Applications of DNA bar coding in molecular systematics of fungi: A review

Gosavi Mahavir C
Department of Botany, SIES College of Arts, Science & Commerce, Sion (W), Mumbai 400022.
mahavirg@sies.edu.in

ABSTRACT
Fungi are a large, diverse and economically important group of organisms. Estimates of the actual number of fungal species vary widely from 1.5 million to 13.5 million, with fewer than 100,000 now known. Some fungi have relative complex and conspicuous morphologies, but others have very simple morphologies. Identification of fungi isolated from samples in laboratories is mainly based on observing morphology under the microscope and use of various cultural techniques. These conventional methods are very time-consuming and laborious. It usually requires several weeks for the fungi to grow sufficiently on culture media, and the identification processes on fungal phenotypic structure. DNA barcoding techniques have provided standardized, reliable and cost-effective methods for fungal species identification. Gene sequencing and phylogenetic analysis targeting the internal transcribed spacer (ITS) region in the fungal genomes are the most commonly used molecular methods for fungal identification. The ITS region has been heavily used in both molecular methods and ecological studies of fungi, due to its high degree of interspecific variability, conserved primer sites and multiple copy nature in the genome. It has been recently accepted as a suitable marker for barcoding fungi. This approach has immensely being used currently for both phylogenetic reconstruction and species recognition.

Key Words: DNA bar coding, Molecular systematics of fungi, rDNA, ITS.

INTRODUCTION
Fungal taxonomy is an essential part of biological research especially in the context of its ecological and economic implications (Shenoy et al., 2007). Traditional methods for fungal species identification require diagnostic morphological characters and are often limited by the availability of fresh fruiting bodies and local identification resources. About 100,000 fungal species have been identified yet, but it has been estimated that there may be from 1.5 to 13.1 million species. According to Hibbett et al., 2011, over the last decade, about 1200 new species of fungi have been described in each year. At that rate, it may take up to 4000 years to describe all species of fungi using current specimen-based approaches. The use of morphology is extremely useful in several cases to
assign organisms to well-defined categories but reliable diagnostic procedure can be time-consuming and require the expertise of different taxonomists (Pereira et al., 2008).

DNA barcoding techniques have provided standardized, reliable and cost-effective methods for fungal species identification. Gene sequencing and phylogenetic analysis targeting the internal transcribed spacer (ITS) region in the fungal genomes are the most commonly used molecular methods for fungal identification (Das & Deb, 2015). The ITS region is now perhaps the most widely sequenced DNA region in fungi. It has typically been most useful for molecular systematics at the species level, and even within species. Because of its higher degree of variation than other genic regions of rDNA that allow selective amplification of fungal sequences (Bruns & Gardes, 1993).

**DNA Barcoding:**

Taxonomy is the science of classifying living organisms according to their morphology. Today identification of organisms has become highly critical as effects of global warming and accelerating habitat destruction. Classical taxonomy falls short in this race to catalog biological diversity before it disappears http://www.dnabarcoding101.org/introduction.html.

DNA barcoding is a method for identifying living organisms to species. It makes use of a short, highly variably regions of the genome that evolves fast enough to differ between closely related species. DNA barcodes function as molecular identifiers for each species, in the same way as the machine-readable black-and-white barcodes are used in the retail industry to identify commercial products. Using Polymerase Chain Reaction (PCR) to amplify a targeted genetic marker and DNA sequencing to determine the content of that sequence, DNA barcoding allows for the comparison of an informative sequence from an unknown fungal sample against a database of identified sequences (Harrington et al., 2014).

In effect, DNA barcoding in its modern form was popularized in a paper by Hebert et al. (2003), who proposed to use the mitochondrial gene COI (cytochrome c oxidase subunit) as the standard barcode for all animals.

A region of the chloroplast gene rbcL – RubisCo large subunit – is used for barcoding plants (Chase et al., 2005). The most abundant protein on earth, RubisCo (Ribulose-1,5-bisphosphate carboxylase oxygenase) catalyzes the first step of carbon fixation. A region of the mitochondrial gene COI is used for barcoding animals. Cytochrome c oxidase is involved in the electron transport phase of respiration. COI in fungi is difficult to amplify, insufficiently variable, and some fungal groups lack mitochondria. Instead, the nuclear internal transcribed spacer (ITS), a variable region that surrounds the 5.8s ribosomal RNA gene, is targeted. Like organelle genes, there are many copies of ITS per genome, and the variability in fungi allows for their identification. During the last 15 years the ITS of nuclear DNA has been used as a target for analyzing fungal diversity in environmental samples, and has recently been selected as the standard marker for fungal DNA barcoding (Bellemain et al., 2010).

**Internal transcribed spacer (ITS) in DNA barcoding of Fungi:**

The nuclear ribosomal coding cistron (rDNA) has been widely utilized for detecting variation among isolates of a fungus. The rDNA is composed of tandemly repeated units, each unit being composed of 5S, 25S and 18S rDNA. Two noncoding regions exits in each repeat: The internal transcribed spacer (ITS 1 and ITS 2) and and intergenic non-transcribed spacer (IGS). It has typically been most useful for molecular systematics at the species level, and even within species. Because of its higher degree of variation than other genic regions of rDNA (for small- and large-subunit rRNA), variation among individual rDNA repeats can sometimes be observed within both the ITS and IGS regions. There are two ITS's in eukaryotes; ITS1 is located between 18S and 5.8S rRNA genes, while ITS2 is between 5.8S and 25S rRNA genes (Lafontaine & Tollervey, 2001). ITS has proven especially useful for elucidating relationships among species and closely related genera in clinically important yeast species (Chen et al., 2011). The ITS region is the most widely sequenced DNA region in molecular ecology of fungi (Peay et al., 2008). Schoch et al. (2012), proposed that among the regions of the ribosomal cistron, ITS region has the highest probability of successful identification for the broadest range of fungi, with the most clearly defined barcode gap between inter- and intraspecific variation. The ITS region has been proposed as the prime fungal barcode species identification (http://www.allfungi.com/its-barcode.php). The ITS region of fungi varies in length from 450bp to 750bp. Approximately 172,000 full-
length fungal ITS sequences are available in Genbank. ITS marker exists in multiple copies in most fungal cells and is retrievable by relatively strong primers with an established record of reliability. From some latest review papers on fungal barcoding, it is observed that in ascomycetes, ITS had the most resolving power for species discrimination (Das & Deb, 2015).

DNA barcoding procedure
The process from field samples to species abundance data involves a long series of steps, from sampling via laboratory handling to bioinformatics treatment. At each step, there is a risk of losing and distorting information.

Lindahl et al., (2013) has reported a standard procedure for DNA Barcoding using ITS region in fungal identification. It involves following steps.

1. Collection of Samples:
Collection of fungal samples vary according to season, mycelia, growth pattern and fruiting bodies. For fungi, use of fruiting bodies is always best since it is easier to obtain DNA from fruiting bodies than mycelia (http://www.dnabarcoding101.org). Multiple samples from same should be collected that appears similar, avoiding the contamination by other fungi. Fresh samples work well for DNA isolation (Kelly et al., 2011).

2. Handling of samples:
Freeze-drying at -20°C enables long-term storage at room temperature, and may also aid later sample homogenization. It also restricts sporulation and rapid growth of opportunists (Lindahl et al., 2013).

3. Homogenization and sub sampling:
Fungal tissue ~10–20 mg should be obtained from the sample. If working with more than one sample, care should be taken not to cross contaminate specimens (Lindahl et al., 2013 & http://www.dnabarcoding101.org). The samples are normally homogenized using appropriate lysis mixtures followed by storing at -20°C.

4. Extraction and purification of DNA:
DNA is extracted from the fungi using standard protocols (http://www.dnabarcoding101.org). Extraction should yield high and uniform amounts of DNA. To achieve this, same extraction protocol should ideally be used for all samples (Tedersoo et al., 2010). The DNA precipitation may have to be further purified by binding of DNA to a silica matrix (Lindahl et al., 2013).

5. Markers and primers: The ideal marker for fungal community studies should have primer sites that are shared by all fungi, be of appropriate length for efficient amplification and sequencing, have high interspecific variation but low intraspecific variation, and be possible to align across all fungi. This is the main reason why the ITS region has been particularly attractive for mycologists (Gazis et al., 2011).

6. Multiplying DNA by PCR: DNA is extracted from the fungi and the barcode portion of the ITS gene is amplified by PCR. After thermal cycling, the amplified DNA is stored on ice or at -20 °C. Before sequencing, PCR products from different samples are mixed in equimolar proportion, so that the DNA sequence output is evenly distributed across all samples. The PCR products have to be purified, to remove primers and short DNA fragments (http://www.dnabarcoding101.org).

7. Analysing PCR Products by Gel Electrophoresis & Sequence alignment: The amplified sequence (amplicon) is submitted for sequencing in one or both directions (Das & Deb, 2015). The PCR samples are subjected to gel electrophoresis and the bands in each lane of the gel are interpreted (http://www.dnabarcoding101.org). ITS primers amplify differently sized products that migrate to different positions on the gel. DNA sequencing of ITS amplicon is required to determine the nucleotide sequence that constitutes the DNA barcode. A single, good-quality barcode from the forward strand is sufficient to identify an organism.

8. Bioinformatics analysis: The sequencing results are then used to search a DNA database. A close match quickly identifies a species that is already represented in the database (Das & Deb, 2015). Novel DNA barcodes can be submitted to GenBank (www.ncbi.nlm.nih.gov). For the fungi primers, the hits should all be to the nuclear internal transcribed spacer of the 5.8s ribosomal RNA gene.

9. Data interpretation: Computer based analysis are applicable for DNA barcoding (Das & Deb, 2015). Barcode Sequences were deposited in GenBank and
The genomic organization of ribosomal RNA (rRNA) genes - responsible for the synthesis of RNA species (the core of ribosomes). Almost all eukaryotes have several copies of each rDNA cluster organized in tandem repeats. In this case, each cluster contains the 18S, 5.8S and 25/28S rRNAs, while the 5S gene is present in separate repeat arrays in the majority of eukaryotes. Where ITS=internal transcribed spacers, ETS=external transcribed spacers.


---

**Fig: Ribosomal RNA genes**

The genomic organization of ribosomal RNA (rRNA) genes - responsible for the synthesis of RNA species (the core of ribosomes). Almost all eukaryotes have several copies of each rDNA cluster organized in tandem repeats. In this case, each cluster contains the 18S, 5.8S and 25/28S rRNAs, while the 5S gene is present in separate repeat arrays in the majority of eukaryotes. Where ITS=internal transcribed spacers, ETS=external transcribed spacers.


---

**Reports of identification of fungal species using DNA barcoding:**

Robideau et al., (2011) had reported use of DNA barcoding in species identification of oomycetes especially in case of genera like Phytophthora and Pythium, they tried DNA barcoding for 1205 isolates representing 23 genera.

Seena, et al., (2010) examined the suitability of ITS1-5.8S-ITS2 rRNA gene region to identify aquatic hyphomycetes, by sequencing and comparing these regions in 94 fungal isolates belonging to 19 species collected in Portuguese streams with different environmental conditions.

Khaund & Joshi (2014) had characterized 10 species of wild mushrooms of Meghalaya, viz. Gomphus floccosus, Lactarius deliciosus, Lactarius volemus, Cantharellus cibarius, Tricholoma viridiolivaceum, Inocybe aff. sphaerospora, Laccaria vinaceaovellanea, Albatrellus ellisi, Ramaria maculatipes and Clavulinia cristata. They also reported that the final species identity generated by the ITS marker matched more accurately with the morphological characteristics/appearance of the specimens indicating the ITS region as a reliable barcode for identifying wild edible mushrooms. Irinyi et al. (2016) depicted standardization of ITS sequence based identification of fungi causing infections in humans and animals.

Khodadadi et al. (2014), in their studies tried to identify the clinically rare yeast isolates. 49 out of 855 (5.7%) yeast isolates which formerly remained unidentified by PCR-RFLP method were subjected to sequence analysis of ITS regions. These species include: Hanseniaspora uvarum, Saccharomyces cerevisiae, Sporidiobolus salmonicolor, Pichia fabianii, Pichia fermentans, Candida famata, Candida inconspicua, Candida maqnoniae, Candida guillermondii, Candida kefyr, Candida rugosa, Candida lusitaniae, Candida orthopsisls, and Candida viswanathii.

Geiser et al. (2007) published complete genome sequences of eight Aspergillus species (Em. nidulans, A. oryzae, A. fumigatus, N. fischeri, A. terreus, A. clavatus, A. niger, A. flavus) based on DNA barcoding.

Wang et al. (2016), evaluated six different loci in their study as potential barcodes in Chaetomium indicated that the 28S large subunit (LSU) nrDNA and the internal transcribed spacer regions and intervening 5.8S nrDNA (ITS) gene regions were unreliable to resolve species, whereas β-tubulin (tub2) and RNA polymerase II second largest subunit (rpb2) showed the greatest promise as DNA barcodes for differentiating Chaetomium species.

---

**REFERENCES**


References


Web References:

http://www.dnabarcoding101.org/introduction.html

© 2016 | Published by IJLSCI

www.ijlsci.in

Int. J. of Life Sciences, Special Issue A7; December, 2015 |115