Molecular Phylogeny of Mitochondrial ATP Synthase subunit six of Mushrooms

A Mani, S Sundaram, D Gupta

Abstract

Purpose: Various modern families of mushrooms are artificial due to morphology based characterization, since these characters may be result of parallel evolution and phenotypic plasticity. The aim of this study was to characterize various mushroom families phylogenetically based on their mitochondrial ATP synthase sub unit six.

Methods: For sequence analysis, inferring the phylogenetic tree and evolutionary characterization the protein sequences of ATP synthase sub unit six from 45 different species were used. The tree was drawn by maximum parsimony method while using the bootstrapping as a test of inferred phylogeny.

Result: A phylogenetic tree was constructed from multiple aligned sequences showing bootstrap values on nodes and species codes on leaves. The analysis of data led to a single most parsimonious tree.

Discussion and conclusion: The inferred tree gives an insight into the evolutionary order of various mushroom species based on their ATP synthase sub unit six. Some of the findings regarding the position of few species in the tree really draw attention and may be of great interest to the taxonomists.

INTRODUCTION

Unavailability of whole genomic sequences of species is a major hurdle in determining molecular phylogeny of organisms. Most of the mushroom species are yet to be sequenced for comprehensive genome based molecular phylogeny determination. The selection of a gene for phylogenetic analysis requires its universal presence in all organisms and easily recognizable conservedness in many species. Small rRNA subunit and mitochondrial sequences carry a great deal for inter-species evolutionary characterization.

Production of energy is the very basic feature of mitochondria so it is called the powerhouse of the cell. ATP synthase, a key multisubunit enzyme carries out the synthesis of ATP from ADP and the subunit six plays a key role of a coupling factor during ATP synthesis. It sits in the inner membrane of the mitochondria and F six unit works as coupling factor during ATP synthesis. ATP synthase subunit six is found in all organisms suggesting that they are ancient and conserved. Sequences of ATP synthase subunit six of various mushrooms are available in biological databases. Mitochondrial ATP synthase subunit six sequences, due to their universal presence, conservedness and large scale sequence availability become the most eligible genes for phylogeny determination of mushrooms. These reasons stimulated the studies to explore their phylogenetic relationships as no such study is available for mushrooms. Choosing the appropriate scoring matrix for sequence alignment is also a key factor in tracking the evolutionary history. Matrices prepared by examining the full range of amino acid substitutions in the family of related protein such as BLOSUM matrix, perform better than matrices based on variations in closely related proteins [1]. To find the evolutionary trees, maximum parsimony, distance and maximum likelihood methods are generally used [1]. Maximum Parsimony method predicts the evolutionary tree that minimizes the number of steps required to generate observed variation in the sequences. Though the method is time consuming but the algorithm is guaranteed to find the best tree.

MATERIALS AND METHOD

In order to find ATP synthase subunit six family members, we performed BLAST [3] searches of the protein database at NCBI [4] using Agaricus campestris ATP synthase subunit six gi|4323024|gb|AAD16168.11 amino acid sequence as query. From the hits, 45 sequences each from different mushroom species were selected for further studies. Out of the 45 sequences selected for the phylogenetic analysis, 15
belong to order Agaricales, 11 belong to boletales, 10 belong to aphyllorales, 4 belong to phallales and one from each of the Canthrellales, Geastrales, Thelephorales, Hymenochichaetales and Nidulariales [5].

Figure 1

Table 1: ATP synthase subunit six Protein sequences with their length, NCBI accession code and tree code. * Genus

All the sequences were taken in FASTA format and were given four letter hypothetical codes. The sequences were examined individually and aligned using CLUSTALW [8]. Multiple sequence alignment, phylogenetic and molecular evolutionary analyses were conducted using MEGA version 4 [10]. For pair wise and multiple alignment gap open penalty was -7 and gap extension penalty was -1 [11]. BLOSUM [12] weight matrix was selected for substitution scoring. Hydrophilic gap penalties were used to increase the chances of a gap within a run (5 or more residues) of hydrophilic amino acids; these are likely to be loop or random coil regions where gaps are more common. The aligned sequences were used to create phylogenetic tree. The evolutionary history was inferred using the maximum parsimony method [13]. The phylogenetic analysis was performed on amino acid sequences with amino: p-distance as substitution model [14]. All the characters were given equal weights. The bootstrap consensus tree inferred from 10000 replicates was taken to represent the evolutionary history of the taxa analyzed [14]. Branches corresponding to partitions reproduced in less than 50% bootstrap replicates were collapsed. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (10000 replicates) are shown next to the branches [14]. The most parsimonious tree was obtained using the Close-Neighbor-Interchange algorithm [10] with search level 3 [12] in which the initial trees were obtained with the random addition of sequences (10 replicates). All positions containing gaps and missing data were eliminated from the dataset (Complete deletion option). There were a total of 99 positions in the final dataset, out of which 65 were parsimony informative. Phylogenetic analyses were conducted in MEGA4 [10].

RESULTS AND DISCUSSION

Figure 2

Figure 1: Multiple alignment of sequences obtained by using ClustalW.
By statistical analysis of multiple aligned sequences (Figure 1), it was observed that leucine, isoleucine, serine, phenyl alanine, valine, threonine and alanine are the most frequently present amino acids with frequency percentage of 17.94, 13.31, 10.27, 9.36, 6.76, 6.10 and 6.09 respectively. While within conserved sites, proline, leucine, phenyl alanine, serine, tyrosine, isoleucine and valine are the most frequently present amino acids with frequency percentage of 12.65, 11.01, 10.91, 9.13, 8.84, 7.90 and 6.59 respectively. The multiple aligned sequence was found with number of conserved sites=62, number of variable sites=194, number of parsimony informative sites=157 and number of singleton sites=36. A phylogenetic tree was constructed from multiple aligned sequences by using Maximum Parsimony method showing bootstrap values on nodes. The analysis of data led to a single most parsimonious tree (Figure 2) with tree length 434, consistency index CI=0.476959, retention index RI=0.661699 and rescaled consistency index RCI=0.315603.
This study presents the first evolutionary analysis of the ATP synthase subunit six family of genes across families of mushrooms. The tree shows that mushroom species of different families; Agaricales, Canthrellales, Aphyllorhophores, Geastrales, Phallales and Boletales are grouped together on the basis of ATP synthase subunit six proteins from mitochondria. The group of order Boletales is supported with 45% bootstrap value, the subgroup of family boletaceae is supported with 81% bootstrap value. Node for order Agaricales is supported by 64% bootstrap value, where as the node for order Phallales is supported by 54% bootstrap value. Order Aphyllorhophores is supported by a very low bootstrap value i.e. 14%. Order Boletales is clearly the original source of mushroom diversity, this explains why they are at the base of the branches within our clade. Order Aphyllorhophores has evolved later with phallales and Agaricales in parallel. Agaricales represents the most modern order of mushrooms. Order Phallales has evolved earlier than agaricales. This phylogeny does not seem to be completely consistent with the current view of taxonomy perhaps due to unavailability of complete genomes and lack of proper numerical representation by species from each order and family. This contradiction with respect to the branching order with high bootstrap values in some cases like that of Clavaria zollingeri and Clavulinopsis laeticolor where they both are from order Agaricales and grouped with order Aphyllorhophores, apparently a case of misplaced species in the taxonomy due to morphology based characterization.

ACKNOWLEDGEMENTS

Financial support from DBT_BTISNet scheme of grants for creation of DBT-Bioinformatics Infrastructure Facility by Department of Biotechnology, Ministry of Science and Technology to Prof. D.K. Gupta. Ashutosh Mani is thankful to UGC New Delhi for a research fellowship.

References

Author Information

Ashutosh Mani, M.Sc.
Center of Bioinformatics, Institute for Interdisciplinary Studies, University of Allahabad

Shanthy Sundaram, Ph.D.
Institute for Interdisciplinary Studies, University of Allahabad

Dwijendra K. Gupta, Ph.D.
Center of Bioinformatics, Institute for Interdisciplinary Studies, University of Allahabad