, DQ124545, FJ439761, DQ124487, DQ124610, FJ444892
L. torreyi A. Gray—Arizona, USA, *Miller & Levin 04-21* (ARIZ); FJ444923, DQ124546, DQ124634, DQ124488, DQ124611, FJ444893
L. villosum Schinz—South Africa, *Venter 652* (BLFU); FJ444924, DQ124547, FJ439762, DQ124489, DQ124612, N/A
L. vimineum Miers—Argentina, *Bernardello & Vesprini 896* (CORD); FJ444925, EF137796, DQ124635, DQ124491, DQ124614, FJ444894

Phrodus Miers
Phrodus microphyllus (Miers) Miers—Chile, *Miller* et al. *04-92* (MASS); FJ444928, EF137801, FJ439765, DQ124495, N/A, FJ189651

Taxa outside of Lycieae
Nolana coelestis Lindl.—Chile, *Miller* et al. *04-98* (MASS); FJ444926, EF137800, FJ439763, DQ124494, DQ124617, FJ189629
Nolana werdermannii I. M. Johnst.—Chile, *Miller* et al. *04-77* (MASS); FJ444927, EF137799, FJ439764, DQ124493, DQ124616, FJ189628

at 94 °C for 5 min; 6 cycles at 94 °C for 1 min, 55–53 °C (decreasing 1 degree every 2 cycles) for 1 min, 72 °C for 2 min; 30 cycles at 94 °C for 1 min, 52 °C for 1 min, 72 °C for 2 min; ending with an extension at 72 °C for 5 min. PCR products were cleaned and sequenced as above. Vector primers were used for sequencing cloned products, with internal primers midF (or occasionally NIAX2F) and midR also used to yield a complete sequence with considerable sequence overlap for this ca. 2 kb region.

2.3. Granule-bound starch synthase (GBSSI)

All GBSSI data included in the present study were previously generated by Levin and Miller (2005) and Levin et al. (2007). The GBSSI region that was amplified and sequenced encompassed exons 2 through 10 (ca. 1800 bp) or occasionally the 3′-end of exon 3 through the 5′-end of exon 8 (ca. 900 bp).

2.4. Internal transcribed spacer (ITS)

Amplification of the internal transcribed spacer region of nuclear ribosomal DNA, composed of ITS1, the 5.8S gene, and ITS2 (Baldwin, 1992; Baldwin et al., 1995) was done following Miller (2002) or using primers ITS5HP (5′-GGA AGG AGA AGT CGT AAC AAG G-3′; Hershkovitz and Zimmer, 1996) or ITSleu1 (5′-GTC CAC TGA ACC TTA TCA TTT AG-3′; Bohs and Olmstead, 2001) and ITS4 (5′-TCC TCC GCT TAT TGA TAT GC-3′; White et al., 1990). Twenty-five microliter reactions contained 1× buffer, 2.5–3.0 mM $MgCl_2$, 0.20 mM dNTPs, 0.40 μM of each primer, 1× Qiagen Q-solution (Qiagen, Inc., Valencia, CA), 0.625 U of Taq polymerase, and 1 μL DNA. Thermal cycler conditions included an initial denaturation at 97 °C for 2 min; 30 cycles at 97 °C for 1 min, 50 °C for 1 min, 72 °C for 1 min; ending with an extension at 72 °C for 7 min. Alternatively, a touchdown procedure was conducted using an initial denaturation at 94 °C for 3 min; 8 cycles at 94 °C for 30 s, 58–52 °C (decreasing 2° every 2 cycles) for 1 min, 72 °C for 1 min; 25 cycles at 94 °C for 30 s, 50 °C for 1 min, 72 °C for 1 min; ending with an extension at 72 °C for 10 min. PCR products were cleaned and sequenced as above, using ITS5HP and ITS4. For a few taxa, sequencing was also done with one or two internal ITS primers (see primer sequences in White et al., 1990).

2.5. Chloroplast trnD (GUC)–trnT (GGU)

Amplification of the intergenic region between the *trnD* and *trnT* genes, including the *trnY* and *trnE* genes, used primers trnD (5′-ACC AAT TGA ACT ACA ATC CC-3′) and trnT (5′-CTA CCA CTG AGT TAA AAG GG-3′) of Demesure et al. (1995). Fifty microliter reactions contained 1× buffer, 3.0 mM $MgCl_2$, 0.20 mM dNTPs, 0.36 μM of each primer, 8.8 ng BSA, 1.25 U of Taq polymerase, and 1–2 μL DNA. Thermal cycler conditions included initial denaturation at 80 °C for 5 min; 40 cycles at 94 °C for 45 s, 50 °C for 30 s, 72 °C for 1 min, ending with an extension at 72 °C for 7 min. Alternatively, a touchdown procedure was used with an initial denaturation at 94 °C for 3 min; 20 cycles at 94 °C for 45 s, 58–

49 °C (decreasing one degree every two cycles) for 30 s, 72 °C for 1 min 30 s; 20 cycles at 94 °C for 45 s, 48 °C for 30 s, 72 °C for 1 min 30 s; with a final extension at 72 °C for 7 min. PCR products were cleaned and sequenced as above, using the same primers as for amplification.

2.6. Chloroplast trnT (UGU)–trnF (GAA)

We amplified the chloroplast region between the *trnT* and *trnF* genes, including the intergenic spacer between *trnT* and *trnL*, the *trnL* intron, the *trnL* 3′-exon (we sequenced only a few bases of the *trnL* 5′-exon), and the intergenic spacer between the *trnL* 3′-exon and *trnF*. For ease of amplification, this piece was amplified with two separate PCRs; one reaction used primers a and b, and the other reaction used primers c and f of Taberlet et al. (1991). This set of primers resulted in a sequence gap of 23 bp at the 3′-end of the *trnL* 5′-exon, and 26 bp at the 5′-end of the *trnL* intron. For both sets of primers, 50 μL reactions contained 1× buffer, 2.0 mM $MgCl_2$, 0.20 mM dNTPs, 0.36 μM of each primer, 8.8 ng BSA, 1.25 U of Taq polymerase, and 1–2 μL DNA. Thermal cycler conditions with primers a and b was a touchdown procedure with an initial denaturing at 94 °C for 4 min; 12 cycles at 94 °C for 1 min, 54–49 °C (decreasing one degree every two cycles) for 1 min, 72 °C for 1 min 30 s; 28 cycles at 94 °C for 1 min, 48 °C for 1 min, 72 °C for 1 min 30 s; ending with an extension at 72 °C for 7 min. The following PCR conditions with primers c and f were used: 94 °C for 4 min; 30–40 cycles at 94 °C for 45–60 s, 52–54 °C for 1 min, 72 °C for 1 min 30 s; 72 °C for 7 min. PCR products were cleaned and sequenced as above, using the same primers as for amplification.

2.7. Sequence alignment and phylogenetic analyses

For each region, sequences from all primers were edited and aligned using either Sequencher v. 4.6 (Gene Codes Corp., 1991–2006) or Autoassembler DNA Sequence Assembly Software v. 1.4.0 (Applied Biosystems, 1989–95). All sequences (usually from two to four sequences depending on the region) for each genomic accession were aligned and combined into a single consensus sequence. For clones, all sequences from a given colony were aligned and combined as a single consensus sequence. Multiple alignment of species (or individual colony) sequences was done manually in SeAl (Rambaut, 2002) and MacClade 4.0 (Maddison and Maddison, 2000).

We used a maximum likelihood (ML) algorithm to analyze each of the four data sets (the chloroplast data were combined as a single data set), each of which included 33 taxa. A maximum parsimony analysis (not shown) of the NIA data using multiple clones for four species, revealed monophyly; thus, the sequence of a single clone per species was included in the analyses presented here. For each data set, ML model parameters were determined using the Akaike information criterion in Modeltest v. 3.7 (Posada and Crandall, 1998). The best-fit model for each data set was used in separate ML analyses in PAUP* (Swofford, 2002). ML settings in PAUP* included the heuristic search option, all most-parsimonious trees from a parsimony analysis of the data set (100 random-addition sequence replicates, MulTrees disabled) as the starting trees (note that not all of the most-parsimonious starting trees are actually used by PAUP*, depending on their ML scores), tree bisection reconnection (TBR) branch-swapping, and the MulTrees option in effect. *Nolana coelestis* and *N. werdermannii* were defined as the monophyletic outgroup, as *Nolana* has been shown to be sister to Lycieae in previous analyses (Olmstead et al., 1999; Levin and Miller, 2005). For each data set all most-likely trees were combined in a strict consensus tree. An ML non-parametric bootstrap (BS) analysis was also conducted for each data set, using the same model parameters as in the original ML analysis, and 100 full heuristic bootstrap replicates, each with 10 random-addition sequence replicates, TBR branch-swapping, and the MulTrees option in effect. Maximum likelihood bootstrap analyses were conducted using PAUP* version 4.0b10 for UNIX (Swofford, 2002) on the Condor (Condor Project, 2005) computer cluster at Amherst College. Bootstrap replicates were parsed for processing using RepMaker (Wilgenbusch, 2003).

2.8. Tests of data congruence

Congruence among data sets was examined using incongruence length difference tests (ILD; Farris et al., 1994, 1995) as implemented by the partition homogeneity test in PAUP*. Two ILD tests were conducted; one compared all four data sets simultaneously, and the other compared only the two low-copy nuclear regions NIA and GBSSI. For each test, 500 ILD replicates were conducted, with 100 random-addition sequence replicates and TBR branch-swapping; the MulTrees option was not in effect. Further, we tested support for topological differences among gene trees using one-tailed Shimodaira–Hasegawa (SH) tests (Shimodaira and Hasegawa, 1999) as conducted in PAUP*. A RELL test distribution was used, with 1000 bootstrap replicates. Specifically, we examined the effect of constraining data sets to show relationships that were well supported by one or more different data sets. We used the same ML parameters as inferred from ModelTest (see above) and compared the likelihoods of ML topologies inferred from unconstrained versus constrained analyses.

3. Results

3.1. Nitrate reductase

Nitrate reductase sequences across the 33 taxa ranged in length from 1610 to 2096 bp [this does not include the six taxa (*L. cooperi*, *L. macrodon*, *L. shockleyi*, *Grabowskia boerhaviaefolia*, *G. obtusa*, and *Nolana coelestis*) for which only partial sequences were obtained], with an aligned length of 2836 bp. Intron 1 ranged in length from 345 to 523 bp, whereas intron 2 was ca. 3 times longer (1042–1354 bp). The number of parsimony-informative (PI) characters was 246 (8.7%), with a rescaled consistency index of 0.60 (Table 2). Likelihood settings from the best-fit model of GTR + I + G (nucleotide frequencies of A = 0.3349, C = 0.1587, G = 0.1551, and T = 0.3513; a substitution rate matrix of A–C: 0.8553, A–G: 3.2166, A–T: 0.5475, C–G: 0.9068, C–T: 2.3405, and G–T: 1; proportion of invariable sites = 0.1877; and gamma distributed variable sites with a shape parameter of 0.8229) were used in an ML analysis in PAUP* (Swofford, 2002). The ML analysis yielded one tree with $-\ln L = 9281.23321$.

The ML bootstrap consensus topology (Fig. 2A) supports the placement of *P. microphyllus* in tribe Lycieae (BS = 100) and sister to *Grabowskia* plus *Lycium* (BS = 62). Within the rest of the tribe, there is a well supported clade of *Grabowskia* plus five species of North American *Lycium* (BS = 100); this lineage is also supported

Table 2
Comparison of the 33 taxa data sets for the three nuclear genomic regions (NIA, GBSSI, ITS) and the combined chloroplast regions (CP: *trnT–trnF* and *trnD–trnT*). PI, parsimony-informative; CI, consistency index.

	NIA	GBSSI	ITS	CP
Total number of characters	2836 bp	1888 bp	692 bp	2785 bp
Number variable characters	638	227	110	64
Number PI characters (%)	246 (8.7%)	107 (5.7%)	66 (9.5%)	42 (1.5%)
CI	0.79	0.95	0.61	0.99
Rescaled CI	0.60	0.91	0.43	0.98

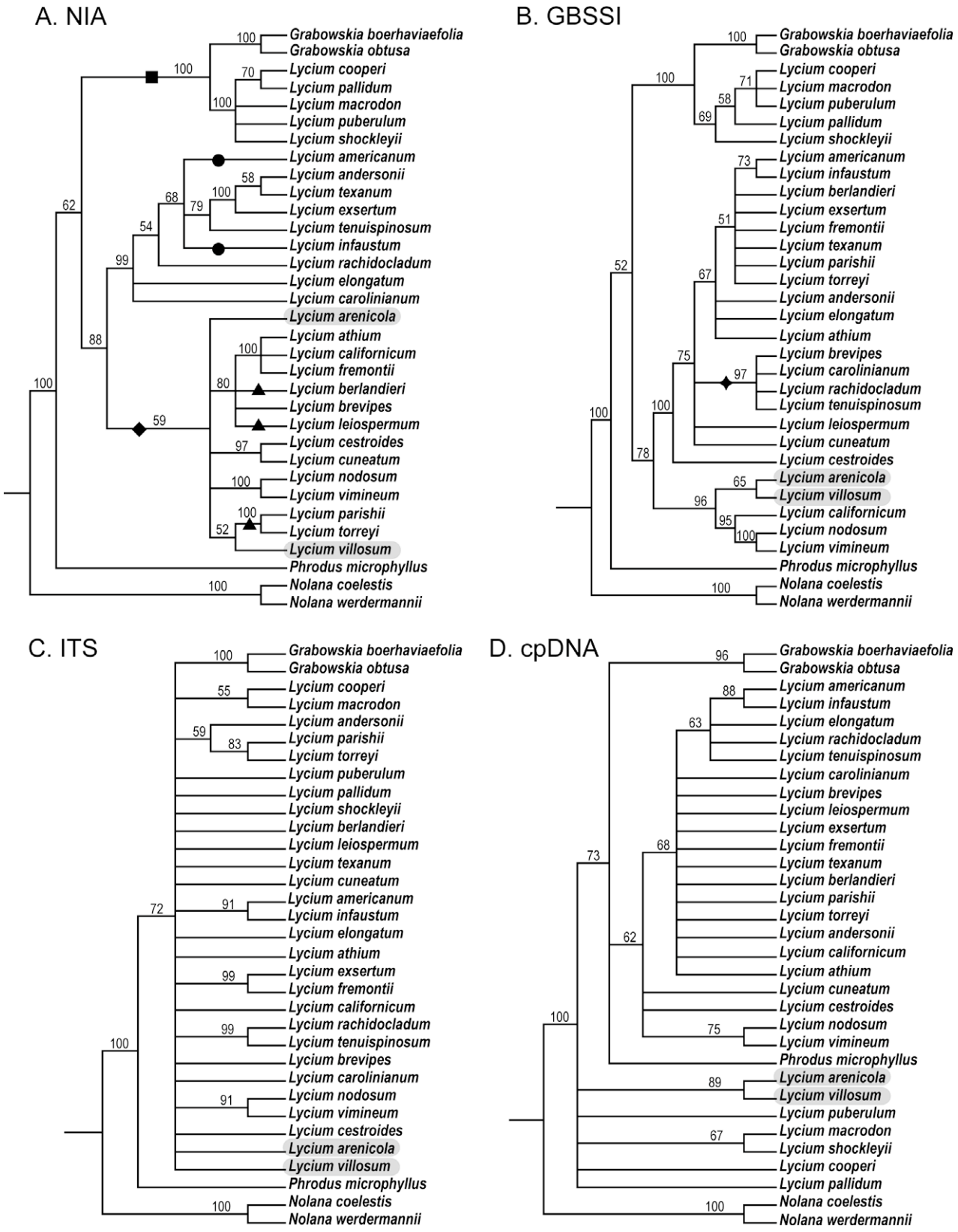

Fig. 2. The maximum likelihood BS consensus topology for the nuclear NIA (A), nuclear GBSSI (B), nuclear ribosomal ITS (C), and the combined chloroplast DNA (D) data sets. The two Old World taxa are shaded; all other taxa are New World. Informative indels noted in the text are indicated with symbols; each symbol represents one shared indel except for circle = two indels and triangle = three indels (note that *L. leiospermum* has only two of the three indels).

by an 18 bp deletion in intron 1. Outside of this *Grabowskia* plus *Lycium* clade is a well supported group (BS = 88) that includes all other sampled *Lycium* species. Although deep relationships within this BS = 88 group have limited support, there are a number of well supported lineages. These include a group of nine American taxa (BS = 99), within which there is a clade of *L. andersonii*, *L. texanum*,

and *L. exsertum* (BS = 100). The other clade of *Lycium* (BS = 59) has a large (ca. 240 bp) insertion in intron 2 (although the sequence is not identical across these taxa). Among these taxa there is a clade containing *L. athium*, *L. californicum*, and *L. fremontii* (BS = 100), and several unplaced sister taxa pairs, including *L. cestroides* and *L. cuneatum* (BS = 97), *L. nodosum* and *L. vimineum* (BS = 100), and *L. parishii* and *L. torreyi* (BS = 100). Indels provide evidence for additional relationships that are not revealed by the ML bootstrap consensus topology (Fig. 2A). A sister taxon relationship of *L. infaustum* and *L. americanum* is supported by two indels (one 7 bp insertion and one 9 bp deletion) in intron 2. In addition, although there is moderate BS support (BS = 80) for a different relationship (see clade including *L. berlandieri* in Fig. 2A), three indels (two in intron 1 and one in intron 2), ranging in size from 2 to 5 bp, support the relationship of *L. berlandieri* with *L. parishii* and *L. torreyi*; two of these indels also support the affinity of *L. leiospermum* to these three taxa.

3.2. Granule-bound starch synthase

Sequences across all 33 taxa ranged from 869 to 1842 (taxa were amplified with a range of primers, resulting in sequences of different lengths; see Levin and Miller, 2005), with an aligned length of 1888 bp. The number of PI characters was 107 (5.7%), with a rescaled consistency index of 0.91 (Table 2). Likelihood settings from the best-fit model of TVM + G (nucleotide frequencies of A = 0.2693, C = 0.1923, G = 0.2030, and T = 0.3354; a substitution rate matrix of A–C: 0.9620, A–G: 2.4896, A–T: 0.5962, C–G: 1.1865, C–T: 2.4896, and G–T: 1; proportion of invariable sites = 0; and gamma distributed variable sites with a shape parameter of 1.0388) were used in an ML analysis in PAUP* (Swofford, 2002). The ML analysis yielded 8 trees with $-\ln L = 4339.97138$. The ML BS consensus topology (Fig. 2B) supports (albeit weakly) the placement of *P. microphyllus* as branching earliest within tribe Lycieae. Further, within the rest of the tribe there is a well supported clade of *Grabowskia* plus five species of North American *Lycium* (BS = 100); this lineage is outside of the rest of *Lycium*. The remaining species are included in two large clades, one (BS = 96) of which includes the two Old World species (*L. arenicola* and *L. villosum*) plus a well supported American clade (BS = 95) of *Lycium californicum*, *L. nodosum*, and *L. vimineum*. The other large clade (BS = 100) includes a strongly supported lineage of *L. brevipes*, *L. carolinianum*, *L. rachidocladum*, and *L. tenuispinosum* (BS = 97), which is also supported by a large 59–62 bp insertion in intron 7.

3.3. Internal transcribed spacer

Sequences across all 33 taxa ranged from 659 to 669, with an aligned length of 692 bp. The number of PI characters was 66 (9.5%), with a rescaled consistency index of 0.43 (Table 2). Likelihood settings from the best-fit model of HKY + I + G (nucleotide frequencies of A = 0.1680, C = 0.3455, G = 0.3041, and T = 0.1824; a transition to transversion ratio of 3.2355; proportion of invariable sites = 0.6794; and gamma distributed variable sites with a shape parameter of 0.6386) were used. The ML analysis yielded 2 trees with $-\ln L = 2149.76051$. The ML BS consensus topology (Fig. 2C) shows *P. microphyllus* as sister to the rest of tribe Lycieae (BS = 72). Within the rest of the tribe there is limited resolution; however, there is strong support for a number of sister group relationships including the two *Grabowskia* species (BS = 100), *L. americanum* + *L. infaustum* (BS = 91), *L. parishii* + *L. torreyi* (BS = 83), *L. exsertum* + *L. fremontii* (BS = 99), *L. nodosum* + *L. vimineum* (BS = 91), and *L. rachidocladum* + *L. tenuispinosum* (BS = 99).

3.4. Chloroplast trnD–trnT and trnT–trnF

Sequences across all 33 taxa ranged from 1106 to 1158 bp for *trnD–trnT* and from 1592 to 1611 bp for *trnT–trnF*, with a total combined aligned length of 2785 bp. The number of PI characters was 42 (1.5%), with a rescaled consistency index of 0.98 (Table 2). Likelihood settings from the best-fit model of TVM + I (nucleotide frequencies of A = 0.3450, C = 0.1629, G = 0.1673, and T = 0.3248; a substitution rate matrix of A–C: 0.5046, A–G: 0.6340, A–T: 0.0647, C–G: 0.4704, C–T: 0.6340, and G–T: 1; proportion of invariable sites = 0.7240; and equal rates of variable sites) were used. The ML analysis yielded 5 trees with $-\ln L = 4251.01042$. There are few supported relationships, although *P. microphyllus* is weakly supported as nested within tribe Lycieae (BS = 73). There is also support for the sister group relationships of the two *Grabowskia* species (BS = 96), *L. americanum* + *L. infaustum* (BS = 88), *L. nodosum* + *L. vimineum* (BS = 75), and the two Old World taxa, *L. arenicola* + *L. villosum* (BS = 89).

4. Discussion

4.1. Phylogenetic utility of NIA

Of the four gene regions examined in this study, NIA has more PI characters than any of the other three regions, with over twice the number of PI characters as GBSSI (Table 2). Further, NIA and ITS clearly have the greatest percentage of PI characters. In species-level studies of *Pistacia* and *Rhus* (Anacardiaceae; Yi et al., 2007, 2008), with similar numbers of taxa as included here, NIA was also found to have the highest number of PI characters. In *Pistacia*, the NIA sequence data had 212 PI characters, in comparison to 173 for ITS. Similarly, in *Rhus* the NIA data had 150 PI characters, a bit higher than the 142 PI characters in ITS. Further, for a much smaller data set of 14 *Betula* species (Li et al., 2007), there were 66 PI characters in the NIA data, compared to 26 PI characters for ITS. In the present study, the number of PI characters in the NIA data (246) was considerably higher than the ITS data (66) (Table 2). This disparity with the findings of Yi et al. (2007, 2008) is likely due in part to differences in the region of the NIA gene that was sequenced. Yi et al. (2007, 2008) and Li et al. (2007) sequenced the third NIA intron, which is less than half as long as the region (including introns 1 and 2) sequenced in the present study (Fig. 1).

Although greater numbers of PI characters are desirable for resolution at fine taxonomic levels, more variable regions may result in increased homoplasy. High levels of homoplasy are often associated with the ITS region, as noted by other researchers for various angiosperm groups (e.g., Buckler et al., 1997; Oh and Potter, 2003; Levin et al., 2005). Álvarez and Wendel (2003) discussed problems with ITS and reasons for the high homoplasy of this region, including incomplete concerted evolution. Our data for ITS also show considerable homoplasy; ITS has both the highest percentage of PI characters and also the lowest consistency index (CI). Álvarez and Wendel (2003) surveyed the literature and conducted pairwise comparisons of the CI for ITS and all other gene regions sequenced for the same taxon set. They found that, on average, ITS had a 14.6% lower CI than other gene regions. Our data concur, with an 18% decrease in CI for ITS when compared to NIA, a 34% decrease in CI for ITS when compared to GBSSI, and a 38% decrease in CI for ITS when compared to the cpDNA data (average of 30% lower CI for ITS; Table 2). Álvarez and Wendel (2003) suggested that high levels of homoplasy can increase the possibility of incorrect inference. We agree, but emphasize that highly variable regions such as ITS may also be expected to yield poor resolution, given conflicting (and likely homoplastic) signal in the data (e.g., Fig. 2C).

Although our data show a higher CI for NIA (0.79) than for ITS (0.61), the level of homoplasy for both ITS and NIA is by far the highest across the four genomic regions examined in this study (Table 2). These two regions also have the greatest percentage of PI characters, which has been shown experimentally to be negatively correlated with CI (Hauser and Boyajian, 1997). It is not surprising that an increase in the percentage of PI characters also decreases CI, especially for large taxon sets, as more parsimony-informative change may result in a concomitant increase in homoplastic change. Thus, although NIA and ITS have roughly the same percentage of PI characters, the extent of homoplasy in NIA is comparatively low given its length and number of PI characters.

NIA was previously suggested as useful for phylogenetic inference in *Scaevola* (Howarth and Baum, 2002, 2005), especially given the high variation in the large intronic regions (i.e., 3.2% variable sites and 2.7% PI sites in NIA intron 3; Howarth and Baum, 2005). However, the magnitude and frequency of indels that are not necessarily homologous may affect the utility of NIA for phylogeny reconstruction. Among the 31 taxa in tribe Lycieae sampled here, there are 19 non-autapomorphic indels (2–242 bp) and 29 autapomorphic indels (2–303 bp). The longest insertion, occurring within intron 2, is shared by half of the *Lycium* species included in this study (clade with BS = 59; Fig. 2A). However, the sequence of this insertion is variable across taxa both in terms of base changes as well as indels (e.g., multiple taxa have autapomorphic or shared insertions or deletions within this large insertion). An analysis excluding this region revealed similar relationships, suggesting that homoplasy in this region does not affect inference of evolutionary relationships. The frequency of indels in introns 1 and 2 of NIA is remarkable, especially given that the indels in the other gene regions (ITS, GBSSI, and cp) combined are only roughly a third of the number of indels in NIA (16 for ITS, GBSSI, and cp regions combined versus 48 in NIA). It is notable that a considerable number of gaps were also required for alignment of the NIA intron 3 in *Rhus* and outgroups, with 38 informative gaps out of 67 total gaps (Yi et al., 2007).

The high number of indels for NIA may be due to a mixture of paralogous and orthologous NIA copies. Hamman-Khalifa et al. (2007) suggest that a duplication of NIA occurred in the ancestor of the domesticated olive (*Olea europaea*), and a second NIA type was observed in one species of *Betula* (Li et al., 2007). It is possible that there are multiple NIA loci in tribe Lycieae, although close relatives (except *Nicotiana tabacum*) have only a single locus (Daniel-Vedele et al., 1989; Salanoubat and Ha, 1993; Harris et al., 2000). Phylogenetic data also do not suggest more than one locus, given that when multiple colonies were included, they were monophyletic. Further, although 60% of NIA amplicons were cloned prior to sequencing, the rest were directly sequenced and yielded clean sequences, suggesting a single locus (although there could be preferential PCR amplification).

4.2. Data incongruence

All four data sets are incongruent (ILD test; $P < 0.01$). In examining individual data sets, the topologies inferred with ITS and cp data were fairly unresolved, resulting in minimal topological conflict with other data sets (Fig. 2). Topological incongruence between the nuclear and cpDNA data is concentrated on the placement of *Grabowskia* and *Phrodus* (see below). Between ITS and the other three data sets, the only supported discordance is the relationship of *L. exsertum* relative to *L. fremontii* (Fig. 2). The nuclear ITS data strongly support *L. exsertum* as sister to *L. fremontii* (Fig. 2C). This sister relationship is not supported, although it is also not refuted, by the GBSSI and cpDNA topologies. However, the BS consensus topology inferred from NIA data (Fig. 2A) conflicts with ITS, placing *Lycium exsertum* sister to *L. andersonii* + *L. texanum* (BS = 100) and *L. fremontii* in a well supported clade with *L. athium* and *L. californicum* (BS = 100). When the NIA data were constrained to place *L. exsertum* as sister to *L. fremontii*, those constrained topologies had a significantly lower likelihood (SH test, $P < 0.001$). It may be that multiple copies of ITS and NIA are the cause of the discordance here, as *L. exsertum* and *L. fremontii* are polyploid (some accessions of *L. californicum* are also polyploid, see below; Chiang-Cabrera, 1981; Yeung et al., 2005). Multiple gene copies in putative hybrids have certainly been reported in other polyploid taxa (nrITS: Obbard et al., 2006; GBSSI: Fortune et al., 2007; Mason-Gamer, 2008). Cloning ITS and sequencing more clones of NIA for these three taxa would be necessary to determine whether multiple copies are indeed present, resulting in topological incongruence among gene regions. A southern blot analysis could also indicate the presence (or absence) of multiple copies.

The topologies based on GBSSI and NIA are the most resolved, resulting in more apparent incongruence; these two data sets are also statistically incongruent (ILD test: $P < 0.01$). A main source of incongruence is the placement of *L. californicum* with *L. nodosum* and *L. vimineum* in the topology inferred from GBSSI (Fig. 2B, BS = 95) versus the placement of *L. californicum* with *L. athium* and *L. fremontii* in the NIA topology (Fig. 2A, BS = 100). The phylogenetic affinities of *L. californicum* have been problematic in the past as well, with gene regions differing in their placement of this species (Miller, 2002; Levin and Miller, 2005). In Miller (2002) nuclear ITS sequence data suggested that *L. californicum* is within a well supported clade with *L. exsertum* and *L. fremontii*, a relationship that is not inconsistent with the ITS topology in the present study. However, we note that the *L. californicum* accessions used in these two studies are from different populations that also differ in ploidy level, with a diploid accession used in Miller (2002) and a tetraploid accession used in the present study. In a phylogenetic analysis based on chloroplast sequence data, Levin and Miller (2005) found that *L. californicum* was in a clade of nine North American species including *L. exsertum*; by contrast, in the same Levin and Miller (2005) paper, *L. californicum* was strongly placed with *L. vimineum* and *L. nodosum* using GBSSI data (as in the present study, Fig. 2B). The disagreement in the phylogenetic placement of *L. californicum* among these three nuclear regions, rather than simply conflicting placement between nrDNA and cpDNA data, suggests that incomplete lineage sorting and/or introgression is the best explanation for this incongruence.

Certainly the biogeography of *L. californicum* is more consistent with the phylogenetic placement of ITS, cp data, and perhaps NIA (but see geographic location of *L. athium*, below). *Lycium fremontii* and *L. exsertum* occur in close proximity to *L. californicum* in southwestern North America, whereas *L. vimineum*, *L. nodosum*, and *L. athium* all occur solely in South America. As no *Lycium* species occur in central America, these groups of taxa are well separated geographically. In terms of morphological affinities, there is some support for the NIA placement of *L. californicum*, given that *L. californicum*, *L. athium*, *L. minimum*, and *L. ameghinoi* (the later two species were not included in the present study) share a distinctive fruit type, with two hardened seeds surrounded by a fleshy mesocarp. The phylogenetic position of *L. californicum* is especially interesting, given that this species has a variable sexual system across its range, with tetraploid gender dimorphic populations and diploid hermaphroditic populations (Yeung et al., 2005). Further studies in progress (Levin et al., unpublished data), with sampling from multiple individuals across multiple populations, should help clarify relationships within this species and among this species and its likely *Lycium* relatives.

Another major conflict between the NIA and GBSSI topologies is the relationship among *L. brevipes*, *L. carolinianum*, *L. rachidocladum*, and *L. tenuispinosum*. Specifically, the topology inferred from GBSSI

data shows a well supported clade of *L. brevipes*, *L. carolinianum*, *L. rachidocladum*, and *L. tenuispinosum* [these taxa also share a 62 (59 bp in *L. brevipes*) bp insertion in GBSSI; Fig. 2B, BS = 97]. However, the NIA data conflict with this relationship, particularly in the placement of *L. brevipes* outside of the well supported clade (Fig. 2A, BS = 99) that includes the other three species. It is possible that this incongruence is due to incomplete lineage sorting, given that *L. brevipes* has the ca. 240 bp insertion in intron 2. When the NIA data are constrained to include *L. brevipes* in a clade with the other three taxa (as supported by the GBSSI data), the topology is significantly less likely (SH test, $P < 0.001$). This group of taxa is of interest biogeographically, given that *L. brevipes* is restricted to North America (California and Mexico), *L. rachidocladum* and *L. tenuispinosum* are South American (Argentina and Chile), and *L. carolinianum* is unique among *Lycium* in occurring in both the Northern hemisphere [coastal North America (mainly Florida, Texas, and Mexico), Hawaii, and a few Japanese islands] and the Southern hemisphere (Easter Island).

Rodriguez and Spooner (2005) investigated the utility of NIA among wild potato species. As in the present study, they report incongruence between NIA and GBSSI data sets. In particular, Rodriguez and Spooner (2005) suggest that intron 2 is the source of the incongruence in *Solanum*, given that when intron 2 is included in analyses of NIA, there is incongruence with both cp restriction site data and GBSSI sequence data. However, when intron 2 was excluded, the phylogeny inferred from nitrate reductase sequence data concurs with the topologies inferred from the cp restriction site data and GBSSI sequence data (Rodriguez and Spooner, 2005). Thus, to explore the effects of intron 2 in the present study, we excluded this intron prior to additional analyses. Comparison of topologies suggests that incongruence was reduced; however, resolution was also greatly reduced, given that considerably fewer characters were included in the analysis without intron 2. When intron 2 was excluded, the bootstrap consensus topology resolved 13 nodes with bootstrap values >50% within tribe Lycieae, whereas the topology including intron 2 (Fig. 2A) resolved 19 nodes with bootstrap values >50%. Much of the resolution lost was at internal nodes, rather than at the tips of the phylogeny. Further, despite the decrease in resolution, the discordant clade of *L. fremontii*, *L. athium*, and *L. californicum* remained.

Previous studies (Yi et al., 2007, 2008) using sequence data from NIA intron 3 (not included in the present study) have also reported incongruence between NIA and other nuclear and chloroplast data. Incongruence between ITS and NIA in *Pistacia*, with the ITS data concurring with cpDNA data, is thought to be due to incomplete lineage sorting or hybridization (Yi et al., 2008). In *Rhus* there was a slight conflict between NIA and ITS data, with the majority of incongruence between the nrDNA and cpDNA data; thus, hybridization is thought to be the primary explanation for incongruence (Yi et al., 2007). In *Betula* (Li et al., 2007) there was limited topological incongruence between ITS and NIA intron 3, given low resolution in the ITS topologies.

4.3. Phylogenetic relationships within tribe Lycieae

Grabowskia species have been consistently placed with five North American species of *Lycium* (see Fig. 2A and B), which are morphologically quite similar to *Grabowskia* (Levin and Miller, 2005; Levin et al., 2007). The topologies inferred from nuclear data in the present study are in accord with these findings; however, the cpDNA data do not group *Grabowskia* species with these five North American species, although *Grabowskia* species are nested within *Lycium* (Fig. 2D). When these cpDNA data are constrained to place *Grabowskia* as sister to these five *Lycium* species, the unconstrained and constrained topologies are equally as likely, although marginally so (SH test, $P = 0.061$). Results of Levin and Miller (2005), using a larger sample of taxa but only cp *trnT–trnF* data, are consistent with either placement of the *Grabowskia* clade (i.e., with or without the five North American *Lycium* species). Further, a recent family-wide study of Solanaceae (Olmstead et al., 2008) based on cp *trnL–trnF* and *ndhF* data is congruent with relationships inferred here from *trnT–trnF* (including *trnL–trnF*) and *trnD–trnT* sequence data (Fig. 2D).

The phylogenetic affinity of *Phrodus* in Levin and Miller (2005) was equivocal with GBSSI and cp *trnT–trnF* data, due to a polytomy at the base of Lycieae. However, Levin et al. (2007), which included GBSSI data from a much larger sampling of taxa, suggested that *P. microphyllus* was sister to the rest of tribe Lycieae, although this conclusion did not have strong support. In the present study, all three nuclear data sets suggest that *P. microphyllus* is sister to the remaining members of tribe Lycieae. Although bootstrap support for this result is limited (bootstrap values range from 52 to 72; Fig. 2A–C), this result is bolstered by its presence in three independent data sets. By contrast, the combined cp *trnT–trnF* and *trnD–trnT* data suggest that *P. microphyllus* is nested within *Lycium* + *Grabowskia* (BS = 73), and results of Olmstead et al. (2008) concur with this relationship. When the cpDNA data are constrained to place *P. microphyllus* at the base of tribe Lycieae, the unconstrained and constrained topologies are equally as likely, although only marginally so (SH test, $P = 0.063$). Furthermore, data from a much larger cpDNA data set including nine spacer regions (Miller et al., in press), strongly suggest that *Phrodus* is nested well within *Lycium*. Thus, the discordant placement of *Phrodus* between the nuclear and cpDNA gene trees indicates that past hybridization has played a role in the history of *Phrodus*. More chloroplast data are needed for *Grabowskia*, although the data suggest that hybridization may also have been important in the evolutionary history of *Grabowskia* or the group of five *Lycium* species that both NIA and GBSSI data strongly support as sister to *Grabowskia* (Fig. 2A–B).

4.4. Conclusions

NIA appears to have considerable phylogenetic signal, especially when compared to the more commonly used nrITS data. However, incongruence between GBSSI and NIA suggests some incomplete lineage sorting, and there are also indications of hybridization, particularly given the conflicting placement of *Phrodus* between the nuclear and cpDNA topologies. Given these conflicts, more sequence data from other independent low-copy nuclear regions appear necessary to fully understand species-level relationships. In Lycieae, multiple conserved orthologous sequences (COSII markers) from the nuclear genome appear promising for understanding relationships among species (Wu et al., 2006; Levin et al., 2008, submitted for publication). Only with further multi-locus nuclear data will we be able to differentiate incomplete lineage sorting in a single gene history from the "true" species history for the closely related American *Lycium* species.

Acknowledgments

This work was supported by National Science Foundation Grant DEB 0343735 to RAL and JSM plus an REU supplement for JB. Additional support was provided by the Department of Biology at Amherst College and the Howard Hughes Medical Institute. The authors thank A. Kwan, N. Feliciano, and A. Savage for laboratory assistance; F. Rodriguez for primer sequences; and two anonymous reviewers for their helpful comments on the manuscript.

References

Álvarez, I., Wendel, J.F., 2003. Ribosomal ITS sequences and plant phylogenetic inference. Mol. Phylogenet. Evol. 29, 417–434.

Baldwin, B.G., 1992. Phylogenetic utility of the internal transcribed spacers of nuclear ribosomal DNA in plants: an example from the Compositae. Mol. Phylogenet. Evol. 1, 3–16.
Baldwin, B.G., Sanderson, M.J., Porter, J.M., Wojciechowski, M.F., Campbell, C.S., Donoghue, M.J., 1995. The ITS region of nuclear ribosomal DNA: a valuable source of evidence on angiosperm phylogeny. Ann. MO Bot. Gard. 82, 247–277.
Bohs, L., Olmstead, R.G., 2001. A reassessment of *Normania* and *Triguera* (Solanaceae). Plant Syst. Evol. 228, 33–48.
Buckler IV, E.S., Ippolito, A., Holtsford, T.P., 1997. The evolution of ribosomal DNA: divergent paralogous and phylogenetic implications. Genetics 145, 821–832.
Cheng, C., Dewdney, J., Nam, H., den Boer, B.G.W., Goodman, H.M., 1988. A new locus (*NIA1*) in *Arabidopsis thaliana* encoding nitrate reductase. EMBO J. 7, 3309–3314.
Chiang-Cabrera, F., 1981. A taxonomic study of the North American species of *Lycium* (Solanaceae). Ph.D. thesis, Department of Botany, University of Texas, Austin, USA.
Clegg, M.T., Zurawski, G., 1992. Chloroplast DNA and the study of plant phylogeny: present status and future prospects. In: Soltis, P.S., Soltis, D.E., Doyle, J.J. (Eds.), Molecular Systematics of Plants. Chapman & Hall, New York, pp. 1–13.
Condor Project, 2005. Condor 6.6.10. Available from: <http://www.cs.wisc.edu/condor/>.
Daniel-Vedele, F., Dorbe, M.-F., Caboche, M., Rouzé, P., 1989. Cloning and analysis of the tomato nitrate reductase-encoding gene: protein domain structure and amino acid homologies in higher plants. Gene 85, 371–380.
Demesure, B., Sodzi, N., 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.
Dillon, M.O., Tu, T., Soejima, A., Yi, T., Nie, Z., Tye, A., Wen, J., 2007. Phylogeny of *Nolana* (Nolaneae, Solanoideae, Solanaceae) as inferred from granule-bound starch synthase I (GBSSI) sequences. Taxon 56, 1000–1011.
Farris, J.S., Källersjö, M., Kluge, A.G., Bult, C., 1994. Testing significance of incongruence. Cladistics 10, 315–319.
Farris, J.S., Källersjö, M., Kluge, A.G., Bult, C., 1995. Constructing a significance test for incongruence. Syst. Biol. 44, 570–572.
Fortune, P.M., Schierenbeck, K.A., Ainouche, A.K., Jacquemin, J., Wendel, J.F., Ainouche, M.L., 2007. Evolutionary dynamics of *Waxy* and the origin of hexaploid *Spartina* species (Poaceae). Mol. Phylogenet. Evol. 43, 1040–1055.
Fortune, P.M., Pourtau, N., Viron, N., Ainouche, M.L., 2008. Molecular phylogeny and reticulate origins of the polyploid *Bromus* species from section *Genea* (Poaceae). Am. J. Bot. 95, 454–464.
Hamat, H.B., Kleinhofs, A., Warner, R.L., 1989. Nitrate reductase induction and molecular characterization in rice (*Oryza sativa* L.). Mol. Gen. Genet. 218, 93–98.
Hamman-Khalifa, A.M., Navajas-Pérez, R., de la Herrán, R., Ruiz Rejón, M., Garrido-Ramos, M.A., Ruiz Rejón, C., Rosúa, J.L., 2007. Establishing the genetic relationships between the wild and cultivated olives using a nuclear intron from nitrate reductase (*nia*-i3). Plant Syst. Evol. 269, 63–79.
Harris, N., Foster, J.M., Kumar, A., Davies, H.V., Gebhardt, C., Wray, J.L., 2000. Two cDNAs representing alleles of the nitrate reductase gene of potato (*Solanum tuberosum* L. cv. Desirée): sequence analysis, genomic organization and expression. J. Exp. Bot. 51, 1017–1026.
Hauser, D.L., Boyajian, G., 1997. Proportional change and patterns of homoplasy: Sanderson and Donoghue revisited. Cladistics 13, 97–100.
Hershkovitz, M.A., Zimmer, E.A., 1996. Conservation patterns in angiosperm rDNA ITS2 sequences. Nucleic Acids Res. 24, 2857–2867.
Howarth, D.G., Baum, D.A., 2002. Phylogenetic utility of a nuclear intron from nitrate reductase for the study of closely related plant species. Mol. Phylogenet. Evol. 23, 525–528.
Howarth, D.G., Baum, D.A., 2005. Genealogical evidence of homoploid hybrid speciation in an adaptive radiation of *Scaevola* (Goodeniaceae) in the Hawaiian Islands. Evolution 59, 948–961.
Hunziker, A.T., 2001. Genera Solanacearum. A.R.G. Gantner Verlag K.-G., Liechtenstein.
Levin, R.A., Miller, J.S., 2005. Relationships within tribe Lycieae (Solanaceae): paraphyly of *Lycium* and multiple origins of gender dimorphism. Am. J. Bot. 92, 2044–2053.
Levin, R.A., Wagner, W.L., Hoch, P.C., Hahn, W.J., Rodriguez, A., Baum, D.A., Katinas, L., Zimmer, E.A., Sytsma, K.J., 2004. Paraphyly in tribe Onagreae: insights into phylogenetic relationships of Onagraceae based on nuclear and chloroplast sequence data. Syst. Bot. 29, 147–164.
Levin, R.A., Watson, K., Bohs, L., 2005. A four-gene study of evolutionary relationships in *Solanum* section *Acanthophora*. Am. J. Bot. 92, 603–612.
Levin, R.A., Myers, N.R., Bohs, L., 2006. Phylogenetic relationships among the "spiny solanums" (*Solanum* subgenus *Leptostemonum*, Solanaceae). Am. J. Bot. 93, 157–169.
Levin, R.A., Bernardello, G., Venter, A.M., Shak, J.R., Miller, J.S., 2007. Evolutionary relationships in tribe Lycieae (Solanaceae). Acta Hort. 745, 225–239.
Levin, R.A., Whelan, A., Miller, J.S., 2008. The utility of nuclear conserved ortholog set II (COSII) genomic markers for species-level phylogenetics in *Lycium*. Meeting abstract, Botany 2008.
Li, J., Shoup, S., Chen, Z., 2007. Phylogenetic relationships of diploid species of *Betula* (Betulaceae) inferred from DNA sequences of nuclear nitrate reductase. Syst. Bot. 32, 357–365.
Lihová, J., Shimizu, K.K., Marhold, K., 2006. Allopolyploid origin of *Cardamine asarifolia* (Brassicaceae): incongruence between plastid and nuclear ribosomal DNA sequences solved by a single-copy nuclear gene. Mol. Phylogenet. Evol. 39, 759–786.
Maddison, W.P., Maddison, D.R., 2000. MacClade 4: Analysis of Phylogeny and Character Evolution. Sinauer, Sunderland, MA, USA.
Mansion, G., Zeltner, L., Bretagnolle, F., 2005. Phylogenetic patterns and polyploid evolution within the Mediterranean genus *Centaurium* (Gentianaceae—Chironieae). Taxon 54, 931–950.
Mason-Gamer, R.J., 2008. Allohexaploidy, introgression, and the complex phylogenetic history of *Elymus repens* (Poaceae). Mol. Phylogenet. Evol. 47, 598–611.
Mason-Gamer, R.J., Weil, C.F., Kellogg, E.A., 1998. Granule-bound starch synthase: structure, function, and phylogenetic utility. Mol. Biol. Evol. 15, 1658–1673.
Mast, A.R., Jones, E.H., Havery, S.P., 2005. An assessment of old and new DNA sequence evidence for the paraphyly of *Banksia* with respect to *Dryandra* (Proteaceae). Aust. Syst. Bot. 18, 75–88.
Miller, J.S., 2002. Phylogenetic relationships and the evolution of gender dimorphism in *Lycium* (Solanaceae). Syst. Bot. 27, 416–428.
Miller, J.S., Kamath, A., Levin, R.A., in press. Comparison of nine noncoding chloroplast regions for phylogenetic interference at the infrageneric level in tribe Lycieae (Solanaceae). Syst. Biol.
Moore, M.J., Jansen, R.K., 2006. Molecular evidence for the age, origin, and evolutionary history of the American desert plant genus *Tiquilia* (Boraginaceae). Mol. Phylogenet. Evol. 39, 668–687.
Mort, M.E., Archibald, J.K., Randle, C.P., Levsen, N.D., O'Leary, T.R., Topalov, K., Wiegand, C.M., Crawford, D.J., 2007. Inferring phylogeny at low taxonomic levels: utility of rapidly evolving cpDNA and nuclear ITS loci. Am. J. Bot. 94, 173–183.
Obbard, D.J., Harris, S.A., Buggs, R.J.A., Pannell, J.R., 2006. Hybridization, polyploidy, and the evolution of sexual systems in *Mercurialis* (Euphorbiaceae). Evolution 60, 1801–1815.
Oh, S.-H., Potter, D., 2003. Phylogenetic utility of the second intron of *LEAFY* in *Neillia* and *Stephanandra* (Rosaceae) and implications for the origin of *Stephanandra*. Mol. Phylogenet. Evol. 29, 203–215.
Olmstead, R.G., Sweere, J.A., Spangler, R.E., Bohs, L., Palmer, J.D., 1999. Phylogeny and provisional classification of the Solanaceae based on chloroplast DNA. In: Nee, M., Symon, D.E., Lester, R.N., Jessop, J.P. (Eds.), Solanaceae IV: Advances in Biology and Utilization. Royal Botanic Gardens, Kew, pp. 111–137.
Olmstead, R.G., Bohs, L., Migid, H.A., Santiago-Valentin, E., Garcia, V.F., Collier, S.M., 2008. A molecular phylogeny of the Solanaceae. Taxon 57, 1159–1181.
Peralta, I.E., Spooner, D.M., 2001. Granule-bound starch synthase (GBSSI) gene phylogeny of wild tomatoes (*Solanum* L. section *Lycopersicon* [Mill.] Wettst. subsection *Lycopersicon*). Am. J. Bot. 88, 1888–1902.
Posada, D., Crandall, K.A., 1998. Modeltest: testing the model of DNA substitution. Bioinformatics 14, 817–818.
Rambaut, A., 2002. Se-Al: sequence alignment editor, version 2.0. Available from: <http://evolve.zoo.ox.ac.uk/>.
Rodriguez, F., Spooner, D.M., 2004. Nitrate reductase phylogeny of wild potatoes (*Solanum* section *Petota*). Abstract, Botany 2004, Snowbird, UT, USA.
Rodriguez, F., Spooner, D.M., 2005. Nitrate reductase phylogeny of wild potatoes (*Solanum* section *Petota*) using all three introns. Abstract, Botany 2005, Austin, TX, USA.
Salanoubat, M., Ha, D.B.D., 1993. Analysis of the petunia nitrate reductase apoenzyme-encoding gene: a first step for sequence modification analysis. Gene 128, 147–154.
Sang, T., Crawford, D.J., Stuessy, T.F., 1997. Chloroplast DNA phylogeny, reticulate evolution, and biogeography of *Paeonia* (Paeoniaceae). Am. J. Bot. 84, 1120–1136.
Shak, J.R., Levin, R.A., Miller, J.S., 2006. Phylogenetic relationships of Old World *Lycium* (Solanaceae): reticulate evolution in the African taxa. Abstract, Botany 2006, Chico, CA, USA.
Shaw, J., Lickey, E.B., Beck, J.T., Farmer, S.B., Liu, W., Miller, J., Sirpun, K.C., Winder, C.T., Schilling, E.E., Small, R.L., 2005. The tortoise and the hare II: relative utility of 21 noncoding chloroplast DNA sequences for phylogenetic analysis. Am. J. Bot. 92, 142–166.
Shaw, J., Lickey, E.B., Schilling, E.E., Small, R.L., 2007. Comparison of whole chloroplast genome sequences to choose noncoding regions for phylogenetic studies in angiosperms: the tortoise and the hare III. Am. J. Bot. 94, 275–288.
Shimodaira, H., Hasegawa, M., 1999. Multiple comparisons of log-likelihoods with applications to phylogenetic inference. Mol. Biol. Evol. 16, 1114–1116.
Small, R.L., 2004. Phylogeny of *Hibiscus* sect. *Muenchhusia* (Malvaceae) based on chloroplast *rpL16* and *ndhF*, and nuclear ITS and GBSSI sequences. Syst. Bot. 29, 385–392.
Small, R.L., Cronn, R.C., Wendel, J.F., 2004. Use of nuclear genes for phylogeny reconstruction in plants. Aust. Syst. Bot. 17, 145–170.
Swofford, D.L., 2002. PAUP*: phylogenetic analysis using parsimony (* and other methods), version 4. Sinauer, Sunderland, MA, USA.
Taberlet, P., Gielly, L., Pautou, G., Bouvet, J., 1991. Universal primers for amplification of three non-coding regions of chloroplast DNA. Plant Mol. Biol. 17, 1105–1109.
Tamura, N., Takahashi, H., Takeba, G., Satoi, T., Nakagawa, H., 1997. The nitrate reductase gene isolated from DNA of cultured spinach cells. Biochim. Biophys. Acta 1338, 151–155.
Vaucheret, H., Vincentz, M., Kronenberger, J., Caboche, M., Rouzé, P., 1989. Molecular cloning and characterization of the two homologous genes coding for nitrate reductase in tobacco. Mol. Gen. Genet. 216, 10–15.
White, T.J., Bruns, T., Lee, S., Taylor, J., 1990. Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. In: Innis, M., Gelfand, D.,

Sninsky, White, T. (Eds.), 2008, PCR Protocols: A Guide to Methods and Applications. Academic Press, San Diego, CA, pp. 315–322.
Wilgenbusch, J., 2003. RepMaker 0.9. Available from: <http://paup.csit.fsu.edu/scripts/RepMaker.tar.gz>.
Wu, F., Mueller, L.A., Crouzillat, D., Pétiard, V., Tanksley, S.D., 2006. Combining bioinformatics and phylogenetics to identify large sets of single-copy orthologous genes (COSII) for comparative, evolutionary and systematic studies: a test case in the euasterid plant clade. Genetics 174, 1407–1420.
Yeung, K., Miller, J.S., Savage, A.E., Husband, B.C., Igic, B., Kohn, J.R., 2005. Association of ploidy and sexual system in *Lycium californicum* (Solanaceae). Evolution 59, 2048–2055.
Yi, T., Miller, A.J., Wen, J., 2007. Phylogeny of *Rhus* (Anacardiaceae) based on sequences of nuclear *Nia*-i3 intron and chloroplast *trnC-trnD*. Syst. Bot. 32, 379–391.
Yi, T., Wen, J., Golan-Goldhirsh, A., Parfitt, D.E., 2008. Phylogenetics and reticulate evolution in *Pistacia*. Am. J. Bot. 95, 241–251.
Zhou, J., Kleinhofs, A., 1996. Molecular evolution of nitrate reductase genes. J. Mol. Evol. 42, 432–442.