Journal of Molecular Biomarkers & Diagnosis

Srivastava et al., J Mol Biomark Diagn 2014, 5:3
DOI: 10.4172/2155-9929.1000172

Review Article Open Access

To Develop Strain Specific Molecular Marker for Easy and Proper Identification of Fungal Species Based on Molecular Characters: A Review

Mukesh Srivastava*, Mohammad Shahid, Anuradha Singh, Vipul Kumar, Sonika Pandey and Y K Srivastava

Biocontrol Laboratory, Department of Plant Pathology, Chandra Shekhar Azad University of Agriculture & Technology, Kanpur, Uttar Pradesh, India

Abstract

The sequencing of fungal genomes is advancing at breakneck-speed, producing voluminous amounts of data. Within the next five years, it is possible that over a couple thousand genomes, representing every major fungal family will be completed and available to the scientific community. In order for this data to have a truly transformative effect on mycological and other research, however, several factors need to be addressed. These include; (1) the establishment of user friendly platforms for examining, sorting, and sifting through the genomes, (2) integration, or at least cross-communication, between the various databases that house the genomic data, and (3) investment in community resources that can act as repositories for and provide materials to researchers, i.e. strains, clones, plasmids, etc. The frameworks for some these needs, e.g. the materials available from the Fungal Genetics Stock Center (FGSC, University of Missouri), are already established and should be reinforced, whereas for others, e.g. data accessibility, the sooner that a plan can be implemented the better. The Fungal Kingdom is considered to contribute greater than 15% of the species richness found in the major groups of organisms .This study is a reflection of the usefulness of sequence analysis of the 28S ribosomal RNA gene in identifying fungal as well as determining fungal diversity. Various techniques that are based on utilizing the 28S rRNA have been discussed. Of critical importance is the manner in which massively parallel sequencing was exploited to correct the under representation of fungal species in compilations of fungal hat were drawn using traditional methods of surveying fungal species from ecosystems.

Keywords: Fungal identification; 28S ribosomal RNA gene; DNA

Introduction

The Fungal Kingdom is considered to contribute greater than 15% of the species richness found in the major groups of organisms [1]. Displaying a wide range of facultative to obligate saprophytic, symbiotic, and pathogenic lifestyles, fungi have colonized all terrestrial ecosystems. Often noticeable in soils only in their fruiting body form, a largely hidden under- or inaccessible world is populated by fungal organisms performing ecological functions that have significant impacts on plant, animal, and even other microbial communities.

Fungal contenders exist for our planet's largest and amongst the oldest organisms, i.e. *Armillaria solidipes* also known as the honey mushroom found in the Malheur National Forest in Oregon, estimated at being between 2000-9,000 years old, spans over several thousand acres. In all, the current total number of fungal species is estimated at over 1.5 million with a little over 100,000 described to date. The described fungi have been placed into approximately 140 orders and over 550 families.

Due to this diversity and their biological and biotechnological importance the sequencing of fungal genomes is well underway [2]. Including yeasts, the total number of fungal genomes sequenced or near completion to date is close to 500 (Genomes Online Database, GOLD, http://www.genomesonline.org/cgi-bin/GOLD/bin/gold.cgi). Although the available fungal genomes are perhaps somewhat biased towards fungi of medical importance, several new whole-genome sequencing initiatives are likely to provide robust and more taxonomically diverse datasets. The Fungal Genome Initiative at the Broad Institute currently houses over 50 fungal genomes, mainly from the Ascomycota, and includes model organisms such as *Neurospora* and *Aspergillus* species, but also extends to the genomes of several basidiomycetes, chytrids, and at least one mucormycete (http://www.broadinstitute.org/scientific-community/science/projects/fungal-genome%20initiative/fungal-genome-initiative). The Mycorrhizal Genome Initiative, designed to sequence and analyze the genomes of ascomycetes and basidiomycetes that function within the context of symbiosis in woody shrubs and trees, has selected 28 species for whole genome sequencing (http://mycor.nancy.inra.fr/IMGC/MycoGenomes/index.html). These data, coupled to the sequencing of a "suite of fungal decayers", including brown and white rot fungi will provide the genomic framework for understanding the biology of these organisms and potentially exploiting them in applications ranging from reforestation and maintenance of ecological biodiversity to bioenergy production.

The Oparin Ocean Scenario states that life arose spontaneously from an "oceanic soup" of chemical elements. In another speculative attempt Charles Darwin formulated a theory of a "warm little pond" in which life began. These two theories (Oparin's and Darwin's) and any scientific theory are not sufficient to explain creation. Whatever was the spark to life and specifically microbial life, human beings have always seeked to trace the present microbial diversity back to the most primitive form of microbial life. Scientific evidence has led us to believe that prokaryotes were first to arise, followed closely by eukaryotes. We learnt that these two groups were later subdivided. Darwin also made various observations from plants to bees and theorized that plants and animals were adapted in the environments

***Corresponding author:** Mukesh Srivastava, Biocontrol Laboratory, Department of Plant Pathology, C. S. A. University of Agriculture & Technology, Kanpur-208002, Uttar Pradesh, India, Tel: 9935240430; E-mail: mukeshcsau@rediffmail.com

Received January 24, 2014; **Accepted** February 04, 2014; **Published** February 27, 2014

Citation: Srivastava M, Shahid M, Singh A, Kumar V, Pandey S, et al. (2014) To Develop Strain Specific Molecular Marker for Easy and Proper Identification of Fungal Species Based on Molecular Characters: A Review. J Mol Biomark Diagn 5: 172. doi:10.4172/2155-9929.1000172

in which they live. Darwin believed that these passed their diversity to their offspring which would then evolve over time. Because of its richness in morphological traits which traditionally formed the basis for classification of organisms, macroscopic life was obviously easier to study than microbial life. In addition to the tedious processes of cultivating microorganisms, microbial morphologies were too simply to provide sufficient information for classification. This difficulty resulted in lack of enthusiasm by students to study microbiology also causing doubt about whether it was worth exploring these unknowingly "riches" of the microbial world.

However, some early microbiologists steadfastly studied microorganisms just as botanists and zoologists studied animal and plants respectively. Microorganisms are probably the most diverse group of organisms but their diversity remains poorly understood. Advent of tools like microscopy helped increased our knowledge of microbial diversity as microbes could then be classified as, among other groups, fungi, bacteria and viruses which were themselves very diverse. Among earlier milestones were the classification of the bacterial genus *Pseudomonas* and the compilation of the Bergey's Manual of Systematic Bacteriology which remained the blueprint of bacterial classification for decades after its adoption. The 28S rRNA gene was a single criterion on which order would be created to reflect fungal diversity. Further study of diversity within microbial groups was also aided by other molecular-based techniques which traversed beyond the limitations of studying morphology to classify microorganisms. However, the 28S rRNA gene remained the primary reference for fungal classification. Various laboratories in the world have generated and deposited 28S rRNA gene sequences in web-based databases. So rapid was this process that DNA sequence database have over a short period been flooded with 28S rRNA gene sequences to afford researchers the opportunity to search these databases to classify fungi. Sequence databases currently contain over a million full-length 28S rRNA sequences representing a broad phylogenetic spectrum that are a useful benchmark for identifying fungal taxa from diverse samples [3].

The purpose of this review is to rearrange all information linked to fungal classification using sequence analysis of the 28S rRNA gene. Collation of the information herein presented has up to now not been done to reflect an evolutionary process of classifying fungal. The author reflected on the use of the 28S rRNA as a tool for studying fungal diversity. The look at the 28S rRNA cuts across various techniques that are used to study fungal diversity.

The 28S rRNA gene

Cytochrome c and ribulose bisphosphate carboxylase are just two examples that have been used for phylogeny for organisms in general but fell away as favourites because not all organisms have these macromolecules and therefore comparison of distant taxa was not possible. The 28S rRNA gene for prokaryotes and the 28S rRNA for eukaryotes have become the new measure for taxonomy. For the purposes of classification up to species level, the 28S rRNA gene is the primary reference for fungi. For strain differentiation within a species other measures beyond the scope of this review may be necessary. Among the three types of RNA that are contained by fungal ribosomes, namely, 28S rRNA, 5S rRNA and 23S rRNA, the 28S rRNA is regarded as the most convenient and reliable measure.

Molecular approaches for detecting and classifying fungi rely on the PCR amplification and sequence analysis of the 28S rRNA gene. The 28S rRNA gene is a suitable parameter for fungal classification because the 28S rRNA gene is universal among fungi and is conserved but has sufficient variation to distinguish between taxa. One PCR primer pair can target the 28S rRNA gene from a wide range fungal species. The supreme advantage of the 28S rRNA gene-based analysis is that it bypasses culturing of bacteria as PCR detection is done on DNA extracted from crude samples. Because of this direct amplification of the 28S rRNA gene from DNA samples, it is possible to detect unculturable fungus which is estimated to exceed 99% of microorganisms' observable in nature [4]. Due to these culture-independent surveys, the number of identifiable bacteria has increased drastically in recent times. In these studies the 28S rRNA gene is PCR-amplified from a DNA sample, the PCR fragment is sequenced, the sequence is queried on a database like the NCBI, sequence hits are pooled from the database, and the sequences are used for phylogenetic analysis. The gene databases have facilitated these studies and as a result fungal diversity surveyors are able to share information eliminating the need for repeated surveys. Various regions of the 28S rRNA gene can be explored to study variation among bacterial phylogenies. Targeting specific regions of the 28S rRNA gene approximately 100 bp instead of longer reads can provide sufficient information for classification [5]. Various PCR primers are available to target these regions [6]. When analyzing metagenomic samples, only a portion (about 500 nucleotides) of the 28S rRNA gene is sufficient for phylogenetic allocation of unknown fungi into respective taxa. This makes massive parallel sequencing convenient for gross analysis of metagenomic samples because of the short read lengths produced by the high-throughput sequencing instruments. However, a longer sequence that covers almost the entire 28S rRNA may be preferable.

28S rRNA Gene Databases

Sequence information of the 28S rRNA gene sequence is being kept to allow fungal investigators the opportunity to undertake comparative studies for classifying fungi. The Ribosomal Database Project (RDP) was established to create a credible bank for 28S rRNA gene information [7]. The RDP is both a collection of sequence information of the 28S rRNA gene as well as a software suit with tools to organise raw sequences into alignments, annotate sequences and organise information to provide a phylogenetic analysis of the data for understanding fungal diversity. The RDP has regular updates of its collection as well as a user account system for users to integrate their data with the existing data. Quality of deposited 28S rRNA gene sequences is assured using the Pintail chimera detection program. Analysis tools of the RDA include the Hierarchy Browser which allows users to browse the dataset by taxonomy, publication or sequenced genome. The RDP also has a SeqCart into which a user can add selected sequences so that analysis can be performed only on the sequences in the cart. The RDP uses the naïve Bayesian classifier to rapidly and precisely classify bacterial 28S rDNA sequences into phylogenies and it is suitable for both single rDNA reads and libraries of thousands of reads generated by high-throughput genome sequencing platforms [8]. This classifier is trained on known and continuously updated sequences of the 28S rRNA gene and a few other sequences of value to fungal diversity. From each query a base word is selected and probed on the database to determine joint probability of assignment to a genus based on the naïve Bayesian assumption using a multiple bootstrap repetition. The RDP Classifier also checks the orientation of the sequences being matched to ensure that reporting is only of those sequences that have the same orientation for both full-length and partial 28S rRNA gene fragments. The RDP Library Compare compares differences between taxa from two sample libraries. Sequence Match of the RDP finds sequences in the database with the closest identity to the query sequence. The query sequence can be matched with sequences either in SeqCart, or with sequences on its page or on sequences on the web through the RDP web service

interface. The RDP Probe Match allows rapid searching of the database to check primer and probe coverage and specificity of probes up to over 60 bases in length. For near-matching strings Probe Match includes ambiguity codons in the search output. The RDP Tree Builder allows users to draw phylogenetic trees with bootstrap confidence values. Another important tool of the RDP is RDP Taxomatic with which users can create interactive heatmaps based on pairwise distances between 28S rRNA gene sequences. RDP Web Services Interfaces allows programmers to link RDP tools with other web interfaces.

The Basic Local Alignment Search Tool (BLAST) of the NCBI is a widely used bioinformatics program. BLAST has wide applications from nucleotide to protein sequences. A query 28S rRNA gene sequence can be compared with all entries in the database to match it with sequences that resemble it above a defined threshold value and emphasize speed in the search for optimal alignment from a huge dataset. Hits of 28S rRNA gene sequences come as the BLAST output and they can be arranged according to coverage or similarity. Hits can be retrieved in FASTA formats which can be used to draw a phylogenetic tree.

The traditional 28S-cloning-and-sequencing approach, provides an in-depth analysis of the richness of fungal species within a sample, however, it remained laborious, costly and would mostly only detect just about a hundred sequences per sample. This warranted alternative assaying techniques, namely terminal Restriction Fragment Length Polymorphisms (t-RFLP) [9-11] and the Denaturing Gradient Gel Electrophoresis (DGGE). These techniques remained relevant and provided useful information for classifying fungi.

Terminal-restriction fragment length polymorphism

The terminal-restriction fragment length polymorphism technique is used to study fungal diversity based on the variation of the 28S rRNA gene. The 28S rRNA gene is PCR-amplified with either one or both of the PCR primers fluorescently labelled. The PCR products are digested with a restriction enzyme. The fragments are separated by electrophoresis and the fragment lengths indicate the diversity that is in the analyzed sample.

Temperature gradient gel electrophoresis/ denaturing gradient gel electrophoresis

These methods exploit the different denaturing properties of double-stranded 28S rRNA gene copies with sequence differences. Small sequence differences between 28S rRNA gene strands can be detected by these methods. Base differences in the 28S rRNA gene determine the stage of electrophoresis at which the double-stranded 28S rRNA gene melts in a gel with either a temperature or a denaturing chemical gradient. Differences in these denaturing differences can correspond to 28S rRNA gene variation.

Metagenomic Approaches

In addition, several specialized platforms already exist and can be used as templates for further expansion. The Saccharomyces Genome Database (http://www.yeastgenome.org/) is one of the most comprehensive of such efforts and includes a trove of tools for examining the data. These include features for querying and analyzing data along biochemical, developmental, and signaling pathways useful to a wide range of researchers without requiring detailed bioinformatic expertise. Extensive (literature) referencing for every gene that has been examined in terms of phenotype, expression, and/or biochemical characterization is invaluable. Additional databases particularly for model organisms also exist [12]. If the identified orthologous genes of the other extant fungal genomes could somehow be linked to these data, a significant contribution to facilitating annotation and research efforts would already be made. Another model to examine is The Arabidopsis Information Resource (TAIR, http://www.arabidopsis.org/), which provides a one-stop resource for the genome sequence, gene structure, product, and expression information, as well as DNA, seed stocks, plasmids, and references dealing with this model plant organism.

Despite their utility, significant efforts will be needed to apply these scaffolds to such ambitious efforts as the 1000 Fungal Genome Project. Whether it would be better to design new interactive platforms essentially *de novo,* with smart features akin to gene-butlers with anticipatory algorithms and on-demand reference linking, or to build upon already established ones is a discussion that should be held given the vast amounts of data that are coming in the pipeline. For example, obtaining the sequence data itself can progress quite rapidly; however, high quality annotation of the genomes to a useful level is likely to be a significant bottleneck. Regarding this latter step, the ongoing development of a system or technology (referred to as The SEED, http://www.theseed.org/wiki/Home_of_the_SEED) for high-throughput high-quality annotation of genomes may be a model worth applying. The idea of the SEED project is to have researchers (not necessarily in bioinformatics) annotate single (or multiple) subsystems within their expertise over a collection of genomes rather than a single genome. This is in contrast to current approaches in which teams, which often have little expertise in the majority of the specific genes that they annotate, analyze whole genomes one at a time; a procedure that can result in significant levels of erroneous annotation. Unless a SEED-like project is coupled to the sequencing efforts, the annotation of the genomes may be cursory and hence limited in use to the wider scientific community.

Five years from now we may well have at our disposal the genomic sequences of several thousand fungi, across all the major families of the Fungal Kingdom. The extent to which these data can revolutionize the field will depend less upon the sequences themselves but more on the tools available for researchers to mine these genomes. The development of an integrated, learning-enabled, expert driven annotation and referenced web-based system, all with a user friendly, intuitive interface and linked to community available repository services, may seem to be a rather tall order. However, investment in building such an infrastructure would be critical for managing the potential data overload and stepping beyond the fungal genomes to yield resources capable of benefiting and amplifying both basic and applied research.

Metagenomics began in the 1980s as the idea of gross extraction of DNA from a sample with a mixture of nucleic acid [13]. PCR was used to selectively amplify the target 28S rRNA gene which would then be cloned and sequenced to reveal the identity of the fungal species from which the 28S rRNA gene came. Sequences representing 28S rRNA gene fragments could be aligned and used to draw a phylogenetic tree. Such a system allowed even relationships between cultured and uncultured fungi to be determined phylogenetically. Although the idea of metagenomics began in the 1980s, it was not until 1998 that Handelsman referred to this gross genomic sample as a “metagenome”

Metagenomic studies inundate data storage bins with massive amounts of data which can be hard to manipulate. This is especially true if the metagenomic analysis is not targeted to a specific gene. For the purposes of fungal classification, the 28S rRNA gene is a suitable target for analysis. Targeting only this gene helps to scale down data generation in metagenomic studies using high-throughput sequencing platforms. The 28S rRNA gene is PCR-amplified from the metagenome and barcoding allows a mixture with hundreds of samples to be combined in one sequencing run.

High-throughput genomic DNA sequencing from gross DNA samples is able to provide a rapid fungal survey with a broader range of captured species. Using high-throughput genomic sequencing platforms, metagenomics has become popular in surveys of plant-associated fungi, animal-associated bacteria, human bacteria and environmental bacteria. Uses of this approach have gained wide applications in the search for novel bacterial species and novel enzymes.

Data Analysis

Metagenomic data may be analyzed using the phylotyping approach. In phylotyping sequences are grouped into bins based on their similarity with reference entries in the database. All the 28S rRNA gene sequences are given identities based on previously cultured and classified bacteria. This method of classification is regarded as stable and reliable. Another approach, the Operational Taxonomic Unit (OTU) approach, does not restrict classification based on already created bins but groups the 28S rRNA gene sequences based on their similarity with each other. The problem with the OTU approach is that cutoffs for assigning a group to the same taxonomic rank e.g. species is not easy to set and therefore placement of 28S rRNA gene sequences remains problematic and the computer algorithms for clustering sequences are slow and demand high amount of computer memory [14-16].

High-throughput genomic sequencing platforms

Sequencing DNA is an attempt to decipher the permutation of its nucleotides. Early attempts to establish a sequencing system included the Maxam-Gilbert sequencing and chain termination methods. The Maxam-Gilbert method or the chemical sequencing method was developed by Walter-Gilbert and Alan Maxam in 1977. Purified DNA could be used directly in the sequencing reaction. End-labelled DNA is cleaved at specific bases using specific reagents. The labelled fragments cleave using dimethyl sulphate which selectively attacks purines, and hydrazine which selectively attacks pyrimidines. The fragments are electrophoresed in a high resolution polyacrylamide gel to deduce the sequence of the DNA molecule. The chain termination method requires a single strand of the DNA molecule whose sequence must be determined, a DNA primer, a DNA polymerase, labelled nucleotides, nucleotides that serve as terminators because of their lack of 3'-OH group which are required to form phosphodiester bonds between two nucleotides. Based on the sequences of the template DNA, DNA strands of various lengths are synthesized. The newly synthesized strands are denatured by heat and run in a polyacrylamide gel. From the various size fragments on the gel, sequence of the DNA molecule can be deciphered.

Studies to assess plant-associated fungi

Analysis of fungal communities associated with the plant root zone benefited tremendously from utilizing sequence analysis of the 28S rRNA gene. No longer are soil microbiologists limited to culturing of soil fungi and identifying the cultured microorganisms. Their scope of bacteria to be studied also extends to unculturable bacteria which were otherwise not detected by traditional means. By PCR using universal primers, the 28S rRNA gene can be amplified from a variety of fungal taxa in the targeted sample. Because of this large-scale generation of sequencing data, soil biologists are able to establish associations between the fungi, the soil and the rhizosphere rapidly. Such knowledge can speed up efforts to manipulate the ecology of the soil for the benefit of the plant.

In eukaryotes, the genes encoding ribosomal RNAs are organized in arrays which contain repetitive transcriptional units involving 16 – 18S, 5.8S and 23 – 18S rRNAs, two transcribed intergenic spacers ITS1 and ITS2 and two external spacer sequences (5′and 3′ ETS). These units are transcribed by RNA polymerase I and separated by non-transcribed Intergenic Spacers (IGS) as represented in Figure 1. The product of RNA polymerase I is processed in the nucleolus, where ITS1 and ITS2 are excised and three types of rRNAs are produced. In eukaryotic genomes, the ITS regions vary greatly in size and sequence. In *S. cerevisiae* the ITS1 spans 361 bp and ITS2 is 232 bp long, the Polymerase Chain Reaction (PCR) and subsequent analysis of amplified rDNA using restriction endonucleases were employed in different studies to achieve efficient interspecies discrimination in medical and food mycology. The ITS regions have important biological meaning in rRNA processing. The structures of analyzed ITS1 and ITS2 contain four or three helical arms. The changes in size and sequence of these regions are then biologically permissible as long as they do not disturb the formation of secondary structures which facilitate the rRNA processing. It thus presents a simple method for determination of inter- and intraspecies variability in fungal isolates.

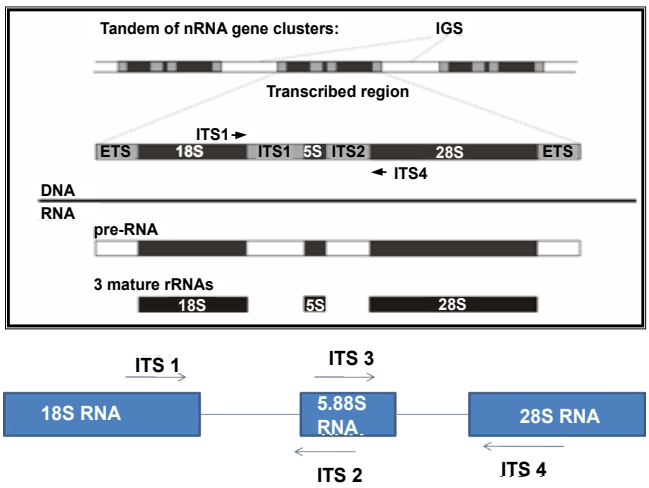

Figure 1: (a) Location of Internal Transcribed Spacer sequence 1 & 2 (b) the position of primers for PCR based amplification of the region.

Molecular phylogenetic analyses of biological control strains of *Trichoderma* (*Ascomycetes*, *Hypocreales*) that have warted conidia are traditionally identified as *T. viride*, the type species of *Trichoderma*. However, two morphologically distinct types of conidial warts (I and II) have been found. Because each type corresponds to a unique mitochondrial DNA pattern, it has been questioned whether *T. viride* comprises more than one species. Combined molecular data (sequences of the internal transcribed spacer 1 ITS-1 and ITS-2 regions and part of the 18S rRNA gene along with results of restriction fragment length polymorphism analysis of the endochitinase gene and PCR fingerprinting), morphology, physiology, and colony characteristics distinguish type I and type II as different species. Type I corresponds to "true" *T. viride*, the anamorphic of *Hypocrea rufa*. Type II represents a new species, *T. asperellum*, which is, in terms of molecular characteristics, close to the neotype of *T. hamatum*.

Analysis of ITS1-5.8S-ITS2 region of the cDNA showed that approximate 600 bp and size variation was observed. Restriction analysis of this region showed that inter and intra -specific polymorphism [17].

Perhaps the most significant near-term development in fungal genomics is the recent DOE-JGI funding of the 1000 Fungal Genomes

Project, which seeks to sequence and assemble a thousand fungal genomes within the next five years (http://1000.fungalgenomes.org/home/). The objectives of this initiative are to sequence at least two representatives from every family or family-level clade of Fungi, thus providing an exceptionally diverse sampling of fungi. These data would benefit scientists across a wide spectrum of research and allow for a level of comparative genomics far deeper than that possible today. Additional international as well as single fungal genome projects exist. At least 23 fungal genomes have been completed or are in progress in various Chinese laboratories and several of these have already been published [18]. It is unclear, however, whether these data (including a number of individual fungal genomes sequenced within the US and throughout the world) have been integrated into any of the databases listed above, a critical omission that needs to be addressed. The impacts of the availability of fungal genomes will depend on their accessibility to the larger scientific community, particularly to those that may not have in-depth expertise in bioinformatics. The National Center for Biotechnology Information (NCBI, http://www.ncbi.nlm.nih.gov/) continues to be one of the major veins by which researchers can access fungal genomic data, and many of the tools available from NCBI address some of the desired functions listed below. However, not all fungal genomes appear to be equally accessible or available and no facile mechanism exists for corrections and annotations. Further investment in the development of platforms that have increased user-friendly interfaces is essential. This should include not only the ability to BLAST the genomes and/or locate a gene, but also to easily perform a range of other functions such as build (genomically) localized comparative analyses between organisms, find neighbouring genes, identify potential regulatory regions, and link genes within biochemical and functional frameworks. An example of such a program, which should be incorporated especially as part of any largescale fungal genome sequencing efforts, is the Comparative Fungal Genomics Platform (CFGP, http://cfgp.riceblast.snu.ac.kr/main.php). The tools available include easy to use links for analyzing, sorting, and storing genes of interest as well as links to DNA, RNA, and protein predictions and functions. A careful survey of the prospective needs of scientists before embarking on building upon these tools would seem to be prudent.

Conclusion

Approaches to identifying and studying fungal diversity often relied on the traditional methods of plating fungi on agar. These approaches are still relevant for culturable fungi but fall short of detecting fastidious and unculturable fungi. Molecular-based techniques like targeted sequencing of the 16S rRNA gene from gross DNA samples have facilitated surveys of fungal diversity. The sequencing and cloning of individual sequences is however tedious and cannot provide a comprehensive survey of a fungal community. The 28S rRNA gene can be amplified from pure fungal colonies or can be amplified directly from a crude sample. Amplified from a crude sample, the 28S rRNA gene can be massively sequenced using high-throughput sequencing instruments. Direct amplification of the 28S rRNA gene and its massive sequencing has corrected the underrepresentation of fungal in many bacterial communities. Analysis of bacterial communities is now made easier by the ample data generated from various bacterial communities' survey projects.

Acknowledgements

The authors are grateful for the financial support granted by the Indian Council of Agriculture Research (ICAR) Govt. of India under the Niche Area of Excellence on "Exploration and Exploitation of *Trichoderma* as an antagonist against soil borne pathogens" running in the Biocontrol Laboratory, Department of Plant Pathology, C.S.A. University of Agriculture and Technology, Kanpur, India.

References

1. Purvis A, Hector A (2000) Getting the measure of biodiversity. Nature 405: 212-219.
2. Grigoriev IV, Cullen D, Goodwin SB, Hibbett D, Jeffries TW, et al. (2011) Fueling the future with fungal genomics. Mycology 2: 192-209.
3. Cole JR, Wang Q, Cardenas E, Fish J, Chai B, et al. (2009) The Ribosomal Database Project: improved alignments and new tools for rRNA analysis. Nucleic Acids Res 37: D141-D145.
4. Amann RI, Ludwig W, Schleifer KH (1995) Phylogenetic identification and in situ detection of individual microbial cells without cultivation. Microbiol Rev 59: 143-169.
5. Monstein H, Nikpour-Badr S, Jonasson J (2001) Rapid molecular identification and subtyping of Helicobacter pylori by pyrosequencing of the 16S rDNA variable V1 and V3 regions. FEMS Microbiol Lett 199: 103-107.
6. Sundquist A, Bigdeli S, Jalili R, Druzin ML, Waller S, et al. (2007) Bacterial flora-typing with targeted, chip-based Pyrosequencing. BMC Microbiol 7: 108.
7. Cole JR, Wang Q, Chai B, Tiedje J (2011). The Ribosomal Database Project: Sequences and Software for High-Throughput rRNA analysis. Handbook of Molecular Microbial Ecology.
8. Wang Q, Garrity GM, Tiedje JM, Cole JR (2007) Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. Appl Environ Microbiol 73: 5261-5267.
9. Muyzer G, de Waal EC, Uitterlinden AG (1993) Profiling of complex microbial populations by denaturing gradient gel electrophoresis analysis of polymerase chain reaction-amplified genes coding for 16S rRNA. Appl Environ Microbiol 59: 695-700.
10. Liu WT, Marsh TL, Cheng H, Forney LJ (1997) Characterization of microbial diversity by determining terminal restriction fragment length polymorphisms of genes encoding 16S rRNA. Appl Environ Microbiol 63: 4516-4522.
11. Fisher MM, Triplett EW (1999) Automated approach for ribosomal intergenic spacer analysis of microbial diversity and its application to freshwater bacterial communities. Appl Environ Microbiol 65: 4630-4636.
12. Claesson MJ, Wang Q, O'Sullivan O, Greene-Diniz R, Cole JR, et al. (2010) Comparison of two next-generation sequencing technologies for resolving highly complex microbiota composition using tandem variable 16S rRNA gene regions. Nucleic Acids Res 38: e200.
13. Pace NR, Stahl DA, Lane DL, Olsen GJ (1986) The analysis of natural microbial populations by rRNA sequences. Adv Microbiol Ecol 9: 1-55.
14. Schloss PD, Handelsman J (2005) Introducing DOTUR, a computer program for defining operational taxonomic units and estimating species richness. Appl Environ Microbiol 71: 1501-1506.
15. Schloss PD, Westcott SL, Ryabin T, Hall JR, Hartmann M, et al. (2009) Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. Appl Environ Microbiol 75: 7537-7541.
16. Sun Y, Cai Y, Liu L, Yu F, Farrell ML, et al. (2009) ESPRIT: estimating species richness using large collections of 16S rRNA pyrosequences. Nucleic Acids Res 37: e76.
17. Latha J, Mukherjee PK (2002) Molecular characterization of ex-type strains of Trichoderma spp. from two Indian type cultures, collections. BARC Newsletter 145-149.
18. An Z, Wang C, Liu X, Bennett JW (2010) China's fungal genomics initiative: a whitepaper. Mycology 1: 1-8.