Available online at www.sciencedirect.com
SCIENCE DIRECT®
Molecular Phylogenetics and Evolution 36 (2005) 484–493
MOLECULAR PHYLOGENETICS AND EVOLUTION
www.elsevier.com/locate/ympev

Incongruence between primary sequence data and the distribution of a mitochondrial *atp1* group II intron among ferns and horsetails

Niklas Wikström [a,*], Kathleen M. Pryer [b]

[a] *Department of Systematic Botany, Evolutionary Biology Centre, Uppsala University, Norbyvägen 18D, 752 36 Uppsala, Sweden*
[b] *Department of Biology, Duke University, Durham, NC 27708, USA*

Received 1 June 2004; revised 17 February 2005
Available online 26 May 2005

Abstract

Using DNA sequence data from multiple genes (often from more than one genome compartment) to reconstruct phylogenetic relationships has become routine. Augmenting this approach with genomic structural characters (e.g., intron gain and loss, changes in gene order) as these data become available from comparative studies already has provided critical insight into some long-standing questions about the evolution of land plants. Here we report on the presence of a group II intron located in the mitochondrial *atp1* gene of leptosporangiate and marattioid ferns. Primary sequence data for the *atp1* gene are newly reported for 27 taxa, and results are presented from maximum likelihood-based phylogenetic analyses using Bayesian inference for 34 land plants in three data sets: (1) single-gene mitochondrial *atp1* (exon + intron sequences); (2) five combined genes (mitochondrial *atp1* [exon only]; plastid *rbcL*, *atpB*, *rps4*; nuclear SSU rDNA); and (3) same five combined genes plus morphology. All our phylogenetic analyses corroborate results from previous fern studies that used plastid and nuclear sequence data: the monophyly of euphyllophytes, as well as of monilophytes; whisk ferns (Psilotidae) sister to ophioglossoid ferns (Ophioglossidae); horsetails (Equisetopsida) sister to marattioid ferns (Marattiidae), which together are sister to the monophyletic leptosporangiate ferns. In contrast to the results from the primary sequence data, the genomic structural data (*atp1* intron distribution pattern) would seem to suggest that leptosporangiate and marattioid ferns are monophyletic, and together they are the sister group to horsetails—a topology that is rarely reconstructed using primary sequence data.

Keywords: *atp1*; *atpB*; Bayesian inference; Euphyllophytes; Ferns; Horsetails; Mitochondrial intron; Monilophytes; Morphology; *rbcL*; *rps4*; SSU rDNA

1. Introduction

The last 10 years of comparative genomics has seen an accelerated increase in the availability of genomic structural characters to phylogenetic studies. These data include the gain and loss of introns, loss or transfer of genes, gene duplication, changes in gene order, and the formation and disruption of gene clusters. The complex nature of these data has provided critical insight into some long-standing questions about the evolution of land plants. For example, the broad phylogenetic relationships that are indicated by the conservative distribution of plastid introns, such as the shared presence of two tRNA introns in the Charophyceae (green algae) and all land plants (Manhart and Palmer, 1990), and by structural rearrangements in the plastid genome, such as the 30-kb inversion shared by all vascular plants except lycophytes (Raubeson and Jansen, 1992a), are largely uncontested. However, recent phylogenetic studies that place Gnetales within conifers suggest that the absence of one of the two inverted repeats in plastids of conifers (Raubeson and Jansen, 1992b; Strauss et al., 1988) may

* Corresponding author. Fax: +46 18 471 6457.
E-mail addresses: Niklas.Wikstrom@ebc.uu.se (N. Wikström), pryer@duke.edu (K.M. Pryer).

1055-7903/$ - see front matter © 2005 Elsevier Inc. All rights reserved.
doi:10.1016/j.ympev.2005.04.008

be more complicated to interpret than initially perceived, given that Gnetales have both inverted repeat copies (Bowe et al., 2000; Chaw et al., 2000; Donoghue and Doyle, 2000; Qiu et al., 1999; Qiu and Palmer, 1999). Data on mitochondrial group I and group II introns (Beckert et al., 1999, 2001; Dombrovska and Qiu, 2004; Malek and Knoop, 1998; Pruchner et al., 2001; Qiu et al., 1998; Vangerow et al., 1999) is also proving critical, and although some mitochondrial patterns appear less conservative than those observed in plastids (Cho et al., 1998; Palmer et al., 2000), others do provide valuable information on relationships among major land plant groups (Dombrovska and Qiu, 2004; Qiu and Lee, 2000; Qiu and Palmer, 1999). Until now, the screening for mitochondrial introns has mainly focused on the relationships among liverworts, hornworts, and mosses (Beckert et al., 2001; Malek and Knoop, 1998; Pruchner et al., 2001; Qiu et al., 1998), and the question of which of these groups is most closely related to vascular plants.

Most phylogenetic analyses (Duff and Nickrent, 1999; Hedderson et al., 1998; Hiesel et al., 1994; Kenrick and Crane, 1997; Kranz and Huss, 1996; Raubeson and Jansen, 1992a) consistently recognize lycophytes (Lycophytina: clubmosses and relatives) as sister group to all remaining vascular plants, or euphyllophytes (Euphyllophytina, sensu Kenrick and Crane, 1997). Euphyllophytes comprise six major monophyletic lineages: Spermatophytata (seed plants), Equisetopsida (horsetails), Polypodiidae (leptosporangiate ferns), Psilotidae (whisk ferns), Marattiidae (marattioid ferns), and Ophioglossidae (ophioglossoid ferns). The relationships among these euphyllophyte lineages have been notoriously difficult to resolve. Phylogenetic analyses based on single genes (Duff and Nickrent, 1999; Hedderson et al., 1998; Hiesel et al., 1994; Kranz and Huss, 1996; Manhart, 1994, 1995), and/or morphology (Kenrick and Crane, 1997; Mishler et al., 1994; Pryer et al., 1995; Rothwell, 1999; Stevenson and Loconte, 1996) have either been incongruent, or only provided weak support for certain relationships. Pryer et al. (2001) partly resolved these difficulties. Their analyses, based on a combined data set of four genes and morphology, provided unequivocal support for the monophyly of all "seed-free" euphyllophyte lineages (Equisetopsida, Psilotidae, Marattiidae, Ophioglossidae, and Polypodiidae), and determined that together they formed the sister group to seed plants. Citing Beck and Stein (1993), Kenrick and Crane (1997) had already suggested the potential monophyly of this group based on a single anatomical character (shared protoxylem distribution), and in their provisional classification, this horsetail-fern clade was named Moniliformopses ("monilophytes," Judd et al., 2002; Pryer et al., 2004a,b). In addition to the morphological and primary sequence data supporting the monophyly of monilophytes, Pryer et al. (2001) documented a 9-bp length increase in the plastid protein coding gene *rps4* that was shared by all moniliophytes. Pryer et al. (2001) also provided strong support for grouping whisk ferns (Psilotidae) and ophioglossoid ferns (Ophioglossidae), a relationship indicated by earlier single-gene analyses (Hedderson et al., 1998; Manhart, 1994; Pryer et al., 1995), but never before unambiguously supported. The clade including Psilotidae and Ophioglossidae was resolved as the earliest-diverging monilophyte lineage, but relationships among horsetails (Equisetopsida), marattioid ferns (Marattiidae), and leptosporangiate ferns (Polypodiidae) were obscure, and there was conflict depending on which optimality criterion (maximum parsimony vs. maximum likelihood) they used (Pryer et al., 2001).

As part of a broader effort to add mitochondrial data to attempt to resolve some of the problems indicated, we came across a previously undocumented group II intron located in the mitochondrial *atp1* gene of leptosporangiate and marattioid ferns. Here, we report on the primary sequence data for both the exon and intron of *atp1*, and we conduct phylogenetic analyses that combine the mitochondrial *atp1* data (newly reported here) with plastid *rbcL*, *atpB*, *rps4*, nuclear SSU rDNA, and morphological data previously reported (Pryer et al., 2001). Our best estimates of phylogeny are used for interpreting the *atp1* intron evolution.

2. Materials and methods

2.1. Taxon sampling

A total of 34 taxa (Table 1) were selected for this mitochondrial *atp1* study to match those taxa considered by Pryer et al. (2001) in the five data sets they made available (morphology; plastid *rbcL*, *atpB*, *rps4*; nuclear SSU rDNA). In many cases, the same DNA or voucher material used in that study was used here to sequence mitochondrial *atp1*. Despite several attempts, using various species, we were unable to amplify mitochondrial *atp1* for *Lygodium* (Schizaeaceae). For some families, a different representative genus was selected. Gleicheniaceae, for example, was represented here by *Sticherus* instead of *Gleichenia*, Salviniaceae by *Azolla* rather than *Salvinia*, and hornworts by *Phaeoceros*, not *Anthoceros*.

2.2. DNA extraction, amplification, and sequencing

Total cellular DNA was extracted using the DNeasy Plant Mini Kit from Qiagen following the protocol specified by the manufacturer. PCR fragments corresponding to bases 99–699 of the *Marchantia polymorpha* mitochondrial *atp1* gene were amplified for 27 species using forward primer F83-*atp1* (5′-ATGAGGTCGGTC GAGTGRT-3′) and reverse primer R725-*atp1* (5′-GGA