

Phylogeography of spruce beetles (*Dendroctonus rufipennis* Kirby) (Curculionidae: Scolytinae) in North America

LUANA S. MAROJA,* STEVEN M. BOGDANOWICZ,* KIMBERLY F. WALLIN,‡
KENNETH F. RAFFA† and RICHARD G. HARRISON*

*Department of Ecology and Evolutionary Biology, Corson Hall, Cornell University, Ithaca, NY 14853, USA, †Department of Entomology, University of Wisconsin-Madison, Madison, WI 53706, USA

Abstract

Tree-feeding insects that are widespread in north temperate regions are excellent models for studying how past glaciations have impacted differentiation and speciation. We used mitochondrial DNA (mtDNA) sequences and allele frequencies at nine microsatellite loci to examine genetic population structure across the current range of the spruce beetle (*Dendroctonus rufipennis*), an economically important insect in North America. Two major haplotype groups occur across northern North America, from Newfoundland to Alaska, on white spruce (*Picea glauca*), and a third distinctive haplotype group occurs throughout the Rocky Mountains on Engelmann spruce (*Picea engelmannii*). The two mtDNA lineages found in northern populations are 3–4% divergent from each other and from the lineages found in the Rocky Mountains. Analyses of microsatellite data also suggest the existence of major population groupings associated with different geographical regions. In the Pacific Northwest, concordant contact zones for genetically distinct populations of spruce beetles and their principal hosts appear to reflect recent secondary contact. Although we could detect no evidence of historical mtDNA gene flow between allopatric population groups, patterns of variation in the Pacific Northwest suggest recent hybridization and introgression. Together with the pollen record for spruce, they also suggest that beetles have spread from at least three glacial refugia. A minimum estimate of divergence time between the Rocky Mountain and northern populations was 1.7 Myr (million years), presumably reflecting the combined effects of isolation during multiple glacial cycles.

Keywords: bark beetles, divergence time, gene genealogy, glaciation, mtDNA, *Picea*, Pleistocene, speciation, spruce

Received 9 October 2006; revision received 23 January 2007; accepted 7 February 2007

Introduction

Fluctuations in climate during the Pleistocene produced extensive changes in the geographical location and continuity of habitats in both Europe and North America (Pielou 1991; Hewitt 1996, 2000, 2004). For many plants and animals, these cycles of habitat expansion and contraction (Webb & Bartlein 1992) resulted in repeated episodes of population subdivision and secondary contact. Geographical subdivision often gave rise to population differentiation (and, in some

cases, speciation; see Klicka & Zink 1997, 1998; Avise & Walker 1998; Avise *et al.* 1998; Johnson & Cicero 2004; for a debate on this issue), either as a result of random drift and lineage extinction or local adaptation to different environments (Hewitt 1996, 2000, 2004; Knowles & Richards 2005). Secondary contact produced hybrid zones, across which alleles could flow, potentially erasing differences that had arisen in allopatry (Harrison 1990; Hewitt 1996).

In many north temperate taxa, patterns of DNA sequence variation reveal a complex history of both diversification and reticulation and potentially allow inferences about the location of Pleistocene refugia and patterns of postglacial colonization (Hewitt 2001). Interpreting patterns of variation requires distinguishing between the 'deep' history that accounts for the origin of observed differences and the

Correspondence: Richard G. Harrison, Fax: (607) 255-8088;
E-mail: rgh4@cornell.edu

‡Present address: Department of Forest Science, Oregon State University, Corvallis, OR 97330, USA.

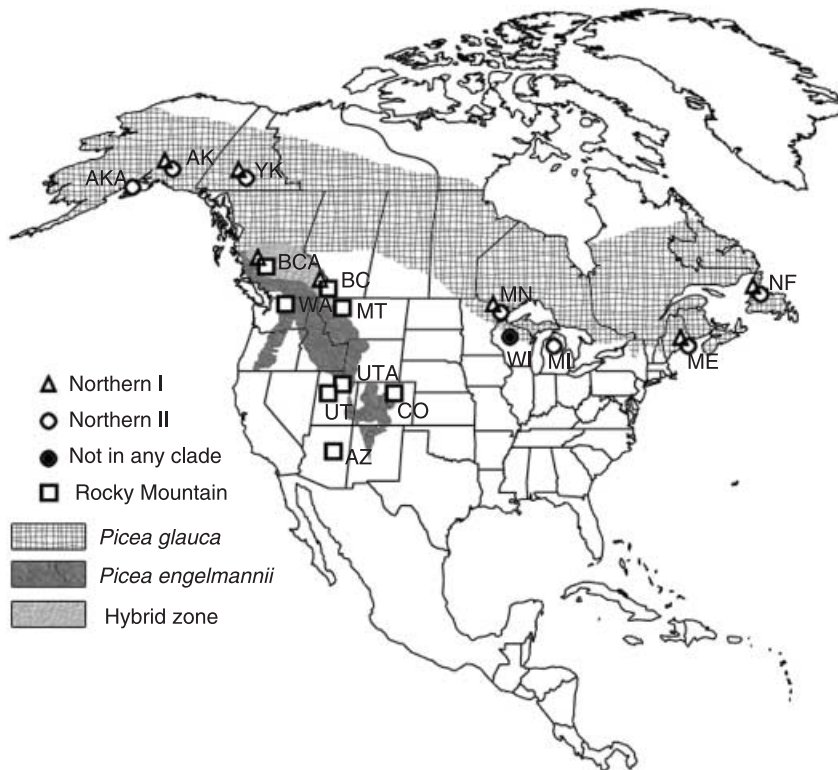


Fig. 1 Geographical distributions of white spruce and Engelmann spruce, together with sample localities of *Dendroctonus rufipennis*. The range of the spruce beetle corresponds to the ranges of the two host trees. Locality names are as given in Table 1. Representation of the three major mitochondrial clades at each locality is indicated by symbols in the key.

more recent history that explains the distribution of existing variation (Demboski & Cook 2001). Current distributions of alleles and haplotypes (and the location of genetic discontinuities or hybrid zones) reveal the signature of recent demographic events (e.g. Demboski & Cook 2001; Good & Sullivan 2001; Zamudio & Savage 2003; Miller *et al.* 2006). In contrast, the origin of many trait or sequence differences will trace to habitat fragmentation and isolation much earlier in the Pleistocene (or before), a history that is more difficult to decipher.

For many tree species in Europe and North America, pollen and microfossil records give direct information on distributions over the past 20 000 years (e.g. Davis 1976; Jackson *et al.* 2000). For phytophagous insects that feed on a limited range of host trees, distribution data for the hosts define likely distributions for the insects. Therefore, tree-feeding insects that are widespread in north temperate regions are ideal models for the study of how past glaciations have impacted differentiation and speciation.

The spruce beetle, *Dendroctonus rufipennis* (Curculionidae: Scolytinae), is distributed in a broad band across northern North America, from Alaska to Newfoundland (Wood 1982). Western populations extend southward into northern California and through the Rocky Mountains to Arizona and New Mexico (Fig. 1). The spruce beetle can feed, reproduce, and develop within many species of *Picea* (spruce trees) (Kelley & Farrell 1998), but the possibility of local or regional preference for individual spruce species has not

been investigated. The spruce beetle is one of the most economically important pest insects affecting North American forests, with outbreaks commonly causing near total mortality of spruce within the susceptible age classes (Werner *et al.* 1996; Holsten *et al.* 1999).

Based on pollen and microfossil records, spruce was distributed in the eastern United States from the ice margin south to Louisiana and was quickly able to occupy recently glaciated terrain, following the retreating Wisconsin ice front very closely (Braun 1950; Pielou 1991; Webb & Bartlein 1992; Jackson *et al.* 2000; see also <http://www.geobrown.edu/georesearch/esh/QE/Research/VegDynam/VegAnima/VegAnima.htm>). Molecular data for black spruce suggest the presence of several eastern refugia south of the glacial margin and possibly an additional refugium in the Canadian Maritime Provinces (Jaramillo-Correa *et al.* 2004). In the western United States, pollen records indicate that spruce persisted in the Pacific Northwest south of the Cordilleran ice sheet (Whitlock & Bartlein 1997); chloroplast DNA haplotype distributions suggest that white spruce also survived in Beringia (Anderson *et al.* 2006). Because *D. rufipennis* can survive well in cold habitats due to a complex life cycle that includes hibernation at both larval and adult stages (Wood 1982), the beetle probably closely tracked the distribution of its hosts during the Pleistocene.

The most abundant *Picea* host species within the current range of *D. rufipennis* are white spruce, *P. glauca*, and

Table 1 Collecting localities ($n = 16$) for *Dendroctonus rufipennis*, together with number of beetles sequenced for mtDNA (n_{mt}) and genotyped for microsatellites (n_{micro}). Also shown are haplotype (gene) diversity (H) \pm standard deviation (Nei 1987), nucleotide diversity (π) \pm standard deviation (Nei 1987) and average number of nucleotide differences within each population (K) for mtDNA. For the microsatellite data, average observed heterozygosity (H_O) and average number of alleles across nine microsatellite loci are shown.

State	Abbr.	Latitude (N)	Longitude (W)	n_{mt}	n_{micro}	$H \pm SD$	$\pi \pm SD$	K	H_O	Alleles
Alaska	AK	62°31'48"	150°3'0"	12	38	0.89 \pm 0.08	0.010 \pm 0.004	11.4	0.34	4.67
Alaska	AKA	60°37'48"	150°19'48"	11	115	0.89 \pm 0.09	0.007 \pm 0.001	8.2	0.36	6.11
Arizona	AZ	35°12'36"	111°39'36"	3	31	0.67 \pm 0.32	0.002 \pm 0.001	2.6	0.32	4.33
British Columbia	BC	54°43'48"	125°0'36"	14	30	0.94 \pm 0.04	0.022 \pm 0.005	25.0	0.44	4.89
British Columbia	BCA	50°15'0"	118°57'36"	2	30	1 \pm 0.5	0.046 \pm 0.022	51.0	0.45	4.44
Colorado	CO	40°31'12"	106°58'48"	3	30	1 \pm 0.07	0.004 \pm 0.001	4.7	0.39	4.89
Maine	ME	44°1'12"	68°-3'0"	5	32	0.9 \pm 0.16	0.022 \pm 0.004	24.4	0.30	2.44
Michigan	MI	45°39'0"	86°19'12"	2	—	1 \pm 0.5	0.003 \pm 0.001	3.0	—	—
Minnesota	MN	47°55'12"	90°33'0"	9	30	0.89 \pm 0.09	0.016 \pm 0.003	17.5	0.30	3.22
Montana	MT	48°19'48"	114°15'0"	2	23	1 \pm 0.5	0.004 \pm 0.002	4.0	0.42	4.89
Newfoundland	NF	49°16'48"	59°27'36"	5	36	0.8 \pm 0.03	0.018 \pm 0.005	19.6	0.31	3.89
Utah	UT	38°40'48"	112°10'12"	7	40	1 \pm 0.07	0.008 \pm 0.002	8.9	0.40	4.78
Utah	UTA	41°47'24"	111°48'0"	6	73	0.93 \pm 0.12	0.006 \pm 0.001	6.9	0.40	5.67
Washington	WA	48°34'12"	120°10'12"	3	22	0	0	0	0.5	5.00
Wisconsin	WI	45°37'36"	89°34'48"	1	—	—	—	—	—	—
Yukon	YK	60°49'48"	136°58'48"	8	26	0.89 \pm 0.11	0.012 \pm 0.004	13.3	0.35	4.78

Engelmann spruce, *P. engelmannii*. White spruce occurs across the northern boreal forest from Newfoundland and Labrador to Alaska. Engelmann spruce occurs along the eastern slope of the coast range and in the Cascades in the Pacific Northwest, and is also an important component of high-altitude forests in the Rocky Mountains, extending south to northern Arizona (Fig. 1). The two *Picea* species overlap in southern British Columbia, where hybrids have been described (Little 1971). Two other spruce species are widely distributed. *Picea mariana* (black spruce) has approximately the same distribution as white spruce, but is a suboptimal and largely nonpreferred host. *Picea pungens* (blue spruce) has a distribution similar to that of Engelmann spruce (Little 1971), but is an infrequent host for spruce beetle (Schmid & Frye 1977). All other spruce species have more restricted geographical distributions (Little 1971).

Here we use mitochondrial DNA (mtDNA) sequence data and allele frequencies at nine microsatellite loci to examine the genetic structure of *D. rufipennis* populations, estimate divergence times among distinct lineages, and infer patterns of postglacial range expansion and secondary contact. Deep divergences in the mtDNA genealogy suggest that lineages currently occupying different geographical regions have been isolated since the Early to Mid-Pleistocene. Patterns of variation for mtDNA and nuclear genes are concordant across much of the range of *D. rufipennis*, but beetles from regions where white spruce and Engelmann spruce overlap differ in their affinities based on the two types of molecular markers. The observed discordance may reflect differential introgression between distinct lineages that have recently come into secondary contact.

Materials and methods

Collection of insects

Cooperators collected *Dendroctonus rufipennis* adults from 16 localities in North America (Table 1, Fig. 1). These localities span the entire distribution of the spruce beetle and range from Alaska east to Newfoundland and south to Arizona. All individuals were collected as adults in 1998–2001 and stored in 100% ethanol until DNA isolation.

DNA isolation and sequencing

To generate mtDNA sequence data, genomic DNA was isolated from 93 whole beetles using the DNeasy tissue kit (QIAGEN). Primers TL2-N-3014 (5'-TCCAATGCACTAATCTGCCATATTA-3') and CI-J-1718 (5'-GGAGGATTTGGAAATTGATTAGTTCC-3') (Simon *et al.* 1994) were used to amplify a fragment of approximately 1.5 kb within the COI gene. Polymerase chain reactions (10 μ L volume) contained 3 mM $MgCl_2$, 0.2 mM dNTPs, 50 mM KCl, 20 mM Tris (pH 8.4), 2.5 ng of each primer and 1 U of *Taq* DNA polymerase (Gibco-BRL) and 1 μ L DNA. PCR amplifications were performed using a thermalcycler (OmniGene, Hybaid) under the following conditions: 35 cycles of 50 s at 95 °C, 60 s at 48 °C and 90 s at 72 °C.

The fragments were sequenced with an ABI PRISM 377 automated sequencer using BigDye terminator labelling (Applied Biosystems). *D. rufipennis* specific primers BBA2365 (5'-CGGATGTGAAGTAGGCTCGGGTAT-3') and BBS2339 (5'-TAGGATTTGTAGTATGGGCCACC-3') were used

for sequencing, in addition to the primers TL2-N-3014 and CI-J-1718. Sequences were aligned with the program LASERGENE MEGALIGN version 5.05 (DNASTAR) using the CLUSTAL method. The program DNASP version 3.14 (Rozas & Rozas 1999) was used to calculate average G + C content, total number of synonymous and segregating sites, the average number of nucleotide differences within and between populations, nucleotide diversity (π), and haplotype (gene) diversity, using Nei's (1987) equations. The neutrality of COI evolution was tested with Tajima's D (Tajima 1989) and Fu and Li's F^* test as implemented by the program DNASP version 3.14 (Rozas & Rozas 1999). Unique mitochondrial sequences used in this study have been deposited in GenBank (Accession nos DQ165383–DQ165449).

Phylogeny reconstruction

Phylogeny reconstruction was carried out using PAUP* version 4.0b10 (Swofford 2002). We reduced the data to a set of unique haplotypes and used maximum parsimony (MP; Fitch 1977) to generate phylogenetic trees. MP trees were generated using multiple heuristic searches, with random addition and (tree-bisection–reconnection) TBR branch swapping. Bootstrap proportions were estimated using 100 bootstrap replicates in PAUP*. We also evaluated clade probabilities using a Bayesian approach. Using the program MRBAYES version 3.0 (Huelsenbeck & Ronquist 2001), searches were run for 10 million generations, sampling every 100 generations and discarding the first 1 million generations (burn-in time). We used default priors and the general time reversible model with invariant sites and gamma rates (GTR + I + G), allowing the rate at each site to change over evolutionary history. Phylogenetic trees were rooted with the COI sequence of the sister species *Dendroctonus murrayanae* (Kelley & Farrell 1998).

Migration and divergence time between populations

We used the program MDIV (Nielsen & Wakeley 2001) with mtDNA data to distinguish between models of isolation with and without gene flow, and to estimate divergence times between *D. rufipennis* populations living on different host trees. MDIV calculates a likelihood function using a Metropolis–Hastings MCMC approach and has been shown to obtain reliable joint estimates of migration rates and divergence times of two populations from single-locus DNA sequence data (Nielsen & Wakeley 2001). The program gives an estimate of migration ($2Nm$) and divergence time from the common ancestral population ($T = t/N$), where N is the effective female population size (for mitochondrial DNA), m is the proportion of individuals from one population that are replaced by individuals from the other population each generation, and t is the divergence time in

generations. We used a finite sites HKY model (there were five sites that did not follow the infinite sites model). The integrated likelihood function surfaces were obtained using 2 million steps in the chain and a burn-in time of 500 000 steps.

We pooled individuals that used white spruce as a host (AK, AKA, ME, MI, MN, NF, WI and YK, total = 53 sequences) and individuals that used Engelmann spruce (AZ, CO, MT, UT, UTA and WA, total = 24 sequences). We excluded populations located within the white and Engelmann spruce hybrid zone, BC and BCA, because the tree species in which the beetles developed was not known.

Microsatellite development and analysis

To develop microsatellite markers, we started with genomic DNA from a pool of adult *D. rufipennis* from the Alaska population (AK) and followed a protocol modified from Hamilton *et al.* (1999). Following digestion of genomic DNA with restriction enzymes and ligation of linkers, we enriched for fragments containing microsatellites using hybridization with biotinylated dimeric and trimeric oligonucleotides followed by magnetic capture. The enriched fraction was ligated into pUC19, and aliquots of the ligation were used to transform *Escherichia coli* DH5 α cells. Colonies were plated out, transferred to nylon membranes, and probed with radiolabelled oligonucleotides (the same used in the enrichment step). Insert sequences from positive colonies were obtained using universal M13 primers that flank the cloning site. Using this protocol we identified 11 genomic regions containing microsatellites that were polymorphic in spruce beetle populations. Repeat motifs, primer sequences, and estimates of variability are shown in Table 2.

For analysis of microsatellite allele frequencies in spruce beetle populations, we isolated DNA from an additional 466 beetles, yielding a total sample of 556 individuals from 14 populations. Fragment analyses were conducted with an ABI PRISM 377 with standard loading and electrophoresis conditions. Alleles were sized relative to an internal size standard and analysed with GENESCAN 2.1 (Applied Biosystems).

Linkage disequilibrium, deviations from Hardy–Weinberg (HW) equilibrium and estimates of F_{ST} were calculated with GENEPOP web version 3.4 (Raymond & Rousset 1995). Linkage disequilibrium between loci was tested with Fisher's exact test using a Markov chain (10 000 dememorizations; 10 000 iterations) under the null hypothesis of no genotype association between pairs of loci. Deviations from HW equilibrium at each locus were tested by a Markov chain approximation (10 000 dememorizations; 10 000 iterations) of unbiased exact P values (Guo & Thompson 1992).

Table 2 Microsatellite repeat motif, primer sequences, GenBank Accession number and estimates of variability [mean expected heterozygosity (H_E) and standard deviation and mean and total number of alleles considering the 14 analysed populations]

Locus	Motif*	Forward primer	Reverse primer	Accession no.	H_E	No. of alleles
Loc1	CCA	TGCCACCGGTCAGGTTTACTC	TGACGACGTCCACAATCACATCTTC	DQ165372	0.42 ± 0.18	4.0 (8)
Loc15	AC	GATGGGCGCTTTTCTGCCACTTA	CATTTTGGGACGGTTTTTCTTG	DQ165373	0.7 ± 0.17	7.4 (12)
Loc17	ACT	AAAGGGTGGAGAGCAACTGGATAC	CGCTTTGGTTCCCTATCTTCTACG	DQ165374	0.37 ± 0.12	2.6 (6)
Loc18	ACT	GATCCACTACTCCTACCCTTG	CAACTGATACAATGTGATAAAAACAGC	DQ165375	0.14 ± 0.11	3.0 (6)
Loc24	AG	CTGGGACGGCCGAGACTCAG	ACAATTGTTTCAGCAGCGGAATCTT	DQ165376	0.79 ± 0.21	18.3 (52)
Loc33	AAG	GGGCCAAATCACA AAAACCAGAAAC	TCGAAATATTGTGGCGGCATCTTTAT	DQ165377	0.66 ± 0.14	6.1 (13)
Loc52	AAG	GGATCACATTGCATGACAGGATACG	CAITTTCAAGCAGAAGCTTCAGATTTCAG	DQ165378	0.06 ± 0.10	1.8 (6)
Loc54	AGG	CTGCCACATCAGCATAACAAAAC	GTGGCCGTGACCCATTC	DQ165379	0.6 ± 0.19	6.2 (11)
Loc56	AAG	TATGCGCTATGGTGACTTGTTCCT	AAGAGGCTACATCGTCAGTGGAT	DQ165380	0.31 ± 0.20	3.4 (11)
Loc65	AC	GAAACATACAATTCGAATGTG	TTCAGCGCAATGAAAAC	DQ165381	0.8 ± 0.08	10.1 (20)
Loc76	AC/AT	ATTATCTTCGCTCTGTCAGTCACTTGA	TCCAATTGATTCACAGTTTAGGATACT	DQ165382	0.22 ± 0.23	2.6 (7)

*Some of the microsatellites have complex motifs — refer to GenBank for motif length and complexity.

Microsatellite population structure

The Wright's F_{ST} analogue (θ) of Weir & Cockerham (1984) was used to assess the level of population differentiation. Significant genotypic differentiation between populations was tested with the log-likelihood G statistic with genotype permutations rather than allele permutations, to avoid the requirement that populations be in HW equilibrium (Goudet *et al.* 1996).

We built an unrooted (neighbour-joining) NJ tree using Cavalli-Sforza chord distances (D_{ce}) (Cavalli-Sforza & Edwards 1967; Takezaki & Nei 1996) with PHYLIP 3.65 (Felsenstein 2005). To calculate bootstrap support, we used PHYLIP 3.65 modules SEQBOOT, GENDIST, NEIGHBOUR and CONSENSE (Felsenstein 2005) with D_{ce} distances and 1000 bootstrap replicates.

We used STRUCTURE 2.1 (Pritchard *et al.* 2000; Falush *et al.* 2003) to identify clusters with distinct allele frequencies and to assign individuals to each of the clusters, without using prior information on collection sites. This program uses a Bayesian approach to assign individuals to clusters based on their genotypes while simultaneously estimating population (cluster) allele frequencies (Pritchard *et al.* 2000). Using the model with admixture, individuals can have membership coefficients (ancestry estimates) in multiple clusters, with the sum of the coefficients equal to one. The program estimates, the proportion of the genome of each individual that comes from cluster K . To choose the appropriate number of clusters for the data set we ran a series of independent runs with K values from 1 to 14 and then calculated the probability for each K [$P(K|X)$] given a flat prior. We used the admixture ancestry model with correlated allele frequencies; this is a more realistic model because migration and mixed ancestry contribute to patterns of variation in the analysed

populations (Falush *et al.* 2003). All runs had a burn-in period of 30 000 iterations and data collection for 5×10^5 iterations. All summary statistics stabilized before the end of the burn-in period, and independent runs always attained the same results.

The program DISTRUCT (Rosenberg 2004) was used for graphical display of population structure. This program depicts each cluster in a different colour and each individual as a line segment partitioned into K coloured components (the lengths of the components being proportional to the membership coefficients in the K clusters).

Isolation by distance

Isolation by distance was calculated for both mtDNA and microsatellite data. Natural logarithms of geographical linear distances (km) between localities were used for the calculation of isolation by distance. The Mantel test (Mantel 1967) was used to determine whether there was a significant relationship between geographical distance and pairwise genetic differentiation. This test was performed with 1000 random iterations to obtain statistical inferences at $\alpha = 5\%$ using the program MANTEL NON-PARAMETRIC CALCULATOR version 2.0 (Liedloff 1999) for the mtDNA data, and the ISOLDE option of GENEPOP web version 3.4 for the microsatellite data.

For the mtDNA data, genetic distances were calculated with PAUP* (Swofford 2002) using the GTR settings and estimating all parameters. If more than one haplotype was present at a given locality, the weighted mean distance between haplotypes from the two localities was used. The genetic distance within each locality was set to zero even if the locality had several distinct haplotypes. For microsatellite data, differentiation statistics were linearized based on the stepping-stone model (Rousset 1997) as $\theta/(1-\theta)$.

Results

mtDNA variation

We sequenced 1114 bp of mtDNA from 93 *Dendroctonus rufipennis* and one individual of the sister species *Dendroctonus murrayanae*. The sequenced region falls entirely within the COI gene, and the first base pair in our sequences corresponds to a first codon position (based on comparisons with the *Drosophila melanogaster* sequence). No stop codons or indels were detected in any sequences. No heterogeneity in base composition was detected. Of the 1114 bp, 132 sites were polymorphic, 103 of which were parsimony informative. Excluding the outgroup, there were 138 inferred mutations. Most of the polymorphic sites were first codon or third codon positions (11.53% and 86.46%, respectively). The majority of inferred substitutions were transitions (413 out of 446), and 93% of the mutations represented synonymous substitutions.

Of the 94 haplotypes (including the outgroup, *D. murrayanae*), 66 were unique. All populations except for WA and WI (for which only one individual was sampled) had more than one haplotype (Table 1). AK, AKA and YK shared one haplotype (10 identical sequences), AK and YK shared a second haplotype (3 identical sequences), and MI and MN shared one haplotype (three identical sequences). All other nonunique haplotypes were shared between individuals from the same locality.

Table 1 reports haplotype diversity (H) and nucleotide diversity (π) (Nei 1987), and the average number of nucleotide differences (K) between haplotypes within populations. The populations with the greatest nucleotide diversity and the largest number of average nucleotide differences were BC and BCA (both located close to or within the spruce hybrid zone), followed by NF and ME. Neither the Tajima's D nor the Fu and Li's F^* tests were significant for the major clades considered separately or for all individuals together.

mtDNA genealogy

A maximum parsimony (MP) heuristic search produced 8417 equally parsimonious trees 360 steps long (Fig. 2). These trees differ only in placement of haplotypes within major groups or clades. Three major haplotype groups were identified (Fig. 2), which we refer to as the 'Northern I' group (individuals from AK, BC, BCA, ME, MN, NF, and YK), the 'Northern II' group (individuals from AK, AKA, ME, MI, MN, NF and YK), and the 'Rocky Mountain' group (individuals from AZ, BC, BCA, CO, MT, UT, UTA, and WA). The Rocky Mountain clade is composed of two distinct subclades identified by all reconstruction methods, one comprising haplotypes found in the AZ, UT, and UTA populations and the second including haplotypes from the

BC, BCA, CO, MT, and WA populations. Figure 2 also includes bootstrap support values (those > 0.8). In the majority rule bootstrap tree, Northern I and Rocky Mountain are both well-supported clades, but Northern II is a paraphyletic assemblage.

The Bayesian posterior probability approach, using the GTR + I + G model, produced a tree with the same major haplotype groups as those obtained by MP (Fig. 3). This tree has a topology in which the Northern II group is basal and paraphyletic with respect to Northern I and Rocky Mountain clades. In the bootstrap MP tree, the Northern I clade is basal, but support for this topology is weak. When considering only strongly supported nodes, the MP and Bayesian approaches produce very similar results.

The mean GTR distance within groups is 0.004 (± 0.001) for Northern I, 0.009 (± 0.003) for Northern II, and 0.013 (± 0.01) for the Rocky Mountain clade. The two Rocky Mountain clades exhibit, on average, 2.4% sequence divergence (uncorrected). Among the major haplotype groups, sequence divergence between the Northern I and Northern II groups is 3%, whereas the Northern groups are 4% and 3.5% divergent from the Rocky Mountain clade, respectively.

The two northern clades are broadly sympatric throughout the range of *Picea glauca*, which is distributed across northern North America from Newfoundland to Alaska (Fig. 1). In contrast, the Rocky Mountain clade is associated with *Picea engelmannii*, which ranges from British Columbia south in the western United States to Arizona and New Mexico (Fig. 1). In southern British Columbia (localities BC and BCA), individuals with Northern I clade haplotypes co-occur with individuals carrying Rocky Mountain haplotypes. These are the only localities where such an overlap has been found.

Migration and divergence time between populations

We used $MDIV$ (Nielsen & Wakeley 2001) with mtDNA data to generate integrated likelihood functions for M ($2Nm$) and T (t/N) (Fig. 4) for currently allopatric beetle populations associated with white and Engelmann spruce. The data are most compatible with a model of no migration ($M=0$) between these two populations and with a value of $T \sim 1.3$. The value of θ calculated using $MDIV$ is 0.03.

To estimate divergence time (t), we assumed one generation per year and used the oft-cited estimate of 1.1% per million years per lineage for arthropod mitochondrial DNA (Brower 1994), which corresponds to a base rate of 1.1×10^{-8} substitutions/site/year or 1.2×10^{-5} substitutions/region/year for the analysed region (1114 bp). This gives an estimate of 1.7 million years (Myr) for the divergence between populations living on white spruce and Engelmann spruce. If we assume one generation every 2 years (as in much of the range of *D. rufipennis*), this estimate doubles to 3.4 Myr. $MDIV$ assumes constant population size

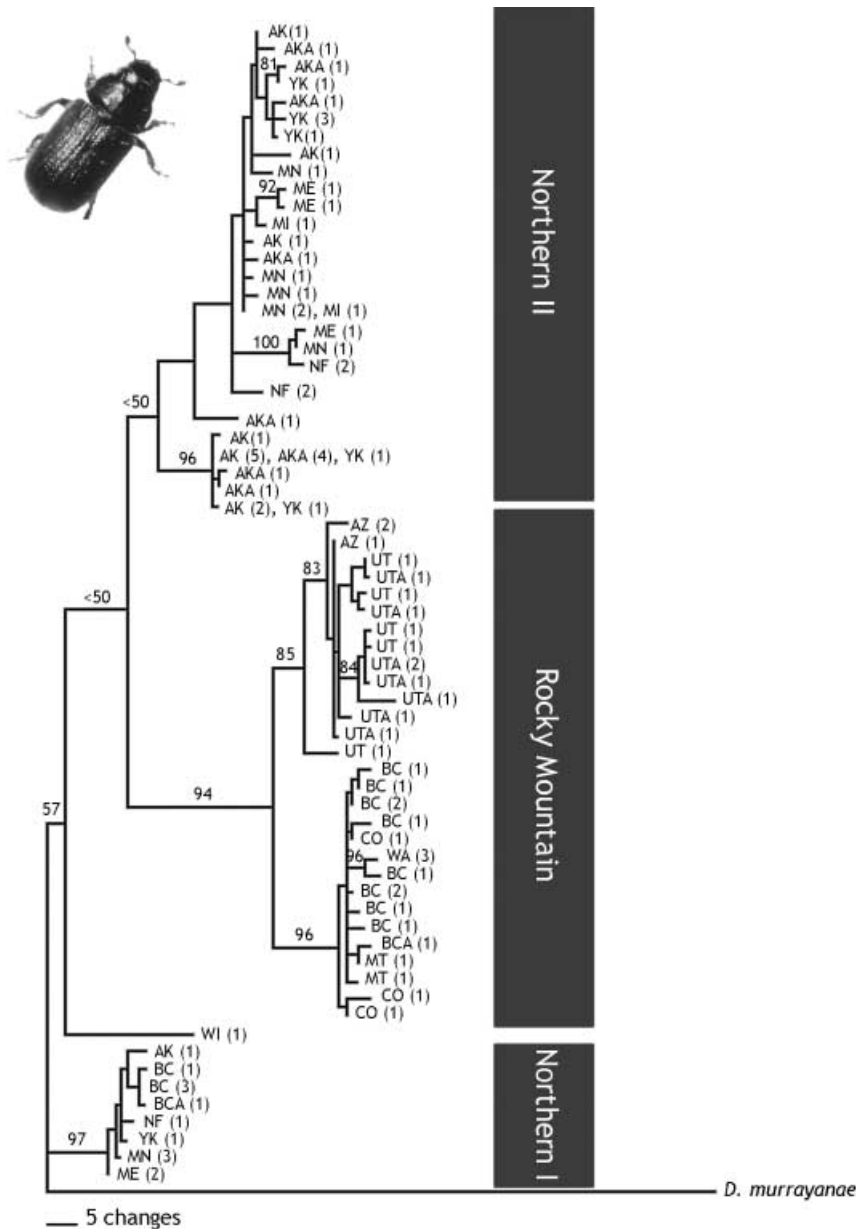


Fig. 2 One of the 8417 most parsimonious trees (360 steps long) based on COI sequence of *Dendroctonus rufipennis*. The many MP trees differ only in how haplotypes are arranged within the three major haplotype groups. Phylogeny reconstruction was carried out with PAUP* (Swofford 2002). Numbers in parentheses show the number of identical haplotypes. Also shown (along the branches) are bootstrap proportions estimated using 100 bootstrap replicates in PAUP*. See text for details.

and yet beetle populations have experienced fluctuations in size due both to climate change and local outbreaks. This violation of assumptions is unlikely to affect qualitative estimates of migration rate (R. Nielsen, personal communication) but could affect estimates of time since divergence.

Variation at microsatellite loci

We estimated allele frequencies in 14 populations for 11 microsatellite loci (MI and WI populations were not analysed due to small sample sizes; Table 1). However, Locus 24 and Locus 33 exhibited rare alleles in homozygous condition and were out of HW equilibrium for at least four

populations (and not only those in which admixture might be expected). We therefore excluded these loci from all analyses. No linkage disequilibrium was observed in the analyses of individual populations or in a global test across all populations.

Of the nine loci used in the analyses, three showed one population out of HW equilibrium even after Bonferroni correction. All three populations out of HW equilibrium exhibited a deficiency of heterozygotes. Populations from ME and MN had the smallest average number of alleles and observed heterozygosity of all populations (Table 1). However, differences among populations in heterozygosity were not significant ($\chi^2 = 0.148$, d.f. = 1, $P > 0.7$).

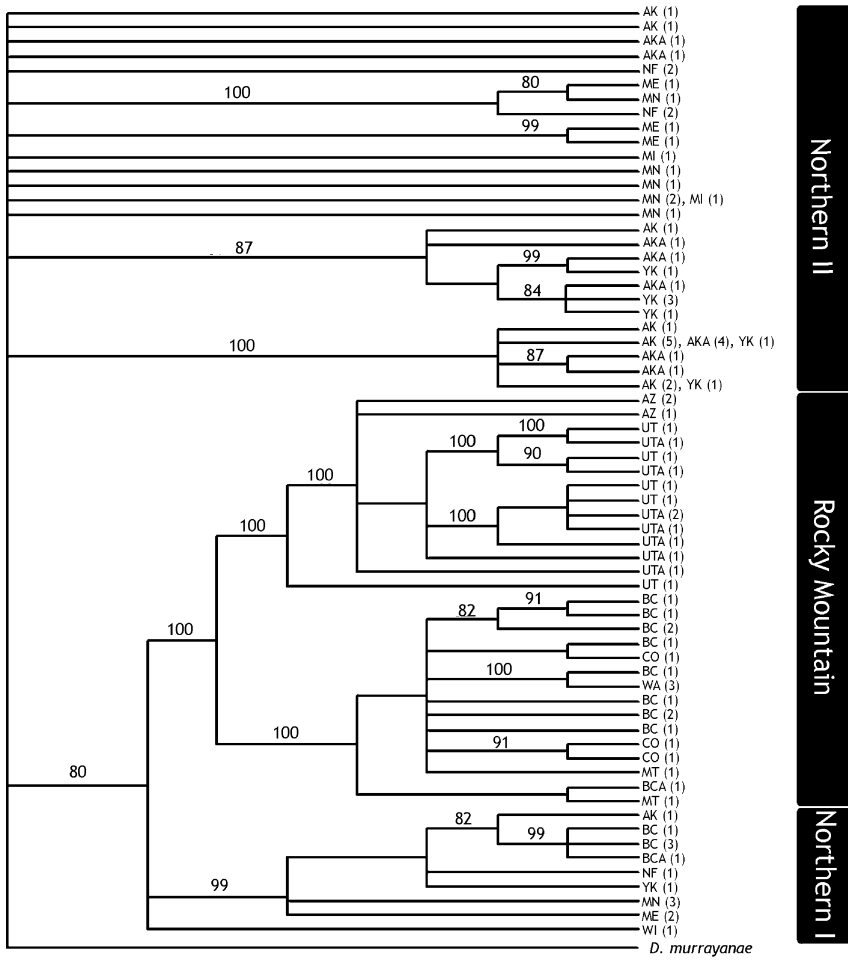


Fig. 3 Posterior probability tree using the coding region (1114 bp) of COI. The search with the program MRBAYES (Huelsenbeck & Ronquist 2001) was run for 10 million generations, discarding the first one million generations. We used default priors, a GTR model, invariant sites and gamma rates.

Table 3 Wright's F_{ST} analogue (θ) of Weir & Cockerham (1984), together with significance value (***) $P < 0.0001$

	AK	AKA	AZ	BCA	BC	CO	ME	MN	MT	NF	UT	UTA	WA
AKA	-0.001												
AZ	0.234***	0.264***											
BCA	0.075***	0.085***	0.253***										
BC	0.076***	0.085***	0.228***	0.012									
CO	0.247***	0.279***	0.016	0.253***	0.233***								
ME	0.133***	0.131***	0.264***	0.178***	0.185***	0.271***							
MN	0.057***	0.056***	0.290***	0.140***	0.160***	0.309***	0.107***						
MT	0.045***	0.052***	0.228***	0.010	0.008	0.230***	0.149***	0.12***					
NF	0.017	0.021***	0.281***	0.091***	0.100***	0.300***	0.126***	0.031***	0.061***				
UT	0.284***	0.311***	0.026	0.285***	0.267***	0.037***	0.310***	0.332***	0.267***	0.329***			
UTA	0.241***	0.270***	0.013	0.254***	0.234***	0.028***	0.254***	0.282***	0.232***	0.282***	0.001		
WA	0.057***	0.067***	0.186***	0.016	0.017	0.184***	0.150***	0.127***	0.005	0.080***	0.228***	0.199***	
YK	0.002	0.013	0.216***	0.055***	0.057***	0.231***	0.109***	0.046***	0.031	0.008	0.265***	0.224***	0.042***

Genetic population structure based on microsatellites

Significant allelic differentiation was found over all loci and populations ($P = 0.0001$), reflecting the presence of three major population groupings, within which F_{ST} values

are small and mostly nonsignificant (Table 3). The three population groupings are groups: (1) AZ, CO, UT, UTA; (2) AK, AKA, YK, NF and (3) BC, BCA, MT, WA. The populations from ME and MN are most closely aligned with the group 2 populations, but pairwise F_{ST} values are higher than

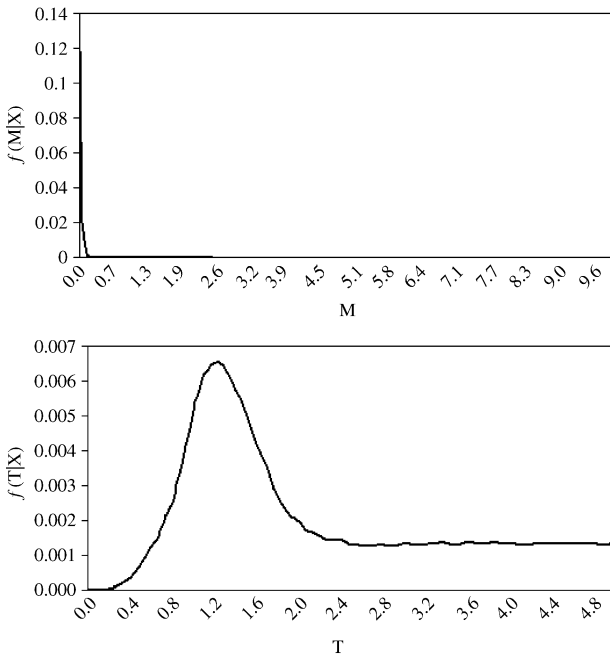


Fig. 4 Posterior distribution of migration ($M = 2Nm$) and divergence time ($T = t/N$) between populations living in white and Engelmann spruce. Posterior distributions were calculated with the program *MDIV* (Nielsen & Wakeley 2001) using an HKY model and 2 million steps (discarding the first 500 000 steps).

comparable values for other pairwise comparisons within that group.

The microsatellite distance tree (Fig. 5) revealed two distinct clusters. One cluster includes only the ‘southern’ Rocky Mountain populations AZ, CO, UT and UTA. The second cluster includes AK, AKA, ME, MN, NF, YK in one subcluster and BC, BCA, MT, WA in a second subcluster. The major difference from population relationships based on mtDNA is that MT and WA are part of the Northern group rather than the Rocky Mountain group. The ‘mixed’

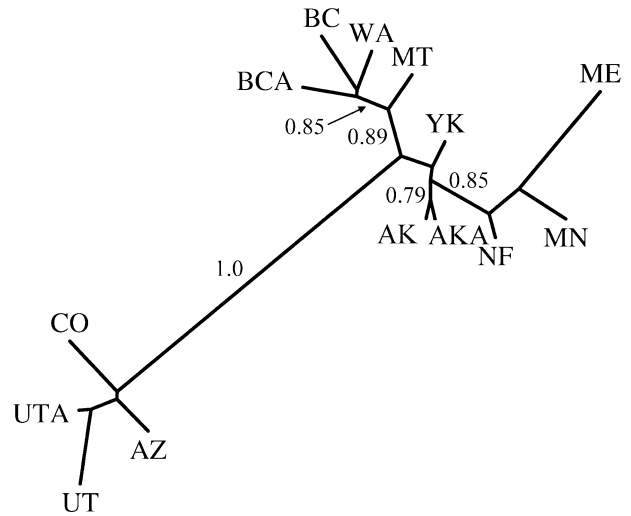


Fig. 5 Microsatellite (neighbour-joining) NJ tree for 14 *Dendroctonus rufipennis* populations. Bootstrap values > 0.75 are shown. The tree is based on Cavalli-Sforza chord distances using nine microsatellite loci. See text for details.

mtDNA populations BC and BCA are most similar to MT and WA and also affiliate with the Northern cluster.

Using *STRUCTURE*, and assuming a flat prior distribution and $K = 1-14$, the model with $K = 5$ had the highest posterior probability ($\Pr(K = 5 | X) \sim 1$). However, following Evanno *et al.* (2005), we also examined the rate of change in the log probability of the data between successive K values, which revealed little change between $K = 3$ and $K = 5$. Figure 6 shows proportional membership in all clusters for each individual beetle, with individuals sorted by populations. The separate panels show these results for $K = 3$, $K = 4$, and $K = 5$. The Rocky Mountain populations (AZ, CO, UT and UTA) are the most uniform and distinctive populations with respect to cluster assignment. Nearly all individuals from these populations have genotypes that

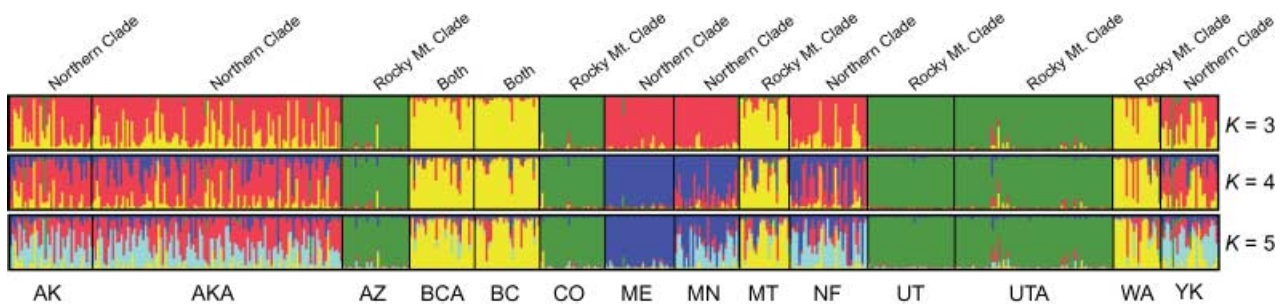


Fig. 6 Estimated population structure based on microsatellite allele frequency variation. Each individual is represented by a vertical line which is partitioned into K coloured components representing the ancestry fractions in $K = 3$, $K = 4$, or $K = 5$ clusters. Individuals are ordered by collection site and black lines separate individuals from different populations. Populations are labelled below the figure and mitochondrial clades for each population are labelled above the figure. For simplicity, Northern Clades I and II are both labelled as Northern Clade.

indicate membership in a single cluster (green), independent of K value. Few individuals within these populations are assigned to multiple clusters. For $K = 5$, individuals from the northern populations AK, AKA, NF and YK are assigned primarily to two clusters (Fig. 6, red and light blue), although NF has many individuals assigned to the dark blue cluster and YK has a larger representation of individuals assigned to the yellow cluster. Individuals from the ME population are assigned mainly to one cluster (dark blue), and those from MN primarily to dark blue and light blue clusters. For $K = 3$, individuals in all of the northern populations are assigned to red and yellow clusters, with the largest membership coefficient in the red cluster. Beetles from the northern Rocky Mountains (MT, WA, BC and BCA) are principally assigned to a single cluster (yellow), although some individuals are also assigned to other clusters (mainly to the 'northern clusters'). Groupings based on STRUCTURE analysis (Fig. 6) are essentially the same as those defined by the distance tree (Fig. 5).

To test if beetles from the sympatric Northern I and Northern II mtDNA groups interbreed at random, we examined allele frequency data for individuals for which both sequence and genotype data were available. Individuals from the Northern I (38 individuals) and Northern II (13 individuals) groups were treated as distinct populations. The microsatellite-based F_{ST} value between these mtDNA haplotype groups was nonsignificant ($F_{ST} = -0.005$), suggesting that Northern I and II groups defined by mtDNA are not differentiated with respect to their nuclear DNA.

Isolation by distance

Using the mtDNA data, we detected a strong signal of isolation by distance when all populations are considered ($g = 4.1$ $P = 0.001$). We also found significant isolation by distance across all populations ($P = 0.012$) based on microsatellite allele frequency data. However, when populations are grouped by region and/or mtDNA clade, significant isolation by distance was found only for the mtDNA data in populations from the Northern II clade ($g = 3.9$, $P = 0.002$). Significant isolation by distance was not found within the Rocky Mountain clade, for populations from the Northern I clade, or when all individuals from northern populations are included (i.e. combining Northern I and II).

Discussion

Genetic structure of spruce beetle populations: Pleistocene refugia, postglacial colonization, and secondary contact

The mtDNA sequence data define three distinct haplotype groups within the widely distributed North American bark beetle, *Dendroctonus rufipennis*. Beetles in the two Northern haplotype groups are broadly sympatric (and

probably randomly interbreeding), but are largely allopatric with respect to beetles in the Rocky Mountain clade. Microsatellite allele frequencies (see Figs 5 and 6) also suggest the existence of three (or more) major population groupings, within each of which pairwise F_{ST} values are small (Table 3). Both mtDNA sequence data and microsatellite allele frequency data show that spruce beetles from Northern and Rocky Mountain populations are genetically distinct. However, mtDNA and microsatellites differ in how they define the affinities of beetles collected from within or adjacent to the region where white and Engelmann spruce co-occur. Mitochondrial DNA places beetles from the WA and MT populations (and most beetles from the two British Columbia populations) with Rocky Mountain beetles to the south, whereas the nuclear gene markers reveal distinctive and very similar allele frequencies in samples from BC, BCA, MT, and WA, and a much closer affiliation of these populations with those to the north.

Southern British Columbia has been identified as an area in which previously allopatric populations or species have come together in secondary contact subsequent to the last glacial maximum. Remington (1968) recognized a distinct 'suture zone' in southern British Columbia, based on the geographical distributions of zones of overlap and hybridization between closely related species. Our population samples BC and BCA are located close to or within suture zone A identified by Remington (1968). However, the 'reality' of many of Remington's suture zones (including suture zone A) has been questioned by Swenson & Howard (2004), and the clustering of hybrid zones may not be as discrete or extensive as originally proposed. Nonetheless, 'hotspots' of hybrid zone clustering do occur in the Pacific Northwest (Swenson & Howard 2004), and many plant and animal species exhibit north-south discontinuities in this region (e.g. Soltis *et al.* 1997; Good & Sullivan 2001; Good *et al.* 2003; Burg *et al.* 2005; Miller *et al.* 2006). In most of these examples, current phylogeographical structure appears to result from secondary contact between populations expanding from refugia south of the glacial boundary and from more northern refugia (e.g. Queen Charlotte Islands, Beringia).

From our mtDNA data, southern British Columbia and the Pacific Northwest of the United States appear to represent a transition zone for spruce beetles with respect to haplotype frequencies. This region coincides with the current zone of overlap and hybridization between the spruce beetles' principal host trees, *Picea glauca* and *Picea engelmannii* (Little 1971). The region of overlap for the spruces extends south of the documented hybrid zone in British Columbia, with stands of white spruce occurring in Washington, Idaho, Montana, and Wyoming, within the range of Engelmann spruce (see <http://plants.usda.gov/java/profile?symbol=PIGL> for distribution maps). At both of our sampled sites in British Columbia, beetles that fall within the Rocky Mountain and Northern I haplotype groups

are found together, whereas individuals from the WA and MT populations all carry Rocky Mountain haplotypes. The Rocky Mountain haplotypes found in the Pacific Northwest and British Columbia are similar to those found in Colorado (see Figs 2 and 3), suggesting that a northward postglacial migration route probably followed the eastern slope of the Rocky Mountains, with no evidence of dispersal from further west (Utah and Arizona beetles have very distinctive mtDNA haplotypes that are not found to the north).

Patterns of microsatellite allele frequency variation in the northern Rocky Mountains and Cascades reveal homogeneity within this region, implying recent population expansion and/or ongoing gene flow. Although beetles from the Pacific Northwest share mtDNA haplotypes with beetles to the north and south, microsatellite allele frequencies suggest that the BC, BCA, MT, and WA populations are distinctive (e.g. for $K = 5$, STRUCTURE reveals that the majority of the nuclear genome of Pacific Northwest beetles is assigned to the 'yellow cluster', whereas no other populations show substantial membership in that cluster). Thus, the Pacific Northwest may represent a region in which beetles from multiple refugia (Beringia, Rocky Mountains, northwest United States, and perhaps eastern United States) have come together.

Pollen data tend to support this hypothesis. Spruce pollen was found abundantly south of the glaciers in the central and eastern United States 21 000 BP, but also was present in both Beringia and the northwestern United States. Indeed, POLLEN VIEWER (<http://www.geo.brown.edu/georesearch/esh/QE/Research/VegDynam/VegAnima/PV31Inst.htm>) suggests that contact between expanding spruce populations from Beringia and the eastern United States (both perhaps associated with white spruce) may have occurred in southwestern Canada or northwestern United States as early as 12–13 000 BP, and that contact with spruce in the Pacific Northwest occurred shortly thereafter. Chloroplast DNA haplotype frequencies in white spruce also argue for secondary contact between populations dispersing from southern and Beringial refugia (Anderson *et al.* 2006).

Field studies indicate that spruce beetles can disperse from 90 to 300 m from their overwintering sites (Werner & Holsten 1997). Given this dispersal ability and their adaptations for cold climates, the range expansion of spruce beetles likely paralleled that of their hosts. Thus, the pollen record, together with the concordant contact zones for genetically distinct populations of spruce beetles and their principal hosts, strongly suggest that these zones are the result of recent secondary contact. Beetles appear to have spread from a minimum of three glacial refugia: (i) from south of the glacial boundary in the central/eastern United States, beetles (carrying one of the Northern mtDNA haplotypes) spread northeast into New England and the Maritime Provinces of Canada and northwest into western

Canada and Alaska; (ii) from Beringia, beetles (carrying the alternative Northern mtDNA haplotype) spread south and east; and (iii) from south of the glacial boundary in the Pacific Northwest and/or from farther south in the Rocky Mountains, beetles (some at least carrying one of the Rocky Mountain haplotypes) spread north.

Although the current distribution of haplotype groups almost certainly reflects recent postglacial colonization, their initial divergence is much older. Our minimum estimate of the divergence time between subpopulations on white and Engelmann spruce is ~1.7 Myr, suggesting that these beetle lineages initially became isolated in the Early Pleistocene (or before). This is a minimum estimate, because it assumes one generation per year, whereas spruce beetles often require 2 years to complete their development (Hansen *et al.* 2001). Repeated episodes of habitat fragmentation have presumably allowed divergent lineages to persist, and regional patterns of differentiation appear to have been maintained in spite of possible opportunities for gene flow to homogenize allele or haplotype frequencies. Isolation for 1–2 my likely provided many opportunities for divergence and local adaptation. Our phylogeographical data provide the appropriate context for examining ecological, behavioural, and physiological differences among beetles from different regions.

Deep divergences within single 'species' have also been found in other taxa (e.g. Hewitt 1996, 2000; Santucci *et al.* 1998; Smith & Farrell 2005), and in these cases, patterns of variation have generally been attributed to the combined effect of several glacial cycles, during each of which allopatric populations diverged. Our data for spruce beetles are similar in some respects to those for another *Dendroctonus* species, *D. brevicomis*, in which eastern (Colorado) and western (California) populations exhibit an mtDNA sequence divergence of 6.9% (Kelley *et al.* 1999). As with spruce beetle, the two distinct *D. brevicomis* clades reflect an apparent disjunction in their host tree (*Pinus ponderosa*), in which two 'varieties' are separated by the Great Basin. Kelley *et al.* (1999) suggest the presence of cryptic species within *D. brevicomis*, and it is possible that reproductively isolated lineages also exist within *D. rufipennis*.

Patterns of gene exchange and barriers to gene flow

The strong signal of isolation by distance when all populations are included in an analysis of mtDNA haplotype data reflects the existence of substantial barriers to gene exchange (either extrinsic or intrinsic) between regional populations. This contrasts with the general absence of isolation by distance within regions, presumably a consequence of the homogenizing effects of dispersal in a beetle that is known to move long distances. Microsatellite data are also consistent with a homogenizing effect of dispersal and the presence of barriers to gene exchange

between regional groups of populations; F_{ST} values between Rocky Mountain and northern populations are high, whereas F_{ST} values are low between very distant populations within population groups (e.g. F_{ST} values between Newfoundland and Yukon or Alaska populations are 0.002–0.013; see Table 3).

MDIV analysis of mtDNA haplotypes also suggests the presence of barriers to gene flow. We cannot reject a hypothesis of no gene flow between two allopatric groups of populations, one in the Rocky Mountains, the other throughout northern spruce forests (Fig. 4). Discordance between mtDNA and microsatellite variation in the Pacific Northwest suggests that local hybridization and introgression have occurred and that mtDNA haplotypes from southern populations have flowed north following secondary contact. Such a pattern of differential (and asymmetrical) introgression of mtDNA has been well documented in other hybrid zones (Harrison 1989, 1990; Ballard & Whitlock 2004).

Although secondary contact subsequent to the most recent glacial retreat has apparently resulted in some gene exchange, our results imply that there has been little gene flow between northern and Rocky Mountain populations subsequent to initial divergence. Either beetle populations from different refugia did not overlap during most or all of the previous interglacial periods and came into secondary contact only very recently, or intrinsic barriers to gene exchange have limited the extent of hybridization and introgression during past episodes of parapatry or sympatry.

Both host specialization and/or differences in pheromone activity might provide bases for intrinsic barriers to gene exchange between previously allopatric populations. Although *D. rufipennis* feeds on more than 60% of the spruce species within its range (Kelley & Farrell 1998), the possibility of local host-plant specialization, commonly found in phytophagous insects, particularly endophytic species, has not been tested. Host-associated allozyme polymorphisms and mtDNA differentiation have been reported for other *Dendroctonus* species (Stock *et al.* 1979; Sturgeon & Mitton 1986; Kelley *et al.* 1999). It has been suggested that bark beetles could quite rapidly evolve specialization within isolated populations (Kelley & Farrell 1998), which is supported by evidence of within- and between-population variance in spruce beetle preferences for phytochemical composition (Wallin & Raffa 2004).

Pheromones can provide a basis for assortative mating in bark beetles (Teale *et al.* 1994). Responses to some pheromone components vary between *D. rufipennis* from Alaska (which colonize *P. glauca*) and British Columbia (which colonize *P. engelmannii*) (Borden *et al.* 1996). Further, the synthesis of and response to bark beetle pheromones are influenced by tree chemistry (Seybold *et al.* 2000). Therefore, host association and pheromone production might act in concert as a reproductive barrier, contributing to the

apparent absence of gene exchange between northern and Rocky Mountain populations.

The existence of very divergent mtDNA lineages in spruce beetle populations living in different geographical regions (and on different host trees) was not anticipated and raises the possibility that these populations also exhibit substantial genetic differentiation for life history or behaviour. Knowing whether they remain distinct in sympatry must await the outcome of more detailed studies of the secondary contact zone in British Columbia. The existence of nonrandom host associations would have important implications for management of this pest. For example, such information would be useful in dissecting the relative strengths of dispersal vs. localized eruptions in regional outbreaks, which currently is a major information gap in formulating landscape-level management plans (Aukema *et al.* 2006).

Acknowledgements

We thank Dr Rasmus Nielsen for help with the program MDIV and other analyses and for suggestions on the manuscript. We thank Ken Gibson, USDA FS, Margaret Skinner, U VT, Michael Bohne U VT, Paul Flanagan, USDA FS, Staffan Lindgren UNBC, Jose Negron USDA FS, Steve Seybold, USDA FS, David Stone, Can. FS, Jon Sweeney, Can. FS, Tom Eager, USDA FS, Dave Leatherman, USDA FS, William Jacobi, CSU, Joel D. McMillin, USDA FS, Richard Werner, USDA FS (retired), Ed Holsten, USDA FS (retired), Ken Zogas, USDA FS, Barbara Bentz, USDA FS, S. Francis, Can. FS, Anne Savoie, Can. FS, Ed VanRanden, Can. FS, and Kirsten Haberkern, UW-Madison for providing us with specimens or otherwise helping with insect collections. KFR was supported by NSF DEB-9629776, DEB-0314215 and DEB-0080609, USDA-NRI WIS04746, McIntire-Stennis, and University WI CALS. RGH was supported by McIntire Stennis Project NYC-183562 and Hatch Project NYC-183401.

References

- Anderson LL, Hu FS, Nelson DM, Petit RJ, Paige KN (2006) Ice-age endurance: DNA evidence of a white spruce refugium in Alaska. *Proceedings of the National Academy of Sciences, USA*, **103**, 12447–12450.
- Aukema BH, Carroll AL, Zhu J *et al.* (2006) Landscape level analysis of mountain pine beetle in British Columbia, Canada: spatiotemporal development and spatial synchrony within the present outbreak. *Ecography*, **29**, 427–441.
- Avise JC, Walker D (1998) Pleistocene phylogeographic effects on avian populations and speciation process. *Proceedings of the Royal Society of London. Series B, Biological Sciences*, **265**, 457–463.
- Avise JC, Walker D, Johns GC (1998) Speciation durations and Pleistocene effects on vertebrate phylogeography. *Proceedings of the Royal Society of London. Series B, Biological Sciences*, **265**, 1707–1712.
- Ballard JWO, Whitlock MC (2004) The incomplete natural history of mitochondria. *Molecular Ecology*, **13**, 729–744.
- Borden JH, Gries G, Chong LJ *et al.* (1996) Regionally-specific bioactivity of two new pheromones for *Dendroctonus rufipennis*

- (Kirby) (Col., Scolytidae). *Journal of Applied Entomology*, **120**, 321–326.
- Braun EL (1950) *Deciduous Forests of Eastern North America*. Blakinson Company, Philadelphia.
- Brower AVZ (1994) Rapid morphological radiation and convergence among races of the butterfly *Heliconius erato* inferred from patterns of mitochondrial DNA evolution. *Proceedings of the National Academy of Sciences, USA*, **91**, 6491–6495.
- Burg TM, Gaston AJ, Winjer K, Friesen VL (2005) Rapid divergence and postglacial colonization in western North American Steller's jay (*Cyanocitta stelleri*). *Molecular Ecology*, **14**, 3745–3755.
- Cavalli-Sforza LL, Edwards AWF (1967) Phylogenetic analysis: models and estimation procedures. *American Journal of Human Genetics*, **19**, 233–257.
- Davis MB (1976) Pleistocene biogeography of temperate deciduous trees. *Geoscience and Man*, **13**, 13–26.
- Demboski JR, Cook JA (2001) Phylogeography of the dusky shrew, *Sorex monticolis* (Insectivora, Soricidae): insight into deep and shallow history in northwestern North America. *Molecular Ecology*, **10**, 1227–1240.
- Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology*, **14**, 2611–2620.
- Falush D, Stephens M, Pritchard JK (2003) Inference of population structure using multilocus genotype data: Linked loci and correlated allele frequencies. *Genetics*, **164**, 1567–1587.
- Felsenstein J (2005) *PHYLIP (Phylogeny Inference Package)*, version 3.6. Department of Genetics, University of Washington, Seattle (program available from: <http://evolution.genetics.washington.edu/phylip.html>).
- Fitch WM (1977) On the problem of discovering the most parsimonious tree. *American Naturalist*, **111**, 223–257.
- Good JM, Demboski JR, Nagosen DW, Sullivan J (2003) Phylogeography and introgressive hybridization: chipmunks (genus *Tamias*) in the northern Rocky Mountains. *Evolution*, **57**, 1900–1916.
- Good JM, Sullivan J (2001) Phylogeography of the red-tailed chipmunk (*Tamias ruficaudus*), a northern Rocky Mountain endemic. *Molecular Ecology*, **10**, 2683–2695.
- Goudet JM, Raymond M, Meeüs T, Rousset F (1996) Testing differentiation in diploid populations. *Genetics*, **144**, 1933–1940.
- Guo SW, Thompson EA (1992) Performing the exact test for Hardy–Weinberg proportions for multiple alleles. *Biometrics*, **48**, 361–372.
- Hamilton MB, Pincus EL, Di Fiore A, Fleischer RC (1999) Universal linker and ligation procedures for construction of genomic DNA libraries enriched for microsatellites. *BioTechniques*, **27**, 500–507.
- Hansen EM, Bentz BJ, Turner DL (2001) Physiological basis for flexible voltinism in the spruce beetle (Coleoptera: Scolytidae). *Canadian Entomologist*, **133**, 805–817.
- Harrison RG (1989) Animal mitochondrial DNA as a genetic-marker in population and evolutionary biology. *Trends in Ecology & Evolution*, **4**, 6–11.
- Harrison RG (1990) Hybrid zones: windows on evolutionary process. In: *Oxford Surveys in Evolutionary Biology* (eds Futuyma DJ, Antonovics J), **7**, 69–128.
- Hewitt GM (1996) Some genetic consequences of the ice ages, and their role in divergence and speciation. *Biological Journal of the Linnean Society*, **58**, 247–276.
- Hewitt GM (2000) The genetic legacy of the Quaternary ice ages. *Nature*, **405**, 907–913.
- Hewitt GM (2001) Speciation, hybrid zones and phylogeography – or seeing genes in space and time. *Molecular Ecology*, **10**, 537–549.
- Hewitt GM (2004) Genetic consequences of climatic oscillations in the Quaternary. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, **359**, 183–195.
- Holsten EH, Their RW, Munson AS, Gibson KE (1999) *Forest Insect Disease Leaflet 127*. US Department of Agriculture Forest Service.
- Huelsenbeck JP, Ronquist F (2001) MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics*, **17**, 754–755.
- Jackson ST, Webb RS, Anderson KH *et al.* (2000) Vegetation and environment in Eastern North America during the Last Glacial Maximum. *Quaternary Science Reviews*, **19**, 489–508.
- Jaramillo-Correa JP, Beaulieu J, Bousquet J (2004) Variation in mitochondrial DNA reveals multiple distant glacial refugia in black spruce (*Picea mariana*), a transcontinental conifer. *Molecular Ecology*, **13**, 2735–2747.
- Johnson NK, Cicero C (2004) New mitochondrial data affirm the importance of Pleistocene speciation in North American birds. *Evolution*, **58**, 1122–1130.
- Kelley ST, Farrell BD (1998) Is specialization a dead end? The phylogeny of host use in *Dendroctonus* bark beetles (Scolytidae). *Evolution*, **52**, 1731–1743.
- Kelley ST, Mitton JF, Paine TD (1999) Strong differentiation in mtDNA of *Dendroctonus brevicornis* on different subspecies of ponderosa pine. *Annals of the Entomological Society of America*, **92**, 193–197.
- Klicka J, Zink RM (1997) The importance of recent ice ages in speciation, a failed paradigm. *Science*, **277**, 1666–1669.
- Klicka J, Zink RM (1998) Response to Pleistocene speciation and the mitochondrial DNA clock. *Science*, **282**, 1995.
- Knowles LL, Richards CL (2005) Importance of genetic drift during Pleistocene divergence as revealed by analyses of genomic variation. *Molecular Ecology*, **14**, 4023–4032.
- Liedloff A (1999) *MANTEL, version 2.0: The Nonparametric Test Calculator* School of Natural Resource Sciences, Queensland, Australia.
- Little EL Jr (1971) *Atlas of United States Trees, Volume 1, Conifers and Important Hardwoods*. US Department of Agriculture Miscellaneous Publication 1146.
- Mantel N (1967) The detection of disease clustering and a generalised regression approach. *Cancer Research*, **27**, 209–220.
- Miller MP, Bellinger R, Forsman ED, Haig SM (2006) Effects of historical climate change, habitat connectivity, and vicariance on genetic structure and diversity across the range of the red tree vole (*Phenacomys longicaudus*) in the Pacific Northwestern United States. *Molecular Ecology*, **15**, 145–159.
- Nei M (1987) *Molecular Evolutionary Genetics*. Columbia University Press, New York.
- Nielsen R, Wakeley J (2001) Distinguishing migration from isolation: a Markov chain Monte Carlo approach. *Genetics*, **158**, 885–896.
- Pielou EC (1991) *After the Ice Age: The Return of Life to Glaciated North America*. The University of Chicago Press, Chicago.
- Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics*, **155**, 945–959.
- Raymond M, Rousset F (1995) GENEPOP (version 12): Population genetics software for exact tests and ecumenism. *Journal of Heredity*, **86**, 248–249.
- Remington C (1968) Suture-zones of hybrid interaction between recently joined biotas. *Evolutionary Biology*, **2**, 321–428.

- Rosenberg NA (2004) DISTRUCT: a program for the graphical display of population structure. *Molecular Ecology Notes*, **4**, 137–138.
- Rousset F (1997) Genetic differentiation and the detection of gene flow from *F*-statistics under isolation by distance. *Genetics*, **145**, 1219–1228.
- Rozas J, Rozas R (1999) DNASP, version 3: an integrated program for molecular population genetics and molecular evolution analysis. *Bioinformatics*, **15**, 174–175.
- Santucci F, Emerson BC, Hewitt GM (1998) Mitochondrial DNA phylogeography of European hedgehogs. *Molecular Ecology*, **7**, 1163–1172.
- Schmid JM, Frye RH (1977) *Spruce Beetle in the Rockies*. USDA Forest Service General Technical Report RM-49.
- Seybold SJ, Bohlmann J, Raffa KF (2000) Biosynthesis of coniferophagous bark beetle pheromones and conifer isoprenoids: an evolutionary perspective and synthesis. *Canadian Journal of Entomology*, **132**, 1–57.
- Simon C, Frati F, Beckenbach A, Crespi B, Liu H, Flook P (1994) Evolution, weighting and phylogenetic utility of mitochondrial gene sequences and a compilation of conserved polymerase chain reaction primers. *Annals of the Entomological Society of America*, **87**, 651–700.
- Smith CI, Farrell BD (2005) Phylogeography of the longhorn cactus beetle *Moneilema appresum* LeCone (Coleoptera: Cerambycidae): was the differentiation of the Maderan sky islands driven by Pleistocene climate changes? *Molecular Ecology*, **14**, 3049–3065.
- Soltis DE, Gitzendanner MA, Strenge DD, Soltis PA (1997) Chloroplast DNA intraspecific phylogeography of plants from the Pacific Northwest of North America. *Plant Systematics and Evolution*, **206**, 353–373.
- Stock MW, Pieman GB, Guenther JD (1979) Genetic differences between Douglas-fir beetles (*Dendroctonus pseudotsugae*) from Idaho and coastal Oregon. *Annals of the Entomological Society of America*, **72**, 394–397.
- Sturgeon KB, Mitton JB (1986) Evolution of bark beetle communities. In: *Bark Beetles in North America Conifers* (eds Mitton JB, Sturgeon KB), pp. 350–384. University of Texas Press, Austin, Texas.
- Swenson NG, Howard DJ (2004) Do suture zones exist? *Evolution*, **58**, 2391–2397.
- Swofford DD (2002) PAUP*. *Phylogenetic Analysis Using Parsimony (*and other methods)*, Version 4. Sinauer Associates, Sunderland, Massachusetts.
- Tajima F (1989) Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics*, **123**, 585–595.
- Takezaki N, Nei M (1996) Genetic distances and reconstruction of phylogenetic trees from microsatellite DNA. *Genetics*, **144**, 389–399.
- Teale SA, Hager BJ, Webster FX (1994) Pheromone-based assortative mating in a bark beetle. *Animal Behavior*, **48**, 569–578.
- Wallin KF, Raffa KF (2004) Feedback between individual host selection behavior and population dynamics in an eruptive herbivore. *Ecological Monographs*, **74**, 101–116.
- Webb T III, Bartlein PJ (1992) Global changes during the last 3 million years: climatic controls and biotic responses. *Annual Reviews of Ecology and Systematics*, **23**, 141–173.
- Weir BS, Cockerham CC (1984) Estimating *F*-statistics for the analysis of population structure. *Evolution*, **38**, 1358–1370.
- Werner RA, Holsten EH (1997) Dispersal of the spruce beetle, *Dendroctonus rufipennis*, and the engraver beetle, *Ips Perturbatus*, in Alaska. US Forest Service Research Paper PNW-RP-501, 1–8.
- Werner RA, Raffa KF, Illman BL (1996) Insect and pathogen dynamics. In: *Alaska's Changing Boreal Forest* (eds Chapin FS III, Oswood M, Van Cleve K, Viereck LA, Verbyla D), pp. 187–207. Oxford University Press, Oxford, UK.
- Whitlock C, Bartlein PJ (1997) Vegetation and climate change in northwest America during the past 125 kyr. *Nature*, **388**, 57–61.
- Wood SL (1982) The bark and ambrosia beetles (Coleoptera: Scolytidae) of North and Central America, a taxonomic monograph. *Great Basin Naturalist Memoirs*, **6**, 1–1359.
- Zamudio KR, Savage WK (2003) Historical isolation, range expansion, and secondary contact of two highly divergent mitochondrial lineages in spotted salamanders (*Ambystoma maculatum*). *Evolution*, **57**, 1631–1652.

This project represents a collaboration between the Raffa Laboratory at University of Wisconsin and the Harrison laboratory at Cornell University. Kenneth Raffa oversees a research program focused on the population dynamics of eruptive forest insects and their interactions with tree chemistry, predators, and symbiotic microorganisms. K.F. Wallin also studies the population dynamics of forest pests. Luana Maroja is a PhD candidate at Cornell University with interests in speciation and mechanisms of reproductive isolation. Steve Bogdanowicz has broad interests in application of molecular markers to evolutionary questions. Rick Harrison is an evolutionary biologist who studies hybrid zones and speciation, primarily in insects.
