Mycol. Res. **108** (7): 766–774 (July 2004). © The British Mycological Society
DOI: 10.1017/S0953756204000334 *Printed in the United Kingdom.*

North American populations of *Entoleuca mammata* are genetically more variable than populations in Europe

Risto KASANEN[1]*, Jarkko HANTULA[1], Michael OSTRY[2], Jean PINON[3] and Timo KURKELA[1]

[1] *Finnish Forest Research Institute, Vantaa Research Station, P.O. Box 18, FIN-01301 Vantaa, Finland.*
[2] *North Central Research Station, 1561 Lindig Avenue, St Paul, MN 55108, USA.*
[3] *UR Pathologie forestière INRA, F-54280 Champenoux, France.*
E-mail: risto.kasanen@metla.fi

Received 20 December 2003; accepted 13 April 2004.

Entoleuca mammata (syn. *Hypoxylon mammatum*) is a damaging pathogen of *Populus tremuloides* and *P. grandidentata* in North America and *P. tremula* in Europe, where the fungus occurs only sporadically in alpine regions and Scandinavia. It has been hypothesized that *E. mammata* was introduced to Europe from North America. In this study, *E. mammata* isolates collected from Europe and North America were compared by a sequence analysis of two DNA markers derived from DNA fingerprints. The objective of the study was to elucidate the relationship between North American and European *E. mammata* populations by testing two hypotheses: (1) North American and European isolates are conspecific; and (2) the fungus was introduced between continents causing both a founder effect and a genetic bottleneck. North American populations were found to be more polymorphic, but no major phylogenetic differences between fungal isolates collected from different continents were found. This result combined with the historical observations of the disease in Europe implies that *E. mammata* was introduced to Europe several centuries ago.

INTRODUCTION

Entoleuca mammata (syn. *Hypoxylon mammatum* is a damaging pathogen of *Populus tremuloides* and to a lesser extent *P. grandidentata* in North America. In Europe, *Populus alba*, *P. trichocarpa*, *P. tremula* and hybrid aspen (*P. tremula* × *P. tremuloides*) are susceptible to the pathogen (Pinon 1979, Terrason, Pinon & Ride 1988, Mazzaglia *et al.* 2001). The taxonomy of *Hypoxylon* has been subject to several rearrangements. Miller (1961) suggested that several species described in the nineteenth century, were synonyms. These included *H. mammiformis*, *H. pauperatum* and *H. macrosporum* which were collected in northern Scandinavia (Lapland). Miller (1961) gave a relatively broad definition of the genus, while Rogers & Ju (1996) defined a stricter structure of *Hypoxylon*. Based on their definition of *Hypoxylon*, *H. mammatum* was moved to the genus *Entoleuca*, the type species of which is *E. callimorpha* (Rogers & Ju 1996). Later Mazzaglia *et al.* (2001) evaluated genetic relationships among *Hypoxylon* and related genera using 5.8S and ITS sequences of rDNA, but since they did not include *E. callimorpha* in the analysis, the transfer is not yet confirmed by molecular evidence.

In North America, the disease is prevalent throughout the north-east, the Great Lakes region, and northwestern prairies. Incidence is fairly low in the central Rocky Mountains, but the disease is noticeably absent in the northern Rocky Mountains and in Alaska despite the abundance of aspen and collections of the fungus (Manion & Griffin 1986). In the Lake States where *P. tremuloides* is the most important pulpwood species *E. mammata* is considered the major cause of aspen mortality.

Recently, large-scale plantations of aspen and hybrid aspen have been established in Nordic and Baltic countries as the forest industry has become interested in aspen fibre for pulp production. As the cultivation of native and hybrid aspen increases, fungal diseases of aspen may gain in both economic and ecological importance. Potentially, *E. mammata* could hinder the development of aspen cultivation. Serious damage in hybrid aspen trials caused by *E. mammata* has been reported in Scandinavia (Illstedt & Gullberg 1993). However, the report did not describe the basis for the identification of the pathogen, so the exact cause of the damage is uncertain.

* Corresponding author.

E. mammata was thought to be restricted to North America, even though European *P. alba* and *P. tremula* were known to be susceptible (Manion & Griffin 1986). The disease was first observed in France in 1975, while the fungus was still a quarantine pest in Europe. At the time, it was thought that the pathogen might have been introduced to Europe in the 1950s on infected flower branches of *P. tremuloides* and *P. grandidentata* used in hybrid aspen breeding (Pinon 1979). However, the geographical distribution of the fungus in France makes it unlikely that the fungus was a recent introduction. The disease was present at several isolated locations in the French Alps where *P. tremula* had been collected for breeding purposes, but where no potentially infected plant material had been introduced. In France, the fungus is widely distributed and damaging in mountain stands, but it is relatively rare and localized in the plains. The distribution of the fungus may depend on suitable hosts. In France two races of *P. tremula* exist, one in mountain regions and the other in the plains. The mountain race is found in Jura, in the Alps (350–1850 m a.s.l.) and in the Pyrénées.

Several European *E. mammata* herbarium specimens examined by Miller (1961) had been collected in the nineteenth century, confirming that *E. mammata* has been present in Europe longer than previously believed. Few occurrences of *E. mammata* are represented in Fennoscandian herbaria (Miller 1961), and the fungus apparently occurs only sporadically in aspen stands. Although the fungus can cause serious mortality of aspen in individual stands, no serious outbreaks have been recorded in Europe. This could be interpreted as a balance between host and pathogen that is usually considered a result of coexistence and coevolution. Based on the information of historical abundance and geographic range of the fungus in Europe, *E. mammata* could be considered a native species. However, the relationship of European and North American populations of the fungus has not been examined, and molecular studies of the population structure of this worldwide pathogen are lacking. The history of *E. mammata* in Europe is consequently poorly known, and open to several interpretations.

The objective of this investigation was to elucidate the relationship between North American, central European and Finnish isolates of *E. mammata* by studying the genetic structure of populations to test the following hypotheses: (1) North American and European isolates are conspecific; and (2) the fungus was introduced between continents causing both a founder effect and a genetic bottleneck. According to hypothesis (1), identical alleles should be found in both continents and no phylogenetic grouping of alleles based on the geographic origin of the isolates should be detected. Hypothesis (2) was tested by measuring genetic differentiation between populations and by analysing genetic variability between continents.

MATERIALS AND METHODS

Fungal isolates

Entoleuca mammata samples were collected from Europe (EU) from 20 locations in France, Finland, Switzerland and Italy (Table 1). The host species included one *Salix*, two *Populus* species, and a *Populus* hybrid. The collections were made during 1975–2002. The French isolates were collected from several locations accross the known distributional area of *E. mammata*. The Finnish populations were collected from two natural aspen stands in Eastern Finland. The size of the sampling area was approx. 5 ha in both stands, which were separated by 10 km. In North America (NA), two populations located in Minnesota were collected in 2001 and 2002. The 2001 population was from along a trail set in natural forest over three miles, and in 2002 the sampling was performed throughout a hybrid aspen plantation several acres in size. The NA populations were geographically separated by over 100 miles. In total, 63 samples were subjected to analysis. Most of the isolations of fungal hyphae were made from the canker margin. Pieces of wood (approx. $10 \times 5 \times 2$ mm) were surface-sterilized by rinsing in 4% NaOCl, 75% ethanol and finally sterile water (10 s each) and incubated on 1.5% malt agar. As fungal growth occurred, the isolates were sub-cultured on 1.5% malt agar. In a few cases, single-ascospore isolations were made by casting spores from moistened perithecia onto 1.5% malt agar (Table 1).

Molecular markers

DNA isolations were made as described by Vainio, Korhonen & Hantula (1998). The fungal tissue was homogenized in 1.5 ml disposable plastic tubes by grinding with a glass rod and quartz sand. Briefly, the DNA isolation procedure included two phenol-chloroform (1:1) extractions, chloroform-phenol-isoamyl alcohol extractions, precipitation with polyethylene glycol (PEG) and drying. The DNA was resuspended into 10 mM Tris HCl buffer containing 1 mM EDTA.

Random amplified microsatellite markers $DBV(CAT)_5$ and $HBH(GAG)_5$ (where B = C/G/T, D = A/G/T, H = A/C/T and V = A/C/G) (RAMS; Hantula *et al.* 1996) were screened in order to find single loci that were amplifiable from a subset of European samples. In the RAMS method, the primers are attached to distal ends of microsatellites, a procedure also utilized in ISSR-PCR (inter short-sequence repeat PCR) (Zhou *et al.* 1999). The PCR protocol by Hantula, Dusabenyagasani & Hamelin (1996) was used. Next, two RAMS markers signed as *H*gag (819 bp) and *H*cat (327 bp) were cloned with TOPO TA cloning kit (Invitrogen, Carlsbad, CA). The sequences were determined with a Li-Cor IR2 DNA sequencer, using M13 universal forward and reverse primers. Based on the sequences obtained, two pairs of primers were designed to amplify fragments. The primer sequences

Table 1. Origin of *Entoleuca mammata* isolates analysed. All isolates are cultured *in vitro* and deposited in the Finnish Forest Research Institute (Vantaa).

Isolate number	Host	Locality	Year of isolation	Isolation
Population samples from Finland				
H (5 isolates)	*Populus tremula*	Punkaharju	2000	CM[b]
K (6 isolates)	*P. tremula*	Punkaharju	2000	CM
Population samples from North America				
US (14 isolates)	*P. tremuloides*	Minnesota	2001	CM
U2 (13 isolates)	*P. tremuloides*	Minnesota	2002	CM
Isolates from France				
JP22	*P. tremula*	Chorges (Hautes-Alpes)	nk[a]	CM
JP24	*P. tremula*	Alps (Les Guibertes)	1975 or 76	
JP26	*P. tremula*	Alps (Les Guibertes)	1975 or 76	CM
JP28	*P. tremula* × *tremuloides*	Velaine-sous-Amance (Lorraine)	1975	CM
JP37	*P. tremula*	Alps (Le Revard)	1976	CM
JP38	*P. tremula*	Alps (Le Revard)	1976	CM
JP39	*P. tremula*	Alps (Albion)	1976	CM
JP41	*P. tremula*	Alps (Revest-du-Bion)	1977	CM
JP43	*P. tremula*	Alps (Couloubron)	1977	CM
JP44	*P. tremula*	Jura	1977	CM
JP84	*P. tremula*	Lorraine (Ludres)	1978	monospore
JP86	*P. tremula*	Lorraine (Ludres)	1978	monospore
JP87	*P. tremula*	Lorraine (Ludres)	1978	monospore
JP88	*P. tremula*	Lorraine (Ludres)	1978	monospore
JP164	*P. trichocarpa*	Orléans	Around 1978	CM
JP178	*P. tremula*	nk	nk	monospore
JP180	*P. tremula*	nk	nk	monospore
JP182	*P. trichocarpa*	West (Nantes)	1985	CM
JP187	*P. trichocarpa*	West (Nantes)	1985	CM
JP192	*P. trichocarpa* '*Fritzi Pauley*'	(Centre) Nogent-sur-Vernisson	1986	CM
JP193	*P. trichocarpa* '*Fritzi Pauley*'	(Centre) Nogent-sur-Vernisson	1986	CM
Isolates from other countries				
JP21	*P. tremula*	Italy (around Cuneo)	1975	CM
JP181	*P. tremula*	Italy (around Cuneo)	1985	CM
JP323	*Salix* sp.	Switzerland	1991	CM
JP271	*P. tremula*	Andorra, Fontenada	1992	CM

[a] nk, not known. [b] CM, canker margin.

were: *H*agF; 5′ gga gat gaa cag cga gga ct, *H*agR; 5′ ggc gaa act gag gag agt ca, *H*catF; 5′ aaa tcc act ccc aaa tgc aa and *H*agR; 5′ ccc tac ctc tag gcc acc tc. These fragments were amplified from all isolates, purified with Hi-Pure PCR Product Purification Kit (Roche, Mannheim) and sequenced with the primers used for amplification. The sequencing was conducted with a LI-COR IR2 automated sequencer according to the manufacturer's recommendations except that annealing temperature of 61 °C was applied.

Testing the hypothesis of conspecificity

To support the hypothesis of conspecificity (1) of the fungus in North America (NA) and Europe (EU), identical alleles should be found from both continents and no phylogenetic grouping of alleles based on geographic origin of isolates should be detected. The alleles were determined by sequencing. To study phylogenetic structures within sequence data, a distance matrix was computed from nucleotide sequences using Kimura's 2-parameter genetic distance model (Kimura 1980). The UPGMA algorithm was used to calculate phylogeny from the distance matrix. Prior to analysis the data was bootstrapped with SEQBOOT program within the PHYLIP software package. The calculations were made with DNADIST and NEIGHBOR programs. Consensus trees were constructed by CONSENSE program (Felsenstein 1989). For the outgroup we chose the sample JP323 collected from *Salix* sp. in Switzerland because in preliminary analyses the sample was shown to form a distant and separate branch in phylogenies. Phenograms were constructed with the TREEDRAW program (Page 1996).

Testing the introduction hypothesis

The introduction hypothesis (2) was tested by measuring genetic differentiation of the fungal populations between continents and analysing genetic variability within populations. In order to measure genetic differentiation caused by a hypothetical founder effect between North American and European populations, we calculated the F_{st} analogue defined by Weir & Cockerham (1984) and G_{st} value (Nei 1987). All calculations were made with FSTAT software (Goudet 2000). As FSTAT has been originally designed for diploid data sets, haploid fungal genotypes were encoded as diploid homozygotes, as suggested by Goudet (2000).

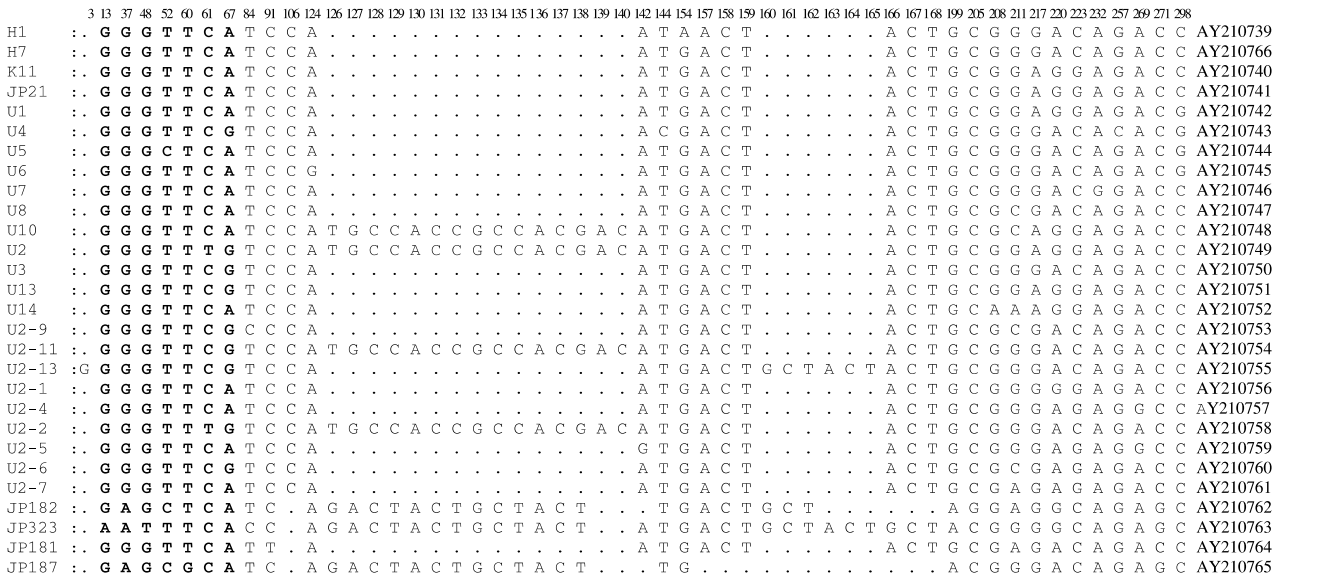


Fig. 1. Polymorphic sites of alleles of marker locus *H*cat (372 bp). Non-polymorphic sites are deleted from the sequences. Complete sequences can be obtained from GenBank (http://www.ncbi.nlm.nih.gov.; accession numbers are given in the end of each line). The nucleotides used for artificial loci are in bold.

If *Entoleuca mammata* was moved between continents, the result would be a genetic bottleneck that could be seen as differences in genetic variability within North American and European populations, especially if only few introduction events had taken place or a limited number of individuals were introduced. To address this question, we calculated Nei's gene diversity index h given by the equation $h = 1 - \sum x_i^2$, where x is the frequency of the i-th allele (Nei 1973). Since the sampling effort differed between NA and EU samples, we used the rarefaction method (Simberloff 1978) in estimating the number of haplotypes (E_{gn}) and alleles per locus in two regions. In rarefaction analysis the data was divided into three groups, North America, Europe and Finland. The Chao-estimator (Chao 1987) was calculated for estimating haplotype and allele richness. The calculations were made with Rarefaction Calculator (http://www2.biology.ualberta.ca/jbrzusto/rarefact.php#Calculator). Recently, the rarefaction method has been shown to be applicable for estimating genotype diversity of populations of microorganisms (Grünwald *et al*. 2003).

Testing the effect of hypervariation

In order to investigate whether the high number of observeable alleles caused bias to our interpretation of population differentiation, we artificially decreased the variation and thus the number of alleles per locus. This was done by taking only few polymorphic bases from the sequence data in calculations. From the *H*gag locus we analysed only five polymorphic nucleotides (Fig. 1) and from the *H*cat locus 7 polymorphic nucleotides (Fig. 2); the loci were named *gag5* and *cat7*. In addition, another pair of artificial loci was created by scoring their length (number of bases). Thus, we analysed four artificial loci in which only part of the variation was used.

RESULTS

Conspecificity hypothesis

Within the *H*cat locus we identified 28 alleles (Fig. 1). In most cases allele polymorphisms were due to single-base-pair point mutations or deletions, but deletions or insertions of six and 12 base-pair elements were also revealed (Fig. 1; note only polymorphic sites are shown). In North American (NA) populations 21 alleles were observed, of which 19 occurred in NA samples only. In European (EU) populations seven out of nine alleles were specific to EU samples. Two alleles were shared with NA and EU samples (Table 2). *H*gag locus had 22 different alleles (Fig. 2), 14 of which were specific to NA samples. European samples had only 5 specific alleles. Three alleles were identical between NA and EU samples in this locus (Table 2). Also in *H*gag locus most polymorphisms were due to single-base-pair mutations but in isolate JP21 an insertion of 590 bp was revealed (Fig. 2). All NA samples had an individual haplotype based on the two markers, but in Europe up to eight isolates shared the same haplotype. The phylogenetic analysis revealed no phylogenetic differences between fungal isolates collected from different continents. In the phenograms, NA and EU isolates were evenly distributed on branches, and no branches with major divergence from phenograms occurred. Also, the bootstrap values for nodes of phylogenies were almost exclusively very low (Figs 3–4).

```
          38 40 88 96 100 106 119 123 160 180 189 191 192 201 203 204 208 216 217 218 223 227 232 233 235 236 240 242 248 249 250 253 254 916
U7    : C G A A T - G A C C A A T T C C C T G G C T C C A G C - T - C - - T  AY210717
U14   : C G A G T - G A C G A A T C C C C T G G C T C C A A C - T - C - - T  AY210718
U9    : T G A A T - C A C G A A T T T C C T G G C T C C A G C - T - C T - T  AY210719
U13   : C G A A T - G A C G A A T T C C C T G G C T C T A A C - T - C - - T  AY210720
K6    : C G A A T - G A C C G A T T C C T T G G T T C C A G C - T - C T - T  AY210721
U5    : C G A G T - A A C G A A T T C C C T G G C T C C A A C - T - C - - T  AY210722
U3    : C G A G T - G A C C G A T T C C T T G G C T C C A G C - T - C T - T  AY210723
H1    : C G A G T - G A C C A T T T C C T T G G C T C C A G C - T - C T - T  AY210724
U1    : C G A A C - G A C C G A T T C C T T G G C T C C A G C - C - C T - T  AY210725
U2-4  : C G A A T - G T T C G A T T C C T T G G C T C C A G C - T - T T - T  AY210726
U2-1  : C G A G T - G A C C G A T T C C T C G G C T C C A G C - T - C T - T  AY210727
U2-11 : C G C A T - G A C G A A T T C C C T G G C T C C A A C - T - C - - T  AY210728
U2-12 : C A A A T - G A C C G A T T C C T T G G C T C C A G C - T - C T - T  AY210729
U2-7  : C G A A T G G A C G A A T T C C T T G G C T C C A G C C T T C - - T  AY210730
U2-9  : C G A A T - G A C C G A T T C C T - - - C T C C A A C - T - C - - T  AY210731
U2-8  : C G A A T - G A C C A A T T C C T T G G C T C C A G C - T - C - - T  AY210732
JP87  : C G A A T - G A C C G A T T C C T T G G C T C C A G C - T - C T - G  AY210733
JP44  : C G A A T - G A C C G A T T C C T T G G C T C C A G C - T - C T - T  AY210734
JP187 : C G A A T - G A C C A A C T C C T T G G T T T C A G C - T - C - - T  AY210735
JP323 : C G A A T - G A C C A A C T C G T T A G T G C C A G C - T - C - - T  AY210736
U8    : C G A G T - G A C C A A T T C C T T G G C T C C C G C - T - C - - T  AY210737
JP21  : C G A A T - G A C C G A T T C C T T G G C T C C A G T - T - C T * T  AY210738
```

Fig. 2. Polymorphic sites of alleles of marker locus *H*gag (819) bp. Non-polymorphic sites are deleted from the sequences. Complete sequences can be obtained from GenBank (http://www.ncbi.nlm.nih.gov.; accession numbers are given in the end of each line). The nucleotides used for artificial loci are in bold. The sample JP21 contained a 590 bp insert which is marked as *.

Table 2. Distribution of alleles in marker loci.

Marker	*n*	Number of alleles	Number of alleles occurring in USA/number of isolates	Number of alleles occurring in Europe/ number of isolates	Number of common alleles
*H*cat	63	28	21/27	9/36	2
*H*gag	63	22	17/27	8/36	3

Introduction hypothesis

Allelic richness defined by Nei's diversity was considerably higher in NA populations than in EU populations (Table 3). Rarefaction curves and the Chao-estimator both suggest that the allelic and haplotype richness is higher in North America than in Europe (Fig. 5, Table 3).

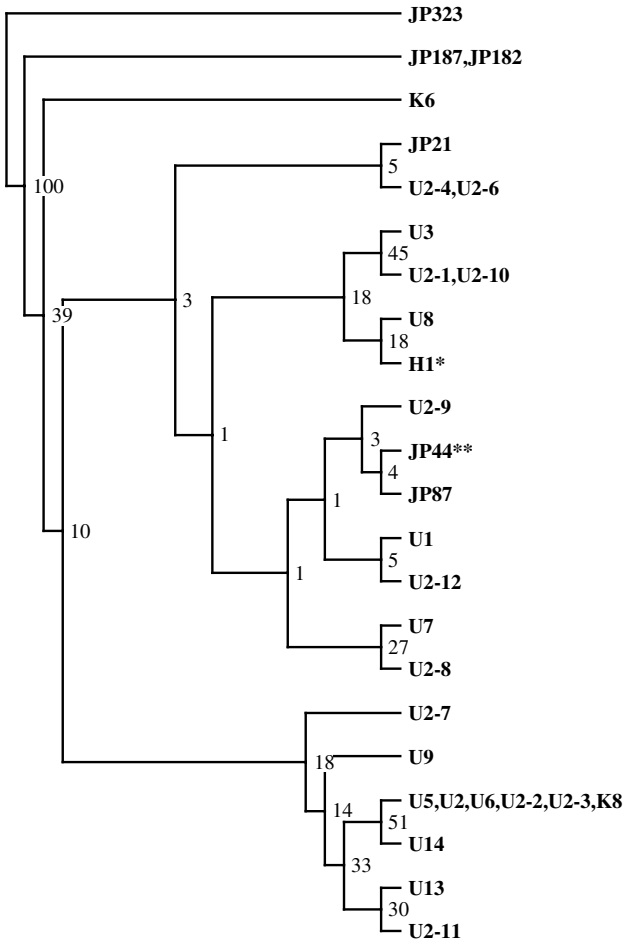
The G_{st} and F_{st} values calculated over *H*gag and *H*cat loci were 0.160 and 0.193, respectively (Table 4). Thus the analysis implies that there is population differentiation between NA and EU samples. Compared to the allele numbers at these loci (28 and 22) the allele numbers in artificial loci were lower; the number of alleles in gag5, cat7, gagL and catL were six, eight, five and seven, respectively. The G_{st} and F_{st} values calculated over the four artificial loci were 0.150 and 0.187, respectively, which are almost identical with those calculated for *H*gag and *H*cat loci (Table 4). Higher G_{st} and F_{st} values were obtained from artificial loci based on point mutations than from those based only on the length. This was obviously due to identical length of alleles with considerably different sequences (homoplasy). This indicates that the high number of alleles is not the source of the high F_{st} and G_{st} values obtained for *H*gag and *H*cat.

DISCUSSION

This is the first molecular analysis of the genetic polymorphism in *Entoleuca mammata* populations. The number of alleles observed in the two loci was relatively high, 22 and 28 alleles per locus. The resolution of the marker system enabled us to identify a unique genotype for all NA samples ($n=27$) with the two markers. A clear methodological conclusion can be drawn: sequence analysis of individual loci obtained originally from genetic fingerprinting is a powerful method of analysing variation. A methodological benefit over genetic fingerprinting systems such as RAMS (Hantula *et al.* 1996) and RAPD (Williams *et al.* 1990) is that the homology of the loci studied between samples is simultaneously verified in sequencing. This is especially valuable if sampling and fungal isolation covers a wide geographical area (different continents) and the fungus has a highly variable *in vitro* morphology such as *E. mammata* (Bagga & Smalley 1974).

The hypothesis of conspecificity of North American and European *E. mammata* was supported by the observation of several identical alleles in the two loci on both continents. The conspecificity was also supported by the lack of phylogenetic grouping of alleles according to their origin, and the absence of significant nodes in the phylogenies.

The second question addressed in this investigation was the direction of introduction. Theoretically,

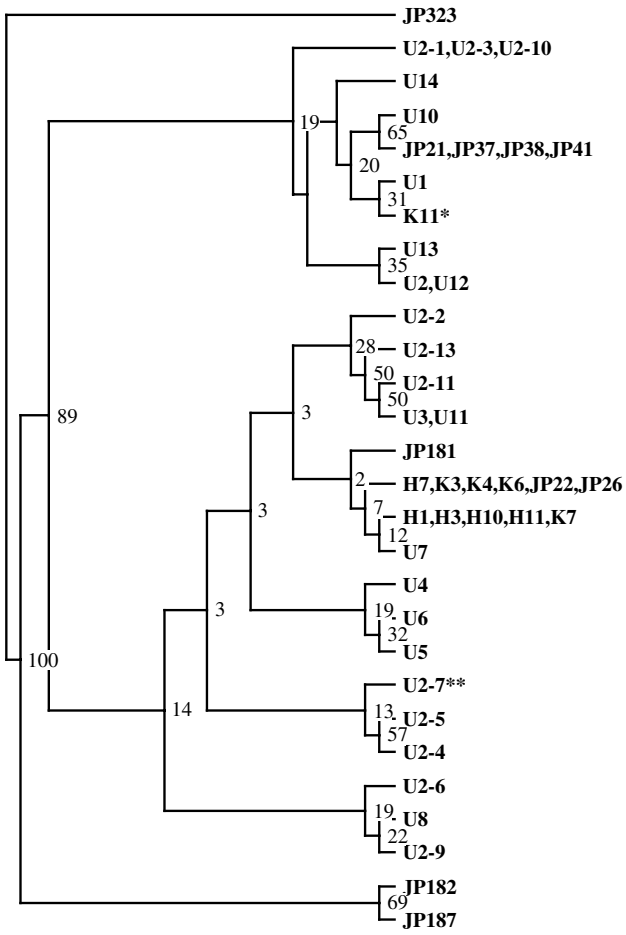


H and K are Finnish isolates; U, North American isolates; JP, European isolates.

* Isolate JP44 was identical to isolates JP22, JP271, JP28, JP39, JP41, JP26, JP37, JP38, JP43, JP84, JP86, JP88, JP329, JP178, JP180, JP192, JP28, JP193, U2-13 and U11.

** Isolate H1 was identical to H3, H7, H10, H11, K3, K4, K7, U4, U10, U12, U2-5, JP24 and JP164.

Fig. 3. Unrooted UPGMA consensus tree of *Entoleuca mammata* isolates. The sequence data from *H*gag marker locus was bootstrapped (100×) and using Kimura's two-parameter genetic distance was calculated (9). Bootstrap values for each node are given. The alignment and the tree can be obtained from TreeBase (http://www.treebase.org; reference numbers: S1066, M1817).



H and K are Finnish isolates; U, North American isolates; JP, European isolates.

* Isolate K11 was identical with isolates K8, U9, U2-12, JP24, JP28, JP87, JP178, JP180, JP193 and JP271.

** Isolate U2-7 was identical with U2-8, JP39, JP43, JP44, JP84, JP86, JP88, JP164 and JP192.

Fig. 4. Unrooted UPGMA consensus tree of *Entoleuca mammata* isolates. The sequence data from *H*cat marker locus was bootstrapped (100×) and using Kimura's two-parameter genetic distance was calculated (9). Bootstrap values for each node are given. The alignment and the tree can be obtained from TreeBase (http://www.treebase.org; reference numbers: S1066, M1818).

organisms introduced from one location to another should have reduced genetic diversity compared to the original population. Good examples of such effects are the reduced diversity of the European race of *Gremmeniella abietina* in North America (Hamelin, Lecours & Laflamme 1998), and the clonal spread of the Irish potato famine fungus, *Phytophtora infestans* (Goodwin, Cohen & Fry 1994). Therefore, low variation can be used as a fingerprint for introduced species. In this study the genetic diversity of *E. mammata* was estimated using Nei's diversity, which was lower in EU than in NA. As this difference could have been due to differences in sample sizes, we calculated rarefaction curves and Chao-estimates, both of which supported our hypothesis. This was also in accordance with our being able to identify a unique genotype for all NA-isolates, whereas groups of identical genotypes were observed among EU-isolates. All these analyses (Nei's diversity, rarefaction, Chao-estimator, and distribution of genotypes) support a hypothesis suggesting introduction of *E. mammata* from NA to EU.

If an introduction from NA to EU had taken place, genetic differentiation should have occurred due to a founder effect and random drift. In accordance with this, both F_{st} and G_{st} values suggested that *E. mammata* populations in the two continents would be genetically differentiated. As it was unclear whether the very high number of alleles per locus could cause bias in the analyses of population differentiation, we tested this possibility by calculating the same indices (G_{st}, F_{st}) from the measured (total) data and from an artificial

Table 3. Expected number of alleles (Chao-estimators) and Nei's genetic diversities calculated for the two loci.

	Chao-estimates		Nei's diversity		
Marker	North America	Europe	North America	Europe	All samples
*H*cat	76	14	0.94	0.81	0.90
*H*gag	29	15	0.92	0.58	0.81

Table 4. Comparison of F_{st} and G_{st} values calculated for the original loci *H*gag and *H*cat and averaged over artificial loci gag5, cat7 and gagL, catL which presented lower diversity with 6, 8, 5 and 7 alleles for locus, respectively.

Marker	F_{st}	G_{st}
*H*cat & *H*gag	0.193	0.160
gag5 & cat7	0.285	0.233
gagL & catL	0.088	0.067
Over all artificial loci	0.187	0.150

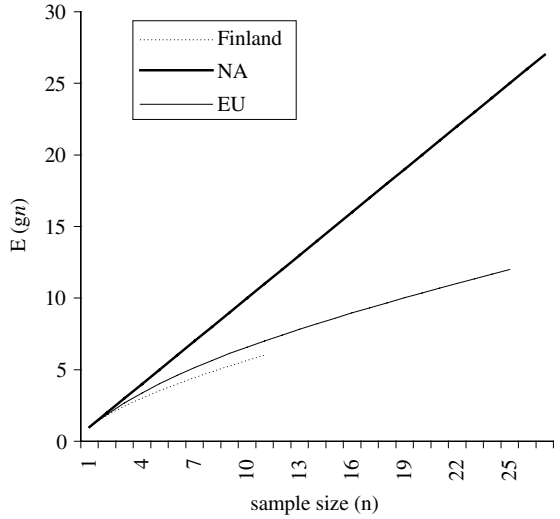


Fig. 5. Richness estimated using the rarefaction method, E(g_n). Richness estimates for haplotypes; North America (USA), Finland, and Central Europe (EU).

data set (created by including only a subset of total variation to reach a more 'normal' number of alleles/locus). Since the F_{st} and G_{st} values for both total data and artificial data sets were in agreement, we concluded that the high number of alleles/locus did not affect the analysis. Therefore, our results indicate true genetic difference between populations of *E. mammata* in the two continents and support the hypothesis that *E. mammata* was introduced from NA to EU.

The number of shared alleles in Europe and North America is relatively small at both loci. This is, however, expected in hypervariable loci when the sampling is not sufficient to cover most of the alleles occurring in the populations (as was the case during this study). In short, assuming that there are no large differences in allele frequencies, the expected numbers of shared alleles (if in EU only a subset of NA alleles occurred) can be calculated from the following equation:

$$ALLshared = \frac{EUest}{NAest} \times EUobs$$

where *ALLshared* is the expected number of alleles occurring both in EU and NA, *EUest* is the estimated total number of alleles in Europe, *NAest* is the estimated total number of alleles in NA and *EUobs* is the observed number of alleles in EU (Table 2). For *EUest* and *NAest* we used the Chao estimators obtained for European and North American populations, respectively (Table 3).

For *H*cat-locus the Chao estimators were 14 and 76 alleles in EU and NA, respectively, and in Europe altogether nine alleles were observed. Based on these figures, 1.7 ($9\times14/76$) shared alleles were expected to be found in our data set. This figure is very close to the two shared alleles actually observed. For *H*gag-locus the expected number of common alleles in EU and NA was 4.1 ($8\times15/29$), which also is close to the three common alleles actually observed. Therefore, these calculations support the introduction hypothesis.

Before concluding that *E. mammata* was introduced by man from North America to Europe, we must consider an alternative explanation for the reduced diversity in Europe. Since the sampling of European isolates was conducted over 27 yr, it is possible (although unlikely) that population shifts occurred over this period. Another possibility is the proportion of latent infections (Chapela 1989), which were not subject to sampling; theoretically this could have skewed the sampling because only aggressive (canker-causing) isolates have been collected. However, our experience is that if a canker is 'contained' or limited by callus growth it is mostly due to resistance of the genotype of the host clone rather than differences in aggressiveness of the isolates. We also consider that selection of aggressiveness should not have had an effect on the results obtained from selectively neutral molecular markers, assuming that the markers and aggressiveness are not genetically linked. If these distant possibilities are excluded, the diversity of European population would reduce if *E. mammata* was native in both North America and Europe and the European population would have gone through a genetic bottleneck after the separation of the two continents. Such a bottleneck could have been caused by glaciations if the Mediterranean sea and the mountains in southern Europe would have hindered the escape of fungi from ice. The outcome of such a bottleneck would have been an acceleration of neutral marker evolution due to fixation of new mutations in small populations. This should have been seen in our data, because unique alleles in Europe would have reduced the number of alleles shared by NA and EU. This, however, was not the case (see above), and therefore we favor the introduction hypothesis over the subdivision + bottleneck

hypothesis. This conclusion is also in accordance with the literature, as there are several native fungi in NA and EU that have evolved to genetic dissimilarity. Examples include *Gremmeniella abietina* (Dusabenyagasani, Laflamme & Hamelin 2002), *Heterobasidion annosum* IS-group S (Johanneson & Stenlid 2003), and *Phlebia gigantea* (Vainio & Hantula 2000).

In NA populations, we observed a high degree of polymorphism which is consistent with the assumption that *E. mammata* is a common, variable and native pathogen in NA. The identification of a unique genotype for each NA isolate strongly suggests that *E. mammata* is an actively outcrossing fungus lacking asexual spread but disseminated primarily by ascospores (Griffin *et al.* 1992). Observations of a complex interaction of *E. mammata* with several wounding agents; woodpeckers and insects such as the poplar borers *Saperda inornata* and *Oberea schaumii*, the periodical cicada *Magicicada septendecim*, the dogday cicada *Tibicen linnei* and the treehopper *Telamona tremulata* (Ostry & Anderson 1998) also suggest a long-term coevolution of the host-pathogen system in North America.

In speculating on possible introduction events, the host range of *E. mammata* must be taken into consideration. The fungus has a wide host range of worldwide genera of trees, including *Acer*, *Alnus*, *Carpinus*, *Fagus*, *Picea*, *Populus*, *Pyrus*, *Salix*, *Sorbus*, and *Ulmus* (Miller 1961). It is therefore possible that the fungus was introduced with some tree species other than poplar or aspen, perhaps some seedlings brought in for ornamental or gardening purposes.

It is possible that in Europe the distribution of the fungus is actually wider than recognized and may include hosts other than *Populus* sp. For example, several species of *Salix* are common, widespread and susceptible (Miller 1961). Being an economically overlooked tree species, there are no reports of the disease on *Salix* in Europe, but herbarium samples have been collected. Unfortunately only a few limited host range studies have been conducted by cross-inoculating hosts, and these have been done only in North America (French, Hodges & Froyd 1969, French & Juzwik 1984). Although little data from other tree species are available, sample JP323 collected from *Salix* sp. (Switzerland) had the most dissimilar alleles at both loci, suggesting that the pathogen in the two tree genera might be distinct although closely related.

Examples of introduced fungi include ones carried by humans and long-distance dispersal by air and other vectors (Brown & Hovmoller 2002). In many cases introduced pathogens such as *Cryptonectria parasitica*, the cause of chestnut blight, and more recently *Phytophthora ramorum*, the cause of sudden oak death, have caused severe outbreaks in new environments or on new hosts (Beattie & Diller 1954, Rizzo *et al.* 2002). Compared to those examples, *E. mammata* has so far caused only minor damage to European aspen stands, although in local situations it can be devastating. In North America, where the pathogen causes more damage, a number of wounding agents (insects and birds) and various environmental variables are believed to play critical roles in spreading the pathogen (Ostry & Anderson 1998). Although the importance of wounding agents or vectors has not been studied in Europe it is possible that an absence of some factor could be limiting the ability of *E. mammata* to spread and cause damage.

In this study we tested two hypotheses: conspecificity of *E. mammata* in NA and EU, and the hypothesis of introduction. The hypothesis of conspecificity was supported by the observation of identical alleles and absence of phylogenetic differentiation. Supporting the introduction hypothesis, analyses of the distribution of genotypes, degree of genetic variation and genetic differentiation all imply that *E. mammata* is a native pathogen in North America and an introduced one in Europe. The introduction may have occurred several centuries ago, and the actual introduction event remains unknown. The degree of genetic variation is relatively high in Europe but considerably less than in North America. This suggests that the introduction was not a single event but most likely occurred in several locations during the last centuries.

ACKNOWLEDGEMENTS

We thank Egbert Beuker for field observations of damage to aspen, Neil Anderson for reviewing the manuscript, and Arlette Schipfer for her participation in survey and isolations. This study was financially supported by the Metsäliitto Group.

REFERENCES

Bagga, D. K. & Smalley, E. B. (1974) Variation of *Hypoxylon pruinatum* in cultural morphology and virulence. *Phytopathology* **64**: 663–667.

Beattie, R. K. & Diller, J. D. (1954) Fifty years of chestnut blight in America. *Journal of Forestry* **52**: 323–329.

Brown, J. K. M. & Hovmoller, M. S. (2002) Aerial dispersal of pathogens on the global and continental scales and its impact on plant disease. *Science* **297**: 537–541.

Chao, A. (1987) Estimating the population size for capture-recapture data with unequal catchability. *Biometrics* **43**: 783–791.

Chapela, I. H. (1989) Fungi in healthy stems and branches of American beech and aspen: a comparative study. *New Phytologist* **113**: 65–75.

Dusabenyagasani, M., Laflamme, G. & Hamelin, R. C. (2002) Nucleotide polymorphisms in three genes support host and geographic speciation in tree pathogens belonging to *Gremmeniella* spp. *Canadian Journal of Botany* **80**: 1151–1159.

Felsenstein, J. (1989) PHYLIP – phylogeny inference package (Version 3.2). *Cladistics* **5**: 164–166.

French, D. W., Hodges, C. S. & Froyd, J. D. (1969) Pathogenicity and taxonomy of *Hypoxylon mammatum*. *Canadian Journal of Botany* **47**: 223–226.

French, D. W. & Juzwik, J. (1984) *Hypoxylon* canker of *Salix* sp. *Phytopathology* **74**: 840.

Goodwin, S. B., Cohen, B. A. & Fry, W. E. (1994) Panglobal distribution of a single clonal lineage of the Irish potato famine fungus. *Proceedings of the National Academy of Sciences, USA* **91**: 11591–11595.

Goudet, J. (2000) *FSTAT: a program to estimate and test gene diversities and fixation indices*. Version 2.9.3. http://www.unil.ch/izea/softwares/fstat.html.

Griffin, D. H., Quinn, K. E., Gilbert, G. S., Wang, C. J. K. & Rosemarin, S. (1992) The role of ascospores and conidia as propagules in the disease cycle of *Hypoxylon mammatum*. *Phytopathology* **82**: 114–119.

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

Hamelin, R. C., Lecours, N. & Laflamme, G. (1998) Molecular evidence of two distinct introductions of the European race of *Gremmeniella abietina* to North America. *Phytopathology* **88**: 582–588.

Hantula, J., Dusabenyagasani, M. & Hamelin, R. C. (1996) Random amplified microsatellites (RAMS) – a novel method for characterizing genetic variation within fungi. *European Journal of Forest Pathology* **26**: 159–166.

Illstedt, B. & Gullberg, U. (1993) Genetic variation in a 26-year old hybrid aspen trial in Southern Sweden. *Scandinavian Journal of Forest Research* **8**: 185–192.

Johannesson, H. & Stenlid, J. (2003) Molecular markers reveal genetic isolation and phylogeography of the S and F intersterility groups of the wood-decay fungus *Heterobasidion annosum*. *Molecular Phylogenetics and Evolution* **29**: 94–101.

Kimura, M. (1980) A simple model for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *Journal of Molecular Evolution* **16**: 111–120.

Manion, P. D. & Griffin, D. H. (1986) Sixty-five years of research on *Hypoxylon* canker of aspen. *Plant Disease* **70**: 803–808.

Mazzaglia, A., Anselmi, N., Vicario, S. & Vannini, A. (2001) Sequence analysis of the 5.8S rDNA and ITS regions in evaluating genetic relationship among some species of *Hypoxylon* and related genera. *Mycological Research* **105**: 670–675.

Miller, J. H. (1961) *A Monograph of the World Species of Hypoxylon*. University of Georgia Press, Athens.

Nei, M. (1973) Analysis of gene diversity in subdivided populations. *Proceedings of the National Academy of Sciences, USA* **70**: 3321–3323.

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

Ostry, M. E. & Anderson, N. A. (1998) Interactions of insects, woodpeckers and *Hypoxylon* canker on aspen. *USDA Forest Service Research Paper* 331.

Page, R. D. M. (1996) TREEVIEW: an application to display phylogenetic trees on personal computers. *Computer Applications in the Biosciences* **12**: 357–358.

Pinon, J. (1979) Origine et principaux caractéres des souches francaises d'*Hypoxylon mammatum* (Wahl.) Miller. *European Journal of Forest Pathology* **9**: 129–142.

Rizzo, D. M., Garbelotto, M., Davidson, J. M., Slaughter, G. W. & Koike, S. T. (2002) *Phytophthora ramorum* as the cause of extensive mortality of *Quercus* spp. and *Lithocarpus densiflorus* in California. *Plant Disease* **86**: 205–214.

Rogers, J. D. & Ju, Y.-M. (1996) *Entoleuca mammata* comb. nov for *Hypoxylon mammatum* and the genus *Entoleuca*. *Mycotaxon* **59**: 441–448.

Simberloff, D. S. (1978) Use of rarefaction and related methods in ecology. In *Biological Data in Water Pollution Assessment: quantitative and statistical analyses* (K. L. Dickson, J. Cairns & R. J. Livingston, eds): 150–165. American Society for Testing and Materials, Philadelphia.

Terrasson, D., Pinon, J. & Ridé, M. (1988) L'infection naturelle de *Populus trichocarpa* par *Hypoxylon mammatum*. *Revue Forets Francaise* **40**: 126–130.

Vainio, E. J. & Hantula, J. (2000) Genetic differentiation between European and North American population of *Phlebiopsis gigantea*. *Mycologia* **92**: 436–446.

Vainio, E. J., Korhonen, K. & Hantula, J. (1998) Genetic variation in *Phlebiopsis gigantea* as determined with random amplified microsatellite (RAMS) markers. *Mycological Research* **102**: 187–192.

Weir, B. S. & Cockerham, C. C. (1984) Estimating F-statistics for the analysis of population structure. *Evolution* **38**: 1358–1370.

Williams, J. G. K., Kubelik, A. R., Livak, K. J., Rafalski, J. A. & Tingey, S. V. (1990) DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. *Nucleic Acids Research* **18**: 6531–6535.

Zhou, Z. H., Miwa, M. & Hogetsu, T. (1999) Analysis of genetic structure of a *Suillus grevillei* population in a *Larix kaempferi* stand by polymorphism of inter-simple sequence repeat (ISSR). *New Phytologist* **144**: 55–63.

Corresponding Editor: G. W. Griffith