British Mycological Society promoting fungal science

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

Secondary structure of ITS2 rRNA provides taxonomic characters for systematic studies — a case in *Lycoperdaceae (Basidiomycota)*

Dirk KRÜGER[†,*], *Andrea GARGAS*[‡]

The Botany Department, University of Wisconsin, 430 Lincoln Drive, Madison, WI 53706, USA

ARTICLE INFO

Article history:
Received 6 April 2006
Received in revised form
24 September 2007
Accepted 30 October 2007
Corresponding Editor: Michael Weiss

Keywords:
Consensus structure
Internal transcribed spacer
ITS2 rDNA
Nomenclature

ABSTRACT

The secondary structure of the ITS2 rDNA transcript (pre-rRNA) could provide information for identifying homologous nucleotide characters useful for cladistic inference of relationships. Such structure data could become taxonomic characters. This work compares the effect of several modern nucleotide alignment strategies, including those making use of structure data, on phylogenetic inference. From both the phylogenetic analyses and comparative secondary structure, implications for taxonomy and evolution of puffball fungi are discussed. *Lycoperdaceae* remain insufficiently resolved with present taxon and data sampling. Neither alignment allows statistically robust phylogenetic hypotheses under any current optimality criterion. The secondary structure data at this time are best used as accessory taxonomic characters as their phylogenetic resolving power and confidence in validity is limited compared with underlying nucleotide characters. We introduce a preliminary nomenclature convention to describe secondary structure for defining consensus features. These consensus structures are illustrated for the clades /Calvatia, /Handkea–Echinatum, /Vascellum, /Morganella, and /Plumbea–Paludosa (*Bovista*).

Introduction

rRNA and secondary structure

The rRNA genes have long been seen as the ultimate tracer of evolutionary relationships. Historically, targeting these regions of the genomes has been eased by primer annealing site conservation and the redundancy of the array, offering a higher number of templates for PCR. Although protein-coding genes are transcribed into mRNA ultimately processed by excision of introns and translation into proteins, RNA is the ultimate phenotypic expression of rDNA. In rRNA, spacers are excised, and the RNA itself is the molecule for selection to act upon. RNA folds up in stretches that are, in fact, double-stranded, while others remain single-stranded.

rRNA genes have been widely used in systematic studies in fungi and beyond, and are common targets for identifying and quantifying phylotypes in medical and environmental samples. The 'coding' (we use coding as coding for RNA) SSU rRNA and LSU rRNA genes are highly conserved. Evidence suggests that secondary structures of the initial transcript play

---

[*] Corresponding author. Tel.: +49 (345) 558 5425.
E-mail address: dirk.krueger@ufz.de

[†] Present address: Helmholtz-Zentrum für Umweltforschung GmbH – UFZ, Department of Soil Ecology, Theodor-Lieser-Str. 4, D-06120 Halle/Saale, Germany.

[‡] Present address: Symbiology LLC, Middleton, WI 53562-1230, USA.

0953-7562/$ – see front matter © 2007 The British Mycological Society. Published by Elsevier Ltd. All rights reserved.
doi:10.1016/j.mycres.2007.10.019

important roles in ribosome assembly (Lalev & Nazar 1999; Lalev *et al.* 2000; Lalev & Nazar 2001), and putative secondary structures have long been recognized and archived for the coding regions of rDNA (van de Peer *et al.* 2000; Cannone *et al.* 2002), and recently for the internal spacers as well (Wolf *et al.* 2005). Bioinformatics programs, such as Mfold (Zuker 2003), Pfold (Knudsen & Hein 2003), and RNAviz (de Rijk *et al.* 2003), compute and depict putative secondary structure of any nucleic acid molecule, and a universal XML-based syntax was proposed for RNA structure bioinformatics (Waugh *et al.* 2002).

Characters from nuclear ribosomal gene sequences have been used for hypotheses of phylogenetic relationships among even distantly related organisms. Molecular phylogenetic analyses require the alignment of homologous sequence characters, and guidance from secondary structure information may aid in the alignment of homologous regions for phylogenetic analysis among plant and animal species (Jobes & Thien 1997; Goertzen *et al.* 2003; Xia *et al.* 2003), even between genomes as evolutionarily distant as of eukaryotic nuclei, prokaryotes, and eukaryotic organelles (Cedergren *et al.* 1988). Improved alignments are thus an alternative to excluding or recoding ambiguous indels (Lutzoni *et al.* 2000; Lee 2001; Young & Healy 2001; Creer *et al.* 2006). A further application of structure information to phylogenetics is recoding structure into new characters states (Caetano-Annolés 2002; Smith *et al.* 2004).

ITS regions have been used for phylogenetic analyses at the species to generic level, yet their primary nucleotide sequence often contains insertions and deletions (indels) making alignment difficult much beyond infraspecific levels. This ITS variability led to the assumption that non-coding ITS1 and ITS2 regions were mere 'junk DNA', whose evolution resulted from the accumulation of chance mutations unfettered by any functional constraints. However, research on plants and green algae suggested that ITS rDNA sequences provide evidence at a super-generic level (Baldwin *et al.* 1995; Hershkovitz & Zimmer 1996; Mai & Coleman 1997) and contain diagnostic characters for deeper divergences (Hershkovitz & Lewis 1996).

Sections of ITS2 transcripts are consistently predicted to form conserved stem-loop structures. Complementary base changes and the recovery of secondary structure motifs independent of primary sequence renders secondary structure highly appropriate for sequence alignment and taxonomy. Wolf *et al.* (2005) compiled a database of predicted secondary structures of ITS2. Incorporating some of these recent advances, we are now able to scrutinize the non-coding spacers and introns for secondary structure features and better employ them for biological systematics.

We have previously shown that the compression of phylogenetic information content inherent in the recoding of structural data *versus* underlying primary DNA sequence data precipitates a decrease of phylogenetic resolution (Krüger & Gargas 2004; Krüger *et al.* 2006). Refining models of rDNA sequence evolution and experimental verification of predicted structures will increase the reliability of rDNA as a phylogenetic marker.

After using secondary structure in polypore fungi (Krüger & Gargas 2004; Krüger *et al.* 2006), in the present study we (1) deduce and compare putative secondary structure features from ITS2 sequences from puffballs (*Lycoperdaceae*); (2) explore the effects of secondary structure incorporation into alignment and, subsequently, on the outcomes of phylogenetic analyses; and (3) define consensus structures as taxonomic characters for clades on phylogenetic trees. Finally, we discuss parsimonious explanations of DNA sequence changes responsible for the predicted secondary structures in *Lycoperdaceae* within a phylogenetic context and compared with previously used morphotaxonomic characters delimiting taxa.

Mutational events include compensatory base changes, one-sided changes (mutations leading to introduction or loss of new features), resizing changes (mutations increasing or decreasing feature sizes) or silent changes (mutations not altering the secondary structure). These trends cannot yet be described as directional, and the timing of putative events relative to clade separation is unknown (an event might be suggested by a homoplasious nucleotide character or we might miss data to correctly infer an event's occurrence). This precludes two useful future uses of pointing out such hypothetical events: (1) possible confirmation with independent data; and (2) possible experimentation on the feasibility of such event to allow formation of a functional ribosome.

Lycoperdaceae

Puffballs (gasteromycetes) are common inhabitants of grasslands, although some are found in forests, e.g. on wood. One example is the giant puffball, *Calvatia gigantea,* which produces millions of spores in its fruit body cavity (Ingold 1971); puffball spores may be potently allergenic (Levetin *et al.* 1992; Horner *et al.* 1993). Within the main family, *Lycoperdaceae*, the value of capillitia (hyphal threads) and spores — two traditional micromorphological characters for delimitation of puffball genera — remain ambiguous. The genera *Bovista* and *Lycoperdon* share common morphological types of capillitia inside their spore-producing cavities (the *Bovista* type, the intermediate type, and the *Lycoperdon* type of capillitium; all occur among *Bovista* species; Krüger *et al.* 2001; Bates 2004), calling into question the usefulness of this character and/or the traditional generic circumscriptions. Currently, available molecular sequence data do not satisfactorily delimit genera *Bovista* and *Lycoperdon* (Krüger *et al.* 2001; Bates 2004). The genus *Handkea* (Kreisel 1989) may be separated from *Calvatia* in micromorphology based on the presence of a distinct capillitium type featuring slit-like depressions in the cell walls, rather than pores.

Puffballs of the order *Lycoperdales,* as currently circumscribed, are believed to be a monophyletic group, although its former enclosing taxon, class *Gasteromycetes*, is now considered an artificial group (Hawksworth *et al.* 1995; Hibbett *et al.* 1997; Moncalvo *et al.* 2002). Morphological characters supporting the DNA-confirmed segregation of *Geastrales* (earthstars) from the order *Lycoperdales* (required to establish monophyly of the *Lycoperdales*) have been highlighted by Iosifidou & Agerer (2002). *Lycoperdaceae* are closely related to agarics (mushrooms such as *Lepiota* spp. and *Agaricus* spp.) (Vellinga 2004) and are perhaps best included in the *Agaricaceae.*