Identification of *Leptographium* species by oligonucleotide discrimination on a DNA microarray

by

Natalie Van Zuydam

Submitted in partial fulfilment of the requirements of Magister Scientiae

Forestry and Agricultural Biotechnology Institute

Department of Genetics

in the

Faculty of Natural and Agricultural Sciences

University of Pretoria

City of Tshwane

Supervisors: **Prof. B. D. Wingfield**

Dr. K. Jacobs

Prof. M. J. Wingfield

July 2009

1

Acknowledgments

Finally yet importantly, I would like to thank my family and friends. I would like to thank my family for their support and patience even when they did not understand what I was doing or why I was doing it. Thank you in particular to my mom who supported me through my thesis. I would also like to thank Rosie Van Zyl, Gayle Heske, Dominique Dye, Louise Roux and Kershney Naidoo for their unfaltering friendship, particularly through the difficult times.

Table of Contents

Preface

The volume of sequencing data available for different organisms has greatly increased with the advances in manual and automated sequencing technology. These advances have dramatically influenced how fungal species are identified and described. This has allowed for intricate molecular identification techniques to be developed and applied to species identification. The genus *Leptographium* presents an outstanding example of a genus that has been rigorously described using morphological characters and molecular techniques. Therefore, extensive amounts of sequence data, available for the majority of species in the genus, as well as meticulous morphological data, for each of the species, are available. The genus is thus an ideal model on which to test the usefulness of some of the newly developed molecular identification techniques. This thesis consists of four chapters that focus on different aspects of species identification, using different molecular techniques, within the genus *Leptographium*. The research chapters investigated novel molecular techniques used to identify species of *Leptographium,* which was presented as a model system.

Chapter 1 represents a review of the literature pertaining to the different DNA based methods of identification for fungi, with special reference to the genus *Leptographium.* The purpose of the review is to summarise molecular identification techniques that are currently available to identify fungi. Species diagnostic arrays and diagnostic PCRs are discussed in detail with reference to identification of *Leptographium* species. The review also includes a short section on the morphology, taxonomy and ecology of *Leptographium* species; presenting the genus as a model for testing the applications of current molecular diagnostic techniques within the ascomycetes.

In the first research chapter, Chapter 2, a prototype array was designed from available sequence data to identify 26 species of *Leptographium.* The array consisted of thirty-four, 20-mer probes that differentiated among 26 species of *Leptographium*. A specific primer pair was designed to amplify a specific target that would bind one of the 34 species-specific probes. The prototype array was challenged with targets from *L. dryocoetidis, L. elegans, L. leptographioides,* and a *Fusarium* species that served as a negative control.

Using the experience gained with the development of the prototype chip, a large array for 56 species of *Leptographium* was designed in Chapter 3. The array was designed from available sequence data using a minimal probe design. Phenograms were constructed from shared and unshared probe characteristics to reveal probes that were common to species grouped together

under a single node and other probes that were branch or species-specific. Thus, a similar design to that used for PhyloChips was adopted for the second array design.

The *in silico* design of a large array with a hierarchical probe structure was transferred to a PCR diagnostic for species of *Leptographium* in Chapter 4 The probes designed for the large *Leptographium* array were transferred to PCR by using them either as a forward or reverse primers. Each "identiprimer" was then coupled with a complementary, universal primer and optimised for amplification in the relevant species. The diagnostic system was thus based on the phenograms produced in the design of the large array, Chapter 3. The diagnostic PCR was tested for repeatability using a blind test as well as identification of undescribed isolates.

The three research chapters presented in this thesis represent individual units that discuss different aspects of identification of *Leptographium* species. It is for this reason that repetition among the chapters has been unavoidable. It is the first time that microarrays and microcoding have been applied to the identification of *Leptographium* species and these results could be used to apply these identification techniques to other genera within the ascomycetes.

1. **Chapter 1**: A review of current DNA based diagnostics for *Leptographium*

Summary

Leptographium is an anamorph genus within the Ophiostomatoid group of fungi and represents a unique case for molecular applications. The genus has a near complete sequence data available for three genes across all known species. This characteristic makes it a perfect test group for investigating applications of new diagnostic techniques within ascomycetes. Probes and primers, for microarrays, are designed from phylogenetically useful gene regions and are fabricated onto a solid substrate using printing technology. The sample is prepared using PCR and is hybridised to the probes under stringent conditions. The resulting fluorescent pattern is rigorously analysed to distinguish species from each other. Diagnostic PCR uses primers that are designed in similar way to the way probes are designed for microarrays and indicate the presence of a species through positive amplification. This research methodology will be applied to *Leptographium* to evaluate the efficacy of microarray technology for discriminating species within that genus. The data gained from this research study will be used in applications for other genera using microarray technology.

Keywords: *Leptographium;* microarray; species identification.

1.1. Introduction

The ophiostomatoid fungi represent an artificial grouping of morphologically similar genera that include *Ophiostoma* H. & P. Sydow*, Ceratocystis* Ell. & Halst*, Sphaeronamemella* Karsten ex Seeler*, Gondwanamyces* Marais & M.J. Wingf. and *Cornuvesica* Viljoen, M.J. Wingf. and Jacobs. Many Ophiostomatoid fungi are economically important as they are pathogenic to plants and cause economic loss through timber damage. They cause sapstain on logs, lumber and pulpwood (Upadhyay, 1993). Sapstain is the discolouration of the sapwood, typically a dark grey or dark blue colour that is caused by the presence of pigmented hyphae growing in the sapwood (Seifert, 1993). Sapstain may cause mortality in trees by blocking the xylem tissue and preventing the uptake of nutrients. The blue stain renders logs unsuitable for export (Seifert, 1993).

Leptographium Crane & Schoknecht is the anamorph genus of *Grosmannia* Goidánich and causes blue stain in a variety of tree hosts (Jacobs, 1999; Zipfel *et al.*, 2006). The genus is commonly associated with bark beetles that act as vectors for transporting the fungi from host to host (Hausner *et al.*, 2005; Lee *et al.*, 2006; Six and Bentz, 2003; Solheim, 1995). The classification of species within the genus remains difficult due to similar morphological features, although classification through molecular methods has resolved some discrepancies (Jacobs *et al.*, 2005; Lee *et al.*, 2003; Lee *et al.*, 2005). Most *Leptographium* species are weak pathogens; however, *L. wageneri* (Kendrick) and *L. procerum* (Kendrick) (Eckhardt *et al.*, 2004b; Wagener and Mielke, 1961) have been associated with diseases of pine. *L. procerum* has been consistently associated with White Pine Root Disease (WPRD) and *L. wageneri* has been identified as the causal agent of Black Stain Root Disease (BSRD) (Schweigkofler *et al.*, 2005). These pathogens are often found growing with morphologically similar, non-pathogenic species of *Leptographium* and this complicates their identification (Jacobs and Wingfield, 2001). In addition, more species are being described in the genus *Leptographium* that share similar morphological characteristics with existing species, making morphology an undesirable characteristic for identification for inexperienced taxonomists.

The available monograph of *Leptographium* integrates morphological and molecular information for each species providing a framework for species identification (Jacobs, 1999). A substantial amount of work has been conducted on the phylogeny of this genus in order to define all the species within the genus (Jacobs *et al.*, 2003; Jacobs, 2004; Jacobs *et al.*, 2001). Three gene regions have been used to draw up phylogenies for *Leptographium*: β-tubulin, Elongation factor-1α and the internal transcribed spacer region (ITS2) (Jacobs *et al.*, 2001). Housekeeping gene sequence data may also be used to design and execute a microarray experiment where unique regions for each

species are selected for a probe sequence (Cleven *et al.*, 2006; Fukushima *et al.*, 2003; Huang *et al.*, 2006). Since this genus is well defined, both taxonomically and phylogenetically, the results obtained from a microarray experiment will be strong indicators of the practical applications of species identification using a microarray platform.

DNA microarrays consist of a solid substrate embedded with an array of molecular probes that can bind to distinctive, prepared targets (Schena and Davis, 2000a). Oligonucleotide probes are embedded in a high-density format at precise, mapped locations (Maughan *et al.*, 2001). These probes may be produced by photolithography or prepared in solution and printed onto the substrate using current printing technology (Theriault *et al.*, 2000). Polymorphisms are detected through the principle of specific DNA base pairing of probes to complementary targets (Conner *et al.*, 1983). The probes are of known sequence and bind to their complementary, fluorescently labelled target (Schena *et al.*, 1995). The targets are prepared through a multiplex PCR and can be either directly or indirectly labelled with fluorescent dyes (Al-Khaldi *et al.*, 2004; Erdogan *et al.*, 2001; Keramas *et al.*, 2003; Myers *et al.*, 2006). Hybridisation of compatible targets to the probe produces a fluorescent signal that is detected using a light source and captured using scanning equipment (Schermer, 2000). Software programs are used to extrapolate and interpret the data from these images (Maughan *et al.*, 2001).

Microarray experiments produce many data points in a single assay making it a desirable system for the identification of multiple species. They have been used as a high throughput method for multiple species identification, reducing the time to make an identification, while simultaneously increasing the accuracy of the result (Cleven *et al.*, 2006; Couzinet *et al.*, 2005; Huang *et al.*, 2006). Traditionally species are recognised using a combination of morphological, molecular and biological characteristics (Taylor *et al.*, 2000). Species identification by a species diagnostic array uses species-specific polymorphisms that are discriminated using stringent hybridisation conditions to discern species (Booth *et al.*, 2003; Burton *et al.*, 2005; Leinberger *et al.*, 2005; Wilson *et al.*, 2002).

Species identification of microbes relies upon the availability of relevant sequence data of marker genes for the design of discriminatory probes (Bodrossy and Sessits, 2004). Species specific probes are embedded in a microarray and exposed to labelled target (Bodrossy and Sessits, 2004). Stable duplexes are formed with perfectly matched target molecules, while mismatched targets form unstable duplexes under specific hybridisation conditions (Wallace *et al.*, 1979). The probes must be well designed from available sequence data and there is a wealth of sequence data available for

Leptographium (Jacobs *et al.*, 2006; Jacobs *et al.*, 2001). This attribute makes it a suitable model genus for exploring the application of this technology to ascomycetes.

1.2. Part I: What is a microarray?

Microarray chips are molecular tools used to answer biological questions pertaining to gene expression, gene interactions as well as questions based on populations of organisms and species identification (Brodie *et al.*, 2006; Lemieux *et al.*, 1998; Li and Stormo, 2001; Loy *et al.*, 2002; Troesch *et al.*, 1999). A microarray physically consists of a solid substrate in which single-stranded probe sequences are embedded (Figure 1). Probe sequences are embedded by printing onto solid substrates (Theriault *et al.*, 2000). The location and base sequence of these probes is known and is exposed to a set of labelled, target molecules of either ssDNA or ss-cDNA producing a coloured hybridisation pattern (Maughan *et al.*, 2001)
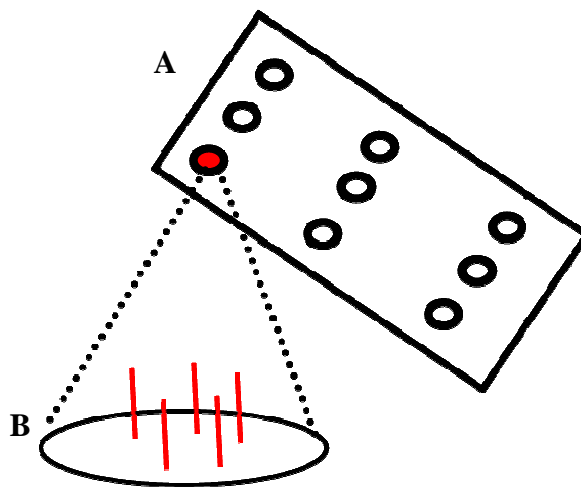
Figure 1. A Schematic representation of a DNA microarray using a microscope slide as a substrate. A: Spots of oligonucleotides of known constitution at documented locations. B: Enlarged spot, showing single stranded oligonucleotides deposited in a known location.

1.2.1. Solid substrates for microarray fabrication

Probes may be embedded onto one of a number of different substrates where selection is based on the detection technique. The choice of solid substrate is important and there are several options

available to researchers (Schena and Davis, 2000a). Substrates have to fulfil several criteria: they have to be non-porous for the deposition of biochemical material in precise locations and to ensure that reagents and samples are not absorbed (Theriault *et al.,* 2000). Excess reagents and sample material must be washed off after the hybridisation reaction or residual reagents and sample material will provide misleading results. The substrate must also have a low fluorescent property so that the detection process is not compromised by a high background signal (Schena and Davis, 2000a).

1.2.1.1. Microscope slides

Glass microscope slides are relatively inexpensive, readily available and can be modified using chemical processes (Beaucage, 2001; Bodrossy and Sessits, 2004). This substrate is convenient and inexpensive to use, as it is compatible with many different fluorescent detection methods and automated fabrication technology (Beaucage, 2001; Loy and Bodrossy, 2006). The slides have a large surface area for embedding numerous probes that capture a large amount of data in one assay (Schena *et al*., 1995). Microscope slides are compatible with existing microscope technology and tools, reducing the cost of additional equipment, required if other substrates are utilised. The slide surface is easily manipulated and modified, and allows for basic hybridisation and washing (Beaucage, 2001). Microscope slides exhibit low background fluorescence simplifying data capture and analysis (Schena *et al*., 1995). However, the substrate does not prevent mechanical damage or dust contamination of the oligonucleotides in the laboratory environment (Schena and Davis, 2000b).

1.2.1.2. Three dimensional microarrays

Most platforms require similar probe thermodynamics so that mass screening can take place under similar conditions (Letowski *et al.*, 2004). However, it is possible to monitor hybridisation in real time so that probe thermodynamics is not a limiting factor (Anderson *et al.*, 2006). Three dimensional microarrays allow for the monitoring of hybridisation in real time, so that not all probes have to have the same melting kinetics (Anderson *et al.*, 2006). This technology was used in to detect *Phytophthora* species and is supported by PamChip microarray platform (KIT Biomedical Research, Amsterdam) (Anderson *et al.*, 2006). The unique feature of this technology is that it is possible to monitor probe-target duplexes over a temperature gradient, thus allowing for the use of a

variable number of targets that have different annealing temperatures. The optimal melting temperatures may also be determined for each probe. The data analysis for this type of experiment is intricate but it does allow for the real time assessment of probe-target duplexes (Anderson *et al.*, 2006).

Particle substrates are alternatives to the microscope slide format and are a common choice for SNP genotyping experiments as they are compatible with closed systems (Cai *et al.*, 2000). Dust and contaminants are less common in a closed system than in the open hybridisation system used for microscope slides.

1.2.1.3. Microparticles

SNP genotyping reactions may involve a primer extension reaction, to identify the SNP, that can be collected by microparticles (Fan *et al.*, 2005; Kaderali *et al.*, 2003). The most common particles used are microspheres or silica beads (Cai *et al*, 2000; Gunderson *et al.*, 2006). Generic tags can be anchored in microparticles and used as probes for capturing tagged, cyclic primer extension products (Cai *et al.*, 2000; Kaderali *et al.*, 2003; Syvånen, 2001). It is common to embed these microspheres with fluorophores that have characteristic emission wavelengths (Landegren et al., 1998).

The Luminex system utilises three fluorophores with the microspheres containing two internal fluorophores, and a third placed on the capture probe, which is used to determine the assay result. The capture probe includes a DNA sequence that acts as a zip code and is complementary to a unique sequence on the microsphere (Chandler and Jarrell, 2002). The capture probe will be included in the cyclic primer extension reaction as a primer and will serve to localise products to a unique microsphere (Syvånen, 2001). Fluorescence of the microsphere indicates the SNP class and the genotype is determined through analysis of the product (Syvånen, 2001). This technology is versatile and lends itself to many different genotyping reactions. Most of these technologies are only cost effective for ultra-high throughput users because of the cost of the instruments, infrastructure and the oligonucleotides. Nano-barcode technology is similar to the microsphere technology but uses nanowires as the substrate (Sha *et al*., 2005).

1.2.1.4. Nano-barcode technology

Nano-barcodes use nanowires as a solid substrate (Figure 2). The structure is very different to the conventional microscope slide format and the newer microparticle technology, as it exploits the properties of nanowires coated with gold or silver (Reiss, 2002). The nanowires are intrinsically encoded by the nature of the coating process so that each adjacent nanowire has a different conductivity (Reiss, 2002). Probes are attached to these nanowires and used to collect target products from the SNP genotyping reactions (Sha *et al.*, 2005). The multiplexing capability of nano-barcodes is high and the technology does not require expensive equipment and can still produce accurate data (Sha *et al*., 2005). The reaction does require a high concentration of DNA that is difficult to obtain through conventional PCR (Sha *et al*., 2005).
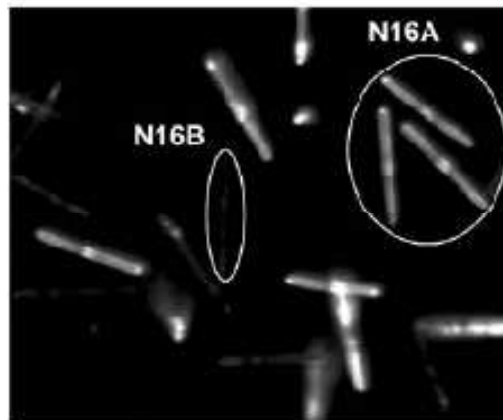
Figure 2. Nanowires are embedded with specific probes that hybridise to specific PCR products. Each nanowire is unique and has a characteristic conductivity that will indicate the SNP class and examination of the product will indicate the genotype of the SNP.

Features of the microarray will reflect the type of data that is captured and follows a sequential design process (Schena and Davis, 2000b). The first steps include target identification, primer design and probe design. The probes for a Species Diagnostic Array (SDA) are designed around polymorphisms in available sequence data. Each polymorphic probe will effectively identify a species (Beaucage 2001, Stenger *et al*., 2002, Volokhov *et al*., 2004).

1.2.2. Species diagnostic microarrays

Traditional identification methods for *Leptographium* species often included the use of morphological identification and general molecular methods (Fujita *et al.*, 2001; Harrington and Cobb, 1986; Jacobs *et al.*, 2000; Lee *et al.*, 2003; Witthuhn *et al.*, 1997). These methods are time consuming in comparison to the single assay achieved by using microarray technology (Burton *et al.*, 2005; Cleven *et al.*, 2006). The species concept is often complicated and multifaceted where ecology, morphology and molecular techniques are used to define and describe a new species (Taylor *et al* 2000.). Describing a new species requires many specialised skills and thorough knowledge of the genus (Taylor *et al*., 2000). *Pythium* and *Leptographium* are two examples of fungal genera where the species within the genus are very similar (Jacobs and Wingfield, 2001, Tambong *et al*., 2006). Identification of species, in such cases, requires taxonomic experts that recognise species according to their own experience and systematic characters (Taylor *et al.*, 2000).

Recently, it has been agreed that for fungi, a phylogenetic species recognition system is more accurate than a morphological species recognition system (Taylor *et al.*, 2000). Changes in the gene sequences of progeny can be tracked more rapidly, before a visible change is noted in phenotypic characters (Taylor *et al.*, 2000). There are many molecular techniques available to identify each species but no multiplex assay existed before the advent of SDAs. Species Diagnostic arrays differentiate species by their hybridisation patterns, which are dependent on a sequence of DNA base pairs.

Microarrays provide a platform for a multiplex assay and reduce the time to achieve identification (Maynard *et al.*, 2005; Roth *et al.*, 2004; Volokhov *et al.*, 2002). However, they must meet certain requirements before they are useful. The resolution of a microarray or its ability to discern species is determined by two very important factors: constitution of the probe sequence and the length of the probe that will anneal to the target (Loy and Bodrossy, 2006; Sergeev *et al.*, 2006). Typical sequences include housekeeping genes or genes that are useful in resolving phylogenies (Anderson *et al.*, 2006; Fukushima *et al.*, 2003; Loy *et al.*, 2002). The LSU is a common choice as well as the SSU and genes that code for proteins involved in DNA synthesis (Fujita *et al.*, 2001; Loy and Bodrossy, 2006). Probes may also be designed from species-specific genes such as unique toxin genes (Burton *et al.*, 2006; Davignon *et al.*, 2005; Sergeev *et al.*, 2006). A diagnostic assay must be practical for routine use, cost effective and sensitive.

The significant advantages of microarrays over PCR diagnostic techniques include the capacity to process many samples simultaneously, reducing diagnosis time, as well as the ability to view the sample composition in its entirety (Al-Khaldi *et al.*, 2004; Cleven *et al.*, 2006; Huang *et al.*, 2006; Maynard *et al.*, 2005; Myers *et al.*, 2006). An SDA format includes an amplification step that increases sensitivity, hybridisation identifies the amplicon and this can be achieved over many samples and markers simultaneously (Call *et al.*, 2001; Leinberger *et al.*, 2005; Sergeev *et al.*, 2006) Changes in sample composition may be used to assess diagnosis and treatment strategies (Sergeev *et al.*, 2006).

An important application of microarrays is the identification of organisms that directly affect human health (Burton *et al.*, 2005; Cleven *et al.*, 2006; Eom *et al.*, 2006; Kostic *et al.*, 2006). These may be organisms that pose a bio-security threat, cause diseases or are advantageous to human life (Volokhov *et al.*, 2004; Xu *et al.*, 2006). The main advantage of a microarray assay is that extracted nucleic acids can be exposed to an infinite number of gene probes in one assay, achieving the same as multiple PCR and sequencing reactions (Burton *et al.*, 2006; Fukushima *et al.*, 2003). The design of robust probes is essential to the success of an SDA.

1.2.3.   Probe design

Probe design is an important and challenging step of the array design as it is the probes that will determine the accuracy of the array (Leinberger *et al.*, 2005; Letowski *et al.*, 2004). Species diagnostic arrays require a very specific design that will differentiate between taxa while taking into consideration that all probes must behave similarly under experimental conditions (Bodrossy and Sessits, 2004). A probe set for species identification must be highly specific to their targets and should not cross-hybridise with other non-target sequences (Loy and Bodrossy, 2006). The constitution and length of the probe will determine the experimental conditions and the accuracy of the microarray. Probes can range in length from 20bp to 70bp, although it is unlikely that an array will be made up of several different probes of varying length but rather of a uniform set of similar length (Kane *et al.*, 2000; Letowski *et al.*, 2004; Loy and Bodrossy, 2006). Probes should all be of similar length to prevent competition for targets and so that they have similar melting temperatures (Bodrossy and Sessits, 2004; Li and Stormo, 2001; Loy and Bodrossy, 2006). Probe length should increase as the genome size becomes larger, to ensure that it is unique; the larger a genome the more likely a shorter probe will match another target region (Loy and Bodrossy, 2006).

Bioinformatics software is an integral part of probe design, since it takes into consideration sequence differences and thermodynamics. The main considerations of probe design are guanine/cytosine content, sequence polymorphisms and low complexity regions (Hyyro *et al.*, 2005). These properties determine the thermodynamic qualities of the probe and determine whether the probes will perform optimally under similar experimental condition (Letowski *et al.*, 2004). New freeware programs such as YODA (Yet another Oligonucleotide Design Application) (Nordberg, 2005) and OligoArray 2.0 (Rouillard *et al.*, 2003) allow the end user to choose the probe design parameters.

Twenty-mer probes are most commonly used in SDAs due to the high homology of probes where differences may consist of single nucleotide variations (Kostic *et al.*, 2006; Letowski *et al.*, 2004; Volokhov *et al.*, 2002). Shorter probes are more sensitive to SNP mismatches and may be printed at much higher density so that more probes may be printed on a single substrate (Cleven *et al.*, 2006; Kostic *et al.*, 2006; Martens *et al.*, 2007; Sergeev *et al.*, 2006; Tambong *et al.*, 2006). The SNP is usually situated toward the middle of the probe where a mismatch will cause the maximum amount of instability in a mismatched duplex (Letowski *et al.*, 2004). This can be further enhanced by including a peptiede nucleic acid (PNA) (Chandler and Jarrell, 2002) or a locked nucleic acid (LNA) (You *et al.*, 2006) at the site of the SNP so that only perfect matches will be accepted as a true duplex. Once the length of the probe has been decided upon, the parameters of probe constitution must be carefully selected.

In the case of closely related taxa, it is advantageous to use conserved genes in conjunction with genes, that are unique to a species, for probe design (Sergeev *et al.*, 2006; Volokhov *et al.*, 2004). It has been found, that using more than one gene as a target sequence increases the focusing ability of the array (Al-Khaldi *et al.*, 2004; Sergeev *et al.*, 2006; Volokhov *et al.*, 2004). In order to identify pathogenic bacteria such as the *Enterobacteriacae* the 16S rRNA sequences are used in conjunction with *cpn60* sequences to differentiate species within this family (Al-Khaldi *et al.*, 2004; Maynard *et al.*, 2005). For the detection of *Bacillus anthracis* unique toxin genes, carried on plasmids, are used (Maynard *et al.*, 2005). A practical characteristic of microarrays is that the probe range can always be extended to include more species as required (Cleven *et al.*, 2006; Sergeev *et al.*, 2006)

1.2.4. Probe synthesis and array fabrication

The integrity of the selected probes and targets is significant but the quality and the intrinsic properties of the microarray determine the quality of the results obtained (Theriault *et al.*, 2000). The quality of a microarray is directly determined by how the array is fabricated. In other words, how the probes are embedded onto the array will directly affect how the array will perform in downstream applications (Rogers *et al.*, 1999). The fabrication of arrays may involve immuno-immobilisation or direct synthesis of oligonucleotides onto the solid support but these are largely ineffective in producing high-quality, high-density arrays (McGall *et al.*, 1996; Theriault *et al.*, 2000). The method favoured for diagnostic arrays is the attachment of pre-synthesised oligonucleotides (probes) via either pre-activated solid supports or passive absorption, an example of this is by using disulphide bonds (Rogers *et al.*, 1999). This type of format is compatible with most printing techniques and coupling chemistries (Rogers *et al.*, 1999).

Probe sequences must be embedded onto the substrate through a mechanical process that draws on technology developed by the printing industry (Theriault *et al.*, 2000). There are three technologies presently available for fabricating microscope slides: photolithography, ink jet printing and microspotting. Photolithography relies on the use of phosphoramidimite DNA bases in solid phase DNA synthesis (McGall *et al.*, 1996). Piezoelectric technologies use a version of 'ink-jet' printing in order to dispense sub-nanolitre volumes of reagents to defined locations (Theriault *et al.*, 2000). Microspotting relies on a contact process where a print head, containing microspotting heads, allows for the transfer of prepared reagents from trays onto a solid surface (Theriault *et al.*, 2000). All the technologies provide sufficient density in order to represent an entire genome on one slide. The choice of technology will depend upon the quality of the data required, throughput, density, cost and flexibility (Theriault *et al.*, 2000).

Probes can be synthesised through photolithography, which is a reaction supported by ink-jet printing technology (Theriault *et al.*, 2000). The piezoelectric device automates local *in situ* DNA synthesis by the phosphoramidimite oligonucleotide synthesis cycle (McGall *et al.*, 1996). Electricity is used to deliver DNA bases, cDNAs and other molecules through tiny delivery jets at defined locations through a non-contact process. The device prints phosphoramidimite DNA bases at specific locations and the coupling reaction is catalysed by a light source (Theriault *et al.*, 2000). Photomasks direct light to specific locations that activate the modified DNA bases resulting in a coupling reaction (McGall *et al.*, 1996). Each coupling step results in the addition of another base to the growing chain (Theriault *et al.*, 2000).

There are printers that are capable of fully automated *in situ* DNA synthesis (McGall *et al.*, 1996; Schena and Davis, 2000b). This technology uses print heads to deliver the four phosphoramidimite bases and tetrazole. The bases have to undergo a 5' OH deprotection step, due to the 5'*O*-dimethoxytrityl protecting group, but this is chemically mediated (McGall *et al.*, 1996). Prior to the availability of this technology, operators were required to print the array and place the slides in coupling and oxidation baths before printing the next set of bases (Schena and Davis, 2000b). Photolithography is not optimal for experiments where high quality data is required as it is does not allow for washing. There is no opportunity for purification steps; therefore, all the products and by-products remain attached to the slide (Schena and Davis, 2000b).

Microspotting is a contact process that involves the delivery of pre-synthesised biological material by ink jet printing technology. In a gene expression pattern experiment, fragments were amplified using PCR and then printed onto glass microscope slides (Theriault *et al.*, 2000). The arrays were processed by chemical and heat treatment to attach the DNA products and to denature them. The material is supplied in an aqueous buffer contained in microtitre plates and is, therefore compatible with ink jet printing technology (Schena *et al.*, 1995). The GeneJet makes use of the piezoelectric technology and is capable of aspirating solutions from microtitre plates and depositing them onto the microscope slide without causing shearing in the DNA. The main obstacle to micro spotting is attaining the correct concentration of probe in the buffer so that the correct concentration of probe is deposited onto the slide (Schena *et al.*, 1995).

1.2.5. Target preparation techniques for diagnostic arrays

1.2.5.1. Target isolation from the genome using PCR

Species diagnostic array systems depend upon an amplification step to isolate and label the target. The sequence content of the array will dictate what kind of PCR method will be used, for example a universal PCR, a specific primer set, or a multiplex PCR (Loy and Bodrossy, 2006). Universal PCRs are used to amplify a target sequence from a number of organisms using conserved primers that reduces the time spent optimizing and designing new primers (Clewley, 2004; Glass and Donaldson, 1995; Maynard *et al.*, 2005; White *et al.*, 1990). However, a specific amplification is most commonly used in cases where the taxa are very closely related, so there is less chance of cross hybridisation. Both types of PCR may be developed into a multiplex that allows for the

amplification of targets using a single reaction (Loy and Bodrossy, 2006). Secondary structure in the target is common with both PCR methods and may affect the results (Lane *et al.*, 2004; Rouillard *et al.*, 2003).

Longer targets tend to form secondary and tertiary structures, thus they are undesirable (Liu *et al.*, 2006). The secondary structure of the targets affects their specificity and efficiency to bind complementary probes during hybridisation, leading to false negative and false positive results (Liu *et al.*, 2006). There are several methods used to reduce the occurrence of secondary structures in the target and one of them is to reduce the length of the target (Liu *et al.*, 2006). It has been shown that hybridisation efficiency can be improved by decreasing amplicon length to between 20-100nt although a slight increase in false positives due to target length reduction have been observed (Liu *et al.*, 2006). Targets can be chosen carefully *in silico* to match specific probes and not to show secondary structure formation. String matching algorithms are useful in identifying targets that meet specific parameters such as length, melting temperature and specificity. Careful *in vitro* preparation using nick translation and random hexamer labelling have been shown to inhibit secondary structure formation in targets and complements the in *silico* design (Lane *et al.*, 2004).

1.2.5.2. Sample labelling

There are two broad categories of methods used to label nucleic acids for microarray analysis: direct and indirect labelling (Do and Choi, 2007). Direct labelling is a strategy that involves the direct labelling of the samples and gives strong hybridisation signals (Do and Choi, 2007). Indirect labelling schemes incorporate epitopes to amino-allyl dUTPs incorporated into a target during Klenow labelling and are stained with Cy3 or Cy5 proteins that bind these epitopes, to give off a fluorescent signal (Do and Choi, 2007). Direct labelling is a less complicated procedure but is not foolproof (Do and Choi, 2007).

The samples are either directly labelled through enzymatic synthesis or through PCR primers. Fluorescent dyes, Cy3 or Cy5 aa-dUTPs, are incorporated along with other dNTPs into the synthesis strand of a single stranded target template (Do and Choi, 2007; Manduchi *et al.*, 2002; Williams *et al.*, 2004). Microarrays are prone to producing false positive or false negative hybridisations. In order to reduce the occurrence of false results, this method of labelling can be made more accurate by including controls in the form of multiple fluors for comparative analysis

(Badiee *et al.*, 2003). Indirect labelling is more laborious but gives results that are more accurate and is the more popular method (Do and Choi, 2007; Schena and Davis, 2000a).

Indirect labelling follows a basic process where an epitope is bound to the nucleic acid sample mixture and then exposed to a stain containing a protein (Do and Choi, 2007; Manduchi *et al.*, 2002). The protein may be fluorescent or may have other proteins associated with them that are involved in a fluorescent reaction. Targets labelled with amino allyl dUTPs, that contain a biotin reactive group, may be exposed to protein conjugates of Cy3 and Cy5 (Do and Choi, 2007). These dye-coupled targets are then hybridised to the array. Controls for both labelling methods are mandatory especially in the case of a two-dye system (Badiee *et al.*, 2003). It must be noted that the fluors do not emit the same signal intensity and these deviations will show discordance or imbalance between the two signals (Do and Choi, 2007). This must be compensated for in the data analysis following hybridisation (Schena and Davis, 2000b).

1.2.6. Hybridisation

Allele specific oligonucleotide (ASO) hybridisation methods developed for single nucleotide polymorphism (SNP) genotyping (Conner *et al.*, 1983; Wallace *et al.*, 1979) are applicable to species diagnostic arrays (Kostic *et al.*, 2006; Wilson *et al.*, 2002). This hybridisation method is capable of discriminating mismatches down to a single nucleotide (Wallace *et al.*, 1979). Allele specific oligonucleotide hybridisation involves immobilizing separated or enzymatically amplified fragments of target DNA by hybridising them to oligonucleotide probes (Saiki *et al.*, 1989; Wallace *et al.*, 1979). SNP genotyping relies on the influence of single base mismatches to destabilise the duplex formed between the probe and the target sample (Conner *et al.*, 1983; Wallace *et al.*, 1979). Accurate discrimination between perfectly hybridised and partially hybridised duplexes can only be achieved under extremely stringent conditions (Landegren *et al.*, 1998; Wallace *et al.*, 1979).

Additives in the hybridisation solution may further enhance discrimination of SNPs. The solution contains a high concentration of targets and may contain other additives that increase the stringency of the hybridisation reaction (Jacobs *et al.*, 1998). Additives such as formamide and DMSO may be included in the hybridisation reaction to increase the affinity of the target for the probes by decreasing secondary structure (Chakrabarti and Schutt, 2001). The conditions for each target probe duplex may be different so it is important to optimize the hybridisation solution as well as the hybridisation temperature for a particular probe and target set.

Specialised platforms have been developed to improve the hybridisation reaction. The duplexed ASO probe and target can be monitored over a temperature gradient (Anderson *et al.*, 2006) or alternatively, an electric field to determine the optimal stringency for discriminating between SNP genotype can be applied to certain platforms (Edman *et al.*, 1997; Nagan and O'Kane, 2001). The hybridisation pattern is detected using lasers and the data is analysed using specific software.

1.2.7. Detection

Successful hybridisation is measured by exciting fluorescent molecules on the sample molecule and collecting these visual signals for interpretation in the form of an image. Fluorescent molecules are excited at different wavelengths; therefore, the light source used to excite them must be variable (Schermer, 2000). Commercial systems use scanning technology with a photomultiplier tube (PMT) to scan the array surface and excite fluorescent molecules (Schermer, 2000). Light emitted from the fluorescent sample is separated from unwanted light using a series of mirrors, filter, and lenses. The light is then converted to an electrical signal with a photon multiplier tube (PMT) (Schermer, 2000).

Scanning technology for fluorescent detection also includes confocal scanning devices and CCD cameras. A confocal scanner utilises laser excitation on a small area of the glass substrate so that the image is gathered in sections (Schermer, 2000; Xu *et al.*, 2006). CCD cameras use many of the same principles but excitation and detection differ on minor points (Schermer, 2000). CCD based imaging involves illumination and detection of a large portion of the substrate simultaneously. A larger area can be viewed, eliminating the need for moveable stages and optics thus reducing costs (Schermer, 2000). CCD systems are much simpler needing only one light source and experience less optical cross talk (Schermer, 2000). Microarrays are scanned using one of the scanning systems and a 16-bit .tiff image generated.

1.2.8. Data analysis

Intensity readings are taken from the image by calculating the sum of the pixel values of each spot and dividing that by the total quantified area. A user defined grid pattern, known as a .gal file, is overlaid on the image and defined areas – circles or squares – that include the spotted probes, these areas are subjected to data mining (Schena and Davis, 2000a). The data from expression arrays may be displayed in a number of ways depending on their complexity. The expression data are

mined and analysed using various image analysis algorithms that are often optimised for individual experiments (Allison *et al.*, 2005; Quackenbush, 2001).

The analysis for species typing is different to that of expression arrays as the relative brightness of the spots is not as important as the strength of the brightness. The brightness or Signal to Noise Ratio (SNR) indicates either a positive or a negative hybridisation (Martens *et al.*, 2007; Maynard *et al.*, 2005). For Diagnostic arrays, it is important to consider SNR thresholds for each species (Martens *et al.*, 2007; Maynard *et al.*, 2005). Thresholds are determined by statistical analysis of control and test probe hybridisation signals and prevent false negative results for less abundant targets (Maynard *et al.*, 2005). In order to analyse microarray data it is necessary to manipulate them so that the effects of inherent variation are minimised.

An empirical microarray experiment contains biological replicates, technical replicates and control spots in order to assist with useful data analysis (Allison *et al.*, 2005). Biological replicates help to measure the variation introduced through differences between biological samples and the effect of measurement (Allison *et al.*, 2005). Technical replicates are necessary for quality assessment and control (Allison *et al.*, 2005). The number of replicates will increase the power of the statistical analyses used but this is not always possible when using expensive technology. The biological samples are often pooled (Quackenbush, 2001) so that the number of microarrays is reduced and this has the effect of reducing variation between slides (Allison *et al.*, 2005). In order for the data between slide replicates to be compared the data must be normalised (Allison *et al.*, 2005).

Normalisation can be performed between and within slide replicates for a one-colour experiment. Global median normalisation is a widely applied method that will prevent outliers from skewing the data, a common problem encountered when taking the average (Quackenbush, 2002; Zien *et al.*, 2001). There is no standard normalisation protocol as different data sets contain different variances that are influenced by different factors (Quackenbush, 2002). It is therefore necessary to use the most suitable normalisation method for the data and then perform an analysis of variance (ANOVA) (Kerr *et al.*, 2000; Zien *et al.*, 2001).

The ANOVA proposed by Edwards (2003), employs a probe specific model that considers each probe individually including the design and treatment factors (Edwards, 2003). In a one-colour experiment the ANOVA value Y is made up of the sum of μ (average intensity of all the spots associated with a specific probe), random array effects, the biological units and measurement error (ε). An ANOVA table makes it possible to estimate the within and between-slide variances that is used to determine the repeatability of the experiment (Edwards, 2003). To identify the probes that have been positively and accurately hybridised, the data are subjected to an F-test (Edwards, 2003;

Kerr *et al.*, 2000). The use of so many varied analysis methods for microarray experiments means that a standard must be set, so that subsequent experiments may be compared with data from previous experiments (Brazma *et al.*, 2001).

Brazma *et al.* (2001) proposed the Minimum Information about a Microarray experiment (MIAME) system to standardise experiment specifications (Brazma *et al.*, 2001). A description of six components of a microarray experiment should be available for public access:

- The experimental design
- The array design that include each array used and every element on the array
- The samples used, the way that they were extracted and labelled; the hybridisation procedure and parameters
- The images, quantification measurements, specifications, and the normalisation controls including the types and values.

1.2.9. Tools used to study fungal genomes and species

Genetic markers identified using molecular techniques can be used to evaluate levels of genetic diversity, phylogenetic relationships and particular races or pathotypes. Markers closely associated with pathogenicity genes are particularly sought after and there are several types of markers available. Isozyme markers are very easy to use and interpret, but are limited in the information that they provide about genetic variation (Zambino and Harrington, 1992). Amplified fragment length polymorphism (AFLP) fingerprinting is a system that is ideal for detecting genetic variation without prior knowledge of the DNA sequence (Majer *et al.*, 1996). Random fragment length polymorphism (RFLP) markers are more sensitive but require specific DNA probes designed from established sequences (Kim *et al*., 2001). More advanced molecular techniques require more sequence knowledge than has been previously available. With the advent of less expensive and faster sequencing technology more sequence data have become available and with it the opportunity to use more advanced molecular techniques.

In order to correctly identify organisms with similar morphology it has been necessary to employ molecular characterisation (Jacobs *et al.*, 2005; Tambong *et al.*, 2006). There are many molecular methods available, mostly based on PCR (Fujita *et al.*, 2001; Lee *et al.*, 2003; Majer *et al.*, 1996). Clinical PCR methods all require the use of pure cultures or enrichment of field samples and have a pre-requisite for knowledge of the organism (Vora *et al*., 2004). These factors influence the amount

of time required to identify a species and also the total number of samples that can be processed simultaneously (Vora *et al.*, 2004).

A new PCR-based technique, reverse line probe assay (PCR-LiPA), has been developed for distinguishing among *Candida, Aspergillus* and *Cryptococcus* species and within genera (Martin *et al.*, 2000). The method employed was to design primers to amplify the ITS (Internal transcribed spacer) region and to anneal the resulting amplicons to species-specific probes developed from regions within the ITS region. The ITS region was chosen for this study as it has shown sufficient variation for species-specific probes to discriminate among species to be developed. Other regions that were previously targeted for PCR assays were the 18S, 28S, actin, heat shock protein and other single copy regions or genes (Kerr *et al.*, 2000). Some interaction was detected between the probes, designed to discriminate within genera, and between PCR products but the results were clear enough to discriminate among species (Martin *et al.*, 2000). The PCR-LiPA relies on the specific annealing of probes to PCR products. Similarly, microarrays rely on the same principle (Martin *et al.*, 2000; Schena and Davis, 2000a).

Unlike other molecular techniques, a hybridisation reaction is limited by the number probes, and not the concentration of target molecules. Therefore, it is possible to screen environmental samples and simultaneously identify small sub-populations of species in a larger microbial population (Brodie *et al.*, 2006; Couzinet *et al.*, 2005; Loy *et al.*, 2002). The hybridisation reaction is highly multiplexed and very sensitive, thus reducing the time, labour and expertise required for full identification of organisms in a sample, and increasing the number of samples that can be processed concurrently (Vora *et al.*, 2004).

1.2.9.1. Multiplex PCR for species identification

Microarrays are comparable to a multiplex PCR but on a more intensive and intricate level. The design parameters are essentially the same for a multiplex PCR and a species diagnostic microarray. Specific, 20-mer oligonucleotides are designed from conserved genes that hybridise to specific targets (Fujita *et al.*, 2001; Martens *et al.*, 2007). The primers bind under similar conditions and are unique, showing no interaction between pairs (Elnifero *et al.*, 2000; Letowski *et al.*, 2004). Multiplex PCRs must amplify regions that are distinctive after gel electrophoresis without interaction between the pairs, so not as many oligonucleotides can be included as those embedded on a solid substrate for microarray experiments.

Optimisation of multiplex PCRs involves specific conditions for the primer pairs included. Thus, what could be achieved over many reactions is achieved in a single reaction reducing costs of reagents and equipment (Edwards and Gibbs, 1994). The limiting factor in such an assay is the number of primers that can be included in a single reaction without observing mispriming, PCR selection and PCR drift. PCR drift is due to temperature fluctuations and reagent interactions during the early cycles of PCR and PCR selection is due to the traits of the template (Wagner *et al*., 1994). This leads to biased amplification of one product or the amplification of an undesired product. Optimisation methods include adding each primer pair subsequent to the optimisation of the previous pair (Edwards and Gibbs, 1994; Markoulatos *et al.*, 2002). It is also possible to improve the specificity of the reaction by optimizing the components of the PCR and using adjuvants (Chakrabarti and Schutt, 2001).

It is necessary to optimise the other PCR reagents along with the primers and the template (Markoulatos *et al.*, 2002). The ratio of $MgCl_2$ to dNTP concentrations is very important in ensuring optimal performance of the *Taq* polymerase and the correct binding of the primers (Markoulatos *et al.*, 2002). The addition of dimethyl sulfoxide (DMSO), bovine serum albumin (BSA) and other adjuvants can often increase the specificity of the multiplex PCR by preventing secondary structure in the primers and the template (Chakrabarti and Schutt, 2001; Markoulatos *et al.*, 2002).

Multiplex PCRs have been used to identify human adenoviruses (Xu *et al.*, 2000), pathogens causing periodontitus (Tran and Rudney, 1996) and human (Fujita *et al.*, 2001) and plant (Hamelin *et al.*, 1996) fungal pathogens. Most studies include a few species associated with a particular host or specific disease; identification is based on sequence polymorphisms. The general method involves designing species-specific primers around conserved genes (Fujita *et al.*, 2001; Hamelin *et al.*, 1996; Jackson *et al.*, 2004; Luo and Mitchell, 2002; Tran and Rudney, 1996; Xu *et al.*, 2000). Wilson *et al.* (2002), recommended the use of PCR in conjunction with microarrays to improve the detection limit for pathogens. Bäckmann *et al.* (1999), amplified the 16S rRNA region from cerebral spinal fluid using universal primers. The amplicons were then used as a template for specific primers to detect pathogens involved in bacterial meningitis. To confirm the PCR result, the amplicons were restricted (Bäckmann *et al*., 1999).

The genus *Leptographium* presents a unique case for fungi that can be attributed to a comprehensive data set across all known species. The sequence data available for *Leptographium* have been used to describe a complete phylogeny for all known species (Jacobs *et al.*, 2006) and therefore would be useful for a diagnostic microarray chip. Microarrays have been used to identify species of many other organisms such as *Candida, Phytophtora, Fusarium* and *Pythium*, based on differences in the ITS region (Anderson *et al.*, 2006; Fujita *et al.*, 2001; Huang *et al.*, 2006 Nicolaisen *et al.*, 2005; Tambong *et al.*, 2006). The ITS region has become a common and useful target for fungal identification studies as it is highly polymorphic (Fujita *et al.*, 2001; Huang *et al.*, 2006). The sequence data available for *Leptographium* is a complete set of ITS, EF1α (Elongation Factor 1 alpha) and β-tubulin (βT) regions (Jacobs *et al.*, 2006). Polymorphisms in these regions may be as small as a single nucleotide difference that differentiates one species from another.

Single nucleotide polymorphisms are single base changes in the nucleotide sequence of an organism and can be useful as molecular markers. They are desirable as genetic markers due to their high abundance, low mutation rates and ease of typing (Erdogan *et al*., 2001). The trend in molecular markers is consistently changing as technology improves and progressively lends itself to more aggressive and accurate markers. SNPs are expected to supersede microsatellites as markers in disease gene mapping in the same manner that microsatellite markers replaced RFLP markers (Landegren *et al*., 1998). SNPs are more prevalent in the genome than microsatellites and are directly inherited.

SNPs are useful in differentiating species if they are unique to species and may be easily included in diagnostic probes. An example is the differentiation of closely related bacterial species based on the 16S rRNA region in Microbial Diagnostic arrays (MDM) (Bodrossy and Sessits, 2004; Keramas *et al.*, 2003; Loy and Bodrossy, 2006; Loy *et al.*, 2002; Maynard *et al.*, 2005). Species-specific polymorphisms in the *Leptographium* sequence data can be used to design species diagnostic probes.

1.3. Part II: A brief overview of *Leptographium*

1.3.1. Taxonomy

Convergent evolution of ascomata and conidiomata that have sticky spore drops suited to insect dispersal in the ascomycetes, has complicated species delimitations of both sexual and asexual states in the Ophiostomatoid fungi (Malloch and Blackwell, 1993; Okada *et al.*, 1998). The generic limits for anamorphic fungi that are dispersed by insects and form droplets of conidia at the apex of synnemata are unclear and complicated (Malloch and Blackwell, 1993; Okada *et al.*, 1998). Until recently, it was generally accepted that *Leptographium* was an anamorph of the teleomorph genus *Ophiostoma* (Hausner *et al.*, 2000). This position has come under considerable scrutiny as *Ophiostoma* species may produce multiple anamorphs from different genera. Zipfel *et al.* (2006), sought to resolve this issue by supporting the division of *Ophiostoma* into different genera using sequence data for the nuclear ribosomal large subunit and β-tubulin. The monophyletic groups correlate to morphological distinctions between species of *Ophiostoma* (Zipfel *et al.*, 2006). These phylogenies place *Leptographium* as the anamorph genus of the teleomorph genus *Grosmannia*. The genus was originally established by Goidánich (1936) but subsequent morphological analysis led to the rejection of this genus in favour of the established genus *Ophiostoma* (Siemaszko, 1939). Although the taxonomy of the groups making up the ophiostomatoid fungi remains largely unresolved, the genus *Leptographium* is well defined according to the phylogenetics and morphological species concepts (Jacobs and Wingfield, 2001; Jacobs *et al.*, 2001; Lu *et al.*, 2008; Masuya *et al.*, 2004; Zhou *et al.*, 2000).

1.3.2. Morphology

Leptographium is characterised by mononematous conidiophores with darkly coloured stipes and a series of branches at the apices (Kendrick, 1962) (Figure 3B). Conidia are hyaline, aseptate and produced in a slimy matrix through holoblastic extension of the conidiogenous cells (Kendrick, 1962). Conidiation results in the accumulation of slimy conidial masses at the apices of conidiophores to facilitate insect dispersal (Harrington, 1988) (Figure 3A). Many *Leptographium* species are tolerant of high concentrations of cycloheximide in culture and have rhamnose in their cell walls. These fungi are suited to insect dispersal and commonly associate with scolytid bark beetles that infest conifers (Lee *et al.*, 2006; Lee *et al.*, 2005; Reay *et al.*, 2002). *Leptographium*

was originally classified based on morphology (Kendrick, 1962) and to a lesser extent allozyme analysis (Zambino and Harrington, 1992), although it is now more common to classify these fungi in accordance with their sequence data (Jacobs, 2004; Jacobs *et al.*, 2005; Jacobs *et al.*, 2001; Lee *et al.*, 2005; Lu *et al.*, 2008; Zhou *et al.*, 2000).

Figure 3. A: Conidiophores of *Leptographium wingfieldii* with conidia in slimy heads in the gallery of a bark beetle (Sha *et al.,* 2005). B: The dark stipes branch into annelids that produce hyaline conidia in a slime mass.

1.3.3.   *Association with bark beetles*

Bark beetles are small insects that typically feed on the phloem tissue of their host trees. Bark beetles are suited to carrying other organisms in mycangia, pits found on the head, prontum or elytral areas, although some fungi may be carried internally (Paine *et al*., 1997; Scott *et al.*, 2008). Bark beetles (*Coleoptera, Scolytidae*) often carry species of *Ophiostoma* as well as yeasts, bacteria, other fungi and in some cases mites (Klepzig *et al.*, 2001; Six and Bentz, 2003). Many bark beetles that infest coniferous trees have been associated with and carry species with *Leptographium* anamorphs (Hausner *et al.*, 2005; Jacobs, 2004; Lee *et al.*, 2006; Lee *et al.*, 2005; Peverieri *et al.*, 2006; Siemaszko, 1939; Six and Bentz, 2003; Zhou *et al.*, 2001; Zhou *et al.*, 2000).

Until recently, it was thought that the bark beetles were simply vectors for blue stain fungi including *Leptographium,* but now a more complex relationship has been partially described (Klepzig *et al.*, 2001; Paine *et al.*, 1997; Scott *et al.*, 2008; Six and Bentz, 2007). The fungi may assist with host colonisation by reducing host defences (Harrington, 1993; Ross and Solheim, 1997) and provide the beetle with nutrition after colonisation of the host tree (Klepzig and Six, 2004; Paine *et al.*, 1997). It has been generally accepted that bark beetles feed on the phloem tissue of

trees causing mechanical damage while the fungi that they carry are able to access and infect the sapwood causing discolouration and in some cases disease (Solheim and Krokene, 1998). However, pathogenicity trials have given mixed results indicating an unclear relationship between the fungi and the bark beetles (Eckhardt *et al.*, 2004b; Harrington and Cobb, 1983; Ross and Solheim, 1997; Wingfield, 1986).

Ophiostomatoid fungi, including *Leptographium* spp., are vectored by bark beetles, infect the host tissue and grow in the galleries of the beetles (Solheim, 1995), but it is not true for all cases, that these fungi are vectored in order to cause disease (Klepzig and Six, 2004). The bark beetles can feed on the spores of fungi to supplement their diet of nutrient deficient phloem (Ayres *et al.*, 2000; Paine *et al.*, 1997). The fungi also concentrate nutrients such as nitrogen (Ayres *et al.*, 2000) and ergol sterols that are not readily available in the nutrient deficient phloem tissue (Bentz and Six, 2006). Feeding on spores of these beneficial fungi encourages the beetles to readily penetrate logs and form galleries (Adams *et al.*, 2008; Eckhardt *et al.*, 2004a).

Aggressive bark beetles, such as *Dendroctonus ponderosae* (Hopkins) (*Scolytidae*), infest trees through pheromone-mediated mass attacks and cause devastation in plantations (Raffa *et al.*, 1993). These bark beetles infest lodgepole pine (*Pinus contorta* var. *latifolia*), a prevalent and commercially important tree. These beetles carry at least two species of *Ophiostomatoid* fungi: *Grosmannia clavigera* (Robinson-Jeffrey and Davidson) and *Ophiostoma montium* (Rumbold). These fungi are part of a symbiotic network of organisms associated with this bark beetle and they provide different benefits (Adams *et al.*, 2008).

Grosmannia clavigera has a long standing association with *D. ponderosa* while *O. montium* appears to be a relatively new associate of the mountain pine beetle (Six and Paine, 1999). *G. clavigera* accelerates the development of associated *D. ponderosa* beetles producing larger adults and bigger clutch sizes when compared to beetles associated with *O. montium* and to those not associated with fungi at all (Adams *et al.*, 2008). The two species of symbiotic fungi associated with *D. ponderosa* are present in different proportions on the beetles depending on the ambient temperature of the beetles' environment (Six and Bentz, 2007).

D. ponderosa occupies a wide geographic range and can tolerate a range of temperatures and the fungi that it carries have different optimal growth temperatures (Six and Bentz, 2007). *O. montium* is more tolerant of lower temperatures than *G. clavigera,* as it sporulates more easily at cooler temperatures (Six and Bentz, 2007). Carrying two species of fungi with different optimal growth temperatures ensures that the beetles will always have a food source despite varying environmental temperatures (Six and Bentz, 2007). However, these beneficial fungi are still vulnerable to rivalry

from other less beneficial fungi. *Dendroctonus ponderosa* carries a bacterium that produces antibiotics against the antagonistic rival fungal species (Scott *et al.*, 2008). A relative of *D. ponderosa, D. frontalis* (Zimmerman), carries a *Streptomyces* bacteria species that protects its food source against antagonists (Scott *et al.*, 2008).

Streptomyces spp. are ubiquitous filamentous bacteria that are known as endophytes, symbionts and antibiotic producers. They are often found in tripartite mutualisms with fungi and insects (Mueller *et al.*, 2008). The prominent example of such an association is the mutualism between Attine ants and *Streptomyces* where the bacteria are essential to maintaining monoculture of the beneficial fungus (Currie *et al*., 1999). Consequently, bark beetles have been found to be associated with different species of *Streptomyces* (Scott *et al.*, 2008). *D. Frontalis* safeguards its food source, *Entomocorticum sp.,* against *O. minus* (Hedgecock) by carrying a bacterium that produces antibiotics against *O. minus* (Scott *et al.*, 2008). In other tripartite mutualisms it is thought that the bacteria benefit by being dispersed and being maintained in a protected niche (Currie, 2001).

In some cases, associated yeast and bacteria can actually stimulate the growth of the beneficial fungi (Adams *et al.*, 2008). In turn the beneficial fungi protect the yeasts and bacteria from tree host endophytes and other antagonistic microorganisms (Adams *et al.*, 2008). *Leptographium* species are found to be associated with specific bark beetles that infest specific tree hosts (Jacobs *et al.*, 2000; Solheim *et al.*, 2001; Zhou *et al.*, 2001) so, this thesis could hold true for other *Leptographium* species. A few species of *Leptographium* are considered true plant pathogens and are carried by aggressive bark beetles.

1.3.4. Major diseases associated with Leptographium species

1.3.4.1. Black stain root disease

Leptographium wageneri is the cause of BSRD and is one of the most important pathogens of conifers in western North America (Harrington and Cobb, 1983; Harrington, 1993). Trees suffering from BSRD will show reduced branch growth, chlorosis, needle retention, reduced needle size and resinous lesions on lower stems (Goheen and Hansen, 1978; Wagener and Mielke, 1961). The disease results in a black staining of colonised sapwood of the roots and lower stems due to presence of pigmented fungal hyphae in infected tissue (Hessburg and Hansen, 1987; Joseph *et al.*, 1998). The fungus kills host trees by colonising water-conducting tissues of roots, root collars and

lower stems. Fungal hyphae block the water movement from roots to foliage causing BSRD symptoms (Joseph *et al.*, 1998).

BSRD kills young saplings and causes slow decline in older trees that are more susceptible to bark beetle attack (Hessburg *et al.*, 1995; Hessburg *et al.*, 2001). *L. wageneri* causes considerable damage to conifer forests in western North America. It has been separated into three host-specialised morphological varieties (Harrington and Cobb, 1986). The distinction is based on differences in conidiophore morphology, maximum growth temperatures, cultural appearances, isozyme variation and host specificity (Harrington and Cobb, 1984; Harrington and Cobb, 1986; Zambino and Harrington, 1989). The three varieties are: *L. wageneri* var. *ponderosum*, *L. wageneri* var. *wageneri* and *L. wageneri* var. *pseudotsugae* (Harrington and Cobb, 1986).

BSRD has displayed a particular infection pattern over wide areas such as plantations that involve bark beetles (Hessburg *et al.*, 2001). A discerning study of the spread of BSRD in plantations was conducted on 10 – 30 year old, Douglas-fir plantations in Oregon (Hessburg *et al.*, 2001). The disease was found to occur in centres spreading from an infected tree to others in close proximity (Hessburg *et al.*, 2001). Decline occurs near the perimeter with dead trees in the interior closer to the initial infection site (Hessburg *et al.*, 2001). Infection centres were prominent in well-stocked stands where the preferred host is dominant and in the presence of stressed trees (Hessburg *et al.*, 2001). Closely associated centres were widespread in areas where substantial tree damage or site disturbance had occurred, especially along roads and skid trails (Hessburg *et al.*, 2001). The use of roads and skid trails causes soil compaction and tree damage (Hessburg *et al.*, 2001). Damage to trees and the depleted drainage capacity of compacted soil causes stress in the trees and make them more susceptible to fungal infection (Hessburg *et al.*, 2001).

Long distance spread of *L. wageneri* involves the root feeding bark beetle *H. nigrinus* (Coleoptera: Scolytidae) as a primary vector in Douglas-fir trees (Harrington *et al.*, 1985; Witcoscky and Hansen, 1985). *H. nigrinus* commonly breed in roots of recently dead, dying or stressed host trees. *L. wageneri* sporulates inside insect galleries in infected roots and root collars forming fruiting bodies on gallery walls (Goheen and Cobb, 1978). The sticky spore droplets protrude into gallery lumens and contaminate young adult beetles as they brush against spore droplets in galleries or pupal chambers (Hessburg *et al.*, 1995). The beetles fly from the original tree of infection and are able to tunnel through the soil toward the roots of healthy or dying trees and deposit spores on the root sapwood exposed during feeding (Hessburg *et al.*, 1995). The fungi gain entry to the sapwood though roots and eventually cause disease symptoms.

1.3.4.2. White pine root decline

White Pine Root Decline (WPRD) is a complex disease that has been persistently associated with *L. procerum,* although pathogenicity tests have shown that the fungus is incapable of causing mortality in trees (Wingfield, 1986). WPRD was first reported in the eastern United States where it is found on many species of pine. The symptoms of WPRD include retention of needles, browning of the needles, resin-soaked, black-streaked wood, basal cankers, extended periods of bud break and retarded shoot growth (Carlson, 1994). In cases where the fungus is present it has been vectored by weevils (Coleoptera: Curculionidae) and not by bark beetles, which are less commonly associated with the fungus (Nevill and Alexander, 1992).

The weevils are often attracted by volatiles such as ethanol and turpenes (Chénier and Philogène, 1989) that are released in response to fungal infection (Kelsey, 2001). While *L. procerum* cannot be directly linked to mortality in trees (Eckhardt *et al.*, 2004b; Wingfield, 1986), it is thought to escalate the infestation of the trees with weevils and bark beetles that feed on the trees causing mechanical damage and death. Weevils are capable of carrying inoculum of *L. procerum* into the inner parts of the tree (Nevill and Alexander, 1992) where the fungus infects the xylem tissue. Blockage of the xylem vessels will reduce the water potential of the tree and result in tree death (Joseph *et al.*, 1998).

The complex nature of WPRD has made it difficult to identify a single source for the disease. Soil moisture and temperature have been found to affect the virulence of *L. procerum* in different ways and are factors that have been linked to more severe cases of WPRD (Kelsey, 2001). Stressed trees are more prone to this disease but whether that is because they are more susceptible targets for the root feeding weevils or that they are better targets for infection by fungi or a combination of both can be debated in favour of any combination (Kelsey, 2001) . Thus, the role of *L. procerum* in WPRD is yet to be fully understood (Jacobs and Wingfield, 2001).

1.4. Research rationale

Traditional forms of identification for *Leptographium* include morphological and molecular identification (Jacobs *et al.*, 2000; Lee *et al.*, 2003; Lee *et al.*, 2005; Lu *et al.*, 2008; Witthuhn *et al.*, 1997). The availability of the complete genome sequences for a number of eukaryotes including *Saccharomyces cerevisiae* and human beings has opened doors for new molecular techniques to be

developed. In the past, many techniques were inconceivable due to the lack of complete sequence data. Microarrays require extensive sequence knowledge and careful design but the data gathered from an array far surpasses that obtained through more conventional molecular methods (Lemieux *et al.*, 1998). A reason why this method is not applied in more cases is that it is expensive; it requires a microarray facility and is intricate in its design and optimisation. These characteristics have made the application of microarrays suitable for large facilities and in very specific cases (Schena and Davis, 2000a). However, this technology is becoming more accessible to smaller groups and can be applied to smaller projects.

Developing a diagnostic microarray for *Leptographium* would have far-reaching applications within the genus and for other ascomycetes. Microarray chips are effective tools for studying population structures (Brodie *et al.*, 2006; Stralis-Pavese *et al.*, 2004) and the composition of plant diseases (Tambong *et al.*, 2006). Multiple *Leptographium* species may infect one host simultaneously and through conventional methods, isolations would have to be done in order to study the infection. Using a microarray it is possible to identify all the *Leptographium* species involved in the infection in one simple experiment in a manner that has been applied to other studies (Bodrossy and Sessits, 2004; Volokhov *et al.*, 2002). Through monitoring the performance of the various probes it is possible to track changes in populations and study how they are changing (Loy and Bodrossy, 2006; Stralis-Pavese *et al.*, 2004). *Leptographium* contains several important pine timber pathogens (Eckhardt *et al.*, 2004b; Wagener and Mielke, 1961) and their rapid identification would be valuable to the disease management process.

There are three gene regions – EF1α, ITS and βT – available for probe design for all the known species of *Leptographium* (Jacobs *et al.*, 2001; Jacobs *et al.*, 2006). The *Leptographium* identification chip will be based on probes designed to identify species by their sequence polymorphisms. All three gene regions were used to draw up a multi-gene phylogeny of *Leptographium* species and will be used to design an accurate set of probes to differentiate all the included species.

Pre-synthesised oligonucleotides will be attached to a vapour coated glass slide using printing technology (Schermer, 2000). Probes will be designed using various software programs and strict parameters. Twenty-mer probes will be designed to have a 50% G+C content with polymorphisms located toward the middle of the probe and in some cases synthesised with LNA nucleotides at the sight of important SNPs (Bodrossy and Sessits, 2004; Letowski *et al.*, 2004; You *et al.*, 2006). Probes will be designed to perform under similar hybridisation temperatures (Letowski *et al.*, 2004).

Hybridisation patterns will be analysed to test the accuracy and efficiency of the chip in identifying species of *Leptographium* (Bates *et al*., 2005).

Data analysis is central to the development of a diagnostic chip. During the development stage it is necessary to evaluate the probes not only for their discriminating power but also for their thermodynamic performance under the experimental conditions (Booth *et al.*, 2003; Letowski *et al.*, 2004). Rigorous statistical testing and verification of data is necessary to develop a microarray that is robust enough to detect the target organisms (Allison *et al.*, 2005; Churchill, 2004; Edwards, 2003). There are many statistical programs available to normalise the data and perform the necessary tests for probe performance (Ferrara *et al.*, 2005; Metfies *et al.*, 2008; Quackenbush, 2001). The data gathered from the development of a diagnostic chip for *Leptographium* may be used in the development of similar arrays for other ascomycetes.

References

Adams, A.S., Six, D.L., Adams, S.M., Holben, W.E., 2008. *In Vitro* interactions between yeasts and bacteria and the fungal symbionts of the mountain pine beetle (*Dendroctonus ponderosae*). Microb Ecol. 56, 460-466.

Al-Khaldi, S.F., Myers, K.M., Rasooly, A., Chizhikov, V., 2004. Genotyping of *Clostridium perfringens* toxins using multiple oligonucleotide microarray hybridization. Mol Cell Probes. 18, 359-367.

Allison, D.B., Cui, X., Page, G.P., Sabripour, M., 2005. Microarray data analysis: from disarray to consolidation and consensus. Nat Rev Genet. 7, 55-66.

Anderson, N., Szemes, M., O'Brien, P., Weerd, M.D., Schoen, C., Boender, P., Bonant, P., 2006. Use of hybridization melting kinetics for detecting *Phytophthora* species using three dimensional microarrays: demonstration of a novel concept for the differentiation of detection targets. Mycol Res. 110, 664-671.

Ayres, M.P., Wilkens, R.T., Ruel, J.J., Lombardero, M.J., Vallery, E., 2000. Nitrogen budgets of phloem-feeding bark beetles with and without symbiotic fungi. Ecology. 81, 2198-2210.

Bäckman, A., Lantz, P.-G., Rådström, P., Olcén, P., 1999. Evaluation of an extended diagnostic PCR assay for detection and verification of the common causes of bacterial meningitis in CSF and other biological samples. Molecular and Cellular Probes. 13, 49-60.

Badiee, A., Eiken, H., Steen, V., Lovlie, R., 2003. Evaluation of five different cDNA labelling methods for microarrays using spike controls. BMC Biotechnol. 3, 23.

Bates, S.R., Baldwin, D.A., Channing, A., GiVord, L.K., Hsu, A., Lu, P., 2005. Cooperativity of paired oligonucleotide probes for microarray hybridization assays. Analytical Biochememistry. 342, 59-68.

Beaucage, S.L., 2001. Strategies in the preparation of DNA oligonucleotide arrays for diagnostic applications. Curr Med Chem 8, 1213-1244.

Bentz, B.J., Six, D.L., 2006. Ergosterol Content of Fungi Associated with *Dendroctonus ponderosae* and *Dendroctonus rufipennis* (*Coleoptera: Curculionidae, Scolytinae*). Ann Entomol Soc Am 99, 189-194.

Bodrossy, L., Sessits, A., 2004. Oligonucleotide microarrays in microbial diagnostics. Curr Opin Microbiol. 7, 245-254.

Booth, S.A., Drebot, M.A., Martin, I.E., Ng, L.K., 2003. Design of oligonucleotide arrays to detect point mutations: molecular typing of antibiotic resistant strains of *Neisseria gonorrhoeae* and Hantavirus infected deer mice. Mol Cell Probes. 17, 77-84.

Brazma, A., Hingamp, P., Quackenbush, J., Sherlock, G., Spellman, P., Stoekert, C., Aach, J., Ansorge, W., Ball, C.A., Causton, H.C., Gaasterland, T., Glenisson, P., Holstege, F.C.P., Kim, I.F., Markowitz, V., Matese, J.C., Parkinson, H., Robinson, A., Sarkans, U., Schulze-Kremer, S., Stewart, J., Taylor, R., Vilo, J., Vingron, M., 2001. Minimum Information about a Microarray Experiment (MIAME) - towards standards for microarray data. Nat Genet 29, 365-371.

Brodie, E.L., De Santis, T.Z., Joyner, D.C., Baek, S.M., Larsen, J.T., Andersen, G.L., Hazen, T.C., Richardson, P.M., Herman, D.J., Tokunaga, T.K., Wan, J.M., Firestone, M.K., 2006. Application of a high-density oligonucleotide microarray approach to study bacterial population dynamics during uranium reduction and reoxidation. Appl Environ Microb. 72, 6288-6298.

Burton, J.E., Oshota, O.J., North, E., Hudson, M.J., Polyanskaya, N., Brehm, J., Lloyd, G., Silman, N.J., 2005. Development of a multipathogen oligonucleotide microarray for detection of *Bacillus anthracis*. Mol Cell Probes. 19, 349-357.

Burton, J.E., Oshota, O.J., Silman, N.J., 2006. Differential identification of *Bacillus anthracis* from environmental *Bacillus* species using microarray analysis. J Appl Microbiol. 101, 754-763.

Cai, H., White, P.S., Torney, D., Deshpande, A., Wang, Z., Marrone, B., Nolan, J.P., 2000. Flow cytometry-based minisequencing: a new platform for high-throughput single-nucleotide polymorphism scoring. Genomics. 66, 135-143.

Call, D.R., Brockman, F.J., Chandler, D.P., 2001. Detecting and genotyping *Escherichia coli* O157:H7 using multiplexed PCR and nucleic acid microarrays. Int J Food Microbiol. 67, 71-80.

Carlson, J.A., 1994. Procerum root disease physiology and disease interactions with ozone, Plant Pathology, Physiology, and Weed Science, Thesis (Ph.D.), Virginia Tech, Virginia, USA., pp. 104.

Chakrabarti, R., Schutt, C.E., 2001. The enhancement of PCR by low molecular weight amides. Nucleic Acids Res. 29, 2377-2381.

Chandler, D.P., Jarrell, A.E., 2002. Enhanced nucleic acid capture and flow cytometry detection with peptide nucleic acid probes and tunable-surface microparticles. Anal Biochem. 312, 182-190.

Chénier, J.V.R., Philogène, B.J.R., 1989. Field responses of certain forest *Coleoptera* to conifer monoterpenes and ethanol J Chem Ecol. 15, 1729-1745.

Churchill, G.A., 2004. Using ANOVA to analyze microarray data. Biotechniques. 37, 173-177.

Cleven, B.E.E., Palka-Santini, M., Gielen, J., Meembor, S., Krönke, M., Krut, O., 2006. Identification and characterization of bacterial pathogens causing bloodstream infections by DNA microarray. J Clin Microbiol. 44, 2389–2397.

Clewley, J.P., 2004. A role for arrays in Clin Virology: fact or fiction? Clin Viro. 29, 2-12.

Conner, B.J., Reyes, A.A., Morin, C., Itakura, K., Teplitz, R.L., Wallace, R.B., 1983. Detection of sickle cell Bˢ-globin allele by hybridization with synthetic oligonucleotides. PNAS USA. 80, 278-282.

Couzinet, S., Jay, C., Barras, C., Vachon, R., Vernet, G., Ninet, B., Jan, I., Minazio, M.-A., Francois, P., Lewc, D., Troesch, A., Schrenzel, J., 2005. High-density DNA probe arrays for identification of *Staphylococci* to the species level. J Microbiol meth. 61, 201-208.

Currie, C.R., 2001. A community of ants, fungi and bacteria: A multilateral approach to studying symbiosis. Annu Rev Microbiol. 55, 357-380.

Currie, C.R., Scott, J.A., Summerbell, R.C., Malloch, D., 1999. Fungus-growing ants use antibiotic-producing bacteria to control garden parasites. Nature 398, 701-704.

Davignon, L., Walter, E.A., Mueller, K.M., Barrozo, C.P., Stenger, D.A., Lin, B., 2005. Use of resequencing oligonucleotide microarrays for identification of *Streptococcus pyogenes* and associated antibiotic resistance determinants. J Clin Microbiol. 43, 5690-5695.

Do, J.H., Choi, D.K., 2007. cDNA labeling strategies for microarrays using fluorescent dyes. Eng Life Sci. 7, 26-34.

Eckhardt, L.G., Goyer, R.A., Klepzig, K.D., Jones, J.P., 2004a. Interactions of *Hylastes* species (*Coleoptera: Scolytidae*) with *Leptographium* species associated with Loblolly pine decline. J Econ Entomol. 97, 468-474.

Eckhardt, L.G., Jones, J.P., Klepzig, K.D., 2004b. Pathogenicity of *Leptographium* species associated with Loblolly pine decline. Plant Dis. 88, 1174-1178.

Edman, C.F., Raymond, D.E., Wu, D.J., Tu, E., Sosnowski, R.G., Butler, W.F., Nerenberg, M., Heller, M.J., 1997. Electric field directed nucleic acid hybridization on microchips. Nucleic Acids Res. 25, 4907–4914.

Edwards, D., 2003. Non-linear normalization and background correction in one-channel cDNA microarray studies. Bioinformatics. 19, 825-833.

Edwards, M.C., Gibbs, R.A., 1994. Multiplex PCR: advantages, development, and applications. Genome Res. 3, 65-75.

Elnifero, E.M., Aashshi, A.M., Cooper, R.J., Klapper, P.E., 2000. Multiplex PCR: optimization and application in diagnostic virology. Clin Microbiol Rev. 13, 559-570.

Eom, H.S., Hwang, B.H., Kimb, D.-H., Lee, I.-B., Kim, Y.H., Cha, H.J., 2006. Multiple detection of food-borne pathogenic bacteria using a novel 16S rDNA-based oligonucleotide signature chip. Biosens. Bioelectron. 22, 845-853.

Erdogan, F., Kirchner, R., Mann, W., Ropers, H.H., Nuber, U.A., 2001. Detection of mitochondrial single nucleotide polymorphisms using a primer elongation reaction on oligonucleotide microarrays. Nucleic Acids Res. 29, e36.

Fan, J.-B., Hu, S.X., Craumer, W.C., Barker, D.L., 2005. BeadArray™-based solutions for enabling the promise of pharmacogenomics. Biotechniques Supplement. 39, 583-588

Ferrara, G.B., Delfino, L., Masulli, F., Rovett, S., Sensi, R., 2005. A fuzzy approach to image analysis in HLA typing using oligonucleotide microarrays. Fuzzy Set. Syst. 152, 37-48.

Fujita, S.-I., Senda, Y., Nakaguchi, S., Hashimoto, T., 2001. Multiplex PCR using internal transcribed spacer 1 and 2 regions for rapid detection and identification of yeast strains. J Clin Microbiol. 39, 3617-3622.

Fukushima, M., Kakinuma, K., Hayashi, H., Nagai, H., Ito, K., Kawaguchi, R., 2003. Detection and identification of *Mycobacterium* species isolates by DNA microarray. J Clin Microbiol. 41, 2605-2615.

Glass, N.L., Donaldson, G.C., 1995. Development of primer sets designed for use with the PCR to amplify conserved genes from filamentous ascomycetes. Appl Environ Microb. 61, 1323-1330.

Goidanich, G., 1936. Il genere di Ascomiceti *Grosmannia* G. Goid. Bollettina della Stazione di Patologia Vegetale di Roma. 16, 26-60.

Goheen, D.J., Cobb, F.W.J., 1978. Occurrence of *Verticicladiella wagenerii* and its perfect state, *Ceratocystis wageneri* sp. nov., in insect galleries. Phytopathology. 68, 1192-1195.

Goheen, D.J., Hansen, E.M., 1978. Black Stain Root Disease in Oregon and Washington. Plant Dis. Rep. 62, 1098-1102.

Gunderson, K.L., Kruglyak, S., Graige, M.S., Garcia, F., Kermani, B.G., Zhao, C., Che, D., Dickinson, T., Wickham, E., Bierle, J., Doucet, D., Milewski, M., Yang, R., Siegmund, C., Haas, J., Zhou, L., Oliphant, A., Fan, J.-B., Barnard, S., Chee, M.S., 2006. Decoding randomly ordered DNA arrays. Genome Res. 14, 870-877.

Hamelin, R.C., Bérubé, P., Gignac, M., Bourassa, M., 1996. Identification of root rot fungi in nursery seedlings by nested multiplex PCR. Appl Environ Microb. 62, 4026-4031.

Harrington, T., Cobb, F.J., 1983. Pathogenicity of *Leptographium* and *Verticicladiella* spp. Isolated from roots of western North American conifers. Phytopathology. 73, 596-599.

Harrington, T., Cobb, F.J., 1984. Host specialization of three morphological variants of *Verticicladiella wageneri*. Phytopathology. 74, 286-290.

Harrington, T.C., Cobb. F.W., Lownsbery, J.W., 1985. Activity of *Hylastes nigrinus*, a vector of *Verticicladiella wageneri*, in thinned stands of Douglas-fir. Can. J. For. Res. 15, 519-523.

Harrington, T.C., Cobb, F.W., 1986. Varieties of *Verticicladiella wageneri*. Mycologia.. 78, 562-567.

Harrington, T.C., 1988. *Leptographium* species, their distributions, hosts, and insect vectors. In: T. C. Haringtom, J. F. W. Cobb (Eds.), *Leptographium* root diseases on conifers, APS Press, St. Paul, Minnesota, USA, pp. 1-39.

Harrington, T.C., 1993. Diseases of conifers caused by *Ophiostoma* and *Leptographium*. In: M. J. Wingfield, Seifert, K. A., and Webber, J. F. (Ed.), *Ceratocystis* and *Ophiostoma*: Taxonomy, Ecology and Pathogenicity, APS Press, Minnesota, USA, pp. 161-172.

Hausner, G., Iranpour, M., J.-J. Kim, C.B., Davis, C.N., Gibb, E.A., Reid, J., Loewen, P.C., Hopkin, A.A., 2005. Fungi vectored by the introduced bark beetle *Tomicus piniperda* in Ontario, Canada, and comments on the taxonomy of *Leptographium lundbergii*, *Leptographium terebrantis*, *Leptographium truncatum*, and *Leptographium wingfieldii*. Can J Bot. 83, 1222-1237.

Hausner, G., Reid, J., Klassen, G.R., 2000. On the phylogeny of members of *Ceratocystis* s.s. and *Ophiostoma* that possess different anamorphic states, with emphasis on the anamorph genus *Leptographium*, based on partial ribosomal DNA sequences. Can. J. Bot. 78, 903-916.

Hessburg, P.F., Goheen, D.J., Bega, R.V., 1995. Black stain root disease of conifers, Forest Insect and Disease Leaflet, Vol. 145, U. S. Department of Agriculture, Washington, D. C.

Hessburg, P.F., Goheen, D.J., Koester, H., 2001. Association of black stain root disease with roads, skid trails, and precommercial thinning in southwest Oregon. West J App For. 16, 127-135.

Hessburg, P.F., Hansen, E.M., 1987. Pathological anatomy of black stain root disease of Douglas fir. Can. J. Bot. 65, 962-971.

Huang, A., Li, J.-W., Shen, Z.-Q., Wang, X.-W., Jin, M., 2006. High-Throughput identification of clinical pathogenic fungi by hybridization to an oligonucleotide microarray. J Clin Microbiol. 44, 3299-3305.

Hyyro, H., Juhola, M., Vihinen, M., 2005. On exact string matching of unique oligonucleotides. Comput Biol Med. 35, 173-181.

Jackson, C.R., Fedorka-Cray, P.J., Barrett, J.B., 2004. Use of a genus- and species-specific multiplex PCR for identification of *Enterococci*. J Clin Microbiol. 42, 3558-3565.

Jacobs, A., Coetzee, M., Wingfield, B.D., Jacobs, K., Wingfield, M.J., 2003. Phylogenetic relationships among *Philocephalai* species and other ascomycetes. Mycologia.. 95, 637-645.

Jacobs, K., 1999. The genus *Leptographium:* A critical Taxonomic analysis, Forestry and Agricultural Biotechnology Institute (FABI) and the Department of Microbiology and Plant Pathology, University of Pretoria, Pretoria, South Africa.

Jacobs, K., Bergdahl, D. R., Wingfield, M. J., Halik, S., Seifert, K. A., Bright, D. E., and Wingfield, B. D., 2004. *Leptographium wingfieldii* introduced into North America and found associated with exotic *Tomicus piniperda* and native bark beetles. Mycol Res. 108, 411-418.

Jacobs, K., Solheim, H., Wingfield, B.D., Wingfield, M.J., 2005. Taxonomic re-evaluation of *Leptographium lundbergii* based on DNA sequence comparisons and morphology. Mycol Res. 10, 1149-1161.

Jacobs, K., Wingfield, B.D., Wingfield, M.J., 2006. *Leptographium:* Anamorphs of *Grosmannia*, The Ophiostomatoid Fungi: Expanding Frontiers, Morton Bay Research Station, North Stradbroke Island, Brisbane, Australia.

Jacobs, K., Wingfield, M.J., 2001. *Leptographium* species: tree pathogens, insect associates and agents of blue stain, APS Press, Minnesota, USA.

Jacobs, K., Wingfield, M.J., Pashenova, N.V., Vetrova, V.P., 2000. A new *Leptographium* species from Russia. Mycol Res. 104, 1524-1529.

Jacobs, K., Wingfield, M.J., Wingfield, B.D., 2001. Phylogenetic relationships in *Leptographium* based on morphological and molecular characters. Can J Bot. 79, 719-732.

Jacobs, K.A., Rudersdorf, R., Neill, S.D., Dougherty, J.P., Brown, E.L., Fritsch, E.F., 1998. The thermal stability of oligonucleotide duplexes is sequence independent in tetraalkylanmonium salt solutions: application to identifying recombinant DNA clones. Nucleic Acids Res. 16, 4638-4652.

Joseph, G., Kelsey, R.G., Thies, W.G., 1998. Hydraulic conductivity in roots of ponderosa pine infected with black-stain (*Leptographium wageneri*) or annosus (*Heterobasidion annosum*) root disease. Tree Physiol. 18, 333-339.

Kaderali, L., Deshpane, A., Nolan, J.P., Scott White, P., 2003. Primer-design for multiplexed genotyping. Nucleic Acids Res. 31, 1796-1802.

Kane, M.D., Jatkoe, T.A., Stumpg, C.R., Lu, J., Thomas, J.D., madore, S.J., 2000. Assessment of the sensitivity and specificity of oligonucleotide (50mer) microarrays. Nucleic Acids Res. 28, 4552-4557.

Kelsey, R.G., 2001. Chemical indicators of stress in trees their ecological significance and implication for forestry in Eastern Oregon and Washington. Northwest Sci. 25, 70-76.

Kendrick, W.B., 1962. The *Leptographium* complex. *Verticicladiella* S. Hughes. Can. J. Bot. 40, 771-797.

Keramas, G., Bang, D.D., Lund, M., Madsen, M., Rasmussen, S.E., Bunkenborg, H., Telleman, P., Christensen, C.B.V., 2003. Development of a sensitive DNA microarray suitable for rapid detection of *Campylobacter* spp. Mol Cell Probes. 17, 187-196.

Kerr, M.K., Martin, M., Churchill, G.A., 2000. Analysis of variance for gene expression Microarray data. J Comput Biol. 7, 819-837.

Kim, J.A., Takahashi, Y., Tanaka, R., Fukushima, K., Nishimura, K., Miyaji, M., 2001. Identification and subtyping of *Trichophyton mentagrophytes* by random amplified polymorphic DNA. Mycoses. 44, 157-1665.

Klepzig, K.D., Moseri, J.C., Lombardero, F.J., Hofstetter, R.W., Ayres, M.P., 2001. Symbiosis and competition: complex interactions among beetles, fungi and mites. Symbiosis. 30, 83-96.

Klepzig, K.D., Six, D.L., 2004. Bark beetle-fungal Symbiosis: context dependency in complex associations. Symbiosis. 37, 189-205.

Kostic, T., Weilharter, A., Rubino, S., Delogu, G., Uzzau, S., Rudi, K., Sessitsch, A., Bodrossy, L., 2006. A microbial diagnostic microarray technique for the sensitive detection and identification of pathogenic bacteria in a background of nonpathogens. Anal Biochem. 360, 244-254.

Landegren, U., Nilsson, M., Kwok, P., 1998. Reading bits of genetic information: methods for single-nucleotide polymorphism Analysis. Genome Res. 8, 769-776.

Lane, S., Evermann, J., Logea, F., Call, D.R., 2004. Amplicon secondary structure prevents target hybridization to oligonucleotide microarrays. Biosens. Bioelectron. 20, 728-735.

Lee, S., Kim, J.-J., Breuil, C., 2006. Pathogenicity of *Leptographium longiclavatum* associated with *Dendroctonus ponderosae* to *Pinus contorta.* Can J For Res. 36, 2864-2872.

Lee, S., Kim, J.-J., Fung, S., Breuil, C., 2003. A PCR-RFLP marker distinguishing *Ophiostoma clavigerum* from morphologically similar *Leptographium* species associated with bark beetles. Can J Bot. 81, 1104-1113.

Lee, S., Kim, J.J., Breuil, C., 2005. *Leptographium longiclavatum* sp. nov., a new species associated with the mountain pine beetle, *Dendroctonus ponderosae*. Mycol Res. 109, 1162-1170.

Leinberger, D.M., Schumacher, U., Autenrieth, I.B., Bachmann, T.T., 2005. Development of a DNA microarray for detection and identification of fungal pathogens involved in invasive mycoses. J Clin Microbiol. 43, 4943–4953.

Lemieux, B., Aharoni, A., Schena, M., 1998. Overview of DNA chip technology. Mol Breeding. 4, 277-289.

Letowski, J., Brousseau, R., Masson, L., 2004. Designing better probes: effect of probe size, mismatch position and number on hybridization in DNA oligonucleotide microarrays. J. Microbiol. Methods. 57, 269-278.

Li, F., Stormo, G.D., 2001. Selection of optimal DNA oligos for gene expression arrays. Bioinformatics. 17, 1067-1076.

Liu, W.-T., Guo, H., Wu, J.-H., 2006. Effects of target length on the hybridization efficiency and specificity of rRNA-based oligonucleotide microarrays. Appl Environ Microb. 73, 73-82.

Loy, A., Bodrossy, L., 2006. Highly parallel microbial diagnostics using oligonucleotide microarrays (Review). Clin Chim Acta. 363, 106-119.

Loy, A., Lehner, A., Lee, N., Adamczy, J., Meier, H., Ernst, J., Schleifer, K.H., Wagner, M., 2002. Oligonucleotide microarray for 16S rRNA gene-based detection of all recognized lineages of sulfate-reducing prokaryotes in the environment. Appl Environ Microb. 68, 5064-5081.

Lu, Q., Decock, C., Zhang, X.Y., Maraite, H., 2008. *Leptographium sinoprocerum* sp. nov., an undescribed species associated with *Pinus tabuliformis-Dendroctonus valens* in northern China. Mycologia.. 100, 275-290.

Luo, G., Mitchell, T.G., 2002. Rapid identification of pathogenic fungi directly from cultures by using multiplex PCR. J Clin Microbiol. 40, 2860-2865.

Majer, D., Mithen, R., Lewis, B.G., Vos, P., Oliver, R.P., 1996. The use of AFLP fingerprinting for the detection of genetic variation in fungi. Mycol Res. 100, 1107-1111.

Malloch, D., Blackwell, M., 1993. Dispersal biology of the ophiostomatoid fungi. In: M. J. Wingfield, K. A. Seifert, J. F. Webber (Eds.), *Ceratocystis* and *Ophiostoma*: taxonomy, ecology and pathogenicity, APS, Minnesota, USA, pp. 185-206.

Manduchi, E., Scearce, L.M., Brestelli, J.E., Grant, G.R., Kaestner, K.H., Stoeckert, C.J., 2002. Comparison of different labeling methods for two-channel high-density microarray experiments. Physiol. Genomics. 10, 169-179.

Markoulatos, P., Siafakas, N., Moncany, M., 2002. Multiplex polymerase chain reaction: A practical approach. J Clin Lab Anal. 16, 47-51.

Martens, M., Weidner, S., Linke, B., de Vos, P., Gillis, M., Willems, A., 2007. A prototype taxonomic microarray targeting the rpsA housekeeping gene permits species identification within the rhizobial genus *Ensifer*. 30, Syst Appl Microbiol, 390-400.

Martin, C., Roberts, D., van der Weide, M., Rossau, R., Jannes, G., Smith, T., Maher, M., 2000. Development of a PCR-based line probe assay for identification of fungal pathogens. J Clin Microbiol. 38, 3735-3742.

Masuya, H., Jacobs, A., Wingfield, M.J., Kubono, T., Ichihara, Y., 2004. *Leptographium pruni,* sp. nov. from bark beetle-infested *Prunus jamasakura* in Japan. Mycologia.. 96, 548 - 557.

Maughan, N.J., Lewis, F.A., Smith, V., 2001. An introduction to microarrays. J Pathol. 195, 3-6.

Maynard, C., Berthiaume, F., Lemarchand, K., Harel, J., Payment, P., Bayardelle, P., Masson, L., Brousseau, R., 2005. Waterborne pathogen detection by use of oligonucleotide-based microarrays. Appl Environ Microb. 71, 8548-8557.

McGall, G., Labadie, J., Brock, P., Wallraff, G., Nguyen, T., Hinsberg, W., 1996. Light-directed synthesis of high-density oligonucleotide arrays using semiconductor photoresists. PNAS USA. 93, 13555–13560.

Metfies, K., Bborsutzki, P., Gescher, C., Medlin, L.K., Frickenhaus, S., 2008. PHYLOCHIPANALYSER — a program for analysing hierarchical probe sets. Mol Ecol Resour. 8, 99-102.

Mueller, U.G., Dash, D., Rabeling, C., Rodrigues, A., 2008. Coevolution between Attine ants and actiomycete bacteria: a reevaluation. Evolution. 62, 2894-2912.

Myers, K.M., Gaba, J., Al-Khaldi, S.F., 2006. Molecular identification of *Yersinia enterocolitica* isolated from pasteurized whole milk using DNA microarray chip hybridization. Mol Cell Probes. 20, 71-80.

Nagan, N., O'Kane, O.J., 2001. Validation of single nucleotide polymorphism genotyping assay for the human serum rasonase gene using electronically active customized microarrays. Clin Biochem. 34, 589-592.

Nevill, R.J., Alexander, S.A., 1992. Transmission of *Leptographium procerum* to eastern white pine by *Hylobius pales* and *Pissodes nemorensis* (Coleoptera: Curculionidae). Plant Dis. 76, 307-310.

Nicolaisen, M., Justesen, A.F., Thrane, U., Skouboe, P., Holmstrom, K., 2005. An oligonucleotide microarray for the identification and differentiation of trichothecene producing and non-producing Fusarium species occurring on cereal grain. J Microbiol Meth. 62, 57-69.

Nordberg, E.K., 2005. YODA: selecting signature oligonucleotides. Bioinformatics. 21, 1365-1370.

Okada, G., Seifert, K.A., Takematsu, A., Yamaoka, Y., 1998. A molecular phylogenetic reappraisal of the *Graphium* complex based on 18S rDNA sequences. Can. J. Bot. 76, 1495-1506.

Paine, T.D., Raffa, K.F., Harrington, T.C., 1997. Interactions among scolytid bark beetles, their associated fungi, and live host conifers. Annu Rev Entomol. 42, 179-206.

Peverieri, G.S., Capretti, P., Tiberi, R., 2006. Associations between *Tomicus destruens* and *Leptographium* spp. in *Pinus pinea* and *P. pinaster* stands in Tuscany, central Italy. Forest Pathol. 36, 14-20.

Quackenbush, J., 2001. Computational analysis of microarray data. Nature Rev. 2, 418-427.

Quackenbush, J., 2002. Microarray data normalization and transformation. Nature Genet Supp. 32.

Raffa, K.F., Phillips, T.W., Salom, S.C., 1993. Strategies and mechanisms of host colonization by bark beetles. In: T. D. Schowalter, G. M. Filip (Eds.), Beetle-pathogen interaction in conifer forests, APS Press, Corvallis, pp. 104-129.

Reay, S.D., Walsh, P.J., Ram, A., Farrel, R.L., 2002. The invasion of *Pinus radiata* seedlings by sapstain fungi, following attack by the Black Pine Bark Beetle, *Hylastes ater* (Coleoptera: Scolytidae). Forest Ecol Manag. 165, 47-56.

Reiss, B.D., Freeman, R.G., Walton, I.D., Norton, S.M., Smith, P.C., Stonas, W.G., Keating, C.D., Natan, M.J., 2002. Electrochemical synthesis and optical readout of striped metal rods with submicron features. J. Electroanayl. Chem. 522.

Rogers, Y.-H., Jiang-Baucom, P., Huang, Z.-J., Bogdanov, V., Anderson, S., Boyce-Jacino, M.T., 1999. Immobilization of oligonucleotides onto a glass support via disulfide bonds: A method for preparation of DNA Microarrays. Anal Biochem. 266, 23-30.

Ross, D.W., Solheim, H., 1997. Pathogenicity to Douglas-fir of *Ophiostoma pseudotsugae* and *Leptographium abietinum*, fungi associated with the Douglas-fir beetle. Can. J. For. Res.27, 39-43.

Roth, S.B., Jalava, J., Ruuskanen, O., Ruohola, A., Nikkari, S., 2004. Use of an oligonucleotide array for laboratory diagnosis of bacteria responsible for acute upper respiratory infections. J Clin Microbiol. 42, 4268-4274.

Rouillard, J.-M., Zuker, M., Gulari, E., 2003. OligoArray 2.0: design of oligonucleotide probes for DNA microarrays using a thermodynamic approach. Nucleic Acids Res. 21, 3057-3062.

Saiki, R.K., Walsh, P.S., Levenson, C.H., Erlich, H.A., 1989. Genetic analysis of amplified DNA with immobilized sequence-specific oligonucleotide probes. PNAS USA. 86, 6230-6234.

Schena, M., and Davis. R. W., 2000a. Genes, genomes and chips. In: M. Schena (Ed.), DNA Microarrays, A practical approach, OUP, New York, pp. 1-15.

Schena, M., and Davis. R. W., 2000b. Technological standard for microarray research. In: M. Schena (Ed.), Microarray Biochip Technology, EP, California, pp. 1-19.

Schena, M., Shalon, D., Davis, R.W., Brown, P.O., 1995. Quantitative monitoring of gene expression patterns with a complementary DNA microarray. Science. 270, 467-470.

Schermer, M.J., 2000. Confocal scanning microscopy in microarray detection. In: M. Schena (Ed.), DNA microarrays, a practical approach, OUP, New York, pp. 16-42.

Schweigkofler, W., Otrosina, W.J., Smith, S.L., Cluck, D.R., 2005. Detection and quantification of *Leptographium wageneri*, the cause of black-stain root disease, from bark beetles (Coleoptera: Scolytidae) in Northern California using regular and real-time PCR. Can. J. For. Res. 35, 1798-1808.

Scott, J.J., Oh, D.-C., Yuceer, M.C., Klepzig, K.D., Clardy, J., Currie, C.R., 2008. Bacterial protection of beetle-fungus mutualism. Science. 322, 63.

Seifert, K.A., 1993. Sapstain of commercial lumber by species of *Ophiostoma* and *Ceratocystis*. In: M. J. Wingfield, K. A. Seifert, J. F. Webber (Eds.), Ceratocystis and Ophiostoma: Taxonomy, Ecology and Pathogenicity, APS, Minnesota, USA, pp. 141-152.

Sergeev, N., Distler, M., Vargas, M., Chizhikov, V., Herold, K.E., Rasooly, A., 2006. Microarray analysis of *Bacillus cereus* group virulence factors. J Microbiol meth. 65, 488– 502.

Sha, M.Y., Walton, I.D., Norton, S.M., Taylor, M., Yamanaka, M., Natan, M.J., Xu, C., Drmanac, S., Huang, S., Borcherding, A., Drmanac, R., Penn, S.G., 2005. Multiplexed SNP genotyping using nanobarcode particle technology. Anal Bioanal Chem. 384, 658-666.

Siemaszko, W., 1939. Fungi associated with bark beetles in Poland [In Polish]. Planta Polonica. 1, 1-54.

Six, D.L., Bentz, B.J., 2003. Fungi associated with the North American spruce beetle, *Dendroctonus rufipennis*. Can. J. For. Res. 33, 1815-1820.

Six, D.L., Bentz, B.J., 2007. Temperature determines symbiont abundance in a multipartite bark beetle-fungus ectosymbiosis. Microb ecol. 54, 112-118.

Six, D.L., Paine, T.D., 1999. Phylogenetic comparison of ascomycete mycangial fungi and *Dendroctonus* bark beetles (Coleoptera: Scolytidae). Ann Entomol Soc Am 92, 159-166.

Solheim, H., 1995. Early stages of blue-stain fungus invasion of lodgepole pine sapwood following mountain pine beetle attack. Can J Bot. 73, 70-74.

Solheim, H., Krokene, P., Langström, B., 2001. Effects of growth and virulence of associated bluestain fungi on host colonisation behaviour of the pine shoot beetles *Tomicus minor* and *T. piniperda*. Plant Pathol. 50, 111-116.

Stenger, D.A., Andreadis, J.D., Vora, G.J., Pancrazio, J.J., 2002. Potential applications of DNA microarrays in biodefense-related diagnostics. Curr Opin Biotech. 13, 208-212.

Stralis-Pavese, N., Sessitsch, A., Weilharter, A., Reichenauer, T., Riesing, J., Csontos, J., Murrell, J.C., Bodrossy, L., 2004. Optimization of diagnostic microarray for application in analysing landfill methanotroph communities under different plant covers. Environ Microbiol. 6, 347-363.

Syvånen, A., 2001. Accessing genetic variation: Genotyping single nucleotide polymorphisms. Nature Genet Supp. 2, 930-944.

Tambong, J.T., de Cock, A.W.A.M., Tinker, N.A., Lévesque, C.A., 2006. Oligonucleotide array for identification and detection of *Pythium* Species. Appl Environ Microb. 72, 2691-2706.

Taylor, J.W., Jacobson, D.J., Kroken, S., Kasuga, T., Geise, D.M., Hibbett, D.S., Fisher, M.C., 2000. Phylogenetic species recognition and species concepts in fungi. Fungal Genet Biol. 31, 21-32.

Theriault, T.P., Winder, S.C., Gamble, R.C., 2000. Application of ink-jet printing technology to the manufacture of molecular arrays. In: M. Schena (Ed.), DNA microarrays, a practical approach, OUP, Oxford, pp. 101-119.

Tran, S.D., Rudney, J.D., 1996. Multiplex PCR using conserved and species-specific 16S rRNA gene primers for simultaneous detection of *Actinobacillus actinomycetemcomitans* and *Porphyromonas gingivalis*. J Clin Microbiol. 34, 2674-2678.

Troesch, A., Nguyen, H., Miyada, C.G., Desvarenne, S., Gineras, T.R., Kaplan, P.M., Cros, P., Mabilet, C., 1999. *Mycobacterium* species identification and rifampin resistance testing with high-density DNA probe arrays. J Clin Microbiol. 37, 49-55.

Upadhyay, H.P., 1993. Classification of the Ophiostomatoid fungi. In: M. J. Wingfield, K. A. Seifert, J. F. Webber (Eds.), *Ceratocystis* and *Ophiostoma*: Taxonomy, ecology and pathogenicity, APS, Minnesota, USA, pp. 7-12.

Volokhov, D., Pomerantsev, A., Violetta Kivovich, Rasooly, A., Chizhikov, V., 2004. Identification of *Bacillus anthracis* by multiprobe microarray hybridization. Diagn Micr Infec Dis. 49, 163-171.

Volokhov, D., Rasooly, A., Chumakov, K., Chizhikov, V., 2002. Identification of *Listeria* species by microarray-based assay. J Clin Microbiol. 40, 4720-4728.

Vora, G.J., Meador, C.E., Stenger, D.A., Andreadis, J.D., 2004. Nucleic Acid Amplification Strategies for DNA Microarray-Based Pathogen Detection. Appl Environ Microb. 70, 3047-3054.

Wagener, W.W., Mielke, J.L., 1961. A staining fungus root disease of ponderosa, Jeffrey and pinyon pines. Plant Dis Rep. 45, 831-835.

Wagner, A., Blackstone, N., Cartwright, P. *et al.*, 1994. Surveys of gene families using polymerase chain reaction: PCR selection and PCR drift. Syst Biol. 43.

Wallace, R.B., Shaffer, J., Murphy, R.F., Bonner, J., Hirose, T., Itakura, K., 1979. Hybridization of synthetic oligodeoxyribonucleotides to 4 X 174 DNA: the effect of single base pair mismatch. Nucleic Acids Res. 6, 3543-3557.

White, T.J., Bruns, T., Taylor, J., 1990. Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. In: A. Innis, D. Gelfand, J. Sninsky, T. White (Eds.), PCR Protocols: a guide to methods and applications, APS, San Diego, USA. pp. 315-322.

Williams, B.A., Gwirtz, R.M., Wold, B.J., 2004. Genomic DNA as a cohybridization standard for mammalian microarray measurements. Nucleic Acids Res 32, e81.

Wilson, W.J., Strout, C.L., DeSantis, T.Z., Stilwell, J.L., Carrano, A.V., Andersen, G.L., 2002. Sequence-specific identification of 18 pathogenic microorganisms using microarray technology. Mol Cell Probes. 16, 119-127.

Wingfield, M.J., 1986. Pathogenicity of *Leptographium procerum* and *L. terebrantis* on *Pinus strobus* seedlings and established trees, European J Forest Path, 16, 299-308.

Witcoscky, J.J., Hansen, E.M., 1985. Root colonising insects recovered from Douglas-fir in various stages of decline due to black-stain root disease. Phytophathology. 75, 399-402.

Witthuhn, R.C., Wingfield, B.D., Wingfield, M.J., Harrington, T.C., 1997. Comparison of three varieties of *Leptographium wageneri* using random amplified polymorphic DNA. S Afr J Bot. 63, 198-200.

Xu, J., Miao, H., Wu, H., Huang, W., Tang, R., Qiu, M., Wen, J., Zhu, S., Li, Y., 2006. Screening genetically modified organisms using multiplex-PCR coupled with oligonucleotide microarray. Biosens Bioelect. 22, 71-77.

Xu, W., McDonough, M.C., Erdman, D.D., 2000. Species-specific identification of human adenoviruses by a multiplex PCR assay. J Clin Microbiol. 38, 4114-4120.

You, Y., Moreira, B.G., Behlke, M.A., Owczarzy, R., 2006. Design of LNA probes that improve mismatch discrimination. Nucleic Acids Res. 34, e60.

Zambino, P.J., Harrington, T.C., 1989. Isozyme variation within and among host-specialized varieties of *Leptographium wageneri*. Mycologia.. 81, 122-133.

Zambino, P.J., Harrington, T.C., 1992. Correspondence of isozyme characterization with morphology in the asexual genus *Leptographium* and taxonomic implications. Mycologia. 84, 12-25.

Zhou, X.D., de Beer, Z.W., Wingfield, B.D., Wingfield, M.J., 2001. Ophiostomatoid fungi associated with three pine-infesting bark beetle in South Africa. Sydowia. 53, 290-300.

Zhou, X.D., Jacobs, K., Morelet, M., Ye, H., Lieutier, F., Wingfield, M.J., 2000. A new *Leptographium* species associated with *Tomicus piniperda* in south-western China. Mycoscience. 41, 573-578.

Zien, A., Aigner, T., Zimmer, R., Lengauer, T., 2001. Centralization: A new method for the normalisation of gene expression data. Bioinformatics. 17, s323-331.

Zipfel, R.D., de Beer, Z.W., Jacobs, K., Wingfield, B.D., Wingfield, M.J., 2006. Multigene phylogenies define Ceratocystiopsis and Grosmannia distinct from Ophiostoma. Stud. Mycology. 55, 75-97.

2. Chapter 2: A prototype species diagnostic oligonucleotide microarray used to identify *Leptographium dryocoetidis*, *L. elegans* and *L. leptographioides*

Natalie R. Van Zuydam[a], Brenda D. Wingfield[a, *], Karin Jacobs[c], E. Steenkamp[b] and Michael J. Wingfield[b]

[a]*Department of Genetics,* [b]*Department of Microbiology and Plant Pathology, Forestry and Agricultural Biotechnology Institute (FABI), University of Pretoria, Pretoria, 0002, South Africa*

[c]*Department of Microbiology, University of Stellenbosch, Private Bag X-1, Matieland, Stellenbosch, South Africa*

Abstract

Leptographium is a relatively small and well-defined genus of the ascomycetes. Morphological identification of *Leptographium* species is complicated and species are thus identified using a combination of molecular and morphological characters. The aim of this study was to test the application of a species diagnostic array to the identification of *Leptographium* species. A prototype array consisting of short oligonucleotides was designed to identify 26 species of *Leptographium*. The 20-mer probes were designed from available sequence data for *Leptographium*, from partial sequences of the internal transcribed spacer (ITS2), β-tubulin (βT) and translational elongation factor 1-α (EF1α). Targets were amplified from *Leptographium dryocoetidis*, *L. elegans* and *L. leptographioides* genomic DNA using species-specific primers and labelled using an indirect labelling method. Targets were hybridised to the array in triplicate experiments and the data were combined and analysed. The prototype array successfully identified the three species included in this study and indicated that microarrays could be a useful tool for identification within the genus and for other ascomycetes.

2.1. Introduction

Fungal species can be identified by sequencing phylogenetically informative genes, conducting amplified fragment length polymorphism (AFLP) analysis, randomly amplified polymorphic DNA (RAPD) and restriction length fragment length polymorphism (RFLP) analysis (De Vos *et al.*, 2007; Kim *et al.*, 2001; Lee *et al.*, 2003; Majer *et al.*, 1996). The most common method for identification is to sequence a phylogenetically informative gene and to use the sequences to draw up a phylogeny (Jacobs and Wingfield, 2001; Lu *et al.*, 2008; Taylor *et al.*, 2000; Zhou *et al.*, 2000). This has increased the amount of sequence data that can now be used in sequence data intensive identification techniques for fungi.

Many different molecular techniques have been used to identify fungi and, until recently, no multiplex assay has been available. Diagnostic microarrays are applicable to the identification of fungi (Nicolaisen *et al.*, 2005) and are capable of highly multiplexed identification of target organisms (Tambong *et al.*, 2006). Diagnostic arrays consist of a solid substrate embedded with a selection of probes that hybridise to distinctive targets (Bodrossy and Sessits, 2004; Kostic *et al.*, 2006; Loy and Bodrossy, 2006; Schena and Davis, 2000b). The probes are designed from the sequence data available for a group of taxa and targets are derived from the organisms that need to be identified by the array (Booth *et al.*, 2003; Letowski *et al.*, 2004). Identification of the target organisms is achieved through mismatch discrimination by the probes during hybridisation (Letowski *et al.*, 2004).

Species-specific probes are designed around unique polymorphisms that discriminate between uncomplimentary and complementary targets during hybridisation (Letowski *et al.*, 2004). The probe sequences include species unique polymorphisms that can be used to identify species when the probes are hybridised to prepared targets. The probes will only bind their complementary targets under highly stringent conditions that are optimised for the microarray platform (Conner *et al.*, 1983; Schena and Davis, 2000b). These targets are prepared from test organisms by PCR and can be either directly or indirectly labelled with fluorescent dyes using a variety of methods (Al-Khaldi *et al.*, 2004; Martens *et al.*, 2007). Hybridisation of the targets to the probes produces a fluorescent signal that may be detected and recorded using scanning equipment (Schermer, 2000). Each microarray experiment produces many data points that are analysed using powerful data analysis software.

The results of microarray experiments are recorded as .tiff images and are analysed using specialised software programs that deal with the numerous data points generated by a single microarray experiment (Maughan *et al.*, 2001; Metfies *et al.*, 2008). The software will measure the intensity of the hybridisation by comparing probe spots to the surrounding background and calculating a single to noise ratio (SNR) (Martens *et al.*, 2007). Positive SNR values indicate successful hybridisation and the SNR value will indicate how many targets have bound to the probe spot. The user can set an SNR threshold that will indicate a true positive result for that experiment (Metfies *et al.*, 2008). The true hybridisations can then be compared to a user-defined matrix that correlates hybridisation data to species probes. Such hybridisation experiments are desirable for species identification as targets are exposed to many probes simultaneously reducing the cost and the amount of time required to identify species.

Microarrays have been used to study communities of microorganisms since they are capable of simultaneously detecting multiple species (Anderson *et al.*, 2006; Brodie *et al.*, 2006; Stralis-Pavese *et al.*, 2004). The microarrays have been used to study methanotrophs in landfill sites, changing microbial communities associated with uranium reduction and reoxidation, and communities of the important plant pathogen *Phytophtora.* The microarrays were capable of detecting target organisms from environmental samples despite the presence of non-target organisms (Anderson *et al.*, 2006; Brodie *et al.*, 2006; Stralis-Pavese et *al.*, 2004). *Leptographium* species are often found growing in mixed communities and are difficult to discern from one another as they share similar morphology (Kendrick, 1962).

Species of *Leptographium* represent asexual states of the Ophiostomatoid genus *Grosmannia* and are morphologically very similar (Jacobs and Wingfield, 2001; Zipfel *et al.*, 2006). They are typically identified based on conidial morphology, patterns of the primary branches, lengths of the conidiophores, and the presence/absence of rhizoids (Kendrick, 1962). The genus *Leptographium* includes more than 56 species that are morphologically similar, typically with erect conidiophores terminating in brush like structures that bear sticky masses of mitospores (Kendrick, 1962). Species within this genus are difficult to identify, even by trained mycologists, as they are morphologically similar so a phylogenetic approach combined with morphological data has been used to discern species. This has led to the generation of a comprehensive sequence data for the majority of the species within the genus. The abundance of molecular and morphological data makes this genus an ideal model for testing molecular diagnostic techniques.

Leptographium species have been well studied in a monograph that includes morphological and molecular data for 56 species (Jacobs *et al.*, 2001). The species limitations have been rigorously

tested and are very well defined making this genus an ideal model for testing new diagnostic techniques. Molecular diagnostic techniques rely on an abundance of sequence data which is available for *Leptographium* over the translational elongation factor 1α (EF1α), β-tubulin (βT) and the rRNA internal transcribed spacer 2 (ITS2) regions has been compiled (Jacobs *et al.*, 2006; Jacobs and Wingfield, 2001). The sequence data were used to infer phylogenies to delineate different species and it was found that phylogenetic groups resemble some of the morphological groupings within the genus (Jacobs and Wingfield, 2001). The phylogenies indicate that sequence polymorphisms are good markers for species identification within *Leptographium* and that it would be possible to design species diagnostic probes from the sequence data. These probes could be used in a diagnostic microarray for *Leptographium.*

Several questions were posed relating to the design parameters of the microarray. The study aimed to identify species of *Leptographium* using an array designed from available sequence data, to assess the use of Locked Nucleic Acids (LNA) at different positions in the probe sequence and to determine whether it is sufficient to use one gene region for probe design or whether multiple regions would provide more accurate species identification.

2.2. Materials and methods

2.2.1. Design of PCR primers and gene-specific oligonucleotide probes

Unique primer and probe sets were designed for each of the 26 *Leptographium* species from the partial BT1, EF1α and ITS2 DNA sequences. Probes and primers were designed using Allele ID 2.0 (Premier Biosoft International, Palo Alto, CA). All the oligonucleotides thus designed were between 18-22 bases in length. Locked nucleic acid (LNA) nucleotides were inserted in the probe sequences at the site of unique SNPs. The specificity of each oligonucleotide was assessed by conducting BLAST searches and only unique oligonucleotides were accepted. The primer sequences BT2a (Glass and Donaldson, 1995) and ITS3 (White *et al.*, 1990) were included as positive controls.

2.2.2. Microarray fabrication

The species-specific 20-mer oligonucleotides were suspended in sterile distilled water to a final concentration of 100pmol/µl (Maynard *et al.*, 2005). The stocks were diluted with dimethyl sulphfoxide (DMSO) (50%, vol/vol) to a final concentration of 50pmol/µl (Maynard *et al.*, 2005). Ten µl of the dilution was transferred into a 384-well microplate and stored at –20˚C, until printing onto Vapour Phase Coated Glass Slides (Amersham Pharmacia Biotech) using a Molecular Dynamics Gen III spotter at the African Centre for Gene Technologies (ACGT) Microarray Facility, University of Pretoria, Pretoria, South Africa (http://fabinet.up.ac.za/microarray). Following printing, the slides were allowed to dry at 45–50% relative humidity overnight. Spotted DNA was then bound to the slides by UV cross-linking at 250 mJ and baking at 80°C for 2 h. Each spot was replicated 16 times at random positions on the array. Microsoft Excel was used to create a random pattern for microarray printing and eight spots of DMSO negative controls were included in the pattern as well as dilution series of the positive controls.

2.2.3.   DNA isolation and PCR conditions for target preparation

Leptographium elegans, L. dryocoetidis and *L. leptographioides* were grown on Potato Dextrose Agar as described in Jacobs *et al.* (2005). The identity of the isolates was confirmed using morphological characters and BLAST searches of partial β-tubulin sequences (Jacobs *et al.*, 2005, Jacobs and Wingfield, 2001). The PCR reaction consisted of 60ng DNA template, 2.5mM MgCl$_2$ concentration, 1x PCR buffer, 0.4mM of each primer, 0.2mM dNTPs and 1U of SuperTherm *Taq* polymerase (Southern Cross, Cape Town) in 20 µl reactions (Jacobs *et al.*, 2005). The cycling conditions for all primer combinations were identical with the exception of annealing temperatures that were optimised individually. The DNA was initially denatured at 94˚C for 5 minutes followed by 30 cycles of denaturation at 94˚C for one minute, primer annealing at an optimised temperature for one minute and extension for 30s at 72˚C, and a final 72˚C extension step for 7 minutes. The primers ITS3 and ITS4 (White, *et al.*, 1990) annealed at 50˚C for all three organisms. The primers Bt2a and Bt2b (Glass and Donaldson, 1995) annealed at 55˚C for *L. dryocoetidis* and *L. elegans* but at 50˚C for *L. leptographioides.* The species-specific primers for *L. dryocoetidis* BT1a+b annealed at 50˚C, EF1a+b annealed at 55˚C and ITS1a+b annealed at 53˚C. The species-specific primers for *L. elegans* EF5a+5b annealed at 52˚C and ITS5a+5b annealed at 45˚C. The species-specific primers for *L. leptographioides* BT4a+4b annealed at 61˚C and EF4a+4b annealed at 50˚C. Amplicons were prepared from a *Fusarium* sp. using ITS3+4 (White *et al.*, 1990), BT2a+2b (Glass and Donaldson, 1995) and EF1F+2R (Jacobs *et al.*, 2004). Targets were prepared from a *Fusarium* sp. isolate using ITS3+4 (White *et al.*, 1990), EF1F+2R (Jacobs, 2004) and Bt2a+2b (Glass and Donaldson, 1995), to serve as a negative control. Amplicons were analysed using agarose gel electrophoresis and the bands visualised using UV irradiation. PCR products were precipitated in 90% ethanol and 0.9mM sodium acetate, pH4.6.   The precipitate was collected by centrifugation at 13 000rpm for 30 minutes.   The pellets were washed twice in 70% ethanol by centrifugation for 2 minutes at 5000rpm. The pellets were then resuspended at ~100ng/µl in deionised water.

2.2.4.   Indirect target labelling

The targets were labelled using a Klenow labelling reaction based on the method used by Martens *et al.*, (2007). Thus, 700ng of clean PCR product was denatured with 0.1µg of random nanomers (Inqaba Biotech, Pretoria) in 16.5µl of deionised water at 94˚C for 5 minutes and placed on ice for 5

minutes. 1U of Klenow enzyme (Roche Diagnostics, Germany), 1x dNTP (45μM dATP, 9μM dTTP, 72μM aa-dUTP, 45μM dGTP, 45μM dCTP, Roche) and 1x Klenow buffer were added to the PCR product and annealed random nanomers in a final volume of 20μl and incubated at 37°C over night. The labelled DNA was separated from the primers and other reagents using the QiaQuick PCR purification kit (Southern Cross Biotechnology, South Africa) using amended wash buffers (5mM KPO$_4$, pH 8.0, 80% EtOH) and the protocol provided by the manufacturers. The product was then eluted twice into 16μl of deionised water and dried down to a pellet. The pellet was re-suspended in 0.5M Sodium bicarbonate (pH9), 0.34μg/μl Cy5 dye in a 10μl reaction for 3 hours. The dye coupling reaction was purified using the QiaQuick PCR purification columns, amended buffers and the manufacturers' protocol (Wang and Seed, 2003).

2.2.5. *Hybridization*

The slides were pre-hybridised at 60°C and hybridised at 50°C for isolates CMW2245 (*L. elegans*) and CMW2803 (*L. leptographioides*), 51°C for CMW442 (*L. dryocoetidis*) and 50°C for *Fusarium* sp. following the protocols described by Maynard *et al*. (2005). Pre-hybridisation was conducted for 1-2hrs at 60°C by adding a pre-hybridisation mixture of 1μg of Herring Sperm DNA (GibcoBRL, U.K.) and 39μl of hybridisation buffer (Amersham Biosciences, U. K.) to the hybridisation chamber. After the pre-hybridisation, the slides were washed for 6 minutes in 2x SSC (Invitrogen) and 0.2% SDS (Sigma) washing solution, rinsed in deionised water and dried at 2000g for 4 minutes. The labelled targets were purified and the dye incorporation was measured. Labelled targets showing less than 30pmols of signal in total were discarded while labelled DNA with an excess of 30pmols of signal were dried down into a pellet. The pellets were resuspended in the 40μl of hybridisation mixture. This mixture was denatured at 95°C for five minutes and placed immediately on ice for a further 5 minutes before being added to the hybridization chamber. Hybridisation took place at the optimal temperature overnight. After hybridization, the slides were washed in three washing solutions at 60°C: twice in 2x SSC (Invitrogen) and 0.2% SDS (Sigma) for 6 minutes, then in 0.2x SSC (Invitrogen) and 0.2% SDS (Sigma) for 2 minutes and finally for 2 minutes in 0.075x SSC (Invitrogen). The slides were rinsed in deionised water and dried by centrifugation at 2000g for 4 minutes.

2.2.6. Scanning, image processing and data analysis

Microarrays were scanned immediately after the hybridisation using a GenePix 4000B scanner (Molecular Dynamics, USA) at 635nm to detect Cy5 labelled targets. The .gal file was overlaid on top of the .tiff image. The blocks were examined for irregular spots and these were manually flagged. The area of the block was adjusted to fit the image correctly and then the signal to noise ratio was calculated as follows: SNR = (Signal median – Background median) x Standard deviation Background^{-1} (Martens *et al*., 2007). The median of each spot was calculated over 4 repeats within a slide and spots were removed from the analysis if they had an SNR value below the median. The sum of the medians was then used to normalise the data across the three replicate slides for each species. The median of the medians for each spot across the three replicate slides was calculated and used to assign a present/absent value to each spot. Each spot was then either assigned a 1 (present, SNR>/= 3.0) or a 0 (absent, SNR<3.0) according to the median SNR value. The best target-probe match was considered to be those with the highest SNR value. Cross hybridisations were investigated from the raw data if they showed an SNR value above 2.5.

2.3. Results

2.3.1.   Primer and probe design

Thirty-seven species-specific primer pairs and probes (Tables 1 and 2) were designed from three gene regions that identify 26 species of *Leptographium*. Three gene regions were used so that multiple diagnostic oligonucleotide probes could be designed for each species. Probe selection was based on alignments generated by ClustalW in order to identify polymorphisms. The oligonucleotide probes included SNPs and longer sequence polymorphisms. The probes and primers were designed to have a melting temperature of 55˚C; they were 20 bases in length and had G+C content of 50%. Oligonucleotide probes were designed so that the unique polymorphisms were placed toward the centre of the probe. The greatest number of unique probes was generated from the EF1α sequence alignments. Analysis of the ITS2 sequence alignments revealed few unique polymorphisms suitable for probe design. Probes were also modified to contain linked nucleic acids (LNAs) at selected SNP sites in order to test whether their performance would be enhanced by such modifications (Table 1).

The use of LNAs at SNP sites was compared within the *L. leptographioides* probes. EF4lep contained an LNA at the SNP site where as ITS4lep and BT4lep did not. The inclusion of an LNA did not show any affect on the SNR value of EF4Lep when compared to ITS4lep and BT4lep (Table 1). In the case of EF19doug where LNAs were inserted at multiple polymorphic sites along the probe length cross hybridisation was observed. Probes with polymorphisms placed toward the centre of the probe such as EF14abiet and EF15brach, were found not to cross hybridise (Table 1).

2.3.2.   Sample preparation

There was a low recovery of short amplicons (~200bp) that showed decreased representation in the final sample. The labelling of these short fragments was less efficient compared to the longer amplicons used in this study. Two test probes BT1 and EF5 failed to hybridise to targets (Table 1). The targets for these two probes were under 200bp in length and may not have been present or sufficiently labelled in the final hybridisation mix.

2.3.3.   Data analysis

The resulting hybridisation patterns showed a positive identification of *L. dryocoetidis, L. elegans* and *L. leptographioides.*   The targets prepared from each species bound to the species-specific probes.   Although, cross hybridisations were observed none of the probes cross hybridised with targets prepared from an unrelated ascomycetous species of *Fusarium.* This indicated that the probes are unique for *Leptographium* species.

The data were analysed to gain insight into the positional affect of the SNP and LNA insertion, the reasons for cross hybridisation, the affinity of test probes for their correct target and the affect of multiple versus single diagnostic probes/species.   SNPs were found to be most disruptive, and thus gave better signal, if they were at a central position. However, if these regions were flanked by strong guanine to cytosine bonds, mismatches were observed by the presence of signal when it was not expected. There was no significant difference between the performance of probes containing LNAs at SNP sites and those that did not.   Fewer cross hybridisations were observed when the probe contained multiple polymorphisms along the probe length, regardless of position or modification at the polymorphic site.

Analysis of the raw data revealed 11 probes that showed cross hybridisation (Table 3) due to their design.   Some probes longer than 20bp formed stable duplexes with non-target sequences despite the single base mismatch.   Shorter probes that showed less homology to the target but were similar at points of guanine-cytosine bonds also formed stable duplexes with non-target sequences.   The strength of the cross hybridisation could be correlated to the number of mismatches between the probe and the target.

Test probes showed high affinity for their targets producing a best match result (Table 1). BT1 and EF5 test probes failed to show any significant hybridisation, but it was possible to identify the test organisms positively due to the presence of multiple diagnostic probes. In the absence of multiple probes and single diagnostic probe failure, the diagnostic array could not be used effectively.

2.4. Discussion

In this study, we have shown that it is possible to identify *Leptographium* species using a microarray platform, if certain design parameters are followed. We have found that the optimal probe length to discern between closely related species of *Leptographium* is 20bp. These *Leptographium* diagnostic probes must be designed from multiple gene regions to compensate for probe failures and cross-hybridisations. Thus, the *Leptographium* probe sequences must contain SNPs or longer polymorphisms in a central position in order to disrupt cross hybridisations as emphasised in a study by Letowski *et al.* (2004). In this study, Locked Nucleic Acids (LNA) were inserted into the probes at these different polymorphic sites and were found to have no effect on probe specificity. This study also showed that the average target lengths were too short and did not compliment the preferred labelling method. For future work, it would be preferable to use a similar probe design but longer targets in conjunction with the same labelling method.

Multiple diagnostic probes for each species were designed from the ITS2, βT and EF1α gene regions to differentiate among the 26 species of *Leptographium* included in this study. Fungal diagnostic probes are typically designed from a single gene region, such as the ITS2 (Anderson *et al.*, 2006; Nicolaisen *et al.*, 2005), but in this study it was evident that a single gene region would be insufficient to differentiate among the 26 species of *Leptographium* included in this study. In other studies, multiple probes have been used to improve the sensitivity of the array (Booth *et al.*, 2003; Burton *et al.*, 2005; Maynard *et al.*, 2005) and were used successfully in this study to differentiate among *Leptographium* species. Probes differentiate among species based on polymorphisms in their sequence and will only bind a completely complementary target thereby identifying a specific species.

The affect of polymorphism identity and position within the probe sequence on hybridisation was investigated in this study to determine the best probe design for species identification. Polymorphism of varying length and position were included in probe sequences across species and across gene regions and it was found that probes that contained polymorphisms towards the centre of the sequence cross-hybridised less often than probes containing polymorphisms at other positions. Probes such as EF25profa and EF18pine that contained multiple polymorphisms situated towards the centre of the probe, did not cross hybridise whereas EF22pini and BT9curvi that had polymorphic sites towards the terminal ends of the probes cross-hybridised. These results are consistent with the findings of Letowski *et al.* (2004), who showed that polymorphisms are most effective at disrupting cross hybridisations if they are located toward the centre of the probe.

However, not all the cross hybridisations encountered in this study could be explained by the position of the polymorphism.

In this study, some of the cross hybridisations were not due to the position of the polymorphism within the probe but other sequence characteristics. Closer inspection of the base identity of the polymorphism and the flanking nucleotides revealed that polymorphisms flanked by multiple G-C bonds were found to be ineffective in disrupting cross hybridisations. It was also found that cross hybridisations occurred if probes showed more than a 95% identity with non-targets, this contradicts other studies that indicate that a single polymorphism within a 20-mer probe is sufficient to identify a species (Letowski *et al.*, 2004 Volokhov *et al.*, 2004). Therefore, probes to differentiate among *Leptographium* species have to show at least a 10% difference to a non-target this equates to at least two polymorphisms along a 20-mer. You *et al.* (2006) showed that the disruptive affect of polymorphisms on a mismatched duplex can be enhanced by inserting LNAs (You *et al.*, 2006) at the polymorphic sites and this was investigated for *Leptographium* diagnostic probes.

LNAs were inserted into several *Leptographium* diagnostic probes at different polymorphic sites to investigate their efficacy in preventing cross hybridisations. The *Leptographium* diagnostic probes containing LNAs at a centre polymorphism behaved comparably to those that did not include an LNA thus LNAs were found to have no affect on probe specificity. In this study single LNAs were used and this may have affected their performance as You *et al.* (2006) found that LNAs are most effective when they are used in triplicate if the centre position is occupied by the SNP (You *et al.*, 2006). However, even when two LNAs were used at two SNPs in EF19doug making the total number of LNAs in the probe 4 and the probe still cross-hybridised, indicating that LNAs have no effect on probe specificity for this particular probe set. Diagnostic microarrays rely on the ability of the probes to discriminate among target and non target amplicons using hybridisation however the target chemistry also has an effect on the hybridisation (Lane *et al.*, 2004; Liu *et al.*, 2006).

Targets were selected to be a certain length and of a particular constitution so that they were complementary to a probe and did not form secondary structures. This is important because secondary structure can prevent the targets from binding to probes and will make the targets less mobile in the hybridisation mixture (Lane *et al.*, 2004). The targets selected for this study were selected to be less than 500bp in length, which is consistent with finding by Liu *et al.* (2007), who found that shorter amplicons showed fewer secondary structures and hybridised probes more successfully. The successful hybridisations of these experiments showed that the target lengths were appropriate to the probe set designed and formed few secondary structures. Another means of decreasing secondary structures in the targets was investigated by Lane *et al.* (2004), who

recommended indirect labelling using a Klenow fragment to decrease secondary structures in the targets.

The labelling method used in this study combined random priming nanomers with Klenow incorporation of amino allyl dUTPs. This labelling method was only partially compatible with the short targets selected for this study, as there were few complementary binding sites for the nanomers. As a result, there were fewer binding sites for the Klenow fragment, which incorporated the aa-dUTPs, affecting the overall signal produced by positive hybridisation. The indirect labelling protocol optimised for this study could be more effective on longer targets and this should be considered for future studies. Targets prepared from the three test *Leptographium* species were successfully hybridised to the array but still cross-hybridised with probes for other *Leptographium* species.

The cross hybridisations observed in this study did not affect species identification detrimentally as multiple probes were included for each species and a stringent single-to-noise ratio was used for data analysis. Positive hybridisations were determined by setting a signal-to-noise ratio (SNR) threshold and comparing it to SNR values obtained in the hybridisations. An SNR value of 3 was selected as the minimum value for a positive hybridisation in this study because this value excluded most cross hybridisations. This SNR value is consistent with the minimum SNR value chosen by Martens *et al.* (2007), to differentiate positive and significant hybridisations for an *Ensifer* prototype array. Cross hybridisations were also counteracted by the presence of multiple diagnostic probes for each *Leptographium* species. Multiple probes compensated for probe failures and cross hybridisations in this study by providing up to three confirmatory results of a species compared to one cross hybridisation.

An optimal array design for *Leptographium* and the successful identification of three test species were the main outcomes of this study. The optimal design consisted of probes that are approximately 20bp in length, containing polymorphisms towards the centre of the probe, which were successful applied in minimizing cross hybridisation. Modification of these polymorphisms by using LNAs at polymorphic sites did not improve specificity during hybridisation. A multiple probe set for each species from multiple gene regions ensured that the identity of the test organisms were deduced despite test probe failures and multiple cross hybridisations. *L. leptographioides, L. elegans* and *L. dryocoetidis* were ultimately successfully identified using the array design and technology indicating that it is a suitable method for species identification in *Leptographium*.

References

Al-Khaldi, S.F., Myers, K.M., Rasooly, A., Chizhikov, V., 2004. Genotyping of *Clostridium perfringens* toxins using multiple oligonucleotide microarray hybridization. Mol Cell Probes. 18, 359-367.

Anderson, N., Szemes, M., O'Brien, P., Weerd, M.D., Schoen, C., Boender, P., Bonant, P., 2006. Use of hybridization melting kinetics for detecting *Phytophthora* species using three-dimensional microarrays: demonstration of a novel concept for the differentiation of detection targets. Mycol Res. 110, 664-671.

Bodrossy, L., Sessits, A., 2004. Oligonucleotide microarrays in microbial diagnostics. Curr Opin Microbiol. 7, 245-254.

Booth, S.A., Drebot, M.A., Martin, I.E., Ng, L.-K., 2003. Design of oligonucleotide arrays to detect point mutations: molecular typing of antibiotic resistant strains of *Neisseria gonorrhoeae* and Hantavirus infected deer mice. Mol Cell Probes. 17, 77-84.

Brodie, E.L., DeSantis, T.Z., Joyner, D.C., Baek, S.M., Larsen, J.T., Andersen, G.L., Hazen, T.C., Richardson, P.M., Herman, D.J., Tokunaga, T.K., Wan, J.M., Firestone, M.K., 2006. Application of a high-density oligonucleotide microarray approach to study bacterial population dynamics during uranium reduction and reoxidation. Appl Environ Microbiol, 72, 6288-6298.

Burton, J.E., Oshota, O.J., North, E., Hudson, M.J., Polyanskaya, N., Brehm, J., Lloyd, G., Silman, N.J., 2005. Development of a multipathogen oligonucleotide microarray for detection of *Bacillus anthracis*. Mol Cell Probes. 19, 349-357.

Conner, B.J., Reyes, A.A., Morin, C., Itakura, K., Teplitz, R.L., Wallace, R.B., 1983. Detection of sickle cell B^S-globin allele by hybridization with synthetic oligonucleotides. PNAS, 80, 278-282.

De Vos, L., Myburg, A.A., Wingfield, M.J., Desjardins, A.E., Gordon, T.R., Wingfield, B.D., 2007. Complete genetic linkage maps from an interspecific cross between *Fusarium circinatum* and *Fusarium subglutinans*. Fung Genet Biol. 44, 701-714.

Glass, N.L., Donaldson, G.C., 1995. Development of primer sets designed for use with the PCR to amplify conserved genes from filamentous ascomycetes. Appl Environ Microbiol 61, 1323-1330.

Jacobs, K., Bergdahl, D. R., Wingfield, M. J., Halik, S., Seifert, K. A., Bright, D. E., and Wingfield, B. D., 2004. *Leptographium wingfieldii* introduced into North America and found associated with exotic *Tomicus piniperda* and native bark beetles. Mycol Res. 108, 411-418.

Jacobs, K., Solheim, H., Wingfield, B.D., Wingfield, M.J., 2005. Taxonomic re-evaluation of *Leptographium lundbergii* based on DNA sequence comparisons and morphology. Mycol Res 10, 1149-1161.

Jacobs, K., Wingfield, B.D., Wingfield, M.J., 2006. *Leptographium:* Anamorphs of *Grosmannia*, The Ophiostomatoid Fungi: Expanding Frontiers, Morton Bay Research Station, North Stradbroke Island, Brisbane, Australia.

Jacobs, K., Wingfield, M.J., 2001. *Leptographium* species: tree pathogens, insect associates and agents of blue stain, APS, Minnesota, USA.

Jacobs, K., Wingfield, M.J., Wingfield, B.D., 2001. Phylogenetic relationships in *Leptographium* based on morphological and molecular characters. Can J Bot. 79, 719-732.

Kendrick, W.B., 1962. The *Leptographium* complex. *Verticicladiella* S. Hughes. Can. J. Bot. 40, 771-797.

Kim, J.A., Takahashi, Y., Tanaka, R., Fukushima, K., Nishimura, K., Miyaji, M., 2001. Identification and subtyping of *Trichophyton mentagrophytes* by random amplified polymorphic DNA. Mycoses. 44, 157-1665.

Kostic, T., Weilharter, A., Rubino, S., Delogu, G., Uzzau, S., Rudi, K., Sessitsch, A., Bodrossy, L., 2006. A microbial diagnostic microarray technique for the sensitive detection and identification of pathogenic bacteria in a background of nonpathogens. Anal Biochem. 360, 244-254.

Lane, S., Evermann, J., Logea, F., Call, D.R., 2004. Amplicon secondary structure prevents target hybridization to oligonucleotide microarrays. Biosens Bioelectron 20, 728-735.

Lee, S., Kim, J.-J., Fung, S., Breuil, C., 2003. A PCR-RFLP marker distinguishing *Ophiostoma clavigerum* from morphologically similar *Leptographium* species associated with bark beetles. Can J Bot. 81, 1104-1113.

Letowski, J., Brousseau, R., Masson, L., 2004. Designing better probes: effect of probe size, mismatch position and number on hybridization in DNA oligonucleotide microarrays. J Microbiol Methods 57, 269-278.

Liu, W.-T., Guo, H., Wu, J.-H., 2006. Effects of target length on the hybridization efficiency and specificity of rRNA-based oligonucleotide microarrays. Appl Environ Microbiol 73, 73-82.

Loy, A., Bodrossy, L., 2006. Highly parallel microbial diagnostics using oligonucleotide microarrays (Review). Clin Chim Acta. 363, 106-119.

Lu, Q., Decock, C., Zhang, X.Y., Maraite, H., 2008. *Leptographium sinoprocerum* sp. nov., an undescribed species associated with *Pinus tabuliformis-Dendroctonus valens* in northern China. Mycologia. 100, 275-290.

Majer, D., Mithen, R., Lewis, B.G., Vos, P., Oliver, R.P., 1996. The use of AFLP fingerprinting for the detection of genetic variation in fungi. Mycol Res. 100, 1107-1111.

Martens, M., Weidner, S., Linke, B., de Vos, P., Gillis, M., Willems, A., 2007. A prototype taxonomic microarray targeting the rpsA housekeeping gene permits species identification within the rhizobial genus Ensifer. Syst Appl Microbiol. 30, 390-400.

Maughan, N.J., Lewis, F.A., Smith, V., 2001. An introduction to microarrays. J Pathol. 195, 3-6.

Maynard, C., Berthiaume, F., Lemarchand, K., Harel, J., Payment, P., Bayardelle, P., Masson, L., Brousseau, R., 2005. Waterborne pathogen detection by use of oligonucleotide-based microarrays. Appl Environ Microbiol 71, 8548-8557.

Metfies, K., Bborsutzki, P., Gescher, C., Medlin, L.K, Frickenhaus, S., 2008. PHYLOCHIPANALYSER — a program for analysing hierarchical probe sets. Mol Ecol Resour. 8, 99-102.

Nicolaisen, M., Justesen, A.F., Thrane, U., Skouboe, P., Holmström, K., 2005. An oligonucleotide microarray for the identification and differentiation of trichothecene producing and non-producing *Fusarium* species occurring on cereal grain. J Microbiol Methods. 62, 57-69.

Schena, M., and Davis. R. W., 2000. Genes, genomes and chips. In: M. Schena (Ed.), DNA Microarrays, A practical approach, OUP, New York, pp. 1-15.

Schermer, M.J., 2000. Confocal scanning microscopy in microarray detection. In: M. Schena (Ed.), DNA microarrays, a practical approach, OUP, New York, pp. 16-42.

Stralis-Pavese, N., Sessitsch, A., Weilharter, A., Reichenauer, T., Riesing, J., Csontos, J., Murrell, J.C., Bodrossy, L., 2004. Optimization of diagnostic microarray for application in analysing landfill methanotroph communities under different plant covers. Environ Microbiol. 6, 347-363.

Tambong, J.T., de Cock, A.W.A.M., Tinker, N.A., Lévesque, C.A., 2006. Oligonucleotide array for identification and detection of *Pythium* Species. Appl Environ Microbiol 72, 2691-2706.

Taylor, J.W., Jacobson, D.J., Kroken, S., Kasuga, T., Geise, D.M., Hibbett, D.S., Fisher, M.C., 2000. Phylogenetic species recognition and species concepts in fungi. Fung Genet Biol. 31, 21-32.

Volokhov, D., Pomerantsev, A., Kivovich, V., Rasooly, A., Chizhikov, V., 2004. Identification of *Bacillus anthracis* by multiprobe microarray hybridization. Diagn. Microbiol. Infectious Dis. 49, 163-17

Wang, X., Seed, B., 2003. Selection of oligonucleotide probes for protein coding sequences. Bioinformatics. 19, 796-802.

White, T.J., Bruns, T., Taylor, J., 1990. Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. In: A. Innis, D. Gelfand, J. Sninsky, T. White (Eds.), PCR Protocols: a guide to methods and applications, APS, San Diego, pp. 315-322.

You, Y., Moreira, B.G., Behlke, M.A., Owczarzy, R., 2006. Design of LNA probes that improve mismatch discrimination. Nucleic Acids Res. 34, e60.

Zhou, X.D., Jacobs, K., Morelet, M., Ye, H., Lieutier, F., Wingfield, M.J., 2000. A new *Leptographium* species associated with *Tomicus piniperda* in south-western China Xu. Mycoscience. 41, 573-578.

Zipfel, R.D., de Beer, Z.W., Jacobs, K., Wingfield, B.D., Wingfield, M.J., 2006. Multigene phylogenies define *Cercatocystiopsis* and *Grosmannia* distinct from *Ophiostoma*. Stud Mycol. 55, 75-97.

Table 1. Thirty six probe sequences ~~v~~ ific gene regions for 26 *Leptographium* species. ~~UNIVERSITEIT VAN PRETORIA / UNIVERSITY OF PRETORIA / YUNIBESITHI YA PRETORIA~~ tubulin, EF- 1α and the ITS2 from 3 different species of *Leptographium* and a *Fusarium* sp. were hybridised to the prototype array and the hybridisation patterns were analysed.

Target Organism	Probe name	Probe sequence	*L. dryocoetidis*	*L. elegans*	*L. leptographioides*	*Fusarium sp.*
L. abieticolens	EF17abiet	GCATGGGTTCTGGACAA[A]CTTA				
L. albopini	EF20albo	CGCTCCGGGTCATCGTGA				
L. antibioticum	BT14anti	GGCACGTCCGATCTCCAG				
	EF14anti	GCCATTCT[T]ATCATTGCCGCT				
L. brachiatum	BT15brach	CTTGAGAGCAACGACAACCTT				
	EF15brach	GCTCGG[T]AAGGGTTCTTTCAAG				
L. brevicollis	BT12brev	GAGGTTCGTATAAATTCGCCCA				
L. calophylli	EF2calo	GATTCTGGACCGCCGCTG				
	ITS2calo	GGAACAGGACGCCAGAGA				
O. chlamydatum	BT8chlam	GGAGCGGATGAGCGTGTA				
	EF8chlam	GGTCACCGTGA[T]TTCATCAAGA				
L. costaricense	BT11cost	[T]GCCGTCCTCGTCGATCT				
	EF11cost	ACTTCTTTCTCTTCGACTTGCC				
L. crassivaginatum	EF26crass	TAACAATCAATCCAGGAAGCCG				

Target Organism	Probe name	Probe sequence	*L. dryocoetidis*	*L. elegans*	*L. leptographioides*	*Fusarium* sp.
L. dryocoetidis	BT1dryo	CGTGCCGTCCTTGTCGAT				
	EF1dryo	CCTTCACTTAGCCTATCTCTGC	+			
	ITS1dryo	TTTGGAGAGGATGCTTT[T]GGC	+++			
L. elegans	EF5ele	CGGTGCCTATTCTCGTGGT				
	ITS5ele	GACGCCAAGCCTCTGTGA		+++		
L. francke-grosmanniae	BT7FraGros	GGCACGATGGACGCTGTC				
	EF7FraGros	ACCGAATCAGGAAGCCGC				
L. fruticetum	BT10frut	AATGGCACCTCTGACCTCC				
L. grandifoliae	BT13grand	TGGATGCTGTCCGTGCTG				
L. leptographioides	BT4lep	GTACAACGGCACCTCCGA			+++	
	EF4lep	CGA[C]ATTGCTCTGTGGAAGTT			+++	
	ITS4lep	CCGTCCGAGTTCCCTGGA			+++	

Target Organism	Probe name	Probe sequence	*L. dryocoetidis*	*L. elegans*	*L. leptographioides*	*Fusarium* sp.
L. procerum	EF24proce	TCCTACTCCATGCTGCTTTATT				
L. profanum	EF25profa	[T]GATAT[T][G]CCCTCTGGAAGTTC				
L. pruni	BT6pru	GACAGGTACAACGGCACG				
L. rubrum	EF16rub	TTCCTTACTCCGTTCTGTCCTT				
L. sibiricum	EF23sibir	CACTCATAACAGGAAGCCGC			+	
L. wageneri var. pseudotsugae	ITS3wagpseu	[T]AGTAACGGCGAGTGAAGCG	+++			
L. yunnanensis	EF21yunn	CTGGTACGGGCGAGTTCG				

Bases in square brackets indicate polymorphic sites where LNAs have been inserted. Median Signal to Noise ratio (SNR) values for each probe were calculated for each experiment and then a median of each probe median was calculated over the triplicate experiments. The median of the median is represented here for the final SNR values indicated by+ signs where +>3 and +++>5.

Table 2. Probe cross hybridisations were observed in the nine hybridisation experiments were investigated if the SNR value> 2.5.

Probe name	Cross hybridisation	Reason for failure	SNR
BT14anti	CAGGTACAACGGCACCTCCGAGCTCCAGCT GGCACGTCCGATCTCCAG	16bp of the probe bound to the Beta tubulin region of *L. leptographioides*	2.5<.>3.0
	CGCCCGGGGCAGGCCCCGAAATCCAGTGGC GGCACG—TCCGATCTCCAG	14bp of the probe bound the ITS1 region of *L. leptographioides*	2.5<.>3.0
BT15brach	CCCCATTCACAAACACACG—CAACATTCCT CTTGAGAGCA—ACGACAACCTT	18bp of the probe binds to Beta tubulin region of *L. leptographioides*	2.5<.>3.0
BT7FraGros	GCCCGGCACCATGGACGCTGTCCGTGCCGG GGCACGATCGACGCTGTC	17bp of the probe bound the Beta tubulin region of *L. leptographioides*	2.5<.>3.0
BT9curvi	GAGGGCCCCCAAAGCGAGTGGCTGGGCCCG CCTTGACAGC—AGTGGCGGT	13bp of the probe bound to the Beta tubulin region of *L. dryocoetidis*	2.5<.>3.0
EF17abiet	AAGTACGCATGGGTTCTTGACAAGCTCAAG GCATGGGTTCTGGACAAACTTA	19bp of the probe binds the Elongation factor 1 α region of *L. leptographioides.* Contains one LNA at an SNP site.	2.5<.>3.0

Probe name	Cross hybridisation	Reason for failure	SNR
EF19doug	GTA<mark>C</mark>GCATGGGTTCTTGACAAGCTCAAGGC <mark>T</mark>GCATGGGTTCTTGACAAGC	The probe binds the Elongation factor 1α region of *L. leptographioides.* Contains two LNA's at two SNP sites.	>3.0
EF25profa	CCATTGATAT<mark>C</mark>GCCCTCTGGAAGTTCGAGA TGATAT<mark>T</mark>GCCCTCTGGAAGTTC	21bp of the probe bind the Elongation factor 1α region of *L. elegans.* This probe contains 3 non-sequential linked nucleic acids (LNA)	2.5<.>3.0
EF8chlam	CCCGGGTCACCGTGA<mark>C</mark>TTCATCAAGAACAT GGTCACCGTGA<mark>T</mark>TTCATCAAGA	21bp of the probe bind the Elongation factor 1α region of *L. leptographioides*. This probe contains an LNA at the site of the unique SNP	2.5<.>3.0

Probe name	Cross hybridisation	Reason for failure	SNR
ITS3wagpseu	CCC**C**AGTAACGGCGAGTGAAGCGGCAACA **T**AGTAACGGCGAGTGAAGCG	The probe binds the Internal transcribed spacer region of *L. dryocoetidis.* This probe contains an LNA at the 5' end.	.>3.0
	GCCT**C**AGTAACGGCGAGTGAAGCGGCACCA **T**AGTAACGGCGAGTGAAGCG	The probe binds the Internal transcribed spacer region of *L. elegans*	2.5<.>3.0
	CCCCAGTAACGGCGAGTGAAGCGGCAACA **T**AGTAACGGCGAGTGAAGCG	The probe binds the Internal transcribed spacer region of *L. leptographioides*	2.5<.>3.0
ITS4lep	GGCGCCGTCCGAGTTCC**T**TGGAACAGGACG CCGTCCGAGTTCC**C**TGGA	17bp of the probe binds the Internal transcribed spacer region of *L. dryocoetidis*	2.5<.>3.0
EF22pini	CACGCG**TTCC**CCTCCTC**CC**CC**A**TTATTAGA G**CTCAA**CCTCCTC**TT**CC—TTATT	16bp of the 21bp probe bound to the EF1 product of *L. Leptographioides*	2.5<.>3.0

The highlights indicate mismatches between probes and non-targets.
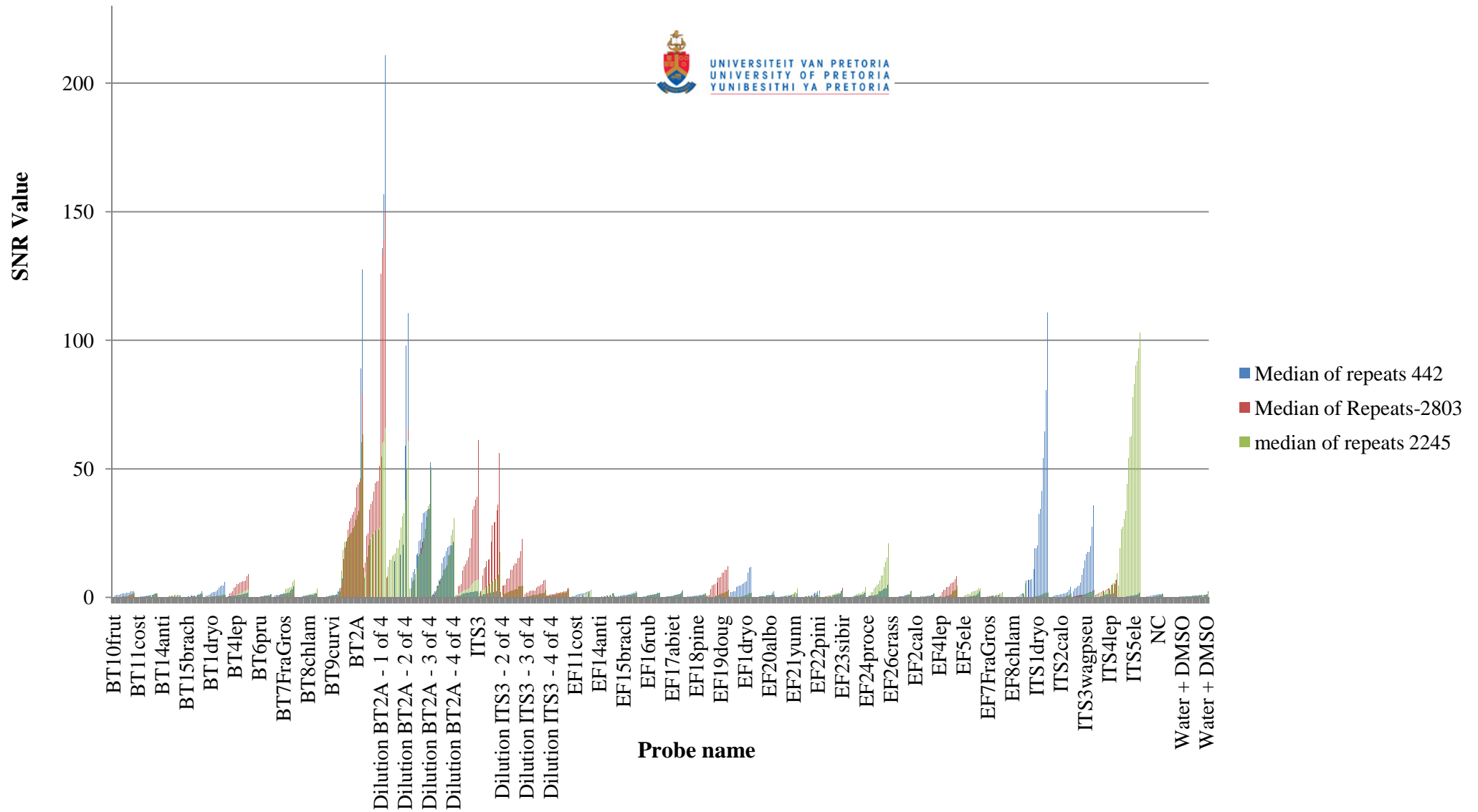
Figure 1. A bar graph showing the median signal intensities from the hybridisation of targets prepared from *L. dryocoetidis* (442), *L. leptographioides* (2803) and *L. elegans* (2245).

3. **Chapter 3:** The Design of a PhyloChip for *Leptographium* species

Natalie R. Van Zuydam[a], Brenda D. Wingfield[a, *], Karin Jacobs[c], and Michael J. Wingfield[b]

[a]Department of Genetics, [b]Department of Microbiology and Plant Pathology, Forestry and Agricultural Biotechnology Institute (FABI), University of Pretoria, Pretoria, 0002, South Africa

[c]Department of Microbiology, University of Stellenbosch, Private Bag X-1, Matieland, Stellenbosch, South Africa

Abstract:

The genus *Leptographium* is a relatively small and well-defined group of fungi that includes species that stain wood and a small number are plant pathogens. Species in the genus have very similar gross morphology and distinguishing between them on this basis is difficult. Thus, other methods of identification have been sought for easier, unambiguous species identification. The aim of this study was to design a PhyloChip for the genus *Leptographium* using available sequence data. DNA sequences from 56 species of *Leptographium* were used in the design of a minimal, 20-mer probe set. The probes included in the chip were designed from the β-tubulin, internal transcribed spacer and translation elongation factor 1-α regions. A total of 89 probes were included in the chip. The probes delineated 54 species, placing *L. guttulatum* and *L. lundbergii* into a 55th taxon. This chip will be useful in identifying known species of *Leptographium* as well as identifying new species of *Leptographium* when they are collected.

3.1. Introduction

The genus *Leptographium* is a small genus when compared to other fungal genera. It consists of species that are anamorphs to the teleomorphic genus *Grosmannia* (Zipfel *et al*., 2006). The species in the genus share very similar gross morphology and are closely related phylogenetically. Despite these characteristics the species are well defined and supported by morphological and molecular data (Jacobs *et al*., 2001). Extensive work has been performed on this genus in the form of a monograph that has led to the generation of a comprehensive sequence data base for 56 species of *Leptographium* (Jacobs *et al.*, 2006; Jacobs and Wingfield, 2001). These characteristics make the genus an ideal model for testing new molecular techniques such as species diagnostic microarrays.

Microarrays represent the application of a hybridisation technology that is based on a solid substrate (Theriault *et al*., 2000). They consist of probes on the array of probes that are printed onto a solid substrate at defined locations. Targets are hybridised to probes on the array. Targets are prepared from samples and labelled using fluorescent dyes (Shapero *et al*., 2004). These targets are then hybridised to an array and bind to complementary probes that are visible as fluorescent spots when the array is exposed to laser light post hybridisation (Schena and Davis, 2000b; Schermer, 2000). Microarrays are most frequently used to measure gene expression where targets are prepared from mRNA and hybridised to 70-mer probes (Schena *et al.*, 1995; Wang *et al.*, 2002). Subsequent to the use of arrays in expression studies, they have been applied to the field of diagnostics (Burton *et al.*, 2005; Huang *et al.*, 2006; Volokhov *et al.*, 2002). Such diagnostic arrays use short 20-mer probes to differentiate among species.

Species diagnostic arrays are frequently made up of 20-mer probes designed from available sequence data for a group of taxa (Roth *et al*., 2004). The 20-mer probes are carefully designed around polymorphic regions that are unique to each taxon (Beaucage, 2001; Stenger *et al.*, 2002; Volokhov *et al.*, 2002). These polymorphisms are centrally located within the probe sequence and consist of words or single nucleotide polymorphisms (SNP) (Kostic *et al*., 2006). The length of the probe and the central position of the taxon unique polymorphisms disrupt hybridisation with uncomplimentary targets. Conversely, hybridisations to completely complementary targets are highly stable and robust enough to remain intact through the stringent washing procedures (Cleven *et al.*, 2006; Kostic *et al.*, 2006; Martens *et al.*, 2007; Sergeev *et al.*, 2006; Tambong *et al.*, 2006). Thus, a species can be identified according to unique polymorphisms within a certain gene region on a species diagnostic microarray. Twenty-mer

probes have also been used in DNA barcoding applications that are based on using unique sequence polymorphisms to identify species (Min and Hickey, 2007; Summerbell *et al.*, 2005).

DNA barcoding may be followed by DNA microcoding makes use of short oligonucleotides in order to fully identify specific species (Summerbell *et al.*, 2005). DNA barcoding has its origins in phylogenetic species identification (Summerbell *et al.*, 2005) that has been used to differentiate morphologically similar species from one another (Jacobs *et al.*, 2000; Taylor *et al.*, 2000; Zhou *et al.*, 2000). Housekeeping genes are commonly used for phylogenetic analysis for many fungal species (Jacobs *et al.*, 2001; Jacobs *et al.*, 2003) and short oligonucleotides containing species specific polymorphisms can be designed from these gene sequences. DNA microarrays have been proposed as a possible platform for microcoding where targets can be exposed to many 20-mer probes simultaneously (Summerbell *et al.*, 2005).

There are many design models available for microarrays that are chosen based on the available sequence data. If the sequence data available for target taxa are diverse then a species-specific probe design can be used (Troesch *et al*., 1999) . The design becomes more complicated for more closely related species where the sequences are more homologous. An array designed to identify closely related species can be based on a minimal probe set. Species share a common group of probes that in different combinations would enable the identification of each species within the group. New species can also be identified by an established array if either new probes are added to the existing array or a hierarchical array design is adopted. PhyloChips are hierarchical in design and are useful in identifying known and previously undescribed species (Loy *et al*., 2002).

The term PhyloChip is used to describe a species diagnostic array that has an intrinsic probe hierarchy (Metfies and Medlin, 2007). The hierarchy is based on the phylogeny of a group of taxa where certain probes will identify nodes of a phylogram. The progression of probes eventually leads to the identification of a known species or a new species (Anderson *et al.*, 2006; Loy *et al.*, 2002).

The genus *Leptographium* is a relatively small and well-defined genus that was chosen to test the applications of microarrays to the identification of ascomycetous fungi. The 56 species included in this study are well defined by sequence data. Sequence data is available for three gene regions β-tubulin (βT), translational elongation factor 1α (EF1α) and the rRNA internal transcribed spacer region 2 (ITS2) that have been used to draw up an extensive phylogeny of the genus. This sequence data could be used in a PhyloChip design so that both known and

unknown species may be identified by the array. The aim of this study was to use the available sequence data to design a hierarchical probe set that could be used to identify the 56 species of *Leptographium* included in this study.

3.2. Materials and methods

3.2.1. Probe design

Twenty-mer probes were designed from partial sequence data of β-tubulin (βT), translation elongation factor 1α (EF1α) and the internal transcribed spacer region 2 (ITS2) (Jacobs *et al.*, 2006). The probes were designed using the default parameters of the minimal probe design, AlelleID 4 (Premier Biosoft). The probes were designed around a melting temperature of 55˚C and were synthesised. Probes were discarded from the set if they showed more than 90% homology to non-targets. Homology was assessed using AnnHyb 4.938 (Oliver Friard©, 1997-2007).

3.2.2. Phenograms

Phenograms were constructed using NTSYSpc21v2.11 (Applied Biostatics) cluster analysis and the neighbour joining algorithms. The probe matrix was then used to match probes to specific nodes and branches of the phenograms manually (Tables 2, 3 & 4). The phenograms were compared to the established phylogenetic trees for *Leptographium* (Jacobs *et al.*, 2006).

3.3. Results

3.3.1.   *Probe design*

A total of 101 probes were designed but only 89 were included in the final minimal probe set (Table 1). Twelve probes were rejected based on percentage similarity to non-target sequences. If a probe sequence was more than 90% similar to a non-target then it was rejected.

3.3.2.   *Phenograms*

The phenograms were constructed separately for each gene region. The phenogram for the ITS2 region separated the species into 4 distinct clades and showed the simplest topology (Figure 1). The species were progressively delineated by the phenograms drawn from β-tubulin (Figure 2) and EF1α (Figure 3) probe sets. The phenograms for β-tubulin and EF1α showed fewer basal probes and more branch probes. Most isolates were conclusively identified by ITS2 and β-tubulin probes while 11 isolates needed to be finally identified by EF1α probes. *Leptographium guttulatum* and *L. lundbergii* could not be distinguished from each other using this probe set.

Jacobs *et al.* (2006) defined 7 distinct clades for the phylogeny of *Leptographium*. The probe phenograms were compared to the phylogeny. Clade 2 of the probe tree corresponded to a mixture of clades 1 and 7 of the phylogeny that included species that were from non-coniferous hosts, or had curved spores. This reflected the phylogeny to an extent where the species with curved spores grouped together in a single clade and those from non-coniferous hosts grouped together in a defined clade. Clade 1 of the probe phenogram corresponded to a clade 2 of the phylogeny that contained species with a central large primary branch. Clade 3 from the phenogram represented a mixture of clades 3, 5 and 6 of the phylogeny. The topology of the ITS2 phenogram resembled the topology of the ITS2 phylogeny and also showed *L. elegans* as an out group for the tree (Jacobs *et al.*, 2006).

3.4. Discussion

A set of 89, 20-mer probes were designed to identify 56 species of *Leptographium*. The probe set included probes designed from the βT, EF1α and ITS2 regions. The probes set completely delineated 54 species of *Leptographium,* placing *L. guttulatum* and *L. lundbergii* into a 55th taxon. The probes were used to construct phenograms that could be used in the data analysis of the array. Comparisons of the phenograms to the phylogenies of *Leptographium* constructed by Jacobs *et al.* (2006) revealed that the phenograms closely resembled those phylogenies. This indicated that the chip could be useful for species identification (Jacobs *et al.*, 2001 ;Jacobs *et al.*, 2006) . The array design was very similar to the design for PhyloChips that have been used to identify other microbial species (Loy *et al.*, 2002; Metfies *et al.*, 2008).

PhyloChips make use of a probe hierarchy usually based on a phylogeny in order to accurately identify species (Metfies *et al.*, 2008). In this study, probes from three different gene regions were used to construct phenograms and positive *in silico* hybridisations were correlated to them. The ITS2 region produced probes that identify many basal nodes, dividing *Leptographium* into 4 main clades. The β-tubulin region produced fewer basal probes, higher order probes and many species-specific probes. The translation EF1α region produced only higher order probes with species-specific probes that provided intricate species delineation for *Leptographium*. This reflects the phylogeny to some extent where the ITS2 region gives only a basic phylogenetic delineation that is refined through the βT and EF1α regions.

Data analysis for PhyloChips has a structured methodology that uses a .gal file and correlates it with a user defined tree (Metfies *et al.*, 2008). The hybridisation patterns are analysed against the tree framework in order to identify the species. The phenograms constructed from *Leptographium* probes could be used as the reference trees where a positive identification would only be considered if all the hierarchical probes showed a positive hybridisation (Metfies *et al.*, 2008).

A PhyloChip for *Leptographium* will be useful in identifying composite populations of multiples species from environmental samples. The genus includes only a few primary plant pathogens that are often found growing amongst non-pathogenic species with similar morphology (Jacobs and Wingfield, 2001). *Leptographium wageneri* causes Black Stain Root Decline (BSRD) and is often found growing amongst non-pathogenic species. The chip designed in this study is capable of identifying pathogenic species as well as differentiating amongst three varieties of *L.*

wageneri. This will be useful as *L. wageneri* is a quarantine organism (EEPO) and rapid identification would help to protect the bio security of many countries.

The probes designed for this large array will be used as selective primers in a PCR diagnostic for 56 species of *Leptographium.* The PCR will be approached in much the same way as the Diagnostic array. Multiplex PCRs will be ordered according to the phenogram hierarchy and results analysed with reference to the hierarchy. The probes are capable of differentiating between 54 taxa placing *L. guttulatum* and *L. lundbergii* into a 55th taxon. Perhaps, if more sequence data were available for the genus, it would be possible to delineate completely all the taxa.

References

Anderson, N., Szemes, M., O'Brien, P., Weerd, M.D., Schoen, C., Boender, P., Bonant, P., 2006. Use of hybridization melting kinetics for detecting *Phytophthora* species using three-dimensional microarrays: demonstration of a novel concept for the differentiation of detection targets. Mycol Res. 110, 664-671.

Beaucage, S.L., 2001. Strategies in the preparation of DNA oligonucleotide arrays for diagnostic applications. Curr Med Chem, 8, 1213-1244.

Burton, J.E., Oshota, O.J., North, E., Hudson, M.J., Polyanskaya, N., Brehm, J., Lloyd, G., Silman, N.J., 2005. Development of a multipathogen oligonucleotide microarray for detection of *Bacillus anthracis*. Mol Cell Probes. 19, 349-357.

Cleven, B.E.E., Palka-Santini, M., Gielen, J.r., Meembor, S., Krönke, M., Krut, O., 2006. Identification and characterization of bacterial pathogens causing bloodstream infections by DNA microarray. J Clin Microbiol. 44, 2389-2397.

Huang, A., Li, J.-W., Shen, Z.-Q., Wang, X.-W., Jin, M., 2006. High-throughput identification of clinical pathogenic fungi by hybridization to an oligonucleotide microarray. J Clin Microbiol. 44, 3299-3305.

Jacobs, A., Coetzee, M., Wingfield, B.D., Jacobs, K., Wingfield, M.J., 2003. Phylogenetic relationships among *Philocephalai* species and other ascomycetes. Mycologia. 95, 637-645.

Jacobs, K., Wingfield, B.D., Wingfield, M.J., 2006. *Leptographium:* Anamorphs of *Grosmannia*, The *Ophiostomatoid* Fungi: Expanding Frontiers, Morton Bay Research Station, North Stradbroke Island, Brisbane, Australia.

Jacobs, K., Wingfield, M.J., 2001. *Leptographium* species: tree pathogens, insect associates and agents of blue stain. APS Press, Minnesota.

Jacobs, K., Wingfield, M.J., Pashenova, N.V., P.Vetrova, V., 2000. A new *Leptographium* species from Russia. Mycol Res. 104, 1524-1529.

Jacobs, K., Wingfield, M.J., Wingfield, B.D., 2001. Phylogenetic relationships in *Leptographium* based on morphological and molecular characters. Can J Bot. 79, 719-732.

Kostic, T., Weilharter, A., Rubino, S., Delogu, G., Uzzau, S., Rudi, K., Sessitsch, A., Bodrossy, L., 2006. A microbial diagnostic microarray technique for the sensitive detection and identification of pathogenic bacteria in a background of nonpathogens. Anal Biochem. 360, 244-254.

Loy, A., Lehner, A., Lee, N., Adamczy, J., Meier, H., Ernst, J., Schleifer, K.-H., Wagner, M., 2002. Oligonucleotide microarray for 16S rRNA gene-based detection of all recognized lineages of sulfate-reducing prokaryotes in the environment. Appl Environ Microb. 68, 5064-5081.

Martens, M., Weidner, S., Linke, B., de Vos, P., Gillis, M., Willems, A., 2007. A prototype taxonomic microarray targeting the rpsA housekeeping gene permits species identification within the rhizobial genus *Ensifer*. Syst Appl Microbiol. 30, 390-400.

Metfies, K., Bborsutzki, P., Gescher, C., Medlin, L.K., Frickenhaus, S., 2008. PHYLOCHIPANALYSER — a program for analysing hierarchical probe sets. Mol Ecol Resour. 8, 99-102.

Metfies, K., Medlin, L.K., 2007. Feasibility of transferring fluorescent in situ hybridization probes to an 18S rRNA gene phylochip and mapping of signal intensities. Appl Environ Microb. 74, 2814–2821.

Min, X.J., Hickey, D.A., 2007. Assessing the effect of varying sequence length on DNA barcoding of fungi. Mol Ecol Notes. 7, 365-373.

Roth, S.B., Jalava, J., Ruuskanen, O., Ruohola, A., Nikkari, S., 2004. Use of an oligonucleotide array for laboratory diagnosis of bacteria responsible for acute upper respiratory infections. J Clin Microbiol. 42, 4268-4274.

Schena, M., and Davis. R. W., 2000. Technological Standard for Microarray Research. In: M. Schena (Ed.), Microarray Biochip Technology, EP, California, pp. 1-19.

Schena, M., Shalon, D., Davis, R.W., Brown, P.O., 1995. Quantitative Monitoring of Gene Expression Patterns with a Complementary DNA Microarray. Science. 270, 467-470.

Schermer, M.J., 2000. Confocal scanning microscopy in microarray detection. In: M. Schena (Ed.), DNA microarrays, a practical approach, OUP, New York, pp. 16-42.

Sergeev, N., Distler, M., Vargas, M., Chizhikov, V., Herold, K.E., Rasooly, A., 2006. Microarray analysis of *Bacillus cereus* group virulence factors. J Microbiol Meth. 65, 488– 502.

Shapero, M.H., Zhang, J., Loraine, A., Liu, W., Di, X., Liu, G., Jones, K.W., 2004. MARA: a novel approach for highly multiplexed locus-specific SNP genotyping using high-density DNA oligonucleotide arrays. Nucleic Acids Res. 32, e181.

Stenger, D.A., Andreadis, J.D., Vora, G.J., Pancrazio, J.J., 2002. Potential applications of DNA microarrays in biodefense-related diagnostics. Curr Opin Biotech. 13, 208-212.

Summerbell, R.C., Lévesque, C.A., Seifert, K.A., Bovers, M., Fell, J.W., Diaz, M.R., Boekhout, T., Hoog, G.S.d., Stalpers, J., Crous, P.W., 2005. Microcoding: the second step in DNA barcoding. Phil Trans R Soc B. 360, 1897-1903.

Tambong, J.T., de Cock, A.W.A.M., Tinker, N.A., Lévesque, C.A., 2006. Oligonucleotide array for identification and detection of *Pythium* species. Appl Environ Microb. 72, 2691-2706.

Taylor, J.W., Jacobson, D.J., Kroken, S., Kasuga, T., Geise, D.M., Hibbett, D.S., Fisher, M.C., 2000. Phylogenetic species recognition and species concepts in fungi. Fungal Genet Biol. 31, 21-32.

Theriault, T.P., Winder, S.C., Gamble, R.C., 2000. Application of ink-jet printing technology to the manufacture of molecular arrays. In: M. Schena (Ed.), DNA microarrays, a practical approach, OUP, Oxford, pp. 101-119.

Troesch, A., Nguyen, H., Miyada, C.G., Desvarenne, S., Gineras, T.R., Kaplan, P.M., Cros, P., Mabilet, C., 1999. *Mycobacterium* species identification and rifampin resistance testing with high-density DNA probe arrays. J Clin Microbiol. 37, 49-55.

Volokhov, D., Rasooly, A., Chumakov, K., Chizhikov, V., 2002. Identification of *Listeria* species by microarray-based assay. J Clin Microbiol. 40, 4720-4728.

Wang, H.-Y., Malek, R.L., Kwitek, A.E., Greene, A.S., Luu, T.V., Behbahani, B., Frank, B., Quackenbush, J., Lee, N.H., 2002. Assessing unmodified 70-mer oligonucleotide probe performance on glass-slide microarrays. Genome Biol. 4, R5.

Zhou, X.D., Jacobs, K., Morelet, M., Ye, H., Lieutier, F., Wingfield, M.J., 2000. A new Leptographium species associated with *Tomicus piniperda* in south-western China Xu. Mycoscience. 41, 573-578.

Zipfel, R.D., de Beer, Z.W., Jacobs, K., Wingfield, B.D., Wingfield, M.J., 2006. Multigene phylogenies define *Ceratocystiopsis* and *Grosmannia* distinct from *Ophiostoma*. Stud. Mycol. 55, 75-97.

Table 1. A total of 89 probes were included in the minimal probes set for a large array to identify 56 *Leptographium* species. The probes were designed from three different gene regions.

Internal Transcribed Spacer Region		Translational Elongation Factor 1a		Beta tubulin	
ITSP1	AATGCTGCTCAAAATGGGAGG	*EF1AP2*	AAACAGGGAATGAAGAATTGCC	*BTP1*	ACAGCAATGGAGTGTAGGT
ITSP2	GGAGCTTCGCAAAGGCCA	*EF1AP3*	AAAGGCAGGGAATGAAGAATTG	*BTP3*	ACAGCATCCATCGTGCCG
ITSP3	ATTGGTTGCTGCAAGCGT	*EF1AP4*	TCGATATTGCTCTGTGGAAGTT	*BTP4*	CCGTCCTTGTGGATCTCG
ITSP4	CAAAGCGAGGGCTAATGCT	*EF1AP5*	TTAAAACCTGACCGCCCAAAA	*BTP5*	CGTCTTCGCCAGGTACAACG
ITSP6	CACAAGGTTGACCTCGGAT	*EF1AP6*	TCGATATTGCTCTGTGGAAGTT	*BTP6*	ACAGCATCCATTGTGCCG
ITSP7	CAGACCGCAGACGCAAGT	*EF1AP7*	AAAGACAGGGAGGAATTGCC	*BTP8*	CAACAAGTACGTGCCTCGC
ITSP8	CCAGCCTTTGTGAAGCTCC	*EF1AP8*	AAAGAGCCCTTGCCGAGC	*BTP10*	AGATTTCTAGCGAGCATGGC
ITSP9	CCCTAAAGACGGCAGACG	*EF1AP9*	TCGCCGCTAACACCCACA	*BTP11*	GACCGTGCTCGCTGGAGATC
ITSP10	CTCCGAGCGTAGTAAGCA	*EF1AP10*	AACGCAAGCAGGTGGAGA	*BTP12*	CAGACGTGCCGTTGTACC
ITSP11	CGAGTCTGTCTCCTTCTCAA	*EF1AP11*	AAGACTTCTCCAACAGGTGG	*BTP13*	CACGGCATCCATCGTACC
ITSP12	CGGTTGGACGCCTAGCCTTT	*EF1AP12*	TCGCCGCTAATACCCAATAC	*BPT14*	AATGGCGTGTAGGTTTCCG
ITSP15	GAGCTTCACAAAGGCTAGGC	*EF1AP13*	AAGGTCCCACAAGGCAGA	*BTP15*	AATGGCGTGTAGGTTTCCG
ITSP17	TGTAATTTGGAGAGGATGCTTT	*EF1AP15*	TCTCCCTTCCTCCCGCCA	*BTP16*	AATGGCGTGTAGGTTTCCG
ITSP18	AAAGGAGGGACAGACTTGC	*EF1AP17*	AATGGGGCGGTCAAATCTAAAG	*BTP17*	ATATGGCGGATTAGATACCACC
ITSP20	AAATCTGAGCTGGTGCCG	*EF1AP18*	TGACACCCACCACAGGAG	*BTP18*	CTAACAGATGTCACAGGCAG
ITSP21-control	AATGCTGCTCAAAATGGGAGG	*EF1AP19*	AAATACAGGGTCCACAGGGC	*BTP20*	CCAGGCAGCAGATTTCCG
ITSP22	AAATGACCGGCAGACGCAA	*EF1AP20*	TGGGCAAGGGCTCTTTCAA	*BTP21*	CATGGATGCCGTCCGTGC

Table 2. Twenty-seven probes were designed from the internal transcribed spacer regions that bind to different targets from each of the species included in the study.

Taxa	Internal Transcribed Spacer region 2																				
	1	2	3	4	6	7	8	9	10	11	12	15	17	18	20	21	22	23	24	25	27
L. abietinum	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0
L. aenigmaticum	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	0	0
L. albopini	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	0	0
L. alethinum	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	1	0	0	0	0	0
L. americanum	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0
L. antibioticum	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0
L. aureum	1	0	0	0	0	0	0	0	0	0	1	1	0	0	0	1	0	0	0	0	0
L. bistatum	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0
L. brachiatum	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0
L. brevicollis	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0
L. bhutannense	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	1	0	0	1	0
L. calophylli	1	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	0	0	0	0
L. chlamydatum	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0
L. clavigerum	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	0	0
L. costaricense	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0
L. crassivaginatum	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0
L. curvisporum	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0
L. douglassi	1	0	0	0	0	0	1	1	0	0	0	0	0	0	0	1	0	0	0	0	0
L. dryocoetidis	1	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	0	0	0	0
L. elegans	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0
L. eucalyptophilum	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0
L. euphyes	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	1	0	0	0	0	0
L. francke-grosmanniae	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	0
L. fruticetum	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	1	0	0	0	0	0
L. grandifoliae	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0
L. guttulatum	1	0	0	0	0	0	0	0	0	0	1	1	0	0	0	1	0	0	0	0	0
L. huntii	1	0	0	0	0	0	0	0	0	1	1	1	0	0	0	1	0	0	0	0	0
L. koreanum	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	0	0

Table 2. Continued

Taxa	Internal transcribed spacer region 2																				
	1	2	3	4	6	7	8	9	10	11	12	15	17	18	20	21	22	23	24	25	27
L. laricis	1	0	0	0	0	0	0	0	0	0	1	1	0	0	0	1	0	0	0	0	0
L. leptographioides	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0
L. longiclavatum	1	0	0	0	0	0	0	0	0	0	1	1	0	0	0	1	0	0	0	0	0
L. neomexicanum	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	1	0	0	0	0	0
L. penicillatum	1	0	0	0	0	0	0	0	0	0	1	1	0	0	0	1	1	0	0	0	0
L. peucophillum	1	0	0	0	0	1	1	0	0	0	0	0	0	0	0	1	0	0	0	0	0
L. piceaperdum	1	0	0	0	0	0	0	0	0	0	1	1	0	0	0	1	0	0	0	0	1
L. pineti	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	1	0	0	0	1	0
L. pini-densiflorae	1	0	0	0	0	0	0	0	0	0	1	1	0	1	0	1	1	0	0	0	0
L. pityophilum	1	0	0	0	0	0	0	0	0	1	1	1	0	0	0	1	0	0	0	0	0
L. procerum	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1	1	0	0	0	0
L. profanum	1	0	0	0	0	0	0	0	0	0	1	1	0	0	0	1	1	0	0	0	0
L. pruni	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0
L. pyrinum	1	0	0	0	0	0	0	0	0	0	1	1	0	0	0	1	0	0	0	0	0
L. reconditum	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	1	0	0	0	0	0
L. robustum	1	0	0	0	0	0	0	0	0	0	1	1	0	0	0	1	0	0	0	0	0
L. rubrum	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0
L. serpens	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1	0	0	0	0	0
L. sibiricum	1	0	0	0	0	0	0	0	0	0	1	1	0	0	0	1	0	1	0	0	0
L. terebrantis	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1	0	0	0	0	0
L. trinacriforme	1	0	0	0	0	0	0	0	0	0	1	1	0	0	0	1	0	0	0	0	0
L. truncatum	1	0	0	0	0	0	0	0	0	0	1	1	0	0	0	1	0	0	0	0	0
L. wageneri v. ponderosa	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
L. wageneri v. pseudotsugae	1	0	0	0	0	0	0	1	1	0	0	1	0	0	0	1	0	0	0	0	0
L. wageneri v. wageneri	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1	0	0	0	0	0
L. wingfieldii	1	0	0	0	0	0	0	0	0	0	1	1	0	0	0	1	0	0	0	0	0
L. yunnanensis	1	0	0	0	0	0	0	0	0	0	1	1	0	0	0	1	0	0	0	0	0

Table 3. A total of 46 probes were designed from the EF1α region. These probes bind to different targets prepared from species of *Leptographium*.

Taxa	Translation Elongation factor 1α																			
	2	3	4	5	6	7	8	9	10	11	12	13	15	17	18	19	20	21	22	24
L. abieticolens	0	0	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
L. abietinum	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
L. aenigmaticum	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
L. albopini	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
L. alethinum	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
L. americanum	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
L. antibioticum	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
L. aureum	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
L. bistatum	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0
L. brachiatum	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
L. brevicollis	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0
L. bhutannense	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
L. calophylli	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0
L. chlamydatum	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0
L. clavigerum	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
L. costaricense	0	0	1	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
L. crassivaginatum	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
L. curvisporum	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
L. douglassi	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
L. dryocoetidis	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
L. elegans	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0
L. eucalyptophilum	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0
L. euphyes	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0
L. francke-grosmanniae	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
L. fruticetum	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
L. grandifoliae	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
L. guttulatum	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
L. huntii	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0
L. koreanum	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
L. laricis	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
L. leptographioides	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
L. longiclavatum	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Table 3. Continued

Taxa	Translation Elongation factor 1α																			
	2	3	4	5	6	7	8	9	10	11	12	13	15	17	18	19	20	21	22	24
L. neomexicanum	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
L. penicillatum	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0
L. peucophillum	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
L. piceaperdum	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
L. pineti	0	0	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
L. pini-densiflorae	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
L. pityophilum	0	0	0	0	0	0	0	1	1	1	0	0	0	0	0	0	0	0	0	0
L. procerum	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0
L. profanum	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
L. pruni	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
L. pyrinum	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
L. reconditum	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
L. robustum	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
L. rubrum	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
L. serpens	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0
L. sibiricum	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
L. terebrantis	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
L. trinacriforme	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
L. truncatum	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
L. wageneri v. ponderosa	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
L. wageneri v. pseudotsugae	0	0	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0
L. wageneri v. wageneri	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
L. wingfieldii	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
L. yunnanensis	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Table 3. Continued

Taxa	Elongation factor 1α																
	25	26	27	28	32	33	34	36	37	38	39	40	41	42	44	45	46
L. abieticolens	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
L. abietinum	0	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0
L. aenigmaticum	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
L. albopini	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
L. alethinum	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
L. americanum	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
L. antibioticum	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
L. aureum	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0
L. bistatum	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
L. brachiatum	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
L. brevicollis	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
L. bhutannense	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
L. calophylli	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
L. chlamydatum	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
L. clavigerum	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0
L. costaricense	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
L. crassivaginatum	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
L. curvisporum	0	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0
L. douglassi	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
L. dryocoetidis	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
L. elegans	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
L. eucalyptophilum	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
L. euphyes	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
L. francke-grosmanniae	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
L. fruticetum	0	0	0	0	0	0	0	0	1	0	0	0	1	0	0	0	0
L. grandifoliae	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
L. guttulatum	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
L. huntii	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
L. koreanum	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0
L. laricis	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
L. leptographioides	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
L. longiclavatum	0	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0
L. lundbergii	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
L. neomexicanum	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
L. penicillatum	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
L. peucophillum	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
L. piceaperdum	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0

Table 3. Continued

Taxa	Elongation factor 1α																
	25	26	27	28	32	33	34	36	37	38	39	40	41	42	44	45	46
L. pineti	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
L. pini-densiflorae	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
L. procerum	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
L. profanum	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
L. pruni	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
L. pyrinum	0	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0
L. reconditum	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
L. robustum	0	0	0	0	0	0	0	1	0	0	1	0	0	0	0	0	0
L. rubrum	0	0	1	1	0	0	0	0	0	1	0	0	0	0	0	0	0
L. serpens	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0
L. sibiricum	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
L. terebrantis	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0
L. trinacriforme	0	0	0	0	0	0	0	1	0	0	1	0	0	0	0	0	0
L. truncatum	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0
L. wageneri v. ponderosa	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0
L. wageneri v. pseudotsugae	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
L. wageneri v. wageneri	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0
L. wingfieldii	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0
L. yunnanensis	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0

Table 4. A total of 37 probes were designed from β-tubulin sequence data. The probes bind different BT targets from different species of *Leptographium.*

Taxa	β-tubulin																
	1	3	4	5	6	8	10	11	12	13	14	15	16	17	18	20	21
L. abieticolens	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
L. abietinum	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
L. aenigmaticum	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0
L. albopini	0	0	0	0	0	0	0	1	1	0	1	0	1	0	0	0	0
L. alethinum	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0
L. americanum	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
L. antibioticum	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
L. aureum	1	1	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0
L. bistatum	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
L. brachiatum	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
L. brevicollis	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
L. bhutannense	0	0	1	0	0	0	1	0	1	0	0	0	0	0	1	0	0
L. calophylli	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
L. chlamydatum	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
L. clavigerum	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
L. costaricense	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
L. crassivaginatum	1	1	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0
L. curvisporum	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
L. douglassi	1	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0
L. dryocoetidis	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
L. elegans	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
L. eucalyptophilum	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
L. euphyes	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0
L. francke-grosmanniae	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
L. fruticetum	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1
L. grandifoliae	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
L. guttulatum	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0
L. huntii	0	0	0	0	0	0	0	1	1	1	0	0	0	0	0	0	0
L. koreanum	0	0	0	0	0	1	0	1	1	0	1	1	1	0	0	0	0
L. laricis	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0
L. leptographioides	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
L. longiclavatum	1	0	0	0	1	0	0	1	1	0	0	0	0	0	0	0	0

Table 4. Continued

Taxa	β-tubulin																
	1	3	4	5	6	8	10	11	12	13	14	15	16	17	18	20	21
L. lundbergii	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0
L. penicillatum	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1	0	0
L. peucophillum	0	0	1	0	0	0	1	0	0	0	0	0	0	1	1	0	0
L. piceaperdum	0	0	0	0	0	0	0	1	1	1	0	0	0	0	0	0	0
L. pineti	0	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0
L. pini-densiflorae	0	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0
L. pityophilum	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
L. procerum	0	0	1	0	0	0	1	0	0	0	0	0	0	1	1	0	0
L. profanum	0	0	1	0	0	0	1	0	0	0	0	0	0	0	1	0	0
L. pruni	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
L. pyrinum	1	0	0	0	1	0	0	1	1	0	0	0	0	0	0	0	0
L. reconditum	0	0	1	0	0	0	0	1	1	0	0	0	0	0	0	0	0
L. robustum	1	0	0	0	1	0	0	1	1	0	0	0	0	0	0	0	0
L. rubrum	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
L. serpens	0	0	1	0	0	0	0	1	1	0	0	0	0	0	0	0	0
L. sibiricum	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
L. terebrantis	1	1	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0
L. trinacriforme	1	0	0	0	1	0	0	1	1	0	0	0	0	0	0	0	0
L. truncatum	0	0	0	0	0	1	0	1	1	1	0	0	0	0	0	0	0
L. wageneri v. ponderosa	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0
L. wageneri v. pseudotsugae	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0
L. wageneri v. wageneri	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
L. wingfieldii	1	0	0	0	1	0	0	1	1	0	0	0	0	0	0	0	0
L. yunnanensis	0	0	0	0	0	1	0	1	1	1	1	1	1	0	0	0	0

Table 4. Continued

Taxa	β-tubulin										
	23	24	25	26	28	29	30	31	33	34	37
L. abieticolens	0	0	0	0	0	0	0	0	0	0	0
L. abietinum	0	0	0	1	0	0	0	0	0	1	0
L. aenigmaticum	0	0	0	0	0	0	0	0	0	0	0
L. albopini	0	0	0	0	0	0	0	0	0	0	0
L. alethinum	0	0	0	0	0	0	0	0	0	0	0
L. americanum	0	0	0	1	0	0	0	0	0	1	0
L. antibioticum	0	0	0	0	0	0	0	0	0	0	0
L. aureum	0	0	0	0	0	0	0	0	0	0	0
L. bistatum	0	0	0	0	0	0	0	0	0	0	0
L. brachiatum	1	0	0	0	0	0	0	0	0	0	0
L. brevicollis	0	1	0	0	0	0	0	0	0	0	0
L. bhutannense	0	0	0	0	0	0	0	0	0	0	0
L. calophylli	0	0	0	0	0	0	0	0	0	0	0
L. chlamydatum	0	0	1	0	0	0	0	0	0	0	0
L. clavigerum	0	0	0	0	0	0	0	0	0	0	0
L. costaricense	0	1	0	0	0	0	0	0	0	0	0
L. crassivaginatum	0	0	0	0	0	0	0	0	0	0	0
L. curvisporum	0	0	0	0	0	0	0	0	0	0	0
L. douglassi	0	0	0	0	0	0	0	0	0	0	0
L. dryocoetidis	0	1	0	0	0	0	0	0	0	0	0
L. elegans	0	0	0	0	0	0	0	0	0	0	0
L. eucalyptophilum	0	0	0	0	0	0	0	0	0	0	0
L. euphyes	0	0	0	0	0	0	0	0	0	0	0
L. francke-grosmanniae	0	0	0	0	0	0	0	0	0	0	0
L. fruticetum	0	0	0	0	0	0	0	0	0	0	0
L. grandifoliae	0	0	0	0	0	0	0	0	1	0	0
L. guttulatum	0	0	0	0	0	0	0	0	0	0	1
L. huntii	0	0	0	0	0	0	0	0	0	0	0
L. koreanum	0	0	0	0	0	0	0	0	0	0	0
L. laricis	0	0	0	0	0	0	0	0	0	0	0
L. leptographioides	0	0	0	0	0	0	0	0	0	0	0
L. longiclavatum	0	0	0	0	0	0	0	0	0	0	0
L. lundbergii	0	0	0	0	0	0	0	0	0	0	1
L. neomexicanum	0	0	0	0	0	0	0	0	0	0	0
L. penicillatum	0	0	0	0	0	0	0	0	0	0	0
L. peucophillum	0	0	0	0	0	0	0	0	0	0	0
L. piceaperdum	0	0	0	0	0	0	0	0	0	0	0
L. pineti	0	0	0	0	0	0	0	0	0	0	0

Table 4. Continued

Taxa	β-tubulin										
	23	24	25	26	28	29	30	31	33	34	37
L. pini-densiflorae	0	0	0	0	0	0	0	0	0	0	0
L. procerum	0	0	0	0	0	0	0	0	0	0	0
L. profanum	0	0	0	0	0	0	0	0	0	0	0
L. pruni	0	1	0	0	0	0	0	0	0	0	0
L. pyrinum	0	0	0	0	0	0	0	1	0	0	0
L. reconditum	0	0	0	0	0	0	0	0	0	0	0
L. robustum	0	0	0	0	0	0	0	0	0	0	0
L. rubrum	1	0	0	0	0	0	0	0	0	0	0
L. serpens	0	0	0	0	0	0	0	0	0	0	0
L. sibiricum	0	0	0	0	0	0	0	0	0	0	0
L. terebrantis	0	0	0	0	0	0	0	0	0	0	0
L. trinacriforme	0	0	0	0	0	0	1	0	0	0	0
L. truncatum	0	0	0	0	0	0	0	0	0	0	0
L. wageneri v. ponderosa	0	0	0	0	1	1	0	0	0	0	0
L. wageneri v. pseudotsugae	0	0	0	0	1	1	0	0	0	0	0
L. wageneri v. wageneri	0	0	0	0	0	0	0	0	0	0	0
L. wingfieldii	0	0	0	0	0	0	0	0	0	0	0
L. yunnanensis	0	0	0	0	0	0	0	0	0	0	0
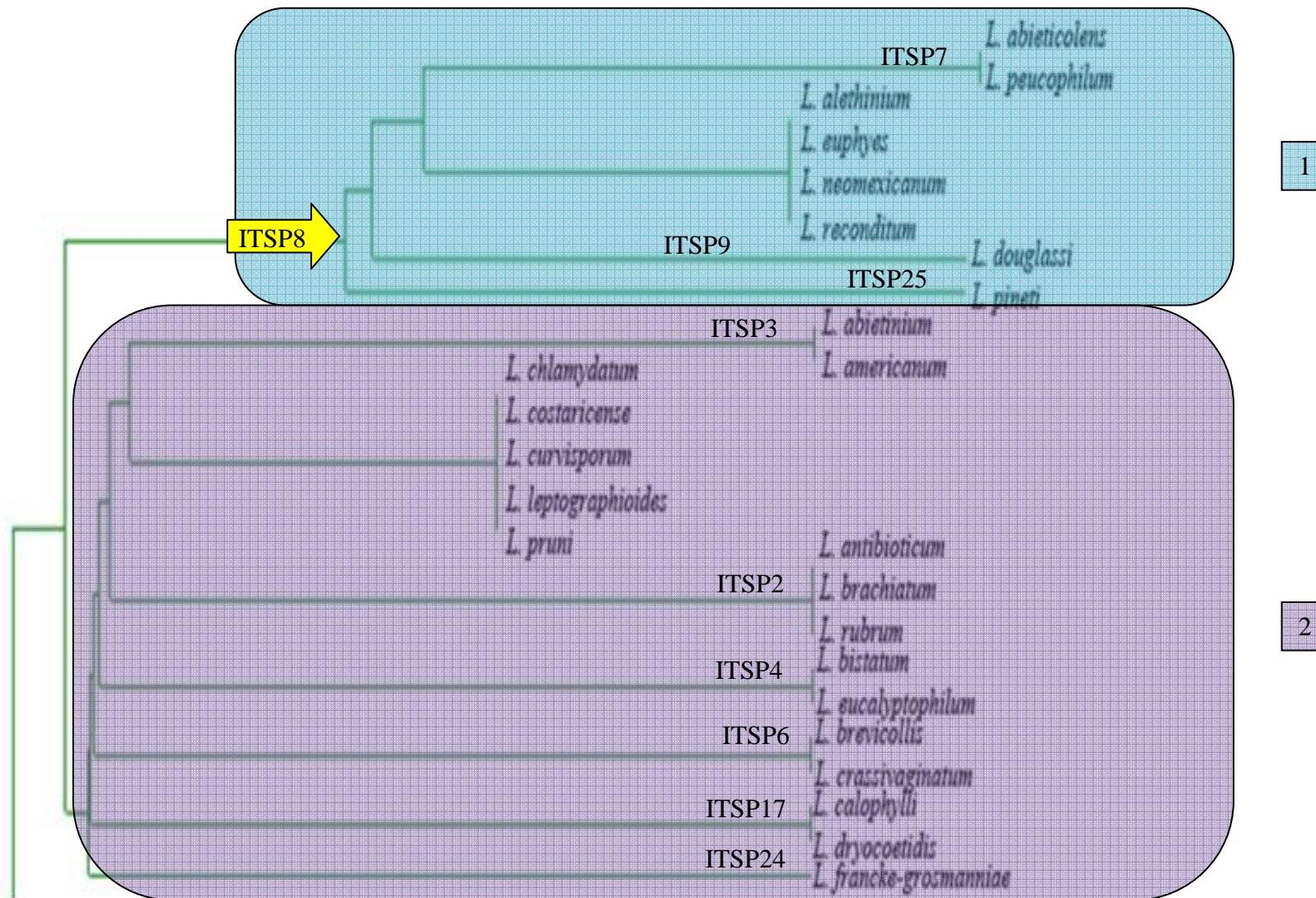
Figure 1. The phenogram for the internal transcribed spacer region 2 shows basic species delineation for *Leptographium* based on the probes designed from this region. The node probes indicated by yellow arrays divide the genus into four clades with 14 taxa specific probes.
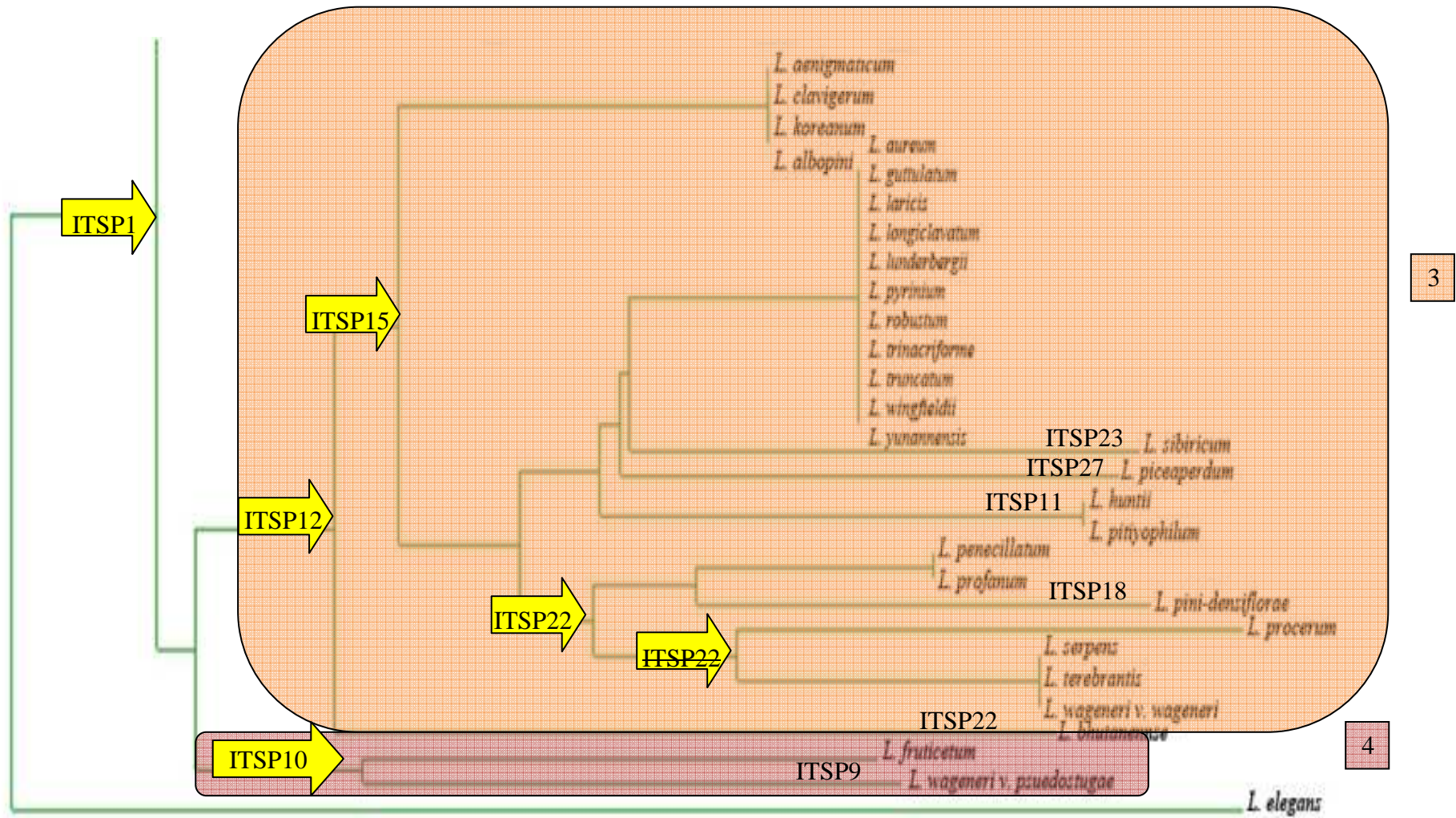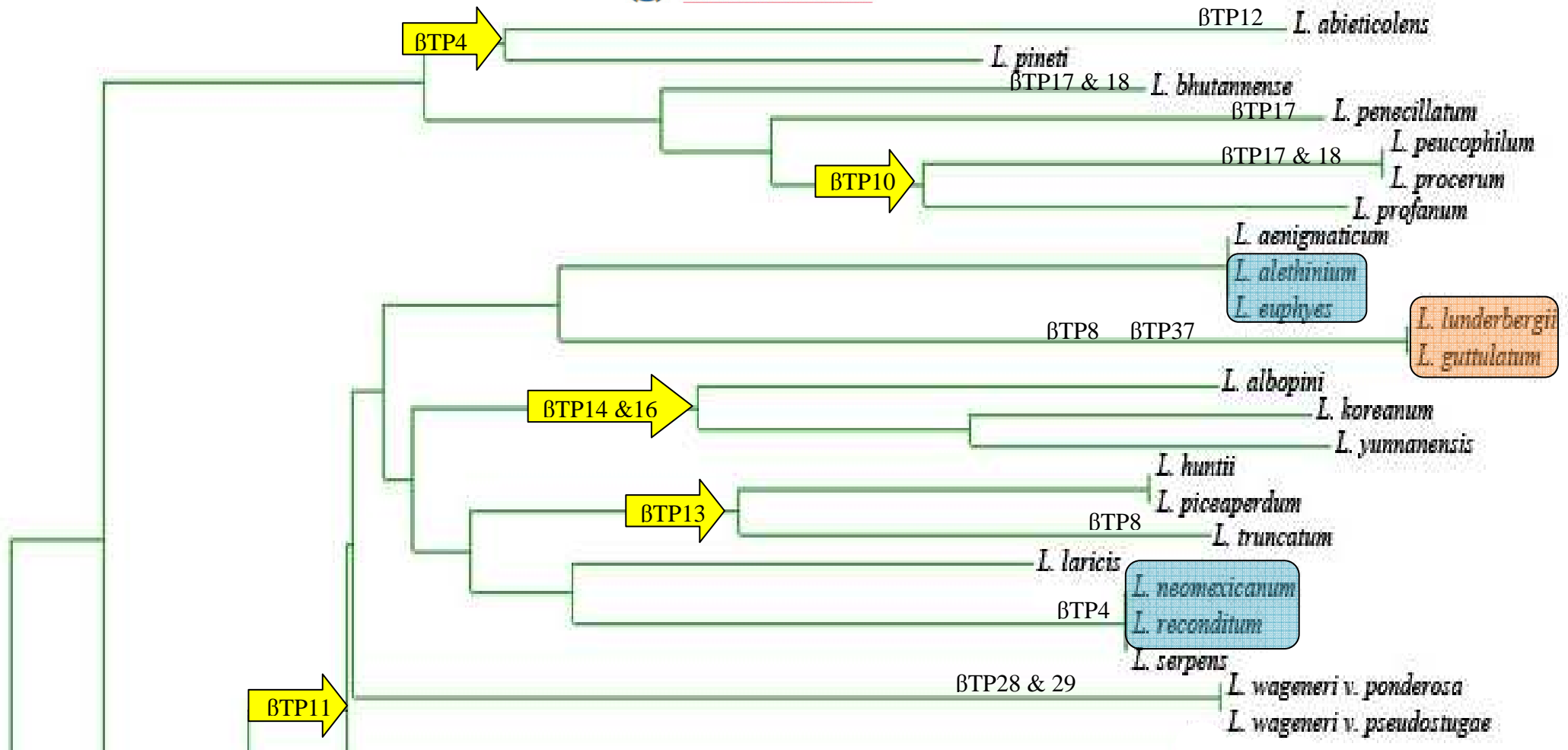
Figure 1. Continued

Figure 2. The phenogram derived from probes designed from the β-tubulin region. Yellow arrows indicate node probes and branch probes that define taxa are indicated on the branch. There are 12 node probes for the βT and 25 taxa specific probes.
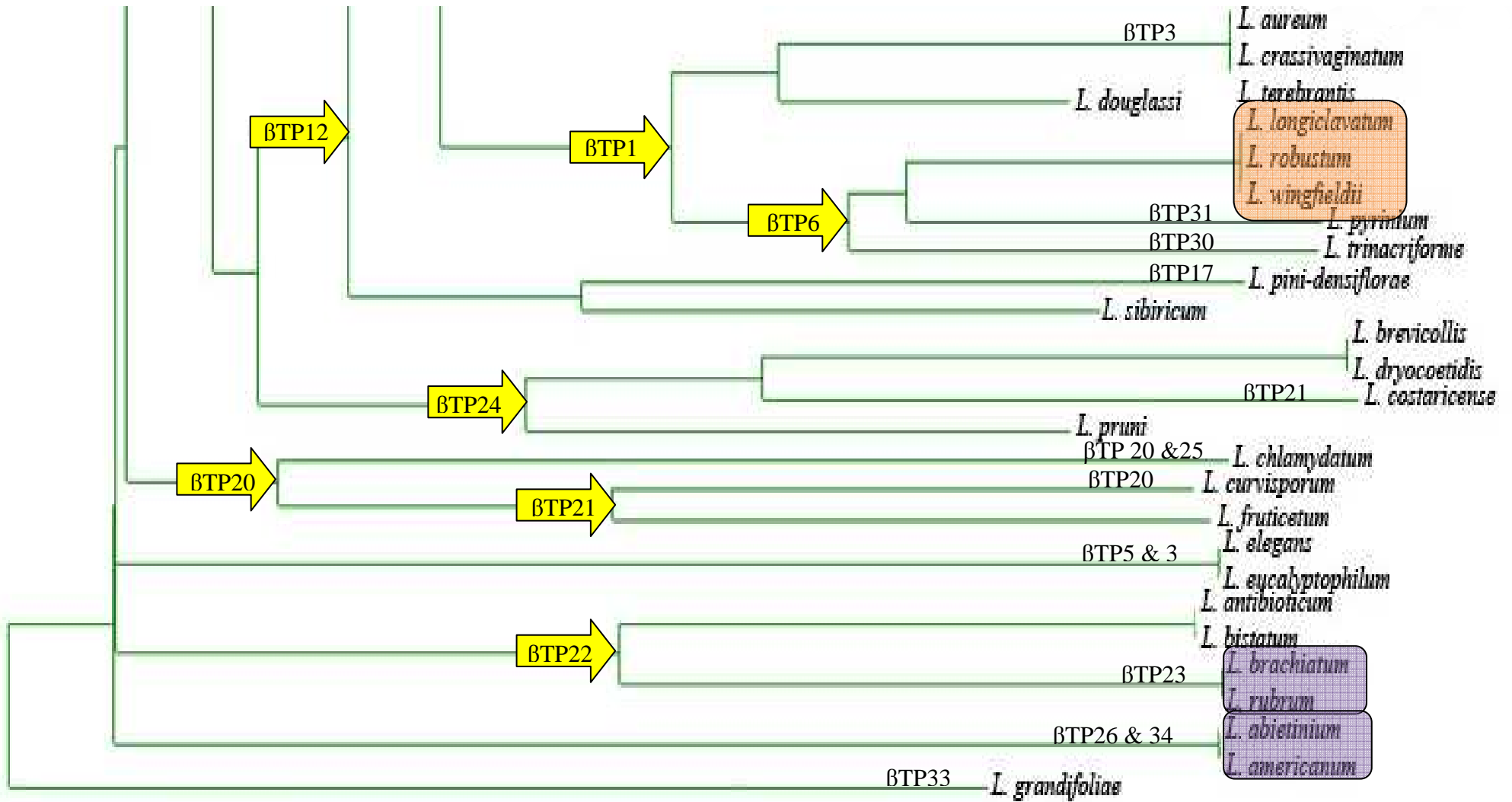
Figure 4. Continued.
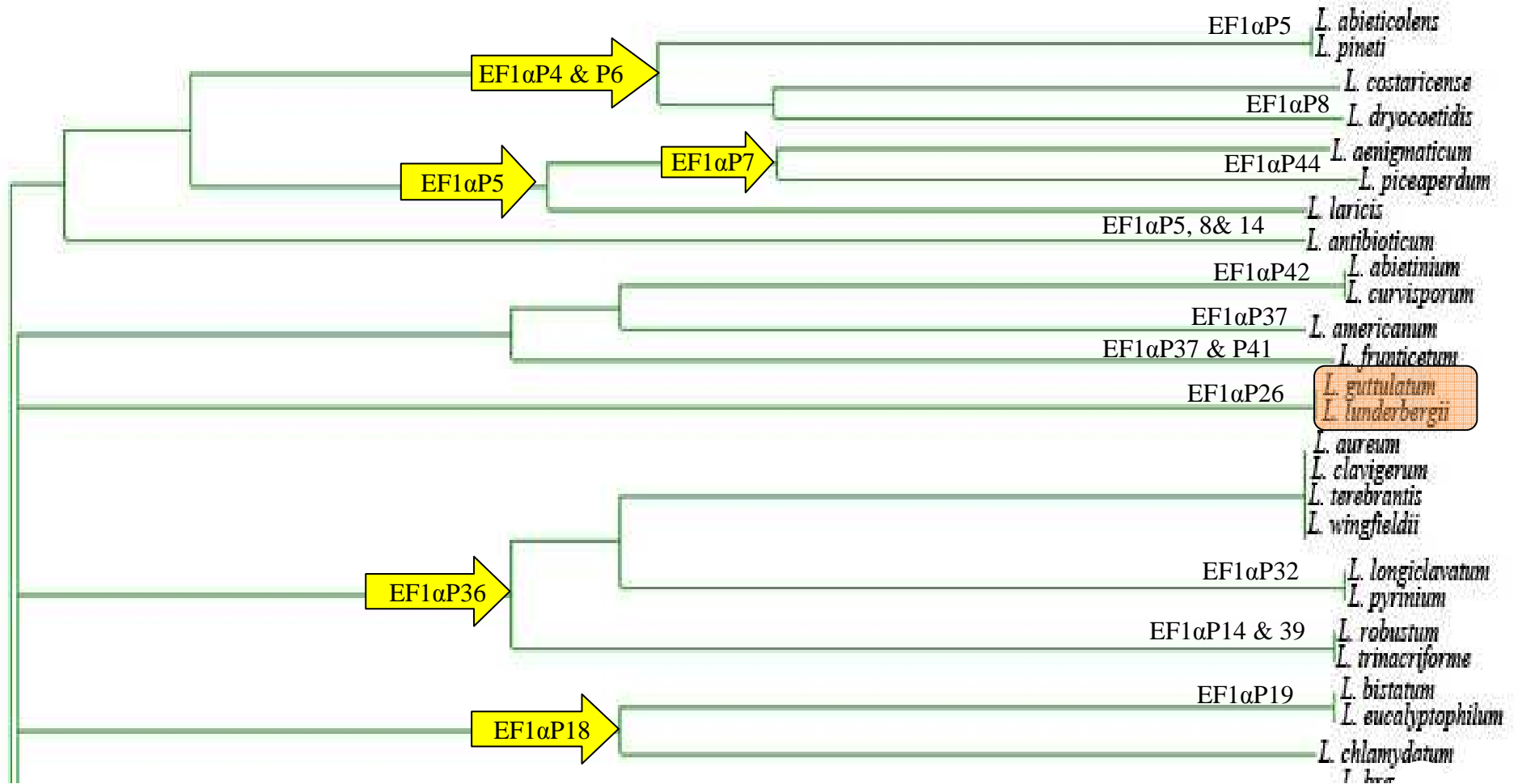
Figure 5. The annotated phenogram derived from the probe set for the translational elongation factor 1α. Yellow arrows indicate a probe that defines a node and probes on branches define taxa. The majority of the probes define taxa for this gene region.
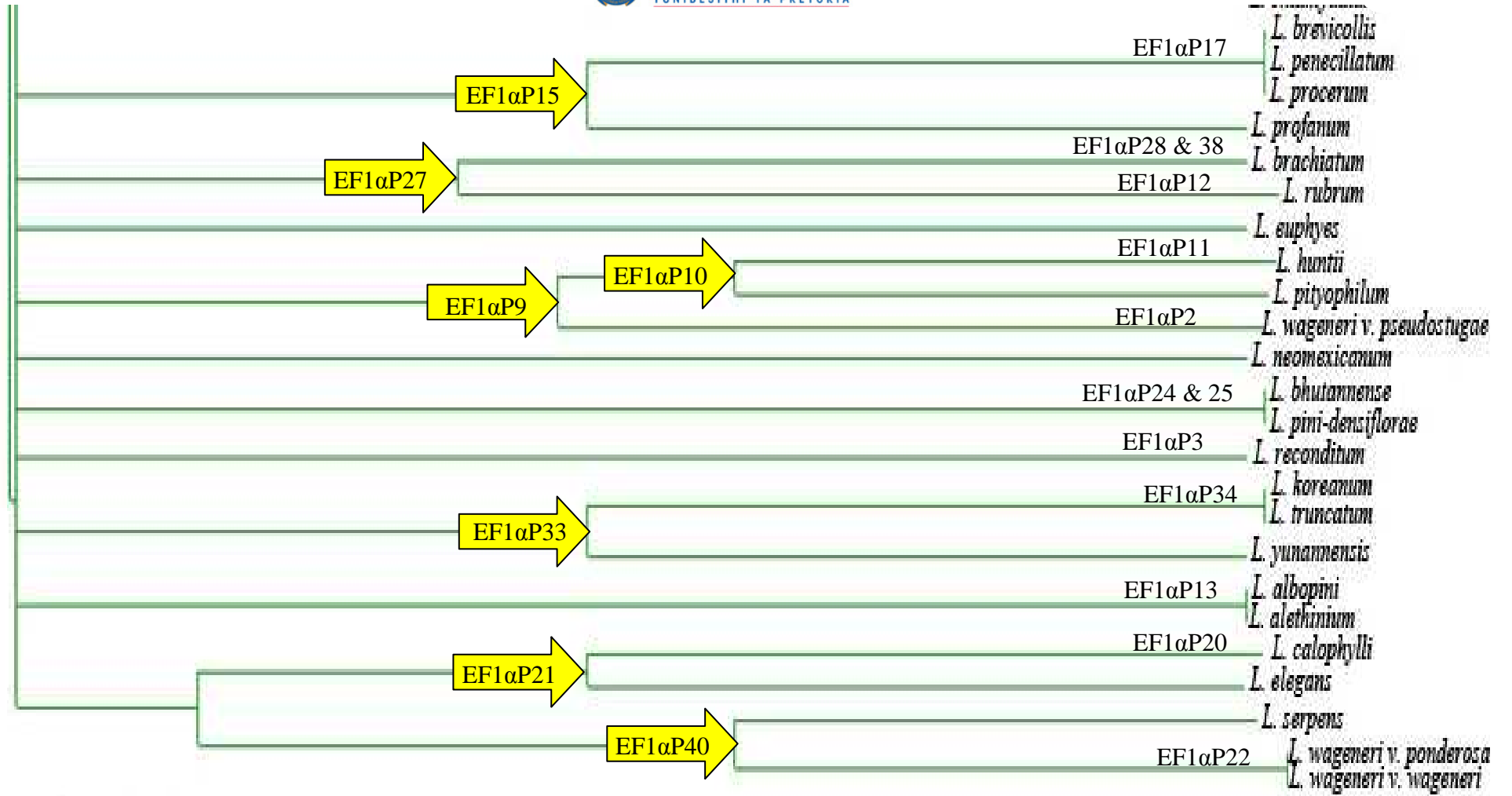
Figure 3. Continued

4. **Chapter 4:** Barcoding and microcoding using 'identiprimers' with *Leptographium* species

Natalie R. Van Zuydam[a], Dina Paciura[b], Karin Jacobs[c], Michael J. Wingfield[a] and Brenda D. Wingfield[a, *]

[a]Department of Genetics, [b]Department of Microbiology and Plant Pathology, Forestry and Agricultural Biotechnology Institute (FABI), University of Pretoria, Pretoria, 0002, South Africa [c]Department of Microbiology, University of Stellenbosch, Private Bag X-1, Matieland, Stellenbosch, South Africa

Abstract

Leptographium species provide an ideal model to test the applications of a PCR microcoding system for other genera of ascomycetes. *Leptographium* species are closely related phylogenetically and share similar gross morphology. Probes designed for a PhyloChip for *Leptographium* have been transferred and tested as primers for a PCR diagnostic against *Leptographium* species. The primers were combined with complementary universal primers to identify known and undescribed species of *Leptographium*. The primer set was optimised for 56 species including the three varieties of *L. wageneri* and then blind tested against 10 random DNA samples. The protocols established in this study successfully identified species from the blind test as well as eight previously undescribed isolates of *Leptographium*. The undescribed isolates were identified as new species of *Leptographium* using the microcoding PCR identification system established in this study. The primers that were positive for each undescribed isolate were used to determine close relatives of these species and some of their biological characteristics. The transfer of oligonucleotides from a microarray platform to a PCR diagnostic was successful and the identification system is robust for both known and unknown species of *Leptographium*.

4.1. Introduction

PCR and the resulting band patterns have been used in a number of studies to identify different species of fungi (Chen *et al.*, 2001; Fujita *et al.*, 2001; Hamelin *et al.*, 1996). A common approach to identification is to amplify a gene region using universal primers, sequence this region and then perform a phylogenetic analysis on the sequence data. This approach has been used to identify new species of *Leptographium* or to confirm the identity of previously described species (Jacobs *et al.*, 2000; Jacobs *et al.*, 2005). The reverse approach is to use available sequence data to design specific primers that amplify a particular amplicon, thus circumventing a sequencing step (Bäckmann *et al.*, 1999; Hamelin *et al.*, 1996). These primers can be present as a pair in a PCR mix or multiplexed with other specific primers (Fujita *et al.*, 2001; Jackson *et al.*, 2004; Redecker, 2000). Primers for species identification in bacteria have been designed around unique polymorphisms that are species specific (Bäckmann *et al.*, 1999; Easterday *et al.*, 2005). In other cases, universal primers are designed to amplify a single amplicon of a particular length that is definitive of a species (Chen *et al.*, 2001; Fujita *et al.*, 2001).

The use of PCR and specific primers is based on the phylogenetic species concept, which uses sequence polymorphisms to delineate fungal species (Taylor *et al.*, 2000). Species can be delineated through shared and unshared sequence characteristics or polymorphisms (Taylor *et al.*, 2000). Primers can be designed around these polymorphisms, so that a single amplicon of a known size will be amplified from a DNA sample, only if the primer sequence is present in the genome that is being used in the amplification (Bäckmann *et al.*, 1999; Hamelin *et al.*, 1996; Tran and Rudney, 1996). This approach has been used for microcoding species of fungi and can be equally as diagnostic as PCR amplification followed by sequencing (Summerbell *et al.*, 2005).

Microcoding has been defined as a specific type of DNA barcoding that allows for the identification of organisms to the genus and species level (Summerbell *et al.*, 2005). DNA barcoding traditionally uses highly conserved genes, like the 18S rDNA gene and the large ribosomal subunit, to assign fungi to higher taxonomic levels such as family and order (Summerbell *et al.*, 2005). Generic and species level gene regions include the internal transcribed spacer region (ITS), β-tubulin (βT) and translational elongation factor (EF1α) as well as the mitochondrial CO1 gene that are less conserved (Seifert *et al.,*2007; Summerbell *et al.*, 2005). The primers used for microcoding are short 20-mer primers that are designed based on variable generic and species level genes and serve to identify organisms to either of these levels. In a previous study, a set of 20-mer primers were designed for

56 species of *Leptographium* (Van Zuydam *et al.,* Chapter 3) to be used on a microarray platform as a PhyloChip.

The PhyloChip for *Leptographium* was designed using a hierarchical set of probes designed from the ITS2, βT and EF1α gene regions available for 56 species. The design consisted of a mixture of common and unique 20-mer probes that identified individual species in different combinations. The ITS2 probes included a single generic probe ITSP1 and other specific probes, which identified particular nodes on a phenogram and delineated species. The ITS2 probes split the genus into five clades that approximated phylogenetic and morphological groups within the genus (Van Zuydam *et al*., Chapter 3). The βT and EF1α probes identified smaller clades and specific species to on the more general identification based on ITS2 primers.

In the current study, we modelled a PCR diagnostic system on a PhyloChip design concept using the probes designed for the *Leptographium* PhyloChip. The system uses the phenograms constructed from the probes for the PhyloChip to define the sequence of diagnostic PCRs that led to species identification (Van Zuydam *et al.,* Chapter 3). If a primer is common to a group of species, it will define a node and if a primer is species-specific, it will define a branch (Van Zuydam *et al.,* Chapter 3). Therefore, PCRs using primers for a node will be conducted before those defining a species. This is similar in organisation to PhyloChips but the primers are combined with either a forward or, a reverse universal primer that allows a dynamic system that can identify known as well as new species. We chose to validate our primers on the fungal genus *Leptographium.*

Leptographium is the anamorph genus of *Grosmannia* and is relatively small when compared to other genera within the *Ophiostomatoid* fungi (Zipfel *et al.*, 2006). *Leptographium* species are characterised by mononematous, branched conidiophores that produce aseptate, hyaline conidia in a slimy matrix (Jacobs, 1999; Kendrick, 1962). *Leptographium* species are similar in appearance, thus making identifications using morphological characters presents a challenge. It is possible to identify *Leptographium* species accurately using molecular techniques and this is achieved through constructing phylogenies from available sequence data (Jacobs *et al.*, 2000; Zhou *et al.*, 2000). The molecular characters are combined with morphological characters to describe new species (Jacobs *et al.*, 2000; Jacobs *et al.*, 2001). As a result, *Leptographium* presents a unique case in the ascomycetes, as there is a comprehensive sequence data set for 56 species across regions of the ITS2 βT and EF1α genes (Jacobs *et al.*, 2006).

The sequence data available for the genus *Leptographium* have been used to design a probe set for a PhyloChip based on shared and unshared sequence polymorphisms. The phenograms constructed from the probe set approximate the phylogenies and morphological groups presented by Jacobs *et*

al., (2006) (Van Zuydam *et al.,* Chapter 3). Thus, in this study, the aim was to microcode 56 known species of *Leptographium* and eight previously undescribed isolates using probes from a PhyloChip as identiprimers combined with a complementary universal primer.

4.2. Materials and methods

4.2.1. DNA isolation and isolates

Isolates used in this study were identified according to morphological characters. The species identification of all isolates had been previously confirmed using DNA sequence comparisons (Van Zuydam *et al,* Chapter 2; Jacobs *et al.,* 2006) DNA was extracted using the Soil Microbe DNA isolation kit (Fermentas, USA) according to the manufacturers' instructions.

4.2.2. Primers

The primers used for this study were designed in a previous study as probes for a species diagnostic microarray (Van Zuydam *et al.,* Chapter 3; Table 5). They were combined with a universal primer that was designed from the opposite strand and was as the name implies identical in DNA sequence for all species in *Leptographium*. Identiprimers for the ITS2 region were combined with either ITS3 (+) or LR3 (-) (White *et al.*, 1990), identiprimers for βT were combined with either Bt2a (+) or BT2b (-) (Glass and Donaldson, 1995) and the identiprimers for EF1α were combined with either EF1F (+) or EF2R (-).

4.2.3. PCR optimization

4.2.3.1. Multiplex PCR

The identiprimers for Clade 1 were combined into a multiplex PCR that consisted of 2.5mM $MgCl_2$, 1x Buffer, 0.4mM dNTPs, and 1U SuperTherm *Taq* polymerase (Southern Cross), 0.4mM of ITSP1, ITSP7, ITSP8 and ITSP9, 1.6mM of LR3, 0.8xV DNA in a 5µl reaction. These primers were optimized against all the species in clade 1 to amplify the correct regions.

4.2.3.2. Standard PCR

The reaction conditions of the PCR using clade-specific probes were optimised so that amplicons were produced only when DNA from isolates within a clade were used in the reaction. The probes were then optimised for DNA from isolates within clades along the sub branches. The PCRs were optimised according to temperature, magnesium chloride concentration and 2-pyrrolidone concentration on each species used in this study (Table 1). A negative control containing no DNA was included in every optimisation step. The stock solution of 2-pyrrilidone was diluted 1 in 10 and further dilutions were made from this working solution. The standard PCR mixture used consisted of 2.5mM $MgCl_2$, 1x Buffer, 1U of SuperTherm *Taq* polymerase (Southern Cross, South Africa), 0.4mM dNTP mix, 0.4mM of each primer and 0.08 x reaction volume of DNA. Five microlitre reactions were used and the entire volume was used to determine amplicon presence and size. Amplicons were separated by gel electrophoresis through a 3% Agarose gel at 80V for 40 minutes and stained with GelRed (Anatech, USA) and visualized using UV light.

4.2.4. Blind test

Ten DNA samples representing 10 species were independently chosen at random from DNA isolated from the 56 species included in this study and relabelled 1-10. These samples were analysed and identified to species level using the protocols established in this study. Positive controls using the DNA from amplicon positive species and a negative control containing no DNA were included in every PCR identification step. The identification process was repeated in triplicate to measure reproducibility.

4.2.5. Identification of new species

Eight previously undescribed (Table 6) isolates of *Leptographium* were included in this study. These species were tested using established protocols from this study, the same positive and negative controls included in the Blind test were included in the PCR identification steps.

4.2.6. Phenogram construction

Phenograms were constructed using NTSYSpc21v2.11 (Applied Biostatics) cluster analysis and the neighbour joining algorithms. The positive amplifications obtained in the blind test and the identification of new species was included in a large matrix. Phenograms were drawn for each of

the gene regions. Smaller phenograms were constructed for the eight new species, *L. bhutannense, L. yunnanensis, L. procerum* and *L. koreanum* using the same method.

4.3. Results

4.3.1.   Primers and PCR optimisation

Individual diagnostic PCRs were optimised for 56 species included in this study. The details of the optimised conditions are summarised in Table 5. A representation of fragments amplified using identiprimers for species from each of the ITS2 clades are shown in Figures 1-8. Non-specific binding was encountered for BTP1, EF1αP32, BTP30 and BTP31 (Table 5) resulting in multiple bands thus these are not useful as identiprimers and must be redesigned. A selection of gels showing optimised PCRs are included in appendix 2.

4.3.2.   Blind test

DNA isolations 3, 4, 6 and 8 from the blind test were accurately identified as *L. procerum, L. pineti, L. pini-densiflorae* and *L. fruticetum* using ITSP2 identiprimers (figure 9). Blind test 1, 5, 7 and 9 were identified as *L. profanum, L. lundbergii/L. guttulatum, L. wageneri* var. *ponderosa* and *L. chlamydatum* using ITSP2 and βT identiprimers (Figures 9 and 10). Blind test 2 was identified as *L. euphyes* using identiprimers from all three gene regions (Figures 9, 10 and 11). Blind test 10 could not be identified to species level due to the failure of BTP30 and BTP31 and is grouped in a large group by the ITSP2 identiprimers (Figure 9). The matrices can be found in appendix 3.

4.3.3.   Undescribed isolates

The isolates included in this study are listed in Table 7 and the results of the PCR are listed in Table 7. All the previously undescribed isolates included were recognised as new *Leptographium* species by the diagnostic technique developed in this study. The species were all positive for the generic ITSP1 primer that indicated that they belong to the genus *Leptographium*. *Leptographium* sp 1, 2, 3, 4, 5 and 8 grouped with *L. elegans* and *L. huntii* (Figure 9). *Leptographium* sp 6 and 7 grouped closely to *L. abieticolens* and *L. peucophillum* in the comprehensive ITS2 tree (Figure 9). The comprehensive βT tree showed that *Leptographium* sp 1 and 4 grouped into a clade with *L. huntii, L. piceaperdum, L. truncatum, L. albopini, L. koreanum, L. yunnanensis, L. guttulatum* and *L. lundbergii. Leptographium* sp 2, 3 and 5 grouped with *L. brevicollis, L. dryocoetidis* and *L. pruni,*

and *Leptographium* sp 6, 7 and 8 grouped with another large clade that included *L. calophylli, L. clavigerum, L. leptographioides, L. francke-grosmanniae, L. pityophilum, L. wageneri* var. *wageneri* and *L. sibiricum* (Figure 10). The comprehensive EF1α tree showed that *Leptographium* sp 1 and 2 grouped with *L. neomexicanum* and *L. reconditum; Leptographium* sp 4 and 8 grouped with *L. reconditum; Leptographium* sp 3 and 5 group with *L. pruni, L. crassivaginatum, L. douglassi, L. francke-grosmanniae, L. leptographioides, L. sibiricum, L. peucophillum* and *L. grandifoliae,* and *Leptographium* sp 6 and 7 grouped with *L. brachiatum* and *L. rubrum* (Figure 11).

Three smaller phenograms were constructed from subsets of the ITS2, βT and EF1α matrices to include the eight undescribed *Leptographium* isolates, *L. yunnanensis, L. bhutannense, L. procerum* and *L. koreanum*. The ITS2 phenogram represents *Leptographium* sp 6 and sp 7 as a single species that is related to *L. procerum* (Figure 12). The ITS2 phenogram also shows that *Leptographium* sp 1 and 2 are closely related as are *Leptographium* sp 3 and 4 (Figure 12). *Leptographium* sp 8 and 5 occupy separate branches and show no close associations with the other *Leptographium* species (Figure 12). The βT phenogram showed that *Leptographium* sp 1 and 4 are closely related to *L. yunnanensis* and *L. koreanum; Leptographium* sp 2 and 3 formed a single taxon that is related to *Leptographium* sp 5, *L. bhutannense* and L. *procerum*, and *Leptographium* sp 6, 7 and 8 formed a single taxon that was related to *Leptographium* sp 2, 3, 5, *L. bhutannense* and *L. procerum* (Figure 13). The EF1α phenogram showed that *Leptographium* sp 1, 4 and 8 grouped away from the other taxa and were related to each other; *Leptographium* sp 2, 3, 5, 7, *L. yunnanensis, L. koreanum, L. bhutannense* and *L. procerum* grouped together with *Leptographium* sp 3 and 5 collapsed into a single taxon with *L. yunnanensis, L. koreanum, L. bhutannense* and *L. procerum* (Figure 14).

4.4. Discussion

This is the first study to apply a microcoding system to an entire genus of ascomycetous fungi. Typically, *Leptographium* species are difficult to identify using morphological characters so the species are differentiated according to both morphological and molecular characters (Jacobs, 1999). In a previous study, phylogenies were constructed from partial sequences of the βT, ITS2 and EF1α regions and revealed that *Leptographium* species are closely related to each other (Jacobs *et al.*, 2006; Jacobs *et al.*, 2001). Probes were designed, for a PhyloChip, from these gene regions to have at least a 10% difference between the primer and similar, but incorrect, target sequences (Van Zuydam *et al.,* Chapter 3). These probes were applied to this study as identiprimers for species identification. In this study, we have achieved species differentiation using identiprimers in PCRs comparable to the differentiation achieved through phylogenetic analysis.

The identification system established in this study is unconventional as primers were designed from multiple gene regions and used in a hierarchical sequence. Identification began with identiprimers from the ITS2 region and then higher order identiprimers from the βT and EF1α regions were used to achieve full delineation of species. This hierarchical system has been adopted for PhyloChip studies (Loy *et al.*, 2002; Metfies *et al.*, 2008), but has not been transferred to a PCR diagnostic. More commonly, for fungi, PCR diagnostics have been designed from a single gene region that only differentiates among a few species (Chen *et al.*, 2001; Fujita *et al.*, 2001; Hamelin *et al.*, 1996). In a study by Fujita *et al.,* 2001, ITS1, ITS3 and ITS4, (White *et al.*, 1990) primers were optimised in a multiplex to amplify the ITS1 and ITS2 regions, in order to type 120 fungal strains consisting of 30 species of yeast. The differences in length of the ITS1 and ITS2 regions among species were used to differentiate the species from each other (Fujita *et al.*, 2001). Our study used a combination of selective identiprimers and amplicon size to identify *Leptographium* species. With closely related taxa, as is the case within the genus *Leptographium,* using one gene region to differentiate species would be not be possible. The results of this study illustrated this fact, we therefore suggest that if this identification technique were applied generally to ascomycetes, it would be essential to use a number of gene regions and associated primers.

This study showed that it is possible to transfer 20-mer probes from a microarray study to a PCR diagnostic application. The design for the microarray was suited to the PCR diagnostic application as the probes were similar in length to PCR primers and multiple probes were designed. It was not possible to design a unique probe for each species within the genus *Leptographium* therefore,

multiple probes from multiple gene regions were designed (Van Zuydam *et al.,* Chapter 3). These probes were transferred to the PCR diagnostic as identiprimers.

The identiprimers were incorporated with complementary universal primers and this allowed for a dynamic identification system rather than a static PCR diagnostic based on a pair of species-specific primers. ITS2 primers were multiplexed in order to categorise DNA samples according to shared sequence characteristics in the ITS2 region using a single reaction. This approach was successful for one subset but not for all the ITS2 primers. However, identification using single plex PCRs was very successful as only four primer failures were encountered despite the large primer set and number of species tested in this study. Primers were determined to have failed if they produced random amplification, or where did not facilitate amplification at all. When the primers had been optimised and interrogated for known species, they were tested on undescribed isolates of *Leptographium* and revealed some intriguing results.

The identiprimers developed in this study support a phenogram that can be compared to an amplification profile to identify described and new species. As the phenograms, approximate the the phylogenies constructed by Jacobs *et al.* (2006), our design also allowed for inferences about phylogenetic relationships to be drawn. The undescribed isolates were all positively identified as representing new species of *Leptographium* and showed interesting cladistic associations and disassociations, indicated by the identiprimers. A dichotomy was observed within the new species according to their primer amplification profiles when they were compared to phenograms constructed in a previous study (Van Zuydam *et al.,* Chapter 3). *Leptographium* sp 1, *Leptographium* sp 2, *Leptographium* sp 4 and *Leptographium* sp 5 associated more closely with species that colonise coniferous hosts and *Leptographium* sp 6, *Leptographium* sp 7 an *Leptographium* sp 8 associated more closely with species that colonise non-coniferous hosts according to the ITS2 primers. These results are supported by the collection data and phylogenies for these species (Paciura *et al.,* unpublished). Higher order ITS2 primers showed that *Leptographium* sp 1, *Leptographium* sp 2, *Leptographium* sp 3, *Leptographium* sp 4 and *Leptographium* sp 5 are related to *L. bhutannense* and that *Leptographium* sp 6, *Leptographium* sp 7 are more closely related to *L. abieticolens* but also share sequence homology with *L. yunnanensis.*

The βT and EF1α primer associations of the undescribed isolates revealed more about their associations with each other and with known species. The primer profiles of *Leptographium* sp 2 and *Leptographium* sp 3 are very similar indicating that they are closely related and this relationship is supported by the phylogeny (Paciura *et al,* unpublished). The same is true for *Leptographium* sp 6 and *Leptographium* sp 7 that have similar profiles to each other but dissimilar to the other new
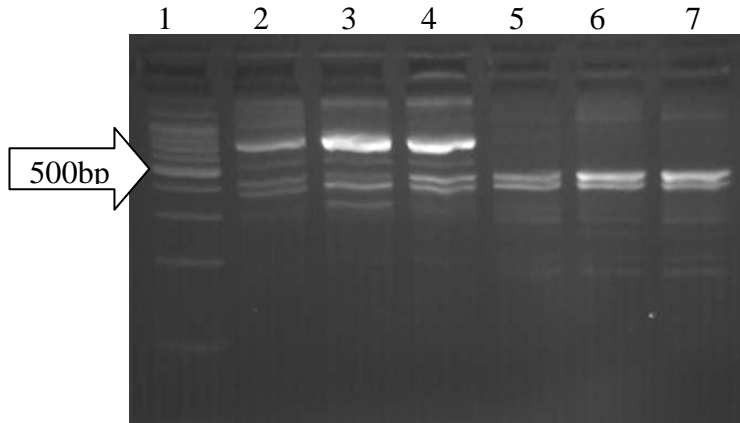
species; they are phylogenetically close to each other and distant to the other new species (Paciura *et al.,* unpublished). The split can be attributed to the difference in the hosts that they colonise. Thus, *Leptographium* sp 7 and *Leptographium* sp 6 were isolated from non-coniferous hosts while the other new species were isolated from coniferous hosts. The βT primers indicate that *Leptographium* sp 1 and *Leptographium* sp 4 are related to *L. yunnanensis.* This is a strong result as the diagnostic for *L. yunnanensis* used two specific primers rather than a specific primer paired with a universal primer. This is also reflected in the phylogenetic relationships of these two species (Paciura *et al,* unpublished).

The results obtained for the new species are interesting in terms of microcoding. Microcoding was proposed as the next step to barcoding by Summerbell *et al*., 2005. Here, the suggestion was that 20-mer oligonucleotides could be used to identify species at the genus and species level. Likewise, this study supports the use of short oligonucleotides in microcoding applications. It supports the use of multiple gene regions to identify species to genus and species level and proposes an identification system based on primer phenograms (Van Zuydam *et al.,* Chapter 3). We found that the relationships between the species based on primer sequence homology roughly resembled biological and phylogenetic relationships. This was true for the known species and the undescribed species of *Leptographium* included in this study. It indicates that DNA microcoding would be successful in identifying known and new species as well as indicating biological and phylogenetic relationships.

References

Bäckmann, A., Lantz, P.-G., Rådström, P., Olcén, P., 1999. Evaluation of an extended diagnostic PCR assay for detection and verification of the common causes of bacterial meningitis in CSF and other biological samples. Mol Cell Probes. 13, 49-60.

Chen, Y.-C., Eisner, J.D., Kattar, M.M., Rassoulian-Barret, S.L., Lafe, K., Bui, U., Limaye, A.P., Cookson, B.T., 2001. Polymorphic internal transcribed spacer region 1 DNA sequences identify medically important yeasts. J Clin Microbiolbiol. 39, 4042-4051.

Easterday, W.R., Van Ert, M.N., Simonson, T.S., Wagner, D.M., Kenefic, L.J., Allender, C.J., Keim, P., 2005. Use of single nucleotide polymorphisms in the plcR gene for specific identification of *Bacillus anthracis*. J. Clin. Microbiol. 43, 1995-1997.

Fujita, S.-I., Senda, Y., Nakaguchi, S., Hashimoto, T., 2001. Multiplex PCR using internal transcribed spacer 1 and 2 regions for rapid detection and identification of yeast strains. J Clin Microbiol. 39, 3617-3622.

Glass, N.L., Donaldson, G.C., 1995. Development of primer sets designed for use with the PCR to amplify conserved genes from filamentous ascomycetes. Appl Environ Microb. 61, 1323-1330.

Hamelin, R.C., Bérubé, P., Gignac, M., Bourassa, M., 1996. Identification of root rot fungi in nursery seedlings by nested multiplex PCR. Appl Environ Microb. 62, 4026-4031.

Jackson, C.R., Fedorka-Cray, P.J., Barrett, J.B., 2004. Use of a genus- and species-specific multiplex PCR for identification of *Enterococci*. J Clin Microbiol. 42, 3558-3565.

Jacobs, K., 1999. The genus *Leptographium:* A critical Taxonomic analysis, Forestry and Agricultural Biotechnology Institute (FABI) and the Department of Microbiology and Plant Pathology, University of Pretoria, Pretoria.

Jacobs, K., Solheim, H., Wingfield, B.D., Wingfield, M.J., 2005. Taxonomic re-evaluation of *Leptographium lundbergii* based on DNA sequence comparisons and morphology. Mycol Res. 10, 1149-1161.

Jacobs, K., Wingfield, B.D., Wingfield, M.J., 2006. *Leptographium:* Anamorphs of *Grosmannia*, The Ophiostomatoid Fungi: Expanding Frontiers, Morton Bay Research Station, North Stradbroke Island, Brisbane, Australia.

Jacobs, K., Wingfield, M.J., Pashenova, N.V., Vetrova, V.P., 2000. A new *Leptographium* species from Russia. Mycol Res. 104, 1524-1529.

Jacobs, K., Wingfield, M.J., Wingfield, B.D., 2001. Phylogenetic relationships in *Leptographium* based on morphological and molecular characters. Can J Bot. 79, 719-732.

Kendrick, W.B., 1962. The *Leptographium* complex. *Verticicladiella* S. Hughes. Can. J. Bot. 40, 771-797.

Loy, A., Lehner, A., Lee, N., Adamczy, J., Meier, H., Ernst, J., Schleifer, K.-H., Wagner, M., 2002. Oligonucleotide microarray for 16S rRNA gene-based detection of all recognized lineages of sulfate-reducing prokaryotes in the environment. Appl Environ Microb. 68, 5064-5081.

Metfies, K., Bborsutzki, P., Gescher, C., Medlin, L.K., Frickenhaus, S., 2008. PHYLOCHIPANALYSER — a program for analysing hierarchical probe sets. Mol Ecol Resour. 8, 99-102.

Redecker, D., 2000. Specific PCR primers to identify arbuscular mycorrhizal fungi within colonized roots. Mycorrhiza. 10, 73-80.

Seifert, K.A., Samson, R.A., de Waard, J.R., Houbraken, J., Lévesque, C.A., Moncalvo, J.-M., Louis-Seize, G., Hebert, P.D.N., 2007. Prospects for fungus identification using CO1 DNA barcodes, with *Penicillium* as a test case. PNAS. 104, 3901–3906.

Summerbell, R.C., Lévesque, C.A., Seifert, K.A., Bovers, M., Fell, J.W., Diaz, M.R., Boekhout, T., de Hoog, G.S., Stalpers, J., Crous, P.W., 2005. Microcoding: the second step in DNA barcoding. Phil Trans Roy Soc B. 360, 1897-1903.

Taylor, J.W., Jacobson, D.J., Kroken, S., Kasuga, T., Geise, D.M., Hibbett, D.S., Fisher, M.C., 2000. Phylogenetic species recognition and species concepts in fungi. Fungal Genet Biol. 31, 21-32.

Tran, S.D., Rudney, J.D., 1996. Multiplex PCR using conserved and species-specific 16S rRNA gene primers for simultaneous detection of *Actinobacillus actinomycetemcomitans* and *Porphyromonas gingivalis*. J Clin Microbiol. 34, 2674-2678.

White, T.J., Bruns, T., Taylor, J., 1990. Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. In: A. Innis, D. Gelfand, J. Sninsky, T. White (Eds.), PCR Protocols: a guide to methods and applications, AP, San Diego, pp. 315-322.

Zhou, X.D., Jacobs, K., Morelet, M., Ye, H., Lieutier, F., Wingfield, M.J., 2000. A new *Leptographium* species associated with *Tomicus piniperda* in south-western China. Mycoscience. 41, 573-578.

Zipfel, R.D., de Beer, Z.W., Jacobs, K., Wingfield, B.D., Wingfield, M.J., 2006. Multigene phylogenies define *Ceratocystiopsis* and *Grosmannia* distinct from *Ophiostoma*. Stud. Mycol. 55, 75-97.

Figure 6. A 3% Agarose gel resolved products of a multiplex PCR for *L. abieticolens* and *L. alethinum* at 100V for 40 minutes. The product for ITSP1 the generic probe amplifies 370bp fragment, ITSP8 that is a node identiprimer amplified a 410bp fragment and ITSP7 a higher order node identiprimer amplifies a 700bp fragment.

Lane 1:   100bp ladder
2:      *L. abieticolens* BTP4 68°C
3:      *L. abieticolens* BTP4 68°C
4:      *L. abieticolens* BTP4 68°C
5:      *L. abieticolens* BTP17 and 18 68°C
6:      *L. abieticolens* BTP17 and 18 68°C
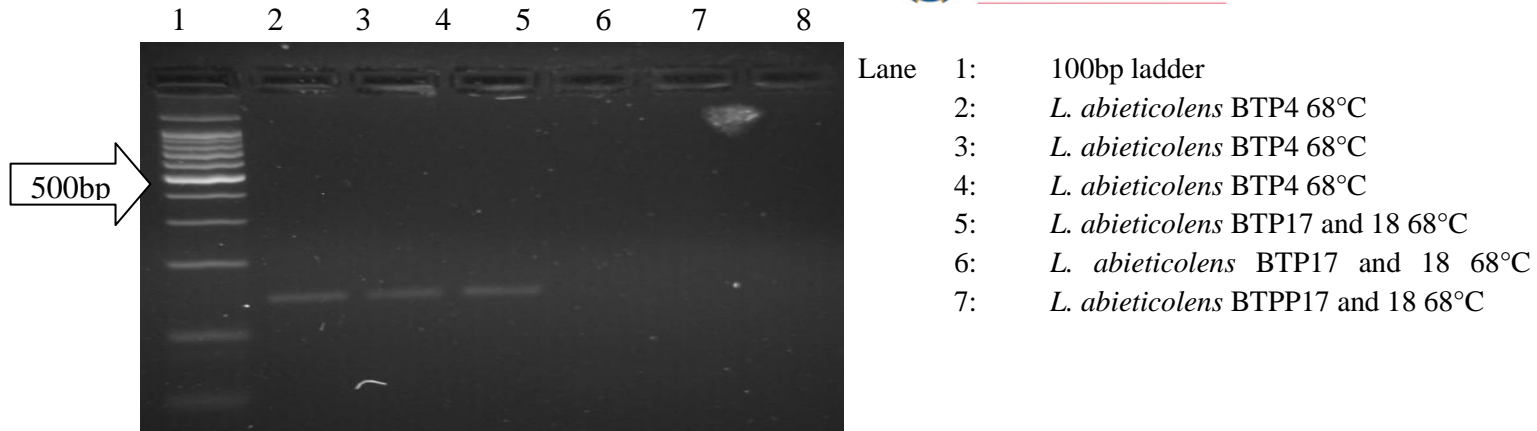7:   *L. abieticolens* BTPP17 and 18 68°C

Figure 7. The 3% Agarose gel resolved the product amplified with species-specific identiprimer BTP4 revealing a 150bp fragment in *L. abieticolens*. The product was separated for 40 minutes at 100V.

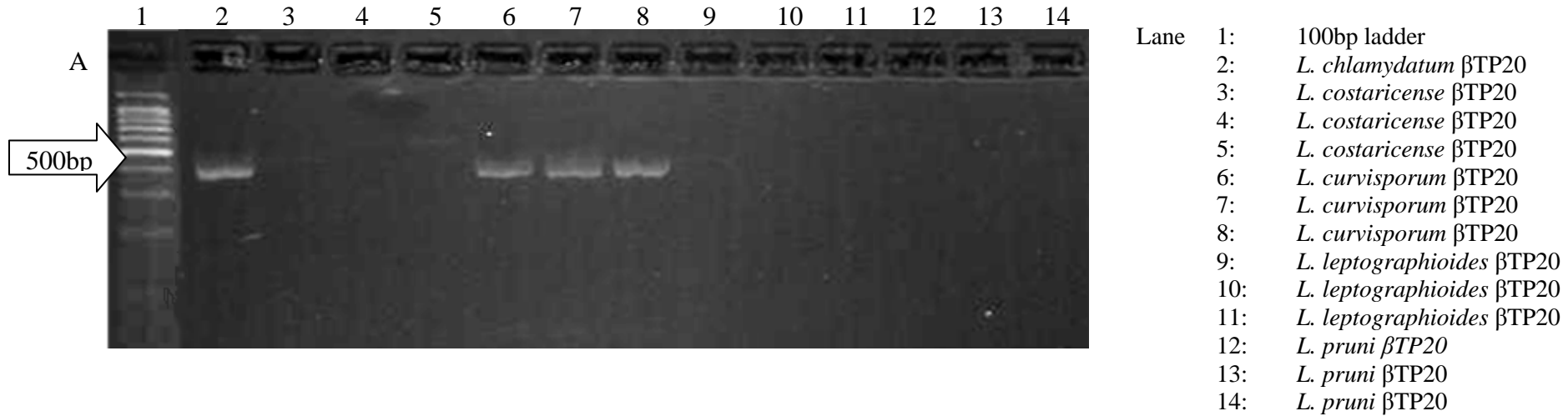**Figure 8.** A 3% Agarose gel resolved the amplicon produced by βTP20 from *L. chlamydatum, L. costaricense, L. curvisporum, L. leptographioides* and *L. pruni.* These species group together under a single clade within the larger clade 2 of the ITS phenogram. They all have the generic ITSP1 generic primer in common but are differentiated from each other using β-tubulin identiprimers.
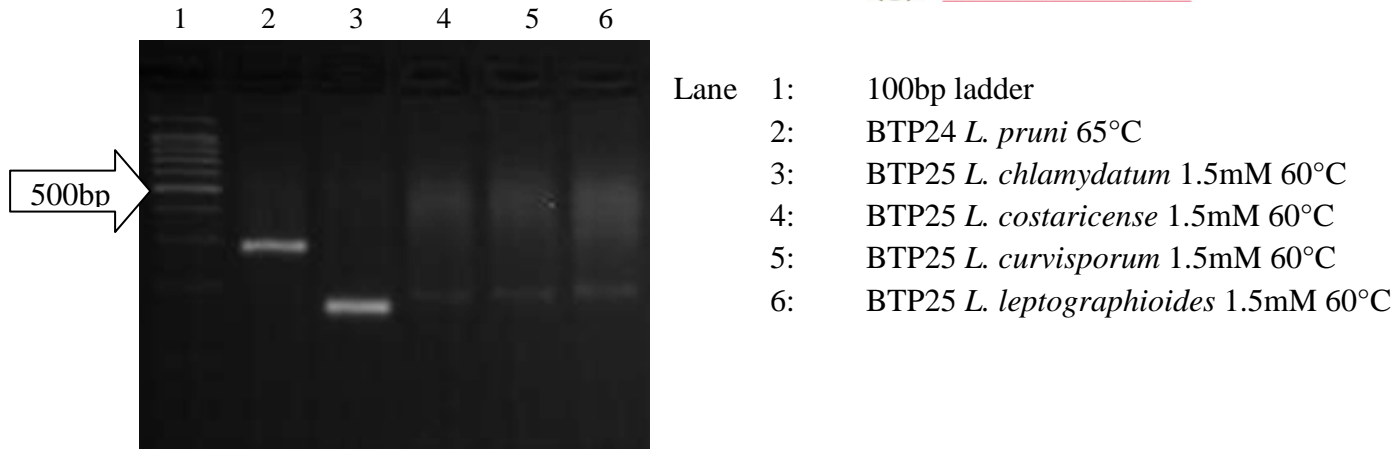
Figure 9. A 3% Agarose gel resolved amplicons from βTP24, that is unique to *L. pruni,* and βTP25, that is unique to *L. chlamydatum* at 100V for 40 Minutes. The other species in the group are delineated by identiprimers EF1α4lep for *L. leptographioides* and βTP21 for *L. costaricense.*

Figure 10. A 3% Agarose gel resolved PCR products from blind test and unknown species amplified using ITSP12 and LR3. The products were separated a 100V for 40 minutes. ITSP12 is a deep node primer for the ITS2 phenogram and is common to many *Leptographium* species.

500bp

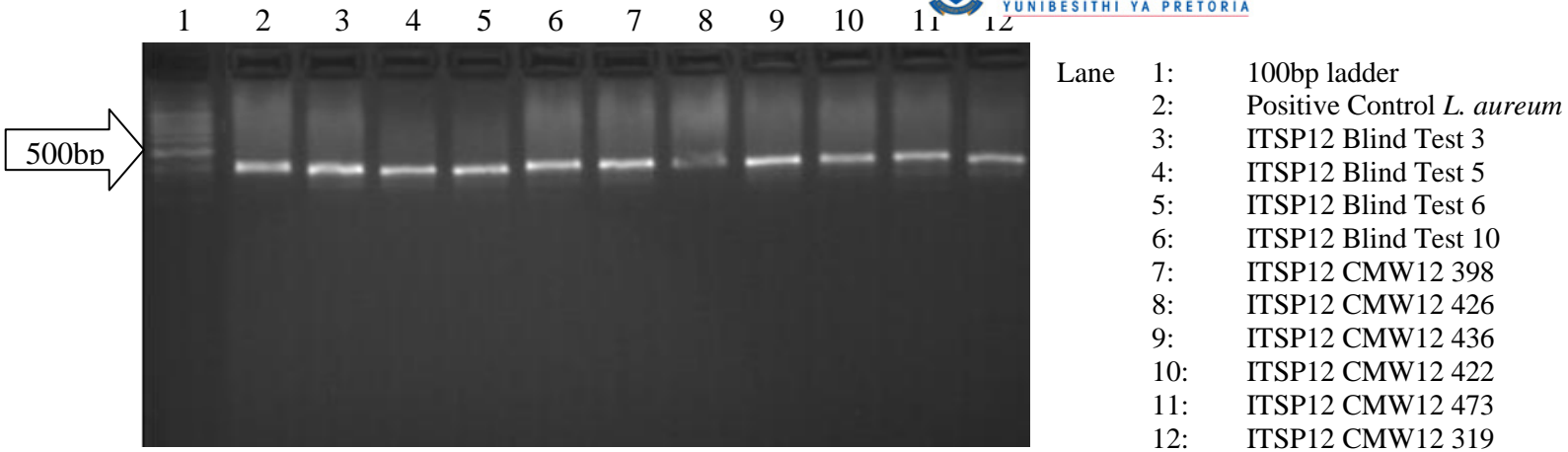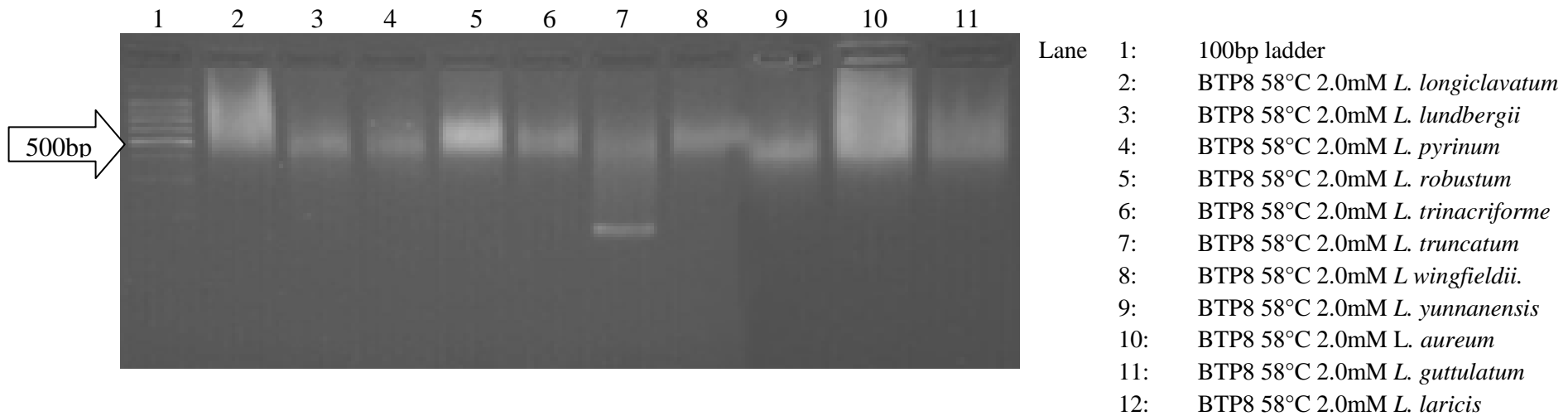| Lane	1:	100bp ladder
	2:	BTP8 58°C 2.0mM *L. longiclavatum*
	3:	BTP8 58°C 2.0mM *L. lundbergii*
	4:	BTP8 58°C 2.0mM *L. pyrinum*
	5:	BTP8 58°C 2.0mM *L. robustum*
	6:	BTP8 58°C 2.0mM *L. trinacriforme*
	7:	BTP8 58°C 2.0mM *L. truncatum*
	8:	BTP8 58°C 2.0mM *L wingfieldii.*
	9:	BTP8 58°C 2.0mM *L. yunnanensis*
	10:	BTP8 58°C 2.0mM L. *aureum*
	11:	BTP8 58°C 2.0mM *L. guttulatum*
	12:	BTP8 58°C 2.0mM *L. laricis*

Figure 11. A 3% gel resolved PCR products, at 100V for 40 minutes, from a species-specific identiprimer BTP8 that separates *L. truncatum* from the other species that are positively amplified by ITSP15 and ITSP12.

Lane 1: 100bp ladder
2: Negative control
3: *L. fruticetum* ITSP10 57°C
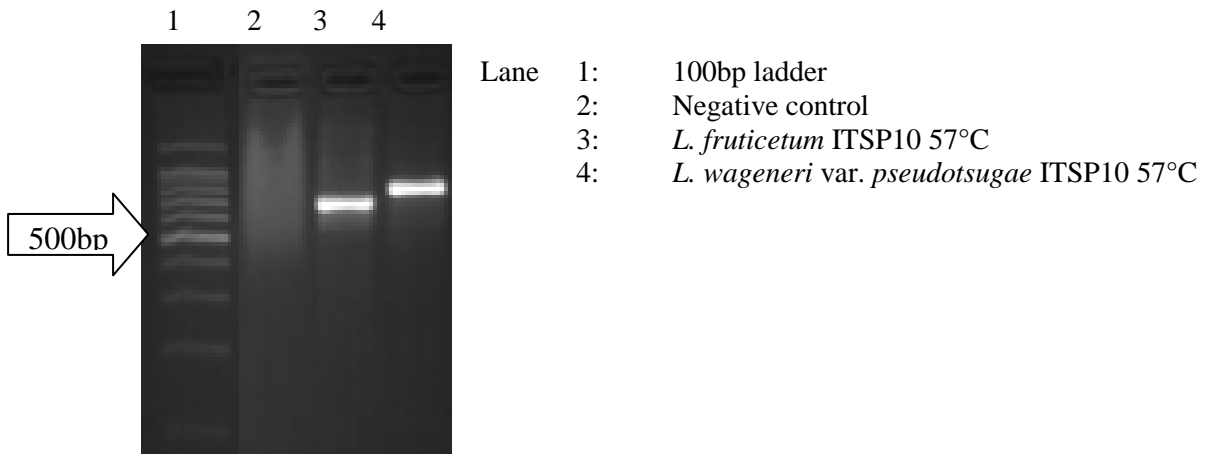4: *L. wageneri* var. *pseudotsugae* ITSP10 57°C

Figure 12. . A. A 3% Agarose gel that resolved products from ITSP10 and LR3. The gel was run at 100V for 40 minutes. *Leptographium fruticetum* groups with *L. wageneri* var. *pseudotsugae* in a 4th clade within the ITS2 phenogram. They share ITSP10 as a common node identiprimer.

Lane 1: 100bp ladder
2: *L. fruticetum* ITSP9 57°C
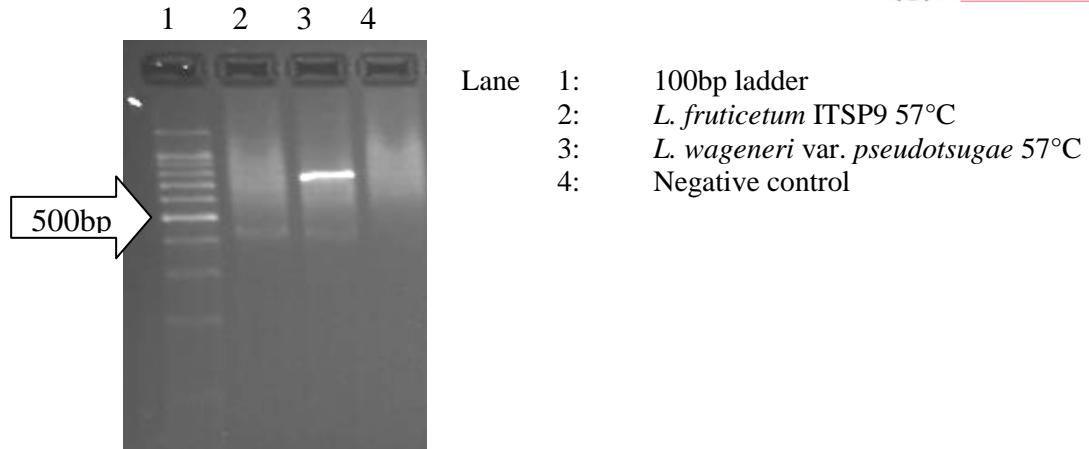3: *L. wageneri* var. *pseudotsugae* 57°C
4: Negative control

Figure 13. A 3% Agarose gel resolved PCR products amplified with ITSP9 at 100V for 40 minutes. The product was separated at 100V for 40 minutes and is unique to *L. wageneri* var. *pseudotsugae*.

Table 5. Fifty-six known species of *Leptographium* were used in this study and tested against a set of primers. The optimised conditions, amplicon size and the primer sequence are detailed below.

Clade/Node primer	Species	Primer	Sequence	Sense	T_A (°C)	Size (bp)	MgCl$_2$ (mM)	2-pyrrolidone
1		ITSP1	AATGCTGCTCAAAATGGGAGG	+	55°C	370		
		7	CAGACCGCAGACGCAAGT	+		700		
		8	CCAGCCTTTGTGAAGCTCC	+		400		
		9	CCCTAAAGACGGCAGACG	+		800		
	L. abieticolens	BTP4	CCGTCCTTGTGGATCTCG	+	68°C	130		
	L. peucophillum	BTP17	ATATGGCGGATTAGATACCACC	-	68°C	200		
		BTP18	CTAACAGATGTCACAGGCAG	+	68°C	250		
	L. alethinum	EF1αP13	AAGGTCCCACAAGGCAGA	-	65°C	200		1/1000
	L. euphyes	EF1αP12	TCGCCGCTAATACCCAATAC	+	65°C	250		1/1000
	L. neomexicanum	EF1αP2	AAACAGGGAATGAAGAATTGCC	-	58°C	900	2.0mM	
		EF1αP3	AAAGGCAGGGAATGAAGAATTG	-	58°C	900	2.0mM	

Clade/Node primer	Species	Primer	Sequence	Sense	T_A (°C)	Size (bp)	MgCl$_2$ (mM)	2-pyrrolidone
	L. reconditum	EF1αP3	AAAGGCAGGGAATGAAGAATTG	-	58°C	900		
	L. douglassi	ITSP9	CCCTAAAGACGGCAGACG	+	55°C	800		
	L. pineti	ITSP25	AAGGAAAGGAGACTTGCGT	-	54°C	380		
2		**ITSP1**	**AATGCTGCTCAAAATGGGAGG**	**+**	**55°C**	**390**		
	L. abietinum	ITSP3	ATTGGTTGCTGCAAGCGT	-	51°C	200		
		EF1αP42	AATGGAAAAGAGGGGCGAGG	-	68°C	900		
	L. americanum	ITSP3	ATTGGTTGCTGCAAGCGT	-	51°C	200		
		EF1αP37	AATGCAGGGTCCCACAGG	-	68°C	490		
	L. antibioticum	ITSP2	GGAGCTTCGCAAAGGCCA	-	55°C	450		
		EF1αP8	AAAGAGCCCTTGCCGAGC	-	68°C	550		
	L. brachiatum	ITSP2	GGAGCTTCGCAAAGGCCA	-	55°C	450		
		EF1αP27	AACAACCAATACAGGAGGCTG	+	68°C	200		

Clade/Node primer	Species	Primer	Sequence	Sense	T_A (°C)	Size (bp)	MgCl$_2$ (mM)	2-pyrrolidone
	L. rubrum	ITSP2	GGAGCTTCGCAAAGGCCA	-	60°C	900		
		EF1αP28	AAACGAGGATGATTTGGGCAA	-	68°C	200		1/10
		EF1αP38	AAACACACACGCCACAACC	+	65°C	400	2.0mM	
	L. bistatum	ITSP4	CAAAGCGAGGGCTAATGCT	-	62°C	150		
		BTP22	ACACGCCATTGCTGTCCA	-	58°C	150		
	L. eucalyptophilum	ITSP4	CAAAGCGAGGGCTAATGCT	-	62°C	150		
		BTP5	CGTCTTCGCCAGGTACAACG	+		250		
		BTP3	ACAGCATCCATCGTGCCG	-		150		
	L. calophylli	ITSP17	TGTAATTTGGAGAGGATGCTTT	+	55°C	200		
		EF1αP20	TGGGCAAGGGCTCTTTCAA	+	68°C	250		
		EF1αP21	AAACGGGCTTTATCTCAGGAC	-	68°C	250		
	L. dryocoetidis	ITSP17	TGTAATTTGGAGAGGATGCTTT	+	55°C	400		

128

Clade/Node primer	Species	Primer	Sequence	Sense	T_A (°C)	Size (bp)	MgCl$_2$ (mM)	2-pyrrolidone
		BTP24	CCTCGTTGAAGTAGACGCTC	-	68°C	100		1/10
	L. chlamydatum	BTP20	CCAGGCAGCAGATTTCCG	+	68°C	490		
		BTP25	CTGGAGATCAGAGTTGCCAT	-	60°C	150	1.5mM	
	L. costaricense	BTP21	CATGGATGCCGTCCGTGC	+	65°C		1.5mM	
	L. leptographioides	EF4lep	CGA[C]ATTGCTCTGTGGAAGTT	+	65°C		1.5mM	
	L. pruni	BTP24	CCTCGTTGAAGTAGACGCTC	-	67°C	400	1.5mM	X2803
	L. curvisporum	BTP20	CCAGGCAGCAGATTTCCG	+	68°C	490		
	L. brevicollis	BTP24	CCTCGTTGAAGTAGACGCTC	-	68°C	100		
	L. crassivaginatum	BTP3	ACAGCATCCATCGTGCCG	-		150		
		ITSP6	CACAAGGTTGACCTCGGAT	+	58°C	550	2.0	
	L. francke-grosmanniae	ITSP24	AACCTTTGAGATAGACTTGCG	-	58°C	400bp		

Clade/Node primer	Species	Primer	Sequence	Sense	T_A (°C)	Size (bp)	MgCl$_2$ (mM)	2-pyrrolidone
3		**ITSP1**	**AATGCTGCTCAAAATGGGAGG**		**55°C**	**370**		
		ITSP12	**CGGTTGGACGCCTAGCCTTT**	**+**				
	L. bhutannense	ITSP22	AAATGACCGGCAGACGCAA	+				
Node primer		**ITSP15**	**GAGCTTCACAAAGGCTAGGC**	**-**	**60°C**			
	L. aenigmaticum	BTP11	GACCGTGCTCGCTGGAGATC	-	55°C	159		
		BTP12	CAGACGTGCCGTTGTACC	-	60°C	250		
	L. albopini	BTP16	AATGGCGTGTAGGTTTCCG	+	68°C	300		
	L. clavigerum	EF1αP36	AAGCAGGTGGGGATGAGATG	-	55°C	260		
	L. koreanum	BTP8	CAACAAGTACGTGCCTCGC	+	68°C	150	1.5mM	
	L. aureum	BTP3	ACAGCATCCATCGTGCCG	-	64°C			
		BTP1	ACAGCAATGGAGTGTAGGT	+	not specific	200		

Clade/Node primer	Species	Primer	Sequence	Sense	T_A (°C)	Size (bp)	MgCl$_2$ (mM)	2-pyrrolidone
	L. guttulatum	BTP37	CGGAAGAGCTGCCCAAAG	-	68°C	380		
	L. laricis	EF1αP5	TTAAAACCTGACCGCCCAAAA	-				
	L. longiclavatum	EF1αP32	AGGCAGAAAGACAGGGAAGAGA	-	66°C not specific	250	2.0mM	
	L. lundbergii	BTP37	CGGAAGAGCTGCCCAAAG	-	68°C	480		
	L. pyrinum	BTP31	AAGAGCGTCTATTGTGGTGT	-		150		
	L. robustum	EF1αP39	AAAGACAGGGAGGAGGATTTG	-	60°C	150		
	L. trinacriforme	BTP30	AGAATTTGTCACTTCAAGCAGA	-				
	L. truncatum	BTP13	CACGGCATCCATCGTACC	-	57°C	450		
		BTP8	CAACAAGTACGTGCCTCGC	+	55°C	200		

131

Clade/Node primer	Species	Primer	Sequence	Sense	T_A (°C)	Size (bp)	$MgCl_2$ (mM)	2-pyrrolidone
	L. wingfieldii	EF1αP36	AAGCAGGTGGGGATGAGATG	-	64°C	280	1.5mM	
	L. yunnanensis	BTP13	CACGGCATCCATCGTACC	-	60°C	195		
		BTP15	AATGGCGTGTAGGTTTCCG	+	60°C			
	L. sibiricum	ITSP23	AAATGACCGGGAAGACGCA	+	65°C	550		
	L. piceaperdum	ITSP27	CCAAAATAAGGGCAGGGCG	-	65°C	700		
	L. huntii	ITSP11	CGAGTCTGTCTCCTTCTCAA	+	65°C	650		
		BTP13	CACGGCATCCATCGTACC	-	58°C	350	1.5mM	
	L. pityophilum	ITSP11	CGAGTCTGTCTCCTTCTCAA	+	65°C	650		
		EF1αP11	AAGACTTCTCCAACAGGTGG	-	58°C	700		
Node primer		**ITSP22**	**AAATGACCGGCAGACGCAA**	**+**	**45**°C	**790**	**2xDNA**	
	L. penicillatum	BTP18	CTAACAGATGTCACAGGCAG	+	69°C	250	2.0	
		BTP10	AGATTTCTAGCGAGCATGGC	+	69°C	900	1.5	

Clade/Node primer	Species	Primer	Sequence	Sense	T$_A$ (°C)	Size (bp)	MgCl$_2$ (mM)	2-pyrrolidone
	L. profanum	BTP18	CTAACAGATGTCACAGGCAG	+	69°C	300	2.0	
	L. pini-densiflorae	ITSP18	AAAGGAGGGACAGACTTGC	-	65°C	900		
	L. procerum	ITSP15	GAGCTTCACAAAGGCTAGGC	-	64°C	500		
	L. serpens	BTP4	CCGTCCTTGTGGATCTCG	+	60°C	250		
	L. terebrantis	BTP3	ACAGCATCCATCGTGCCG	-	63°C	350		
	L. wageneri var. wageneri	EF1αP22	AAAGGAAACACGGAGAGCATCG	+		600		
4		**ITSP10**	**CTCCGAGCGTAGTAAGCA**	+				
	L. fruticetum	ITSP10		+	55°C	600		
	L. wageneri var. pseudotsugae	ITSP10			55°C	500		
		ITSP9	CCCTAAAGACGGCAGACG	+	57°C	750		
5	*L. elegans*	ITSP1	AATGCTGCTCAAAATGGGAGG	+	55°C	370		

Clade/Node primer	Species	Primer	Sequence	Sense	T_A (°C)	Size (bp)	MgCl$_2$ (mM)	2-pyrrolidone
		EF5ele	CGGTGCCTATTCTCGTGGT	+	60°C	400		
Only ITS1								
	L. wageneri var. *ponderosa*	BTP28	AATCATGCACAGAGAGCTAACA	+	55°C	290	1.5mM	
		BTP29	ATCGCAGCTCGGGTAGATC	-		300		
		EF1αP22	AAAGGAAACACGGAGAGCATCG	+		600		
	L. grandifoliae	BTP33	AAACCTTCCGAGATGTCCAC	+		150		

The general primers are indicated as clade or node primers in bold and specific primers are included adjacent to the species that they identify.

Those highlighted indicate primer failures where primers need to be redesigned.

Table 6. Eight previously undescribed species of *Leptographium* were collected from various geographical regions and various hosts.

CMW No.	CountryID	HostID
12346	Seychelles	*Calophyllum*
12398	Tanzania	*Eucalyptus* spp
12326	Chile	*Pinus radiata*
12422	Chile	*Araucaria araucana*
12319	Chile	*Eucalyptus globulus*
12425	China	*Unknown*
12471	China	*Picea koraiensis*
12473	USA	*Pinus thunbergii*

Table 7. Eight previously undescribed isolates were tested against the primer set. The table includes the primers that showed positive amplification, for all the other primers the value is indicated as zero, no amplification.

Probes/Species	CMW12346 *Leptographium* sp 1	CMW12398 *Leptographium* sp 2	CMW12326 *Leptographium* sp 3	CMW12422 *Leptographium* sp 4	CMW12319 *Leptographium* sp 5	CMW12425 *Leptographium* sp 6	CMW12471 *Leptographium* sp 7	CMW12473 *Leptographium* sp 8
ITSP1	1	1	1	1	1	1	1	1
ITSP6	1	1	1	1	1	0	0	1
ITSP7	0	0	0	0	0	1	1	0
ITSP8	1	0	0	0	0	0	0	0
ITSP11	1	1	1	0	1	0	0	0
ITSP12	1	1	1	1	1	0	0	1
ITSP15	0	0	0	0	0	1	1	0
ITSP17	0	0	1	1	0	0	0	0
ITSP18	0	1	0	0	0	0	0	0
ITSP23	0	1	1	0	1	0	0	0
ITSP24	0	1	0	0	0	0	0	0
ITSP25	1	1	0	0	0	0	0	0
EF1αP2	1	1	0	0	0	0	0	0
EF1αP3	1	0	0	1	0	0	1	1
EF1αP12	1	0	0	0	0	0	0	0

Probes/Species	CMW12346	CMW12398	CMW12326	CMW12422	CMW12319	CMW12425	CMW12471	CMW12473
	Leptographium sp 1	*Leptographium* *s*p 2	*Leptographium* sp 3	*Leptographium* sp 4	*Leptographium* sp 5	*Leptographium* sp 6	*Leptographium* sp 7	*Leptographium* *sp 8*
EF1αP36	1	0	0	1	0	0	0	1
EF1αP38	0	0	0	0	0	1	1	0
EF1αP42	0	0	0	0	0	1	1	0
BTP4	0	1	1	0	0	0	0	0
BTP13	1	0	0	1	0	0	0	0
BTP15	1	0	0	1	0	0	0	0
BTP24	0	1	1	0	1	0	0	0
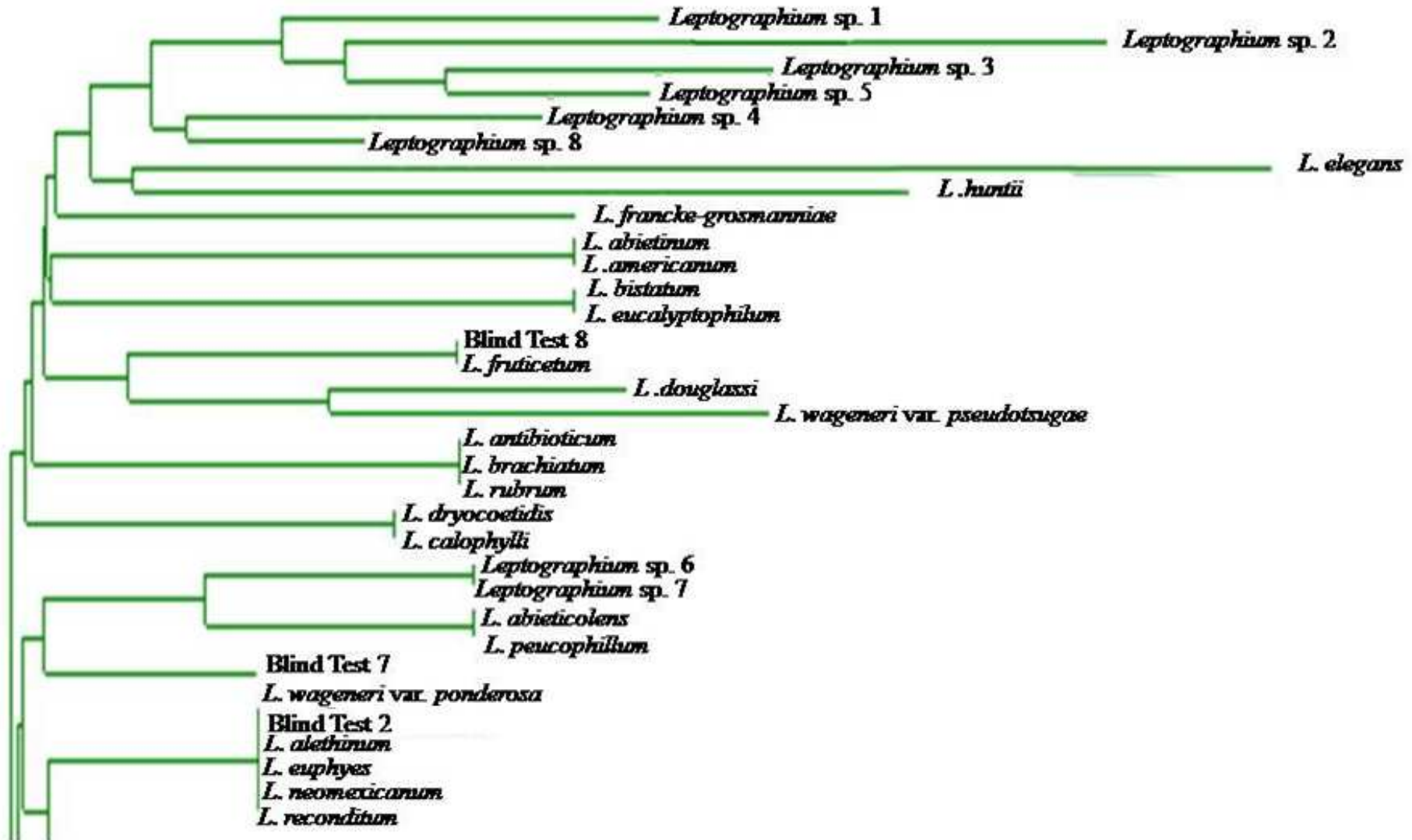
Figure 14. Probe phenograms were constructed for the blind test species and for the undescribed species of *Leptographium* from a 1 0 matrix representing the absence and presence of amplicons.
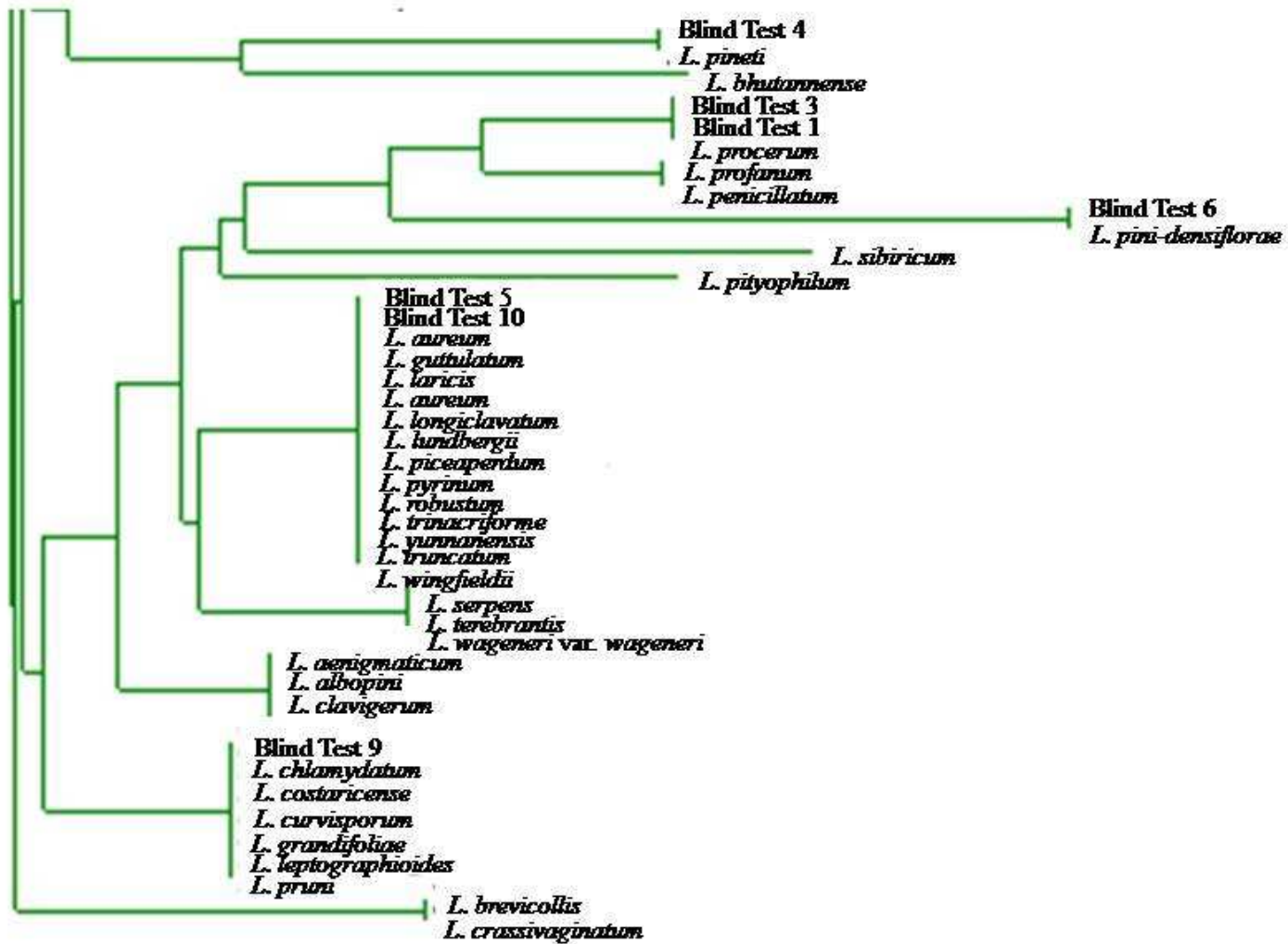
Figure 9. Continued
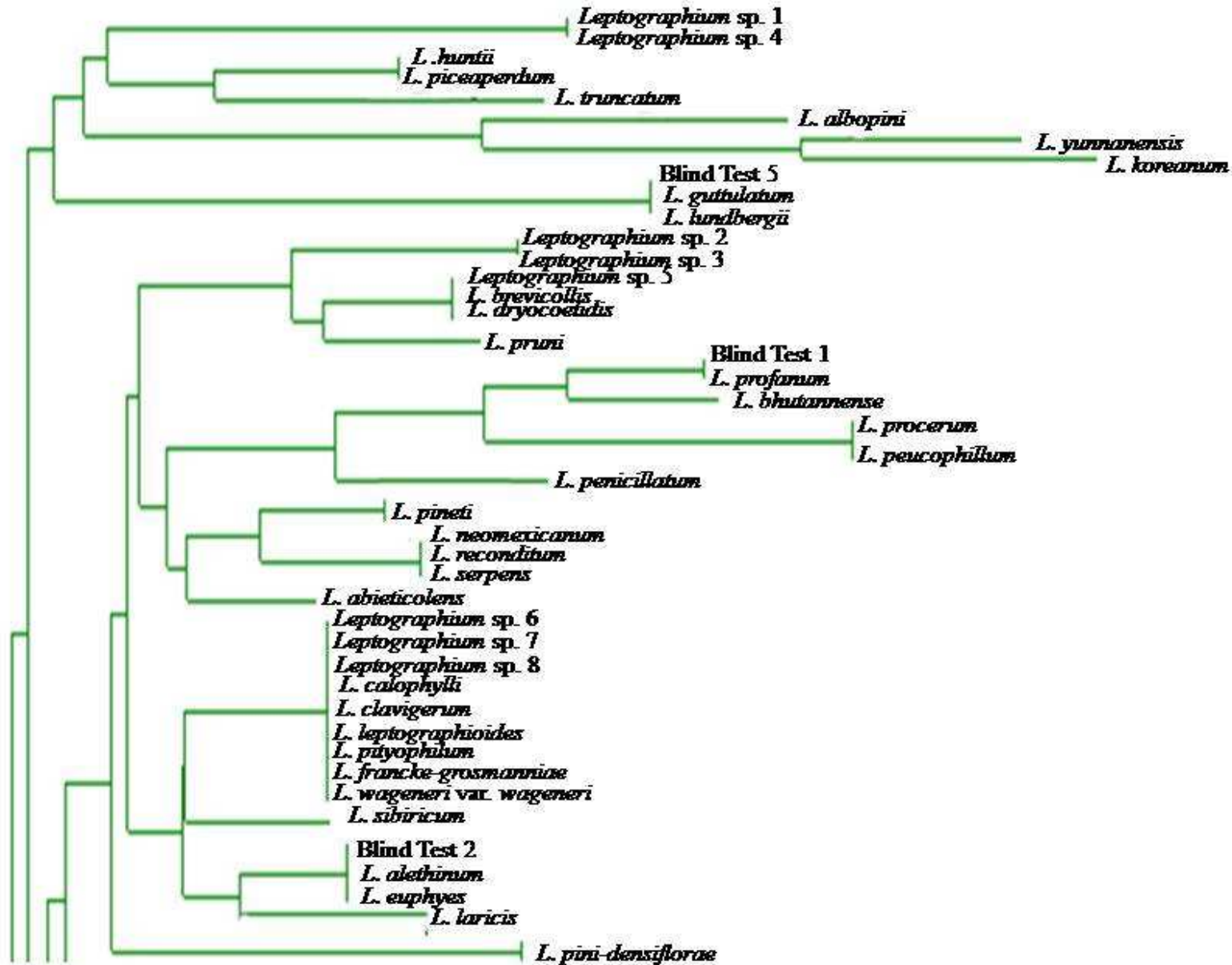
Figure 15. β-tubulin identiprimer phenogram constructed for the blind test species and undescribed species of *Leptographium from* a 1 0 matrix of present and absent amplicons.
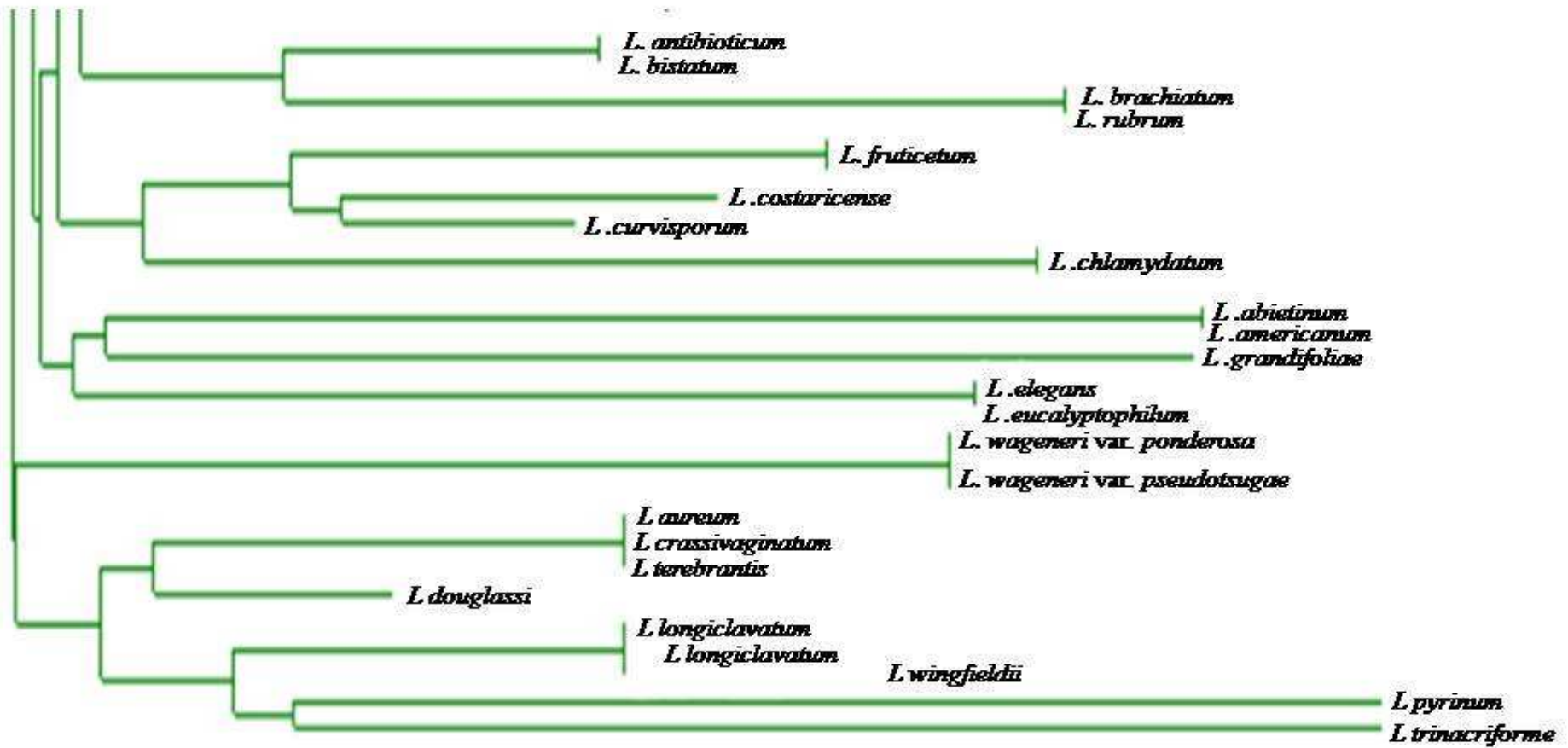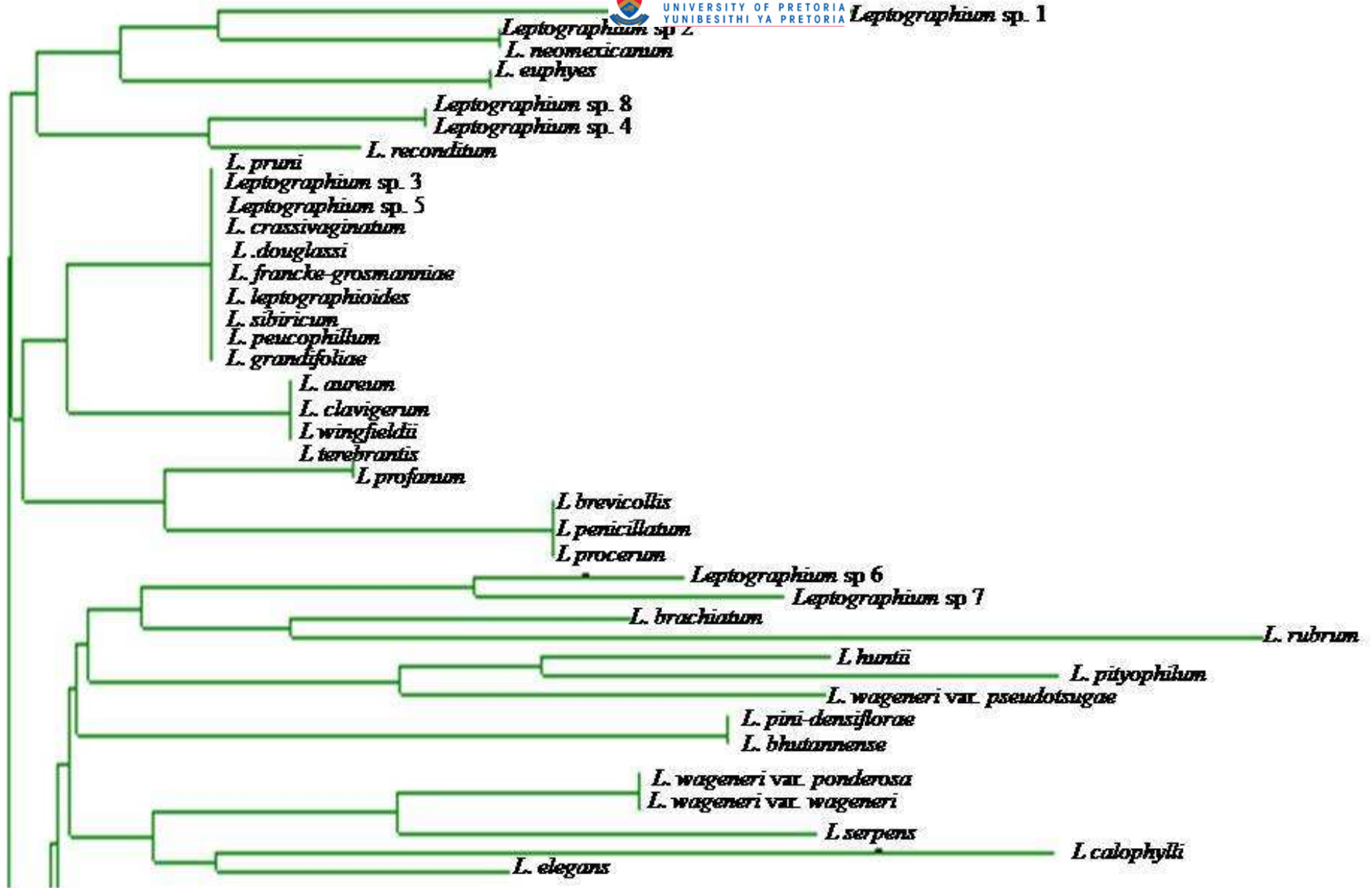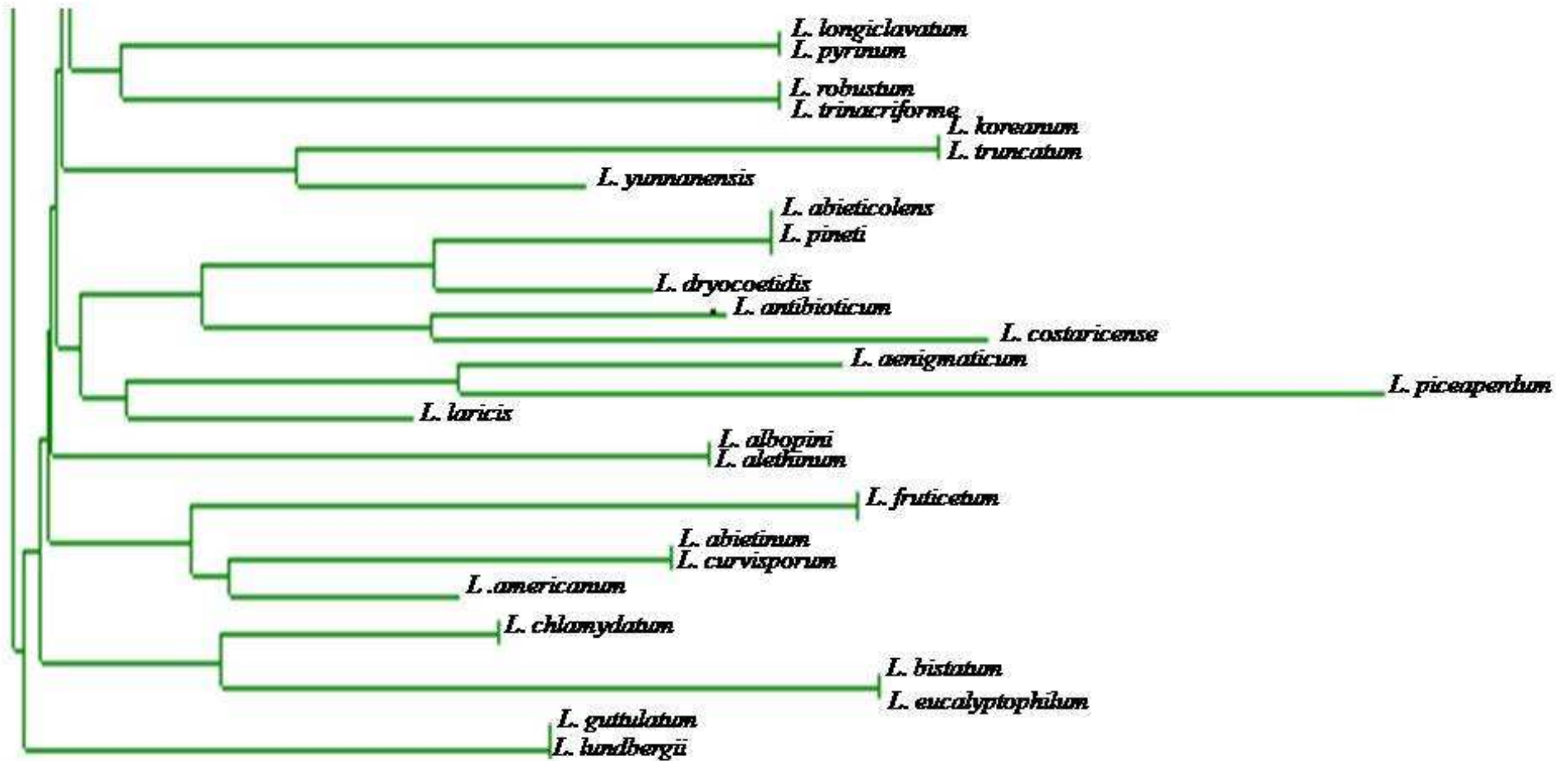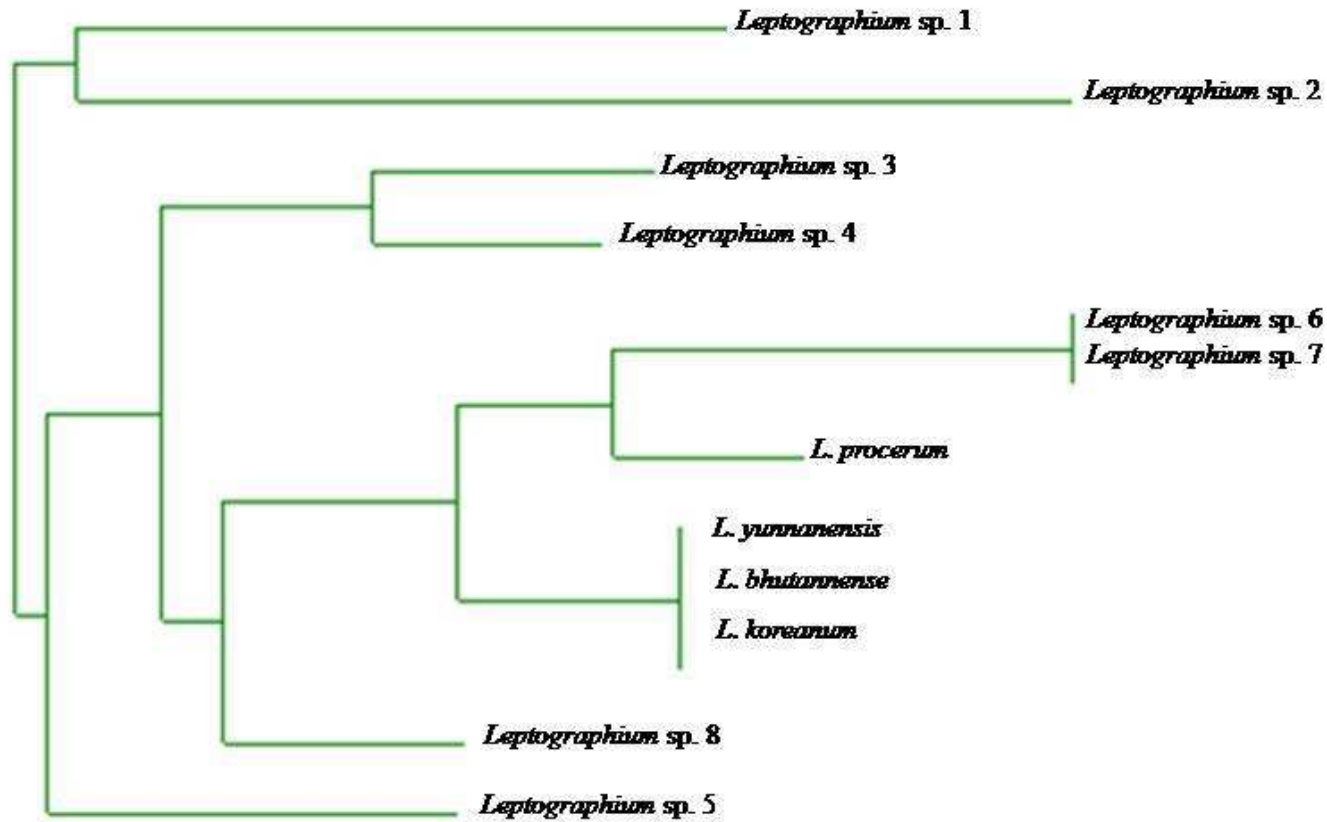
Figure 10. Continued

Figure 16. Elongation factor 1α identiprimer phenogram constructed for the blind test species and undescribed species of *Leptographium from* a 1 0 matrix of present and absent amplicons.

Figure 11. Continued

Figure 17. A phenogram constructed from a small matrix of ITS2 identiprimers for 8 undescribed species of *Leptographium* and related, described *Leptographium* species.
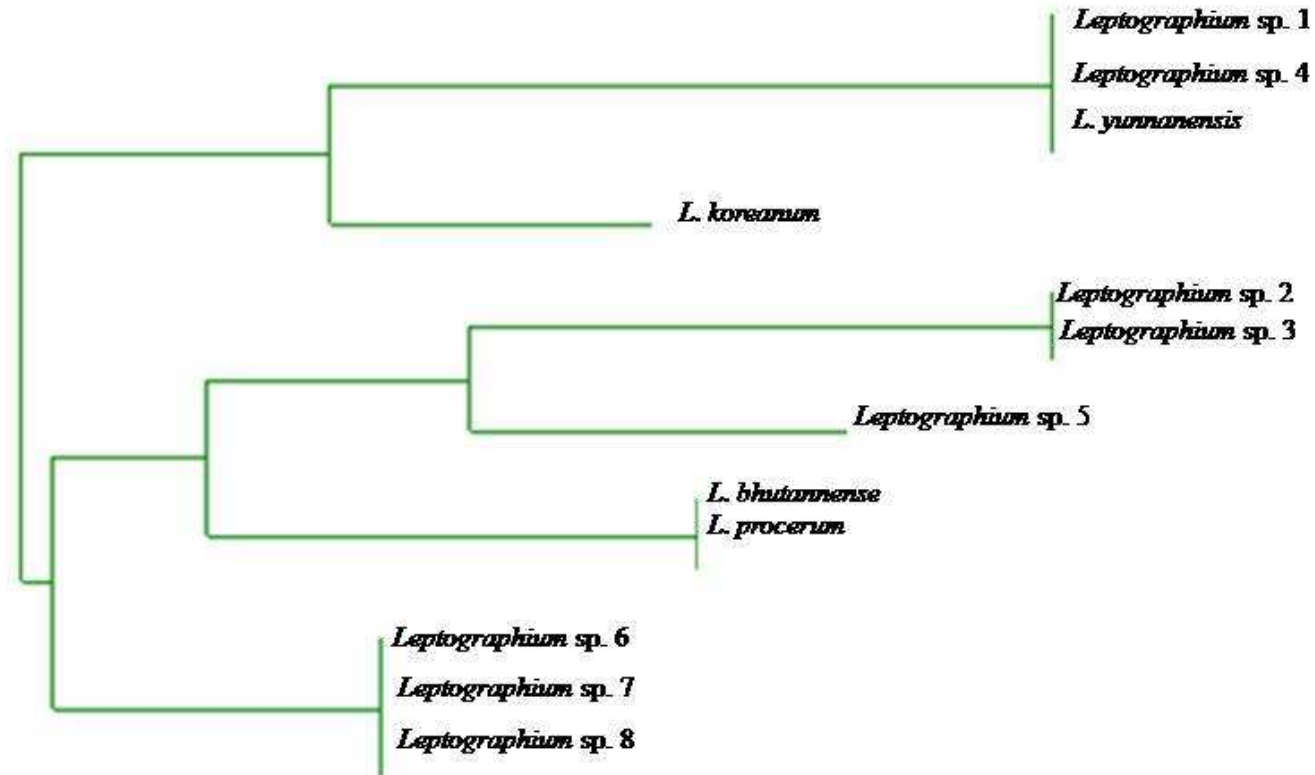
Figure 18. A phenogram constructed from a subset of β-tubulin identiprimers for 8 undescribed species of *Leptographium* and related, described *Leptographium* species.
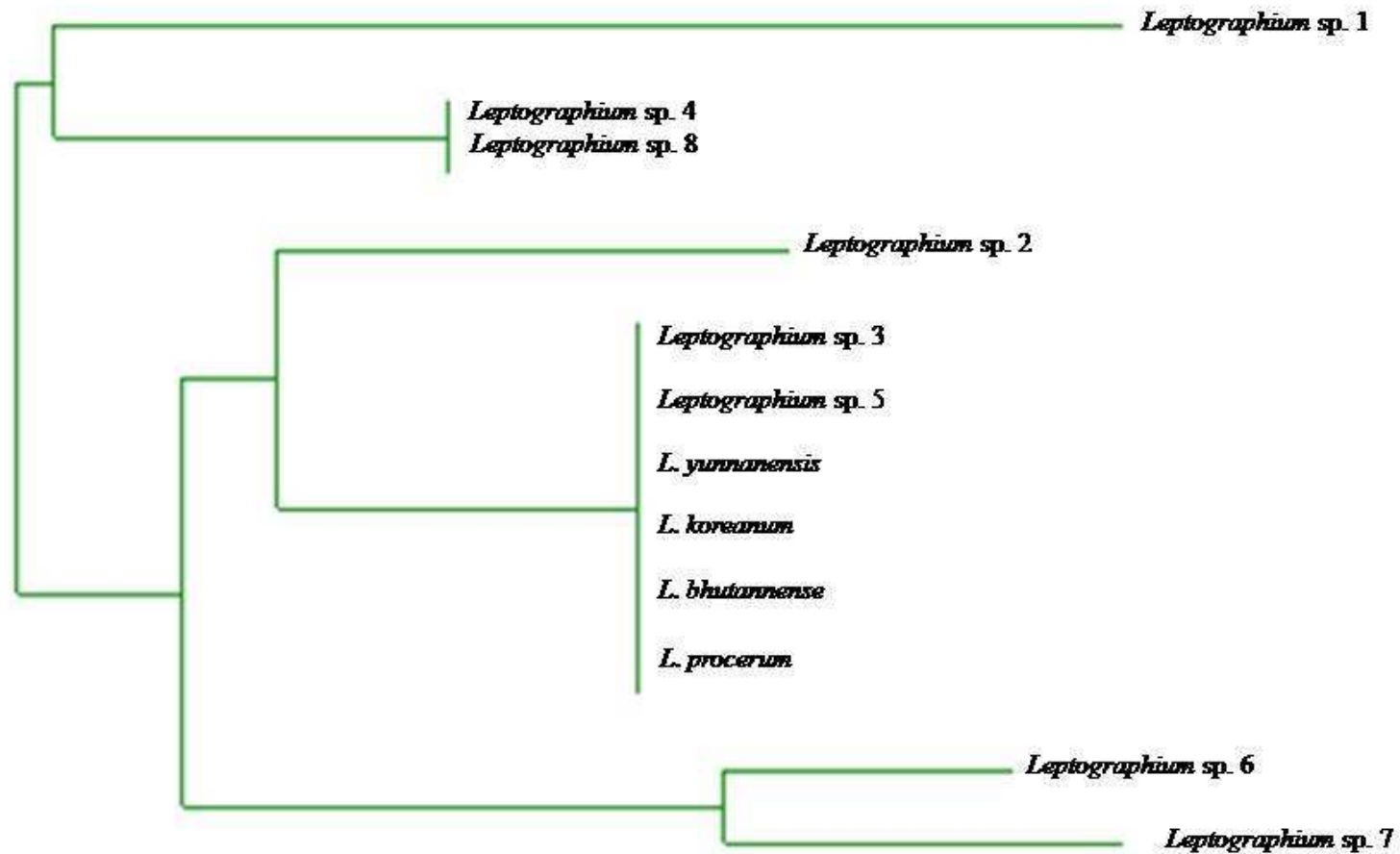
Figure 19. A phenogram constructed from a subset of translation elongation factor 1α identiprimers for 8 undescribed species of *Leptographium* and related, described *Leptographium* species

Summary

The genus *Leptographium* was chosen as a model to test the applications of two different diagnostic techniques. The genus is unique in the ascomycetes in that it is relatively small and well defined morphologically and phylogenetically. This presents a unique case for the ascomycetes, as there are sequence data available for 56 species of *Leptographium*, which constitute the majority of taxa in the genus. These characteristics make the genus an ideal model for testing the applications of microarray and diagnostic PCR techniques to the ascomycetes. A pilot study that was performed to test diagnostic microarray applications, included twenty-six species of *Leptographium*, and was based on a species-specific probe design. At least one and at the most three species-specific, 20-mer probes bound specifically amplified targets from one of the twenty-six species. The prototype array was used successfully to identify three test species of *Leptographium* i.e. *L. dryocoetidis, L. elegans* and *L. leptographioides.* The experiments revealed criteria that could be used to improve design parameters for a comprehensive diagnostic array.

Using the experience gained from the prototype array, a large array for fifty-six *Leptographium* species was designed using a minimal probe design combined with targets amplified using available universal primers. The array was designed in a similar way to a PhyloChip that uses probes to identify nodes and branches on a phylogeny. Probes were designed from three gene regions and were used to construct phenograms from a matrix of shared and unshared probe characteristics. The 89 probes designed from the large array were transferred to a PCR diagnostic by combining them with either a forward or a reverse universal primer.

The PCR diagnostic technique was optimised against the 56 known species of *Leptographium* and then tested against them in a blind test. The "identiprimers" were then applied to eight undescribed isolates of *Leptographium* in order to test the aptitude of the microcoding system to detect new species. The system was successful in identifying both known and previously undescribed isolates of *Leptographium.* The system is also capable of indicating the phylogenetic and biological relationships of the undescribed isolates.

This study supports the use of short oligonucleotides in microcoding applications and particularly the use of multiple gene regions to identify species to genus and species level. It also provides an identification system based on primer phenograms from multiple gene regions. The research in this thesis has shown that it is possible to transfer oligonucleotides between technologies and use them

in different ways. These technologies were both successful and could be applied to other genera of ascomycetes.