10 SPPADBASE: The First On-Line Searchable Database of PCR Primers for Phytopathogenic Fungi

Stefano Ghignone[1] and Quirico Migheli[2]

[1] Plant Protection Institute, Sect. Turin – CNR, V.le P.A. Mattioli, 25, I-10125 Turin, Italy.
[2] Department of Plant Protection - Center for Biotechnology Development and Biodiversity Research, University of Sassari, Via E. De Nicola 9, I-07100 Sassari, Italy.

INTRODUCTION

The fast and unambiguous identification of microbial pathogens affecting plants or plant products is an essential prerequisite for obtaining high-quality and safe production. Ecologically friendly practice of the modern agriculture requires the adoption of diagnostic techniques able to detect minimum inoculum levels of pathogens in soil, seeds, transplants or crops, to limit the raise of epidemics and to address the adoption of rational and efficient control means. Moreover, there is an increasing public and official awareness of the potential threat of bio-terrorism directed against food and agriculture (Monke, 2004). Rapid detection techniques for bioweapon agents are a critical need for the first-responder community.

Among the nucleic acid-based diagnostic techniques, those involving the Polymerase Chain Reaction (PCR; Mullis and Faloona, 1987) are the most suited for early detection of phytopathogenic agents, due to their high sensitivity and the potential for automation. Many sequence source types could be selected and used as target for specific primer design. These may include, for instance, Random Amplified Polymorphic DNAs (Williams *et al*., 1990; Welsh and McClelland, 1990), internal transcribed spacer (ITS) regions of the ribosomal RNA genes (White *et al*., 1990) or other specific gene sequences. Primer sets can be designed to target specificity at the genus, species, or physiological race level, to distinguish a particular pathogen from closely related organisms.

A common and tedious task for researchers and technicians is to search for and retrieve bibliographic references of published and validated specific primer sets for a given pathogen querying the Internet, abstract collections and monthly journals' tables of contents. Very few examples of specific primer set collections for phytopathogenic agents have been released: a summary of primers

for the diagnostic characterization of phytopathogenic bacteria seems to be the only one printed so far (Louws *et al*., 1999). Moreover, among 719 molecular biology databases publicly available recorded by Galperin (2006) or among the 2470 BMC biomedical databases catalog available at http://databases.biomedcentral.com/, no online repository of primer sets of this kind is accessible. To overcome this lack of information, we released the first online searchable database of primer sets useful for the detection and identification of plant pathogenic fungi.

ACCESSIBILITY

"*The Database of PCR Primers for Phytopathogenic Fungi*" can be accessed at http://www.sppadbase.com (Fig. 10.1). SPPADBASE is an acronym that stands for "Specific Primers for Phytopathogenic Agents Data Base" and is a logical choice to host primer sets specific to other phytopathogens, such as viruses or bacteria, including microbial herbicides and biological agents relevant in agroterrorism.

Primer sets can be searched through a quick (Fig. 10.2(A)) or an advanced (Fig. 10.2(B)) search field. In the first case, users can search for primers typing the genus or species name of the target organism, the forward or the reverse primer name, the template DNA, the PCR technique, an author name or the year of publication; the use of word parts is allowed. In the second case, users can focus their searches by typing or selecting an appropriate search term; with this interface it is possible to retrieve primer sets using as search term the database ID (known from a previous search), the primer sequence, the GenBank accession numbers and the reference journal. The organism definition fields (with grey background) are combinable and allow retrieving primers for a specific pathogenic fungal agent, defining its subspecies rank, the subspecies epithet or other name modifiers.

The database records matching the query criteria are then tabled in a search results page (Fig. 10.3). For each set, the forward and the reverse primer couple, the target DNA, the PCR technique (if the data are available), and the bibliographic reference are displayed. Clicking on the reference, a pop-up window will open, showing the publication details in brief; if a PubMed ID is available, a link will open the corresponding page at NCBI. Clicking on the database ID of a selected set, will show the complete details (Fig. 10.4). The primer set detail pages are organized in the following sections:

1. **Internet resources**: This section allows queries with external search engines (Google, Altavista, Alltheweb) and other reference databases: the "NCBI Taxonomy Browser" shows the taxonomical lineage of the organism of interest, as it appears in the GenBank database, and provides direct access to the Entrez-deposited sequences obtained from each organism; the "CABI Bioscience Index Fungorum" provides the correct organism nomenclature based on a database of fungal names containing over 350,000 names of fungi at species level and below, derived from a number of published lists; the "Centraal Bureau voor Schimmelcultures" database provides the Anamorph/Teleomorph connections of the organism.
2. **Primer set**: This section shows the name and the sequences of the forward/reverse primers of the selected set, the amplicon size (if available), and a tool to perform BLAST similarity searches with the selected oligonucleotide sequence. If the primer set was designed for Real Time PCR experiments based on TaqMan™ chemistry, the probe name and sequence are also displayed.

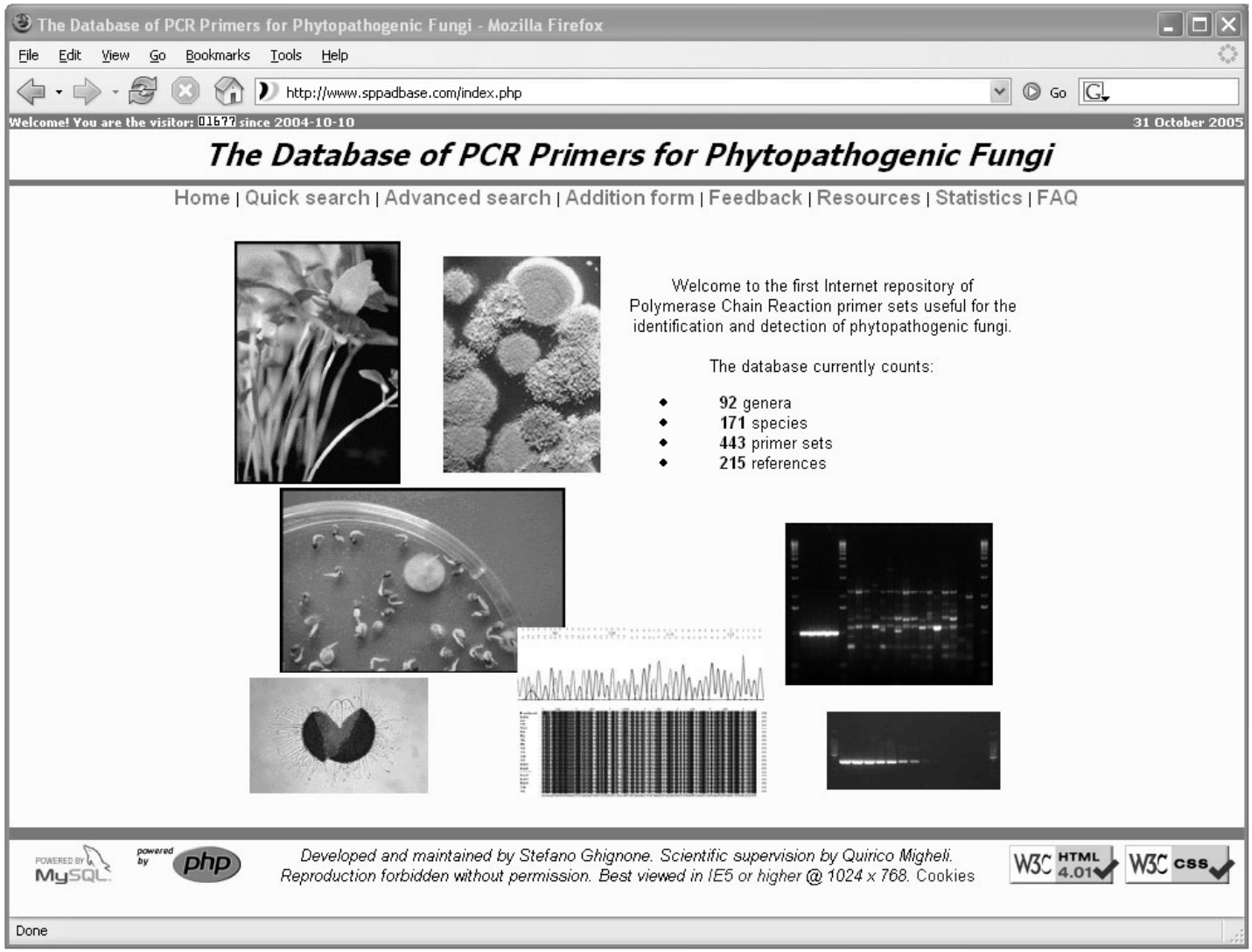

Figure 10.1. "*The Database of PCR Primers for Phytopathogenic Fungi*" home page. Live database statistics are shown. Site pages are reachable trough the navigation bar.

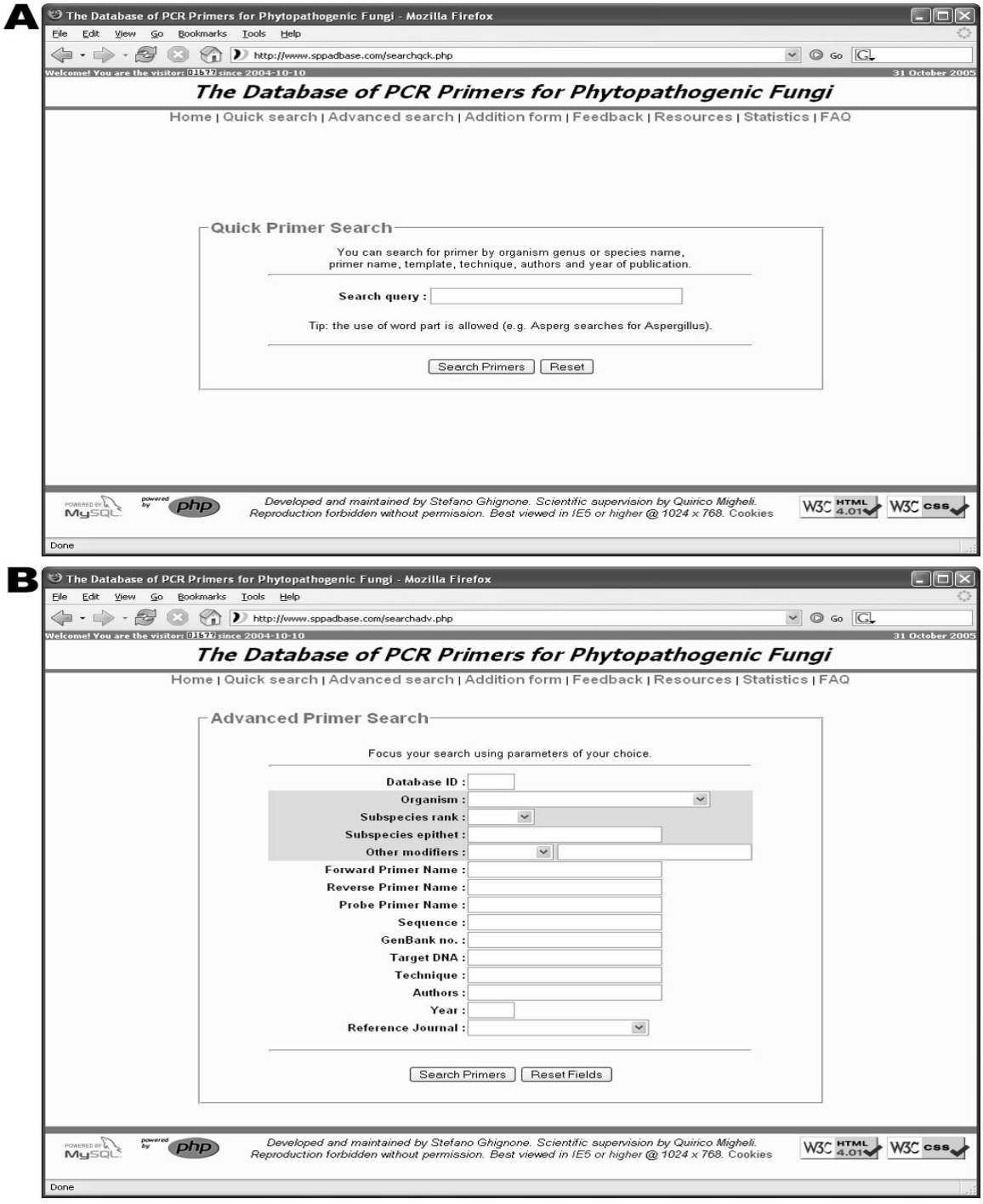

Figure 10.2. SPPADBASE **(A)** Quick search interface and **(B)** Advanced search interface.

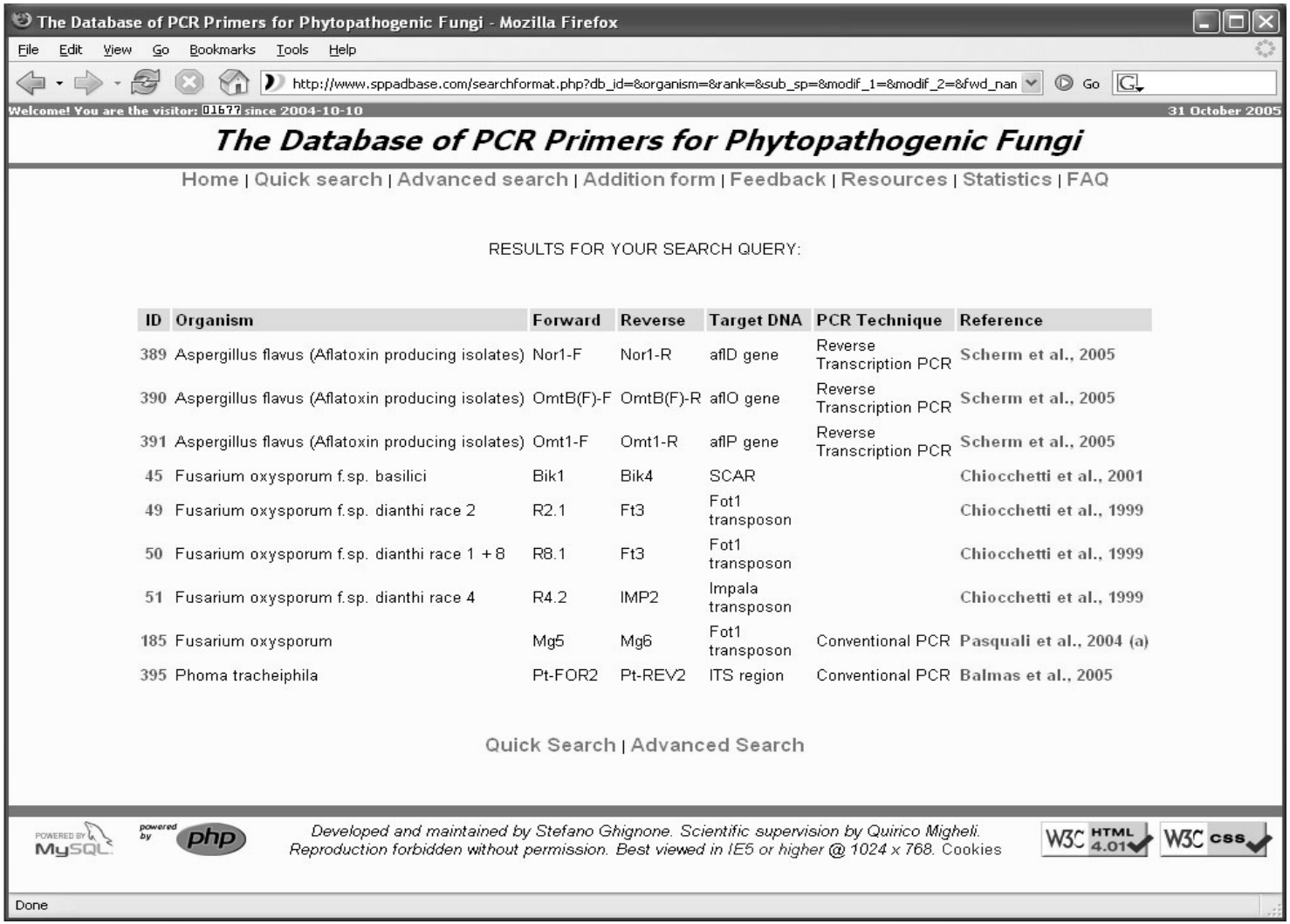

Figure 10.3. Example of a search result page. Target organisms are listed in alphabetical order. For each record, database ID, complete organism definition, forward and reverse primer name, target DNA, PCR technique and reference are given. Clicking on the reference, a pop up will open showing the publication details in brief. Clicking on the database ID of a selected set, the complete details will be shown.

Figure 10.4. Example of a primer set details page. Primers (and probe, in case) sequences, link to source sequence, target DNA and amplicon size, PCR technique and publication details are shown, ordered in section. Query systems to search the Internet with the organism definition, to check the current organism nomenclature, to perform homology BLAST searches with primer/probe sequences are also available. In the lower side of the page, submitter details are displayed along with the record submission and modification dates.

3. **Source Sequence**: If the GenBank accession number of the primer's source sequence is available, it is displayed and linked directly to the NCBI GenBank flat file.
4. **Reference**: Complete reference details are shown here. If available, the article abstract and the link to the corresponding PubMed ID are displayed.
5. **Submitter**: Users are invited to submit basic data (authors, reference) of missing primer sets. Submitter details will appear in this section.

DATA SOURCES

Only published primer sets that do not require further amplification product analyses for the identification of the target organism have been considered for the construction of the live database. Primer set reference collection has been performed scanning the tables of contents of the most widely diffused phytopathology-related journals from the early 1990s to date. More than 30 reference sources are now monitored to maintain the reference database (Table 10.1).

Table 10.1. List of the phytopathology-related journals currently monitored to maintain the primers' database updated.

African Journal of Biotechnology	Journal of Plant Diseases and Protection
Applied and Environmental Microbiology	Journal of Plant Pathology
Applied Microbiology and Biotechnology	Letters in Applied Microbiology
Australasian Plant Pathology	Molecular Ecology
Biological Control	Molecular Genetics and Genomics
BMC Microbiology	Molecular Plant Pathology
Canadian Journal of Plant Pathology	Molecular Plant-Microbe Interactions
Canadian Journal of Botany	Mycologia
European Journal of Forest Pathology	Mycological Research
European Journal of Plant Pathology	Mycoscience
FEMS Microbiology Letters	Physiological and Molecular Plant Pathology
Forest Pathology	Phytopathology
Fungal Genetics and Biology	Plant Disease
International Journal of Food Microbiology	Plant Pathology
Journal of Applied Microbiology	Plant Pathology Bullettin
Journal of Food Protection	Systematic and Applied Microbiology
Journal of Microbiological Methods	Weed Science
Journal of Phytopathology	

The accuracy of the organism nomenclature, the taxonomical position and anamorph/teleomorph connection, as well as GenBank accession numbers reported can be checked querying external databases accessible from the primer set details page (Fig. 10.4). However, we encourage users to send us notification regarding missing or erroneous data. Furthermore, being this list far from exhaustive, users are invited to submit missing primer sets through an appropriate interface, providing detailed literature reference information and/or uploading an electronic copy of the published article. Submitters also are asked to provide personal details that will appear linked to the newly added primer set, following the system administration review, in the live database.

IMPLEMENTATION

"*The Database of PCR Primers for Phytopathogenic Fungi*" is a Web service implemented with PHP (http://www.php.net/), an open-source server-side scripting language, designed for generating HTML contents. MySQL (http://www.mysql.com/), an open-source database, is used as the data-management system.

The complete application is currently hosted on a Linux based machine, running an Apache 2.0 Web server.

A single script generates each HTML page code; each page code compliancy to the current HTML 4.01 Specification has been validated with the W3C MarkUp Validation Service (http://validator.w3.org/), as well as the Cascading Style Sheet (CSS) employed and the site links has been validated with the W3C CSS Validation Service (http://jigsaw.w3.org/css-validator/) and the Link Checker (http://validator.w3.org/checklink/), respectively.

The application has been optimized for the Microsoft Internet Explorer 5 browser or higher and it is best viewed at 1024x768 dpi screen resolution.

Application functionality has also been tested with other widely used Internet browsers such as the Gecko codebase class-based ones (e.g., Netscape, Mozilla), Firefox and Opera, running in Windows, Linux or Mac environment.

REFERENCES

Galperin M.Y. (2006). The Molecular Biology Database Collection: 2006 update. *Nucleic Acids Research* 34 (Database issue): D3-D5.

Louws F.J., Rademaker J.L.W. and de Bruijn F.J. (1999). The three Ds of PCR-based genomic analysis of phytobacteria: Diversity, detection, and disease diagnosis. *Annual Review of Phytopathology* 37: 81-125.

Monke J. (2004). Agroterrorism: threats and preparedness, Congressional Research Service Report RL32521 for US Congress. http://www.fas.org/irp/crs/RL32521.pdf

Mullis K. and Faloona F.A. (1987). Specific synthesis of DNA *in vitro* via a polymerase-catalyzed chain reaction, *Methods in Enzymology* 155: 35-40.

Welsh J. and McClelland M. (1990). Fingerprinting genomes using PCR with arbitrary primers. *Nucleic Acids Research* 18: 7213-7218.

White T.J., Bruns T., Lee S. and Taylor J. (1990). Amplification and direct sequencing of fungal ribosomal RNA Genes for phylogenetics. In: PCR - Protocols and Applications - A Laboratory Manual, Innis N., Gelfand D., Sninsky J. and White T. Eds., Academic Press, New York, 315-322.

Williams J.C.K., Kubelik A.R., Livak K.J., Rafalski J.A. and Tingey V. (1990). DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. *Nucleic Acids Research* 18: 6531-6535.