Received: 7 July 2021 | Revised: 27 January 2022 | Accepted: 31 January 2022

DOI: 10.1111/mec.16384

ORIGINAL ARTICLE

MOLECULAR ECOLOGY WILEY

Long evolutionary history of an emerging fungal pathogen of diverse tree species in eastern Asia, Australia and the Pacific Islands

Olga Kozhar[1] | Mee-Sook Kim[2] | Jorge Ibarra Caballero[1] | Ned B. Klopfenstein[3] | Phil G. Cannon[4] | Jane E. Stewart[1]

[1]Colorado State University, Fort Collins, Colorado, USA

[2]USDA Forest Service, Pacific Northwest Research Station, Corvallis, Oregon, USA

[3]USDA Forest Service, Rocky Mountain Research Station, Moscow, Idaho, USA

[4]USDA Forest Service, Forest Health Protection, Vallejo, California, USA

Correspondence
Olga Kozhar and Jane E. Stewart, Colorado State University, Fort Collins, CO, USA.
Emails: olga.kozhar@colostate.edu (O.K.); jane.stewart@colostate.edu (J.E.S.)

Mee-Sook Kim, USDA Forest Service, Pacific Northwest Research Station, Corvallis, OR, USA.
Email: meesook.kim@usda.gov

Funding information
USDA Forest Service, State & Private Forestry, Forest Health Protection, Special Technology Development Program, Grant/Award Number: R5-2015-01 and R5-2019-02; Joint Venture Agreements, Grant/Award Number: 15-JV-11221633-160, 19-JV-11221633-093, 19-JV-11261957-078 and 20-JV-11221633-141; Colorado State University

Handling Editor: Sara Branco

Abstract

Emerging plant pathogens have been increasing exponentially over the last century. To address this issue, it is critical to determine whether these pathogens are native to ecosystems or have been recently introduced. Understanding the ecological and evolutionary processes fostering emergence can help to manage their spread and predict epidemics/epiphytotics. Using restriction site-associated DNA sequencing data, we studied genetic relationships, pathways of spread and the evolutionary history of *Phellinus noxius*, an emerging root-rotting fungus of unknown origin, in eastern Asia, Australia and the Pacific Islands. We analysed patterns of genetic variation using Bayesian inference, maximum-likelihood phylogeny, population splits and mixtures measuring correlations in allele frequencies and genetic drift, and finally applied coalescent-based theory using Approximate Bayesian computation (ABC) with supervised machine learning. Population structure analyses revealed five genetic groups with signatures of complex recent and ancient migration histories. The most probable scenario of ancient pathogen spread is movement from an unsampled population to Malaysia and the Pacific Islands, with subsequent spread to Taiwan and Australia. Furthermore, ABC analyses indicate *P. noxius* spread occurred thousands of generations ago, contradicting previous assumptions that this pathogen was recently introduced to multiple geographical regions. Our results suggest that recent emergence of *P. noxius* in eastern Asia, Australia and the Pacific Islands has probably been driven by anthropogenic and natural disturbances, such as deforestation, land-use change, severe weather events and/or introduction of exotic plants. This study provides a novel example of applying genome-wide allele frequency data to unravel the dynamics of pathogen emergence under changing ecosystem conditions.

KEYWORDS
brown root rot disease, disturbed ecosystems, evolutionary history, gene flow, genetic diversity, pathogen emergence, *Phellinus noxius*, population structure

1 | INTRODUCTION

New opportunities for organismal evolution can be created by human-induced changes within diverse ecosystems. Early spread of pathogens and other ecosystem changes are associated with the migration of *Homo sapiens* out of central Africa and subsequent human migrations (Cavalli-Sforza & Feldman, 2003). The development of agriculture, animal/plant domestication and trade among global regions has provided more opportunities to create anthropogenic environments that reflect the changing demands of human societies (Pingali, 2012; Santini et al., 2018). Land usage for general infrastructure and food production is increasing at an alarming rate to support the rapid growth of human populations. While such drastic adaptations of wild ecosystems satisfy human needs in the short term, these ecosystem changes often lead to the ecological imbalances that are associated with long-term, negative consequences, including climate change, epidemics/epiphytotics and severe decreases of biodiversity. Over the last few decades, an unprecedented number of plant diseases caused by fungi and other pathogens have caused severe damage and/or extirpation of native and domesticated species, which has also threatened food security worldwide (Corredor-Moreno & Saunders, 2019; Fisher et al., 2012).

Forests are among the most biologically complex terrestrial ecosystems, and also contain numerous and diverse pathosystems (Richardson et al., 2005). Within healthy and well-balanced forest ecosystems, individual tree species are assumed to live in quasi-equilibrium with pathogens (Otrosina, 2005; Otrosina and Ferrell, 1995). However, altering of such equilibria can lead to disease emergence due to multiple causes, including the introduction of invasive pathogens (Goss et al., 2011; Hamelin, 2005; Potter & Urquhart, 2017; Stewart et al., 2018), physical disturbances of forest land that result in increased damage via interactions between native pathogens and host species (Gilbert & Hubbell, 1996; Hansen & Goheen, 2000; Rachowicz et al., 2005; Speight & Woodward, 2016), the introduction of exotic plant species (Drenth & Guest, 2016; Speight & Woodward, 2016), as well as changes in environmental conditions (Ayres & Lombardero, 2000; Mildrexer et al., 2019; Ritokova et al., 2016; Woods et al., 2016). The underlying causes of plant pathogen emergence are not always clear from observations and historical records alone. Studying patterns of genetic variation and spatiotemporal dynamics in plant pathogen populations can, in turn, shed light on the origin of epiphytotics and/or pathogen emergence.

Populations of recently introduced organisms generally exhibit decreased genetic variation due to "founder effects" (Clegg et al., 2002; Nei et al., 1975; Slatkin & Excoffier, 2012). Such populations are often represented by a limited number of genetic lineages ("founders") from either the same or different source populations (Nichols & Hewitt, 1994; Ray et al., 2003; Slatkin, 1987). In contrast, populations of organisms with a long evolutionary history, in general, have high levels of genetic variation that accumulated over long time periods, and often can establish geographical structure across a regional scale (Slatkin, 1987; Wang et al., 2016). Mode of reproduction, ongoing and historical gene flow, genetic drift, mutation and selection influence the genetic patterns of extant populations. Applying different sets of contemporary analytical methods to genome-wide allele frequency data allows us to make inferences of evolutionary histories and mechanisms driving pathogen emergence with a greater confidence. Such methods have been utilized to unravel various evolutionary histories of well-monitored, economically important plant and animal pathogens (Dutech et al., 2012; Maside et al., 2015; Rosenblum et al., 2013; Silva et al., 2012). However, few case studies are available that differentiate between native and non-native origin hypotheses of pathogen emergence when historical and spatiotemporal data for these organisms are scarce (one such example is *Tracheliastes polycolpus*, an emerging ectoparasite of freshwater fishes; Rey et al., 2015).

In the present work, we demonstrate the utility of the classical population genetics, phylogenetic and simulation-based (e.g., Approximate Bayesian computation [ABC] with supervised machine learning) frameworks to unravel the evolutionary history of an emerging pathogen for which data on native origin are lacking. We specifically focused on the population structure and evolutionary history of *Phellinus noxius* (Corner) G. H. Cunningham, an emerging destructive tree pathogen of unknown origin (Ann et al., 1999). *P. noxius* is a fast-growing, diploid basidiomycete with a wide range of woody hosts in tropical/subtropical areas, including eastern Asia, Oceania, Central America and Africa (Pegler & Waterston, 1968). *P. noxius* causes brown root rot disease of more than 200 woody plant species from 59 families with little host specificity (Ann et al., 2002), including economically important crops, such as breadfruit (*Artocarpus altilis*), mango (*Mangifera indica*), cacao (*Theobroma cacao*), coffee (*Coffea arabica*), rubber (*Hevea brasiliensis*), as well as diverse fruit and nut crops (Ann et al., 1999; Bolland, 1984; Brooks, 2002; Chang, 1995; Ivory, 1996; Nandris et al., 1987a). Brown root rot causes tree mortality within one to several years following initial infection. Mycelia of the pathogen can survive in dead tree roots for more than 10 years, providing inoculum for new infections years after the last disease outbreak (Cannon et al., 2022; Chang, 1996). The primary dissemination method of *P. noxius* is via root-to-root contact that allows spread from infected roots to healthy roots of neighbouring trees, but the fungus is also capable of producing basidiospores that can travel longer distances via air (Chung et al., 2015). Early observations of *P. noxius* were reported in Japan (Yasuda, 1916) and Taiwan (Sawada, 1928), but it was not recognized as an important tree pathogen until it was identified in Singapore in 1932 (Corner, 1932). In the 1980s, outbreaks of *P. noxius*-caused wood rot were reported in tree plantations of the Northern Mariana Islands, Australia, Taiwan, Japan, Hainan Island of China, Vanuatu and West Africa (Abe et al., 1995; Ann et al., 2002; Bolland, 1984; Hodges & Tenorio, 1984; Kobayashi et al., 1991; Nandris et al., 1987a; Neil, 1986; Tai, 1979). Based on recent timing of its emergence in the Pacific regions, it has been hypothesized that *P. noxius* may have been introduced to some areas within the last century (Hodges & Tenorio, 1984). Recent *P. noxius* emergences have also been hypothesized to be induced by rapid environmental changes, such as typhoons in the Pacific (Akiba et al., 2015), reforestation (Carnegie, 2007) and land-use change

(Wairiu, 2017), and through human-mediated movement of plant material among islands (Cannon et al., 2022).

Recent studies with molecular markers of *P. noxius* populations reported high levels of genetic diversity and gene flow among populations in Taiwan (Chung et al., 2015) and Japan (Akiba et al., 2015). Phylogenetic and population genetics analyses using protein-coding loci revealed three genetic lineages of *P. noxius* in eastern Asia, Australia and the Pacific Islands (Stewart et al., 2020). These findings suggest a complex evolutionary history of *P. noxius* that has yet to be determined. Additionally, the results of bioclimatic modelling revealed that *P. noxius* represents a threat to many tropical/subtropical regions worldwide, including the Hawaiian Islands, southeastern USA, eastern Mexico, eastern Central America, parts of Africa and Madagascar, with different genetic lineages representing distinct threats (Stewart et al., 2020). Herein, using restriction site-associated DNA sequencing (RADseq) data on 107 *P. noxius* isolates from 15 locations, we address four questions: (i) What is the genetic structure of *P. noxius* populations in regions of eastern Asia, Australia and the Pacific Islands? (ii) What are the recent migration patterns of *P. noxius* across the studied regions? (iii) What are the most likely historical colonization routes of *P. noxius* across the studied regions? (iv) How old are the populations of *P. noxius* in eastern Asia, Australia and the Pacific Islands? Answering these questions will provide insights into the processes leading to the pathogen's emergence in contemporary ecosystems. Specifically, our study illustrates how the combination of population genetics, phylogenetics and ABC methods can be used to reconstruct evolutionary history, unravel original pathways of colonization, and support hypotheses of native and/or introduced origins of emerging pathogens, for which historical and ecological data are limited.

2 | MATERIALS AND METHODS

2.1 | Fungal isolates

A total of 107 *Phellinus noxius* isolates were collected from 15 locations: eastern Asia (Hong Kong, Malaysia, Taiwan, Japan [Ogasawara, Okinawa, Miyakojima and Taketomi]); Australia; and the Pacific Islands (Micronesia [Palau, Yap, Pohnpei and Kosrae], Mariana Islands [Guam and Saipan] and American Samoa) (see Table S1 for collection locations and isolate details).

2.2 | DNA extraction, RADseq and bioinformatics analyses

Genomic DNA was extracted as described in Stewart et al. (2020). RADseq libraries were prepared following Graham et al. (2015) with slight modifications. From each sample, 25 ng of DNA was digested using 0.33 mM 1× Cutsmart Buffer (New England Biolabs; NEB), 5.0 μl of dH_2O, and 0.5 μl of each enzyme: *Msp*I (20 units; NEB), *Bam*HI (20 units; NEB), and Cla (10 units; NEB). Samples were incubated for 1 h at 37°C for digestion. Subsequently, 0.75 mM ATP (NEB), 0.25× Ligase Buffer (NEB) and DNA Ligase (100 units; NEB) were added, and samples were incubated at 22°C for 20 min, 37°C for 10 min, followed by 65°C for 20 min to halt enzyme activity. Samples were purified with NaCl-PEG diluted 1.2× SpeedBeads (Cytiva Malborough) and resuspended with 20 μl of dH_2O.

Each library (one isolate per individual) was PCR (polymerase chain recation)-amplified in a 25-μl PCR containing 1× HiFi Buffer (Kapa Biosystems; Kapa), 0.375 mM dNTPs (Kapa), 0.5 units HiFi DNA Polymerase (Kapa), 0.5 mM iTru5 and iTru7 external primers (Integrated DNA Technologies; IDT), and 5 μl of internal adapter-ligated template from each library following Glenn et al. (2019). The PCR cycle was as follows: 95°C for 2 min, followed by 12 cycles of 98°C for 20 s, 60°C for 15 s and 72°C for 30 s, then a step of 72°C for 5 min, and lastly a hold at 15°C. Libraries were purified with SpeedBeads, quantified on a Qubit 2.0 Fluorometer (Life Technologies) and pooled. The pooled library was sized selected at 550 (±10%) with a LabChip XT Caliper (PerkinElmer by Macrogen and sequenced on an Illumina NEXTSEQ version 2 300-cycle kit to obtain paired-end sequences of 101 bp. Sequencing, raw data cleaning (adapter removal) and quality control were performed at Macrogen.

*Msp*I and *Bam*HI barcodes were removed from demultiplexed reads with TRIMMOMATIC version 0.39 (Bolger et al., 2014), with parameters HEADCROP, SLIDINGWINDOW:5:18 and MINLEN 75 bp. The quality of the trimmed sequence reads was assessed using FASTQC version 0.11.9 (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/). The *process_radtag* program from STACKS version 2.55 (Catchen et al., 2011, 2013) was used to ensure cut-sites remained intact and perform quality trimming with parameters -r (rescue barcodes and RAD-tags), -c (clean data, remove any read with an uncalled base) and -q (discard reads with low-quality scores). The reads were aligned to a reference genome of isolate P919-02W.7 (Ibarra Caballero et al., 2020; GCA_010367325.1) using BOWTIE2 version 2.3.2 (Langmead & Salzberg, 2012) with default settings. SAM files were sorted and converted to BAM files with SAMTOOLS version 1.7 (Li et al., 2009). The STACKS *gstacks* program was used to build a catalogue of loci from reference aligned/sorted reads and call single nucleotide polymorphisms (SNPs) with default settings. The STACKS *populations* program was used to filter and extract SNPs at polymorphic loci present in ≥80% of individuals across the whole data set. Three data sets were extracted with the *populations* program: *full* (all SNPs/locus recorded) for estimation of population genetics statistics, *thinned* (single SNP/locus recorded) for analyses of population genetic structure and *fixed* (a *phylip* format file containing SNPs differentially fixed within and variant among populations) for phylogenetic analysis.

2.3 | Genetic diversity and population structure

Number of private alleles (*A*), observed heterozygosity (H_O), expected heterozygosity (H_E), nucleotide diversity (π), inbreeding

coefficient (F_{IS}) and haplotype diversity (Hap) were calculated for locations with a minimum of 10 isolates with the *populations* program in STACKS using polymorphic (variant) loci in the *full* data set (Catchen et al., 2011, 2013).

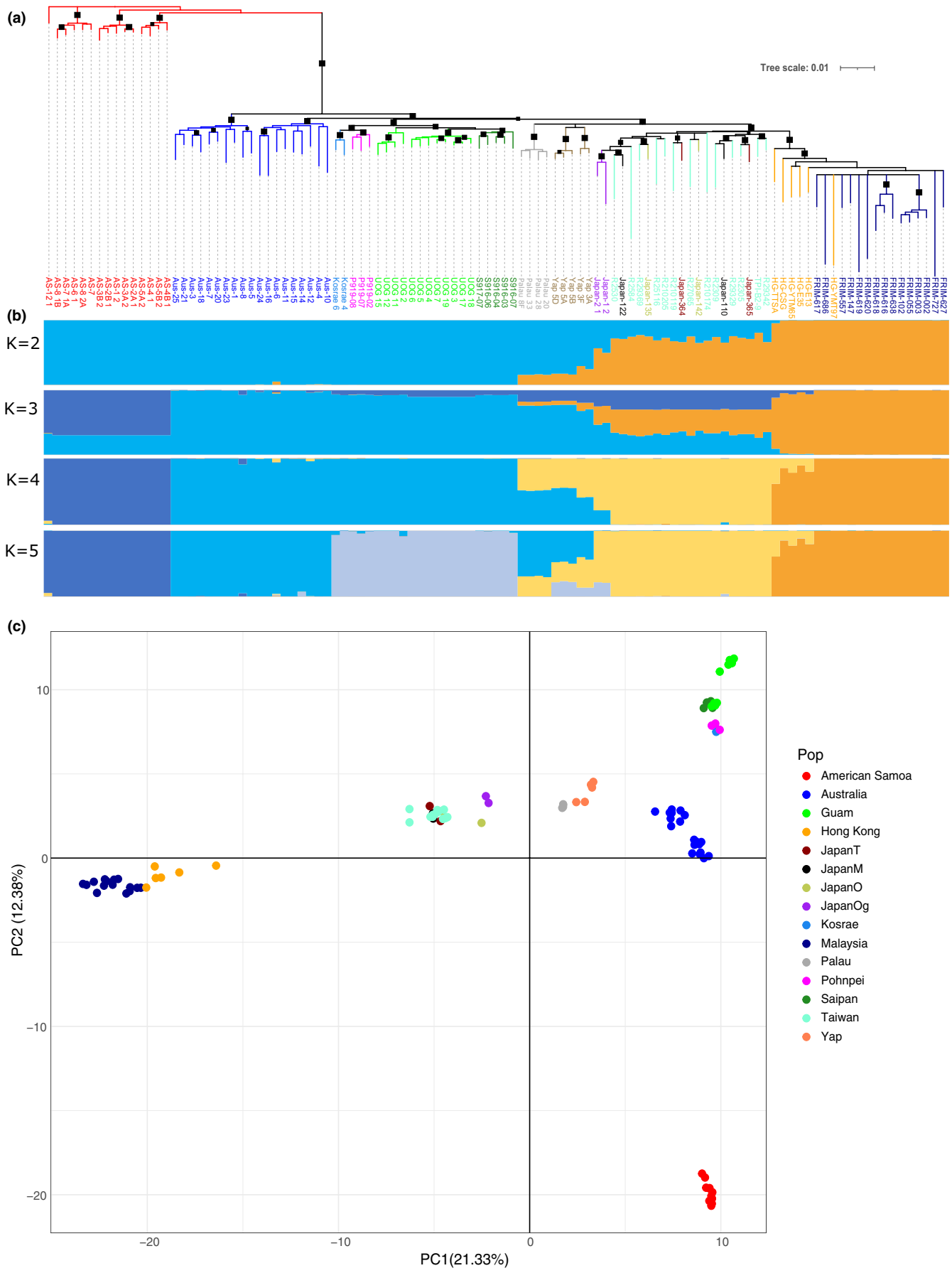
To assess genetic structure of populations of *P. noxius*, three independent approaches were implemented. First, relationships among sampled isolates were determined by building a maximum-likelihood phylogenetic tree with IQ-TREE version 2.0.3 (Nguyen et al., 2015) using the *fixed* data set. In this analysis, all loci were treated as a single partition, and the best-fit model TVMe+ASC+R4, selected by MODELFINDER (Kalyaanamoorthy et al., 2017), was used for the ascertainment bias correction (+ASC) model (Lewis, 2001). Tree branch support was assessed via 1000 ultrafast bootstrap replicates (Hoang et al., 2018) with hill-climbing nearest neighbour interchange search bootstrap optimization to reduce the risk of overestimating branch supports (-bnni) due to severe violations of the model. The concatenated tree was visualized with the ITOL online tool version 5.7 (itol.embl.de).

Using the *thinned* data set, isolates were assigned to genetic groups using Bayesian inference implemented in STRUCTURE version 2.3.4 (Pritchard et al., 2010) with Python utility StrAuto for analysis parallelization (Chhatre & Emerson, 2017) using a model with admixture (settings: 10,000 burn-in period and 100,000 Markov chain Monte Carlo iterations). The number of groups (*K*) varied from 1 to 12 and this analysis was replicated 10 times. The optimal number of *K* was determined using the methods of Evanno et al. (2005) in STRUCTURE HARVESTER (Earl & vonHoldt, 2012) (Figure S1). Membership coefficients at the optimal *K* across the 10 independent runs were averaged with CLUMPP version 1.1.2 (Jakobsson & Rosenberg, 2007). Using the same *thinned* data set, these data were also subjected to principal component analyses (PCAs) within the R package Adegenet version 2.1.3 (Jombart, 2008). The significance of population differentiation was assessed with pairwise F_{ST} calculations using ARLEQUIN version 3.5.1.2 (Excoffier & Lischer, 2010).

2.4 | Demographic analysis

We applied ABC with supervised machine learning implemented in DIYABC Random Forest (DIYABC-RF) version 1.0 (Collin et al., 2021) to further investigate the evolutionary history of *P. noxius* in the studied regions. Since ABC framework requires populations without continuous gene flow, we chose five nonadmixed, well-populated locations that represent five genetic groups identified in population structure analyses (see Results). These populations were American Samoa, Australia, Guam, Malaysia and Taiwan. Due to the lack of historical records on the invasion history of *P. noxius*, we incorporated the results of phylogenetic analysis (see Results, Figure 1a) as the starting point to construct hypothetical evolutionary scenarios for DIYABC-RF analyses. Since our phylogenetic tree is unrooted, the directionality of the topology cannot be determined. Hence, we initially determined relationships among some populations that were located within or closer to the continent (Malaysia and Taiwan, and the next clade in the phylogenetic tree: Guam). We performed five sequential DIYABC-RF analyses (Figure 2) that are described hereafter. For each topology, we used an approach implementing the classic invasion scenarios where each derived population passed through a bottleneck reflected in changes in effective population sizes (N_e, effective population sizes of each population; and N_b, effective population sizes of bottlenecks for each population). For analysis 1, we designed 28 scenarios where either of these three populations gave rise to (the origin of) the other two populations. Due to the large area of potential *P. noxius* distribution in eastern Asia and the Pacific Islands that was not covered in our study (Philippines, Indonesia, etc.), we also included scenarios that tested an unsampled (referred to here as "ghost") population as the origin, including one scenario indicating that all three populations originated from a ghost population (Figure S2, Supplementary text 1a). These scenarios were analysed individually, as well as combined in groups (Collin et al., 2021). For the combined group analysis, scenarios were combined into four groups with group 1 consisting of scenarios considering the Taiwan population as the origin (scenarios 11, 12, 17, 19, 25 and 27), group 2 consisting of scenarios considering the Malaysia population as the origin (scenarios 10, 13, 20, 21, 23 and 24), group 3 consisting of scenarios considering the Guam population as the origin (scenarios 14, 15, 16, 18, 22 and 26) and group 4 consisted of scenarios considering a ghost population as the origin (scenarios 1, 2, 3, 4, 5, 6, 7, 8, 9 and 28). To improve the robustness of the first analysis (see Results), we performed a second analysis whereby we constructed nine scenarios with the same three populations where a ghost population was considered as the origin (Figure S3a, Supplementary text 1b). Again, scenarios were analysed individually, as well as in groups, with group 1 consisting of scenarios without admixture (scenarios 1, 2, 4, 5, 7 and 8) and group 2 consisting of scenarios with admixture (scenarios 3, 6 and 9). Next, the best-supported topology from analyses 1 and 2 was used in analysis 3 to infer the probable origin of the population from American Samoa (opposite end in the phylogenetic tree, Figure 1a). In this analysis, six different scenarios were

FIGURE 1 Maximum-likelihood phylogenetic tree (a), STRUCTURE Bayesian clustering (b) and principal component analysis (PCA) (c) of *Phellinus noxius* populations from 15 locations. (a) Different colours represent samples from different locations reflected in the PCA key (c). Branches with ≥95% likelihood support are indicated with black squares. Branch support was calculated with 1000 ultrafast bootstrap replicates. (b) STRUCTURE results from $K = 2$–5; each vertical bar corresponds to the sample in the phylogenetic tree above STRUCTURE graphs. The height of the bar represents the posterior probability of sample assignment to colour-coded genetic groups. (c) PCA representing distinct genetic groups within the studied data set. Each dot indicates one individual. Percentages in parentheses represent the variance explained by each principal component (PC). For abbreviations of isolates and locations, see Table S1


(a)
Tree scale: 0.01
(b)
K=2
K=3
K=4
K=5
(c)
PC2 (12.38%)
PC1(21.33%)
10
0
−10
−20
−20
−10
0
10
Pop
American Samoa
Australia
Guam
Hong Kong
JapanT
JapanM
JapanO
JapanOg
Kosrae
Malaysia
Palau
Pohnpei
Saipan
Taiwan
Yap

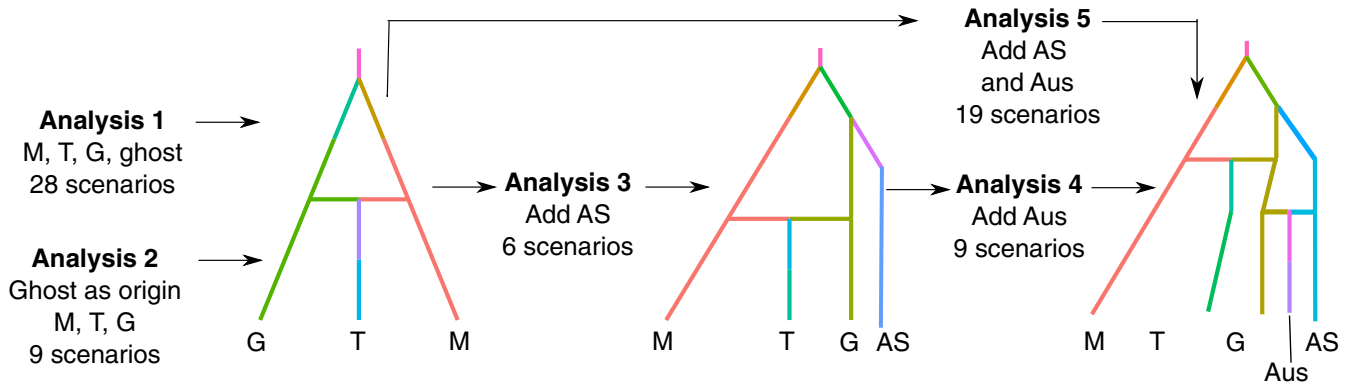


FIGURE 2 Flowchart of Approximate Bayesian computation and Random Forest analyses implemented in DIYABC-RF version 1.0. M—Malaysia, T—Taiwan, G—Guam, AS—American Samoa, Aus—Australia. Ghost – unsampled population. For analyses details, see Figures S2–S4, Supplementary text and main text

compared reflecting hypotheses that the American Samoa population diverged either from a ghost population or from the Guam population (Guam is the closest clade to American Samoa among Guam, Malaysia and Taiwan in the phylogenetic tree and closest population geographically among these three locations; Figure S3b, Supplementary text 1c). These scenarios were analysed individually, as well as in groups with group 1 consisting of scenarios in which the American Samoa population diverged from a ghost population (scenarios 1, 2, 3 and 4) and group 2 consisting of scenarios in which the American Samoa population diverged from the Guam population (scenarios 5 and 6). The best-supported topology from analysis 3 was then used to infer the probable origin of the Australia population (clade in the phylogenetic tree between American Samoa and Guam). Nine demographic scenarios were compared in analysis 4, reflecting hypotheses that the Australia population diverged either from Guam, American Samoa or a ghost population (Figure S3c, Supplementary text 1d). Scenarios were analysed individually as well as in groups for group 1 consisting of scenarios whereby the Australia population diverged from the Guam population (scenarios 1, 2 and 3), group 2 scenarios tested if the Australia population diverged from the American Samoa population (scenarios 4 and 5), group 3 scenarios tested whether the Australia population was the result of admixture (scenarios 6 and 7) and, finally, group 4 scenarios tested whether the Australia population diverged from a ghost population (scenarios 8 and 9). Lastly, in analysis 5, populations from American Samoa and Australia were both added to the best-supported topology identified in analyses 1 and 2 to allow building more flexible models of divergence events for American Samoa and Australia populations (e.g., relaxing the predefined divergence point of the American Samoa population in analysis 3 before adding the Australia population in analysis 4). A total of 19 scenarios were compared in these analyses (Figure S4, Supplementary text 1e). Again, scenarios were analysed individually as well as in groups. Group 1 consisted of scenarios testing whether the American Samoa and Australia populations originated from the Guam population (scenarios 1, 2,

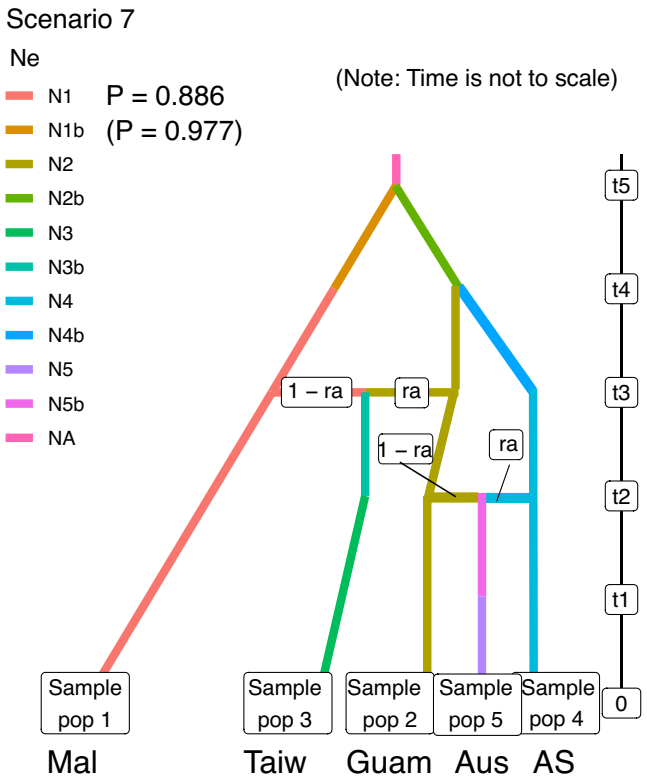


FIGURE 3 Winning hypothetical evolutionary scenario of the invasion routes of *Phellinus noxius* in eastern Asia, Australia and the Pacific Islands based on the results of five sequential DIYABC-RF analyses (see Supplementary materials for details about all tested scenarios in each analysis). P, posterior probability averaged across five independent runs for analysis 4. Values in parentheses are posterior probability of scenarios group containing the winning scenario (see Methods for details). Mal—Malaysia, Taiw—Taiwan, Aus—Australia, AS—American Samoa. For topology building, we used an approach implementing the classic invasion scenario where each derived population passed through a bottleneck ("founder effect"). N_e key represents effective population sizes for each population (N) and the effective population sizes of bottlenecks (N_b) for each population

TABLE 1 Genetic diversity indices of five *Phellinus noxius* populations with sample sizes >10 isolates

Population	*N*	*A*	H_O	H_E	π	F_{IS}	Hap
American Samoa	15	4490	0.043	0.044	0.046	0.009	0.462
Australia	19	8438	0.029	0.073	0.076	0.160	0.763
Malaysia	14	7985	0.034	0.078	0.081	0.168	0.829
Taiwan	13	4782	0.034	0.090	0.094	0.201	0.948
Guam	12	618	0.048	0.043	0.046	0.002	0.459

Abbreviations: *N*, number of isolates; *A*, number of private alleles; H_O, observed heterozygosity; H_E, expected heterozygosity; π, nucleotide diversity; F_{IS}, inbreeding coefficient; Hap, haplotype diversity. For population information, see Table S1.

3, 10 and 17), group 2 scenarios tested if the American Samoa population originated from the Australia population or vice versa (scenarios 4, 5, 11, 12 and 16), group 3 scenarios were conducted considering admixture (scenarios 6, 7, 13, 14 and 15) and group 4 scenarios contained a ghost population (scenarios 8, 9, 18 and 19).

Prior-scenario checking was performed with PCA representing the simulated data set from the training and observed data set as described by Collin et al. (2021). According to Collin et al. (2021), the observed data set must be reasonably located within the cloud of simulated data sets on the PCA plot. In all five analyses, 3000 training data sets were generated for each tested scenario with uniform prior distribution of parameters and divergence times being constructed in descending order from $t_5 > t_4 > t_3 > t_2 > t_1$ to enforce topological congruence. Priors were set to 100–10,000 for effective population sizes (*N*) and 10–10,000 for divergence times of populations (*t*). The choice of scenario and estimation of posterior probability of the best-supported scenario were performed using a random forest algorithm module in DIYABC-RF, which produced 1000 trees for each of the analyses (Collin et al., 2021). To assess whether the forest produced enough trees, error/accuracy metrics were plotted as a function of the number of trees in the forest (Collin et al., 2021). The choice of scenario and the quality of predictions in DIYABC-RF analyses were assessed by the global and local error/accuracy indices, simultaneous projection of data sets of the training database, and observed data set on the first linear discriminant analysis (LDA) axes, as suggested by Collin et al. (2021). Five independent runs were performed for each analysis.

To estimate divergence times among the populations, a training data set included 20,000 data sets simulated under scenario 7 of analyses 4 and 5 (Figure 3) with axes of a partial least squares (PLS) regression analysis. For each *t* parameter, inferred point estimates and computed global and local accuracy indices, which correspond to global and local normalized mean absolute error (NMAE) using out-of-bag estimators from a sample of 10,000 data sets, were randomly selected in the training set, with the mean and the median as the point estimates (Collin et al., 2021). We determined that 20,000 data sets in the training set were sufficient for parameter estimation by evaluating the stability of the global accuracy indices on subsets of 19,000, 18,000 and 17,000 data sets of the training set. Additionally, we simulated 100,000 data sets under the same scenario and determined that global accuracy indices remained similar (e.g., for parameter t_4, global NMAE at 100,000 and 20,000 was 0.089 and 0.093, respectively). The number of trees in the random forest was set to 1000. For each parameter, five independent runs of random forest analyses were performed with the same training set.

2.5 | Admixture and migration

Relationships among sampled populations and possible admixture were investigated with TREEMIX version 1.13 (Pickrell & Pritchard, 2012). The *f* coefficient was used to identify the information contribution from each migration edge that was added to the tree. Up to 20 possible migration edges were tested. The migration edges were added to the tree until 99.8% of the variance in ancestry was explained by the model. Five independent runs of TREEMIX were conducted. To ensure robustness of modelled migration events, *f*3 and *f*4 statistics were computed from 15 *P. noxius* populations using threepop and fourpop programs from TREEMIX (Patterson et al., 2012). The three- and four-population tests are less-parameterized methods to test for treeness (Reich et al., 2009). For the three-population test (*f*3(A;B,C)), significantly negative values of *f*3 statistics implies that population A is admixed. For the four-population test (*f*4(A,B;C,D)), a significantly nonzero value indicates gene flow in the tree (Reich et al., 2009). Since some of the populations had small sample sizes (e.g., populations of islands of Japan with two isolates compared to populations from Australia, American Samoa, Taiwan, Malaysia and Guam with >10 isolates), we evaluated whether uneven sample sizes could bias estimates of allelic variation between populations in migration analyses following the approach from Soghigian et al. (2020). For this evaluation, we resampled, without replacement, each population of *P. noxius* with two individuals, using the R function dplyr. For each resampling, we conducted the *f*3 and *f*4 tests as described above. We resampled our populations 100 times, ran TREEMIX on each resampled data set, summarized the resampled mean estimates of the *f*3 and *f*4 statistics, and compared them to the results from the main TREEMIX analysis.

3 | RESULTS

Following the initial cleaning, the total length of each locus was 75 bp. The cleaned *Phellinus noxius* data set contained 272,316 genotyped

TABLE 2 Results of DIYABC-RF for scenario choice in five sequential analyses detailed in Figures S2–S4

Analysis	Type of treatment	Global error rate	Local error rate	Votes scenario: 1	2	3	4	5	6
				15	16	17	18	19	20
1	Groups	0.311 (0.002)	0.19 (0.012)	108 (2.07)	134 (5.03)	41.8 (4.21)	716 (8.9)		
	Individual	0.498 (0.007)	0.46 (0.004)	60.2 (3.27)	0.2 (0.45)	8.4 (1.52)	34.8 (5.26)	6 (1.87)	422 (14.7)
				79.6 (8.79)	2.6 (0.89)	0 (0.00)	2.4 (1.52)	48.6 (7.37)	0.8 (1.10)
2	Groups	0.174 (0.001)	0.099 (0.027)	302 (12.6)	698 (12.6)				
	Individual	0.340 (0.001)	0.144 (0.026)	53.4 (8.91)	0.6 (0.55)	9 (2.35)	40 (4.06)	3.2 (0.84)	874 (11.7)
3	Groups	0.076 (0.001)	0.038 (0.006)	54.8 (5.81)	945 (5.81)				
	Individual	0.371 (0.001)	0.092 (0.011)	177 (18.6)	44.4 (7.13)	54.2 (7.89)	23.6 (6.11)	507 (18.71)	194 (10.9)
4	Groups	0.156 (0.001)	0.023 (0.003)	409 (16.7)	20.8 (3.49)	569 (16.3)	1 (1.22)		
	All individual	0.300 (0.001)	0.114 (0.011)	176 (4.44)	4.4 (1.52)	222 (9.45)	15 (5.34)	15.2 (5.4)	151 (9.63)
5	Groups	0.209 (0.001)	0.101 (0.009)	337 (22.9)	48.4 (4.83)	599 (22.4)	16 (4.9)		
	Individual	0.396 (0.001)	0.221 (0.017)	171 (9.58)	3.4 (2.07)	206 (7.4)	10.2 (3.83)	10.2 (4.02)	171 (7.22)
				14.6 (4.22)	28 (4.8)	6.6 (1.52)	0.2 (0.45)	46 (5.24)	

Note: RF analyses included 3000 simulated data sets per scenario and the number of trees was set to 1000. Global (prior) and local (posterior) error rates were estimated using out-of-bag estimators from a sample of 10,000 data randomly chosen in each training set. The values in the table are mean values over the five replicates for each of the analyses with standard deviation over the five replicates provided in parentheses for each metric. Details about the groups of the scenarios are provided in the Materials and methods.

loci, with the final data sets containing 7554 polymorphic loci, after filtering. Mean effective coverage per sample was 42.2× (*SD* = 8.5×, min = 11.2×, max = 65.9×). The *full*, *fixed* and *thinned* data sets contained 75,782, 11,600 and 7554 SNPs, respectively.

3.1 | Genetic diversity of populations

Genetic diversity of *P. noxius* fluctuated across locations, with overall higher diversity (π) in some populations located within or closer to the Asian continent (Malaysia, Taiwan and Australia) (Table 1). For all polymorphic loci in populations from Australia, Malaysia and Taiwan, levels of observed heterozygosity (H_O) were ~1.5–2.5 times lower than levels of expected heterozygosity (H_E), while having the highest levels of inbreeding coefficient (F_{IS}), suggesting assortative mating and/or the presence of population structure ("Wahlund effect") in these locations. Populations from Australia, Malaysia and Taiwan had the highest haplotype diversity (Hap), whereas the population from Guam had the least number of private alleles (*A*) (Table 1).

3.2 | Population structure

Based on the results of phylogenetic analysis, *P. noxius* isolates from American Samoa formed a well-supported (ultrafast bootstrap [UFB] = 100%) clade distinct from the other isolates (Figure 1a). The American Samoa clade was further split into three clusters (UFB = 100% for each) suggesting the presence of genetic structure within the island. Isolates from Australia formed two distinct clades (UFB = 100% for each), and isolates from each island location within the Mariana Islands and Micronesia (Kosrae, Pohnpei, Guam, Saipan, Palau and Yap) also formed distinct clades. Two clusters were also identified within the Guam population, with one closely related to the Saipan population. Isolates sampled from Yap and Palau showed closer relationships to populations from Japan and Taiwan, compared to populations from other islands of Micronesia. Isolates from the four islands of Japan clustered within the clade containing isolates from Taiwan, suggesting gene flow or shared ancestry among these locations. Isolates from the Japanese island of Ogasawara (JapanOg; Japan-1 2 and Japan-2 1) formed a distinct cluster within the Taiwan–Japan clade (Figure 1a). Finally, isolates from Hong Kong and Malaysia formed another distinct clade (UFB = 100%), signalling possible gene flow or shared ancestry between these two locations.

Based on the results of the Bayesian clustering analyses, the model with two genetic groups (*K* = 2) produced the highest likelihood score, and the model with five genetic groups (*K* = 5) had the second highest likelihood (Figure 1b; Figure S1). In the *K* = 2 results, *P. noxius* isolates clustered into generally distinct western and eastern genetic groups. Isolates from American Samoa, Australia, Saipan, Guam, Pohnpei and Kosrae were assigned to the eastern group with almost no admixture, while isolates from Malaysia and Hong Kong were clearly assigned to the western group, with only one isolate from Hong Kong showing 4% admixture with the eastern group. The remaining isolates within this western group were represented by one dominant genetic group with varying degrees of admixture. Isolates from Yap and Palau showed 83% group assignment to the eastern group and 17% group assignment to the western group. The western genetic group was dominated by isolates from Japan

7	8	9	10	11	12	13	14	
21	22	23	24	25	26	27	28	Posterior probability
								0.806 [group 4] (0.013)
9.2 (2.95)	0.8 (0.84)	2 (1.41)	19.6 (4.93)	0.4 (0.55)	7.2 (1.30)	10 (2.24)	0.8 (0.45)	0.538 [scen. 6] (0.004)
3.4 (2.30)	3.4 (2.79)	2 (2.00)	185 (18.70)	7.4 (2.19)	0.2 (0.45)	2.8 (1.10)	80.6 (8.41)	
								0.901 [group 2] (0.027)
14.4 (3.21)	0.6 (0.89)	4.8 (2.59)						0.876 [scen. 6] (0.032)
								0.962 [group 2] (0.006)
								0.908 [scen. 5] (0.011)
								0.977 [group 3] (0.003)
413 (4.83)	1.6 (0.55)	1.4 (1.14)						0.886 [scen. 7] (0.011)
								0.901 [group 3] (0.014)
288 (8.41)	1.4 (1.67)	0.4 (0.55)	3.2 (1.48)	14.4 (2.3)	17.6 (1.14)	4.4 (2.3)	4.8 (3.42)	0.779 [scen. 7] (0.017)

Taketomi (JapanT; Japan-364 and Japan-365), Japan Miyakojima (JapanM; Japan-110 and Japan-122) and Japan Okinawa (JapanO; Japan-135 and Japan-142), with 65%–70% group assignment, while isolates from JapanOg were assigned approximately equally to both eastern and western groups. Taiwan isolates had ~30% admixture with the eastern group. When *K* was set to 3, isolates from American Samoa showed admixture with other eastern populations (Australia and Micronesia), populations from all Japanese islands, Taiwan and Hong Kong. Isolates from Yap, Palau, Japan and Taiwan, and one isolate from Hong Kong showed admixture with all three groups at varying levels. When $K = 4$, isolates from American Samoa formed a distinct group with only one individual showing 5% admixture with some western populations. Isolates from Australia, Guam, Pohnpei, Kosrae and Saipan formed a distinct group with low-level admixture with other groups. Isolates from Yap, Palau and JapanOg showed admixture with populations from Australia–Micronesia and Japan–Taiwan. Hong Kong isolates primarily grouped with the Malaysia population, forming a genetic group with signatures of admixture from the Taiwan–Japan group. Isolates from Malaysia formed a distinct group, with no admixture. With $K = 5$, isolates from Australia and Micronesia were further split into two groups with little admixture, except for isolates from Yap and Palau that showed admixture with Australia, Micronesia and Japan–Taiwan populations. Approximately 20% of JapanOg isolates were assigned to the Micronesia group.

First (21.33%) and second (12.38%) principal components in the PCA separated the *P. noxius* isolates into distinct clusters (Figure 1c), confirming results from the phylogenetic analysis (Figure 1a) and STRUCTURE groupings (Figure 1b). Pairwise F_{ST} showed significant pairwise differentiation among the five main genetic groups (pairwise F_{ST} ranged from 0.26 to 0.63, with all *p*-values <.0001, Table S2).

3.3 | Demographic history of *P. noxius* populations

Prior-scenario checking with PCA indicated that the observed data set was located within clouds of simulated data sets. This suggested that defined conditions for each of the five analyses were suitable for the random forest analyses (Figure S5).

3.3.1 | Analyses 1 and 2: An unsampled (ghost) population gave rise to populations from Malaysia and Guam, and the Taiwan population was the result of their admixture

In the DIYABC-RF analysis 1 that tested 28 hypothetical evolutionary scenarios (Figure S2), group 4 (consisting of scenarios considering a ghost population as the origin [scenarios 1, 2, 3, 4, 5, 6, 7, 8, 9 and 28]) and scenario 6 (when all scenarios were considered individually) were selected as the best fit, based on the classification votes and posterior probabilities (Table 2; Figure S2) with a mean posterior probability of 0.538 and global and local errors of 0.498 and 0.460, respectively. When the scenarios were analysed in groups, posterior probability and global/local errors improved, with mean posterior probability of 0.806 accompanied by global and local errors of 0.311 and 0.190, respectively (Table 2). The projections of data sets from the training set on the first two LDA axes indicated low power to discriminate among the tested scenarios in the analyses (Figure S6a,b). To improve prediction quality, we ran analysis 2 consisting of nine hypothetical evolutionary scenarios considering a ghost population as the origin (based on the results of grouped analysis 1 in which groups with scenarios containing a ghost population as the origin

TABLE 3 Results for DIYABC-RF estimation of population divergence times under scenario 7 detailed in Figure 3

Parameter	Posterior point estimates of:			Global (prior) NMAE computed from:		Local (posterior) NMAE computed from:	
	Mean	Median	90% CI	Mean	Median	Mean	Median
*t*2	3412.626 (40.779)	3267.680 (28.614)	1623.400–5671.132 (64.972–124.368)	0.2123 (0.0004)	0.1956 (0.0030)	0.2641 (0.0044)	0.2378 (0.0138)
*t*3	3487.788 (52.052)	3317.380 (65.913)	1778.398–5803.462 (43.104–65.243)	0.1752 (0.0001)	0.1675 (0.0008)	0.2189 (0.0086)	0.2081 (0.0118)
*t*4	4985.848 (18.361)	4808.980 (44.769)	2389.390–7971.990 (66.649–93.213)	0.1499 (0.0001)	0.1465 (0.0009)	0.3328 (0.0081)	0.3133 (0.0314)
*t*5	7583.080 (49.039)	7688.000 (70.1106)	4871.094–9772.200 (114.648–19.447)	0.1082 (0.0000)	0.1067 (0.0006)	0.1777 (0.0074)	0.1845 (0.0075)

Note: RF analysis included 20,000 simulated data sets and the number of trees was set to 1000. Global (prior) and local (posterior) error rates were estimated using out-of-bag estimators from a sample of 10,000 data randomly chosen in a training set. The values in the table are mean values over the five replicates for each of the analyses with standard deviation over the five replicates provided in parentheses for each metric. CI, credibility interval; NMAE, normalized mean absolute error.

received the best support). The results of analysis 2, when analysing scenarios individually, indicated that scenario 6 was the best fit based on the classification votes and posterior probability (Table 2; Figure S3a) with a mean posterior probability of 0.876. The global and local errors were 0.340 and 0.144, respectively. When scenarios were analysed in groups, group 2 (consisting of scenarios with admixture [scenarios 3, 6 and 9]) was the best fit with a posterior probability of 0.901 accompanied by global and local errors of 0.174 and 0.099, respectively. The projection of data sets from the training set on the first two LDA axes indicated substantial power to discriminate among the tested scenarios (Figure S6c—the observed data set was clearly located within the cloud of scenario 6, and Figure S6d—the observed data set was located within the area of group 2 in the density plot).

3.3.2 | Analysis 3: Population from American Samoa originated from the Guam population prior to the origin of the Taiwan population

In analysis 3, group 2 (containing scenarios 5 and 6, Figure S3b) and scenario 5 (when all scenarios were considered separately) had the highest classification votes and posterior probabilities, with mean posterior probabilities of 0.962 and 0.908 for group 2 and scenario 5, respectively (Table 2; Figure S3b). Global and local errors were 0.076 and 0.038, and 0.371 and 0.092 for grouped and individual analyses, respectively (Table 2). The projections of data sets from the training set on the first two LDA axes indicated overall substantial power to discriminate among the tested scenarios, with stronger prediction quality for grouped analysis (Figure S6f—the observed data set was located within the area of group 2 in the density plot) compared to individual analyses (Figure S6e—the observed data set was located within reasonable proximity to clouds of scenario 5).

3.3.3 | Analyses 4 and 5: The population from Australia was the result of admixture of the Guam and American Samoa populations

In analysis 4, group 3 (containing scenarios 6 and 7) and scenario 7 (Figure S3c) were chosen as the best fit with posterior probabilities of 0.977 and 0.886, respectively. The global and local errors were 0.156 and 0.023, and 0.300 and 0.114 for grouped and individual analyses, respectively (Table 2). The projections of data sets from the training set on the first two LDA axes indicated overall substantial power to discriminate among tested scenarios (Figure S6g—the observed data set located within the cloud of scenario 7 when scenarios were analysed individually; and Figure S6h—the observed data set located within the cloud of group 3 when scenarios were analysed in groups). Finally, in analysis 5, group 3 (containing scenarios 6, 7, 13, 14 and 15) and scenario 7 (Figure S4) were chosen as the best fit with posterior

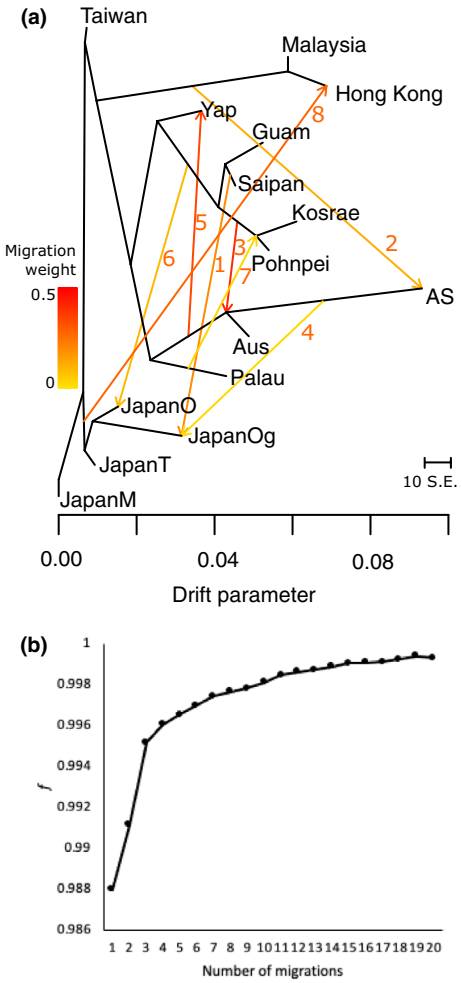


FIGURE 4 (a) Phylogenetic network inferred by TREEMIX of the relationships among *Phellinus noxius* populations with eight migration edges that explain 99.8% of the variance in the model. The migration edges are numbered from 1 to 8, reflecting the order in which they were added to the model. (b) The *f* index representing the fraction of the variance in the sample covariance matrix accounted for by the model covariance matrix as a function of the number of modelled migration events. For population abbreviations, see Table S1

probabilities of 0.901 and 0.779, respectively. The global and local errors were 0.209 and 0.101, and 0.396 and 0.221 for grouped and individual analyses, respectively (Table 2). The projections of data sets from the training set on the first two LDA axes indicated overall substantial power to discriminate among tested scenarios (Figure S6i—the observed data set located within reasonable proximity to the cloud of scenario 7 when scenarios were analysed individually; and Figure S6j—the observed data set located within the cloud of group 3 when scenarios were analysed in groups).

In the random forest analyses, 1000 trees were determined to be sufficient to ensure a stable estimation of the global accuracy metrics by plotting error/accuracy metrics as a function of the number of trees in the forest (Figure S7). Increasing the number of trees to 2000 did not improve the prediction error (results not shown).

Overall, results from these five sequential demographic analyses and estimation of population divergence times suggested a long history of invasion of *P. noxius* in the studied regions. In particular, Malaysia and Guam populations were predicted to have diverged from a ghost population 7583 generations ago (confidence interval [CI] 4871.094–9772.200). The Guam *P. noxius* population, in turn, is ancestral to the population from American Samoa, which was estimated to have arisen 4985 generations ago (CI 2389.390–7971.990) (Table 3, Figures 3 and 5a). The Taiwan population appears to be the result of admixture between Malaysia and Guam populations that is predicted to have occurred roughly 3487 generations ago (CI 1778.398–5803.462) (Table 3, Figures 3 and 5a), and the Australia population was the result of admixture among American Samoan and Guam populations that probably occurred 3412 generations ago (CI 1623.400–5671.132) (Table 3, Figures 3 and 5a). According to these results, the admixture event leading to the establishment of the Taiwan population of *P. noxius* occurred slightly before the admixture event leading to the establishment of the population in Australia.

3.4 | Migration pathways of *P. noxius*

The maximum = likelihood tree models produced similar topologies with and without migration (Figure 4a; Figure S8). The tree model allowing for migration clustered geographical populations into a few main branches that were consistent with the phylogenetic, STRUCTURE and PCA analyses of *P. noxius* (Figure 4a). The four populations from the Japanese islands were grouped together, with Taiwan as a sister group. Populations from Malaysia and Hong Kong formed a second branch, while the third branch comprised populations from Micronesia (except Palau). The Australia and American Samoa populations formed a fourth branch, with American Samoa having the largest genetic drift, suggesting its divergence from the Australia population.

The first three migration edges (gene flow events) accounted for more than half of the total variance explained by the model based on *f*-statistics (Figure 4b), with the first migration event explaining 98.8% of the variance in relatedness among *P. noxius* populations. At migration events (m) = 8, the model explained 99.8% of the variance in relatedness among populations, which accounted for 85.4% of the total variance explained by the model (Figure 4b). This model also generated improved residuals among populations (Figure S9). Overall, TREEMIX analyses suggested a complex evolutionary history and migration patterns among *P. noxius* populations from eastern Asia, Australia and the Pacific Islands (Figures 4 and 5b), and confirmed several possible admixture events detected by STRUCTURE (Figure 1b). These results indicate that the JapanOg population is the result of admixture between Japan–Taiwan, Micronesia (Saipan) and American Samoa, with 20% and 3% of its ancestry coming from Saipan and American Samoa, respectively. JapanO traces about 10% of its ancestry to populations from Micronesia. TREEMIX indicated that the Hong Kong

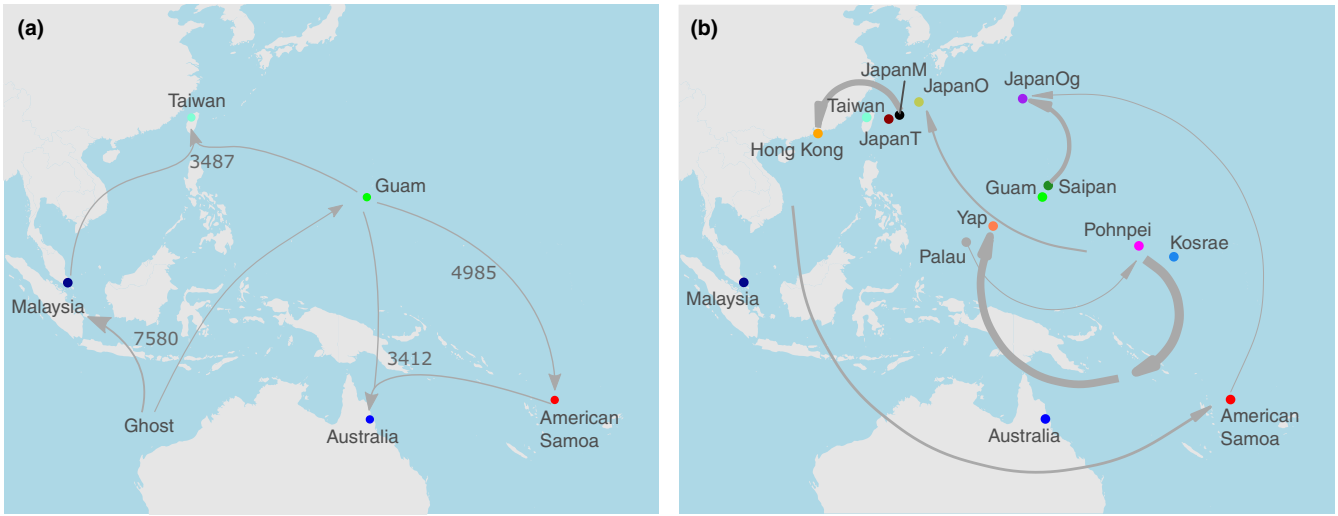


FIGURE 5 (a) Invasion routes of *Phellinus noxius* determined with DIYABC-RF analyses. The analyses were performed among populations representing five genetic groups identified in this study. Numbers next to the lines represent number of generations. "Ghost" indicates the ghost (unsampled) population with unknown location that gave origin to populations from Malaysia and Guam. (b) Migration among *P. noxius* populations sampled from 15 locations detected with TREEMIX. The width of the lines reflects the percentage of ancestry (migration weight) in the recipient from the donor population (see Figure 4 and main text). Migration to American Samoa (AS) traces back to the ancestral origins in Malaysia and/or Hong Kong, migration to Yap traces back to the ancestral origins in Australia and/or AS, migration to the ancestor of Australia and AS traces back to the ancestral origins in Pohnpei and Kosrae, and migration to Hong Kong traces back to the ancestral populations sampled from islands of Japan

population is the result of admixture from Malaysia and Japan–Taiwan populations, tracing 28% of its ancestry to Japan–Taiwan populations. The analysis also inferred migration from Micronesia to more southern populations of Australia and American Samoa migration weight (w = 41%), as well as migration in another direction, from Australia–American Samoa to Yap (w = 35%). About 12% of American Samoa ancestry also traces to the Malaysia–Hong Kong group; by contrast, the Palau population was placed separately from other populations from Micronesia, with ~4% of its ancestry tracing to populations from Kosrae and Pohnpei, suggesting the presence of gene flow among islands of Micronesia (Figure 4a). The robustness of migration edges inferred by TREEMIX was confirmed with the significant Z-scores using *f*3 and *f*4 population tests for *P. noxius* (Table S3). The results of *f*3 and *f*4 population tests with a data set where populations were subsampled to include only two individuals (and resampled 100 times) yielded qualitatively similar conclusions (Table S4), confirming eight admixture events among populations detected by TREEMIX. It is worth noting that a larger sample size in terms of SNPs and individuals per population used in the main analyses had higher power based on *f*4 and *f*3 statistics and *Z*-scores.

4 | DISCUSSION

As the human population intrudes into wild ecosystems and globalization increases, rapid declines of biodiversity and numerous outbreaks of emerging infectious diseases have been threatening both plants and animals (Alroy, 2017; Schmeller et al., 2020). Characterizing processes leading to pathogen emergence is critical to understand pathogen evolution, establish effective management strategies and forecast/prevent future epidemics/epiphytotics. In this paper, we provide a novel example showing the integrated application of population genetics and an ABC model-based framework with supervised machine learning to reveal the possible origins of an emerging pathogen with limited available data, focusing on a recently emerged tree-root pathogen, *Phellinus noxius*, in eastern Asia, Australia and the Pacific Islands. Our analyses revealed a complex colonization history of *P. noxius* that began over 7000 generations ago. We identified five distinct genetic groups of *P. noxius* from Malaysia/Hong Kong, Taiwan/Japan, Micronesia (Guam, Saipan, Pohnpei, Palau, Yap and Kosrae), Australia and American Samoa, with the signatures of multiple migration events among these locations. ABC analyses indicated initial spread of *P. noxius* from a ghost population that gave rise to populations in Malaysia and the Pacific Islands (Guam and American Samoa), subsequently giving rise to populations in Taiwan and Australia. According to the ABC analyses, the major spread of *P. noxius* occurred thousands of generations ago, contradicting previous hypotheses that this pathogen had been introduced to many areas within the last century.

4.1 | Hypothesis of recent introduction, high polymorphism and geographically structured populations

Over the last several decades, *P. noxius* has caused extensive damage to diverse tree species in tropical and subtropical regions (Abe

et al., 1995; Ann et al., 2002; Bolland, 1984; Hodges & Tenorio, 1984; Kobayashi et al., 1991; Nandris et al., 1987a, 1987b; Neil, 1986; Tai, 1979). Because *P. noxius* and the disease it causes (brown root rot disease) have emerged over recent decades, it was hypothesized to have been recently introduced to many areas from an unknown centre of origin (Ann et al., 2002; Hodges & Tenorio, 1984; Sahashi et al., 2007; Tsai et al., 2017). Recent population genetics studies with microsatellites and SNP markers identified highly polymorphic *P. noxius* populations from Japan and Taiwan (Akiba et al., 2015; Chung et al., 2015, 2017) that seem to contradict hypotheses of recent introductions of *P. noxius* to these regions. Similarly, analysed populations in this study also had high levels of polymorphisms, with haplotype diversities of between 0.459 and 0.948 (Table 1). High levels of polymorphisms are usually detected in populations with a long evolutionary history and/or that were established from multiple distinct source populations (Guillemaud et al., 2010; Kolbe et al., 2004). Some populations also exhibit large numbers of private alleles (Table 1), suggesting the presence of gene flow barriers among these locations. Populations sampled in Guam and American Samoa had lower genetic diversity compared to the other three analysed populations (Malaysia, Taiwan and Australia). These two islands are geographically distant from the Asian continent, and even if these populations have long evolutionary histories, geographical isolation may have influenced recovery from the "founder effect," which typically results in lower genetic diversity. Another plausible explanation of lower genetic diversity in these populations could be attributed to different frequencies of sexual reproduction occurring in these locations compared to the other populations. Additional studies are required to assess these factors.

Over time, restricted gene flow facilitates population differentiation (Slatkin, 1987). Five genetically distinct and geographically distributed groups of *P. noxius* were identified using Bayesian population structure analyses, PCA and phylogenetic analyses (Figure 1). Interestingly, populations from American Samoa and Malaysia showed almost no admixture with other genetic groups, suggesting limited or absent ongoing gene flow, and isolates from American Samoa formed the most genetically distinct clade in the phylogenetic tree (Figure 1a) and TREEMIX graph (Figure 4a), concurring with previous phylogenetic analyses with protein-coding loci (Stewart et al., 2020) and reflecting its isolation from other populations.

4.2 | Initial routes and timing of colonization

Implementation of sequential ABC analyses with supervised machine learning in combination with population structure analyses allowed us to reconstruct the most probable evolutionary scenario of initial pathogen colonization without prior knowledge about the place of origin or historical data on its spread. It is important to note that our data set, which was obtained from areas were *P. noxius* was noticeably emerging, covers only a portion of the total geographical distribution of this pathogen. Collections of *P. noxius* were not obtained from areas within Southeast Asia and Oceania (e.g., Singapore, Philippines, Indonesia, Fiji and others) where this pathogen probably occurs. With our limited data set, we cannot determine the centre of origin of *P. noxius*, and additional populations from other geographical areas could potentially affect the divergence topology of the five populations analysed here. Nevertheless, our data set provides meaningful information about the evolutionary history of *P. noxius*, the cause of brown root rot disease, in the regions covered in our current study. ABC Random Forest (RF) analyses support initial spread of *P. noxius* from a ghost population to Malaysia and the Pacific Islands (Guam and American Samoa), with subsequent spread to Taiwan and Australia. These findings are partially congruent with human settlement on the Pacific Islands (Matisso-Smith, 2015) and/or the distribution of wind-dispersed species to the Pacific Islands (Ito, 1998). ABC RF analyses generated overall low global and local error rates, as well as low standard deviations among replicate analyses (Tables 2 and 3), suggesting the accuracy of our inferences. Global and local error rates were consistently lower when scenarios were analysed in groups, which is consistent with Collin et al. (2021). While analyses 1 and 2 testing the divergence of Malaysia, Taiwan and Guam populations, and analyses 3, 4 and 5 testing the divergence of American Samoa and Australia populations, differed in prediction quality, they produced consistent results on the choice of scenario (scenario 6 in analyses 1 and 2, and scenario 7 in analyses 4 and 5). On the parameter estimation (*t*) stage, ABC RF analyses demonstrated that initial spread of *P. noxius* probably occurred thousands of generations ago, and probably moved to Malaysia and Guam more than 7500 generations ago, probably reaching Taiwan and Australia more than 3400 generations ago. It was shown previously that migration can affect the estimation of divergence times in coalescent analysis (Gronau et al., 2011). DIYABC-RF does not allow migration to be incorporated in the model, and hence in a probable case of multiple introductions, divergence times of *P. noxius* populations may deviate slightly from those estimated with DIYABC-RF, as was shown with a coalescent model allowing for migration in analyses of the demographic history of human populations (Gronau et al., 2011). *P. noxius*, similar to other basidiomycete tree pathogens, can reproduce both sexually, via basidiospore dissemination and infection of hosts via wounds, and asexually, infecting trees via mycelial spread during root-to-root contact. Previous studies have shown that *P. noxius* can also survive in dead roots for at least a decade (Chang, 1996). Basidiocarps of *P. noxius* are observed at variable frequencies in different locations (Akiba et al., 2015; Ann, Lee, & Tsai, 1999; Chung et al., 2015; Hodges & Tenorio, 1984), and the abundance of basidiocarps depends on the environmental conditions (Bolland, 1984; Chung et al., 2015), similar to other closely related species (Lim et al., 2007). Irregular sexual cycles make it impossible to translate the number of generations determined with ABC RF into an exact number of years since the initial colonization events; however, the number of generations can provide an estimate of the relative timing of these events. Even if we consider one or several sexual cycles per year, *P. noxius* must have been present in eastern Asia, Australia and the Pacific Islands for millennia.

4.3 | Complex migration history

We identified eight possible migration events using genome-wide allele frequency data and a Gaussian approximation to genetic drift (as described in Pickrell & Pritchard, 2012; Figure 4), supporting some of the admixture results in the population structure analyses (Figure 1b). These results reflect a complex migration history of *P. noxius* within this global region, which is not unusual for a plant pathogen (Moller & Stukenbrock, 2017). Even though some populations in our data set were represented with low sample sizes (e.g., islands of Japan and some islands in Micronesia), nonparametric *f*3 and *f*4 tests supported the presence of admixture among locations identified by TREEMIX with data sets representing all samples and resampling of two individuals per population (Tables S3 and S4). Because the relative *f*3 and *f*4 statistics from the resampled analysis were slightly lower than those estimated with the full data set, it remains possible that some potential migration events may have been missed due to underrepresented populations. Nevertheless, both analyses produced consistent qualitative results, supporting eight migration events detected by TREEMIX. It was suggested that TREEMIX performs best with larger sample sizes (Pickrell & Pritchard, 2012), but our results indicate that, probably due to a relatively large number of SNPs, even small sample sizes can detect admixture that is consistent with the results from the full data set. Similar conclusions on the robustness of TREEMIX analysis with small population sample sizes were reported previously in demographic analysis of yellow fever mosquito, *Aedes aegypti* (Soghigian et al., 2020). In population structure analyses, the isolates from islands of Japan showed signatures of admixture with Micronesia populations, and TREEMIX analyses indicated migration events from Saipan to JapanOg and from Pohnpei to JapanO. This supports a previously stated hypothesis of potential northward movement of *P. noxius* from the Mariana Islands to JapanOg by typhoons or other factors (Akiba et al., 2015). All other migration events are reported for the first time in this study, which can be explained by other hypotheses suggesting that in addition to severe weather events, such as typhoons, complex migration patterns of *P. noxius* in the region may be the result of human activities, including migration and trade among islands (Firth, 2007; Matisso-Smith, 2015), movement of potentially infested wooden material (Hodges & Tenorio, 1984) and cultural traditions (Cannon et al., 2022). Although each of these factors probably plays a role in pathogen spread and may be a contributing factor driving emergence in multiple areas within last few decades, they do not appear to explain the origin of *P. noxius* in these diverse regions. Until recently, no sanitary measures against *P. noxius* were implemented (Cannon et al., 2022), so it is plausible that further pathogen spread could have coincided with movement of equipment, wood, planting stock, etc. Note, however, that the presence of continuous gene flow violates one of the assumptions of the TREEMIX model and can lead to unclear results (Pickrell & Pritchard, 2012). Because the possibility of continuous gene flow was not evaluated in this study, the TREEMIX migration results should be interpreted with a degree of caution. In addition to TREEMIX results, population structure analyses also detected signs of admixture among some genetic groups, suggesting potential movement of genetic material among distinct groups. Such movement can have a direct effect on pathogen and disease emergence. For example, the genetic groups of *P. noxius* determined in Stewart et al. (2020) have different suitable climate space, and potential hybridization among these groups may favour their adaptation to wider range of climatic/environmental conditions. Also, Ibarra Caballero et al. (2020) showed that the *P. noxius* genome encodes a large arsenal of wood-degrading enzymes, and exchange and accumulation of such genetic regions among different genetic groups could contribute to its virulence and wide host range.

4.4 | Evolution of emerging pathogens in disturbed ecosystems

Though migration has occurred among populations, tree mortality as a result of *P. noxius* infection is strongly associated with plantings of diverse trees/shrubs, rather than natural forest ecosystems, suggesting that disturbance can contribute to the emergence of this pathogen (Ann et al., 2002; Ann, Lee, & Tsai, 1999; Bolland, 1984; Brooks, 2002; Chang, 1995; Hodges & Tenorio, 1984). In the Mariana Islands, *P. noxius*-infested stands occur on abandoned agricultural lands on the island of Rota, and *P. noxius* is present in areas devastated during WWII on Saipan (Hodges & Tenorio, 1984). Similar to Saipan, large forest areas were also destroyed in Malaysia during WWII (Farid et al., 2009). In the 1980s, the acreage of forest plantations grew rapidly in tropical regions due to the forecasted shortage of timber from natural forests (Chong, 1979), and as a result, a high prevalence of root diseases was commonly observed in these plantations (Lee, 1997). Thus, it appears that such severe deforestation and reforestation activities, along with the introduction of exotic plant species for cultivation, have contributed to the emergence of *P. noxius* in these disturbed landscapes during the 20th and 21st centuries.

Physical disturbances of forest land have long been recognized as major contributors to forest diseases (Gilbert & Hubbell, 1996; Groot et al., 2019). In western North America, a general increase in root disease incidence caused by native pathogens, *Coniferiporia sulphurascens* (reported as *P. weirii*), *Armillaria* and *Heterobasidion* spp., has been observed as a result of logging and further replacement of root disease-tolerant *Pinus* with root disease-susceptible Douglas-fir (*Pseudotsuga menziesii*) and true fir (*Abies*) species (Hansen & Goheen, 2000). These root pathogens often thrive with forest thinning practices, because of their ability for rapid colonization of root systems associated with cut stumps and subsequent spread to neighbouring trees. As a root pathogen, *P. noxius* has a similar life cycle and is capable of rapid colonization of injured and/or overcrowded trees.

Forest ecosystems in subtropical/tropical regions have undergone drastic changes over the last century (Burgess & Wingfield, 2002; Carnegie, 2007). For example, 95% of Australian forests are

dominated by diverse *Eucalyptus* and *Corymbia* species, but large portions of Australia's forested lands were heavily logged and are currently managed as regrowth forests (Carnegie, 2007). Between 1995 and 1998, hardwood plantations in Australia almost doubled (Love et al., 1999; Turnbull, 2000). These disturbances resulted in localized outbreaks of native pathogens (Burgess & Wingfield, 2002). Similar increases in disease incidence have been observed in plantations in Tasmania that were established on sites previously occupied by natural forests (Carnegie, 2007). A recent study demonstrated that forest disturbances alter the diversity and composition of the soil microbiome (Bowd et al., 2021). For example, the diversity of saprotrophic fungi was positively associated with the number of fire events. Even though this study did not examine the diversity and abundance of forest pathogens, we hypothesize that if a facultative saprotroph can maintain its pathogenicity (as with *P. noxius*), physical disturbances of forest ecosystems may lead to epiphytotic outbreaks caused by such organisms.

Additionally, severe weather events and/or climate change can contribute to forest disease epidemics/epiphytotics. Tropical typhoons regularly cause structural damage to trees on many of the Pacific Islands, and these events are thought to contribute to the emergence of *P. noxius* on affected islands within the Pacific region (Akiba et al., 2015; Cannon et al., 2022). Strong winds probably facilitate spread of *P. noxius*, the resulting tree wounds provide *P. noxius* with infection portals, and heavy rainfall during typhoon seasons provides a suitable environment for pathogen growth. Similar outbreaks have also been associated with severe weather events in temperate forests (Hansen & Goheen, 2000).

4.5 | Conclusions

A growing body of evidence suggests that human impact on ecosystems plays a major role in the emergence of newly evolved, introduced or native pathogens in both animals and plants (Altizer et al., 2013; Drenth & Guest, 2016; Fisher et al., 2012). Emerging pathogens are generally characterized by increasing incidence and fast growing ecosocial impact. Emerging pathogens constitute a major challenge for evolutionary biologists due to the limited availability of data on their history, prevalence and occurrence, population structure, and the processes contributing to their emergence. Development of analytical frameworks that help to unravel population demographic processes with limited data is a key step toward overcoming this challenge. In this paper, we demonstrate the utility of integrating population genetics with statistical and simulation tools to determine the spatiotemporal dynamics and evolutionary history of an emerging tree pathogen, *P. noxius*, in eastern Asia, Australia and the Pacific Islands. Based on the limited historical information and genetic evidence presented here, we hypothesize that emergence of *P. noxius* as a pathogen and the increasing reports of brown root rot disease are probably being driven by anthropogenic and natural disturbances, such as deforestation, subsequent land-use change, severe weather events and climate change, that markedly influence the health of forest and horticultural trees. The incidence of emerging infectious diseases has been rising over the last few decades (Fisher et al., 2012; Jones et al., 2008). Similar patterns of pathogen emergence due to contemporary activities and climate change have been reported for other pathogenic fungi, such as *Ceratocystis montia* and *Europhium clavigerum* associated with mountain pine beetle (*Dendroctonus ponderosae*) causing mortality of pine trees in temperate North America (Fisher et al., 2012), *Batrachochytrium dendrobatidis*, a cause of worldwide amphibian declines due to chytridiomycosis in the last 20 years (Rosenblum et al., 2013), and *Aspergillus sydowii*, one of the drivers of coral extinction (Harvell et al., 1999), to name just a few. Rapid changes in the environment have led to ecological imbalances in diverse ecosystems that contribute to outbreaks of infectious diseases in both animals and plants. Herein, using genome-wide allele frequency data and contemporary analytical methods combined with supervised machine learning, we provide evidence of a novel case where anthropogenic and/or environmental changes in tropical areas rather than recent pathogen introductions have probably led to the emergence of the fungal pathogen, *P. noxius*, that causes brown root rot disease. Unfortunately, because of its wide range of hosts and suitable habitats, *P. noxius* threatens food security and forest health in tropical and subtropical regions worldwide. The findings presented here provide a foundation for further research investigating the evolution of emerging pathogens, evaluation of the risks of epiphytotic outbreaks due to forest disturbances and climate change, and the development of strategies for sustainable management of forest ecosystems. With the development and integration of computational resources in molecular biology, our ability to test complex evolutionary scenarios with limited eco-evolutionary information about emerging organisms has significantly improved, making it possible to determine parameters associated with pathogen emergences.

ACKNOWLEDGEMENTS

The authors thank Y. Ota, N. Sahashi, M. Akiba, T. Hattori, N. Atibalentja, F. Brooks, C.-L. Chung, H.-H. Lee, J.-N. Tsai, Y.-C. Huang, R.-F. Liou, A. M. C. Tang, R. Y. C. Lam, M. W. K. Leung, L. M. Chu, H. S. Kwan, E. K. Dann, A. Mohd Farid, S. S. Lee, G. S. Pegg, R. L. Schlub and L. S. Shuey for sharing their *P. noxius* isolates for this study, and the USDA Forest Service, State & Private Forestry, Forest Health Protection, Special Technology Development Program (R5-2015-01, R5-2019-02), and Joint Venture Agreements (15-JV-11221633-160, 19-JV-11221633-093, 19-JV-11261957-078, 20-JV-11221633-141) to Colorado State University (J.E.S.) for partial funding of this study.

AUTHOR CONTRIBUTION

M.S.K., N.B.K., P.G.C. and J.E.S. designed the research. J.I.C., M.S.K., J.E.S., N.B.K. and P.G.C. performed the research. O.K. and J.I.C. analysed the data. O.K. wrote the manuscript. J.E.S., M.S.K., J.I.C., N.B.K. and P.G.C. provided edits.

CONFLICT OF INTEREST

Authors declare no conflict of interests.

DATA AVAILABILITY STATEMENT

Read data are publicly available in the Sequence Data Archive (BioProject ID: PRJNA735036). Three data sets used in this paper and configuration files from five sequential DIYABC-RF analyses are available on Dryad: https://doi.org/10.5061/dryad.c866t1g89.

ORCID

Olga Kozhar https://orcid.org/0000-0002-0976-7334

REFERENCES

Abe, Y., Onuki, M., Hattori, T., & Tsurumachi, M. (1995). Brown root rot caused by *Phellinus noxius* in windbreaks on Ishigaki Island, Japan. Incidence of disease, pathogen and artificial inoculation. *Japanese Journal of Phytopathology*, *61*, 425–433. https://doi.org/10.3186/jjphytopath.61.425

Akiba, M., Ota, Y., Tsai, I. J., Hattori, T., Sahashi, N., & Kikuchi, T. (2015). Genetic differentiation and spatial structure of *Phellinus noxius*, the causal agent of brown root rot of woody plants in Japan. *PLoS One*, *10*, e0141792. https://doi.org/10.1371/journal.pone.0141792

Alroy, J. (2017). Effects of habitat disturbance on tropical forest biodiversity. *Proceedings of National Academy of Sciences*, *114*(23), 6056–6061. https://doi.org/10.1073/pnas.1611855114

Altizer, S., Ostfeld, R. S., Johnson, P. T. J., Kutz, S., & Harvell, C. D. (2013). Climate change and infectious diseases: from evidence to a predictive framework. *Science*, *341*, 514–519. https://doi.org/10.1126/science.1239401

Ann, P. J., Chang, T. T., & Ko, W. H. (2002). *Phellinus noxius* brown root rot of fruit and ornamental trees in Taiwan. *Plant Disease*, *86*, 820–826. https://doi.org/10.1094/PDIS.2002.86.8.820

Ann, P. J., Lee, H. L., & Huang, T. C. (1999). Brown root rot of 10 species of fruit trees caused by *Phellinus noxius* in Taiwan. *Plant Disease*, *83*, 746–750. https://doi.org/10.1094/PDIS.1999.83.8.746

Ann, P. J., Lee, H. L., & Tsai, J. N. (1999). Survey of brown root disease of fruit and ornamental trees caused by *Phellinus noxius* in Taiwan. *Plant Pathology Bulletin*, *8*, 51–60.

Ayres, M. P., & Lombardero, M. J. (2000). Assessing the consequences of global change for forest disturbance from herbivores and pathogens. *The Science of the Total Environment*, *262*, 263–286. https://doi.org/10.1016/S0048-9697(00)00528-3

Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: A flexible trimmer for Illumina Sequence Data. *Bioinformatics btu170*, *30*(15), 2114–2120 https://doi.org/10.1093/bioinformatics/btu170

Bolland, L. (1984). *Phellinus noxius*: cause of a significant root-rot in Queensland hoop pine plantations. *Australian Forestry*, *47*, 2–10. https://doi.org/10.1080/00049158.1984.10675972

Bowd, E. J., Banks, S. C., Bissett, A., May, T. W., & Lindenmayer, D. B. (2021). Disturbance alters the forest soil microbiome. *Molecular Ecology*, *00*, 1–29. https://doi.org/10.1111/mec.16242

Brooks, F. (2002). Brown root rot disease in American Samoa's tropical rain forest. *Pacific Science*, *56*, 377–387.

Burgess, T., & Wingfield, M. J. (2002). Impact of fungal pathogens in natural forest ecosystems: a focus on eucalypts. In K. Sivasithamparam, K. W. Dixon, & R. L. Barrett (Eds.), *Microorganisms in plant conservation and biodiversity* (pp. 285–306). Kluwer Academic Publishers.

Cannon, P. G., Klopfenstein, N. B., Kim, M. S., Stewart, J. E., & Chung, C. L. (2022). *Brown root rot disease caused by Phellinus noxius in USA-Affiliated Islands of the Western Pacific*. General Technical Report PNW-GTR-XXX. Portland, OR: USA. Dept. of Agriculture, Forest Service, Pacific Northwest Research Station. *in press*.

Carnegie, A. J. (2007). Forest health condition in New South Wales, Australia, 1996–2005. II. Fungal damage recorded in eucalypt plantations during forest health surveys and their management. *Australasian Plant Pathology*, *36*, 225–239. https://doi.org/10.1071/AP07021

Catchen, J. M., Amores, A., Hohenlohe, P., Cresko, W., & Postlethwait, J. H. (2011). Stacks: building and genotyping loci de novo from short-read sequences. *G3: Genes, Genomes, Genetics*, *1*, 171–182. https://doi.org/10.1534/g3.111.000240

Catchen, J., Hohenlohe, P. A., Bassham, S., Amores, A., & Cresko, W. A. (2013). Stacks: an analysis tool set for population genomics. *Molecular Ecology*, *22*, 3124–3140. https://doi.org/10.1111/mec.12354

Cavalli-Sforza, L. L., & Feldman, M. W. (2003). The application of molecular genetic approaches to the study of human evolution. *Nature Genetics*, *33*, 266–275. https://doi.org/10.1038/ng1113

Chang, T. T. (1995). Decline of nine tree species associated with brown root rot caused by *Phellinus noxius* in Taiwan. *Plant Disease*, *79*, 962–965. https://doi.org/10.1094/PD-79-0962

Chang, T. T. (1996). Survival of *Phellinus noxius* in soil and roots of dead host plants. *Phytopathology*, *86*, 272–276. https://doi.org/10.1094/Phyto-86-272

Chhatre, V. E., & Emerson, K. J. (2017). StrAuto: Automation and parallelization of STRUCTURE analysis. *BMC Bioinformatics*, *18*, 192. https://doi.org/10.1186/s12859-017-1593-0

Chong, P. W. (1979). The growing domestic demand for timber and its influence on forest management. *Malaysian Forester*, *42*, 378–389.

Chung, C.-L., Huang, S.-Y., Huang, Y.-C., Tzean, S.-S., Ann, P.-J., Tsai, J.-N., Yang, C.-C., Lee, H.-H., Huang, T. W., Huang, H.-Y., Chang, T.-T., Lee, H.-L., & Liou, R.-F. (2015). The genetic structure of *Phellinus noxius* and dissemination patterns of brown root rot disease in Taiwan. *PLoS One*, *10*, e0139445. https://doi.org/10.1371/journal.pone.0139445

Chung, C.-L., Lee, T.-J., Akiba, M., Lee, H.-H., Kuo, T.-H., Liu, D., Kr, H.-M., Yokoi, T., Roa, M. B., Lu, M.-Y.- J., Chang, Y.-Y., Ann, P.-J., Tsai, J.-N., Chen, C.-Y., Tzean, S.-S., Ota, Y., Hattori, T., Sahashi, N., Liou, R.-F. … Tsai, I. J. (2017). Comparative and population genomic landscape of *Phellinus noxius*: a hypervariable fungus causing root rot in trees. *Molecular Ecology*, *26*, 6301–6316. https://doi.org/10.1111/mec.14359

Clegg, S. M., Degnan, S. M., Kikkawa, J., Moritz, C., Estoup, A., & Owens, I. P. (2002). Genetic consequences of sequential founder events by an island-colonizing bird. *Proceedings of the National Academy of Sciences*, *99*(12), 8127–8132.

Collin, F.-D., Durif, G., Raynal, L., Lombaert, E., Gautier, M., Vitalis, R., Marin, J.-M., & Estoup, A. (2021). Extending approximate Bayesian computation with supervised machine learning to infer demographic history from genetic polymorphisms using DIYABC Random Forest. *Molecular Ecology Resources* Early View, *21*(8), 2598–2613, https://doi.org/10.1111/1755-0998.13413

Corner, E. J. H. (1932). The identification of the brown root fungus. *Gardeners Bulletin Straits Settlements*, *5*, 317–350.

Corredor-Moreno, P., & Saunders, D. G. O. (2019). Expecting the unexpected: factors influencing the emergence of fungal and oomycete plant pathogens. *New Phytologist*, *225*, 118–125. https://doi.org/10.1111/nph.16007

Drenth, A., & Guest, D. I. (2016). Fungal and oomycete diseases of tropical tree fruit crops. *Annual Review of Phytopathology*, *54*, 373–395. https://doi.org/10.1146/annurev-phyto-080615-095944

Dutech, C., Barres, B., Bridier, J., Robin, C., Milgroom, M. G., & Ravigne, V. (2012). The chestnut blight fungus world tour: successive introduction events from diverse origins in an invasive plant fungal pathogen. *Molecular Ecology*, *21*(16), 3931–3946. https://doi.org/10.1111/j.1365-294X.2012.05575.x

Earl, D. A., & vonHoldt, B. M. (2012). STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources*, *4*, 359–361. https://doi.org/10.1007/s12686-011-9548-7

Evanno, G., Regnaut, S., & Goudet, J. (2005). Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology*, *14*, 2611–2620. https://doi.org/10.1111/j.1365-294X.2005.02553.x

Excoffier, L., & Lischer, H. E. L. (2010). Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources*, *10*, 564–567. https://doi.org/10.1111/j.1755-0998.2010.02847.x

Farid, A. M., Lee, S. S., Maziah, Z., & Patahayah, M. (2009). Pathogenicity of *Regidoporus microporus* and *Phellinus noxius* against four major plantation tree species in Peninsular Malaysia. *Journal of Tropical Forest Science*, *21*, 289–298.

Firth, S. (2007). Pacific Islands trade, labor, and security in an era of globalization. *The Contemporary Pacific*, *19*, 111–135.

Fisher, M., Henk, D., Briggs, C., Brownstein, J. S., Madoff, L. C., McCraw, S. L., & Gurr, S. J. (2012). Emerging fungal threats to animal, plant and ecosystem health. *Nature*, *484*, 186–194. https://doi.org/10.1038/nature10947

Gilbert, G. S., & Hubbell, S. P. (1996). Plant diseases and conservations of tropical forests. *BioScience*, *2*, 98–106. https://doi.org/10.2307/1312812

Glenn, T. C., Nilsen, R. A., Kieran, T. J., Sanders, J. G., Bayona-Vásquez, N. J., Finger, J. W., Pierson, T. W., Bentley, K. E., Hoffberg, S. L., Loucha, S., Garcia-De Leon, F. J., del Rio Portilla, M. A., Reed, K. D., Anderson, J. L., Meece, J. K., Aggrey, S. E., Rekaya, R., Alabady, M., Belanger, M., & Faircloth, B. C. (2019). Adapterama I: universal stubs and primers for 384 unique dual-indexed or 147,456 combinatorially-indexed Illumina libraries (iTru & iNext). *PeerJ*, *7*, e7755. https://doi.org/10.7717/peerj.7755

Goss, E. M., Larsen, M., Vercauteren, A., Werres, S., Heungens, K., & Grunwald, N. J. (2011). *Phytophtora ramorum* in Canada: evidence for migration within North America and from Europe. *Phytopathology*, *101*, 166–171. https://doi.org/10.1094/PHYTO-05-10-0133

Graham, C. F., Glenn, T. C., McArthur, A. G., Boreham, D. R., Kieran, T., Lance, S., Manzon, R. G., Martino, J. A., Pierson, T., Rogers, S. M., Wilson, J. Y., & Somers, C. M. (2015). Impacts of degraded DNA on restriction enzyme associated DNA sequencing (RADSeq). *Molecular Ecology Resources*, *15*, 1304–1315. https://doi.org/10.1111/1755-0998.12404

Gronau, I., Hubisz, M. J., Gulko, B., Danko, C. G., & Siepel, A. (2011). Bayesian inference of ancient human demography from individual genome sequences. *Nature Genetics*, *43*, 1031–1035. https://doi.org/10.1038/ng.937

Groot, M., Diaci, J., & Ogris, N. (2019). Forest management history is an important factor in bark beetle outbreaks: Lessons for the future. *Forest Ecology and Management*, *433*, 467–474. https://doi.org/10.1016/j.foreco.2018.11.025

Guillemaud, T., Beaumont, M. A., Ciosi, M., Cornuet, J. M., & Estoup, A. (2010). Inferring introduction routes of invasive species using approximate Bayesian computation on microsatellite data. *Heredity*, *104*, 88–99.

Hamelin, R. C. (2005). Forest pathology in the era of genomics. In: J. E. Lundquist, & R. C. Hamelin (Eds.), *Forest Pathology - From genes to landscapes* (pp. 1–7). APS Press.

Hansen, E. M., & Goheen, E. M. (2000). *Phellinus weirii* and other native root pathogens as determinants of forest structure and process in western North America. *Annual Review of Phytopathology*, *38*, 515–539. https://doi.org/10.1146/annurev.phyto.38.1.515

Harvell, C. D., Kim, K., Burkholder, J. M., Colwell, R. R., Epstein, P. R., Grimes, D. J., Hofmann, E. E., Lipp, E. K., Osterhaus, A. D., Overstreet, R. M., Porter, J. W., Smith, G. W., & Vasta, G. R. (1999). Review: Emerging marine diseases – Climate links and anthropogenic factors. *Science*, *285*, 1505–1510. https://doi.org/10.1126/science.285.5433.1505

Hoang, D. T., Chernomor, O., von Haeseler, A., Minh, B. Q., & Vinh, L. S. (2018). UFBoot2: Improving the ultrafast bootstrap approximation. *Molecular Biology and Evolution*, *35*, 518–522. https://doi.org/10.1093/molbev/msx281

Hodges, C. S., & Tenorio, J. A. (1984). Root disease of *Delonix regia* and associated tree species in the Mariana Islands caused by *Phellinus noxius*. *Plant Disease*, *68*, 334–336. https://doi.org/10.1094/PD-68-334

Ibarra Caballero, J. R., Ata, J. P., Leddy, K. A., Glenn, T. C., Kieran, T. J., Klopfenstein, N. B., Kim, M.-S., & Stewart, J. E. (2020). Genome comparison and transcriptome analysis of the invasive brown root rot pathogen, *Phellinus noxius*, from different geographic regions reveals potential enzymes associated with degradation of different wood substrates. *Fungal Biology*, *124*, 144–154. https://doi.org/10.1016/j.funbio.2019.12.007

Ito, M. (1998). Origin and Evolution of endemic plants on the Bonin (Ogasawara) Islands. *Researchers on Population Ecology*, *40*, 205–212. https://doi.org/10.1007/BF02763405

Ivory, M. H. (1996). Diseases of forest trees caused by the pathogen *Phellinus noxius*. In S. P. Raychaudhuri, & K. Maramorosch (Eds.), *Forest trees and palms: Diseases and control* (pp. 111–133). Science Publishers.

Jakobsson, M., & Rosenberg, N. A. (2007). CLUMPP: A cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics*, *23*, 1801–1806. https://doi.org/10.1093/bioinformatics/btm233

Jombart, T. (2008). adegenet: A R package for the multivariate analysis of genetic markers. *Bioinformatics*, *24*, 1403–1405. https://doi.org/10.1093/bioinformatics/btn129

Jones, K. E., Patel, N. G., Levy, M. A., Storeygard, A., Balk, D., Gittleman, J. L., & Daszak, P. (2008). Global trends in emerging infectious diseases. *Nature*, *451*, 990–993. https://doi.org/10.1093/bioinformatics/btn129

Kalyaanamoorthy, S., Minh, B. O., Wong, T. K. F., von Haeseler, A., & Jermiin, L. S. (2017). ModelFinder: Fast model selection for accurate phylogenetic estimates. *Nature Methods*, *14*, 587–589. https://doi.org/10.1038/nmeth.4285

Kobayashi, T., Abe, Y., & Kawabe, Y. (1991). Brown root rot disease: A new threat found in windbreaks in Okinawa Prefecture, Japan. *Forest Chemistry*, *118*, 1–7. (in Japanese).

Kolbe, J. J., Glor, R. E., Schettino, L. R. G., Lara, A. C., Larson, A., & Losos, J. B. (2004). Genetic variation increases during biological invasion by a Cuban lizard. *Nature*, *431*, 177–181. https://doi.org/10.1038/nature02807

Langmead, B., & Salzberg, S. (2012). Fast gapped-read alignment with Bowtie 2. *Nature Methods*, *9*, 357–359. https://doi.org/10.1038/nmeth.1923

Lee, S. S. (1997). Diseases of some tropical plantation Acacias in Peninsular Malaysia. In K. M. Old, S. S. Lee, & J. K. Sharma (Eds.), *Diseases of Tropical Acacias. Proceedings of the International Workshop. Subanjeriji (South Sumatra). 28 April – 3 May 1996. CIFOR Special Publication. Center for International Forestry Research, Jakarta, 53–61*.

Lewis, P. O. (2001). A likelihood approach to estimating phylogeny from discrete morphological character data. *Systematic Biology*, *6*, 913–925. https://doi.org/10.1080/106351501753462876

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R. & 1000 Genome Project Data Processing Subgroup (2009). The Sequence alignment/map (SAM) format and SAMtools. *Bioinformatics*, *25*, 2078–2079. https://doi.org/10.1093/bioinformatics/btp352

Lim, Y. W., Sturrock, R., Leal, I., Pellow, K., Yamaguchi, T., & Breuil, C. (2007). Distinguishing homokaryons and heterokaryons in *Phellinus sulphurascens* using pairing tests and ITS polymorphisms. *Antonie*

van Leeuwenhoek, *93*, 99–110. https://doi.org/10.1007/s10482-007-9184-0

Love, G., Yainshet, A., & Dlugoz, J. (1999). Forests products: Outlook in regional markets to 2003–04. In *National stakeholder timber pest conference. 20–22 April 1999* (pp. 3–15).

Maside, X., Gómez-Moracho, T., Jara, L., Martín-Hernández, R., De la Rúa, P., Higes, M., & Bartolomé, C. (2015). Population genetics of *Nosema apis* and *Nosema ceranae*: One host (*Apis mellifera*) and two different histories. *PLoS One*, *10*(12), e0145609. https://doi.org/10.1371/journal.pone.0145609

Matisso-Smith, E. (2015). Ancient DNA and the human settlement of the Pacific: an overview. *Journal of Human Evolution*, *79*, 93–104. https://doi.org/10.1016/j.jhevol.2014.10.017

Mildrexler, D. J., Shaw, D. C., & Cohen, W. B. (2019). Short-term climate trends and the Swiss needle cast epidemic in Oregon's public and private coastal forestlands. *Forest Ecology and Management*, *432*, 501–513. https://doi.org/10.1016/j.foreco.2018.09.025

Moller, M., & Stukenbrock, E. H. (2017). Evolution and genome architecture of fungal plant pathogens. *Nature Reviews Microbiology*, *15*, 756–771. https://doi.org/10.1038/nrmicro.2017.76

Nandris, D., Nicole, M., & Geiger, J. P. (1987a). Root rot diseases of rubber trees. *Plant Disease*, *71*, 298–306. https://doi.org/10.1094/PD-71-0298

Nandris, D., Nicole, M., & Geiger, J. P. (1987b). Variation in virulence among *Rigidoporus lignosus* and *Phellinus noxius* isolates from West Africa. *European Journal of Forest Pathology*, *17*, 271–281. https://doi.org/10.1111/j.1439-0329.1987.tb01026.x

Nei, M., Maruyama, T., & Chakraborty, R. (1975). The bottleneck effect and genetic variability in populations. *Evolution*, *29*(1), 1–10.

Neil, P. E. (1986). A preliminary note on *Phellinus noxius* root rot of *Cordia alliodora* plantings in Vanuatu. *European Journal of Forest Pathology*, *16*, 274–280. https://doi.org/10.1111/j.1439-0329.1986.tb00192.x

Nguyen, L. T., Schmidt, H. A., von Haeseler, A., & Minh, B. Q. (2015). IQ-TREE: A fast and effective stochastic algorithm for estimating maximum likelihood phylogenies. *Molecular Biology and Evolution*, *32*, 268–274. https://doi.org/10.1093/molbev/msu300

Nichols, R. A., & Hewitt, G. M. (1994). The genetic consequences of long distance dispersal during colonization. *Heredity*, *72*(3), 312–317. https://doi.org/10.1038/hdy.1994.41

Otrosina, W. J. (2005). Exotic ecosystems: where root disease is not a beneficial component of temperate conifer forests. In: J. E. Lundquist, & R. C. Hamelin (Eds.), *Forest Pathology. From genes to landscapes*. (pp. 121–126). APS Press.

Otrosina, W. J., & Ferrell, G. T. (1995). Root diseases: primary agents and secondary consequences of disturbance. In: L. G. Eskew, comp. *Forest health through silviculture: proceedings of the 1995 National Silviculture Workshop, Mescalero, New Mexico, May 8-11, 1995. Gen. Tech. Rep. RM-GTR-267*. Fort Collins, CO: U.S. Department of Agriculture, Forest Service, Rocky Mountain Forest and Range Experiment Station, 87 – 92.

Patterson, N., Moorjani, P., Luo, Y., Mallick, S., Rohland, N., Zhan, Y., Genschoreck, T., Webster, T., & Reich, D. (2012). Ancient admixture in human history. *Genetics*, *192*, 1065–1093. https://doi.org/10.1534/genetics.112.145037

Pegler, D. N., & Waterston, J. M. (1968). Phellinus noxius. In: *C. M. I. Descriptions of Pathogenic Fungi and Bacteria, No. 195*. Commonwealth Mycological Institute.

Pickrell, J. K., & Pritchard, J. K. (2012). Inference of population splits and mixtures from genome-wide allele frequency data. *PLoS Genetics*, *8*, https://doi.org/10.1371/journal.pgen.1002967

Pingali, P. L. (2012). Green revolution: impacts, limits, and the path ahead. *Proceedings of the National Academy of Sciences of the United States of America*, *109*, 12302–12308. https://doi.org/10.1073/pnas.0912953109

Potter, C., & Urquhart, J. (2017). Tree diseases and pest epidemics in the Anthropocene: a review of the drivers, impacts and policy responses in the UK. *Forest Policy and Economics*, *79*, 61–68. https://doi.org/10.1016/j.forpol.2016.06.024

Pritchard, J. K., Stephens, M., & Donnelly, P. (2010). Inference of population structure using multilocus genotype data. *Genetics*, *155*, 945–959.

Rachowicz, L. J., Hero, J. M., Alford, R. A., Taylor, J. W., Morgan, J. A. T., Vredenburg, V. T., Collins, J. P., & Briggs, C. J. (2005). The novel and endemic pathogen hypotheses: Competing explanations for the origin of emerging infectious diseases of wildlife. *Conservation Biology*, *19*, 1441–1448. https://doi.org/10.1111/j.1523-1739.2005.00255.x

Ray, N., Currat, M., & Excoffier, L. (2003). Intra-deme molecular diversity in spatially expanding populations. *Molecular Biology and Evolution*, *20*(1), 76–86. https://doi.org/10.1093/molbev/msg009

Reich, D., Thangaraj, K., Patterson, N., Price, A. L., & Singh, L. (2009). Reconstructing Indian population history. *Nature*, *461*(7263), 489–494.

Rey, O., Fourtune, L., Paz-Vinas, I., Loot, G., Veyssiere, C., Roche, B., & Blanchet, S. (2015). Elucidating the spatio-temporal dynamics of an emerging wildlife pathogen using approximate Bayesian computation. *Molecular Ecology*, *24*, 5348–5363. https://doi.org/10.1111/mec.13401

Richardson, B., Klopfenstein, N. B., & Peever, T. L. (2005). Assessing forest-pathogen interactions at the population level. In: J. E. Lundquist, & R. C. Hamelin (Eds.), *Forest Pathology. From genes to landscapes*. (pp. 21–30). APS Press.

Ritokova, G., Shaw, D. C., Filip, G., Kanaskie, A., Browning, J., & Norlander, D. (2016). Swiss needle cast in Western Oregon Douglas-fir plantations: 20-year monitoring results. *Forests*, *7*, 155. https://doi.org/10.3390/f7080155

Rosenblum, E. B., James, T. Y., Zamudio, K. R., Poorten, T. J., Ilut, D., Rodriguez, D., Eastman, J. M., Richards-Hrdlika, K., Joneson, S., Jenkinson, T. S., Longcore, J. E., Parra Olea, G., Toledo, L. F., Arellano, M. L., Medina, E. M., Restrepo, S., Flechas, S. V., Berger, L., Briggs, C. J., & Stajich, J. E. (2013). Complex history of the amphibian-killing chytrid fungus revealed with genome resequencing data. *Proceedings of the National Academy of Sciences of the United States of America*, *23*, 9385–9390. https://doi.org/10.1073/pnas.1300130110

Sahashi, N., Akiba, M., Ishihara, M., Abe, Y., & Morita, S. (2007). First report of the brown root rot disease caused by *Phellinus* noxius, its distribution and newly recorded host plants in the Amami Islands, southern Japan. *Forest Pathology*, *37*, 167–173. https://doi.org/10.1111/j.1439-0329.2007.00491.x

Santini, A., Liebhold, A., Migliorini, D., & Woodward, S. (2018). Tracing the role of human civilization in the globalization of plant pathogens. *ISME Journal*, *12*, 647–652. https://doi.org/10.1038/s41396-017-0013-9

Sawada, K. (1928). Camphor tree decline. *Description Catalog Formosan Fungi*, *4*, 86–91.

Schmeller, D. S., Courchamp, F., & Killeen, G. (2020). Biodiversity loss, emerging pathogens and human health risks. *Biodivers and Conservation*, *29*, 3095–3102. https://doi.org/10.1007/s10531-020-02021-6

Silva, D. N., Talhinhas, P., Cai, L., Manuel, L., Gichuru, E. K., Loureiro, A., Varzea, V., Paulo, O. S., & Batista, D. (2012). Host-jump drives rapid and recent ecological speciation of the emergent fungal pathogen *Colletotrichum kahawae*. *Molecular Ecology*, *21*(11), 2655–2670.

Slatkin, M. (1987). Gene flow and the geographic structure of natural populations. *Science*, *236*(4803), 787–792. https://doi.org/10.1126/science.3576198

Slatkin, M., & Excoffier, L. (2012). Serial founder effects during range expansion: a spatial analog of genetic drift. *Genetics*, *191*(1), 171–181.

Soghigian, J., Gloria-Soria, A., Robert, V., Goff, G. L., Failloux, A.-B., & Powell, J. R. (2020). Genetic evidence for the origin of *Aedes*

aegypti, the yellow fever mosquito, in the southwestern Indian Ocean. *Molecular Ecology*, *29*, 3593–3603.

Speight, M. R., & Woodward, S. (2016). Pest management in tropical forestry. In L. Pancel, & M. Köhl (Eds.), *Tropical Forestry Handbook*. Springer.

Stewart, J. E., Kim, M.-S., Ota, Y., Sahashi, N., Hanna, J. W., Akiba, M., Ata, J. P., Atibalentja, N., Brooks, F., Chung, C.-L., Dann, E. K., Mohd Farid, A., Hattori, T., Lee, S. S., Otto, K., Pegg, G. S., Schlub, R. L., Shuey, L. S., Tang, A. M. C., … Klopfenstein, N. B. (2020). Phylogenetic and population genetic analyses reveal three distinct lineages of the invasive brown root-rot pathogen, *Phellinus noxius*, and bioclimatic modeling predicts differences in associated climate niches. *European Journal of Plant Pathology*, *156*, 751–766. https://doi.org/10.1007/s10658-019-01926-5

Stewart, J. E., Ross-Davis, A. L., Graça, R. N., Alfenas, A. C., Peever, T. L., Hanna, J. W., Uchida, J. Y., Hauff, R. D., Kadooka, C. Y., Kim, M.-S., Cannon, P. G., Namba, S., Simeto, S., Perez, C. A., Rayamajhi, M. B., Lodge, D. J., Arguedas, M., Medel-Ortiz, R., Lopez-Ramirez, M. A., … Klopfenstein, N. B. (2018). Genetic diversity of *Puccinia psidii* in the Americas and Hawaii: Potential global implications. *Forest Pathology*, *48*, e12378. https://doi.org/10.1111/efp.12378

Tai, F. L. (1979). *Sylloge fungorum sinicorum*. Science Press.

Tsai, J.-N., Ann, P.-J., Liou, R.-F., Hseih, W.-H., & Ko, W.-H. (2017). *Phellinus noxius*: molecular diversity among isolates from Taiwan and its phylogenetic relationship with other species of *Phellinus* based on sequences of the ITS region. *Botanical Studies*, *58*, 9. https://doi.org/10.1186/s40529-017-0162-1

Turnbull, J. W. (2000). Economic and social importance of eucalypts. In P. J. Keane, G. A. Kile, F. D. Podger, & B. N. Brown (Eds.), *Diseases and pathogens of eucalypts* (pp. 1–10). CSIRO Publishing.

Wairiu, M. (2017). Land degradation and sustainable land management practices in Pacific Island Countries. *Regional Environmental Change*, *17*, 1053–1064. https://doi.org/10.1007/s10113-016-1041-0

Wang, J., Fernandez-Pavia, S. P., Larsen, M. M., Garay-Serrano, E., Gregorio-Cipriano, R., Rodriguez-Alvarado, G., Grunwald, N. J., & Goss, E. M. (2016). High levels of diversity and population structure in the potato late blight pathogen at the Mexico centre of origin. *Molecular Ecology*, *26*(4), 1091–1107. https://doi.org/10.1111/mec.14000

Woods, A. J., Marti-Garcia, J., Bulman, L., Vasconcelos, M. W., Boberg, J., La Porta, N., Peredo, H., Vergara, G., Ahumada, R., Brown, A., & Diez, J. J. (2016). Dothistroma needle blight, weather, and possible climate triggers for the disease's recent emergence. *Forest Pathology*, *46*, 443–452. https://doi.org/10.1111/efp.12248

Yasuda, A. (1916). Miscellaneous note of Fungi (56). *Botanical Magazine*, *Tokyo*, *30*, 350. (in Japanese).

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

How to cite this article: Kozhar, O., Kim, M.-S., Ibarra Caballero, J., Klopfenstein, N. B., Cannon, P. G., & Stewart, J. E. (2022). Long evolutionary history of an emerging fungal pathogen of diverse tree species in eastern Asia, Australia and the Pacific Islands. *Molecular Ecology*, 31, 2013–2031. https://doi.org/10.1111/mec.16384