The molecular characterization of interaction between *Fusarium circinatum* and *Pinus patula*

Eduard Venter

Submitted in partial fulfilment of the requirements for the degree

Philosophiae Doctor

in the

Faculty of Natural and Agricultural Science,

Department of Genetics,

at the

University of Pretoria.

January 2004

Promotor: **Prof. A-M. Botha-Oberholster**

Co-Promotors: **Prof. B.D. Wingfield**

Prof. M.J. Wingfield

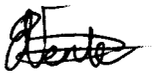
eo quod in multa sapientia multa sit
indignatio et qui addit scientiam addat et laborem

Dedicated to Riana,
And my parents

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

DECLARATION

I, the undersigned, hereby declare that the thesis submitted herewith for the degree *Philosophiae Doctor* to the University of Pretoria, contains my own independent work. This work has not been submitted to any other university faculty for any degree.



Eduard Venter
January 2004

The results presented in this thesis are original and were obtained from research in the Department of Genetics at the University of Pretoria, under the supervision of Prof. Anna-Maria Botha-Oberholster.

The following articles and presentations have been published/submitted for publication from the results of this study:

1. **E Venter,** B D Wingfeld, M J Wingfeld, A-M Botha 2003. Infection of *Pinus patula* with *Fusarium circinatum* does not induce detectable levels of chitinase. PMPP03/66 (submitted).

2. **E Venter**, B D Wingfeld, M J Wingfeld, A-M Botha 2003. Transcript derived fragments from a compatible interaction between *Pinus patula* and the pitch canker pathogen, *Fusarium circinatum*. (in preparation).

3. **E Venter**, B D Wingfeld, M J Wingfeld, A-M Botha 2003. Comparison of nucleotide binding site leucine rich repeat genes in *Pinus*. (submitted).

4. **E Venter**, A-M Botha, B D Wingfield, M J Wingfield 2000. Development of a differential display library from *Pinus patula* upon infection with *Fusarium circinatum*. Poster at the Forest Genetics for the Next Millennium, IUFRO Working Party 2.08.01. Durban, South Africa.

5. **E Venter**, A-M Botha, B D Wingfield, M J Wingfield 2000. Identification of a resistance response in *Pinus patula*. Poster at the Durable Disease Resistance Symposium, Ede-Wageningen, The Netherlands.

6. **E Venter**, A-M Botha, B D Wingfield, M J Wingfield 2001. Induced sequences linked to a resistance response in *Pinus patula* infected with *Fusarium circinatum*. Poster at Plant Biology 2001, Annual meeting of the American Society of Plant Biologists, Providence, Rhode Island, USA.

7. **E Venter**, B D Wingfield, M J Wingfield, A-M Botha 2003. Induction of gene sequences in *Pinus patula* after inoculation with the pitch canker pathogen, *Fusarium circinatum*. Poster at the 8th International Congress of Plant Pathology. Christchurch, New Zealand. (Selected for an oral presentation of poster)

8. **E Venter**, A Jacobs, B D Wingfield, M J Wingfield, A-M Botha 2002. What drives *Fusarium circinatum*? Analysis of pathogenicity at the molecular level. Presentation at the 40th Annual SASPP Congress. South Africa.

9. **E Venter**, B D Wingfield, M J Wingfield, A-M Botha 2002 *Fusarium circinatum* induced mRNA sequences in *Pinus patula*. Poster at the 40th Annual SASPP Congress. South Africa.

ACKNOWLEDGEMENTS

Almighty God, for all the grace and strength that He has granted me.

My parents and family, for their love, sacrifices and support throughout my life and my studies. They have made this journey possible.

Riana, for her love and support through the last eight years. She has been the rock on which I could stand and my voice of reason. Thank you for your input into this thesis.

Riana's parents for their love and support and for remembering me in their prayers.

Anna-Maria, for guidance and support throughout my scientific career.

Brenda and Mike, for input into this thesis.

The Microarray lab and Department of Biochemistry, for providing equipment without which I would not have been able to complete this study.

Albé, Danie, and Fourie for general assistance with various computer- and Linux related questions. Without your help this would have been tremendously difficult.

Marike Myburgh for technical assistance.

The students from the labs of FABI, especially Franco, for advice and discussions that went into this thesis.

Paul, Steve, and Ian of the Scottish Crop Research Institute, for making my stay there a pleasant and educational experience.

The Wednesday group for remembering me in their prayers.

The National Research Foundation, for a Ph.D. scholarship.

The Mellon Foundation, for a chance to broaden my horizons by funding my attendance to international conferences and my stay at the Scottish Crop Research Institute.

Komatiland Forests Research [previously Southern African Forestry Company Ltd. (SAFCOL)], for providing the *Pinus patula* trees; and the Tree Pathology Cooperative Programme, for the fungal cultures used in this study.

My friends, for listening to my various theories...far-fetched and otherwise.

TABLE OF CONTENTS

CHAPTER IV

CHAPTER V

LIST OF FIGURES

CHAPTER 1

CHAPTER 2

CHAPTER 3

CHAPTER 4

CHAPTER 5

LIST OF TABLES

CHAPTER 2

CHAPTER 3

CHAPTER 4

CHAPTER 5

ABBREVIATIONS

ACRE	*Avr9/Cf-9* rapidly elicitated genes
AFLP	Amplified fragment length polymorphism
AMV	Avian Myeloblastosis Virus
AOS	Active oxygen species
[α^{32}P] dCTP	The radioisotope of dCTP labelled at the Alpha 32 Phosphate
BAC	Bacterial artificial chromosome
BLAST	Basic local alignment search tool
bp	Basepairs
BCB	Blue copper protein
°C	Degree centigrade
CDD	Conserved Domain Database
cDNA	Complementary DNA (refers to DNA synthesised from mRNA)
CODEHOP	Consensus-degenerate hybrid oligonucleotide primers
COGEME	Phytopathogenic Fungi and Oomycete EST Database
CTAB	Hexadecyltrimethylammonium bromide
DD	Differential display
DEPC	Diethyl pyrocarbonate
DIG	Digoxygenin-11-dUTP
DMSO	Dimethyl sulfoxide
DNA	Deoxyribose nucleic acid
EMBOSS	European Molecular Biology Open Software Suite
dCTP	Deoxycytosine triphophate (one of the dNTPs)
dNTP	Deoxynucleoside triphosphate
DTT	Dithiotreitol
EDTA	Ethylene diaminotetraacetic acid
ELISA	Enzyme linked immunosorbent assay
EST	Expressed sequence tag
f.sp.	*Forma specialis*
g	Gram
HCl	Hydrochloric acid
HMM	Hidden Markov model
Indel	Insertion and deletion
kb	kilobasepairs (1kb = 10^3 basepairs)
λ	Lambda
l	Litre
LiCl	Lithium chloride
LRR	Leucine rich repeat
M	Molar
Mb	Megabases (1Mb = 10^6 basepairs)
MAST	Motif Alignment and Search Tool
MEME	Multiple Expectation of Maximisation for Motif Elicitation
$MgCl_2$	Magnesium chloride
min	Minutes
ml	Millilitre
mm	Millimetre
mM	Millimolar
MPSS	Massively parallel signature sequencing
mRNA	Messenger RNA
m/v	Mass/volume
n	Haploid number
NaCl	Sodium chloride
NBS	Nucleotide binding site
NCBI	National Centre for Biotechnology Information
NIL	Near isogenic line
ng	Nanogram
ORF	Open reading frame

PAGE	Polyacrylamide gel electrophoresis
PCR	Polymerase chain reaction
PEG	Polyethylene glycol
PES	Polyethersulfone
pmol	Pico mole
PVP	Polyvinyl pyrrolidone
QTL	Quantitative trait loci
RACE	Rapid amplification of cDNA ends
RAP-PCR	RNA arbitrarily primed polymerase chain reaction
RDA	Representational difference analysis
RFLP	Restriction fragment length polymorphism
RGA	Resistance gene analogue
RIL	Recombinant inbred line
RNA	Ribonucleic acid
ROS	Reactive oxygen species
Rpm	Revolutions per minute
rRNA	Ribosomal RNA
RT	Reverse transcriptase
RT-PCR	Reverse transcription PCR
s	Seconds
SAGE	Serial analysis of gene expression
SDS	Sodium Dodecyl Sulphate
SNP	Single nucleotide polymorphism
SSC	Sodium citrate buffer
SSPE	Sodium phosphate buffer
SSH	Suppression subtractive hybridisation
TAE	Tris-acetate-EDTA buffer
TDF	Transcript derived fragment
TIR	Toll/Interleuken-1 like receptor
T_m	Melting temperature
Tris	2-Amino-2-(hydroxymethyl)-1,3-propandiol
TE	Tris-EDTA
µ	Micro
Unit	i) Polymerase: One unit is defined as the amount of enzyme that will incorporate 10 nmoles of dNTP's into acid insoluble material in 30 minutes at 74°C under specific assay conditions. ii) Restriction Endonuclease: One unit is the enzyme activity that cleaves 1 µg λ DNA in 1 h at 37°C in SuRE/Cut buffer in a total volume of 25 µl. iii) T4 DNA Ligase: One unit of T4 DNA Ligase is the amount of enzyme activity that converts 1 nmol [^{32}P] from pyrophosphate into Norit-absorbable material in 20 minutes at 37°C.
µg	Microgram
µl	Microlitre
UTR	Untranslated region
UV	Ultraviolet
V	Volt
v/v	Volume per volume
w/v	weight per volume
x g	Relative centrifugal force

PREFACE

The pitch canker fungus, *Fusarium circinatum*, presents a major threat to exotic *Pinus* seedlings in South Africa. As *Pinus* spp. have a very long life cycle, it is not feasible to breed for resistance against the fungus. However, little knowledge exists regarding the molecular responses of *Pinus* trees to infection by the pitch canker fungus. Furthermore, the biology of *F. circinatum* is also not well characterised at the molecular level.

The major research objective of this thesis was to study the interaction of *P. patula* and *F. circinatum* at the molecular level. This was achieved through analyses of the induction of known resistance linked enzymes. A further attempt was made to identify genes linked to the defence response in *Pinus patula* after infection by the fungus has taken place. The next objective was to study the molecular biology of *F. circinatum* in comparison to other *Fusarium* spp. The study was to identify factors that might be linked to pathogenicity in *Pinus*. It was postulated that the nucleotide binding site leucine rich repeat (NBS-LRR) genes in plants are involved in defence responses. Thus, another research objective was to isolate and study the structure and relationships of *Pinus* NBS-LRR with previously reported NBS-LRR analogues. To accomplish these objectives a compilation of chapters that each deals with different approaches and techniques to answer specific questions relating to this interaction, are presented in this thesis.

Several reviews have been published that compare different techniques for gene discovery. These reviews include several of the newest techniques available and focus on high throughput. This leaves a void about information regarding less expensive technologies that can be applied to specific questions. Furthermore, only a few of these reviews focus specifically on PCR-based techniques and link them with the field of molecular plant pathology. A literature review that deals with this issue is, therefore, presented in **Chapter One**.

Several enzymes are linked with the broad defence response in plants. The first defence responses that are usually identified in many plant-pathogen interactions are the increase of chitinase and glucanase enzymes. These enzymes have been linked to the breakdown of

fungal cell walls. Therefore, the aim of **Chapter Two** was to isolate and study two chitinase homologues after inoculation with *F. circinatum*. This was achieved under three objectives. Firstly, enzyme activities of both chitinase and glucanase were assessed. Secondly, two chitinase homologues were studied with regard to their induction after infection by *F. circinatum*. Thirdly, partial fragments of the chitinase homologues were amplified from *P. patula* and compared to known chitinase genes from several plant species.

Several studies between the early responses of plants to pathogens have been reported. These studies most often focus on genes that are expressed during the hypersensitive response (HR) in plants. In the interaction between *P. patula* and *F. circinatum* there is no detectable HR, and only a general defence response is observed. This creates the need to study the response in conifers to pathogen interaction in a non-HR environment. The aim of **Chapter Three** was to further increase this knowledge on basic plant-pathogen interactions in *Pinus*, and also to characterise the specific response to the pitch canker fungus. This was accomplished by isolating several transcripts that are up-regulated after infection of *P. patula* with *F. circinatum*.

During serial sub-culturing of several fungal species, culture degeneration is observed. This is also evident in *Fusarium* species as an alteration in their morphology and a loss in pathogenicity. Several studies have been reported about the morphological aspects that are linked to cultural degeneration. No molecular study has been reported for this phenomenon. The *Fusarium subglutinans sensu lato* species complex accommodates several species previously separated on host preference. Several of these species are closely related plant pathogens. This creates the ideal opportunity to study differences under adverse culture conditions. Therefore, the aim of **Chapter Four** was the isolation of differentially expressed genes in *F. circinatum* compared to a degenerated version of this culture; and compared to *F. subglutinans*, *F. guttiforme*, and *F. opheodes prov. nom.*

The most diverse range of *R* genes isolated to date, belongs to the nucleotide binding site leucine rich repeat (NBS-LRR) group. This group of genes are considered to be a good source for the identification of candidate resistance genes in plants. This may also be true for conifers, as several sequences from this group have been detected from conifer ESTs. Therefore, **Chapter Five** reports the isolation of homologous sequences from a range of *Pinus* species and comparison of these sequences against reported conifer NBS homologues.

This thesis is a global study into the host-pathogen interaction between *Pinus patula* and *Fusarium circinatum*. As this study was approached from both sides of the interaction, it necessarily means that chapters with different focuses and angles have been grouped together. Therefore, each of the chapters represents an independent unit and will be submitted for publication to various journals in their respective fields.

CHAPTER 1

PCR-based methods used to identify differentially expressed genes in host-pathogen interactions

Introduction

Assigning function to genes and identifying which genes are playing a role under different physiological conditions are of great importance. This is especially true in plant-pathogen interactions where genes linked with pathogenicity are more readily detectable at certain stages. Most of the plant species' genomes will not be sequenced until sequencing becomes less expensive. In comparison to *Arabidopsis* (125 Mbp [11; 73]), the large sizes and complexities of genomes from economically viable crops make them extremely difficult to sequence, for example, *Oryza* = 415 Mbp, *Populus* = 550 Mbp, *Eucalyptus* = 600 Mbp, *Zea* = 2700 Mbp, *Hordeum* = 5000 Mbp, *Triticum* = 16000 Mbp, *Pinus* genomes >28000 Mbp [80; 27; 25; 33; 58; 49; 81]. Therefore, the only option to identify genes of interest is to study the genes that are expressed in certain interactions or stress conditions. These identified genes can then be compared to those in plants that are being characterised at the genomic level, for example against *Arabidopsis*, *Oryza*, and *Populus*. Identification of such genes will enhance our understanding of defence and stress responses in non-model interactions.

The choice of model plant to study genomics is based on several criteria. The plant should have a short life cycle, produce large numbers of offspring, have a small nuclear genome, and it must be easily transformed and propagated clonally [11; 69; 66]. The plant that fits these criteria the best is *Arabidopsis*, hence it has been chosen as the model dicotyledonous plant [73]. However, extrapolations from this plant have indicated that several model systems are needed for different classes of plants. Therefore, *Oryza* has been chosen as the model monocotyledonous plant [31; 86]. Although both of these plants provide excellent models for the study of herbaceous plant anatomy and gene expression, they are not well suited for further analysis of woody plants. For this reason, *Populus* has been selected as the model for trees [11; 69]. The choice of *Populus* was based on a sufficiently small genome size, and it can be easily clonally propagated [11; 49; 69]. However, it has longer generation spans than *Arabidopsis* and *Oryza*, and it can take several years to flower [11; 69]. In contrast with other tree species, like *Pinus*, these are not severe limitations, and *Populus* should be an outstanding model for the study of trees.

Several comparative studies between the annotated *Arabidopsis* and *Oryza* genomes have been reported [31; 38; 63; 66; 81]. Data from several of these studies indicate that large numbers of genes in *Arabidopsis* and *Oryza* differ. When comparing predicted genes longer

than 300 bp (H_{300}) from *Oryza*, it was found that no homology in *Arabidopsis* exists for 3886 rice genes [31; 81]. These genes represent 6 % of the predicted number of genes in the rice genome. The assumption is that the differences in gene content are linked to differences between monocotyledonous and dicotyledonous plants [81].

In the absence of genome sequences, genes in non-model organisms are usually discovered through mapping techniques. This is complicated by the reported inconsistencies in macro- and microcolinearity mapping [27; 28; 5; 34; 81]. Completion of sequencing of the two rice genomes provides an excellent opportunity to study microcolinearity and macrocolinearity (synteny) within a species. Han and Xue [34], showed that there is a high level of microcolinearity in rice, but that high numbers of rearrangements and polymorphisms exist between the two sequenced genomes. Although few contiguous sequence data stretches exist in other grass genomes, these data can be extrapolated to mapping between rice and grass genomes [5]. Within the cereals conservation is seen at the syntenic level, but it is less conserved at the microcolinearity level [81]. This makes it extremely difficult to clone genes through a map-based approach based on maps from related species [5].

Comparisons between rice and *Arabidopsis* genomes indicate that there is little conservation in colinearity. Studies indicate that there is little or no synteny between these two genomes [26; 66]. However, small numbers of genes grouped in blocks are found that display conservation between the two genomes [63]. This indicates that there is less synteny between the two genomes of these two model plants than previously predicted. Most of the synteny studies that have been undertaken are also linked with flowering in angiosperms [38]. These data can be used in predictions when comparing plants, but care should be taken and preferably colinearity of genes within macrocolinear stretches should be ascertained before any similarity is assumed [5].

Several comparisons relating to wood formation in *Arabidopsis* and tree species have been reported [1; 68; 88; 17; 39]. In these studies, the effective number of unidentified genes ranged from 10 – 55 %. This indicates that there is a large amount of gene conservation between species of seed plants, i.e. between *Pinus* and *Arabidopsis*. This indicates that *Arabidopsis* could be used for comparative genomics in gymnosperms [39]. However, the genes that could not be assigned any homology will most probably not be identified unless specific transcriptome changes are studied in gymnosperms [6]. Furthermore, no study

comparing resistance gene expression or host-pathogen interaction between a gymnosperm and *Arabidopsis* has been attempted.

Techniques to Study Gene Expression

Several different techniques exist to study and identify genes of interest in non-model plants. These techniques includes the classical approaches such as, subtractive cloning [35]; or more recently developed techniques such as, microarray analysis [64; 47], expressed sequence tag (EST) sequencing, massively parallel signature sequencing (MPSS) [12], and serial analysis of gene expression (SAGE) [78]), as well as different forms of subtractive hybridisation [36; 87; 21; 42]. With the advent of PCR, the sensitivity and throughput of gene discovery increased. The first of these techniques that made use of PCR were differential display (DD) and RNA arbitrarily primed PCR (RAP-PCR) [43; 82]. They were followed by cDNA amplified fragment length polymorphism (cDNA-AFLP) [4; 53] and subtraction hybridisations that incorporated a PCR step [36; 21; 42]. The most recent additions in techniques for gene discovery are the high throughput systems like microarray, SAGE, and MPSS.

Using most high throughput systems for gene discovery is expensive [23]. This hampers routine utilisation in laboratories that are studying specific interactions not linked to model systems. As a high throughput technique, microarrays provide a very useful tool to study gene expression. For example, it has been useful in studying and comparing changes taking place in the plant transcriptome under different types of stress [65; 19]. The main disadvantage is the cost involved to set up and make use of a microarray laboratory [60]. Similar situations exist for SAGE and MPSS where in excess of 100 000 ESTs can be identified, albeit at high cost [23]. The use of these two high throughput sequencing methods is hampered by the need for comprehensive databases against which to screen the identified tags [40; 23]. Therefore, they are not readily available for laboratories that study non-model systems and that have relatively small budgets.

Subtractive hybridisation is an uncomplicated and relatively inexpensive alternative to some of the higher throughput techniques. However, subtraction hybridisation lacks the sensitivity to identify genes that are expressed in low copy numbers, several micrograms of starting RNA

is needed for a successful subtraction, and it is more time consuming [45; 23]. This necessitates the use of techniques that can detect slight changes in gene expression during events such as pathogen infection.

Methods that incorporate a PCR step represent very sensitive approaches for gene expression studies. They usually make use of smaller quantities of RNA and are less time consuming [4; 40; 23]. These methods are also straightforward to implement in most laboratories that can perform PCR reactions on a regular basis. Furthermore, the power of enrichment through subtractive hybridisations and PCR can be combined to enhance the sensitivity of the method used [74]. Several caveats are associated with any PCR associated technique. The most important of these, is the possibility of amplification of false positive fragments [87]. This leads to issues surrounding reproducibility [71].

The most common PCR-based methods are cDNA-AFLP [4; 53], differential display (DD) and RNA arbitrarily primed PCR (RAP-PCR) [43; 82], suppression subtractive hybridisation (SSH) [21], and cDNA representational difference analysis (cDNA-RDA) [36] (Fig. 1). These four methods can be classed into two groups, those that incorporate a subtraction hybridisation to lessen the complexity of the sample (cDNA-RDA and SSH), and those that lessen the complexity directly through the PCR step (cDNA-AFLP, DD, and RAP-PCR). This review will focus on these four methods and how they have been successfully used to study interactions between plants and pathogens.

PCR-Based Transcriptome Analysis

Differential Display of RNA (DD and RAP-PCR)

This technique was first reported by two research groups [43; 82]. The two techniques are quite similar and are both based on previous reports of random amplification of polymorphic DNA [83; 84]. The technique can be divided into DD [43] or RAP-PCR [82] (Fig. 1). First RNA is extracted from the samples that are going to be displayed against each other. It is not necessary to separate the poly adenylated RNA from the total RNA. However, purification of mRNA from the total RNA does enhance the performance of DD over RAP-PCR [62; 48]. Then first strand synthesis is done with either an anchored poly(dT) primer (DD), or with random primers (RAP-PCR). The single stranded cDNA is then primed in a PCR reaction

with the addition of poly(dT) primers and 10-12 mer random oligonucleotides, as found in DD. In RAP-PCR all the PCR products are formed from priming with 18-20 mer random oligonucleotides and there is no addition of poly(dT) oligonucleotides. Once the PCR has been completed, the samples are compared with each other through denaturing polyacrylamide gel electrophoreses (PAGE). These displays can be visualised using silver staining, radioactive or fluorescent labelling. The fluorescent and radioactive labelling can be through incorporation of a labelled dNTP, or through end labelling of one of the primers.

As the first methods to be developed after subtraction hybridisation, use of DD and RAP-PCR has several advantages. The most important of these, is that they allow simultaneous comparison of several different samples on one gel and they require small quantities of RNA for each display [9; 71]. Recent advances in reverse transcriptase enzymes have made it possible to use lower quantities of RNA for DD [10]. Furthermore, the up- or down regulation of transcripts can be visualised in one display reaction [9]. This can be seen in bands present in either the sample being studied, or in the sample against which it is displayed. The simplicity and high speed of these two techniques makes them favourites in studying multiple samples and they have surpassed the use of subtraction hybridisation for gene discovery [71].

All PCR reactions primed through arbitrary primers have inherent problems. The first of these is that a high number of primers is required to screen a transcriptome [71; 52]. The success of detecting novel transcripts expressed in low copy numbers is regarded as a function of arbitrary primer annealing. This signifies that the amplification of an individual transcript is the function of the initial concentration of the transcript, as well as the specificity with which the arbitrary primer can bind the transcript [51]. Therefore, it is postulated that the amplification of a novel transcript can be outperformed by amplification of cDNAs that are in excess, although they do not have such specificity to the primer [4]. This is manifested in both DD and RAP-PCR as a reproducibility problem [71; 52].

Both DD and RAP-PCR have the disadvantage of producing several non-specific fragments (false positives) [9]. These fragments can be due to co-migration of fragments in the gel or due to incomplete amplification of fragments during the elongation step of the PCR [29]. Furthermore, as DD makes use of anchored poly(dT) oligonucleotides, the display is biased to identify transcript derived fragments (TDF) from the 3`-untranslated region of the genes that

are expressed [4; 8; 44]. This makes it laborious to identify the fragments of interest when not working in a model system like *Arabidopsis* or *Oryza* [48; 44].

Amplified Fragment Length Polymorphism of cDNA (cDNA-AFLP)

As with DD, cDNA-AFLP is a direct extension of a previously described technique. Here, AFLP has been applied to study the differences in gene expression levels [4] (Fig. 1). First RNA is extracted from the samples; the mRNA is purified from this total RNA and converted into double stranded cDNA. The cDNA is then double digested with a frequent and rare cutter restriction enzyme that leaves overhangs on the cDNA ends. The enzymes are chosen in such a way that they provide the optimum coverage of the transcriptome [4]. Thereafter, adaptors, specific to the overhangs produced by the restriction enzymes, are ligated onto these fragments. These modified fragments serve as the template for a primary PCR, known as pre-amplification, with primers that are specific to the adaptor sequences. The pre-amplification serves to increase the amount of template that is going to be used in the subsequent amplifications. The addition of extra bases onto the ends of the pre-amplification primers makes them more selective and hence the cDNA pools are simplified. This lowers the number of bands in successive amplifications and makes the detection of polymorphic bands simpler. Using selective primers in the pre-amplification will depend on the complexity of the transcriptome studied. Following pre-amplification, the reactions are used in secondary amplification reactions. These reactions are all primed with primers specific to the adaptor sequences. Here, selective bases are added to the primers to reduce the complexity of the profile. These profiles are compared to each other on PAGE. The visualisation of the fragments is through silver staining, radioactive, or fluorescent labelling. The fluorescent and radioactive labelling can be through incorporation of a labelled dNTP, or through end labelling of one of the primers.

The high throughput of cDNA-AFLP makes it an excellent technique to study several samples concurrently [4]. These samples can be from different individuals or a few individuals with several different treatments. The use of selective primers in the pre-amplification is unnecessary in some cases (i.e. fungi and bacteria) as the transcriptomes are reduced when compared to those of higher eukaryotes [2; 20]. This negates the reduction of the sample complexity.

With cDNA-AFLP, all the primer combinations will amplify only DNA that derives from expressed genes. This makes any marker identified from cDNA-AFLP very useful in mapping, because their map position should also be the location of the gene for which it is a marker [15]. This can have great advantages when combining the identified markers with a map based cloning approach or a genomic map [15].

cDNA-AFLP is more labour intensive than differential display and requires greater technical experience. This and the use of a high number of primers to screen for differential expression are the greatest disadvantages of cDNA-AFLP. The choice of restriction enzymes is also imperative, as different combinations will screen different regions of the transcriptome. Although cDNA-AFLP does not require prior sequence knowledge, some knowledge helps in deciding which enzymes to use. Some programs are also available for designing experiments. For example GenEST [59] can project the coverage obtained with different enzymes based on EST data.

Subtraction Hybridisation Techniques Combined with PCR

Representational Difference Analysis of cDNA (cDNA-RDA)

Representational difference analysis (RDA) is a technique devised by Lisitsyn *et al.* [45; 46] to detect low copy DNA, for example cancer linked genes or viruses, within complex genomes. In 1994, Hubank and Schatz [36] adapted RDA for use with cDNA (Fig. 1). When performing cDNA-RDA, both the target population (tester) and the non-target population (driver) cDNA are digested by a four base recognition sequence restriction enzyme. *Dpn*II is one of the most commonly used restriction enzymes for RDA, but any enzyme can be used. Thereafter, a set of adaptors, specific to the overhangs produced by the restriction digest, is ligated onto the ends of these fragments. Primers specific to these adaptors are used in a PCR reaction to generate amplicons of a set size. During this PCR cycle there is a size selection for smaller fragment sizes that are termed the 'representation' [45; 36]. After completion of the PCR the adaptors are removed by enzyme digestion and a new set of adaptors is ligated onto the ends of only cDNA from the tester pool. Then the first subtraction is performed on the tester amplicons with the driver cDNA in excess. Primers specific to the new adaptor sequences are used in a PCR cycle to enrich for the novel fragments. These fragments are called the first difference product and separated on an agarose gel. They can be excised from

the gel and cloned or for added sensitivity they can be reintroduced and the RDA cycle repeated.

cDNA-RDA is a sensitive method to identify nominal differences in gene expression between two sample sets [36]. This technique is several times more sensitive than previous subtraction hybridisation methods. In the same time period required for classic subtraction hybridisation methods, cDNA-RDA can isolate several gene transcripts that are expressed at extremely low levels. Furthermore, the identification of false positives is extremely low when compared with more classical techniques [36]. This level of false positives can also be regulated by the number of repetitions of RDA [18].

Using cDNA-RDA limits researchers to search for minute differences between a restricted number of samples. Due to the technical complexity associated with cDNA-RDA, it is not feasible to compare several samples with each other. Furthermore, to gain a true representation of the whole set of differentially expressed genes in a sample, several repetitions with different enzymes may be needed. The cDNA-RDA technique is also sensitive to differences in expression level between differentially expressed genes [18]. This implies that highly expressed differential genes would be favoured after the first three to four repetitions of RDA. To identify larger pools of differentially expressed genes, including those that are nominally expressed, the RDA should only be repeated for one to two rounds. However, this results in the co-amplification of non-specific sequences, which leads to tedious screening of clones in the enriched library. These factors make it a tedious task to identify several genes of interest in a specific interaction.

Suppression Subtractive Hybridisation (SSH)

In 1996 Diatchenko *et al.* [21] proposed a novel method that combines the power of suppression PCR [32] with the normalisation and enrichment of subtraction hybridisation to enrich for low abundance mRNA transcripts (Fig. 1). Combination of these techniques resulted in a thousand-fold enrichment system. This procedure is based on combining enrichment for low abundance transcripts with a powerful PCR step that allows for the suppression of background molecules. In brief, the method involves restricting cDNA from two populations with a four base restriction enzyme, usually *Rsa*I, which produces blunt-ends. The cDNA population being studied (tester) is split into two subsets and two different adaptors are selectively ligated onto these cDNAs. This is accomplished by not

phosphorylating the adaptors so that they can only be ligated onto the 5'-ends of the cDNAs. These two subsets are then combined with an excess of the driver cDNA in two separate hybridisation reactions. The first hybridisation serves to equalise the abundant and less abundant transcript [21]. Then the two separate hybridisation reactions are combined in the presence of excess digested driver and allowed to hybridise further, during which an enrichment of the low abundance molecules occurs [21]. Since the adaptors are designed with two conserved regions of high GC content, all sequences that have two similar adaptors at each end will form a panhandle structure. These fragments cannot be amplified during the PCR step due to the suppression properties of this structure [32]. This means that the only molecules that are exponentially amplified are the ones with different adaptors at each end [21]. The resulting fragments are separated on a normal agarose gel and visualised through ethidium bromide staining.

SSH can identify gene transcripts that are expressed in exceptionally low amounts [41]. This makes it an efficient method to study small changes in expression of gene transcripts between samples that are quite similar. This is due to a predicted 1000 fold enrichment achieved after one round of SSH [21]. This is made possible by the normalisation that takes place during the first subtraction which proceeds according to second order hybridisation kinetics [21; 32]. Another robust feature of SSH is that no separation of the fragments of interest from the rest of the pool is needed. This increases the sensitivity for fragments of low abundance. Furthermore, the identification of false positives is less pronounced than for differential display methods [41].

As with other PCR-based subtraction techniques, SSH has several limitations. The use of SSH to identify novel transcripts is limited by the number of samples that can be processed. It has also been reported that the complexity of the sample can influence the subtraction efficiency in certain samples [61]. Restriction enzyme digest of the samples can also lead to very short fragments. This in turn leads to fragments that are inordinately short to match the fragment to a known EST or gene sequence [54]. The pool of identified transcripts after a subtraction leads to a banding profile with significant smearing. This necessitates a high level of screening to identify clones in the subsequent enriched library.

Application of Techniques to Plant-Pathogen Interactions

Since the development of these techniques to study gene expression, they have been used extensively to study interactions among plants and fungal pathogens. These studies can be divided into race specific resistance, and non-host studies in plants; as well as studies of changes in gene expression of fungal pathogens. Although cDNA-RDA was developed in 1994, it has not yet been used to study the interaction between plants and pathogens. Therefore, the following section of this review includes studies using DD, cDNA-AFLP and SSH.

Race Specific Resistance Linked Studies

To study the mode of action and genes linked to signal transduction early (0-4 hrs) in the interaction between *Nicotiana tabacum*, transgenic for *Cf-9* gene, and *Cladosporium fulvum*, Durrant *et al.* [24] made use of cDNA-AFLP analysis. In this approach they studied the changes that occur in cell cultures, in contrast to leaves, after addition of the Avr9 protein. This model was used, as it is easier to synchronise and reproduce the experimental conditions under study. To disseminate induction of defence genes in reaction to the production of active oxygen species (AOS), they first treated the cells with diphenylene-iodonium to inhibit the production of AOS. This made it possible to isolate 260 fragments of which 37 up- and 5 down-regulated fragments shared homology to known genes. Eighteen of the up-regulated sequences were then analysed with cDNA-AFLP over the first four hours after exposure to the Avr-9 protein. This was possible as cDNA-AFLP can be used for semi-quantitative induction analysis that is proportional to induction analysis by Northern blot [4].

The gene transcripts analysed in this interaction are of great interest. As the induction of defence response genes by AOS has been eliminated, the identification of transcripts expressed early in the interaction can be attributed to function as regulatory elements or in signalling. This study also identified several *ACRE* genes (*Avr9/Cf-9* rapidly elicited genes) that were expressed over a range of stress conditions. The results indicates that there is significant similarity between responses over mechanical and biological stress [24]. More importantly, the study focused on the genes active during the early stages of pathogen infection. These identified genes should be expressed during the signal transduction pathway and not the defence genes that are induced through this pathway.

SSH and cDNA macroarrays have been used in a study to identify the altered gene expression induced by six quantitative trait loci (QTL) in the *Glycine max* and *Fusarium solani* f.sp. *glycines* interaction [37]. The aim of this study was to identify differences in gene expression levels between a recombinant inbred line (RIL) and a susceptible soybean variety, 14 days after infection of the roots with *F. solanum* f.sp. *solani.* After the SSH subtraction 259 clones were sequenced. Of these 135 non-redundant clones were selected and spotted onto nylon membranes. These clones were analysed by reverse Northern blot with RNA from the infected and non-infected plant roots. Hybridisation to cDNA from the non-resistant soybean displayed down-regulation, whereas the resistant RIL displayed a high level of up-regulated cDNAs. The results of this study indicate that the pyramiding of genes helps with effective maintenance of gene expression levels during pathogen infection [37]. It further shows that there is an additive level of gene expression that is correlated to the number of QTLs found in a specific genotype.

Recently, DD analysis has been used to study the role that mitochondria play during senescence and pathogen attack [50]. Here, cell cultures from *N. tabacum* have been treated with antimycin A, which inhibits mitochondrial electron transport [76], and gene expression from these cells was compared with untreated cell cultures. Using this technique, seven cDNAs were identified that are up-regulated. Six of these seven cDNAs shared homology to genes that are involved in programmed cell death. These can be linked to either senescence or to pathogen attack. These transcripts were also significantly induced when the cell cultures were expressed to H_2O_2 and salicylic acid. These two oxidative stress linked chemicals had not previously been linked to inhibition of mitochondrial electron transport [50].

When the mitochondrial electron transport system is inhibited by antimycin A there is a marked increase in reactive oxygen species (ROS). This increase, as well as induction of the seven transcripts, is inhibited when pre-treating the cell cultures with an antioxidant. This suggests that the formation of ROS species is linked to stress-induced signalling between the mitochondria and the nucleus. These data collected using DD isolated fragments support data reported from animal studies that the mitochondria play a significant role in programmed cell death.

Non-Host Interaction Studies

Two elegant studies on *Agrobacterium tumefaciens* induced genes in plants, *N. tabacum* and *Ageratum conyzoides* have been reported [22; 77]. Although *Agrobacterium* is not a fungal plant pathogen, these two studies are included as *Agrobacterium* is continuously used to generate transgenic plants through *Agrobacterium*-mediated transformation. Ditt *et al.* [22], used cDNA-AFLP analysis to study gene expression in *Ag. conyzoides* by comparing *Agrobacterium*- and mock-inoculated plants. The cDNA-AFLP analysis identified 75 strongly regulated fragments after 48 hrs, of which 56 were already highly regulated after 24 hrs. Further characterisation of subsets of these 56 fragments revealed homology to genes linked to signal transduction and pathogen response. Analysis of the defence response genes indicated that non-pathogenic bacteria, i.e. *Escherichia coli*, similarly regulate them. One of the signal transduction-linked genes, receptor kinase *Xa21*, was only regulated upon *Agrobacterium* infection. This gene was down-regulated by the presence of *Agrobacterium*. Infection by *Agrobacterium* further induced the expression of a nodulin- and lectin-like gene. Identification of these genes reflects the relationship between *Agrobacterium* and *Rhizobium*. These results indicate that *Agrobacterium* has evolved to adapt to and counter the defence responses launched by the plant that it is infecting [22].

Veena *et al.* [77], undertook a similar study by isolating differentially expressed genes in *N. tabacum*. Here they used SSH, combined with macroarray screening, to further understand plant responses upon transfer-competent and incompetent *Agrobacterium* strain infection. The results were similar to those reported by Ditt *et al.* [22]. Defence responses in *Nicotiana* were also severely down-regulated after 48 hrs. Further indications were that *Agrobacterium* makes use of the host plant cellular machinery to assist in the infection progress. Together these studies could be used to assist in developing new strategies for transformation of recalcitrant plants [77].

Taylor and Harrier [70] used DD to study changes in expression levels of *Lycopersicon esculentum* cells while they are colonised by *Glomus mosseae*, an arbuscular mycorrhizal (AM) fungus. To circumvent fungal cDNA isolation from infected plants, DD was done on cDNA isolated from the tomato leaves. This is possible as infection by AM fungi changes the physiology of the whole plant [70]. This analysis identified 5 differentially expressed fragments, of which four were down-regulated in leaf tissue and only one displayed up-regulation in both the leaf and root tissues. One fragment displayed tissue-specific regulation,

as it was up-regulated in root tissue, but down-regulated in leaf tissue. This study has led to the identification of genes that can be used to compare pathogen and symbiotic infection in plants.

Pathogen Derived Transcript Analyses

To study the induction of genes in *Peronospora parasitica* during infection of *Arabidopsis thaliana*, van der Biezen *et al.* [75], used cDNA-AFLP analysis. Here they inoculated the *eds1* strain of *A. thaliana* with *P. parasitica*. During this interaction they analysed the combined pool of plant and fungal cDNA. From a total of 16000 fragments, sixty fragments were only present in the infected samples. Twenty-three fragments were selected from the original sample set and they were sequenced. This revealed that seven of the chosen fragments were of *Arabidopsis* origin. Ten of the remaining sixteen fragments were further analysed. Three of these fragments shared high homology to stress linked sequences, and the rest were linked to housekeeping genes, as identified by database searching.

P. parasitica is a biotrophic pathogen. To study gene expression in this oomycete is extremely difficult as isolated RNA will be from both the host and the pathogen. However, cDNA-AFLP proved to be an effective method to identify differentially expressed transcripts from mixed pools of cDNA [75].

Use of SSH to study gene expression in the biotrophic pathogen, *Puccinia triticina*, has also been reported [72]. Study of genes expressed during the infection of *Triticum aestivum* by *P. triticina* is extremely difficult. To enrich for genes of fungal origin, SSH was conducted with cDNA from mock-inoculated control plants as driver, and cDNA from plants inoculated with urediniospores as tester. The subtraction yielded 2800 cloned fragments, of which 350 were randomly picked and sequenced. This led to the identification of 104 non-redundant clones. Of the 104 clones, forty-four were of ribosomal origin. Most of these were from the fungus, leading to the speculation that a heightened level of protein synthesis is occurring. This conclusion is justified because samples were taken two days before sporulation [72]. A further fifty-six clones were further analysed by reverse Northern blot analysis against cDNA from *Puccinia* germinating spores, cDNA from inoculated wheat, as well as mock-inoculated wheat control. Twenty-six of the cloned fragments displayed induction in the cDNA from the inoculated wheat plants. Thirteen of these fragments did not hybridise to either the control cDNA or the germinating spores. Their fungal origin was confirmed by Southern

hybridisation to DNA extracted from germinating spores. These results show that the thirteen fragments have an expression specific to infection stage of the pathogen.

In the study of Thara *et al.* [72], several of the isolated fragments shared homology to known genes expressed during fungal infection. The most significant of these was a fragment that shared homology with cyclophilin, a virulence determining protein expressed by *Magnaporthe grisea* when infecting rice [79]. Three of the isolated fragments shared high homology to genes that are expressed in other plant pathogenic fungi upon infection. This suggests that gene expression in different pathogens during infection is similar. Furthermore, the high level of enrichment for transcripts of fungal origin in the library supports the use of SSH to study the induction of genes during an obligatory biotrophic interaction between a host and pathogen.

Muñoz *et al.* [55] used DD to study the induction of genes in *Phytophthora capsici* during infection of *Capsicum annuum*. Mycelial mats of a virulent and non-virulent strain of *P. capsici* were co-cultured with, and allowed to infect, three-week-old pepper seedlings. They were then physically removed from the host, RNA was extracted and displayed against RNA obtained from mycelial mats that were cultured in the absence of pepper seedlings. Eight differentially induced fragments were isolated. Homology to known genes could be identified for only one of the isolated fragments. This fragment shared homology with cutinase genes from other fungal pathogens.

The comparison of a non-virulent strain against a virulent strain by Muñoz *et al.* [55], identified a novel cutinase gene from an oomycete pathogen during infection of pepper seedlings. The role of this gene in pathogenesis has been well characterised in ascomycetes but has never before been reported for any oomycete [55]. This fragment was isolated after the first two to four hours of infection, indicating that it may play a significant role during the infection process. This result underscores the sensitivity of DD in this system to identify differentially induced transcripts during pathogen infection.

Novel Uses Not Linked To Host-Pathogen Interactions

Brugmans *et al.* [15], used segregating populations of *A. thaliana* and *Solanum tuberosum* to derive transcriptome maps. The 214 markers from the female and 219 markers from the male chromosome maps from *Solanum* were combined with previously published data to derive the

final transcriptome maps. Data produced from this study indicate that the identified polymorphisms in transcript fingerprints are not due to expression levels, but linked more to single nucleotide polymorphisms (SNP) and indels (insertions or deletions) [15]. These maps will be very useful in map-based cloning, as the markers placed on the map will be directly linked to coding regions of the genome. Hence they predict that current BAC landing through genomic markers will be more efficient through the use of 'gene landing' derived from transcriptome maps [15].

Breyne *et al.* [13], reported the use of quantitative genome-wide expression analysis as a viable alternative to microarray profiling. Here, they studied the cell cycle in synchronised *N. tabacum* cells over 11 hrs, by inhibiting DNA polymerase α. To improve the ratio of transcript identification, three changes were made to the original cDNA-AFLP protocol [4]. First, choosing the most 3'-end fragment reduced the number restriction fragments per transcript. Secondly, an *in silico* analysis was done to obtain the enzyme combination that scanned the highest percentage of the transcriptome. These enzymes were further selected based on the size of the fragment produced and the fact that this fragment must span across some of the coding region of the transcript. Thirdly, the optimal number of selective nucleotides was determined to enhance the amplification of transcripts expressed in low levels. By comparing the expression of several histone genes, it was possible to evaluate the quality of the data produced using this technique. The refinements to cDNA-AFLP reported here, have revealed that the emerging data can be applied to any biological system to study global gene expression [13]. Furthermore, it is possible to analyse the data in a similar manner to microarray data. Hence it provides a useful alternative to more expensive microarray analysis.

Comparison of the Methods

A preliminary literature search on the Elsevier Publishing Group's Scirus search engine indicates that DD and then cDNA-AFLP are the preferred techniques to study genetic responses in plant-pathogen interactions. This indicates an obvious bias towards the use of DD and cDNA-AFLP to study plant-pathogen interactions. This bias is not dependent on the time since these methods have been developed. The first use of cDNA-AFLP was reported approximately four years after DD [43; 4]. During this time both RAP-PCR and cDNA-RDA were invented [82; 36], and already the use of cDNA-AFLP has surpassed the use of these

two methods. This biased use can be attributed to several facts. The two most prevalent are the inherent complexity of certain procedures, and the development of commercial kits for certain techniques.

The preference of one method over another is well illustrated by the use of cDNA-RDA and SSH. cDNA-RDA has not been used to study host-pathogen interactions. This is most probably because of the emergence of SSH, shortly after the application of RDA to study changes at the transcriptome level. SSH has several advantages over cDNA-RDA, of which the difference in technical difficulty is the most apparent. SSH incorporates several steps less than the cDNA-RDA procedure (Fig. 1). Therefore, it is technically less challenging to perform, without any loss in sensitivity. As with cDNA-RDA, it is possible to increase the sensitivity by re-introducing the products of the first subtraction into the driver pool and to use this in a second hybridisation. There is no size selection before the actual subtraction reactions in SSH. Therefore, there is no need for multiple enzymes to 'represent' the pool of cDNAs. This implies that SSH is faster to accomplish than cDNA-RDA. Furthermore, no PCR bias is introduced before the actual subtraction hybridisation steps in SSH. However, the suppression effect of the primary PCR in SSH adds tremendous value to this technique. Here the formation of temperature stable structure cancels out the effect of amplification from background molecules that contain similar adaptors on either end [32].

cDNA-AFLP has several advantages over DD, and other methods. As it makes use of specific primers that bind to the adaptor sequence, it does not suffer from any of the difficulties encountered by arbitrarily primed reactions [4]. The semi-quantitative nature of cDNA-AFLP also increases its value as method of choice to use when analysing differential gene expression. It has been reported in several studies that this feature of cDNA-AFLP can be correlated to analysis by Northern blots [4; 14; 15]. This adds value to the use of this technique, as preliminary data obtained from the analysis can be used for prediction of the best clones to use for further study.

A direct comparison of cDNA-AFLP and DD in identifying differential gene induction linked to a phytopathogen has been reported [29]. These authors report on the use of both techniques to identify genes involved in the detoxification of brassinin, an alkaloid produced by *Brassica* species. They also recorded the efficacy of each method to identify genes expressed in fungi. The fungal material used was cultured in minimal media that was supplemented with *p*-

methylbenzyldithiocarbamate, an analogue of brassinin. Comparison between the two methods indicated that cDNA-AFLP displayed an average number of 6 differential fragments per primer pair, for DD this number was two differential fragments per primer pair. The DD analysis displayed short fragments with a high level of background noise. This necessitated the use of cDNA-AFLP for further analysis. From this analysis 11 of 14 fragments were identified as being differentially expressed. In this instance, it was not possible to isolate any differentially induced fragments using DD without extensive further screening.

Concluding Remarks

The identification and understanding of gene expression in host-pathogen interactions is of great importance to plant pathology and ultimately global food security. Knowledge gained from gene expression studies would be useful in developing strategies for durable resistance in crop plants. This research is complex and requires reliable techniques to study the interactions in plants, particularly where limited sequence data are available. PCR-based techniques provide a rapid and reliable option when compared with the more expensive high throughput techniques such as microarrays, SAGE, and MPSS [15]. Increasing the throughput of techniques such as cDNA-AFLP and DD is possible. This became possible with the increase in automatic gel analysis systems, for example, Saga^{MX} AFLP® Analysis Software (LI-COR), AFLP®-Quantar™ (*Pro*) (Keygene N.V.), GeneScan® Analysis Software (Applied Biosystems), and CEQ™ AFLP® Dominant Scoring Software (Beckman Coulter). For high throughput screening it is necessary to use an RNA extraction technique that is rapid, provides pure RNA and that is repeatable. Furthermore, high throughput amplification systems, and automatic scoring software are required [56].

Choosing between the different techniques depends on the question that needs to be answered. The implementation of cDNA-AFLP and DD is similar, as is that between cDNA-RDA and SSH. However, subtraction based techniques cannot achieve the throughput produced by non-subtraction-based techniques. A further advantage of non-subtraction techniques is that they consider both the up and down-regulated fragments in one display [9]. This negates the need for the reverse repetition of the experiment. Therefore, when several samples have to be compared over time or treatments, the option would be for DD and/or cDNA-AFLP. If the opposite is true and a low sample number is used, then cDNA-RDA or SSH can be used.

These two techniques incorporate a high level of sensitivity that is not necessarily matched by differential display techniques [36; 21; 41].

To increase the functionality of the different PCR-based techniques, it is possible to combine several of these. Birch *et al.* [7], reported the successful use of cDNA-AFLP to test the differential expression of fragments identified from a SSH reaction in *Phytophthora*-infected potato. The quantitative nature of cDNA-AFLP makes this an efficient method to screen a large number of clones generated by subtraction techniques. The combination of SSH with cDNA macro- and microarray has also efficiently enhanced the throughput obtained in several studies [85; 37; 77]. This gives the same results as with screening through cDNA-AFLP, however, higher concentrations of RNA are needed in comparison with cDNA-AFLP. Similar studies have been reported for cDNA-RDA [30]. Other combinations that have also resulted in higher sensitivity were reported by Nakata *et al.* [57], who used subtraction hybridisation and differential screening of an SSH generated library; and Uhl and Wang [74], who screened a subtraction hybridisation library through DD.

The strength of gene discovery techniques that are based on PCR is their sensitivity. However, this sensitivity can also be a disadvantage as existing methods, such as Northern blot analysis, may not have adequate sensitivity to verify expression levels of the identified transcripts [21]. This is because the expression level is below the sensitivity level of conventional Northern blots or reverse Northern blots [21; 67]. However, the use of quantitative real-time RT-PCR can alleviate this problem. This technique combines the speed and sensitivity of PCR with the strength of Northern blots (For review see [16]). This is reflected in a recent report by Avrova *et al.* [3], where the sensitivity of real-time RT-PCR detected up-regulation of genes during the biotrophic phase of the oomycete pathogen *P. infestans*. These genes were previously only detected in the necrotrophic phase using Northern blot analysis.

Detection of molecular signalling during early phases of host-pathogen or biotrophic plant-fungus interactions is important. This will aid in understanding the pathogenicity of pathogens and the resistance response in plants. To understand these early interactions, sensitive techniques for gene discovery will become increasingly necessary. These studies can then be compared with host-pathogen interactions studied in model systems. This is most probably where techniques such as microarray and large-scale sequencing will predominate [44]. However, several projects with lower levels of financial support will focus on more specific

questions. Here, PCR-based techniques will most likely remain the methods of choice, as they are reliable, sensitive and relatively affordable.

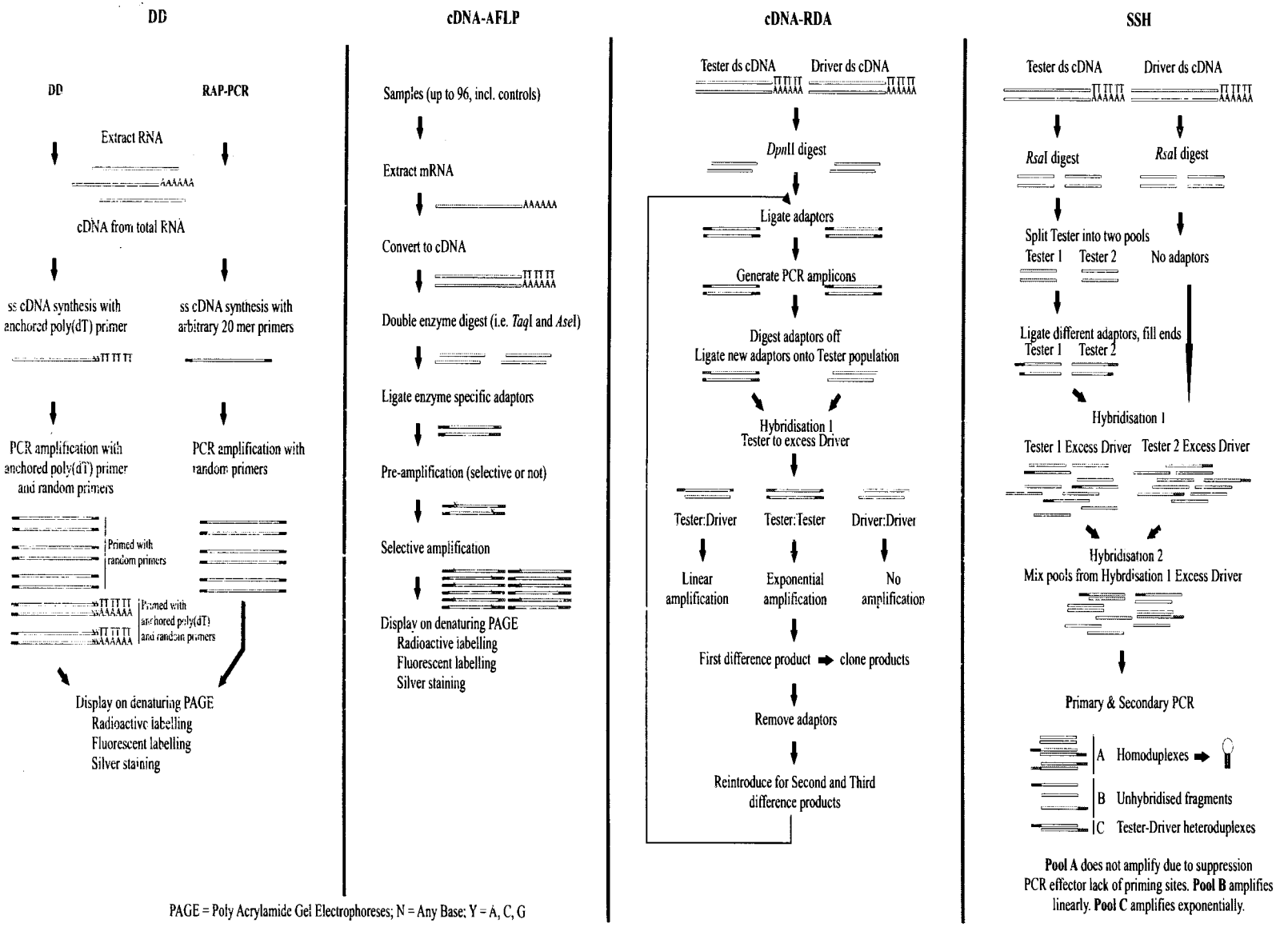


Figure 1: Comparative diagram of DD [43], RAP-PCR [83], cDNA-AFLP [4], cDNA-RDA [36], and SSH [21].

References

1. **Allona I, Quinn M, Shoop E, Swope K, St. Cyr S, Carlis J, Riedl J, Retzel E, Campbell MM, Sederoff R, Whetten RW** (1998) Analysis of xylem formation in pine by cDNA sequencing. Proceedings of the National Academy of Sciences, USA **95:** 9693-9698

2. **Avrova AO, Stewart HE, De Jong W, Heilbronn J, Lyon GD, Birch PR** (1999) A cysteine protease gene is expressed early in resistant potato interactions with *Phytophthora infestans*. Molecular Plant-Microbe Interactions **12:** 1114-1119

3. **Avrova AO, Venter E, Birch PRJ, Whisson SC** (2003) Profiling and quantifying differential gene transcription in *Phytophthora infestans* prior to and during the early stages of potato infection. Fungal Genetics and Biology **40:** 4-14

4. **Bachem CWB, van der Hoeven RS, de Bruijn SM, Vreugdenhil D, Zabeau M, Visser RGF** (1996) Visualization of differential gene expression using a novel method of RNA fingerprinting based on AFLP: Analysis of gene expression during potato tuber development. Plant Journal **9:** 745-753

5. **Bennetzen JL, Ma J** (2003) The genetic colinearity of rice and other cereals on the basis of genomic sequence analysis. Current Opinion in Plant Biology **6:** 128-133

6. **Bhalerao R, Nilsson O, Sandberg G** (2003) Out of the woods: forest biotechnology enters the genomic era. Current Opinion in Biotechnology **14:** 206-213

7. **Birch PJ, Avrova AO, Duncan JM, Lyon GD, Toth RL** (1999) Isolation of potato genes that are induced during an early stage of the hypersensitive response to *Phytophthora infestans*. Molecular Plant-Microbe Interactions **12:** 356-361

8. **Birch PRJ, Toth IK, Avrova AO, Dellagi A, Lyon GD, Pillay D** (2000) Recent advances in the identification of differentially expressed genes. South African Journal of Science **96:** 83-85

9. **Bos T, Hadman M** (1998) Improvements to differential display. *In* PD Siebert, JW Larrick, eds, Gene cloning and analysis by RT-PCR. Eaton Publishing, Natick, USA, pp 145-158

10. **Bosch I, Melichar H, Pardee AB** (2000) Identification of differentially expressed genes from limited amounts of RNA. Nucleic Acids Research **28:** e27

11. **Bradshaw Jr HD, Ceulemans R, Davis J, Stettler R** (2000) Emerging model systems in plant biology: Poplar (*Populus*) as a model forest tree. Journal of Plant Growth and Regulation **19:** 306-313

12. **Brenner S, Johnson M, Bridgham J, Golda G, Lloyd DH, Johnson D, Luo S, McCurdy S, Foy M, Ewan M, Roth R, George D, Eletr S, Albrecht G, Vermaas E, Williams SR, Moon K, Burcham T, Pallas M, DuBridge RB, Kirchner J, Fearon K, Mao JI, Corcoran K** (2000) Gene expression analysis by massively parallel signature sequencing (MPSS) on microbead arrays. Nature Biotechnology **18:** 630-634

13. **Breyne P, Dreesen R, Cannoot B, Rombaut D, Vandepoele K, Rombauts S, Vanderhaeghen R, Inzé D, Zabeau M** (2003) Quantitative cDNA-AFLP analysis for genome-wide expression studies. Molecular Genetics and Genomics **269:** 173-179

14. **Breyne P, Zabeau M** (2001) Genome-wide expression analysis of plant cell cycle modulated genes. Current Opinion in Plant Biology **4:** 136-142

15. **Brugmans B, Fernandez del Carmen A, Bachem CWB, van Os H, van Eck HJ, Visser RGF** (2002) A novel method for the construction of genome wide transcriptome maps. Plant Journal **31:** 211-222

16. **Bustin SA** (2000) Absolute quantification of mRNA using real-time reverse transcription polymerase chain reaction assays. Journal of Molecular Endocrinology **25:** 169-193

17. **Chaffey N, Cholewa E, Regan S, Sundberg B** (2002) Secondary xylem development in *Arabidopsis*: a model for wood formation. Physiologia Plantarum **114:** 594-600

18. **Chang D, Denny C** (1998) Isolating differentially expressed genes by representational difference analysis. *In* PD Siebert, JW Larrick, eds, Gene cloning and analysis by RT-PCR. Eaton Publishing, Natick, USA, pp 193-202

19. **Cheong YH, Chang HS, Gupta R, Wang X, Zhu T, Luan S** (2002) Transcriptional profiling reveals novel interactions between wounding, pathogen, abiotic stress, and hormonal responses in *Arabidopsis*. Plant Physiology **129:** 661-677

20. **Dellagi A, Birch PRJ, Heilbronn J, Lyon GD, Toth IK** (2000) cDNA-AFLP analysis of differential gene expression in the prokaryotic plant pathogen *Erwinia caratovora*. Microbiology **146:** 165-171

21. **Diatchenko L, Lau YC, Campbell AP, Chenchik A, Moqadam F, Huang B, Lukyanov S, Lukyanov K, Gurskaya N, Sverdlov ED, Siebert PD** (1996) Suppression subtractive hybridization: A method for generating differentially regulated or tissue-specific cDNA probes and libraries. Proceedings of the National Academy of Sciences, USA **93:** 6025-6030

22. **Ditt RF, Nester EW, Comai L** (2001) Plant gene expression response to *Agrobacterium tumefaciens*. Proceedings of the National Academy of Sciences, USA **98:** 10954-10959

23. **Donson J, Fang Y, Espiritu-Santo G, Xing W, Salazar A, Miyamoto S, Armendarez V, Volkmuth W** (2002) Comprehensive gene expression analysis by transcript profiling. Plant Molecular Biology **48:** 75-97

24. **Durrant WE, Rowland O, Piedras P, Hammond-Kosack KE, Jones JDG** (2000) cDNA-AFLP reveals a striking overlap in race-specific resistance and wound response gene expression profiles. Plant Cell **12:** 963-977

25. **Elsik CG, Williams CG** (2000) Retroelements contribute to the excess low-copy-number DNA in pine. Molecular and General Genetics **264:** 47-55

26. **Feng Q, Zhang Y, Hao P, Wang S, Fu G, Huang Y, Li Y, Zhu J, Liu Y, Hu X, Jia P, Zhang Y, Zhao Q, Ying K, Yu S, Tang Y, Weng Q, Zhang L, Lu Y, Mu J, Lu Y, Zhang LS, Yu Z, Fan D, Liu X, Lu T, Li C, Wu Y, Sun T, Lei H, Li T, Hu H, Guan J, Wu M, Zhang R, Zhou B, Chen Z, Chen L, Jin Z, Wang R, Yin H, Cai Z, Ren S, Lv G, Gu W, Zhu G, Tu Y, Jia J, Zhang Y, Chen J, Kang H, Chen X, Shao C, Sun Y, Hu Q, Zhang X, Zhang W, Wang L, Ding C, Sheng H, Gu J, Chen S, Ni L, Zhu F, Chen W, Lan L, Lai Y, Cheng Z, Gu M, Jiang J, Li J, Hong G, Xue Y, Han B** (2002) Sequence and analysis of rice chromosome 4. Nature **420:** 316-320

27. **Feuillet C, Keller B** (1999) High gene density is conserved at syntenic loci of small and large grass genomes. Proceedings of the National Academy of Sciences, USA **96:** 8265-8270

28. **Freeling M** (2001) Grasses as a single genetic system. Reassessment 2001. Plant Physiology **125:** 1191-1197

29. **Gellatly KS, Ash GJ, Taylor JL** (2001) Development of a method for mRNA differential display in filamentous fungi: comparison of mRNA differential display reverse transcription polymerase chain reaction and cDNA amplified fragment length polymorphism in *Leptosphaeria maculans*. Canadian Journal of Microbiology **47:** 955-960

30. **Geng M, Wallrapp C, Müller-Pillasch F, Frohme M, Hoheisel JD, Gress TM** (1998) Isolation of differentially expressed genes by combining representational difference analysis (RDA) and cDNA library arrays. Biotechniques **25:** 434-438

31. **Goff SA, Ricke D, Lan TH, Presting G, Wang R, Dunn M, Glaszebrook J, Sessions A, Oelier P, Varma H, Hadley D, Hutchison D, Martin C, Katagiri F, Lange BM, Moughamer T, Xia Y, Budworth P, Zhong J, Miguel T, Paszkowski U, Zhang S, Colbert M, Sun WL, Chen L, Cooper B, Park S, Wood TC, Mao L, Quail P, Wing R, Dean R, Yu Y, Zharkikh A, Shen R, Sahasrabudhe S, Thomas A, Cannings R, Gutin A, Pruss D, Reid J, Tavtigian S, Mitchell J, Eldredge G, Scholl T, Miller RM, Bhatnagar S, Adey N, Rubano T, Tusneem N, Robinson R, Feldhaus J, Macalma T, Oliphant A, Briggs S** (2002) A draft sequence of the rice genome (*Oryza sativa* L. ssp. *japonica*). Science **296:** 92-100

32. **Gurskaya NG, Diatchenko L, Chenchik A, Siebert PD, Khaspekov GL, Lukyanov KA, Vagner LL, Ermolaeva OD, Lukyanov SA, Sverdlov ED** (1996) Equalizing cDNA subtraction based on selective suppression of polymerase chain reaction: Cloning of Jurkat cell transcripts induced by phytohemaglutinin and phorbol 12-myristate 13-acetate. Analytical Biochemistry **240:** 90-97

33. **Hall SE, Dvorak WS, Johnston JS, Price HJ, Williams CG** (2000) Flow cytometric analysis of DNA content for tropical and temperate new world pines. Annals of Botany **86:** 1081-1086

34. **Han C, Xue Y** (2003) Genome-wide intraspecific DNA-sequence variations in rice. Current Opinion in Plant Biology **6:** 134-138

35. **Hedrick SM, Cohen DI, Nielsen EA, Davis MM** (1984) Isolation of cDNA clones encoding T cell-specific membrane-associated proteins. Nature **308:** 149-153

36. **Hubank M, Schatz DG** (1994) Identifying differences in mRNA expression by representational difference analysis of cDNA. Nucleic Acids Research **22:** 5640-5648

37. **Iqbal M, Yaegashi S, Njiti V, Ahsan R, Cryder K, Lightfoot D** (2002) Resistance locus pyramids alter transcript abundance in soybean roots inoculated with *Fusarium solani* f.sp. *glycines*. Molecular Genetics and Genomics **268:** 407-417

38. **Izawa T, Takahashi Y, Yano M** (2003) Comparative biology comes into bloom: genomic and genetic comparison of flowering pathways in rice and *Arabidopsis*. Current Opinion in Plant Biology **6:** 113-120

39. **Kirst M, Johnson AF, Baucom C, Ulrich E, Hubbard K, Staggs R, Paule C, Retzel E, Whetten R, Sederoff RR** (2003) Apparent homology of expressed genes from wood-forming tissues of loblolly pine (*Pinus taeda* L.) with *Arabidopsis thaliana*. Proceedings of the National Academy of Sciences, USA **100:** 7383-7388

40. **Kozian DH, Kirschbaum BJ** (1999) Comparative gene-expression analysis. Trends in Biotechnology **17:** 73-78

41. **Kuang WW, Thompson DA, Hoch RV, Weigel RJ** (1998) Differential screening and suppression subtractive hybridization identified genes differentially expressed in an estrogen receptor-positive breast carcinoma cell line. Nucleic Acids Research **26:** 1116-1123

42. **Lavery DJ, Lopez-Molina L, Fleury-Olela F, Schibler U** (1997) Selective amplification via biotin- and restriction-mediated enrichment (SABRE), a novel selective amplification procedure for detection of differentially expressed mRNAs. Proceedings of the National Academy of Sciences, USA **94:** 6831-6836

43. **Liang P, Pardee AB** (1992) Differential display of eukaryotic messenger RNA by means of the polymerase chain reaction. Science **257:** 967-971

44. **Lievens S, Goormachtig S, Holsters M** (2001) A critical evaluation of differential display as a tool to identify genes involved in legume nodulation: looking back and looking forward. Nucleic Acids Research **29:** 3459-3468

45. **Lisitsyn N, Lisitsyn N, Wigler M** (1993) Cloning the differences between two complex genomes. Science **259:** 946-951

46. **Lisitsyn NA, Segre JA, Kusumi K, Lisitsyn NM, Nadeau JH, Frankel WN, Wigler MH, Lander ES** (1994) Direct isolation of polymorphic markers linked to a trait by genetically directed representational difference analysis. Nature Genetics **6:** 57-63

47. **Lockhart DJ, Dong H, Byrne MC, Follettie MT, Gallo MV, Chee MS, Mittmann M, Wang C, Kobayashi M, Horton H, Brown EL** (1996) Expression monitoring by hybridization to high-density oligonucleotide arrays. Nature Biotechnology **14:** 1675–1680

48. **Lowe K** (2000) Recent advances in fluorescent differential display. *In* RA Leslie, HA Robertson, eds, Differential display. Oxford University Press, Oxford, p 256

49. **Mann CC, Plummer ML** (2002) Forest biotech edges out of the lab. Science **295:** 1626-1629

50. **Maxwell DP, Nickels R, McIntosh L** (2002) Evidence of mitochondrial involvement in the transduction of signals required for the induction of genes associated with pathogen attack and senescence. Plant Journal **29:** 269-279

51. **McClelland M, Mathieu-Daude F, Welsh J** (1995) RNA fingerprinting and differential display using arbitrarily primed PCR. Trends in Genetics **11:** 242-246

52. **Medhurst AD, Chambers D, Gray J, Davis JB, Shale JA, Mason I, Jenner P, Newton RA** (2000) Practical aspects of the experimental design for differential display of transcripts obtained from complex tissue. *In* RA Leslie, HA Robertson, eds, Differential display. Oxford University Press, Oxford, p 256

53. **Money T, Reader S, Qu LJ, Dunford RP, Moore G** (1996) AFLP-based mRNA fingerprinting. Nucleic Acids Research **24:** 2616-2617

54. **Mooney JL, Kayne PS, O'Brien SP, Debouck CM** (2000) Construction and application of gene microarrays on nylon membranes. *In* RA Leslie, HA Robertson, eds, Differential display. Oxford University Press, Oxford, p 256

55. **Muñoz CI, Bailey AM** (1998) A cutinase-encoding gene from Phytophthora capsici isolated by differential-display RT-PCR. Current Genetics **33:** 225-230

56. **Myburg AA, Remington DL, O'Malley DM, Sederoff RR, Whetten RW** (2001) High-throughput AFLP analysis using infrared dye-labeled primers and an automated DNA sequencer. Biotechniques **30:** 348-357

57. **Nakata PA, McConn MM** (2002) Sequential subtractive approach facilitates identification of differentially expressed genes. Plant Physiology and Biochemistry **40:** 307-312

58. **Plomion C, Leprovost G, Stokes A** (2001) Wood formation in trees. Plant Physiology **127:** 1513-1523

59. **Qin L, Prins P, Jones JT, Popeijus H, Smant G, Bakker J, Helder J** (2001) GenEST, a powerful bidirectional link between cDNA sequence data and gene expression profiles generated by cDNA-AFLP. Nucleic Acids Research **29:** 1616-1622

60. **Ramonell KM, Somerville S** (2002) The genomics parade of defense responses: to infinity and beyond. Current Opinion in Plant Biology **5:** 291-294

61. **Rebrikov DV, Britanova OV, Gurskaya NG, Lukyanov KA, Tarabykin VS, Lukyanov SA** (2000) Mirror orientation selection (MOS): a method for eliminating false positive clones from libraries generated by suppression subtraction hybridization. Nucleic Acids Research **28:** e90

62. **Rothschild CB, Bowden DW, Shelness GS** (1998) DD/AP-PCR: differential display and arbitrarily primed PCR of oligo(dT) cDNA. *In* PD Siebert, JW Larrick, eds, Gene cloning and analysis by RT-PCR. Eaton Publishing, Natick, USA, pp 159-169

63. **Sasaki T, Sederoff RR** (2003) Genome studies and molecular genetics. The rice genome and comparative genomics in higher plants. Current Opinion in Plant Biology **6:** 97-100

64. **Schena M, Shalon D, Davis RW, Brown PO** (1995) Quantitative monitoring of gene expression patterns with a complementary DNA microarray. Science **270:** 467-470

65. **Schenk PM, Kazan K, Wilson I, Anderson JP, Richmond T, Somerville SC, Manners JM** (2000) Coordinated plant defense responses in Arabidopsis revealed by microarray analysis. Proceedings of the National Academy of Sciences, USA **97:** 11655-11660

66. **Schoof H, Karlowski WM** (2003) Comparison of rice and *Arabidopsis* annotation. Current Opinion in Plant Biology **6:** 106-112

67. **Simões-Araújoa JL, Rodriguesa RL, de A. Gerhardta LB, Mondegoa JMC, Alves-Ferreiraa M, Rumjanekb NG, Margis-Pinheiro M** (2002) Identification of differentially expressed genes by cDNA-AFLP technique during heat stress in cowpea nodules. FEBS Letters **515:** 44-50

68. **Sterky F, Regan S, Karlsson J, Hertzberg M, Rohde A, Holmberg A, Amini B, Bhalerao R, Larsson M, Villarroel R, van Montagu M, Sandberg G, Olsson O, Teeri TT, Boerjan W, Gustafsson P, Uhlén M, Sundberg B, Lundeberg J** (1998) Gene discovery in the wood-forming tissues of poplar: Analysis of 5,692 expressed sequence tags. Proceedings of the National Academy of Sciences, USA **95:** 13330-13335

69. **Taylor G** (2002) *Populus*: *Arabidopsis* for forestry. Do we need a model tree? Annals of Botany **90:** 681-689

70. **Taylor JL, Harrier LA** (2003) Expression studies of plant genes differentially expressed in leaf and root tissues of tomato colonised by the arbuscular mycorrhizal fungus *Glomus mosseae*. Plant Molecular Biology **51:** 619-629

71. **Terryn N, Rouzé P, van Montagu M** (1999) Plant genomics. FEBS Letters **452:** 3-6

72. **Thara VK, Fellers JP, Zhou J-M** (2003) *In planta* induced genes of *Puccinia triticina*. Molecular Plant Pathology **4:** 51-56

73. **The *Arabidopsis* Genome Initiative** (2000) Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. Nature **408:** 796-815

74. **Uhl GR, Wang XB** (2000) Subtracted differential display and antisense approaches to analysing brain signalling. *In* RA Leslie, HA Robertson, eds, Differential display. Oxford University Press, Oxford, p 256

75. **van der Biezen EA, Juwana H, Parker JE, Jones JDG** (2000) cDNA-AFLP display for the isolation of *Peronospora parasitica* genes expressed during infection in *Arabidopsis thaliana*. Molecular Plant-Microbe Interactions **13:** 895-898

76. **Vanlerberghe GC, McIntosh L** (1994) Mitochondrial electron transport regulation of nuclear gene expression. Plant Physiology **105:** 867-874

77. **Veena, Jiang H, Doerge RW, Gelvin SB** (2003) Transfer of T-DNA and Vir proteins to plant cells by *Agrobacterium tumefaciens* induces expression of host genes involved in mediating transformation and suppresses host defense gene expression. Plant Journal **35:** 219-236

78. **Velculescu VE, Zhang L, Vogelstein B, Kinzler KW** (1995) Serial analysis of gene expression. Science **270:** 484–487

79. **Viaud MC, Balhadère PV, Talbot NJ** (2002) A *Magnaporthe grisea* cyclophilin acts as a virulence determinant during plant infection. Plant Cell **14:** 917-930

80. **Wakamiya I, Newton RJ, Johnston JS, Price HJ** (1993) Genome size and environmental factors in the genus *Pinus*. American Journal of Botany **80:** 1235-1241

81. **Ware D, Stein L** (2003) Comparison of genes among cereals. Current Opinion in Plant Biology **6:** 121-127

82. **Welsh J, Chada K, Dalal SS, Cheng R, Ralph D, McClelland M** (1992) Arbitrarily primed PCR fingerprinting of RNA. Nucleic Acids Research **20:** 4965-4970

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

84. **Williams JGK, Kubelik AR, Livak KJ, Rafalski JA, Tingey SV** (1990) DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. Nucleic Acids Research **18:** 6531-6535

85. **Yang GP, Ross DT, Kuang WW, Brown PO, Weigel RJ** (1999) Combining SSH and cDNA microarrays for rapid identification of differentially expressed genes. Nucleic Acids Research **27:** 1517-1523

86. **Yu J, Hu S, Wang J, Wong GK, Li S, Liu B, Deng Y, Dai L, Zhou Y, Zhang X, Cao M, Liu J, Sun J, Tang J, Chen Y, Huang X, Lin W, Ye C, Tong W, Cong L, Geng J, Han Y, Li L, Li W, Hu G, Huang X, Li W, Li J, Liu Z, Li L, Liu J, Qi Q, Liu J, Li L, Li T, Wang X, Lu H, Wu T, Zhu M, Ni P, Han H, Dong W, Ren X, Feng X, Cui P, Li X, Wang H, Xu X, Zhai W, Xu Z, Zhang J, He S, Zhang J, Xu J, Zhang K, Zheng X, Dong J, Zeng W, Tao L, Ye J, Tan J, Ren X, Chen X, He J, Liu D, Tian W, Tian C, Xia H, Bao Q, Li G, Gao H, Cao T, Wang J, Zhao W, Li P, Chen W,**

Wang X, Zhang Y, Hu J, Wang J, Liu S, Yang J, Zhang G, Xiong Y, Li Z, Mao L, Zhou C, Zhu Z, Chen R, Hao B, Zheng W, Chen S, Guo W, Li G, Liu S, Tao M, Wang J, Zhu L, Yuan L, Yang H (2002) A draft sequence of the rice genome (*Oryza sativa* L. ssp. *indica*). Science **296:** 79-92

87. **Zeng J, Gorski RA, Hamer D** (1994) Differential cDNA cloning by enzymatic degrading subtraction (EDS). Nucleic Acids Research **22:** 4381-4385

88. **Zhang Y, Sederoff RR, Allona I** (2000) Differential expression of genes encoding cell wall proteins in vascular tissues from vertical and bent loblolly pine trees. Tree Physiology **20:** 457-466

CHAPTER 2

Infection of *Pinus patula* with *Fusarium circinatum* does not induce detectable levels of chitinase

Submitted to Physiological and Molecular Plant Pathology as:
Infection of *Pinus patula* with *Fusarium circinatum* does not induce detectable levels of chitinase
E Venter, B D Wingfeld, M J Wingfeld, A-M Botha
PMPP03/66

Abstract

Although plants lack chitin, several chitinase genes have been isolated from them. Chitinases hydrolyse the ß-1,3 bonds in glucosamine polymers. It has been suggested that these enzymes are linked to early responses to invasion by fungal pathogens, and that they act by breaking down the pathogen cell walls that contain chitin. We have studied the induction of chitinase and ß-1,3 glucanase in *Pinus patula* after the infection with the pine pathogen *Fusarium circinatum*. The studies of chitinase induction were done by enzyme activity assays, Western blot assays and quantitative real time RT-PCR. The ß-1,3 glucanase enzyme induction was also assayed in conjunction with the chitinase assays. Our results showed that the ß-1,3 glucanase followed the same trends as those seen in other plant-pathogen and pest interactions. There was an induction of the ß-1,3 glucanase in the inoculated plants after 24 hrs, whereas the controls showed a response that peaked at 5 days after inoculation. In contrast, the enzyme assays, RT-PCR and Western blot analysis showed no detectable induction of chitinases after exposure to the pathogen. We amplified fragments of the *LP6* and *PSCHI4* homologues from *P. patula* proving that these chitinase genes are present in *P. patula*. These homologues displayed high degrees of similarity to homologues that have been identified in *P. taeda* and *P. strobus*. The lack of detectable levels of chitinase induction in *P. patula* might explain the high levels of susceptibility to the pitch canker fungus observed in seedlings of this tree.

Keywords

Chitinase, Patula pine, pitch canker fungus, quantitative real time RT-PCR, stress response

Nucleotide and/or amino acid sequence data are to be found at GenBank as Accession No. CF660335 & CF660336.

Introduction

The South African forestry industry relies heavily on *Pinus patula* as one of its dominant soft wood species. Studies on this species have shown that it shows a high degree of susceptibility to the pitch canker fungus, *Fusarium circinatum*. Thus, it is considered to be one of the most important pathogens of plantation tree species in South Africa [26]. *F. circinatum* causes severe losses in nurseries and also hampers the establishment of trees on new sites.

The pitch canker fungus *F. circinatum* [19; 8] was first associated with serious canker diseases in the South Eastern United States [17; 15]. In recent years, this important pine pathogen has been introduced into California where it has devastated native *P. radiata*. It has also been found in other parts of the world including Mexico, where it is considered to be native, and Chile [7; 27]. The pitch canker fungus was first detected in South African nurseries in the early 1990's [25]. Since then it has spread to all pine seedling nurseries in the country, but there have been no reports of the fungus infecting mature trees.

On *P. patula* and *P. radiata* seedlings in South Africa *F. circinatum* behaves as a cambial canker pathogen [26]. The seedlings show severe wilting three weeks after inoculation and they die within six weeks. Studies on mature trees have shown that the fungus stimulates pine trees to exude resin to reduce the spread of the pathogen [17].

F. circinatum is a necrotrophic pathogen that thrives on dead cells and this negates the effect of pitch formation [3; 14]. This forces the tree to rely on other measures of defence against infection. One of these strategies used by plants is to produce chitinases [13]. Chitinases are a group of enzymes that hydrolyse chitin, which is a major component of the cell walls of many fungal plant pathogens, but is not found in plants. Chitinases reside in four sub-classes I, II, III, and IV separated on their function and localisation in the plant cell [13]. The chitinases in groups I, II, and IV have similar structures and can be found either in the vacuoles or as extracellular enzymes. The chitinases in group III are generally characterised as being lysozymes, which have an added chitinase activity.

Production of chitinases in plants has been recognised as an early response to pathogen infection. These enzymes can be induced by a wide variety of environmental factors, including pathogens and pests. For example, Botha *et al.* [5] reported that in wheat, Russian wheat aphid (*Diuraphis noxia*) feeding, exogenous ethylene, and wounding induces chitinase enzymes. Likewise, in dicotyledonous *Arabidopsis thaliana* and *Nicotiana tabacum* [20] chitinases have been isolated and shown to respond to pathogen infection [18; 20; 21]. In pines it has been shown that chitinases can be induced by elicitors [28], wounding [29], and pathogens [14]. Chitinase expression is also down-regulated by environmental stresses [12]. The widespread induction of chitinases in angiosperms and gymnosperms suggests that this response is conserved throughout the plant kingdom.

The aim of this study was to investigate the expression of chitinase in *P. patula* after infection with the pitch canker fungus *F. circinatum*. Chitinase enzyme assays and Western blot analysis were done to determine the level of *P. patula* response to infection by *F. circinatum*. Expression of two known chitinase homologues (a Class I chitinase, *LP6*, and a Class II chitinase, *PSCHI4*) from pine trees were also characterised for *P. patula*.

Materials and Methods

Plant Materials

The plants used in this study were seedlings obtained from open-pollinated *Pinus patula* trees propagated by Komatiland Forests Research, Sabie, South Africa. These seedlings were cloned so that they could be used in multiple inoculation trials. Clones were planted in clonal hedges under shade netting on the research farm of the University of Pretoria.

Fungal Isolate and Inoculation Trials

An isolate of *Fusarium circinatum* (CMW13229) previously shown to be highly pathogenic [25] was used in this study. It is maintained in the Forestry and Agricultural Biotechnology Institute culture collection (FABI, University of Pretoria, Pretoria, South Africa). Spores were collected from cultures of the test isolate and stored at – 70°C to maintain pathogenicity of the isolate. For inoculations, spores were plated on ½ strength potato dextrose agar (PDA)

(NT-Merck) plates and the fungus was allowed to grow for one week. Fungal plugs were then inserted into manually created wounds (5 x 2 mm in size) on the stem or side branches of the seedlings. These wounds and agar blocks were sealed with laboratory film to prevent contamination and desiccation of the inoculums. Control plants were mock inoculated (wounding control) with sterile agar plugs.

Protein Extraction

Total cellular and extracellular protein was extracted at time zero and at 24 hour intervals for seven days. The phloem tissue surrounding the inoculation point was ground to a fine powder in liquid nitrogen. Then, 10 ml extraction buffer (50 mM Tris-HCl pH 7.8, 0.5 mM PMSF, 5 mM 2-mercaptoethanol, 10 % (w/v) insoluble polyvinylpyrrolidone) was added and the material was further ground. The suspension was centrifuged for 10 minutes at 10000 x g at 4°C. Total protein concentration was determined following the method of Bradford [6] using bovine γ-globulin as a standard. The protein extracts were divided into aliquots and stored at –20°C.

Enzyme Activity Assays and Protein Blot Analysis

Chitinase activity was assayed using a modification of the method of Boller *et al.*[4]. Controls with serial dilutions (0, 0.01, 0.02, 0.03, 0.04, 0.05, 1, 3, and 5 U) of commercial chitinase from *Serratia marcescens* (Sigma) were also performed to verify the assay conditions used (not shown). The reaction mixture contained 10 µl total cellular protein (1 µg), 490 µl 50 mM Sodium-acetate (pH 6.5), and 1.5 mg colloidal chitin. This was mixed and incubated at 37°C for 30 minutes. After incubation it was centrifuged at 1000 x g for 1 minute to precipitate the chitin. Three hundred microlitres of supernatant was then added to 20 µl 1.5 % (w/v) cytohelicase (Sigma) and 30 µl 1 M phosphate buffer (pH 7.1) and incubated at 37°C for 30 minutes. Then 250 µl was added to 50 µl 800 mM tetraborate buffer (pH 9.1), heated in boiling water for 3 minutes, cooled and 1500 µl 1 % (m/v) 4-dimethylaminobenzaldehyde added. This mixture was incubated at 37°C for 20 minutes and cooled under tap water. The absorbance was immediately read at 585 nm.

This enzyme assay was confirmed in a separate inoculation trial. In this trial, protein extraction and the chitinase assay were performed under standard conditions with chitin

Azure (Sigma) as substrate following standard protocols. The viability of the proteins in the samples were determined by a ß-1,3-glucanase assay (0.5 µg per assay) [9].

The extracted proteins were separated on 12 % polyacrylamide gels with a 6 % stacking gel (ratio acrylamide N,N'-methylene-bisacrylamide 100:1), transferred to PVDF membranes (Micron Separation Inc.) with a Bio-Rad Trans-Blot SD cell and probed using antibodies directed against tobacco chitinase (PR-Q) [22]. Detection was done with the ECL-Plus system (Amersham Life Science). The time trial experiment was done over seven days because results of other studies have shown that levels of chitinase accumulated after seven days [14]. Wheat protein from *Triticum aestivum* cultivar 'Tugela DN' (*Dn1*, SA1684/Tugela*5) infested by the Russian wheat aphid was used as positive control for chitinase induction [5].

PCR Detection of Chitinase Homologues

PCR was used to detect genes homologous to chitinase in the *P. patula* genome. Specific primers were designed from the chitinase homologues of *P. taeda* and *P. strobus* sequences from GenBank (Accession no U31309 and U57410) that amplified fragments smaller than 200 base pairs (bp). These primers (pschi4F 5-TTCTGCACAACAAGGAGTCG, pschi4R 5-GTAACCCCACGCATATGGAC; tchit 5-ATCAAGTGATGGTGGGGAAG, tchit3 5-TTGTCACCAGCATTTTCTCG) were then used to amplify potential homologues using the following reaction conditions: 94°C for 1 min, 54°C for 30 seconds, and 72°C for 1 min, for 30 cycles. The reaction had an initial denaturing step at 94°C for 3 min and a final elongation step at 72°C for 5 min. The PCR consisted of 1X Expand reaction buffer with 1.5 mM $MgCl_2$, 5 µM of each primer, 200 µM of each dNTP, and 1 U Expand polymerase mix (Roche).

Chitinase mRNA Induction Analysis

The PCR primers designed to amplify the short fragment were also used in reverse transcription PCR (RT-PCR) and real time PCR analysis of RNA extracted from two *P. patula* seedlings that display different levels of tolerance after pathogen inoculation. RNA was extracted at the time of inoculation, and after 1, 3, 6, 12, 24, and 48 hrs, using the CTAB method of Chang *et al.* [10]. The resulting RNA was treated with DNase I and converted to cDNA with AMV reverse transcriptase (Promega). Equal concentrations of template cDNA

were used in the RT-PCR and Lightcycler reactions using conditions, previously described. To ensure equal loading of template cDNA, RT-PCR was performed with Rubisco specific primers (rubiscoF 5- AATGTCTCCGTGGTGGACTC, rubiscoR 5- TAATTTCACCCGTC-TCAGCC).

Genomic Chitinase Gene Isolation and Computational Analysis

To verify the identity of the two chitinase homologues, longer sequences of the chitinase genes were obtained through PCR. Primers were designed from the ends of the *PSCHI4* (pscgen 5-CCCGCAATCATAAATGTGTG, pscgenR 5-GTCTGGCGTTTCTTCATATC, pscFull 5-ATGGCGTACACGAATATGAAG, pscFullR 5-TCGAGGTTGGATCCCAC, pschgenF 5-CCCGCAATCATAAATGTGTG, pschgenR 5-GTCTGGCGTTTCT-TCATATC) and *LP6* (chitlarge 5-GGCACGAGTTTACAAAAG, chitlargeR 5-CATCGAAACATTGTATGTC) homologues and used for amplification in a PCR reaction containing 5 μM of each primer, 1X PCR reaction buffer with 1.5 mM $MgCl_2$, 200 μM of each dNTP, and 1 U Expand polymerase mix (Roche). The PCR conditions were: 94°C for 1 min, 54°C for 30 seconds, and 72°C for 2 min, for 30 cycles. The reaction had an initial denaturing step at 94°C for 3 min and a final elongation step at 72°C for 5 min. The PCR products were cleaned by ethanol precipitation and subjected to cycle sequencing with BigDye™ Dye Terminator v3.0 and run on an ABI™ 3100 (Applied Biosystems). The sequences were aligned to *P. taeda* and *P. strobus* chitinase sequences in GenBank through the BLAST two sequences option (Accession No. U31309 and U57410) [1].

To investigate the grouping of the *LP6* and *PSCHI4* genes, the amino acid sequences were used in a phylogenetic analysis. These sequences were combined with data from available amino acid sequences used by Araki and Torikata [2], comprising the class I and II chitinases from several organisms. The sequences were downloaded from GenBank (Table 1) and then aligned with ClustalX [24]. The resulting alignments were used to derive protein genealogies by neighbour-joining analysis using PAUP* 4.0b10 [23].

Results

Enzyme Activity Assay and Protein Level Detection

ß-1,3-glucanase activity was induced 24 hrs after inoculation. This was followed by a decrease in activity and a second increase after 4 days (Fig. 1). In the control, wounding alone also induced the ß-1,3-glucanase activity. This induction reached its highest level after 5 days. Wounding alone resulted in a higher enzyme activity than wounding together with fungal inoculation.

The enzyme activity assay for the induction of chitinase was negative for all samples tested. No chitinase activity was detected, even in the presence of chitin azure as substrate (Fig. 1) [16].

Western blot analysis failed to detect any chitinase subunits in the *P. patula* samples, whereas the wheat control showed detection of the chitinase subunits. This was true for both the seedlings inoculated with the pitch canker fungus and the wounded control plants (Fig1).

Chitinase mRNA Induction Analysis

Constitutive low levels of transcripts of the two chitinase homologues were present even at the time of inoculation (Fig. 2). The levels of the *LP6* transcripts decreased sharply after wounding, but after 5 hrs no significant differences were detected between the wounded control and inoculated tissue. The concentration of the *PSCHI4* homologue increased after one hour, followed by a decrease and thereafter, it reached much higher levels in both the inoculated and wounded samples. The levels of the *LP6* homologue showed much higher steady state transcript levels than those found for *PSCHI4*.

Genomic Chitinase Gene Isolation and Computational Analysis

These PCRs amplified several retrotransposable elements and we used internal primers to amplify segments only from the open reading frames (ORF). For the *PSCHI4* homologue we were able to amplify 469 bp, and a 520 bp fragment for the *LP6* homologue. The *P. patula* fragment was homologous to *PSCHI4* from 798 – 1265 bp, and for the *LP6* homologue from 929 – 1448 bp. The homologues displayed high homology to the known *PSCHI4* and *LP6*

gene homologues (Fig. 3). The *LP6* alignment displayed only differences in the sequences, where the primers had annealed. In the *PSCHI4* homologue there was a 50 bp insertion that was not evident in the *P. strobus* gene.

The phylogenetic analysis showed seven distinct clades (Fig. 4). The two clades referred to as clade A and clade B are the class I low molecular weight (I-L) and class II low molecular weight (II-L) chitinases. All the class II chitinases are indicated by a closed square. The class II chitinases was dispersed throughout the phylogenetic tree and only the subclass II-L had a distinct phylogenetic clade. This group included the known pathogenesis related chitinases PR-P and PR-Q [18]. The *PSCHI4* gene is basal to this group. The only members of the class I chitinases that shows a phylogenetic clade was the subclass I-L. The *LP6* gene displayed higher similarity to the *Urtica dioica* class I chitinase than with the *Picea glauca* class I chitinase.

Discussion

In this study we examined the primary defence response of *P. patula* after inoculation with the pitch canker fungus. This was done in an effort to understand the early responses after pathogen infection in *P. patula*. To this end, we studied the levels of enzyme activity for both chitinases and ß-1,3-glucanase. No chitinase activity was detected over a period of 7 days.

We were not able to detect chitinase activity at any level using enzyme assays and Western blot analysis in *P. patula*. This is in contrast to the findings of Davis *et al.* [14] who observed an induction of PSCHI4 protein when *P. elliottii* var. *elliottii* and *P. taeda* were challenged with *F. circinatum*. This response was more pronounced in susceptible as opposed to resistant plants. Our studies considered chitinase levels at different time intervals up to 7 days after inoculation and we were unable to detect any chitinase activity at any time. In the resistant reaction examined by Davis *et al.* [14], the response was lower but highest accumulation of PSCHI4 occurred three days after inoculation.

Real time RT-PCR assays of the chitinase homologues displayed different levels of induction for the *LP6* and *PSCHI4* homologues after wounding or inoculation. The *LP6* transcripts in *P. patula* followed the same trend after wounding, as those reported by Chang *et al.* [11]. This is in contrast to the expected accumulation in transcripts that would be expected after inoculation with *F. circinatum*.

We were able to detect levels of the *PSCHI4* transcripts very early in the experiments by using real time RT-PCR. Davis *et al.* [14] and Wu *et al.* [29] report the induction of the *PSCHI4* transcripts by Northern blots at around 3 days and for the wounding study at 21 hours after wounding. However, there is no difference in *P. patula*'s response upon wounding or pathogen attack in the first 24 hours. This could indicate that the induction of *PSCHI4* in response to *F. circinatum* infection only takes place after 24 hrs.

The lack of any detectable chitinase activity in *P. patula* is an unusual and interesting result of this study. This is particularly because production of chitinase is generally recognised as one of the early responses by plants to pathogen infection [13]. Our results have shown that *P. patula PSCHI4* and *LP6* homologues and those in their *P. taeda* and *P. strobus* counterparts share a sequence similarity of 80 % and 99 % respectively, for the sections that we have sequenced. Comparison of the *P. patula* homologue to the *P. strobus PSCHI4* shows a distinct insertion into the first intron of the *P. patula* homologue [28]. The other nucleotide substitutions result in the change of 9 amino acids (not shown). Thus, the lack of chitinase activity is not due to the absence of the genes but possibly a deficiency in chitinase expression or up-regulation.

In the phylogenetic analysis several different class I and II chitinases were studied to identify the groups that should accommodate the *LP6* and *PSCHI4* homologues. The *PSCHI4* gene grouped basal to the subclass II-L clade. This is congruent with the placing of this chitinase as a class II based on its domain structures [28]. The *LP6* homologue had an unusual grouping with a flowering plant, *Urtica dioica* (stinging nettle). The relationship between chitinases in *U. dioica* and *P. patula* is closer than that between the chitinases in *P. taeda* and *Picea glauca* chitinases. This could be due to differences in genera, but one would expect that the chitinases in conifers would be most closely related.

Results of this study have shown not induced to detectable levels in *P. patula* challenged by *F. circinatum*. These low levels of chitinase activity could be linked to the susceptibility to pathogen attack seen in *P. patula*. This apparent lack of induction is interesting as the chitinase group is considered to be part of a general defence response to pathogen attack [13]. If chitinases are not induced in *P. patula* after wounding or pathogen attack, this would imply that the trees rely more on other defence responses, such as ß-1,3- glucanase induction, after challenge. Enhancing the induction of the chitinase genes in *P. patula* might increase resistance to pathogen attack. This could be achieved by placing the *PSCHI4* homologue under control of a pathogen-induced promoter to enhance the expression levels upon pathogen infection.

Acknowledgements

We thank Komatiland Forests Research, the National Research Foundation and the Mellon Foundation for financial assistance and Komatiland Forests Research for providing the plant material used in this experiment. We also thank A. Jacobs whom provided assistance with the phylogenetic analyses.

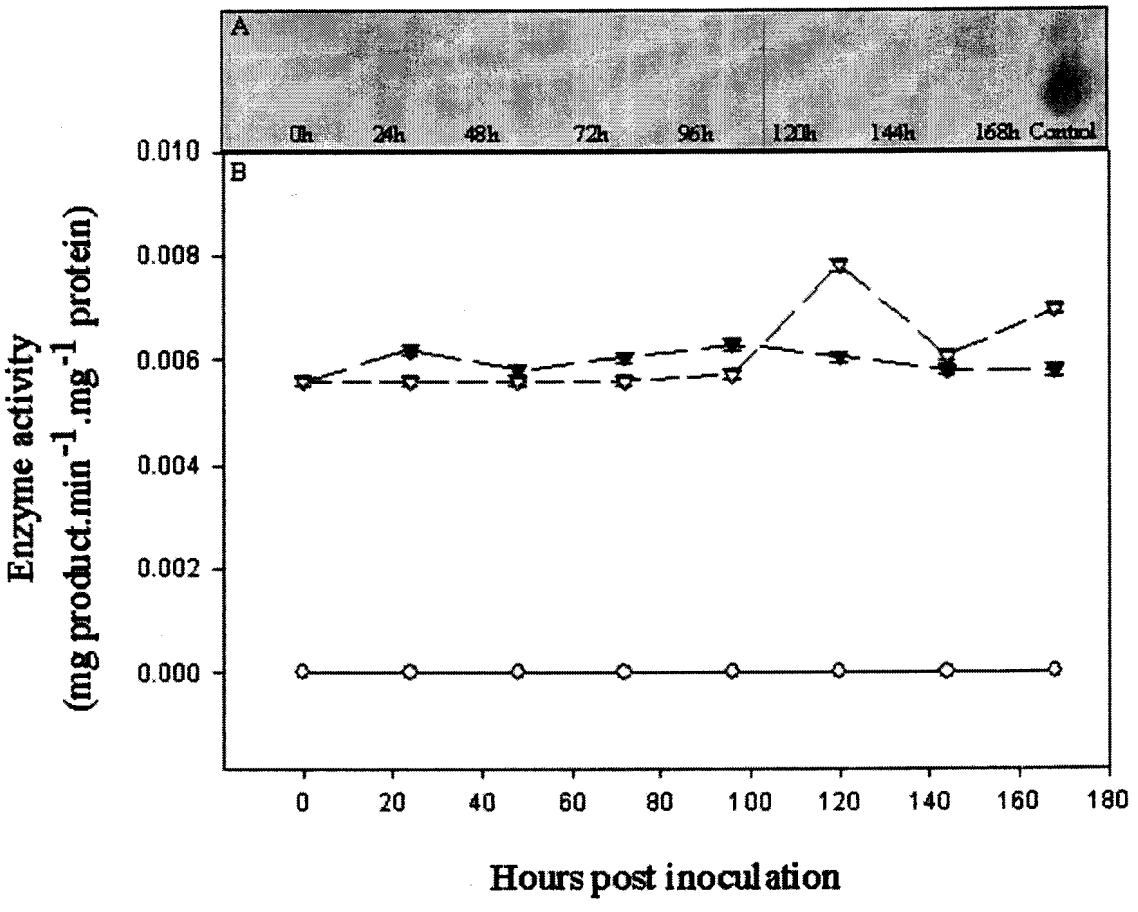


Figure 1: Detection of chitinases through Western Blot analysis and enzyme assays. (A) Western Blot analysis displaying no detectable levels of chitinase in the inoculated samples as compared to wheat positive control. (B) ß-1,3-glucanase activity (∇-wounding control, ▼-inoculated plants) displays induction after both inoculation and wounding control. No chitinase activity (○-wounding control, ●-inoculated plants, indicators overlap and only the ○ are visible) is detectable in either the wounding control or inoculated samples.

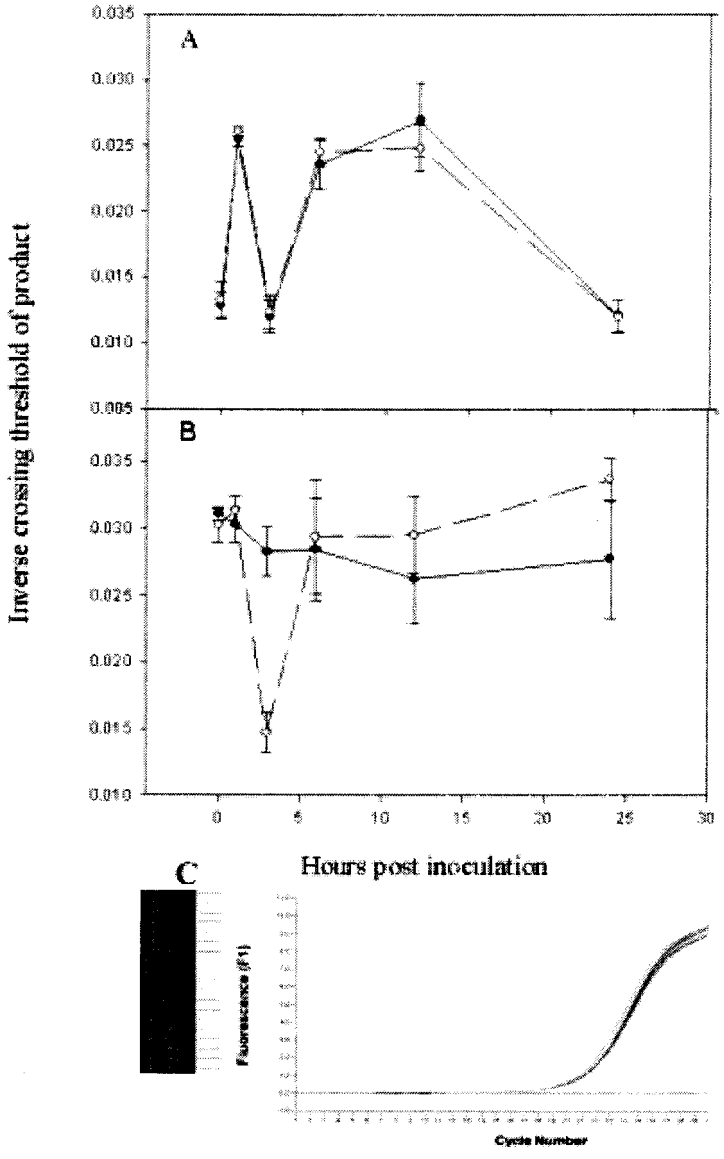


Figure 2: Real time quantitative RT-PCR analysis of the *PSCHI4* and *LP6* homologues in *P. patula*. (A) Analysis of *PSCHI4* homologue. (B) Analysis of *LP6* homologue. (C) Fluorescence of Rubisco RT-PCR product to indicate equal loading of template cDNA. (○) Wounding control, and (●) Inoculated sample.

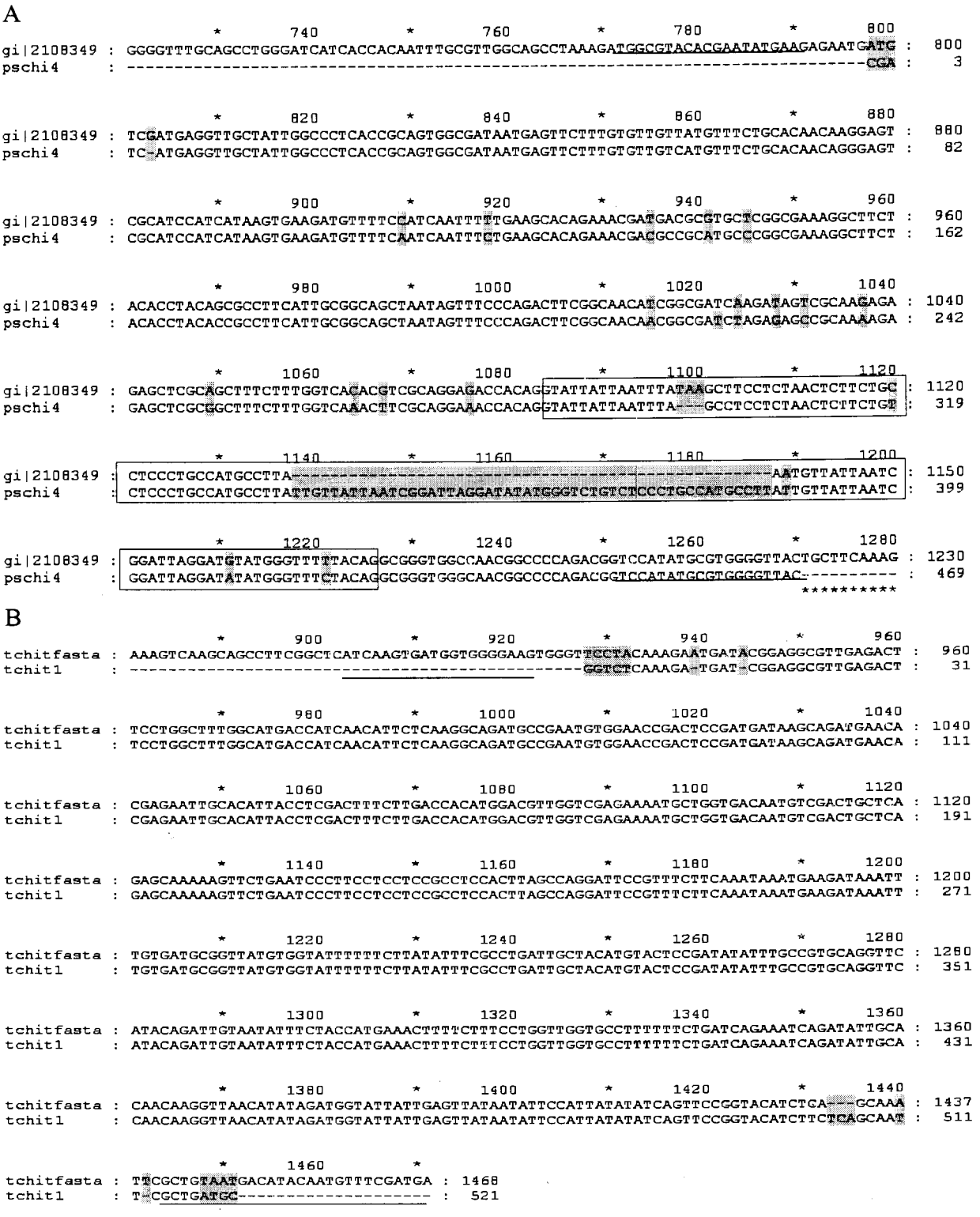


Figure 3: Pairwise alignment between the *P. patula* homologues and (A) *P. strobus PSCHI4* gene, and (B) *P. taeda LP6* gene. The boxed-in area represents an intron and differences between sequences are shaded. Underlined text indicates primer-binding sites.

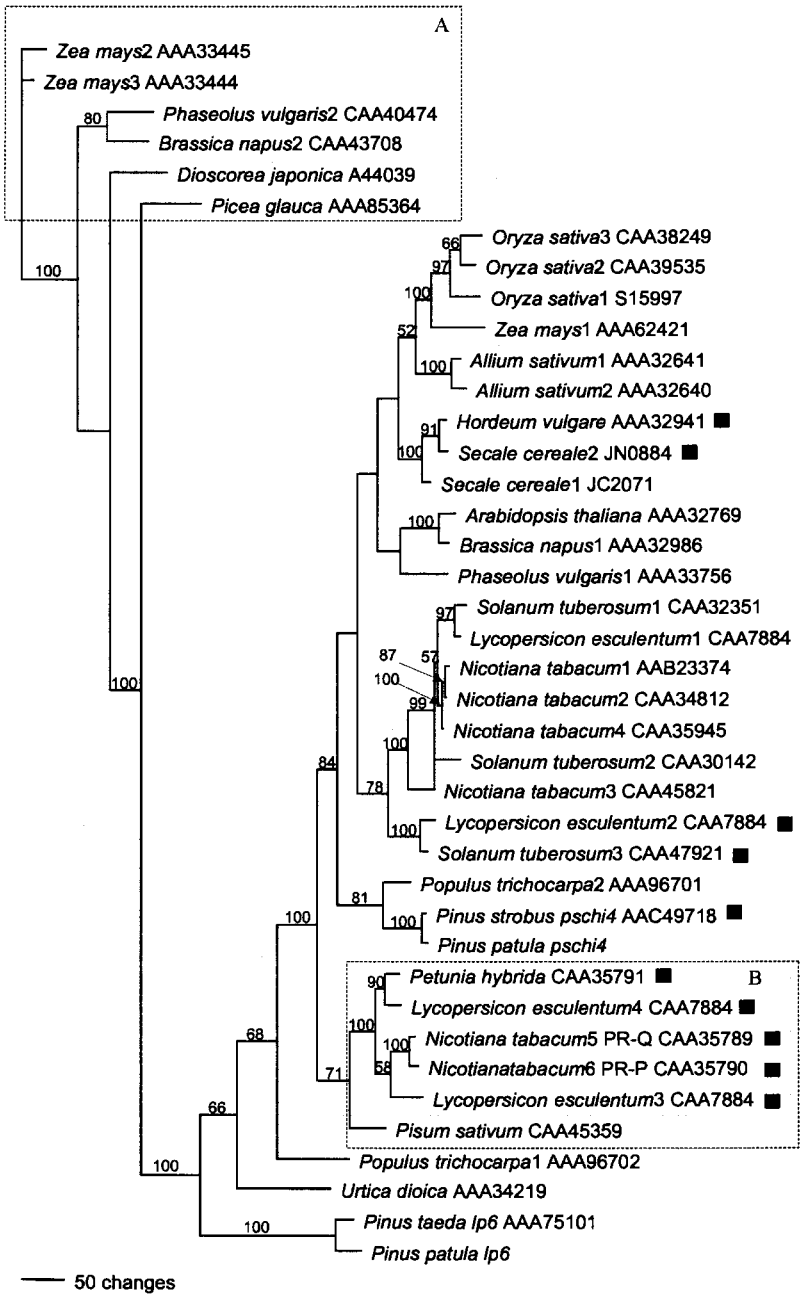


Figure 4: Neighbour-joining analysis of class I and class II chitinases. Class I-L chitinases (A) and class II-L chitinases (B) are boxed. (■) Designates all class II chitinases. The *PSCHI4* gene groups basal to the class II-L clade, whereas the *LP6* gene groups with *U. dioica*. Both the pine chitinases are by association of the high molecular weight subclass of chitinase. Bootstrap values are indicated above the branches.

Table 1: Chitinases used in the phylogenetic analysis.

Accession Number*	Organism	Name as used in Analysis
AAA32641	*Allium sativum*	*Allium sativum* 1
AAA32640	``	*Allium sativum* 2
AAA32769	*Arabidopsis thaliana*	*Arabidopsis thaliana*
AAA32986	*Brassica napus*	*Brassica napus* 1
CAA43708	``	*Brassica napus* 2
A44039	*Dioscorea japonica*	*Dioscorea japonica*
AAA32941	*Hordeum vulgare*	*Hordeum vulgare*
CAA78845	*Lycopersicon esculentum*	*Lycopersicon esculentum* 1
CAA78843	``	*Lycopersicon esculentum* 2
CAA78846	``	*Lycopersicon esculentum* 3
CAA78844	``	*Lycopersicon esculentum* 4
AAB23374	*Nicotiana tabacum*	*Nicotiana tabacum* 1
CAA34812	``	*Nicotiana tabacum* 2
CAA45821	``	*Nicotiana tabacum* 3
CAA35945	``	*Nicotiana tabacum* 4
CAA35789	``	*Nicotiana tabacum* 5 PR-Q
CAA35790	``	*Nicotiana tabacum* 6 PR-P
S15997	*Oryza sativa*	*Oryza sativa* 1
CAA39535	``	*Oryza sativa* 2
CAA38249	``	*Oryza sativa* 3
CAA35791	*Petunia hybrida*	*Petunia hybrida*
AAA33756	*Phaseolus vulgaris*	*Phaseolus vulgaris* 1
CAA40474	``	*Phaseolus vulgaris* 2
AAC49718	*Picea glauca*	*Picea glauca*
CF660335	*Pinus patula*	*Pinus patula pschi4*
CF660336	``	*Pinus patula lp6*
AAA85364	*Pinus strobus*	*Pinus strobus pschi4*
AAA75101	*Pinus taeda*	*Pinus taeda lp6*
CAA45359	*Pisum sativum*	*Pisum sativum*
AAA96702	*Populus trichocarpa*	*Populus trichocarpa* 1
AAA96701	`	*Populus trichocarpa* 2
JC2071	*Secale cereale*	*Secale cereale* 1
JN0884	“	*Secale cereale* 2
CAA32351	*Solanum tuberosum*	*Solanum tuberosum* 1
CAA30142	``	*Solanum tuberosum* 2
CAA47921	``	*Solanum tuberosum* 3
AAA34219	*Urtica dioica*	*Urtica dioica*
AAA62421	``	*Zea mays* 1
AAA33445	``	*Zea mays* 2
AAA33444	``	*Zea mays* 3

*Class II chitinases are underlined and designated with a solid square in the Fig. 3

References

1. **Altschul SF, Thomas LM, Alejandro AS, Jinghui Z, Zheng Z, Webb M, Lipman DJ** (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Research **25:** 3389-3402

2. **Araki T, Torikata T** (1995) Structural classification of plant chitinases: two classes in class I and class II chitinases. Bioscience, Biotechnology, and Biochemistry **59:** 336-338

3. **Barrows-Braoddus J, Dwinell LD** (1983) Histopathology of *Fusarium moniliforme* var. *subglutinans* in four species of southern pines. Phytopathology **73:** 882-889

4. **Boller T, Gehri A, Mauch F, Vögeli U** (1983) Chitinase in bean leaves: induction by ethylene, purification, properties, and possible function. Planta **157:** 22-31

5. **Botha A-M, Nagel MAC, van der Westhuizen AJ, Botha FC** (1998) Chitinase isoenzymes in near-isogenic wheat lines challenged with the Russian wheat aphid, exogenous ethylene, and mechanical wounding. Botanical Bulletin of Academia Sinica **39:** 99-106

6. **Bradford MM** (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Analytical Biochemistry **72:** 248-253

7. **Britz H, Coutinho TA, Gordon TR, Wingfield MJ** (2001) Characterisation of the pitch canker fungus, *Fusarium circinatum*, from Mexico. South African Journal of Science **67:** 609-614

8. **Britz H, Coutinho TA, Wingfield MJ, Marasas WFO** (2002) Validation of the description of *Gibberella circinata* and morphological differentiation of the anamorph *Fusarium circinatum*. Sydowia **54:** 9-22

9. **Cabib E, Bowers B** (1971) Chitin and yeast budding: Localization of chitin in yeast bud scars. Journal of Biological Chemistry **246:** 152-159

10. **Chang S, Puryear J, Cairney J** (1993) A simple and efficient method for isolating RNA from pine trees. Plant Molecular Biology Reporter **11:** 113-116

11. **Chang S, Puryear J, Dias MADL, Newton RJ, Funkhouser EA, Cairney J** (1996) Gene expression under water deficit in loblolly pine (*Pinus taeda*): Isolation and characterization of cDNA clones. Physiologia Plantarum **97:** 139-148

12. **Chang S, Puryear J, Funkhowser EA, Newton RJ, Cairney J** (1996) Cloning of a cDNA for a chitinase homologue which lacks chitin-binding sites and is down-regulated by water stress and wounding. Plant Molecular Biology **31:** 693-699

13. **Collinge DB, Kragh KM, Mikkelsen JD, Nielsen KK, Rasmussen U, Vad K** (1993) Plant chitinases. Plant Journal **3:** 31-40

14. **Davis JM, Wu H, Cooke JEK, Reed JM, Luce KS, Michler CH** (2002) Pathogen challenge, salicylic acid, and jasmonic acid regulate expression of chitinase gene homologs in pine. Molecular Plant-Microbe Interactions **15:** 380-387

15. **Gordon TR, Storer AJ, Wood DL** (2001) The pitch canker epidemic in California. Plant Disease **85:** 1128-1139

16. **Hackman RH, Goldberg M** (1964) New substrates for use with chitinases. Analytical Biochemistry **8:** 397-401

17. **Hepting GH, Roth ER** (1946) Pitch canker, a new disease of some southern pines. Journal of Forestry **44:** 742-744

18. **Linthorst HJ, van Loon LC, van Rossum CM, Mayer A, Bol JF, van Roekel JS, E.J. M, Cornelissen BJ** (1990) Analysis of acidic and basic chitinases from tobacco and petunia and their constitutive expression in transgenic tobacco. Molecular Plant-Microbe Interactions **3:** 252-258

19. **Nirenberg HI, O'Donnell K** (1998) New *Fusarium* species and combinations within the *Gibberella fujikuroi* species complex. Mycologia **90:** 434-458

20. **Samac DA, Hironaka CM, Yallaly PE, Shah DM** (1990) Isolation and characterization of the genes encoding basic and acidic chitinase in *Arabidopsis thaliana*. Plant Physiology **93:** 907-914

21. **Samac DA, Shah DM** (1991) Developmental and pathogen-induced activation of the *Arabidopsis* acidic chitinase promoter. Plant Cell **3:** 1063-1072

22. **Stintzi A, Heitz T, Prasad V, Wiedemann-Merdinoglu S, Kauffmann S, Geoffroy P, Legrand M, Fritig B** (1993) Plant 'pathogenesis-related' proteins and their role in defense against pathogens. Biochimie **75:** 687-706

23. **Swofford D** (2000) PAUP*, Phylogenetic analysis using parsimony. MA: Sinauer, Sunderland

24. **Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG** (1997) The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Research **25:** 4876-4882

25. **Viljoen A, Wingfield MJ** (1994) First report of *Fusarium subglutinans* f. sp. *pini* on pine seedlings in South Africa. Plant Disease **78:** 309-312

26. **Wingfield MJ, Coutinho TA, Roux J, Wingfield BD** (2002) The future of exotic plantation forestry in the tropics and southern Hemisphere: Lessons from pitch canker. South African Forestry Journal **195:** 79-82

27. **Wingfield MJ, Jacobs A, Coutinho TA, R. A, Wingfield BD** (2002) First report of the pitch canker fungus, *Fusarium circinatum*, on pines in Chile. Plant Pathology **51:** 397

28. **Wu H, Echt CS, Popp MP, Davis JM** (1997) Molecular cloning, structure and expression of an elicitor-induced chitinase gene from pine trees. Plant Molecular Biology **33:** 978-987

29. **Wu H, Michler CH, LaRussa L, Davis JM** (1999) The pine *Pschi4* promoter directs wound-induced transcription. Plant Science **142:** 199-207

CHAPTER 3

Transcript derived fragments from a compatible interaction between *Pinus patula* and the pitch canker pathogen, *Fusarium circinatum*.

Prepared for submission to Molecular Plant Pathology as:
Transcript derived fragments from a compatible interaction between *Pinus patula* and the pitch canker pathogen, *Fusarium circinatum*.

Abstract

The pitch canker pathogen, *Fusarium circinatum*, represents one the most important threats to exotic plantation forestry in the southern hemisphere. In South Africa, *Pinus patula*, which is the most widely planted species, is highly susceptible to infection by this pathogen. Despite its importance, very little is known regarding the molecular responses of trees to infection by *F. circinatum*. In this study we considered the interaction between *F. circinatum* and *P. patula* using suppression subtractive hybridisation (SSH) to enrich for genes expressed 24 hrs after inoculation. The induction profiles of some of the transcript derived fragments (TDF) were followed through real time quantitative RT-PCR. The TDFs from pine had homology to several stress linked genes. One fragment with homology to phytocyanin proteins, that are associated with primary defence responses in plants, was selected for further characterisation. The induction profiles for this sequence were faster and reached higher levels when compared to plants mock inoculated as controls. There were also significant differences in plants that were more susceptible to infection than the more tolerant plants. Several of the isolated TDFs were of fungal origin and one of these shared homology with *GRG1*, a fungal gene known to be up-regulated under conditions of carbon starvation and when a pathogen is infecting its host. Although all of the sequences identified in this study have been previously linked to stress, most have not been linked with specific plant-pathogen interactions. The interaction between the susceptible *P. patula* and the pitch canker fungus provides a unique opportunity to study both the fungal infection and the host responses.

Keywords

Suppression subtractive hybridisation, pitch canker fungus, stress response

Nucleotide and/or amino acid sequence data are to be found at GenBank as Accession No. CB074804- CB074822, CF660355.

Introduction

Pitch canker is the most important disease of *Pinus* spp. in South African seedling nurseries [42]. *Fusarium circinatum* was first detected in South Africa in the early 1990s [39]. The origin of the pitch canker fungus is unknown, but it is well known in the south eastern United States and is possibly native to Mexico [18]. The disease is found on several pine species, but seems to be the most severe on *P. elliottii*, *P. patula* and *P. radiata*, which forms the basis of the South African forestry industry [4; 39; 19; 34].

The most distinctive symptoms of pitch canker on mature trees are sunken resinous cankers found on branches and stems [21]. Infection results in copious resin exudation and the so-called pitch penetrate and stain the wood surrounding the canker. Cankers can girdle and kill branches and when these occur on the main stems whole trees are killed [34]. *Fusarium circinatum* can also infect and destroy the flowers, seeds and mature cones [13; 34]. Pitch production is stimulated by the pathogen and this appears not to be detrimental to the survival of the fungus [21; 4; 11]. On seedlings the symptoms are wilting within four weeks and these seedlings usually die within six weeks.

Genetic responses of conifers to infection by pathogens in endemic forest pathosystems appears to be quantitative [41]. To identify quantitative resistance usually requires the use of inbred lines. This has been extremely difficult to achieve in conifers due to the long life cycles and the high genetic load carried by their genomes [41; 12]. However, naturally occurring pitch canker infections suggests that the response to this disease is under genetic control [4; 14; 5].

Artificial inoculations with *F. circinatum* in the field and greenhouse reflect those of natural infections [5]. This suggests that selection for disease resistant material could be achieved at the seedling stage [4]. However, studies have also shown that there is a correlation between the size of the seedling and the tree's ability to tolerate the pathogen [22; 18]. Therefore, the interaction between *F. circinatum* and *Pinus* spp. could be used as a model to study the interaction between conifers and pathogens [6].

Repeated inoculations of trees with the pitch canker fungus results in increased disease tolerance for *P. radiata* [17; 6]. This tolerance appears not to be due to systemic acquired

resistance (SAR) but more likely through systemic induced resistance (SIR) [6]. Systemic induced resistance differs from the SAR in that no chemical has been associated with this induction [1]. In *Picea abies* similar results were observed after infection with *Ceratocystis polonica*, but the acquired resistance appeared localised and due to anatomical changes [26]. The responses to pitch canker in pines seem to follow known routes of host responses in other plants and could provide an opportunity to study these interactions at the molecular level.

The most important pine species grown in South Africa is *P. patula*. This species is also amongst the most susceptible to *F. circinatum*. Almost nothing is known regarding the interaction between the fungus and *P. patula*, particularly at the molecular level. Therefore, the responses by *P. patula* 24-hrs post inoculation with *F. circinatum* were studied using suppression subtractive hybridisation (SSH). Some of the isolated transcript derived fragments (TDF) were further evaluated in *P. patula* plants with different tolerance levels.

Materials and Methods

Plant Material and Fungal Inoculations

The plants used in this study were seedlings obtained from open-pollinated *P. patula* trees supplied by Komatiland Forests Research, Sabie, South Africa. These seedlings were cloned for multiple inoculation trials and the clones were planted in clonal hedges under shade netting on the research farm of the University of Pretoria.

An isolate of *F. circinatum* (CMW13229), previously shown to be highly pathogenic to pine trees [39], was chosen for this study. This isolate is maintained in the culture collection of the Forestry and Agricultural Biotechnology Institute (FABI, University of Pretoria, Pretoria, South Africa). It was maintained on half strength (20 g/l) Potato Dextrose Agar (PDA) (NT-Merck) during this study. Conidia were collected from cultures of the test isolate and aliquots were stored at – 70°C to maintain pathogenicity of the isolate for successive inoculation experiments. The three *P. patula* clones displaying highest, and the three displaying the lowest tolerance levels in an initial inoculation trial performed for Komatiland Forestry Research, were selected for RNA extractions (Fig. 1). The highest level of tolerance was designated tolerant and the lowest tolerance level designated susceptible for this study. For inoculations, the fungus was cultured for one week on half strength PDA plates. Fungal plugs

(5 x 2 mm) taken from the actively growing margins of the isolate were then inserted into wounds of equal size on stems or side branches of seedlings. The inoculated wounds were sealed with laboratory film to prevent contamination and desiccation of the inoculum. Control plants were inoculated with sterile PDA plugs.

RNA Extraction

Directly before infection RNA was isolated from all plants, this was repeated 24-hrs post inoculation. RNA was isolated with a modified method of Hughes and Galau [23]. Phloem tissue was ground in liquid nitrogen and transferred into a 50 ml Oakridge tube containing 10 ml cold extraction buffer (200 mM Tris-HCl pH 8.5, 300 mM LiCl, 10 mM Na_4EDTA, 5 mM Thiourea, 1 mM Aurintricarboxylic acid, 10 mM DTT, 1.5 % (w/v) N-Lauroylsarcosine, 1.0 % (v/v) Nonidet P-40, 1.0 % (w/v) Sodium deoxycholic acid). This was homogenised using a vortex mixer and rapidly frozen at –80°C. The extraction mixture was then thawed to 4°C in a 37°C water bath and 1/3 (w/v) 8.5 M KAc was added and mixed. After 15 min on ice, the extraction mixtures were centrifuged and the supernatant passed through 1 layer PES 40/22 membrane (Polymon). After addition of 3.3 M NaAc to 1/9 of the total volume, the RNA was precipitated with 0.5 volumes acetone for 2 hrs at –20°C, centrifuged at 14 000 rpm and the supernatant discarded. The pellets were dried under reduced pressure, resuspended in TE buffer (10 mM Tris-HCl pH 7.6, 1 mM EDTA), and precipitated overnight with 2 M LiCl at 4°C. After centrifugation the pellet was dried and dissolved in 400 µl TE. To this was added 1.5 volumes of 5 M KAc, incubated for 3 hours on ice and centrifuged. The pellet was resuspended in 400 µl TE and extracted with chloroform: butanol (4:1). The upper phase was precipitated for 2 hrs with 1/9 volume 3.3 M NaAc (pH 6.1) and two volumes ethanol. This was centrifuged and the pellet resuspended in 60 µl of ddH_2O and stored at –80°C.

The mRNA fraction was separated from the total RNA with Oligotex™ beads (Qiagen) and cDNA synthesised using the cDNA Synthesis System (Roche). For real time quantitative RT-PCR, RNA was isolated at 0, 1, 3, 6, 12, and 24 hrs post inoculation with the method of Chang and co-workers [7]. RNA from *F. circinatum* used in the SSH reactions was extracted with TRIzol™ reagent (GibcoBRL) following the manufacturer's specifications.

Suppression Subtractive Hybridisation

Suppression subtractive hybridisation (SSH) was performed with the PCR-Select™ cDNA Subtraction Kit (Clontech). cDNA from the samples was designated as tester and the cDNA that was used for subtraction, the driver. Four subtractions were performed. In the first of these the cDNA from the inoculated susceptible individuals was used as driver against cDNA from the inoculated tolerant trees. In the second subtraction, cDNA from the un-inoculated susceptible trees was used as driver against cDNA from the inoculated tolerant individuals. For the third subtraction cDNA from the un-inoculated tolerant trees were used as driver against cDNA from the inoculated tolerant trees and in the fourth subtraction, cDNA from the un-inoculated susceptible trees was used as driver against cDNA from the inoculated susceptible trees. Pools of cDNA from *F. circinatum* grown in culture were added to all the driver samples to eliminate excess fungal sequences. After subtraction and two successive rounds of PCR, the amplicons were size fractionated on a 2 % agarose gel containing 0.5 µg/ml ethidium bromide, extracted with the MinElute™ Reaction Cleanup kit (Qiagen), and cloned into pGEM-T Easy vector (Stratagene). The efficiency of the subtractions was tested through Southern blot analysis with DIG label (Roche) [38].

Screening and analyses of fragments

All colonies containing inserts were amplified by PCR. These products were dried in the 96-well plates and resuspended in 50 % (v/v) dimethyl sulfoxide (DMSO) [20]. The denatured products were blotted in duplicate onto nylon membranes with a 96-pin replicator (Nunc). One set of membranes was hybridised with DIG-labelled (Roche) cDNA from *F. circinatum* and the other with probes consisting of different combinations of the cDNA from the different trees. Clones that hybridised with cDNA from *F. circinatum* and the susceptible pools of cDNA were discarded. All clones that displayed differential hybridisation were selected, sequenced and analysed using the BLAST algorithm at GenBank (http://www.ncbi.nlm.nih.gov) [2]. These transcript-derived fragments (TDFs) were further analysed with PCR and Southern blots to determine whether they were of fungal or plant origin. The TDFs were analysed using real time quantitative RT-PCR, as the amounts of RNA sampled were insufficient to make use of Northern blots. Primers were designed (Table 1) to amplify the different fragments with the LightCycler® using the LightCycler-DNA FastStart Master SYBRGreen I kit (Roche). These PCRs were conducted on first strand cDNA from RNA that was reverse transcribed with AMV reverse transcriptase (Promega).

Three sample sets were analysed over time and sample extraction. These sample sets were inoculated tolerant and susceptible trees, and the control inoculations. A PCR with primers specific for the large subunit of coniferous Rubisco were also run to prove equal loading of sample cDNA.

Results

Suppression Subtractive Hybridisation

The SSH reaction yielded several well-defined PCR amplicons (Fig. 2). The negative control contained some smearing, but no dilution was done on this sample after the primary PCR, and included 58 consecutive PCR cycles. The efficiency of the subtractions was verified through Southern blot analysis and a difference in intensity between the lanes containing tester and driver cDNA was observed (Not shown). The cloned PCR products that displayed induction were selected for further study after differential screening. The 16 transcript derived fragments obtained from this study are summarized in Tables 2 and 3. From the 16 different sequences identified most could not be assigned a putative function. Six of these fragments shared homology with retrotransposable elements and were discarded from the study. One TDF appeared to be a chimeric sequence composed of an unknown sequence linked to 114 bp of a 25S ribosomal fragment. This fragment was not used in further analysis as the unknown part of the sequence only contained remnants of the *Rsa*I restriction site on one side.

Analyses of TDFs

Several TDFs (CB074810, CB074811, CB074820, and CB074823) had significant homology to unknown *Pinus* ESTs (E-values in excess of e^{-42}), but a putative function could not be assigned to them. These TDFs were highly homologous to ESTs that fall under the 'Molecular Basis of Wood Formation in the Pine Megagenome'-project undertaken by the Forest Biotechnology group at North Carolina State University.

Fragment CF660355 shared homology to UDP-glucose pyrophosphorylase at the genomic and amino acid level. This fragment was isolated from an SSH subtraction (Chapter 4) to enrich for fungal transcripts expressed during the infection process. In the study from Chapter

4 we were interested in the isolation of fungal transcripts and therefore, the *Pinus* fragments were included in this study. No induction analysis was performed on this fragment.

Fragment CB074822 had homology to a blue copper protein precursor (AAF18529) (E-value = $4e^{-14}$) and also displayed high similarity to a *P. taeda* EST from dbEST at GenBank. This protein resides in the phytocyanin protein family. The induction of this TDF was studied through real time quantitative RT-PCR and induction was evident after inoculation (Fig. 4). This induction differed between the control plants, susceptible inoculated plants, and tolerant inoculated plants. The tolerant individuals displayed an induction between 1-3 hrs, whereas the susceptible individuals and control display an induction only between 3-6 hrs (Fig. 4A). Induction of this TDF in the tolerant individuals reached a maximum after 12 hrs, whereas, in the other two samples there was a decrease in the expression of the TDF. Differences of expression between the wounding control and the susceptible plant were evident between 6-12 hrs after inoculation with the fungus. In the control inoculation there was a decrease in the expression of the TDF that was not evident in the two inoculated samples. At 24 hrs there was no difference between the levels of the TDF in the three samples.

Equal loading of cDNA template was confirmed by amplification of a segment from the Rubisco gene (Fig. 4B). Rubisco was chosen for amplification over the ribosomal internal transcribed spacer (ITS) region, as these primers also amplified cDNA that was of fungal origin and these primers are therefore not suitable for quantification of pine cDNA concentrations. This was also reported by Saar *et al.* [36].

One TDF (CB078409) was isolated through both the subtractions between the un-inoculated and inoculated susceptible as opposed to the inoculated tolerant plants. This TDF shared homology with phytochelatin synthase from *Typhia latifolia* at the genomic level (E-value = $9e^{-31}$). With the quantitative RT-PCR, levels of the TDF were detected only in the tolerant inoculated and control inoculations samples. However, this TDF was not detected at all the time intervals. In the tolerant individual it was detected at 1 and 6 hrs after inoculation, and for the control inoculation at 1 and 24 hrs after inoculation. There was no difference between the levels detected in the samples.

Two TDFs (CB074812 and CB074818), were detected only in *F. circinatum* genomic DNA (Fig. 5). These fragments were isolated from the *F. circinatum* isolate used for the

inoculations. These sequences shared homology with folic acid synthase from *Neurospora crassa* (E = $2e^{-36}$) and the glucose repressible gene (*GRG1*) from *Xanthophyllomyces dendrorhous* (E = $3e^{-18}$), respectively. The time of induction of these sequences could not be determined as fungal cDNA was not present at all times of sampling. These sequences could be linked to genes that are expressed in the host-pathogen interaction as the subtraction reactions were fortified with pooled RNA samples from cultures of *F. circinatum*.

Discussion

In this study we used SSH to gain understanding of the interaction between *P. patula* and *F. circinatum* after infection has taken place. We were able to identify 16 TDFs that are up-regulated after pathogen infection. Of these fragments, two originated from the pathogen and the remaining from the host. Six of these fragments were of unknown function and had no homology to known proteins, although they displayed homology to unidentified ESTs from cDNA libraries that have been developed to study wood formation in *P. taeda*. Only one fragment was linked with glucose and carbon metabolism. Of the 16 TDFs, two were further analysed using quantitative RT-PCR to determine the differences in induction levels for the susceptible and tolerant individuals.

The fact that several TDFs identified in this study shared homology with ESTs identified in *P. taeda* libraries developed to study wood formation was of interest. The detection of these TDFs through SSH indicates that they are differentially expressed between the susceptible and tolerant individuals 24 hrs after inoculation. This could indicate that tolerant trees respond more rapidly to pathogen ingress than more susceptible trees. Wood formation has been linked to resistance of conifers to fungal invasion where lignification and shedding of infected parenchyma cells contribute to resistance [4; 26]. Histological studies on *P. taeda*, *P. elliottii*, *P. virginiana*, and *P. serotina* seedlings have likewise shown that several layers of parenchyma cells are produced and lignified in response to *F. circinatum* infection [4]. These observations were the same as those that occur in early wound reactions, where differentiation of the parenchyma cells is visible at an early stage [31]. In *Picea abies* this differentiation is visible after infection with *Ceratocystis polonica* between 6-9 days [31]. In *P. taeda* the formation of compression wood to bending stress was found to be visible after three days [44]. Transcript derived fragments linked to wood formation isolated in this study,

indicates that this process is detectable within 24 hrs after inoculation with a pathogen and might represent one of the mechanisms of pathogen resistance in *P. patula*. However, the true function of these ESTs in this interaction remains to be discovered.

The UDP-glucose pyrophosphorylase (CF660355) isolated in this study plays a role in cell metabolism under normal and cold stress conditions [9]. UDP-glucose pyrophosphorylase has been implicated in the synthesis of cellulose [25; 3; 9]. The role that this gene plays in the synthesis of cellulose can be in the formation of the new parenchyma cells in response to wounding. This gives an indication that UDP-glucose pyrophosphorylase is not only a part of the metabolic pathways under normal conditions, but also provides the plant with a means to respond to a plethora of stress situations. We were able to isolate this gene at a very early stage in the infection process. Several studies have shown that wood formation is linked to resistance responses in trees. However, these studies reported this formation after time intervals of days [31; 44]. In the present study we were able to isolate a putative gene that is linked with cellulose production and stress responses 24 hours after inoculation.

The identification of genes linked to deposition of lignin and new parenchyma cells in this study, might indicate that this is one of the major defence responses of *P. patula* to infection by the pitch canker fungus. This conclusion is supported by previous studies of the morphological changes that take place in several pine species (*P. elliottii*, *P. taeda*, *P. serotina*, and *P. virginiana*) after infection with *F. circinatum* [4]. This pathogen response is also consistent with responses of *Picea abies* to pathogens, pests and mechanical wounding by changes in the anatomical structure [26; 16; 31; 15]. This might be part of constitutive defence responses in conifers in response to environmental stress and pathogen attack.

The blue copper protein precursor (BCB) isolated in this study displays clear induction after artificial inoculation of *P. patula* with *F. circinatum*. This induction differed from that observed in the wounded control and was more pronounced in the tolerant clone than in the susceptible *P. patula* clone. This family of proteins is believed to represent redox-active proteins and they have been found only in plants [32]. They have been linked to the primary plant defence response [8; 32], oxidative stress responses [35], and in lignin formation in the cell wall [44]. The induction of BCB has also been observed in a susceptible and resistant *Arabidopsis* pathogen response to infection by *Xanthomonas campestris* pv. *campestris* [33]. Two previous studies on *P. taeda* subjected to stress conditions have resulted in the isolation

of members of this protein family [8; 44]. It seems likely that this gene homologue is also active in the defence reaction of *P. patula* after infection by *F. circinatum*.

Two separate subtraction reactions identified the same transcript (CB078409). This transcript potentially encodes for a protein which is homologous to a protein functioning as a light-dependent plastocyanin-ferredoxin oxidoreductase that transfers electrons from plastocyanin to ferredoxin [37]. This transcript also displayed homology at the nucleotide level to phytochelatin synthase. Phytochelatin synthases are responsible for the detoxification of heavy metals, especially cadmium [10]. There are two types of heavy metal detoxification proteins in plants. They are the phytochelatin synthases, which have not been found in animals, and the metallothioneins that are found in animals [10]. Kim *et al.* [24] identified metallothioneins as the most abundant genes identified in *Oryza sativa* after infection with *Magnaporthe grisea*. Therefore, it was interesting that two putative genes linked to metal stress were identified in *P. patula* after infection with *F. circinatum* and that they could be linked with other host-pathogen interactions.

The two fungal genes isolated in this study are necessary to the fungus for survival in the host plant. Folic acid is necessary for the production of purine nucleotides, methionine, and for other enzyme catalysed reactions involving the transfer, oxidation and reduction of single carbon units [27]. The *GRG1* gene is induced upon carbon starvation in *Neurospora crassa* [28]. *GRG1* is responsible for switching on many other genes that are necessary for the survival of the fungus in the host. The kinetics of induction for this gene has been studied for fungal isolates grown in culture. However, this has not been done for levels of the gene expressed during the host-pathogen interaction. Further analysis of *GRG1* in the host-pathogen interaction may provide a target gene for fungicides that can assist in the containment of fungal spread [28].

In this study we discarded the sequence that appears to be a chimeric cDNA clone because its formation and origin could not be determined. The formation of chimeric sequences during SSH has previously been documented and is expected to be quite low [43; 29]. These sequences are formed during the first subtraction reaction when a fragment from the driver cDNA pool link end to end with adaptor ligated cDNA from the tester pool. During the second subtraction reaction, the driver fragment can anneal to a complementary sequence from the tester cDNA that contains the second adaptor. After filling of the ends, this fragment

will be exponentially amplified during the suppression PCR. To screen for chimeric clones it is assumed that the *Rsa*I restriction site is restored. However, in this clone there was no restored *Rsa*I site. This could indicate that the origin of the unknown fragment is either from a non-full-length cDNA, or it is derived from the 5' end of a full-length cDNA. In both cases, these clones should not contain an *Rsa*I site to the 5' end.

In this study SSH was useful in detecting sequences early in the interaction between *P. patula* and *F. circinatum*. As expected, there were no transcripts linked to a definitive hypersensitive response (HR), as HR is not evident in this interaction as described by other researchers [19; 6]. All of the identified fragments were linked to stress responses except for UDP-glucose pyrophosporylase, which was associated with cell metabolism. Making use of the interaction between a susceptible *Pinus* sp. and the pitch canker fungus provided a unique opportunity to study both the fungal infection and the host response. To further identify genes of interest the use of other techniques such as cDNA-AFLP should complement SSH very well as additional time and sampling points can be accommodated.

Acknowledgements

We thank Komatiland Forests Research, the National Research Foundation and the Mellon Foundation for financial assistance and Komatiland Forests Research for providing the plant material used in this experiment.

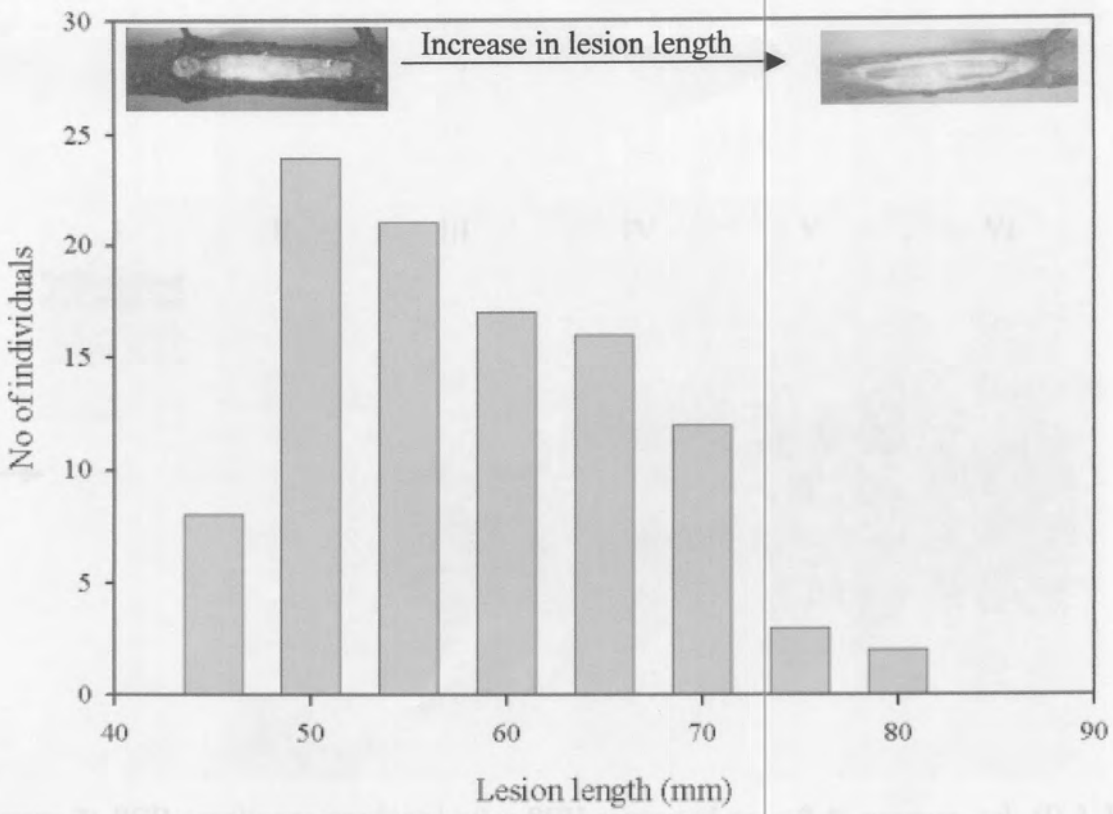


Figure 1: Variability in susceptibility of *P. patula* seedlings from one family to *F. circinatum*. Three trees from the ends of the bell curve were used for the SSH reactions. They displayed lesion sizes of 45 mm for the more tolerant (tolerant) plants and 80 mm for the less tolerant (susceptible) plants. Inserts indicate the differences in lesions observed.

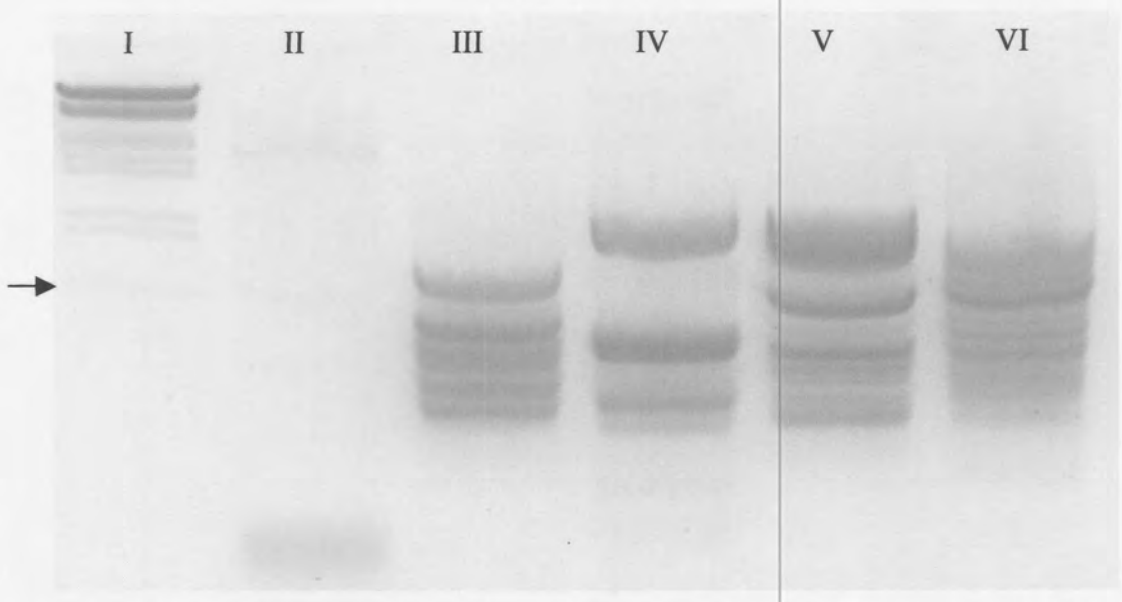


Figure 2: PCR-amplicons produced after SSH separated on a 2 % agarose gel. (I) λ DNA digested with *Eco*RI and *Hin*dIII, arrow indicates 564 bp fragment. (II) Negative control. Faint bands are present in the negative control, but this represents 58 rounds of PCR without dilution as in the subtracted samples. (III) Subtraction between inoculated susceptible and inoculated tolerant. (IV) Subtraction between un-inoculated susceptible and inoculated tolerant. (V) Subtraction between un-inoculated tolerant and inoculated tolerant. (VI) Subtraction between un-inoculated susceptible and inoculated susceptible.

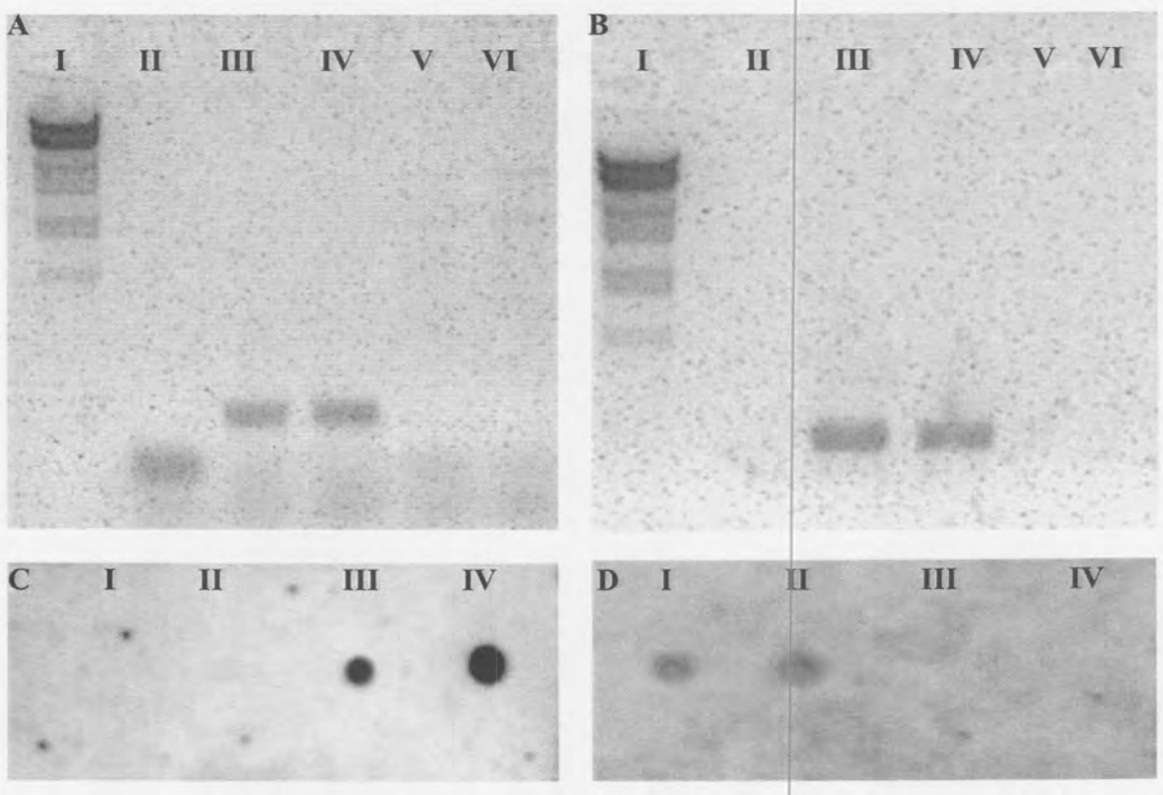


Figure 3: Induced fragments screened with PCR and Southern blot to determine origin. (A) Fragments CB078409 and CB074822 amplified from *Pinus* DNA (Lanes III and IV) and no amplification from *Fusarium* DNA (Lanes V and VI). (B) Fragments CB074812 and CB074818 amplified from *Fusarium* DNA (Lanes III and IV) and no amplification from *Pinus* DNA (Lanes V and VI). In both A and C, Lanes I and II are marker III and negative control (C) Southern blot analysis with *Pinus*-labelled DNA of CB074812 (I), CB074818 (II), CB078409 (III), and CB074822 (IV) to indicate that CB078409 (III) and CB074822 (IV) are from *Pinus patula*. (D) Southern blot analysis with *Fusarium*-labelled DNA of CB074812 (I), CB074818 (II), CB078409 (III), and CB074822 (IV) to indicate that CB074812 (I) and CB074818 (II) are from *Fusarium circinatum*.

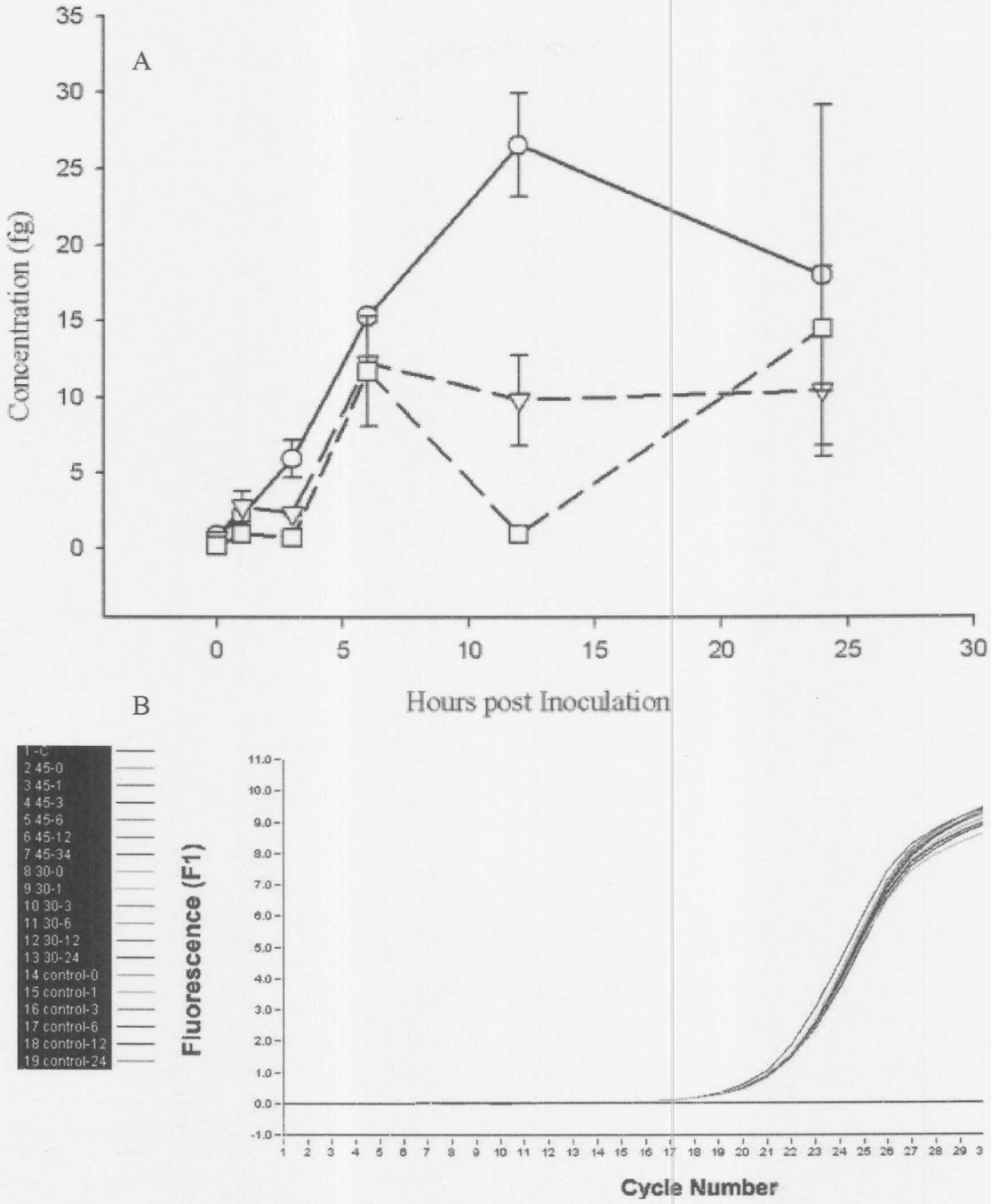


Figure 4: (A) Real time quantitative analysis of CB074822 displaying homology to a small blue copper protein. The student's t-test (p=0.013) was used to test if the data from different runs can be combined [30]. (○) Inoculated tolerant plant, (▽) Inoculated susceptible plant (□) Wounding control. (B) RT-PCR with conifer specific Rubisco primers to determine equal loading of cDNA samples. The Y-axis reflects the relative fluorescence of the PCR product. The X-axis indicates the cycle number when the PCR product crosses the threshold value and comes out of the background.

Table 1: Primers used in this study.

Name	TDF	Sequence (5'-3')	T_M (°C)	Amplicon length (bp)	Reference
RubiscoF	--	AATGTCTCCGTGGTGGACTC	55	134	This study
RubiscoR	--	TAATTTCACCCGTCTCAGCC	``		``
A85-1	CB078409	CATCTTGGATTTGGTCTCCC	60	152	``
A85-2	``	GGTGAATCTCGACCCATCTC	``		``
B12-1	CB074818	CAGGCTACCTCTGGTGCTTC	59	133	``
B12-2	``	TCTTGTCGTGAGTGGTCTCG	``		``
B51-1	CB074822	CATACACCAGTGGCAACGAC	60	243	``
B51-2	``	GAGTGGTGGATGGTGGAGAG	``		``
ITS3	--	GCATCGATGAAGAACGCAGC	54	355-450	[40]
ITS4	--	TCCTCCGCTTATTGATATGC	``		[40]

Table 2: Summary of subtractions performed.

	Subtraction 1	Subtraction 2	Subtraction 3	Subtraction 4
Driver cDNA	Inoculated susceptible	Un-inoculated susceptible	Un-inoculated tolerant	Un-inoculated susceptible
Tester cDNA	Inoculated tolerant	Inoculated tolerant	Inoculated tolerant	Inoculated susceptible
Identified transcript derived fragments	CB074809 CB074812 CB074814 CB074816	CB074804 CB074806 CB074809 CB074810 CB074811 CB074815 CB074819 CB074820	CB074817	CB074818 CB074822

Table 3: Sequences identified through SSH from pines after infection with *F. circinatum*.

TDF No.	Database match[a]			
	BLASTN homology	E-value	BLASTX homology	E-value
CB074823	No hit; xylem side wood EST (*Pinus taeda*)	0.0	No hit	--
CB074821	AE012082, complete genome sequence (*Xanthomonas axonopodis*)	$1e^{-103}$	No hit	--
CB074812	No hit; BE60493, unknown EST (*Sorghum bicolor*)	0.0	CAB91336, folic acid synthase (*Neurospora crassa*)	$2e^{-36}$
CB074816	AY103740, mRNA sequence (*Zea mauys*)	$8e^{-28}$	BAB17749, coatmer ß-subunit (*Oryza sativa*)	$3e^{-65}$
CB074817	L47672, embryo abundant protein (*Picea glauca*)	$2e^{-29}$	AAB01567, embryo abundant protein (*Picea glauca*)	$3e^{-18}$
CB074818	AJ291824, glucose repressible gene (*Xanthophyllomyces dendrorhous*)	$1e^{-10}$	CAC24572, glucose repressible gene (*Xanthophyllomyces dendrorhous*)	$3e^{-18}$
CB074822	No hit; AW290812, xylem wood EST (*Pinus taeda*)	$2e^{-30}$	AAF18529, blue copper protein (*Arabidopsis thaliana*)	$4e^{-14}$
CB074811	No hit, BQ654489, xylem wood EST (*Pinus taeda*)	$3e^{-42}$	No hit	--
CB074819	AF216198, retrotransposon-like fragment (*Pinus taeda*)	$2e^{95}$	AAK52561, retroelement pol protein (*Oryza sativa*)	$4e^{-19}$
CB074820	No hit, BQ290639, xylem root wood EST	$1e^{-109}$	No hit	--
CB074804	AJ489202, ty1-copia-like pol pseudogene (*Beta nana*)	$1e^{-29}$	No hit	--
CB074809	AF308658, phytochelatin synthase (*Typha latifolia*)	$9e^{-31}$	No hit	--
CB074814	No hit	--	No hit	--
CB074810	No hit, BQ290639 xylem root wood EST (*Pinus taeda*)	$1e^{-104}$	No hit	--
CB074807	No hit, CD424352, salicylic acid treated EST (*Sorghum bicolor*)	$2e^{-58}$	No hit	--
CF660355[b]	AF425969, UDP-glucose pyrophosphorylase (*Amorpha fructicosa*)	$7e^{-19}$	AAL33919, UDP-glucose pyrophosphorylase (*Amorpha fructicosa*)	$8e^{-59}$

[a]Where no similarity (designated No hit;) was found on the non-redundant nucleic acid database, BLASTN was performed on EST database.

[b]This sequence was isolated in Chapter 4.

References

1. **Agrawal AA, Tuzun S, Bent E**, eds (1999) Induced plant defences against pathogens and herbivores. APS Press, St. Paul MN

2. **Altschul SF, Thomas LM, Alejandro AS, Jinghui Z, Zheng Z, Webb M, Lipman DJ** (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Research **25:** 3389-3402

3. **Amor Y, Haigler CH, Johnson S, Wainscott M, Delmer DP** (1995) A membrane-associated form of sucrose synthase and its potential role in synthesis of cellulose and callose in plants. Proceedings of the National Academy of Sciences, USA **92:** 9353-9357

4. **Barrows-Braoddus J, Dwinell LD** (1983) Histopathology of *Fusarium moniliforme* var. *subglutinans* in four species of southern pines. Phytopathology **73:** 882-889

5. **Blakeslee G, Rockwood D** (1998) Variation in resistance to pitch canker in slash and loblolly pines. *In* MC Devey M, Gordon T, ed, Current and potential impacts of pitch canker in radiata pine. Proceedings of the IMPACT Monterey workshop. Commonwealth Scientific and Industrial Research Organisation, pp 70-75

6. **Bonello P, Gordon TR, Storer AJ** (2001) Systemic induced resistance in Monterey pine. Forest Pathology **31:** 99-106

7. **Chang S, Puryear J, Cairney J** (1993) A simple and efficient method for isolating RNA from pine trees. Plant Molecular Biology Reporter **11:** 113-116

8. **Chang S, Puryear J, Dias MADL, Newton RJ, Funkhouser EA, Cairney J** (1996) Gene expression under water deficit in loblolly pine (*Pinus taeda*): Isolation and characterization of cDNA clones. Physiologia Plantarum **97:** 139-148

9. **Ciereszko I, Johansson H, Kleczkowski LA** (2001) Sucrose and light regulation of a cold-inducible UDP-glucose pyrophosphorylase gene via a hexokinase-independent and abscisic acid-insensitive pathway in *Arabidopsis*. Biochemical Journal **354:** 67-72

10. **Cobbett CS** (2000) Phytochelatins and their roles in heavy metal detoxification. Plant Physiology **123:** 825-832

11. **Davis JM, Wu H, Cooke JEK, Reed JM, Luce KS, Michler CH** (2002) Pathogen challenge, salicylic acid, and jasmonic acid regulate expression of chitinase gene homologs in pine. Molecular Plant-Microbe Interactions **15:** 380-387

12. **Dvorak WS, Jordan AP, Hodge GR, Romero JL** (2000) Assessing evolutionary relationships of pines in the *Oocarpae* and *Australes* subsections using RAPD markers. New Forests **20:** 163-192

13. **Dwinell D** (1998) Association of the pitch canker fungus with cones and seeds of pines. *In* MC Devey M, Gordon T, ed, Current and potential impacts of pitch canker in radiata pine. Proceedings of the IMPACT Monterey workshop. Commonwealth Scientific and Industrial Research Organisation, pp 35-39

14. **Dwinell LD, Barrows-Broaddus JB, Kuhlman EG** (1985) Pitch canker: A disease complex of southern pines. Plant Disease **69:** 270-276/277

15. **Franceschi VR, Krekling T, Christiansen E** (2002) Application of methyl jasmonate on *Picea abies* (Pinaceae) stems induces defense-related responses in phloem and xylem. American Journal of Botany **89:** 578-586

16. **Franceschi VR, Krokene P, Krekling T, Christiansen E** (2000) Phloem parenchyma cells are involved in local and distant defense responses to fungal inoculation or bark-beetle attack in Norway spruce (Pinaceae). American Journal of Botany **87:** 314-326

17. **Gordon T, Wikler K, Clark S, Storer A, Bonello P** (1998) Resistance to pitch canker in Monterey pine. *In* MC Devey M, Gordon T, ed, Current and potential impacts of pitch canker in radiata pine. Proceedings of the IMPACT Monterey workshop. Commonwealth Scientific and Industrial Research Organisation, pp 87-89

18. **Gordon TR, Storer AJ, Wood DL** (2001) The pitch canker epidemic in California. Plant Disease **85:** 1128-1139

19. **Guerra-Santos J** (1998) Pitch canker on Monterey pine in Mexico. *In* MC Devey M, Gordon T, ed, Current and potential impacts of pitch canker in radiata pine. Proceedings of the IMPACT Monterey workshop. Commonwealth Scientific and Industrial Research Organisation, pp 58-61

20. **Hegde P, Qi R, Abernathy K, Gay C, Dharap S, Gaspard R, Earle-Hughes J, Snesrud E, Lee N, Quackenbush J** (2000) A concise guide to cDNA microarray analysis. Biotechniques **29:** 548-562

21. **Hepting GH, Roth ER** (1946) Pitch canker, a new disease of some southern pines. Journal of Forestry **44:** 742-744

22. **Hodge GR, Dvorak WS** (2000) Differential responses of Central American and Mexican pine species and *Pinus radiata* to infection by the pitch canker fungus. New Forests **19:** 241-258

23. **Hughes DW, Galau G** (1988) Preparation of RNA from cotton leaves and pollen. Plant Molecular Biology Reporter **6:** 253-257

24. **Kim S, Ahn I-P, Lee Y-H** (2001) Analysis of genes expressed during rice-*Magnaporthe grisea* interactions. Molecular Plant-Microbe Interactions **14:** 1340-1346

25. **Kleczkowski LA** (1994) Glucose activation and metabolism through UDP-glucose pyrophosphorylase in plants. Phytochemistry **37:** 1507-1515

26. **Krokene P, Christiansen E, Solheim H, Franceschi VR, Berryman AA** (1999) Induced resistance to pathogenic fungi in Norway spruce. Plant Physiology **121:** 565-569

27. **Mathews C, van Holde KE** (1990) Biochemistry. The Benjamin/Cummings Publishing Company, Redwood City

28. **McNally MT, Free SJ** (1988) Isolation and characterization of a *Neurospora* glucose-repressible gene. Current Genetics **14:** 545-551

29. **Morales P, Thurston CF** (2003) Efficient isolation of genes differentially expressed on cellulose by suppression subtractive hybridization in *Agaricus bisporus*. Mycological Research **107:** 401-407

30. **Muller PY, Janovjak H, Miserez A, Dobbie D** (2002) Processing of gene expression data generated by quantitative real-time RT-PCR. Biotechniques **32:** 1372-1379

31. **Nagy NE, Franceschi VR, Solheim H, Krekling T, Christiansen E** (2000) Wound-induced traumatic resin duct development in stems of Norway spruce (Pinaceae): anatomy and cytochemical traits. American Journal of Botany **87:** 302-313

32. **Nersissian AM, Immoos C, Hill MG, Hart PJ, Williams G, Herrmann RG, Valentine JS** (1998) Uclacyanins, stellacyanins, and plantacyanins are distinct subfamilies of phytocyanins: plant-specific mononuclear blue copper proteins. Protein Science **7:** 1915-1929

33. **O'Donnell PJ, Schmelz EA, Moussatche P, Lund ST, J.B. J, Klee HJ** (2003) Susceptible to intolerance – a range of hormonal actions in a susceptible *Arabidopsis* pathogen response. Plant Journal **33:** 245-257

34. **Owen D, Adams D** (1998) Overview of pitch canker in California. *In* M Devey, C Matheson, T Gordon, eds, Current and potential impacts of pitch canker in radiata pine. Proceedings of the IMPACT Monterey workshop. Commonwealth Scientific and Industrial Research Organisation, pp 21-24

35. **Richards KD, Schott EJ, Sharma YK, Davis KR, Gardner RC** (1998) Aluminum induces oxidative stress genes in *Arabidopsis thaliana*. Plant Physiology **116:** 409-418

36. **Saar DE, Polans NO, Sørensen PD, Duvall MR** (2001) Angiosperm DNA contamination by endophytic fungi: Detection and method of avoidance. Plant Molecular Biology Reporter **19:** 249-260

37. **Salisbury FB, Ross CW** (1992) Plant physiology, Ed 4. Wadsworth Publishing Company, Belmont

38. **Sambrook J, Fritsch EF, Maniatis T** (1989) Molecular cloning: a laboratory manual, Ed 2. Cold Spring Harbor Press, Cold Spring Harbor

39. **Viljoen A, Wingfield MJ** (1994) First report of *Fusarium subglutinans* f. sp. *pini* on pine seedlings in South Africa. Plant Disease **78:** 309-312

40. **White TJ, Bruns T, Lee S, Taylor JW** (1990) Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. *In* MA Innis, DH Gelfland, JJ Sninsky, TJ White, eds, PCR protocols: a guide to methods and applications. Academic Press, Inc., San Diego, California, pp 315-322

41. **Wilcox PL, Amerson HV, Kuhlman EG, Liu B-H, O'Malley DM, Sederoff RR** (1996) Detection of a major gene for resistance to fusiform rust disease in loblolly pine by genomic mapping. Proceedings of the National Academy of Sciences, USA **93:** 3859-3864

42. **Wingfield MJ, Wingfield BD, Coutinho T, Viljoen A, Britz H, Steenkamp E** (1998) Pitch canker: A South African perspective. *In* MC Devey M, Gordon T, ed, Current and potential impacts of pitch canker in radiata pine. Proceedings of the IMPACT Monterey workshop. Commonwealth Scientific and Industrial Research Organisation, pp 62-69

43. **Zhang J, Underwood LE, D'Ercole AJ** (2000) Formation of chimeric cDNAs during suppression subtractive hybridization and subsequent polymerase chain reaction. Analytical Biochemistry **282:** 259-262

44. **Zhang Y, Sederoff RR, Allona I** (2000) Differential expression of genes encoding cell wall proteins in vascular tissues from vertical and bent loblolly pine trees. Tree Physiology **20:** 457-466

CHAPTER 4

Transcript derived fragments from *Fusarium circinatum* subtracted against other species in the *Gibberella fujikuroi* species complex.

Abstract

The production of mycotoxins by fungi can aid in plant pathogenesis. In some cases these toxins pose serious health risks for humans and animals. However, in certain *Fusarium* spp. mutating the genes linked to toxin production does not lessen the pathogenicity of the fungus. In this study we investigated the molecular characteristics that underlie pathogenicity in *Fusarium circinatum* using suppression subtractive hybridisation (SSH). *F. circinatum* was subtracted against a degenerate culture of the same fungus as well as, *F. guttiforme*, *F. subglutinans*, and *Fusarium opheodes prov. nom.* The transcripts that were identified were linked to several known stress and pathogen related genes. Several ESTs from plant cDNA libraries (*Triticum aestivum*, *Hordeum vulgare*, *Sorghum bicolor*, *Zea mays*, *Gossypium hirsutum*) were also identified as being of fungal origin, possibly from *Fusarium* spp. One transcript had high homology to a polyketide synthase gene (*PKS1*) found in *Gibberella fujikuroi*. This prompted us to study fumonisin production in *F. circinatum* using RT-PCR analysis with degenerate primers that detects *FUM5*, a polyketide synthase associated with fumonisin production. This transcript was detected in *F. circinatum* under a variety of conditions. However, an ELISA detected no fumonisin production. This indicates that *FUM5* plays another role in *F. circinatum*. The identification of these transcripts should provide a foundation for more intensive gene discovery studies in *F. circinatum* and other *Fusarium* spp.

Keywords

Polyketide synthase, fumonisin, culture degeneration

Nucleotide and/or amino acid sequence data are to be found at GenBank as Accession No. CF660337 – CF660355.

Introduction

Fusarium circinatum Nirenberg & O'Donnell (=*Fusarium subglutinans* (Wollenweb. & Reinking) Nelson, Toussoun & Marasas f. sp. *pini*) is the causal agent of pitch canker that damages several *Pinus* spp. Hepting first observed the occurrence of this fungus on *P. virginia* in the southern United States in 1946 [20]. Since then, this disease has spread to many parts of the world and has reached epidemic levels in several parts of the United States, where it is annihilating native, susceptible pine species [1; 17; 33; 44]. The pitch canker fungus was first detected in South African nurseries by the early 1990's and is now one of the most serious pathogens of pine seedlings in the country [46].

The *Fusarium subglutinans sensu lato* species complex accommodates a number of species that have previously been separated based on host preferences. Nirenberg and O'Donnell [31] elevated several of these *formae speciales* to species level. They included *F. guttiiforme* Nirenberg and O'Donnell (formerly *F. subglutinans* f. sp. *ananas*), the causal agent of disease in pineapple, and *F. circinatum*, which causes pitch canker on pines [38; 18]. An unknown saprophytic *Fusarium* sp. has recently been isolated from natural ecosystems in South Africa. The fungus belongs to the *F. subglutinans* complex and the name *F. opheodes* has been proposed for it (A. Jacobs, unpublished data). The close relatedness of these three species provides a unique opportunity to identify underlying pathogenicity factors in *Fusarium*.

Production of secondary metabolites or mycotoxins contributes to pathogenicity in various fungi [24; 28]. These mycotoxins can either be host-selective or non-selective. The host-selective toxins are produced primarily by species of two genera, *Alternaria* and *Cochliobolus*, whereas, several genera of fungi have been shown to produce the non-selective toxins [26; 47; 28]. In *Fusarium* spp. (section *Liseola*) the most important toxins are the fumonisins [16; 26]. Several other toxins are also produced by *Fusarium* spp. and have been linked to pathogenicity in plants [24].

The genes that encode for the production of secondary metabolites in fungi are usually found in clusters reminiscent of bacterial pathogenicity islands [48]. This implies that these gene clusters could be acquired and transferred via horizontal gene transfer, between different fungal species within and between genera. In *Fusarium*, this introduction could be simplified as this fungal genus is known for its promiscuous uptake of DNA in culture [35]. If this were to occur in nature, it would be important to identify and track the genes in various pathogenic *Fusarium* spp. Identification of these genes would also contribute to a better understanding of the plant-pathogen interactions that takes place after infection.

Fusarium spp. are known to be relatively unstable in culture with morphological, pathogenicity, and toxigenicity changes being observed [45; 49; 42; 8; 25; 14; 6; 50]. So-called degenerate cultures have also been reported for other fungi, specifically for *Aspergillus* species [7; 21]. The loss of toxigenicity in *Aspergillus flavus* was postulated to be as a lack of competition when the fungus is grown in pure culture [7]. However, Horn and Dorner [21] found that this is not so and adverse cultural and environmental conditions tends to stabilise this degeneration. This could be indicative of underlying change in gene expression linked with pathogenicity. If this were true, it would be less time consuming to study the genes that are changed due to serial transfers, than by making targeted mutants through gene knockout methods. However, mutating genes linked to toxin production in *Fusarium* has not resulted in reduced pathogenicity as demonstrated in other fungi (For review see [24]).

Identification of the underlying factors that play a role in pathogenicity of *Fusarium circinatum* and other species are important. Therefore, the aim of this study was to extend existing knowledge regarding the underlying differences between *F. circinatum* and other pathogenic and non-pathogenic *Fusarium* species. This was achieved by comparisons of transcripts between *F. circinatum* and a degenerate culture of the same fungus, as well as three other related *Fusarium* spp. including *F. guttiforme*, *F. subglutinans*, and *Fusarium opheodes prov. nom.* As it is known that degenerate cultures of *Fusarium* can have reduced pathogenicity, we studied the differences between a degenerated culture and the wild type isolate to enrich for any changes taking place in gene expression levels. This was done to ascertain if degeneration of these cultures could be detected at the molecular level. We further

studied the induction of *F. circinatum* genes early in the infection process by inoculating the fungus into *Pinus patula* and isolating transcripts that are expressed *in planta*.

Materials and Methods

Fungal Isolates, Culture Degeneration and Inoculations

The *Fusarium circinatum* strain used in this experiment was isolated from diseased *Pinus radiata* seedlings found in forestry nurseries in South Africa [46]. Spores were collected from cultures of the test isolate and stored at – 70°C to maintain pathogenicity. The other *Fusarium* spp. used in this study were *Fusarium guttiforme* from pineapple; *F. opheodes prov. nom*, commonly isolated from grass and reeds, *Fusarium subglutinans* from maize, and a serially sub-cultured strain of *F. circinatum* (Table 1).

In order to produce a degenerated of *F. circinatum* isolate, isolate CMW13229 was subjected to 15 rounds of sub-culturing, performed on malt extract agar (18 g/l, MEA) (NT-Merck), until the culture displayed morphological dissimilarities to the parent strain. These differences were observed on MEA, potato dextrose agar (18 g/l, PDA) (NT-Merck), and complete medium [11]. All strains are maintained in the *Fusarium* collection of the Forestry and Agricultural Biotechnology Institute (FABI, University of Pretoria, Pretoria, South Africa).

For inoculations, conidia of *F. circinatum* and the degenerate isolate were plated out onto half strength PDA and incubated at room temperature in the dark for 7 days. Mycelial plugs were removed from the edges of the cultures and inoculated into *Pinus patula* seedlings. Three plants were inoculated by making holes through the bark (5 x 2 mm) to expose the cambium. Mycelial plugs were then placed in contact with the cambium and the wounds were closed with Parafilm (American National Can™). Sampling point for RNA extraction was at 24 hrs post inoculation.

RNA extraction

For fungal subtractions, RNA was isolated from cultures grown in liquid malt extract for 48 hrs. Mycelium was collected by centrifugation and total RNA was extracted with TRIzol® reagent (GibcoBRL). Total RNA was extracted from the pine seedlings surrounding the inoculation point using a modified version of the method of Hughes and Galau [22] as in Chapter 3. All the mRNA fractions were separated from the total RNA with Oligotex™ (Qiagen) and cDNA synthesised using the cDNA Synthesis System (Roche).

Suppression Subtractive Hybridisation to Identify Novel Transcripts

Suppression subtractive hybridisation (SSH) was performed with the PCR-Select™ cDNA Subtraction Kit (Clontech). Three separate subtractions were performed with the fungal cultures. In the first subtraction, the low copy mRNAs were enriched for in the *F. circinatum* strain by subtracting against the degenerated culture. This subtraction was also done in reverse. In the second subtraction, enrichment was done against either one of the *F. subglutinans* (a), *F. guttiforme* (b), *F. opheodes prov. nom* (c), or all three combined (d). *F. opheodes prov. nom* was included in this study as it is occasionally isolated from *F. circinatum*-infected pine seedlings in nurseries. It is believed that its host in nature is grass or reeds and that it is not pathogenic (A. Jacobs, unpublished data). In the third subtraction, the expression of fungal genes *in planta* was evaluated by enriching for fungal mRNA from the infected trees at 24 hrs after inoculation. This was done against a combined RNA pool of mRNA from *F. circinatum* grown on half strength PDA and mRNA from the same pine seedlings that were inoculated into prior to inoculation and after 24 hrs. After subtraction two successive rounds of PCR amplified the novel fragments. All transcript-derived fragments (TDF) were separated on a 2 % agarose gel, extracted with the MinElute™ Reaction Cleanup kit (Qiagen), and cloned into pGEM-T Easy (Promega) for transformation to competent cells. The efficiency of the subtractions was verified through Southern blot analysis with either radioactive or DIG label (Roche) [40].

Screening and Analyses of Differentially Induced Fragments

All colonies containing inserts were amplified using PCR, dried and resuspended in 50 % v/v dimethyl sulfoxide (DMSO) [19], and spotted onto Nylon membranes with a 96 well plate replicator (Nunc). The membranes were hybridised, according to standard protocols, to [α^{32}P] dCTP labelled or DIG labelled (Roche) subtraction products, cDNA from the other *Fusarium* spp. and from *Pinus patula* [40]. All colonies that contained fragments linked to *F. circinatum*-induced sequences were sequenced and analysed with BLAST against the EST, nucleic non-redundant, and translated databases at GenBank (http://www.ncbi.nlm.nih.gov) [2]. All sequences were also searched against the EST database at the Phytopathogenic Fungi and Oomycete EST Database (COGEME: http://cogeme.ex.ac.uk/) [43] and the *F. graminearum* full genome sequence found at the Whitehead Institute (http://www.genome.wi.mit.edu/annotation/fungi/fusarium/ index.html).

Detection of Fumonisin Production by F. circinatum

The one TDF identified through the subtraction between *F. circinatum* and its degenerated analogue displayed significant similarity with the polyketide synthase type 1 gene (*PKS1*) (AJ315370) from *Gibberella fujikuroi*. Polyketide synthase genes have been linked with fumonisin production and this fragment was further analysed [35]. Fumonisin production by *F. circinatum*, *F. guttiforme*, *F. subglutinans*, the degenerate culture, and *F. opheodes prov. nom* were assayed for on GYAM media (0.24 M glucose, 0.05% w/v yeast extract, 8.0 mM L-asparagine, 5.0 mM malic acid, 1.7 mM NaCl, 4.4 mM K_2HPO_4, 2.0 mM $MgSO_4$, and 8.8 mM $CaCl_2$, pH 3 with phosphoric acid) [35]. The cultures were grown for four days in Erlenmeyer flasks shaken at 200 rpm in the dark at 25°C. RNA was sampled at 24 hr intervals after inoculation according to previously described method. After 96 hrs the cultures were filtered through Whatman 1 paper (Merck) to harvest fungal biomass and then through Cameo 25GSS 0.45 µm filters (Separations) to obtain a pure culture filtrate. The fungal biomass was ground to a fine powder in liquid nitrogen and resuspended in 70 % methanol. The culture filtrate and ground samples were analysed for presence of fumonisins by the Ridascreen® Fumonisin EXPRESS kit (Biopharm GmBH). This kit detects the presence of fumonisin B1, B2, and B3 through ELISA. The colorimetric changes were determined on a Multiskan Ascent v1.24 plate reader (Thermo Labsystems) at an absorbance of 620 nm with a 5s slow shaking step before the measurements were taken. Measurements were taken in duplicate for

the culture filtrates and single measurements for the ground samples. The standards included in the kit were 1, 2, 5 mg/kg.

To detect expression of the *PKS1* gene in *F. circinatum*, the previously reported KS primer set (KS-1: 5'-GGRTCNCCIARYTGIGTICCIGTICCRTGIGC, KS-2: 5'-MGIGARGCIYTIGCIAT-GGAYCCICARCARMG; M = A/C, N = A/C/G/T, I = Inosine, R = A/G, Y = C/T) was used in a RT-PCR reaction [35]. These reactions consisted of 100 ng cDNA and 5µM of each of the KS primers as previously described [35]. The PCR was conducted as 25 cycles of 20 s at 94°C, 20 s at 53°C, and 60 s at 72°C with an initial denaturation of 30 s at 94°C and a final elongation step of 60 s at 72°C. All products were separated as for the SSH subtractions.

Results

Morphology of Cultures

Cultural degeneration was visible as morphological changes after fifteen cycles of sub-culturing (Fig. 1). The colour of the wild type *F. circinatum* was white through salmon (9'd) [36] on all MEA, PDA and complete media. This closely resembled the morphology as previously described for *F. subglutinans* [30]. The degenerated culture displayed significant differences in culture morphology and colour when compared to the wild type *F. circinatum* isolate. The degenerated culture had a vinaceous (1''d) centre and the rest of the colony was white on MEA. On PDA it was pale vinaceous grey (59''''d) in the centre becoming vinaceous (1''d), with a white margin [36]. On complete medium it was buff (19''f) with a white margin [36]. This culture also produced less aerial mycelium than the wild type on all three media used (Fig. 1). Cultures transferred from those with abundant aerial mycelium displayed the most variation. Once a culture exhibited suppressed mycelial growth, there was less variation in subsequent sub-cultures. The degenerate cultures displayed several sectors that reverted to wild type morphology when grown on PDA plates (Fig. 1). This was not commonly observed on MEA and was never observed on complete medium.

SSH Analysis of F. circinatum

Several transcript-derived fragments (TDFs) were obtained after the three rounds of SSH (Fig. 2, Tables 2, 3, and 4). The efficiency of the subtractions was verified with Southern blot analysis and the differential induction of the TDFs was evaluated through reverse Northern blots with first strand cDNA (Fig. 3 and 4). The similarities between the TDFs and known fungal sequences from GenBank are given in Table 3, and the COGEME database in Table 4.

The subtraction between *F. circinatum* and the other *Fusarium* spp. gave rise to profiles that contained more bands than the subtraction reactions between the degenerate isolate, *F. circinatum*, and from the inoculated pine tree. The most TDFs were identified in the subtraction between *F. circinatum* and *F. opheodes prov. nom*, 10 out of the 16 identified TDFs (Table 2). The remaining three subtractions between the *Fusarium* spp. gave only six differentially induced fragments. Most of these sequences were either unknown or linked to stress responses by fungi (Tables 3 and 4). The *in planta* subtraction identified only one predominant sequence that has homology to the UDP-glucose pyrophosporylase ($E=8e^{-59}$) of *Amorpha fructicosa*. Screening of the subtracted cDNA library only yielded clones with pine inserts. All the sequenced clones shared homology to the UDP-glucose pyrophosphorylase. This fragment (CF660355) was used in a study presented in Chapter 3 as the plant and fungal material used in both experiments were the same. The subtraction between *F. circinatum* and its degenerate analogue identified a TDF with significant homology ($E = 5e^{-45}$ and $E = 5e^{-99}$) to a polyketide synthase (PKS) gene (*FUM5*) that has been linked to fumonisin production in *G. fujikuroi* [35].

Homology to proteins with known function could only be discerned for seven of the TDFs (CF660342, CF660344, CF660343, CF660339, CF660337, CF660348, CF660349) after searching both the NCBI and COGEME databases. Two TDFs shared (CF660348, CF660343) homology with stress response proteins expressed in *Fusarium* and *Gibberella*, one of the teleomorphs associated with *Fusarium*. CF660342 shared significant homology with a thioredoxin ($E = 3e^{-22}$) from *Triticum*. E17 shared significant homology ($E = 2e^{-37}$) with flavohemoglobin identified from the databases at NCBI and COGEME. CF660337 shared homology ($E = 3e^{-6}$) with an EST expressed under nitrogen starvation conditions and

with a glucanase precursor (E = $3e^{-5}$). Homology with the glucanase precursor was rather weak as the cut-off expectancy considered significant is any E-value that is smaller than e^{-5}. CF660339 could only be identified from the COGEME database where it shared homology ($2e^{-68}$) with ribulokinase. CF660349 shared significant homology (E = $2e^{-99}$) to the polyketide synthase gene (*PKS1*) identified in *G. fujikuroi*.

Several of the TDFs had significant homology to *Triticum*, *Sorghum*, *Hordeum*, *Zea*, and *Gossypium* sequences at the NCBI EST database (Table 5). A cut-off point of 90 % was used for the similarity analysis between the *F. circinatum* sequences and the NCBI ESTs. The sequence lengths and the matches between them were also taken into account. The sequences in question detected only fungal ESTs in conjunction with those mentioned in Table 5.

FUM5 Analysis

Identification of the *PKS1* homologue prompted the identification and analyses of a *FUM5* homologue with RT-PCR. With the fumonisin assay, no levels of fumonisin B1, B2, and B3 were detected in the culture filtrate or in the ground mycelia (Not shown). This was expected as *F. circinatum* is not known to produce fumonisin B1, B2, or B3 [15]. However, with the KS primer set, two fragments were detected in all the samples (Fig. 5). In the RT-PCR reactions the expected band of ca. 650 bp was observed together with a slightly larger fragment that co-amplified with the *FUM5* homologue. Sequencing of the two fragments confirmed that the 650 bp fragment is homologous to the *FUM5* gene. However, the larger fragment had homology to fungal type I polyketide synthase, but did not match the *FUM5* homologue.

Discussion

In this study we made use of SSH for the identification of TDFs from *F. circinatum* to gain an improved understanding of gene expression from this fungus. The isolates used in the subtraction reactions were chosen to identify TDFs that can be linked to differences in known pathogens from the section *Liseola*. The reed isolate, *F. opheodes prov. nom.*, was included in the study as it is sometimes isolated from *Pinus* seedlings that are infected with *F. circinatum*. This fungus is also closely related to the pitch canker fungus, but not pathogenic to pines (A. Jacobs, unpublished data). Nineteen TDFs from different subtraction events were isolated in this study. These TDFs shared significant homology to several stress-linked ESTs and unidentified proteins in different databases (Tables 3 and 4). The two best results from these subtractions were the homology to thioredoxin (CF660342) and flavohemoglobin (CF660344). These two proteins have been well studied in several organisms [41; 4; 27].

Flavohemoglobins in fungi are associated with denitrifying conditions and nitrosative stress [41; 27]. It would seem that flavohemoglobins do not reduce nitrate for energy gain, but rather as a detoxifying pathway [27]. In *Saccharomyces cerevisae* the expression of Yhb1, a flavohemoglobin, protein protects against nitrosylation of proteins by detoxifying nitric oxide (NO) under both anaerobic and aerobic conditions. It would seem likely that during culturing of *F. circinatum* on MEA that flavohemoglobin expression is up-regulated to protect against nitrosative stress (stress caused by excess nitrogen).

Thioredoxin is found in all organisms [4]. This group of proteins are ubiquitous and, therefore, have many functions. These can include DNA synthesis, stress responses, and regulation of apoptosis in mammals ([4] and references therein). It is to be expected that a protein with such diverse function would be up regulated during stress and can thus be identified when *F. circinatum* was experiencing stress during this study. The main stress in which thioredoxin is detected is oxidative stress [4]. Here it plays a role in the neutralisation of H_2O_2. It is clear that *F. circinatum* is under stress when cultivated on MEA (Table 3), this could be the reason that the thioredoxin is induced. However, it would seem that the stress experienced by *F. circinatum* is not oxidative, but rather linked with nitrogen and carbon

starvation as two fragments (CF660337 and CF660345) shared homology to proteins expressed under these conditions.

Several TDFs gave BLAST results against ESTs submitted as plant cDNA sequences. The high levels of homology displayed (Table 5), suggest that these cDNA libraries are potentially contaminated with cDNA originating from *Fusarium* species. Species of the genus *Fusarium* are known to be non-obligate pathogens or endophytes of grass species [12; 37; 51]. Endophytes interfere with cDNA sequencing projects, as it is difficult to ensure that the cDNA library is devoid of fungal or bacterial sequences until sequencing analysis has commenced [39]. High sequence homology of the resulting clones to fungal genes may result from this type of contamination.

Morphological degeneration of *F. circinatum* was easily achieved through sub-culturing. The morphological changes that occurred were consistent with those reported previously [49; 42; 8; 25; 14; 6; 50]. Previous studies have attempted to characterise degenerate cultures based on morphology and not at the molecular level. In this study the only gene expression difference identified was the polyketide synthase gene (*PKS1*). This may be due to the high similarity between the two cDNA populations produced for the subtraction, or that the screening of the subsequent screening of the clones failed to identify the mRNAs that are differentially regulated.

The isolation of the *PKS1* gene homologue from *F. circinatum* and the similarly identified *FUM5* homologue found in *G. fujikuroi* [35], might indicate fumonisin production. However, we did not detect any fumonisin with an ELISA that screened specifically for the three forms of fumonisins. This result is consistent with those of a study conducted to detect mycotoxins from 15 different *Fusarium* spp., where no fumonisin production was detected from *F. circinatum* grown on maize kernels [15]. The results obtained in this study, combined with those of Proctor *et al.* [35], support the view that *F. circinatum* does not produce fumonisin.

Detection of the *FUM5* homologue using RT-PCR in *F. circinatum* grown on GYAM medium was interesting. This gene is linked to fumonisin production in *G. fujikuroi* and can be detected after both 64 and 88 hrs of growth on GYAM medium [35]. Our results are similar to those of Proctor *et al.* [35] in that we also detected the transcripts of *FUM5* at 64 and 88 hrs of growth on GYAM. However, no fumonisin production was detected at 88 hrs. This would indicate the *FUM5* homologue in *F. circinatum* plays a role in other polyketide biosynthesis reactions. Polyketide synthases are enzymes that polymerise simple fatty acids into small molecules known as polyketides [23; 34]. These molecules are produced by prokaryotes and eukaryotes as secondary metabolites that have a myriad of functions [23; 34]. To detect the function of this homologue in *F. circinatum*, disruption of the *FUM5* homologue would have to be performed and the subsequent mutant strains investigated for loss of function.

Failure to detect any fungal expressed transcripts through the *in planta* subtraction is not unexpected. Several previous studies, using different techniques and studying different interactions, have been reported where low numbers of fungal transcripts have been encountered when studying an *in planta* interaction [29; 5]. We postulate that the level of fungal expressed genes was moderately low at our sampling point (24 hrs). This problem might be rectified by allowing for the interaction between the fungus and host to continue for a prolonged time. However, it would then not be possible to study key genes that are expressed early on in the pathogen during the infection process.

The differences observed in transcript levels isolated in the subtractions were unexpected. It can be argued that the subtractions between *F. circinatum* and *F. subglutinans* or *F. guttiforme* yielded fewer transcripts as these fungi are very similar to each other [31]. All three of these fungi were previously classified as *F. subglutinans* and only separated into *forma specialis* based on host specificity. They were described as different species based on phylogenetic differences [32]. However, *F opheodes prov. nom* is morphologically, pathogenetically, and phylogenetically different from *F. circinatum* (A. Jacobs, unpublished data).

In these experiments we attempted to identify the differences between several pathogens residing in the section *Liseola*. None of the fragments isolated were linked with known pathogenicity factors, although several of them were homologous to ESTs of unknown function found at the Phytopathogenic Fungi and Oomycete EST Database (COGEME). These sequences could not be identified for either one of two reasons. They could be from other parts of genes or ESTs for which the full-length sequences have not been isolated to date. The other explanation can be that they are from novel gene transcripts and no homologous sequences exist in any of the public databases. These fragments should therefore be studied to ascertain their role in pathogenicity of *F. circinatum*.

Acknowledgements

We thank Komatiland Forests Research, the National Research Foundation and the Mellon Foundation for financial assistance.

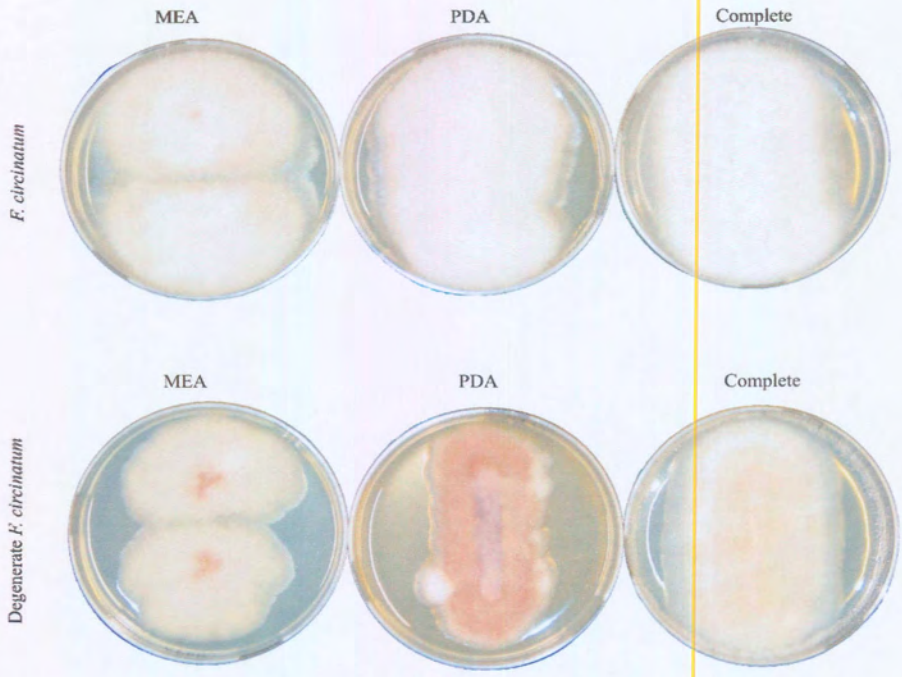


Figure 1: Cultural degeneration of *F. circinatum* after 15 cycles of sub-culturing on MEA.

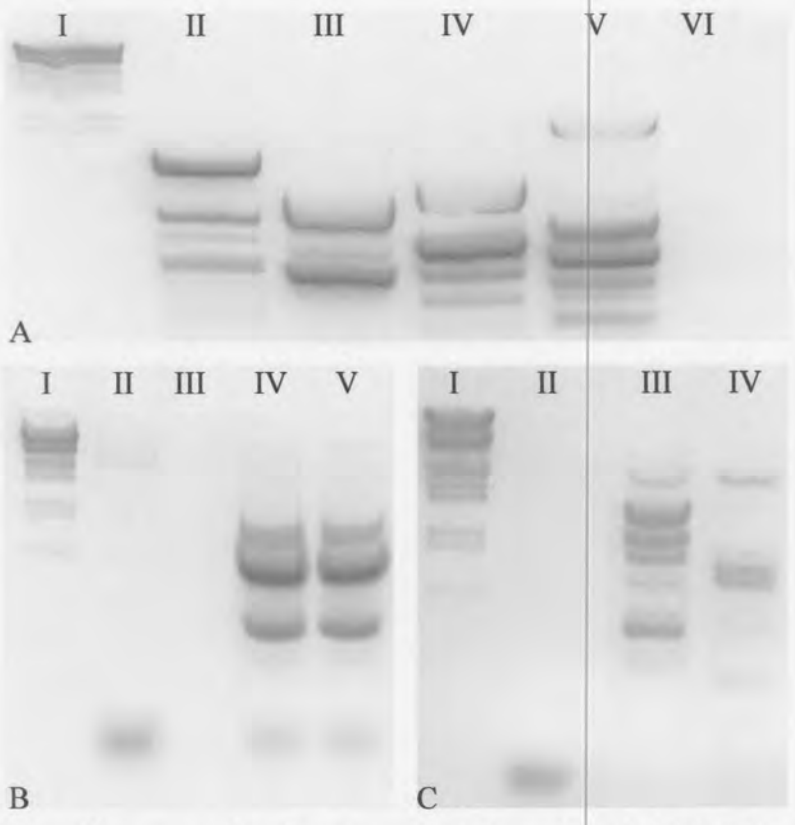


Figure 2: PCR amplicons produced after the subtraction hybridisation reactions. (A) Subtracted profiles after SSH between *F. circinatum* and other *Fusarium* species. (I) Marker III, λ DNA digested with *Eco*RI and *Hin*dIII, (II) subtracted against *F. opheodes prov. nom*, (III) subtracted against *F. subglutinans* (IV) subtracted against *F. guttiforme* (V) subtracted against *F. guttiforme*, *F. subglutinans*, and *F. opheodes prov. nom*, combined in equal concentrations, (VI) negative water control. (B) Profiles of subtracted amplicons after *in planta* SSH. (I) Marker III, (II) negative water control, (III) open, (IV and V) *in planta* PCR amplicons profiles. (C) PCR amplicons produced after subtraction between *F. circinatum* and its degenerated culture. (I) Marker III, (II) Negative water control, (III) forward subtraction where *F. circinatum* was used as tester, (IV) reverse subtraction where *F. circinatum* was used as driver. All the water controls were not diluted when subjected to secondary PCR, hence they have passed through *ca.* 50 PCR cycles.

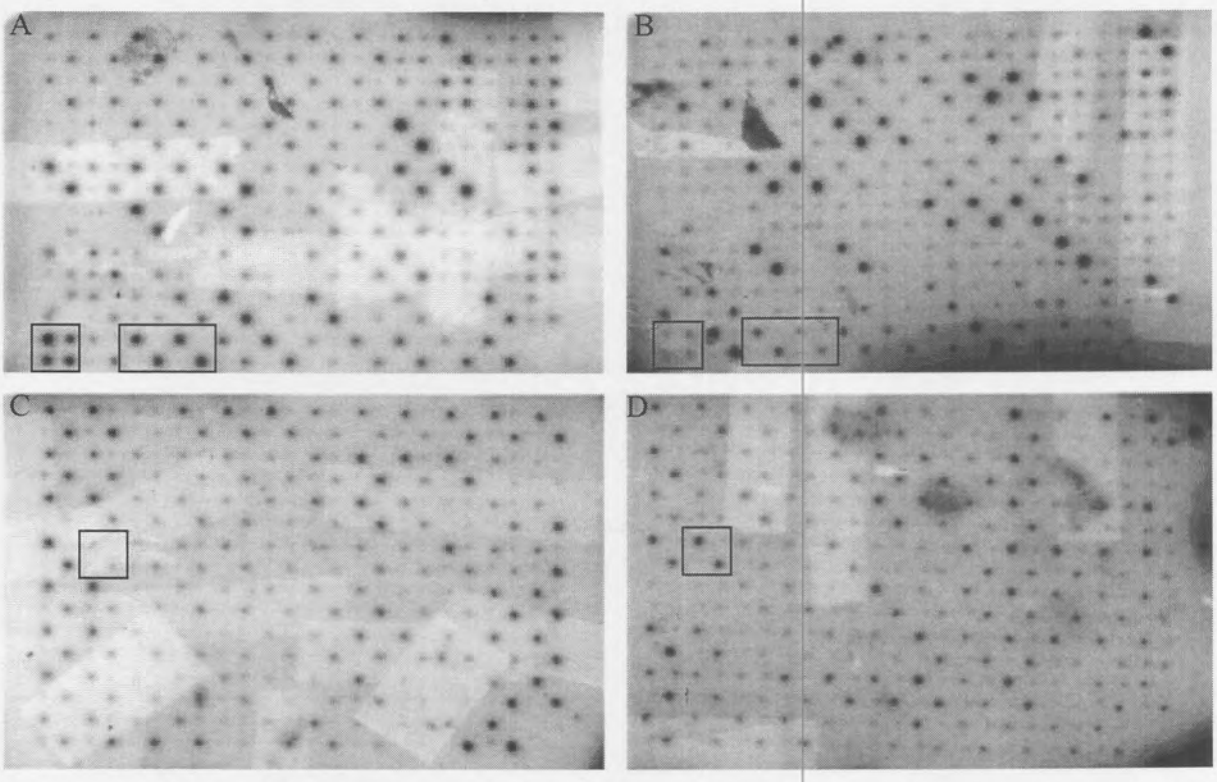


Figure 3: Differential screening of subtracted amplicons. The membranes were spotted in duplicate (A and B, C and D). Clones were spotted in duplicate and the bottom right corner four spots represent negative PCR controls. All membranes were exposed for 24 hrs. (A, C) were hybridised to cDNA from the tester *F. circinatum* and (B, D) were hybridised to cDNA pools from the driver populations. Squares are an example of differential induction between two of the clones between the two membranes.

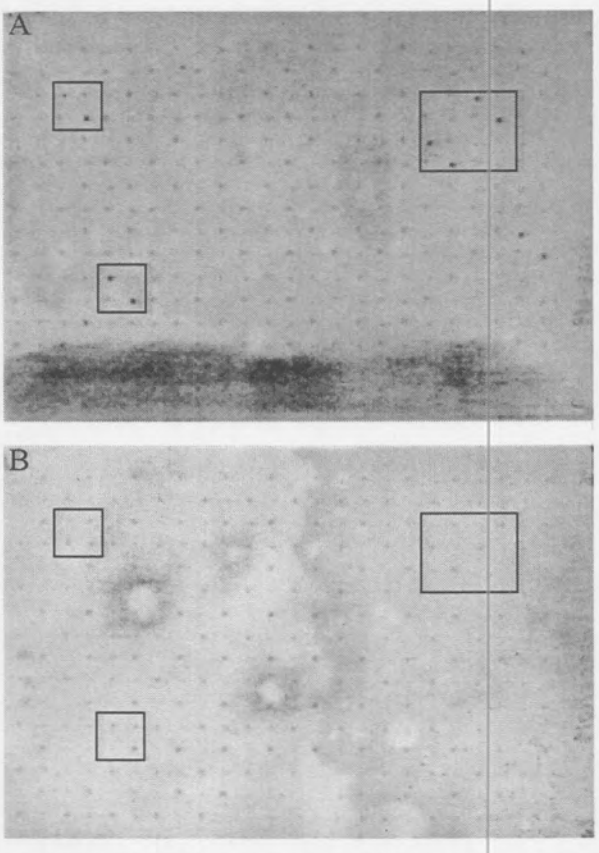


Figure 4: The subtracted amplicons after *in planta* SSH. Duplicates of each amplicons were spotted onto duplicate membranes (A, B), hybridised to DIG labelled cDNA and developed overnight. These membranes were hybridised to either cDNA from *F. circinatum* or a combined pool of the other *Fusarium* isolates. This last pool also contained cDNA from inoculated *Pinus patula* trees.

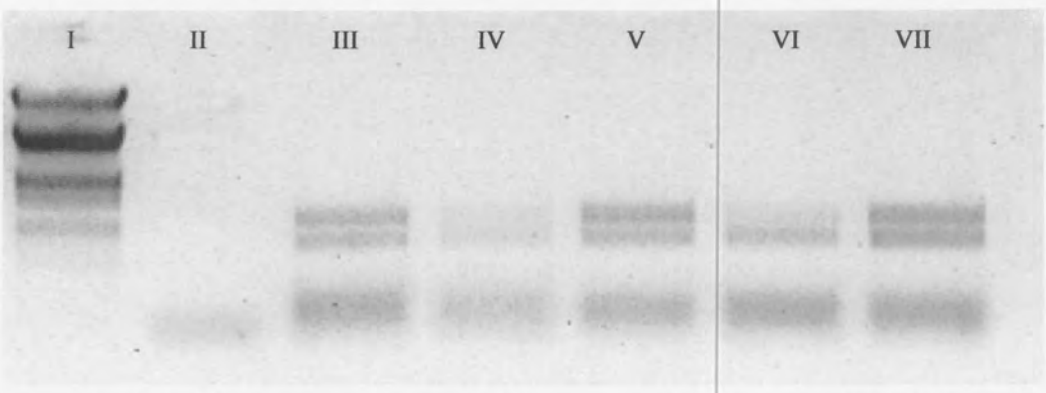


Figure 5: RT-PCR detection of the polyketide synthase homologue. RT-PCR reaction was performed on RNA samples extracted from GYAM grown cultures at 64 and 88 hrs post inoculation. (I) Marker III, (II) negative water control, (III) *F. circinatum* – 64 hrs, (IV) *F. verticillioides* – 64 hrs, (V) – *F. circinatum* – 88 hrs, (VI) *F. verticillioides* – 88 hrs, (VII) degenerated culture – 88 hrs. The smaller fragment of *ca.* 650 bp is the *FUM5* homologue and the larger fragment is a type I fungal polyketide synthase homologue but does not match *G. fujikuroi FUM5*.

Table 1: *Fusarium* isolates used in this study.

Isolate[a]	Host	Accession number[c]	Reference
F. circinatum (H)	*Pinus radiata*	CMW13229	[9]
F. circinatum (H)[b]	*Pinus radiata*	CMW13232	Present study
F. subglutinans (E)	*Zea mays*	CMW13231	[31]
F. guttiforme	*Ananas comosus*	MRC7539	``
F. opheodes prov. nom	Grass and reeds	CMW13230	--
F. verticillioides (B)	*Zea mays*	PPRI 6525	[31]

[a]Letters in brackets indicates the mating population of the fungus. [b]The degenerated culture of *F. circinatum* CMW13229. [c]CMW indicates the culture collection of the Forestry and Agricultural Biotechnology Institute (FABI, University of Pretoria, Pretoria, South Africa). MRC denotes the culture collection of the Medical Research Council (MRC, Cape Town, South Africa). PPRI denotes the culture collection of the Agricultural Research Council Plant Protection Research Institute (ARC-PPRI, Pretoria, South Africa).

Table 2: Summary of subtractions performed and the fragments identified in each.

	Subtraction 1	Subtraction 2				Subtraction 3
Driver cDNA	Degenerate culture	*F. subglutinans*	*F. guttiforme*	*F. opheodes*	Combination	*F. circinatum* and *P. patula*
Tester cDNA	*F. circinatum*	*F. circinatum*	*F. circinatum*	*F. circinatum*	*F. circinatum*	*F. circinatum*
Identified transcript derived fragments	CF660349	CF660337 CF660338	CF660339 CF660340	CF660342 CF660343 CF660344 CF660345 CF660346 CF660347 CF660348 CF660350 CF660352 CF660353	CF660351 CF660354	CF660355

Table 3: Sequences identified after SSH of *F. circinatum* with several *Fusarium* species. These sequences were analysed against the GenBank databases.

TDF No.	Database match[a]			
	BLASTN homology	E-value	BLASTX homology	E-value
CF660342	AJ005840, thioredoxin (*Triticum aestivum*)	$3e^{-62}$	AJ005840, thioredoxin (*Triticum aestivum*)	$3e^{-22}$
CF660351*	No hit; BI949398, *Fusarium* infected - EST (*Hordeum vulgare*)	$2e^{-43}$	No hit	-
CF660343*	No hit; BM351682, maize - EST (*Zea mays*)	$1e^{-96}$	No hit	-
CF660337	No hit; BU066029, nitrogen starved mycelia - EST (*Gibberella zeae*)	$3e^{-6}$	No hit	-
CF660341	No hit; CD460565, unknown - EST (*Fusarium graminearum*)	$3e^{-16}$	No hit	-
CF660344	AB016807, flavohemoglobin (*Fusarium oxysporum*)	$4e^{-96}$	AB016807, flavohemoglobin (*Fusarium* oxysporum)	$2e^{-37}$
CF660340	Unknown EST	$1e^{-92}$	EAA27671, hypothetical protein (*Neurospora crassa*)	$6e^{-47}$
CF660345*	No hit; CD456222, unknown – EST (*Fusarium graminearum*)	$4e^{-36}$	AL112087, unknown C and N deprivation (*Botrytis cinerea*)	$4e^{-20}$
CF660339	No hit; CD460858, unknown – EST (*Fusarium graminearum*)	$2e^{-94}$	CAD11374, hypothetical protein (*Neurospora crassa*)	$5e^{-49}$
CF660346*	No hit; BU808671, unknown - EST (*Triticum aestivum*)	$1e^{-114}$	No hit	-
CF660349	AJ315370, *PKS1* gene (*Gibberella fujikuroi*)	$2e^{-99}$	PKS1 protein (*Gibberella fujikuroi*)	$3e^{-45}$
CF660354*	No hit; BI201466, unknown – EST (*Fusarium sporotrichioides*)	$3e^{-58}$	EAA31224, predicted protein (*Neurospora crassa*)	$2e^{-18}$
CF660338*	No hit; CD456222, unknown – EST (*Fusarium graminearum*)	$1e^{-51}$	EAA31224, predicted protein (*Neurospora crassa*)	$3e^{-14}$
CF660350	No hit	-	EAA31111, hypothetical protein (*Neurospora crassa*)	$3e^{-22}$
CF660347*	No hit; BU808671, unknown - EST (*Triticum aestivum*)	$1e^{-109}$	No hit	-
CF660352	No hit	-	EAA26662, predicted protein (*Neurospora crassa*)	$4e^{-39}$
CF660348*	No hit; BI200866, unknown EST (*Fusarium sporotrichioides*)	-	No hit	-
CF660353	No hit	-	EAA31111, hypothetical protein (*Neurospora crassa*)	$4e^{-22}$

[a]For species complexes, the *forma specialis* were not added to the name. Where no similarity (designated No hit;) was found on the non-redundant nucleic acid database, BLASTN was performed on EST database. *The BLAST gave highest similarity to *Hordeum*, *Gossypium*, *Triticum*, and *Zea* ESTs found at GenBank.

Table 4: Sequence similarity search (TBLASTX) against the EST COGEME database for phytopathogenic fungi.

TDF No.	Database match[a]	E-value
CF660341	No hit	-
CF660346	No hit	-
CF660342	No hit	-
CF660345	GzCon[1165],unknown protein (*Gibberella zeae*)	$2e^{-49}$
CF660344	GzCon[0025], flavohemoglobin (*Gibberella zeae*)	$1e^{-36}$
CF660343	GzCon[0238], stress response protein (*Gibberella zeae*)	$3e^{-7}$
CF660337	MagCon[0071], glucanase precursor (*Magnaporthe grisea*)	$3e^{-5}$
CF660339	GzCon[3853], ribulokinase (*Gibberella zeae*)	$2e^{-68}$
CF660351	MagCon[7351], unknown protein (*Magnaporthe grisea*)	$2e^{-22}$
CF660354	GzCon[1165], unknown protein (*Gibberella zeae*)	$2e^{-49}$
CF660349	No hit	-
CF660338	GzCon[1165], unknown protein (*Gibberella zeae*)	$3e^{-40}$
CF660350	WOAA030ZH05C1, unknown (*Botryotinia fuckeliana*)	$5e^{-12}$
CF660347	No hit	-
CF660352	Ct1907410, unknown	$1e^{-28}$
CF660348	FsCon[0216], stress response protein (*Fusarium sporotrichioides*)	$3e^{-8}$
CF660353	Mgc04g03f, unknown (*Mycosphaerella gramini*)	$1e^{-13}$

[a]Accession numbers are unisequence identifiers for COGEME database.

Table 5: Obtained sequences believed to be of fungal origin and designated as plant ESTs in GenBank.

TDF No.	EST Similarity[a]	% similarity	Length[b]	E-value	Reference[c]
CF660346	BU808671, *Triticum aestivum*	97	225/230	$1e^{-114}$	Unpublished
``	BJ235436, *Triticum aestivum*	93	216/230	$2e^{-92}$	``
``	BQ608005, *Triticum aestivum*	92	208/226	$4e^{-81}$	[10]
``	BE402448, *Triticum aestivum*	92	208/226	$4e^{-81}$	[3]
``	BU984347, *Hordeum vulgare*	91	205/223	$1e^{-78}$	Unpublished
CF660345	CD427543, *Sorghum bicolor*	96	377/390	0.0	``
CF660343	BM351682, *Zea mays*	93	233/249	$1e^{-96}$	``
``	CD485852, *Gossypium hirsutum*	90	139/154	$2e^{-37}$	``
CF660351	BI949398, *Hordeum vulgare*	85	191/224	$2e^{-43}$	``
CF660354	CD427543, *Sorghum bicolor*	96	377/390	0.0	``
CF660348	BM351682, *Zea mays*	93	234/249	$5e^{-99}$	``
``	CD485852, *Gossypium hirsutum*	90	139/154	$2e^{-37}$	[13]
CF660338	CD427543, *Sorghum bicolor*	94	361/382	1e-153	Unpublished
``	CD427646, *Sorghum bicolor*	98	115/117	$1e^{-51}$	``

[a]Only the first five hits are shown when there were more than five hits. Similarity cut-off was selected at 90 % homology, except for E65, which was 85 %. [b]Indicates the amount of matching base pairs between the fungal ESTs and ESTs deposited in GenBank. [c]References as cited on GenBank.

References

1. **Adams D, Frankel S, Tidwell T** (1998) Chronicle of pitch canker in California. *In* MC Devey M, Gordon T, ed, Current and potential impacts of pitch canker in radiata pine. Proceedings of the IMPACT Monterey workshop. Commonwealth Scientific and Industrial Research Organisation, pp 49-53

2. **Altschul SF, Thomas LM, Alejandro AS, Jinghui Z, Zheng Z, Webb M, Lipman DJ** (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Research **25:** 3389-3402

3. **Anderson OA, Appels R, Bailey P, Blake T, Close T, Cloutier S, Dubcovsky J, Feuillet C, Gale M, Graner A, Gustafson P, Herrmann RG, Holton T, Jacquemin JM, Jia J, Joudrier P, Langridge P, Lazo GR, Lin JJ, McGuire P, Ogihara Y, Pecchioni N, Qualset C, Schuch W, Selvaraj G, Shariflou M, Sorrells M, Warburton M, Wenzel G** (2000) International Triticeae Mapping Initiative (ITMI) Workshop Report: 2000.

4. **Arnér ESJ, Holmgren A** (2000) Physiological functions of thioredoxin and thioredoxin reductase. European Journal of Biochemistry **267:** 6102-6109

5. **Avrova AO, Venter E, Birch PRJ, Whisson SC** (2003) Profiling and quantifying differential gene transcription in *Phytophthora infestans* prior to and during the early stages of potato infection. Fungal Genetics and Biology **40:** 4-14

6. **Awuah RT, Lorbeer JW** (1988) Nature of cultural variability in *Fusarium oxysporum* f. sp. *apii* race 2. Phytopathology **78:** 385-389

7. **Bilgrami KS, Sinha SP, Jeswal P** (1988) Loss of toxigenicity of *Aspergillus flavus* strains during subculturing - a genetic interpretation. Current Science **57:** 551-552

8. **Bolton AT, Donaldson AG** (1972) Variability in *Fusarium solani f. pisi* and *F. oxysporum f. pisi*. Canadian Journal of Plant Science **52:** 189-196

9. **Britz H, Coutinho TA, Wingfield MJ, Marasas WFO** (2002) Validation of the description of *Gibberella circinata* and morphological differentiation of the anamorph *Fusarium circinatum*. Sydowia **54:** 9-22

10. **Clarke B, Lambrecht M, Rhee SY** (2003) *Arabidopsis* genomic information for interpreting wheat EST sequences. Functional and Integrative Genomics **3:** 33-38

11. **Correll JC, Klittich CJR, Leslie JF** (1987) Nitrate nonutilizing mutants of *Fusarium oxysporum* and their use in vegetative compatibility tests. Phytopathology **77:** 1640-1646

12. **Desjardins AE, Plattner RD** (1998) Distribution of fumonisins in maize ears infected with strains of *Fusarium moniliforme* that differ in fumonisin production. Plant Disease **82:** 953-958

13. **Dowd C, Wilson I, McFadden H** (2003) Different gene expression responses in cotton root and hypocotyl tissues during infection with Fusarium wilt disease. Plant Cell **Submitted**

14. **Duncan JS, Bu'Lock JD** (1985) Degeneration of zearalenone production in *Fusarium graminearum*. Experimental Mycology **9:** 133-140

15. **Fotso J, Leslie JF, Smith JS** (2002) Production of beauvericin, moniliformin, fusaproliferin, and fumonisins B1, B2, and B3 by fifteen ex-type strains of *Fusarium* species. Applied and Environmental Microbiology **68:** 5195-5197

16. **Gelderblom WCA, Jaskiewicz K, Marasas WFO, Thiel PG, Horak RM, Vleggaar R, Kriek NPJ** (1988) Fumonisins: novel mycotoxins with cancer-promoting activity associated with *Fusarium moniliforme*. Applied and Environmental Microbiology **54:** 1806-1811

17. **Gordon T, Wikler K, Clark S, Storer A, Bonello P** (1998) Resistance to pitch canker in Monterey pine. *In* MC Devey M, Gordon T, ed, Current and potential impacts of pitch canker in radiata pine. Proceedings of the IMPACT Monterey workshop. Commonwealth Scientific and Industrial Research Organisation, pp 87-89

18. **Gordon TR, Storer AJ, Wood DL** (2001) The pitch canker epidemic in California. Plant Disease **85:** 1128-1139

19. **Hegde P, Qi R, Abernathy K, Gay C, Dharap S, Gaspard R, Earle-Hughes J, Snesrud E, Lee N, Quackenbush J** (2000) A concise guide to cDNA microarray analysis. Biotechniques **29:** 548-562

20. **Hepting GH, Roth ER** (1946) Pitch canker, a new disease of some southern pines. Journal of Forestry **44:** 742-744

21. **Horn BW, Dorner JW** (2002) Effect of competition and adverse culture conditions on aflatoxin production by *Aspergillus flavus* through successive generations. Mycologia **94:** 741-751

22. **Hughes DW, Galau G** (1988) Preparation of RNA from cotton leaves and pollen. Plant Molecular Biology Reporter **6:** 253-257

23. **Hutchinson CR** (1999) Microbial polyketide synthases: More and more prolific. Proceedings of the National Academy of Sciences, USA **96:** 3336-3338

24. **Idnurm A, Howlett BJ** (2001) Pathogenicity genes of phytopathogenic fungi. Molecular Plant Pathology **2:** 241-255

25. **Kirksey JW, Cole RJ** (1974) Screening for toxin-producing fungi. Mycopathologia et Mycologia Applicata **54:** 291-296

26. **Leslie JF, Plattner RD, Desjardins AE, Klittich CJR** (1992) Fumonisin B1 production by strains from different mating populations of *Gibberella fujikuroi* (*Fusarium* section *Liseola*). Phytopathology **82:** 341-345

27. **Liu L, Zeng M, Hausladen A, Heitman J, Stamler JS** (2000) Protection from nitrosative stress by yeast flavohemoglobin. Proceedings of the National Academy of Sciences, USA **97:** 4672-4676

28. **Markham JE, Hille J** (2001) Host-selective toxins as agents of cell death in plant-fungus interactions. Molecular Plant Pathology **2:** 229-239

29. **Muñoz CI, Bailey AM** (1998) A cutinase-encoding gene from *Phytophthora capsici* isolated by differential-display RT-PCR. Current Genetics **33:** 225-230

30. **Nelson JC, Tousson TA, Marasas WFO** (1983) *Fusarium* species: an illustrated manual for identification. Pennsylvania State University Publishers

31. **Nirenberg HI, O'Donnell K** (1998) New *Fusarium* species and combinations within the *Gibberella fujikuroi* species complex. Mycologia **90:** 434-458

32. **O'Donnel K, Cigelnik E, Nirenberg HI** (1998) Molecular systematics and phylogeography of the *Gibberella fujikuroi* species complex. Mycologia **90:** 465-493

33. **Owen D, Adams D** (1998) Overview of pitch canker in California. *In* M Devey, C Matheson, T Gordon, eds, Current and potential impacts of pitch canker in radiata pine. Proceedings of the IMPACT Monterey workshop. Commonwealth Scientific and Industrial Research Organisation, pp 21-24

34. **Pfeifer BA, Khosla C** (2001) Biosynthesis of polyketides in heterologous hosts. Microbiology and Molecular Biology Reviews **Vol. 65:** 106-118

35. **Proctor RH, Desjardins AE, Plattner RD, Hohn TM** (1999) A polyketide synthase gene required for biosynthesis of fumonisin mycotoxins in *Gibberella fujikuroi* mating population A. Fungal Genetics and Biology **27:** 100-112

36. **Rayner RW** (1970) A mycological colour chart. Commonwealth Mycological Institute, Kew, Surrey

37. **Riley RT, Norred WP** (1999) Mycotoxin prevention and decontamination - a case study on maize. Food, Nutrition and Agriculture **23:** 25-32

38. **Rohrbach KG** (1994) Fusariosis. *In* RC Ploetz, GA Zentmyer, WT Nishijima, KG Rohrbach, HD Ohr, eds, Compendium of tropical fruit diseases. APS Press, St. Paul, p 49

39. **Saar DE, Polans NO, Sørensen PD, Duvall MR** (2001) Angiosperm DNA contamination by endophytic fungi: Detection and method of avoidance. Plant Molecular Biology Reporter **19:** 249-260

40. **Sambrook J, Fritsch EF, Maniatis T** (1989) Molecular cloning: a laboratory manual, Ed 2. Cold Spring Harbor Press, Cold Spring Harbor

41. **Shoun H, Tanimoto T** (1991) Denitrification by the fungus *Fusarium oxysporum* and involvement of cytochrome P-450 in the respiratory nitrite reduction. Journal of Biological Chemistry **266:** 11078-11082

42. **Snyder WC, Hansen HN** (1954) Variation and speciation in the genus *Fusarium*. Annals of the New York Academy of Sciences **60:** 16-23

43. **Soanes DM, Skinner W, Keon J, Hargreaves J, Talbot NJ** (2002) Genomics of phytopathogenic fungi and the development of bioinformatic resources. Molecular Plant-Microbe Interactions **15:** 421-427

44. **Storer A, Wood D, Gordon T** (1998c) Epidemiology of pitch canker disease in California. *In* MC Devey M, Gordon T, ed, Current and potential impacts of pitch canker in radiata pine. Proceedings of the IMPACT Monterey workshop. Commonwealth Scientific and Industrial research Organisation, pp 30-32

45. **Tu C** (1929) Physiologic specialization in *Fusarium* spp. causing headblight of small grains. Phytopathology **51:** 145-162

46. **Viljoen A, Wingfield MJ** (1994) First report of *Fusarium subglutinans* f. sp. *pini* on pine seedlings in South Africa. Plant Disease **78:** 309-312

47. **Walton JD** (1996) Host-selective toxins: agents of compatibility. Plant Cell **8:** 1723-1733

48. **Walton JD** (2000) Horizontal gene transfer in the evolution of secondary metabolite gene clusters in fungi: An hypothesis. Fungal Genetics and Biology **30:** 167-171

49. **Wellman FL, Blaisdell DJ** (1941) Pathogenic and cultural variation among single-spore isolates from strains of the tomato wilt *Fusarium*. Phytopathology **31:** 103-120

50. **Wing N, Burgess LW, Bryden WL** (1995) Cultural degeneration in two *Fusarium* species and its effect on toxigenicity and cultural morphology. Mycological Research **99:** 615-620

51. **Yates IE, Hiett KL, Kapczynski DR, Smart W, Glenn AE, Hinton DM, Bacon CW, Meinersmann R, Liu S, Jaworski AJ** (1999) GUS transformation of the maize fungal endophyte *Fusarium moniliforme*. Mycological Research **103:** 129-136

CHAPTER 5

Comparison of nucleotide binding site leucine rich repeat genes in *Pinus*.

Prepared for submission to Theoretical and Applied Genetics as:
Comparison of nucleotide binding site leucine rich repeat genes in *Pinus*.

Abstract

Resistance gene analogues have been reported from a diverse set of plant species. These sequences are being used in comparative and evolutionary analyses to identify their relationships with each other. This knowledge will contribute towards isolating the genes that code for resistance linked proteins. In *Pinus* both classes of NBS-LRR genes (Toll/Interleuken-1-receptor like nucleotide binding site leucine rich repeats, *TNL*; and coiled-coil nucleotide binding site leucine rich repeats, *CNL*) have been detected. To further our understanding of these classes of resistance gene analogues, degenerate PCR was used to isolate TIR-NBS-LRR like resistance gene analogues from several *Pinus* spp. The sequences derived from the analogues were further characterised by comparing them with resistance gene analogues previously reported from *Pinus*. The DNA fragments sequenced in this study grouped with the previously identified TNL class *TNL* genes. Through motif analysis several of the known conserved motifs found in NBS domains were identified. The P-loop and GLPL motifs displayed a high level of conservation at the amino acid level with other plant NBS motifs. However, slight differences in several of the conserved regions were observed when the *Pinus* analogues were compared with *Arabidopsis thaliana* analogues. The most notable of these was a conserved motif identified in the same position as the RNBS-B motif but the amino acid sequence of this motif differed significantly from that found in other plants. This is the first study of resistance gene analogues in *Pinus*.

Keywords

TIR-NBS-LRR, *TNL*, *CNL*, kinase-3, P-loop, pines

Introduction

Several resistance (*R*) genes have been cloned [9; 15]. These *R* genes can be classified into six distinct classes, the largest of which contain nucleotide binding site leucine-rich repeat domains (NBS-LRR) [14; 9]. These NBS-LRR homologues encode for structural proteins and they play a role in signal transduction pathways during recognition of pathogens [27]. They are also believed to recognise several different pathogens such as viruses, bacteria, fungi, nematodes, as well as insects, through very specific interactions [34]. Hundreds of NBS-LRR homologues or resistance gene analogues (RGA) have been identified based on sequence data. Although the exact function of these genes has not been identified, they are related to *R* genes. This is based on the fact that they cluster together with known *R* genes in phylogenetic analyses, and several have also been mapped to functional *R* gene loci [16; 42; 23].

Known NBS-LRR genes can be divided into two classes based on the amino terminal domains. They are the *Drosophila* Toll and mammalian Interleuken-1-receptor like (Toll/Interleuken-1-receptor or TIR) and the coiled-coil (CC) motifs [40; 21; 11; 23; 26; 29; 28]. In monocotyledonous plants the CC-NBS-LRR genes (*CNL*) are the dominant class of genes with only a few TX and TN like sequences identified. No TIR-NBS-LRR genes (*TNL*) have been found for this group of plants [28; 2; 25]. In dicotyledonous plants the *TNL* genes are encountered in higher numbers than the *CNL* genes [23; 8; 25; 7; 24]. Both classes of these genes have been detected in conifers, mainly in *Pinus taeda*, but the only described members are *TNL* [25].

The *TNL* genes encode the TIR domain at the N terminus. This domain differs from the animal TIR, where it is found at the carboxyl end [13]. The TIR plays an essential role in pathogen triggered signal transduction [40]. In flax it has also been shown that polymorphisms in the TIR region can alter resistance specificity of the *L* resistance gene [12]. This, combined with the multiple TIR domains found in the *Arabidopsis* genome that are not linked with an NBS or LRR region [25; 24], would indicate that the TIR domain has several functions in pathogen recognition and defence response.

CC-NBS-LRR genes (*CNL*) contains a coiled-coil domain (CC) that is the prevalent domain linked to the NBS domains in monocotyledonous cereals [28]. This domain comprises 2–5 helices with amino acids packed into a 'knob-hole' conformation, where the side chain of one amino acid fits into a hole arising from four other side chains on the other helix [21]. The true function of CC

domains is not yet understood. It is believed that they may function similarly to the TIR domain [13]. However, they are also found at the carboxyl terminus of a transmembrane domain of the *RPW8* gene [41]. Here, it is suspected that they are directly involved in pathogen recognition or that they act as helper molecules for another recognition protein or molecule [13; 41].

The cores of NBS-LRR *R* genes encode a nucleotide binding domain [39; 16]. The NBS domain is characterised by several short conserved regions [33; 39]. These regions can be conserved among NBS family members and different between the different NBS family members [24]. The most common of these regions is the phosphate-binding loop or P-loop [33]. This region is found in both the ATP- and GTP- binding families of NBS genes [39]. Other conserved regions are the kinase-2, kinase-3, and GLPL regions. These regions are sufficiently conserved to be used for classifying sequences into the NBS class [39; 22]. They are also exploited to design degenerate primers to detect a wide range of NBS homologues [16; 19; 42].

Leucine rich repeat (LRR) domains are found in proteins with different functions. They are located mainly at the carboxyl terminus of the protein. However, all of these proteins interact with other proteins and several function in signal transduction pathways [17; 27]. The LRRs are characterised by serially repeated leucine residues (xxLxLxx), or other aliphatic amino acids [10; 14]. The only LRR-containing protein that has been described through crystallography is the porcine ribonuclease A inhibitor [18]. From this structure, it is inferred that the leucine residues are buried and do not take part in protein interactions, but impart conformational structure to the protein [17; 18]. Most LRR regions do not match a consensus sequence and this would indicate that they do not have a highly regular structure of the xxLxLxx conserved repeats [6]. This might be due to the fact that many conformational structures are needed to recognise a diverse set of pathogen-derived molecules.

The rearrangement of the leucine residues in LRR regions indicates that they are involved in the recognition of different pathogens [37; 10]. This was shown experimentally using the ten alleles of the *L* gene in flax that imparts rust resistance [12; 20]. Functional analysis on different combinations of the LRR repeats for the ten alleles, conferred different resistance specificities to susceptible isolates when they were transformed. The combination of more than one allele in these experiments conferred new specificities to the transformants that differed from those from which the combined constructs were derived [12].

The best defined relationships in the *TNL* and *CNL* genes are from studies on *Arabidopsis* and *Medicago truncatula* [8; 25; 43; 24]. These studies have shown that the TIR and non-TIR sequences group into different clades in phylogenetic analyses [43; 24]. However, these genes are intermingled when the mapping data are analysed for *Arabidopsis* [24]. Preliminary data suggest that this is not the case in the *M. truncatula* and *Glycine max* genomes, where non-TIR sequences and TIR sequences are found on separate chromosomes [42; 43]. It is suspected that larger genomes with more retrotransposons and higher ploidy levels, i.e. *Pinus* and *Triticum*, might have more complex clusters of *CNL* and *TNL* genes [24]. Phylogenetic analysis between RGA of *M. truncatula* and *M. sativa*; and between *Lycopersicon esculentum* and *Solanum tuberosum* have shown that RGAs from different species group to similar branches [27; 43]. This indicates that orthologous genes exist or that these genes are derived from an orthologous ancestor [43]. However, the TIR and non-TIR are always separated into different clades [43; 24]. Therefore, they represent two very distinct classes of genes.

Meyers *et al.* [25] have identified new combinations of TIR sequences that are related to the *TNL* genes (Table 1). This analysis has identified TIR domains that are not linked to either the NBS or the LRR. These TIR domains could be necessary adaptor proteins that link with other TNL proteins to facilitate in the signal transduction pathway [25] Through an updated annotation of the *Arabidopsis* genome, they have also found that *TNL* genes can have multiple TIR and NBS sequences [24]. It is suspected that the TIR, CC, NBS, and LRR domains may have evolved separately, but at some stage they combined to form multi-domain proteins that have an evolutionary advantage over the separate domains [24].

Determining the evolutionary relationships between the *TN/TNL* and *CN/CNL* genes is complicated. An analysis of the introns and exons in these genes shows that there are considerable differences between them. The *TN/TNL* genes contain on average four exons and the *CN/CNL* class usually only has one [24]. The assortment and positions (phases as 0, +1 and +2) of these introns in the *CN/CNL* genes suggest that they are the more ancient of the two gene types [24]. The branch lengths on the phylogenetic trees of the *CNL* genes are much longer than those derived from TNL sequences [8]. This also indicates a more ancient origin for the former gene type [24]. Based on molecular dating, the split between the *TNL* and *CNL* predates the split between gymnosperms and angiosperms [24]. This makes it difficult to deduce evolutionary relationships of these genes based on the data from sequenced genomes of angiosperms [24].

Several studies have been done on *TNL* and *CNL* genes in dicotyledonous and monocotyledonous plants [27; 28; 2; 8; 25; 24]. However, only two studies have included conifer species [8; 25]. The current understanding of *TNL* and *CNL* genes is thus limited in conifers. To further the understanding of the structure and relationships of *TNL* genes in the genus *Pinus*, we amplified and sequenced segments of the NBS regions of these genes with degenerate primers. These sequences were used in a phylogenetic analysis to assign these RGAs to a specific class of *TNL* genes. These sequences were also analysed for known motifs to elucidate the structure of *Pinus TNL* genes.

Materials and Methods

Plant Material and DNA Extraction

The pine isolates used in this study were *P. elliottii*, *P. engelmannii*, *P. greggii*, *P. patula*, *P. radiata*, and *P. taeda*. Needle material for DNA extraction was obtained from Komatiland Forest Products (*P. patula*, *P. radiata*, *P. elliottii*), Global Forest Products (*P. taeda*), and Mondi Ltd. (*P. greggii*, *P. engelmannii*), South Africa.

DNA was extracted using a modification of the CTAB method [35]. Needle material from the different pine isolates was ground in liquid nitrogen. This powder was added to 750 µl of hot extraction buffer (5 % w/v CTAB, 1.4 M NaCl, 0.1 M Tris-HCl pH 8, 1 % w/v polyvinyl pyrrolidinone, 0.2 % v/v 2-mercaptoethanol). This was mixed thoroughly and incubated at 60°C for one hour. One volume of chloroform: isoamyl alcohol (24:1) was added, homogenised with a vortex mixer and centrifuged for 10 min at 10000 rpm. The upper phase was transferred to a new tube and the step repeated until the interface was clean. The volume of the upper phase was determined and a tenth volume 3 M NaOAc (pH 4.6) and twice the total volume of ethanol were added. DNA that precipitated immediately was spooled out, washed with 70 % ethanol and resuspended in 100 µl water. All DNA samples were further cleaned through polyethylene glycol (PEG) precipitation, as no PCR amplification was successful unless this step was performed. A volume of 20 µl from each of the DNA samples was added to 20 µl 3M NaCl and 120 µl 14 % PEG 8000. This was incubated on ice for 30 min and then centrifuged for 30 min at 14000 rpm at 4°C. The pellet was washed with 70 % ethanol, dried and resuspended in 20 µl water. Concentration of the samples was determined spectrophotometrically and all sample concentrations were adjusted to 5 ng/µl for use in subsequent PCR reactions. All samples were stored at –20°C for the duration of the study.

Primer Design, PCR Amplification and Sequencing

Pinus taeda NBS-LRR amino acid sequences deposited by Meyers *et al.* [25] were obtained from GenBank. These sequences were aligned to each other with ClustalX [38]. The conserved domain database (CDD) at GenBank was used to identify conserved domains for primer design. Two domains similar to the KTTL (P-loop) (+55 aa) and GLPL (+245 aa) regions of the NBS consensus sequence at CDD were identified. These regions were selected for design of degenerate primers as consensus-degenerate hybrid oligonucleotide primers (CODEHOP) (http://blocks.fhcrc.org/codehop.html) [31].

Degenerate primers and primers designed by Yu *et al.* [42] were used to amplify putative NBS-LRR homologues using the following reaction conditions on a GeneAmp® 9700 (Perkin Elmer): 94°C for 1min, 60°C for 40 s – 1°C/cycle, 72°C for 1 min for 6 cycles. This was followed by 30 cycles of the same conditions except for an annealing temperature of 55°C and 1 second was added to the extension step after each cycle. The reaction had an initial denaturation of three minutes and a final elongation period of five minutes. The ramp speed of the GeneAmp® was set to 50 % between the annealing step and the amplification step for amplification of the target region as no amplicons were observed in the absence of this adjustment. The reaction conditions were 1X Reaction buffer (10 mM Tris-HCl, pH 9.0, 50 mM KCl, 0.1 % Triton® X-100), 1.5 U *Taq* polymerase (Promega), 1.5 mM $MgCl_2$, 200 µM each dNTP, 25 µM each forward and reverse primer. All PCR amplicons were separated on 2 % agarose gels. Amplicons of approximately 570 bp and 340 bp were expected after amplification. Fragments of this size were punched from the gel and extracted with the MinElute Reaction Cleanup kit (Qiagen). The amplicons were cloned into pGEM-T Easy vector (Promega) for transformation to competent JM109 *E. coli* cells according to the manufacturer's instructions.

Positive colonies were screened using colony PCR with the T7 (5-TAATACGACTCACTATAGGG) and Sp6 (5-ATTTAGGTGACACTATAG) primer sets and amplicons of the expected size range were purified through ethanol precipitation. The PCR conditions were 30 s at 94°C, 30 s at 50°C, and 30 s at 72°C, with an initial denaturation and breaking of the cells at 94°C for 1 minute, and a final extension at 72°C for 5 min. The reaction conditions were 1X Reaction buffer, 0.5 U *Taq* polymerase (Promega), 0.3 mM $MgCl_2$, 40 µM each dNTP, 1 µM of each primer. The amplicons were subjected to cycle sequencing of both strands with the BigDye™ Dye Terminator v3.0 using the same primer sets as before. The sequencing reactions were run on an ABI™ 3100 (Applied Biosystems).

Computational Analyses

The different classes of NBS sequences were abbreviated using the descriptions reported by Meyers *et al.* [25] (Table 1). Homology of the sequenced fragments to known NBS regions were determined by screening against the database at GenBank through the BLASTX algorithm (http://www.ncbi.nlm.nih.gov) [1]. The NBS-like sequences were translated to amino acid sequences by the European Molecular Biology Open Software Suite (EMBOSS) [30], and aligned to each other and previously identified NBS homologues with ClustalX (Table 2). Sequences that were not homologous to NBS were included in a tree derived from nucleotide alignments to determine their relationship with each other.

DNA and amino acid sequence data were analysed using the neighbour joining programme in PAUP* 4.4b10 [36]. In the analyses all uninformative characters were excluded and a 1000 pseudo-replicate bootstrap analysis was performed to determine the confidence in branching points. Groups with a bootstrap value higher than 70 % were retained. The trees containing NBS homologues were rooted to NBS homologues from *Physcomitrella patens* (BJ59673, BJ198288) [25], and the tree containing the nucleotide sequences was rooted to NBS homologues from *Cryptomeria japonica* (AU084676).

To further our understanding of the grouping and structure of the isolated NBS homologues a motif analysis was performed on these sequences. The sequences were analysed using Multiple Expectation of Maximisation for Motif Elicitation (MEME) and the Motif Alignment and Search Tool (MAST) (http://meme.ucsd.edu/) [3; 4; 5]. The parameters were set to identify ten conserved regions. In the MEME analysis, motifs are identified and their specific position recorded on each of the sequences that were analysed. The MAST program then makes use of these motifs and arranges them on each of the sequences. The sequences are ordered according to the E-values that are allocated accordingly to the strength of the match of the sequence to the motif pattern.

Results

Primer Design and PCR Amplification

The alignment between the *P. taeda* NBS sequences and the NBS domain found at the CDD indicated that the P-loop (KTTL) and the LPL region found in the NBS of other plants were also conserved in *Pinus* (Fig. 1). These conserved regions allowed for the design of degenerate primers

(Table 3) that would potentially allow for the amplification of a range of NBS homologues in conifers. The degenerate primers were designed to contain a degenerate core and a consensus clamp region that should stabilise the binding of the primer [31]. The primers that were designed in this study included approximately 60 amino acids more in the amplicons than the NBS-set. They were effective in amplifying NBS homologues from several *Pinus* species.

The primers designed for this study were different to those reported by Yu *et al.* [42]. Their primers targeted smaller fragments of the NBS region, which stretched from the 180th amino acid to the LPL region (Fig. 1). The forward primer of the NBS set was designed to bind to the amino acid sequence SRIIITTR. This sequence was not conserved in the *Pinus* sequences that were amplified in this study or obtained from the web. The use of this primer set did not result in the amplification of any resistance gene analogues (RGA) and all the detected fragments were from retrotransposable elements.

PCR with the degenerate primers produced multiple bands in both reaction sets. Fragments that were larger than the expected 570 bp proved to be retrotransposable elements when sequenced. The KTTL-F and LPL2 primers gave less smearing and less non-specific amplicons than the NBS-set, which did not amplify any RGAs. Several of the fragments isolated by the CODEHOP-designed degenerate primers, which fell in the expected size range, were not NBS homologues. These fragments gave no significant homology to any identified protein when searched against the GenBank database. Not all of the sequences generated in this study contained open reading frames (ORF) as they were interspersed with stop codons.

Phylogenetic Analyses

Three phylogenetic analyses were performed in this study. The first included only the sequences identified by Meyers *et al.* [25] from *P. taeda* and *C. japonica*. This provided a basis for comparison to the full dataset (Fig. 2). The second analysis contained all the NBS homologue data points (Fig. 3), and the last data set contained the sequences that displayed little or no homology to NBS regions or any identified proteins (Fig. 4). The first two data sets were constructed from alignments between amino acid sequences of the NBS homologues, and the last tree between nucleotide alignments of the NBS homologues and the remaining sequences. Nucleotide sequences were used, as some of them contained several stop codons.

The different classes of NBS homologues grouped into two distinct clades (Fig. 2). The division of these clades was supported by confidence intervals of 70 % and 100 % for clade I and II, respectively. Clade I included the TNL, TX, TPL, and TN classes of NBS homologues. The NBS homologue from *C. japonica* was also included in this clade. The grouping of one of the TNL (AAM28917) and the TN (AAM28913) class was supported by a 93 % confidence interval. Clade II included three separate sub-groups supported by high confidence intervals (98 %, 89 %, and 100 % respectively). The two TNL homologues (AAM28908, AY091556) in this clade grouped together with an NL (AAM28909). This sub-clade was separated from the rest of the NL homologues (AAM28912, AAM28911, AAM28915, AAM28916). Although all the NL homologues grouped into this clade, they were not equally related (Fig. 2). The TNL class appeared polyphyletic as it was found in both the major clades of the phylogram.

The relationships between the different NBS homologues that were obtained in this study with known gymnosperm NBS homologues are shown in Figure 3. The topography of this tree was similar to the first tree (Fig. 1), which was generated by aligning previously published NBS homologues (Fig. 2). Three well-separated clades with strong bootstrap support emerged (Fig. 3). Clade I included the NL homologues and two of the TNL homologues as in Fig. 2. The two *P. pinaster* sequences (BX249943, BX251712) also resided in this clade. However, they were further removed from the other sequences and resided in a discrete sub-group. Clade II contained the TNL, TX, TPL, and TN classes of NBS homologues. The resolution in this clade was better than that observed in the tree produced from the first dataset (Fig. 2). However, the TNL (AAM28917) and TN (AAM28913) class sequences were not present in clade III as would have been expected from the first dataset. These two sequences shared greater homology with sequences obtained in this study and grouped in clade II. The NBS homologue from *P. radiata*, isolated with the KTTL-F and LPL2 primer set, grouped separately from the other sequences isolated with this primer set.

The relationship between the NBS homologues and the unidentified sequences obtained in this study are shown in Figure 4. These relationships were based on nucleotide alignments of the sequences generated with the KTTL-F and LPL primer set. This analysis was based on nucleotide alignments, as several fragments could not be translated to uninterrupted amino acid sequences. The groups observed in this analysis were not well resolved. The NBS homologues included in this tree had the same designations as those in Fig. 3.

The NBS homologues grouped into two sub-clades (clades I and IV) supported by confidence intervals of 78 % and 94 % respectively (Fig. 4). Based on nucleotide sequence, the *P. patula* 2 and 3 sequences were more distant from the other NBS homologues. In this analysis they grouped closer to the *P. radiata* sequence in clade IV. Several of the sequences without known function (*P. elliottii* N1, *P. taeda* N1 and N2, and *P. patula* N1 and N2) were clustered together to form clade II. However, not all the sequences found in this cluster could be translated into uninterrupted amino acid sequences. A second cluster (clade III) was observed that was supported by an 83 % bootstrap value. This cluster contained sequences from *P. greggii* and *P. elliottii*.

Structural Analyses

The MEME analysis identified 10 conserved motifs for the NBS sequences (Fig. 5 and Table 4). The consensus amino acid sequences for these motifs were conserved when analysed in the *Pinus* sequences (Fig. 6 and Table 4), but they were different from the proposed *Arabidopsis* motifs [24]. For the proposed *Arabidopsis* sequences encoding the GLPL motif the consensus sequence is EVAxLAGGLPLKVL, for the kinase-2 motif it is RLKDKKVLIVLDDVD, and for the RNBS-B motif the consensus is SRIIVTTEDK. In *Pinus*, the consensus sequences were KFVEACGGLPLSLK, QLSSFHALIILDDVD, and SLILITSRDR, respectively (Fig. 6 and Table 4). It was possible to putatively assign five other motifs based on their location and shared homology in the analysed sequences. Of these four sequences two were found in the TIR, two in the LRR and one corresponded to the RNBS-B motif's position, but the sequence was not similar to the conserved *Arabidopsis* sequence. The P-loop motif was not detected as a conserved motif as only two of the sequences isolated in this study contained the KTTL amino acid sequence when translated from the primer sequences.

Several other motifs were also identified (Table 4). These motifs corresponded closely to the previously described *Arabidopsis* motifs [24]. The existence of four domains has been reported in the TIR region of *Arabidopsis* (TIR-1 through TIR-4). The MEME analysis in this study identified two motifs that corresponded to the TIR-1 and TIR-2. The TIR-2 region was high in tryptophan residues (Table 4). This was not reported for the TIR motifs of *Arabidopsis*. In *Pinus* the RNBS-A, RNBS-D and MHD regions were also detected. The RNBS-A was interspaced between the P-loop and kinase-2. The RNBS-D followed the GLPL region. The MHDV region corresponded to a mixture of the reported *CNL* and *TNL* MHD motifs (TIR: MH[ndkr]LLQQ; CC: MHD[vm][vlim]R[ed]) reported by Bai *et al.* [2] (Table 4). The LRR motifs were recognisable through the serial repetition of the xLxxLx motif. Motif eight contained two repeats

(LPDPFGNLKNLRHINMSFCPKLKQLPDSFGNLTHLQHIDMH) not previously reported for this region. The two most conserved motifs were the kinase-2 (LIILDDVD) and GLPL (GLPL) regions (Fig. 6 and Table 4). In *TNL* genes the kinase-2 motif encodes a D at the end and for the *CNL* genes a W [24].

Based on the sequence of the kinase-2 regions all the sequences isolated in this study were considered to be *TNL* and not *CNL* homologues. Placing NBS homologues into either the *CNL* or *TNL* genes can be based on the last amino acid found in the kinase-2 motif [23; 25]. In these studies it was reported that in most *CNL* genes the last amino acid found in the kinase-2 domain is a tryptophan. For *TNL* genes, this amino acid is an aspartic acid.

The MEME analysis was also performed on the non-NBS homologues (Fig. 7 and Table 5). Several motifs were found in the NBS homologues as well as the non-NBS sequences. The two motifs numbered one (GLPL) and three (P-loop) could not be used in this analysis as they encoded for the primer sequences used and were found in all the sequences. The P-loop motif (motif three) was duplicated in several of the sequences (Fig. 7). Translation of motifs five and nine indicated homology with the RNBS-B and RNBS-A motifs found in *Pinus* NBS amino acid sequences. Motif two indicated homology with the kinase-2 motif found in *Pinus* NBS amino acid sequences. These were the only motifs that shared significant homology with known NBS conserved motifs.

Several other motifs were identified in the non-NBS sequences. When translated, they had homology to amino acid sequences that were highly conserved in all the *Pinus* NBS sequences analysed. However, these sequences are not conserved motifs found in NBS sequences. For example, motif four shared homology to the GIEDSSIYNLTGLNTQ stretch between the RNBS-B and GLPL regions of the NBS homologues. It was situated between the kinase-2 region and the RNBS-B region. These two motifs were inverted in the non-NBS homologues in comparison with the NBS homologues included in this analysis.

Two motifs (eight and ten) were identified that shared no homology to any of the included NBS sequences. However, motif eight was conserved in the non-NBS sequences (Fig. 7). The motif order in these sequences also differed from that found for the NBS sequences. Translations to identify homology of these motifs to the NBS amino acid sequences introduced up to three stop codons into the sequence.

Discussion

In this study degenerate primers were used to amplify NBS-LRR homologues from several *Pinus* spp. commonly grown in plantations. Based on the conserved aspartic acid found in the kinase-2 domain (Table 4), the sequences were assigned as *TNL* homologues. These homologues were used in several analyses to determine the relationship between them at amino acid level, and to determine the structural differences in the repeats that they contain. As the *TNL* genes comprise a super family of RGA genes that contains various subgroups [8; 25; 7; 24], the sequences were further analysed and compared with known NBS-LRR regions from *P. taeda*.

Primer Design and NBS Sequence Analysis

The two most conserved regions of the NBS are the P-loop and GLPL regions. Of these two, the *Pinus* P-loop shares the highest homology to the known consensus sequence (GGxGKTTLAK for *Pinus*) that is found at GenBank. The GLPL region of *Pinus* (CxGLPLxLKV) is also highly homologous to the consensus found at GenBank (CKGLPLALKVLGGLL). This makes it possible to design degenerate primers that will detect NBS domains in a host of plant species. The other region that has been used previously, the NBS primer set used in this study (Table 3), is designed around the RNBS-B domain [42; 24]. This domain's position is conserved in *Pinus* but the sequence differs completely from that of the reported consensus sequence. In the CDD consensus sequence it is encoded by SRVIVTTR and in *Arabidopsis* by SRIIVTTE. The identified *Arabidopsis* and *G. max* NBS domains correspond closely to the consensus. In *Pinus* the consensus sequence from the isolated sequences were DSLILITSR (Fig. 6 and Table 4). The degenerate primer designed by Yu *et al.* [42] targets the RNBS-B domain with the reverse primer. We were not able to amplify any RGA sequences from *Pinus* with this primer, as this specific amino acid sequence is not conserved in *Pinus*.

The sequences that were detected in this study all group in one cluster. This suggests that the degenerate primers might be biased. This is also reflected in the grouping of sequences into one cluster in the motif analysis (clade II in Fig. 2, Fig. 5 and 7). In future, primers could be redesigned from published data to isolate more NBS-like sequences [24]. However, designing primers more specific for NBS homologues found in gymnosperms is currently not feasible from the existing databases on gymnosperms. The kinase-2 domain in *Pinus* serves as a good example. In our MEME analysis we identified the consensus sequence as LIILDDVD. Only two of the *Pinus* sequences used has the same sequence as the *Arabidopsis* kinase-2 consensus (LIVLDDVD).

Compensating for the codon differences for these two amino acids would raise the degeneracy of a primer eight times. A preferable option would be to design several primers for this region that have a lower level of degeneracy.

No gymnosperm genome sequence data are available to enable the identification of all possible NBS analogues. To understand the full range of these analogues in gymnosperms, it is therefore necessary to supplement the available sequences that are found in public databases. Several degenerate primers have been designed to amplify NBS homologues from diverse species [16; 19; 42]. In this study, we have made use of primers designed specifically from known *Pinus* sequences. Analysis of the resulting amplified fragments can now be used further in comparative analyses to design several primer sets that will be able to target all the known NBS variants from gymnosperms. While this will result in a more complete understanding in gymnosperms, the full spectrum of RGAs present in the genome will only emerge when a full genome sequence is available.

Comparative Analysis of NBS Homologues

Phylogenetic analysis for the *Pinus* NBS homologues was performed on putative amino acid sequences. The region stretching between the P-loop and the GLPL region of the NBS was used to infer the genetic relationship of these sequences to *R* genes. It is feasible to make use of the NBS region for a phylogenetic analysis as they are related to known TIR, CC and LRR regions. It can, therefore, be assumed that they represent components of *R* genes [27]. They also contain several conserved domains that can be utilised to optimise alignment between different sequences. The genetic relationships were based on a neighbour joining distance analysis of the P-loop to the GLPL region. Neighbour joining analysis was used as it is comparable to more computational intensive analysis when used on smaller datasets [32].

Phylogenetic analysis grouped the different classes of *TNL* genes into three distinct clades (Fig. 3). Clade II contained all the sequences that were isolated in this study. In this clade the observed sequence divergence was less than expected (Fig. 6). This was also seen in the short branch lengths that are evident when this clade is compared with clades I and II. In these two clades the observed sequence divergence, was much more apparent. This is similar to other studies where the branch lengths reflect the *TNL* and *CNL* sequence diversity [23; 8; 24].

Clade III grouped away from clades I and II (Fig. 3). This is also obvious from the motif analysis of the *TNL* genes (Fig. 5). Clade III contained the sequences designated by Meyers *et al.* [25] as the TX and TL. None of these sequences contained any NBS or LRR motifs and were comprised of only the TIR regions. The only exception to this clade is the sequence previously designated as TPL (*P. taeda* AAM28914). This sequence contains the two TIR motifs that are followed by P-loop and the LRR motifs (Fig. 5 and Table 4). This sequence contained none of the other NBS motifs. The P-loop was not detected in our dataset during the MEME analysis. This could be due to specifying that the maximum number of motifs to identify should be ten. As the P-loop motif was not present in several of the sequences that were analysed, the preset number was reached before it was called.

The *P. taeda* AAM28907 sequence was previously designated as being of the TNL class. However, we could not find any of the NBS motifs prevalent in this sequence (Fig. 4). It contained only the two TIR motifs in series as TIR-1, TIR-2, and again TIR-1. Therefore, based on the assembled information there is insufficient evidence to class it into the TNL group and it would be preferable to refer to it as TX.

A vast difference was observed between the conservation of the RNBS-B of *Pinus* compared with several other plant species. However, the changes in amino acid sequence observed did not have a great effect on the polarity of this region. The conserved sequence for other plant species is effectively neutral over the whole motif, with a double negative charge (ED) to the carboxyl side. In *Pinus*, the amino acids also code for a neutral motif. The only difference observed here is that there is one positively charged amino acid (R) followed by a negatively charged amino acid (D). Upon analysis of the several types of NBS genes that have been described and compared, it is interesting that the kinase domains contribute to the nucleotide binding capacity of the domain. We believe that the slight change in the RNBS-B domain in *Pinus* would not be detrimental to the function of this domain.

The non-NBS homologue sequences obtained in this study shared several similarities to the NBS homologues (Fig. 7). These sequences were isolated from the *Pinus* genome as they had sufficient shared similarity to the P-loop and GLPL regions for the degenerate primers to detect them. Several of these sequences contained the known kinase-2 and RNBS-B motifs, and shared homology to the regions between the RNBS-B and GLPL region. Several other motifs present in the NBS homologues were also found in these sequences in different combinations. Although several of

them are interspersed with stop codons it could be that they are remnants of NBS homologues, as this has been previously reported [8; 24]. To further understand these sequences it would be feasible to assemble the motifs present in them into a hidden Markov model (HMM) to identify similar sequences from public databases. This might prove useful in identifying older versions of the *TNL* that might differ significantly from the current consensus for NBS genes.

NBS homologues are common in *Pinus*. Several sequences were isolated from different *Pinus* species through the use of degenerate primers. Several NBS homologues have also been isolated in different studies by different groups. When comparing these NBS homologues to those in *Arabidopsis*, it was found that they formed a separate monophyletic group closely related to the *Physcomitrella patens* homologues [25]. Therefore, it is assumed that in conifers, these analogues have developed independently [25]. Thus, a study on *Pinus* homologues would be valuable in the search for *TNL* and *CNL* ancestors found in older plants. Unfortunately, the current drive for sequencing of a specific genome is determined by economic benefit. This, coupled with the costs involved in sequencing a conifer genome, has meant that a full genome sequence for conifer is not yet available. This makes the identification of conserved regions in ancient plants, and degenerate primer design from these, the only current option to study these genes. However, the need exists to extend the primer design to the conserved regions found in the TIR as well as the MHD. Amplification of the sequence between these two domains would enable better analysis of the NBS region and heighten the inferences that can be drawn from this domain that is found in several identified *R* genes. This might also accelerate gene discovery and contribute towards breeding for disease tolerance in conifers.

There are several opportunities to study the *TNL* and *CNL* genes that are present in the *Pinus* genome. To further our understanding of *Pinus* and more ancient *TNL* and *CNL* genes it will be necessary to obtain additional full length sequences. If this is achieved, it would be possible to analyse the full complement of *TNL* and *CNL* genes as has been done for the *Arabidopsis* Col-0 genome [24]. It is evident that the division between the *TNL* and *CNL* genes occurred before the divergence of gymnosperms and angiosperms [8]. Therefore, the isolation of *TNL* and *CNL* genes from the *Pinus* genome will further assist in inferring evolutionary relationships of these genes. This also implies that the need exists to find and study these homologues in more ancient plants than the gymnosperms. Knowledge gained from these plants will heighten our understanding of the evolution of *R* genes in economically important plants.

Acknowledgements

We thank Komatiland Forests Research, the National Research Foundation and the Mellon Foundation for financial assistance. Komatiland Forests Research, Global Forest Products, and Mondi Ltd. supplied plant material used in this experiment.

```
                      10        20        30        40        50        60
                ....*....|....*....|....*....|....*....|....*....|....*....|
consensus         1 EVNTVRSS-GLVPDESTVVGREDMVEAVIEKLLEMS---ENLGVVGIVGMGGVGKTTLAK 56
Pinus           186 vkgvlekvkrplnvstyptgldekvaeferamllnqqssretrVIGIVGLGGAGKTTLAK 245
gi 2258315      153 KQETRTPS-TSLVDDSGIFGRKNEIENLVGRLLSMDtkrKNLAVVPIVGMGGMGKTTLAK 211
gi 862904       220 DIWSHISKeNLILETDELVGIDDHITAVLEKLSLDS---ENVTMVGLYGMGGIGKTTTAK 276
gi 548086       138 TDGGSIQVtCREIPIKSVVGNTTMMEQVLEFLSEe----EERGIIGVYGPGGVGKTTLMQ 193
gi 2443884      137 SQPPPRSE-VEERPTQPTIGQEDMLEKAWNRLME-----DGVGIMGLHGMGGVGKTTLFK 190
gi 1513144     1078 YVAPSFSA-YTQRANEEMEGFQDTIDELKDKLLGGS---PELDVISIVGMPGLGKTTLAK 1133
gi 699495       196 DNGTDRWS-SPVYDHTQVVGLEGDKRKIKEWLFRSNd--SQLLIMAFVGMGGLGKTTIAQ 252
gi 2258317      154 KLETRTPS-TSLIDEPDIFGRQSEIEDLIDRLLSEGasgKNLTVVPIVGMGGLGKTTLAK 212
gi 1931650      124 EQKEIRQT-FANSSESDLVGVEQSVEALAGHLVEND----NIQVVSISGMGGIGKTTLAR 178

                      70        80        90       100       110       120
                ....*....|....*....|....*....|....*....|....*....|....*....|
consensus        57 QIYNDFS-VGGHFDSVAWVVVSKTYTE----FDLQKT-ILQELGSED---ESW------- 100
Pinus           246 EIFNRR---RSNYKQIYFQSDVRENAA----RKSLKS-LQKLILKHLakvDVG------- 290
gi 2258315      212 AVYNDER-VQKHFGLTAWFCVSEAYDA----FRITKG-LLQEIGSTD---LKAddnlnql 262
gi 862904       277 AVYNK---ISSCFDCCCFIDNIRETQEkdgvVVLQKKlVSEILRIDS---GSV------- 323
gi 548086       194 SINNELItKGHQYDVLIWVQMSREFGE----CTIQQA-VGARLGLS------W------- 235
gi 2443884      191 KIHNKFAeIGGTFDIVIWIVVSKGVMI----SKLQED-IAEKLHLCD---DLW------- 235
gi 1513144     1134 KIYNDPE-VTSRFDVHAQCVVTQLYSW----RELLLT-ILNDVLEPsd----------- 1175
gi 699495       253 EVFNDKE-IEHRFERRIWVSVSQTFTE----EQIMRS-ILRNLGDAS---V--------- 294
gi 2258317      213 AVYNDES-VKNHFDLKAWFCVSEAYNA----FRITKG-LLQEIGSID---L--------- 254
gi 1931650      179 QVFHHDM-VQRHFDGFAWVFVSQQFTQ----KHVWQR-IWQELQPQN---GDI------- 222

                     130       140       150       160       170       180
                ....*....|....*....|....*....|....*....|....*....|....*....|
consensus       101 -----DHKNEGE-LAVKLKELLLRKRFLLVLDDVWEKE--DWDDIGVPFPD-GENGSRVI 151
Pinus           291 -----IDSTDDG-IEKLRRYLQSSHAFLI-LDDVDHFQ--QVEALLLPVRDvVRQGSLIL 341
gi 2258315      263 qvklkADDNLNQ-LQVKLKEKLNGKRFLVVLDDVWNDNypEWDDLRNLFLQ-GDIGSKII 320
gi 862904       324 -----GFNNDSG-GRKTIKERVSRFKILVVLDDVDEKF--KFEDMLGSPKD-FISQSRFI 374
gi 548086       236 -----DEKETGEnRALKIYRALRQKRFLLLLDDVWEEI--DLEKTGVPRPD-RENKCKVM 287
gi 2443884      236 -----KNKNESD-KATDIHRVLKGKRFVLMLDDIWEKV--DLEAIGIPYPS-EVNKCKVA 286
gi 1513144     1176 ----rNEKEDGE-IADELRRFLLTKRFLILIDDVWDYK--VWDNLCMCFSD-VSNRSRII 1227
gi 699495       295 ------GDDIGT-LLRKIQQYLLGKRYLIVMDDVWDKNlsWWDKIYQGLPR-GQGGS-VI 345
gi 2258317      255 -----VDDNLNQ-LQVKLKERLKEKKFLIVLDDVWNDNynEWDELRNVFVQ-GDIGSKII 307
gi 1931650      223 -----SHMDEHI-LQGKLFKLLETGRYLVVLDDVWKEE--DWDRIKAVFPR-KR-GWKML 272

                     190       200       210       220       230       240
                ....*....|....*....|....*....|....*....|....*....|....*....|
consensus       152 VTTRSESVAGRMG-GTSKPHEVESLEPEESWELFSNKVFPKNLP-SEHPELEEVAKEIVE 209
Pinus           342 VTSRNKDILKRSGiAESFIFTVTGLSPPQSRELFCCHAFDQRYP---PIEFEQVVEEFVH 398
gi 2258315      321 VTTRKESVALMMD---SGAIYMGILSSEDSWALFKRHSLEHKDP-KEHPEFEEVGKQIAD 376
gi 862904       375 ITSRSMRVLGTLNenqCKLYEVGSMSKPRSLELFSKHAFKKNTP---PSYYETLANDVVD 431
gi 548086       288 FTTRSIALCNNMGa--EYKLRVEFLEKKHAWELFCSKVWRKDl--LESSSIRRLAEIIVS 343
gi 2443884      287 FTTRSREVCGEMGd--HKPMQVNCLEPEDAWELFKNKVGDNTl--SSDPVIVELAREVAQ 342
gi 1513144     1228 LTTRLNDVAEYVKc-eSDPHHLRLFRDDESWTLLQKEVFQGE---SCPPELEDVGFEISK 1283
gi 699495       346 VTTRSESVAKRVQardDKTHRPELLSPDNSWLLFCNVAFAANDGtCERPELEDVGKEIVT 405
gi 2258317      308 VTTRKDSVALMMG---NEQISMGNLSTEASWSLFQRHAFENMDP-MGHSELEEVGRQIAA 363
gi 1931650      273 LTSRNEGVGIHADp-kSFGFKTRILTPEESWKLCEKIVFHRRDEtGTLSDMEAMGKEMVT 331

                     250       260       270       280       290       300
                ....*....|....*....|....*....|....*....|....*....|....*....|
consensus       210 KCKGLPLALKVLGGLLASKS-TV-PEWKHVLEQLNNELAG-----RNLNEVLSILSLSYD 262
Pinus           399 ACADLPLSLKVLGGLLRGKD-LK-YGKE-LLHDISKNLPQ-----D----IMATLKISYD 446
gi 2258315      377 KCKGLPLALKALAGMLRSKS-EV-DEWRNILRSEIWELPs------CSNGILPALMLSYN 428
gi 862904       432 TTAGLPLTLKVIGSLL-FKQ-EI-AVWEDTLEQLRRTL--------NLDEVYDRLKISYD 480
gi 548086       344 KCGGLPLALITLGGAMAHRE-TE-EEWIHASEVLTRFPAem---kgM-NYVFALLKFSYD 397
gi 2443884      343 KCRGLPLALNVIGETMSSKT-MV-QEWEHAIHVFNTSAAef---sdMQNKILPILKYSYD 397
gi 1513144     1284 SCRGLPLSVVLVAGVLKQKKkTL-DSWKVVEQSLSSQRIg------sLEESISIIGFSYK 1336
```

```
gi 699495    406 KCKGLPLTIKAVGGLLLCKD-HVyHEWRRIAEHFQDELRgn---tsETDNVMSSLQLSYD 461
gi 2258317   364 KCKGLPLALKTLAGMLRSKS-EV-EEWKCILRSEIWELR--------DNDILPALMLSYN 413
gi 1931650   332 CCGGLPLAVKVLGGLLATKH-TV-PEWKRVYDNIGPHLAgrssldNLNSIYRVLSLSYE 389

                      310       320       330       340
                 ....*....|....*....|....*....|....*....|...
consensus    263 NLPM-HLKRCFLYLALFPEDYDIPVKQLISLWIAEGFVEAENE 304
Pinus        447 SLGK-IEKQIFLDIVCFFIGED--KDTAIRIWDGtewggslrf 486
gi 2258315   429 DLPA-HLKQCFAYCAIYPKDYQFRKEQVIHLWIANGLVHQFHS 470
gi 862904    481 ALNP-EAKEIFLDIACFFIGQN--KEEPYYMWTDCNFYPASNI 520
gi 548086    398 NLESdLLRSCFLYCALFPEEHSIEIEQLVEYWVGEGFLTSSHG 440
gi 2443884   398 SLGDeHIKSCFLYCALFPEDGEIYNEKLIDYWICEGFIGEDQV 440
gi 1513144  1337 NLPH-YLKPCFLYFGGFLQGKDIHVSKMTKLWVAEGFVQANNE 1378
gi 699495    462 ELPS-HLKSCILTLSLYPEDCVIPKQQLVHGWIGEGFVMWRNG 503
gi 2258317   414 DLPA-HLKRCFSFCAIFPKDYPFRKEQVIHLWIANGLVPVEDE 455
gi 1931650   390 NLPM-CLKHCFLYLAHFPEYYEIHVKRLFNYLAAEGIITSSDD 431
```

Figure 1: Amino acid alignment of *P. radiata* NBS region to the conserved NB-ARC sequence (pfam00931.8) at GenBank's Conserved Domain Database (CDD). Bold text represents the conserved P-loop (KTTL) and GLPL regions found in most NBS-sequences. The consensus sequence is derived from deposited data found at CDD. *Pinus* is the *P. radiata* NBS sequence. Cursive bold text indicates the region (SRIIITTR) that the reverse primer (NBS-R1) binds to [42].

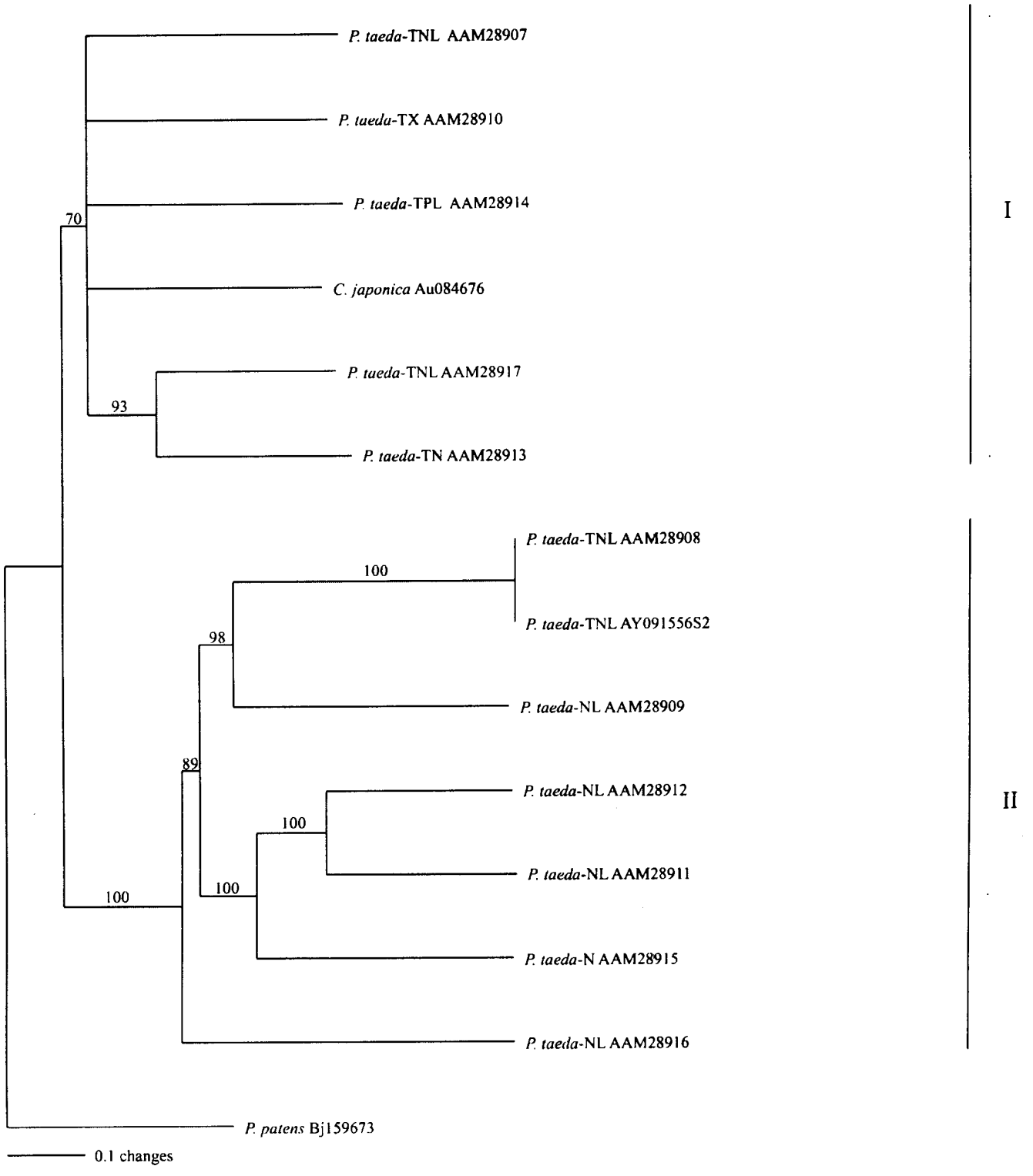


Figure 2: Tree resulting from the first analysis, using the neighbour joining algorithm, of amino acid sequences from previously classified *TNL* genes [25]. Bootstrap values are calculated from 1000 replicates and values higher than 70 % were retained. Abbreviations used for the different classes can be found in Table 1. The tree is rooted to the *Physcomitrella patens* NBS homologue.

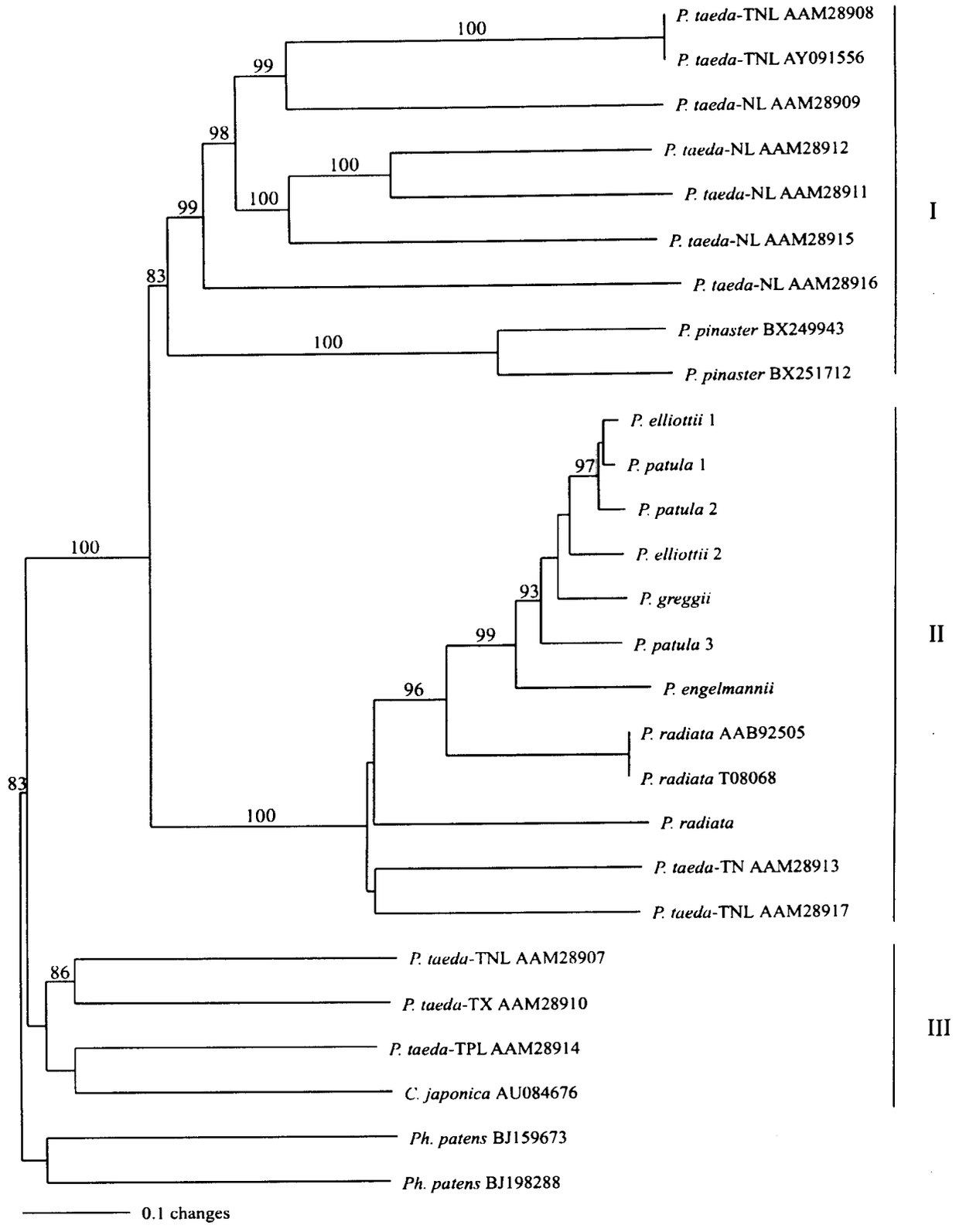


Figure 3: Distance tree produced from a neighbour joining analysis of the aligned *Pinus* amino acid sequences. The tree includes sequences from *Cryptomeria japonica* and *Physcomitrella patens* as they were identified by Meyers *et al.* [25] to be closely related. Bootstrap values are calculated from 1000 replicates and values higher than 70 % were retained. Abbreviations used for the different classes can be found in Table 1. The tree is rooted to *Ph. patens* NBS homologues.

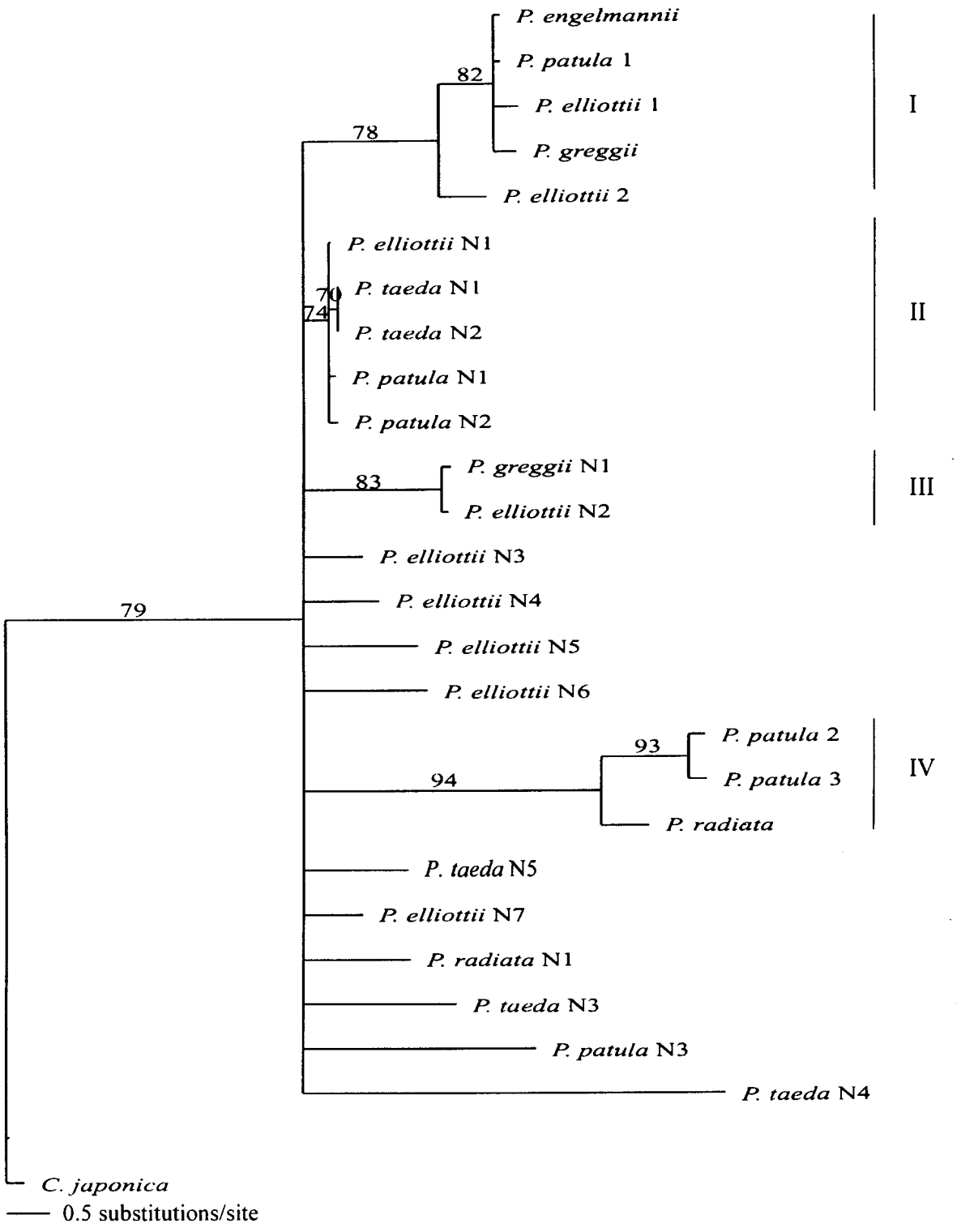


Figure 4: Distance tree produced from the analysis of DNA sequences of the different fragments amplified in this study. Several of these sequences displayed no homology to known sequences. Bootstrap values are calculated from 1000 replicates where only branches that had values higher than 70 % were kept. The letter N followed by a number indicates the sequences with no known homology and the number of sequences identified from a specific species. The tree is rooted to *C. japonica* (AU084676).

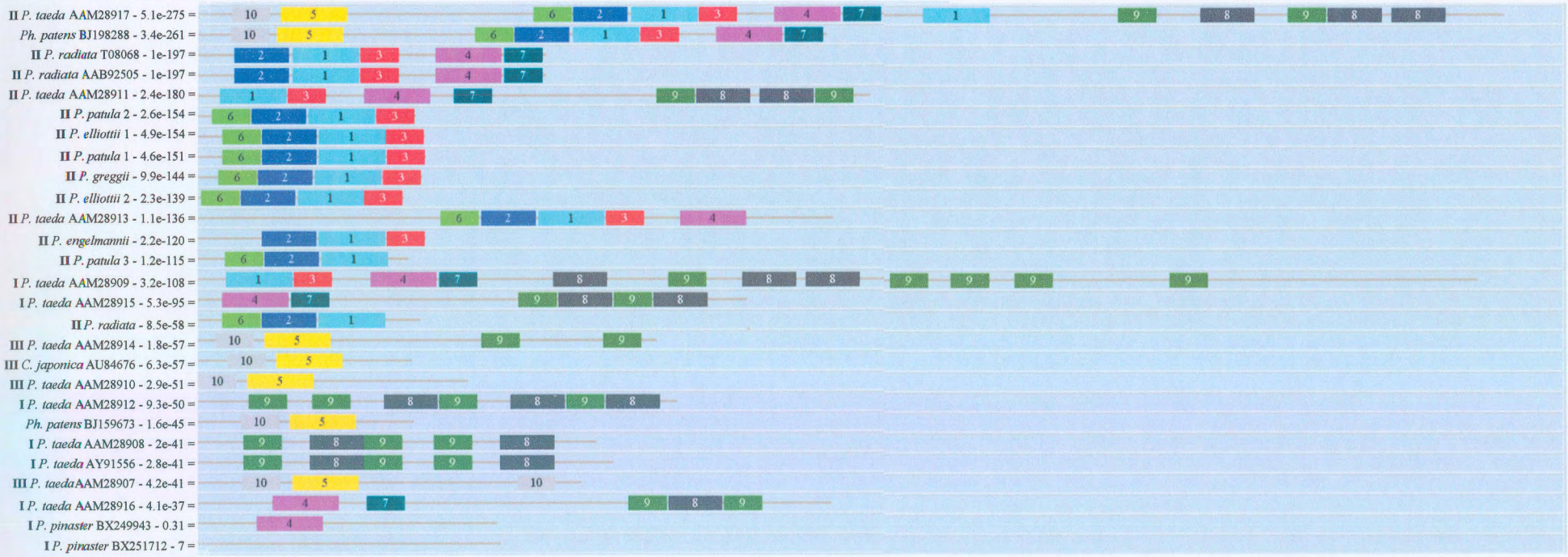


Figure 5: Motif patterns identified in conifer NBS homologues. Different numbered boxes are the different motifs identified by the MEME analysis [4] and they are displayed by MAST [5]. The ten motifs and their sequences are summarised in Table 4. E-values indicate the strength of the match of the sequence to all the motifs. Bold numbers indicate the clade a sequence belongs to in Figure 3.

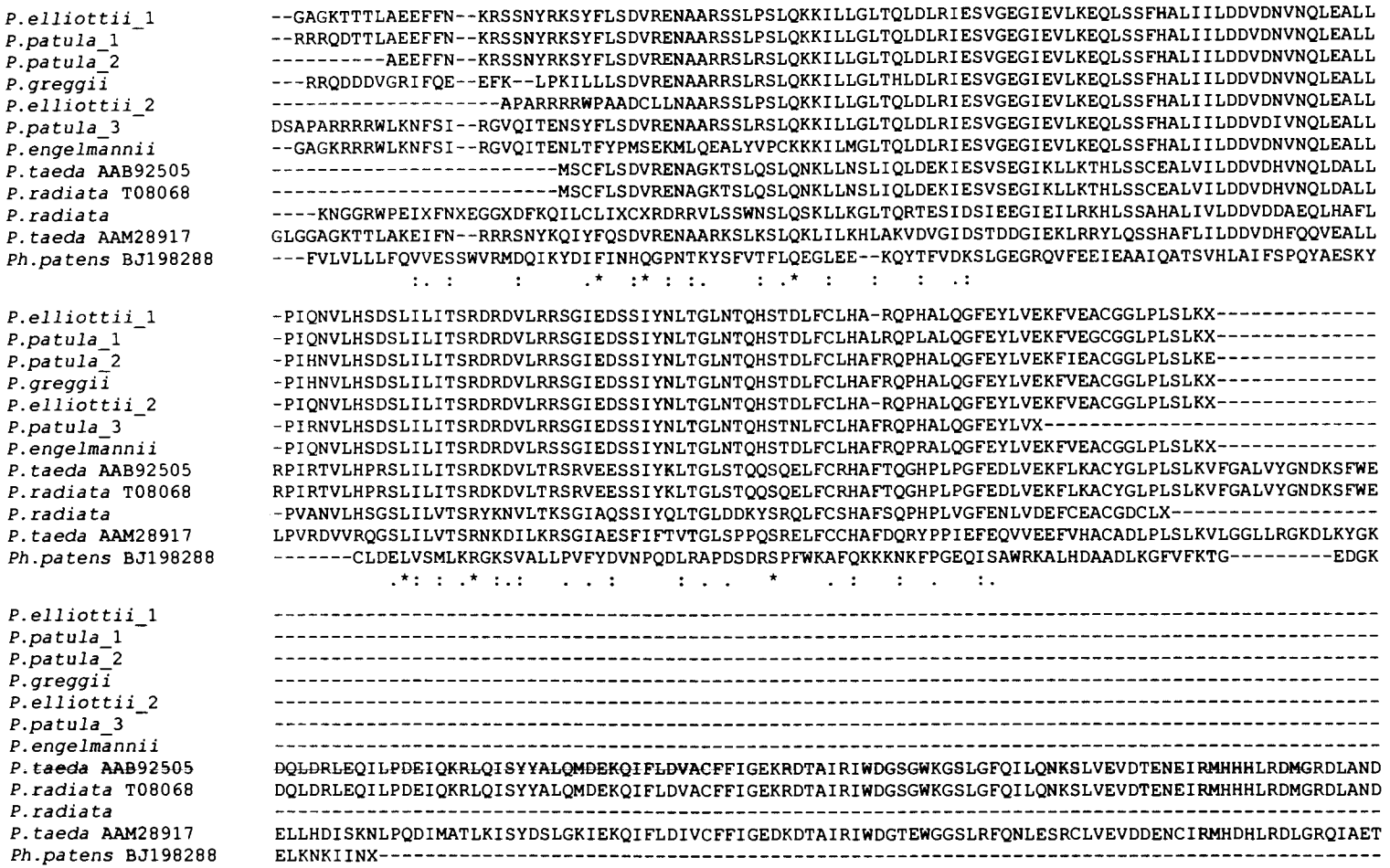


Figure 6: Amino acid alignment of sequences in clade II as presented in Figure 3. The *Ph. patens* sequence was also included in this alignment, as it had the same motifs as found in this clade.

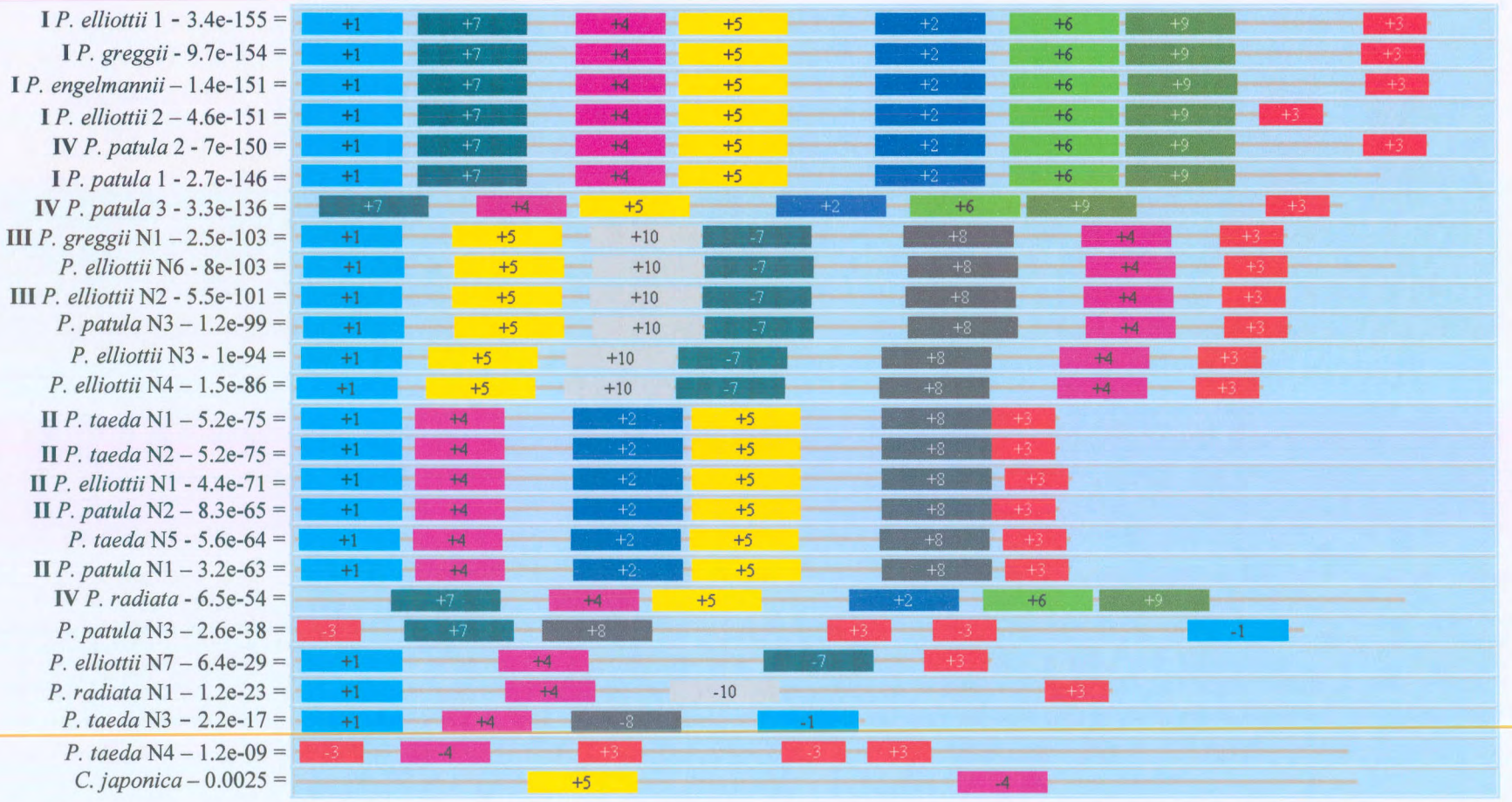


Figure 7: Motif patterns identified in nucleotide sequences of NBS and non-NBS sequences. Different numbered boxes are the different motifs identified by the MEME analysis [4] and they are displayed by MAST [5]. The +/- in front of the number indicates the orientation of the motif. The ten motifs and their sequences are summarised in Table 5. E-values indicate the strength of the match of the sequence to all the motifs. Bold numbers indicate the clade a sequence belongs to in Figure 4.

Table 1: Classes of NBS genes as described by Meyers *et al.* [25].

Abbreviation	Class
NL	Nucleotide binding site Leucine rich repeat
TNL	Toll/Interleukin-1 receptor Nucleotide binding site Leucine rich repeat
TN	Toll/Interleukin-1 receptor Nucleotide binding site
TPL	Toll/Interleukin-1 receptor P-loop Leucine rich repeat
TX	Toll/Interleukin-1 receptor

Table 2: NBS sequences used in the phylogenetic analysis.

Accession No.	Identity	Species	Reference
BX249943	--	*Pinus pinaster*	--
BX251712	--	''	--
AAM28908	TNL	*P. taeda*	[25]
AY091556	TNL	''	''
AAM28912	NL	''	''
AAM28911	NL	''	''
AAM28915	NL	''	''
AAM28909	NL	''	''
AAM28916	NL	''	''
	Putative TNL	*P.elliottii*	Present study
	''	*P. patula*	''
	''	*P. patula*	''
	''	*P. elliottii*	''
	''	*P. greggii*	''
	''	*P. patula*	''
	''	*P. engelmannii*	''
AAB92505	Kinase –P-loop	*P. radiata*	--
T08068	N-gene-homologue	''	--
BJ198288	--	*Physcomitrella patens*	[25]
	--	*P. radiata*	Present study
AAM28913	TN	*P. taeda*	[25]
AAM28917	TNL	''	''
	--	*P. patula*	Present study
AAM28907	TNL	*P. taeda*	[25]
AAM28910	TX	''	''
AAM28914	TPL	''	''
AU084676	--	*Cryptomeria japonica*	''
BJ159673	--	*Ph. patens*	''

Table 3: Degenerate primers designed to amplify NBS homologues. The primers in this study were designed from the P-loop (KTTL) and GLPL region of the *Pinus* NBS homologous region. This primer design incorporates a 5` GC-rich consensus clamp that stabilises the binding of the primer, with a more degenerate core to allow for better stringency during amplification [31].

Name	Region	Sequence*	Reference
KTTL-F	GGAGKTTLAK	5`-cggcgccgccaaraciaciaciytigc	Present study
LPL2	VKLSLPLGACA	5`-cctcagagacagaggcagtcciccrcaigc	``
NBS-F1	GMGGVKT	5`-ggaatgggnggngtnggnaarac	[42]
NBS-R1	SRIIITTR	5`-yctagttgtraydatdayyytrc	``

*D = A/G/T, I = A/C/T, N = A/C/G/T, R = A/G, Y = C/T

Table 4: Amino acid motifs aligned and displayed in Figure 5 using MAST analysis. Putative names assigned to the different motifs. Amino acids that are similar to the annotation provided for the *Arabidopsis* consensus sequences [24] are underlined.

Motif no.	Motif width (amino acids)	Motif similarity	Best possible match
1	50	RNBS-B	PIRNVLHPDSLILITSRDRDVLRRSGIEDSSIYNLTGLNTQHSTELFCCH
2	41	Kinase-2	TQLDERIESVGEGIEVLKEQLSSFHALIILDDVDHVNQLEA
3	29	GLPL	FRQPHPLPGFEYLVEKFVEACGGLPLSLK
4	50	RNBS-D	IQKRLQISYYALQMDEKQIFLDVACFFIGEKRDTAIRIWDGSGWKGSLGF
5	50	TIR-2	GEYIWPQIEAAIRTAWVHIAIFSPGYAESMWCLDELLWMFKSGGTIIPVF
6	29	RNBS-A	YRKSYFLSDVRENAARSSLPSLQKKILLG
7	29	MHDV	LQNRCLVEVDTENEIRMHHHLRDMGRDLA
8	41	LRR (end)	LPDPFGNLKNLRHINMSFCPKLKQLPDSFGNLTHLQHIDMH
9	29	LRR	CWQLKYLPDAFGNLNNLQHINMSGCPGLE
10	29	TIR-1	FINHRGPDTKKTFAKHLYRRLREHGLRVF

Table 5: Nucleotide motifs aligned and displayed in Figure 7 using MAST analysis.

Motif	Motif width (bp)	Best possible match
1	46	TCAGAGACAGAGGCAGTCCGCCACAGGCTGCCGCAAACTCATAAGC
2	50	ATTATCTACATCATCTAAAATTATTAAAGCATGAAAGGATGAGAGTTGTT
3	29	TTCAGCCAACGTCGTCGTCTTGCCGGCGC
4	50	AGGCCTGTTAGATTGTGAATTGATGAATGCCGGATTCCTGA
5	50	CAAAACATTGCTGTCCCGAGATGTAATCAAGATTACAATATCGGAGTTGA
6	50	TCTATGCCTTCACCAACGCTCTCTATTCGAAGATCCAATTGAGTCAAACC
7	50	TCGAACCCTTGAAGTGCGAGGGGTTGTCTGAAGGCATGCAAGCAAAAGAG
8	50	TTCGGTAGGTTTGACGGTCAACAAACCCGGCGTGCAGCCCAACTCATAGT
9	50	AAGAATTTTTTTTTGCAGGGAACGTAGAGAGCTTCTTGCAGCATTTTCTC
10	50	ACGGCTCTCCAGCGCCACCTGCACCGCGTCTTCCTTGCTGGTCACGGTTT

References

1. **Altschul SF, Thomas LM, Alejandro AS, Jinghui Z, Zheng Z, Webb M, Lipman DJ** (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Research **25:** 3389-3402

2. **Bai J, Pennill LA, Ning J, Lee SW, Ramalingam J, Webb CA, Zhao B, Sun Q, Nelson JC, Leach JE, Hulbert SH** (2002) Diversity in nucleotide binding site-leucine-rich repeat genes in cereals. Genome Research **12:** 1871-1884

3. **Bailey TL, Elkan C** (1994) Fitting a mixture model by expectation maximization to discover motifs in biopolymers. *In* Proceedings of the Second International Conference on Intelligent Systems for Biology. AAAI Press, pp 28-36

4. **Bailey TL, Elkan C** (1995) The value of prior knowledge in discovering motifs with MEME. *In* Proceedings of the Second International Conference on Intelligent Systems for Biology 3. AAAI Press, pp 21-29

5. **Bailey TL, Gribskov M** (1998) Combining evidence using p-values: Application to sequence homology searches. Bioinformatics **14:** 48-54

6. **Bent AF** (1996) Plant disease resistance genes: function meets structure. Plant Cell **8:** 1757-1771

7. **Cannon SB, Young ND** (2003) The genomic architecture and evolution of NBS-LRRs. *In* G Stacey, NT Keen, eds, Plant-microbe interactions, Ed 6 Vol 6. APS Press, St. Paul

8. **Cannon SB, Zhu H, Baumgarten AM, Spangler R, May G, Cook DR, Young ND** (2002) Diversity, distribution, and ancient taxonomic relationships within the TIR and non-TIR NBS-LRR resistance gene subfamilies. Journal of Molecular Evolution **54:** 548-562

9. **Dangl JL, Jones JD** (2001) Plant pathogens and integrated defence responses to infection. Nature **411:** 826-833

10. **Ellis J, Dodds P, Pryor T** (2000) The generation of plant disease resistance gene specificities. Trends in Plant Science **5:** 373-379

11. **Ellis J, Jones J** (1998) Structure and function of proteins controlling strain-specific pathogen resistance in plants. Current Opinion in Plant Biology **1:** 288-293

12. **Ellis JG, Lawrence GJ, Luck JE, Dodds PN** (1999) Identification of regions in alleles of the flax rust resistance gene *L* that determine differences in gene-for-gene specificity. Plant Cell **11:** 495-506

13. **Fluhr R** (2001) Sentinels of disease. Plant resistance genes. Plant Physiology **127:** 1367-1374

14. **Hammond-Kosack K, Jones JDG** (2000) Responses to plant pathogens. *In* B Buchanan, W Gruissem, R Jones, eds, Biochemistry and molecular biology of plants. American Society of Plant Physiologists, pp 1102-1156

15. **Hulbert SC, Webb CA, Smith SM, Sun Q** (2001) Resistance gene complexes: Evolution and utilization. Annual Review of Phytopathology **39:** 285-312

16. **Kanazin V, Marek LF, Shoemaker RC** (1996) Resistance gene analogs are conserved and clustered in soybean. Proceedings of the National Academy of Sciences, USA **93:** 11746-11750

17. **Kobe B, Deisenhofer J** (1994) The leucine-rich repeat: a versatile binding motif. Trends in Biochemical Science **19:** 415-421

18. **Kobe B, Deisenhofer J** (1995) A structural basis of the interactions between leucine-rich repeats and protein ligands. Nature **374:** 183-186

19. **Leister D, Ballvora A, Salamini F, Gebhart C** (1996) A PCR-based approach for isolating pathogen resistance genes from potato with potential for wide application in plants. Nature Genetics **14:** 421-429

20. **Luck JE, Lawrence GJ, Dodds PN, Shepherd KW, Ellis JG** (2000) Regions outside of the leucine-rich repeats of flax rust resistance proteins play a role in specificity determination. Plant Cell **12:** 1367-1378

21. **Lupas A** (1996) Coiled coils: new structures and new functions. Trends in Biochemical Science **21:** 375-382

22. **Meyers BC, Dickerman AW, Michelmore RW, Sivaramakrishnan S, Sobral BW, Young ND** (1999) Plant disease resistance genes encode members of an ancient and diverse protein family within the nucleotide-binding superfamily. Plant Journal **20:** 317-332

23. **Meyers BC, Dickerman AW, Michelmore RW, Sivaramakrishnan S, Sobral BW, Young ND** (1999) Plant disease resistance genes encode members of an ancient and diverse protein family within the nucleotide-binding superfamily. Plant Journal **20:** 317-332

24. **Meyers BC, Kozika A, Griegoa A, Kuanga H, Michelmore RW** (2003) Genome-wide analysis of NBS-LRR–encoding genes in *Arabidopsis*. Plant Cell **15:** 809-834

25. **Meyers BC, Morgante M, Michelmore RW** (2002) TIR-X and TIR-NBS proteins: two new families related to disease resistance TIR-NBS-LRR proteins encoded in *Arabidopsis* and other plant genomes. Plant Journal **32:** 77-92

26. **Nimchuk Z, Rohmer L, Chang JH, Dangl JL** (1999) Knowing the dancer from the dance: *R*-gene products and their interactions with other proteins from host and pathogen. Current Opinion in Plant Biology **4:** 288-294

27. **Pan Q, Liub Y-S, Budai-Hadriana O, Sela M, Carmel-Gorenc L, Zamirc D, Fluhr R** (2000) Comparative genetics of nucleotide binding site-leucine rich repeat resistance gene homologues in the genomes of two dicotyledons: tomato and *Arabidopsis*. Genetics **155:** 309-322

28. **Pan Q, Wendel J, Fluhr R** (2000) Divergent evolution of plant NBS-LRR resistance gene homologues in dicot and cereal genomes. Journal of Molecular Evolution **50:** 203-213

29. **Qureshi ST, Gros P, Malo D** (1999) Host resistance to infection: genetic control of lipopolysaccharide responsiveness by Toll-like receptor genes. Trends in Genetics **15:** 291-294

30. **Rice P, Longden I, Bleasby A** (2000) EMBOSS: The European Molecular Biology Open Software Suite. Trends in Genetics **16:** 276-277

31. **Rose TM, Schultz ER, Henikoff JG, Pietrookovski S, McCallum CM, Henikoff S** (1998) Consensus-degenerate hybrid oligonucleotide primers for amplification of distantly related sequences. Nucleic Acids Research **26:** 1628-1635

32. **Rzhetsky A, Nei M** (1993) Theoretical foundation of the minimum-evolution method of phylogenetic inference. Molecular Biology and Evolution **10:** 1073-1095

33. **Saraste M, Sibbalda PR, Wittinghofer A** (1990) The P-loop – a common motif in ATP- and GTP-binding proteins. Trends in Biochemical Science **15:** 430-434

34. **Staskawicz BJ, Mudgett MB, Dangl JL, Galan JE** (2001) Common and contrasting themes of plant and animal diseases. Science **292:** 2285-2289

35. **Steenkamp ET, Wingfield BD, Coutinho TA, Wingfield MJ, Marasas WFO** (1999) Differentiation of *Fusarium subglutinans* f. sp. *pini* by histone gene sequence data. Applied and Environmental Microbiology **65:** 3401-3406

36. **Swofford D** (2000) PAUP*, Phylogenetic analysis using parsimony. MA: Sinauer, Sunderland

37. **Thomas CM, Jones DA, Parniske M, Harrison K, Balint-Kurti PJ, Hatzixanthis K, Jones JDG** (1997) Characterization of the tomato *Cf-4* gene for resistance to *Cladosporium fulvum* identifies sequences that determine recognitional specificity in *Cf-4* and *Cf-9*. Plant Cell **9:** 2209-2224

38. **Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG** (1997) The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Research **25:** 4876-4882

39. **Traut TW** (1994) The functions and consensus motifs of nine types of peptide segments that form different types of nucleotide-binding sites. European Journal of Biochemistry **222:** 9-19

40. **Whitham S, Dinesh-Kumar SP, Choi D, Hehl R, Corr C, Baker B** (1994) The product of the tobacco mosaic virus resistance gene N: similarity to toll and the interleukin-1 receptor. Cell **78:** 1101-1115

41. **Xiao S, Ellwood S, Calis O, Patrick E, Li T, Coleman M, Turner JG** (2001) Broad-spectrum mildew resistance in *Arabidopsis thaliana* mediated by *RPW8*. Science **291:** 118-120

42. **Yu YG, Buss GR, Maroof MAS** (1996) Isolation of a superfamily of candidate disease-resistance genes in soybean based on a conserved nucleotide-binding site. Proceedings of the National Academy of Sciences, USA **93:** 11751-11756

43. **Zhu H, Cannon SB, Young ND, Cook DR** (2002) Phylogeny and genomic organization of the TIR and non-TIR NBS-LRR resistance gene family in *Medicago truncatula*. Molecular Plant-Microbe Interactions **15:** 529-539

Summary

The main objective of this thesis was the elucidation of the host-pathogen interaction between *Pinus patula* and *Fusarium circinatum*. This was accomplished by studying differential gene expression at the molecular level. Therefore, the first chapter reports the use of PCR-based methods in gene discovery and transcriptome analysis. The use of these techniques in the identification of novel transcripts in host-pathogen interactions is addressed. These examples illustrate the differences and strong features of each technique.

Chitinases are linked to defence responses in plants. In chapter two, the induction of chitinases in *P. patula* was assessed at both the protein and genetic level. Western blot analysis and enzyme activity assays indicate that chitinase enzyme is not detected as part of the defence response by *P. patula* after infection by *F. circinatum*. This was further confirmed by the lack of significant induction of two *Pinus* chitinase genes, *LP6* and *PSCHI4*, as determined by RT-PCR analysis. Partial DNA sequence homologues for the *LP6* and *PSCHI4* genes were determined and compared with a variety of plant chitinases. The low levels of detectable chitinase induction in *P. patula* might explain the high levels of susceptibility to the pitch canker fungus observed in seedlings of this tree.

Pinus patula, the most widely planted species in South Africa, is highly susceptible to infection by *F. circinatum*. In chapter three, suppression subtractive hybridisation was used to elucidate the changes taking place at the molecular level early on in this interaction. Most of the identified transcripts shared homology to both biotic and abiotic stress in plants. The induction of one fragment, displaying homology to phytocyanin proteins, was followed through RT-PCR. Induction levels for this fragment differed significantly between less and more susceptible plants. Although most of the sequences isolated in this study can be linked to stress, most have not been linked with specific plant-pathogen interactions. This raises questions in regard to the function of these genes in host-pathogen interactions. Further research to identify the function of these sequences in the defence response will be needed. These sequences can also be tested against a family of *Pinus* trees to ascertain if they will be useful in marker assisted selection.

A molecular analysis of culture degeneration and pathogenicity of *F. circinatum* was attempted in chapter four. In this chapter, the differential induction of transcripts in *F. circinatum* was determined against several other *Fusarium* spp. Several of the identified fragments shared homology with stress related proteins. One transcript shared homology to a polyketide synthase, *FUM5*, that could be linked to fumonisin production in other *Fusarium* spp. ELISA detected no fumonisin production, although the *FUM5* transcripts were detected. The identification of all the transcripts could provide a basis for more intensive gene discovery studies in *F. circinatum* and other *Fusarium* spp. The induction of these sequences in different isolates needs to be studied to prove their function in *F. circinatum*. This study also complements several other studies that studied the morphological characteristics of culture degeneration.

Resistance gene analogues have been reported from a diverge set of plant species. In chapter five, degenerate PCR amplification was used to isolate TIR-NBS-LRR like resistance gene analogues from a range of *Pinus* species. These sequences were further characterised through comparative analysis with previously reported *Pinus* resistance gene analogues. Through motif analysis, several of the known conserved motifs found in NBS domains were identified and conservation with other plant NBS motifs is indicated. The P-loop and GLPL motifs displayed a high level of conservation on amino acid level with other plant NBS motifs. However, slight differences in several of the conserved regions were observed when the *Pinus* analogues were compared with *Arabidopsis thaliana*. The identification of differences between angiosperm and gymnosperm NBS sequences indicates that design of new degenerate probes and primers for the isolation of more ancient NBS sequences is needed. Further, phylogenetic and structural analyses of these sequences will also aid in understanding the relationship between angiosperm and gymnosperm NBS sequences. The knowledge gained from such a study will highlight the similarities between angiosperm and gymnosperm defence responses. This study represents the first detailed report on RGA in *Pinus*.