Diverse sources of infection and cryptic recombination revealed in South African *Diplodia pinea* populations

Wubetu Bihon[†], Bernard Slippers[†], Treena Burgess[†*], Michael J. Wingfield[†] and Brenda D. Wingfield[†]

[†]Department of Genetics, Forestry and Agricultural Biotechnology Institute, University of Pretoria, Lunnon Road, 0002 Pretoria, South Africa; [*]School of Biological Sciences and Biotechnology, Murdoch University, Perth 6159, Australia.

Corresponding author: wubetu.bihon@fabi.up.ac.za, Phone: +27 12 420 3938; Fax: +27 12 420 3960

Abstract

This study considers the population diversity and structure of *D. pinea* in South Africa at different spatial scales from single trees to plantations, as well as comparing infections on healthy and diseased trees. A total of 236 isolates were characterized using thirteen microsatellite markers. Analysis of these markers confirmed previous results that *D. pinea* has a high level of gene and genotypic diversity in South Africa, with the latter values ranging from 6 % to 68 % for the different plantations. The data also reflect a fungus with randomly associated alleles in populations at local plantation scales and for the population as a whole. These results suggest that recombination is occurring in *D. pinea* and that it most likely has cryptic sexual state. The study also reveals the sources of endophytic infection and stress related disease out-breaks as diverse infections that have occurred over a long time period. In contrast, wound-associated die-back appears to be caused by clones of the pathogen occurring in narrow time frames.

Keywords: Population genetics, forest pathology, invasive fungal pathogen, cryptic sexual reproduction, SSR markers

Introduction

A number of pine species were introduced into South Africa from Europe in the early 1700's and to the area known as the Western Cape Province (Lundquist 1987; Burgess and Wingfield 2001). Since that time, pines have been extensively planted in the KwaZulu-Natal, Mpumalanga and Northern Province, as well as the Eastern and Western Cape Provinces. Concomitant with the expansion of pine plantations from south to north along the east coast, there have been reports of disease caused by the pathogen *Diplodia pinea* (=*Sphaeropsis sapinea*). Thus, Diplodia shoot blight and die-back, the disease attributable to this fungus, was first reported in 1909 from the Eastern Cape, then in the 1927 in the Northern Province (Lundquist 1987) and subsequently throughout the country (Swart and Wingfield 1991).

Diplodia pinea is a haploid, opportunistic plant pathogen of mostly *Pinus* spp., but also occasionally found on other coniferous trees. It has been reported as causing significant losses to pine stands in many parts of the world (Burgess and Wingfield 2001, 2002). However, it owes its notoriety to the extensive damage that it causes after hail damage in South Africa (Swart *et al*. 1985; Swart and Wingfield 1991).

Diplodia species like their relatives in the Botryosphaeriaceae (Crous *et al*. 2006; Slippers and Wingfield 2007) are latent pathogens, mainly causing disease when trees are subjected to environmental or other stress conditions (Swart and Wingfield 1991; Smith *et al*. 1996; Stanosz *et al*. 1997). The predisposing factors favouring infection by *D. pinea* include wounding through hail and pruning, water stress and other environmental extremes (Swart and Wingfield 1991; Smith *et al*. 2002; Blodgett and Bonello 2003). In South Africa, substantial loss of timber has resulted from disease caused by *D. pinea* following hail damage. *Pinus radiata* is particularly susceptible and has been abandoned as a plantation species in areas receiving

summer rainfall where hail is common (Lundquist 1987; Zwolinski *et al*. 1990; Swart and Wingfield 1991).

Previously, the population diversity of South African *D. pinea* (morphotype A) isolates from seed cones has been studied using Vegetative Compatibility Groups (VCG's) (Smith *et al*. 2000). Subsequently, simple sequence repeats (SSR) markers and VCGs were used to characterize the diversity of 26 and 54 isolates of this species from South Africa, respectively (Burgess *et al*. 2001; Burgess *et al*. 2004). Because *D. pinea* is an introduced pathogen in South Africa, it was surprising that these studies showed a higher genetic diversity in *D. pinea* populations in South Africa than those from elsewhere in the world (Burgess *et al* 2004). However, the sample sizes used in the SSR analyses were insufficient to establish whether there was any structure to the South African population. While VCGs were useful in these studies to understand genetic diversity, it was not possible to consider gene flow and relative relationships between populations (Glass *et al*. 2000). Most recently, Bihon *et al*. (2010a) used 13 microsatellite markers to show that a high level of diversity was also evident in individual, healthy, mature trees.

In populations of an organism where a sexual state has not been found, it is difficult to explain the presence of high genotypic diversity (Kohli and Kohn 1998). Populations that reproduce only asexually are expected to exhibit a higher degree of clonality, with few genotypes at higher frequencies (Chen and McDonald 1996; Kohli and Kohn 1998). The explanation for the observed high levels of genotypic diversity, despite expected exclusive asexual reproduction of *D. pinea* in South Africa, was that there had been multiple introductions from a wide variety of sources, consistent with forestry practice in the country (Smith *et al*. 2000; Burgess *et al* 2004). The possibility that the genotypic diversity could also result from cryptic sexual reproduction

has, however, not been tested. In this regard, genetic recombination can be estimated directly by showing the presence of sexual structures or indirectly by measuring non-random associations among loci and using genealogical approaches (Moore *et al*. 2009).

The aim of this study was to address three key questions regarding the genotypic diversity of *D. pinea* in South Africa. Firstly, to determine how the diversity of the fungus is distributed spatially, within a single plantation and between geographically separated plantations. A second aim was to consider where the fungus reproduces exclusively asexually or whether and undetected, a cryptic sexual cycle might also occur. Finally, we considered whether there is a distinction between the population diversity of isolates associated with disease outbreaks and those collected as endophytes from healthy pine tissue.

Materials and Methods

Sample collection and isolation

Sampling of asymptomatic and symptomatic (die-back) *P. patula* trees was conducted in the two major pine-growing provinces or regions of South Africa, namely KwaZulu-Natal and Mpumalanga (Fig. 1). Samples were collected from branches on 3 to 5 year-old trees as well as branches, stems and cones on mature trees older than 12 years of age. Samples from Balgown were all from asymptomatic trees and those from the Boston plantation were from trees displaying die-back after water stress. One segment of a plantation at Seven Oaks (KwaZulu-Natal province) had suffered hail damage before sampling and die-back symptoms were obvious on most of the trees. Samples at this site were thus collected from die-back symptoms after hail damage. Others from the Mpumalanga province were collected from a mixture of die-back and

asymptomatic trees. The samples were maintained at 4 ^{0}C in the laboratory and isolations were made within three days after collection.

Isolations were made from stems and branches as described previously by Stanosz *et al* 1997. Cultures were incubated under continuous light at 25 ^{o}C. After 4 to 6 days of incubation, cultures with a white and fluffy mycelium typical of *D. pinea* were sub-cultured onto 2 % Water Agar (2 % m/v Biolab agar) with two autoclaved pine needles on the agar surface to stimulate the production of pycnidia. After two to three weeks, single spore isolates or isolates from hyphal tips were made and each of the cultures is maintained in the culture collection (CMW) of the Forestry and Agricultural Biotechnology Institute (FABI), University of Pretoria.

DNA Extraction, PCR amplification and separation of SSR loci

Cultures were grown on MEA in Petri dishes for two weeks and mycelium scraped from the surface of the plates for DNA extraction as described by de Wet *et al*. (2003). Thirteen SSR loci were amplified for all isolates as described previously (Bihon *et al*. 2010a & b; Santana *et al*. 2009). Thirteen fluorescently labelled SSR-PCR products were multiplexed and one micro-litre of these multiplexed PCR products was separated on ABI Prism 3100 Genetic analyzer. The amplicon peaks were determined based on the four fluorescent dyes used and the sizes of the DNA fragments. The mobilities of SSR products were compared to those of internal size standards (LIZ-500) and allele sizes were estimated by GeneScan 2.1 and GeneMapper 3.7 computer software (Applied Biosytems).

Gene and genotypic diversity

For each of the loci, individual alleles were assigned a different letter. For each isolate, a data matrix of 13 multistate characters (one for each locus) was compiled (eg. AABDCGDDABDDE). The frequency of each allele at each locus for the entire and clone-corrected populations was calculated, and allele diversity determined using the program POPGENE (Yeh *et al*. 1999) and the equation $H = 1 - \sum_k x^2_k$, where x_k is the frequency of the k^{th} allele (Nei 1973). Chi-square test was conducted to evaluate the differences in the frequencies of alleles at each locus after removing the duplicated multilocus haplotypes (clone correction). Genotypic diversity (G) was estimated using the equation $G = {}^{1}\!\Big/\!_{\sum pi^2}$ where *pi* is the observed frequency of the i^{th} genotype in the population (Stoddart and Taylor 1988).

Population differentiation and clustering

Clone corrected populations were analyzed using the program Multilocus (Agapow and Burt 2000), with an estimate of Wright's, F_{ST} as Θ = Q-q/1-q to calculate population differentiation theta (Θ) where Q is the probability that two alleles from the same population are the same and q is the probability that two alleles from different populations are the same (Weir 1997). When Θ = 0, allele frequencies between populations are equal and thus these populations are interpreted as being identical. However, when Θ = 1, frequencies of alleles are unique to a population and the populations are isolated from each other sharing no common alleles. The level of difference between the observed value and 1000 times randomized data sets were determined by the probability value at $P < 0.05$. The null-hypothesis that there is no population differentiation was rejected when significant difference between populations at alpha (P) ≤ 0.05 was observed. Gene

flow between populations were also determined using the formula $M = [(1/\Theta)-1]/2$ (Cockerham and Weir 1993) where M = gene flow. Population differentiation or genetic distance and approximate geographic distances were compared between populations in order to consider whether there was a correlation between genetic diversity and physical distance.

Population structure was inferred and assigned in STRUCTURE 2.2 that clusters individuals into *K* distinct populations (clusters) and permits mixed ancestry (Pritchard *et al*. 2000). The program was run for 100,000 replicates of Monte Carlo Markov Chain (MCMC) analysis after an initial burn-in of 20,000 for *K* ranging from 1 to 15 at 20 iterations. A Cluster identity of >75 % was used to assign clusters.

Index of Association

Association of alleles for a clone-corrected *D. pinea* population was inferred by calculating the Index of Association (I_A) and rBarD ($\bar{r}D$) using the program Multilocus (Agapow & Burt 2000). The observed values of I_A and $\bar{r}D$ were compared to 1000 times randomized data sets at a $P \leq 0.05$ level of significance. When the observed I_A and $\bar{r}D$ value was significantly different from the randomized data sets at ($P \leq 0.05$), the null hypothesis that the alleles are randomly associated and the population is expected to freely undergo recombination, was rejected. Such a result indicates linkage between the loci, which is unlikely in this case, or a clonally reproducing population. In contrast, if the observed I_A and $\bar{r}D$ fell within the distribution range, the null hypothesis of random association of alleles would be supported.

Analysis of molecular variance (AMOVA)

AMOVA, using 9999 permutations, was conducted to differentiate the sources of variation between and within populations, as well as between regions or provinces, using the software GeneAlEx version 6.2 (Peakall and Smose 2006). The null hypothesis was that variation within populations, between populations and among regions is equally responsible for the total genetic diversity in South African *D. pinea* strains.

Results

Sample collection and isolation

A total of 238 isolates were obtained from seven locations (Fig. 1). Fifty-nine isolates were collected from Seven Oaks pine plantation (KwaZulu-Natal province) where most trees had been damaged by a hail storm and die-back was evident. In addition, isolates were collected from die-back on trees after drought stress at Boston (56) and from asymptomatic trees at Balgown (26) and Vryheid (13). In Mpumalanga province, isolates were obtained from samples collected in Sabie (47), Ngodwana (28), and Lothair (9), representing a mixture from trees with die-back symptoms and from asymptomatic trees.

Gene and genotypic diversity

The 13 SSR markers produced a total of 43 alleles among 236 individuals with a minimum of 18 alleles (Ngodwana) and a maximum of 29 alleles (Boston). Isolates from Balgown had 28 alleles and those from Sabie 25 alleles. Most of the loci were polymorphic within plantations (Fig. 2). The percentage of polymorphic loci ranged from 30.8 % in Ngodwana to 92.3 % in the Boston collections. There were three unique alleles in Balgown, three in Seven-Oaks, two in Lothair-

Vryheid and one each in Boston, Sabie and Ngodwana (Fig. 2). Only 9 of the 43 alleles were shared between all the plantations (Table 1). The thirteen loci each contained two to five alleles when considering all the isolates. Two alleles from each locus were dominant within populations, while the rest were rarely found.

There was high genotype diversity in most of the South African *D. pinea* populations sampled and genotype diversity differed significantly between these populations (Table 1). The percentage diversity ranged from 6 % at Seven-Oaks to 68 % at Balgown. The lowest genotype diversity was obtained for the population acquired from hail-damaged and diseased trees in the Seven Oaks plantation, while the highest genotype diversity was obtained for a population from asymptomatic trees at Balgown. Isolates from Boston that had been collected from die-back symptoms following drought stress showed 43.7 % diversity. In total, all the isolates were represented by 122 haplotypes, of which 36.7 % were detected only once (Table 1). From the clone corrected populations, Chi-square values indicated that there were highly significant differences (alpha $\leq$ 0.001) for gene diversity at most SSR loci between populations, with the exception of locus SS8 which was not significantly different (alpha $\leq$ 0.05) (Table 2). The Lothair and Vryheid populations were subsequently considered collectively because the gene and genotype diversities were not significantly different at any of the SSR loci and the population differentiation (Θ) was not significantly different between them.

Population differentiation

Population differentiation (Θ) values differed significantly (alpha $\leq$ 0.05) between the six South African populations of *D. pinea* (Table 3). Moreover, there was no correlation between the genetic and geographic distances with $r^2 = 0.07$. High levels of genetic distance were found even

between the most closely situated populations (e.g. Sabie and Ngodwana, Θ = 0.631), but there were also data points with average genetic distances at greater physical distances (e.g. Sabie and Boston, Θ = 0.212) (Table 3).

All isolates of *D. pinea* from South Africa resided in six clusters when analyzed using the program STRUCTURE 2.2 and *LnK* had a maximum of K = 6 (Fig. 3a & b). This implied that entire population represents six sub-populations. These six groups also reflect the geographic origins of most of the isolates. However, there were some isolates residing in a population that did not match their geographic origin. For example, the Boston population of isolates was separated into two clusters, one of which included isolates from the Lothair-Vryheid population.

Index of Association

In the assessment of random mating for the entire South African *D. pinea* population, the Index of Association (I_A) and rbarD ($\bar{r}D$) values were 0.373 and 0.031, respectively, and they all fell within the values obtained from 1000 randomized datasets (P = 0.973 for all isolates combined) (Fig. 4). The I_A and $\bar{r}D$ for each of the six populations were also not significantly different from those of the randomized datasets (Table 4), suggesting that the null hypothesis there is recombination of genotypes is supported.

Molecular variance

Genetic variation for *D. pinea* isolates was partitioned within and among populations and provinces using analysis of molecular variance (AMOVA). The result indicated that 55 % of the variation was due to that among populations and 45 % due to variation within populations. There was no contribution to total genetic variation resulting from variation among regions (Table 5).

Discussion

The results of this study revealed high levels of genetic and genotypic diversity among six spatially separated populations of *D. pinea* in South Africa with little gene flow between these populations (Table 3). The results also suggest that the high level of genotypic diversity is, at least in part, due to the presence of a cryptic sexual stage, another form of recombination, or even possibly new alleles arising through mutation. Interestingly, levels of diversity were significantly lower in a population of *D. pinea* associated with disease after hail damage, compared to a population derived from endophytic infections on healthy trees.

The total number of alleles, unique alleles and percent polymorphic loci reflect a high level of gene diversity for *D. pinea* in South Africa. Nei's (1973) mean gene diversity (*H*) for all of populations was also high, ranging from 0.194 to 0.443. Similarly, there was high genotypic diversity in most of the populations and overall in the South Africa *D. pinea* population, ranging from 6 % to 68 %. This is surprising for an introduced pathogen that is believed to reproduce exclusively asexually (McDonald and McDermott 1993; Milgroom 1996; Taylor *et al*. 1999). These high levels of gene and genotypic diversity in *D. pinea* populations are consistent with those found previously (Smith *et al*. 2000; Burgess *et al*. 2001; Burgess *et al*. 2004). For example, Burgess *et al*. (2001) reported 71 % genotypic diversity in South African populations. Similar studies using VCG and SSR markers also suggested very high levels of diversity for *D. pinea* in South Africa (Smith *et al*. 2000; Burgess *et al*. 2004). These results raise the question as to how such high diversity could have arisen in the non-mating populations of *D. pinea*.

Mutation could have played a role in increasing genetic diversity in *D. pinea* populations in South Africa over the past century. However, this source of variation would be expected to be small in terms of the total diversity observed and relatively short period of time that the fungus

has been present in South Africa (McDonald and McDermott 1993). Potential evidence for mutation is derived from the private or unique alleles found in the geographically defined populations of *D. pinea*. These alleles might also represent remnants of diversity introduced since the establishment of plantations into these regions. Genetic drift and selection would have been expected to eliminate many such unique alleles from the populations over time (Halliburton 2004), unless they confer a fitness benefit to the fungus (FitzSimmons *et al*. 1997; Milgroom and Cortesi 1999; Burgess *et al*. 2004).

In all the local populations considered in this study, as well as the South African population as a whole, alleles at the different loci were randomly associated. This result is considered evidence of recombination as has been shown in other fungi (Geiser *et al*. 1994; Arie *et al*. 2000; Groenewald *et al*. 2008; Pérez *et al*. 2010). No sexual structures have been observed in *D. pinea*, despite the fact that it has been studied intensively for more than a century (Swart and Wingfield 1991). However, spermatia like spores were observed in this fungus (Wingfield and Knox-Davies 1980; Palmer *et al*. 1987). Parasexuality, one of the mechanisms of mitotic genotype recombination in fungi (Milgroom 1996, Taylor *et al*. 1999) cannot be considered as a potential reason for the allelic recombination in *D. pinea*, because all reported parasexual recombination has been between individuals of the same VCGs. In *D. pinea* there is a high diversity of VCGs (Burgess *et al*. 2001) and parasexual recombination between them is unlikely. Hence, our result suggests that *D. pinea* may have a rare, cryptic sexual state in its life cycle. This would be reasonable explanation for the high levels of genotypic diversity observed in South Africa, and elsewhere, in this apparently asexual (Burgess *et al*. 2004) fungus.

A higher level of genetic variation was observed in endophytic populations of *D. pinea,* compared to populations obtained from die-back symptoms after hail damage. Isolates with the

highest level of genotype diversity (68.4 %) were from asymptomatic *P. patula* trees at Balgown, while the lowest levels of diversity were observed for isolates obtained from die-back trees after a hail storm (e.g. in Seven Oaks at 5.8 %). Only one site was sampled after hail damage and more sampling should be done to substantiate this observation. However, these results can be explained by the fact that endophytic communities of *D. pinea* result from numerous individual infections over the life-time of trees. Furthermore, this diverse endophytic community is involved in causing disease when plants are under stress in the absence of wounding. When wide-scale concurrent wounding occurs such as during hail storms, additive infections from asexual conidia leading to disease could occur resulting in a more uniform population. Consistent with our results, Dakin *et al*. (2010) reported that all eight isolates of *Neofusicoccum australe* from asymptomatic *Agonis flexuosa* were in different vegetative compatibility groups (VCGs), while 26 isolates from a cankered tree included eleven VCG's. This would suggest that wounding is involved in the infection process because it provides an easy entry for *D. pinea* (Flowers *et al*. 2006). An alternative explanation for the lower diversity of the fungus after hail damage or in cankers could be competition and selection among the endophytic strains. Strains that are relatively active at the time when trees are stressed (Smith *et al* 2002) would grow faster and take over large portions of the tree than strains that are in an opposite manner.

Genetic diversity was significantly differentiated between all populations. STRUCTURE analyses indicated that, isolates resided in six distinct populations, which were mostly consistent with the geographic origins of the isolates. Furthermore, AMOVA analysis showed that 55 % of the variation was distributed amongst populations. These data suggest that little gene flow occurs between populations in plantations, even those that are geographically relatively close to each other (approximately 65 km). This low rate of gene flow between populations might be

explained by the fact that the conidia of *D. pinea* are spread mainly by rain splash (Swart *et al*. 1987; Swart and Wingfield 1991) and hence have limited potential for long distance dispersal. This main mode of spread was inferred from the fact that in months of the year where there was no rainfall, virtually no conidia were collected inside plantations (Swart *et al*. 1987). There was, however, some evidence of movement of genotypes between populations, such as those from Boston and Lothair and Vryheid. These exceptions most likely reflect occasional long distance dispersal due to human intervention by moving infected wood, cones or live plants.

This study has shown that new genotypes of *D. pinea* are generated via cryptic recombination that results higher genotypic diversity of South African populations. Furthermore, the study has provided new insights into the role of endophytic infections and wounds in disease outbreaks. The exact mechanism giving rise to recombination in *D. pinea* will require intensive further study. We suggest that there is a need to support the evidence of natural recombination by characterizing genes associated with mating and presence of sexual structure in nature.

Acknowledgements

This research was supported financially by the DST/NRF Centre of Excellence in Tree Health Biotechnology (CTHB) and members of the Tree Protection Co-operative Program (TPCP), South Africa.

References

Agapow PM, Burt A, 2000. 'Multilocus 1.2' (Department of Biology, Imperial College: Ascot, UK).

Arie T, Kaneko I, Yoshida T, Noguchi M, Nomura Y, Yamaguchi I, 2000. Mating-type genes from asexual phytopathogenic ascomycets *Fusarium oxysporum* and *Alternaria alternata*. *Molecular Plant-Microbe Interactions* **13**: 1330 – 1339.

Bihon W, Burgess T, Slippers B, Wingfield MJ, Wingfield BD, 2010a. Distribution of *Diplodia pinea* and its genotypic diversity within asymptomatic *Pinus patula* trees. *Australasian Plant Pathology* (in press).

Bihon W, Slippers B, Burgess T, Wingfield MJ, Wingfield BD, 2010b. Sources of *Diplodia pinea* endophytic infections in *Pinus patula* and *P. radiata* seedlings in South Africa. *Forest Pathology* doi: 10.1111/j.1439-0329.2010.00691.x.

Blodgett JT, Bonello P, 2003. The aggressiveness of *Sphaeropsis sapinea* on Austrian pine varies with isolate group and site of infection. *Forest Pathology* **33**: 15 – 19.

Burgess T, Wingfield MJ, 2001. Exotic pine forestry in the Southern Hemisphere: A brief history of establishment and quarantine practices. *South African Forestry Journal* **192**: 79 – 84.

Burgess T, Wingfield BD, Wingfield MJ, 2001. Comparison of genotypic diversity in native and introduced populations of *Sphaeropsis sapinea* isolated from *Pinus radiata*. *Mycological Research* **105**: 1331 – 1339.

Burgess T, Wingfield MJ, 2002. Quarantine is important in restricting the spread of exotic seed-borne tree pathogens in the southern hemisphere. *International Forestry Review* **4**: 56 – 65.

Burgess TI, Wingfield MJ, Wingfield BD, 2004. Global distribution of *Diplodia pinea* genotypes revealed using simple sequence repeat (SSR) markers. *Australasian Plant Pathology* **33**: 513 – 519.

Chen RS, McDonald BA, 1996. Sexual reproduction plays a major role in the genetic structure of populations of the fungus. *Mycosphaerella graminicola*. *Genetics* **142:** 1119 – 1127**.**

Cockerham CC, Weir BS, 1993. Estimation of gene flow from *F*-statistics. *Evolution* **47**: 855 – 863.

Crous PW, Slippers B, Wingfield MJ, Rheeder J, Marasas WFO, Philips AJL, Alves A, Burgess T, Barber P, Groenewald JZ, 2006. Phylogenetic lineages in the *Botryosphaeriaceae*. *Studies in Mycology* **55**: 235 – 253.

Dakin N, White D, Hardy GE, Burgess TI, 2010. The opportunistic pathogen, *Neofusicoccum australe*, is responsible for crown dieback of peppermint (*Agonis flexuosa*) in Western Australia. *Australasian Plant Pathology* **39**: 202 – 206.

De Wet J, Burgess T, Slippers B, Preisig O, Wingfield BD, Wingfield MJ, 2003. Multiple gene genealogies and microsatellite markers reflect relationships between morphotypes of *Sphaeropsis sapinea* and distinguish a new species of *Diplodia*. *Mycological Research* **107**: 557 – 566.

FitzSimmons NN, Moritz C, Limpus CJ, Pope L, Prince R, 1997. Geographic structure of mitochondrial and nuclear gene polymorphisms in Australian green turtle populations and male-biased gene flow. *Genetics* **147**: 1843 – 1854.

Flowers JL, Hartman JR, Vaillancour LJ, 2006. Histology of *Diplodia pinea* in diseased and latently infected *Pinus nigra* shoots. *Forest Pathology* **36**: 447 – 459.

Geiser DM, Arnold ML, Timberlake WE, 1994. Sexual origins of British *Aspergillus nidulans* isolates. *Proceedings of the National Academy of Sciences*, USA **91**: 2349 – 2352.

Glass NL, Jacobson DJ, Shiu PKT, 2000. The genetics of hyphal fusion and vegetative incompatibility in filamentous ascomycete fungi. *Annual Review of Genetics* **34**: 165 – 168.

Groenewald M, Linde CC, Groenewald JZ and Crous PW, 2008. Indirect evidence for sexual reproduction in *Cercospora beticola* populations from sugar beet. *Plant Pathology* **57**: 25 – 32.

Halliburton R (2004) Introduction to population genetics. Pearson Education, Inc., USA.

Kohli Y, Kohn LM, 1998. Random association among alleles in clonal populations of *Sclerotinia sclerotiorum*. *Fungal Genetics and Biology* **23**: 139 – 149.

Lundquist JE, 1987. A history of five forest diseases in South Africa. *South African Forestry Journal* **140**: 51 – 59.

McDonald BA, McDermott JM, 1993. Population Genetics of Plant Pathogenic Fungi: Electrophoretic markers give unprecedented precision to analyses of genetic structure of populations. *BioScience* **43**: 311 – 319.

Milgroom MG, 1996. Recombination and multilocus structure of fungal populations. *Annual Review of Phytopathology* **34**: 457 – 477.

Milgroom MG, Cortesi P, 1999. Analysis of population structure of the chestnut blight fungus based on vegetative incompatibility genotypes. *Proceeding of the National Academy of Sciences*, USA **96**: 10518 – 10523.

Moore GG, Singh R, Horn BW, Carbone I, 2009. Recombination and lineage-specific gene loss in the aflatoxin gene cluser of *Aspergillus flavus*. *Molecular Ecology* **18**: 4870 – 4887.

Nei M, 1973. Analysis of gene diversity in subdivided populations. *Proceedings of the National Academy of Sciences* USA **70**: 3321 – 3323.

Palmer MA, Stewart EL, Wingfield MJ, 1987. Variation among isolates of *Sphaeropsis sapinea* in the north central United States. Phytopathology **77**: 944 – 948.

Peakall R, Smouse PE, 2006. GENALEX 6: Genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes* **6**: 288 – 295.

Pérez G, Slippers B, Wingfield BD, Huntera GC, Wingfield MJ, 2010. Micro- and macrospatial scale analyses illustrates mixed mating strategies and extensive gene flow in populations of an invasive haploid pathogen. *Molecular Ecology* **19**: 1801 – 1813.

Pritchard JK, Stephens M, Donnelly P, 2000. Inference of Population Structure Using Multilocus Genotype Data. *Genetics* **155**: 945 – 959.

Santana QC, Coetzee MPA, Steenkamp ET, Mlonyeni OX, Hammond GNA, Wingfield MJ, Wingfield BD 2009. Microsatellite discovery by deep sequencing of enriched genomic libraries. *Biotechniques* **46**, 217–223.

Slippers B, Wingfield MJ, 2007. *Botryosphaeriaceae* as endophytes and latent pathogens of woody plants: Diversity, ecology and impact. *Fungal Biology Reviews* **21**: 90 – 106.

Smith H, Wingfield MJ, Cours PW, Coutinho TA, 1996. *Sphaeropsis sapinea* and *Botryosphaeria dothidea* endophytic in Pinus spp. and Eucalyptus spp. in South Africa. *South African Journal of Botany* **62**: 86 – 88.

Smith H, Wingfield MJ, Coutinho TA, 2002. The role of latent *Sphaeropsis sapinea* infections in post-hail associated die-back of *Pinus patula*. *Forest Ecology and Management* **164**: 177 – 184.

Smith H, Wingfield MJ, de Wet J, Coutinho TA, 2000. Genotypic diversity of *Sphaeropsis sapinea* from South Africa and Northern Sumatra. *Plant Disease* **84**, 139 – 142.

Stanosz GR, Smith DR, Guthmiller MA, Stanosz JC, 1997. Persistence of *Sphaeropsis sapinea* on or in asymptomatic red and jack pines. *Mycologia* **89**: 525 – 530.

Stoddart JA, Taylor JF, 1988. Genotype diversity: Estimation and prediction in samples. *Genetics* **118**: 705 – 711.

Swart WJ, Knox-Davis PS, Wingfield MJ, 1985. *Sphaeropsis sapinea,* with special reference to its occurrence on *Pinus* spp. in South Africa. *South Africa Forestry Journal* **35**: 1 – 8.

Swart WJ, Wingfield MJ, 1991. Biology and control of *Sphaeropsis sapinea* on *Pinus* species in South Africa. *Plant Disease* **75**: 761 – 766.

Swart WJ, Wingfield MJ, Knox-Davies PS, 1987. Conidial dispersal of *Sphaeropsis sapinea* in three climatic regions of South Africa. *Plant Disease* **71**: 1038 – 1040.

Taylor JW, Jacobson DJ, Fisher MC, 1999. The evolution of asexual fungi: Reproduction, speciation and classification. *Annual Review of Phytopathology* **37**: 197 – 246.

Weir BS, 1997. Genetic data analysis II. Sinauer Associates Inc.: Sunderland, MA.

Wingfield MJ, Knox-Davies PS, 1980. Association of *Diplodia pinea* with a root disease of pines in South Africa. Plant Disease **64**: 221 – 223.

Yeh FC, Yang RC, Boyle T, 1999. POPEGENE version 1.31 (Microsoft windows based freeware for population genetic analysis, Alberta).

Zwolinski JB, Swart WJ, Wingfield MJ, 1990. Economic impact of a post-hail outbreak of dieback induced by *Sphaeropsis sapinea*. *European Journal of Forest Pathology* **20**: 405 – 411.

Figure 1. Locations where samples were collected, MP is Mapumalanga and KZN is Kwa-Zulu Natal provinces

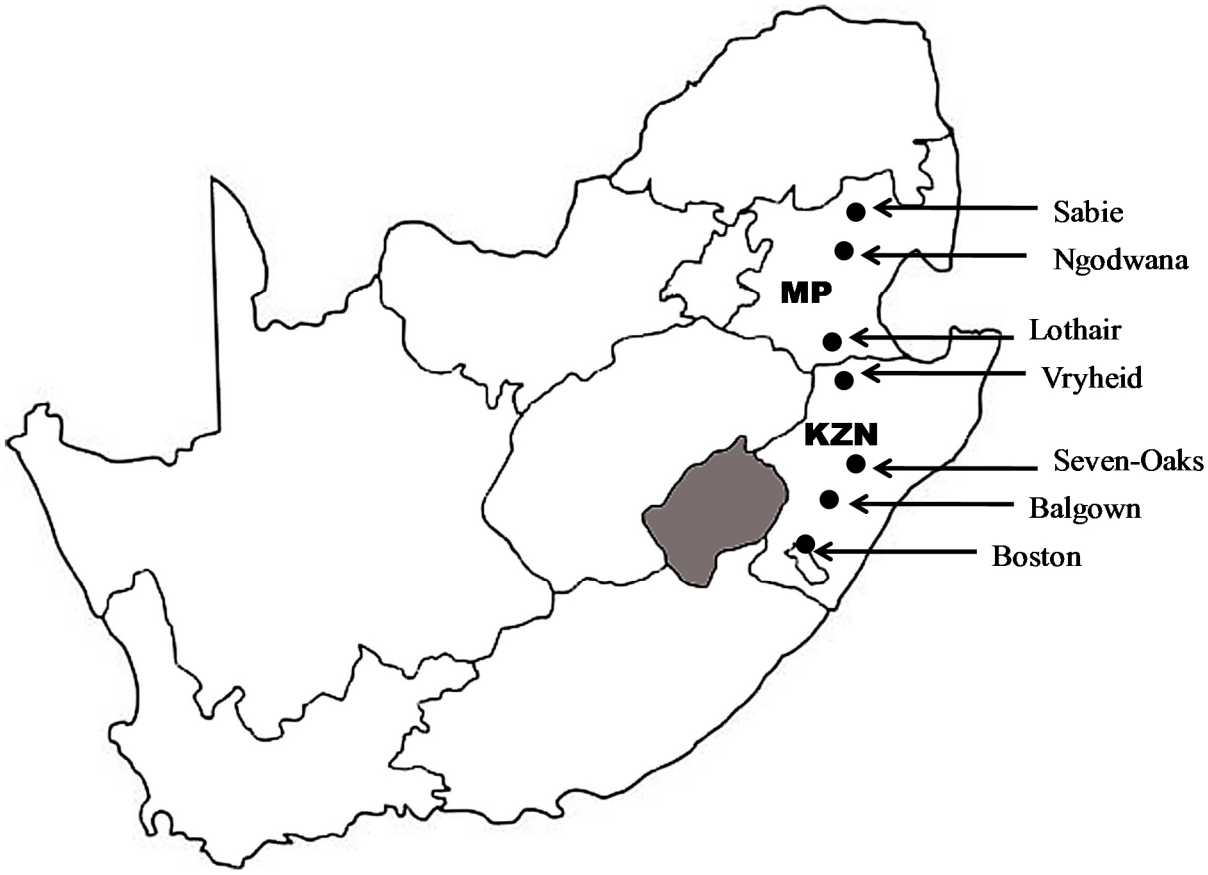


Figure 2. Differences in genetic structure including number of alleles, polymorphic loci and unique alleles for isolates from five plantations. Lot-Vry represents the population Lothair and Vryheid.

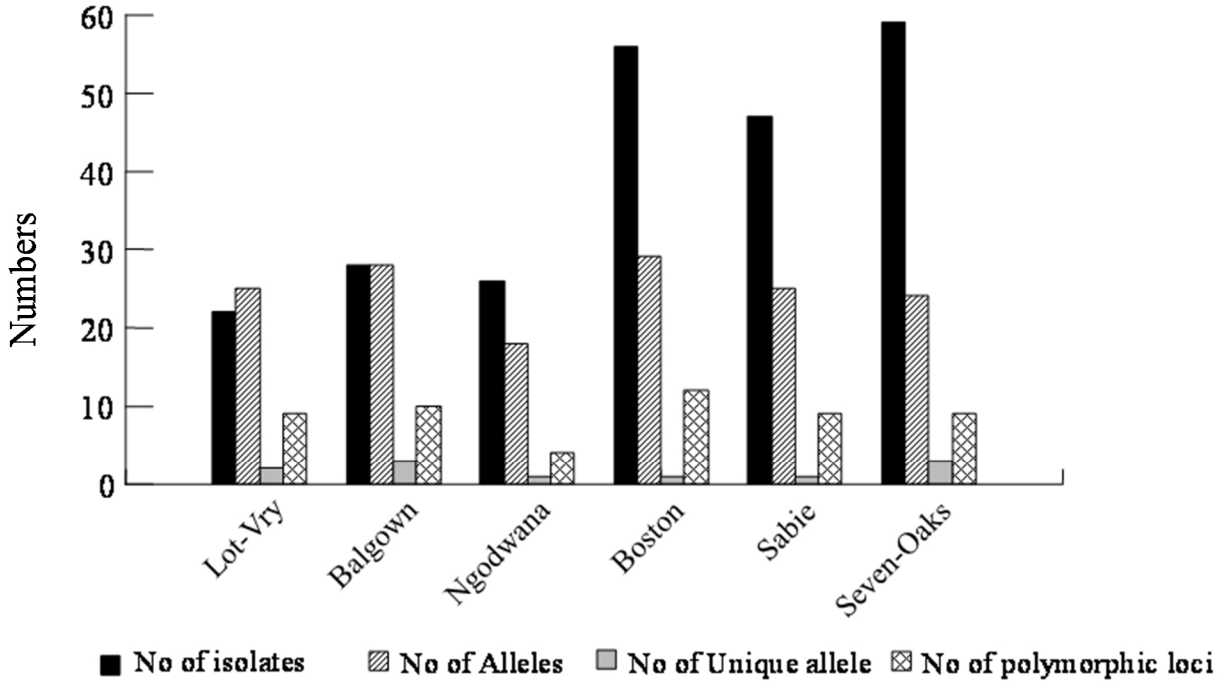
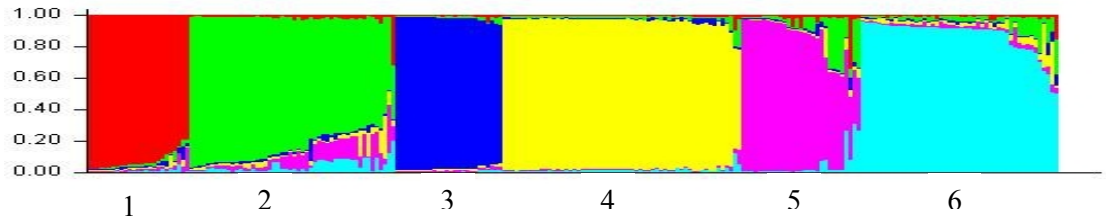


Figure 3. Affiliation of individual genotypes of *D. pinea* as assessed using Structure 2.2 and separated into six discrete vertical bars that are organized by sampling groups (a) and different K vs LnK values (b). Differences in colour within a vertical bar (a) indicate a multi-population affiliation of an individual genotype. The higher K vs LnK values represent the most likely number of clusters of individuals (b).



Different colours indicated different populations: 1 = Balgown, 2 = Boston, Lot-Vry, 3 = Ngodwana, 4 = 7-Oaks, 5 = Boston and 6 = Sabie

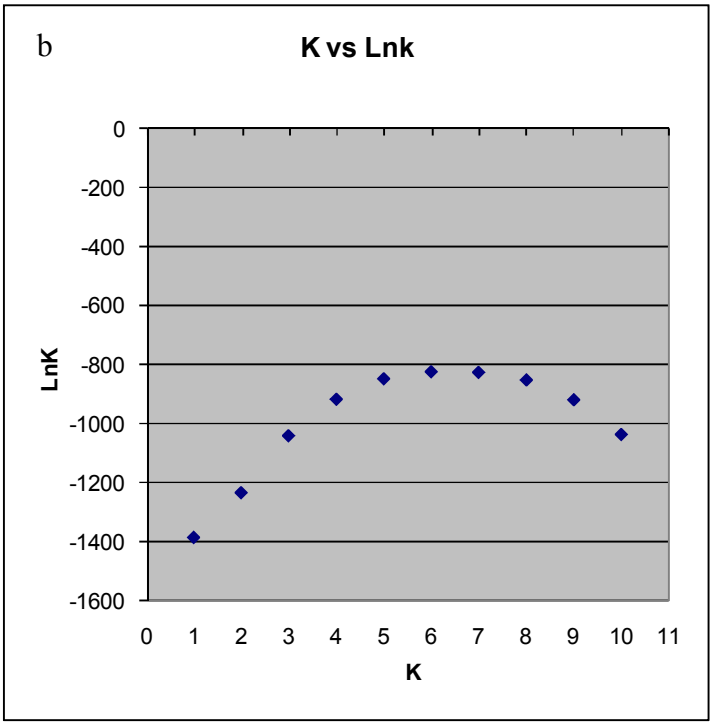


Figure 4. Index of Association (I_A) of clone corrected populations of *D. pinea* in South Africa. Arrows indicated where the observed I_A found from the 1000 times randomized data sets. P indicated the probability level we accept the differences. When $P < 0.05$, the H_o was rejected.

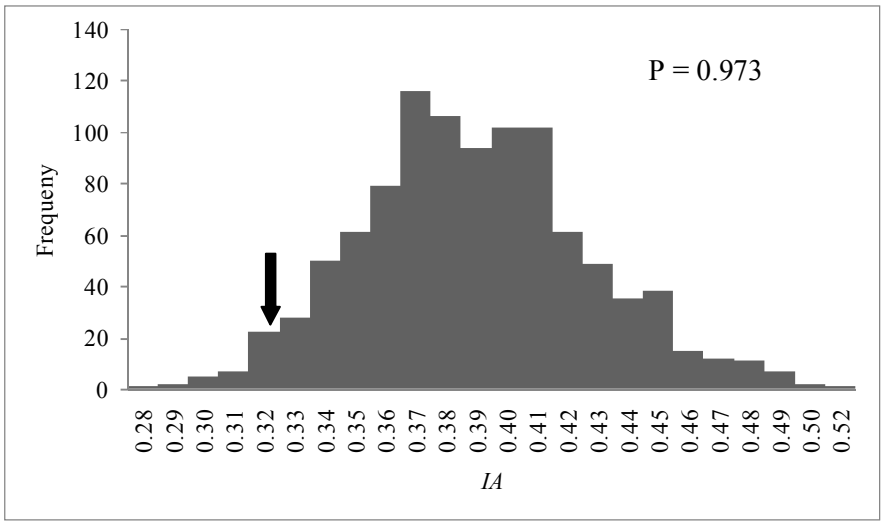


Table 1. Allele frequencies and genotype diversity by population

Locus	*Allele*	*Balgown*	*Boston*	*7-Oaks*	*Sabie*	*Ngodwana*	*Lot-Vry*
SS1	A	..	..	..	0.191	..	..
	B	0.577	0.036	..	..	..	..
	C	0.423	0.964	1.000	0.809	1.000	0.727
	D	..	..	..	..	..	0.273
SS2	A	0.962	0.839	0.017	0.979	..	0.727
	B	0.038	0.161	0.983	0.021	1.000	0.227
	C	..	..	..	..	..	0.045
SS5	A	0.231	0.268	..	..	..	0.091
	B	..	0.268	..	0.021	..	0.045
	C	0.769	0.250	1.000	0.723	..	0.773
	D	..	0.214	..	0.255	1.000	0.091
SS7	A	..	..	0.034	..	..	..
	B	..	0.232	..	0.106	..	0.045
	C	1.000	0.768	0.966	0.894	1.000	0.955
SS8	A	1.000	0.821	0.966	1.000	1.000	1.000
	B	..	0.179	0.034	..	..	..
SS9	A	1.000	0.857	..	0.128	1.000	0.864
	B	..	0.143	1.000	0.872	..	0.136
SS10	A	0.154	0.214	0.136	0.064	..	0.455
	B	0.769	0.643	0.864	0.489	0.692	0.545
	C	0.077	0.143	..	0.447	0.308	..
SS11	A	..	..	0.017	..	..	0.091
	B	0.923	0.018	0.203	..	1.000	..
	C	0.077	0.982	0.780	1.000	..	0.909
SS12	A	0.038	..	0.017	..	..	..
	B	..	..	0.017	..	..	..
	C	0.038	0.018	0.966	0.255	..	..
	D	0.846	0.982	..	0.723	0.846	1.000
	E	0.077	..	..	0.021	0.038	..
	F	..	..	..	..	0.115	..
SS13	A	0.808	1.000	0.797	1.000	0.192	1.000
	B	0.192	..	0.203	..	0.808	..
SS14	A	..	..	0.017	..	..	..
	B	0.385	0.964	0.983	0.936	1.000	0.955
	C	0.615	0.018	..	0.064	..	0.045
	D	..	0.018	..	..	..	..
SS15	A	0.231	..	..	..	..	..
	B	0.077	0.089	0.153	0.383	0.423	0.136
	C	0.692	0.911	0.847	0.617	0.577	0.864
SS16	A	0.038	..	..	..	..	..
	B	0.923	0.179	..	..	..	..
	C	0.038	..	..	..	..	..
	D	..	0.821	1.000	1.000	1.000	1.000
G		*17.789*	*24.500*	*3.436*	*24.275*	*6.145*	*11.52*
*G**		*68.421*	*43.750*	*5.824*	*51.648*	*23.636*	*52.380*
MLH		*21*	*34*	*20*	*34*	*10*	*15*

G = Genotype diversity (Stoddart & Taylor 1988); G* is percentage maximum diversity = (G/N)*100, where N is total number of isolates per population and MLH = Multilocus Haplotypes.

Table 2. Gene diversities (*H*) and contingency chi-square tests for differences in allele frequencies for 13 SSR loci across clone corrected populations of *D. pinea*.

Locus	*Gene diversity (H)*						*Chi-Sq*	*df*
	Balgown	Boston	7-Oaks	Sabie	Nogdwana	Lot-Vry		
SS1	0.499	0.057	0.000	0.337	0.000	0.320	138.14***	15
SS2	0.091	0.291	0.095	0.057	0.000	0.320	87.107***	10
SS5	0.363	0.741	0.000	0.372	0.000	0.516	164.646***	15
SS7	0.000	0.360	0.180	0.208	0.000	0.124	132.66***	10
SS8	0.000	0.208	0.180	0.000	0.000	0.000	9.738NS	5
SS9	0.091	0.327	0.180	0.251	0.000	0.320	102.39***	5
SS10	0.390	0.535	0.420	0.576	0.500	0.480	37.476***	10
SS11	0.172	0.057	0.455	0.000	0.000	0.124	114.88***	10
SS12	0.331	0.057	0.095	0.455	0.340	0.000	106.12***	25
SS13	0.363	0.000	0.455	0.000	0.320	0.000	57.12***	5
SS14	0.472	0.113	0.095	0.161	0.000	0.124	61.37***	15
SS15	0.490	0.291	0.420	0.457	0.500	0.320	36.91***	10
SS16	0.177	0.251	0.000	0.000	0.000	0.000	110.11***	15
Mean	0.443	0.438	0.314	0.365	0.194	0.338		

*Significant difference at alpha (P) ≤ 0.05, ** Significant difference at P ≤ 0.01, *** highly significant difference at P ≤ 0.001, NS = non-significant difference at P ≤ 0.05, *df* = degree of freedom = (Number of alleles – 1) * (Number of populations – 1).

Table 3. Population differentiations (Θ) (above the diagonal) and gene flow (below the diagonal) among *D. pinea* populations from different locations in South Africa

	Balgown	*Boston*	*7-Oaks*	*Sabie*	*Ngodwana*	*Lot-Vry*
Balgown	--	0.404***	0.527***	0.491***	0.541***	0.291***
Boston	0.738	--	0.528***	0.212***	0.555***	0.203***
7-Oaks	0.449	0.447	--	0.504***	0.643***	0.470***
Sabie	0.518	1.858	0.492	--	0.631***	0.361*
Ngodwana	0.424	0.401	0.277	0.292	--	0.664***
Lot-Vry	1.218	1.963	0.564	0.885	0.253	--

* indicates significant difference at P ≤ 0.05 and *** indicated highly significant different at P ≤ 0.01.

Table 4. Observed I_A and $\bar{r}_d$ values of each populations

Population	I_A	$\bar{r}_d$	*P*
7-Oaks	-0.111	-0.012	0.727
Boston	-0.012	-0.001	0.525
Balgown	0.00002	0.0000002	0.468
Sabie	-0.063	-0.008	0.684
Ngodwana	-0.259	-0.087	1.000
Lot-Vry	-0.213	-0.027	0.837

Table 5. Analysis of molecular variance of South African *D. pinea* populations hierarchically partitioned.

Source of Variation	*df*	*SS*	*MS*	*Est. Var.*	*% Variation*
Among Regions	1	21.003	21.003	0.000	0%
Among Pops	4	230.530	57.632	1.516	55%
Within Pops	230	280.544	1.220	1.220	45%
Total	235	532.076			

df = degree of freedom, SS = Sum of square, MS = Mean square and Est. Var. = Estimated variance