Wachowiak, Witold; Boratynska, Krystyna; Cavers, Stephen. 2013. **Geographic patterns of nucleotide diversity and population differentiation in three closely related European pine species from the Pinus mugo complex**. *Botanical Journal of the Linnean Society*, 172 (2). 225-238. 10.1111/boj.12049

Geographic patterns of nucleotide diversity and population differentiation in three closely related European pine species from the *Pinus mugo* complex

WITOLD WACHOWIAK[1,2*], KRYSTYNA BORATYŃSKA[1], STEPHEN CAVERS[2]

[1]*Institute of Dendrology, Polish Academy of Sciences, Parkowa 5, PL-62-035 Kórnik, Poland*
[2]*Centre for Ecology and Hydrology Edinburgh, Bush Estate, Penicuik, Midlothian EH26 0QB, UK*

***Corresponding author:** Witold Wachowiak (witoldw@man.poznan.pl), Institute of Dendrology, Polish Academy of Sciences, Parkowa 5, PL-62-035 Kórnik, Poland

Abstract

Nucleotide polymorphism at twelve nuclear loci and two mitochondrial gene fragments was studied in three closely related pine species from the *Pinus mugo* complex in populations across the species distribution range in Europe. Despite large differences in census size of the populations, high and similar levels of nucleotide diversity (θ_{sil}=~0.013-0.017) were found at nuclear loci in the three pine species. More rapid decay of overall linkage disequilibrium and recombination to diversity ratio (ρ/θ) was observed across the species distribution range in *P. mugo* (ρ=0.0369±0.0028; ρ/θ=~2.2) as compared to *P. uncinata* (ρ =0.0054±0.0011; ρ/θ=~0.4) and *P. uliginosa* (ρ = 0.0051±0.0010, ρ/θ=~0.4). However, regional groups of *P. mugo* showed similar levels of LD and ρ/θ ratio to the other species. An excess of rare nucleotide variants was found in *P. mugo* at four loci, but overall the allelic frequency spectrum in the three species did not significantly deviate from neutrality (multilocus Tajima's *D*= -0.681, *D*= -0.118 and *D*= -0.266, p>0.05, respectively). Bayesian clustering methods showed no clear correspondence of clusters to species or geographical regions. Some differences between populations and species were found in a hierarchical AMOVA and in the distribution of the *mt*DNA haplotypes suggesting rather limited gene flow between the taxa and ongoing divergence. As all three pine taxa have very similar genetic backgrounds they form an excellent system for searching for loci involved in adaptive variation and speciation.

ADDITIONAL KEYWORDS: adaptation – natural selection – divergence – speciation – *mt*DNA variation.

INTRODUCTION

Many closely related species can be identified based on detailed morphological characteristics, however much less is usually known about their corresponding genetic divergence. At the interspecific level, the pattern of nucleotide diversity is influenced by several factors including the time of speciation, demographic history, natural selection and hybridization. Loci involved in species differentiation show increased divergence relative to background variation as a result of directional selection (Excoffier, 2004; Coyne & Orr, 2004). Deviations from the standard neutral model of evolution at specific loci may indicate their role in adaptation and speciation. In contrast, deviation from the neutral model at many loci is expected when species have undergone significant demographic fluctuations, for example during their range shifts. Hence, multilocus studies of nucleotide diversity and among-population divergence in closely related species can resolve evolutionary and demographic signals and help to explain speciation.

The three pine species investigated here, *Pinus mugo* Turra *s.l.* (dwarf mountain pine), *P. uncinata* Ramond (mountain pine) and *P. uliginosa* Neumann (peat-bog pine), differ significantly from each other in phenotype, geographical distribution and ecology. *Pinus mugo* is a high-altitude polycormic European pine of up to few meters in height and *P. uncinata* and *P. uliginosa* are trees of up to 20 m height adapted to mountain environments and lowland peatbogs, respectively. Morphological and ecological divergence among these taxa are likely due to adaptation to different habitats resulting from the disjunction of species ranges and the isolation of populations in different Pleistocene refugia (Christensen, 1987a; Christensen, 1987b). Taxonomically they are treated as independent species or as subspecies in the *P. mugo* Turra complex (Christensen, 1987b; Businský & Kirschner, 2006). They share karyotypic features with other pine species from the subgenus *Pinus* (Bogunić *et al.*, 2011). Moreover, reproductive isolation between the three taxa is not complete and they hybridize naturally (Businský, 1998; Wachowiak & Prus-Głowacki, 2008).

At present, *P. mugo* forms large, shrubby populations above the tree line up to an altitude of about 2700 m with major populations in mountainous regions of Western Europe (the Alps), Central and Eastern Europe (Sudetes, Tatars, Carpathians), and south-east through Bosnia and Herzegovina, Montenegro, Serbia and Romania to the Rila and Pirin Mts of Bulgaria, as well as several outlier populations (Critchfield & Little, 1966). *Pinus uncinata* is a forest

forming component in the upper mountain regions at altitudes of 1400-2700 m in the Pyrenees, the Massif Central, Western Alps and several marginal populations in the Iberian Peninsula. *Pinus uliginosa* forms several isolated populations in lowland peatbogs in Central Europe (Hamerník & Musil, 2007). Given that the species currently form mostly geographically disjunct populations and occupy specific habitats of considerable ecological diversity, it seems likely that interglacial range shifts to their present day distribution have potentially influenced standing nucleotide variation.

In this study we used data on nucleotide polymorphism from multiple genomic regions to look at the patterns of divergence and underlying evolutionary processes occurring during speciation of the taxa from the *P. mugo* complex. We analysed nucleotide and haplotype diversity, allele frequency and linkage disequilibrium in multiple nuclear gene fragments in samples from multiple populations across the species ranges. We looked also at differentiation at two gene fragments from the mitochondrial genome which in pines is maternally inherited and seed dispersed and therefore allows assessment of population structure more effectively than pollen-dispersed *n*DNA markers. Using this data we aimed to 1) investigate the level of nucleotide and haplotype diversity in closely related taxa 2) identify among-species variation due to selection or demographic processes related to geographical isolation and migration and 3) identify the primary drivers of among-species differentiation.

MATERIALS AND METHODS

SAMPLING AND DNA EXTRACTION

Seed samples from 31 populations of the three pine taxa were included in this study (Figure 1, Table S1). Seeds from twelve to fifteen individual trees were sampled from each population. Seeds were germinated for 4-7 days in moist petri dishes. Genomic DNA was extracted from haploid megagametophyte (maternal tissue that surrounds the embryo in the seed) from a single seed from each mother tree following a standard Qiagen DNeasy Plant Mini Kit protocol. As DNA samples were haploid, the haplotypes were determined by direct sequencing.

NUCLEAR AND MITOCHONDRIAL LOCI STUDIED

In total, 12 nuclear loci were analyzed. Those include ABI3-interacting protein (*a3ip2*), abscissic acid-responsive protein (*abaR*), chalcone synthase (*chcs*), caffeoyl CoA *O*-methyltransferase (*ccoaomt*), early response to dehydration 3 protein (*erd3*), abscissic acid, water dehydrative stress and ripening induced gene family members 1 and 3 (*lp3-1, lp3-3*) and several dehydrin genes including some that were identified in expression studies in Scots pine (Joosen *et al.*, 2006). Variation in the mitochondrial genome was determined for the *nad*1 intron using internal PCR primers *nad*1E-G (Soranzo *et al.*, 2000) and for *nad*7 intron 1 (Jaramillo-Correa, Beaulieu & Bousquet, 2004). A full list of genes used in the study and PCR amplification conditions are presented in Table S2.

PCR AMPLIFICATION AND SEQUENCING

PCR-amplification was performed with Thermo MBS thermal cyclers and carried out in a total 25µl containing about 10ng of haploid template DNA, 50µM of each of dNTP, 2µM of each primer and 0.25U *Taq* DNA polymerase (BioLabs) with the respective 1 x PCR buffer. PCR followed standard amplification procedures with $MgCl_2$ concentration optimized for each primer pair as described in Table S2. Standard amplification procedures were used with initial denaturation at 94°C for 3min followed by 35 cycles with 30sec. denaturation at 94°C, 30sec. annealing and 1 min. extension at 72°C, and a final 5min. extension at 72°C. PCR fragments were purified using *Exo–Sap* (exonuclease, alckaline-phosphatase) enzymatic treatment. About 20ng of PCR product was used as a template in 10µl sequencing reactions with the Big Dye Terminator DNA Sequencing Kit (Applied Biosystems, Carlsbad, CA, USA) performed by the GenePool sequencing service, University of Edinburgh, UK. Haploid DNA samples were sequenced in one direction for each locus. For mitochondrial gene fragments, several samples from different populations and species were sequence-characterised to check for presence of single nucleotide polymorphisms or potential fragment length variations. Codon-Code Aligner software ver. 3.7.1 (Codon Code Corporation, Dedham, MA, USA) was used for editing of the chromatograms, visual inspection of all polymorphic sites detected and assembly to produce alignments on the basis of the nucleotide sequence.

158 SEQUENCE ANALYSIS

159 For each species, ~5kb of *n*DNA was aligned across genes excluding PCR primer sites (Table
160 S3). Nucleotide sequence alignments were constructed in Codon-Code Aligner and were
161 further manually adjusted using GENEDOC (Nicholas & Nicholas, 1997). Coding and non-
162 coding regions (untranslated regions) were annotated on the basis of alignment with known
163 sequences identified in the international databases (National Center for Biotechnology
164 Information (http://www.ncbi.nlm/nih.gov/gorf/gorf.html) and ExPASy (http://www.expasy.
165 ch/tools/dna.html). Sequence data from all twelve nuclear genes were used for the nucleotide
166 diversity estimates at the between species level. For nucleotide diversity estimates in
167 individual populations and population structure analysis within species, we used sequence
168 data from seven genes (including *a3ip2, abaR, chcs, ccoaomt, dhn2PP, dhn3, erd3*) that were
169 successfully amplified in at least half samples in each population. Considering the large and
170 patchy distribution of *P. mugo* populations in Europe, we also divided the populations into
171 five geographical groups, namely Central Europe (population M1.1, M1.2, M2, M3),
172 Carpathians (M4, M5, M17), Balkans (M6-M9), Alps (M10-M14) and Apennines (M15,
173 M16). Unique haplotype sequences of each locus reported in this paper are deposited in the
174 NCBI sequence database under accession numbers JQ027729-JQ028129. The mitochondrial
175 sequences were compared to each other and available GeneBank sequence data for *nad1*
176 intron and *nad7* intron 1 in Scots pine (Soranzo et al., 2000; Naydenov *et al.*, 2007;
177 Pyhäjärvi, Salmela & Savolainen, 2008). Sequence contigs were assembled using
178 CodonCode Aligner software (CodonCode Corporation). A polymorphic 31 bp indel in *nad1*
179 intron observed in the three pine species was scored using diagnostic primers *nad1*H-I
180 (Soranzo et al., 2000). To score the size differences in the *nad*7 intron 1 caused by 5bp indel,
181 the amplified fragment were digested with 0.5 U of *Dra*II restriction enzyme. The PCR
182 products (~5µl) of both polymorphic regions were electrophoretically separated on 2%
183 agarose gel and scored for indel variation.

184 NUCLEOTIDE DIVERSITY AND NEUTRALITY TESTS AT NUCLEAR GENES

185 The number of shared, exclusive and fixed polymorphic sites and their distribution for each
186 nuclear locus among species was determined using SITES 1.1. software
187 (http://lifesci.rutgers.edu/~heylab). Indels were excluded from the analysis. Nucleotide
188 diversity at twelve loci for each species was estimated using population mutation parameter θ
189 (equals to $4N_e\mu$, were N_e is the effective population size and µ is the mutation rate per

nucleotide site per generation; Watterson (1975)). Multilocus estimates of silent theta were computed on basis of the number of segregating synonymous and non-coding sites and the length of each locus using the MCMC approach as previously reported (Pyhäjärvi *et al.*, 2007). The estimates of multilocus nucleotide diversity were conducted at 12 genes combined for each species separately and also for samples from each of the five geographical regions defined for *P. mugo*.

Deviations of particular genes from the frequency distribution spectrum expected under the standard neutral model of evolution were assessed with Tajima's *D* (Tajima, 1989). The distribution of Tajima's *D* was investigated for each locus at all samples combined from each species, and separately for all individual populations in each species and the groups of populations defined for *P. mugo*. Statistically significant deviation from neutral model of evolution was evaluated by coalescence simulations as implemented in DnaSp v.5 software (Librado & Rozas, 2009). Orthologous sequences from the outgroup species *P. pinaster* (subgenus *Pinus*) were used in two heterogeneity tests, the McDonald-Kreitman (MDK) test (Thornton, 2005) and the Hudson-Kreitman-Aguadé (HKA) test (Jiggins *et al.*, 2008). The significance of multilocus estimates of the Tajima's *D* and HKA test statistics were evaluated by comparison to a distribution generated by 1000 coalescent simulations using the HKA program (http://lifesci.rutgers.edu/~heylab). The MDK test was conducted in DnaSp v.5.

LINKAGE DISEQUILIBRIUM AND HAPLOTYPE DIVERSITY AT NUCLEAR LOCI

The level of linkage disequilibrium for each species was measured as the correlation coefficient r^2 (Hill & Robertson, 1968) using informative sites. The decay of linkage disequilibrium with physical distance was estimated using non-linear regression of r^2 between polymorphic sites and the distance in base pairs (*bp*) between sites. The non-linear least-squares estimate of ρ ($\rho=4N_e c$, were N_e is effective population size, c is the recombination rate) was fitted by the *nls*-function implemented in R statistical package (http://www.r-project.org). For *P. mugo*, the level of linkage disequilibrium was also measured in the five geographical groups. The number of haplotypes (N_e) and haplotype diversity (H_d) were computed for each gene using DnaSP v.5. The number and frequency of unique and shared haplotypes in pairwise comparisons between species was calculated with Arlequin v.3 (Excoffier, Laval & Schneider, 2005). To further check relationships between the samples at each locus, the haplotypes of combined datasets of all species were clustered in neighbour-

joining analyses using genetic distances corrected for multiple hits as implemented in MEGA v.4 software (Tamura *et al.*, 2007). Genealogical relationships among sequences were estimated by statistical parsimony using the method of Templeton *et al*. (1992) as implemented in TCS software (Clement, Posada & Crandall, 2000). Gaps in sequence alignments were coded in TCS analysis as missing data.

COMPARISON OF POLYMORPHISM AND DIVERGENCE AT NUCLEAR LOCI

Nucleotide polymorphism based on non-synonymous (π_a) and synonymous (π_s) sites and nucleotide divergence based on non-synonymous (K_a) and synonymous (K_s) sites relative to the outgroup *P. pinaster* species was calculated using DnaSP v.5 software. Interspecific differentiation between species was measured by the difference between the average pairwise divergence between species and the average intraspecific pairwise variation (net divergence; Nei (1987)), which should increase with time since speciation. Locus-by-locus estimates of net between-species divergence per site were obtained by correcting for within-species diversity using SITES 1.1. Silent divergence in pairwise comparisons between species at each locus was calculated in DnaSP v.5.

POPULATION STRUCTURE AT NUCLEAR AND *MT*DNA LOCI

To measure differentiation among populations and groups of populations within species a fixation index F_{ST} (Weir & Cockerham, 1984) was calculated for each locus and tested for significance by 1000 permutations as implemented in Arlequin v.3. As all species were closely related and may hybridize naturally, F_{ST} statistics were also estimated at the between species level. Differentiation between populations was also measured as a weighted average over all polymorphic sites. The hierarchical distribution of genetic variation among taxa and populations within taxa, among 5 geographical regions defined in *P. mugo* and among populations within regions was estimated at seven nuclear loci using an analysis of molecular variance (AMOVA) in Arlequin v.3. The genetic structure of the populations at nuclear loci was explored using the genetic mixture analysis of linked haploid sequences data as implemented in BAPS software (Corander & Tang, 2007). The MLST-format as a separate fasta file was used for each locus and ten independent runs were conducted for each K (1-15) to estimate the number of clusters for all samples combined and also separately for each species. The spatial distribution of mitochondrial haplotypes was mapped using the GIS software ArcMap in ArcGIS v9.2 (ESRI Ltd).

RESULTS

NUCLEOTIDE DIVERSITY AND NEUTRALITY TESTS

Across all species, the genetic background was very homogenous at the twelve nuclear loci. No fixed differences were found between the species; they shared ~60-74% of all polymorphic sites detected and they showed very similar and low average silent (~0.02) and net divergence (~0.001) (Figure 2, Table S4-S5).

Overall, the average nucleotide diversity (π) at twelve loci was slightly higher in *P. mugo* as compared to *P. uncinata* and *P. uliginosa* at total (π=0.0118 vs. 0.0113 and 0.0091), silent (π=0.0183 vs. 0.0178 and 0.0141) and nonsynonymous sites (π=0.0070 vs 0.0058 and 0.0048) (Table 1). This high nucleotide diversity was mostly due to the variation at six loci (*chcs, dhn2, dhn7, dhn9, lp3-1* and *lp3-3*) with π>0.01. Multilocus estimates of silent theta (θ) were high for all species with θ = ~0.017 for *P. mugo* and slightly lower values for the other species (θ = ~0.013), however the estimates for all taxa showed broad and partly overlapping credibility intervals (Table 2, Figure 3). Very similar nucleotide divergence was found in each of the five geographical regions defined for *P. mugo* (θ= ~0.011-0.014) (Supplementary Table S6). In individual populations the average total nucleotide diversity at seven loci varied between π=0.0032 (M8) to π=0.0094 (UN4) (Table S7).

A tendency towards an excess of singleton mutations across genes was detected at twelve loci in the three pine species (*D*= -0.681, *D*= -0.118 and *D*= -0.266, respectively) (Table 1) and in each of five geographical regions of *P. mugo* (Table S7), but the difference was not statistically significant (P>0.05). At individual loci and all samples combined for each species, a significantly negative value of Tajima's *D* was found only in *P. mugo* at 4 loci (*a3iP, ccoaomt, dhn3, erd;* Table S3). At most individual loci and populations, *D* was not significantly different from zero with few exceptions including D<0 at locus *a3ip* (population UN9, *P. mugo* from Alps), *ccoaomt (P. mugo* from Carpathians, Balkans and Alps), *dhn3* (M8,12,13,15,UN2,5,8, *P. mugo* from Central Europe, Balkans and Apennines), *erd* (M8), *dhn9* (UN2,6), *lp33* (UN 3) and D>0 at *abaR* (M6,16, UN7), *chcs* (M6,14,15, UN5,7, UG1, *P. mugo* from Alps and Apennines), *dhn2PP* (M8), *dhn3* (UN4). Average values of *D* in individual populations at 7 loci were not significantly different from zero (Table S7).

The taxa showed very similar divergence from the outgroup *P. pinaster* at all (Ks=~0.04) and silent nucleotide sites (Ks=~0.05). Positive correlation between within species polymorphism and divergence from the outgroup *P. pinaster* was found in the multilocus HKA test. No deviation from the standard neutral model of evolution was found in the McDonald-Kreitman test.

LINKAGE DISEQUILIBRIUM AND HAPLOTYPE DIVERSITY

Rapid decay of linkage disequilibrium between pairs of parsimony informative sites was found across the species (Figure S1) with observed r^2 values of ~0.2 at a distance of about 100 to 500 bp. The decay of linkage disequilibrium measured by non-linear least-squares estimate of ρ was more rapid for all samples combined in *P. mugo* (ρ= 0.0369±0.0028) as compared to *P. uncinata* (ρ =0.0054±0.0011) and *P. uliginosa* (ρ = 0.0051± 0.0010). The rate of decay of LD and the relative level of recombination to diversity (ρ / θ ratio) were about five times higher in *P. mugo* (ρ/θ=~2.2) as compared to the other taxa (ρ/θ=~0.4) (Table 2). The estimates of both LD and ρ / θ ratio in five geographical regions of *P. mugo* analysed separately were similar or slightly lower than for the other species (Supplementary Table S6).

The average number of haplotypes per gene (*Ne*) differed between the species; however the overall haplotype diversity was similar with average H_d = ~0.6-0.7 (Table 1). About 30% of all haplotypes detected were shared in pairwise comparisons between species (Table S8) and the haplotypes did not cluster by taxon (or population) at any locus. In the network analysis, the haplotype with the highest outgroup probability was shared between all three taxa at each locus studied (see examples presented in Supplementary Figure S2). No major clade specific to one species or geographical location was detected in any of the haplotype networks. Unique haplotypes were found at each locus in each species except for loci *a3ip, erd* and *lp3-3* in *P. uliginosa* (Table S8). Across seven loci, every population in each of the three species contained at least one unique haplotype (except population M10, Table S7).

POPULATION STRUCTURE AT NUCLEAR AND MITOCHONDRIAL GENE FRAGMENTS

Average among population differentiation within the species at all polymorphic sites combined at seven loci was low but statistically significant (average F_{ST}=~0.06; ~0.03; ~0.03 for *P. mugo*, *P. uncinata* and *P. uliginosa*, respectively). Significant differentiation was

present over the five *P. mugo* geographical groups (F_{ST}=~0.04, P<0.05, all polymorphic sites, 7 loci). Significant pairwise differentiation was found between Central European vs. Carpathian (0.033) and Alpine (0.041) populations and also between Balkan and Apennine sites (0.043, Table S9). However, these values were largely due to populations M1, M4 and M13 that were significantly differentiated from most of the other populations (overall F_{ST} at 15 populations excluding M1, M4 and M13 was ~0.016, P>0.05). In *P. uncinata*, this variation was due to populations UN4 and UN10, which were differentiated from UN1,2,5,10 and most other populations, respectively (overall F_{ST} at 8 populations excluding UN4,10 was ~0.018, P>0.05). In *P. uliginosa* populations UG1 and UG2 were significantly differentiated at only the locus *erd* (overall F_{ST} at 6 loci ~0.02, P>0.05). Pairwise F_{ST} at all polymorphic sites showed significant differentiation between each pair of taxa. At individual loci, significant differentiation was found between *P. mugo* and *P. uncinata* at 9 loci (except *ccoaomt, dhn3, dhn7*), between *P. mugo and P. uliginosa* at 5 loci (including *a3ip, abaR, dhn2pp, erd, lp3-3*) and between *P. uliginosa and P. uncinata* at 4 loci (including *chcs, dhn2, erd, lp3-1*).

In a hierarchical AMOVA, *P. mugo* and *P. uncinata* showed weak but significant differentiation from each other (F_{CT} = 0.022, P< 0.05). In total, about 90% of the variation was present within populations that showed significant within species differentiation (F_{SC}=~0.056-0.083, p<0.001). No differentiation was found among the 5 geographical regions of *P. mugo*.

The clustering analysis of all samples combined across the three pine species and also analysed separately for each of the species suggests in each case the presence of three genetic clusters (K = 3). However the clusters included mixtures of individuals from all species and populations and there was no clear correspondence with species or geographical region. In the three taxa, the samples from clusters 1 and 2 were present in all populations and geographical regions except for population M7, which contained only samples assigned to cluster 1, and population UN4, which contained only samples assigned to cluster 2. Cluster 3 was the smallest, containing only 9 samples of *P. mugo* (from population M12 (1), M13(3),M15(1) and M17(4)), 5 samples of *P. uncinata* (populations UN4(4) and UN7(1)) and 8 of *P. uliginosa* (populations UG1(2), UG2(3), UG3(3)).

In the mitochondrial genome three mitotypes were detected including 5bp insertion in *nad7.1* and 31bp deletion in *nadHI* regions (mitotype 1) and 5bp and 31bp deletions (mitotype 2)/ insertions (mitotype 3) in *nad7.1* and *nadHI* regions, respectively. Mitotype 1 was the most

common in *P. mugo* and *P. uliginosa* populations. Mitotype 2 was found only in single *P. mugo* populations from Poland, Slovakia and Slovenia and the *P. uliginosa* population from Batorów reserve (Table S1). Mitotype three was detected in *P. uncinata* except one population of the species from Italy fixed for mitotype 1 (Figure1, Table S1).

DISCUSSION

GENETIC VARIATION AT NUCLEAR LOCI

We used nucleotide polymorphisms at a set of nuclear loci to investigate patterns of nucleotide divergence in a complex of three European pine species. The loci selected are potentially related to dehydrative stress and/or cold tolerance in plants and may therefore play a role in local adaptation to mountain *vs.* lowland environments. Among these habitats, adaptive divergence might be generated by directional or balancing selection. For loci under directional selection, we would expect a loss of shared polymorphism and accumulation of fixed differences between taxa (Excoffier, 2004, Coyne & Orr, 2004). Alternatively, if balancing selection had acted we would expect maintenance of shared polymorphisms across populations and species. However, our data provide no evidence of selection as the loci showed very low variance in net divergence and for each taxon analyzed we observed a positive overall correlation between intraspecific polymorphism and divergence from the outgroup, *P. pinaster*. Therefore, we assume that nucleotide polymorphism at the loci was mostly due to neutral evolutionary processes and population history of the species.

Despite large differences in census size of the populations and distribution range, the species showed similar levels of nucleotide polymorphism. The multilocus silent nucleotide diversity was high (θ = ~0.013-0.017) and similar to estimates for closely related Scots pine (*P. sylvestris*) at similar loci in Scotland (θ_{sil}=0.011, Wachowiak *et al.* (2011b) but higher than estimates for continental Scots pine populations (~0.005, Pyhäjärvi *et al.* 2007 *;* ~0.0089, Wachowiak *et al.*, 2009). Silent nucleotide diversity in the *P. mugo* complex was also higher than in other pine species (e.g. *P. taeda* (González-Martinez *et al.*, 2006) and *P. pinaster* (Eveno *et al.*, 2008)). Haplotype diversity was high (H_d = ~0.7) and similar to that observed in *P. sylvestris* (H_d = 0.68; Wachowiak *et al.*, 2009). High nucleotide and haplotype diversity was accompanied by a very similar genetic background across the study species, shown by

383 the high number of shared polymorphisms and haplotypes (Tables S4, S8), lack of fixed
384 differences, and lack of monophyly at any of the nuclear loci studied. Furthermore, the
385 analysis of genealogical relationships among sequences indicated the highest outgroup
386 probability for those haplotypes shared between all three taxa and no major clade specific to
387 one species or geographical location was detected in the haplotype networks (Figure S2).
388
389 The high genetic similarity between taxa from *P. mugo* complex was in line with previous
390 karyotype and molecular data. For instance no differences in genome organization (number
391 of chromosomes 2n=24) or genome size (~40×10^3 Mbp) were observed between *P. mugo* and
392 *P. uncinata* (Bogunić et al., 2011). No differentiation of the taxa was observed using RAPD
393 markers (Monteleone, Ferrazzini & Belletti, 2006) or *cp*SSR loci (Heuertz *et al.*, 2010). Such
394 genetic similarity can be accounted for by several factors that influence the level of
395 nucleotide polymorphism including recent speciation history and interspecific gene flow.
396 Similar levels of nucleotide diversity may suggest speciation has occurred more recently than
397 the time needed for accumulation of fixed differences in neutrally evolving genomic regions.
398 Relatively recent speciation (less than ~5mya) was suggested between closely related *P.*
399 *sylvestris* and *P. mugo* / *P. uliginosa* (Wachowiak, Palme & Savolainen, 2011a). Considering
400 the large effective population sizes, longevity, long generation intervals and efficient gene
401 flow in pines, the time since divergence in the *P. mugo* complex has likely been too short for
402 genetic drift to have generated significant interspecific differences relative to ancestral
403 populations.
404
405 INTERSPECIFIC GENE FLOW AND ONGOING SPECIES DIVERGENCE
406 Retention of ancestral polymorphism could also be due to interspecific gene flow considering
407 the lack of reproductive isolation between the taxa found in controlled crosses (Lewandowski
408 & Wisniewska, 2006) and evidence for natural hybridization in contact zones between *P.*
409 *mugo* and *P. uliginosa*, and between *P. uncinata* and closely related *P. sylvestris* (Wachowiak
410 & Prus-Głowacki, 2008; Jasińska *et al.*, 2010; Kormutak *et al.*, 2008). By comparing spatial
411 differentiation at biparentally and maternally inherited markers, the difference between long
412 distance (pollen) and short distance (seed) gene flow can be assessed. In our study low
413 differentiation across populations was found in cluster analysis at all nuclear gene loci; there

was no clear correspondence of clusters with particular species or geographical regions. Low differentiation between populations of the three taxa was found previously at other pollen mediated markers, e.g. isozymes (Lewandowski, Boratyński & Mejnartowicz, 2000), RAPDs (Monteleone et al., 2006) and *cp*SSR loci (Heuertz et al., 2010). This suggests the homogenizing effects of gene flow on species genetic diversity, as might be expected for highly outcrossing, wind-pollinated species. However, despite overall high genetic similarity between the species, limited seed mediated gene flow is suggested by the clear division at mitochondrial markers over a large geographical area, in particular between *P. uncinata* and *P. mugo* populations (Figure 1). Furthermore, our data provide some evidence of interspecific differentiation at individual nuclear loci, which suggests gene flow may not be unrestricted between the taxa. Unique haplotypes and polymorphic sites were found in most populations corresponding to significant differentiation between the taxa at several individual nuclear gene loci and significant differentiation between *P. mugo* and *P. uncinata*.

Although more markers would be needed to improve resolution of the level of genetic difference among species our data suggest ongoing divergence and limited (or restricted to narrow contact zones) present day gene flow between the taxa. Ongoing divergence was suggested by karyotype observations of the heterochromatin patterns among *P. mugo* and *P. uncinata* (Bogunić et al., 2011) and despite similar genetic background, the taxa investigated here show a considerable amount of phenotypic divergence and can be clearly identified as morphological species. For instance they show significant differentiation in biometric traits including cone morphology (Marcysiak, 2004; Marcysiak & Boratyński, 2007), and morphological and anatomical traits of needles (Boratyńska & Bobowicz, 2001). However, the loci that drive phenotypic differentiation and adaptive divergence of the taxa remain unknown. Previous studies indicated that, although some genomic regions of large phenotypic effect may exist in forest tree species (Kinloch, Parks & Fowler, 1970; Howe *et al.*, 2003), the majority of quantitative traits in trees are controlled by many genes with small individual effects. New methods for studies of the genetic basis of complex traits including genome-wide association genetics and genomic selection approaches will certainly shed new light on the genes under selection and will help to evaluate their effects in shaping phenotypic and physiological differences between closely related yet highly differentiated species.

DEMOGRAPHIC PATTERNS OF NUCLEOTIDE VARIATION

At present, *P. mugo* and *P. uncinata* form mostly allopatric and discontinuous populations in mountainous regions of Europe whereas *P. uliginosa* is restricted to some postglacial peatbogs in Central Europe. Macrofossil data indicates an early Holocene presence of conifers at high elevations (>2000 m) in the northwestern Alps and Central Europe (Blarquez *et al.*, 2010; Cheddadi *et al.*, 2006) suggesting that some populations of cold tolerant species could have survived the last glacial maximum in mountain refugia located north of well known species hotspots in the Iberian, Apennine or Balkan peninsulas (Petit *et al.*, 2003). For species like *P. mugo* and *P. uncinata*, post-glacial expansion at their lower altitudinal limit was followed, with climate warming, by gradual retreat to higher elevations (Burga, 1988; Ramil-Rego *et al.*, 1998; Benito *et al.*, 2007). Especially for *P. mugo,* our data provide molecular signatures that support the hypothesis of species range shifts rather than bottlenecks due to reduction of a previously bigger distribution. *Pinus mugo* showed a significant excess of singleton mutations at four loci, the only one of the taxa to do so, and the highest overall excess of low frequency nucleotide polymorphisms. This pattern of nucleotide diversity suggests range shifts related to recolonisation events, similar to findings in other pines (e.g. in continental European populations of *P. sylvestris* (Wachowiak, Balk & Savolainen, 2009) *P. pinaster* (Eveno et al., 2008) and other conifers (Brown *et al.*, 2004, Beaulieu, Perron & Bousquet, 2004; Krutovsky & Neale, 2005; Heuertz *et al.*, 2006)). As for *P. mugo*, range shifts and gradual retreat with climate warming were postulated for *P. uncinata* populations in modelling studies, based on distribution on the Iberian Peninsula during the last glacial maximum and Mid-Holocene (Benito et al., 2007). However, an excess of singleton mutations was observed in half of the *P. uncinata* and one of the *P. uliginosa* populations (Table S7) and there was a lower overall excess of low frequency variants in both species, accompanied by slower decay of LD relative to *P. mugo* (Table 1, Figure S1). This could be due to severe population size reduction: *P. uliginosa* in particular has a currently very limited and fragmented distribution and has experienced significant decline in recent decades (Wachowiak et al., 2011a; Wachowiak & Prus-Głowacki, 2009).

A similar excess of singleton mutations was found across all geographical regions of *P. mugo* (Table S6). Pooling samples had no effect on the overall observed nucleotide frequency spectra. However, an overall decrease of LD in *P. mugo* was likely due to combining samples

with different demographic histories, as both the decay of linkage disequilibrium and the relationship between recombination and mutation parameter estimates in regional groups of the species (Table S6) were similar to the values observed for *P. uncinata* and *P. uliginosa* (Table 2). The signature of population expansion, marked by an excess of low frequency derived mutations and the presence of unique polymorphic sites and haplotypes, was found in most *P. mugo* populations. Large amounts of private variation were also reported in populations of *P. mugo* and other taxa from the *P. mugo* complex at chloroplast microsatellite loci (Dzialuk *et al.*, 2009; Heuertz et al., 2010). Regardless of the timing of the expansion event, it seems that the signal of population growth has persisted in *P. mugo* populations despite recent (late Holocene) range contraction.

GENETIC STRUCTURE AT INTRA-SPECIFIC LEVEL

Our nuclear gene study suggests that there is little population substructure within taxa. In *P. mugo*, evidence of population differentiation was found between Central European and Carpathian and Alpine populations, and also between Balkan and Apennine populations. As discussed previously, some evidence for sampling across populations with different demographic histories was found, namely the difference in patterns of nucleotide diversity among regional groups relative to species wide variation (Tables 2, S6). Some heterogeneity among Carpathian populations of *P. mugo* were found in biometric investigations of needles and cones (Boratyńska, Marcysiak & Boratyński, 2005). This study suggested a common origin of *P. mugo* from central and southern Europe including one Carpathian and two Sudeten populations and a very distant and isolated population from the Abruzzian Apennines (Boratyńska et al., 2005). No significant interpopulation differentiation was reported at allozyme loci in Bulgarian populations of *P. mugo* (Slavov & Zhelev, 2004) and four *P. mugo* populations from Austria and France at *cp*SSRs loci (Heuertz et al., 2010). In our sample of *P. uncinata*, only two populations from Central Pyrenees were distinct from other populations. However, the marginal southernmost population from Sierra de Gúdar, which was one of the most differentiated *P. uncinata* populations at *cp*SSR loci (Dzialuk et al., 2009, Heuertz et al., 2010) showed no significant differentiation at the nuclear loci studied. In other *P. uncinata* studies, low differentiation among some Pyreneean populations was reported at allozyme loci (Lewandowski et al., 2000) and sclerenchyma cell types of needles (Boratyńska & Boratyński, 2007).

Some evidence for sub-structuring was evident in the *mt*DNA data. All *P. mugo* and two *P. uliginosa* populations were fixed for, or contained, mitotype 1 (Figure 1, Table S1). This mitotype was present in only one population of *P. uncinata* species that was otherwise fixed for mitotype 3. This suggests a possible hybrid origin for this *P. uncinata* population from Italy as it grows in close proximity to *P. mugo* and both species are known to hybridize in contact zones in the Alps (Christensen, 1987a; Businský & Kirschner, 2006; Marcysiak & Boratyński, 2007). Alternatively, the mitotype 1 may simply be present in *P. uncinata* at low frequency. Therefore, considering the maternal inheritance and seed mediated dispersal of mitochondrial genomes in pines, denser sampling of *P. uncinata* would be needed to confirm whether mitotype 1 is species-diagnostic for *P. mugo* vs. *P. uncinata* and if, together with mitotype 3, could be used in hybridization studies in contact zones of the species. Similarly, mitotype 2 was found only in three *P. mugo* populations from central and eastern part of the species distribution and in one *P. uliginosa* population, which suggests independent demographic and/or evolutionary histories for these populations. However, a larger number of mitochondrial markers would be required to detect population substructure and reconstruct postglacial recolonisation routes.

CONCLUSIONS

Our study shows very high similarity in genetic variation between morphologically distinct pine species from the *P. mugo* complex. Lack of clear assignment of samples to species or geographical regions suggests a common evolutionary history and recent divergence but also potential phenotypic differentiation of the taxa due to likely limited interspecific gene flow. Larger numbers of markers are needed to improve resolution of genetic differences among species and populations. High morphological and ecological differentiation between the species, despite low genetic differentiation, suggests that natural selection has likely acted on very limited areas of the genome. Therefore this pine species complex will provide a valuable model for ongoing studies of the genetic basis of adaptive variation and speciation.

ACKNOWLEDGMENTS

This work was financially supported by the Polish National Science Centre (grant nr. 2011/01/B/NZ8/01634) and NERC (grant nr. NE/H003959/1).

CONFLICT OF INTEREST STATEMENT

The authors declare that there are no conflicts of interest.

REFERENCES

Beaulieu J, Perron M, Bousquet J. 2004. Multivariate patterns of adaptive genetic variation and seed source transfer in *Picea mariana*. *Canadian Journal of Forest Research* **34:** 531-545.

Benito Garzon M, Sanchez de Dios R, Sainz Ollero H. 2007. Predictive modelling of tree species distributions on the Iberian Peninsula during the Last Glacial Maximum and Mid-Holocene. *Ecography* **30:** 120-134.

Blarquez O, Carcaillet C, Bremond L, Mourier B, Radakovitch O. 2010. Trees in the subalpine belt since 11 700 cal. BP: origin, expansion and alteration of the modern forest. *Holocene* **20:** 139-146.

Bogunić F, Siljak-Yakovlev S, Muratović E, Pustahija F, Medjedović S. 2011. Molecular cytogenetics and flow cytometry reveal conserved genome organization in *Pinus mugo* and *P. uncinata*. *Annals of Forest Science* **68:** 179-187.

Boratyńska K, Bobowicz MA. 2001. *Pinus uncinata* Ramond taxonomy based on needle characters. *Plant Systematics and Evolution* **227:** 183-194.

Boratyńska K, Boratyński A. 2007. Taxonomic differences among closely related pines *Pinus sylvestris*, *P. mugo*, *P. uncinata*, *P. rotundata* and *P. uliginosa* as revealed in needle sclerenchyma cells. *Flora* **202:** 555-569.

Boratyńska K, Marcysiak K, Boratyński A. 2005. *Pinus mugo* (*Pinaceae*) in the Abruzzi Mountains: high morphological variation in isolated populations. *Botanical Journal of the Linnean Society* **147:** 309-316.

Brown GR, Gill GP, Kuntz RJ, Langley CH, Neale DB. 2004. Nucleotide diversity and linkage disequilibrium in loblolly pine. *Proceedings of the National Academy of Sciences of the United States of America* **101:** 15255-15260.

Burga CA. 1988. Swiss vegetation history during last 18,000 years. *New Phytologist* **110:** 581-602.

Businský R. 1998. *Pinus mugo* agg. in former Czechoslovakia - taxonomy, distribution, hybrid populations and endangering. *Zprávy Československé Botanické Společnosti. Praha* **33:** 29-52.

Businsky´ R, Kirschner J. 2006. Nomenclatural notes on the *Pinus mugo* complex in Central Europe. *Phyton***:** 129–139.

Cheddadi R, Vendramin G, Litt T, François L, Kageyama M, Lorentz S, Laurent J, Beaulieu J, Sadori L, Jost A, Lunt D. 2006. Imprints of glacial refugia in the modern genetic diversity of *Pinus sylvestris*. *Global Ecology and Biogeography* **15:** 271-282.

Christensen KI. 1987a. A morphometric study of the *Pinus mugo* Turra complex and its natural hybridization with *P. sylvestris* L. (*Pinaceae*). *Feddes Repertorium* **98:** 623-635.

Christensen KI. 1987b. Taxonomic revision of the *Pinus mugo* complex and *P. x rhaetica* (*P. mugo* x *sylvestris*) (*Pinaceae*). *Nordic Journal of Botany***:** 383-408.

Clement M, Posada D, Crandall KA. 2000. TCS: a computer program to estimate gene genealogies. *Molecular Ecology* **9:** 1657-1659.

Corander J, Tang J. 2007. Bayesian analysis of population structure based on linked molecular information. *Mathematical Biosciences* **205:** 19-31.

Coyne JA, Orr HA 2004. *Speciation* Sunderland, Massachusetts: Sinauer Associates Inc.

588 **Critchfield WB, Little E. 1966.** Geographic distribution of the Pines of the world. U.S. Department of
589 Agriculture.
590 **Dzialuk A, Muchewicz E, Boratyński A, Montserrat J, Boratyńska K, Burczyk, J. 2009.** Genetic
591 variation of *Pinus uncinata* (*Pinaceae*) in the Pyrenees determined with *cp*SSR markers. *Plant*
592 *Systematics and Evolution* **277:** 197-205.
593 **Eveno E, Collada C, Guevara MA, Leger V, Soto A, Diaz L, Leger P, Gonzalez-Martinez SC, Cervera**
594 **MT, Plomion C, Garnier-Gere PH. 2008.** Contrasting patterns of selection at *Pinus pinaster*
595 Ait. drought stress candidate genes as revealed by genetic differentiation analyses.
596 *Molecular Biology and Evolution* **25:** 417-437.
597 **Excoffier L. 2004.** Patterns of DNA sequence diversity and genetic structure after a range expansion:
598 lessons from the infinite-island model. *Molecular Ecology* **13:** 853-864.
599 **Excoffier L, Laval G, Schneider S. 2005.** Arlequin ver. 3.0: An integrated software package for
600 population genetics data analysis. *Evolutionary Bioinformatics Online* **1:** 47-50.
601 **González-Martinez SC, Ersoz E, Brown GR, Wheeler NC, Neale DB. 2006.** DNA sequence variation
602 and selection of tag single-nucleotide polymorphisms at candidate genes for drought-stress
603 response in *Pinus taeda* L. *Genetics* **172:** 1915-1926.
604 **Hamerník J, Musil I. 2007.** The *Pinus mugo* complex – its structuring and general overview of the
605 used nomenclature. *Journal of Forest Science* **53:** 253-266.
606 **Heuertz M, De Paoli E, Kallman T, Larsson H, Jurman I, Morgante M, Lascoux M, Gyllenstrand N.**
607 **2006.** Multilocus Patterns of Nucleotide Diversity, Linkage Disequilibrium and Demographic
608 History of Norway Spruce [*Picea abies* (L.) Karst]. *Genetics* **174:** 2095-2105.
609 **Heuertz M, Teufel J, González-Martínez SC, Soto A, Fady B, Alía R, Vendramin GG. 2010.** Geography
610 determines genetic relationships between species of mountain pine *Pinus mugo* complex in
611 western Europe. *Journal of Biogeography* **37:** 541-556.
612 **Hill WG, Robertson A. 1968.** Linkage disequilibrium in finite populations. *Theoretical and Applied*
613 *Genetics* **38:** 226-231.
614 **Howe GT, Aitken SN, Neale DB, Jermstad KD, Wheeler NC, Chen THH. 2003.** From genotype to
615 phenotype: unraveling the complexities of cold adaptation in forest trees. *Canadian Journal*
616 *of Botany* **81:** 1247-1266.
617 **Jaramillo-Correa JP, Beaulieu J, Bousquet J. 2004.** Variation in mitochondrial DNA reveals multiple
618 distant glacial refugia in black spruce (*Picea mariana*), a transcontinental North American
619 conifer. *Molecular Ecology* **13:** 2735-2747.
620 **Jasińska AK, Wachowiak W, Muchewicz E, Boratyńska K, Montserrat JM, Boratyński A. 2010.**
621 Cryptic hybrids between *Pinus uncinata* and *P. sylvestris*. *Botanical Journal of the Linnean*
622 *Society* **163:** 473-485.
623 **Jiggins CD, Salazar C, Linares M, Mavarez J. 2008.** Hybrid trait speciation and *Heliconius* butterflies.
624 *Philosophical Transactions of the Royal Society B: Biological Sciences* **363:** 3047-3054.
625 **Joosen RVL, Lammers M, Balk PA, Bronnum P, Konings MCJM, Perks M, Stattin E, van Wordragen**
626 **MF, van der Geest AHM. 2006.** Correlating gene expression to physiological parameters and
627 environmental conditions during cold acclimation of *Pinus sylvestris*, identification of
628 molecular markers using cDNA microarrays. *Tree Physiol* **26:** 1297-1313.
629 **Kinloch BB, Parks GK, Fowler CW. 1970.** White Pine Blister Rust: Simply Inherited Resistance in
630 Sugar Pine. *Science* **167:** 193-195.
631 **Kormutak A, Vookova B, Manka P, Salaj J, Camek V, Gömöry D. 2008.** Abortive embryogenesis in
632 hybrid swarm populations of *Pinus sylvestris* L. and *Pinus mugo* Turra. *Trees - Structure and*
633 *Function* **22:** 657-662.
634 **Krutovsky KV, Neale DB. 2005.** Nucleotide diversity and linkage disequilibrium in cold-hardiness-
635 and wood quality-related candidate genes in Douglas fir. *Genetics* **171:** 2029-2041.

Lewandowski A, Boratyński A, Mejnartowicz L. 2000. Allozyme investigations on the genetic differentiation between closely related pines *Pinus sylvestris*, *P. mugo*, *P. uncinata* and *P. uliginosa* (*Pinaceae*). *Plant Systematics and Evolution* **221:** 15-24.

Lewandowski A, Wisniewska M. 2006. Crossability between *Pinus uliginosa* and its putative parental species *Pinus sylvestris* and *Pinus mugo*. *Silvae Genetica* **55:** 52-54.

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

Marcysiak K. 2004. Interpopulational variability of *Pinus uncinata* Ramond ex DC. in Lam. and DC. (*Pinaceae*) cone characters. *Dendrobiology* **51:** 41-49.

Marcysiak K, Boratyński A. 2007. Contribution to the taxonomy of *Pinus uncinata* (*Pinaceae*) based on cone characters. *Plant Systematics and Evolution* **264:** 57-73.

Monteleone I, Ferrazzini D, Belletti P. 2006. Effectiveness of neutral RAPD markers to detect genetic divergence between the subspecies *uncinata* and *mugo* of *Pinus mugo* Turra. *Silva Fennica* **40:** 391-406.

Naydenov K, Senneville S, Beaulieu J, Tremblay F, Bousquet J. 2007. Glacial vicariance in Eurasia: mitochondrial DNA evidence from Scots pine for a complex heritage involving genetically distinct refugia at mid-northern latitudes and in Asia Minor. *BMC Evolutionary Biology* **7:** 233.

Nei M. 1987. Molecular Evolutionary Genetics. Columbia University Press, New York. ed.

Nicholas K, Nicholas H. 1997. Genedoc: A tool for editing and annotating multiple sequence alignments. *Distributed by the authors.*

Petit RJ, Aguinagalde I, de Beaulieu J-L, Bittkau C, Brewer S, Cheddadi R, Ennos R, Fineschi S, Grivet D, Lascoux M, Mohanty A, Muller-Starck G, Demesure-Musch B, Palme A, Martin JP, Rendell S, Vendramin GG. 2003. Glacial Refugia: Hotspots But Not Melting Pots of Genetic Diversity. *Science* **300:** 1563-1565.

Pyhäjärvi T, Garcia-Gil MR, Knürr T, Mikkonen M, Wachowiak W, Savolainen O. 2007. Demographic History Has Influenced Nucleotide Diversity in European *Pinus sylvestris* Populations. *Genetics* **177:** 1713-1724.

Pyhäjärvi T, Salmela MJ, Savolainen O. 2008. Colonization routes of *Pinus sylvestris* inferred from distribution of mitochondrial DNA variation. *Tree Genetics & Genomes* **4:** 247-254.

Ramil-Rego P, Munoz-Sobrino C, Rodriguez-Guitian CM, Gomez-Orellana L. 1998. Differences in the vegetation of the North Iberian Peninsula during the last 16, 000 years. *Plant Ecology* **138:** 41-62.

Slavov GT, Zhelev P. 2004. Allozyme variation, differentiation, and inbreeding in populations of *Pinus mugo* in Bulgaria. *Canadian Journal of Forest Research* **34:** 2611-2617.

Soranzo N, Alia R, Provan J, Powell W. 2000. Patterns of variation at a mitochondrial sequence-tagged-site locus provides new insights into the postglacial history of European *Pinus sylvestris* populations. *Molecular Ecology* **9:** 1205-1211.

Tajima F. 1989. Statistical-method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* **123:** 585-595.

Tamura K, Dudley J, Nei M, Kumar S. 2007. MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) Software Version 4.0. *Molecular Biology and Evolution* **24:** 1596-1599.

Thornton K. 2005. Recombination and the properties of Tajima's D in the context of approximate likelihood calculation. *Genetics* **171:** 2143-2148

Wachowiak W, Balk P, Savolainen O. 2009. Search for nucleotide diversity patterns of local adaptation in dehydrins and other cold-related candidate genes in Scots pine (*Pinus sylvestris* L.). *Tree Genetics & Genomes* **5:** 117-132.

Wachowiak W, Palme AE, Savolainen O. 2011a. Speciation history of three closely related pines *Pinus mugo* (T.), *P. uliginosa* (N.) and *P. sylvestris* (L.). *Molecular Ecology* **20:** 1729-1743.

Wachowiak W, Prus-Głowacki W. 2008. Hybridisation processes in sympatric populations of pines *Pinus sylvestris* L., *P. mugo* Turra and *P. uliginosa* Neumann. *Plant Systematics and Evolution* **271:** 29-40.

Wachowiak W, Prus-Głowacki W. 2009. Different patterns of genetic structure of relict and isolated populations of endangered peat-bog pine (*Pinus uliginosa* Neumann). *Journal of Applied Genetics* **50(4):** 329-339.

Wachowiak W, Salmela MJ, Ennos RA, Iason G, Cavers S. 2011b. High genetic diversity at the extreme range edge: nucleotide variation at nuclear loci in Scots pine (*Pinus sylvestris* L.) in Scotland. *Heredity* **106:** 775-787.

Watterson GA. 1975. On the number of segregating sites in genetical models without recombination. *Theor Popul Biol* **7:** 256 - 276.

Weir BS, Cockerham CC. 1984. Estimating F-Statistics for the Analysis of Population Structure. *Evolution* **38:** 1358-1370.

Figure Legend:

Figure 1.

Geographical location of the analysed populations of the pine species from the *Pinus mugo* complex and distribution of mitotypes detected at the *mt*DNA regions (see Table S1 and text for details).

[Text under the Figure 1]

Colours represent different mitotypes detected based on the presence/absence of 5bp and 31bp insertion/deletion in the *nad7.1* and *nadHI mt*DNA regions, respectively. Mitotype 1 (green): 5bp insertion and 31bp deletion; mitotype 2 (yellow): 5bp and 31bp deletions, mitotype 3 (red): 5bp and 31bp insertions, respectively. Sizes of circles correspond to number of samples analysed in each population ranging from 12 to 4 (UG2).

Figure 2.

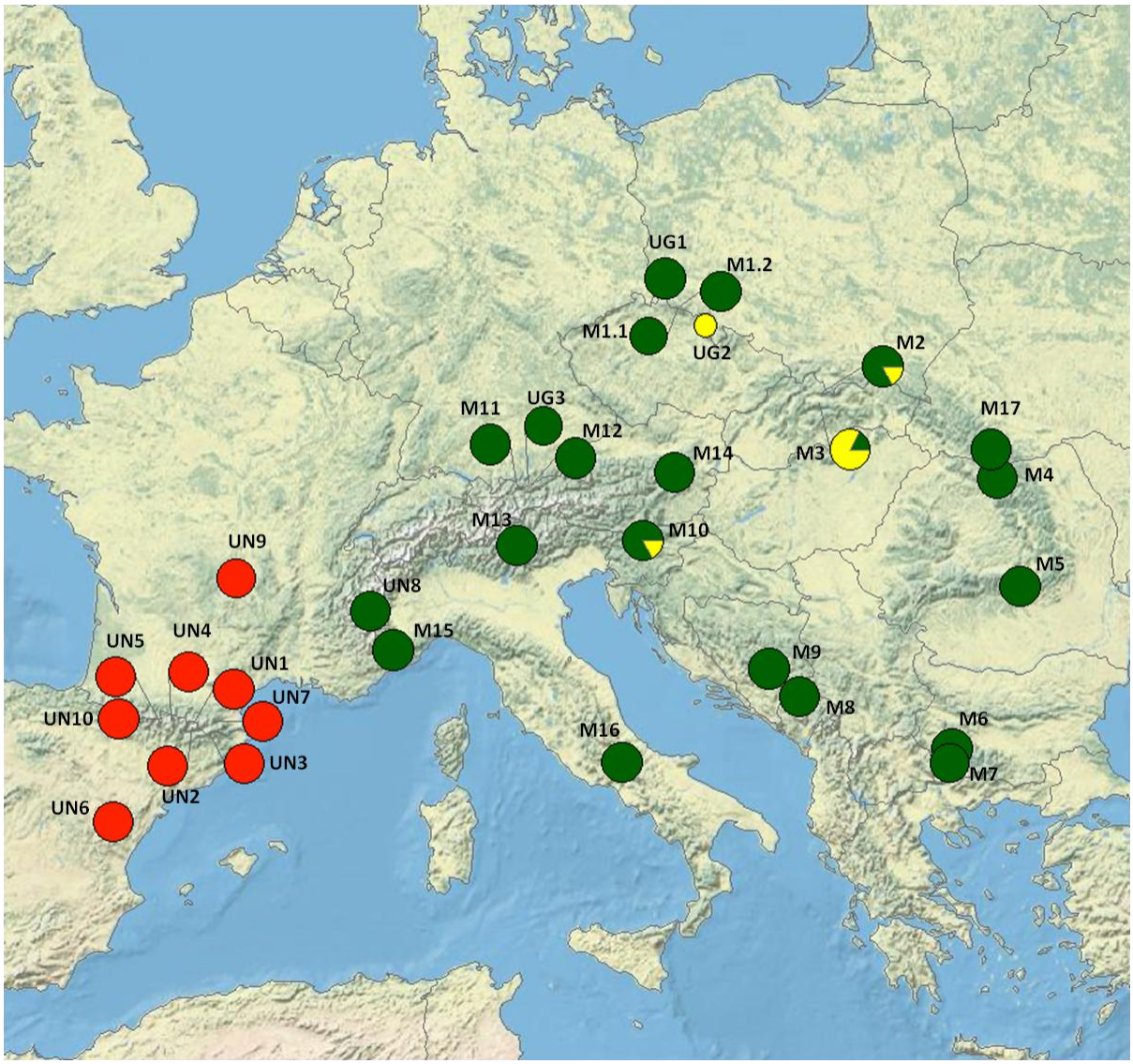
Distribution of polymorphic sites detected at the nuclear loci in pairwise comparisons between species. M – *P. mugo*, PUN – *P. uncinata*, PUG – *P. uliginosa*.

Figure 3.

Posterior densities of multilocus θ (per base pair) for the analysed taxa at 12 nuclear loci. Credibility intervals are reported in Table 2.

Figure 1.

Geographical location of the analysed populations of the pine species from the *Pinus mugo* complex and distribution of mitotypes detected at the *mt*DNA regions (see Table S1 and text for details).



Colours represent different mitotypes detected based on the presence/absence of 5bp and 31bp insertion/deletion in the *nad7.1* and *nadHI mt*DNA regions, respectively. Mitotype 1 (green): 5bp insertion and 31bp deletion; mitotype 2 (yellow): 5bp and 31bp deletions, mitotype 3 (red): 5bp and 31bp insertions, respectively. Sizes of circles correspond to number of samples analysed in each population ranging from 12 to 4 (UG2).

Figure 2.
Distribution of polymorphic sites detected at the nuclear loci in pairwise comparisons between species. M – *P. mugo*, PUN – *P. uncinata*, PUG – *P. uliginosa*.

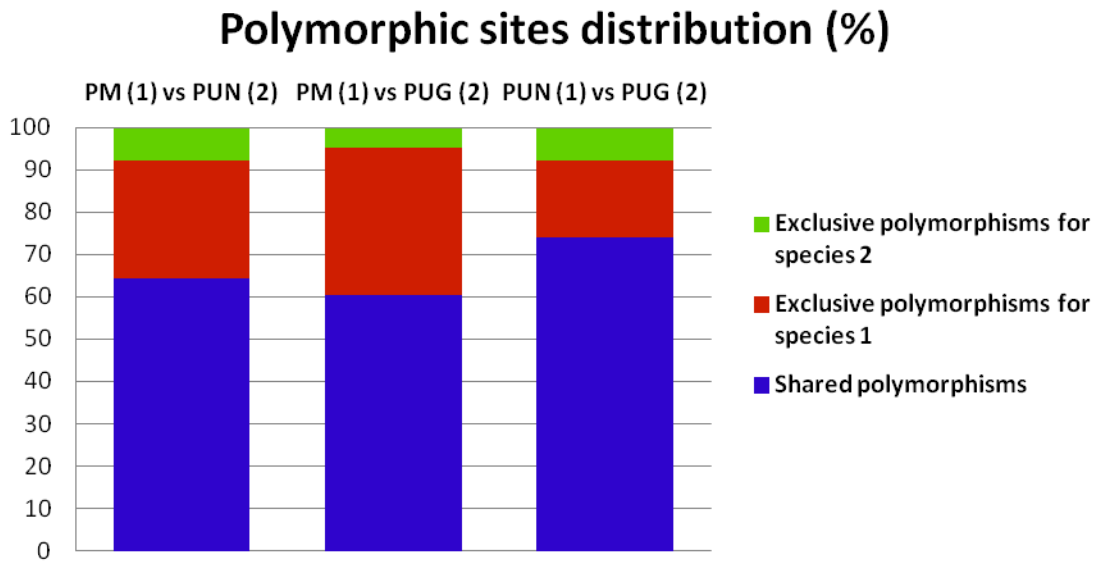
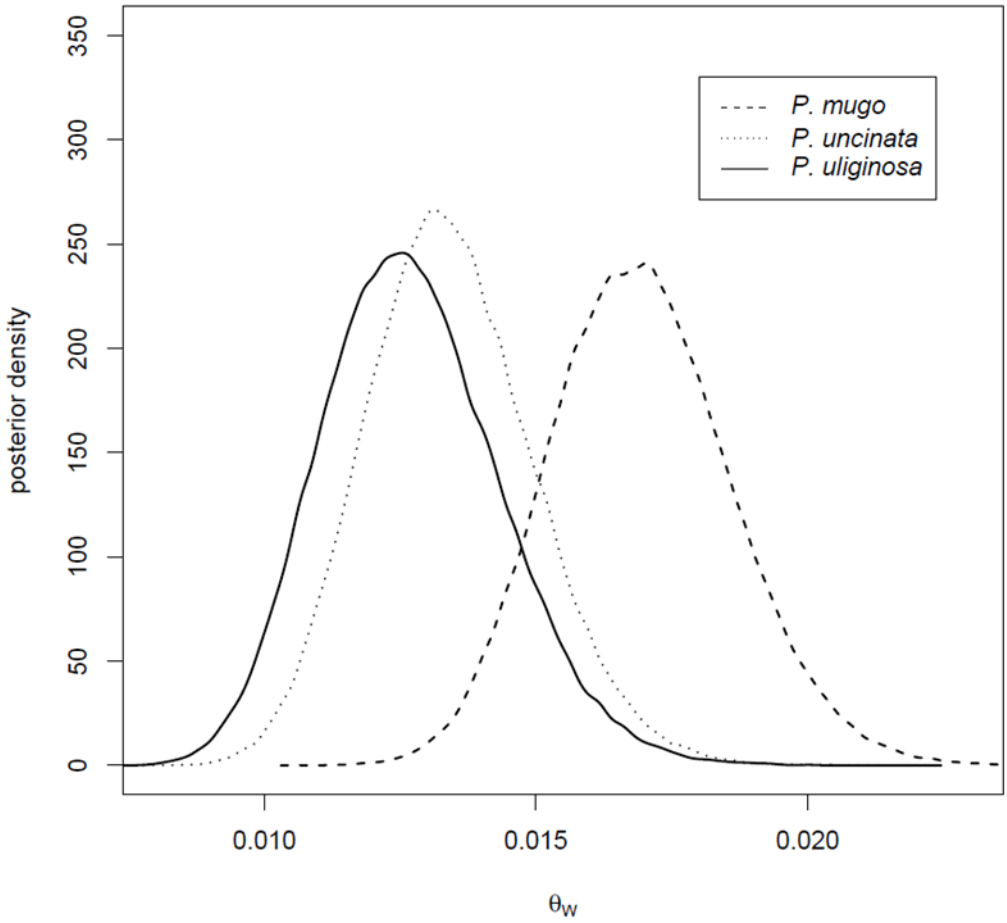


Figure 3.
Posterior densities of multilocus θ (per base pair) for the analysed taxa at 12 nuclear loci. Credibility intervals are reported in Table 2.



Tables:

Table 1. Average nucleotide and haplotype diversity and divergence at 12 nuclear loci

				Nucleotide diversity							Haplotype diversity		Divergence	
Species	N	L_{total}	S_{total}	π_{total}	L_{silent}	$S_{sil.}$	π_{silent}	S_{nonsyn}	π_{nonsyn}	*D*	**Ne**	H_d (SD)	$\mathbf{K_{total}}$	$\mathbf{K_{silent}}$
P. mugo	**1362**	**4851**	**321 (108)**	**0.0118**	**2351**	**17.3**	**0.0185**	**9.5**	**0.0070**	**-0.681**	**289**	**0.730 (0.026)**	**0.043**	**0.051**
P. uncinata	**747**	**4851**	**213 (41)**	**0.0113**	**2350**	**12.3**	**0.0178**	**5.5**	**0.0058**	**-0.118**	**150**	**0.687 (0.042)**	**0.044**	**0.053**
P. uliginosa	**404**	**4851**	**173 (65)**	**0.0091**	**2379**	**9.16**	**0.0141**	**4.3**	**0.0048**	**-0.266**	**96**	**0.623 (0.055)**	**0.041**	**0.050**

N – total number of sequences analysed; L – total length of sequence in base pairs excluding indels; S- total number of polymorphic sites detected (number of singleton mutations); π – nucleotide diversity (Nei 1987); *D* – Tajima's *D* test (Tajima 1989); Ne – total number of haplotypes; H_d – haplotype diversity (standard deviation); K – average pairwise divergence per site to the outgroup *P. pinaster*

Table 2. Nucleotide diversity estimates at 12 nuclear loci

Species	**θ[a]**	**C.I. (95%)[b]**	**ρ [c]**	**ρ / θ[d]**
P. mugo	0.0169	0.0139 - 0.0204	0.0369 (0.0028)	2.18
P. uncinata	0.0134	0.0106 - 0.0166	0.0054 (0.0011)	0.403
P. uliginosa	0.0126	0.0098 - 0.0162	0.0051 (0.0010)	0.405

[a] median for silent sites;
[b] 95% credibility intervals for θ;
[c] recombination rate parameter ρ (standard error in parenthesis),
[d] the ration between least-squares estimate of ρ and silent θ;

Supporting Information

Table S1. Geographic origin of the samples and corresponding haplotypes at the two *mt*DNA fragments found in each population.

No	Species	Population	Acronym	Latitude N	Longitude E	Altitude	*Mitotype*[1]
1	*P. mugo*	Poland, Sudety Mts, Czarny Kocioł Jagniątkowski	M1.1	50^{0}47’05’’	15^{0}35’30’’	1350	1
2	*P. mugo*	Poland, Sudety Mts, Śląskie Kamienie	M1.2	50°46′35″	15°36′08″	1400	1
3	*P. mugo*	Poland, Tatra Mts., Wołowiec	M2	49^{0}13’07’’	19^{0}45’50’’	1620	1,2
4	*P. mugo*	Slovakia, Lower TatraMts., Łysa	M3	49^{0}00’43’’	19^{0}39’35’’	1500	1,2
5	*P. mugo*	Romania, Easter Carpathians, Prislop Pass	M4	47^{0}34’03’’	24^{0}48’00’’	1720	1
6	*P. mugo*	Romania, Southern Carpathians, Busteni	M5	45^{0}25’55’’	25^{0}27’06’’	2070	1
7	*P. mugo*	Bulgaria, Rila Mts, Belitsa	M6	42^{0}04’01’’	23^{0}30’00’’	2100	1
8	*P. mugo*	Bulgaria, Pirin Mts.,Vikhren	M7	41^{0}46’07’’	23^{0}25’22’’	2000	1
9	*P. mugo*	Montenegro, Durmitor Mts., Zabljak	M8	43^{0}09’33’’	19^{0}05’27’’	2100	1
10	*P. mugo*	Bosnia and Herzegovina, Bjelasnica Mts.	M9	43^{0}45’00’’	18^{0}13’08’’	2120	1
11	*P. mugo*	Slovenia, Kamnik – Savinja Alps	M10	46^{0}21’10’’	14^{0}34’43’’	1600	1,2
12	*P. mugo*	Germany, Bavarian Alps, Kreuzspitze	M11	47^{0}31’30’	10^{0}55’12’’	1870	1
13	*P. mugo*	Austria, Karwendel Mts., Scharnitz	M12	47^{0}22’42’’	11^{0}17’45’’	1400	1
14	*P. mugo*	Italy, Prealpi Venete, Lago di Tovel	M13	46^{0}15’39’’	10^{0}56’46’’	1200	1
15	*P. mugo*	Austria, Steirisch-Niederoserreichische, Turntaler-Kogel	M14	47^{0}41’18’’	15^{0}29’32’’	1600	1
16	*P. mugo*	Italy, Alps, Col de Tende	M15	44^{0}08’00’’	7^{0}22’30’’	2000	1
17	*P. mugo*	Italy, Abruzzi, La Maiella	M16	41^{0}46’20”	13^{0}58’30’’	2200	1
18	*P. mugo*	Ukraine, Carpathians, Charnokhora	M17	48^{0}08’00’’	24^{0}37’30’’	1250	1
19	*P. uncinata*	Andorra, Eastern Pyrenees, Vall de Ransol	UN1	42^{0}35’02’’	1^{0}38’21’’	2025	3
20	*P. uncinata*	Andorra, Eastern Pyrenees, San Miguel de Engolasters	UN2	42^{0}31’28’’	1^{0}34’12’’	2000	3
21	*P. uncinata*	Spain, Eastern Pyrenees, Vall de Nuria	UN3	42^{0}20’45’’	2^{0}06’15’’	2200	3
22	*P. uncinata*	Spain, Central Pyrenees, Port de la Boniagua	UN4	42^{0}39’48’’	0^{0}57’44’’	2100	3
23	*P. uncinata*	Spain, Western Pyrenees, Castiello de Jaca	UN5	42^{0}41’19’’	-0^{0}32’12’’	1720	3
24	*P. uncinata*	Spain, Sierra de Gudar	UN6	40^{0}28’49’’	-0^{0}41’51’	2000	3
25	*P. uncinata*	France, Eastern Pyrenees, Col de Jau	UN7	42^{0}39’19’’	2^{0}15’22’’	1520	3
26	*P. uncinata*	Italy, Alps, Claviere	UN8	44^{0}56’20’’	6^{0}44’00’’	1300	1
27	*P. uncinata*	France, Col de la Croix de Morand	UN9	45^{0}35’58’’	2^{0}50’44’’	1200	3
28	*P. uncinata*	Spain, Benasque	UN10	42^{0}37’19’’	0^{0}39’51’’	1450	3
29	*P. uliginosa*	Poland, Low Silesian Pinewood, Węgliniec	UG1	51^{0}17’50’’	15^{0}14’20’’	190	1
30	*P. uliginosa*	Poland, Wielkie Torfowisko Batorowskie reserve	UG2	50^{0}27’32’	16^{0}23’01’’	750	2*
31	*P. uliginosa*	Germany, Mittenwald	UG3	47^{0}28’50’’	11^{0}16’27’’	856	1

[1] – mitotypes detected based on the presence/absence of 5bp and 31bp insertion/deletion in the *nad7.1* and *nad1HI mt*DNA regions, respectively. Mitotype 1: 5bp insertion and 31bp deletion; mitotype 2: 5bp and 31bp deletions, mitotype 3: 5bp and 31bp insertions, respectively.

* only four samples from that population were successfully amplified at *nad7.1* gene fragment.

Table S2. Nuclear loci included in the nucleotide diversity analyses and PCR amplification conditions.

Gene	**Protein / Function**	**Ta(°C)**	**[$MgCl_2$]**
a3ip2	ABI3-interacting protein 2	61	1.5
abaR	abscisic acid responsive protein	53	1.5
ccoaomt	caffeoyl CoA *O*-methyltransferase	60	1.5
chcs	chalcone synthase	61	1.5
dhn2	dehydrin 2 – dehydrative stress response	64	1.5
dhy2PP	dehydrin – dehydrative stress response	61	1.5
dhn3	dehydrin 3 – dehydrative stress response	60	1.5
dhn7	dehydrin 7 – dehydrative stress response	56	1.5
dhn9	dehydrin 9 – dehydrative stress response	54	1.5
erd3	early responsive to dehydration 3	57	1.5
lp3-1	ABA and WDS induced gene-1	61	1.5
lp3-3	ABA and WDS induced gene-3	63	1.5
nad1EG	NADH dehydrogenase subunit 1 intron 1	60	1.5
nad7.1	NADH dehydrogenase subunit 7 intron 1	55	1.5

Table S3. Summary statistics of nucleotide and haplotype variation and frequency distribution spectrum of polymorphism at analysed nuclear loci in the three taxa of the *P. mugo* complex.

			Nucleotide diversity									**Haplotype diversity**		**Divergence**	
Locus	*Species*	N	L_{total}	S_{total}	π_{total}	L_{silent}	$S_{sil.}$	π_{silent}	S_{nonsyn}	π_{nonsyn}	*D*	Ne	H_d (SD)	$\mathbf{K}_{total}$	$\mathbf{K}_{silent}$
a3ip2	*P. mugo*	162	304	8 (8)	0.0003	295.0	8	0.0003	-	-	-2.107*	7	0.073 (0.028)	0.0217	0.0217
	P. uncinata	82	304	4(2)	0.0010	295.0	5	0.0010	-	-	-1.586	6	0.247 (0.062)	0.0221	0.0221
	P. uliginosa	24	304	2 (0)	0.0015	297.0	2	0.0015	-	-	-0.354	3	0.42 (0.0121)	0.0223	0.0223
abaR	*P. mugo*	169	406	24 (12)	0.0078	162.7	11	0.0119	13	0.0047	-0.842	26	0.873 (0.016)	0.0473	0.0417
	P. uncinata	78	406	14 (4)	0.0063	162.8	6	0.0103	8	0.0027	-0.652	15	0.839 (0.033)	0.0468	0.0421
	P. uliginosa	39	406	14 (4)	0.0064	164.7	6	0.0073	8	0.0057	-0.829	14	0.834 (0.052)	0.0454	0.0370
ccoaomt	*P. mugo*	121	385	14 (6)	0.0021	276.3	13	0.0028	1	0.0002	-1.860*	11	0.435 (0.054)	0.0188	0.0274
	P. uncinata	90	385	7 (3)	0.0012	273.3	7	0.0016	0	0.0000	-1.609	8	0.285 (0.062)	0.0186	0.0273
	P. uliginosa	45	385	5 (3)	0.0015	276.3	5	0.0020	0	0.0000	-1.252	5	0.311 (0.084)	0.0186	0.0271

chcs	*P. mugo*	151	306	23(11)	0.0141	148.4	16	0.0283	7	0.0007	0.129	26	0.866 (0.015)	0.0226	0.0469
	P. uncinata	83	306	17 (1)	0.0160	148.3	14	0.0306	4	0.0022	1.021	16	0.857 (0.019)	0.0260	0.0470
	P. uliginosa	45	306	14 (3)	0.0141	148.3	14	0.0290	0	0.0000	1.059	15	0.909 (0.025)	0.0225	0.0470
dhn2	*P. mugo*	65	535	25(8)	0.0102	231.9	21	0.0136	5	0.0059	-0.711	28	0.917 (0.023)	0.0401	0.0440
	P. uncinata	37	535	18(3)	0.0111	236.9	12	0.0144	6	0.0068	0.291	19	0.91 (0.031)	0.0400	0.0438
	P. uliginosa	11	535	16(5)	0.0125	242.9	12	0.0160	4	0.0078	-0.096	10	0.982 (0.046)	0.0403	0.0451
dhn2PP	*P. mugo*	153	426	27(10)	0.0072	169.3	18	0.0165	10	0.0011	-1.124	39	0.921(0.012)	0.0309	0.0243
	P. uncinata	93	426	19 (6)	0.0087	168.8	15	0.0197	4	0.0013	-0.054	18	0.894(0.015)	0.0323	0.0338
	P. uliginosa	45	426	13 (6)	0.0064	169.7	10	0.0145	3	0.0009	-0.203	11	0.877 (0.022)	0.0300	0.0285
dhn3	*P. mugo*	169	271	23(6)	0.0054	95.4	15	0.0064	8	0.0047	-1.820*	16	0.313(0.047)	0.0421	0.0581
	P. uncinata	83	271	18(5)	0.0067	95.4	12	0.0081	6	0.0059	-1.518	10	0.453(0.067)	0.0418	0.0578
	P. uliginosa	45	271	12(2)	0.0106	95.4	8	0.0128	5	0.0095	0.0896	3	0.275 (0.074)	0.0375	0.0524
dhn7	*P. mugo*	40	262	15(3)	0.0124	122.2	9	0.0218	6	0.0112	0.866	9	0.812(0.04)	0.0700	0.0764
	P. uncinata	16	262	12(2)	0.0181	124.2	7	0.0244	5	0.0133	1.664	5	0.767(0.066)	0.0676	0.0690
	P. uliginosa	28	262	12(2)	0.0156	122.2	6	0.0229	6	0.0101	1.453	6	0.646(0.074)	0.0700	0.0782
dhn9	*P. mugo*	61	569	63(22)	0.0244	189.4	28	0.0306	35	0.0210	-0.487	27	0.939(0.015)	0.1457	0.1572
	P. uncinata	46	569	39(4)	0.0160	188.4	22	0.0204	17	0.0136	-0.369	10	0.666(0.074)	0.1436	0.1509
	P. uliginosa	35	569	40(15)	0.0200	189.3	22	0.0267	18	0.0162	0.074	9	0.701 (0.072)	0.1436	0.1531
erd3	*P. mugo*	117	586	25(16)	0.0022	288.2	20	0.0042	5	0.0003	-2.120*	22	0.719(0.036)	0.0063	0.0090
	P. uncinata	49	586	3(2)	0.0010	289.2	3	0.0020	0	0.0000	-0.266	4	0.548(0.031)	0.0060	0.0087
	P. uliginosa	45	586	4(3)	0.0004	289.2	4	0.0007	0	0.0000	-1.764	5	0.207(0.079)	0.0053	0.0072
lp3-1	*P. mugo*	89	357	32(4)	0.0191	253.1	25	0.0205	6	0.0098	0.046	60	0.984(0.006)	0.0345	0.0429
	P. uncinata	48	357	28(6)	0.0195	246.4	25	0.0241	3	0.0058	0.057	29	0.965 (0.012)	0.0368	0.0485
	P. uliginosa	15	357	13(4)	0.0113	262.4	11	0.0139	2	0.0032	-0.092	10	0.924 (0.053)	0.0297	0.0389
lp3-3	*P. mugo*	65	444	31(2)	0.0367	118.7	23	0.0650	9	0.0171	1.853	18	0.909(0.018)	0.0368	0.0581
	P. uncinata	42	444	26(3)	0.0305	121.8	19	0.0565	7	0.0120	1.598	10	0.81(0.037)	0.0487	0.0852
	P. uliginosa	12	444	24(18)	0.0199	121.7	18	0.0349	6	0.0092	-1.199	5	0.667(0.141)	0.0612	0.1123
Average/ Total	***P. mugo***	**113.5**	4851	**25.8 (9)**	**0.0118**	2351	**17.3**	**0.0185**	**9.5**	**0.0070**	**-0.681**	**24.1**	**0.730 (0.026)**	**0.0431**	**0.0507**
	P. uncinata	**62.3**	4851	**17.1(3.4)**	**0.0113**	2350	**12.3**	**0.0178**	**5.5**	**0.0058**	**-0.118**	**12.5**	**0.687 (0.042)**	**0.0442**	**0.0530**
	P. uliginosa	**29.4**	4851	**13.1(5.4)**	**0.0091**	2379	**9.16**	**0.0141**	**4.3**	**0.0048**	**-0.266**	**7.8**	**0.623(0.055)**	**0.0407**	**0.0497**

N – number of individual samples (alleles) analysed; S – number of polymorphic sites detected (singletons); L – length of sequence in base pairs including indels; π – nucleotide diversity (Nei 1987); Ne– number of haplotypes; H_d – haplotype diversity (standard deviation); *D* test (Tajima 1989); $^*P<0.05$; K_s –pairwise divergence per site to the outgroup *P. pinaster* using Jukes-Cantor correction.

Table S4. Number of shared and exclusive polymorphisms and fixed differences among the three pine taxa at 12 nuclear loci.

	P. mugo – P. uncinata				***P. mugo – P. uliginosa***				***P. uncinata– P . uliginosa***			
	shared	**exclusive**		**fixed**	**shared**	**exclusive**		**fixed**	**shared**	**exclusive**		**fixed**
Locus		**S1**	**S2**			**S1**	**S2**			**S1**	**S2**	
a3ip2	1	7	3	0	1	7	1	0	2	2	0	0
abaR	11	13	4	0	11	13	4	0	10	5	5	0
ccoaomt	4	10	3	0	5	9	0	0	4	3	1	0
chcs	11	12	6	0	11	12	3	0	10	7	4	0
dhn2	17	11	3	0	15	13	2	0	14	6	3	0
dhn2PP	13	14	6	0	9	18	4	0	12	7	1	0
dhn3	15	9	4	0	12	12	1	0	12	7	1	0
dhn7	11	4	1	0	10	5	2	0	11	1	1	0
dhn9	36	27	3	0	35	28	5	0	34	5	6	0
erd3	2	23	1	0	3	22	0	0	1	2	2	0
lp3-1	24	9	4	0	11	22	2	0	11	17	2	0
lp3-3	27	10	3	0	26	11	0	0	22	8	4	0
Total	**172**	**149**	**41**	0	**149**	**172**	**24**		**143**	**70**	**30**	
%	**64.4**	**27.9**	**7.7**	**0**	**60.3**	**34.8**	**4.9**	**0**	**74.1**	**18.1**	**7.8**	**0**

* one SNP is a replacement change between hydrophobic Valine and Phenylalanine and the second one is in a noncoding

Table S5. Net divergence per base pair and silent divergence between the three pine taxa at 12 nuclear loci.

	Net divergence			**Silent divergence**		
Locus	***P. mugo* - *P. uncinata***	***P. mugo* - *P. uncinata***	***P. mugo* -*P. uliginosa***	***P. mugo* - *P. uncinata***	***P. mugo* - *P. uncinata***	***P. mugo* -*P. uliginosa***
a3ip2	0.0000	0.0000	0.0010	0.0007	0.0010	0.0012
abaR	0.0002	0.0002	0.0000	0.0117	0.0084	0.0072
ccoaomt	0.0000	0.0000	0.0000	0.0023	0.0021	0.0016
chcs	0.0012	0.0001	0.0020	0.0324	0.0296	0.0352
dhn2	0.0008	0.0005	0.0004	0.0154	0.0151	0.0154
dhn2PP	0.0003	0.0002	0.0001	0.0190	0.0161	0.0174
dhn3	0.0000	0.0005	0.0003	0.0073	0.0096	0.0104
dhn7	0.0007	0.0003	0.0007	0.0248	0.0231	0.0256
dhn9	0.0015	0.0002	0.0004	0.0283	0.0298	0.0247
erd3	0.0003	0.0001	0.0003	0.0038	0.0027	0.0020
lp3-1	0.0010	0.0010	0.0018	0.0229	0.0176	0.0223
lp3-3	0.0034	0.0087	0.0015	0.0682	0.0708	0.0508
Average	**0.0008**	**0.0010**	**0.0007**	**0.0197**	**0.0188**	**0.0178**

Table S6. Nucleotide diversity estimates at 12 nuclear loci for the 5 geographical groups of *P. mugo* studied

Groups	θ^a	C.I. (95%)[b]	ρ [c]	ρ / θ^d	D^e
Central Europe	0.0113	0.0083 - 0.0150	0.00036 (0.00057)	0.0317	-0.180
Carpathians	0.0141	0.0107 - 0.0186	0.00243 (0.00060)	0.1723	-0.288
Balkans	0.0128	0.0098 – 0.0166	0.00297 (0.00078)	0.2320	-0.523
Alps	0.0133	0.0017 - 0.0103	0.00520 (0.00096)	0.3910	-0.259
Apennines	0.0110	0.0017 - 0.0204	0.00052 (0.00079)	0.0473	-0.273

[a] median for silent sites;
[b] 95% credibility intervals for θ;
[c] recombination rate parameter ρ (standard error in parenthesis)
[d] the ration between least-squares estimate of ρ and silent θ
[e] *D* test (Tajima 1989);

Table S7. Nucleotide and haplotype variation and frequency distribution spectrum of polymorphism at seven nuclear loci in the analysed populations of taxa from the *P. mugo* complex. Populations acronyms as in Table 1.

Population	L_{total}	S_{total}	π_{total}	*D*	N	Population	L_{total}	S_{total}	π_{total}	*D*	N
M1_1	2636	25	0.0040	-0.441	20 (3)	M16	2634	29	0.0047	0.142	25 (3)
M1_2	2640	27	0.0036	-0.643	19 (1)	M17	2640	41	0.0084	0.253	24 (9)
M2	2638	28	0.0043	-0.427	22 (5)	UN1	2640	32	0.0045	-0.161	26 (3)
M3	2639	29	0.0055	-0.336	22 (2)	UN2	2641	33	0.0049	-0.245	29 (2)
M4	2636	39	0.0055	0.044	32 (10)	UN3	2639	31	0.0054	-0.204	21 (1)
M5	2635	30	0.0045	0.178	26 (6)	UN4	2056	46	0.0094	0.084	23 (3)
M6	2639	31	0.0054	0.574	24 (5)	UN5	2640	48	0.0068	-0.122	30 (8)
M7	2635	35	0.0054	0.321	31 (8)	UN6	2636	36	0.0049	0.043	27 (4)
M8	2637	26	0.0032	-0.166	25 (4)	UN7	2371	25	0.0053	0.574	17 (1)
M9	2631	30	0.0051	0.003	25 (3)	UN8	2049	33	0.0069	0.210	18 (3)
M10	2643	27	0.0038	-0.821	22 (0)	UN9	2637	30	0.0047	0.154	25 (2)
M11	2635	34	0.0058	-0.278	25 (2)	UN10	2055	34	0.0052	-0.409	16 (2)
M12	2640	46	0.0063	-0.299	36 (4)	UG1	2638	50	0.0062	-0.145	35 (5)
M13	2059	31	0.0064	-0.147	19 (5)	UG2	2641	36	0.0060	0.320	29 (4)
M14	2642	31	0.0048	-0.194	24 (6)	UG3	2254	34	0.0049	0.053	29 (3)
M15	2637	39	0.0062	-0.410	21 (4)						

L – length of sequence in base pairs including indels; S – number of polymorphic sites detected; π – nucleotide diversity (Nei 1987); *D* test (Tajima 1989); N – number of haplotypes including indels (unique haplotypes in parenthesis).

Table S8. Number of shared and exclusive haplotypes at 12 nuclear loci between the three pine taxa

	P. mugo – P. uncinata			***P. mugo – P. uliginosa***			***P. uncinata – P . uliginosa***		
	shared	**exclusive**		**shared**	**exclusive**		**shared**	**exclusive**	
Locus		**S1**	**S2**		**S1**	**S2**		**S1**	**S2**
a3ip2	2	5	4	2	5	1	3	3	0
abaR	9	17	7	9	17	6	8	8	7
ccoaomt	3	8	5	4	7	1	4	4	1
chcs	8	17	8	9	16	7	7	9	9
dhn2	4	17	12	2	19	7	0	16	9
dhn2PP	6	23	13	7	23	5	3	17	9
dhn3	5	10	5	2	13	1	2	8	1
dhn7	3	6	2	4	5	2	3	2	3
dhn9	6	21	4	6	21	3	6	4	3
erd3	2	20	2	4	18	0	2	2	2
lp3-1	7	54	22	2	59	8	1	28	9
lp3-3	8	15	5	5	18	0	2	11	3
Total	**63**	**213**	**89**	**56**	**221**	**41**	**41**	**112**	**56**
%	**29.3**	**50.0**	**20.7**	**30.0**	**59.0**	**11.0**	**32.8**	**44.8**	**22.4**

Table S9. Pairwise F_{ST} values at all polymorphic sites combined across seven nuclear loci between five geographical regions of *P. mugo*.

Group	1	2	3	4
1. Central Europe	-			
2. Carpathians	0.033*	-		
3. Balkans	0.019	-0.008	-	
4. Alps	0.041*	-0.006	0.018	-
5. Apennines	0.025	0.020	0.043*	-0.002

***p<0.05**

Table S10. F_{ST} values at individual loci in pairwise comparisons between the three pine taxa.

	P. mugo – P. uncinata	***P. mugo – P. uliginosa***	***P. uncinata – P . uliginosa***
a3ip2	0.034**	0.177***	0.009
abaR	0.028**	0.029*	0.000
ccoaomt	0.012	-0.005	0.002
chcs	0.084***	0.004	0.126***
dhn2	0.109***	0.010	0.092**
dhy2PP	0.033**	0.029**	0.009
dhn3	-0.002	0.023	0.006
dhn7	0.042	0.029	0.064
dhn9	0.056*	-0.001	0.034
erd3	0.153***	0.031*	0.271***
lp3-1	0.042**	0.035	0.072**
lp3-3	0.063*	0.219**	0.052

***p<0.05, ****p<0.01, *****p<0.001**

Figure S1. Scatter plot of the squared correlation coefficient of allele frequencies (r^2) as a function of distance in base pairs between pairs of polymorphic sites in *P. mugo* (A), *P. uncinata* (B) and *P. uliginosa* (C) at all nuclear loci combined. Decline in linkage disequilibrium is shown by nonlinear fitting curve of the mutation-recombination-drift model (see material and methods section for details). Recombination rate parameter ρ (standard error in parenthesis) for *P. mugo* is ρ= 0.0369 (0.0028), for *P. uncinata* is ρ =0.0054 (0.0011) and ρ = 0.0051 (0.0010) for *P. uliginosa.*

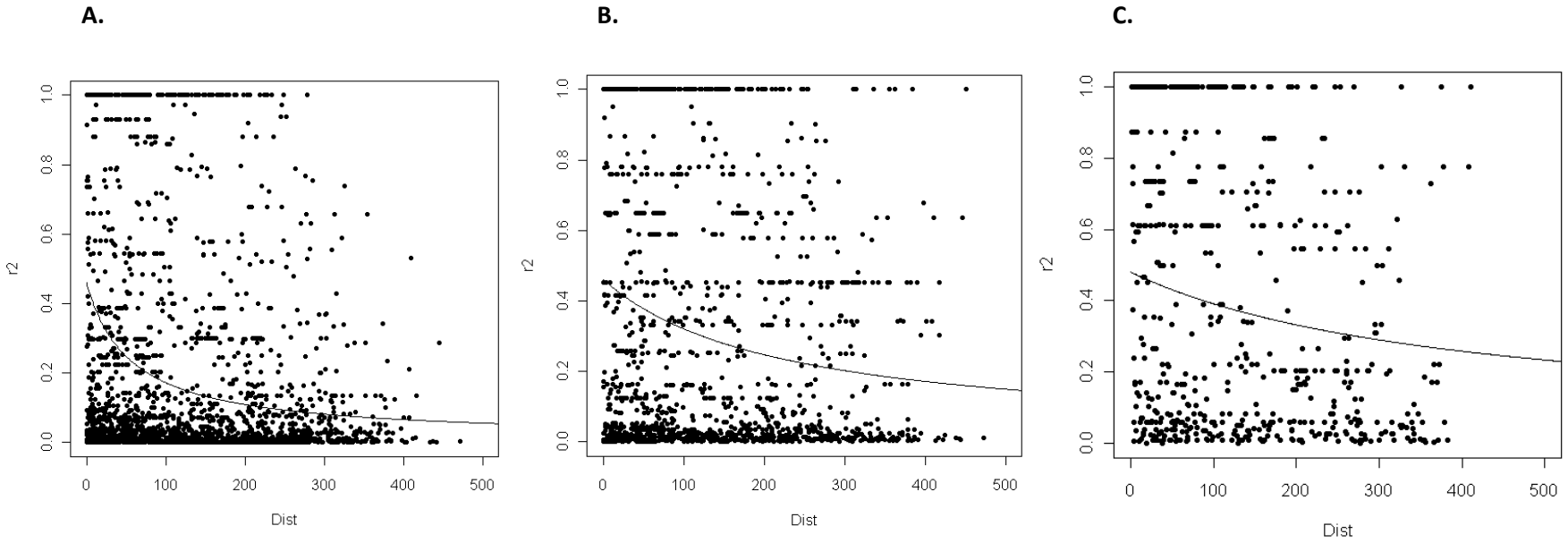
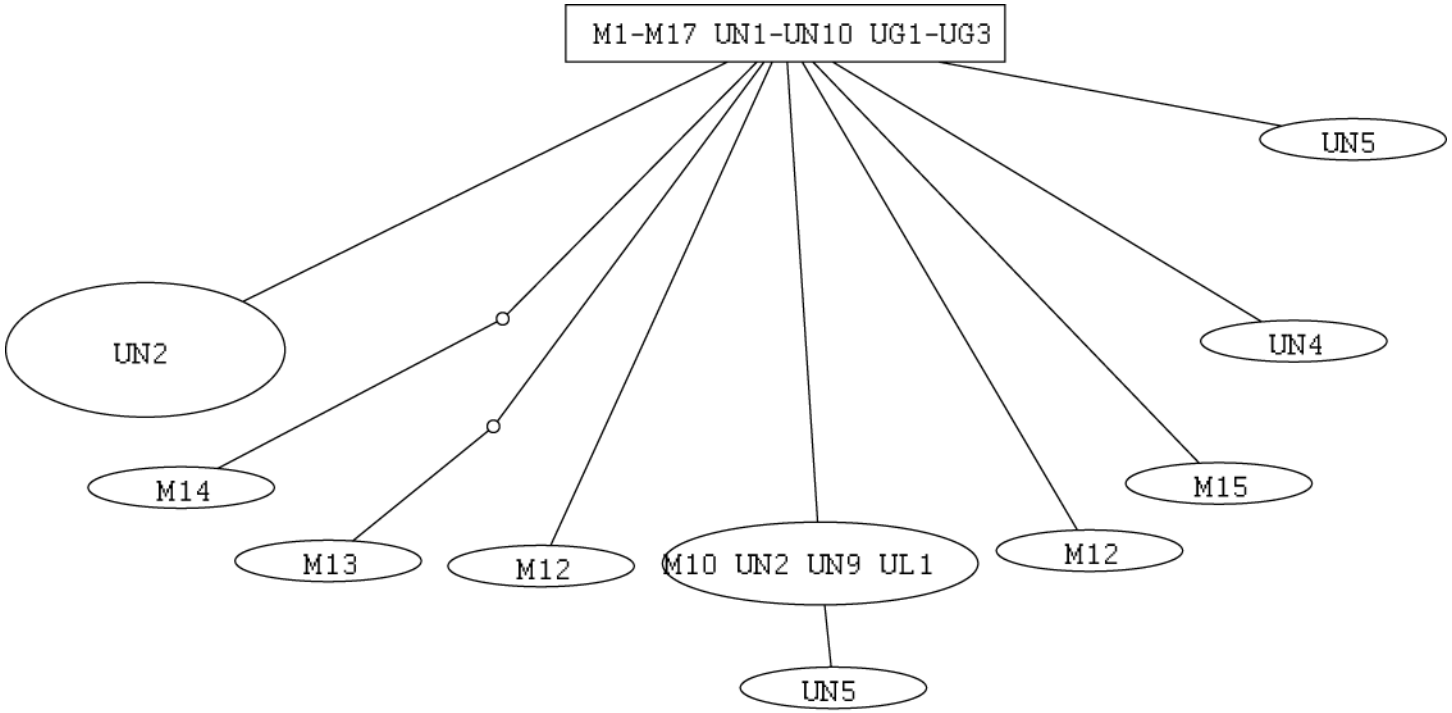
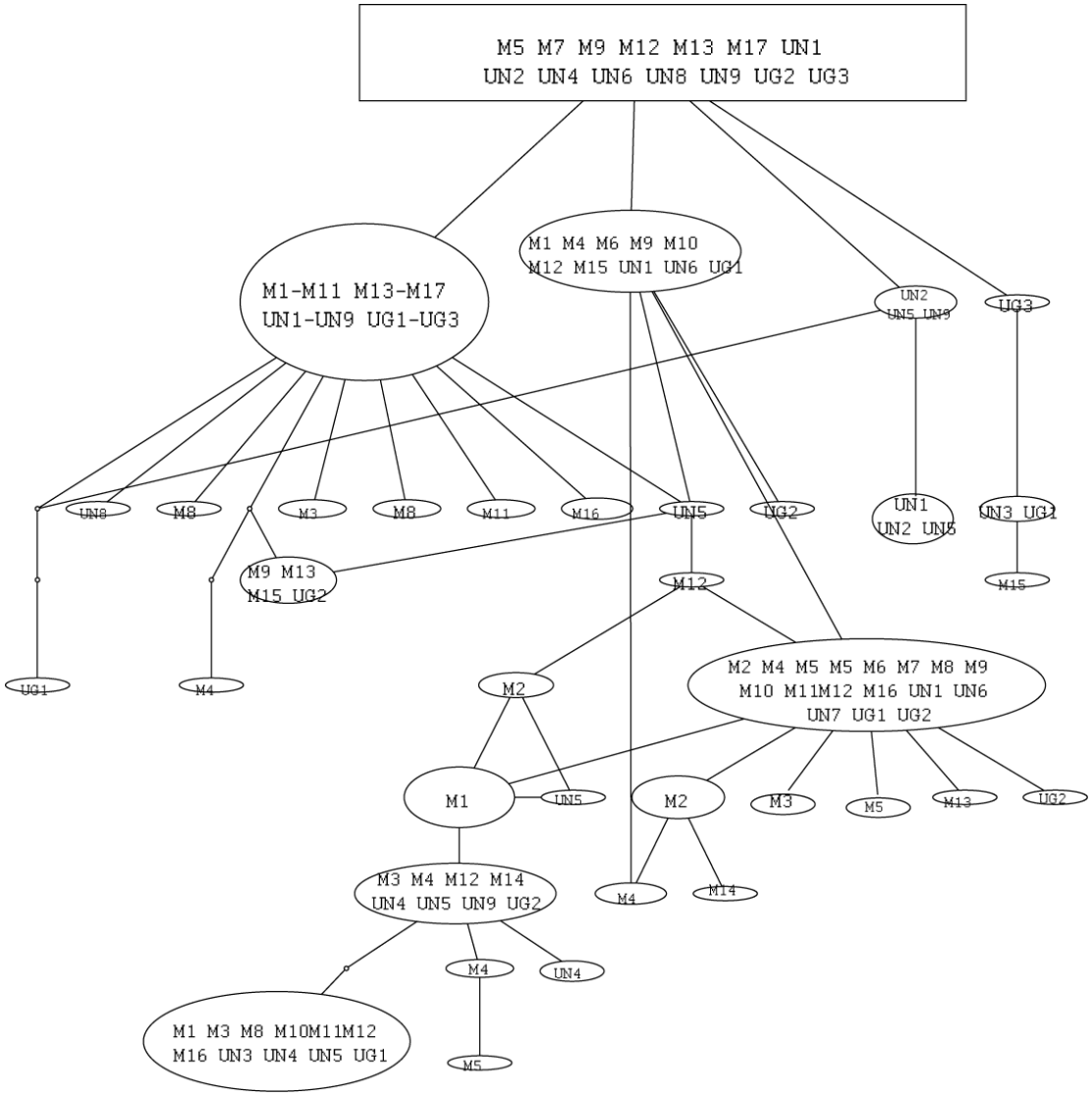


Figure S2 A-D. An example of TCS haplotype network of the four genes (*a3ip, abaR, ccoaomt, erd3*, respectively) as calculated by statistical parsimony with gaps in the data set coded as missing. The haplotype with the highest outgroup probability is displayed as a square, while other haplotypes are displayed as ovals. Haplotype frequencies are proportional to the size of the oval and small empty circles on branches are representative of haplotypes missing from the data set. Corresponding samples are represented by the names of the populations they belong to listed in Table S1. As evident from the genes genealogies, the outgroup haplotypes are shared between all three taxa. There was any major clade detected in the haplotype network specific to any species or geographical location.

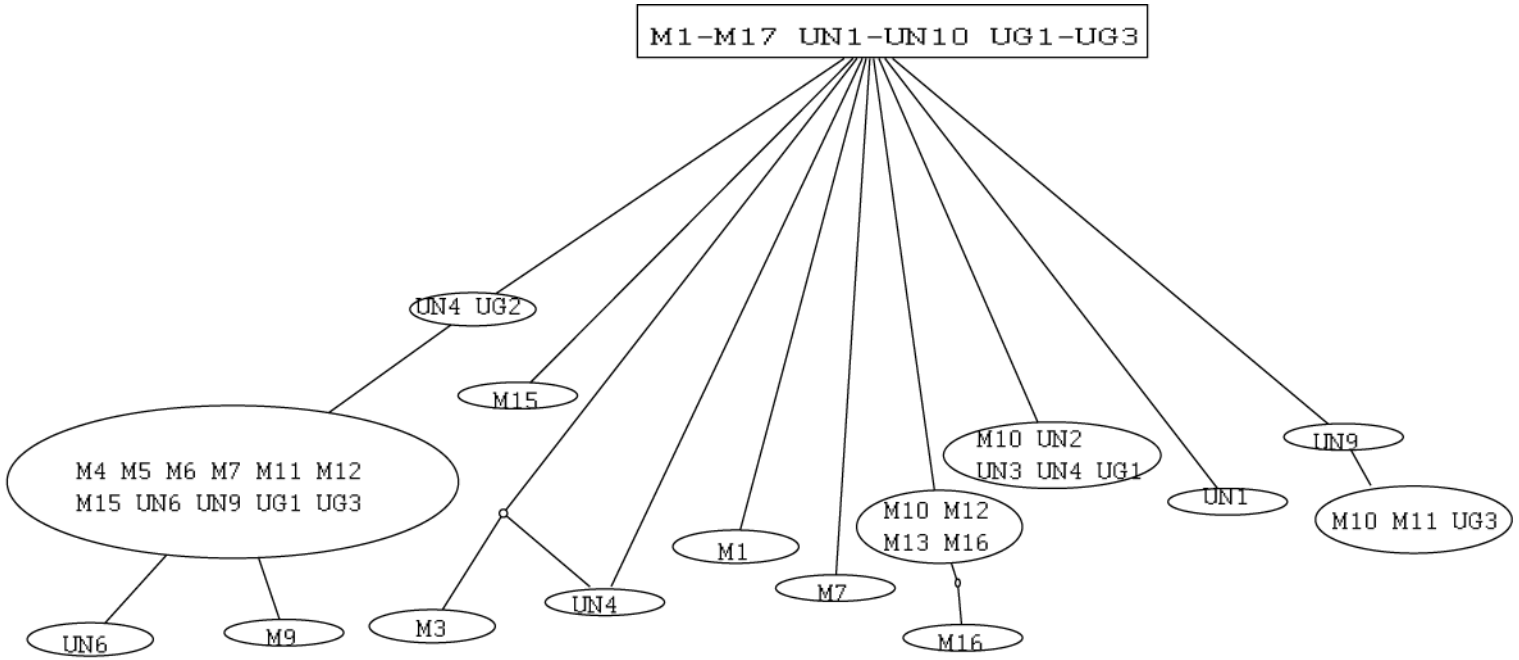
A. *a3ip2*



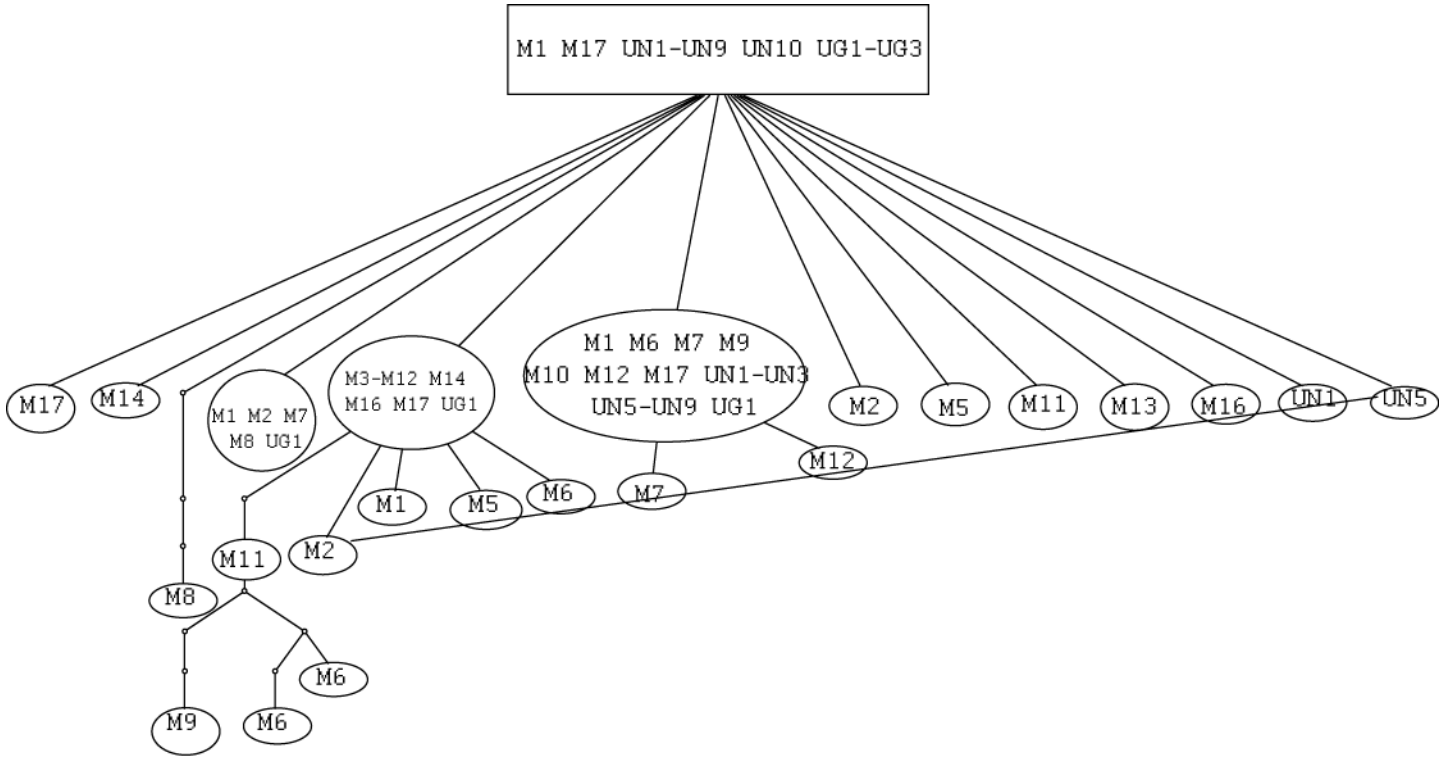
B. *abaR*


M5 M7 M9 M12 M13 M17 UN1
UN2 UN4 UN6 UN8 UN9 UG2 UG3
M1 M4 M6 M9 M10
M12 M15 UN1 UN6 UG1
M1-M11 M13-M17
UN1-UN9 UG1-UG3
UN2
UN5 UN9
UG3
UN8
M8
M3
M8
M11
M16
UN5
UG2
UN1
UN2 UN5
UN3 UG1
M15
M9 M13
M15 UG2
M12
UG1
M4
M2
M2 M4 M5 M5 M6 M7 M8 M9
M10 M11M12 M16 UN1 UN6
UN7 UG1 UG2
M1
UN5
M2
M3
M5
M13
UG2
M3 M4 M12 M14
UN4 UN5 UN9 UG2
M4
M14
M4
UN4
M1 M3 M8 M10M11M12
M16 UN3 UN4 UN5 UG1
M5

C. *ccoaomt*


M1-M17 UN1-UN10 UG1-UG3
UN4 UG2
M15
M4 M5 M6 M7 M11 M12
M15 UN6 UN9 UG1 UG3
M10 UN2
UN3 UN4 UG1
UN9
UN1
M10 M11 UG3
M10 M12
M13 M16
M1
M7
UN4
M3
UN6
M9
M16

D. *erd3*


M1 M17 UN1-UN9 UN10 UG1-UG3
M17
M14
M1 M2 M7
M8 UG1
M3-M12 M14
M16 M17 UG1
M1 M6 M7 M9
M10 M12 M17 UN1-UN3
UN5-UN9 UG1
M2
M5
M11
M13
M16
UN1
UN5
M12
M1
M5
M6
M7
M11
M2
M8
M6
M9
M6