Journal of Advanced Bioinformatics Applications and Research ISSN 0976-2604
Vol 2, Issue 3, 2011, pp 200-205 http://www.bipublication.com

PHYLOGENETIC RELATIONSHIP IN FOUR *ASPERGILLUS* SPECIES BASED ON THE SECONDARY STRUCTURE OF INTERNAL TRANSCRIBED SPACER REGION OF rDNA

M. Thangaraj[1], B. Valentin Bhimba[2], J. Meenupriya[2]
[1]Centre of Advanced Study in Marine Biology, Faculty of Marine Sciences,
Annamalai University, Parangipettai, Tamil Nadu, India - 608 502.
[2]Department of Biotechnology, Sathyabama University, Chennai. India-119.

ABSTRACT

Recent studies show that, the secondary structure of the RNA transcript could provide useful information for systematic studies of fumgi. These morphological informations are not found in primary sequences. In this study, four *Aspergillus* species internal transcribed spacer region 1 (ITS1) secondary structure was predicted and compared. By observing the structures, some topologies were conserved in all the species and some were varied. These characters could be used for the identification fungi at generic level. Secondary structure topologies such as first stalk and junctions were varied between species. The phylogenetic trees based on both sequence and secondary structure showed similar clad in the four *Aspergillus* species.

Keywords: ITS, secondary structure, marine fungi, *Aspergillus*, phylogeny

INTRODUCTION

Currently most of the phylogenetic studies have focused on the primary DNA sequence information. However, RNA secondary structures are also used in systematic studies because they include characteristics, not found in the primary sequence, that give a distinct 'morphological' information. The rRNA morphometrics approach has been employed in the present study for comparisons between primary and secondary structure information. It is an established fact that rRNA structure is highly conserved throughout evolution as most of the folding is functionally important despite primary sequence divergence [1]. The novel approach of that relies both on traditional morphological comparison and on molecular sequence comparison by measuring the structural parameters of the ITS1 secondary structure homologies (geometrical features, bond energies, base composition etc.) is recently being used to study the phylogenetic relationships of varies species [2]. The secondary-structure elements of the RNA molecule, i.e., the helices, loops, bulges, and separating single-stranded portions, can be considered as phylogenetic characters [3, 4, 5].

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 SSU rRNA and LSU rRNA genes are highly conserved [6]. 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 intraspecific levels. This ITS variability led to the assumption that non-coding ITS1 and ITS2 regions, 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 [7] and contain diagnostic characters for deeper divergences [8].

Bioinformatics programs, such as Mfold [9] and Pfold [10] compute and depict putative secondary structure of any nucleic acid molecule, and a universal XML- based syntax was proposed for RNA structure bioinformatics [11]. Sections of ITS1 and 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. [12] 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 systematic. The present study was undertaken to demonstrate the sequence analysis of the ribosomal RNA (ITS1) of four selected marine fungi which are having potent antimicrobial properties. We also construct its phylogeny using rRNA secondary structures supplementing the primary sequence analysis.

MATERIALS AND METHODS

The fungi, *Aspergillus terreus*, *A. ochraceus*, *A. flavus* and *A. melleus* were cultured in the medium of potato dextrose agar broth at room temperature in dark for 72 hrs. The genomic DNA was isolated and the ITS region was amplified using primers, ITS1 (5'-TCCGTAGGTGAACCTGCGG-3') and ITS4 (5'-TCCTCCGCTTATTGATATGC-3') as per the standard protocol [13]. The cleaned up PCR product was sequenced by a commercial sequencing facility (Genei, Bangalore, India). The sequences were submitted to NCBI and got the accession numbers as GU815344, HQ449676, HQ449677 and HQ449678.

The Initial alignments were compiled using ClustalX with manual minimization of gaps using BioEdit. The ITS1 rDNA regions were identified by comparison to the published sequences of *Aspergillus novoparasiticus* [14]. ITS1 sequences were arranged with MEGA for construction of phylogenetic trees that were inferred using distance method like neigbour joining and character state maximum parsimony methods. Test of phylogenetic accuracy was done by high bootstrap values. After delimitation of their ITS1- proximal stems, initial folding for each ITS1 rDNA sequence was based on models predicted by Mfold version 3.1 [9] with default conditions set to: linear RNA sequence, folding temperature of 37^0C, 20% suboptimality, upper bound of 50 foldings, no limit to the maximum physical distance between paired bases, no constraint information, maximum number of nucleotides in a bulge loop=30, maximum asymmetry of an interior/bulge loop=30). The phylogenetic analysis was carried out using the Bayesian approach in Mr. Bayes V3.1.2. The input data sets were the secondary-structure elements of the RNA molecule, i.e., the helices, loops, bulges, and separating single-stranded portions.

RESULTS AND DISCUSSION

Phylogenetic tree using maximum parsimony method was obtained by comparing the ITS1 sequence of *Aspergillus* species (Fig.1). Based on the ITS1 primary sequence similarity test, *A. terreus* and *A. ochraceus* are closely related species than the other two isolates. The evolutionary divengence between sequences was also estimated and the number of base substitutions per site between sequences is

shown in Fig.2. The pairwise distance and nucleotide frequency (%) of ITS1 sequence in four *Aspergillus* species are given in Table 1 and 2 respectively.

The RNA secondary structures were predicted to provide the basic information for phylogenetic analysis are given in Fig. 3. The secondary structural features of ITS1 regions as shown in figures were analyzed based on the conserved stems and loops, which in order of preference were interior loop, hairpin loop and exterior loop in all the isolates. We obtained similar topologies for the four species on the basis of traditional primary sequence analysis using MEGA and the secondary structure matrix. The secondary structures yielded homologous models that grouped the conserved features (i.e. first stem and junctions). Generally RNA secondary structure prediction programs rely on free energy minimization using nearest-neighbour parameters for predicting the overall structural stability in terms of Gibbs free energy at 37^0C. The observed similarities at the secondary structural level are further reflected at the energy level (-ΔG). The difference in their topology, however, is due to differences in nucleotide sequence length. The ΔG energy composition of interior loop, hairpin loop, bulge loop, multi loop and exterior loop and stack in all the isolates are given in Table 3.

In the present study the phylogenetic studies involvement of secondary structure analysis as a tool, RNA folding is used for refining the alignment. The measurable structural parameters of the molecules are directly used as specific characters to construct a phylogenetic tree. These structures are inferred from the sequence of the nucleotides, often using energy minimization [15]. The primary sequence analysis revealed the closeness of these four *Aspergillus* species. On the basis of ITS1 primary sequence, *A. terreus* and *A. ochraceus* are the closely related species than the other two isolates. The same type of relationship was also observed in the phylogenetic tree based on the secondary structure (Fig.4).

Molecular morphometrics appears to be complementary to classical primary sequence analysis in phylogenetic studies as it takes into consideration only the size variation of homologous structural segments and this choice implies that the overall architecture of the molecule remains same among the observed taxa. This method helps in taking into account the regions where multiple alignments are barely reliable because of large number of indel operations. Besides, the secondary structures are built on each sequence separately thereby making it unnecessary for a sharp computational sequence analysis. Thus homologous recognizable characters on the secondary structures are easily traced out contrary to finding right counterpart for each nucleotide in every other sequence [16]. The secondary structures of ITS sequences of these fungal species provided us additional information for their relationship. The secondary structure analysis of the same data also confirmed the results of the primary sequence analysis.

CONCLUSION

The molecular study of the genus *Aspegillus*, which are the most important fungi having potent antimicrobial compounds is very important for their identification. In this context ITS motifs can be considered as a promising tool for *Aspergillus* species identification. RNA secondary structure analysis could be a valuable molecular tool for distinguishing new species and tuning up of fungal systematic because ITS1 secondary structure contains more information than the usual primary sequence alignment.

ACKNOWLEDGEMENTS

The authors thank The Dean, Faculty of Marine Sciences, CAS in Marine Biology and The

Authorities of Annamalai University for encouragement and facilities.

REFERENCES

1. Caetano-Anolles. 2002. Tracing the evolution of RNA structure in ribosomes. Nucleic acid Res., 30: 2575-2587

2. Billoud, B., M.A. Guerrucci, M. Masselot, and J.S. Deutsch. 2000. Cirripede phylogeny using a novel approach: molecular morphometrics. Mol Biol Evol., 17: 1435-1445

3. Zwieb, C., C. Glotz, and R. Brimacombe. 1981. Secondary structure comparisons between small subunit ribosomal RNA molecules from six different species. Nucleic Acids Res., 9: 3621-3640

4. Schultz, J., S. Maisel, D. Gerlach, T. Muller, and M. Wolf. 2005. A common core of secondary structure of the internal transcribed spacer 2 (ITS2) throughout the Eukaryota. RNA, 11: 361-364

5. Grajales, A., C. Aguilar, and J.A. Sanchez. 2007. Phylogenic reconstruction using secondary structures of Internal Transcribed Spacer 2 (ITS2): finding the molecular and morphological gap in Caribbean gorgonian corals. BMC Evol. Biol., 7: 90

6. Lalev, A.I., P.D. Abeyrathne, and R.N. Nazar. 2000. A RAC protein-binding site in the internal transcribed spacer 2 of pre-rRNA transcripts from *Schizosaccharomyces pombe*. J Mol Biol., 302: 65-77

7. Mai, J.C., and A.W. Coleman. 1997. The internal transcribed spacer 2 exhibits a common secondary structure in green algae and flowering plants. J Mol Evol., 44: 258-271

8. Hershkovitz, M.A., and L.A. Lewis. 1996. Deep-level diagnostic value of the rDNA-ITS region. Mol Biol Evol., 13: 1276-1295

9. Zuker, M. 2003. Mfold web server for nucleic acid folding and hybridization prediction. Nucleic Acid Res., 31: 3406-3415

10. Knudsen, B., and J.J. Hein. 2003. Pfold: RNA secondary structure prediction using stochastic context-free grammars. Nucleic Acids Res., 31: 3423-3428

11. Waugh, A., P. Gendron, R. Altman, J.W. Brown, D. Case, D. Gautheret, S.C. Harvey, N. Leontis, J. Westbrook, E. Westhof, M. Zuker, and F. Major. 2002. RNAML: a standard syntax for exchanging RNA information. RNA, 8: 707-717

12. Wolf, M., M. Achtziger, J. Schultz, T. Dandekar, and T. Muller. 2005. Homology modeling revealed more than 20,000 rRNA internal transcribed spacer 2 (ITS2) secondary structures. RNA, 11: 1616-1623

13. White, T. J., T. Bruns, S. Lee, and J. W. Taylor. 1990. Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. In: *PCR Protocols: A Guide to Methods and Applications*, eds. Innis, M. A., D. H. Gelfand, J. J. Sninsky, and T. J. White. Academic Press, Inc., New York. 315-322

14. Goncalves, S.S., A.M. Stchigel, J.F. Cano, P.C. Godoy-Martinez, A.L. Colombo, and J. Guarro. 2011. *Aspergillus novoparasiticus*: a new clinical species of the section Flavi. Med Mycol., 49: 1-9

15. Zuker, M. 1994. Prediction of RNA secondary structure by energy minimization. Methods Mol Biol., 25: 267-294

16. Prasad, P.K., V. Tandon, D.K. Biswal, L.M. Goswami, and A. Chatterjee. 2009. Phylogenetic reconstruction using secondary structures and sequence motifs of ITS2 rDNA of *Paragonimus westermani* (Kerbert, 1878) Braun, 1899 (Digenea: Paragonimidae) and related species. BMC Genomics, 10(Suppl 3.): S25

Table 1. The pairwise distance of ITS1 sequence in four *Aspergillus* species.

Species	*A. terreus*	*A. ochraceus*	*A. flavus*	*A. melleus*
A. terreus	*****			
A. ochraceus	0.030	*****		
A. flavus	0.184	0.175	******	
A . melleus	0.210	0.172	0.197	******
Overall	0.161			

Table 2. The nucleotide frequency (%) of ITS1 sequence in four *Aspergillus* species

Species	U	C	A	G
A. terreus	22.1	33.7	16.3	27.9
A. ochraceus	22.1	34.9	15.7	27.3
A. flavus	23.3	34.3	17.4	25.0
A. melleus	22.7	32.6	18.6	26.2
Average	22.5	33.9	17.0	26.6

Table 3. ΔG Energy composition in ITS1 secondary structure of four *Aspergillus* species

	Hairpin loop	Bulge loop	Interi or loop	Externa l loop	Multi loop	Stack
A. terreus	26.50	3.60	-6.50	-3.00	2.90	-93.3
A. ochraceus	24.90	3.60	-8.60	-3.50	2.90	-93.1
A. flavus	18.20	3.60	-4.60	-1.70	4.00	-90.8
A. melleus	20.60	2.80	6.00	-4.60	-4.6	-84.6

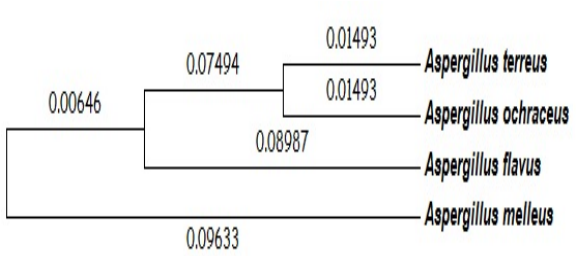

Fig.1. Maximum Parsimony tree of ITS1 sequence of four *Aspergillus* species

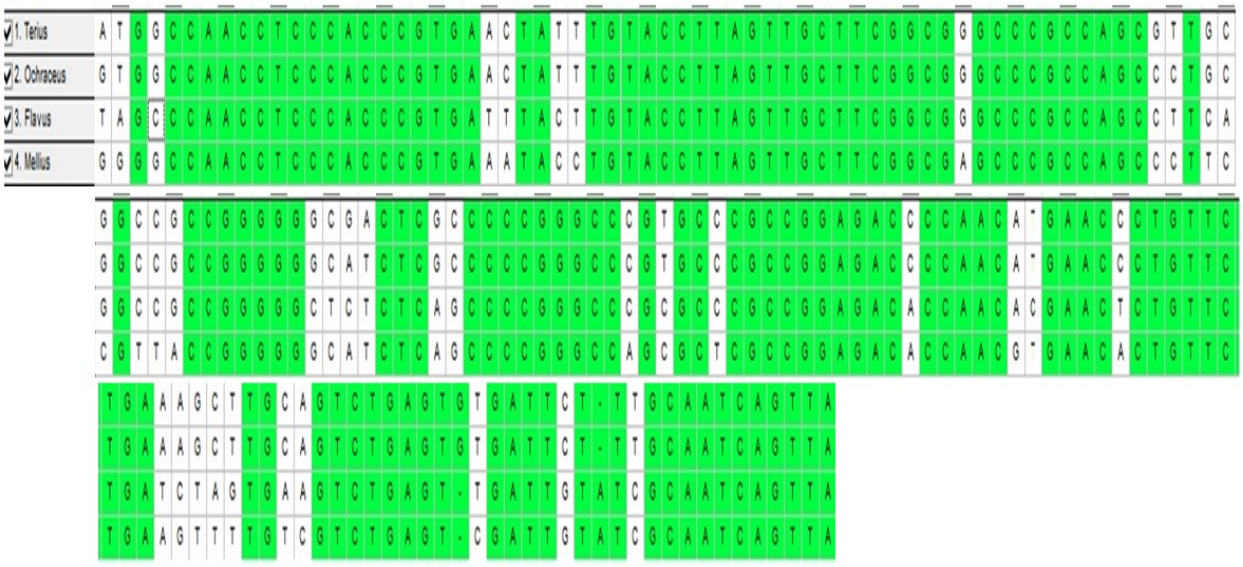

Fig. 2. The conserved bases of ITS1 sequence in four *Aspergillus* species

Fig. 3. Predicted secondary structures and their structure formation enthalpies according to MFOLD of four fungi isolates. (a) *A. terreus*, ΔG=-64.22 kal/mole; (b) *A. ochraceus*, ΔG=-66.12 kal/mole; (c) *A. flavus*, ΔG=-59.12 kal/mole; *A. melleus*, ΔG=-55.00 kal/mole

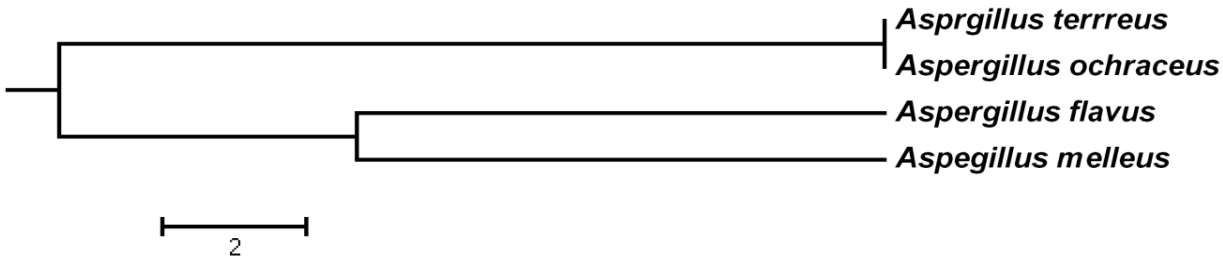

Fig.4. Phylogenetic relationship of four *Aspergillus* species based on the secondary structure of ITS1.