1. Genetic analyses suggest separate introductions of the pine pathogen *Lecanosticta acicola* into Europe

2. Authors:

Josef Janoušek, Michael J. Wingfield, José G. Marmolejo Monsivais, Libor Jankovský§, Christian Stauffer§, Adam Konečný§, Irene Barnes§

§ Equally contributing senior authors

3. Affiliations of authors:

Josef Janoušek, Department of Forest Protection and Wildlife Management, Faculty of Forestry and Wood Technology, Mendel University in Brno, Brno 613 00, Czech Republic

Michael J. Wingfield, Department of Genetics, Forestry and Agricultural Biotechnology Institute (FABI), University of Pretoria, Pretoria 0002, South Africa

José G. Marmolejo Monsivais, Facultad de Ciencias Forestales, UANL, Nuevo León 67700, Mexico

Libor Jankovský, Department of Forest Protection and Wildlife Management, Faculty of Forestry and Wood Technology, Mendel University in Brno, Brno 613 00, Czech Republic

Christian Stauffer, Department of Forest and Soil Sciences, University of Natural Resources and Life Sciences, Vienna 1190, Austria

Adam Konečný, Department of Botany and Zoology, Faculty of Science, Masaryk University, Brno 625 00, Czech Republic

Irene Barnes, Department of Genetics, Forestry and Agricultural Biotechnology Institute (FABI), University of Pretoria, Pretoria 0002, South Africa

3. Corresponding author: Josef Janoušek, Department of Forest Protection and Wildlife Management, Faculty of Forestry and Wood Technology, Mendel University in Brno, Zemědělská 3, Brno 613 00, Czech Republic, +420 545 134 103, janousek.jose@gmail.com

4. Keywords: Approximate Bayesian Computation, fungus, haploid, reproductive mode, *Mycosphaerella dearnessii*, population genetics

ABSTRACT

Janoušek, J., Wingfield, M.J., Marmolejo Monsivais, J. G., Jankovský, L., Stauffer, C., Konečný, A., Barnes, I., 2015, Genetic analyses suggest separate introductions of the pine pathogen *Lecanosticta acicola* into Europe, Phytopathology

Lecanosticta acicola is a heterothallic ascomycete that causes brown spot needle blight on native and non-native *Pinus* spp. in many regions of the world. In this study we investigated the origin of European *L. acicola* populations and estimated the level of random mating of the pathogen in affected areas. Part of the Elongation Factor 1-α gene was sequenced, eleven microsatellite regions were screened, and the mating type idiomorphs were determined for 201 isolates of *L. acicola* collected from three continents and 17 host species. The isolates from Mexico and Guatemala were unique, highly diverse and could represent cryptic species of *Lecanosticta*. The isolates from East Asia formed a uniform and discrete group. Two distinct populations were identified in both North America and Europe. Approximate Bayesian Computation analyses strongly suggest independent introductions of two populations from North America into Europe. Microsatellite data and mating type distributions indicated random recombination in the populations of North America and in Europe. Its inter-continental introduction can most likely be explained as a consequence of the movement of infected plant material. In contrast, the spread of *L. acicola* within Europe appears to be primarily due to conidial dispersion and probably also ascospore dissemination.

Increasing human activity related to globalisation and climate change is increasing the risks of biological invasions by plant pathogens (Grunwald and Goss 2011; Wingfield et al. 2011). Such biological invasions by fungal pathogens have already resulted in diseases that have modified environments and reduced natural biodiversity (Desprez-Loustau et al. 2007; Fisher et al. 2012; Garbelotto 2008). In this regard, there is considerable concern for the long-term preservation of sensitive forest ecosystems (Boyd et al. 2013; Wingfield et al. 2015) especially since invasions by forest pathogens have been increasing exponentially in Europe during the last four decades (Santini et al. 2013). Serious consequences of such introductions into new environments are well documented, e.g. ash dieback in Europe caused by *Hymenoscyphus fraxineus* (Gross et al. 2014) and *Phytophthora ramorum* that causes sudden oak death of oak and tanoak in native forests in North America, and a rampant nursery disease known as Ramorum blight (Eyre and Garbelotto 2015). It is, therefore, important to correctly identify fungal pathogens and to understand the processes underpinning introductions into new environments.

Population genetics can provide a reliable means to indentify and understand the invasion history of plant pathogens (Giraud et al. 2008; Grunwald and Goss 2011). In this regard, a statistical method receiving increasing attention is Approximate Bayesian Computation (ABC; Beaumont et al. 2010; Beaumont et al. 2002). ABC is a statistical framework that calculates the relative probabilities of complex, competing models of evolutionary history of populations and estimates the demographic parameters underlying a given model (Bertorelle et al. 2010; Sunnåker et al. 2013). This approach can be used to reconstruct the demographic history of invasive species (Boissin et al. 2012; Guillemaud et al. 2010; Konečný et al. 2013), including fungal plant pathogens (e.g. Dilmaghani et al. 2012; Dutech et al. 2012).

Mode of reproduction plays an essential role in the genetic diversity of ascomycetes and in their infection biology (McDonald and Linde 2002; Milgroom 1996). The presence of sexual

reproduction in a population can reflect the ability of the pathogen to become invasive in a new environment because it has an influence on pathogen overwintering, dissemination, infection biology and its ability to adapt to environmental conditions (Barrett et al. 2008; Bazin et al. 2014; Giraud et al. 2008). It is, however, commonly difficult to detect sexual recombination in a fungal population where this mode of reproduction is cryptic or facultative (Milgroom 1996; Saleh et al. 2012). To overcome this obstacle, the distribution of the mating type idiomorphs can be determined in a given population. The presence of both idiomorphs at approximately equal frequencies suggests that regular sexual reproduction is occurring, while presence of single idiomorph would indicated the population is undergoing asexual reproduction (Barnes et al. 2014; Linde et al. 2003; Taylor 1999).

Lecanosticta acicola (Thüm.) Syd. (sexual state: *Mycosphaerella dearnessii* M. E. Barr) is a heterothallic (Janoušek et al. 2014) ascomycete that causes brown spot needle blight on pines. The pathogen is reported to infect more than 30 pine species, variable levels of susceptibility being reported across and within pine species (Sinclair and Lyon 2005; Tainter and Baker 1996). Severe infection of needles by *L. acicola* can lead to retardation of growth and death of trees (Kais 1975). The pathogen reproduces asexually via conidia that are dispersed predominantly by rain splash and dew. Sexual ascospores are dispersed primarily in air currents over long distances (Kais 1971; Siggers 1944; Wolf and Barbour 1940). Little is known, however, regarding the extent to which sexual reproduction contributes to the life cycle or to the spread of brown spot needle blight in affected regions.

Brown spot needle blight has been known on *Pinus palustris* and other pine species in the south and south-east of the USA since the 19th century where it has been an important constraint to *P. palustris* seedling regeneration (Sinclair and Lyon 2005; Thümen 1878). In the northern parts of North America, brown spot needle blight occurs predominantly on native *P. strobus* and non-native *P. sylvestris* (Boyce 1959; Laflamme et al. 2010; Stanosz 1990). In

Central America and Mexico, *L. acicola* has been described from sea-level tropical forests up to high altitude forests (Evans, 1984; Alonso and Perez 1987) on native pine species. In South America, it is has been reported from only non-native pine plantations in Colombia (Gibson, 1980). In Europe, *L. acicola* was first reported in Spain in 1940's (Martínez 1942) and later, in former Yugoslavia, France and in Central Europe (Jankovský et al. 2009; Lévy and Lafaurie 1994; Milatović 1976). It is currently listed as a quarantine pathogen by the European Plant Protection Organisation (EPPO; Pehl et al. 2015). *L. acicola* also causes severe defoliation of non-native pines in China (Huang et al. 1995) where it was suggested, based on RAPD marker analyses, that the pathogen was introduced from southern USA. However, the origin of the pahogen in Europe and other regions of the world remains unknown.

The objective of this study was to consider the origin of European populations of *L. acicola* and to estimate the opportunities for its sexual reproduction in screened populations. More specifically, (i) we investigated the intra-specific variability of *L. acicola* on a global scale, (ii) determined the genetic structure and diversity of populations, (iii) deciphered the historico-demographical relationships between North American and European populations and (iv) determined and compared level of random mating in studied populations of *L. acicola*. Part of the elongation factor (EF) gene was sequenced, eleven microsatellites were screened and the mating type idiomorphs were determined for populations of isolates collected from three continents and 17 pine species.

MATERIALS AND METHODS

Sample collection, isolation, DNA extraction and isolate identification

Needle samples with developed fruiting bodies of *L. acicola* were collected randomly from individual trees by authors or obtained from colleagues (for details about location and number of isolates, see Table 1 and Supplementary Table S1). Samples were collected from the Americas (Colombia, Guatemala, Mexico, USA and Canada), eleven European countries, China, South Korea and Japan.

Needle samples were stored at -80°C to prevent their deterioration and to kill mite and insect contaminants. Needles with developed fruiting bodies and showing *Lecanosticta*-like discolouration were selected for isolation and conidia were isolated as previously described by Barnes et al. (2004). All isolates were cultivated on 2% malt-extract agar media with yeast extract (5g/l) for 3-6 weeks. Isolates showing typical *Lecanosticta*-like morphology (Pehl et al. 2015) were chosen for DNA extraction and DNA-based species identification.

Fungal mycelium was transferred to Eppendorf tubes and lyophilised for several hours to facilitate subsequent homogenisation. DNA was extracted using the PowerSoil® DNA Isolation Kit (12888; MoBio; Carlsbad, CA, USA), eluted with 100 µl elution buffer following the manufacturer's instructions and stored at -20°C.

DNA-based identification was performed for each isolate with species-specific primers LAtef-F/R as described by Ioos et al. (2010). Isolates testing positive for *L. acicola* amplified a product of 237bp with this primer set. PCR products were visualised on 2% agarose gel stained with GelRed™ (Hayward, CA, USA).

DNA sequencing, intra-specific variation and phylogenetic analyses

The elongation factor (EF) region of 87 selected isolates from different locations was PCR amplified using elongation factor primers EF1/2 (Jacobs et al. 2004). PCR was performed using 2 mM $MgCl_2$, 100 µM dNTPs, 0.2 µM of the primers, 0.2U Taq polymerase (Fermentas, Vilnius, Lithuania), 1x $(NH_4)_2SO_4$ buffer (Fermentas), 2.0 µl of genomic DNA (gDNA) and H_2O to 20 µl volume. The PCR profile was as follows: 94°C 10 min, 35 cycles 94°C 30s, 58°C 45s, 72°C 60s and 72°C for 10 min. Correct amplification was verified by gel electrophoresis as described above and the amplicons were custom-sequenced at the Cancer Research Centre DNA Sequencing Facility (University of Chicago, Chicago, IL, USA).

Sequence data were edited using BIOEDIT v 7.2.0. (Hall 1999) and aligned using MAFFT v. 7 (Katoh et al. 2002; Katoh and Standley 2013). Additional EF sequences were retrieved from GenBank and included in the analyses: JX901650 (*L. guatemalensis*; Guatemala; IMI 281598), JX901651 and JX901652 (*L. longispora*; Mexico; CBS 133789, CPC 17940), JX901648 (*L. acicola*; Mexico; CBS 133789), JX901647 (*L. acicola*; France; LNPV 243), KT737239 (*L. acicola*; France; CBS 871.95) and KC013002 (USA; CBS 133791) - epitype of *L. acicola* (Quaedvlieg et al. 2012). Sequence datasets were compiled using MEGA 5.2 (Tamura et al. 2011) and EF haplotypes were defined using TCS 1.21 software (Clement et al. 2000). Sequences of isolates representing each EF haplotype were submitted to GenBank (Supplementary Table S1) and their alignment to TreeBASE (http://purl.org/phylo/treebase/phylows/study/TB2:S18403).

Nucleotide diversity (Pi) was calculated for selected groups of EF haplotypes using DnaSP v5 (Librado and Rozas 2009). Median-joining haplotype network was constructed using program Network 4.613 (http://www.fluxus-engineering.com/sharenet.htm; Bandelt et al. 1999) to depict relationships amongst different EF haplotypes.

Phylogenetic analyses were done using maximum parsimony (MP), maximum likelihood (ML) and Bayesian Inference (BI) and results visualised in MEGA 5.2. MP analyses were performed using PAUP 4.0b10 (Swofford 2003). Random stepwise addition heuristic searches were performed with tree-bisection-reconnection branch-swapping. Alignment gaps were considered as a fifth character state and confidence was estimated by performing 1,000 bootstrap replications (Felsenstein 1985) with simple sequence addition.

For the ML analyses, the nucleotide substitution model with the best likelihood for the dataset was selected with jModelTest v. 2.1.1 (Darriba et al. 2012; Guindon et al. 2003) using the Akaike Information Criterion. ML analyses were carried out in PhyML 3.0 (Guindon et al. 2010). Confidence levels were estimated from 1,000 bootstrap replicates.

BI analysis was performed with MrBayes 3.1.2 (Ronquist and Huelsenbeck 2003) running $3 \cdot 10^6$ Markov Chain Monte Carlo (MCMC) generations. Four runs were performed and trees were sampled every 100^{th} generation. The generalised time reversible substitution model with gamma-distributed rate variation across sites and the proportion of invariable sites was selected. The burn-in value was determined with TRACER v. 1.5 (Rambaut and Drummond 2007; http://tree.bio.ed.ac.uk/software/tracer/) and the log-likelihood scores of sampled trees were plotted against the generation time to compare the results of each run. Nodes with a posterior probability ≥ 0.95 were considered to be significantly supported by the data.

Microsatellite genotyping

All isolates except those from Asia, South America, Guatemala and Mexico were screened with eleven microsatellite markers. PCR amplification of the microsatellite regions was performed as described in Janoušek et al. (2014) with the exception of the annealing temperature for primer set MD6 and MD8 being increased to 63°C to reduce stutter bands. Annealing temperatures were decreased for the isolates originating from Mexico and

Guatemala as follows: 54°C for MD1, 5, 9, 10 primer sets and to 56°C for primer set MD8. Correct amplification was verified by gel electrophoresis on a sub-set of samples. Amplicons were pooled into two panels (Janoušek et al. 2014) for fragment analysis on an ABI 3730XL (Applied Biosystems) and sized with LIZ 500 size standard (Applied Biosystems). Lengths of alleles were scored using GeneMapper^{TM} 4.1 (Applied Biosystems). Isolates from the same location with identical multilocus haplotypes (MLHs) were considered as clones and excluded for selected analyses (clone-corrected data set).

Analyses of population structure

The program STRUCTURE 2.3.4 was used to identify genetically different sub-populations within and between North-American and European isolates. Using a model-based clustering method, the most likely number of genetic groups (K), was determined by employing a Bayesian MCMC clustering algorithm (Falush et al. 2003; Hubisz et al. 2009; Pritchard et al. 2000).

The full clone-corrected haploid data set was analysed running $2 \cdot 10^5$ burn-in iterations followed by $6 \cdot 10^5$ MCMC iterations. The model with correlated allele frequencies and allowing admixture was selected. Thirty replicates for each K (1-6) were performed to increase the precision of the parameter estimates and to reduce the effect of stochasticity of the MCMC algorithm (Excoffier and Heckel 2006). The analysis was repeated on a subset of isolates representing a cluster identified in the first run (marked by blue in Fig. 3), employing nine polymorphic markers (two loci were monomorphic within blue group). STRUCTURE HARVESTER (Earl and vonHoldt 2012) was used on-line to generate input files of each K for CLUMPP v1.1.2 (Jakobsson and Rosenberg 2007). CLUMPP was used to identify potential dissimilar solutions among the results of individual STRUCTURE runs for each K. For this purpose, the Greedy algorithm implemented in CLUMPP was used to calculate the

pair-wise 'symmetric similarity coefficient' (SSC) to identify similar runs (SSC > 0.9) of each K. The averaged results for similar runs from CLUMPP were then used to generate summary bar plots for each K with DISTRUCT version 1.1 (Rosenberg 2004).

Complementary to the Bayesian analysis, Principal Component Analysis (PCA) was conducted on the clone-corrected data set. PCA presents genotypes in a multivariate space described by the principal components and it does not rely on any population genetic model (Jombart et al. 2009; McVean 2009). The analysis was performed using GENETIX v. 4.0.5.2 (Belkhir et al. 1996–2004; http://kimura.univ-montp2.fr/genetix/). As in the STRUCTURE analyses, the full dataset with all microsatellite loci was analysed first and then repeated using only the isolates identified in the northern cluster with nine microsatellite loci. PCA graphics were visualised by two principal components as recommended by Jombart et al. (2009).

Genetic differentiation was calculated between pairs of genetic STRUCTURE and PCA-defined populations (CE, central and north Europe; MS, Mississippi; NA, north-east North America; SE, south-west Europe) using FSTAT v2.9.3.2 (Goudet, 2002). Additionally, hierarchical analyses of molecular variance (AMOVA; Excoffier et al. 1992) implemented in GenAlEx 6.5 (Peakall and Smouse 2006, 2012) were performed to investigate the relative contributions of host species and geographic origin to the partitioning of genetic variance. The non clone-corrected microsatellite data set of the European isolates was used for AMOVA and isolates originating from unknown pine species were excluded. All data were treated as haploid-SSR and permutation tests were performed with 9,999 random permutations of haplotypes.

Genetic diversity in populations

The isolates were divided into four STRUCTURE and PCA-defined populations for genetic diversity analyses (CE; MS; NA; SE; Table 2). Furthermore, isolates were divided based on

their geographic origin: isolates more than 100 km apart were divided into separate groups (Table 2 and Fig.1). Although in some cases, only one or two isolates represented a single location, these isolates were included in the analyses of populations representing larger geographic areas. Isolates from Asia and South America were excluded as the sample size was inordinately low. Isolates from Mexico and Guatemala amplified poorly at most loci and thus were excluded from the analyses.

The number of MLHs was calculated for each population and location. The Clonal Fraction (CF) was calculated as 1-[(number of different haplotypes)/(total number of isolates)] (Zhan et al. 2003). Genotypic diversity (*G*), defined as the probability that two individuals taken at random have different haplotypes, was calculated on non clone-corrected datasets using MULTILOCUS v1.3 (Agapow and Burt 2001). Number of alleles (N_a) and allelic richness (A_R) was calculated using FSTAT. Gene diversity was estimated over all loci for each population and location by calculating unbiased expected heterozygosity (H_e, Nei 1978) in GENETIX.

Migration scenarios

The Approximate Bayesian Computation (ABC) framework was used to elucidate the demographic history of the North-American and European populations of *L. acicola*. All ABC analyses were performed on clone corrected microsatellite data using the program DIYABC v.2.0.4 (Cornuet et al. 2014), which allows inferences to be made on the demographic history of populations of haploid species. Only the markers constituting perfect microsatellite repeat motifs were used for the ABC analyses. These included seven makers, namely MD1, MD2, MD4, MD7, MD8, MD9 and MD12.

The evolutionary scenarios were drawn based on four STRUCTURE and PCA-defined genetic clusters (CE; MS; NA; SE). In total, six different scenarios were tested to compare the

different evolutionary scenarios (graphically represented in Fig. 4). Prior distributions of the demographic parameters were defined as uniform and with a broad range, due to a very limited knowledge of *L. acicola* population history (see Supplementary Table S4 for details of priors).

The default settings for the mutational parameters of microsatellites, rates and modalities of mutation were used in DIYABC software, as these settings are commonly used for most Eukaryotes (Cornuet et al. 2014). Each locus was assumed to follow a generalised stepwise mutation model (GSM; Estoup et al. 2002) with a possible range of 40 contiguous allelic states. The allelic range was extended to 63 allelic states for locus MD8. The genetic variation within and between populations was summarised using a set of 'one sample summary statistics' (mean number of alleles, mean size variance) and 'two sample summary statistics' (mean genic diversity, F_{ST}, classification index, $(\delta\mu)^2$ distance) – as described in the DIYABC manual available at http://www1.montpellier.inra.fr/CBGP/diyabc/index.php. For each scenario, 10^6 data sets were simulated.

All scenarios tested were first compared by their relative posterior probabilities using polychotomous logistic regression from 1% of the closest simulated data sets to the observed one in a multidimensional space of summary statistics, which were transformed by linear discriminant analysis (LDA; Estoup et al. 2012). The scenario with the significantly highest posterior probability value (95% confidence interval) was selected as best. Secondly, posterior distributions of parameters were estimated for the most likely scenario by the logit transformation of parameters and linear regression on 1% of the closest simulated data sets.

Confidence in scenario choice was evaluated based on sets of 200 pseudo-observed data sets (pods), obtained by simulations for each scenario with parameter values taken from given distributions. The same number of loci and individuals as real data set was tested. The relative

posterior probabilities of each competing scenario were estimated for each pod using LDA-transformed summary statistics. These probabilities were further used to calculate the type I and II errors for the scenarios chosen with our real data set. Type I error (false positives) is the probability of excluding the selected scenario when it is the true scenario. Type II error (false negatives) is the probability of selecting the scenario when it is not the true scenario (mean of type II error was calculated over the competing scenarios).

The goodness-of-fit of the best scenario was evaluated using 'model checking' option. This analysis allows an evaluation of the extent to which the selected scenario and associated posterior distribution are supported by the observed *L. acicola* genetic data. The summary statistics, which were not used in the previous analyses, were used for model checking as recommended by Cornuet et al. (2010).

Reproductive mode

The mating type idiomorph for each isolate was determined using multiplex PCR runs in 6 µl volumes consisting of 2.5 mM $MgCl_2$, 100 µM dNTPs, 0.2 µM of each MAT specific primer (Janoušek et al. 2014), 0.2U Taq polymerase (Fermentas), 1x $(NH_4)_2SO_4$ buffer (Fermentas), 0.8 µl of gDNA and H_2O. The PCR products were visualised using gel electrophoresis and the idiomorph of each isolate was determined based on the expected size of 634 bp for MAT1-1-1 and 323 bp for MAT1-2. The distribution of the mating type idiomorphs in the populations was calculated and an exact binomial test was performed using the programme STATISTICA 10 (StatSoft CR s.r.o.) to determine whether populations / sampled locations deviated from the null hypothesis of a 1:1 ratio of random mating (McDonald 2009).

Clone-corrected data-based parsimony tree-length permutation tests (PTLPT; Burt et al. 1996) and multilocus linkage disequilibrium analyses were performed to test for random mating amongst isolates from selected locations. In the populations where random mating occurred,

the expected parsimony tree length would be substantially longer. In contrast, a clonally reproducing population would produce one well-resolved, significantly shorter tree. The data sets were generated using MULTILOCUS v1.3 (Agapow and Burt 2001) and analysed using PAUP 4.0b10 (Swofford 2003). To assess the statistical significance associated with the null hypothesis of random mating, 1,000 randomisations were performed. MULTILOCUS v1.3 was used to calculate the standardised index of association, expressed by *rBarD*, which is sample size independent (Agapow and Burt 2001). Datasets were randomised 10,000 times. The observed values of *rBarD* were compared with the values of the randomised datasets. If the observed value was significantly different from the randomised dataset ($P \leq 0.05$), the null hypothesis was rejected. In contrast, if observed and randomised datasets were not significantly different, this indicated that the analysed population is undergoing random mating.

RESULTS

Sample collection, isolations, DNA extractions and isolate identification

Isolates that showed morphologies typical of *Lecanosticta* were obtained from infected needles and DNA was extracted from these isolates. In total, 201 isolates were identified as *L. acicola*. Twenty two isolates originated from the south-east of Canada, 88 isolates from six states in the USA, 78 isolates from 11 countries in Europe, three isolates from Asia (China, South Korea and Japan), eight isolates from Mexico and Guatemala, and two isolates from South America (Colombia). Overall, the isolates originated from 17 pine species collected from three continents (Table 1 and Supplementary Table S1, Fig. 1). All isolates are maintained in the culture collection of the Department of Forest Protection and Wildlife Management, Mendel University in Brno, Czech Republic and/or in the culture collection (CMW) of the Forestry and Agricultural Biotechnology Institute (FABI), University in Pretoria, South Africa (Supplementary Table S1).

DNA sequencing, intra-specific variation and phylogenetic analyses

Partial EF amplicons of about 900 bp were sequenced and fourteen haplotypes were identified in 87 isolates originating from America, Europe and Asia (Fig. 1; Supplementary Table S1). All eight isolates from Mexico and Guatemala revealed a nucleotide diversity of Pi = 0.0356. For all other regions, the nucleotide diversity was considerably lower with the next highest diversity found in Mississippi (Pi = 0.0039) where four haplotypes were identified from 19 sequenced isolates. One of these haplotypes (GenBank Access. no. KJ938451; designated in yellow in Fig. 1 and Supplementary Fig. S1; Fig.2; Supplementary Table S1) was present in 14 isolates from Mississippi and all isolates from France, Spain and Colombia. A single haplotype (KJ938438; designated in blue in Fig. 1 and Supplementary Fig. S1; Fig. 2; Supplementary Table S1) was identified in all the isolates from five states in the northern

USA (n = 27), Quebec in Canada (n = 7) and from nine countries in Central and North Europe (n = 23). The three isolates from China, South Korea and Japan shared the same unique EF haplotype (KJ938450; designated as dark red in Fig. 1 and Supplementary Fig. S1; Fig. 2; Supplementary Table S1).

Haplotype network (Supplementary Fig. S1) revealed high level of diversity for isolates from Guatemala and Mexico. All haplotypes detected in Europe, Asia, Colombia, Canada and USA are closely related to those originating from Mexico.

The phylogenies generated with ML, BI and MP analyses were congruent and are represented as a single MP tree with corresponding bootstrap support values for the nodes indicated on the branches (Fig. 2). For the BI analysis, the burn-in period was determined for the first 10 000 generations and all runs produced trees of the same log-likelihood.

In the phylogenies, two major clades were obtained. One clade included only the isolates originating from Mexico and Guatemala also revealing high diversity (Fig. 2). The isolates from Mexico and all the isolates Europe, North America and Asia clustered together into one clade having high bootstrap support. Phylogenetic analyses suggested that the Mexican & Guatemalan clade is ancestral to the second clade (Fig. 2).

Microsatellite genotyping

All isolates were successfully amplified with the 11 microsatellite markers (Supplementary Table S2). The only exceptions were the isolates from Mexico and Guatemala that amplified with lower amplification success (80%) despite several PCR optimisation attempts. In addition, marker MD6 was monomorphic among all eight of these isolates. Isolates from Colombia and Asia were not included in some of the analyses due to small sample sizes in these populations.

365 **Analyses of population structure**

366 From 30 runs, STRUCTURE consistently identified two major groups (K=2) for the 129
367 isolates analysed. The first group included all the isolates from Mississippi (MS) and the
368 south-west of Europe (SE) and collectively, is designated as the 'southern lineage'. The second
369 group, designated as the 'northern lineage', included all the isolates from north-east of North
370 America and those from central, south-eastern and northern Europe (Supplementary Fig. S2).
371 At K=3, the isolates from SE (France and Spain) separated out from the MS population in the
372 southern lineage. K=4 and above revealed a collapse of genetic structure within the MS
373 population. In the northern cluster, 33% of 30 runs at K=5 and 63% of 30 runs at K=6
374 distinguished a European (CE) group from a North American group (NA) cluster
375 (Supplementary Fig. S2). Separate STRUCTURE analysis of the northern cluster at K=2
376 confirmed the presence of these two distinct groups (Supplementary Fig. S3).

377 The principal component analysis (PCA) also identified two major groups across the whole
378 dataset, corresponding to those obtained with STRUCTURE at K=2. The SE population
379 formed a cluster, distinct from the MS population (Fig. 3). When only the northern cluster
380 obtained in STRUCTURE was analysed using PCA, a faint distinction was evident between
381 the European and the North American group (Fig. 3).

382 Genetic differentiation between pairs of STRUCTURE and PCA-defined clusters was highest
383 between the populations from the 'southern lineage' and those from the 'northern lineage' (F_{ST}
384 = 0.662 - 0.471; Supplementary Table 3). Lower population genetic differentiation was seen
385 between populations of the same lineage (MS vs. SE: F_{ST} = 0.351; NA vs. CE: F_{ST} = 0.286).

386 Analysis of molecular variance (AMOVA) based on φPT value shows that only a low
387 percentage (12.18%) of molecular variance could be attributed to host species in Europe.

Geographic origin of European isolates contributed to molecular variance only marginally (3.2%; Supplementary Table S3).

Genetic diversity in populations

A total of 43 multilocus haplotypes (MLHs) from 67 isolates were detected in the NA population from collections made in seven locations in the north-east of North America (CF = 0.358; Table 2). The population from a single location in Mississippi contained 34 MLHs from 40 isolates analysed (CF = 0.150). The population originating from two locations in south-west Europe consisted of nine isolates in which five MLHs were determined (CF = 0.444). The CE population, spanning 10 locations, was comprised of 69 isolates in which 48 MLHs were obtained (CF = 0.289; Table 2).

The population from MS had the highest level of allelic richness (7.1±0.89), highest number of alleles (8.81), highest genotypic diversity (0.992) and highest genetic diversity (0.46±0.32) of all the populations analysed (Table 2). The NA population, representing isolates from a considerably larger area than the MS population, revealed a lower allelic richness (1.59±0.21), number of alleles (3.55), genotypic diversity (0.955) and genetic diversity (0.32±0.30). The SE population revealed the lowest values of the indices calculated (Table 2). The CE population had an allelic richness of 3.65±0.71, 4.54 number of alleles, a genotypic diversity of 0.931 and a genetic diversity of 0.35±0.27.

Migration scenarios

ABC analysis showed that the introduction of *L. acicola* into Europe was most likely from North America (scenario 6, relative posterior probability: 0.505; Fig. 4 and 5). This scenario assumed an unknown ancestral population that gave rise to the NA population. The MS population split from the NA and these two American populations gave rise to the two European populations, SE and CE, respectively. Scenarios assuming the opposite direction of migration, i.e. from Europe to North America, had the least support (relative posterior

probability of scenario 2 and 3: 0.000, 0.0087; Fig. 4). Relative posterior probabilities were higher for scenarios assuming the presence of an ancestral population having given rise to one of the American populations (Fig. 4).

Posterior distributions of parameters for the best-supported scenarios are shown in Supplementary Table S4. Posterior distributions of effective population sizes for all sampled populations were relatively narrow and thus informative. The mode of the MS population was determined at 5,530, whereas it was 88 for the SE population. The mode of the NA population was identified at 829 and at 1,410 for the European CE population. The posterior distribution for the time of the split was determined as follows: divergence of CE from NA (t1c) occurred 127 generations ago, the divergence of SE from MS (t1h) occurred 654 generations ago, the divergence of NA+CE populations from the MS+SE populations occurred probably 3,410 generations ago.

Power analyses revealed that the type I error (false positive) associated with the best-supported scenario (scenario 6) was relatively high (0.34). The mean of the type II errors (false negatives) associated with the best-supported scenario was low (0.078), indicating that the probability of selecting the best scenario when the data were simulated with an alternative scenario, was low. Twenty-six summary statistics that had not been previously used for model selection were used for model checking. Only one of the 26 statistics had a low probability value when the model was checked (Supplementary Table S5), indicating that the selected scenario fitted the observed data well.

Reproductive mode

Frequency of the mating type idiomorphs for the isolates of *L. acicola* differed for different regions (Table 2 and Supplementary Table S1). Both mating type idiomorphs were identified at equal ratios (MAT1-1/2: 4/4) in Guatemala and Mexico. Similarly, the mating types did not

differ significantly from a 1:1 ratio in the population from Mississippi but they did differ from this ratio for the NA population (Table 2). At a smaller geographical scale, isolates from CAP and VMW showed a similar skewed ratio for the MAT1-1/MAT1-2 at 17/1 and 22/1, respectively. Both mating type idiomorphs were found for isolates collected from the same needle originating from Canada. In Europe, the mating type ratio was 4/5 in SE and 23/44 in CE population (Table 2). For individual locations such as GBU (Germany) and ATN (Austria), the ratios did not significantly differ from 1:1. Only the MAT1-2 idiomorph was identified in isolates from the Czech Republic site Červená Blata (25 isolates), whereas only the MAT1-1 idiomorph was identified in isolates (three) collected in Borkovická Blata. Both mating type idiomorphs were found in isolates from a single needle collected in France. Only the MAT 1-2 idiomorph was found in isolates from South America (Colombia; two isolates) and Asia (China, South Korea and Japan; three isolates; Supplementary Table S1).

Parsimony tree-length permutation tests (PTLPT) revealed relatively long trees for the MS dataset and did not significantly differ from randomised trees (Table 2). The isolates originating from two locations in the north-eastern part of North America, CAP and VMH, had shorter tree lengths (22 and 18 steps, respectively), but they also did not differ significantly from randomised trees (Table 2). The same scenario was true for the isolates from two European locations (GBU, ATN) that produced relatively short trees at 12 and 23 steps, respectively (Table 2). In contrast, isolates from location CZB in the Czech Republic, produced a tree of 22 steps, which was significantly different from randomised trees (Table 2).

Another measure of random mating in the populations, *rBarD*, did not reveal significant differences from a randomised dataset (Table 2) in the MS population or in the isolates from CAP (Quebec, Canada). Measure of linkage disequilibrium was significantly different for the isolates from VMW (Vermont, USA; Table 2). The three European locations showed

consistent results with the PTLPT analyses: isolates from GBU (Germany) and ATN (Austria) were not significantly different from randomised data simulating a recombining population. CZB (Czech Republic) showed significant linkage disequilibrium and therefore, did not show any evidence for sexual recombination (Table 2).

DISCUSSION

This is the first investigation to consider the global movement of the pine needle pathogen *L. acicola* in different regions of the world. We detected high diversity amongst the isolates from Guatemala and Mexico and some of these most likely represent cryptic species closely related to, but distinct, from *L. acicola*. The isolates from East-Asia formed a unique and discrete group. Furthermore, two distinct populations in North America (southern and northern one) were identified and both populations have very likely been introduced into Europe independently. Interestingly, these two European populations of *L. acicola* reflect the geographic distribution of North-American populations, which are probably related to climatic and host adaption for each of the lineages. In addition, evidence was provided for sexual recombination within the pathogen populations in parts of Europe and in North American. It was thus clear that *L. acicola* is an invasive alien in Europe that reproduces asexually and very likely also sexually.

Phylogenetic relationships within *L. acicola*, population structure and genetic diversity

Haplotype and nucleotide diversity, determined from the EF sequences, revealed high diversity amongst the isolates from Guatemala and Mexico. These results are consistent with the view of Evans (Evans 1984), who attributed substantial morphological variation to the existence of a species complex in *L. acicola*. The high level of genetic diversity for *L. acicola* in this region could be attributed to its long-term or native presence in Central America and Mexico, where it occurs from sea-level tropical forests to high altitude rain forests (Evans 1984). The pathogen is omnipresent on native pines but it was occasionally also found to cause a serious needle blight (Evans 1984; I. Barnes, unpublished data). This suggests a high level of host resistance, probably as a result of long-term plant-pathogen co-evolution (Barrett et al. 2008). High diversity amongst the isolates originating from Mexico is probably also

related to the high diversity of pine species found in Mexico that is also the highest at global scale (Farjon 1996; Gernandt and Pérez-de la Rosa 2014).

The fact that the microsatellite markers, designed from isolates from USA, Europe and Japan, amplified poorly in the Central American isolates adds credence to the view that *L. acicola* represents a species complex in Central America and Mexico. Phylogenetic analyses indicated that the Central American and Mexican haplotypes of *L. acicola* are ancestral to those found elsewhere in the world. Phylogenetic analyses indicated that Mexico is probably the area of origin of *L. acicola* populations in the USA and Canada.

Haplotypic and nucleotide diversity for isolates of *L. acicola* from North America was lower than that found in Mexico and Guatemala. This corresponds to a more uniform conidial morphology observed by Evans (1984) in collections from North America. The population from Mississippi encompassed the highest genetic diversity and allelic richness of all populations analysed. Two distinct lineages ('southern' and a 'northern' lineage) reported by Huang et al. (1995) from their RAPD data, were also observed in this study using both EF sequence data and microsatellite markers. These two lineages have been reported to differ in culture morphology, conidial germination and pathogenicity to various pine species (Huang et al. 1995; Kais 1972). All these findings suggest a level of adaption to climatic conditions and host for isolates in the two lineages, a view also proposed by Huang et al. (1995). These two lineages could represent two distinct but cryptic species (Restrepo et al. 2014). This is likely given the fact that many cryptic species are being discovered in ascomycete fungi where DNA sequence data (e.g. Sakalidis et al. 2013; Walker et al. 2011) or microsatellite markers (e.g. Pérez et al. 2012; Schoebel et al. 2013) are applied to taxonomic and population genetic studies.

514 The same two lineages of *L. acicola* present in North America were also found in collections
515 of isolates from Europe and they had similar geographical distributions. The southern lineage
516 was identified in Spain and France, and the northern lineage in several other parts of Europe.
517 Their distinct geographical distribution is very likely not associated with pine species but
518 could be explained by their different climatic requirements (especially to temperature). The
519 presence of the more virulent (Kais 1972) southern lineage in France could explain a serious
520 outbreak of brown spot needle blight in the 1990's (Lévy and Lafaurie 1994). This epidemic
521 resulted in the total loss of about 270 ha *Pinus attenuata* x *radiata* plantations (Lévy 1996). It
522 is also interesting, that in both North America and Europe, neither the southern or northern
523 lineages have overlapping geographical distributions.

524 The southern lineage of *L. acicola* was identified in Colombia where the pathogen has caused
525 severe defoliation in pine plantations in the past (Evans 1984). Pines are not native in South
526 America and it is highly likely that *L. acicola* was introduced into this area with plant
527 material. A similar human-mediated introduction has been documented for the related pine
528 needle pathogen, *Dothistroma septosporum,* from Chile into Ecuador (Barnes et al. 2014).
529 However, in order to determine the origin of *L. acicola* in Colombia, additional investigations
530 will be required.

531 The *L. acicola* lineage identified in East Asia was unique and found only in this region. Its
532 introduction was suggested to have occurred from the southern or south-eastern USA (Huang
533 et al. 1995). The present study shows clearly that the isolates sampled in East Asia do not
534 form part of the same southern lineage that is present in the southern USA or Europe.
535 However, the introduction of *L. acicola* into Asia from an un-sampled population in the
536 south-eastern USA cannot be ruled out because only isolates from Mississippi were
537 considered in this study.

Evolutionary relationships between North-American and European populations

Analyses of demographic history revealed that both North-American lineages of *L. acicola* were most likely introduced into Europe independently after the divergence of the southern and northern lineages in North America. This pattern of multiple introductions from different sources seems to be common for worldwide invasive species as has been documented for various organisms using Approximate Bayesian Computation (ABC) approaches (Barrès et al. 2012; Konečný et al. 2013). Historically (based on ABC analyses), the southern lineage was introduced into Europe first, followed by the northern lineage. This is, however, difficult to verify because there could be many interfering parameters including differing mean generation times in each lineage that could be influenced by climatic conditions. The time at which the southern lineage split from northern lineage was estimated to have been about 3400 generations ago. Although this is an estimated value, it shows that the divergence of these lineages is relatively ancient. If we consider life cycle of 1 year, this would correspond to period of approximately 3400 years ago. Although it is clear that the European isolates of *L. acicola* originated in North America, determination of the exact origin of these isolates would require more intensive and precise sampling in North America as well as in other parts of the world.

Reproductive mode

Both mating type idiomorphs were identified in the isolates of *L. acicola* originating from Mexico and Guatemala. This confirms that the fungus is heterothallic and that it probably undergoes sexual reproduction. This would be consistent with EF sequence data and with the fact that Evans (Evans 1984) observed sexual structures in *Lecanosticta* spp. from Central America.

Mating type idiomorph distribution and microsatellite data analyses provided strong indication that the population of *L. acicola* from Mississippi is undergoing sexual

563 reproduction. This result was expected, because sexual structures of the fungi were commonly
564 found on infected needles in the south-eastern USA (Kais 1971; Siggers 1944) and ascospores
565 have been detected throughout the whole year (Henry 1954). Results of this study also
566 indicated that random mating occurs in the population from north-eastern North America.
567 Moreover, the presence of both mating type idiomorphs in isolates from the same needle in
568 Canada would clearly increase opportunities for individuals of opposite mating type to
569 interact and reproduce sexually (Barnes et al. 2011; Linde et al. 2003). Neither asci, nor
570 ascospores have, however, been observed in the northern regions of North America (Evans
571 1984; Nicholls and Hudler 1972). Although windblown ascospores would be formed only
572 occasionally in this population, they could play an important role in causing rapid outbreaks
573 of *L. acicola*, such as those recently been observed in the north-east of USA (Munck et al.
574 2012). Mating type idiomorph frequencies as well as microsatellite data analyses suggest that
575 sexual reproduction is most likely occurring in *L. acicola* in Austria and Germany. In contrast,
576 data obtained from the isolates from the Czech Republic showed that the pathogen reproduces
577 predominantly asexually.

578 This study provides the first indication for sexual reproduction of *L. acicola* in Europe. Sexual
579 reproduction, which gives rise to windborne ascospores, could explain the current epidemics
580 of brown spot needle blight in Switzerland, Germany and Austria (Angst 2011; Blaschke
581 2002; Hintsteiner et al. 2012). Ascospore dissemination thus probably plays a major role in
582 the long distance dispersal of *L. acicola* within Europe (Wingen et al. 2013) as has also been
583 seen for the wheat pathogen *Phaeosphaeria nodorum* (Sommerhalder et al. 2010). Sexual
584 reproduction allows for the formation of new haplotypes, while asexual reproduction can
585 enhance fast multiplication of advantageous haplotypes. This could lead to local adaptation
586 (Milgroom 1996) of *L. acicola*.

587 **Conclusions**

We have shown that at least two introductions of *L. acicola* have occurred from North America into Europe in the past. This is another example of an inter-continental introduction of a plant pathogen into a new environment, most likely as a consequence of human activity. In addition, results showed that *L. acicola* most probably also reproduces sexually. The pathogen therefore has a substantial potential to adapt to new environments and infect new pine host species. Its presence in geographical areas with variable climates illustrates its high level of ecological tolerance and ability to adapt to new environments. It can thus be expected, that *L. acicola* will spread rapidly from infected trees to surrounding pine stands if climatic conditions were to become favourable for infection to occur. *L. acicola* is a serious pathogen and it remains on the European A2 list of quarantine pathogens. It is, therefore, important to monitor its spread and movement in future, including impact of a potential contact between isolates of the two lineages in Europe.

ACKNOWLEDGEMENTS

We are most grateful to our colleagues listed in Table S1 for providing needle samples or fungal isolates without which this study would not have been possible. We also thank Susanne Krumböck, Dagmar Palovčíková, Louise Innes, Dana Nelson, Martin Kostovčík, Coralie Bertheau, Hannes Schuller and Martina Kloudová for their contribution to this study. The project was supported financially by COST CZ LD12031 (DIAROD), the FPS COST Action FP1102 (DIAROD), Project Indicators of trees vitality Reg. No. CZ.1.07/2.3.00/20.0265 co-financed by the European Social Fund and the state budget of the Czech Republic, the Scholarship Foundation of the Republic of Austria (OeAD-GmbH, Austria) for Josef Janoušek. The research was supported with funding from the Austrian Science Fund FWF and the European Union Seventh Framework Programme FP7 2007–2013 (KBBE 2009-3) under grant agreement 245268 ISEFOR. Access to computing and storage facilities owned by parties and projects contributing to the National Grid Infrastructure MetaCentrum, provided under the programme "Projects of Large Infrastructure for Research, Development, and Innovations" (LM2010005), is greatly appreciated.

LITERATURE CITED

Agapow, P.-M., and Burt, A. 2001. Indices of multilocus linkage disequilibrium. Mol. Ecol. Notes 1:101-102.

Alonso, M.R., and Perez, P. 1987. Asociacion *Lecanosticta acicola*, *Lophodermium australe* y *Pestalozzia* sp. sobre *Pinus* spp. en Cuba. Rev. For. Baracoa 17:99–103.

Angst, A. 2011. Braune Föhren in Garten und Parks. Wald und Holz 4:41-42.

Bandelt H-J., Forster P., Röhl A. (1999) Median-joining networks for inferring intraspecific phylogenies. Mol. Biol. Evol. 16:37-48.

Barnes, I., Crous, P. W., Wingfield, B., and Wingfield, M. J. 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., Kirisits, T., Wingfield, M. J., and Wingfield, B. D. 2011. Needle blight of pine caused by two species of *Dothistroma* in Hungary. For. Pathol. 41:361-369.

Barnes, I., Wingfield, M. J., Carbone, I., Kirisits, T., and Wingfield, B. D. 2014. Population structure and diversity of an invasive pine needle pathogen reflects anthropogenic activity. Ecol. Evol. 4:3642-3661.

Barrès, B., Carlier, J., Seguin, M., Fenouillet, C., Cilas, C., and Ravigné, V. 2012. Understanding the recent colonization history of a plant pathogenic fungus using population genetic tools and Approximate Bayesian Computation. Heredity 109:269-279.

Barrett, L. G., Thrall, P. H., Burdon, J. J., and Linde, C. C. 2008. Life history determines genetic structure and evolutionary potential of host-parasite interactions. Trends Ecol. Evol. 23:678-685.

Bazin, É., Mathé-Hubert, H., Facon, B., Carlier, J., and Ravigné, V. 2014. The effect of mating system on invasiveness: some genetic load may be advantageous when invading new environments. Biol. Invasions 16:875-886.

Beaumont, M. A., Zhang, W., and Balding, D. J. 2002. Approximate Bayesian Computation in Population Genetics. Genetics 162:2025-2035.

Beaumont, M. A., Nielsen, R., Robert, C., Hey, J., Gaggiotti, O., Knowles, L., and Estoup, A. 2010. In defence of model-based inference in phylogeography. Mol. Ecol. 19:436-446.

Belkhir K., Borsa P., Chikhi L., Raufaste N. and Bonhomme F. 1996-2004. GENETIX 4.05, logiciel sous Windows TM pour la génétique des populations. Laboratoire Génome, Populations, Interactions, CNRS UMR 5171, Université de Montpellier II, Montpellier (France).

Bertorelle, G., Benazzo, A., and Mona, S. 2010. ABC as a flexible framework to estimate demography over space and time: some cons, many pros. Mol. Ecol. 19:2609-2625.

Blaschke, M. 2002. Die Nadeln der Latsche (*Pinus mugo*), ein Habitat fur zwei Pilzarten der Gattung *Mycosphaerella*. Mycol. Bavarica 5:56-60.

Boissin, E., Hurley, B., Wingfield, M. J., Vasaitis, R., Stenlid, J., Davis, C., de Groot, P., Ahumada, R., Carnegie, A., Goldarazena, A., Klasmer, P., Wermelinger, B., and Slippers, B. 2012. Retracing the routes of introduction of invasive species: the case of the *Sirex noctilio* woodwasp. Mol. Ecol. 21:5728-5744.

Boyce, J.S. 1959. Brown spot needle blight on eastern white pine. Plant Disease Reporter 43: 420.

Boyd, I. L., Freer-Smith, P. H., Gilligan, C. A., and Godfray, H. C. 2013. The consequence of tree pests and diseases for ecosystem services. Science 342:1235773.

Burt, A., Carter, D. A., Koenig, G., White, T. J., and Taylor, J. W. 1996. Molecular markers reveal cryptic sex in the human pathogen *Coccidioides immitis*. Proc. Natl. Acad. Sci. U.S.A. 93:770-773.

Clement, M., Posada, D., and Crandall, K. A. 2000. TCS: a computer program to estimate gene genealogies. Mol. Ecol. 9:1657-1659.

666 Cornuet, J. M., Ravigné, V., and Estoup, A. 2010. Inference on population history and model
667 checking using DNA sequence and microsatellite data with the software DIYABC (v1.0).
668 BMC Bioinformatics 11:401.
669 Cornuet, J. M., Pudlo, P., Veyssier, J., Dehne-Garcia, A., Gautier, M., Leblois, R., Marin, J.
670 M., and Estoup, A. 2014. DIYABC v2.0: a software to make approximate Bayesian
671 computation inferences about population history using single nucleotide polymorphism,
672 DNA sequence and microsatellite data. Bioinformatics 30:1187-1189.
673 Darriba, D., Taboada, G., Doallo, R., and Posada, D. 2012 jModelTest 2: more models, new
674 heuristics and parallel computing. Nat. Methods 9:772.
675 Desprez-Loustau, M. L., Robin, C., Buée, M., Courtecuisse, R., Garbaye, J., Suffert, F.,
676 Sache, I., and Rizzo, D. M. 2007. The fungal dimension of biological invasions. Trends
677 Ecol. Evol. 22:472-480.
678 Dilmaghani, A., Gladieux, P., Gout, L., Giraud, T., Brunner, P. C., Stachowiak, A., Balesdent,
679 M. H., and Rouxel, T. 2012. Migration patterns and changes in population biology
680 associated with the worldwide spread of the oilseed rape pathogen *Leptosphaeria*
681 *maculans*. Mol. Ecol. 21:2519-2533.
682 Dutech, C., Barrès, B., Bridier, J., Robin, C., Milgroom, M. G., and Ravigné, V. 2012. The
683 chestnut blight fungus world tour: successive introduction events from diverse origins in an
684 invasive plant fungal pathogen. Mol. Ecol. 21:3931-3946.
685 Earl, D. A., and vonHoldt, B. M. 2012. STRUCTURE HARVESTER: a website and program
686 for visualizing STRUCTURE output and implementing the Evanno method. Conserv.
687 Genet. Resour. 4:359-361.
688 Estoup, A., Jarne, P., and Cornuet, J. M. 2002. Homoplasy and mutation model at
689 microsatellitte loci and their consequences for population genetics analysis. Mol. Ecol.
690 11:1591-1604.

691 Estoup, A., Lombaert, E., Marin, J. M., Guillemaud, T., Pudlo, P., Robert, C. P., and Cornuet,
692 J. M. 2012. Estimation of demo-genetic model probabilities with Approximate Bayesian
693 Computation using linear discriminant analysis on summary statistics. Mol. Ecol.
694 Resour.12:846-855.
695 Evans, H. C. 1984. The genus *Mycosphaerella* and its anamorphs. Mycol. Papers 153:1-103.
696 Excoffier, L., and Heckel, G. 2006. Computer programs for population genetics data analysis:
697 a survival guide. Nat. Rev. Genet. 7:745-758.
698 Excoffier, L., Smouse, P. E., and Quattro, J. M. 1992. Analysis of Molecular Variance
699 Inferred From Metric Distances Among DNA Haplotypes: Application to Human
700 Mitochondrial DNA Restriction Data. Genetics 131:479-491.
701 Eyre, C. A., and Garbelotto, M. 2015. Detection, Diversity, and Population Dynamics of
702 Waterborne *Phytophthora ramorum* Populations. Phytopathology 105:57-68.
703 Falush, D., Stephens, M., and Pritchard, J. K. 2003. Inference of Population Structure Using
704 Multilocus Genotype Data: Linked Loci and Correlated Allele Frequencies. Genetics
705 164:1567-1587.
706 Farjon, A. F. L. S. 1996. Biodiversity of *Pinus* (*Pinaceae*) in Mexico: speciation and paleo-
707 endemism. Bot. J. Linn. Soc. 121:365-384.
708 Felsenstein, J. 1985. Confidence Limits of Phylogenies: An Approach Using the Bootstrap.
709 Evolution 39:783-791.
710 Fisher, M. C., Henk, D. A., Briggs, C. J., Brownstein, J. S., Madoff, L. C., McCraw, S. L.,
711 and Gurr, S. J. 2012. Emerging fungal threats to animal, plant and ecosystem health.
712 Nature 484:186-194.
713 Garbelotto, M. 2008. Molecular analysis to study invasions by forest pathogens.
714 Phytopathologia Mediterranea 47:183–203.
715 Gernandt, D. S., and Pérez-de la Rosa, J. A. 2014. Biodiversidad de *Pinophyta* (coníferas) en

México. Revista Mexicana de Biodiversidad 85:126-133.

Gibson, I.A.S. 1980. Two pine needle fungi new to Colombia. Trop. Pest Manag. 26:38–40.

Giraud, T., Enjalbert, J., Fournier, E., Delmotte, F., and Dutech, C. 2008. Population genetics of fungal diseases of plants. Parasite 15:449-454.

Goudet, J. 2002. FSTAT, a programme to estimate and test gene diversities and fixation indices (version 2.9.3).

Gross, A., Hosoya, T., and Queloz, V. 2014. Population structure of the invasive forest pathogen *Hymenoscyphus pseudoalbidus*. Mol. Ecol. 23:2943-2960.

Grünwald, N. J., and Goss, E. M. 2011. Evolution and population genetics of exotic and re-emerging pathogens: novel tools and approaches. Annu. Rev. Phytopathol. 49:249-267.

Guillemaud, T., Beaumont, M. A., Ciosi, M., Cornuet, J. M., and Estoup, A. 2010. Inferring introduction routes of invasive species using approximate Bayesian computation on microsatellite data. Heredity 104:88-99.

Guindon, S. and Gascuel, O. 2003. A Simple, Fast, and Accurate Algorithm to Estimate Large Phylogenies by Maximum Likelihood. Syst. Biol. 52:696-704.

Guindon, S., Dufayard, J. F., Lefort, V., Anisimova, M., Hordijk, W., and Gascuel, O. 2010. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. Syst. Biol. 59:307-321.

Hall, T. A. 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symposium Series 41:95-98.

Henry, B. W. 1954. Sporulation by the brown spot fungus on longleaf pine needles. Phytopathology 44:385-386.

Hintsteiner, M., Cech, T. L., Halmschlager, E., Stauffer, C., Kirisits, T., and Sieber, T. 2012. First report of *Mycosphaerella dearnessii* on *Pinus nigra* var. *nigra* in Austria. For. Pathol. 42:437-440.

741 Huang, Z.-Y., Smalley, E. B., and Guries, R. P. 1995. Differentiation of *Mycosphaerella*
742 *dearnessii* by Cultural Characters and RAPD Analysis. Phytopathology 85:522-527.
743 Hubisz, M. J., Falush, D., Stephens, M., and Pritchard, J. K. 2009. Inferring weak population
744 structure with the assistance of sample group information. Mol. Ecol. Resour. 9:1322-
745 1332.
746 Ioos, R., Fabre, B., Saurat, C., Fourrier, C., Frey, P., and Marçais, B. 2010. Development,
747 Comparison, and Validation of Real-Time and Conventional PCR Tools for the Detection
748 of the Fungal Pathogens Causing Brown Spot and Red Band Needle Blights of Pine.
749 Phytopathology 100:105-114.
750 Jacobs, K., Bergdahl, D. R., Wingfield, M. J., Halik, S., Seifert, K. A., Bright, D. E., and
751 Wingfield, B. D. 2004. *Leptographium wingfieldii* introduced into North America and
752 found associated with exotic *Tomicus piniperda* and native bark beetles. Mycol. Res.
753 108:411-418.
754 Jakobsson, M., and Rosenberg, N. A. 2007. CLUMPP: a cluster matching and permutation
755 program for dealing with label switching and multimodality in analysis of population
756 structure. Bioinformatics 23:1801-1806.
757 Jankovský, L., Palovčíková, D., and Tomšovský, M. 2009. Brown spot needle blight
758 associated with *Mycosphaerella dearnessii* occurs on *Pinus rotundata* in the Czech
759 Republic. Plant Pathol. 58:398.
760 Janoušek, J., Krumböck, S., Kirisits, T., Bradshaw, R. E., Barnes, I., Jankovský, L., and
761 Stauffer, C. 2014. Development of microsatellite and mating type markers for the pine
762 needle pathogen *Lecanosticta acicola*. Australas. Plant Pathol. 43:161-165.
763 Jombart, T., Pontier, D., and Dufour, A. B. 2009. Genetic markers in the playground of
764 multivariate analysis. Heredity 102:330-341.
765 Kais, A. G. 1971. Dispersal of *Scirrhia acicola* spores in southern Mississippi. Plant Disease

Reporter 55:309-311.

Kais, A. G. 1972. Variation between southern and northern isolates of *Scirrhia acicola*. Phytopathology 62:768.

Kais, A. G. 1975. Environmental Factors Affecting Brown-Spot Infection on Longleaf Pine. Phytopathology 65:1389-1392.

Katoh, K., Misawa, K., Kuma, K., and Miyata, T. 2002. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. Nucleic Acids Res. 30: 3059–3066.

Katoh, K., and Standley, D. M. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol. Biol. Evol. 30:772-780.

Konečný, A., Estoup, A., Duplantier, J. M., Bryja, J., Bâ, K., Galan, M., Tatard, C., and Cosson, J. F. 2013. Invasion genetics of the introduced black rat (*Rattus rattus*) in Senegal, West Africa. Mol. Ecol. 22:286-300.

Laflamme, G., Côté, C., and Innes, L. 2010. White pine needle diseases in Eastern Canada. Phytopathology 101:S260.

Lévy, A. 1996. Le point sur le champignon de quarantine *Scirrhia acicola* dans le sud-ouest de la France. La Santé des Forêts: 28-30.

Lévy, A., and Lafaurie, C. 1994. Découverte de *Scirrhia acicola*. Phytoma - La défense des végétaux 463:33-35.

Librado, P., and Rozas, J. 2009. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. Bioinformatics 25:1451-1452.

Linde, C. C., Zala, M., Ceccarelli, S., and McDonald, B. A. 2003. Further evidence for sexual reproduction in *Rhynchosporium secalis* based on distribution and frequency of mating-type alleles. Fungal Genet. Biol. 40:115-125.

Martínez, J. B. 1942. Las micosis del *Pinus insignis* en Guipuzcoa.

McDonald, J. H. 2009. Handbook of Biological Statistics. Sparky House Publishing.

McDonald, B. A., and Linde, C. 2002. Pathogen population genetics, evolutionary potential, and durable resistance. Annu. Rev. Phytopathol. 40:349-379.

McVean, G. 2009. A genealogical interpretation of principal components analysis. PLoS Genet. 5:e1000686.

Milatović, I. 1976. Needle cast of pines caused by fungi *Scirrhia pini* Funk et Parker and *S. acicola* (Dearn.) Siggers in Yugoslavia. Poljopr. Znan. Smotra 39:511-513.

Milgroom, M. G. 1996. Recombination and the multilocus structure of fungal populations. Annu. Rev. Phytopathol. 34:457-477.

Munck, I. A., Burns, B., Lombard, K., Weimer, J., and Ostrofsky, W. D. 2012. What is causing white pine needle damage in northeastern North America? Phytopathology 102:S4.84.

Nei, M. 1978. Estimation of average heterozygosity and genetic distance from a small number of individuals. Genetics 89:583-590.

Nicholls, T. H., and Hudler, G. W. 1972. Red pine - a new host for brown spot (*Scirrhia acicola*). Plant Disease Reporter 56:712-713.

Peakall, R. and Smouse, P. E. 2006. GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. Mol. Ecol. Notes 6:288–295.

Peakall, R., and Smouse, P. E. 2012. GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research - an update. Bioinformatics 28:2537-2539.

Pehl, L., Cech, T. L., and Ioos, R. 2015. PM 7/46 (3) *Lecanosticta acicola* (formerly *Mycosphaerella dearnessii*), *Dothistroma septosporum* (formerly *Mycosphaerella pini*) and *Dothistroma pini*. EPPO Bulletin 45:163-182.

Pérez, G., Slippers, B., Wingfield, M. J., Wingfield, B. D., Carnegie, A. J., and Burgess, T. I. 2012. Cryptic species, native populations and biological invasions by a eucalypt forest

pathogen. Mol. Ecol. 21:4452-4471.

Pritchard, J. K., Stephens, M., and Donelly, P. 2000. Inference of Population Structure Using Multilocus Genotype Data. Genetics 155:945-959.

Quaedvlieg, W., Groenewald, J. Z., Yáñez-Morales, M De J., and Crous, P. W. 2012. DNA barcoding of *Mycosphaerella* species of quarantine importance to Europe. Persoonia 29:101-115.

Rambaut, A. and Drummond, A. 2007. Tracer v. 1.5. Computer program and documentation distributed by the authors at http://beast.bio.ed.ac.uk/Tracer.

Restrepo, S., Tabima, J. F., Mideros, M. F., Grünwald, N. J., and Matute, D. R. 2014. Speciation in fungal and oomycete plant pathogens. Annu. Rev. Phytopathol. 52:289-316.

Ronquist, F., and Huelsenbeck, J. P. 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19:1572-1574.

Rosenberg, N. A. 2004. DISTRUCT: a program for the graphical display of population structure. Mol. Ecol. Notes 4:137-138.

Sakalidis, M. L., Slippers, B., Wingfield, B. D., Hardy, G. E. S. J., Burgess, T. I., and Austin, J. 2013. The challenge of understanding the origin, pathways and extent of fungal invasions: global populations of the *Neofusicoccum parvum-N. ribis* species complex. Divers. Distrib. 19:873-883.

Saleh, D., Xu, P., Shen, Y., Li, C., Adreit, H., Milazzo, J., Ravigné, V., Bazin, É., Notteghem, 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.

Santini, A., Ghelardini, L., De Pace, C., Desprez-Loustau, M. L., Capretti, P., Chandelier, A., Cech, T., Chira, D., Diamandis, S., Gaitniekis, T., Hantula, J., Holdenrieder, O., Jankovský, L., Jung, T., Jurc, D., Kirisits, T., Kunca, A., Lygis, V., Malecka, M., Marçais, B., Schmitz, S., Schumacher, J., Solheim, H., Solla, A., Szabo, I., Tsopelas, P., Vannini,

A., Vettraino, A. M., Webber, J., Woodward, S., and Stenlid, J. 2013. Biogeographical patterns and determinants of invasion by forest pathogens in Europe. New Phytol. 197:238-250.

Schoebel, C. N., Jung, E., and Prospero, S. 2013. Development of new polymorphic microsatellite markers for three closely related plant-pathogenic *Phytophthora* species using 454-pyrosequencing and their potential applications. Phytopathology 103:1020-1027.

Siggers, P. V. 1944. The brown spot needle blight of pine seedlings (Technical Bulletin No. 870). Pages 1-36.

Sinclair, W., and Lyon, H. H. 2005. Diseases of Trees and Shrubs. 2nd ed. Cornell University Press.

Sommerhalder, R. J., McDonald, B. A., Mascher, F., and Zhan, J. 2010. Sexual recombinants make a significant contribution to epidemics caused by the wheat pathogen *Phaeosphaeria nodorum*. Phytopathology 100:855-862.

Stanosz, G. 1990. Premature needle drop and symptoms associated with brown spot needle blight on *Pinus strobus* in Northcentral Pennsylvania. Phytopathology 80:124.

Sunnåker, M., Busetto, A. G., Numminen, E., Corander, J., Foll, M., and Dessimoz, C. 2013. Approximate Bayesian computation. PLoS Comput. Biol. 9:e1002803.

Swofford, D.L. 2003. PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods). Version 4. Sinauer Associates, Sunderland, Massachusetts.

Tainter, F. H., and Baker, F. A. 1996. Brown Spot. Pages 467-492 in: Principles of Forest Pathology.

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., and Kumar, S. 2011. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol. Biol. Evol. 28:2731-2739.

Taylor, J. W. 1999. The Evolutionary Biology and Population Genetics Underlying Fungal Strain Typing. Clinical Microbiology Reviews 12:126–146.

Thümen, F. 1878. Fungorum Americanorum. Flora 61:177-184.

Walker, A. S., Gautier, A. L., Confais, J., Martinho, D., Viaud, M., Le, P. C. P., Dupont, J., and Fournier, E. 2011. *Botrytis pseudocinerea*, a new cryptic species causing gray mold in French vineyards in sympatry with *Botrytis cinerea*. Phytopathology 101:1433-1445.

Wingen, L. U., Shaw, M. W., and Brown, J. K. M. 2013. Long-distance dispersal and its influence on adaptation to host resistance in a heterogeneous landscape. Plant Pathol. 62:9-20.

Wingfield, M. J., Slippers, B., Roux, J., and Wingfield, B. 2011. Fifty Years of Tree Pest and Pathogen Invasions, Increasingly Threatening World Forests. Pages 89-99 in: Fifty Years of Invasion Ecology: The Legacy of Charles Elton. D. M. Richardson, ed. Blackwell Publishing Ltd.

Wingfield, M. J., Brockerhoff, E. G., Wingfield, B., and Slippers, B. 2015. Planted forest health: The need for a global strategy. Science 349:832-836.

Wolf, F. A., and Barbour, W. J. 1940. Brown spot needle disease of pines. Phytopathology 31:61-73.

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 Genet. Biol. 38:286-297.

888 **Table 1.** Information on *Lecanosticta acicola* isolates used in this study: number of isolates

889 per location, geographic origin of the isolates and host species.

Location number	No. of isolates	Country	Sampling site/State/Region	Pine host
1	18	Canada	Lake Pinseault / Quebec	*P. strobus*
		Canada	Lake Aberdeen / Quebec	*P. strobus*
		Canada	Fort William / Quebec	*P. strobus*
		Canada	Demers-Centre / Quebec	*P. strobus*
		Canada	Waltham / Quebec	*P. strobus*
2	4	Canada	Lake Drummond / Quebec	*P. strobus*
		Canada	Montréal / Quebec	*P. strobus*
3	23	USA	Washington, Waterbury / Vermont	*P. strobus*
		USA	Windsor, Bethel / Vermont	*P. strobus*
		USA	Orange, Brookfield / Vermont	*P. strobus*
4	3	USA	Merrimack / New Hampshire	*P. strobus*
		USA	Hillsboro / New Hampshire	*P. strobus*
		USA	Hopkinton-Everett / New Hampshire	*P. strobus*
5	10	USA	York, Lyman / Maine	*P. strobus*
6	6	USA	Androscoggin, Leeds / Maine	*P. strobus*
7	3	USA	Piscataquis, Sangerville / Maine	*P. strobus*
8	2	USA	Wexford County / Michigan	*P. sylvestris*
9	1	USA	Merrillan / Wisconsin	*P. sylvestris*
10	40	USA	Harrison County / Mississipi	*P. palustris*
		USA	Harrison County / Mississipi	*P. taeda*
11	5	Mexico	Galeana / Nuevo León	*P. arizonica* var. *stormiae*
		Mexico	Iturbide / Nuevo León	*P. halepensis*
12	1	Mexico	Piñal de los Amoles / Nuevo León	*Pinus* sp.
13	2	Guatemala	Santa Cruz Verapaz, near Tactíc	*P. oocarpa*
14	2	Colombia	Villanueva, Casanare	*P. caribaea*
15	1	Spain	San Sebastián de Garabandal / Cantabria	*P. radiata*
16	8	France	Pyrénées-Atlantiques	*P. radiata*
		France	Landes	*P. attenuata* x *radiata*
		France	Gironde	*P. muricata*
17	4	Switzerland	Zürich, Nordheim	*P. mugo*
		Switzerland	Zürich, Honggerberg	*P. mugo*
		Switzerland	Walensee	*P. mugo*
		Switzerland	Cham, Hammergut	*P. mugo*
18	1	Italy	Gardone / Brescia	*P. mugo*
19	2	Germany	Grassau	*P. mugo*
20	12	Germany	Murnau	*P. mugo*
		Germany	Murnauer Filze	*P. mugo*
		Germany	Pfrühlmoos	*P. mugo*
		Germany	Untersedlhof	*P. mugo*
21	1	Croatia	Zadar	*P. halapensis*
22	2	Slovenia	Bled	*P. mugo*
23	15	Austria	Gmunden	*P. nigra*
		Austria	Weyer	*P. mugo*
		Austria	Steyer, Pestalozzistraße	*P. mugo*
		Austria	Saimannslehen	*Pinus* sp.
		Austria	Sankt Gallen	*Pinus* sp.
		Austria	Hollenstein an der Ybbs	*P. mugo*
		Austria	Opponitz	*P. mugo*
		Austria	Sankt Gallen	*P. mugo*
		Austria	Waidehofen an der Ybbs	*P. mugo*
24	28	Czech Republic	Borkovická Blata	*P. uncinata* subsp. *ulliginosa*
		Czech Republic	Červená Blata	*P. uncinata* subsp. *ulliginosa*
25	2	Lithuania	Juodkrante for. distr. / Klaipėdský kraj	*P. mugo*
		Lithuania	Smiltyné forest distr. / Klaipėdský kraj	*P. mugo*
26	2	Estonia	Tallin / Harju maakond	*P. ponderosa*

890

Table 2. Population characteristics based on microsatellite data and mating type idiomorphs for *Lecanosticta acicola* isolates from North America and Europe.

ID	Continent	Country	region	N	MLHs	CF	MAT1-1/MAT1-2	L(PTLPT)	r_d	A_R	N_a	G	H_e
NA	North America north-east			67	43	0.358	61/6**			1.59±0.21	3.55	0.955	0.32±0.30
CAP	"	Canada (Québec)	Pontiac	18	13	0.278	17/1**	22^{ns}	-0.011^{ns}	2.02±0.71	2.54	0.954	0.32±0.24
CAM	"	"	Montréal	4	4		$3/1^{nd}$						
VMW	"	USA (Vermont)	Washington County	23	14	0.391	22/1**	18^{ns}	0.067*	1.83±0.59	2.55	0.944	0.26±0.21
NHM	"	USA (New Hampshire)	Merrimack County	3	1		$3/0^{nd}$						
MEY	"	USA (Maine	York	10	5		10/0**						
MEL	"	"	Leeds	6	4		$5/1^{ns}$						
MES	"	"	Sangerville	3	2		$1/2^{nd}$						
MS	North America	USA (Mississippi)	Harrison County	40	34	0.150	$22/18^{ns}$	146^{ns}	0.002^{ns}	7.1±0.89	8.81	0.992	0.46±0.32
SE	Europe south-west			9	5	0.444	$4/5^{ns}$			1.00±0.59	1.27	0.888	0.09±0.14
ESS	"	Spain	San Sebastián de Garabandal	1	1		0/1						
FSW	"	France	South-West regions	8	4		$4/4^{nd}$						
CE	Europe Central and North			69	48	0.289	23/44*			3.65±0.71	4.54	0.931	0.35±0.27
CHN	"	Switzerland	northern Cantons	4	2		$3/0^{nd}$						
ITG	"	Italy	Gardone	1	1		$0/1^{nd}$						
SLB	"	Slovenia	Bled	2	2		$0/2^{nd}$						
CRZ	"	Croatia	Zadar	1	1		$0/1^{nd}$						
GBU	"	Germany	Upper Bavaria	12	9	0.25	$6/5^{ns}$	12^{ns}	-0.132^{ns}	1.66±0.48	2.09	0.878	0.26±0.23
GBG	"	"	Grassau	2	2		$0/2^{nd}$						
ATN	"	Austria	northern regions	15	14	0.067	$8/7^{ns}$	23^{ns}	0.028^{ns}	2.19±0.35	2.72	0.942	0.30±0.25
CZB	"	Czech Republic	Southern Bohemia	28	13	0.536	3/25**	22**	0.076*	2.18±0.94	2.72	0.793	0.23±0.21
LTC	"	Lithuania	Curonian Spit	2	2		$2/0^{nd}$						
EET	"	Estonia	Tallin	2	2		$0/1^{nd}$						

N, number of isolates; MLHs, number of multilocus haplotypes; CF, clonal fraction; MAT1-1/MAT1-2 is the ratio of the mating type idiomorphs; the null hypothesis that the ratio is 1:1 was tested using exact binomial test; L(PTLPT) is the length of the observed tree in number of steps; *rBarD*, multilocus linkage disequilibrium; for the L(PTLPT) and *rBarD* tests the null hypothesis of random mating was tested; nd, not determined; ns, $P > 0.05$; *$P < 0.05$; **$P < 0.001$; N_a, number of alleles; A_R, allelic richness averaged across loci (±standard deviation); G, genotypic diversity; H_e, gene diversity (mean± standard deviation, Nei 1978).

902 **Table 3.** Genetic differentiation (F_{ST}) between pairs of genetically-defined clusters of
903 *Lecanosticta acicola* populations. Abbreviations of populations: CE, central and north-
904 European population; MS, Mississippi population; NA, north-east North-American
905 population; SE, south-west European population.

		MS	NA	CE
NA	0.481			
CE	0.471	0.286		
SE	0.351	0.662	0.642	

906

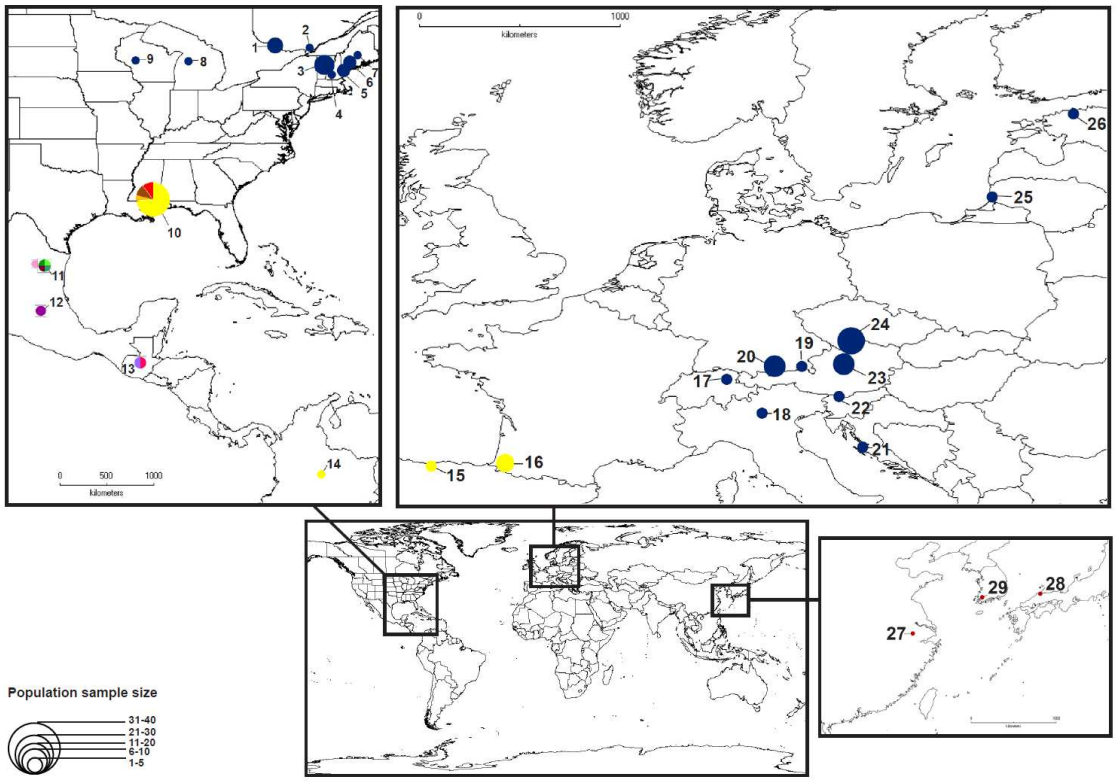
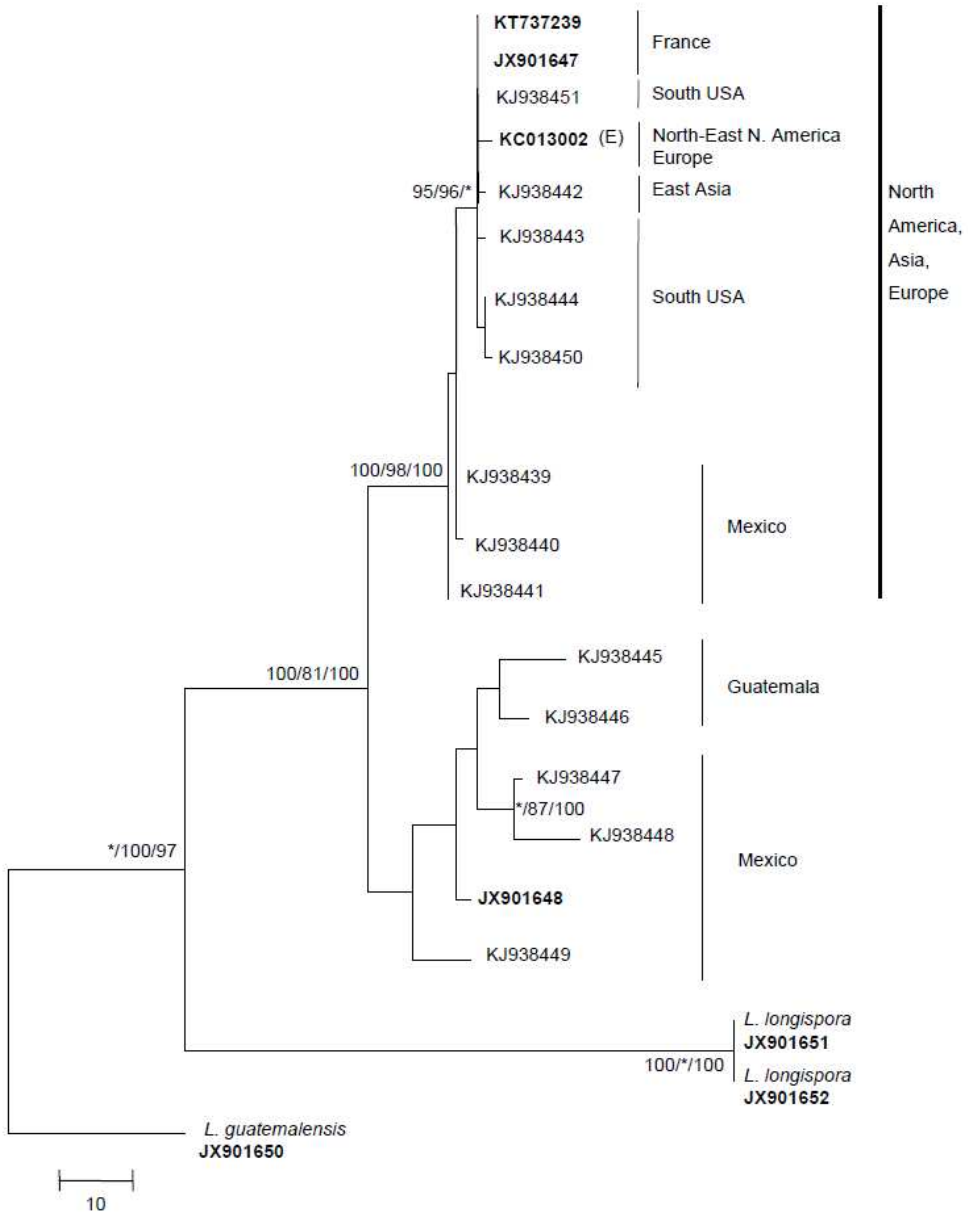


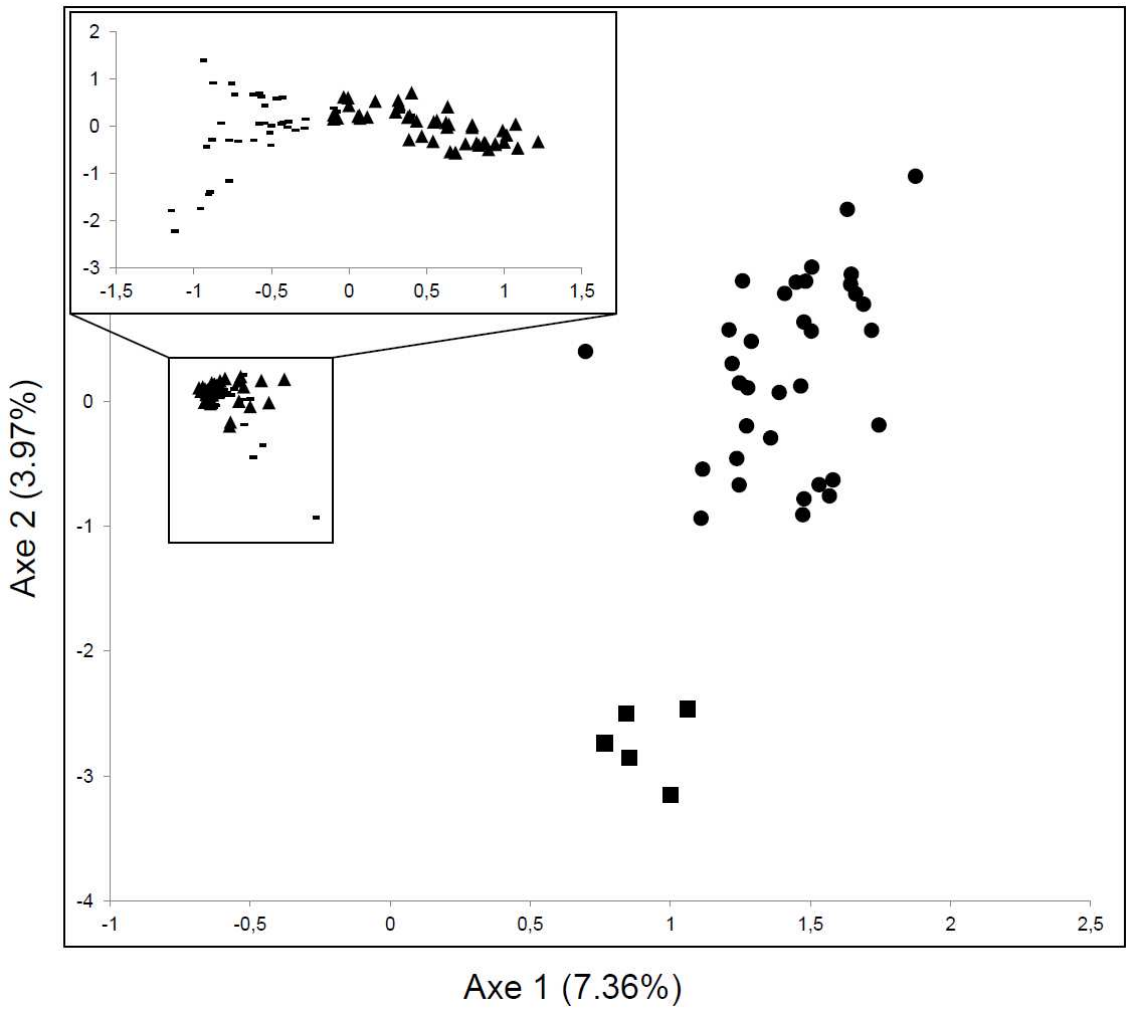
Figure 1. Geographic representation of 14 EF haplotypes of *Lecanosticta acicola* in America, Europe and East Asia (enlarged areas). Numbers are the codes of locations corresponding to Table 1; each colour represents one haplotype (colours correspond to Fig. S1); size of circles represents number of isolates used in this study; scale bar indicate 500 km and 1000 km distance in each area.



915

916

917 **Figure 2.** Phylogram based on maximum parsimony analyses of EF sequences. The
918 representative isolates from GenBank are indicated in bold and the isolate indicated by "(E)"
919 represents the epitype of *L. acicola*. Maximum parsimony (MP), maximum likelihood (ML)
920 and Bayesian inference (BI) bootstrap support values (1000 replicates) are indicated at the
921 nodes (support values: MP > 75%; ML > 75%; BI > 95%; * - not significant). Scale bar
922 indicates 10 nucleotide mutations.



923

924 **Figure 3.** Results of PCA analysis of the clone-corrected data of *L. acicola* isolates
925 represented on two principal component axes. The isolates originating from Mississippi are
926 designated with a dot, isolates from south-west Europe by a square, isolates from north-
927 eastern N. America ('NA' population) by a triangle and isolates representing 'CE' population
928 (central, northern and south-eastern Europe) are indicated with a dash. Results of separate
929 PCA analysis of NA and CE (northern lineage) populations are presented in the nested
930 rectangle (axe 1: 7.11%; axe 2: 5.60%).

931

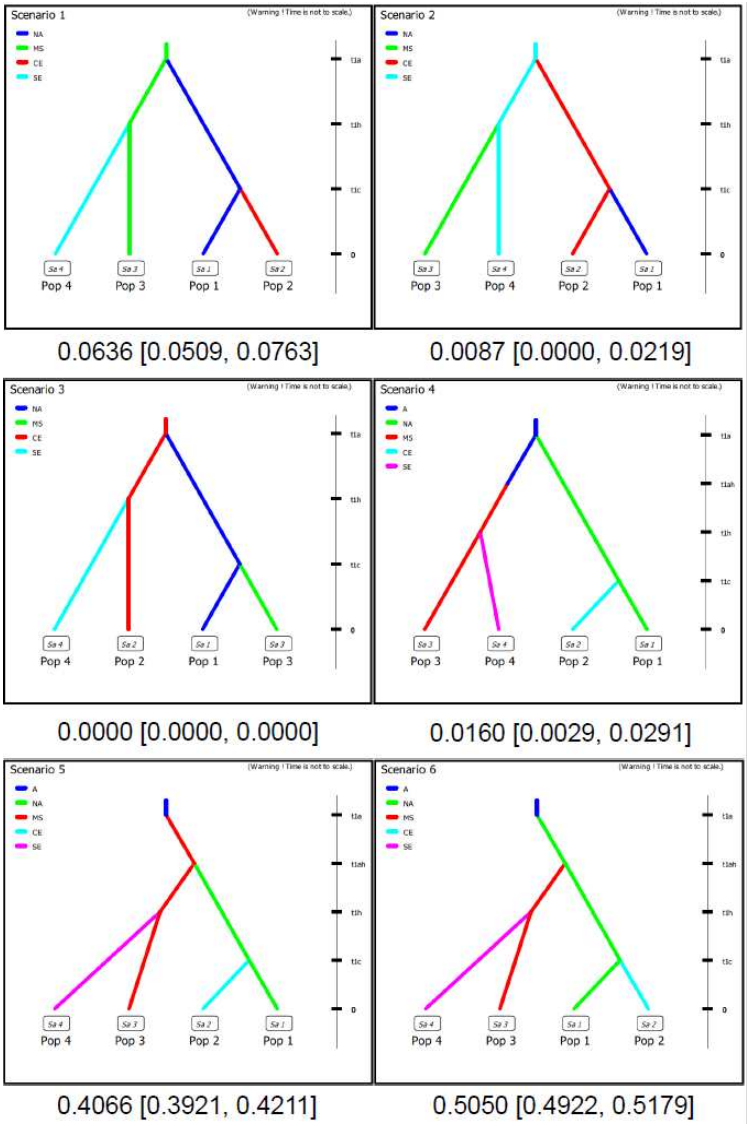


Figure 4. Graphical representation of six scenarios of the demographic history and their relative posterior probabilities. Abbreviations used on time scales refer to time parameters used during simulations (description of each parameter is provided in Table S4). Abbreviations of populations: A, assumed ancestral population; CE, central and north-European population; MS, Mississippi population; NA, north-east North-American population; SE, south-west European population.

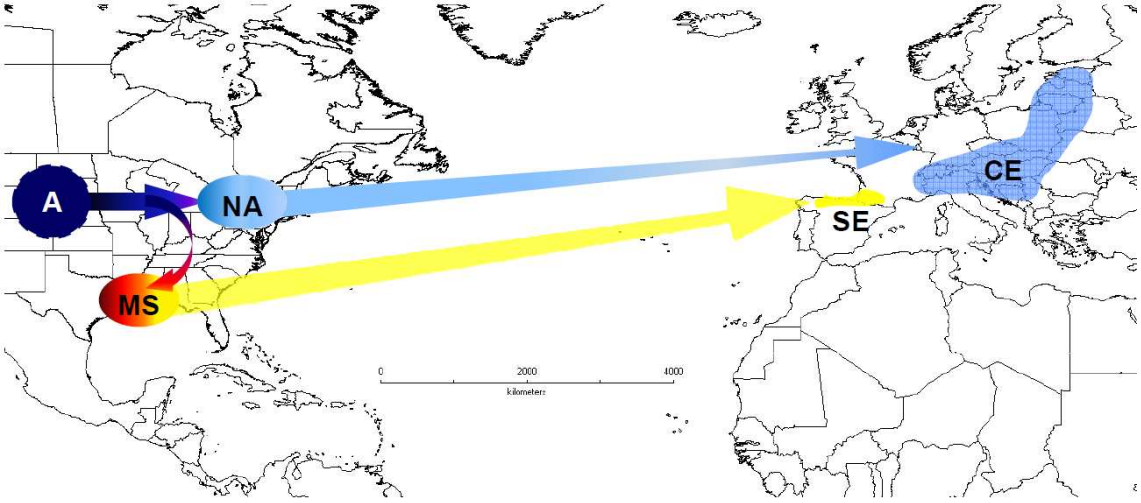


Figure 5. Graphical representation of the most supported evolutionary scenario of *Lecanosticta acicola* invasion from North America to Europe. Ancestral population (A) gave origin of the blue lineage that gave arise MS population and parts of the two populations were introduced from North America to Europe independently. Abbreviations of populations: A, assumed ancestral population (its geographic location is unknown and does not correspond to its placement on this map); CE, central and north-European population; MS, Mississippi population; NA, north-east North-American population; SE, south-west European population.

948 **Supplementary Tables**

949 **Table S1.** Details for the *Lecanosticta acicola* isolates (N=201) obtained from needles of
950 *Pinus* spp., used in this study including culture collection and GenBank accession numbers,
951 mating type idiomorph, geographical origin and coordinates, description of site/stand type,
952 host species, date of collection, collector/supplier and, where available, the altitude of the
953 collection site.

954 **Table S2.** Summary table of microsatellite genotyping results. Each isolate is characterised by
955 the multilocus haplotype generated from eleven microsatellite markers and represented as the
956 length of the amplified fragment. No amplification is indicated by a star.

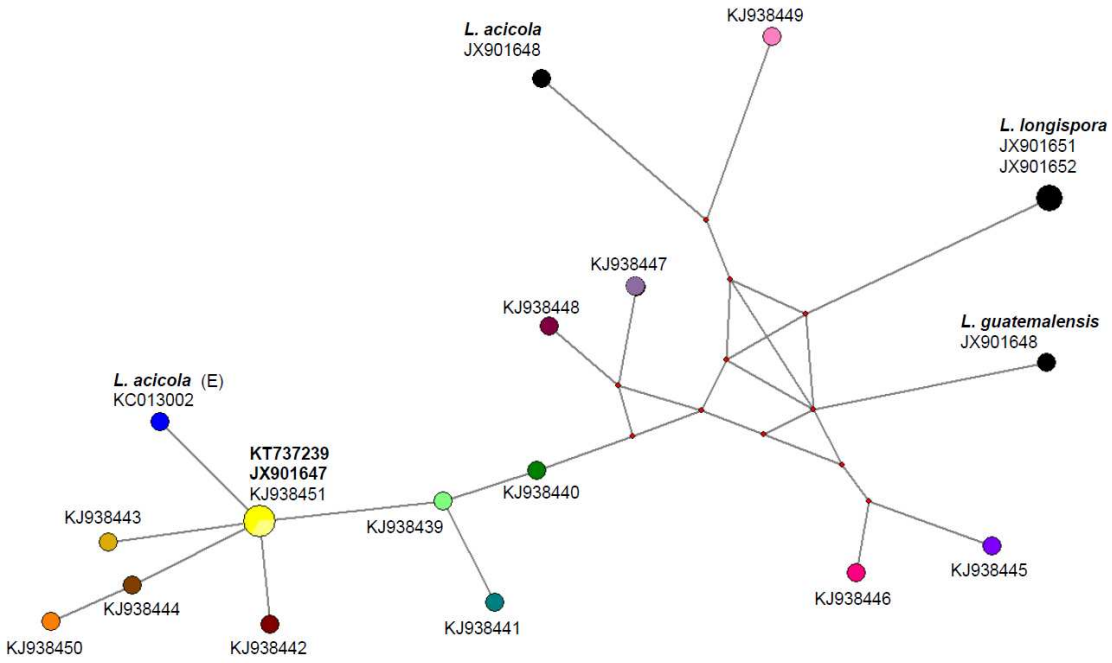
957 **Table S3.** Hierarchical analyses of molecular variance (AMOVA) of European isolates of *L.*
958 *acicola* according to (1.) host species and (2.) geographic origin.

959 **Table S4.** Prior and posterior distributions of demographic, historic and mutation parameters
960 estimated and used in the ABC analyses.

961 **Table S5.** Model checking using 26 summary statistics not used for the previous ABC model
962 selection in Table S4.

963

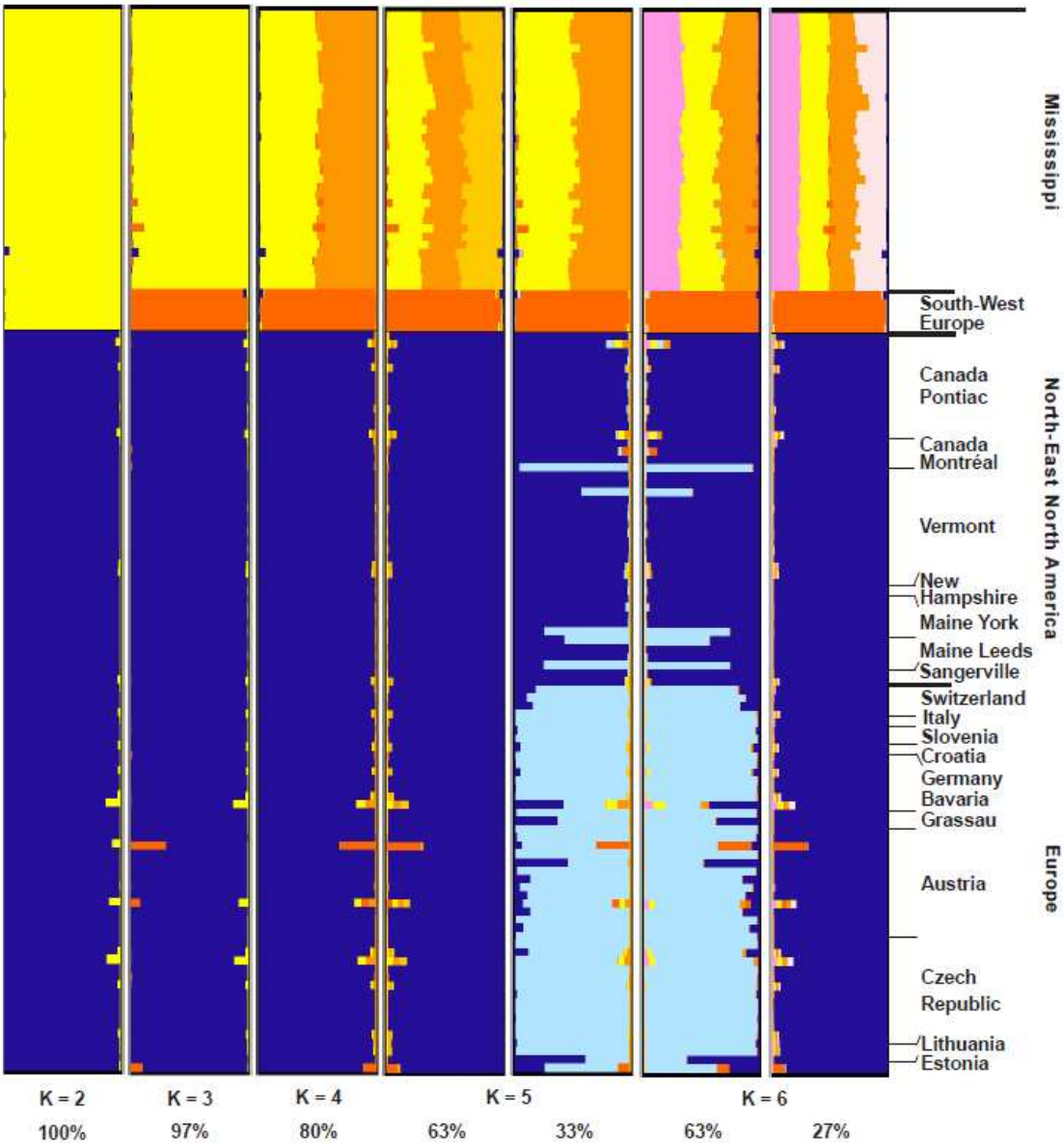
964 **Supplementary Figures**



965

966 **Figure S1.** Median-joining haplotype network constructed based on partial EF gene sequence
967 data of *L. acicola* isolates. Each haplotype, presented as a node, is coloured according to Fig.
968 1. The isolate indicated by "(E)" represents the epitype of *L. acicola*. The representative
969 isolates from GenBank are indicated in bold.

970



971

972 **Figure S2.** STRUCTURE results for K=2-6 presented as summary bar plots of 129

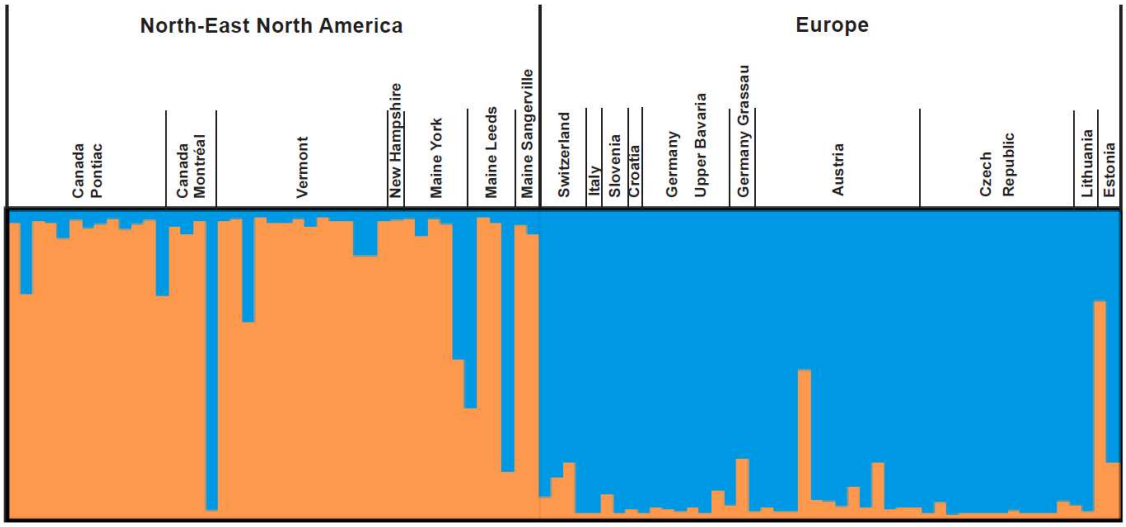
973 *Lecanosticta acicola* isolates (clone-corrected data). Each isolate is represented by a single

974 horizontal line divided into K clusters (different colours). The percentage indicates the

975 proportion of independent STRUCTURE runs (from a total of thirty) that correspond to the

976 significantly similar clustering pattern (SSC > 0.9).

977



978

979 **Figure S3.** Refined results of STRUCTURE analysis for K=2 of the northern cluster observed

980 in Fig. S2. 90 *L. acicola* isolates are presented as bar plots (clone-corrected data). Each isolate

981 is represented by one vertical line.

982

983

984