ACADEMIC PRESS

Molecular Phylogenetics and Evolution 23 (2002) 525–528

MOLECULAR PHYLOGENETICS AND EVOLUTION

www.academicpress.com

Short Communication

Phylogenetic utility of a nuclear intron from nitrate reductase for the study of closely related plant species

Dianella G. Howarth[a,*] and David A. Baum[b]

[a] *Department of Organismic and Evolutionary Biology, Harvard University Herbaria, 22 Divinity Avenue, Cambridge, MA 02138, USA*
[b] *Department of Botany, University of Wisconsin, 530 Lincoln Drive, Madison, WI 53706, USA*

Received 13 September 2001; received in revised form 17 December 2002

Molecular sequence data have become increasingly used for examining the evolutionary history of plants at scales ranging from relationships among the major lineages of land plants to relationships within individual genera. Phylogenetic systematics has progressed in part through the development of new molecular markers suited to particular classes of phylogenetic problems (e.g., Mathews et al., 2000; Olsen and Schaal, 1999; Palmer et al., 1988; Taberlet et al., 1991; Whitlock and Baum, 1999). Nonetheless, there are few genes that may be easily amplified from a wide range of taxa that have levels of sequence variation appropriate for studies of closely related species or studies at the population level (Schaal et al., 1998). Consequently, species and population-level phylogenies, especially involving recent radiations, most commonly utilize techniques such as inter-simple sequence repeats (Wolfe and Randle, 2001), amplified fragment length polymorphisms (Albertson et al., 1999; Bakkeren et al., 2000), and microsatellites (Billotte et al., 2001; Dayanandan et al., 1997). Unfortunately, these techniques do not provide gene genealogies, which may be useful for clarifying species limits and patterns of gene flow (Avise and Ball, 1990; Baum and Shaw, 1995), but rather provide a summation across numerous, possibly discordant, gene genealogies.

The most commonly used marker for sequence analysis at low phylogenetic levels is the internal transcribed spacer (ITS) region of 18S–5.8S–28S rDNA (Baldwin et al., 1995). While ITS has been useful in elucidating some island lineages (reviewed in Baldwin et al., 1998), it is often invariant in clades that have undergone rapid or recent radiations (Ganders et al., 2000; Vargas et al., 1998). This is particularly true in island radiations, which are consequently often assessed with morphological data (Buss et al., 2001; Wagner and Funk, 1995). Additionally, ITS is present in multiple copies per genome and subject to varying degrees of concerted evolution, which can occasionally result in a distorted picture of evolutionary history (Buckler et al., 1997, Wendel et al., 1995).

Palumbi and Baker (1994) described a method to find more variable regions of the genome, which they called exon-primer, intron-crossing (EPIC). The method entails using data from two or more model species to find conserved exons that flank variable introns. Primers can then be designed within these flanking exons that will amplify the intron in a broader array of taxa. Strand et al. (1997) used such a technique to design primers for amplifying introns from a diversity of plant taxa. However, they focused on markers with multiple small introns. For example, glyceraldehyde-3-phosphate dehydrogenase (G3PDH), which has been used successfully at the population level (Olsen and Schaal, 1999), has introns that are typically only 100 bp in length. Longer introns would provide more phylogenetic information per unit sequence because less exon sequence would need to be generated. Therefore, we set out to use EPIC to find longer nuclear introns that could be amplified from a diversity of plants and could provide phylogenetic information at low taxonomic levels.

This paper describes one target of this effort, the nitrate reductase (NIA) gene, which is a promising candidate for phylogenetic reconstruction. The gene (sometimes abbreviated NR) has been isolated from fungi, algae, and land plants (Zhou and Kleinhofs, 1996). NIA from most land plants acts as a homodimer and, with NADH as a cofactor, catalyzes the first reaction in the uptake of nitrogen from the soil, the reduction of

* Corresponding author. Fax: +617-495-9484.
E-mail addresses: dhowarth@oeb.harvard.edu (D.G. Howarth); dbaum@facstaff.wisc.edu (D.A. Baum).

PII: S1055-7903(02)00035-0

Table 1
Nitrate reductase genes published in GenBank and the size of their 3rd introns

Species	Intron size (bp)	Publication	GenBank accession no.
Arabidopsis thaliana	439	Wilkinson and Crawford (1993)	Z19050
Cichorium intybus	1411	Palms et al. (unpublished)	X84103
Lotus japonicus	596	Waterhouse et al. (unpublished)	X80670
Lycopersicon esculentum	436	Daniel-Vedele et al. (1989)	X14060
Nicotiana tabacum	652, 788	Vaucheret et al. (1989)	X14058, X14059
Petuna x hybrida	497	Salanoubat and Ha (1993)	L11563
Phaseolus vulgaris	677, 1646	Hoff et al. (1991); Jensen et al. (1994)	X53603, U01029
Scaevola procera	1285	This paper	AF460205
Spinacia oleracea	1374	Tamura et al. (1997)	D86226
Zea mays	85[a]	Campbell et al. (unpublished)	U20450

Note. Two different copies were identified in *Nicotiana* (which is tetraploid) and *Phaseolus*.
[a] This is only one of three loci for which the sequence of NIA is reported.

nitrate to nitrite (Hoff et al., 1994). The gene has three introns in conserved positions in all angiosperms studied (Salanoubat and Ha, 1993). The location of the first two introns is also shared with green algae (nonangiosperm land plants have not been studied), whereas the third intron (NIA-i3) has been found only in angiosperms (Zhou and Kleinhofs, 1996). This third intron varies in length (Table 1), but is generally large for a plant intron, attaining 1.6 kb in *Phaseolus* (Jensen et al., 1994).

We downloaded and aligned published GenBank NIA protein and nucleotide sequences from *Petuna*, *Cichorium*, and *Lotus*. The intron positions were determined, allowing degenerate primers to be designed within conserved regions of exons III and IV (Table 2). There was, theoretically, enough conservation in the exons for primers to amplify NIA-i3 from both rosids and asterids. Primers were also designed for the second intron of NIA (Table 2), but these primers did not amplify as consistently. Successfully cloned sequences from NIA-i2 indicate that it is actually longer than NIA-i3 in *Scaevola* (approximately 1.7 kb).

To evaluate the phylogenetic utility of NIA-i3 we here focus on the genus *Scaevola* (Goodeniaceae). *Scaevola* is of interest in this case because it includes a number of rapidly evolving island clades, including a radiation of seven species endemic to Hawaii. We amplified NIA-i3 from genomic DNA of four divergent *Scaevola* species (*S. aemula*, *S. nitida*, *S. plumieri*, and *S. taccada*) and seven species endemic to Hawaii (*S. chamissoniana*, *S. coriacea*, *S. gaudichaudiana*, *S. gaudichaudii*, *S. kilaueae*, *S. mollis*, and *S. procera*). These taxa were chosen because they differ in their degree of divergence, and sequences from ITS are already available (unpublished data), allowing a comparison of sequence evolution between these regions. Additionally, partial G3PDH sequences are also available for *S. taccada* and the Hawaiian endemics (unpublished data).

PCR amplification was carried out on a PTC-200 Peltier Thermal Cycler (MJ Research, Waltham, MA). NIA-i3 was amplified using a stepdown PCR protocol (Hecker and Roux, 1996) with the annealing temperature beginning at 62 °C and ending at 49 °C. The reactions were performed using either PfuTurbo Hotstart DNA Polymerase (Stratagene, La Jolla, CA) or *Taq* DNA Polymerase (QIAGEN, Valencia, CA) in 25 μL, with final concentrations of 2.5 mM $MgCl_2$, 0.5 μM of each primer, 0.8 mM dNTPs, and 1× Q solution (QIAGEN, Valencia, CA). The reactions were then gel purified using the QIAquick gel extraction protocol (QIAGEN, Valencia, CA) and eluted with 10 μL of ddH_2O. This product was then cloned using the pGEM-T Easy TA cloning kit (Promega, Madison, WI). At least five colonies were screened for each taxon to check for multiple loci. Clones were purified with the Sigma GenElute Plasmid Miniprep Kit (Sigma, St. Louis, MO) and were then sequenced using the original primers with BigDye termination mix (PE Applied Biosystems, Foster City, CA) and run on an automated DNA sequencer (ABI 377 or 3100). To verify the sequences, additional sequencing reactions were conducted using *Scaevola*-specific internal primers.

The primers NIA3F and NIA3R amplified a 1.2- to 1.5-kb fragment in all the *Scaevola* species used. A smaller band of approximately 0.3 kb was also amplified

Table 2
Primer sequences and locations for the second and third introns of nitrate reductase

	Forward primer	Reverse primer
NIA-i2	5′-TCBGTGATTACGACGCCGTGTCATGA-3′ 68–93 bases 5′ of intron	5′-GACCARAARCACCARCACCARTAYT-3′ 104–128 bases 3′ of intron
NIA-i3	5′-AARTAYTGGTGYTGGTGYTTYTGGTC-3′ 107–132 bases 5′ of intron	5′-GAACCARCARTTGTTCATCATDCC-3′ 0–24 bases 3′ of intron

Table 3
Pairwise sequence divergence (proportion of differing sites) of NIA-i3 in *Scaevola*

	S. aemu	*S. gaud*	*S. niti*	*S. plum*	*S. proc*	*S. tacc*
S. aemu	0	**0.1031 (2.30)**	**0.1352 (4.13)**	**0.1276 (2.47)**	**0.1073 (2.32)**	**0.1600 (2.88)**
S. gaud	0.0449	0	**0.0905 (5.39)**	**0.0475 (1.27)**	**0.0106 (2.72)**	**0.1486 (2.66)**
S. niti	0.0327	0.0168	0	**0.1235 (4.33)**	**0.0975 (5.02)**	**0.0850 (1.72)**
S. plum	0.0517	0.0375	0.0285	0	**0.0537 (1.43)**	**0.1522 (2.10)**
S. proc	0.0463	0.0039	0.0194	0.0375	0	**0.1427 (2.45)**
S. tacc	0.0556	0.0558	0.0493	0.0725	0.0583	0

Note. Numbers above the diagonal (in boldface) are the third intron of NIA. Values below the diagonal are for the ITS region (both ITS1 and ITS2). The values in parentheses are the proportional difference of NIA over ITS. Only two of the seven Hawaiian species are included in this table. Specific epithets are abbreviated to the first four letters: aemu, *S. aemula*; gaud, *S. gaudichaudiana*; niti, *S. nitida*; plum, *S. plumieri*; proc, *S. procera*; tacc, *S. taccada*.

across all taxa. The cloned product of this smaller band corresponds to a separate, apparently nonfunctional copy of NIA (there were multiple missense base mutations in the exon sequence) with a much smaller intron (96 bp), which will not be discussed further.

Within the larger fragment no more than two distinct but almost identical sequences (putative alleles) were found from a single individual except in *Scaevola glabra*, which is tetraploid. Therefore, it appears that the diploid species contain only one NIA locus with the full-length intron.

The percentage sequence divergence among taxa was 1.3- to 5.4-fold greater in NIA-i3 than in ITS in all comparisons (Table 3). A similar difference was found for G3PDH. For example, sequence divergence at G3PDH was 3.1 times less than that of NIA-i3 in a comparison between *S. taccada* and *S. procera*.

Within the seven Hawaiian species, the matrix for NIA-i3 was 1287 characters, of which 33 (2.6%) were parsimony informative. For these same taxa, an ITS matrix was 786 bases in length with only 11 (1.4%) parsimony-informative characters. The G3PDH matrix was 915 bp with only 6 (0.7%) parsimony-informative sites. Unfortunately, these numbers are not easily comparable due to differential sampling of accessions and the presence of multiple alleles in NIA-i3.

Separate phylogenetic analyses of the ITS and NIA-i3 sequences (ignoring indels) were conducted using parsimony. In both cases only two clades were supported with >70% bootstrap—suggesting that phylogenetic signal is roughly equivalent between these two data sets. However, unlike ITS, NIA-i3 also contained four parsimony-informative indel regions which, when coded as binary characters, led to one additional clade attaining a bootstrap >70%. Additionally, multiple accessions of a species had identical ITS sequences, unlike multiple accessions in NIA-i3 which varied within species. Of the 13 Hawaiian individuals sampled from individuals with unambiguous species affiliation (based on morphology), 4 were heterozygous. Nonetheless, the alleles from each named species clustered on the phylogeny. In contrast, 5 putatively hybrid individuals were found to be heterozygous with one of each of the putative parental species alleles. Therefore, it appears that NIA-i3 yields greater low-level phylogenetic information than ITS.

The primers given in Table 2 appear to be broadly applicable. For example, unpublished research in *Tilia* (J. Li and Y. Chen), *Antirrhinum* (R.K. Oyama and D.A. Baum), and *Ribes* (L. Schultheis and M.J. Donoghue) has involved use of these primers and has yielded high levels of molecular variation, as we found for *Scaevola*. Although multiple loci are sometimes present, these generally give products of different sizes, allowing them to be separated by gel purification (Jensen et al., 1994). Consequently, we are optimistic that NIA-i3 will have general utility for the study of the history and population biology of rapidly evolving species and complexes.

Acknowledgments

We thank R.K. Oyama, J. Li, Y. Chen, L. Schultheis, and M.J. Donoghue for sharing unpublished data. Helpful comments on the manuscript were provided by B.G. Baldwin, S.J. Brunsfeld, M.P. Dunn, L. Hileman, R.K. Oyama, and B.A. Whitlock. This work was supported by an ASPT Graduate Research Award and a Harvard University Student Research Fund, both to D.G. Howarth.

References

Albertson, R.C., Markert, J.A., Danley, P.D., Kocher, T.D., 1999. Phylogeny of a rapidly evolving clade: The cichlid fishes of Lake Malawi, East Africa. Proc. Natl. Acad. Sci. USA 96, 5107–5110.

Avise, J.C., Ball, R.M. (Eds.), 1990. Principles of Genealogical Concordance in Species Concepts and Biological Taxonomy. Oxford University Press, New York.

Bakkeren, G., Kronstad, J.W., Levesque, C.A., 2000. Comparison of AFLP fingerprints and ITS sequences as phylogenetic markers in Ustilaginomycetes. Mycologia 92, 510–521.

Baldwin, B.G., Crawford, D.J., Francisco-Ortega, J., Kim, S.C., Sang, T., Stuessy, T.F., 1998. In: Soltis, D.E., Soltis, P.S., Doyle, J.J. (Eds.), Book Molecular Phylogenetic Insights on the Origin and Evolution of Oceanic Island Plants, vol. II. Kluwer Academic Publishers, Boston, pp. 410–441.

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.

Baum, D.A., Shaw, K.L., 1995. In: Hoch, P.C., Stephenson, A.G. (Eds.), Book Genealogical Perspectives on the Species Problem, vol. 53. Missouri Botanical Garden Press, St. Louis, pp. 89–303.

Billotte, N., Risterucci, A.M., Barcelos, E., Noyer, J.L., Amblard, P., Baurens, F.C., 2001. Development, characterisation, and across-taxa utility of oil palm (*Elaeis guineensis* Jacq.) microsatellite markers. Genome 44, 413–425.

Buckler, E.S., Ippolito, A., Holtsford, T.P., 1997. The evolution of ribosomal DNA: Divergent paralogues and phylogenetic implications. Genetics 145, 821–832.

Buss, C.C., Lammers, T.G., Wise, R.R., 2001. Seed coat morphology and its systematic implication in *Cyanea* and other genera of Lobelioideae (Campanulaceae). Am. J. Bot. 88, 1301–1308.

Daniel-Vedele, F., Dorbe, M.F., Caboche, M., Rouze, P., 1989. Cloning and analysis of the tomato nitrate reductase encoding gene-protein domain structure and animo acid homologies in higher plants. Gene 85, 371–380.

Dayanandan, S., Bawa, K.S., Kesseli, R., 1997. Conservation of microsatellites among tropical trees (Leguminosae). Am. J. Bot. 84, 1658–1663.

Ganders, F.R., Berbee, M., Pirseyedi, M., 2000. ITS base sequence phylogeny in *Bidens* (Asteraceae): Evidence for the continental relatives of Hawaiian and Marquesan. Bidens. Syst. Bot. 25, 122–133.

Hecker, K.H., Roux, K.H., 1996. High and low annealing temperatures increase both specificity and yield in touchdown and stepdown PCR. BioTechniques 20, 478–485.

Hoff, T., Stummann, B.M., Henningsen, K.W., 1991. Cloning and expression of a gene encoding a root specific nitrate reductase in bean *Phaseolus vulgaris*. Physiol. Plant. 82, 197–204.

Hoff, T., Truong, H.N., Caboche, M., 1994. The use of mutants and transgenic plants to study nitrate assimilation. Plant Cell Environ. 17, 489–506.

Jensen, P.E., Hoff, T., Moller, M.G., Stummann, B.M., Henningsen, K.W., 1994. Indentification and characterization of a nitrate reductase gene from bean (*Phaseolus vulgaris*) containing four introns. Physiol. Plant. 92, 613–623.

Mathews, S., Tsai, R.C., Kellogg, E.A., 2000. Phylogenetic structure in the grass family (Poaceae): Evidence from the nuclear gene phytochrome B. Am. J. Bot. 87, 96–107.

Olsen, K.M., Schaal, B.A., 1999. Evidence on the origin of cassava: Phylogeography of *Manihot esculenta*. Proc. Natl. Acad. Sci. USA 96, 5586–5591.

Palmer, J.D., Jansen, R.K., Michaels, H.J., Chase, M.W., Manhart, J.R., 1988. Chloroplast DNA variation and plant phylogeny. Ann. Mo. Bot. Gard. 75, 1180–1206.

Palumbi, S.R., Baker, C.S., 1994. Contrasting population structure from nuclear intron sequences and mtDNA of humpback whales. Mol. Biol. Evol. 11, 426–435.

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.

Schaal, B.A., Hayworth, D.A., Olsen, K.M., Rauscher, J.T., Smith, W.A., 1998. Phylogeographic studies in plants: problems and prospects. Mol. Ecol. 7, 465–474.

Strand, A.E., Leebens-Mack, J., Milligan, B.G., 1997. Nuclear DNA-based markers for plant evolutionary biology. Mol. Ecol. 6, 113–118.

Taberlet, P., Gielly, L., Pautou, G., Bouvet, J., 1991. Universal primers for amplification of 3 noncoding 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. BBA-Protein Struct. Mol. Enzymol. 1338, 151–155.

Vargas, P., Baldwin, B.G., Constance, L., 1998. Nuclear ribosomal DNA evidence for a western North American origin of Hawaiian and South American species of *Sanicula* (Apiaceae). Proc. Natl. Acad. Sci. USA 95, 235–240.

Vaucheret, H., Kronenberger, J., Rouze, P., Caboche, M., 1989. Complete nucleotide sequence of the 2 homologous tobacco nitrate reductase genes. Plant Mol. Biol. 12, 597–600.

Wagner, W.L., Funk, V.A. (Eds.), 1995. Hawaiian Biogeography: Evolution on a Hot Spot Archipelago. Smithsonian Inst. Press, London House, New York.

Wendel, J.F., Schnabel, A., Seelanan, T., 1995. An unusual ribosomal DNA sequence from *Gossypium gossypioides* reveals ancient, cryptic, intergenomic introgression. Mol. Phylogenet. Evol. 4, 298–313.

Whitlock, B.A., Baum, D.A., 1999. Phylogenetic relationships of *Theobroma* and *Herrania* (Sterculiaceae) based on sequences of the nuclear gene Vicilin. Syst. Bot. 24, 128–138.

Wilkinson, J.Q., Crawford, N.M., 1993. Identification and characterization of a chlorate resistant mutant of *Arabidopsis thaliana* with mutations in both nitrate reductase structural genes NIA1 and NIA2. Mol. Gen. Genet. 239, 289–297.

Wolfe, A.D., Randle, C.P., 2001. Relationships within and among species of the holoparasitic genus *Hyobanche* (Orobanchaceae) inferred from ISSR banding patterns and nucleotide sequences. Syst. Bot. 26, 120–130.

Zhou, J., Kleinhofs, A., 1996. Molecular Evolution of Nitrate Reductase Genes. J. Mol. Evol. 42, 432–442.