Received: 29 October 2023 | Accepted: 20 January 2024

DOI: 10.1111/ppa.13876

Plant Pathology WILEY

ORIGINAL ARTICLE

Cryptic sexual reproduction in an emerging *Eucalyptus* shoot and foliar pathogen

Nam Q. Pham[1] | Hiroyuki Suzuki[2,3] | Tuan A. Duong[2] | Brenda D. Wingfield[2] | Irene Barnes[2] | Alvaro Durán[4] | Michael J. Wingfield[1]

[1]Department of Plant and Soil Sciences, Forestry and Agricultural Biotechnology Institute, University of Pretoria, Pretoria, South Africa

[2]Department of Biochemistry, Genetics and Microbiology, Forestry and Agricultural Biotechnology Institute, University of Pretoria, Pretoria, South Africa

[3]Faculty of Agro-Food Science, Niigata Agro-Food University, Niigata, Japan

[4]Plant Health Program, Research and Development, Asia Pacific Resources International Holdings Ltd, Riau, Indonesia

Correspondence
Nam Q. Pham, Department of Plant and Soil Sciences, Forestry and Agricultural Biotechnology Institute (FABI), University of Pretoria, Pretoria 0028, South Africa.
Email: nam.pham@fabi.up.ac.za

Abstract

Eucalyptus scab and shoot malformation is an emerging disease and a serious threat to the global plantation forestry industry. The disease appeared in North Sumatra (Indonesia) in the early 2010s and the causal agent was recently described as a novel species, *Elsinoe necatrix*. Nothing is known regarding its possible origin or why it emerged rapidly to cause a serious local epidemic. To investigate its population biology, we developed 15 polymorphic microsatellite markers as well as mating-type markers using genome sequences for two *E. necatrix* isolates. Isolates of the pathogen were collected from different host varieties at four locations in the Lake Toba region of North Sumatra and characterized using these markers. A high level of genotypic diversity was observed for all populations with little to no genetic differentiation between sampling areas. Discriminant analysis of principal components, genotype networks and analysis of molecular variance all showed a lack of population structure and a high level of gene flow among sampling regions. Mating-type ratios and linkage disequilibrium analyses suggest that sexual recombination is likely to be occurring, although a sexual state has not been found for the pathogen. The results of this study highlight the fact that new genotypes of *E. necatrix*, probably arising from cryptic sexual recombination, will challenge efforts to manage the disease, and that breeding and selection for tolerance will require substantial host genetic diversity.

KEYWORDS
Elsinoe necatrix, fungal pathogen, plantation forestry, population genetics, scab disease, sexual recombination

1 | INTRODUCTION

Eucalyptus, native to Australia and some neighbouring islands, is widely cultivated in the tropics and Southern Hemisphere to produce solid timber, pulp and other products (Turnbull, 1999). These plantation forestry programmes, based on monocultures of fast-growing and high-yielding genotypes, have expanded rapidly during the course of the last few decades, covering more than 20 million ha worldwide (Iglesias-Trabado & Wilstermann, 2008). As has been the case for other crops grown as non-natives in new environments, there has been a gradual but persistent emergence of insect pests and pathogens, placing the sustainability of these

planted forests at risk (Burgess & Wingfield, 2017; Wingfield, 2003; Wingfield et al., 2008).

Eucalyptus scab and shoot malformation is an emerging *Eucalyptus* disease and a serious threat to the global forestry industry. The disease first appeared in North Sumatra, Indonesia, in the early 2010s and the causal agent was recently described as the novel species *Elsinoe necatrix* (Pham et al., 2021). This aggressive shoot and foliar pathogen causes black necrotic spots on young tissues that progress to scab-like wounds as the lesions age. Infected trees respond to infection by producing shoots with small leaves that commonly appear feathered, and susceptible genotypes exhibit severely retarded growth (Pham et al., 2021). Currently, *E. necatrix* is known only in North Sumatra, Indonesia.

The reproductive biology of pathogens plays a significant role in shaping their population diversity and structure. *E. necatrix* is a heterothallic fungus, a characteristic probably shared by all species in the genus (Pham et al., 2023). Thus, sexual reproduction in these fungi requires contact between two individuals of opposite mating type (Ni et al., 2011; Turgeon & Yoder, 2000). The presence of sexually compatible strains would provide the opportunity for the production of new genotypes, leading to a more genetically diverse population with a greater potential for more aggressive genotypes to evolve (Drenth et al., 2019; McDonald & Linde, 2002). Asexual spores of *E. necatrix* have been observed on infected plant tissues (Pham et al., 2021). However, despite careful examination of infected *Eucalyptus* tissues, the sexual state of the fungus has not been found.

E. necatrix poses a considerable threat to *Eucalyptus* plantation forestry in Indonesia and probably other countries into which it could be accidentally introduced. It is thus necessary to gain an understanding of its genetic diversity and population structure, which could provide clues to its possible origin as well as guidance in the development of disease management strategies. To achieve this goal, we developed and used microsatellite and mating-type markers to investigate a large collection of *E. necatrix* isolates from different regions of North Sumatra.

2 | MATERIALS AND METHODS

2.1 | Sampling, isolations, DNA extraction and identification of isolates

Sampling was conducted during June and July 2021 from epidemic outbreaks of *E. necatrix* in four different eucalypt planting regions in the Toba Lake region of North Sumatra. These included Aek Nauli (AEN), Aek Raja (AER), Habinsaran (HAB) and Tele (TEL) (Figure 1). The four regions covered a total area of approximately 40,000 ha of plantations, which consisted of multiple compartments of different eucalypt genotypes propagated from cuttings.

Symptomatic branchlets (approximately 15 cm long) were randomly collected from different compartments, placed in sealed plastic bags, and transferred to the laboratory for fungal isolation. The pathogen was isolated from necrotic lesions on young leaves, shoots or petioles as described by Pham et al. (2021), where each individual lesion resulted in a single fungal isolate. To prevent repeated isolation of the same individual, only one fungal isolate was retained per branchlet and up to three isolates were maintained per compartment. Wherever possible, samples were taken from different host genotypes or species so as to ensure a wide diversity of samples.

In addition to a broad sampling strategy across the different regions, we applied hierarchical sampling in a single compartment of *Eucalyptus pellita* in AEN to consider whether there was localized spread of genotypes. This compartment consisted of trial plantings comprising multiple provenances of *E. pellita* propagated from seed. A total of 15 branchlets were collected from 15 randomly selected trees of five *E. pellita* provenances each and isolations were made from three randomly collected scab lesions on each branchlet. The isolates collected from this single *E. pellita* compartment were compared with those obtained from different compartments in the same region to test if there was long-distance spread of genotypes. Details of the isolates obtained in this study and their hosts are provided in Table S1.

DNA was extracted from 10-day-old isolates grown on potato dextrose agar (PDA; BD Difco) at 25°C, using PrepMan Ultra Sample Preparation Reagent (Thermo Fisher Scientific) following the manufacturer's protocols. The identity of the isolates was confirmed by sequencing the rRNA internal transcribed spacer regions 1 and 2 (ITS), including the 5.8S rRNA region, as described by Pham et al. (2021).

2.2 | Microsatellite marker development and amplification

The genomes of two *E. necatrix* isolates, CMW 56129 (JANZYI000000000; Pham et al., 2023) and CMW 56134 (JANZYH000000000; Wingfield et al., 2022), were used to identify polymorphic microsatellite regions. Microsatellite repeats were mined from the genome of isolate CMW 56134 using Krait v. 1.3.3 (Du et al., 2018). The search for microsatellite regions was set to target tri-, tetra-, penta- and hexanucleotide repeat units, with at least 4 to 10 repeats, and contained the sequence 250 bp upstream and downstream flanking the microsatellite region.

The genome of *E. necatrix* isolate CMW 56134 was annotated using the fungal version of GeneMark-ES (Ter-Hovhannisyan et al., 2008). The annotation was then mapped to the search results in Krait to identify the regions not occurring within predicted coding regions. Primer pairs were designed for candidate regions using Primer3 (Untergasser et al., 2012), integrated within Krait. Parameters were customized to have an optimal annealing temperature of approximately 60°C and to have G or C present in the first two nucleotides at the 3′ end. Primers pairs were chosen that would produce fragment sizes ranging from 100 to 450 bp, and further selected based on their presence on different scaffolds of the genome. Those regions were then extracted and subjected to a local BLAST analysis using Geneious Prime 2023.0.3 (https://www.geneious.com) against

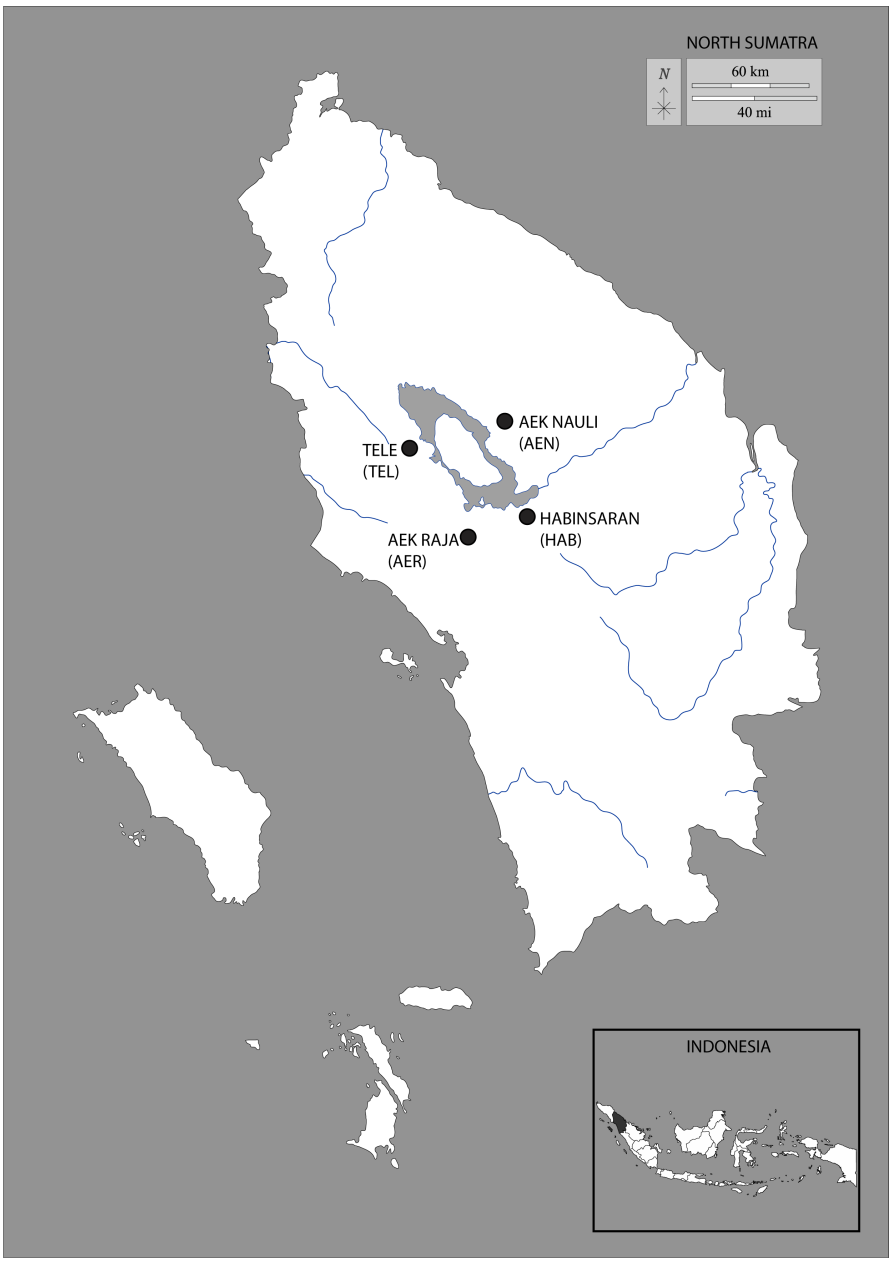

FIGURE 1 The four sampling locations of *Elsinoe necatrix* in North Sumatra, Indonesia. [Colour figure can be viewed at wileyonlinelibrary.com]

the genome of isolate CMW 56129 to identify polymorphisms in the microsatellite region. Sequences with multiple homologies to different locations in the genome were also eliminated from the pipeline to avoid nonspecific binding of the primers.

PCRs to test the amplification of selected primers were performed in a total volume of 13 µL containing 1 µL of genomic DNA, 0.25 µL of each primer (10 µM), 2.5 µL 5× MyTaq buffer (Bioline), 0.25 µL MyTaq DNA polymerase (Bioline) and 8.75 µL sterile deionized water. The thermal cycling included an initial denaturation at 95°C for 5 min; 10 primary amplification cycles of 30 s at 95°C, 30 s at 60°C, and 60 s at 72°C; then 30 additional cycles of the same reaction cycle, with a 5 s increase in the annealing step per cycle. Reactions were completed with a final extension at 72°C for 30 min. Amplicons obtained were sequenced to confirm that the primers amplified the correct loci.

For fragment analysis, the selected microsatellite markers were divided into two panels based on their expected amplicon sizes, and the forward primer of each set was labelled with a fluorescent dye (6-FAM, PET, VIC and NED; Thermo Fisher Scientific) (Table 1). Single-plex PCRs were performed as described above and 1 µL of each PCR product from the same panel was pooled and diluted with sterile deionized water to a total volume of 150 µL. Subsequently, 1 µL of pooled product was mixed with 0.14 µL GeneScan LIZ-500 size standard (Thermo Fisher Scientific) and 8.86 µL Hi-Di formamide (Thermo Fisher Scientific) and run on an ABI 3500xl Genetic Analyser (Thermo Fisher Scientific) at the Sequencing Facility of the Faculty of Natural and Agricultural Sciences, University of Pretoria. Genome-sequenced isolates (CMW 56129 and CMW 56134) were used as positive control samples for each run. Allele calls were determined using GeneMapper software v. 4.0 (Thermo Fisher Scientific).

TABLE 1 Polymorphic microsatellite markers developed for *Elsinoe necatrix*.

	Primer (5′–3′)					Alleles				
Locus	Forward	Reverse	Dye	Panel	Motif	Number	Observed alleles (bp)	GenBank accession	H_{exp}	E_5
MSEN01	TCTGAGGAATCCTGGAGGGG	GCTCGCATCTCTACCAGACC	VIC	1	GTG	2	200, 206	PP199424, PP199425	0.17	0.56
MSEN02	GTTGTGAGGTGCGTTGATGG	GATCGATACCACCCTTCCGG	PET	1	GGA	2	266, 269	PP199426, PP199427	0.47	0.93
MSEN03	ATCTGAAGAGCCACCTGTGC	GTCGTTGACGTTCACAGTGC	6-FAM	1	TGAG	2	296, 300	PP199428, PP199429	0.30	0.70
MSEN04	TGTTCAGCCAGCTACCTTCG	CAGGGACTCGATCATCGTGG	VIC	1	GCAA	2	330, 334	PP199430, PP199431	0.19	0.58
MSEN05	CGAACTCTATCAGGGGTGGC	GGATCCCCATCATGAGCTCG	PET	1	CTTTT	2	347, 357	PP199432, PP199433	0.39	0.81
MSEN06	CTTCGAGAATGCTTCGCTGC	AACCAGAGAGACGAACAGCC	6-FAM	1	ACA	2	356, 359	PP199434, PP199435	0.47	0.94
MSEN07	TTCGTCCAGGTGACTTGTCG	TAGTCGCGCTTCAAGGAGC	VIC	1	TCA	2	376, 379	PP199436, PP199437	0.28	0.68
MSEN08	TGTGTCGATTCGGTTGAGGG	GATCGAGATGTGTCGGTCCC	6-FAM	1	CAG	2	397, 400	PP199438, PP199439	0.35	0.75
MSEN09	TTGTCACTGCGCTATCTCCC	CGACGGAACTGTCATCCAGG	PET	1	TGG	2	408, 412	PP199440, PP199441	0.44	0.89
MSEN10	AGTGCCTGATTATGGAGCGG	TTCGGCGGGTTTACAGTAGC	6-FAM	2	TGCTTG	2	223, 235	PP199442, PP199443	0.24	0.64
MSEN11	ATGGAGGTGATTGATGCCCG	GATTCGGCCAATCAATGCCC	NED	2	GAA	3	280, 292, 295	PP199444, PP199445, PP199446	0.47	0.89
MSEN12	TTTGTCCGAATGTCCTGGGG	GTGAGGCTCAGCTCTCAACC	6-FAM	2	GAG	2	324, 327	PP199447, PP199448	0.09	0.47
MSEN13	CTTTTTGACCTCGCGTGTCG	ATTGGGAGAGTACAAGCGGC	NED	2	GAC	2	365, 371	PP199449, PP199450	0.46	0.92
MSEN14	GATCGATGCGGGATAGGTGG	GCGAGACGATATGGGAGACG	NED	2	AGAAGC	2	397, 403	PP199451, PP199452	0.42	0.86
MSEN15	AGCCCTCGTCGTGTTTAACC	CGATTTGCCATAGTGTGCGG	VIC	2	CCA	2	412, 415	PP199453, PP199454	0.23	0.62

Abbreviations: E_5, Pielou's evenness index; H_{exp}, Nei's unbiased gene diversity.

2.3 | Population genetic analysis

All population genetic analyses were performed in R v. 4.0.2 (R Core Team, 2020) using the package *poppr* (Kamvar et al., 2014). For each individual locus, allele diversity and abundance, including number of observed alleles, Nei's gene diversity (H_{exp}; Nei, 1978) and Pielou's evenness index (E_5; Pielou, 1975) were determined. A genotype accumulation curve (Kamvar et al., 2014), calculated by randomly sampling *n* loci 1000 times, was plotted to determine whether the number of markers developed were sufficient to capture the expected genotypic diversity.

In addition to H_{exp} and E_5 indices, to describe genotypic richness and diversity, number of multilocus genotypes (MLG) observed, number of estimated MLG based on smallest sample size calculated using rarefaction (eMLG; Grünwald et al., 2003), Simpson's index (λ; Simpson, 1949), Stoddard and Taylor's index (*G*; Stoddart & Taylor, 1988), maximum percentage of genotypic diversity (G_{max}; $G_{max}=G/N\times100$), and Shannon–Wiener index (*H*; Shannon, 2001) were all considered.

Minimum spanning networks (MSN) of all MLGs were constructed based on Bruvo's genetic distance (Bruvo et al., 2004) to assess their relatedness. A discriminant analysis of principal components (DAPC) was constructed to consider the possible presence of any population cluster. In addition, an analysis of molecular variance (AMOVA; Excoffier et al., 1992) was calculated to detect differentiation between and within populations, as well as between and within sampling hierarchies.

The mating-type idiomorph (*MAT1-1* or *MAT1-2*) in each of 186 isolates was identified using primers M1-EIF/M1-EIR and M2-EIF/M2-EIR, as described by Pham et al. (2023). Mating-type ratio, with the null hypothesis of a 1:1 ratio for the occurrence of the two *MAT1* idiomorphs in sexually reproducing populations, was tested using Pearson's chi-square goodness-of-fit test (Pearson, 1900). In addition, the standardized index of association ($\bar{r}_d$; Agapow & Burt, 2001) was calculated to verify any signature of random mating, using the clone-corrected and non-clone-corrected datasets.

3 | RESULTS

3.1 | Isolates and identification

A total of 186 isolates were obtained from leaves and shoots showing symptoms of *E. necatrix* infection, from four geographic regions in North Sumatra (Figure 1, Table S1). The collection included 17 isolates from AER, 26 from HAB, 64 from TEL and 79 isolates from AEN. Samples were obtained from different species of *Eucalyptus*, that is, *E. grandis*, *E. urophylla*, *E. pellita*, *E. dunnii*, *E. resinifera* and their hybrids. The fungus was also isolated from *Corymbia* spp., that is, *C. citriodora* and its hybrid *Corymbia torelliana*×*citriodora* (Table S1). Of the 79 isolates from AEN, 44 were obtained from a hierarchal sampling of *E. pellita* trees in a single compartment (Table S1). The identity of all isolates was confirmed as *E. necatrix* based on their ITS sequences.

3.2 | Microsatellite marker development and amplification

Genome mining of isolate CMW 56134 using Krait resulted in a total of 1054 perfect microsatellite repeat regions, of which 700 fulfilled the search criteria. Of the 700, 547 were tri-, 55 were tetra-, 42 were penta- and 56 were hexanucleotide repeats. A total of 319 intergenic microsatellite regions were extracted after mapping to the genome annotation in Krait. These potential regions were mapped to the genome of isolate CMW 56129 to identify polymorphic simple-sequence repeats (SSRs) between two genomes, and a total of 15 regions were retained after further filtering based on the selective criteria described above. Primer pairs were designed to amplify these 15 regions (Table 1). When used in PCRs, all 15 sets of primers resulted in the amplification of a single product that contained the targeted microsatellite region consistently and were confirmed as being polymorphic based on amplicon sequence analyses. The 15 labelled primer sets were divided into two panels (Table 1), according to their expected amplicon size, and they resulted in the successful amplification of their targeted microsatellite regions in all 186 isolates.

Variations were observed between the allele sizes obtained from GeneScan fragment analysis and fragment lengths obtained through Sanger sequencing. In particular, the actual size for MSEN01 obtained from sequencing was 3 bp shorter (−3) than that scored from fragment analysis. Similarly, this was observed in the case of MSEN02 (−2), MSEN03 (+2), MSEN04 (+0), MSEN05 (−2), MSEN06 (+1), MSEN07 (+0), MSEN08 (+4), MSEN09 (+1), MSEN10 (−1), MSEN11 (+2), MSEN12 (+1), MSEN13 (+2), MSEN14 (+2), MSEN15 (+3). The actual allele sizes obtained from sequencing were used to readjust those obtained from GeneScan fragment analysis and used in all subsequent analyses. This was done in order to facilitate the transferability of data generated by different systems in future studies (Table 1; Table S1).

3.3 | Population genetic analyses

Across the 15 loci, a total of 31 alleles were detected from 186 isolates, and two or three alleles were obtained per locus with a mean of 2.067 (Table 1). Locus MSEN11 had the greatest number of observed alleles (three). Loci MSEN02, MSEN06, MSEN11 were the most diverse ($H_{exp}=0.47$) and among those having the most even distribution of alleles ($E_5=0.89$–0.94). Locus MSEN12 had the lowest diversity ($H_{exp}=0.09$) and the most uneven allele distribution ($E_5=0.47$). No significant pairwise linkage was observed between loci based on the linkage disequilibrium test, and all 15 loci were thus included in subsequent analyses. The genotype accumulation curve showed that the number of loci used in this study were sufficient to capture the genotypic diversity of these isolates (Figure S1).

Genetic diversity statistics for *E. necatrix* populations collected at the four sampling locations are presented in Table 2. There was

high genotypic diversity and a very low number of shared genotypes at each of the four sites. In total, 152 MLGs were detected among the 186 isolates, 136 of which were represented by a single isolate. Only 10 MLGs occurred in at least two populations. MLG87 was represented by eight individuals and was the only MLG present in three populations (AEN, HAB and TEL). All populations had a similar proportion of MLGs (eMLGs). Of all 31 alleles detected across 15 markers, 30 were observed in all four populations and one unique allele (locus MSEN11; allele 295) was observed in AEN. The unbiased gene diversity (H_{exp}) was relatively similar between AEN, AER and HAB (H_{exp} = 0.33–0.35), whereas TEL had a slightly smaller value (0.29). The maximum percentage of genetic diversity was reached in HAB (G_{max} = 100) and lowest in TEL (G_{max} = 52.46). The Simpson indices (λ) for the four sites showed a high probability (93%–98%) that two randomly selected samples had a different genotype. High evenness indices (E_5 = 0.95–1.00 in AEN, AER, and HAB; E_5 = 0.76 in TEL) suggested that sites were not dominated by a single genotype as is the case in clonal populations.

The lack of population structure related to geographic location was evident in the DAPC, where overlap was observed in MLGs between samples from different locations (Figure 2). The minimum spanning network inferred from the SSR data showed no indication of any cluster specifically related to sampling locations or host genotypes (Figure 3). These results were statistically supported by AMOVA calculations, with only 2.13% of the variation present between regional populations, while 97.87% of the genetic variation was attributed to variability between individuals (p = 0.002; Table 3).

The mating-type PCR assay revealed that both *MAT1-1* and *MAT1-2* idiomorphs were detected in isolates from all four locations sampled in North Sumatra (Table 4). There was also no significant deviation from the expected 1:1 ratio of mating type in isolates in any of the four sampled areas, indicating the occurrence of random mating in the populations. This was supported by the index of association analyses using both non-clone-corrected and clone-corrected datasets for the four locations presented in Table 4. All the $\bar{r}_d$ values approached 0 and fell within the normal distribution expected under no linkage (p > 0.01). The results showed that none of the four regional populations were in linkage disequilibrium, suggesting a random recombination of alleles during outcrossing.

No significant differences were observed in genetic diversity from different hierarchal spatial scales of sampling in AEN (Table S2). The Simpson indices predicted at least 97% chance of randomly sampling two different genotypes, irrespective of sampling being obtained from a single compartment or from different compartment (Table S2). Multiple genotypes were also found for isolates from a single branchlet. Clonal genotypes were occasionally detected on the same branchlet, but the number was particularly low (Table S1). Mating-type ratios and linkage disequilibrium analyses were indicative of sexual recombination in both sampling hierarchies (Table S2). No population differentiation was detected between samples from a single compartment compared with those obtained from different compartments in AEN (Table S2, Figures 4 and 5).

4 | DISCUSSION

E. necatrix is a newly emerged *Eucalyptus* leaf and shoot pathogen that has, thus far, been found only in Indonesia. The development of polymorphic microsatellite markers in this study, and their application on a large and diverse collection of isolates of this aggressive pathogen, revealed a lack of population structure. These isolates, collected from four distinct locations and from different eucalypt species or genotypes also displayed a substantial level of gene flow among all the subpopulations examined. The high levels of genotypic diversity, a lack of dominant clonal genotypes, the absence of strong linkage disequilibrium and a broadly equal occurrence of both mating types in isolates was surprising given that the disease caused by *E. necatrix* emerged rapidly and was unknown prior to 2014.

Our results showed that the population of *E. necatrix* is neither defined by geographical location nor eucalypt host species. This is despite the fact that *E. necatrix* has a wide range of *Eucalyptus* hosts in North Sumatra. Initially, the pathogen was reported from *E. grandis*, *E. urophylla* and hybrids of these trees (Pham et al., 2021). Our collections showed that at least three other species of *Eucalyptus* and their hybrids can be infected. Interestingly, infections caused by *E. necatrix* were also found on *Corymbia* spp. (previously accommodated in *Eucalyptus*; Hill & Johnson, 1995), and their hybrids. The fact that *E. necatrix* is able to infect various host species is intriguing

TABLE 2 Indices of genetic diversity based on microsatellite data for *Elsinoe necatrix* from North Sumatra, Indonesia.

Region	*N*	MLG	eMLG	Number of alleles/ private alleles	*H*	*G*	G_{max}	λ	E_5	H_{exp}
Aek Nauli (AEN)	79	70	16.6 ± 0.59	31/1	4.21	64.34	81.44	0.98	0.95	0.35
Aek Raja (AER)	17	16	16.0 ± 0.00	30/0	2.75	15.21	89.47	0.93	0.97	0.35
Habinsaran (HAB)	26	26	17.0 ± 0.00	30/0	3.26	26.00	100.00	0.96	1.00	0.33
Tele (TEL)	64	51	15.5 ± 1.15	30/0	3.78	33.57	52.46	0.97	0.76	0.29
Total	186	152	16.5 ± 0.70	31	4.91	108.79	58.49	0.99	0.80	0.33

Abbreviations: E_5, Pielou's evenness index; eMLG, number of estimated MLG based on smallest sample size calculated using rarefaction; *G*, Stoddard and Taylor's index; G_{max}, maximum percentage of genotypic diversity; *H*, Shannon–Wiener index; H_{exp}, Nei's unbiased gene diversity; MLG, number of multilocus genotypes observed; *N*, number of individuals; λ, Simpson's Index.

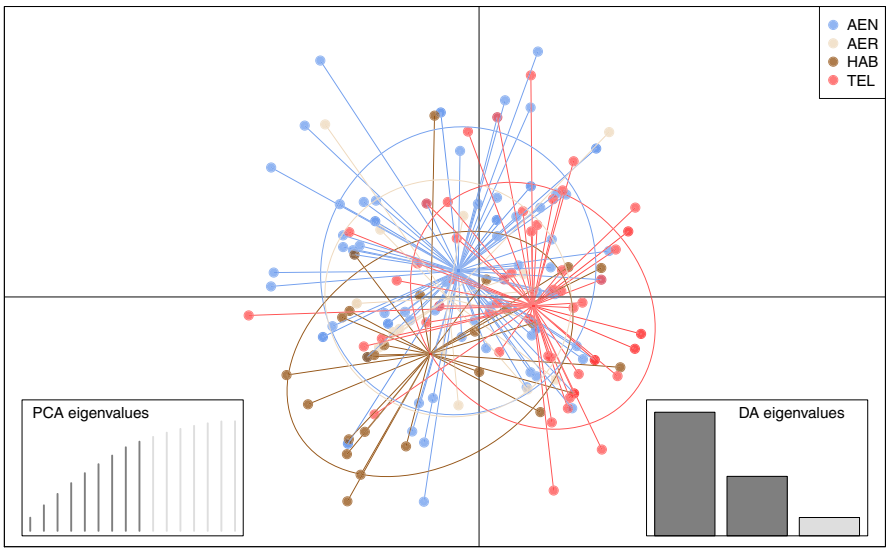

FIGURE 2 Scatterplot from discriminant analysis of principle components (DAPC) of the first two principle components discriminating *Elsinoe necatrix* populations. Individuals are indicated as dots and the four estimated populations, based on geographic locations (Figure 1), are shown as inertia ellipses. [Colour figure can be viewed at wileyonlinelibrary.com]

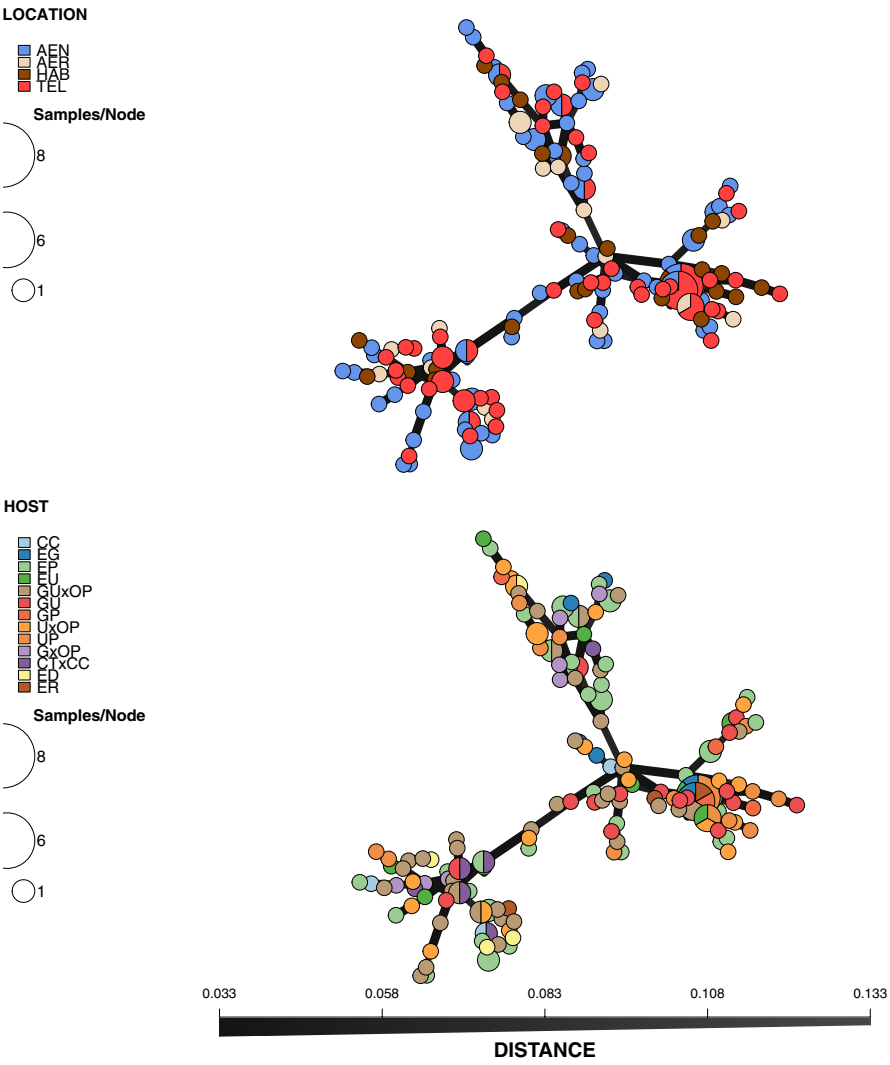

FIGURE 3 Minimum spanning network (MSN) showing the relationship of *Elsinoe necatrix* genotypes based on Bruvo's distance. Each node represents one multilocus genotype (MLG) and the size of the node is proportional to the number of individuals with that MLG. Nodes are coloured according to sampling location and host. Location: AEN, Aek Nauli; AER, Aek Raja; HAB, Habinsaran; TEL, Tele; Host: CC, *Corymbia citriodora*; CTxCC, *Corymbia torelliana*×*citriodora*; ED; *Eucalyptus dunnii*; EG, *Eucalyptus grandis*; EP, *Eucalyptus pellita*; ER; *Eucalyptus resinifera*; EU, *Eucalyptus urophylla*; GP, *E. grandis*×*pellita*; GU, *E. grandis*×*urophylla*; GUxOP, *E. grandis*×*urophylla* (open pollination); GxOP, *E. grandis* (open pollination); UP, *E. urophylla*×*pellita*; UxOP, *E. urophylla* (open pollination). [Colour figure can be viewed at wileyonlinelibrary.com]

in light of the relatively narrow host range reported for other species in this genus (Fan et al., 2017).

The origin of *E. necatrix* remains unknown. One interpretation is that the population of the pathogen in North Sumatra could have originated on trees in native forests, which occur in close proximity to the plantation areas. In this case, it is possible that several strains of the pathogen underwent a host shift from native *Myrtaceae* to infect susceptible *Eucalyptus* in the plantations. Future surveys to

TABLE 3 Analysis of molecular variance (AMOVA) of microsatellite data for *Elsinoe necatrix* isolates separated based on geographic location.

	df	Sum of squares	Mean squares	Estimate of variance	Total variation (%)	φ	*p*
Between four populations	3	28.02	9.34	0.11	0.90	0.02	0.002
Within populations	182	890.91	4.90	4.90	99.10		

TABLE 4 Mating-type ratios and tests for linkage disequilibrium of *Elsinoe necatrix* isolates collected from four regions in North Sumatra.

Region	MAT1-1:MAT1-2 (*p*-value χ^2)[a]	Before clone correction		After clone correction	
		$\bar{r}_d$	*p*	$\bar{r}_d$	*p*
AEN	37:42 (0.57)	0.002	0.295	−9.25e−06	0.496
AER	11:6 (0.23)	0.026	0.029	0.021	0.072
HAB	16:10 (0.24)	−0.010	0.853	−0.010	0.858
TEL	40:24 (0.05)	0.016	0.011	−0.002	0.637
Total	104:82 (0.11)	0.005	0.061	0.001	0.323

Note: $\bar{r}_d$, the standardized index of association.

[a]Pearson's χ^2 goodness-of-fit test for the null hypothesis of 1:1 ratio for the occurrence of the two *MAT1* idiomorphs in sexually reproducing populations; $p > 0.05$ indicates that null hypothesis was not rejected.

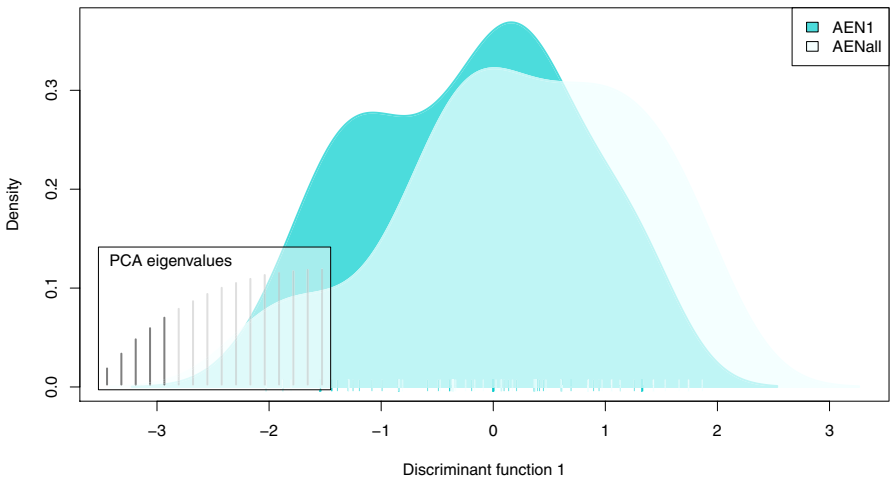

FIGURE 4 Density plot of the first axis of discriminant analysis of principal components (DAPC) for the different sampling levels in Aek Nauli (AEN). AEN1, isolates from one compartment; AENall, isolates from multiple compartments. [Colour figure can be viewed at wileyonlinelibrary.com]

consider this questions need to include native *Myrtaceae* in natural forests of Indonesia.

An alternative view to *E. necatrix* having originated on native trees in Indonesia would be that the pathogen is an alien invasive fungus recently introduced into plantations of North Sumatra. The rapid emergence of the disease, coupled with its sudden and destructive impact would support this hypothesis. Furthermore, our microsatellite analyses revealed signals of a genetic founder effect, where most (14) of the polymorphic microsatellite loci had only two alleles. Although the locus MSEN11 had three alleles, the third allele was found in only a single individual. This could be a signal of mutation, or a shift in allele frequencies after a founder event. Furthermore, unhindered sexual outcrossing could result in considerable genetic variation not typical of a founder effect, such as observed in the populations of some sexually reproducing invasive tree pathogens (Bengtsson et al., 2012; Gross et al., 2014; Pérez et al., 2010). Our results thus hint at the possibility that *E. necatrix* may have been accidentally introduced to North Sumatra, initially by a small number of individuals with compatible mating types.

The fact that Eucalyptus scab and shoot malformation is apparently caused by panmictic populations of *E. necatrix* is surprising, given that the disease has only recently emerged in North Sumatra. This pattern is different to that of many other newly invasive pathogens of *Eucalyptus* that are usually characterized by an uncomplicated clonal population structure and low genetic diversity. For example, clonal populations have been observed for numerous Teratosphaeria leaf blight pathogens (Aylward et al., 2021; Havenga et al., 2020; Pérez et al., 2012; Taole et al., 2015), the Teratosphaeria stem canker pathogens (Jimu et al., 2016a, 2016b), and in the case of the well-known Chrysoporthe stem canker pathogen (Nakabonge et al., 2007). Where populations of such

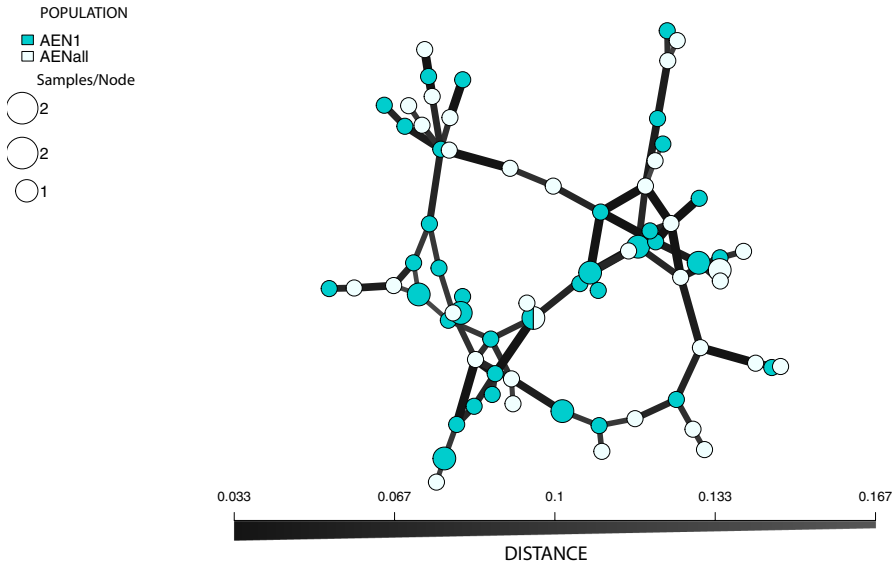

FIGURE 5 Minimum spanning network (MSN) showing the relationship of *Elsinoe necatrix* genotypes within Aek Nauli (AEN). Each node represents one multilocus genotype (MLG) and the size of the node is proportional to the number of individuals with that MLG. Nodes are coloured according to different hierarchical sampling scales. AEN1, isolates from one compartment; AENall, isolates from multiple compartments. [Colour figure can be viewed at wileyonlinelibrary.com]

pathogens have high levels of genetic diversity, as was found for *E. necatrix* in the present study, these are interpreted as having either arisen from multiple introductions of different genotypes or due to sexual outcrossing. Our results suggest that sexual outcrossing is occurring in *E. necatrix*, characterized by high genotypic diversity, random associations of alleles at different loci and the existence of equal proportions of mating types.

Suggesting that sexual reproduction has led to the highly diverse population of *E. necatrix* in Sumatra is bedevilled by the fact that a sexual state has not been found for the pathogen. The only infective propagules that are known for the pathogen are conidia. These have mostly been observed using scanning electron microscopy of infected leaves, although they have occasionally also been found in cultures subsequent to the taxonomic description of *E. necatrix*. Sexual states are known for other *Elsinoe* spp. such as *E. panici* or *E. veneta*, where ascospores are typically wind dispersed (Anderson, 1956; Gabel & Tiffany, 1987; Jones, 1924). Assuming, as seems likely in the case of *E. necatrix*, that a cryptic sexual state is present in the eucalypt-growing environment of North Sumatra, wind-dispersed ascospores could arise from ascocarps on infected debris. These ascospores would have the potential to spread freely to different trees across the plantations and be responsible for primary infections as has been shown for some other species of *Elsinoe* (Anderson, 1956; Gabel & Tiffany, 1987; Jones, 1924).

In most cases, different genotypes of *E. necatrix* were found to be present on the same tree. This indicates that the populations of *E. necatrix* in North Sumatra are highly diverse, with multiple and independent infections occurring on individual plants. We hypothesize that gene flow is primarily achieved through the dispersal of spores produced by sexual reproduction over medium to long distances, because of a lack of genetic coherence among individuals within the same tree or compartment. Genetically similar *E. necatrix* individuals were found on the same branchlet, but only at very low levels. This implies that localized vertical dispersion or repeated infections of clonal genotypes is occurring sporadically.

Many eucalypt-infecting fungi in their native habitats tend to be inconsequential where they occur in diverse native ecosystems due to long periods of co-evolution with their hosts (Burgess & Wingfield, 2017; Crous et al., 2019). However, in the setting of large-scale monoculture plantations, they can have a high epidemic potential (Andjic et al., 2019; Pham et al., 2021; Roux et al., 2024). The question arises as to why *E. necatrix* has been so successful in North Sumatra. This could be due to a relatively low host genetic diversity and low resistance levels in clonal plantations. That, coupled with favourable climatic conditions such as high rainfall and cooler temperatures in North Sumatra, have provided an environment conducive to a rapid build-up of outcrossing genotypes of *E. necatrix*.

This study provides insights into the genetic diversity and reproductive potential of *E. necatrix* that affords opportunities for efficient control measures. Future studies should include a comparative investigation of the relative susceptibility of various provenances of *Eucalyptus* to *E. necatrix*, and to identify genetic markers associated with the pathosystem and host defence mechanisms. The data presented here suggest that *E. necatrix* is actively recombining and thus is likely to overcome narrow host genetic resistance, which is an effective strategy in many groups of pathogenic fungi in the 'arms race' (Clay & Kover, 1996) with their hosts (Drenth et al., 2019; McDonald & Linde, 2002). Thus, breeding programmes will need to include a substantially high level of host genetic diversity (Burdon, 2001; Burdon et al., 2006) to contend with the disease caused by *E. necatrix*.

ACKNOWLEDGEMENTS

This study was initiated through the bilateral agreement between the Forestry and Agricultural Biotechnology Institute (FABI), University of Pretoria and the Royal Golden Eagle (RGE), Indonesia. We are grateful to both institutions for providing financial support. We are also grateful to Omar Syaref Purba who assisted us in collecting samples.

CONFLICT OF INTEREST STATEMENT

The authors declare that there is no conflict of interest.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request.

ORCID

Nam Q. Pham https://orcid.org/0000-0002-4938-9067
Hiroyuki Suzuki https://orcid.org/0000-0002-9301-9207
Irene Barnes https://orcid.org/0000-0002-4349-3402

REFERENCES

Agapow, P.M. & Burt, A. (2001) Indices of multilocus linkage disequilibrium. *Molecular Ecology Notes*, 1, 101–102.
Anderson, H.W. (1956) Diseases of fruit crops. In: *Diseases of fruit crops*. New York, USA: McGraw-Hill.
Andjic, V., Carnegie, A.J., Pegg, G.S., Hardy, G.E.S.J., Maxwell, A., Crous, P.W. et al. (2019) 23 years of research on Teratosphaeria leaf blight of *Eucalyptus*. *Forest Ecology and Management*, 443, 19–27.
Aylward, J., Havenga, M., Dreyer, L.L., Roets, F., Wingfield, B.D., Pérez, C. et al. (2021) Genetic diversity of *Teratosphaeria pseudoeucalypti* in *eucalyptus* plantations in Australia and Uruguay. *Australasian Plant Pathology*, 50, 639–649.
Bengtsson, S.B., Vasaitis, R., Kirisits, T., Solheim, H. & Stenlid, J. (2012) Population structure of *Hymenoscyphus pseudoalbidus* and its genetic relationship to *Hymenoscyphus albidus*. *Fungal Ecology*, 5, 147–153.
Bruvo, R., Michiels, N.K., d'Souza, T.G. & Schulenburg, H. (2004) A simple method for the calculation of microsatellite genotype distances irrespective of ploidy level. *Molecular Ecology*, 13, 2101–2106.
Burdon, J.J., Thrall, P.H. & Ericson, L. (2006) The current and future dynamics of disease in plant communities. *Annual Review of Phytopathology*, 44, 19–39.
Burdon, R.D. (2001) Genetic diversity and disease resistance: some considerations for research, breeding, and deployment. *Canadian Journal of Forest Research*, 31, 596–606.
Burgess, T.I. & Wingfield, M.J. (2017) Pathogens on the move: A 100-year global experiment with planted eucalypts. *Bioscience*, 67, 14–25.
Clay, K. & Kover, P.X. (1996) The red queen hypothesis and plant/pathogen interactions. *Annual Review of Phytopathology*, 34, 29–50.
Crous, P.W., Wingfield, M.J., Cheewangkoon, R., Carnegie, A., Burgess, T., Summerell, B. et al. (2019) Foliar pathogens of eucalypts. *Studies in Mycology*, 94, 125–298.
Drenth, A., McTaggart, A.R. & Wingfield, B.D. (2019) Fungal clones win the battle, but recombination wins the war. *IMA Fungus*, 10, 18.
Du, L., Zhang, C., Liu, Q., Zhang, X. & Yue, B. (2018) Krait: an ultrafast tool for genome-wide survey of microsatellites and primer design. *Bioinformatics*, 34, 681–683.
Excoffier, L., Smouse, P.E. & Quattro, J. (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics*, 131, 479–491.
Fan, X., Barreto, R., Groenewald, J.Z., Bezerra, J., Pereira, O., Cheewangkoon, R. et al. (2017) Phylogeny and taxonomy of the scab and spot anthracnose fungus *Elsinoë* (Myriangiales, Dothideomycetes). *Studies in Mycology*, 87, 1–41.
Gabel, A. & Tiffany, L. (1987) Life history of *Elsinoë panici*. *Proceedings of the Iowa Academy of Sciences*, 94, 121–127.
Gross, A., Hosoya, T. & Queloz, V. (2014) Population structure of the invasive forest pathogen *Hymenoscyphus pseudoalbidus*. *Molecular Ecology*, 23, 2943–2960.
Grünwald, N.J., Goodwin, S.B., Milgroom, M.G. & Fry, W.E. (2003) Analysis of genotypic diversity data for populations of microorganisms. *Phytopathology*, 93, 738–746.
Havenga, M., Wingfield, B.D., Wingfield, M.J., Dreyer, L.L., Roets, F., Chen, S. et al. (2020) Low genetic diversity and strong geographic structure in introduced populations of the *Eucalyptus* foliar pathogen *Teratosphaeria destructans*. *Plant Pathology*, 69, 1540–1550.
Hill, K.D. & Johnson, L.A.S. (1995) Systematic studies in the eucalypts. 7. A revision of the bloodwoods, genus *Corymbia* (Myrtaceae). *Telopea*, 6, 185–504.
Iglesias-Trabado, G. & Wilstermann, D. (2008) *Eucalyptus* universalis: Global cultivated eucalypt forests map 2008, version 1.0.1. *GIT Forestry Consulting's Eucalyptologics: Information Resources on Eucalyptus Cultivation Worldwide*. Accessed 23rd January 2024: http://git-forestry-blog.blogspot.com/2008/09/eucalyptus-global-map-2008-cultivated.html
Jimu, L., Chen, S., Wingfield, M.J., Mwenje, E. & Roux, J. (2016a) Three genetic groups of the eucalyptus stem canker pathogen *Teratosphaeria zuluensis* introduced into Africa from an unknown source. *Antonie Van Leeuwenhoek*, 109, 21–33.
Jimu, L., Chen, S., Wingfield, M.J., Mwenje, E. & Roux, J. (2016b) The *Eucalyptus* stem canker pathogen *Teratosphaeria gauchensis* represents distinct genetic groups in Africa and South America. *Forest Pathology*, 46, 229–239.
Jones, L.K. (1924) Anthracnose of cane fruits and its control on black raspberries in Wisconsin. *Agricultural Experiment Station Bulletin*, 59, 1–26.
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.
McDonald, B.A. & Linde, C. (2002) Pathogen population genetics, evolutionary potential, and durable resistance. *Annual Review of Phytopathology*, 40, 349–379.
Nakabonge, G., Gryzenhout, M., Wingfield, B.D., Wingfield, M.J. & Roux, J. (2007) Genetic diversity of *Chrysoporthe cubensis* in eastern and southern Africa. *South African Journal of Science*, 103, 261–264.
Nei, M. (1978) Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics*, 89, 583–590.
Ni, M., Feretzaki, M., Sun, S., Wang, X. & Heitman, J. (2011) Sex in fungi. *Annual Review of Genetics*, 45, 405–430.
Pearson, K. (1900) X. On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 50, 157–175.
Pérez, G., Slippers, B., Wingfield, B.D., Hunter, G.C. & Wingfield, M.J. (2010) Micro- and macrospatial scale analyses illustrates mixed mating strategies and extensive geneflow in populations of an invasive haploid pathogen. *Molecular Ecology*, 19, 1801–1813.
Pérez, G., Slippers, B., Wingfield, M.J., Wingfield, B.D., Carnegie, A.J. & Burgess, T.I. (2012) Cryptic species, native populations and biological invasions by a eucalypt forest pathogen. *Molecular Ecology*, 21, 4452–4471.
Pham, N.Q., Duong, T.A., Wingfield, B.D., Barnes, I., Durán, A. & Wingfield, M.J. (2023) Characterisation of the mating-type loci in species of *Elsinoe* causing scab diseases. *Fungal Biology*, 127, 1484–1490.
Pham, N.Q., Marincowitz, S., Solís, M., Duong, T.A., Wingfield, B.D., Barnes, I. et al. (2021) Eucalyptus scab and shoot malformation: a new and serious foliar disease of *Eucalyptus* caused by *Elsinoe necatrix* sp. nov. *Plant Pathology*, 70, 1230–1242.
Pielou, E.C. (1975) *Ecological diversity*. New York, NY: John Wiley & Sons.
R Core Team. (2020) *R: a language and environment for statistical computing*. R Foundation for Statistical Computing. Available from: https://www.R-project.org [Accessed 31st January 2024].

Roux, J., Wingfield, M.J., Marincowitz, S., Solís, M., Phungula, S. & Pham, N.Q. (2024) Eucalyptus scab and shoot malformation: a new disease in South Africa caused by a novel species, *Elsinoe masingae*. *Forestry*, 97, cpad031.

Shannon, C.E. (2001) A mathematical theory of communication. *ACM SIGMOBILE Mobile Computing and Communications Review*, 5, 3–55.

Simpson, E. (1949) Measurement of diversity. *Nature*, 163, 688.

Stoddart, J.A. & Taylor, J.F. (1988) Genotypic diversity: estimation and prediction in samples. *Genetics*, 118, 705–711.

Taole, M., Bihon, W., Wingfield, B.D., Wingfield, M.J. & Burgess, T.I. (2015) Multiple introductions from multiple sources: invasion patterns for an important *Eucalyptus* leaf pathogen. *Ecology and Evolution*, 5, 4210–4220.

Ter-Hovhannisyan, V., Lomsadze, A., Chernoff, Y.O. & Borodovsky, M. (2008) Gene prediction in novel fungal genomes using an ab initio algorithm with unsupervised training. *Genome Research*, 18, 1979–1990.

Turgeon, B.G. & Yoder, O. (2000) Proposed nomenclature for mating type genes of filamentous ascomycetes. *Fungal Genetics and Biology*, 31, 1–5.

Turnbull, J.W. (1999) Eucalypt plantations. *New Forests*, 17, 37–52.

Untergasser, A., Cutcutache, I., Koressaar, T., Ye, J., Faircloth, B.C., Remm, M. et al. (2012) Primer3—new capabilities and interfaces. *Nucleic Acids Research*, 40, e115.

Wingfield, B.D., Berger, D.K., Coetzee, M.P., Duong, T.A., Martin, A., Pham, N.Q. et al. (2022) IMA genome-F17: draft genome sequences of an *Armillaria* species from Zimbabwe, *Ceratocystis colombiana*, *Elsinoë necatrix*, *Rosellinia necatrix*, two genomes of *Sclerotinia minor*, short-read genome assemblies and annotations of four *Pyrenophora* teres isolates from barley grass, and a long-read genome assembly of *Cercospora zeina*. *IMA Fungus*, 13, 19.

Wingfield, M.J. (2003) 2003 Daniel McAlpine memorial lecture increasing threat of diseases to exotic plantation forests in the southern hemisphere: lessons from Cryphonectria canker. *Australasian Plant Pathology*, 32, 133–139.

Wingfield, M.J., Slippers, B., Hurley, B., Coutinho, T., Wingfield, B.D. & Roux, J. (2008) Eucalypt pests and diseases: growing threats to plantation productivity. *Southern Forests*, 70, 139–144.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Pham, N.Q., Suzuki, H., Duong, T.A., Wingfield, B.D., Barnes, I., Durán, A. et al. (2024) Cryptic sexual reproduction in an emerging *Eucalyptus* shoot and foliar pathogen. *Plant Pathology*, 73, 1145–1155. Available from: https://doi.org/10.1111/ppa.13876