Antonie van Leeuwenhoek (2016) 109:21–33
DOI 10.1007/s10482-015-0606-0

ORIGINAL PAPER

Three genetic groups of the *Eucalyptus* stem canker pathogen *Teratosphaeria zuluensis* introduced into Africa from an unknown source

Luke Jimu · ShuaiFei Chen · Michael J. Wingfield · Eddie Mwenje · Jolanda Roux

Received: 15 July 2015 / Accepted: 5 October 2015 / Published online: 24 October 2015

Abstract The *Eucalyptus* stem canker pathogen *Teratosphaeria zuluensis* was discovered in South Africa in 1988 and it has subsequently been found in several other African countries as well as globally. In this study, the population structure, genetic diversity and evolutionary history of *T. zuluensis* were analysed using microsatellite markers to gain an enhanced understanding of its movement in Africa. Isolates were collected from several sites in Malawi, Mozambique, Uganda and Zambia. Data obtained were compared with those previously published for a South African population. The data obtained from 334 isolates, amplified across eight microsatellite loci, were used for assignment, differentiation and genetic diversity tests. STRUCTURE analyses, θ_{st} and genetic distances revealed the existence of two clusters, one dominated by isolates from South Africa and the other by isolates from the Zambezi basin including Malawi, Mozambique and Zambia. High levels of admixture were found within and among populations, dominated by the Mulanje population in Malawi. Moderate to low genetic diversity of the populations supports the previously held view that the pathogen was introduced into Africa. The clonal nature of the Ugandan population suggests a very recent introduction, most likely from southern Africa.

Keywords Admixture · Evolutionary history · Multi-locus genotype · Population structure · Genetic diversity

L. Jimu (✉)
Department of Plant Production and Soil Science, Forest Science Postgraduate Programme, Forestry and Agricultural Biotechnology Institute (FABI), University of Pretoria, Private Bag X20, Hatfield, Pretoria 0028, South Africa
e-mail: jolanda.roux@fabi.up.ac.za

S. Chen · M. J. Wingfield · J. Roux
Department of Microbiology and Plant Pathology, FABI, University of Pretoria, Private Bag X20, Hatfield, Pretoria 0028, South Africa

E. Mwenje
Bindura University of Science Education (BUSE), P. Bag 1020, Bindura, Zimbabwe

Introduction

Eucalyptus species are widely cultivated in plantations outside Australia where most of these trees are native. In Africa, they are the most widely cultivated group of trees, covering approximately 22.4 % of the continent's total plantation area (Chamshama and Nwonwu 2004). *Eucalyptus* species were introduced to the continent in the second half of the nineteenth century and the beginning of the twentieth century (FAO 2011). Approximately 11.2 % (2.2 million ha) of the global plantation area under *Eucalyptus* species occurs in Africa (Rejmánek and Richardson 2011). The domestication of *Eucalyptus* has seen a rapid movement of germplasm, particularly seed and eucalypt

timber products, across borders. This global trade has been identified as a major conduit for the spread of important eucalypt pathogens globally (Wingfield et al. 2008, 2015; Montesclaros declaration, http://www.iufro.org/science/divisions/division-7/70000/publications/montesclaros-declaration/).

Teratosphaeria stem canker is one of the most important *Eucalyptus* diseases to have recently emerged outside Australia. The disease was first reported in the Zululand region of South Africa in 1988 (Wingfield et al. 1997). At that time, the causal pathogen was described as *Coniothyrium zuluense* Wingf., Crous & T.A. Cout. Numerous taxonomic arrangements have led to name changes for this pathogen (Cortinas et al. 2006a, b, c), of which *Teratosphaeria zuluensis* (Crous et al. 2009a, b) is most current. Subsequent to the description of *T. zuluensis*, another closely related species was reported to give rise to the same symptoms on *Eucalyptus* species in South America (Cortinas et al. 2006c). That pathogen was described as *Colletogloeopsis* (now *Teratosphaeria*) *gauchensis* M.N. Cortinas, Crous & M.J. Wingf (Cortinas et al. 2006c). *T. zuluensis* is known to cause disease in South Africa (Wingfield et al. 1997), Mexico (Roux et al. 2002), Thailand (van Zyl et al. 2002), Vietnam (Old et al. 2003), Malawi, Mozambique (Roux et al. 2005), China (Cortinas et al. 2006b), Zambia (Chungu et al. 2010) and Uganda (Jimu et al. 2014). *Teratosphaeria gauchensis* has been reported from Hawaii-USA (Cortinas et al. 2004), Ethiopia (Gezahgne et al. 2005), Uganda (Roux et al. 2005), Argentina, Uruguay (Cortinas et al. 2006c) and recently Portugal (Silva et al. 2014) and Zimbabwe (Jimu et al. 2015).

Teratosphaeria stem canker is characterised by discrete black or dark brown measles-like lesions that merge to form large necrotic patches on stems of *Eucalyptus* (Wingfield et al. 1997; Roux et al. 2002; van Zyl et al. 2002). At the onset of infection, small, discrete necrotic lesions are formed on the green stem tissues of twigs, branches and stems in the upper parts of trees (Cortinas et al. 2006c). These lesions expand, forming elliptical cankers and small black pycnidia are found on the dead bark. Lesions caused by *T. zuluensis* penetrate deeply into the wood and stimulate the formation of pockets filled with kino (Wingfield et al. 1997). The kino exudes from the wood, staining timber during harvesting and sawing (Gezahgne et al. 2003). The kino pockets seriously reduce the quality and strength of sawn timber (Wingfield et al. 1997). Stem infections by the pathogen make peeling of the bark difficult, thereby increasing the costs associated with pulping (Wingfield et al. 1997; Roux et al. 2002; van Zyl et al. 2002). Affected trees display stunted height growth, which is sometimes accompanied by deformed stems, dead tree tops and brush-like flattened crowns on trees that have been repeatedly infected (Wingfield et al. 1997). The stressed trees respond by developing epicormic shoots around the cankers. In severe cases, girdled trees eventually die (Wingfield et al. 1997; van Zyl et al. 2002; Gezahgne et al. 2003; Old et al. 2003).

At the time of its discovery, and for the subsequent decade, *T. zuluensis* was thought to be native to South Africa. This was based on the fact that the pathogen had not been reported elsewhere in the world. Discovery of Teratosphaeria stem canker in many other countries of the world in the early 2000s (van Zyl et al. 2002; Roux et al. 2002; Gezahgne et al. 2003; Cortinas et al. 2006b) cast doubt on a South African origin. Population genetic studies on *T. zuluensis* have also suggested that the pathogen was most likely introduced into South Africa from an unknown source (Cortinas et al. 2010; Chen et al. 2011).

The aim of this study was to consider the population structure and genetic diversity of *T. zuluensis* in southern and eastern Africa. Of particular interest were previous results from Malawi that showed very high diversity values for a population in that country (Cortinas et al. 2010). An additional aim was to obtain a regional perspective for the pathogen, as opposed to the previous country-specific approaches. This was done to gain a better understanding of the introduction and movement of *T. zuluensis* in Africa. We utilised eight polymorphic microsatellite markers designed by Cortinas et al. (2006a, 2010) to determine the population structure, genetic diversity and mode of reproduction for *T. zuluensis* populations collected from various countries in Africa.

Materials and methods

Populations

Six populations of *T. zuluensis* were collected from *Eucalyptus* plantations and woodlots in three countries

of Africa. Three of these were from the Mangochi, Mulanje and Thyolo districts in Malawi, two areas (Gurue and Manica) in Mozambique, and Jinja in Uganda (Table 1). An additional two populations were previously collected in South Africa, by Cortinas et al. (2010), and Zambia by Chungu et al. (2010).

Isolates were collected following a hierarchical sampling procedure. In each *Eucalyptus* plantation showing signs of Teratosphaeria stem canker, a central reference tree with canker symptoms was chosen as a starting point. All symptomatic trees occurring in two diagonal transects intersecting this central tree were sampled based on the presence of visible fungal fruiting bodies on typical lesions. The length of transects ranged from about 100 to 700 m for all sites sampled. Bark and branch samples were placed in separate paper bags and transported to the laboratory for isolation of the pathogen.

Bark pieces were incubated in moist chambers for 1–2 days to induce sporulation of *Teratosphaeria* species. Spore masses were then transferred to Petri dishes containing 2 % MEA (20 g/L malt extract, 15 g/L agar, Biolab, Midrand, South Africa and 1 L deionised water). After 5–7 days, cultures were purified by transferring single hyphal tips to fresh 2 % MEA plates. The plates were incubated at 25 °C for 4 weeks. Representative isolates from each country were deposited with the culture collection (CMW) of the Forestry and Agricultural Biotechnology Institute (FABI), University of Pretoria, South Africa.

Isolate identification

Data were obtained from a total of 334 isolates, including 96 previously genotyped by Cortinas et al. (2010) and 238 genotyped in the current study. Of the 238 isolates genotyped in this study, 56 had previously been collected from Zambia by Chungu et al. (2010). The remaining 182 isolates were collected during this study from six sites in Malawi, Mozambique and Uganda (Table 1). To verify the identity of the collected isolates as *T. zuluensis*, cultures were grouped based on morphology and origin. Three representatives were selected from each group and they were subjected to DNA sequence comparisons. Mycelium was scraped from the surfaces of selected, actively growing cultures using sterilised surgical blades and transferred to 2 mL Eppendorf tubes, freeze dried and ground to a fine powder using a Retsch Mixer MM 301. The method described by Möller et al. (1992) was used to extract DNA. Each sample was treated with 3 µL of RNase A (1 mg/mL) and left overnight to digest RNA. Final DNA working concentrations were adjusted to ~75 ηg/µL, using a Thermo Scientific NanoDrop® ND-1000 Spectrophotometer (NanoDrop Technologies, Wilmington, DE, USA).

The ITS regions (Internal Transcribed Spacer) of the ribosomal DNA operon, including the ITS 1, 5.8S gene and ITS 2 regions, of each sample was amplified using primers ITS 1 and ITS 4 (White et al. 1990). The PCR mix and cycling conditions, as well as product

Table 1 Geographical location and details of *T. zuluensis* isolates used in this study. Details for isolates collected from South Africa were summarised by Cortinas et al. (2010)

Country	Site	Host	Year	Collector	NT[a]	NI[b]
Malawi	MN	*E. grandis*	2011	Jimu L, Roux J	12	23
	ML	*E. grandis*	2011	Jimu L, Roux J	26	35
	TY	*E. grandis*	2011	Jimu L, Roux J	11	21
Mozambique	GR	*E. grandis*, GU, GC	2011	Jimu L, Roux J	27	42
	MC	*E. grandis*, GC	2011	Jimu L, Roux J	33	43
Uganda	J	*E. grandis*	2012	Jimu L	18	18
Zambia[c]	CT & K	*E. grandisE. cloeziana*	2007	Roux J, Chungu D	–	56

CT Chati, *GR* Gurue, *J* Jinja, *MC* Manica, *ML* Mulanje, *MN* Mangochi, *TY* Thyolo

[a] Number of trees

[b] Number of isolates

[c] Isolates previously collected by Chungu et al. (2010)

clean-up and sequence PCRs were the same as those described by Jimu et al. (2014). Sequences obtained were compared with those in GenBank (http://blast.ncbi.nlm.nih.gov/) using a Blast search. The sequences obtained in this study were compared with sequences (GenBank No. DQ240207, DQ240214) for epitype strains (CMW 17321, CMW 17322) of *T. zuluensis* (Cortinas et al. 2006c).

Simple sequence repeats (SSR) PCR and gene scan analyses

DNA for simple sequence repeat (SSR) PCR was extracted as described above for all isolates confirmed to represent *T. zuluensis*. Eight primer pairs previously developed for *T. zuluensis* (Cortinas et al. 2006a, 2010) were used to study the populations obtained. The forward primer of each pair was fluorescently labelled with FAM, NED, PEP or VIC dyes for filter set G5 (Applied Biosystems, Foster City, CA, USA). The reverse primers were retained in their original forms. PCR reactions and conditions were the same as those described by Cortinas et al. (2006a, 2010). Fragment analysis reactions for each sample were set as follows: 1 µL of PCR product was added to a mixture of 8.8 µL Hi-Di Formamide and 0.2 µL GenScan 500 LIZ size standard (Applied Biosystems, Foster City, CA, USA). The mixture was denatured at 95 °C for 3 min and then cooled on ice. PCR amplicons were size-separated on an ABI PRISM™ 3100 Automated DNA Sequencer (Applied Biosystems, Foster City, CA, USA) together with the internal size standard GenScan LIZ 500.

Amplicon size analyses were done using GeneMapper v4.3 software (Applied Biosystems, Foster City, CA, USA). Each product length (rounded to the nearest base pair) was considered a different allele. Every allele was assigned an alphabetical designation. Alleles for each isolate provided the multilocus genotype (MLG). Isolates with the same MLG were considered to be clones. At this stage, data generated by Cortinas et al. (2010) were combined with data from the present study. This was possible because both data sets had been generated using an ABI PRISM™ Automated DNA Sequencer (Applied Biosystems, Foster City, CA, USA) and GENEMAPPER (Applied Biosystems).

Fresh DNA was extracted from 25 isolates used by Cortinas et al. (2010) and re-analysed. This allowed confirmation that it was possible to combine data from that study, with those generated in the current study. The data were subjected to further analyses after several rounds of quality control. Whether there were sufficient numbers of isolates and loci to support statistical significance was assessed by plotting mean genotypic diversity against the number of loci using MULTILOCUS v1.3 (Agapow and Burt 2001).

Locus neutrality

The Ewens-Watterson's test in Popgene v3.2 (Yeh et al. 1999) was used to consider the selective neutrality of loci. This test is used to assess the selective neutrality of loci, under the infinite allele model (Ewens 1972; Watterson 1978). For all loci, F-values were computed between the standardized range of U95 and L95, at a 95 % confidence level.

Population structure

The Bayesian clustering algorithm in STRUCTURE v2.2 (Pritchard et al. 2000; Falush et al. 2003) was used to assess structure in the *T. zuluensis* populations. Individuals were assigned into 'K' clusters based on allelic frequencies. The method excludes prior information regarding location. The analyses were tested in two stages. The first was done to determine the optimal K using an admixture ancestry model and an independent allele frequency model. According to the Admixture model, individual *T. zuluensis* isolates can have mixed ancestry, i.e. isolate *i* inherited a certain percentage of its genome from ancestors in population *K* (Pritchard et al. 2000). The independent allele frequency model requires that allele frequencies in different populations should be reasonably different from each other (Pritchard et al. 2000). The runs were set at 20 for $K = 1$–8. The algorithm was executed at 300,000 Markov Chain Monte Carlo (MCMC) repetitions following a burn-in of 30 000 iterations. The most likely value for K was estimated using Evanno's ΔK method (Evanno et al. 2005) and the log likelihood of values of *K*, *L*(*K*) (Pritchard et al. 2000) using STRUCTURE HARVESTER (Earl and vonHoldt 2012).

The second stage of the analysis involved assigning individual *T. zuluensis* isolates to the optimal K obtained in the first stage. This was done at 1,000,000

MCMC repetitions following a burn-in of 100 000 iterations. Results from 20 runs for each K were visualized in CLUMPAK (Kopelman et al. 2015). Each isolate was assigned to a given cluster when the proportion of its genome in that cluster (q_K) was higher than an arbitrary cut-off value of 80 % (Liu et al. 2003; Fukunaga et al. 2005; Isenegger et al. 2008; Vigouroux et al. 2008). Isolates that could not meet the arbitrary cut-off were assigned to the 'admixed' group. Principal component analysis (PCA) was conducted using XLSTAT (Microsoft Corporation). Pearson's correlation coefficient was used to perform multidimensional scaling.

Analysis of molecular variance

An analysis of molecular variance (AMOVA) test was run to determine hierarchical partitioning of molecular variation among groups using Arlequin 3.11 (Excoffier et al. 2005). Groups were based on STRUCTURE (Pritchard et al. 2000; Falush et al. 2003) outputs and geographical origin of isolates. Three statistics of interest, F_{ST}, F_{SC} and F_{CT} were computed. F_{ST} is the amount of variation contained in a sub-population, relative to the total. F_{SC} measures the total genetic variation present in sub-populations between groups whilst F_{CT} describes the variation found within groups relative to the total variation. Population subdivision is, therefore, measured within populations (F_{ST}), among populations within groups (F_{SC}), and among groups (F_{CT}).

Population differentiation

Relatedness among *T. zuluensis* populations was assessed using a modified version of Wright's F_{st} for haploids, called theta (θ_{st}) (Weir 1996), in MULTILOCUS v1.3 (Agapow and Burt 2001). Theta was computed using the equation

$$\theta_{st} = Q - q/1 - q,$$

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. Q and q were summed across the evaluated loci for multiple loci. To test whether observations deviated significantly from the hypothesis of no linkage disequilibrium among loci, the observed value was compared with results of 1000 randomised data sets. Genetic distances (Nei 1973) among populations were estimated using POPGENE v3.2 (Yeh et al. 1999).

Genetic diversity

Measures of genetic diversity, including allelic distribution, polymorphism, average and effective numbers of alleles, private alleles and Nei's gene diversity were computed using POPGENE v3.2 (Yeh et al. 1999). Nei's gene diversity was based on the equation

$$H = 1 - \sum x_k^2,$$

where x_k is the kth allele frequency (Nei 1973).

Random mating

The Index of Association (I_A), a statistical tool in MULTILOCUS v1.3 (Agapow and Burt 2001), was used to assess the reproductive modes of the clone-corrected populations. I_A is a function of the variance in the number of heterozygous loci found in pair-wise comparisons of all haploid individuals in a population (Milgroom 1996). This tests the probability that two isolates sharing the same allele at a locus will have a common allele at another locus. I_A values were computed through a comparison of observed values and expected values of 10,000 randomly mating datasets (Taylor et al. 1999). The likelihood of each *T. zuluensis* population undergoing recombination was ascertained when the observed values were within the distribution range of the recombined values. Conversely, a population was considered to be clonally propagated when the observed values were outside the distribution range of the recombined values.

Results

Populations and isolate identification

Bark samples, showing lesions typical of Teratosphaeria stem canker, were collected from a total of 127 trees from Malawi (49), Mozambique (60) and Uganda (18). One to five fungal isolates were obtained from each tree, giving a total of 182 resembling those belonging to the *Teratosphaeriaceae*. BLAST results of the ITS sequences identified all the isolates as *T.*

zuluensis. Of the 182 isolates, 79 were from Malawi, 85 from Mozambique and 18 from Uganda (Table 1). An additional 56 isolates confirmed as *T. zuluensis* from Zambia (Chungu et al. 2010), were incorporated into this study. Microsatellite data for 96 *T. zuluensis* isolates from South Africa (Cortinas et al. 2010) were also included for reference purposes.

Locus neutrality

A total of eight polymorphic loci were amplified for all the isolates used in this study. All loci were neutral according to the Ewens–Wetteurson test (results not shown). The loci had F-values between the standardized range of U95 and L95 at 95 % confidence level.

Population structure

Population structure analysis inferred using a cluster–based method in STRUCTURE, revealed the existence of two genetic clusters ($K = 2$) based on Evanno's ΔK method (Fig. 1a) and five ($K = 5$) clusters based on the likelihood $L(K)$ method (Fig. 1b). Evanno's method showed the highest peak of ΔK to correspond to $K = 2$, with the next largest ΔK peak corresponding to $K = 5$. For the $L(K)$ method, the likelihood increased from $K = 1$ until $K = 5$ after which it decreased rapidly. The K value with the highest likelihood ($K = 5$) was chosen to represent the number of clusters. We chose to consider $K = 2$ only because no structure was evident in the populations at $K = 5$ (Fig. 2).

With the arbitrary cut-off value of 80 % ancestry for assignment at $K = 2$, 151 (of 192) *T. zuluensis* MLGs (79 %) were assigned to either of the two clusters. We refer to the two clusters as the Zambezi-Basin and Zululand clusters because they were dominated by MLGs from these regions. The Zululand cluster consisted of 73 MLGs whilst the Zambezi-Basin cluster had 78. The Zambezi-Basin cluster was dominated by MLGs from Gurue (15 of 17 MLGs; $q_K = 83$–99 %), Mangochi (all eight MLGs; $q_K = 84$–99 %), Manica (23 of 27 MLGs; $q_K = 81$–99 %), Mulanje (5 of 20 MLGs; $q_K = 84$–97 %), Thyolo (all nine MLGs; $q_K = 88$–99 %), South Africa (7 of 96 MLGs; $q_K = 81$–94 %) and Zambia (11 of 14 MLGs; $q_K = 83$–99 %). The Zululand cluster was dominated by MLGs from South Africa (65 of 96; $q_K = 80$–99 %) with very few from Manica (two MLGs; $q_K = 83$ & 86 %) and Mulanje (6 MLGs; $q_K = 83$–95 %). A relatively smaller number of MLGs (41) appeared to have ancestry from both clusters, having q_K values of less than 80 % for either cluster. Multilocus genotypes with mixed-ancestry were from Gurue (two MLGs), Manica (two MLGs), Mulanje (nine MLGs), South Africa (24 MLGs), Uganda (one MLG) and Zambia (three MLGs). The two clusters and the admixed group obtained through STRUCTURE were similarly observed in PCA (Fig. 3), in which two major components explained 37.5 % of the variability.

Analysis of molecular variance

Analysis of molecular variance (AMOVA) showed that most of the variance in the *T. zuluensis* collection could be attributed to within population variation (80 % total variation, $F_{ST} = 0.20$, $P = 0.001$) (Table 2). The remaining variation was attributable to among population variation within clusters (7.14 % total variation, $F_{SC} = 0.08$, $P = 0.001$) and among populations (12.9 % total variation, $F_{CT} = 0.13$, $P = 0.001$).

Population differentiation

Pairwise comparisons of *T. zuluensis* populations showed that the Zambezi-Basin cluster was significantly different ($\theta_{st} = 0.21$; $P = 0.001$) to the Zululand cluster. The admixed group was significantly different to both the Zambezi-Basin ($\theta_{st} = 0.09$; $P = 0.001$) and the Zululand ($\theta_{st} = 0.08$; $P = 0.001$) clusters. The same trends shown by θ_{st} were revealed through pairwise comparisons of genetic distances (D). A low genetic distance ($D = 0.11$) was obtained between the admixed group and both the Zambezi-Basin and the Zululand clusters. A higher genetic distance ($D = 0.34$) was obtained between the Zambezi-Basin and the Zululand clusters.

Genetic diversity

A total of 49 alleles were recovered from eight polymorphic microsatellite loci amplified in all isolates (Tables 3, 4). The distribution of alleles and the proportions of polymorphic loci were not significantly different among the Zululand and Zambezi-Basin clusters or the admixed group. Similarly, the average (Na) and effective (Ne) numbers of alleles did not vary among the Zambezi-Basin (Na = 4.25, 1.49 SD; $Ne = 2.39$,

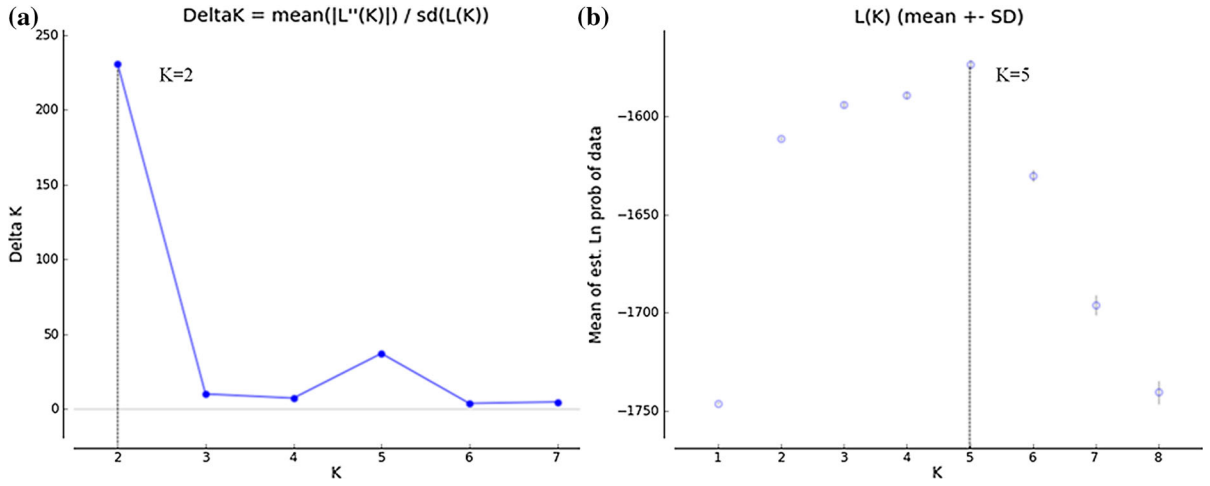

Fig. 1 Computation of the number of clusters (K) in *T. zuluensis* populations. **a** K = 2 had the highest ΔK peak height (Evanno et al. 2005), **b** mean posterior probabilities of eight runs for each K, K = 1 to K = 8 (Pritchard et al. 2000). The value with the highest *L* (*K*) corresponded to K = 5

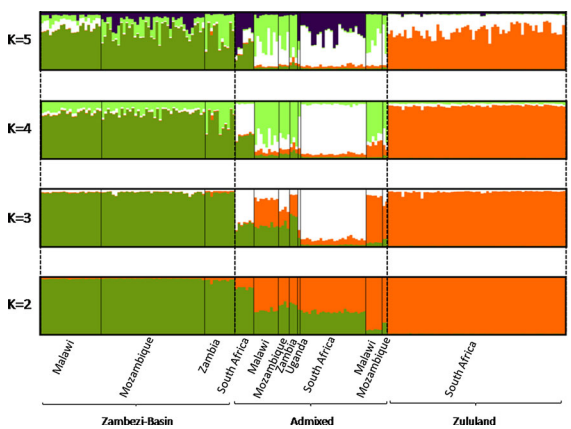

Fig. 2 Population structure of *T. zuluensis*. Each individual is represented by a *thin vertical line*, which is partitioned into K segments that represent the proportion of each cluster. *Black lines* separate sites from which individuals were collected

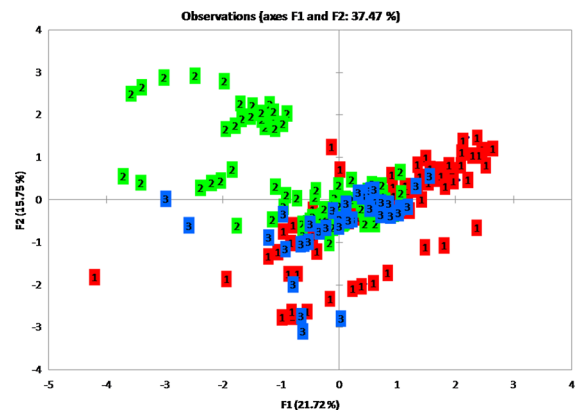

Fig. 3 Principal component analysis (PCA) of the 191 haplotypes representing STRUCTURE defined clusters: the Zambezi-Basin (*red*) and Zululand (*green*) clusters as well as the admixed group (*blue*)

1.12 SD), Zululand (*Na* = 4.00, 1.31 SD; *Ne* = 2.30, 0.84 SD) and the admixed groups (*Na* = 4.25, 1.58SD; *Ne* = 2.24, 0.81 SD). Nei's gene diversity was higher in the Zululand cluster (*h* = 0.52, 0.13 SD) compared to the Zambezi-Basin cluster (*h* = 0.48, 0.25 SD) and the admixed group (*h* = 0.47, 0.27 SD).

A total of 192 MLGs were recovered from all isolates. The MLGs were approximately equally distributed between the Zambezi-Basin cluster (73) and the Zululand cluster (78) with the admixed group (41) containing the remainder of the MLGs. Some MLGs were shared by at least two populations within each cluster. The most common MLG was shared by isolates in four populations belonging to the Zambezi-Basin cluster i.e. Gurue, Manica, Thyolo and Zambia. Three MLGs were shared by some sites in the same cluster i.e. Gurue and Manica, Gurue and South Africa as well as Mangochi and Manica. In the Zululand cluster, one MLG was shared between Manica and Mulanje. The admixed group had two MLGs shared between Mulanje and South Africa as well as Mulanje and Zambia. Single MLGs were obtained in all cases where multiple isolates were collected per individual tree.

Table 2 Hierarchical partitioning of the molecular variation for the *T. zuluensis* dataset

Source of variation	d.f.	Sum of squares	Variance components	Percentage of variation
Among clusters	2	51.33	0.31 Va	12.87
Among populations within clusters	12	40.20	0.17 Vb	7.14
Within populations	176	335.62	1.91 Vc	80.0
Total	190	427.14	2.38	

Fixation indices: FSC = 0.08197 (P < 0.001); FST = 0.20009 (P < 0.001); FCT = 0.12866 (P < 0.001)

Table 3 Allele frequencies of the Zambezi-Basin, Zululand and the Admixed clusters recovered from eight microsatellite markers

Locus	Allele size	Zambezi-Basin cluster	Zululand cluster	Admixed cluster	Locus	Allele size	Zambezi-Basin cluster	Zululand cluster	Admixed cluster
Czulu1	150	0.1111	0.1644	0.2051	*Kzulu10*	305[a]	0.1111		
	152	0.7917	0.3836	0.6667		326		0.2055	0.0256
	154	0.0417	0.3425	0.0769		332[a]		0.0137	
	156	0.0139	0.0959	0.0256		335	0.6250	0.2329	0.4359
	158	0.0417		0.0256		346		0.0685	0.0256
	160[a]		0.0137			365		0.0959	0.0256
Czulu2	173[a]			0.0256		415	0.1528	0.3836	0.4359
	178	0.1944		0.0769		425	0.1111		0.0513
	186	0.6806	0.3836	0.5385	*Kzulu12*	273	0.0139		0.0256
	188	0.1250	0.6164	0.3590		275[a]		0.1644	
Czulu3	156[a]		0.2740			296[a]	0.1528		
	169	0.0139	0.0548	0.0256		298	0.8333	0.7534	0.9744
	170	0.9861	0.6712	0.9744		300[a]		0.0685	
Czulu6	310[a]		0.0274			304[a]		0.0137	
	321[a]	0.2361			*Kzulu13*	123	0.1944	0.0274	0.2051
	324	0.1528	0.1781	0.2821		127	0.0139	0.5753	0.2821
	330	0.2639	0.7671	0.4872		129[a]			0.0256
	333	0.3056		0.0513		131	0.0417		0.0256
	336	0.0278	0.0274	0.1795		137	0.2083	0.3425	0.4359
	339[a]	0.0139				141[a]		0.0274	
Czulu7	213	0.3472	0.0548	0.1795		149[a]		0.0274	
	219	0.0417		0.0513		151[a]	0.4583		
	221	0.3472	0.6575	0.5385		153[a]	0.0833		
	223	0.2083	0.2877	0.1795		155[a]			0.0256
	237	0.0556		0.0513					

[a] Private allele

Random mating

The observed I_A values for the Zambezi-Basin cluster (0.19), the Zululand cluster (0.06) and the admixed group (0.12) were within the distributions of I_A values for the 10 000 randomised data sets (Fig. 4). The observed I_A value (0.18) for the combined data set comprising of the two clusters and the admixed group, coincided with the highest value of the randomised data set. However, the I_A values for the Zululand cluster and the admixed group were not significant (P = 0.05). Thus, the I_A results support the hypothesis that recombination is occurring in *T. zuluensis* populations.

Table 4 Genetic diversity parameters for STRUCTURE identified clusters and the sites within clusters

Cluster	Country	Site	NNC[a]	NCC[b]	MLGs[c]	ML[d]	PL[e]	Na[f]	Np[g]	H[h]
Zambezi-Basin	Malawi	MN	23	12	8	2	6	17		0.34
		ML	8	5	5	2	6	19		0.39
		TY	21	11	9	3	5	17		0.32
	Mozambique	GR	39	24	15	1	7	24		0.41
		MC	39	29	23	1	7	25		0.45
	South Africa		–	–	7	0	8	19		0.44
	Zambia		47	20	11	2	6	20		0.40
	Total			126	73	0	8	34	6	0.49
Zululand	Malawi	ML	10	10	6	2	6	14		0.32
	Mozambique	MC	2	2	2	3	5	13		0.31
	South Africa		–	–	65	0	8	23		0.52
	Total			78	78	0	8	32	9	0.52
Admixed Group	Malawi	ML	17	11	9	2	6	19		0.39
	Mozambique	GR	3	3	2	4	4	12		0.25
		MC	2	2	2	4	4	12		0.25
	South Africa		–	25	24	0	8	27		0.44
	Uganda		18	18	1	8	0	8		–
	Zambia		9	5	3	4	4	13		0.25
	Total			46	41	0	8	34	3	0.47
Total				249	184	0	8	49		0.56

GR Gurue, *MC* Manica, *ML* Mulanje, *MN* Mangochi, *TY* Thyolo

[a] Number of non-clone corrected isolates

[b] Number of clone corrected isolates

[c] Multilocus genotypes

[d] Monomorphic loci

[e] Polymorphic loci

[f] Number of alleles

[g] Number of private alleles

[h] Gene diversity

Discussion

Analyses of population structure and differentiation on isolates of *T. zuluensis* from southern and eastern Africa showed that the populations could be grouped into three genetic clusters, with an origin likely to be outside the continent. This would be consistent with the results of previous studies by Cortinas et al. (2010) and Chen et al. (2011). The three clusters that we have designated the Zululand, Zambezi-Basin and an admixed group, were characterised by moderate to low genetic diversities, which is consistent with expectation for introduced pathogen populations (McDonald 1997; Stukenbrock et al. 2007). The origin of *T. zuluensis* seems likely to be Australia, where most *Eucalyptus* spp. planted outside their native range have originated, even though the pathogen has not been found there. This would be similar to the closely related *Teratosphaeria nubilosa*, a eucalypt leaf pathogen shown to consist of three genetic groups and low to moderate genetic diversity in areas outside its native range (Pérez et al. 2010, 2012). Alternatively, the centre of origin of the pathogen could be south-east Asia, based on very high genetic diversities characteristic of populations in China (Chen et al. 2011). In this situation, the origin would likely be from a native species of Myrtaceae and resulting from host shifts, which are increasingly commonly being found for *Eucalyptus* pathogens (e.g. Slippers et al. 2005; Graça et al. 2013; van der Merwe et al. 2013).

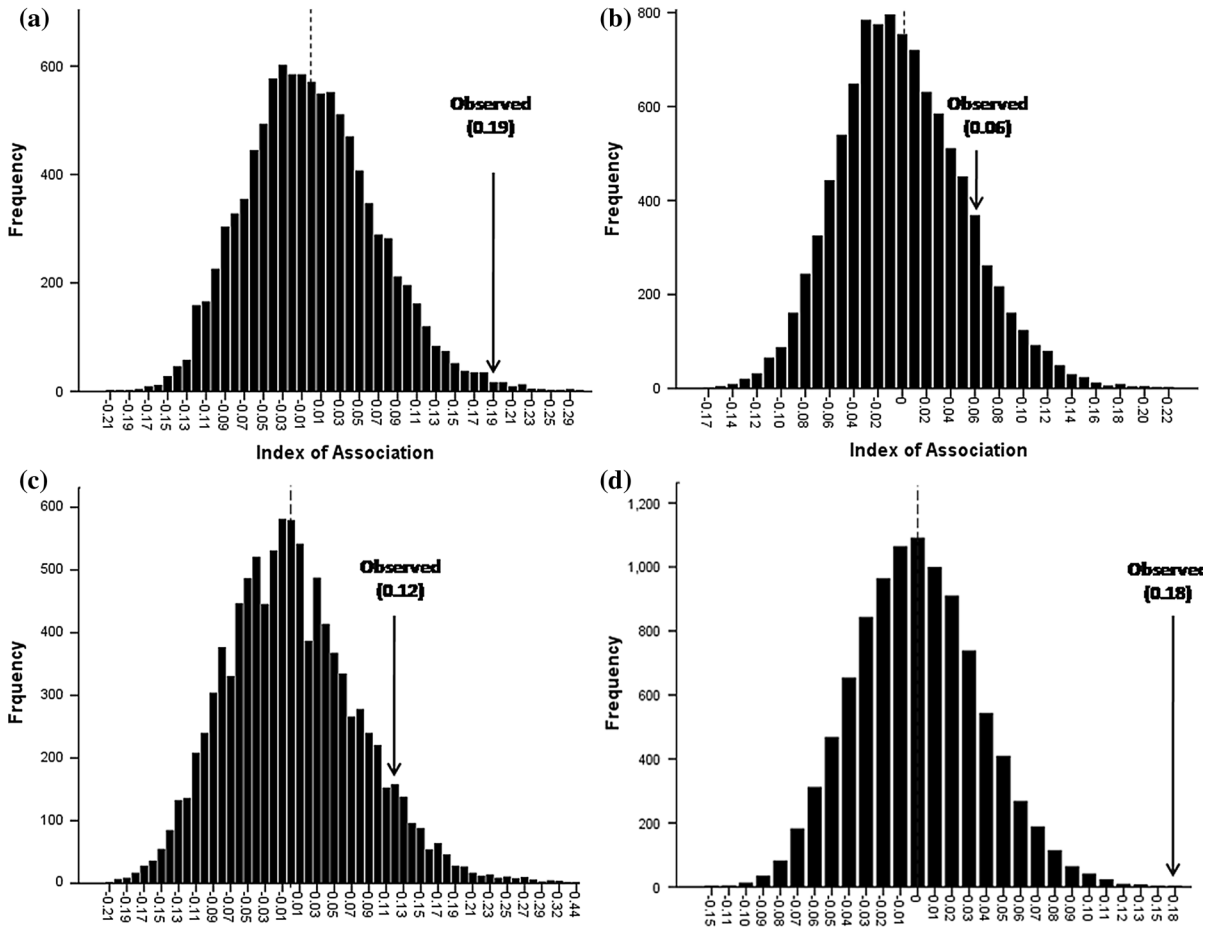

Fig. 4 Position of the observed I_A relative to I_A values for 10,000 artificially recombined data sets whose mean is 0 and is represented by the *dashed line*. **a** Zambezi-Basin cluster, **b** Zululand cluster, **c** the admixed group and **d** all the three data sets combined

The occurrence of common genotypes for isolates of *T. zuluensis* within the Zambezi-Basin countries (Malawi, Mozambique and Zambia) as well as between the Zululand cluster (South Africa) and the admixed group dominated by Mulanje (Malawi) isolates provides evidence that the pathogen has moved on the continent. The spread of genotypes could be attributed to the post-introduction exchange of infected germplasm (particularly seed and cuttings), trade in timber products (Wingfield et al. 2008, 2015) and wind/water dispersal of spores within more limited areas. A recent study by Jimu et al. (2015b) showed the occurrence of *T. zuluensis* in *E. grandis* seed, suggesting that exchange in seed is playing an important role in the spread of the pathogen. This would also be consistent with the history of *Eucalyptus* seed production and trade in Africa. For many decades, South Africa has been the source of *Eucalyptus* seed for Mulanje and parts of central, eastern and southern Africa. Indeed, some tea companies in the Mulanje area have confirmed that *Eucalyptus* seed was introduced from South Africa to establish plantations used for drying tea leaves. Thus, the high levels of admixture observed in the Mulanje population could have been a result of the introduction of genotypes from the Zululand cluster, followed by recombination, leading to high genetic diversity. A similar scenario has previously been suggested by Cortinas et al. (2010).

The existence of a clonal population of *T. zuluensis* from a single plantation in Uganda suggests a very recent introduction into that area. This is supported by the recent report of the pathogen (Jimu et al. 2014) in a country where only *T. gauchensis* was previously

known to cause Teratosphaeria stem canker on *Eucalyptus* (Roux et al. 2005). The clonal nature of the population is consistent with a single introduction of a naturally selected genotype (Ficetola et al. 2008; Stukenbrock and McDonald 2008) from an unknown source population. Although the source population is not known, population structure and allelic distribution results from this study strongly suggest that *T. zuluensis* could have moved from the Southern African Development Community (SADC) region to Uganda. Given the history of *Eucalyptus* seed importation into Uganda, it is probable that *T. zuluensis* could have been introduced from South Africa. Indeed, a forestry company in the Jinja area has confirmed that the origin of the seed that was used to regenerate the compartment where *T. zuluensis* was first reported came from South Africa.

The fact that the sexual state of *T. zuluensis* has never been observed in nature (Wingfield et al. 1997) does not preclude its existence. The possible existence of the sexual state in *T. zuluensis* is strongly supported by the moderate genetic diversities recorded in this study as well as high diversities that have previously been recorded (Cortinas et al. 2010; Chen et al. 2011). Furthermore, I_A tests conducted in this study as well as previous studies (Cortinas et al. 2010; Chen et al. 2011) suggest random mating in some of the populations. The existence of the sexual state in *T. zuluensis* would be consistent with other related *Eucalyptus* pathogens where recombination is thought to have contributed to high genetic diversities (Pérez et al. 2010, 2012).

Conclusions

T. zuluensis is becoming an increasingly important pathogen of *Eucalyptus* in Africa. Previous studies on the pathogen have had a continental perspective. In contrast, this study has considered *T. zuluensis* from a regional perspective, by including populations from several countries in Africa, among them, previously un-sampled areas such as Malawi, Mozambique and Uganda. Results revealed two fronts of invasion for *T. zuluensis*, both from unknown sources. This was coupled with high levels of admixture, probably as a result of recombination and exchange of seed within the continent.

The clonal nature of the Ugandan population suggests a very recent introduction from the SADC region. This underpins the need for great care to be taken by African countries when establishing new plantation areas and thus to avoid the spread of one of the most important pathogens of *Eucalyptus* .

Acknowledgments This work is based on research supported by the National Research Foundation of South Africa (Grant specific unique reference number (UID83924). The grant holders acknowledge that opinions, findings and conclusions or recommendations expressed in any publication generated by the NRF supported research are that of the authors and that the NRF accepts no liability what so ever in this regard. We thank the members of the Tree Protection Co-operative Program (TPCP), the THRIP initiative of the Department of Trade and Industry and the DST/NRF Centre of Excellence in Tree Health Biotechnology (CTHB) of the Forestry and Agricultural Biotechnology Institute (FABI), University of Pretoria, South Africa for financial support. We further thank forestry companies and private farmers in Malawi, Mozambique and Uganda for permission to collect samples in their plantations; the Forestry Research Institute of Malawi (FRIM) and the Sawlog Production Grant Scheme (SPGS) of Uganda for assisting with sample collection.

References

Agapow PM, Burt A (2001) Indices of multilocus linkage disequilibrium. Mol Ecol Notes 1:101–102

Chamshama SAO, Nwonwu FOC (2004) Forest plantations in sub-saharan Africa. A report prepared for the project: lessons learnt on sustainable forest management in Africa. Report, KSLA, AFORNET, FAO

Chen SF, Barnes I, Chungu D, Roux J, Wingfield MJ, Xie YJ, Zhou XD (2011) High population diversity and increasing importance of the *Eucalyptus* stem canker pathogen, *Teratosphaeria zuluensis*, in South China. Australas Plant Pathol 40:407–415

Chungu D, Muimba-Kankolongo A, Wingfield MJ, Roux J (2010) Identification of fungal pathogens occurring in eucalypt and pine plantations in Zambia by comparing DNA sequences. Forestry 83:507–515

Cortinas MN, Barnes I, Wingfield BD, Wingfield MJ (2006a) Polymorphic microsatellite markers for the *Eucalyptus* fungal pathogen *Colletogloeopsis zuluensis*. Mol Ecol Notes 6:780–783

Cortinas MN, Burgess T, Dell B, Xu D, Crous PW, Wingfield BD, Wingfield MJ (2006b) First record of *Colletogloeopsis zuluense* comb.nov., causing a stem canker of *Eucalyptus* in China. Mycol Res 110:229–236

Cortinas MN, Barnes I, Wingfield MJ, Wingfield BD (2010) Genetic diversity in the *Eucalyptus* stem pathogen *Teratosphaeria zuluensis*. Australas Plant Pathol 39:383–393

Cortinas MN, Crous PW, Wingfield BD, Wingfield MJ (2006b) Multigene phylogenies and phenotypic characters distinguish two species within the *Colletogloeopsis zuluensis* complex associated with *Eucalyptus* stem cankers. Stud Mycol 55:133–146

Cortinas MN, Koch N, Thain J, Wingfield BD, Wingfield MJ (2004) First record of the *Eucalyptus* stem canker pathogen, *Coniothyrium zuluense* from Hawaii. Australas Plant Pathol 33:309–312

Crous PW, Groenewald JZ, Summerel BA, Wingfield BD, Wingfield MJ (2009a) Co-occurring species of *Teratosphaeria* on *Eucalyptus*. Persoonia 22:38–48

Crous PW, Summerell BA, Carnegie AJ, Wingfield MJ, Groenewald JZ (2009b) Novel species of *Mycosphaerellaceae* and *Teratosphaeriaceae*. Persoonia 23:119–146

Earl D, vonHoldt B (2012) STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. Conserv Genet Resour 4:359–361

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

Ewens WJ (1972) The sampling theory of selectively neutral alleles. Theor Popul Biol 3:87–112

Excoffier L, Laval G, Schneider S (2005) Arlequin ver. 3.0: an integrated software package for population genetics data analysis. Evol Bioinf Online 1:47–50

Falush D, Stephens M, Pritchard JK (2003) Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. Genetics 164:1567–1587

FAO (2011) *Eucalyptus* in East Africa: Socio-economic and environmental issues. Working Paper FP46/EFAO, Rome

Ficetola GF, Bonin A, Miaud C (2008) Population genetics reveals origin and number of founders in a biological invasion. Mol Ecol 17:773–782

Fukunaga K, Hill J, Vigouroux Y, Matsuoka Y, Sánchez GJ, Liu K, Buckler ES, Doebley J (2005) Genetic diversity and population structure of teosinte. Genetics 169:2241–2254

Gezahgne A, Cortinas MN, Wingfield MJ, Roux J (2005) Characterisation of the Coniothyrium stem canker pathogen on *Eucalyptus camaldulensis* in Ethiopia. Australas Plant Pathol 34:85–90

Gezahgne A, Roux J, Wingfield MJ (2003) Diseases of exotic *Eucalyptus* and *Pinus* species in Ethiopian plantations. S Afr J Sci 99:29–33

Graça RN, Ross-Davis AL, Klopfenstein N, Kim M-S, Peever TL, Cannon PG, Aun CP, Mizubuti ESG, Alfenas AC (2013) Rust disease of eucalypts, caused by *Puccinia psidii*, did not originate via host jump from guava in Brazil. Mol Ecol 22:6033–6047

Isenegger DA, Macleod WJ, Ford R, Taylor PWJ (2008) Genotypic diversity and migration of clonal lineages of *Botrytis cinerea* from chickpea fields of Bangladesh inferred by microsatellite markers. Plant Pathol 57:967–973

Jimu L, Kemler M, Wingfield MJ, Mwenje E, Roux J (2015b) The *Eucalyptus* stem canker pathogen *Teratosphaeria zuluensis* detected in seed samples. Forestry. doi:10.1093/forestry/cpv037

Jimu L, Wingfield MJ, Mwenje E, Roux J (2014) First report of *Teratosphaeria zuluensis* causing stem canker of *Eucalyptus grandis* in Uganda. For Pathol 44:242–245

Jimu L, Wingfield MJ, Mwenje E, Roux J (2015) Diseases of *Eucalyptus* species in Zimbabwean plantations and woodlots. South For. doi:10.2989/20702620.2014.1001682

Kopelman NM, Mayzel J, Jakobsson M, Rosenberg NA, Mayrose I (2015) CLUMPAK: a program for identifying clustering modes and packaging population structure inferences across K. Mol Ecol Resour. doi:10.1111/1755-0998.12387

Liu J, Goodman M, Muse S, Smith JS, Buckler E, Doebley J (2003) Genetic structure and diversity among maize inbred lines as inferred from DNA microsatellites. Genetics 165:2117–2128

McDonald BA (1997) The population genetics of fungi: tools and techniques. Phytopathology 87:448–453

Milgroom MG (1996) Recombination and the multilocus structure of fungal populations. Annu Rev Phytopathol 34:457–477

Möller EM, Bahnweg G, Sandermann H, Geiger HH (1992) A simple and efficient protocol for isolation of high molecular weight DNA from filamentous fungi, fruit bodies, and infected plant tissues. Nucleic Acids Res 20:6115–6116

Nei M (1973) Analysis of gene diversity in subdivided populations. Proc Natl Acad Sci USA 70:3321–3323

Old KM, Wingfield MJ, Yuan ZQ (2003) A manual of diseases of eucalypts in South-East Asia. CIFOR, Jakarta

Pérez G, Slippers B, Wingfield BD, Hunter GC, Wingfield MJ (2010) Micro- and macro spatial scale analyses illustrates mixed mating strategies and extensive gene flow in populations of an invasive haploid pathogen. Mol Ecol 19:1801–1813

Pérez G, Slippers B, Wingfield MJ, Wingfield BD, Carnegie AJ, Burgess TI (2012) Cryptic species, native populations and biological invasions by a eucalypt forest pathogen. Mol Ecol 21:4452–4471

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. Genetics 155:945–959

Rejmánek M, Richardson DM (2011) Eucalypts. In: Simberloff Daniel, Rejmánek Marcel (eds) *Encyclopedia* of biological invasions. University of California Press, Berkeley

Roux J, Meke G, Kanyi B, Mwangi L, Mbaga A, Hunter GC, Nakabonge G, Heath RN, Wingfield MJ (2005) Diseases of plantation forestry tree species in eastern and southern Africa. S Afr J Sci 101:409–413

Roux J, Wingfield MJ, Cibrian D (2002) First report of Coniothyrium canker of *Eucalyptus* in Mexico. Plant Pathol 51:382

Silva MRC, Diogo E, Bragança H, Machado H, Phillips AJL (2014) *Teratosphaeria gauchensis* associated with trunk, stem and foliar lesions of *Eucalyptus globulus* in Portugal. For Pathol. doi:10.1111/efp.12160

Slippers B, Stenlid J, Wingfield MJ (2005) Emerging pathogens: fungal host jumps following anthropogenic introduction. Trends Ecol Evol 20:420–421

Stukenbrock EH, Banke S, Javan-Nikkhah M, McDonald BA (2007) Origin and domestication of the fungal wheat pathogen *Mycosphaerella graminicola* via sympatric speciation. Mol Biol Evol 24:398–411

Stukenbrock EH, McDonald BA (2008) The origins of plant pathogens in agro-ecosystems. Annu Rev Phytopathol 46:75–100

Taylor JW, Jacobson DJ, Fisher MC (1999) The evolution of asexual fungi: reproduction, speciation and classification. Annu Rev Phytopathol 37:197–246

van der Merwe NA, Steenkamp ET, Rodas C, Wingfield BD, Wingfield MJ (2013) Host switching between native and non-native trees in a population of the canker pathogen *Chrysoporthe cubensis* from Colombia. Plant Pathol 62:642–648

van Zyl LM, Couthinho TA, Wingfield MJ, Pongpanich K, Wingfield BD (2002) Morphological and molecular relatedness of geographically diverse isolates of *Coniothyrium zuluense* from South Africa and Thailand. Mycol Res 106:51–59

Vigouroux Y, Glaubitz JC, Matsuoka Y, Goodman MM, Sánchez GJ, Doebley J (2008) Population structure and genetic diversity of new world maize races assessed by DNA microsatellites. Am J Bot 95:1240–1253

Watterson GA (1978) The homozygosity test of neutrality. Genetics 88:405–417

Weir BS (1996) Genetic data analysis II. Sinauer Associates, Sunderland

White TJ, Bruns T, Taylor J (1990) Amplification and direct sequencing of fungal ribosomal RNA genes for phylognetics. In: Innis MA, Gelfand DH, Sninsky JJ, White JW (eds) A guide to molecular methods and applications. Academic Press, New York, pp 315–322

Wingfield MJ, Crous PW, Coutinho TA (1997) A serious canker disease of *Eucalyptus* in South Africa caused by a new species of *Coniothyrium*. Mycopathologia 136:139–145

Wingfield MJ, Slippers B, Hurley BP, Coutinho TA, Wingfield BD, Roux J (2008) Eucalypt pests and diseases: growing threats to plantation productivity. South For 70:139–144

Wingfield MJ, Brockerhoff EG, Wingfield BD, Slippers B (2015) Planted forest health: the need for a global strategy. Science 349:832–836

Yeh FC, Yang RC, Boyle T (1999) POPGENE. Version 1.31. Microsoft Windows based freeware for population genetic analysis. University of Alberta: Alberta