Population structure and genetic diversity suggest recent introduction of *Dothistroma pini* in Slovakia

Katarína Adamčíková[1*], Zuzana Jánošíková[1], Ariska van der Nest[2], Slavomír Adamčík[3], Emília Ondrušková[1], Irene Barnes[2]

[1]Department of Plant Pathology and Mycology, Institute of Forest Ecology Slovak Academy of Sciences, Nitra, Slovak Republic

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

[3]Institute of Botany, Plant Science and Biodiversity Center, Slovak Academy of Sciences, Bratislava, Slovak Republic

* Email: katarina.adamcikova@ife.sk

KEYWORDS ascomycete fungus, multiplex PCR, clonality, microsatellites, sexual recombination

Abstract

Dothistroma pini is one of two pathogens causing the foliar disease of pines: Dothistroma needle blight (DNB). The species was re-defined in 2004 and subsequently recorded in several European countries. In Slovakia, the first report of the pathogen was in 2013. In this study, the population structure, genetic diversity and reproductive mode of 105 isolates collected from 10 localities and seven hosts were determined in Slovakia. Species-specific mating type markers, ITS haplotype determination, and sixteen microsatellite markers were used to characterise and genotype the isolates. Overall, 15 unique multilocus haplotypes (MLH) based on microsatellite markers, and three ITS haplotypes were identified. Three independent methods (DAPC, STRUCTURE, EDENetwork) separated the isolates into two distinct population clusters corresponding with ITS haplotypes. A high level of clonality was recorded suggesting that conidia are the primary source of pathogen dispersal. The low genetic diversity, predominantly asexual reproductive mode of the pathogen, and the fact that most isolates were collected from introduced tree species, and native species in artificially planted urban greenery, supports the hypothesis that *D. pini* has been recently introduced into Slovakia.

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1111/ppa.13428

1 INTRODUCTION

Dothistroma needle blight (DNB) is a serious needle disease of conifers that primarily affects pine species (*Pinus* spp.) in several countries across most continents (Drenkhan *et al*., 2016). The disease is characterized by red bands surrounding black, erumpent conidiomata that split the epidermal layers of infected pine needles. These needles become necrotic, are cast, and after successive defoliation, the disease can result in stunted tree growth, and in extreme cases tree death. Needles of all ages are commonly affected (Kowalski & Jankowiak, 1998).

Comparison of DNA sequence data has clearly shown that this disease is caused by two phylogenetically distinct species: *Dothistroma septosporum* (Dorog.) Morelet and *Dothistroma pini* Hulbary (Barnes *et al*., 2004; 2016). *Dothistroma septosporum* is reported to have a worldwide distribution, broad host range (Drenkhan *et al*., 2016) and numerous population genetics studies using microsatellite markers have shown high levels of gene and haplotypic diversity, especially in certain parts of Eurasia and North America (Drenkhan *et al*., 2016). Evolutionary modelling using microsatellite markers currently supports a Eurasian origin of the pathogen (Mullett *et al*., 2021). Analyses using whole genome sequencing are also improving genetic classification and population analyses. The calculation of genomic nucleotide diversity was especially essential to distinguish very closely related groups in the case of *D. septosporum* populations from Scotland and western Canada and helped to clarify the population´s history (Ennos *et al*., 2020; Capron *et al*., 2021).

Since 2004, when *D. pini* was recognised as a separate species to *D. septosporum* (Barnes *et al*., 2004), the pathogen has been reported in 14 European countries (Drenkhan *et al*., 2016; Ondrušková *et al*., 2017; Matsiakh *et al*., 2018). The long-term presence of *D. pini* in Europe has also been confirmed by re-evaluation of historical herbarium specimens. For example, Fabre *et al*. (2012) detected this pathogen in herbarium material dating back to 1907 and 1965 in France, indicating that this species is not a recent invader, at least in some parts of Europe. Besides Europe, *D. pini* was recorded in North America, where it was confirmed in north-central USA (Michigan, Nebraska, Minnesota, South and North Dakota, Indiana) affecting five pine species (Barnes *et al*., 2004; 2014) as well as in the southern USA in Arkansas (Mullett *et al.,* 2018).

The Internal Transcribed Spacer (ITS) region remains the barcode of choice for species identification when sequencing *Dothistroma* isolates (Barnes *et al*., 2016). In this regard, variation has been observed in this region with six different haplotypes being recorded for *D. pini* based on

worldwide collections. The six haplotypes designated as Dp_HAP.1, Dp_HAP.2, Dp_HAP.3, Dp_HAP.4, Dp_HAP.5 and Dp_HAP.6 differ either based on point mutations or single nucleotide insertions (Barnes *et al*., 2016; Mullett *et al*., 2018). Only three of them, Dp_HAP.1, Dp_HAP.2 and Dp_HAP.4 have been reported in Europe (Barnes *et al*., 2016; Mullett *et al*., 2018).

In heterothallic fungal species, such as *D. pini*, mating type genes play an important part in their biology and evolution (McDonald & Linde, 2002). Thus, knowledge of these genes (its presence and frequency) can provide insight into the potential prevalence of the reproduction mode (sexual vs asexual) in different species and affect the genetic and genotypic diversity of species population by sexual recombination (Milgroom, 1996). While the presence of both mating type idiomorphs up to equal frequencies indicates the occurrence of sexual reproduction, the presence of a single idiomorph in a population suggests asexual reproduction is taking place (Barnes *et al*., 2014). Both mating types of *D. pini* were detected in Slovakia (Ondrušková *et al*., 2018), Ukraine, France (Siziba *et al*., 2016), Switzerland (Queloz *et al*., 2014), Slovenia (Drenkhan *et al*., 2016) and Spain (Ortíz de Urbina *et al*., 2017) within Europe. In addition, in Switzerland, both mating types have been detected on the same needle (Queloz *et al*., 2014). Despite these reports, the sexual state of *D. pini* has not yet been described (Drenkhan *et al*., 2016).

Microsatellite markers for *D. pini* have been developed only recently and could be used to better understand the possible origin and patterns of pathogen spread (Siziba *et al*., 2016). The markers were tested on 24 isolates both from France and Slovenia. The findings indicated that the *D. pini* population from La Fert-Imbault in France was genetically more diverse than the Slovenian population from Pivka and that the latter population was completely clonal (Siziba *et al*., 2016). No other population studies, including determining the contribution of sexual recombination on the genetic diversity of *D. pini,* have been conducted.

Presently DNB is widespread in Slovakia, occurring in all regions and affecting 11 pine and two spruce species in various types of artificial (urban greenery, arboretum, Christmas tree plantations, nursery) and naturally regenerated stands (Ondrušková *et al*., 2017; 2018; 2020; Jánošíková-Hečková *et al*., 2018).

Both species of *Dothistroma* have been confirmed in Slovakia (Ondrušková *et al*., 2017; 2018; Jánošíková-Hečková *et al*., 2018). *Dothistroma septosporum* was identified across the entire country with a six-fold higher frequency than *D. pini* while the host species range was comparable. Both mating types of *D. septosporum* were identified in Slovakia, with different distributions and ratios in individual localities (Barnes *et al*., 2014; Ondrušková *et al*., 2017). The

even distribution of both mating types and the high genotypic diversity suggests sexual recombination plays a significant role in the pathogen's life cycle in Slovakia (Jánošíková *et al*., 2021). The high genetic diversity of *D. septosporum* suggests the pathogen has been established in Slovakia for a long period of time (Jánošíková *et al*., submitted). Such information has not yet been determined for *D. pini*.

Dothistroma pini was first identified in Slovakia using molecular methods from samples collected in 2013-2015 (Ondrušková *et al*., 2017; 2018; Jánošíková-Hečková *et al*., 2018). The Slovakian records expanded the known worldwide host range of this pathogen from 12 reported host taxa (Drenkhan *et al*., 2016) to 18, including *Picea abies* (Jánošíková-Hečková *et al*., 2018). Due to its similarity in symptoms with *D. septosporum*, it is not known whether the pathogen is also well established in the country or if may have been recently introduced as was suggested by Ondrušková *et al*. (2018) due to the presence of only one mating type per locality studied. The aim of this study was, therefore, to expand on the collections made by Ondrušková *et al*. (2018) with more intensive sampling across the country, and then to determine the genetic diversity, population structure and primary reproductive mode of the pathogen in Slovakia.

2 MATERIALS AND METHODS

2.1 Sample collection and fungal isolation

The selected locations for sample collection were sites where *D. pini* were previously identified (Jánošíková – Hečková *et al*., 2018; Ondrušková *et al*., 2018). In addition, the samples and isolates collected from the studies of Jánošíková – Hečková *et al*. (2018) and Ondrušková *et al*. (2018) were also included in this study. The sampling of pine needles with typical DNB symptoms was carried out from 2013 to 2017. Samples were collected with the same sampling strategy as published by Jánošíková – Hečková *et al*. (2018). Each sample consisted of at least 30 needles that were collected randomly from different parts of the tree crown. One sample was collected from each tree. When five or fewer affected trees were present per stand, needles were collected from each symptomatic tree and treated as a separate sample. When more than five trees were present in a stand, a minimum of five randomly selected trees were sampled. In the case of arboreta and urban greenery (trees planted as ornamentals in towns), attention was paid to register and collect samples from all affected host species. Samples were placed in a −20°C freezer until further processing.

Fungal cultures were obtained as described by Ondrušková *et al*. (2017). A minimum of one isolate was obtained per sample/tree, and a maximum of five isolates per sample depending on how well developed the conidiomata were for spore release, and the isolation of pure cultures from the isolation procedure. Cultures used in this study are maintained in the culture collection of IFE SAS, Department of Plant Pathology and Mycology, Nitra, Slovakia, and some duplicate representatives are simultaneously held in the culture collection of the Forestry and Agricultural Biotechnology Institute, Pretoria, South Africa (Table S1). Herbarium specimens of samples were deposited in the Plant Pathology Herbarium (NR) of the Institute of Forest Ecology of Slovak Academy of Sciences (IFE SAS), Nitra, Slovakia. In cases where the location consisted of a big sampling area, needles of the same host and date of sampling within each stand were pooled together in one submission (Table 1).

2.2 Investigation of historical samples in herbaria

To determine if there were any historical records of DNB, or deposits of infected material in herbaria, the herbarium specimens of Slovak origin were checked in three local herbaria: SAV (Institute of Botany, Slovak Academy of Sciences Herbarium, Bratislava), BRA (Slovak National Museum, Bratislava), NR (Plant Pathology Herbarium of the IFE SAS), and in three international ones: BP (Hungarian Natural History Museum, Budapest, Hungary), BRNM (Moravian Museum, Brno, Czech Republic) and PRM (National Museum, Praha, Czech Republic). During the search, all names associated with both DNB pathogens (current, old names, synonyms etc. as summarised in Barnes *et al*. (2016)) were used.

2.3 DNA extraction, sequencing and ITS haplotype identification

DNA was extracted from pure fungal cultures after incubation for ca. 3 weeks at 20°C in 3% MEA using the E.Z.N.A. Fungal DNA Mini-Kit (Omega Bio-Tek Inc., Norcross, GA, USA), following the manufacturer´s instructions. Identification of isolates to species level was done by sequencing the ITS region of the ribosomal DNA with primers ITS1F and ITS4 (White *et al*., 1990). PCR reactions and conditions were set up according to Ondrušková *et al*. (2017).

The PCR products were visualized and purified as described by Ondrušková *et al*. (2017, 2018) and sequenced by the Seqme company (Dobříš, Czech Republic).

Dothistroma ITS haplotypes were determined using manually improved alignments generated by the Geneious Alignment function in Geneious vs. 10.2.6 (Kearse *et al*., 2012) and comparing

sequences to the deposited reference sequences of Barnes *et al*. (2016) for ITS haplotypes Dp_HAP.1 - Dp_HAP.5, and Mullett *et al*. (2018) for ITS haplotype Dp_HAP.6.

2.4 Microsatellite amplification and determination of multilocus haplotypes

Sixteen microsatellite markers that were developed for *D. pini* (Siziba *et al*., 2016), were used in this study. An additional marker (Doth_A) developed by Barnes *et al*. (2008), was included as an identification control to distinguish between *D. pini* and *D. septosporum*. PCR reactions were performed in 15 µl reaction volumes. The reactions consisted of 1 µl template DNA (20 ng/µl concentration), 0.15 µl of 100 nM of each of the forward and reverse primers, 0.12 µl Faststart Taq DNA polymerase (Roche Diagnostics, Indianapolis, IN), 200 µM of each dNTP, 1.5 µl of a 2.5mM $MgCl_2$, 1.5 µl 10x PCR reaction buffer and the volume was adjusted to 15 µl with sterile SABAX water (Adcock Ingram, Midrand, South Africa).

PCR reactions were carried out on an Applied Biosystems® Veriti® 96 well Thermal cycler (Thermo Fisher Scientific, Waltham, MA). The cycling conditions for all microsatellite fragments included an initial denaturation step at 95 °C for 4 min, 10 cycles consisting of 94 °C for 20 s (denaturation), a 45 s annealing step according to the primer pair annealing temperature (Table S2) and an elongation step of 45 s at 72 °C. This was followed by a further 25 cycles of 94 °C for 20 s, 45 s with a 5 s extension step per cycle at the indicated annealing temperature, a 72 °C extension for 45 s and a final extension step of 60 °C for 30 min. The amplified products were visualized by staining 5 µl of each product with 1 µl GelRed™ nucleic acid gel stain (Biotium, Fremont, CA). The fragments were separated on 2% SeaKem® LE agarose gel (Lonza, Rockland, ME) for 15 min at 90 V and viewed under a UV light using the GelDoc™ EZ Imager (BioRad, Hercules, CA). PCR products were pooled into two panels for fragment analysis according to Siziba *et al*. (2016) and with adjusted dilutions (Table S2). The panels were prepared for analysis by adding 1 µl of the pooled products to 12 µl formamide and 0.14 µl GENESCAN™ -500 LIZ® (Life Technologies, Applied Biosystems, Warrington, UK) size standard. The prepared reactions were run on an ABI PRISM 3500xl capillary auto sequencer (Thermo Fisher Scientific). Alleles sizes were scored using GENEMAPPER® Software version 5.0 (Applied Biosystems, Foster City, CA).

Multilocus haplotypes (MLH) for each isolate was obtained by combining the alleles determined for each marker for each individual. If two individuals had the same combination of alleles at all the markers, they were considered to be clones. The number of MLHs was determined using the poppr R package (Kamvar *et al*., 2014).

2.5 Datasets

Two datasets were used for the analyses; one consisted of all individuals (non-clone corrected dataset [non-CC]) and the other consisted of a single representative of each clone per population (clone-corrected dataset [CC]). Populations were defined from geographic regions and from host species. As some locations were represented by low sample size (< 6 sampled trees), the geographically close locations were grouped into one region to get a sufficient number of individuals for analyses. Isolates were grouped in four regions as follows: Region 1 = Košice and Sečovce, Region 2 = Banská Belá, Kováčová, Ľubochňa and Zvolen, Region 3 = Mlyňany Arborétum, and Region 4 = Jahodná, Gabčíkovo and Trstice (Table S1). In the populations that consisted of host species, only those that contained six or more isolates were used.

2.6 Population structure and differentiation

Population structure was assessed using two different methods for CC datasets; discriminant analysis of principal components (DAPC) (Jombart *et al*., 2010) and STRUCTURE v. 2.3.4 (Falush *et al*., 2003). For the DAPC, the number of clusters (K) of genetically related individuals was established by a multivariate technique that forms no inferences regarding the population model or data structure (Jombart *et al*., 2010). Optimal number of clusters was assessed by Bayesian information criterion (BIC) using the 'find.cluster' function. All these procedures were conducted in the R package adegenet (Jombart, 2008; Jombart *et al*., 2010). The a-score function was used to determine the optimal number of principal components retained in the analyses. The a-score can serve as a criterion for choosing the optimal number of principal components (PC) in the principal component analyses (PCA) step of DAPC, i.e. the number of PC maximizing the a-score (closest to one) (Jombart *et al*., 2010).

Secondly, the program STRUCTURE v. 2.3.4 (Falush *et al*., 2003) that implements a Bayesian clustering algorithm was used to assign individuals to a specified number of clusters (K). The optimal number of K was determined using the ΔK method of Evanno *et al*. (2005) in CLUMPAK (Kopelman *et al*., 2015). To define the optimal K, 10 independent runs of K=1-5 were carried out. Each run had a burn in of 80,000 iterations followed by a run of 1,000,000 iterations, using a model of correlated allele frequencies and with admixture among populations allowed. MLHs with a membership coefficient of greater than 0.5 in a particular STRUCTURE cluster were allocated to that cluster for further analyses.

Hierarchical analysis of molecular variance (AMOVA) was carried out on the clone-corrected dataset to test the hypotheses of population differentiation using ´adegenet´ (Jombart, 2008). AMOVA was used to test for differentiation between isolates grouped into the same STRUCTURE/DAPC clusters. The isolates were also grouped by host species and by regions. Host species for analyses were restricted to isolates from the most abundant hosts (n>6). The null hypothesis of no genetic difference between populations was rejected at $p < 0.05$. The significance was tested by 1000 permutations of the dataset.

The network analysis was carried out using the software package EDENetworks v 2.18 (Kivelä *et al*., 2015). EDENetworks is constructed for visualizing and investigating networks in order to study relationships in a dataset, without prior assumptions on the clustering of individuals, populations or genetic groups (Kivelä *et al*., 2014). Weighted networks, where networks are derived from genetic distance matrices by thresholding out the largest distances, were constructed from the data, in which each node represents a MLH. Networks were predicted from a genotype matrix of individual-centred samples using a distance measure based on the proportion of shared alleles. The threshold distance value was set slightly below the percolation threshold, such that the network still remained connected to identify the most related genotypes (Kivelä *et al*., 2015). The nodes represent fundamental units of the system and the links between them represent their interactions or relationships. The strength of such relationships can be taken into account in the form of link weights – a genetic distance is related with a link. The weights represent genetic similarity (Kivelä *et al*., 2014; 2015). The betweenness centrality was calculated for each node, describing the importance of the node (number of shortest paths between other nodes passing through a node) (Kivelä *et al*., 2014; Kiel, 2016).

Pairwise genetic differentiation (Fst) was calculated in hierfstat version 0.5-7 (Goudet & Jombart, 2020) and Fst value calculations were done according to Nei (1987) between Clusters and Regions.

2.7 Genetic diversity

The measures to evaluate the genetic diversity were calculated for populations grouped by DAPC/STRUCTURE clusters and by regions.

The CC dataset was used to calculate Nei's unbiased gene diversity and H_{exp} (Nei, 1978), and the non-CC dataset was used to evaluate genotypic diversity, richness and evenness, using the package poppr (Kamvar *et al*., 2014) of the R 3.6.3 environment for statistical computing (R Core

Team, 2020). The following measures were calculated: i) Genotypic richness, eMLH, the expected number of MLHs (eMLH) calculated by rarefaction to the smallest sample size, ii) Shannon-Wiener index, H (Grünwald *et al*., 2003), iii) Stoddart and Taylor's index, G (Stoddart & Taylor, 1988), iv) Simpson's index λ (Simpson, 1949), v) Genotypic evenness, E5, that is an estimation of evenness equal to 1 when all genotypes are equally represented and reduces toward 0 as a single genotype becomes more dominant (Grünwald *et al*., 2003). The clonal fraction (CF), which is the proportion of isolates derived from clones, or asexual reproduction, was calculated according to the method of Zhan *et al*. (2003).

2.8 Mating types and sexual recombination

The mating type of each isolate was determined using the mating type specific primers of Groenewald *et al*. (2007). The composition of PCR mixture and thermal cycling conditions were as those described by Ondrušková *et al*. (2017). The set of primers amplify a fragment of 820 bp for the *MAT1-1-1* and 480 bp for the *MAT1-2* idiomorph.

To investigate the possibility of sexual recombination, the following tests were used: (i) an exact binomial test using two tailed P-values to determine whether groups differed significantly from the null hypothesis of a 1:1 ratio of mating type (McDonald, 2014), and (ii) the index of association (I_A) together with its associated measure r_d calculated using the package poppr (Kamvar *et al*., 2014). Clonal populations are expected to have significant linkage disequilibrium due to the non-random association of alleles at different loci, while sexual populations are expected to be in linkage equilibrium. The I_A and r_d from the observed data were compared to values obtained after 1,000 permutations to simulate random mating. The null hypothesis of no linkage among loci was rejected at $p > 0.05$ and indicate sexual reproduction can occur in population.

The isolates for both tests were grouped by DAPC/STRUCTURE clusters and by regions. Both, the clone-corrected and non-clone-corrected datasets were analysed .

3 RESULTS

3.1 Studied locations and fungal isolates

Dothistroma pini was previously confirmed in 31 locations in Slovakia (Jánošíková - Hečková *et al*., 2018), however, these were mostly through methods of DNA extraction and PCR amplification (Groenewald *et al*., 2007) directly from needles. The isolation of *D. pini* was unsuccessful from some of these samples, either due to low levels of infections, or symptomatic

needles that did not contain conidiomata, during the collection period. In this study, isolation of *D. pini* cultures was successful from 53 needle samples from 10 different localities (Table 1.) Four locations were situated in urban greenery, four locations were forests (plantations and/or natural regeneration) and two *D. pini* records came from arboreta/botanical gardens.

Trees sampled in urban greenery were limited in numbers. Isolates either came from a single tree or shrub, or, in addition, were mixed infections with the more dominating *D. septosporum*. The three plantations consisted of *P. nigra* in areas ranging in size from 1 to 3 ha, and situated in south west Slovakia. These trees were heavily infected allowing for a greater collection of samples from these locations (Jahodná, Trstice and Gabčíkovo). The natural regenerated forest plantation is a mixed plantation (mainly *P. sylvestris*) with low DNB infection and related symptoms situated in the most northern region of the country and with the highest altitude (615 m.a.s.l.).

The largest collection of *D. pini* isolates was obtained from the Arboretum in Mlyňany (18 samples and 41 cultures in total) due to the large area and broad host range affected. In Košice Botanical garden, the samples were collected from two newly planted young *P. ponderosa* trees.

Altogether, isolates from seven pine species were analysed. The majority of the samples came from *P. nigra* (54 isolates) collected in five locations, followed by *P. ponderosa* (34 isolates). A limited number of *D. pini* isolates were obtained from native pine species; two isolates from *P. mugo* artificially planted in urban greenery and one isolate from *P. sylvestris* from natural regeneration forest plantation.

3.2 Investigation of historical samples in herbaria

The search for pine needle specimens or cultures that could represent any of the *Dothistroma* pathogens was negative for all six herbaria that were checked. No historical records of DNB exist in SAV, SLO, NR, BP, BRNM or PRM.

3.3 ITS haplotypes

A total of 105 isolates were identified as *D. pini* based on ITS sequences (Table S1). ITS sequences of all investigated isolates were designated to one of three already described ITS haplotypes (Barnes *et al*., 2016; Mullett *et al*., 2018); no new ITS haplotype was recorded. Dp_HAP.1 was generally widespread and the most dominant haplotype, being present in seven of the 10 localities, comprising more than 90% of the isolates (95 isolates) and found on six host species. ITS haplotypes Dp_HAP.2 and Dp_HAP.4, were represented by seven and three isolates

respectively. Dp_HAP.2 was recorded within samples from the Banská Belá, Kováčová, Mlyňany Arboretum and Sečovce locations while Dp_HAP.4 was found in Zvolen and Mlyňany Arboretum (Figure 1, Table 2).

3.4 Determination of multilocus haplotypes

The microsatellite analysis of 105 isolates yielded a total of 35 different alleles. Nine of the 16 microsatellite markers were polymorphic ranging from 2 at DP_4, DP_15, DP_16, DP_8, to 7 different alleles for the most variable marker, DP_12. The other seven loci were monomorphic (Table S2).

Based on the 16 microsatellite loci a total of 15 unique MLHs were detected in the 105 isolates. More than half of the MLHs (n=8) were represented by a single isolate, two MLHs (13 and 15) were represented by 36 and 45 isolates, respectively. These two most dominant MLHs were found in four locations (Mlyňany Arboretum, Gabčíkovo, Jahodná and Trstice locations) while both MLHs were present at each of these four locations (Figure 1). Eleven MLHs were found at single locations, two (MLH06 and MLH12) at two locations and two (MLH 13 and 15) at four different locations.

The most diverse location in terms of MLHs was Mlyňany Arboretum, where six MLHs were determined within 41 isolates, followed by Jahodná with five MLHs detected within 26 isolates (Figure 1).

3.5 Population structure and differentiation

The DAPC sequential K-means analyses and the BIC clustering on the clone corrected dataset separated the data into two clusters (Figure S1). Similarly, STRUCTURE analysis indicated an optimum of two clusters when both methods, ΔK method of Evanno *et al*. (2005) and ln(Pr(X | K)) by Pritchard *et al*. (2000, 2010), were used to choose the best K (Figure S2).

Both methods, STRUCTURE and DAPC, classified the individual isolates into the same two clusters (Figure 2). Ten isolates grouped in Cluster 2 and all were either ITS haplotypes Dp_HAP.2 and Dp_HAP.4 (Table S1, Figure 1). The rest of the 95 isolates grouped into Cluster 1 and all of these were ITS Dp_HAP_1 (Table S1, Figure 1). Only two locations (Zvolen and Mlyňany Arboretum) had MLHs identified in both clusters (Figure 3).

The AMOVA showed that grouping the isolates by STRUCTURE and DAPC clusters was supported and highly significant for the CC dataset. Conversely, no significant population

differentiation was found when populations were grouped by Region or by host species (p=0.117 and p=0.8 respectively, Table 3).

For the network analyses, the sharp break between the two cluster divergences is seen most clearly in the EDENetwork, net of genetic relationship allele sharing statistics (Figure 4). It showed that the node which joined the two genetically related groups is MLH07 at automatic percolation threshold level 0.29. It has the highest value of betweenness centrality, which is the measure of importance of a node. There is a clear direct genetic connection between the two clusters via MLH07 and four MLHs in cluster 1 (MLH02, 03, 10, 12). The blue colour and longer length of links between these nodes indicates a weak genetic relationship between the two clusters. The genetic relatedness of nodes in the larger Cluster (STRUCTURE/DAPC cluster 1) are stronger than those in the smaller Cluster (STRUCTURE/DAPC Cluster 2).

Fst calculated between the two Clusters was very high at 0.76, and there were few shared alleles. Thus, genetic variation could be explained by the population structure. The level of divergence between Regions (Table S3), although lower than between Clusters, was still relatively high ranging from 0.325 (between Region 1 and Region 2)-0.540 (between Region 2 and Region 4).

3.6 Genetic diversity

The genotypic and gene diversity were different for the two clusters (Table 4a). Congruently all genetic diversity measures were higher for Cluster 2 despite the lower sample size. Both genotypic richness (eMLH) and genotypic evenness (E5) were higher for cluster 2, and genotypic diversity indexes (H, G, λ) as well, with the difference ranging from 0.132 for λ to 1.48 for G (Table 4a). Likewise, at allelic level, gene diversity (H_{exp}) was over three times higher for Cluster 2 (0.157 vs 0.050). Overall, the values for gene diversity were low. The clonal fraction was also different for the two clusters; high for Cluster 1, where almost 90% of isolates were clones, while in Cluster 2 the proportion of clonal isolates was lower (CF=0.5, Table 4a). The overall Slovakian *D. pini* population was clonal (CF = 0.857).

For Regions, the highest genotypic richness after rarefaction, genotypic diversity and evenness, and gene diversity were estimated for Region 2 located in the middle of Slovakia. This is also the region where the lowest proportion of clones was determined (Table 4b). The diversity in the other three Regions was similar with slight difference in range between them. The clonal fraction was highest for Region 3 and 4, in the south west part of the country (0.854 and 0.890), in accordance with the lowest gene diversity 0.090 and 0.027 (Table 4a, b).

3.7 Mating types and reproductive mode

The mating type of 103 isolates from 105 were successfully characterised at nine of the 10 locations using the species-specific mating type primers. For two isolates the identification of mating type failed (Table S1).

Overall, the *MAT1-2* idiomorph (92%) was the most dominant. The *MAT1-1-1* idiomorph was only present in two localities (Banská Belá and Košice), while the rest of the seven localities all contained isolates of *MAT1-2*.

As only a single mating type was identified per location, random mating was not studied at the level of locations. For the geographically close localities grouped into Regions, the mating type ratio did not differ significantly from 1:1 for the non-CC as well as the CC dataset for Regions 1 and 2 (Table 5a). When analysing isolates grouped by STRUCTURE clusters, Cluster 2 did not significantly differ from a 1:1 ratio in spite of the slight dominance of MAT1-2 in both datasets (non-CC and CC). When the whole Slovakian population was taken into account, the mating type significantly differed from a ratio of 1:1.

The test of index of association I_A and r_d supported random mating for both Clusters but only in case of the CC dataset. Linkage equilibrium was found within samples grouped by Clusters indicating sexual reproduction could be expected. For the non-CC as well as the CC dataset for Regions, significant disequilibrium was determined, indicating only asexual reproduction is expected (Table 5b).

4 DISCUSSION

This study represents the first population study of *D. pini* in Slovakia using the 16 microsatellite markers designed by Siziba *et al.* (2016). Analyses of 105 isolates from 10 locations and on seven hosts indicated the populations to have a high level of clonality and low level of genetic diversity, and with little support for random mating. These data, together with the lack of any historical samples in herbaria support the hypothesis that *D. pini* is a recently introduced pathogen in Slovakia.

Three of the six known ITS haplotypes for *D. pini* were obtained among the 105 isolates in our study; no new ITS haplotypes were revealed. In Spain, the same three ITS haplotypes (Dp_HAP.1, Dp_HAP.2 and Dp_HAP.4) were found within 16 isolates (Mullett *et al*., 2018), and these were also identified in France (Barnes *et al.*, 2016). In other European countries (Czech Republic,

Hungary, Romania, Slovenia, Russia, Ukraine) only single ITS haplotype from the three previously mentioned was recorded (Barnes *et al.*, 2016). The most frequent ITS haplotype, Dp_HAP.1, is reported from Europe as well from USA, while ITS haplotypes Dp_HAP.3, Dp_HAP.5 and Dp_HAP.6 were reported only from USA and seem to be lacking in Europe (Barnes *et al.*, 2016; Mullett *et al.*, 2018).

Three clustering methods (DAPC, STRUCTURE, EDENetwork) supported the division of the isolates from Slovakia into two main clusters. Clusters were not congruent with geographic location or host species but corresponded to ITS haplotypes where all isolates in Cluster 1 contained ITS haplotype Dp_HAP_1 while those in Cluster 2 were made up of isolates with either Dp_HAP.2 or Dp_HAP.4. In EDENetwork for the multilocus haplotype, MLH07 from Cluster 2, the largest betweenness centrality (in Figure 4 indicated by size of node) was calculated based on connections with other MLH. Moreover, it represents the connection between two clusters suggesting that MLH07 should play a crucial role in *D. pini* populations in Slovakia. To identify its role in genetic structure and diversity, further detailed studies and data analyses are needed.

The population studied from Slovakia (105 isolates) only contained 35 alleles. This is relatively low compared to the previously analysed dataset of 32 *D. pini* isolates from seven countries, where 66 alleles were obtained when the newly developed 16 microsatellite loci were tested across different *D. pini* isolates (Siziba *et al.*, 2016). The gene diversity calculated for two European *D. pini* populations from France and Slovenia, with 24 isolates respectively, showed different results. The population from France was genetically diverse (Hexp= 0.35) with both mating types, while the Slovenian population was completely clonal (Hexp= 0.00) (Siziba *et al*., 2016). The diversity in the Slovakian population is more similar to the diversity observed in the Slovenian population, with low diversity and only one mating type present per location. When the microsatellite markers for *D. pini* were tested on a limited number of isolates, higher gene diversity values were observed in the included populations (USA, France, Ukraine), where both mating types were identified (Siziba *et al*., 2016).

The overall genotypic diversity of the Slovakian population was relatively low, however, when smaller datasets in terms of Clusters and Regions were analysed, slight differences were recorded. Cluster 2 and Region 2 showed higher values of genotypic diversity measures. Each MLH assigned to Cluster 2 was only detected from a single tree growing in urban greenery at a single location, except for MLH06 that was detected in Zvolen and Mlyňany Arboretum. As these MLHs have a restricted incidence, it should be supposed that they were introduced recently due to

anthropogenic activities. Alternatively, the single occurrence could be explained as a result of mutation which is a source of all genetic variation and has an impact on diversity (Milgroom, 2015).

When comparing the diversity measures for genotypic and gene diversity of the two clusters, the gene diversity of Cluster 2 is three times higher compared to Cluster 1, while the difference in genotypic indexes between the two clusters is lower. Higher gene variability could lead to increase of genotypic diversity by recombination through random mating (Milgroom, 1996). This has likely not yet occurred in the studied populations.

The distribution, and presence of only a single mating type of *D. pini* per location in Slovakia, implies that sexual reproduction is not occurring in the studied locations. Therefore, the impact of sexual recombination on the genetic and genotypic diversity of *D. pini* populations in the country can be excluded. The lack of sexual reproduction in Slovakia is furthermore supported by the overall high fraction of clones (more than 85%). These findings suggest that asexual reproduction is the main mode of reproduction for the fungus and that conidia are the primary source of dispersal.

Random mating was only supported when analysing the data in clusters, and only in Cluster 2. Cluster 2 had a lower fraction of clones (50%) compared to cluster 1 (89%) and most of the isolates came from Region 2 (locations: Banská Belá, Kováčová, Ľubochňa and Zvolen). However, the closest localities with the opposite mating type were at least 15 km apart (for Banská Belá with MAT1-1 and Zvolen with MAT1-2), or more (45 km for Banská Belá and Mlyňany Arboretum or 30.5 km for Košice and Sečovce). This physical distance would act as a natural barrier making it highly unlikely for the mixing of individuals with opposite mating types to occur. In a study on natural spore dispersal, Mullett *et al*. (2016) recorded the pathogen spores at a maximum distance of 1436 m from any infection source. The sexual state of *D. pini* has still not been described (Drenkhan *et al*., 2016) and the lack of long-distance dispersal of the pathogen by wind dispersed ascospores could be one of the reasons for the restricted *D. pini* geographical distribution.

Isolates from seven pine hosts were analysed in the present study. Only one sample was collected in a naturally regenerated plantation of native *P. sylvestris*; however, the sampling place was an abandoned forest nursery (I. Podhorec – Forests of the Slovak Republic state enterprise, Ľubochňa, Slovakia, 23 March 2014, personal communication). The rest of the locations were artificially planted trees either in urban greenery and arboreta or for Christmas tree purposes. Most

isolates were collected from introduced tree species and native species in artificially planted urban greenery, which could indicate the non-native origin of the pathogen, and further supports the notion that the pathogen was introduced into Slovakia. In addition, the lack of any historical records of DNB in the local and neighbouring herbaria, excludes the long-term presence, and favours the short term presence and recent emergence of *D. pini* in Slovakia.

The first report of *D. pini* in Slovakia was in 2013 (Ondrušková *et al*., 2018), where samples were collected in the south west part of the country close to the Hungarian border in the localities of Jahodná and Gabčíkovo. A third of all the MLHs detected in Slovakia were present in these locations. *Pinus nigra* seedlings planted there for Christmas tree purposes in an area of 2.8 ha were most likely imported from Hungary (J. Habara – Forests of the Slovak Republic state enterprise, Gabčíkovo, Slovakia, 28 November 2013, personal communication). These plantations are about 30-35 years old and were planted around 1987-1990.

DNB in neighbouring Hungary has a longer history and was already reported in 1990, and again during 1994-1995, when large disease outbreaks were occurring throughout the country (Szabó, 1997; Barnes *et al*., 2011). Both *D. pini* and *D. septosporum* were detected in Hungary (Barnes *et al*., 2011). The seedlings imported into south west Slovakia from Hungary could be the primary source of infection into Slovakia from where the pathogen has subsequently spread. Alternatively, natural spread of the pathogen into Slovakia could have also occurred as there are no significant geographical barriers between Hungary and the south part of Slovakia which would prevent pathogen spread. Future studies should confirm this hypothesis by comparing the genetic similarity of the Slovakian and Hungarian isolates.

The trees infected in the urban greenery have an unknown place of origin, but the import of infected plant material from foreign countries could be considered as an introductory pathway of *D. pini* into Slovakia. Especially, in the east part of country (Sečovce location), from where the isolate´s (MLH05) ITS sequence (ITS haplotype Dp_HAP.2) was identical to Ukrainian *D. pini* isolates (Ondrušková *et al*., 2018). In the Arboreta (Mlyňany and Košice) trees infected with *D. pini* have a local origin. Pine seedlings used for planting were produced in local nurseries (P. Vrábel – Košice Urban Greenery Administration, Slovakia, 15 June 2015; I. Tábor - Mlyňany 12 December 2019, personal communication). As these Arboreta are frequently visited places, tourism may play an important role in the anthropogenic spread of the pathogen (Jankovský *et al*., 2009). The highest genetic variability and the highest number of MLHs recorded were from Mlyňany Arboretum and this supports this elucidation. Furthermore, the high concentration of

susceptible pine species and appropriate climatic and environmental conditions in Arboretum Mlyňany favour the pathogen life cycle.

In conclusion, the gene and genotypic diversity and all diversity indexes of the populations studied in Slovakia were low. The low diversity with increasing levels of clonality is typical for asexually reproducing populations (Milgroom, 1996). In all aspects, results of this study, including the presence of only a single mating type per location, indicates a recent introduction of *D. pini* into Slovakia.

Population analyses of this pathogen are lacking worldwide. The increasing number of *D. pini* reports indicates its rapid and recent spread, or, an increased awareness of this pathogen. These increasing records, do however, offer an opportunity for larger population structure and genetic diversity studies, which could enable the tracing of historic origin and dispersal of the pathogen. Furthermore, knowledge of its virulence and dispersal mechanisms could help to identify effective control measures to reduce its spread.

ACKNOWLEDGEMENT

Authors thank Katarina Pastirčáková, the curator of herbarium NR, for her help with inspection regarding herbarium specimens and communications with the herbaria BRNM, PRM and BP regarding their collections. This work was supported by the Scientific Grant Agency VEGA of the Slovak Academy of Sciences, No. 2/0077/18. Funding for the laboratory component of the work was provided by the Tree Protection Co-operative Programme (TPCP) at the Forestry and Agricultural Biotechnology Institute (FABI), University of Pretoria.

CONFLICTS OF INTEREST

The authors declare no conflict of interest.

DATA AVAILABILTY STATEMENT

All data has been provided in Supplementary Tables 1 and 2

ORCID

Katarína Adamčíková 0000-0001-5407-5233

Ariska van der Nest 0000-0002-0640-0609

Slavomír Adamčík 0000-0003-2156-5767

Emília Ondrušková 0000-0002-5082-8709
Irene Barnes 0000-0002-4349-3402
Zuzana Jánošíková 0000-0001-8746-389X

REFERENCES

Barnes, I., Crous, P.W., Wingfield, M.J. and Wingfield, B.D. (2004) Multigene phylogenies reveal that red band needle blight of *Pinus* is caused by two distinct species of *Dothistroma*, *D. septosporum* and *D. pini*. *Studies in Mycology,* 50, 551–565.

Barnes, I., Cortinas, M.N., Wingfield, M.J. and Wingfield, B.D. (2008) Microsatellite markers for the red band needle blight pathogen, *Dothistroma septosporum*. *Molecular Ecology Resources*, 8, 1026–1029.

Barnes, I., Kirisits, T., Wingfield, M.J. and Wingfield, B.D. (2011) Needle blight of pine caused by two species of *Dothistroma* in Hungary. *Forest Pathology*, 41, 361–369.

Barnes, I., van der Nest A., Mullett, M.S, Crous, P.W, Drenkhan, R., Musolin, D.L. and Wingfield, M.J. (2016) Neotypification of *Dothistroma septosporum* and epitypification of *D. pini*, causal agents of Dothistroma needle blight of pine. *Forest Pathology*, 46, 388–407.

Barnes, I., Wingfield, M.J., Carbone, I., Kirisits, T. and Wingfield, B.D. (2014) Population structure and diversity in an invasive pine needle pathogen reflects anthropogenic activity. *Ecology and Evolution*, 18, 3642–3661.

Capron, A., Feau, N., Heinzelmann, R., Barnes, I., Benowicz, A., Bradshaw, R.E. et al. (2021) Signatures of post-glacial genetic isolation and human-driven migration in the Dothistroma needle blight pathogen in western Canada. *Phytopathology.* doi:10.1094/PHYTO-08-20-0350-FI

Drenkhan, R., Tomešová-Haataja, V., Fraser, S., Bradshaw, R.E., Vahalík, P., Mullett, M.S. et al. (2016) Global geographic distribution and host range of *Dothistroma* species: a comprehensive review. *Forest Pathology*, 46, 408–442.

Ennos, R.A., Sjökvist, E.I., Piotrowska, M.J., Riddell, C. and Hoebe, P.N. (2020) Using genome resequencing to investigate racial structure, genetic diversity, sexual reproduction and hybridisation in the pine pathogen *Dothistroma septosporum*. *Fungal Ecology*, 46, 100921.

Evanno, G., Regnaut, S. and Goudet, J. (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology*, 14, 2611–2620.

Fabre, B., Ioos, R., Piou, D. and Marçais, B. (2012) Is the emergence of Dothistroma needle blight of pine in France caused by the cryptic species *Dothistroma pini*? *Phytopathology*, 102, 47–54.

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

Goudet, J. and Jombart T. (2020) hierfstat: Estimation and Tests of Hierarchical F-Statistics. R package version 0.5-7. Available at: https://CRAN.R-project.org/package=hierfstat

Groenewald, M., Barnes, I., Bradshaw, R.E., Brown, A., Dale, A., Groenewald, J.Z. et al. (2007) Characterization and worldwide distribution of the mating type genes in the Dothistroma needle blight pathogens. *Phytopathology*, 97, 825–834.

Grünwald, N.J., Goodwin, S.B., Milgroom, M.G. and Fry, W.E. (2003) Analysis of genotypic diversity data for populations of microorganisms. *Phytopathology*, 93, 738–746.

Jankovský, L., Palovčíková, D., Dvořák, M. and Tomšovský, M. (2009) Records of brown spot needle blight related to *Lecanosticta acicola* in the Czech Republic. *Plant Protection Science,* 45, 16–18.

Jánošíková - Hečková, Z., Ondrušková, E., Barta, M., Ostrovský R., Kádasi Horáková, M., Pastirčáková, K. et al. (2018) The hosts and geographic range of Dothistroma needle blight in Slovakia. *Forest Pathology*, 48, e12421.

Jánošíková, Z., Dutech, C., Ondrušková, E., Adamčíková, K., Mullett, M.S. (2021) Population structure and genetic diversity of *Dothistroma septosporum* in Slovakia.*European Journal of Plant Pathology,* doi: 10.1007/s10658-021-02266-z

Jombart, T. (2008) Adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics*, 24, 1403–1405.

Jombart, T., Devillard, S. and Balloux, F. (2010) Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. *BMC Genetics*, 11, 94.

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

Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S. et al. (2012) Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics*, 28, 1647–1649.

Kiel, S. (2016) A biogeographic network reveals evolutionary links between deep-sea hydrothermal vent and methane seep faunas. *Proceedings of the Royal Society B, Biologial Sciences*, 283, 20162337.

Kivelä, M., Arenas, A., Barthélemy, M., Gleeson, J.P., Moreno, Y. and Porter, M.A. (2014) Multilayer networks. *Journal of Complex Networks*, 2, 203–271.

Kivelä, M., Arnaud-Haond, S. and Saramäki, J. (2015) EDENetworks: A user friendly software to build and analyse networks in biogeography, ecology and population genetics. *Molecular Ecology Resources*, 15, 117–122.

Kopelman, N.M., Mayzel, J., Jakobsson, M., Rosenberg, N.A. and Mayrose, I. (2015) CLUMPAK: a program for identifying clustering modes and packaging population structure inferences across K. *Molecular Ecology Resources*, 15, 1179–1191.

Kowalski, T. and Jankowiak, R. (1998) First record of *Dothistroma septospora* (Dorog.) Morelet in Poland: a contribution to the symptomology and epidemiology. *Phytopatologia Polonica*, 16, 16–29.

Matsiakh I., Doğmuş-Lehtijärvi, H. T., Kramarets, V., Aday Kaya, A.G., Oskay, F., Drenkhan, R. et al. (2018) *Dothistroma* spp. in Western Ukraine and Georgia. *Forest Pathology*, 48, e12409.

McDonald, B.A. and Linde, C. (2002) Pathogen population genetics, evolutionary potential, and durable resistance. *Annual Review of Phytopathology*, 40, 349–379.

McDonald, J.H. (2014) *Handbook of Biological Statistics*. Baltimore, MD: Sparky House Publishing.

Milgroom, M.G. (1996) Recombination and the multilocus structure of fungal populations. *Annual Review of Phytopathology*, 34, 457–477.

Milgroom, M.G. (2015) Population Biology of Plant Pathogens. Genetics, Ecology, and Evolution. Itaca, New York: APS Press.

Mullett, M.S., Adamson, K., Bragança, H., Bulgakov, T., Georgieva, M., Henriques, J. et al. (2018) New country and regional records of the pine needle blight pathogens *Lecanosticta acicola*, *Dothistroma septosporum* and *Dothistroma pini*. *Forest Pathology*, 48, e12440.

Mullett, M.S., Drenkhan, R., Adamson, K., Boroń, P, Lenart-Boroń, A., Barnes, I., Tomšovský M. et al. (2021) Worldwide genetic structure elucidates the Eurasian origin and invasion pathways of *Dothistroma septosporum*, causal agent of Dothistroma needle blight. *Journal of Fungi*, 7, 111. Doi. /10.3390/ jof7020111

Mullett, M.S., Tubby, K.V., Webber, J.F. and Brown, A.V. (2016) A reconsideration of natural dispersal distances of the pine pathogen *Dothistroma septosporum*. *Plant Pathology*, 65, 1462–1472.

Nei, M. (1978) Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics*, 89, 583–590.

Nei, M. (1987) Molecular Evolutionary Genetics. New York: Columbia University Press

Ondrušková, E., Hečková, Z., Kádasi Horáková, M., Koltay, A., Ostrovský, R., Pažitný J. *et al*. (2017) Distribution and characterization of Dothistroma needle blight pathogens on *Pinus mugo* in Slovakia. *European Journal of Plant Pathology*, 148, 283–294.

Ondrušková, E., Jánošíková, Z., Adamčík, S., Kádasi Horáková, M., Rakúsová-Sládková, D. and Adamčíková, K. (2018) Needle blight caused by *Dothistroma pini* in Slovakia: distribution, host range and mating types. *Scandinavian Journal of Forest Research*, 33, 650–656.

Ondrušková, E., Ostrovský, R., Jánošíková, Z., Adamčíková, K. and Kobza, M. (2020) Selected climatic variables in Slovakia are favourable to the development of Dothistroma needle blight. *Folia oecologica*, 47, 144–152.

Ortíz de Urbina, E., Mesanza, N., Aragonés, A., Raposo, R., Elvira-Recuenco, M., Boqué, R. et al. (2017) Emerging Needle Blight Diseases in Atlantic Pinus Ecosystems of Spain. *Forests*, 8, 18.

Pritchard, J.K., Stephens, M. and Donnelly, P. (2000) Inference of population structure using multilocus genotype data. *Genetics*, 155, 945–959.

Pritchard, J.K., Wen, X. and Falush, D. (2010) Documentation for STRUCTURE Software: Version 2.3. Available at: https://www.ccg.unam.mx/~vinuesa/tlem09/docs/structure_doc.pdf

Queloz, V., Wey, T. and Holdenrieder, O. (2014) First record of *Dothistroma pini* on *Pinus nigra* in Switzerland. *Plant Disease*, 98, 1744.

R Development Core Team (2020) R: A language and environment for statistical computing. R Foundation for Statistical Computing. Vienna, Austria. Available at: http://www.R-project.org/

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

Siziba, V.I., Wingfield, M.J., Sadiković, D., Mullett, M.S., Piškur, B. and Barnes, I. (2016) Development of microsatellite markers for the pine blight pathogen, *Dothistroma pini*. *Forest Pathology*, 46, 497–506.

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

Szabó, I. (1997) Occurrence of *Dothistroma septospora* in Austrian pine plantation. *Erdészeti és Gazdászati Lapok*, 132, 44–45. [In Hungarian].

White, T.J., Bruns, T.D., Lee, S. and Taylor, J. (1990) Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. In: Innis, M.A., Gelfand, D.H., Sninsky, J.J. and White, T.J. (Eds.) *PCR Protocols: A guide to methods and applications*. London, UK: Academic Press. pp.315–322.

Zhan, J., Pettway, R.E. and McDonald, B.A. (2003) The global genetic structure of the wheat pathogen *Mycosphaerella graminicola* is characterized by high nuclear diversity, low mitochondrial diversity, regular recombination, and gene flow. *Fungal Genetics and Biology*, 38, 286–297.

FIGURE LEGENDS

Figure 1

Distribution of ITS haplotypes (HAP) and multilocus haplotypes (MLH) of *Dothistroma pini* in Slovakia. Each pie chart represents a single location, the sectors in pie charts represent individual MLHs and circles around the charts correspond to the ITS haplotype of isolates.

Figure 2

DAPC (a) and STRUCTURE (b) clustering of *Dothistroma pini* multilocus haplotypes (MLH), for the clone-corrected dataset. Each MLH is represented by a vertical line, and colours represent the estimated membership in each cluster.

Figure 3

Arrangement of *Dothistroma pini* multilocus haplotypes (MLH) in clusters and their presence in the studied locations and Regions. The colour scale for the MLHs correspond with Figure 1.

Figure 4

EDENetwork genetic relationship network for *Dothistroma pini* MLHs with an automatic threshold of 0.29.

The colour of the nodes is according to membership of the MLHs to the DAPC / STRUCTURE clusters, white nodes = Cluster 1, red nodes = Cluster 2. The node size by betweenness centrality. The thickness, length and colour of links indicate their weight. Longer and thin links = weak relationships; thick and short links = strong relationships. The colour bar from left to right indicates the increase of relationships between the nodes.

SUPPORTING INFORMATION LEGENDS

Supporting Information Figure 1

DAPC sequential K-means (a), and the BIC clustering (b) indicating the division of the *Dothistroma pini* isolates from Slovakia into two Clusters.

Supporting Information Figure 2

Delta K (a) plot of the STRUCTURE analysis of the *Dothistroma pini* isolates from Slovakia, showing optimal K=2 by Evanno et al.(2005) as the most probable number of populations in the dataset, and probability by K graph (b) using median values of Ln(Pr Data) the k for which Pr(K=k) is highest: 2.

Supporting Information Table S1

Details of the *Dothistroma pini* isolates from Slovakia used in this study.

Supporting Information Table S2

Details of PCR amplification using 16 microsatellite markers designed for *Dothistroma pini*.

Supporting Information Table S3

Pairwise genetic differentiation, Fst, calculated between genetic populations of *Dothistroma pini* grouped by regions.

Table 1. Details of localities sampled for *Dothistroma pini*, including host species, and herbarium numbers for samples collected.

No	Locality	Plantation type	Coordinates	Altitude (m.a.s.l.[a])	Host species	Date collected	Herbarium item	No of samples collected
1	Banská Belá	urban greenery	N48.46611111, E18.95250000	498	*P. nigra*	December-2015	NR5479	1
2	Gabčíkovo	forest plantation	N47.89388889, E17.60250000	115	*P. nigra*	November-2013, March-2014, December-2017	NR5467, NR5468, NR5882	4
3	Jahodná	forest plantation	N48.04277778, E17.72694444	103	*P. nigra*	November-2013, September-2014, December-2017	NR5469, NR5470, NR5471, NR5883	17
4	Košice	arboretum	N48.73555556, E21.23527778	268	*P. ponderosa*	June-2015	NR5907, NR5908	2
5	Kováčová	urban greenery	N48.60805556, E19.10472222	307	*P. mugo*[b]	February-2017	NR5906	1
6	Ľubochňa	natural regeneration forest plantation	N49.0485, E19.1501	615	*P. sylvestris*[b]	March-2014	NR5472	1

No	Locality	Plantation type	Coordinates	Altitude (m.a.s.l.[a])	Host species	Date collected	Herbarium item	No of samples collected
7	Mlyňany Arboretum	arboretum	N48.31944444, E18.36888889	190	*P. coulteri*	May-2015, June-2017	NR5799, NR5215, NR5216	3
					P. jeffreyi	May-2014	NR5210	1
					P. nigra	May-2015	NR5214	1
					P. ponderosa	November 2014, May-2015	NR5828, NR5211, NR5212, NR5213, NR5217, NR5218, NR5219	12
					P. schwerinii	May-2015	x[c]	1
8	Sečovce	urban greenery	N48.702689, E21.65416667	139	*P. mugo*[b]	May-2015	NR5399	1
9	Trstice	plantation/nursery	N48.01750000, E17.7916667	109	*P. nigra*	October-2015, July-2017	NR5814, NR5886	6
10	Zvolen	urban greenery	N48.57666667, E19.12388889	293	*P. jeffreyi*	March-2015, December-2017	NR5885	2

[a] m.a.s.l. = meters above sea level

[b] native pine species in Slovakia

[c] x = herbarium item not deposited

Table 2. *Dothistroma pini* ITS haplotypes and mating type identified per locality and host species.

No	Locality	Host species	NoI[a]	ITS haplotypes	Mating type
1	Banská Belá	*P. nigra*	3	Dp_HAP.2	MAT1
2	Gabčíkovo	*P. nigra*	6	Dp_HAP.1	MAT2
3	Jahodná	*P. nigra*	26	Dp_HAP.1	MAT2
4	Košice	*P. ponderosa*	5	Dp_HAP.1	MAT1
5	Kováčová	*P. mugo*	1	Dp_HAP.2	neg.r.[c]
6	Ľubochňa	*P. sylvestris*	1	Dp_HAP.1	MAT2
7	Mlyňany	*P. coulteri*	7	Dp_HAP.1 (5)[b], Dp_HAP.4 (2)[b]	MAT2
	Arboretum	*P. jeffreyi*	2	Dp_HAP.1	MAT2
		P. nigra	2	Dp_HAP.1	MAT2
		P. ponderosa	29	Dp_HAP.1 (27)[b], Dp_HAP.2 (2)[b]	MAT2
		P. schwerinii	1	Dp_HAP.1	MAT2
8	Sečovce	*P. mugo*	1	Dp_HAP.2	MAT2
9	Trstice	*P. nigra*	17	Dp_HAP.1	MAT2
10	Zvolen	*P. jeffreyi*	4	Dp_HAP.1 (3)[b], Dp_HAP.4 (1)[b]	MAT2
	Total	7 pine hosts	105	Dp_HAP.1, Dp_HAP.2, Dp_HAP.4	

[a]NoI= number of isolates.

[b]The number in parenthesis indicates the number of isolates containing the specific haplotype per location and host species.

[c]neg. r. = negative results obtained using the mating type specific primers of Groenewald et al. (2007).

Table 3. Analysis of molecular variance (AMOVA) of *D. pini* isolates grouped into structure groups, regions and hosts.

Clone corrected dataset		Degree of freedom	Percentage of variation	Φ (total)	p-value
Structure clusters	Variation among clusters	1	71.74	0.717	**0.002*** **
	Variation within clusters	13	28.26		
Regions	Variation among regions	3	17.19	0.172	0.117
	Variation within regions	14	82.81		
Hosts	Variation among hosts	3	-12.24	-0.122	0.8
	Variation within hosts	13	112.24		

Bold p-values indicate the null hypothesis of no population differentiation can be rejected, p-value <0.01 (***)

Table 4. Genotypic and gene diversity measures for populations of *Dothistroma pini* from Slovakia. Genetic diversity statistics calculated for two clusters suggested by DAPC and STRUCTURE (a), and for the isolates grouped by region (b).

Pop	N[a]	MLH[b]	eMLH[c]	SE[d]	H[e]	G[f]	λ[g]	E5[h]	Clonal fraction	Hexp[i]
(a)										
Cluster 1	95	10	3.34	0.947	1.27	2.69	0.628	0.655	0.895	0.050
Cluster 2	10	5	5	0	1.50	4.17	0.76	0.904	0.500	0.157
Total	105	15	4.11	1.149	1.61	3.26	0.693	0.564	0.857	0.116
(b)										
Region 1	6	2	2	0	0.451	1.38	0.278	0.676	0.667	0.167
Region 2	9	5	5	0	1.465	3.86	0.741	0.859	0.445	0.225
Region 3	41	6	2.6	0.954	0.722	1.44	0.306	0.416	0.854	0.090
Region 4	49	5	2.55	0.708	0.814	1.78	0.438	0.62	0.890	0.027

[a]N = Total number of isolates.

[b]Number of multilocus haplotypes.

[c]The number of expected MLH calculated by rarefaction to the smallest sample size

[d]Standard error

[e]Shannon-Wiener Index of MLH diversity

[f]Stoddart and Taylor's index of MLH diversity

[g]Simpson's index

[h]Genotypic evenness

[i]Nei's gene diversity

Table 5. Mating type distribution of *Dothistroma pini* isolates and tests for random mating, (a) exact binomial two tailed test, (b) index of association

a

	Non-clone corrected dataset			Clone corrected dataset		
	MAT 1-1	MAT 1-2	p value	MAT 1-1	MAT 1-2	p value
Region						
1	5	1	**0.219**	1	1	**1.5**
2	3	5	**0.727**	1	3	**0.625**
3	0	40	0.000	0	6	**0.31**
4	0	46	0.000	0	5	**0.063**
DAPC/Structure						
Cluster 1	5	82	0.000	1	9	0.022
Cluster 2	3	6	**0.508**	1	3	**0.625**

For $p < 0.05$ mating type ratio significantly deviated from the null hypothesis, bold p-values ($p > 0.05$) did not differ significantly from the null hypothesis of a 1:1 mating type ratio

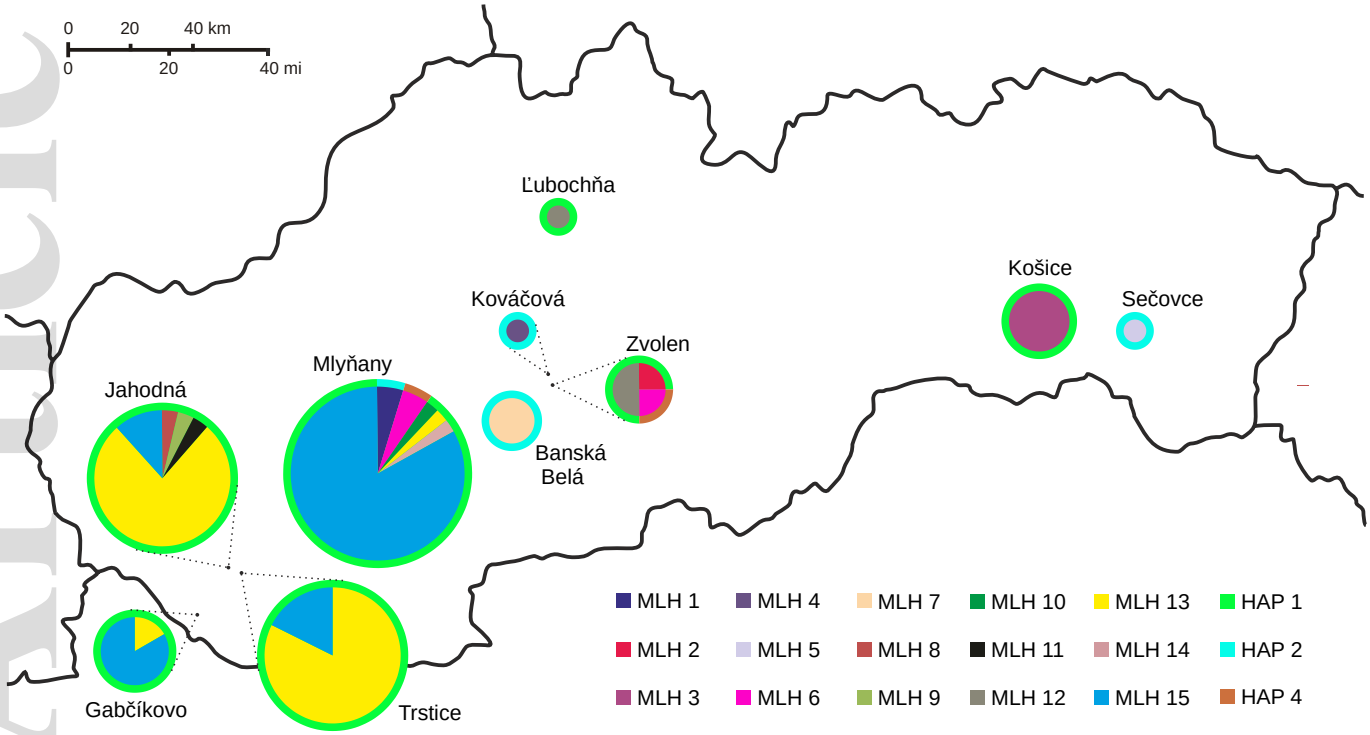
b

	Non-clone corrected dataset			Clone corrected dataset		
	I_A	r_d	p value	I_A	r_d	p value
Region						
1	7	1.00	0.000	na	na	na
2	2.72	0.55	0.000	2.15	0.44	0.001
3	5.05	0.74	0.000	2.59	0.37	0.00010
4	0	na	1	0	na	1
DAPC/Structure						
Cluster 1	0.197	0.107	0.003	0.023	0.01	**0.529**
Cluster 2	0.613	0.123	0.014	0.20	0.04	**0.341**

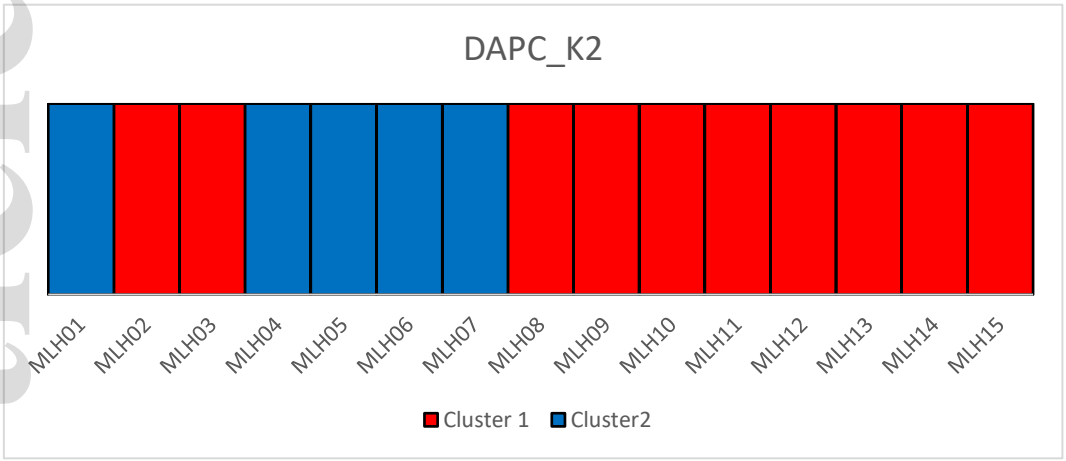
Bold p-values ($p > 0.05$) indicate random mating is supported by the test

I_A = index of association

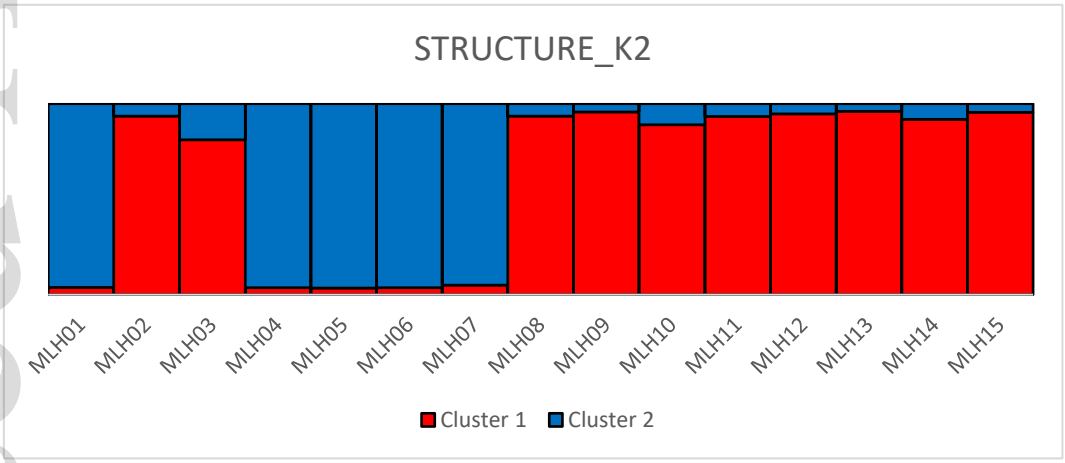
r_d = standardized index of association

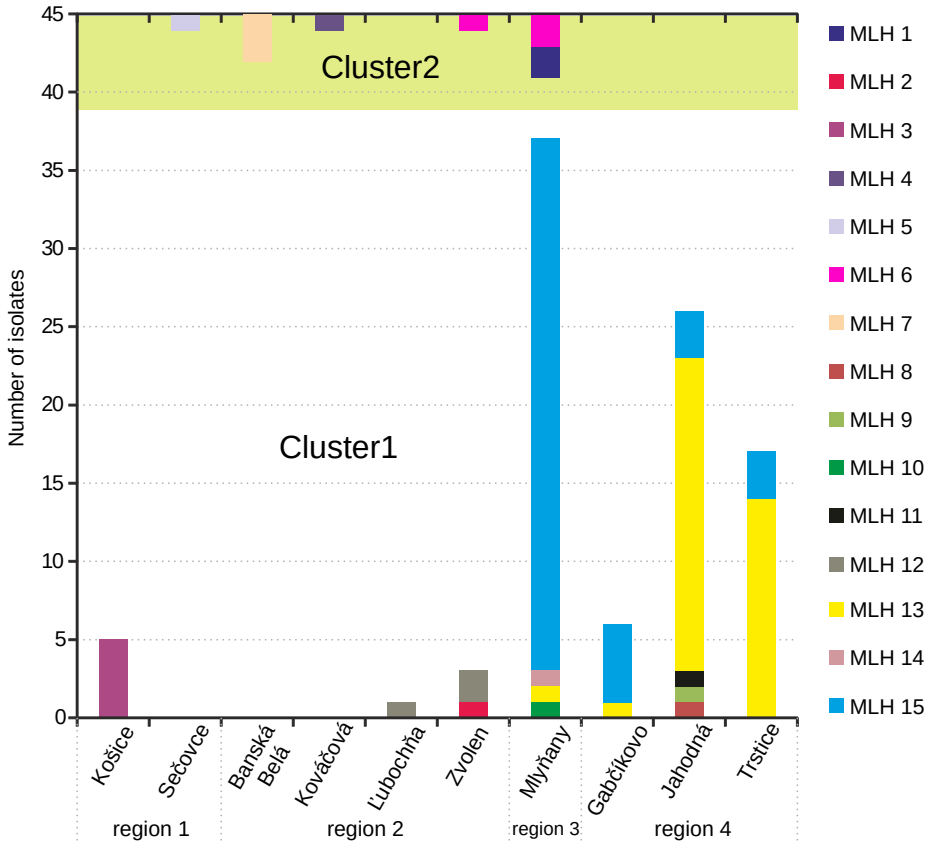


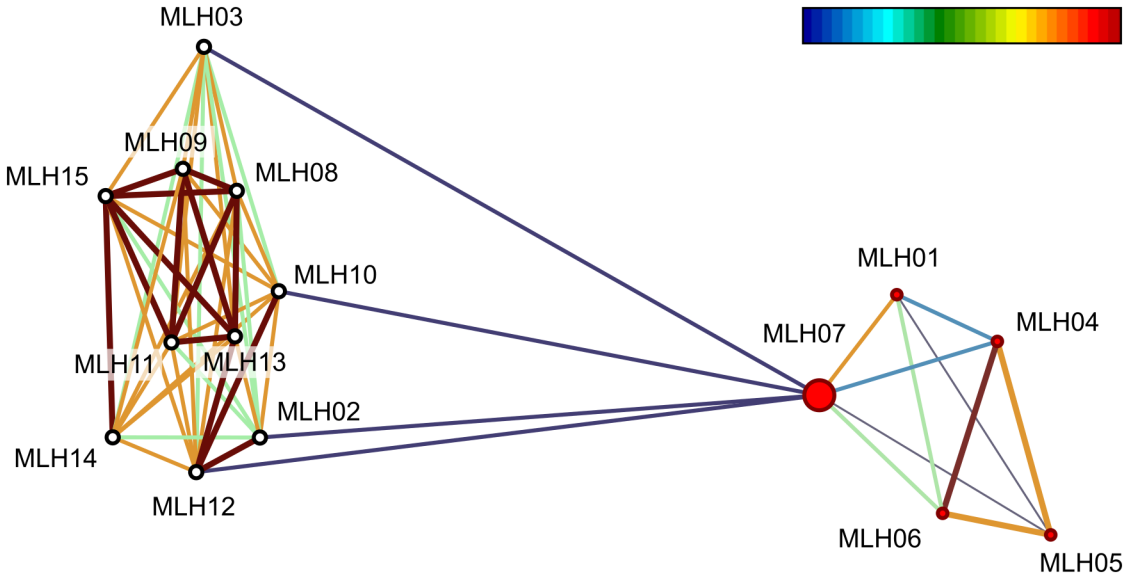
a



b







ppa_13428_f4.tiff