Infection, Genetics and Evolution

# A possible centre of diversity in South East Asia for the tree pathogen, *Ceratocystis manginecans*

Arista Fourie[a], Michael J. Wingfield[a], Brenda D. Wingfield[a], Pham Q. Thu[b], Irene Barnes[a]*

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

[b] Forest Protection Research Centre, Vietnamese Academy of Forest Sciences, Dong Ngac, Tu Liem, Hanoi, Vietnam

Short title: Genetic diversity of *C. manginecans* populations

*Corresponding author: Irene Barnes

Email: irene.barnes@fabi.up.ac.za

Tel. no.: +27(12) 420 5142

Fax no.: +27 (12) 420 3960

## Highlights

• In Oman and Pakistan, *C. manginecans* from native legumes are genetically identical to the clonal population from *M. indica*.

• The data suggest a possible area of origin for *C. manginecans* in South East Asia, particularly Vietnam.

• Microsatellite primers were designed, from the genome of *C. manginecans*, that can be used in a multiplex PCR approach.

## ABSTRACT

The fungal pathogen, *Ceratocystis manginecans,* has caused serious canker and wilt disease on *Mangifera indica* (mango), on legume tree species in Oman and Pakistan and on *Acacia* spp. in Indonesia. A *Ceratocystis* species, with similar morphology to *C. manginecans*, has recently been reported in Vietnam, causing severe disease of *Acacia* trees. Previous population genetic studies on isolates from *M. indica* in Oman and Pakistan have shown that the pathogen represents a single clonal haplotype, indicative of an introduced pathogen. The aim of this study was to investigate the genetic diversity and

population structure of 160 *C. manginecans* isolates, from four host-associated populations from Oman, Pakistan, Indonesia and Vietnam. This was done by applying a combination of 14 previously developed microsatellite markers and a new set, designed in this study from two different *C. manginecans* genomes. Sequence data confirmed that the isolates in Vietnam are the same species as those in Indonesia and were thus identified as *C. manginecans*. Unlike the populations in Oman and Pakistan, relatively high levels of genetic variation were found for the isolates from Indonesia and Vietnam. The Vietnam population was significantly differentiated from the other populations and isolates from this area had the highest level of genetic diversity thus far encountered for the pathogen.

**Keywords** *Acacia* spp. · *Ceratocystis manginecans* · mango trees · population genetic diversity · Vietnam

## 1. Introduction

*Ceratocystis manginecans* is a fungal pathogen, first described from Oman and Pakistan, as the causal agent of severe mango wilt disease (Al Adawi et al., 2013b; Al Adawi et al., 2006; Van Wyk et al., 2007). The species resides in *Ceratocystis sensu stricto* (De Beer et al., 2014) that includes numerous cryptic species, the majority of which are aggressive tree pathogens (Al Adawi et al., 2013b; Engelbrecht and Harrington, 2005; Ocasio-Morales et al., 2007; Roux and Wingfield, 2013; Wingfield et al., 2013). *C. manginecans* has also been reported as a serious pathogen of *Acacia mangium* trees in Indonesia, identified at the time as the new species *C. acaciivora* (Tarigan et al., 2011). These two species have, however, recently been synonymised as *C. manginecans* (Fourie et al., 2014).

Since the first report of mango wilt in Oman in 1999 (Al Adawi et al., 2003), *C. manginecans* has spread to 12 districts of Oman and to three in Pakistan (Al Adawi et al., 2013a; Al Adawi et al., 2006; Van Wyk et al., 2007). Concern regarding the pathogen has

increased in recent years, as it has been found on the native legume trees *Prosopis cineraria* (Ghaf) in Oman and *Dalbergia sissoo* (Shisham) in Pakistan (Al Adawi et al., 2013b). This species has also been reported on the widely planted *A. mangium* in South East Asian countries, including Indonesia (Tarigan et al., 2011) and Malaysia (Brawner et al., 2015), and on *Eucalyptus* species in South China (Chen et al., 2013). Isolates of *Ceratocystis* spp*,* which could represent the same species, have also been reported from *Eucalyptus* trees in Brazil and *Punica granatum* (pomegranate) in India (Harrington et al., 2014).

Disease symptoms associated with *C. manginecans* infection on *M. indica* and legume trees include staining of vascular tissue, stem cankers and wilt (Al Adawi et al., 2013b; Tarigan et al., 2011; Van Wyk et al., 2007). These symptoms can result in tree death within six months of infection (Al Adawi et al., 2006; Tarigan et al., 2011). *A. mangium* trees are a primary resource in the paper and pulp industry in Indonesia and its neighbouring countries (FSIV, 2009; Maslin, 2013). Oman and Pakistan are also highly reliant on mango production for fruit consumption and exports (Al Adawi et al., 2006; FAOSTAT, 2011; Kumar et al., 2011). The disease is thus a serious threat to fruit crop and plantation tree industries in the Middle East and South East Asia.

In a recent survey of *Acacia* plantations in Vietnam, a species of *Ceratocystis* was identified causing stem cankers and crown wilt on *Acacia auriculiformis, A. mangium* and a hybrid of the two species in eight provinces in the country (Thu et al., 2012). This was the first report of a *Ceratocystis* spp. causing disease on *Acacia* in Vietnam. The geographic proximity of Vietnam to Indonesia and the similarity of the disease on the same host plant (Tarigan et al., 2011) suggested that the pathogen could be *C. manginecans.*

A population genetic study was recently performed on *C. manginecans* isolates collected from *M. indica* trees from Oman and Pakistan (Al Adawi et al., 2014). Isolates from the wood-boring beetle *Hypocryphalus mangiferae*, the major vector of *C. manginecans* on *M. indica* in Oman and Pakistan (Al Adawi et al., 2013a), were included. The data showed that the isolates from *M. indica* and *H. mangiferae* represent a single clonal population,

suggesting that the fungus is an introduced pathogen in these countries. This introduction could have come from South East Asia where the fungus occurs on *A. mangium* (Tarigan et al., 2011), or from South-America, as suggested by Harrington et al. (2014). The rapid distribution of the pathogen in Oman and Pakistan was most likely facilitated by *H. mangiferae* (Al Adawi et al., 2013a; Al Adawi et al., 2013b; Masood et al., 2010). This insect is thought to be native to southern Asia but has not yet been connected to outbreaks of *C. manginecans* in this region (Atkinson and Peck, 1994; Masood et al., 2010; Masood et al., 2008).

The aim of this study was to determine whether the isolates from Vietnam represented the same species as those causing disease on *A. mangium* in Indonesia. The species identity was determined using a sequence based approach. Additionally, the population genetic structure of *C. manginecans* isolates from the Middle East (Oman and Pakistan) and South East Asia (Indonesia and Vietnam) was investigated. From Oman and Pakistan, the study included new isolates from the legume trees *P. cineraria* and *D. sissoo*, respectively, as these were not considered in terms of their genetic structure in the study of Al Adawi et al. (2014). The genetic diversity of the populations was investigated using a combination of previously developed microsatellite markers (Barnes et al., 2001; Steimel et al., 2004), as well as a new set of markers designed in this study using the recently published genomes of two *C. manginecans* isolates (Van der Nest et al., 2014). A small representative collection of the clonal isolates from *M. indica* and *H. mangiferae* from Oman and Pakistan, used in a previous population study (Al Adawi et al., 2014), were included for comparison purposes and evaluation of the new microsatellite markers.

## 2. Materials and methods

### *2.1. Fungal isolates and DNA extraction*

Isolates from Oman and Pakistan, used in this study, included those that were isolated from legume tree species (Table 1). The isolates from *P. cineraria* trees were collected in Sohar (Oman; Fig. 1A) and those from *D. sissoo* trees in Faisalabad and Shorkot (Pakistan; Fig. 1B) (Al Adawi et al., 2013b). In addition, a subset of 21 clonal isolates from *M. indica* and *H. mangiferae*, as used in the population study by Al Adawi et al. (2014) (Fig. 1A and 1B), were included in the current study for comparative purposes. In Indonesia, isolates were collected from wilting *A. mangium, A. crassicarpa* and *A. auriculiformis* x *A. mangium* hybrid tree species in seven plantation areas in and around the Riau province of Sumatra (Fig. 1C). Isolates were collected from diseased *A. mangium, A. auriculiformis* or hybrids of these two tree species in Vietnam during a previous disease survey across seven provinces (Thu et al., 2012; Fig. 1D). All isolates are maintained in the culture collection (CMW) of the Forestry and Agricultural Biotechnology Institute (FABI), Pretoria, South Africa.

All cultures were grown at room temperature (23-25°C) on 2% malt extract agar (MEA; Biolab, Midrand, South Africa) supplemented with 50 mg/l streptomycin (Sigma-Aldrich, Germany) and 100 µg/l thiamine (Sigma-Aldrich, Germany). Mycelium was collected from the surface of the MEA plates and freeze dried. Dried mycelium was crushed in a Mixer Mill type MM 301 Retsch® tissue lyser (Retsch, Germany), with metal beads. DNA was extracted using a phenol/chloroform method (Goodwin et al., 1992) and quantified using a Nanodrop ND_1000 instrument (Nanodrop, Wilmington, Delaware). The quality of the DNA was then analysed by means of agarose gel electrophoresis (AGE) on a 1% agarose gel by combining a volume of 5 µl DNA with 2 µl GelRed™ (Biotium, California). The DNA was visualised with UV light.

**Table 1**

Details of the *Ceratocystis manginecans* isolates collected in Oman, Pakistan, Indonesia and Vietnam, used in this study, which includes the host, geographic location and sample sizes obtained from different regions.

| Country | CMW number[a] | Host | Location | Collector and date collected | Sample size |
|---|---|---|---|---|---|
| Oman | 15377[e], 15353[e] | *Mangifera indica* | Shinas | A. O. Al Adawi (2003-2004)[b] | 2 |
| | 15366[e], 15369[e], 15385[e], 15391[e] | " | Liwa | " | 4 |
| | 15313-15316[e] | " | Sohar | " | 4 |
| | 15371[e] | " | Quariyat | " | 1 |
| | 15317[e] | *Hypocryphalus mangiferae* | Sohar | " | 1 |
| | 15384[e] | " | Liwa | " | 1 |
| | 15381[e], 15382[e] | " | Quariyat | " | 2 |
| | 17225, 17568, 17570 | *Prosopis cineraria* | Sohar | A. O. Al Adawi (2004)[c] | 3 |
| | | | | | **Total 18** |
| Pakistan | 17567[e] | *M. indica* | Faisalabad | A. O. Al Adawi (2003-2004)[b] | 1 |
| | 23637[e], 23642[e], 23643[e] | " | Multan | " | 3 |
| | 23628[e], 23630[e] | *H. mangiferae* | Faisalabad | " | 2 |
| | 23623 , 23624 | *Dalbergia sissoo* | " | A. O. Al Adawi (2004)[c] | 2 |
| | 23625 | " | Shorkot | " | 1 |
| | | | | | **Total 9** |
| Indonesia | 22560, 22561, 22563-22580 | *Acacia mangium* | Pelalawan | M. Tarigan (2006) | 20 |
| | 37866-37868 | *A. crassicarpa* | " | M. Tarigan (2011) | 3 |
| | 37869 | *A. mangium* | | " | 1 |
| | 37870, 37871 | *A. mangium x A. auriculiformis* hybrid | | " | 2 |
| | 22581-22585 | *A. mangium* | Logas | M. Tarigan (2006) | 5 |
| | 37860, 37861 | " | " | M. Tarigan (2010) | 2 |
| | 22587-22591, 22593-22619, 22621-22623,22625,22626 | " | Teso | M. Tarigan (2006) | 37 |

| | | | | | |
|---|---|---|---|---|---|
| | 37857-37859 | " | " | M. Tarigan (2010) | 3 |
| | 37825-37828, 37925, 37926 | *A. mangium x A. auriculiformis* hybrid | " | " | 6 |
| | 37829-37836 | " | Kerinci | " | 8 |
| | 37837-37855 | *A. mangium* | " | " | 18 |
| | 37856 | " | Baserah | " | 1 |
| | 37862 | | Langgam | " | 1 |
| | 37863-37865 | *A. mangium* wildling | Sei Kebaro, North Sumatra | M. Tarigan (2011) | 3 |
| | | | | | **Total 110** |
| Vietnam | 38000-38004 | *A. auriculiformis* | Thua Thien Hue | P. Q. Thu (2008)[d] | 5 |
| | 38005, 38012-38019 | *A. mangium* | " | | 9 |
| | 38020 | " | Quang Ninh | | 1 |
| | 38023 | " | Tuyen Quang | | 1 |
| | 38006 | *Acacia hybrid* | Binh Duong | | 1 |
| | 38007 | " | Thua Thien Hue | | 1 |
| | 38008 | " | Binh Phuoc | | 1 |
| | 38009-38011 | " | Dong Nai | | 3 |
| | 38021 | " | Lam Dong | | 1 |
| | | | | | **Total 23** |

[a] All the *C. manginecans* isolates are maintained in the CMW culture collection of FABI, University of Pretoria, South Africa.

[b] Al Adawi et al., 2014.

[c] Al Adawi et al., 2013b.

[d] Thu et al., 2012.

[e] Isolates used in a previous study (Al Adawi et al., 2013b).

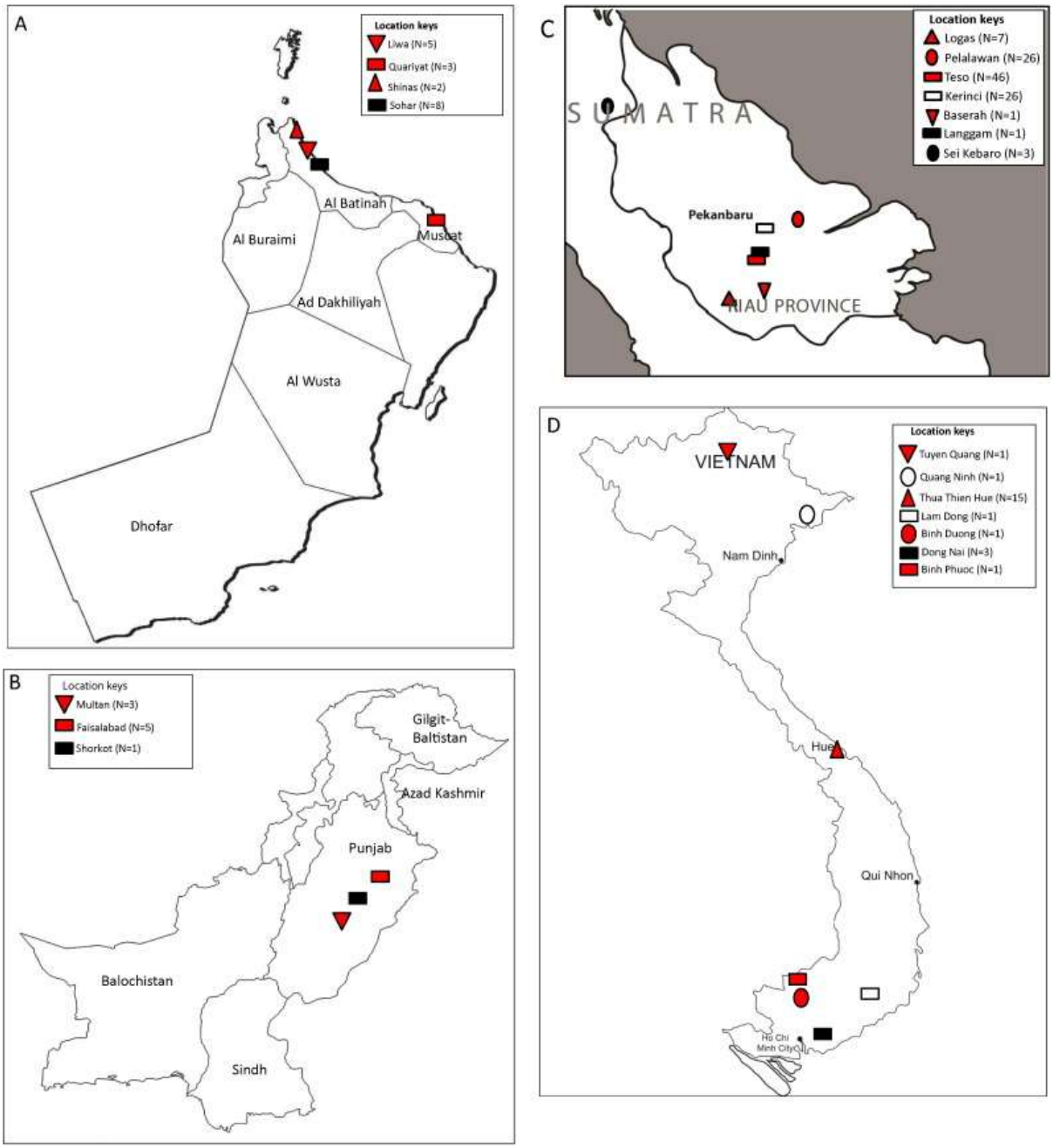


**Fig. 1.** Maps of Oman (A), Pakistan (B), the Riau province in the Sumatra island of Indonesia (C) and Vietnam (D), indicating the collection sites and the number of *C. manginecans* isolates ("N") used in this study. The different collection sites are indicated by different keys on the maps.

### *2.2. Confirmation of isolate identity*

#### *2.2.1. PCR amplification and sequencing*

Isolates from Oman and Pakistan, including those from *M. indica, H. mangiferae, D. sissoo and P. cineraria,* had been identified as *C. manginecans* in previous studies (Al Adawi et al., 2013b; Al Adawi et al., 2014). Those from Indonesia and Vietnam were identified as a species of *Ceratocystis*, based on morphology. Subsets of ten isolates from both countries were subjected to DNA sequence analyses to confirm the species identity. The ITS gene region was amplified in the isolates, using primers ITS1 and ITS4 (White et al., 1990). Additionally, the β-tubulin 1 (βT 1) region was amplified to confirm the ITS results, using the primers βT1a and βT1b (Glass and Donaldson, 1995).

The PCR reaction mixture, used for the amplification of both gene regions, consisted of 30 ng DNA, 2.5 mM $MgCl_2$, 250 µM dNTP, 1 unit FastStart *Taq* DNA polymerase (Roche Applied Science, South Africa), 2.5 µl 10x FastStart *Taq* DNA polymerase PCR buffer (including 1.5 mM $MgCl_2$), 0.2 µM of the forward and reverse primer each and adjusted to a total reaction volume of 25 µl with autoclaved distilled $H_2O$. The PCR cycler program for both gene regions consisted of: 95 °C for 5 min, 10 cycles of 95 °C for 30 s, 60 °C for 45 s, 72 °C for 90 s, another 30 cycles of 95 °C for 30 s, 60 °C for 45 s and 72 °C for 90 s (5 s increase per cycle at 72 °C) with a final step at 72 °C for 10 min. PCR and sequence amplicons were purified with 6% Sephadex G-50 in Centricep filter columns (Sigma-Aldrich, Germany). The sequence of the amplicons was determined on an ABI PRISM™ 3500xl Autosequencer (Applied BioSystems, Foster City, California) according to the manufacturer protocol.

#### *2.2.2. Sequence alignment and phylogenetic analysis*

The raw sequence reads for each of the isolates were assembled in CLC Main Workbench V. 6.0 (CLC Bio, www.clcbio.com) to obtain consensus sequences for every

isolate. The sequences obtained were aligned to previously generated data sets of the ITS and βT 1 regions, respectively, of *C. fimbriata sensu stricto* (Fourie et al., 2014). Alignments were performed using MUSCLE in MEGA v. 5 (Tamura et al., 2011) and used to construct phylogenetic trees of ITS and βT 1 based on Maximum parsimony (MP). The MP trees were constructed in PAUP 4.0, based on a heuristic search option, with random addition of sequences and Tree Bisection Reconstruction (TBR) branch swopping method (Swofford, 2002). Statistical support for the tree branches was confirmed with a 1000 bootstrap replicates.

### *2.3. Development of microsatellite markers and fragment analysis*

Microsatellite markers, available for species of *Ceratocystis* (Barnes et al., 2001; Steimel et al., 2004), were used in this study. However, these markers were developed for the species *C. cacaofunesta* and *C. fimbriata*. In order to increase the rigor of population genetic analyses, specifically for *C. manginecans*, additional microsatellite or simple sequence repeat (SSR) markers were developed from *C. manginecans* isolates. This was achieved using the fully sequenced genomes of two *C. manginecans* isolates, CMW 17570 (Van der Nest et al., 2014) and CMW 22621 (unpublished). Microsatellite regions were identified in the genome of CMW 22621 by analysing the assembled Illumina sequence contigs in Msatfinder v. 2.0 on-line (Thurston and Field, 2005). Parameters were set to detect tri, tetra, penta or hexa nucleotide motifs with a minimum of 9 repeat units. Contigs containing microsatellite regions were imported into CLC Genomics workbench v. 6.0 (CLC Bio, www.clcbio.com). To identify polymorphic SSR regions, the contigs that contained perfect microsatellites, optimal for primer design, were selected and compared with the orthologous region in CMW 17570 by means of BLASTn analysis. For the SSR regions that were polymorphic between the two isolates, primers were designed to flank these regions. The parameters for primer design included a primer length between 18 bp and 22 bp, a G/C content between 0.4 and 0.6 and an annealing temperature between 55°C and 62°C.

Primers were also designed to produce fragment lengths that ranged between 180 and 500 bp to allow for multiplexing during fragment analyses.

The PCR amplification conditions of the 27 available *Ceratocystis* SSR primer sets (Barnes et al., 2001; Steimel et al., 2004), as well as the newly designed primers, was optimised using three *C. manginecans* isolates (CMW 22563, CMW 38022, CMW 38002). PCR reaction mixtures for the previously developed primers were identical to those for the ITS region, described in 2.2.1., with the exception of the addition of 2 mM $MgCl_2$ per reaction. A multiplex PCR procedure was developed for the new SSR markers by combining a maximum of three primer sets in a single PCR reaction. The PCR reaction mixture consisted of 30 ng DNA, 1.5 units MyTaq™ DNA Polymerase (Bioline Ltd., UK), 5 µl 5x MyTaq™ reaction buffer and a total primer concentration of 0.32 µM for both the forward and reverse primers. Reaction volume was a total of 25 µl. The PCR cycler was programmed as follows: 95 °C for 5 min, 10 cycles of 95 °C for 30 s, 55 °C for 45 s, 72 °C for 1 min, followed by 30 cycles of 95 °C for 30 s, 55 °C for 45 s (with 5 s/cycle time increase), 72 °C for 1 min and a final step at 60 °C for 35 min. PCR amplification success was evaluated using AGE. Individual PCR amplicons, obtained from the new set of SSR primers, were sequenced for all three isolates to confirm that the allelic variation observed is due to mutations within the microsatellite repeat region and not as a result of indels in the flanking region.

The primer sets that amplified successfully in all three isolates were resynthesized and the forward primer labelled with either PET, VIC, NED or FAM fluorescent dyes from the G5 labelling kit (Applied Biosystems, Warrington, UK). These primers were screened on all isolates used in the study, including those from Oman, Pakistan, Indonesia and Vietnam. Three panels were designed for fragment analyses using GeneScan (Applied Biosystems, Thermo Fisher Scientific, Carlsbad, USA), by taking amplicon size and fluorescent dye colour into consideration, such that the amplicons of different primer sets could be pooled and run together in a single lane. A 1/200 dilution of each microsatellite PCR product in a particular panel was pooled. Of the pooled PCR reaction mix, 1 µl was then combined with 0.2 µl GeneScan Liz500(-250) size standard (Applied Biosystems, Thermo Fisher Scientific,

Carlsbad, USA) and 10 µl formamide and run on an ABI PRISM$_{TM}$ 3500xl Autosequencer (Thermo Fisher Scientific, Carlsbad, USA). GeneScan data were analysed with GeneMapper® v.4.1 software (Applied Biosystems, Thermo Fisher Scientific, Carlsbad, USA) to score allele fragment sizes. The 21 isolates from *M. indica* and *H. mangiferae*, considered by Al Adawi et al. (2014) and also used in the present study, were re-run on the ABI PRISM$_{TM}$ 3500xl Autosequencer and re-scored to avoid shifts in allele sizes experienced when using different instruments. Additionally, the newly developed markers were considered for this population and were tested in all 21 isolates.

All unique alleles (an allele present only in a single isolate) were sequenced to confirm the existence of the allele. In addition, various different alleles per locus were sequenced to confirm that the variation observed was due to the expansion or contraction in the microsatellite region and not in the flanking regions. PCR and sequencing reactions were performed with identical conditions as described above, except that non-fluorescent primers were used.

### *2.4. Statistical analyses of SSR data*

All 160 isolates analysed in the study were provisionally grouped into four populations based on host and/or geographic origin. The first population included isolates from *D. sissoo* and *P. cineraria* in Oman and Pakistan, the second from *M. indica* and *H. mangiferae* in the same countries, the third from *Acacia* spp. in Indonesia and the fourth from *Acacia* spp. in Vietnam. Two sets of data were generated: non clone-corrected data where all individuals were considered in the analyses and clone-corrected data where individuals with identical haplotypes in a population were removed. The data were clone-corrected using GenAlEx v.6.2 (Peakall and Smouse, 2006). Clone-corrected data were used for all analyses except where otherwise mentioned.

The analogue for Wright's $F_{ST}$ value in haploid individuals, $\phi_{PT}$ (Excoffier et al., 1992), was calculated for the individuals in GenAlEx v.6.2 (Peakall and Smouse, 2006) using an analysis of molecular variance (AMOVA). The $\phi_{PT}$ values were compared in a pair-wise manner between each of the host and/or geographic populations to determine if the differentiation between any of the populations is significant. The AMOVA analysis was used to compare the variation within and between populations. Statistical significance of the differentiation in both analyses was assessed based on 1 000 permutations. A null hypothesis of no genetic variation among the populations was assumed, at a P value of 0.001.

Analyses with hierarchical sub-structuring were only considered for Indonesia as the other countries contained very few isolates per region. The non clone-corrected dataset was used for analysis. Two of the collection sites in Indonesia (Baserah and Langgam) were represented by a single isolate and these were incorporated with the isolates of the nearest collection site (Logas and Teso, respectively), for this part of the analyses. AMOVA and $\phi_{PT}$ values were then determined based on the five collection sites in the country.

To determine whether population structure was present among the investigated isolates, the program STRUCTURE v 2.2 (Pritchard et al., 2000) was used. Analysis in the STRUCTURE program uses a Bayesian clustering method to assign each individual to a potential population without prior information of its origin. To determine the most likely number of clusters or populations (K), K-values were ranged from one to seven. A total of 1 million Markov Chain Monte Carlo (MCMC) generations were run, including a burn-in value of 100 000 generations and 20 iterations per K-value. The entire run was repeated. The best parameters for analysis of the dataset were determined by evaluating the effect on the K-value when considering both models of admixture and no admixture as well as correlated and non-correlated allele frequencies among populations.

To determine the most likely K-value from the data obtained, the statistic $\Delta K$ (Evanno et al., 2005) was calculated using software from Clumpak (Kopelman et al., 2015). This is

based on the rate of variation between the LnP(D) of successive K-values. A model of admixture and allelic correlation was assumed. The consensus figure from all 20 iterations of the selected K-value were also constructed using Clumpak.

Results from the AMOVA and STRUCTURE analyses indicate that the isolates from the legume trees and those from *M. indica* in Oman and Pakistan cannot be differentiated as distinct populations and were, therefore, considered as one population for subsequent analyses. The existence of private alleles in each of the three populations were determined using GenAlEx v.6.2 (Peakall and Smouse, 2006). The number of alleles, allele frequencies and the gene diversity (H) (Nei, 1973) of each microsatellite locus in all three populations were calculated using the non clone-corrected dataset with POPGENE v. 1.32 (Yeh et al., 1999). The gene diversity of each population was also calculated for the clone-corrected dataset.

The genotypic diversity (G) of each population was calculated for the non clone-corrected isolates following the method of Stoddard and Taylor (1988). Both the number of haplotypes and G for each population was determined using the statistical program R v. 3.1.2 (R Development Core Team, Vienna) and the R package poppr (Kamvar et al., 2014). Due to the significant difference in population sample sizes, the expected number of haplotypes and the estimated genotypic diversity (eG) was calculated with poppr using rarefaction, based on the smallest population size of 23 (Grünwald et al., 2003).

In order to calculate the genetic distance between the isolates, the proportion of shared alleles (Chakraborty and Jin, 1993) was used as a distance measure and a distance matrix was constructed using the software Populations v. 1.2.32 (Langella, 1999). This genetic distance is based on the average proportion of alleles shared between and within populations when taking all individuals sampled into consideration (Chakraborty and Jin, 1993). Subsequently, a neighbor joining (NJ) tree was constructed in the same program to visualise the calculated distances. One thousand bootstrap replicates were performed to ensure the results are statistically significant. The NJ tree was viewed and edited in MEGA v. 5 (Tamura et al., 2011).

## 3. Results

### *3.1. Fungal isolates and DNA extraction*

A total of 160 isolates were collected and used in this study (Table 1). These included 110 isolates from seven *Acacia* plantation areas in Indonesia and 23 isolates from seven regions in Vietnam. For the Oman collection, 10 isolates were from *M. indica*, four from *H. mangiferae* infesting these trees and three were from *P. cineraria* (Table 1). Isolates from Pakistan included four from *M. indica* from two locations, two from *H. mangiferae* and three from *D. sissoo,* collected from two different locations in Pakistan. All isolates used in this study had morphologies typical of *C. manginecans.*

### *3.2. Confirmation of isolate identity*

#### *3.2.1. PCR amplification and sequencing*

PCR amplification of the ITS and βT 1 gene regions in the isolates resulted in the expected 550 bp and 600 bp fragment size, respectively, and high quality sequences for most isolates. However, the sequence chromatograms of ITS for some isolates from Indonesia and Vietnam had clear peaks up to ~120 bp after which the base calling showed conflicting sequence data due to two ITS types present (Al Adawi et al., 2013b; Naidoo et al., 2013). The poor quality sequence data for these isolates were not used for identification purposes.

#### *3.2.2. Sequence alignment and phylogenetic analysis*

Alignment of all the high quality ITS sequences showed that the isolates had two different ITS haplotypes. The isolates contained either haplotype 1 of *C. manginecans*

(previously defined as the species *C. acaciivora* (Al Adawi et al., 2013b) or haplotype 2 (defined as *C. manginecans*). The βT 1 region was identical for all the isolates used in this study and they all grouped as a single clade in the phylogenetic tree as *C. manginecans* (results not shown). Because *C. acaciivora* has been reduced to synonymy with *C. manginecans* (Fourie et al., 2014), all isolates in this study were considered to represent *C. manginecans*.

### *3.3. Development of microsatellite markers and fragment analysis*

Twelve polymorphic SSR regions were identified from the genome sequences of the two isolates of *C. manginecans*. The sequences generated for each SSR region confirmed the presence of the microsatellite polymorphisms for 10 of the 12 markers. Nine of the markers were tri-nucleotide repeats and one was a tetra-nucleotide repeat (marker AF5). These 10 markers were used for further population analyses (Table 2).

Of the 27 available SSR primers (Barnes et al., 2001; Steimel et al., 2004), 19 amplified successfully in all 160 isolates (Table S1). The markers GACA650, CF13/14 and AG15/16 were monomorphic in all isolates investigated. CAT1 resulted in a fingerprint profile for some isolates for which the allele size could not be determined and CF11/12 produced allele sizes differing by a single nucleotide. These SSR's were thus excluded from further analyses. The remaining 14 SSR markers and the 10 newly designed markers, together, resulted in 70 alleles ranging from two to eight alleles per locus (Table 3). In three of the markers, CAA10, CAA15 and CAT3K, the most dominant allele was present in more than 90% of the isolates, indicating the low level of variation displayed by these markers (Table 3).

A total of 27 of the 70 alleles were private. They were either observed at high frequency but were unique to a single population or they occurred only once in that specific population (Table 3). The population from Oman and Pakistan included a private allele (166) in marker CAA9, which occurred in a single *D. sissoo* isolate. The *M. indica* isolates also included one allele (306) unique to the host, in marker CAA38, which occurred at a frequency of 0.095. In

**Table 2**

Details of the new polymorphic microsatellite markers, developed from the genomes of two *C. manginecans* (CMW 17570 and CMW 22621) isolates.

| SSR locus | SSR motif in *C. manginecans CMW 17571/ CMW22621* | Forward primer sequence 5'-3' Reverse primer sequence 5'-3' | Fluores-cent label | Tm (°C) | Multiplex PCR no.[a] | GenBank acc. no. |
|---|---|---|---|---|---|---|
| AF2 | (AGA)10/ (AGA)12 | CATTCTCGAAACACTAGCG AGGAGAGAAAGGATGGTGG | VIC | 57 | 1 | KJ601490 |
| AF3 | (CTG)12/ (CTG)13 | AAGAAGAAGGAAAGCATCCG ACATCAACATCGTTTCTAGCCA | NED | 57 | 3 | KJ601491 |
| AF4 | (ACA)10/ (ACA)8 | CTGTTTGACGGCTTTGGA TGCTAATGGAGGTCGGTG | PET | 57 | 1 | KJ601492 |
| AF5 | (GTCA)9/ (GTCA)6 | TGTTCTTCTGATTGTGCACT GAGGTTGGCGTTGGTTAG | FAM | 57 | 5 | KJ601493 |
| AF6 | (GAG)13/ (GAG)11 | CTATTGCGAGTTCAAGGC ACCCCTCATGATTCACTTAC | VIC | 55 | 2 | KJ601494 |
| AF7 | (AGC)10/ (AGC)9 | CCTACATCTTCTTTGAGCCCTT GTTGTGGCTGCTGGGTTT | NED | 60 | 1 | KJ601495 |
| AF8 | (GAG)12/ (GAG)11 | CTATCTGTCCTTGCCCCT CGGGCCTTTCTTTTGTTTCTT | PET | 59 | 5 | KJ601496 |
| AF9 | (GCA)9/ (GCA)11 | ACTCTACTACCCTCACAC GACTAGGCCTCCATTGAA | FAM | 55 | 2 | KJ601497 |
| AF11 | (ACA)10/ (ACA)7 | TTGGACATACTTGGACGGG ATTTAGTGGGAATCTGCGG | VIC | 57 | 3 | KJ601498 |
| AF12 | (ACA)17/ (ACA)12 | ACAGGTAAGAAGGGACAGAA GATAAGGAGAGTGGGAAAGG | PET | 60 | 4 | KJ601499 |

[a] PCR amplifications of the microsatellite regions were performed as multiplex reactions by pooling two to three primer sets in a single reaction. Primer sets with the same multiplex number were combined in a reaction. Therefore, for a set of 10 markers, five PCR reactions are needed.

marker AF12, one allele (438) occurred in all isolates from Oman and Pakistan but was unique to the population. The population from Indonesia included seven private alleles with two alleles occurring in single isolates (Table 3). Isolates from Vietnam included a total of 18 private alleles with four alleles occurring only once in the population. The DNA sequences of the private alleles confirmed that the variation observed between alleles was within the microsatellite repeat region, following a stepwise mutation pattern (Fig. S1).

**Table 3**

Summary statistics, including the number of alleles, frequencies of private and predominant alleles, allelic range and gene diversity for each microsatellite marker screened on 160 *C. manginecans* isolates.

| | | Private alleles [Frequency in population][b] | | | | | |
|---|---|---|---|---|---|---|---|
| SSR locus[a] | Number of alleles | Oman and Pakistan | Indonesia (*Acacia*) | Vietnam (*Acacia*) | Allelic range | Highest allele frequency | Gene diversity[c] |
| **AF2** | 3 | | | **198** [0.043] | 198-206 | 0.851 | 0.248 |
| **AF3** | 3 | | | **233** [0.043] | 218-233 | 0.503 | 0.506 |
| **AF4** | 3 | | | **252** [0.304] | 243-252 | 0.857 | 0.254 |
| **AF5** | 3 | | | **270** [0.043] | 250-270 | 0.721 | 0.402 |
| **AF6** | 2 | | **309** [0.009] | | 294-300 | 0.596 | 0.485 |
| **AF7** | 2 | | | | 322-325 | 0.658 | 0.45 |
| **AF8** | 2 | | | | 346-349 | 0.665 | 0.45 |
| **AF9** | 2 | | | | 412-418 | 0.839 | 0.264 |
| **AF11** | 3 | | | **450** [0.391] | 447-455 | 0.87 | 0.227 |
| **AF12** | 6 | **438** [1.000] | **456** [0.303], **459** [0.028] | **444** [0.043], **447** [0.348] | 438-459 | 0.55 | 0.626 |
| AG7/AG8 | 3 | | | **293** [0.087] | 284-293 | 0.528 | 0.51 |
| CF15/CF16 | 2 | | | | 474-476 | 0.5432 | 0.498 |
| CF23/CF24 | 2 | | | | 157-160 | 0.5217 | 0.499 |
| AAG8 | 2 | | | **180** [0.870] | 177-180 | 0.87 | 0.219 |
| AAG9 | 3 | | **400** [0109] | | 397-403 | 0.503 | 0.564 |
| CAA9 | 4 | **166** [0.037] | | **172** [0.652] | 166-223 | 0.8571 | 0.246 |
| CAA10 | 3 | | | **135** [0.261], **138** [0.435] | 126-138 | 0.901 | 0.185 |
| CAA15 | 2 | | | **320** [0.478] | 320-323 | 0.926 | 0.128 |
| CAA38 | 8 | **306** [0.074] | **288** [0.009], **324** [0.136], **369** [0.027] | **207** [0.565], **234** [0.391] | 207-369 | 0.603 | 0.598 |
| CAA80 | 3 | | | **316** [0.478] | 304-316 | 0.696 | 0.452 |
| CAT3K | 2 | | | **306** [0.478] | 306-310 | 0.932 | 0.128 |
| CAT9X | 3 | | | **279** [0.826] | 279-286 | 0.851 | 0.252 |
| CAT1200 | 2 | | | | 376-379 | 0.857 | 0.237 |
| CAG5 | 2 | | | | 319-322 | 0.894 | 0.19 |

[a] Microsatellite markers indicated in bold were developed in this study.

[b] Private alleles are indicated in bold. These refer to alleles observed only once in a population and to alleles unique to a specific population, regardless of the frequency of the allele.

[c] Gene diversity was calculated for each microsatellite marker, based on Nei (1973) gene diversity calculations.

*3.4. Statistical analyses of SSR data*

Analysis of the population differentiation, based on AMOVA, supported the existence of genetically distinct populations with an overall $\phi_{PT}$=0.787 (P=0.001). Of the total variation observed for all the isolates, 79% was due to genetic differences among populations and only 21% due to variation within the population (Table 4). Pairwise comparisons showed that

**Table 4**

Summary of the analysis of molecular variance (AMOVA) of the *C. manginecans* isolates collected from Oman, Pakistan, Indonesia and Vietnam to evaluate the genetic variation within and between populations.

| Source | df[a] | Estimated variance | Proportion of variance components (%) | Statistic | Value | Probability |
|---|---|---|---|---|---|---|
| Among populations | 3 | 3776.117 | 79% | | | |
| Within populations | 37 | 1019.719 | 21% | φPT | 0.787 | 0.001 |
| Total | 40 | 4795.837 | 100% | | | |

[a] Degrees of freedom.

**Table 5**

Pairwise comparison of population differentiation ($\phi_{PT}$) among the four populations investigated in this study.

| Sampling group | Oman and Pakistan Legumes | Oman and Pakistan Mango | Indonesia *Acacia* |
|---|---|---|---|
| Oman and Pakistan Legumes | - | - | - |
| Oman and Pakistan Mango | 0.000* | - | - |
| Indonesia *Acacia* | 0.242* | 0.219* | - |
| Vietnam *Acacia* | 0.726 | 0.682 | 0.842 |

*Value is not statistically significant, P-value > 0.05.

most of this variation can be ascribed to the *Acacia* populations from Vietnam and Indonesia. The Vietnam population was genetically well differentiated from the other populations with

$\phi_{PT}$ values ranging from 0.682 to 0.842 (P=0.001) (Table 5). The Indonesia population was less differentiated from the *M. indica* and legume populations ($\phi_{PT}$= 0.219 and 0.242, respectively) and no differentiation was observed between the isolates from *M. indica* and legume trees from Oman and Pakistan. This suggests the possible existence of three populations.

Possible sub-structuring in Indonesia was not evident as AMOVA analyses indicated a $\phi_{PT}$ of 0.262 (P=0.009) and the majority of variation observed seemed to be within collection sites (74%) and not between collection sites (26%). Pairwise comparisons of $\phi_{PT}$ values, however, did support the region of Kerinci to be significantly differentiated from Teso and Logas ($\phi_{PT}$= 0.415 and 0.462, respectively).

Analysis of population structure was based on the clone-corrected sample set of four populations, consisting of 40 individual haplotypes. The Bayesian clustering method used for analysis indicated the most likely number of genetic clusters (K) to be three (Fig. S2). For these K-values all 20 iterations were congruent. This K-value is in agreement with the expected number of clusters, when considering the gene diversity between the populations, together with the results from the AMOVA analyses. Preliminary runs indicated no significant difference in the calculation of K whether a model assuming admixture was considered or not or whether correlation between alleles was considered. When isolates were considered to consist of three genetic clusters (Fig. 2), the isolates from Oman and Pakistan were assigned to a single, distinct cluster but the individuals from *Acacia* from Indonesia had 30% membership to this cluster and those from Vietnam had 22%. The individuals from Indonesia primarily belonged to a distinct cluster (64% of their identity) and the Vietnam isolates had 75% unique genetic identity to another cluster with very little admixture between the two clusters. As no population differentiation could be seen between the *M. indica* and legume populations from Oman and Pakistan, based on pairwise comparisons and STRUCTURE analyses, these isolates were combined into a single population for subsequent analyses.

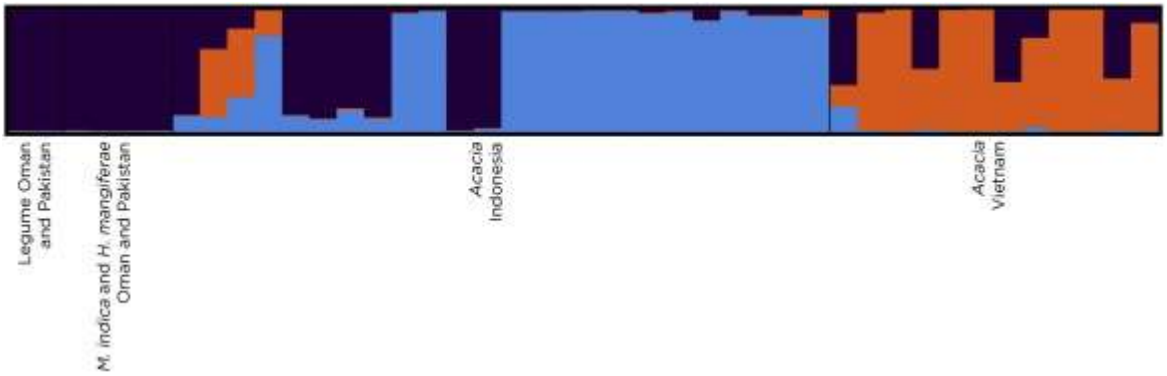


**Fig. 2.** Analysis of population structure (K) of a clone-corrected dataset of the *C. manginecans* isolates collected from legume trees and *M. indica* in Oman and Pakistan and from *Acacia* trees in Indonesia and Vietnam. Each individual is represented by a single vertical line and the colours indicate the relatedness of an isolate to a specific cluster. STRUCTURE analysis suggested the isolates most likely consist of three clusters (K=3) and this data is represented in the figure.

Analysis of the data obtained from the microsatellite markers, in POPGENE v. 1.32 (Yeh et al., 1999), showed a gene diversity ranging from 0.127 to 0.624 per locus with an average gene diversity of 0.437 (Table 3). The genetic variation in the different populations, including the number of alleles present and the gene diversity in each population, is presented in Table 6. The number of individuals, used for the clone-corrected dataset of each population, were equal to the number of haplotypes in that population (Table 6). The population from Oman and Pakistan had the lowest average gene diversity (H=0.01 and clone-corrected H=0.05) and the *Acacia* population from Vietnam had the highest (H=0.44 and clone-corrected H=0.494).

A total of 39 haplotypes were identified in all the *C. manginecans* isolates screened. The greatest number of haplotypes (23) and unique haplotypes (12) were found in the Indonesian population (Table 6), but this was also the population with the largest sample size. The expected numbers of haplotypes, based on rarefaction of 23 individuals, were 9.77 in the Indonesian population and 11 in the Vietnam population (Table 6), suggesting that the genotypic richness might be higher in Vietnam. The isolates from Oman and Pakistan had the lowest genotypic diversity with rarefaction, based on 23 individuals (eG=1.49), and the isolates from *Acacia* from Vietnam had the highest (eG=7.25) (Table 6).

**Table 6**

Summary statistics of isolates collected from Oman, Pakistan, Indonesia and Vietnam indicating the genetic diversity of the populations.

| Population | No. of isolates | No. of alleles | No. of alleles/ locus | No. of private alleles | No. of haplotypes (Priv. haplotypes[a]) | Expected no. of haplotypes[b] | Most frequent haplotype[c] | H[d] | H (clone corrected)[d] | G[e] | eG[e] |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Oman and Pakistan | 27 | 27 | 1.13 | 3 | 5 (3) | 4.54 | 82% | 0.01 | 0.05 | 1.48 | 1.49 |
| Indonesia (*Acacia*) | 110 | 50 | 2.08 | 7 | 23 (12) | 9.77 | 31% | 0.23 | 0.287 | 6.4 | 5.69 |
| Vietnam (*Acacia*) | 23 | 60 | 2.5 | 18 | 11 (5) | 11 | 21% | 0.44 | 0.494 | 7.25 | 7.25 |

[a] Refers to haplotypes that are unique to the population and occurs only once in the population.

[b] Calculated using rarefaction, based on the smallest population size of 23 individuals.

[c] Indicates the percentage of isolates in the population that contain the most frequent haplotype.

[d] The average gene diversity (H), obtained from all the microsatellite markers combined, was calculated in each population and on the clone-corrected populations, based on Nei (1973) gene diversity calculations.

[e] Genotypic diversity (G) was calculated using the method of Stoddard and Taylor (1988) and the estimated G (eG) was obtained using rarefaction, based on the smallest sample size of 23 individuals.

The genetic distance calculated between all the isolates (non clone-corrected dataset), displayed in the NJ tree (Fig. 3), showed distinct groupings, clustering the isolates from Oman and Pakistan in a single clade with little variation. The isolates from Indonesia grouped together but this group was also sub-divided into smaller clusters, indicating a higher level of variation in the population. Some of the clusters corresponded with collection sites (Teso and Pelalawan, respectively) and likely consist of a high number of clonal isolates. Isolates with identical haplotypes were found in five different regions (Teso, Kerinci, Baserah, Pelalawan and Sei Kebaro), indicating significant movement of the pathogen

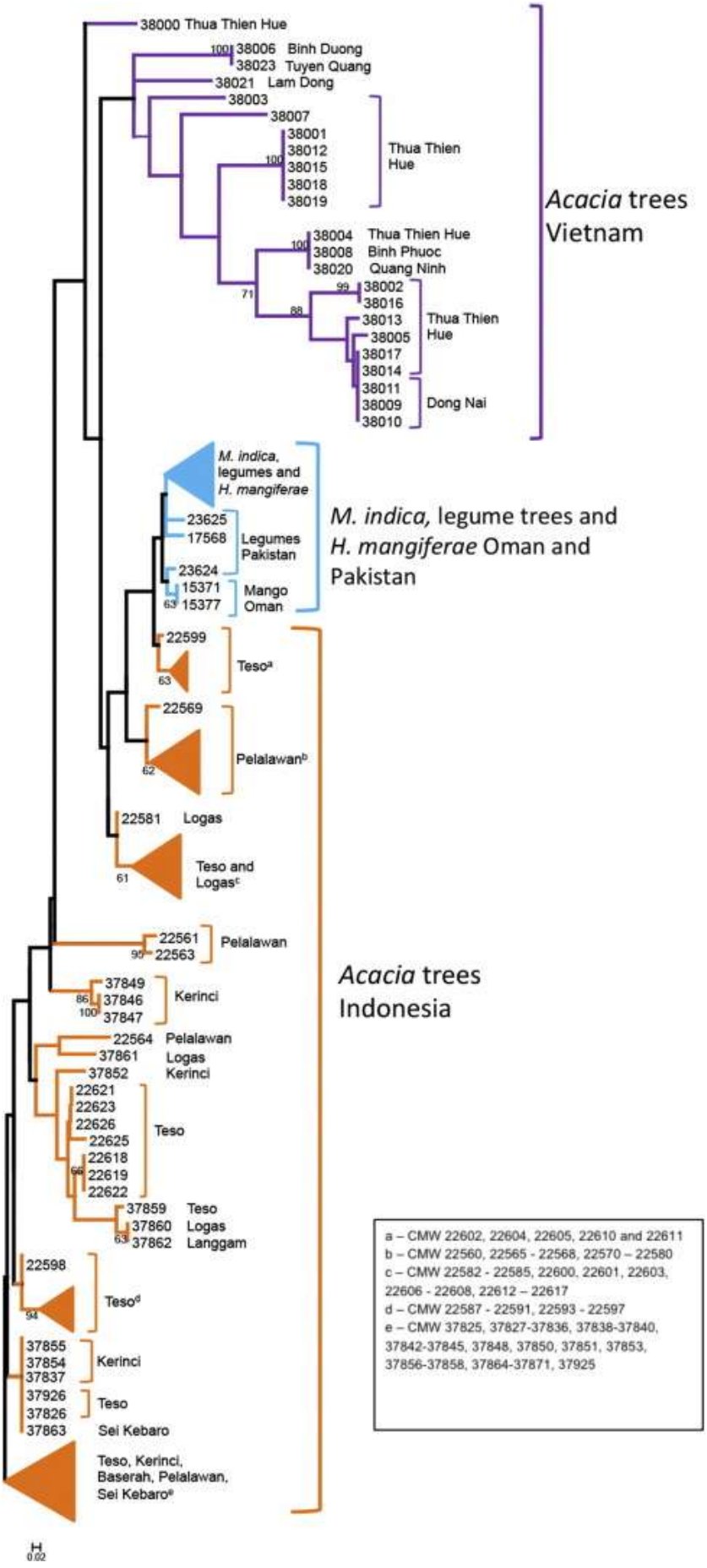


**Fig. 3.** A neighbor joining tree that represents the genetic distance between the isolates investigated in this study. Distances were determined based on the proportion of shared microsatellite alleles (Chakraborty and Jin, 1993) and the statistical support for each branch was determined by 1000 bootstrap replicates. Bootstrap values higher than 60% are indicated in the figure. Each of the four pre-defined populations, based on host-association and country of origin, is indicated in a different colour. Distinct groupings could be observed among the isolates, with all the isolates from *Acacia* from Indonesia forming one cluster (orange), and those from Vietnam forming another cluster (purple). All the isolates from Oman and Pakistan were very closely related (blue).

throughout the country. One of the clusters, consisting of isolates from Teso, had a high genetic relatedness to the isolates from Oman and Pakistan. Comparison of the original microsatellite data (Table S1) showed that these isolates differed at only one or two alleles from the Oman/Pakistan population. The Vietnam isolates grouped in a cluster significantly distinct from all other populations. The haplotypes of the Vietnam isolates showed no distinct clustering correlated to the collection sites with a single haplotype identified in three regions which are 800 km and 1700 km apart (Fig. 1D).

## 4. Discussion

The microsatellite markers applied in this study showed significantly higher genetic diversity among the isolates from South East Asia, compared to those from the Middle East. The highest genetic diversity was found in the isolates from Vietnam. This suggests that the fungal pathogen has existed in that area for a long period of time and could possibly be considered indigenous to this area. Based on DNA sequence comparisons, this study confirmed that the isolates associated with severe wilt of *Acacia* spp. in Vietnam represent the same species as that killing *A. mangium* in Indonesia and *M. indica* in Oman and Pakistan. Thus, all of the isolates investigated were confirmed to represent the single species *C. manginecans*. When species concepts alternative to the genealogical concordance phylogenetic concept are considered, these isolates could be seen as different haplotypes of *C. fimbriata* (Oliveira et al., 2015). However, based on previous phylogenetic analyses including five gene regions (Fourie et al., 2014) we consider *C. manginecans* as a species distinct from the sweet potato black rot fungus which includes the type of *C. fimbriata*.

The new microsatellite markers designed in this study detected similar values of genetic variation for the *C. manginecans* isolates as the previously designed (Barnes et al., 2001; Steimel et al., 2004), broad range markers. The advantage of the new markers is that they were designed to be amplified in multiplex, reducing costs and time. It has been

suggested that the cross-species transfer of microsatellite markers could show less polymorphism in a different species (Barbará et al., 2007). However, the gene diversity for the entire set of samples ranged from 0.254 to 0.624 for each of the new markers with an average gene diversity of 0.393, which is only slightly higher than for the markers developed previously (Barnes et al., 2001; Steimel et al., 2004). This provides support for the fact that either the new or the previously designed markers would give a fair representation of the actual diversity in the populations being studied.

The low level of genetic diversity identified in the isolates of *C. manginecans* from legume trees in Oman and Pakistan was similar to that found for isolates from *M. indica* and the insect vector *H. mangiferae*. This supports the hypothesis that the species was introduced to those areas and that a single population has spread widely throughout the countries to different hosts (Al Adawi et al., 2014). The same haplotype present in the isolates from *M. indica* and *H. mangiferae* was found on the native legume trees (Fig. 3). This raises the prospect that *C. manginecans* has undergone a host jump/shift from non-native and commercially propagated fruit trees to native trees or vice versa. The introduction of a pathogen along with an introduced, commercially planted host could easily result in a host jump onto native hosts as seen, for example, with the globally distributed *Neofusicoccum parvum* (Sakalidis et al., 2013).

The genetic diversity observed for *C. manginecans* in Vietnam was higher than that observed for the same pathogen in Indonesia. This is substantiated by the high proportion of unique haplotypes (48% of the isolates) and private alleles (18) found in the isolates from Vietnam, compared to Indonesia (21% and 7, respectively). This might, at least in part, be explained by the larger geographical distribution of sites investigated in Vietnam that included various ecological zones in the country. This is in contrast to the regions in Indonesia that were in closer proximity to each another.

The *C. manginecans* population from *Acacia* spp. in Vietnam had the highest levels of genetic diversity observed for the species to date. The values observed in these isolates (H=0.44 and eG=7.25) are comparable with those for other populations of *Ceratocystis* spp.

that are considered native. For example, highly diverse populations of *C. albifundus* in Africa had a gene diversity of H=0.41 (Barnes et al., 2005) and this continent is believed to be the area of origin of the pathogen. Similarly, high values of genetic diversity were found for the native *Theobroma cacao* pathogen, *C. cacaofunesta* (H=0.198), in Brazil (Ferreira et al., 2010).

The high level of genetic diversity observed for isolates from Vietnam suggests that the pathogen has been present in South East Asia for a long period of time. In this regard, they would have accumulated variation through events such as mutation or recombination (McDonald, 1997). The possible existence of an indigenous pathogen would explain the serious disease that it has begun to cause on a non-native *Acacia* spp., introduced from Australia (Roux and Wingfield, 2009; Sein and Mitlöhner, 2011). The pathogen could have existed on a native host and adapted to infect these *Acacia* trees after they were introduced to South East Asia. This is not an uncommon phenomenon and has, for example, been seen for introduced *Eucalyptus* plantations, where native pathogens such as C*hrysoporthe austroafricana* and *C. cubensis*, present on native Melastomataceae, adapted to cause serious disease on non-native *Eucalyptus* trees (Nakabonge et al., 2006; Wingfield et al., 2015). The existence of a native pathogen would also increase the possibility of an association emerging between the fungus and insect vectors, thus resulting in the emergence of new disease problems (Atkinson and Peck, 1994; Doran and Skelton, 1982; Moran et al., 1989; Mukherjee, 1972).

All the isolates used in this study, collected from the Middle East and South East Asia, could reliably be separated into three distinct clusters. The population from Vietnam was supported to be genetically distinct from Indonesia and from the isolates collected from the Middle East, based on the pairwise comparisons of $\phi_{PT}$ values (Table 5). However, the population from Indonesia appeared less differentiated from the populations in Oman and Pakistan. The haplotype of one of the isolates found in Indonesia differed from the Oman and Pakistan isolates by only a single allele. This difference might have been due to a recent

mutation event in a single isolate. It is also possible that the introduction of *C. manginecans* into Oman and Pakistan could have originated from one of the other countries considered in this study.

Analyses of the genetic sub-structuring for the Indonesian isolates, based on $\phi_{PT}$ pairwise comparisons, showed that those from Kerinci, Indonesia, were genetically differentiated from the Teso and Logas isolates. However, the NJ tree (Fig. 3) indicated definitive admixture with identical haplotypes observed in isolates from all three of these collection sites. Similarly, admixture could be seen in the Vietnam collection of isolates with no clear structure linked to geographical location. This could be due to anthropogenic dispersal of wood products, transfer of vegetative cuttings to different plantation areas or dispersal by insects. The involvement of insects in the population admixture of *C. manginecans* is supported by the identical haplotype found on *H. mangifera* and its tree hosts in Oman and Pakistan (Al Adawi et al., 2013a). Similar involvement of tree-infesting bark beetles in the distribution of a fungus has, for example, been observed for the conifer pathogens *Ophiostoma ips* (Zhou et al., 2007) and *Grosmannia clavigera* (Tsui et al., 2012).

Results of this study have substantially expanded available knowledge regarding the geographic distribution of the important tree pathogen, *C. manginecans*. The pathogen was previously known only on *M. indica, H. mangiferae* (as a vector) and legume trees in the Middle East and on non-native *Acacia* spp. in Indonesia (Al Adawi et al., 2013b; Tarigan et al., 2011). Results of this study confirmed that *C. manginecans* also occurs on *Acacia* spp. in Vietnam, of which the species *A. auriculiformis* is a newly reported host. *Ceratocystis manginecans* has become a serious threat to fruit and forest plantation industries in the Middle East and South East Asian countries. In addition, the same fungus has been reported on *Acacia* spp. in Malaysia (Brawner et al., 2015) and on *Eucalyptus* trees in South China (Chen et al., 2013). The SSR markers used in this study should, in future, be screened on isolates from these countries to determine the genetic structure of additional populations.

This information could assist in determining the global distribution patterns of the pathogen as well as in establishing effective quarantine and trade regulations between countries.

**Acknowledgements**

This work was enabled by the financial support of the members of the Tree Protection Co-operative programme (TPCP) at the Forestry and Agricultural Biotechnology Institute (FABI), University of Pretoria. We acknowledge the National Research Foundation (NRF) for financial support and the Thuthuka Foundation (grant no 80670) for a student bursary grant for Arista Fourie. We thank Carlos Batista for his technical assistance in this study. We also thank Nguyen M. Chi from the Forest Protection Research Centre, Vietnamese Academy of Forest Sciences, and Tran T.T. Tam from Thai Nguyen University, Vietnam, for collection and isolation of fungal samples in Vietnam. For collections and isolations in Indonesia, we thank Mr. Mathin Tarigan from PT Riau Andalan Pulp and Paper, Pekanbaru, Indonesia.

**References**

Al Adawi, A.O., Al Jabri, R.M., Deadman, M.L., Barnes, I., Wingfield, B., Wingfield, M.J., 2013a. The mango sudden decline pathogen, *Ceratocystis manginecans*, is vectored by *Hypocryphalus mangiferae* (Coleoptera: Scolytinae) in Oman. Eur. J. Plant Pathol. 135, 243-251.

Al Adawi, A.O., Barnes, I., Khan, I.A., Al Subhi, A.M., Al Jahwari, A.A., Deadman, M.L., Wingfield, B.D., Wingfield, M.J., 2013b. *Ceratocystis manginecans* associated with a serious wilt disease of two native legume trees in Oman and Pakistan. Australas. Plant Pathol. 42, 179-193.

Al Adawi, A.O., Barnes, I., Khan, I.A., Deadman, M.L., Wingfield, B.D., Wingfield, M.J., 2014. Clonal structure of *Ceratocystis manginecans* populations from mango wilt disease in Oman and Pakistan. Australas. Plant Pathol. 43, 393-402.

Al Adawi, A.O., Deadman, M.L., Al Rawahi, A.K., Al Maqbali, Y.M., Al Jahwari, A.A., Al Saadi, B.A., Al Amri, I.S., Wingfield, M.J., 2006. Aetiology and causal agents of mango sudden decline disease in the Sultanate of Oman. Eur. J. Plant Pathol. 116, 247-254.

Al Adawi, A.O., Deadman, M.L., Al Rawahi, A.K., Khan, A.J., Al Maqbali, Y.M., 2003. *Diplodia theobromae* associated with sudden decline of mango in the Sultanate of Oman. Plant Pathol. 52, 419-419.

Atkinson, T.H., Peck, S.B., 1994. Annotated Checklist of the Bark and Ambrosia Beetles (Coleoptera: Platypodidae and Scolytidae) of Tropical Southern Florida. Fla. Entomol. 77, 313-329.

Barbará, T., Palma-Silva, C., Paggi, G.M., Bered, F., Fay, M.F., Lexer, C., 2007. Cross-species transfer of nuclear microsatellite markers: potential and limitations. Mol. Ecol. 16, 3759-3767.

Barnes, I., Guar, A., Burgess, T., Roux, J., Wingfield, B.D., Wingfield, M.J., 2001. Microsatellite markers reflect intra-specific relationships between isolates of the vascular wilt pathogen *Ceratocystis fimbriata*. Mol. Plant Pathol. 2, 319-325.

Barnes, I., Nakabonge, G., Roux, J., Wingfield, B.D., Wingfield, M.J., 2005. Comparison of populations of the wilt pathogen *Ceratocystis albifundus* in South Africa and Uganda. Plant Pathol. 54, 189-195.

Brawner, J., Japarudin, Y., Lapammu, M., Rauf, R., Boden, D., Wingfield, M.J., 2015. Evaluating the inheritance of *Ceratocystis acaciivora* symptom expression in a diverse *Acacia mangium* breeding population. South. For. 77, 83-90.

Chakraborty, R., Jin, L., 1993. A unified approach to study hypervariable polymorphisms: Statistical considerations of determining relatedness and population distances, in: Pena, S.D.J., Chakraborty, R., Epplen, J.T., Jeffreys, A. (Eds.), DNA Fingerprinting: State of the Science. Birkhäuser Basel, pp. 153-175.

Chen, S., Van Wyk, M., Roux, J., Wingfield, M.J., Xie, Y., Zhou, X., 2013. Taxonomy and pathogenicity of *Ceratocystis* species on *Eucalyptus* trees in South China, including *C. chinaeucensis* sp. nov. Fungal Divers. 58, 267-279.

De Beer, Z.W., Duong, T.A., Barnes, I., Wingfield, B.D., Wingfield, M.J., 2014. Redefining *Ceratocystis* and allied genera. Stud. Mycol. 79, 187-219.

Doran, J.C., Skelton, D.J., 1982. *Acacia mangium* seed collections for international provenance trials, Forest Genetics Resource Information, Food and Agriculture Organization, Rome, Italy, pp. 47-53.

Engelbrecht, C.J.B., Harrington, T.C., 2005. Intersterility, morphology and taxonomy of *Ceratocystis fimbriata* on sweet potato, cacao and sycamore. Mycologia 97, 57-69.

Evanno, G., Regnaut, S., Goudet, J., 2005. Detecting the number of clusters of individuals using the software structure: a simulation study. Mol. Ecol. 14, 2611-2620.

Excoffier, L., Smouse, P.E., Quattro, J.M., 1992. Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. Genetics 131, 479-491.

FAOSTAT, 2011. FAOSTAT On-line. Rome. United Nations Food and Agriculture Organization. Available at: http://faostat3.fao.org/faostat-gateway/go/to/home/E. Accessed: 9 September 2013.

Ferreira, E.M., Harrington, T.C., Thorpe, D.J., Alfenas, A.C., 2010. Genetic diversity and interfertility among highly differentiated populations of *Ceratocystis fimbriata* in Brazil. Plant Pathol. 59, 721-735.

Fourie, A., Wingfield, M., Wingfield, B., Barnes, I., 2014. Molecular markers delimit cryptic species in *Ceratocystis sensu stricto*. Mycol. Progress 14, 1-18.

Forestry Science Institute of Vietnam, 2009. Asia-Pacific Forestry Sector Outlook Study II: Vietnam Forestry Outlook Study. Food and Agriculture Organization of the United Nations, Bangkok.

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

Goodwin, S.B., Drenth, A., Fry, W.E., 1992. Cloning and genetic analysis of two highly polymorphic, moderately repetitive nuclear DNAs from *Phytophthora infestans*. Curr. Genet. 22, 107-115.

Grünwald, H.J., Goodwin, S.B., Milgroom, M.G., Fry, W.E., 2003. Analysis of genotypic diversity for populations of microorganisms. Phytopathology 93, 738-746.

Harrington, T.C., Kazmi, M.R., Al-Sadi, A.M., Ismail, S.I., 2014. Intraspecific and intragenomic variability of ITS rDNA sequences reveals taxonomic problems in *Ceratocystis fimbirata sensu stricto*. Mycologia 106, 224-242.

Kamvar, Z.N., Tabima, J.F., Grünwald, N.J., 2014. Poppr: an R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction. PeerJ 2, e281.

Kopelman, N.M., Mayzel, J., Jakobsson, M., Rosenberg, N.A., Mayrose, I., 2015. Clumpak: a program for identifying clustering modes and packaging population structure inferences across K. Mol. Ecol. Res. 15, 1179-1191.

Kumar, B., Mistry, N.C., Singh, B., Gandhi, C.P., 2011. Indian Horticulture Database National Horticulture Board, pp. 1-296.

Langella, O., 1999. Populations 1.2.30. Available at: http://www.bioinformatics.org/populations. Accessed: 2 October 2015.

Maslin, B., 2013. KEW Royal Botanical Gardens (*Acacia mangium*). Available at: http://www.kew.org/science-conservation/plants-fungi/acacia-mangium-brown-salwood. Accessed: 9 July 2013.

Masood, A., Saeed, S., Erbilgin, N., Jung Kwon, Y., 2010. Role of stressed mango host conditions in attraction of and colonization by the mango bark beetle *Hypocryphalus mangiferae* Stebbing (Coleoptera: Curculionidae: Scolytinae) and in the symptom development of quick decline of mango trees in Pakistan. Entomol. Res. 40, 316-327.

Masood, A., Saeed, S., Sajjad, A., 2008. Characterization and damage patterns of different bark beetle species associated with mango sudden death syndrome in Punjab, Pakistan. Pakistan Entomol. 30, 163-168.

McDonald, B.A., 1997. The population genetics of fungi: Tools and techniques. Phytopathology 87, 448-453.

Moran, G.F., Muona, O., Bell, J.C., 1989. *Acacia mangium*: A tropical forest tree of the coastal lowlands with low genetic diversity. Evolution 43, 231-235.

Mukherjee, S.K., 1972. Origin of mango (*Mangifera indica*). Econ. Bot. 26, 260-264.

Naidoo, K., Steenkamp, E.T., Coetzee, M.P.A., Wingfield, M.J., Wingfield, B.D., 2013. Concerted Evolution in the Ribosomal RNA Cistron. PLoS ONE 8, e59355.

Nakabonge, G., Roux, J., Gryzenhout, M., Wingfield, M.J., 2006. Distribution of Chrysoporthe canker pathogens on *Eucalyptus* and *Syzygium* spp. in eastern and southern Africa. Plant Disease 90, 734-740.

Nei, M., 1973. Analysis of Gene Diversity in Subdivided Populations. Proc. Natl. Acad. Sci. 70, 3321-3323.

Nei, M., 1987. Molecular evolutionary genetics. Columbia University Press, New York.

Ocasio-Morales, R.G., Tsopelas, P., Harrington, T.C., 2007. Origin of *Ceratocystis platani* on native *Platanus orientalis* in Greece and its impact on natural forests. Plant Disease 91, 901-904.

Oliveira, L.S.S., Harrington, T.C., Ferreira, M.A., Damacena, M.B., Al-Sadi, A.M., Al-Mahmooli, I.H.S., Alfenas, A.C., 2015. Species or genotypes? Reassessment of four recently described species of the Ceratocystis wilt pathogen, *Ceratocystis fimbriata*, on *Mangifera indica*. Phytopathology 105, 1229-1244.

Peakall, R.O.D., Smouse, P.E., 2006. Genalex 6: genetic analysis in Excel. Population genetic software for teaching and research. Mol. Ecol. Notes 6, 288-295.

Pritchard, J.K., Stephens, M., Donnelly, P., 2000. Inference of population structure using multilocus genotype data. Genetics 155, 945-959.

Roux, J., Wingfield, M.J., 2009. *Ceratocystis* species: emerging pathogens of non-native plantation *Eucalyptus* and *Acacia* species. South. For. 71, 115-120.

Roux, J., Wingfield, M.J., 2013. *Ceratocystis* species on the African continent, with particular reference to *C. albifundus*, an African species in the *C. fimbriata sensu lato* species complex, in: Seifert, K.A., De Beer, Z.W., Wingfield, M.J. (Eds.), Ophiostomatoid Fungi: Expanding Frontiers. CBS-KNAW Fungal Biodiversity Centre, AD Utrecht, The Netherlands, pp. 131-138.

Sakalidis, M.L., Slippers, B., Wingfield, B.D., Hardy, G.S.J., Burgess, T.I., 2013. The challenge of understanding the origin, pathways and extent of fungal invasions:

global populations of the *Neofusicoccum parvum–N. ribis* species complex. Divers. Distrib. 19, 873-883.
Sein, C.C., Mitlöhner, R., 2011. *Acacia mangium* Willd. Ecology and silviculture in Vietnam. Center for International Forestry Research (CIFOR), pp. 1-26.
Steimel, J., Engelbrecht, C.J.B., Harrington, T.C., 2004. Development and characterization of microsatellite markers for the fungus *Ceratocystis fimbriata*. Mol. Ecol. Notes 4, 215-218.
Stoddart, J.A., Taylor, J.F., 1988. Genotypic diversity: estimation and prediction in samples. Genetics 118, 705-711.
Swofford, D.L., 2002. PAUP*. Phylogenetic Analysis Using Parsimony (*and other methods), 4 ed. Sinauer Associates, Sunderland, MA.
Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., Kumar, S., 2011. MEGA5: Molecular Evolutionary Genetics Analysis Using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony Methods. Mol. Biol. Evol. 28, 2731-2739.
Tarigan, M., Roux, J., Van Wyk, M., Tjahjono, B., Wingfield, M.J., 2011. A new wilt and die-back disease of *Acacia mangium* associated with *Ceratocystis manginecans* and *C. acaciivora sp*. nov. in Indonesia. SA J. Bot. 77, 292-304.
Thu, P.Q., Qynh, D.N., Dell, B., 2012. *Ceratocystis* sp. causes crown wilt of *Acacia* spp. planted in some ecological zones of Vietnam. J. Plant Prot. N5 (245), p24-29.
Thurston, M.I., Field, D., 2005. Msatfinder on-line v. 2.0. Available at: http://www.genomics.ceh.ac.uk/msatfinder/. Accessed: 15 January 2013.
Tsui, C.K.M., Roe, A.D., El-Kassaby, Y.A., Rice, A.V., Alamouti, S.A.M., Sperling, F.A.H., Cooke, J.E.K., Bohlmann, J., Hamelin, R.C., 2012. Population structure and migration pattern of a conifer pathogen, *Grosmannia clavigera*, as influenced by its symbiont, the mountain pine beetle. Mol. Ecol. 21, 71-86.
Van der Nest, M.A., Bihon, W., Vos, L.D., Naidoo, K., Roodt, D., Rubagotti, E., Slippers, B., Steenkamp, E.T., Wilken, P.M., Wilson, A., Wingfield, M.J., Wingfield, B.D., 2014. IMA Genome-F 2: *Ceratocystis manginecans*, *Ceratocystis moniliformis*, *Diplodia sapinea*. IMA Fungus 5, 135–140.
Van Wyk, M., Al Adawi, A.O., Khan, I.A., Deadman, M.L., Al Jahwari, A.A., Wingfield, B.D., Ploetz, R., Wingfield, M.J., 2007. *Ceratocystis manginecans* sp. nov., causal agent of a destructive mango wilt disease in Oman and Pakistan. Fungal Divers. 27, 213-230.
White, T.J., Bruns, T., Lee, S., Taylor, J., 1990. Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics, in: Innis, M.A., Gelfand, D.H., Sninsky, J.J., White, T.J. (Eds.), PCR Protocols: A sequencing guide to methods and applications. Academic Press, San Diego, USA, pp. 315-322.
Wingfield, M.J., Brockerhoff, E.G., Wingfield, B.D., Slippers, B., 2015. Planted forest health: The need for a global strategy. Science 349, 832-836.
Wingfield, M.J., Roux, J., Wingfield, B.D., Slippers, B., 2013. *Ceratocystis* and *Ophiostoma*: international spread, new associations and plant health, in: Keith A. Seifert, De Beer, W., Wingfield, M.J. (Eds.), Ophiostomatoid Fungi: Expanding Frontiers. CBS-KNAW Fungal Biodiversity Centre, Netherlands, pp. 191-200.
Yeh, F.C., Yang, R., Boyle, T., 1999. POPGENE Version 1.31: Microsoft window-based freeware for population genetic analysis. University of Alberta and Centre for International Forestry Research.
Zhou, X., Burgess, T.I., De Beer, Z.W., Lieutier, F., Yart, A., Klepzig, K., Carnegie, A., Portales, J.M., Wingfield, B.D., Wingfield, M.J., 2007. High intercontinental migration rates and population admixture in the sapstain fungus *Ophiostoma ips*. Mol. Ecol. 16, 89-99.