

A northern glacial refugium for bank voles (*Clethrionomys glareolus*)

Petr Kotlík^{*†‡}, Valérie Deffontaine^{§¶}, Silvia Mascheretti^{||}, Jan Zima^{**}, Johan R. Michaux^{§¶}, and Jeremy B. Searle[†]

^{*}Department of Vertebrate Evolutionary Biology and Genetics, Institute of Animal Physiology and Genetics, Academy of Sciences of the Czech Republic, CZ-27721 Liběchov, Czech Republic; [†]Department of Biology (area 2), University of York, P.O. Box 373, York YO10 5YW, United Kingdom; [§]Unité de Recherches Zoogéographiques, Université de Liège, Bâtiment B22, Boulevard du Rectorat, B-4000 Liège, Belgium; [¶]Centre de Biologie et de Gestion des Populations, Unité Mixte de Recherche 1062, Campus International de Baillarguet, CS 30016, F-34988 Montferrier/Lez (Montpellier), France; ^{||}First Department of Medicine, Christian Albrechts University, Schittenhelmstrasse 12, D-24105 Kiel, Germany; and ^{**}Institute of Vertebrate Biology, Academy of Sciences of the Czech Republic, CZ-60365 Brno, Czech Republic

Edited by David B. Wake, University of California, Berkeley, CA, and approved August 14, 2006 (received for review April 21, 2006)

There is controversy and uncertainty on how far north there were glacial refugia for temperate species during the Pleistocene glaciations and in the extent of the contribution of such refugia to present-day populations. We examined these issues using phylogeographic analysis of a European woodland mammal, the bank vole (*Clethrionomys glareolus*). A Bayesian coalescence analysis indicates that a bank vole population survived the height of the last glaciation ($\approx 25,000$ – $10,000$ years B.P.) in the vicinity of the Carpathians, a major central European mountain chain well north of the Mediterranean areas typically regarded as glacial refugia for temperate species. Parameter estimates from the fitted isolation with migration model show that the divergence of the Carpathian population started at least 22,000 years ago, and it was likely followed by only negligible immigration from adjacent regions, suggesting the persistence of bank voles in the Carpathians through the height of the last glaciation. On the contrary, there is clear evidence for gene flow out of the Carpathians, demonstrating the contribution of the Carpathian population to the colonization of Europe after the Pleistocene. These findings are consistent with data from animal and plant fossils recovered in the Carpathians and provide the clearest phylogeographic evidence to date of a northern glacial refugium for temperate species in Europe.

climate change | colonization | northern refugia | phylogeography | Quaternary

Climate change has had major impacts on geographic distributions, demography, and, thus, evolution of species. It is now well established that in the northern hemisphere many temperate species retreated from large continental areas during the height of the last glaciation ($\approx 25,000$ – $10,000$ years B.P.) and were only able to survive in sheltered refugia, which provided suitable conditions. However, controversy and uncertainty remain regarding the number and location of glacial refugia that contributed to modern populations. Temperate-adapted woodland species would have been particularly strongly affected because the unglaciated areas at this time in northern and central Europe, Asia, and North America were largely treeless. Numerous distributional and phylogeographic patterns of fauna and flora are consistent with the process of colonization from refugia at lower latitudes after the Pleistocene (1–3). In Europe, in particular, there has been much emphasis on the Mediterranean peninsulas of Iberia, Italy, and the Balkans as possible locations for glacial refugia (4, 5). However, there is evidence that refugia for temperate species may also have existed further north in central and western Europe (6), locations described as “northern refugia” by Stewart and Lister (7). We follow the terminology of Stewart and Lister. The most compelling evidence supporting the existence of northern refugia comes from the vicinity of the Carpathians, a mountain chain $>1,500$ km long in central Europe. In particular, plant pollen and macrofossils recovered from various places in the Carpathians showed that coniferous and broad-leaved trees were part of the local full-glacial envi-

ronment (8, 9). This finding is supported by evidence from animal fossils, which provide records of woodland species of small mammals in deposits from the last glaciation in the Carpathians (10, 11). Furthermore, divergent mtDNA sequences from extant Carpathian populations of several temperate vertebrate species, including amphibians (12, 13) and fish (14, 15), indicate the maintenance of separate lineages that may have persisted in the Carpathians during the last glaciation. Various lines of evidence thus point to the existence of a glacial refugium for temperate species in the Carpathians. As a result, modern populations of species in central and northern Europe might have not been derived exclusively from southern refugial populations, but also from populations that survived in much more northerly regions. A proper realization of the importance of northern refugia thus has the potential to change dramatically the way that postglacial colonization of Europe is viewed, providing a different perspective on the communities of temperate species that currently occur there.

This phylogeographic study examines the genetic evidence for a Carpathian refugium in a European woodland mammal, the bank vole (*Clethrionomys glareolus*). The bank vole is a small rodent that is widely distributed in Europe between the British Isles and northern Spain in the west to central Siberia in the east. It is found abundantly in a wide variety of woodland habitats, such as temperate broad-leaved and mixed forests (16), and its glacial distribution most likely tracked this type of environment (11). The bank vole is an excellent model to test the relative contribution of traditionally recognized refugia and those that may have existed farther north, because its current wide distribution includes all three Mediterranean peninsulas. The bank vole has already been the subject of phylogeographic analyses that suggested that glacial refugia located in central and eastern Europe, as opposed to the Mediterranean refugia, made a major contribution to the modern population of this species in Europe (6, 17). The Carpathian Mountains have been proposed as a possible refugial area for this species based on the fossil record. Bank vole fossils have been identified in deposits from the last glaciation in the Carpathian Mountains together with typical glacial assemblages and suggest that the species may have had a continuous presence in this region (10, 11, 18). We have collected mtDNA sequences from the Carpathians and adjacent regions (Fig. 1) and present the clearest evidence yet that a bank vole population survived in a central European refugium

Conflict of interest statement: No conflicts declared.

This paper was submitted directly (Track II) to the PNAS office.

Abbreviation: IM, isolation with migration.

Data deposition: The sequences reported in this paper have been deposited in the GenBank database (accession nos. DQ472230–DQ472348).

[†]To whom correspondence should be addressed at: Department of Vertebrate Evolutionary Biology and Genetics, Institute of Animal Physiology and Genetics, Academy of Sciences of the Czech Republic, CZ-277 21 Liběchov, Czech Republic. E-mail: kotlik@iapg.cas.cz.

© 2006 by The National Academy of Sciences of the USA

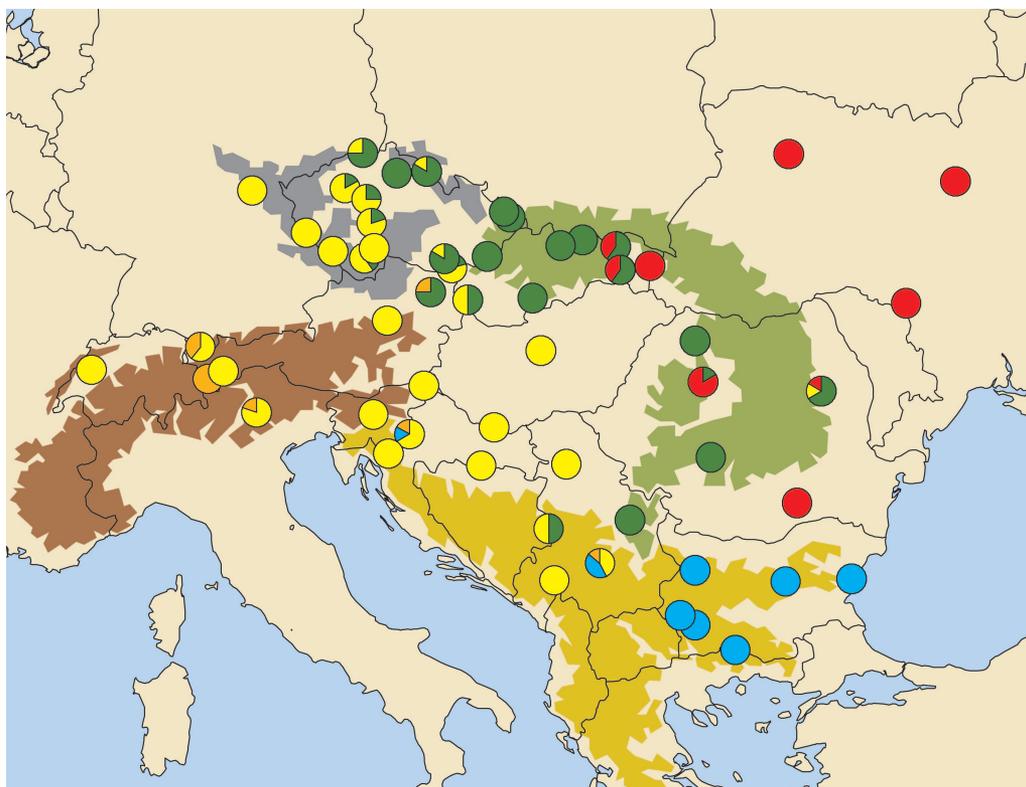


Fig. 1. Map showing the collection sites and the distribution of haplotype lineages. Pie charts show the proportions of haplotypes belonging to each lineage within each population. The colors equate to the clades identified in Fig. 2. Approximate distribution of major mountain regions is shown in green for the Carpathians, in gray for the Bohemian Massif, in brown for Alps, and in yellow for the Balkans.

through the height of the last glaciation. Together with recent findings in North America (19), our results give substantial credence to the occurrence of glacial refugia for temperate species further north than traditionally recognized.

Results and Discussion

Phylogenetic Analyses. The reconstruction of phylogenetic relationships among the 119 haplotypes identified in 224 bank voles demonstrates multiple divergent clades (Fig. 2). Three of these clades were described in an earlier range-wide phylogeographic survey that revealed five European *C. glareolus* clades but failed to resolve finer scale divisions (17). With the enhanced sampling and longer sequences, we discovered a “Carpathian clade” defined by four synonymous nucleotide changes (0.0086–0.0192 net divergence from other clades), which was supported by a bootstrap value of 77% (Fig. 2) and had a localized geographic distribution. The 36 haplotypes in this clade were recovered only from sites in the Carpathians and their close vicinity, and the majority of bank voles from the Carpathians carried haplotypes from this clade (Fig. 1). The remaining 83 haplotypes in our study represented the western, eastern, and Balkan clades identified earlier. There were also seven distinct haplotypes that fell outside these major clades, basal to the Carpathian and western clades and geographically embedded within the distribution of the western clade (Figs. 1 and 2). Despite the dense sampling, the major clades did not show extensive geographic overlap except where their distributions come in contact. These results thus demonstrate that the Carpathian clade is a geographically localized and monophyletic lineage, indicating that ancestors of bank voles from this clade have been isolated in the past from the ancestors of the other clades.

Population Divergence Analysis. If bank voles of the Carpathian clade descended from a population that survived in a Carpathian refugium at the height of the last glaciation, the split from voles of the western population should have occurred earlier than that time. Fig. 3A shows the estimated posterior probability distribution of the divergence time between the two populations obtained from the isolation with migration (IM) model. The divergence time is clearly resolved, with posterior distribution that has a single narrow peak and bounds that fall within the prior distribution. Given the range of mutation rate estimates (see *Materials and Methods*), the position of the peak ($t = 2.11$) corresponds to 30,153–108,887 years. The IM analysis thus suggests that the divergence of the two populations occurred during the height of the last glaciation ($\approx 25,000$ – $10,000$ years B.P.) or earlier, and the bank voles from the Carpathian clade must have survived in a separate refugium. If, on the contrary, the Carpathian population had been founded from the same southern refugial source as the western population and their divergence was postglacial, then the estimate of t should be much more recent. However, any divergence time $< 22,000$ years has very low probability in our IM analysis (Table 1), making such a scenario highly unlikely.

The effective population size for the Carpathian population is >2 times smaller than that for the western population, and the size of the estimated ancestral population was ≈ 5 and 15 times smaller than the two recent populations, respectively (Table 1 and Fig. 3B). This finding suggests that the Carpathian and western populations survived in separate refugia during the height of the last glaciation but most likely expanded from the same ancestral population at the end of an earlier glaciation.

The estimated rate of mtDNA gene flow into the Carpathian population after its separation is at or near zero, whereas gene

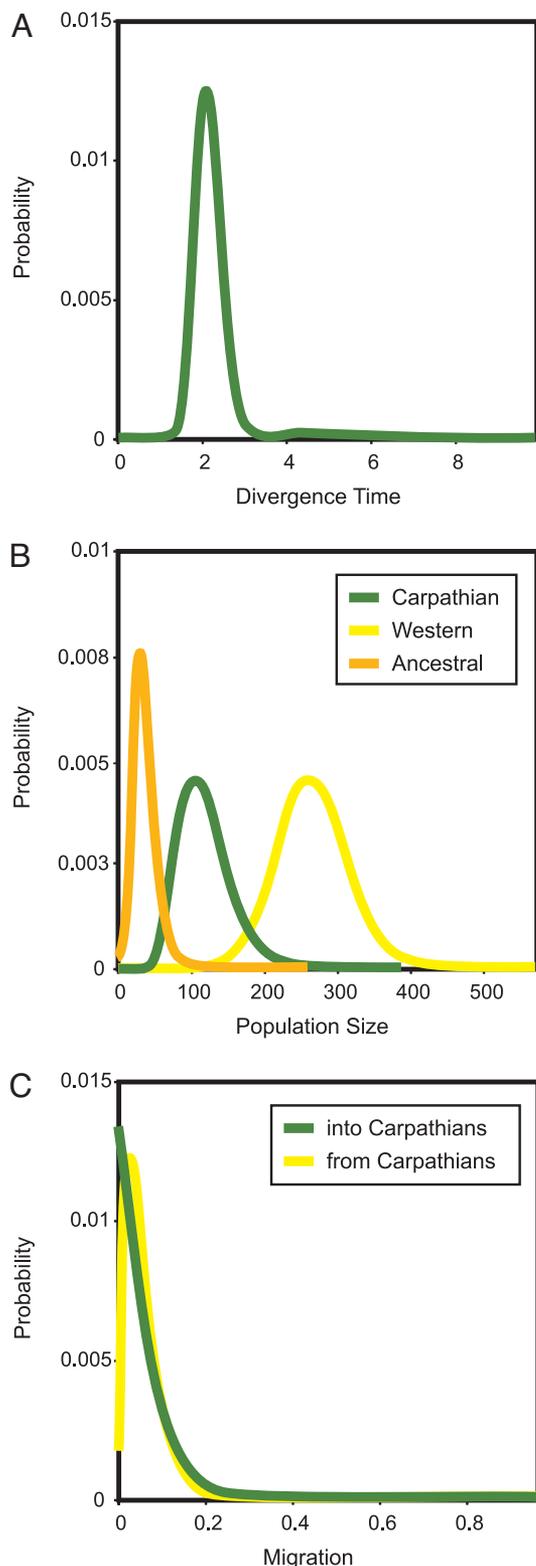


Fig. 3. Posterior probability distributions (scaled by the mutation rate) estimated for divergence time (A), effective population sizes (B), and migration rates (C).

refugia in relatively southern and relatively northern locations, it may have had an impact on the genetic and ecological properties of the species. From which type of refugium a particular population derives may have significance with regards to physiological

traits such as cold-tolerance or dispersal capacity. If postglacial colonization from northern refugia was common, the concept that species responded to climate change mainly by undergoing large-scale distribution shifts to track suitable conditions might be erroneous. Instead, maintaining small isolated populations that persisted because of a locally favorable climate might be the major mechanism by which temperate species responded to the climate change (7, 26). Proper realization of the importance of northern refugia thus has the potential to change dramatically the way that postglacial colonization of Europe is viewed, providing a different perspective on how temperate species respond to changing environmental conditions.

Materials and Methods

Samples. A total of 224 bank voles were collected from 56 localities across the Carpathians and adjacent geographic regions in central and southeastern Europe in Austria, Bulgaria, Czech Republic, Germany, Croatia, Hungary, Switzerland, Italy, Romania, Serbia and Montenegro, Slovakia, Slovenia, and Ukraine (Fig. 1 and Table 2, which is published as supporting information on the PNAS web site). A single individual of the Tien Shan red-backed vole (*Clethrionomys centralis*), phylogenetically a sister species to the bank vole, was sequenced and used as outgroup. Tissue samples of liver, spleen, or toe clips were obtained from mammalogists conducting fieldwork or were collected for the purpose of this study, and were stored in 95% ethanol at 4°C.

Molecular Biological Techniques. Genomic DNA was extracted by the Qiagen (Valencia, CA) DNeasy Tissue kit. A 1,074-bp fragment of the mtDNA *cytb* gene was amplified by PCR by using primers located within the *cytb* (5'-CCCTCTAAT-CAAATCATCAA-3') and Thr tRNA genes (5'-TTTCATT-TCTGGTTTACAAGAC-3'). The primers were designed on the basis of published *cytb* sequences (27), and sequences that were generated with the primer (5'-TGGTGGGGGAAGAGTC-CTT-3'), designed within the Pro tRNA gene by using sequences published by Stacy *et al.* (28). The PCR conditions followed standard methods described for arvicolid rodents (17, 23). Negative extraction and PCR controls with no tissue and no template DNA, respectively, were used in each experiment. The resulting PCR products were purified by using the Qiagen QIAquick PCR Purification kit or the Millipore (Bedford, MA) Montage PCR centrifugal filter devices and were directly cycle-sequenced with the ABI PRISM BigDye Terminator Cycle Sequencing kit (Applied Biosystems, Foster City, CA), the sequencing primer LCLE2 (17), and a newly designed reverse sequencing primer (5'-GTTGGGTTGTTGGATCCTG-3'). The extension products were run on ABI 3730 automated sequencers. Sequences were aligned manually, and any ambiguity was resolved by sequencing the complementary strand.

Phylogenetic Reconstruction. The phylogenetic relationships among the sequences were reconstructed by using the maximum likelihood optimality criterion. The analyses were performed by the algorithm in PHYLML 2.4.4 that simultaneously adjusts tree topology and branch lengths to maximize tree likelihood (29), and using the HKY evolutionary model (30) with the base frequencies A, 0.31; C, 0.29; and G, 0.13; a transition/transversion ratio of 18.07; the proportion of invariable sites set at 0.68; and γ -distributed rates across sites with the shape parameter α equaling to 1.00. This model was determined to be the most appropriate for our dataset by the hierarchical likelihood ratio test of goodness of fit of 56 different nested models to the data, as implemented in Modeltest 3.06 (31). To quantify the confidence in the partitioning within the trees, we performed the nonparametric bootstrap test as applied to phylogenetics by Felsenstein (32) using 1,000 replications.

IM Model Analysis. We analyzed the data under the IM model of population divergence (33, 34). The model assumes that an ancestral population splits into two descendant populations with gene flow possibly continuing between the diverging populations. To fit the IM model to the bank vole data, we used a Bayesian coalescent method that integrates over all possible genealogies by using a Markov chain Monte Carlo (MCMC) approach. This method estimates posterior probability distributions for six demographic parameters including divergence time, two-directional gene flow rates, and effective population sizes of two current populations and the ancestral population (34). We have taken the fossil evidence suggesting that the Carpathians provided glacial refugia for bank voles (11, 17, 18) as a hypothesis of the presence of a distinct and isolated population. The Carpathian population was defined as 43 bank voles collected in the Carpathians; this population excludes 14 individuals that represent recent admixture of the eastern clade (ref. 17; Fig. 1) to eliminate their negative effect on the parameter estimates (35). The western population included 138 bank voles from regions westerly adjacent to the Carpathians and principally characterized by the western clade but carrying also haplotypes of the Carpathian clade and haplotypes basal to both these clades (Figs. 1 and 2). Four bank voles carrying haplotypes of the highly divergent Balkan clade were excluded as representing admixture from the eastern Balkans (Fig. 1). We used the IM program (34) to run the MCMC simulations assuming the HKY model of sequence evolution and uniform prior distributions of parameter ranges, which were empirically determined to ensure that the posterior distributions fell completely within the prior distributions (36). The peaks of the posterior distributions were thus taken as maximum likelihood estimates of the parameters (33, 37). For credibility intervals we recorded for each parameter the 90% highest posterior density (HPD) interval, i.e., the shortest

span that includes 90% of the probability density of a parameter. The analysis was done using several independent runs, each with 4 to 15 chains under the Metropolis coupling to improve mixing (36). Each chain was initiated with a burn-in period of 100,000 updates, and the total length of each analysis was between 5 and 30 million updates. The analysis was considered to have converged on a stationary distribution if the independent runs generated similar posterior distributions (38). To convert the parameter estimates scaled by the mutation rate to calendar years, we used a wide range of plausible divergence rates available for *cytb* of arvicolid rodents of 3.6% to 13% per million years (17, 39, 40). These divergence rates equate to 1.9×10^{-5} to 7.0×10^{-5} mutations per year for the gene region studied (1,074 base pairs). To explore whether natural selection at *cytb* may have acted to limit gene flow, we tested whether the bank vole variation conforms to expectations under selective neutrality using the McDonald–Kreitman test (41) implemented in DnaSP 4.10.8 (42). For this analysis, we used Fisher's exact test to compare the ratio of nonsynonymous to synonymous polymorphisms within the bank vole (overall and for each clade separately) to the ratio of the number of nonsynonymous to synonymous fixed differences between the bank vole and the Tien Shan red-backed vole.

We thank the many colleagues who assisted with field collections, including A. Belancic, V. Bjedov, I. Coroiu, J. Farkas, J. Hausser, G. Horvath, L. Choleva, J. Margaletić, S. Marková, N. Martinková, P. Miklós, E. Mikolášková, P. Mikulíček, A. Mishra, P. Munclinger, D. Murariu, P. Nová, D. Peshev, F. Sedláček, F. Spitzenberger, M. Stanko, P. Suchomel, M. Šandera, L. Tomović, P. Vogel, J. Uhlíková, V. Vohralík, and P. Zupančič. This work was funded by the Royal Society of London/North Atlantic Treaty Organization Postdoctoral Fellowship grant (to P.K. and J.B.S.), Czech Science Foundation Grant 206/05/P032 (to P.K.), and Academy of Sciences of the Czech Republic Grant IRP IAPG AV0Z 50450515 (to P.K.).

1. Lessa EP, Cook JA, Patton JL (2003) *Proc Natl Acad Sci USA* 100:10331–10334.
2. Hewitt GM (2004) *Philos Trans R Soc London Ser B* 359:183–195.
3. Runck AM, Cook JA (2005) *Mol Ecol* 14:1445–1456.
4. Taberlet P, Fumagalli L, Wust-Saucy AG, Cosson JF (1998) *Mol Ecol* 7:453–464.
5. Hewitt GM (1999) *Biol J Linn Soc* 68:87–112.
6. Bilton DT, Mirol PM, Mascheretti S, Fredga K, Zima J, Searle JB (1998) *Proc R Soc London Ser B* 265:1219–1226.
7. Stewart JR, Lister AM (2001) *Trends Ecol Evol* 16:608–613.
8. Jankovská V, Chromy P, Nižnianská M (2002) *Acta Palaeobotanica* 42:39–50.
9. Willis KJ, van Andel TH (2004) *Q Sci Rev* 23:2369–2387.
10. Nadachowski A (1989) *Acta Theriologica* 34:37–53.
11. Horáček I (2000) *Geolines* 11:103–107.
12. Babik W, Branicki W, Crnobrnja-Isailovic J, Cogalniceanu D, Sas I, Olgun K, Poyarkov NA, Garcia-París M, Arntzen JW (2005) *Mol Ecol* 14:2475–2491.
13. Voros J, Alcobendas M, Martínez-Solano I, Garcia-Paris M (2006) *Mol Phylogenet Evol* 38:705–718.
14. Kotlík P, Berrebi P (2002) *Mol Phylogenet Evol* 24:10–18.
15. Janko K, Culling MA, Ráb P, Kotlík P (2005) *Mol Ecol* 14:2991–3004.
16. Raczynski J (1983) in *Ecology of the Bank Vole*, ed Petruszewicz K (Polish Scientific, Warsaw, Poland), pp 3–10.
17. Deffontaine V, Libois R, Kotlík P, Sommer R, Nieberding C, Paradis E, Searle JB, Michaux JR (2005) *Mol Ecol* 14:1727–1739.
18. Nadachowski A, Miekina B, Garapich A (2003) in *Oblazowa Cave: Human Activity, Stratigraphy, and Palaeoenvironment*, eds Valde-Nowak P, Nadachowski A, Madeyska T (Institute of Archaeology and Ethnology, Polish Academy of Sciences, Krakow, Poland), pp 134–140.
19. Rowe KC, Heske EJ, Brown PW, Paige KN (2004) *Proc Natl Acad Sci USA* 101:10355–10359.
20. Björkman L, Feurdean A, Wohlfarth B (2003) *Rev Palaeobot Palynol* 124:79–111.
21. Petit RJ, Brewer S, Bordsacs S, Burg K, Cheddadi R, Coart E, Cottrell J, Csaikl UM, van Dam B, et al. (2002) *For Ecol Manage* 156:49–74.
22. Heuertz M, Funeschi S, Anzidei M, Pastorelli R, Salvini D, Paule L, Frascaria-Lacoste N, Hardy OJ, Vekemans X, et al. (2004) *Mol Ecol* 13:3437–3452.
23. Jaarola M, Searle JB (2002) *Mol Ecol* 11:2613–2621.
24. Rendell S, Ennos RA (2002) *Mol Ecol* 11:69–78.
25. Pinceel J, Jordaens K, Pfenninger M, Backeljau T (2005) *Mol Ecol* 14:1133–1150.
26. Pearson RG (2006) *Trends Ecol Evol* 21:111–113.
27. Cook JA, Runck AM, Conroy CJ (2004) *Mol Phylogenet Evol* 30:767–777.
28. Stacy JE, Jorde PE, Steen H, Ims RA, Purvis A, Jakobsen KS (1997) *Mol Ecol* 6:751–759.
29. Guindon S, Gascuel O (2003) *Syst Biol* 52:696–704.
30. Hasegawa M, Kishino H, Yano T (1985) *J Mol Evol* 22:160–174.
31. Posada D, Crandall KA (1998) *Bioinformatics* 14:817–818.
32. Felsenstein J (1985) *Evolution (Lawrence, Kans)* 39:783–791.
33. Nielsen R, Wakeley J (2001) *Genetics* 158:885–896.
34. Hey J, Nielsen R (2004) *Genetics* 167:747–760.
35. Slatkin M (2005) *Mol Ecol* 14:67–73.
36. Won Y-J, Sivasundar A, Wang Y, Hey J (2005) *Proc Natl Acad Sci USA* 102:6581–6586.
37. Won Y-J, Hey J (2005) *Mol Biol Evol* 22:297–307.
38. Hey J (2005) *PLoS Biol* 3:1–11.
39. Conroy CJ, Cook JA (2000) *J Mammal* 81:344–359.
40. Galbreath KE, Cook JA (2004) *Mol Ecol* 13:135–148.
41. McDonald JH, Kreitman M (1991) *Nature* 351:652–654.
42. Rozas J, Sanchez-Delbarrio JC, Messeguer X, Rozas R (2003) *Bioinformatics* 19:2496–2497.