Population Genomics Trace Clonal Diversification and Intercontinental Migration of an Emerging Fungal Pathogen of Boxwood

Nicholas LeBlanc,^{1,2,3} Marc A. Cubeta,³ and Jo Anne Crouch^{1,†}

¹ Mycology and Nematology Genetic Diversity and Biology Laboratory, United States Department of Agriculture–Agricultural Research Service, Beltsville, MD

² Oak Ridge Institute for Science and Education, ARS Research Participation Program, Oak Ridge, TN

³ Department of Entomology and Plant Pathology, North Carolina State University, Center for Integrated Fungal Research, Raleigh, NC

Accepted for publication 9 October 2020.

ABSTRACT

Boxwood blight was first documented in Europe, prior to its recent colonization of North America, where it continues to have significant negative impacts on the ornamental industry. Due to near genetic uniformity in the two sister species of fungal plant pathogens that cause boxwood blight, understanding historical disease emergence and predicting future outbreaks is limited. The goal of this research was to apply population genomics to understand the role of pathogen diversification and migration in disease emergence. Specifically, we tested whether the primary pathogen species *Calonectria pseudonaviculata* has remained genetically isolated from its European-limited sister species *C. henricotiae*, while diversifying into clonal lineages that have migrated among continents. Whole-genome sequencing identified 1,608 single-nucleotide polymorphisms (SNPs) in 67 *C. pseudonaviculata* isolates from four continents and 1,017 SNPs in 13 *C. henricotiae* isolates from Europe. Interspecific genetic differentiation and an absence of shared polymorphisms indicated lack of gene flow between the

sister species. Tests for intraspecific genetic structure in *C. pseudonaviculata* identified four genetic clusters, three of which corresponded to monophyletic phylogenetic clades. Comparison of evolutionary divergence scenarios among the four genetic clusters using approximate Bayesian computation indicated that the two *C. pseudonaviculata* genetic clusters currently found in the United States were derived from different sources, one from the first genetic cluster found in Europe and the second from an unidentified population. Evidence for multiple introductions of this pathogen into the United States and intercontinental migration indicates that future introductions are likely to occur and should be considered in plant disease quarantine regulation.

Keywords: boxwood blight, *Calonectria*, clonality, emerging diseases, evolution, fungal pathogens, genomics, population biology, single-nucleotide polymorphisms,

Emerging plant diseases pose a global threat to the health and productivity of plants and the biodiversity of native ecosystems (Fisher et al. 2012). Global dispersal of microorganisms responsible for these diseases is often mediated by human activity, which can lead to rapid range expansions, even in the presence of significant natural barriers (Anderson et al. 2004; Brasier 2008; Liebholt et al. 2012). Human-mediated movement not only leads to the introduction of nonnative invasive plant pathogens into disease-free regions but can also have significant impacts on pathogen evolution, giving rise to novel population genetic and phenotypic diversity

(Gladieux et al. 2015; McDonald and Linde 2002; McDonald and Stukenbrock 2016). Consequently, reconstructing routes of pathogen introduction is essential for developing trade regulatory strategies and understanding evolutionary processes associated with successful establishment of introduced microorganisms (Gladieux et al. 2015; Sakai et al. 2001).

Boxwood blight is an emerging fungal disease of broadleaf evergreen shrubs and small trees in the genus *Buxus* (commonly referred to as boxwood). Boxwood blight poses a serious economic and ecological threat due to the widespread commercial production and cultivation of boxwood as ornamental plants and the significant role of these plants in natural ecosystems (LeBlanc et al. 2018). The disease is caused by two sister species of fungi in the phylum Ascomycota, family Nectriaceae: *Calonectria pseudonaviculata* and *C. henricotiae* (Gehesquière et al. 2016). *C. pseudonaviculata* is geographically widespread across all areas where boxwood blight is found and, to date, *C. henricotiae* only occurs in continental Europe and the United Kingdom (Gehesquière et al. 2016; LeBlanc et al. 2019). Despite having a common host and overlapping geographic range, *C. henricotiae* and *C. pseudonaviculata* display clear genetic and phenotypic differences. Genomic data have shown that each species has a single mating type idiomorph (Malapi-Wight et al. 2019), and *C. pseudonaviculata* has a loss-of-function mutation in one of three *CYP51* paralogs, compared with *C. henricotiae* (Stravoravdis et al. 2020). There are also phenotypic differences between the species. *C. henricotiae* can grow at higher temperatures and is more insensitive toward certain fungicide classes (Gehesquière et al. 2016; Miller et al. 2018). Because fungicides are currently one of the primary approaches used to manage boxwood blight disease, resistance by *C. henricotiae* represents a significant concern to the ornamental industry (Daughtrey 2019).

†Corresponding author: J. A. Crouch; joanne.crouch@usda.gov

Funding: This work is supported by the United States Department of Agriculture (USDA) National Institute of Food and Agriculture award number 2019-67012-29631 to N. LeBlanc, USDA Agricultural Research Service (ARS) project 8042-22000-298-00D, USDA-ARS Floriculture Nursery Research Initiative, and by an appointment of N. LeBlanc to the ARS Research Participation Program administered by the Oak Ridge Institute for Science and Education (ORISE) through an interagency agreement between the U.S. Department of Energy (DOE) and the USDA. ORISE is managed by Oak Ridge Associated Universities under DOE contract number DE579AC05-06OR23100.

Current address of N. LeBlanc: United States Department of Agriculture–Agricultural Research Service, Crop Improvement and Protection Research Unit, Salinas, CA 93905.

*The e-Xtra logo stands for “electronic extra” and indicates that supplementary figures and supplementary materials are published online.

The author(s) declare no conflict of interest.

This article is in the public domain and not copyrightable. It may be freely reprinted with customary crediting of the source. The American Phytopathological Society, 2021.

The global origins of the boxwood blight pathogens are unknown, although the two species are estimated to have diverged approximately two million years ago (Malapi-Wight et al. 2019). Considering the geographic and phenotypic differences between the two boxwood blight pathogen species, understanding how *C. pseudonaviculata* became geographically widespread is critical for informing future efforts to prevent further range expansion of these pathogens. Following initial discovery of boxwood blight in the United Kingdom and New Zealand in the 1990s, the disease rapidly spread across continental Europe and throughout native ecosystems around the Black Sea (LeBlanc et al. 2018). In 2011, boxwood blight was identified in North America, with *C. pseudonaviculata* identified on the east and west coasts of the United States and Canada (Daughtrey 2019; LeBlanc et al. 2018). Since then, the disease has spread across the United States, with reports of *C. pseudonaviculata* from at least 30 states (Castroagudín et al. 2020).

Multiple studies have shown that *C. pseudonaviculata* and *C. henricotiae* lack genetic diversity and are clonal in their current geographic ranges. Lack of genetic diversity in populations of *C. pseudonaviculata* and *C. henricotiae* was first demonstrated from multilocus phylogenetic analyses and amplified fragment length polymorphisms (Gehesquière et al. 2016; Henricot and Culham 2002). A survey of global pathogen populations using 11 simple sequence repeat (SSR) loci also found limited genetic diversity in *C. henricotiae* and *C. pseudonaviculata* but provided evidence of more than one genotype in each species (LeBlanc et al. 2019). In particular, 14 genotypes were identified from 270 isolates of *C. pseudonaviculata* collected during the global expansion of boxwood blight (LeBlanc et al. 2019). Further analysis of these data found no evidence for random association among the 11 loci (Castroagudín et al. in press; LeBlanc et al. 2019), rejecting the hypothesis of random mating and supporting known mating type distributions in the two pathogen species (Castroagudín et al. in press; Malapi-Wight et al. 2019).

The goal of this research was to trace the global emergence of boxwood blight. To overcome the challenge of near genetic uniformity in *C. pseudonaviculata* and *C. henricotiae* when assessed using SSRs and phylogenetic markers, we applied population genomic approaches based on whole-genome sequencing to evaluate evidence of pathogen diversification and intercontinental migration as drivers of disease emergence. Specifically, we tested the hypotheses that (i) *C. pseudonaviculata* is genetically isolated from its rare sister species, *C. henricotiae*; (ii) *C. pseudonaviculata* has diversified into multiple clonal lineages; and (iii) *C. pseudonaviculata* clonal lineages have migrated among continents. This study provides novel insight into the emergence of boxwood blight and demonstrates the potential for using similar approaches to understand and prevent the spread of newly emerged clonal plant pathogens.

MATERIALS AND METHODS

Isolate collection and genome sequencing. Isolates of *C. henricotiae* and *C. pseudonaviculata* ($n = 80$) used for this study were selected to represent the emergence and initial occurrence of boxwood blight in Europe, Asia, New Zealand, and the United States. Where possible, isolates from these continents were selected to represent initial disease detection and more recent collection of pathogen isolates (see Supplementary Table S1 for a complete list of isolates and associated metadata). The isolates included European-limited *C. henricotiae* ($n = 13$) and more geographically widespread *C. pseudonaviculata* ($n = 67$). Among *C. pseudonaviculata* isolates were representatives of the four most frequently sampled SSR genotypes (G1, G2, G3, and G4) sampled and reported previously (LeBlanc et al. 2019). Isolates representing less frequently sampled SSR genotypes (G5, G6, G7, and G12) were also included in this study. The remaining *C. pseudonaviculata* isolates ($n = 26$)

originated from recent collections in the United States (Castroagudín et al. in press). All isolates were regularly maintained on potato dextrose agar (Becton Dickinson, Franklin Lakes, NJ, U.S.A.) and used to derive pure-culture, single-conidial isolates for extraction of genomic DNA.

Fungal tissue for DNA extractions was generated by growing each isolate in clarified liquid V8 media (200 ml of Campbell's V8 Juice + 800 ml of water) at 125 rpm in 250-ml flasks for 5 days at room temperature under constant fluorescent illumination on a rotary shaker. Genomic DNA was extracted using the OmniPrep for Fungi extraction kit (G-Biosciences, St. Louis, MO, U.S.A.). Integrity of extracted DNA was verified on a 1% agarose gel prior to purification using the Genomic DNA Clean and Concentrator-10 kit (Zymo Research, Irvine, CA, U.S.A.). DNA was quantified using a Qubit 2.0 fluorometer (Invitrogen, Carlsbad, CA, U.S.A.) and genome libraries were constructed using TruSeq DNA Nano kits (Illumina Inc., San Diego, CA, U.S.A.). Libraries were quantified using the KAPA Library Quantification Kit (Kapa Biosystems, Wilmington, MA, U.S.A.) and sequenced using a 600-cycle cartridge (300 cycles each of paired forward and reverse reads) on the Illumina MiSeq platform (Illumina Inc.). Sequence data are deposited in the NCBI Sequence Read Archive under BioProject PRJNA631235.

Single-nucleotide polymorphism identification and genotyping.

A hybrid PacBio/Illumina genome assembly contained within 27 scaffolds from *C. pseudonaviculata* isolate CBS 139707 (also known as isolate cpsCT1) was used as the *C. pseudonaviculata* reference genome assembly (Malapi-Wight et al. 2019). Genes were predicted from the CBS 139707 assembly using the Funannotate v1.5.3 pipeline (Palmer 2017). Additional genome assemblies in this study were generated from paired-end reads using SPAdes v3.12.0 (Bankevich et al. 2012) (Supplementary Table S2).

To identify single-nucleotide polymorphisms (SNPs) in the 67 *C. pseudonaviculata* and 13 *C. henricotiae* isolates, paired-end reads were aligned to the CBS 139707 genome sequence using bwa v0.7.17 (Li and Durbin 2009). Paired-end reads from *C. henricotiae* isolates were also aligned to the *C. henricotiae* P-105865 genome assembly to evaluate potential impact of using a *C. pseudonaviculata* versus *C. henricotiae* genome as a reference for SNP identification. SNPs were identified from aligned paired-end reads using GATK v4.0 (McKenna et al. 2010) following recommendations outlined in the GATK best-practices guidelines (DePristo et al. 2011). Joint genotyping was performed for each species separately, except for estimation of interspecific genetic divergence. The R package vcfr (Knaus and Grünwald 2017; R Development Core Team 2015) and vcftools v0.1.17 (Danecek et al. 2011) were used to evaluate the distribution and quality of SNPs and determine optimal criteria for identification and removal of low-quality data. SNPs with the following criteria were filtered from the data: QUAL < 250, QD < 2, MQ < 50, and DP < 5. Because variant sites identified in the two species showed an order of magnitude difference in coverage, a maximum DP filter of 2,000 was used for *C. pseudonaviculata* and 800 for *C. henricotiae*. Additional filtering removed nonbiallelic SNPs and SNPs with missing data in at least one isolate for each species (see Supplementary Table S3 for summary of SNP type prior to final filtering step). SNP location and effects relative to predicted genes in the *C. pseudonaviculata* CBS 139707 genome were determined using SnpEff 4.3 (Cingolani et al. 2012).

Additional genome assemblies that were generated from paired-end sequence data with SPAdes (Supplementary Table S2) were used for SSR genotyping. Allelic variation at 11 previously identified SSR loci (LeBlanc et al. 2019) was identified as described by Castroagudín et al. (in press) or by aligning SSR locus sequence data (GenBank accessions KJ396041 to KJ396051) to genome assemblies using BLASTn (Altschul et al. 1997).

Summary statistics of *C. pseudonaviculata* and *C. henricotiae*. Summary statistics were calculated using vcftools v0.1.17 (Danecek et al. 2011). SNP density, Tajima's D, and nucleotide diversity (π) were calculated using 100-kb nonoverlapping bins. Statistics were calculated from final filtered SNP datasets that only included biallelic SNPs without missing data. Differentiation between the two species and genetic clusters in *C. pseudonaviculata* was estimated using the fixation index (F_{ST}) (Hudson et al. 1992) with the program DnaSP v6.12.03 (Rozas et al. 2017). Shared polymorphisms between species were identified using bcftools (Li et al. 2009). The standardized index of association ($\bar{r}_d$) was used to test for random association among loci using the package poppr v2.8.3 (Kamvar et al. 2014). The $\bar{r}_d$ estimates were compared with a null distribution generated from 999 replications of the data. The pairwise homoplasy index (PHI) (Bruen et al. 2006) was calculated to test for evidence of recombination using SplitsTree v4.16.1 (Huson and Bryant 2006). Phylogenetic networks were generated to visualize reticulations among *C. pseudonaviculata* and *C. henricotiae* isolates using the neighbor-net algorithm implemented through SplitsTree. All calculations were made for each species and geographic populations of *C. pseudonaviculata*.

Genetic structure and phylogenetic clades in *C. pseudonaviculata*. Additional population genetic analyses focused on *C. pseudonaviculata*, the most widely geographically distributed of the two boxwood blight species. Discriminate analysis of principal components (DAPC) was used to test for genetic clusters independent of geographic population, with the R package adegenet v2.1.1 (Jombart 2008). The optimal number of clusters was determined based on the Bayesian information criterion (BIC) calculated for 1 to 20 clusters. The optimal number of principal components to retain was determined using DAPC cross validation. Final posterior probabilities of group assignments were based on retaining 10 axes in the principal component analysis and two axes in the discriminate analysis. Maximum-likelihood phylogenies were inferred from SNPs using RAxML v8.0.0 (Stamatakis 2014). Trees were estimated using the general time reversible model of substitution with a γ model of rate heterogeneity. Because the alignment contained only variable sites, an ascertainment bias correction implemented through RAxML was used for likelihood calculations. Confidence in clades was assessed based on 100 bootstrap replicates.

Historical divergence scenarios of *C. pseudonaviculata* genetic clusters. Models of historical divergence among *C. pseudonaviculata* clusters were conducted using approximate Bayesian computation implemented in the program DIYABC v2.1.0 (Cornuet et al. 2014). Analyses of SNP data were conducted without subdivision by the 27 supercontigs in the CBS 139707 assembly. To obtain optimal overlap of simulated with observed data, all summary statistics were calculated for each cluster and cluster combination. Single-cluster statistics included gene diversity (Nei 1987) and pairwise distance metrics included F_{ST} (Weir and Cockerham 1984) and Nei's genetic distance (Nei 1972). In addition, maximum-likelihood coefficients of admixture were calculated using the maximum-likelihood method adapted from Choisy et al. (2004) and implemented in DIYABC. Admixture statistics were calculated for each pairwise combination of genetic clusters as parents to a third admixed cluster.

Specific divergence scenarios were chosen to test hypotheses regarding the origin of the primary genetic cluster (cluster B as determined by DAPC) found in the United States. In total, 22 scenarios were evaluated to test five main hypotheses (Supplementary Fig. S3). The hypotheses included (i) the four genetic clusters diverged at the same time (scenarios 1 and 2); (ii) cluster B was parent to clusters A, C, and D (scenarios 3 and 4); (iii) cluster A, C, or D was parent to the three remaining clusters (scenarios 5 and 6, 11 and 12, and 17 and 18); (iv) cluster A, C, or D was parent to the three remaining clusters and cluster B diverged earlier from the same

parent cluster (scenarios 7 and 8, 13 and 14, and 19 and 20); and (v) cluster A, C, or D was parent to two genetic clusters and an unsampled population was parent to genetic cluster B (scenarios 9 and 10, 15 and 16, and 21 and 22). All scenarios assumed an unidentified ancestral population. The timing and relative order of cluster divergence in the scenarios were tested with and without population bottlenecks.

Reference tables were generated from a total 2,000,000 simulated datasets using uniform priors (see Supplementary Table S5 for additional information on scenario parameters and prior distributions). Weighted polychotomous logistic regression with linear discriminate analysis of the summary statistics was used to calculate posterior probabilities and rank the scenarios (Cornuet et al. 2008). Confidence in identification of the most likely divergence scenario was estimated using posterior-based error calculations using 1,000 test pseudo-observed datasets.

RESULTS

Genomewide SNPs in *C. pseudonaviculata* and *C. henricotiae*. Genomewide SNPs were identified in 80 isolates of *C. henricotiae* and *C. pseudonaviculata* collected from known geographic locations (Supplementary Table S1). On average, 8.2 million reads per isolate were aligned to the *C. pseudonaviculata* CBS 139707 genome assembly (range = 3.7 million to 21.8 million reads), with an average genomewide coverage of 44 $\times$ (range = 21 $\times$ to 116 $\times$). In total, 1,608 SNPs that met our established quality threshold were identified from the 67 *C. pseudonaviculata* isolates when read data were aligned to the *C. pseudonaviculata* CBS 139707 reference genome assembly (Table 1). Alignment of *C. henricotiae* read data from 13 isolates to the *C. pseudonaviculata* CBS 139707 assembly identified 1,017 SNPs that were polymorphic in *C. henricotiae* isolates. All of the 1,017 *C. henricotiae* SNPs were invariant in the 67 *C. pseudonaviculata* isolates. Alignment of *C. henricotiae* data to the *C. henricotiae* P-10-5865 genome assembly identified 450 SNPs. Multiple SNPs in each species were predicted to have functional effects through generation or loss of stop codons and nonsynonymous mutations (Supplementary Table S4). Analysis of SNP density of the 67 *C. pseudonaviculata* isolates showed an average of three SNPs per 100 kb across 22 contigs of the CBS 139707 assembly, whereas SNPs from the 13 *C. henricotiae* isolates spanned 23 contigs with an average density of three SNPs per 100 kb (Supplementary Fig. S1). Comparison of SNP datasets for the two species showed no overlap of sites segregating in either species. Based on the 398,223 SNPs identified from the 80 isolates using the *C. pseudonaviculata* CBS 139707 genome, *C. pseudonaviculata* and *C. henricotiae* showed genetic differentiation with an F_{ST} value of 0.99.

Average nucleotide diversity (π) was higher in *C. henricotiae* than *C. pseudonaviculata* (Wilcoxon rank sum test; $W = 141,480$, $P < 0.001$) (Table 1). To account for potential bias from identifying SNPs in *C. henricotiae* using a *C. pseudonaviculata* genome assembly, data were also aligned to a genome assembly from isolate *C. henricotiae* P-10-5865 (573 scaffolds, $N_{50} = 211,275$). This analysis yielded 450 SNPs, and also showed that nucleotide diversity was significantly greater in *C. henricotiae* than *C. pseudonaviculata* (Wilcoxon rank sum test; $W = 119,500$, $P < 0.001$) (Table 1). Nucleotide diversity was significantly different among Asian, European, and U.S. populations of *C. pseudonaviculata* (Kruskal-Wallis test; $\chi^2 = 259.74$, $P < 0.001$), with the highest diversity found in the Asian population.

Tajima's D as well as measurement of random association among alleles and clonality were estimated at the species level and for geographic populations of *C. pseudonaviculata*, defined by continent of origin. Based on these statistics, both species and *C. pseudonaviculata* populations had negative Tajima's D values, indicating an excess of rare SNPs, but these values did not differ from a null expectation of neutral evolution based on 95%

confidence limits around $D = 0$ (Tajima 1989) (Table 1). The standardized index of association ($\bar{r}_d$) was calculated for each species and geographic populations of *C. pseudonaviculata* and compared with a null distribution to test for random association among alleles. Estimated values were all significantly different

from null distributions (Table 1), indicating a lack of evidence of random mating in both species. The PHI test did not reject the null hypothesis of clonality for *C. pseudonaviculata*. However, the PHI test rejected the null hypothesis of clonality for *C. henricotiae* (Table 1). Visualization of phylogenetic networks showed greater

TABLE 1. Genomewide summary statistics for *Calonectria pseudonaviculata* and *C. henricotiae*

Species, population	Isolates	SNPs ^a	Nucleotide diversity ^b	Tajima's D ^c	$\bar{r}_d$ ^d	PHI ^e
<i>C. pseudonaviculata</i>	67	1,608	0.0000018	-1.34	0.008*	0.05
Asia	6	210	0.0000057	-0.18	0.208*	0.022
Europe	20	644	0.0000024	-1.20	0.008*	0.1
United States	39	843	0.0000019	-1.20	0.022*	0.062
New Zealand	2	109	...	...	...	...
<i>C. henricotiae</i> ^f	13	1,017	0.0000088	-0.57	0.006*	0.594**
<i>C. henricotiae</i> ^g	13	450	0.0000029	-0.9	0.0123*	0.108**

^a Single-nucleotide polymorphisms (SNPs).
^b Nucleotide diversity (π) averaged over 100-kb nonoverlapping windows (Nei 1987).
^c Tajima's D averaged over 100-kb nonoverlapping windows (Tajima 1989).
^d Standardized index of association testing null hypothesis of random mating, where an asterisk (*) indicates $P < 0.01$.
^e Observed pairwise homoplasmy index (PHI) testing null hypothesis of clonality, where asterisks (**) indicate $P < 0.001$.
^f SNPs identified using *C. pseudonaviculata* CBS 139707 genome assembly (contigs = 27, N₅₀ = 3,534,399).
^g SNPs identified using *C. henricotiae* P-10-5865 genome assembly (contigs = 573, N₅₀ = 211,275).

Fig. 1. *Calonectria pseudonaviculata* genetic clusters, based on discriminate analysis of principal components. **A**, Scatter plot showing distance among and within the four genetic clusters identified using 1,608 single-nucleotide polymorphisms. **B**, Bayesian Information Criterion (BIC, y-axis) plotted for increasing number of clusters; lowest BIC values ($k = 4$) are highlighted in red. **C**, Posterior probabilities (y-axis) of each isolate (x-axis) being assigned to a given cluster ($k = 4$) arranged by isolate continent of origin (AS = Asia, EU = Europe, NZ = New Zealand, and US = United States). Absence of multiple colors per isolate indicates high (>0.99) posterior probabilities of group assignment.

population subdivision and relatively fewer reticulations within phylogenetic networks of *C. pseudonaviculata* compared with *C. henricotiae* (Supplementary Fig. S2A to C).

TABLE 2. Genetic differentiation and number of single-nucleotide polymorphisms (SNPs) among genetic clusters in *Calonectria pseudonaviculata*

Comparison	F_{ST}^a	SNPs ^b
Cluster A versus cluster B	0.310	113 (19)
Cluster A versus cluster C	0.377	104 (18)
Cluster A versus cluster D	0.185	99 (19)
Cluster B versus cluster C	0.511	130 (13)
Cluster B versus cluster D	0.212	97 (16)
Cluster C versus cluster D	0.396	117 (11)

^a Fixation index.

^b Average number of SNPs between genetic clusters. Standard deviations are shown in parentheses.

Genetic structure and phylogenetic clades in *C. pseudonaviculata*. Evidence of genetic structure in *C. pseudonaviculata* was tested using nonparametric DAPC. Comparison of BIC values for 1 to 20 clusters showed the greatest support for four genetic clusters (A to D), with 10 principal components retained. Scatter plots showed clear separation among the four clusters and posterior probabilities for each isolate showed high confidence for group assignments (>0.99) (Fig. 1). Cluster A was represented by 15 isolates from Europe, New Zealand, and the United States, including the first collections made from disease outbreaks occurring in France and Spain in 2009 and 2007, respectively. Also within this cluster were two isolates from the first detection of boxwood blight in 2011 in Oregon (ODA-HN1 and ODA1), and during 2016 in Washington state (JAC17-06) and California (CDFA666). Additional U.S. isolates in cluster A included collections made from the east coast, from Pennsylvania (326664BT) and Virginia (1218VA), both collected in 2018. Isolates from cluster B were collected from 2011 to 2018 and

Fig. 2. Heatmap showing number of pairwise single-nucleotide polymorphisms (SNPs) among *Calonectria pseudonaviculata* isolates. The SNP legend shows range from 0 to 168 SNPs between given isolate pairs. Dendrogram represents hierarchical clustering among isolates based on SNP data. Diagonal of the heatmap represents comparison of an isolate to itself. Cluster legend shows classification of isolates to genetic clusters identified using discriminate analysis of principal components.

represented the predominant group in the United States, including the first isolates sampled from the original U.S. boxwood blight outbreaks in North Carolina (NCBB1) and Connecticut (cpsCT12, cpsCT10, and CBS 139707) in 2011. Although cluster B was primarily composed of U.S. isolates, two isolates from Europe belonged to this group. Cluster C was represented by four isolates collected from native ecosystems around the Black Sea in Turkey and Iran. Cluster D contained two representative isolates from the first boxwood blight disease outbreaks that occurred in the United Kingdom in the 1990s (isolates RHS 10183 and RHS 9934). Additional isolates in cluster D were collected in New Zealand and from native

ecosystems in Turkey; no isolates from the United States were members of this cluster.

Pairwise comparisons among isolates showed variation in genetic differentiation and average number of SNPs among the four genetic clusters (Table 2). Differentiation was lowest between clusters A and D ($F_{ST} = 0.185$) and highest between clusters B and C ($F_{ST} = 0.511$). Overall, pairwise comparisons showed an average of 91 SNPs among isolates. The fewest number of SNPs between an isolate pair was 9 and the greatest was 168 (0.00002 to 0.00031% of the 54-Mb genome). Isolates within genetic clusters generally differed by fewer SNPs than comparisons between genetic clusters (Fig. 2).

Fig. 3. Maximum-likelihood phylogenetic tree of *Calonectria pseudonaviculata* isolates. Tree was inferred based on 1,608 single-nucleotide polymorphisms. Thickened branches indicate >80% bootstrap support. Rectangular legend shows colors that correspond to differing levels of bootstrap support. Black stars indicate isolates linked to first emergence of boxwood blight in the United Kingdom in 1990s and the United States in 2011. Isolates sampled from boxwood outside the United States are highlighted in gray text and country and year of origin are in parentheses. Isolates from the United States are highlighted in black text and country, state, and year of origin are in parentheses. Circular legends show regional origin of U.S. isolates. The four genetic clusters identified using discriminant analysis of principal components are highlighted by dashed ellipses.

Allelic variation at 11 polymorphic SSR loci indicated that two *C. pseudonaviculata* SSR multilocus genotypes (MLGs), SSR-G1 and SSR-G3, were shared between clusters A and D (Supplementary Table S1). All other SSR MLGs were members of a single cluster. *C. pseudonaviculata* isolates representing the second most frequent global SSR genotype, SSR-G2, were only classified in cluster B and the four SSR-G4 isolates were only classified in cluster C. Additional singleton SSR genotypes were classified in clusters B and D.

To determine the genetic relationships among *C. pseudonaviculata* isolates, SNP data were used to infer a maximum-likelihood phylogeny. Multiple well-supported clades were present in the phylogeny (Fig. 3). Isolates classified in DAPC clusters A, B, and C represented monophyletic clades with 100% support. Isolates in cluster D did not entirely fall within a single well-supported clade but were separate from clusters A, B, and C. Within the clade that corresponded to cluster B there were also multiple subclades supported by high-confidence branch support. These data provide additional evidence of genetic differentiation in *C. pseudonaviculata* beyond the four clusters identified using DAPC.

Divergence scenarios among *C. pseudonaviculata* genetic clusters. Approximate Bayesian computation was used to compare alternate scenarios of divergence among the four genetic clusters identified by DAPC and phylogenetic analysis. A primary goal in this analysis was to identify the possible origin of cluster B, the predominant group occurring in the United States (Figs. 1 and 3). Estimates of confidence in identifying the correct scenario showed that the logistic approach identified the correct scenario in approximately 88% of the 1,000 test datasets (posterior predictive error 0.12).

In all, 21 of the 22 tested scenarios had posterior probabilities < 0.12, and 13 had a posterior probability of 0 (Supplementary Fig. S3). Scenario 9 was the single high-scoring scenario, with a posterior probability of 0.867 (confidence interval = 0.675–1). In this scenario, cluster B diverged from an unsampled source population, while cluster D was parent to clusters A and C. This scenario also included changes in effective population sizes for cluster B and D following divergence from the unidentified ancestral population (Fig. 4). Further evaluation of effective population sizes indicated a greater decrease in effective population size of cluster B compared with cluster D (Supplementary Table S5).

Fig. 4. Most probable historical divergence scenarios among *Caloneectria pseudonaviculata* genetic clusters based on approximate Bayesian computation. Letters below lineages correspond to the four *C. pseudonaviculata* genetic clusters identified using discriminate analysis of principal components. Thinner lines indicate potential change in effective population size following divergence of cluster B and D from an unidentified ancestral population. Scale bar represents time from present (t_0) to past divergence (t_1 and t_2) and change in effective population size (t_{db}) events.

In this study, we employed population genomics to characterize genetic diversity and differentiation in the two closely related species that cause the globally emergent boxwood blight disease. Genomewide SNPs showed clear genetic differentiation between the two sister species *C. pseudonaviculata* and *C. henricotiae* that cause this disease. These findings support prior work that show interspecific differentiation using housekeeping genes (Gehesquière et al. 2016), SSR loci (LeBlanc et al. 2019), and divergence dating (Malapi-Wight et al. 2019). Focusing on *C. pseudonaviculata*, which is distributed globally in all areas of the world where the disease is known, we found that four clonal lineages evolved and migrated across continental and oceanic boundaries. Presence of two distinct *C. pseudonaviculata* lineages in the United States at the time of the initial disease outbreaks in 2011 provided compelling and strong evidence for multiple pathogen introductions. Comparison of divergence scenarios among clonal lineages indicated that the most common lineage found in the United States (cluster B) diverged from an unsampled population and the less frequent lineage in the United States (cluster A) diverged from a primarily European lineage. The outcomes of this research provide insight into boxwood blight emergence and a conceptual framework for high-resolution molecular epidemiology studies for these and other clonal fungi that cause emerging plant diseases.

The presence of multiple differentiated genetic lineages (i.e., genetic clusters and well-supported phylogenetic clades) in *C. pseudonaviculata* provides the first population genomics data to critically examine the emergence of boxwood blight disease in the United States. Based on these genetic data, two divergent clonal lineages were responsible for initial U.S. outbreaks of boxwood blight in 2011. One lineage was found on the west coast in Oregon (cluster A) and one lineage was found on the east coast in Connecticut and North Carolina (cluster B). Because both lineages are also found on other continents, the alternative scenario (i.e., introduction of a single lineage into the United States, followed by rapid divergence) is extremely unlikely. Instead, the most plausible scenario is that members of two distinct *C. pseudonaviculata* lineages were introduced into the United States prior to the first detections of boxwood blight in 2011. Additional sampling of *C. pseudonaviculata* since 2011 showed that isolates from cluster A first sampled in Oregon are now established on the east coast and isolates from cluster B first detected in North Carolina and Connecticut are now on the west coast in California. Due to small sample sizes and lack of comprehensive information documenting interstate pathogen movement, these data do not enable comprehensive analyses of domestic pathogen migration (Goss et al. 2009). However, considering the limited nonanthropogenic dispersal distance of pathogen propagules (Gehesquière et al. 2013), these data provide additional evidence for significant movement of the pathogen on infected boxwood material within the U.S. ornamental industry (Daughtrey 2019).

Evidence for at least two introductions of *C. pseudonaviculata* into the United States raises questions about the origins of these clonal lineages. Historical data provide evidence and suggests that boxwood blight first became established in Europe (including the United Kingdom) prior to its detection elsewhere in the world. This suggests that European pathogen populations may have served as the source for subsequent disease emergence on other continents. In support of a European origin of boxwood blight, the clonal lineages first detected in Europe were subsequently found in Asia, New Zealand, and the United States. However, simulation analysis of historical divergence scenarios showed that the primary clonal lineage found in the United States, cluster B, was derived from an unsampled population. Though this putative unsampled population could be in Europe, it could also be from the original unknown source population that served as the source for emergence of

boxwood blight in the United Kingdom in the 1990s. These results implicate European *C. pseudonaviculata* populations as sources of pathogen introductions but also suggest that populations of unknown origins may be responsible for historical expansion of boxwood blight disease and pathogen introduction and movement into new continents. From a practical perspective, evidence for multiple introductions and geographic origins of *C. pseudonaviculata* in the United States suggests that future introductions are likely to occur.

In agreement with prior studies (Castroagudín et al. in press; LeBlanc et al. 2019; Malapi Wight et al. 2019), genomewide SNPs show that populations of *C. pseudonaviculata* and *C. henricotiae* do not undergo random mating and have limited genetic diversity in their known geographic ranges. Strictly or predominantly asexual reproduction combined with lack of genetic diversity in plant pathogens is generally associated with reduced risk for disease management (McDonald and Linde 2002) and is common among introduced fungi (Gladieux et al. 2015). However, the level of risk associated with *C. pseudonaviculata* and *C. henricotiae* has to be understood from a broader perspective. As shown in this research, genetic diversity has arisen in these introduced species, and multiple clonal lineages of *C. pseudonaviculata* have evolved or been introduced since discovery of this pathogen in the 1990s (Henricot and Culham 2002). Phenotypic diversification has been documented during the range expansion of other introduced clonal fungi (Forsythe et al. 2018; Robin et al. 2017), which has yet to be determined for the boxwood blight fungi. Another consideration is the historical absence of selective pressures placed on *C. pseudonaviculata* and *C. henricotiae* populations due to a lack of host genetic resistance. The majority of boxwood plants grown in the landscape are the highly susceptible *Buxus sempervirens* cultivars, particularly the widely grown *Suffruticosa* (Batdorf 2004). Ongoing efforts to identify and deploy new sources of genetic resistance in host populations (Van Laere et al. 2019) are currently in progress and likely to introduce strong selective pressures on pathogen populations, and may influence the evolution of *C. pseudonaviculata* and *C. henricotiae*, even in the absence of sexual reproduction.

Only a single mating type idiomorph has been identified from the *MAT1* loci of *C. henricotiae* and *C. pseudonaviculata*, indicating limited opportunities for sexual recombination (Malapi-Wight et al. 2019) and consistent with the minimal genetic diversity and evidence for clonality found in this study. All known isolates of *C. henricotiae* have the *MAT1-1* mating type, whereas all known isolates of *C. pseudonaviculata* are mating type *MAT1-2* (Castroagudín et al. in press; Malapi-Wight et al. 2019). Departures from a 1:1 ratio of mating type idiomorphs is often observed in populations of heterothallic (i.e., nonselfing) fungi (Brasier 2001; Dung et al. 2013; Havenga et al. 2020; Kosanke et al. 2018; Milgroom et al. 2008). However, it is possible that populations of *C. henricotiae* and *C. pseudonaviculata* in their native ranges could harbor both mating types and undergo random mating (Saleh et al. 2012). Another question that has arisen based on the presence of opposite mating types in *C. henricotiae* and *C. pseudonaviculata* is whether or not the two species are reproductively isolated, because they currently share a common host in Europe and the United Kingdom. Based on an absence of shared polymorphisms and estimates of genetic differentiation, we found no evidence for genetic exchange between the *C. pseudonaviculata* and *C. henricotiae* isolates included in this study. Despite this lack of evidence, the potential for hybridization should not be ignored in future research. Even in the case where hybridization is rare and progeny may have reduced fitness, such events can have large impacts on the evolution of emerging pathogens. For example, rare hybridization events between fungal *Ophiostoma* spp. that cause Dutch elm disease have been linked to introgression of mating type loci across species boundaries (Brasier et al. 1998; Paoletti et al. 2006). Additional studies on mating type loci using phylogenetics

and heterologous expression suggest that transfer across species boundaries may also occur among other ascomycete fungi (Martin et al. 2013). If similar events were to occur in the boxwood blight pathogens, this could lead to the presence of both mating types in the same species in introduced ranges and increase the probability of sexual recombination.

To date, there is no evidence that *C. henricotiae* has expanded its range beyond continental Europe and the United Kingdom. This conclusion is supported by prior genotyping of boxwood pathogen isolates collected from 1998 to 2014 on multiple continents (LeBlanc et al. 2019) and ongoing sampling across the United States between 2014 and 2019 (Castroagudín et al. in press). Recent disease surveys in native ecosystems in Iran also failed to recover *C. henricotiae* (Khazaeli et al. 2018). However, considering the evidence for long-distance migration of *C. pseudonaviculata* found in this study, it seems unlikely that *C. henricotiae* will remain confined to Europe. Further range expansion of *C. henricotiae* even within Europe will have significant consequences on the ability to manage boxwood blight disease due to the reduced fungicide sensitivity of this pathogen species (Gehesquière et al. 2016). Range expansion of *C. henricotiae* would also influence boxwood blight epidemiology in the landscape due to differences in temperature growth optima of these species and associated disease development (Avenot et al. 2017). Altogether, the risk associated with *C. henricotiae* migration warrants ongoing surveys of this boxwood blight pathogen species in the U.S. ornamental industry.

Because introduced fungal pathogens often display reduced population genetic diversity and differentiation due to genetic bottlenecks and lack of sexual recombination (Gladieux et al. 2015), traditional population genetic markers (e.g., multilocus sequence typing or SSRs) may fail to capture informative mutations in pathogen populations. In this work, SNPs represented <0.003% of an estimated 54-Mb genome in the primary pathogen species. Despite being rare, these mutations identified at least four differentiated clonal lineages that have emerged since initial detection of boxwood blight disease in the 1990s. These results provide a striking contrast to previous SSR genotyping, which found no evidence of genetic structure and near genotypic uniformity in *C. pseudonaviculata*, the primary boxwood blight pathogen (Castroagudín et al. in press; LeBlanc et al. 2019). The broader implications of this research are that population genomic data are informative for understanding the dynamics and evolution of highly clonal or recently introduced fungal pathogens. As shown in this and other research, these data can inform historical scenarios of disease emergence as well as verify appropriate species boundaries (Gladieux et al. 2018; Goss 2015) that have significant practical applications for regulating global movement of emerging fungal plant pathogens.

ACKNOWLEDGMENTS

We thank individuals who contributed to boxwood blight sample collections. Mention of trade names or commercial products in this publication is solely for the purpose of providing specific information and does not imply recommendation or endorsement by the U.S. Department of Agriculture. The USDA is an equal opportunity provider and employer.

LITERATURE CITED

- Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, Z., Miller, W., and Lipman, D. 1997. Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. *Nucleic Acids Res.* 25:3389-3402.
- Anderson, P. K., Cunningham, A. A., Patel, N. G., Morales, F. J., Epstein, P. R., and Daszak, P. 2004. Emerging infectious diseases of plants: Pathogen pollution, climate change and agrotechnology drivers. *Trends Ecol. Evol.* 19:535-544.
- Avenot, H. F., King, C., Edwards, T. P., Baudoin, A., and Hong, C. X. 2017. Effects of inoculum dose, temperature, cultivar, and interrupted leaf wetness period on infection of boxwood by *Calonectria pseudonaviculata*. *Plant Dis.* 101:866-873.

- Bankevich, A., Nurk, S., Antipov, D., Gurevich, A. A., Dvorkin, M., Kulikov, A. S., Lesin, V. M., Nikolenko, S. I., Pham, S., Prjibelski, A. D., Pyshtkin, A. V., Sirotkin, A. V., Vyahhi, N., Tesler, G., Alekseyev, M. A., and Pevzner, P. A. 2012. SPAdes: A new genome assembly algorithm and its applications to single-cell sequencing. *J. Comput. Biol.* 19:455-477.
- Batdorf, L. 2004. Boxwood: An Illustrated Encyclopedia. American Boxwood Society, Boyce, VA, U.S.A.
- Brasier, C. M. 2001. Rapid evolution of introduced plant pathogens via interspecific hybridization: Hybridization is leading to rapid evolution of Dutch elm disease and other fungal plant pathogens. *Bioscience* 51:123-133.
- Brasier, C. M. 2008. The biosecurity threat to the UK and global environment from international trade in plants. *Plant Pathol.* 57:792-808.
- Brasier, C. M., Kirk, S. A., Pipe, N. D., and Buck, K. W. 1998. Rare interspecific hybrids in natural populations of the Dutch elm disease pathogens *Ophiostoma ulmi* and *O. novo-ulmi*. *Mycol. Res.* 102:45-57.
- Bruen, T. C., Philippe, H., and Bryant, D. 2006. A simple and robust statistical test for detecting the presence of recombination. *Genetics* 172:2665-2681.
- Castroagudín, V. L., Weiland, J. E., Baysal-Gurel, F., Cubeta, M., Daughtrey, M., Ward, N. A., LaMondia, J., Luster, D. G., Hand, F. P., Shishkoff, N., Williams-Woodward, J., Yang, X., LeBlanc, N., and Crouch, J. A. One clonal lineage of *Calonectria pseudonaviculata* is primarily responsible for the boxwood blight epidemic in the United States. *Phytopathology* In press. doi.org/10.1094/PHYTO-04-20-0130-R
- Castroagudín, V. L., Yang, X., Daughtrey, M. L., Luster, D. G., Pscheidt, J. W., Weiland, J. E., and Crouch, J. A. 2020. Boxwood blight disease: A diagnostic guide. *Plant Health Prog.* 21:291-300.
- Choisy, M., Franck, P., and Cornuet, J. M. 2004. Estimating admixture proportions with microsatellites: Comparison of methods based on simulated data. *Mol. Ecol.* 13:955-968.
- Cingolani, P., Platts, A., Wang, L. L., Coon, M., Nguyen, T., Wang, L., Land, S. J., Lu, X., and Ruden, D. M. 2012. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff. *Fly* 6:80-92.
- Cornuet, J. M., Pudlo, P., Veyssier, J., Dehne-Garcia, A., Gautier, M., Leblois, R., and Estoup, A. 2014. DIYABC v2.0: A software to make approximate Bayesian computation inferences about population history using single nucleotide polymorphisms, DNA sequence and microsatellite data. *Bioinformatics* 30:1187-1189.
- Cornuet, J. M., Santos, F., Beaumont, M. A., Robert, C. P., Marin, J. M., Balding, D. J., Guillemaud, T., and Estoup, A. 2008. Inferring population history with DIY ABC: A user-friendly approach to approximate Bayesian computation. *Bioinformatics* 24:2713-2719.
- Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., Handsaker, R. E., Lunter, G., Marth, G. T., Sherry, S. T., McVean, G., Durbin, R., and 1000 Genomes Project Analysis Group. 2011. The variant call format and VCFtools. *Bioinformatics* 27:2156-2158.
- Daughtrey, M. 2019. Boxwood blight: Threat to ornamentals. *Annu. Rev. Phytopathol.* 57:189-209.
- DePristo, M. A., Banks, E., Poplin, R., Garimella, K. V., Maguire, J. R., Hartl, C., Philippakis, A. A., del Angel, G., Rivas, M. A., Hanna, M., McKenna, A., Fennell, T. J., Kernysky, A. M., Sivachenko, A. Y., Cibulskis, K., Gabriel, S. B., Altshuler, D., and Daly, M. J. 2011. A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat. Genet.* 43:491-498.
- Dung, J. K. S., Peever, T. L., and Johnson, D. A. 2013. *Verticillium dahlia* populations from mint and potato are genetically divergent with predominant haplotypes. *Phytopathology* 103:445-459.
- Fisher, M. C., Henk, D. A., Briggs, C. J., Brownstein, J. S., Madoff, L. C., McCraw, S. L., and Gurr, S. J. 2012. Emerging fungal threats to animal, plant and ecosystem health. *Nature* 484:186-194.
- Forsythe, A., Giglio, V., Asa, J., and Xu, J. 2018. Phenotypic divergence along geographic gradients reveals potential for rapid adaptation of the white-nose syndrome pathogen, *Pseudogymnoascus destructans*, in North America. *Appl. Environ. Microbiol.* 84:e00863-18.
- Gehesquière, B., Crouch, J. A., Marra, R. E., Van Poucke, K., Rys, F., Maes, M., Gobin, B., Höfte, M., and Heungens, K. 2016. Characterization and taxonomic reassessment of the box blight pathogen *Calonectria pseudonaviculata*, introducing *Calonectria henricotiae* sp. nov. *Plant Pathol.* 65:37-52.
- Gehesquière, B., D'Haeyer, S., Pham, K. T. K., Van Kuik, A. J., Maes, M., Höfte, M., and Heungens, K. 2013. qPCR assays for the detection of *Cylindrocladium buxicola* in plant, water, and air samples. *Plant Dis.* 97:1082-1090.
- Gladieux, P., Condon, B., Ravel, S., Soanes, D., Maciel, J. L. N., Nhani, A., Chen, L., Terauchi, R., Lebrun, M.-H., Tharreau, D., Mitchell, R., Pedley, K. F., Valent, B., Talbot, N. J., Farman, M., and Fournier, E. 2018. Gene flow between divergent cereal- and grass-specific lineages of the rice blast fungus *Magnaporthe oryzae*. *MBio* 9:e01219-17.
- Gladieux, P., Feurtey, A., Hood, M. E., Snir, A., Clavel, J., Dutech, C., Roy, M., and Giraud, T. 2015. The population biology of fungal invasions. *Mol. Ecol.* 24:1969-1986.
- Goss, E. M. 2015. Genome-enabled analysis of plant-pathogen migration. *Annu. Rev. Phytopathol.* 53:121-135.
- Goss, E. M., Larsen, M., Chastagner, G. A., Givens, D. R., and Grünwald, N. J. 2009. Population genetic analysis infers migration pathways of *Phytophthora ramorum* in US nurseries. *PLoS Pathog* 5:e1000583.
- Havenga, M., Wingfield, B. D., Wingfield, M. J., Roets, F., Dreyer, L. L., Tatham, C. T., Duong, T. A., Wilken, P. M., Chen, S. C., and Aylward, J. 2020. Mating strategy and mating type distribution in six global populations of the *Eucalyptus* foliar pathogen *Teratosphaeria destructans*. *Fungal Genet. Biol.* 137:103350.
- Henricot, B., and Culham, A. 2002. *Cylindrocladium buxicola*, a new species affecting *Buxus* spp., and its phylogenetic status. *Mycologia* 94:980-997.
- Hudson, R. R., Slatkin, M., and Maddison, W. P. 1992. Estimation of levels of gene flow from DNA sequence data. *Genetics* 132:583-589.
- Huson, D. H., and Bryant, D. 2006. Application of phylogenetic networks in evolutionary studies. *Mol. Biol. Evol.* 23:254-267.
- Jombart, T. 2008. *adeigenet*: A R Package for the multivariate analysis of genetic markers. *Bioinformatics* 24:1403-1405.
- Kamvar, Z. N., Tabima, J. F., and Grünwald, N. 2014. *Poppr*: An R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction. *PeerJ* 2:e281.
- Khazaeli, P., Rezaee, S., Mirabolfathy, M., Zamanizadeh, H., and Kiadaliri, H. 2018. Genetic and phenotypic variation of *Calonectria pseudonaviculata* isolates causing boxwood blight disease in the Hyrcanian forest of Iran. *Agric. Res. Technol. Open Access J.* 19:556081.
- Knaus, B. J., and Grünwald, N. 2017. vcfr: A package to manipulate and visualize variant call format data in R. *Mol. Ecol. Resour.* 17:44-53.
- Kosanke, S., Hamann, L., Kupsch, C., Garcia, S. M., Chopra, A., and Gräser, Y. 2018. Unequal distribution of the mating type (*MAT*) locus idiomorphs in dermatophyte species. *Fungal Genet. Biol.* 118:45-53.
- LeBlanc, N., Gehesquière, B., Salgado-Salazar, C., Heungens, K., and Crouch, J. A. 2019. Limited genetic diversity across pathogen populations responsible for the global emergence of boxwood blight identified using SSRs. *Plant Pathol.* 68:861-868.
- LeBlanc, N., Salgado-Salazar, C., and Crouch, J. A. 2018. Boxwood blight: An ongoing threat to ornamental and native boxwood. *Appl. Microbiol. Biotechnol.* 102:4371-4380.
- Li, H., and Durbin, R. 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 25:1754-1760.
- Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., and 1000 Genome Project Data Processing Subgroup. 2009. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 25:2078-2079.
- Liebold, A. M., Brockerhoff, E. G., Garrett, L. J., Parke, J. L., and Britton, K. O. 2012. Live plant imports: The major pathway for forest insect and pathogen invasions of the US. *Front. Ecol. Environ.* 10:135-143.
- Malapi-Wight, M., Veltri, D., Gehesquière, B., Heungens, K., Rivera, Y., Salgado-Salazar, C., and Crouch, J. A. 2019. Global distribution of mating types shows limited opportunities for mating across populations of fungi causing boxwood blight disease. *Fungal Genet. Biol.* 131:103246.
- Martin, S. H., Steenkamp, E. T., Wingfield, M. J., and Wingfield, B. D. 2013. Mate-recognition and species boundaries in the ascomycetes. *Fungal Divers.* 58:1-12.
- McDonald, B. A., and Linde, C. 2002. Pathogen population genetics, evolutionary potential, and durable resistance. *Annu. Rev. Phytopathol.* 40:349-379.
- McDonald, B. A., and Stukenbrock, E. H. 2016. Rapid emergence of pathogen in agro-ecosystems: Global threats to agricultural sustainability and food security. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 371:20160026.
- McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernysky, A., Garimella, K., Altshuler, D., Gabriel, S., Daly, M., and DePristo, M. A. 2010. The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* 20:1297-1303.
- Milgroom, M. G., Sotirovski, K., Spica, D., Davis, J. E., Brewer, M. T., Milev, M., and Cortesi, P. 2008. Clonal population structure of the chestnut blight fungus in expanded ranges in southeastern Europe. *Mol. Ecol.* 17:4446-4458.
- Miller, M. E., Shishkoff, N., and Cubeta, M. A. 2018. Thermal sensitivity of *Calonectria henricotiae* and *Calonectria pseudonaviculata* conidia and microsclerotia. *Mycologia* 110:546-558.
- Nei, M. 1972. Genetic distance between populations. *Am. Nat.* 106:283-292.
- Nei, M. 1987. *Molecular Evolutionary Genetics*. Columbia University Press, New York, NY, U.S.A.
- Palmer, J. M. 2017. Funannotate: A fungal genome annotation and comparative genomics pipeline. <https://funannotate.readthedocs.io/en/latest/>
- Paoletti, M., Buck, K. W., and Brasier, C. M. 2006. Selective acquisition of novel mating type and vegetative incompatibility genes via interspecies gene transfer in the globally invading eukaryote *Ophiostoma novo-ulmi*. *Mol. Ecol.* 15:249-262.

- R Development Core Team. 2015. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.
- Robin, C., Andanson, A., Saint-Jean, G., Fabreguettes, O., and Dutech, C. 2017. What was old is new again: Thermal adaptation within clonal lineages during range expansion in a fungal pathogen. *Mol. Ecol.* 26:1952-1963.
- Rozas, J., Ferrer-Mata, A., Sánchez-DelBarrio, J. C., Guirao-Rico, S., Librado, P., Ramos-Onsins, S. E., and Sánchez-Gracia, A. 2017. DnaSP 6: DNA sequence polymorphism analysis of large data sets. *Mol. Biol. Evol.* 34:3299-3302.
- Sakai, A. K., Allendorf, F. W., Holt, J. S., Lodge, D. M., Molofsky, J., With, K. A., Ellstrand, N. C., McCauley, D. E., O'Neil, P., Parker, I. M., Thompson, J. N., and Weller, S. G. 2001. The population biology of invasive species. *Annu. Rev. Ecol. Evol. Syst.* 32:305-332.
- Saleh, D., Xu, P., Shen, Y., Li, C., Adreit, H., Milazzo, J., Ravigné, V., Bazin, E., Nottéghem, J. L., Fournier, E., and Tharreau, D. 2012. Sex at the origin: An Asian population of the rice blast fungus *Magnaporthe oryzae* reproduces sexually. *Mol. Ecol.* 21:1330-1344.
- Stamatakis, A. 2014. RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 30:1312-1313.
- Stravoravdis, S., LeBlanc, N. R., Marra, R. E., Crouch, J. A., and Hulvey, J. P. 2020. Widespread occurrence of a CYP51A pseudogene in *Calonectria pseudonaviculata*. *Mycobiology* 48:44-50.
- Tajima, F. 1989. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* 143:1457-1465.
- Van Laere, K., Heungens, K., Gehesquière, B., Leus, L., Hermans, D., and Van Huylenbroeck, J. 2019. Breeding and selection of *Buxus* for resistance to *Calonectria pseudonaviculata*. *J. Phytopathol.* 167:363-370.
- Weir, B. S., and Cockerham, C. C. 1984. Estimating F-statistics for the analysis of population structure. *Evolution* 38:1358-1370.