

## Genetic differentiation of *Sorbus torminalis* in Eastern Europe as determined by microsatellite markers

Veronika KUČEROVÁ<sup>1</sup>, Martin HONEC<sup>1</sup>, Ladislav PAULE<sup>1</sup>, Petar ZHELEV<sup>2</sup> & Dušan GÖMÖRY<sup>1\*</sup>

<sup>1</sup>Technical University in Zvolen, Faculty of Forestry, TG Masaryka 24, 960 53 Zvolen, Slovakia;  
 e-mail: vkucerova@vsld.tuzvo.sk, paule@vsld.tuzvo.sk, gomory@vsld.tuzvo.sk

<sup>2</sup>University of Forestry, 10 Kliment Ohridsky blvd., 1756 Sofia, Bulgaria

**Abstract:** The genetic variation in fourteen *Sorbus torminalis* (L.) Crantz. populations distributed over the eastern and south-eastern part of its range was studied using seven nuclear microsatellite loci. The differentiation level was relatively high ( $F_{ST} = 0.228$ ), as expected for a species with a fragmented range. The distance-based approach to the analysis of differentiation patterns (neighbour-joining tree based on pairwise coefficients of differentiation) did not reveal a clear geographical structure. On the other hand, model-based Bayesian methods (BAPS and STRUCTURE) gave geographically continuous clusters of populations. The occurrence of populations deviating strongly from the general pattern is attributed to founder effect. In spite of a generally high differentiation, a significant isolation-by-distance pattern was found, which might be a consequence of postglacial migration and gene flow among descendants of different refugia.

**Key words:** *Sorbus torminalis*; genetic differentiation; isolation by distance; nuclear microsatellites

### Introduction

The wild service tree, *Sorbus torminalis* (L.) Crantz (Rosaceae), is a forest tree of the lower elevations in the temperate and Mediterranean zone of Europe, western Asia and north-western Africa. Although the range of this species is extensive, population densities are generally very low. As a light-demanding post-pioneer species, it is frequently outcompeted by climax species such as oak and beech in forest stands (Demesure-Musch & Oddou-Muratorio 2004). However, due to efficient seed dispersal by birds and small mammals, and the ability of vegetative spread by root suckers, the tree can easily colonize appropriate sites, such as large gaps, clearings, low-density forests and abandoned agricultural land surrounded by forests, remaining permanently present in the landscape.

The species is diploid ( $2n = 34$ ), higher ploidy levels have not been reported. It forms allopolyploid as well as probably homoploid hybrids with *S. aucuparia* and especially *S. aria* (Nelson-Jones et al. 2002). Hybrid taxa of the *S. latifolia* group are generally agamospermous, but sexual populations exist as well (Aas et al. 1994). Nevertheless, no backcrosses have been reported, thus hybrids are reproductively isolated from *S. torminalis* and no gene-pool contamination through introgression occurs (Oddou-Muratorio et al. 2001).

Geographical distribution of the levels of genetic diversity and differentiation patterns of *S. torminalis*

have been extensively studied using various types of gene markers (allozymes, RAPD, SSR, cpPCR-RFLP) on the local or regional level (Bednorz et al. 2006; Belletti et al. 2008; Demesure et al. 2000; Rasmussen & Kollmann 2007). However, no rangewide study has been conducted. Marker studies contributed significantly to the knowledge of the reproductive system, extent of clonal spread and gene dispersal through pollen and seed flow (Angelone et al. 2007; Hoebee et al. 2006, 2007; Oddou-Muratorio et al. 2004; Rasmussen & Kollmann 2007), *S. torminalis* even became a sort of model species for the study of gene flow processes (Klein et al. 2008; Oddou-Muratorio et al. 2006).

The wood of wild service tree belongs to the most high-priced among the native European tree species (Demesure-Musch & Oddou-Muratorio 2004), so that despite a rare occurrence the species cannot be considered economically unimportant. Improvement programs including establishing seed orchards have been initiated in several European countries, which on the one hand is a positive development as they contribute to the preservation of the most valuable gene pool, but on the other hand, there is a risk that such seed orchards with a relatively narrow gene pool will be used as exclusive source of reproductive material for any afforestation. Moreover, the species suffers from range fragmentation and reduction of population sizes due to prevailing agricultural use of land at low elevations and contemporary forestry practices favouring closed-

\* Corresponding author

Table 1. List of the analyzed *Sorbus torminalis* populations.

Name	Country	Longitude	Latitude	Sample size
Kakanj-1	Bosnia and Herzegovina	18°06'E	43°11'N	30
Kakanj-2	Bosnia and Herzegovina	18°05'E	44°10'N	31
pooled population	Bulgaria	~26°23'E	~42°40'N	9
Dulovo	Bulgaria	27°09'E	43°49'N	41
Ucka	Croatia	14°12'E	45°19'N	9
Tbilisi	Georgia	44°47'E	41°43'N	10
Piaski u Poznania	Poland	17°09'E	51°49'N	24
Bejan	Romania	22°51'E	45°57'N	16
Medias	Romania	24°21'E	46°10'N	20
Bountin	Romania	25°47'E	44°39'N	16
Kšinná	Slovakia	18°21'E	48°49'N	9
Babky	Slovakia	19°38'E	49°11'N	41
Muránska planina	Slovakia	20°03'E	48°45'N	10
Koštabona	Slovenia	13°45'E	45°29'N	17

canopy high forests (Hoebee et al. 2007). Therefore, both *in situ* and *ex situ* gene conservation are highly important. The information on geographical patterns of genetic variation is essential for designing efficient conservation measures both on national and international levels. The objective of our study was thus describing genetic variation patterns of *S. torminalis* in the eastern part of the species' distribution range which before has been understudied.

### Material and methods

Fourteen *S. torminalis* populations distributed over the eastern and southeastern parts of its distribution range were sampled (Table 1). Sample size varied from 9 to 41 adult trees per population, depending on population size. Populations from southeastern Bulgaria were represented by very small samples (2–3 trees), therefore, they were pooled and treated as a single population. To avoid sampling of ramets of the same clone, minimum distance between trees was 20 m. The collected leaves were immediately dried in silica gel. To extract genomic DNA, two leaf discs with a diameter 8 mm were cut and ground to a fine powder in a mixer mill using tungsted beads. Genomic DNA was extracted using the CTAB method according to Doyle and Doyle (1987) which was modified for a smaller amount of starting material.

For genotyping, seven microsatellite loci (MSS1, MSS5, MSS6, MSS9, MSS13, MSS16) according to Oddou-Muratario et al. (2005) and MS6g (Nelson-Jones et al. 2002) were used. Primers were multiplexed into two groups, group A containing primers MSS6, MS6g, MSS9, MSS13, MSS16 and group B containing primers MSS1, MSS5. All PCRs were performed in a MJ Research PTC 200 or Perkin Elmer GeneAmp 9700 thermal cycler in 10 µL reactions. PCR Master mix contained approximately 50 ng of DNA, 5 µL Qiagen Multiplex PCR Kit and in multiplex A primers in following amounts: 0.20 µL MSS6, 0.1 µL MS6g, 0.15 µL MSS9, 0.40 µL MSS13, 0.15 µL MSS16; in multiplex B primers in following amounts: 0.10 µL MSS1, 0.15 µL MSS5. Amplification profile consisted of an initial denaturation step at 95°C for 15 min and followed by 35 cycles with the following profile: 30 s denaturation at 94°C, 90 s annealing step (see below) and 1 min extension step at 72°C. The annealing temperature in the first cycle was 59°C for multiplex A (62°C for multiplex B), it was subsequently reduced in each cycle by 0.5°C (0.3°C in multiplex B) for the next 10 cycles and was

continued at 54°C (59°C in multiplex B) for remaining 24 cycles. Final extension was 30 min at 60°C. Amplification products were separated on an ABI 3100 DNA sequencer, and genotypes were determined using the GeneMapper software v. 3.7.

To assess the trends of genetic differentiation, we used both distance-based and model-based approaches. First, pairwise coefficients of differentiation  $F_{ST}$  (Weir & Cockerham 1984) between populations were calculated using the program GENEPOP v. 3.3 (Raymond & Rousset 1995). An unrooted cladogram was then constructed using the neighbour-joining method employing the program PHYLIP v. 3.61 (Felsenstein 2004). Relationship between  $F_{ST}$  and geographical distances among populations were tested using simple Mantel tests, the significance of Mantel correlation was derived from 100,000 random permutations. The calculations were performed using the program *zt* (Bonnet & van de Peer, Ghent University, Belgium). Furthermore, the dataset was tested for the presence of a significant isolation by distance pattern as described by Rousset (1997). As recommended for a two-dimensional case, regression of estimates of  $F_{ST}/(1 - F_{ST})$  against logarithm of distance was evaluated using GENEPOP v.3.1b (Raymond & Rousset 1995). As the paths of colonization of distribution range are unknown, simple air distances between populations were used.

Second, we used model-based Bayesian clustering procedures. The procedure following Pritchard et al. (2000) requires unlinked genetic markers. Therefore, we tested linkage disequilibria within our dataset using GENEPOP v. 3.3. The clustering procedure uses Bayesian inference to assign multilocus genotypes of the sampled trees to a predefined number of clusters  $K$  in a way that Hardy-Weinberg and linkage disequilibria within clusters were maximally explained. The program STRUCTURE (Pritchard et al. 2000) was run 10 times for each  $K = 1$  to 10, with a burn-in period of 10,000 and subsequent 50,000 iterations to determine the optimum number of clusters using the criterion of Evanno et al. (2005). As the increment of average likelihood for more than 5 groups became negligible and the criterion  $\Delta K$  exhibited the highest peak at  $K = 5$  (Fig. 1), a further run for  $K = 5$  with 100,000 iterations after a burn-in period of 20,000 was used to assign individuals to clusters. Since we assumed the existence of gene exchange among populations, we used an admixture model estimating the fraction of the genome that each individual draws from each cluster. The fractions of the gene pools of both clusters were

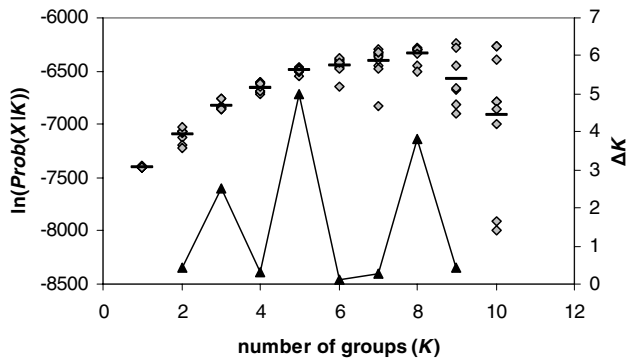


Fig. 1. Estimated log posterior probabilities of the data (grey diamonds) and the Evanno's criterion  $\Delta K$  (black triangles) as a function of the number of clusters  $K$ .

subsequently averaged over individuals within populations.

The method of Corander et al. (2003) allows an estimation of posterior probabilities of pairwise equality of sampled populations (i.e., their origin from a single population), as well as that of different groupings of the sampled populations (i.e., population structure). The program BAPS v.1.1 (Corander et al. 2003) was used for the calculation. The number of sampled populations was too high for an exact enumeration of probabilities, a Markov chain Monte Carlo approximation (100,000 iterations) was used instead.

## Results and discussion

The overall differentiation was quite high ( $F_{ST} = 0.228$ ), as expected for a species with a fragmented range. This value is similar to differentiation levels observed in studies based on allozymes (Bednorz et al. 2006; Demesure et al. 2000; Rasmussen & Kollmann 2007). Our observation is slightly higher, reflecting probably a larger area covered by sampling, as we investigated populations from south-eastern Europe and partly central Europe. On the other hand, Belletti et al. (2008) observed even higher differentiation levels among Italian populations using RAPD markers.

The neighbour-joining tree (Fig. 2) did not reveal a clear geographical structure, although there are some indications of the existence of geographical patterns. Slovak populations, together with the Romanian population Bejan, are placed in one branch, the other one is composed of the populations from south-eastern Balkans (Romania and Bulgaria) and contains also the Georgian population. The third branch contains populations from the western Balkans (Slovenia, Croatia, Bosnia), but also the pooled Bulgarian sample and the Polish population. Despite unclear spatial patterns, differentiation is significantly correlated with geographical distance (Mantel's  $r = 0.757$ ;  $P = 0.015$ ), what seems to be the consequence of isolation by distance, as the Rousset's model yielded a regression  $F_{ST}(1 - F_{ST}) = -0.1071 + 0.0195 \ln d$ , which is highly significant ( $P > 0.9999$ ).

Model-based approach gave more interpretable outcomes in terms of geography, although the proce-

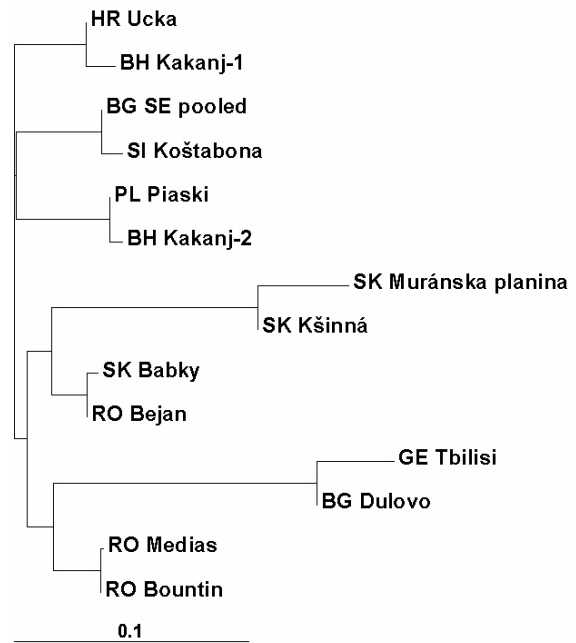


Fig. 2. Nearest-neighbour tree based on pairwise  $F_{ST}$  among the investigated *Sorbus torminalis* populations.

dures do not take into account the geographical positions of individual samples. The procedure BAPS yielded two partitions with almost equal posterior probabilities ( $P = 0.547$  and  $0.441$  for 6 and 7 clusters, respectively), posterior probabilities for all other partitions were far below 0.01. Clusters were geographically continuous: Poland with Slovakia (except the population Muránska planina), eastern Balkans (Romania with Bulgaria), Slovenia with Croatia, Bosnia and Herzegovina, and Georgia. The Slovak population Muránska planina formed a separate group. The second partition is almost identical, only the Polish and Slovak populations are separated into two distinct clusters (Fig. 3).

The number of clusters found was lower in STRUCTURE, five groups were identified as the most probable number. The distribution of gene pools of these groups shows a quite clear geographical pattern and corresponds well with the outcomes of BAPS (Fig. 4). Genes of the first group are highly represented in Central-European populations (Poland, Slovakia) and their proportions decrease quite smoothly towards the Balkans. The second group prevails in Bosnian populations, and gradually disappears both towards the North and the East. The third group predominates in south-eastern Balkans (Bulgaria), and its frequency decreases smoothly along the Carpathian arc. As expected, the geographically distant Georgian population contains almost exclusively genes of a separate fourth group. Again, the Slovak population Muránska planina represents a separate group, against all expectations.

According to Hamrick et al. (1992), in plant species with life-history traits similar to *S. torminalis*, differentiation levels are generally moderate (on average,  $F_{ST} \approx 0.1$  and  $0.05$  for outcrossing insect-pollinated

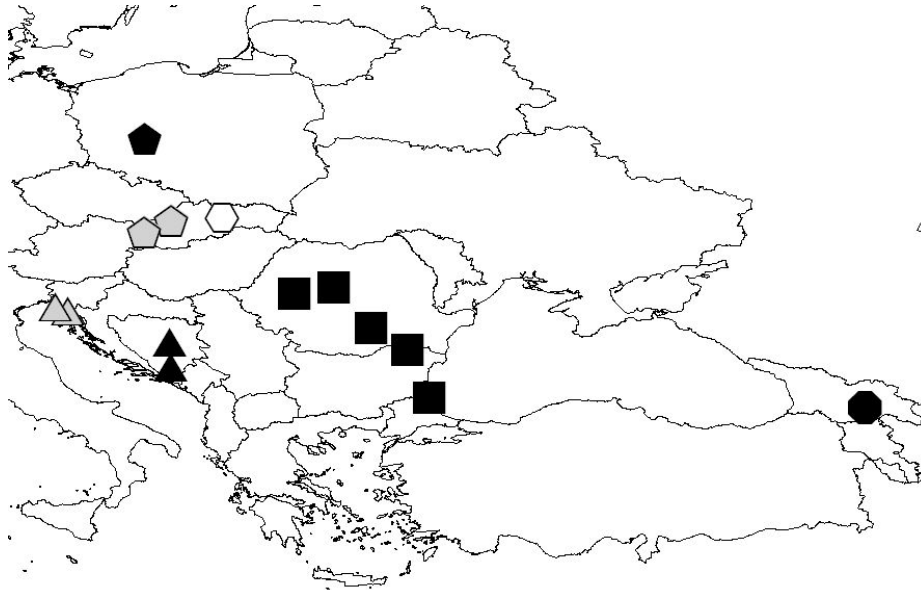


Fig. 3. Model-based analysis of population structure in *S. torminalis*: Distribution of groups identified by BAPS (7-groups partition).



Fig. 4. Model-based analysis of population structure in *S. torminalis*: Proportions of individual gene pools identified by STRUCTURE in each population.

and animal-dispersed (ingested seeds) species, respectively, which is much lower than the value we observed). *S. torminalis* occurs usually in small populations with scattered distribution, and because of low levels of gene flow between populations (Bednorz et al. 2006), genetic drift caused by recurrent bottlenecks due to fluctuating population sizes can be an important factor shaping their gene pools. This causes loss of genetic diversity and increases differentiation between populations. A high level of differentiation between populations can be explained also by founder effect. Populations of *S. torminalis* are characterized by metapopulation dynamics, they undergo frequent extinction and subsequent recolonization (Demesure et al. 2000; Oddou-Muratorio et al. 2001; Angelone et al. 2007). In such cases, the number of individuals involved in the founding effect

and the number of populations from which they come from is the crucial factor influencing genetic differentiation between young populations (Wade & McCauley 1988). Founder effect probably had an important part in the evolution of the population Muránska planina, which was found outstanding (as regards gene pools) among the populations of the Poland-Slovak group.

Potentially, such local deviations from general patterns may arise through interspecific hybridization and introgression, which is common in Rosaceae. *S. torminalis* is known to form hybrids with *S. aria*, designated as *S. latifolia*, which are generally diploid, sexual and able of backcrossing (Aas et al. 1994; Nelson-Jones et al. 2002). However, hybridization is generally unidirectional (*S. aria* as the pollen donor), hybrids are morphologically distinguishable and the backcrosses tend to re-

semble more *S. aria* (Aas et al. 1994; Oddou-Muratorio et al. 2001; Chester et al. 2007), so that introgression seems to be very improbable cause of the occurrence of populations not fitting into general geographic patterns.

In spite of a high overall differentiation, a significant isolation by distance was identified. Geographically proximate populations tend to be genetically more similar than expected under random distribution. This is partly contrasting with previous observations, since a strong isolation by distance was reported at the intrapopulation scale (Oddou-Muratorio et al. 2004), but at the regional scale, the observations differ: Oddou-Muratorio et al. (2001) or Angelone et al. (2007) identified only a weak IBD pattern, whereas Bednorz et al. (2006) or Belletti et al. (2008) found clear inter-regional differences. The observed patterns of geographical structure visible from model-based analysis may reflect glacial history and postglacial migration. However, no rangewide phylogeographic studies are available and our data are too scarce to test this hypothesis explicitly.

#### Acknowledgements

The study was supported by the grant of the Slovak Grant Agency for Science (VEGA) No. 1/3262/06 and the Centre of Excellence "Adaptive Forest Ecosystems" funded by the Operational Programme Research and Development financed from the European Regional Development Fund. The assistance of Z. Sarvašová, D. Krajmerová, I. Romšáková, and R. Brus is greatly appreciated.

#### References

- Aas G., Maier J., Baltisberger M. & Metzger S. 1994. Morphology, isozyme variation, cytology, and reproduction of hybrids between *Sorbus aria* (L.) Crantz and *S. torminalis* (L.) Crantz. *Botanica Helvetica* **104**: 195–214.
- Angelone S., Hilfiker K., Holderegger R., Bergamini A. & Hoebee S.E. 2007. Regional population dynamics define the local genetic structure in *Sorbus torminalis*. *Mol. Ecol.* **16**: 1291–1301.
- Bednorz L., Myczko L. & Kosinski P. 2006. Genetic variability and structure of the wild service tree (*Sorbus torminalis* (L.) Crantz) in Poland. *Silvae Genetica* **55**: 197–202.
- Belletti P., Monteleone I., Ferrazzini D. 2008. A population genetic study in a scattered forest species, wild service tree [*Sorbus torminalis* (L.) Crantz], using RAPD markers. *Eur. J. Forest Res.* **127**: 103–114.
- Chester M., Cowan R.S., Fay M.F. & Rich T.C.G. 2007. Parentage of endemic *Sorbus* L. (Rosaceae) species in the British Isles: evidence from plastid DNA. *Bot. J. Linn. Soc.* **154**: 291–304.
- Corander J., Waldmann P. & Sillanpää M.J. 2003. Bayesian analysis of genetic differentiation between populations. *Genetics* **163**: 367–374.
- Demesure B., Le Guerroue B., Lucchi G., Prat D. & Petit R.J. 2000. Genetic variability of a scattered temperate forest tree: *Sorbus torminalis* L-(Crantz). *Ann. Forest Sci.* **57**: 63–71.
- Demesure-Musch B. & Oddou-Muratorio S. 2004. EUFORGEN Technical Guidelines for genetic conservation and use for wild service tree (*Sorbus torminalis*). Inter. Plant Genetic Resources Inst., Rome, Italy, 6 pp.
- Doyle J.J. & Doyle J.L. 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochem. Bull.* **19**: 11–15.
- Evanno G., Regnaut S. & Goudet J. 2005. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Mol. Ecol.* **14**: 2611–2620.
- Felsenstein J. 2004. PHYLIP (Phylogeny Inference Package) version 3.6. Department of Genome Sciences, University of Washington, Seattle.
- Hamrick J.L., Godt M.J.W. & Sherman-Broyles S.L. 1992. Factors influencing levels of genetic diversity in woody plant species. *New Forests* **6**: 95–124.
- Hoebee S.E., Arnold U., Duggelin C., Gugerli F., Brodbeck S., Rotach P. & Holderegger R. 2007. Mating patterns and contemporary gene flow by pollen in a large continuous and a small isolated population of the scattered forest tree *Sorbus torminalis*. *Heredity* **99**: 47–55.
- Hoebee S.E., Menn C., Rotach P., Finkeldey R. & Holderegger R. 2006. Spatial genetic structure of *Sorbus torminalis*: The extent of clonal reproduction in natural stands of a rare tree species with a scattered distribution. *Forest Ecol. Manag.* **226**: 1–8.
- Klein E.K., Desassis N. & Oddou-Muratorio S. 2008. Pollen flow in the wildservice tree, *Sorbus torminalis* (L.) Crantz. IV. Whole interindividual variance of male fecundity estimated jointly with the dispersal kernel. *Mol. Ecol.* **17**: 3323–3336.
- Nelson-Jones E.B., Briggs D. & Smith A.G. 2002. The origin of intermediate species of the genus *Sorbus*. *Theor. Appl. Genet.* **105**: 953–963.
- Oddou-Muratorio S., Aligon C., Decroocq S., Plomion C., Lamant T. & Musch-Demesure B. 2005. Microsatellite primers for *Sorbus torminalis* and related species. *Mol. Ecol. Notes* **1**: 297–299.
- Oddou-Muratorio S., Demesure-Musch B., Pelissier R. & Gouyon P.H. 2004. Impacts of gene flow and logging history on the local genetic structure of a scattered tree species, *Sorbus torminalis* L. Crantz. *Mol. Ecol.* **13**: 3689–3702.
- Oddou-Muratorio S., Klein E.K., Demesure-Musch B. & Austerlitz F. 2006. Real-time patterns of pollen flow in the wildservice tree, *Sorbus torminalis* (Rosaceae). III. Mating patterns and the ecological maternal neighborhood. *Am. J. Bot.* **93**: 1650–1659.
- Oddou-Muratorio S., Petit R.J., Le Guerroue B., Guesnet D. & Demesure B. 2001. Pollen- versus seed-mediated gene flow in a scattered forest tree species. *Evolution* **55**: 1123–1135.
- Pritchard J.K., Stephens M. & Donnelly P. 2000. Inference of population structure from multilocus genotype data. *Genetics* **155**: 945–959.
- Rasmussen K.K. & Kollmann J. 2007. Genetic diversity, spatial patterns, and growth of root sprouts in a temperate tree at the northern distribution limit. *Ecoscience* **14**: 250–258.
- Raymond M. & Rousset F. 1995. GENEPOP (Version 1.2): population genetics software for exact tests and ecumenicism. *J. Hered.* **86**: 248–249.
- Rousset F. 1997. Genetic differentiation and estimation of gene flow from *F*-statistics under isolation by distance. *Genetics* **145**: 1219–1228.
- Wade M.J. & McCauley D.E. 1998. Extinction and recolonization: their effects on the genetic differentiation of local populations. *Evolution* **42**: 995–1005.
- Weir B.S. & Cockerham C.C. 1984. Estimating *F*-statistics for the analysis of population structure. *Evolution* **38**: 1358–1370.