Phylogeographic patterns in *Drosophila montana*

P. M. MIROL,* M. A. SCHÄFER,† L. ORSINI,‡ J. ROUTTU,† C. SCHLÖTTERER,† A. HOIKKALA‡ and R. K. BUTLIN§

*School of Biology, The University of Leeds, Leeds LS2 9JT, UK, †Institut für Tierzucht und Genetik, Josef Baumann Gasse 1, Universität Innsbruck, Institut für Okologie, 1210 Wien, Austria, ‡Department of Biological and Environmental Sciences, PO Box 35, 40014 University of Jyväskylä, Finland, §Animal and Plant Sciences, The University of Sheffield, Sheffield S10 2TN, UK

Abstract

The *Drosophila virilis* species group offers valuable opportunities for studying the roles of chromosomal re-arrangements and mating signals in speciation. The 13 species are divided into two subgroups, the montana and virilis ‘phylads’. There is greater differentiation among species within the montana phylad in both karyotype and acoustic signals than exists among members of the virilis phylad. *Drosophila montana* is a divergent species which is included in the montana phylad. Here, we analyse the phylogeography of *D. montana* to provide a framework for understanding divergence of acoustic signals among populations. We analysed mitochondrial sequences corresponding to the cytochrome oxidase I and cytochrome oxidase II genes, as well as 16 microsatellite loci, from 108 lines of *D. montana* covering most of the species’ range. The species shows a clear genetic differentiation between North American and Scandinavian populations. Microsatellite allele frequencies and mitochondrial DNA haplotypes gave significant $F_{ST}$ values between populations from Canada, USA and Finland. A Bayesian analysis of population structure based on the microsatellite frequencies showed four genetically distinct groups, corresponding to these three populations plus a small sample from Japan. A network based on mitochondrial haplotypes showed two Finnish clades of very different shape and variability, and another clade with all sequences from North America and Japan. All *D. montana* populations showed evidence of demographic expansion but the patterns inferred by coalescent analysis differed between populations. The divergence times between Scandinavian and North American clades were estimated to range from 450 000 to 900 000 years with populations in Canada and the USA possibly representing descendants of different refugial populations. Long-term separation of *D. montana* populations could have provided the opportunity for differentiation observed in male signal traits, especially carrier frequency of the song, but relaxation of sexual selection during population expansion may have been necessary.

Keywords: demography, *Drosophila montana*, microsatellites, mitochondrial DNA, phylogeography

Received 9 July 2006; revision received 26 September 2006; accepted 30 October 2006

Introduction

The *Drosophila virilis* group comprises 13 species and subspecies divided into two clades, the virilis and montana phylads. The montana phylad is further subdivided into three lineages, montana, littoralis and kanekoi (Spicer 1992; Spicer & Bell 2002). Although most studies published to date support the existence of the four groups, the phylogenetic relationships between species within the groups are more contentious, varying with the marker used to infer the phylogeny (Spicer 1991, 1992; Nurminsky et al. 1996; Spicer & Bell 2002; Adrianov et al. 2003). Within the virilis phylad, the Palearctic endemics *D. virilis* and *D. lummei* are clearly defined. However, relationships among North American members of the phylad are confounded by shared ancestral polymorphisms in mitochondrial DNA (mtDNA) sequences (Caletka & McAllister 2004), nuclear
genes (Hilton & Hey 1996, 1997) and microsatellites (Orsini et al. 2004).

Throckmorton (1982) suggested that the phylads diverged in the Early Miocene, or not later than the Pliocene, when both of them entered the New World by way of Beringia. Divergence times between the phylads have been calculated using different markers, and vary from 7 million years ago (Ma) using restriction fragments of mitochondrial genes (Ostrega 1985), to 9 Ma according to sequences of the Adh gene (Nurminsky et al. 1996) and 11 Ma based on mitochondrial 12S and 16S ribosomal RNA genes (Spicer & Bell 2002). In general, species within the montana phylad have evolved more in terms of chromosomal rearrangements, and are also more variable regarding the number of inversions segregating within populations than members of the virilis phylad (Throckmorton 1982), although species in this group show a higher number of fusions. The primitive karyotype is found in D. virilis which, in contrast to other species of the group, shows no inversion polymorphisms. This contrast is interesting in terms of the possible role of chromosomal inversions in speciation (Butlin 2005). Also, species of the virilis phylad have higher crossabilities among taxa within the phylad than species of the montana phylad (Throckmorton 1982).

All species of the D. virilis group differ in the acoustic mating signals (‘songs’) produced by males during courtship. These male courtship songs, produced by wing vibration, play an important role both in species recognition (Liimatainen & Hoikkala 1998) and in sexual selection within species in the wild (Aspi & Hoikkala 1995). Song characteristics vary much more among species of the montana phylad than they do among virilis phylad species. Also, the importance of courtship song varies among species: Drosophila montana females rarely accept the courtship of a ‘mute’ (wingless) male, whereas D. virilis females readily accept this kind of courtship (Hoikkala 1988; Hoikkala et al. 2005).

Divergence in mating signals and responses between populations contribute to prezygotic reproductive isolation. There is abundant evidence that this type of barrier to gene flow is critical for the existence of many animal and plant species and that it evolves early in the process of speciation (Coyne & Orr 2004). Divergence in signals and responses may be due to sexual selection (Panhuis et al. 2001) or to reinforcement (Servedio & Noor 2003) but it could also be an incidental by-product of divergence due to ecological selection pressures or to drift. Progress in understanding the evolution of signals and responses requires documentation of patterns of variation both within and among species, and of the genetic basis of this variation. The D. virilis group has been a productive model system for the analysis of the contribution of acoustic mating signals to reproductive isolation (Liimatainen & Hoikkala 1998; Saarikettu et al. 2005), the genetic basis of interspecific differences (e.g. Hoikkala & Lumme 1987), variation in male song traits and female responses within the species (e.g. Ritchie et al. 2005), and sexual selection exercised by females on male song (e.g. Aspi & Hoikkala 1995; Ritchie et al. 1998).

The pattern of genetic variation within and among populations of a species is strongly affected by its phylogeographic history. Particularly strong signatures might be expected for species occupying formerly glaciated regions (Hewitt 2001) and for domesticated species (Cymbron et al. 2005; Larson et al. 2005; Pedrosa et al. 2005). It is important to know this history if one wishes to interpret patterns of variation in traits such as mating signals and responses (cf. Tregenza et al. 2000). Analyses based on putatively neutral markers such as mtDNA sequences or microsatellites provide a baseline against which the effects of selection can be tested as well as providing data from which past events, such as population expansions and colonizations, can be inferred. These events may have been responsible for changes in selection pressures that underlie variation in mating behaviour, for example. In this study, we propose phylogeographic and demographic scenarios for one of the species of the virilis group, D. montana, as a basis for studies of the evolutionary history of songs and preferences. We also provide a preliminary analysis of song variation within this species.

Materials and methods

Drosophila stocks and sampling

In total, 108 Drosophila montana strains, covering the species’ range, were selected for analysis (see Supplementary material). The stocks were collected during a time period covering 50 years, from 1947 to flies sampled in Finland in 2001/2002. In addition, three new wild populations of D. montana were sampled in 2003, in Colorado (USA), Vancouver (Canada) and Oulanka (Finland). New isofemale strains were established from wild caught individuals from these populations. A single individual from each strain, either from laboratory stocks or freshly caught, was used to extract DNA and for polymerase chain reaction (PCR) amplification of the COI and COII mitochondrial genes. A different individual from a subset of strains was used to extract DNA and amplify microsatellite markers, under the assumption that individuals within each strain are genetically homogeneous.

Amplification and sequencing of mitochondrial DNA

DNA was extracted from ethanol-preserved flies following a standard protocol (Sambrook et al. 1989), where the samples were homogenized in buffer and proteinase k, and DNA was extracted with chloroform-isooamyl alcohol and precipitated with isopropanol. The amplification of
mitochondrial DNA was carried out with primers flanking the COI gene in the tRNA_LYS and tRNA_LEU (Liu & Beckenbach 1992; TL2: 5'-ATGGCAGATTATGCAATGG-3', TKN: 5'-GTTTAGAGACCATCTTTG-3'), which amplify an 850-bp fragment that includes the 688 bp COII gene. Three other fragments corresponding to the mitochondrial genes ND5 and COI, and to the AT-rich control region were also amplified in a subsample of strains in order to assess their variability and select one of them for inclusion in the analysis. The variability found in COI was considered suitable and this second fragment was amplified and sequenced for the entire sample. The primers used were COI-1460-F: 5'-ATCTATCGCTAAAACTTCAGCC-3' and COI-2195-R: 5'-ACTTCAGGGTGACCAAAAAATC-3' (Simon et al. 1994; de Brito et al. 2002) which amplify the complete 670 bp corresponding to the COI gene. PCRs were performed in 50 µL volumes including 0.5 µm of each primer, 200 µm dNTPs, 1.5 mM MgCl₂ and 1 U Taq polymerase (Bioline) in reaction buffer. Initial denaturation was for 7 min at 94 °C followed by 35 cycles of 1 min at 94 °C, 1 min at the annealing temperature (54 °C for COI and 56 °C for COII) and 1 min at 72 °C, and a final incubation of 5 min at 72 °C. The products were purified using QIAquick columns (QIAGEN) and sequenced using the forward primer. Sequences (GenBank Accession nos DQ 426717 to DQ 426799) were aligned with the forward primer. Sequences (GenBank Accession nos DQ 426717 to DQ 426799) were aligned with

**Microsatellite typing**

Individuals were genotyped for a total of 16 microsatellite markers. For PCR amplification, DNA was extracted from a single individual from each strain using a high-salt extraction protocol (Miller et al. 1988). PCR reactions were performed in a final volume of 10 µL containing 50–100 ng of genomic DNA, 1 µm of each primer (forward end-labelled with 32P), 200 µm dNTPs, 1.5 mM MgCl₂ and 1 U Taq polymerase following standard protocols (Schlötterer 1998). The amplification consisted of 30 cycles with 50 s at 94 °C, 50 s at 45–62 °C (depending on locus), and 50 s at 72 °C. We applied an initial denaturing step of 3 min at 94 °C and a terminal extension of 45 min at 72 °C, allowing for a quantitative terminal transferase reaction. PCR products were separated on a 7% denaturing polyacrilamide gel (32% formamide, 5.6 µm urea) and visualized by autoradiography after 12–24 h. Allele sizes were determined running a ‘PCR slippage ladder’ and a known size standard adjacent to the samples (Schlötterer & Zangerl 1999). Detailed information on primer sequences, repeat motifs and PCR conditions for the loci Mon17a, 21, 23, 25, 26, 29, 30a, 30b and 31 can be found in Orsini *et al.* (2004). The remaining seven loci have been developed recently and are not yet published (Mon17b, 33a, 34, 36, 37, 39). Detailed information on these loci will be provided on request.

**Mitochondrial DNA analysis**

The partition homogeneity test (PHT), as implemented in PAUP 4.0, was used to test for incongruence between the COI and COII data sets. The test is based in the incongruence-length difference test by Farris *et al.* (1995). The null hypothesis is that the two loci are no more incongruent than two randomly generated partitions of equal size. One hundred replicates were generated and the P value obtained was 0.23, which indicates congruence between the data sets. Therefore, the two fragments were combined for all subsequent analyses.

Pairwise distances between haplotypes were estimated in ARLEQUIN 2.0 (Schneider *et al.* 1997). Due to the low level of diversity, no corrections were made for multiple substitutions. Analysis of population genetic structure was carried out using analysis of molecular variance (AMOVA) in ARLEQUIN. AMOVA takes into account the number of molecular differences between haplotypes in an analysis of variance framework equivalent to F statistics, with significance tested by permutation.

The frequency distributions of the numbers of segregating sites in all possible pairwise comparisons, known as mismatch distributions, were calculated in ARLEQUIN. Slatkin & Hudson (1991) demonstrated that the mismatch distributions of a stable population have a ragged profile due to stochastic lineage loss. In contrast, an exponentially growing population has a smooth unimodal distribution approaching a Poisson distribution. This reflects a star-like genealogy in which all of the coalescent events occurred in a short period of time. ARLEQUIN allows comparisons of observed mismatch distributions with those expected at equilibrium in a stable population or after a sudden expansion at scaled time T (= 2 μT generations) from scaled size θ₁ (= 2N_1 μ females) to θ₀ (= 2N_0μ females).

ARLEQUIN was also used to conduct tests of the standard neutral model for a demographically stable population. Tajima’s D-test (1989) compares two estimators of the population parameter θ = 2Nµ, θπ and θS. θS is based on the average pairwise number of differences between sequences (Tajima 1989) and θ₀ is estimated from the number of segregating sites in a population (Waterson 1975). The F-test of selective neutrality by Fu and Li (1993) evaluates the probability of observing a random sample with a number of alleles similar to, or smaller than the observed value, given the observed number of pairwise differences, taken as an estimator of θ. The F statistic, especially, is very sensitive to recent fluctuations in effective population sizes. In general, negative values of Tajima’s D and Fu’s F significantly different from zero indicate a population demographic expansion while positive values indicate contraction. However, selection may lead to similar patterns.
Population history and phylogeography

The program FLUCTUATE (Kuhner et al. 1998) was used to make simultaneous estimates of present-day θ and the population growth rate g, assuming an exponential model of growth and using a maximum-likelihood approach. The parameters used for the simulations were obtained by running a hierarchy of likelihood-ratio tests in MODELTEST 3.0 (Posada & Crandall 1998) to choose the model of evolution with the best fit to the data. Skyline plots were constructed using GENIE version 3.0 (Pybus et al. 2000). The starting trees were obtained using maximum likelihood with molecular clock enforced. GENIE was also used to calculate the fit to different models of population growth, with fit assessed using the corrected Akaike information criterion (AIC).

The phylogenetic relationships between species were inferred in PAUP 4.0b10 using maximum parsimony, distance-based methods and maximum likelihood. The best-fitting model of nucleotide substitution for the maximum-likelihood analyses was selected using MODELTEST 3.0, as above. Maximum parsimony and maximum-likelihood heuristic searches were conducted with 1000 random sequence addition replicates. Because the classic phylogenetic methods are not directed toward analysis of intraspecific data, we constructed networks based on statistical parsimony using the program tcs 1.06 (Clement et al. 2000). Phylogenetic methods assume that ancestral haplotypes are no longer present; yet coalescent theory predicts that ancestral haplotypes may be the most frequent sequences sampled in a population level study. Statistical parsimony is particularly useful to estimate robust networks when few nucleotide differences exist among haplotypes and it assigns outgroup weights to haplotypes, allowing hypothesis testing about geographical origin.

Microsatellite analysis

Genetic differentiation between D. montana populations was calculated using F statistics according to Weir & Cockerham (1984) where F statistics are weighted computations of the F coefficients of Wright (1978). Statistical significance of FST values was tested by 10 000 permutations of genotypes among populations. This conservative procedure does not assume Hardy–Weinberg equilibrium and allows for linkage among loci. We applied the sequential Bonferroni correction procedure to account for multiple testing (Sokal & Rohlf 1995). These calculations were performed with version 3.12 of the MICRO SATellite-ANALYSER software (Dieringer & Schlötterer 2003).

In order to quantify the amount of genetic variation resulting from differentiation between continents relative to that resulting from geographical separation within continents, we performed an AMOVA using ARLEQUIN 2.0 (Schneider et al. 1997). Variance components were tested for significance by permuting genotypes among populations and among groups (FCT), by permuting genotypes among populations but within groups (FSC) and by permuting populations among groups (FST).

We used BAPS 2.0 (Corander et al. 2003) to test for population substructure. BAPS 2.0 estimates hidden population substructure based on multilocus genotypes and on the geographical sampling information in a Bayesian statistical framework. We ran 10 000 updates after a burn-in phase of 5000 iterations. Three independent runs were performed to test the robustness of our results.

Song recording and analysis

For song recording, a virgin sexually mature male and female (18–24 days old) of the same strain were transferred into a mating chamber (a Petri dish with diameter 5.5 cm, height 1.3 cm). The roof of the chamber was made of nylon mesh, and the floor covered with a moistened filter paper. Male songs were recorded with a JVC condenser microphone, which was kept above the courting flies. Recordings were made with a Sony TC-FX33 cassette recorder between 0800 and 1200 at a temperature of 20 ± 1 °C.

Song analysis was carried out with the SIGNAL Sound Analysis System. Male songs were analysed by measuring the lengths of the pulse trains (PTL) and by counting the number of pulses per train (PN) from the oscillograms. Pulse length (PL), interpulse interval (IPI; the time from the beginning of one pulse to the beginning of the next pulse) and the number of cycles in a pulse (CN) were measured for the fourth sound pulse of each pulse train. Carrier frequency (FRE) was measured from the Fourier spectra of the pulse trains. For each strain, the means of different song traits were calculated over the songs of five males per strain (three pulse trains per male). Mantel tests were used to investigate the relationship among genetic distances between strains, as measured by mtDNA haplotype and song characters. Analysis of variance was used to test for song differentiation among geographic regions. We did not attempt to establish a correlation between microsatellite and song data because the number of samples for which we have both kind of information was small.

Results

Species-level analysis

A neighbour-joining tree (Fig. 1) based on the combined COI and COII sequence data for Drosophila montana and three other species of the virilis group, Drosophila virilis (five strains from the UK, China, Japan and USA), Drosophila littoralis (isofemale lines established from populations in Finland and in Portugal) and Drosophila borealis (one isofemale line established from flies captured.
in Manitoba, Canada), is in agreement with topologies described for these species previously (Throckmorton 1982; Nurminsky et al. 1996; Spicer & Bell 2002). *D. borealis* and *Drosophila montana* are sister species. Within *D. littoralis*, there is no detectable differentiation between the localities included, Portugal and Finland. The *D. montana* samples from Finland occupy an ancestral position within the species. Maximum likelihood and parsimony topologies were similar to the neighbour-joining tree.

**Mitochondrial variation in Drosophila montana**

We examined a total of 108 *D. montana* lines, including many laboratory strains and new lines established from wild individuals captured in 2002 and 2003. The samples covered North America [Canada (C) and USA (U)], Finland (F) and Japan (J). The analysis is based on a total of 1358 base pairs, 670 corresponding to the COI and 688 to the COII. The number of different haplotypes obtained was 72. There were no significant differences between laboratory strains established before 2002 and the recently established lines, or between years for regions where direct comparisons were possible (F 2002 vs. F 2003: $F_{ST} = 0.02479$, $P = 0.592$; F laboratory strains vs. F 2002/2003: $F_{ST} = 0.00055$, $P = 0.356$; U laboratory strains vs. U 2003: $F_{ST} = 0.00906$, $P = 0.195$; C laboratory strains vs. C 2003: $F_{ST} = 0.03864$, $P = 0.085$). Therefore, all individuals from the same geographic region were pooled for subsequent

Fig. 1 Neighbour joining tree of mtDNA COI and COII haplotypes. F: Finland, P: Portugal, NA: North America, J: Japan. Numbers at the nodes are bootstrap values based on 1000 replicates. Only a random sample of sequences from *Drosophila montana* was included.
analyses. Note that three lines from Alaska were included within the Canadian sample following a geographic rather than political criterion.

Genetic differentiation among lines from Europe (Oulanka and Kemi), North America (Utah, Colorado and Vancouver) and Japan (Kawasaki) was tested using AMOVA (Table 1). The $\Phi_{ST}$ value was significant (0.147, $P < 0.0001$). The differentiation between regions was also significant ($\Phi_{CT}$: 0.12506, $P = 0.02151$). C, continents (Finland, Japan, North America); P, populations; *$P < 0.05$; **$P < 0.001$.

Table 1 AMOVA results for the genetic differentiation among Drosophila montana populations between and within continents based on 16 microsatellite loci (upper lines) and mitochondrial DNA (lower line)

<table>
<thead>
<tr>
<th>Source of variation</th>
<th>d.f.</th>
<th>SS</th>
<th>Variance components</th>
<th>Percentage of variation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Among C</td>
<td>2</td>
<td>95.262</td>
<td>0.848*</td>
<td>16.53</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>4.188</td>
<td>0.0612</td>
<td>12.51</td>
</tr>
<tr>
<td>Among P</td>
<td>3</td>
<td>38.866</td>
<td>0.450***</td>
<td>8.78</td>
</tr>
<tr>
<td>within C</td>
<td>3</td>
<td>1.606</td>
<td>0.011</td>
<td>2.17</td>
</tr>
<tr>
<td>Within P</td>
<td>138</td>
<td>528.469</td>
<td>3.829</td>
<td>74.69</td>
</tr>
<tr>
<td></td>
<td>87</td>
<td>36.604</td>
<td>0.421</td>
<td>85.33</td>
</tr>
</tbody>
</table>

Fixation indices — microsatellites: $F_{ST} = 0.248, F_{SC} = 0.100$; mitochondrial DNA: $\Phi_{ST}$: 0.14672, $P < 0.0001$; $\Phi_{SC}$: 0.02476, $P = 0.12805$; $\Phi_{CT}$: 0.12506, $P = 0.02151$. C, continents (Finland, Japan, North America); P, populations; *$P < 0.05$; **$P < 0.001$.

Figure 2 shows the network created with statistical parsimony. The lines from Finland are clearly differentiated from the other populations, while lines from the remaining regions are mixed together. Despite the significant differentiation detected by AMOVA, haplotypes from the USA and Canada do not occupy separate clades in the network. The Finnish samples constitute two differentiated clades. One of them (Fig. 2) contains a sequence constituting the central node, which is also the most frequent haplotype — present in 24 lines out of 54 examined — and the one with the highest outgroup weight (0.12). We need to be careful about this result since there are many more samples from Oulanka than from any other single population. However, the position of the Finnish samples in the species-level tree (Fig. 1) also confirms the basal placement of these haplotypes. This is the only clade with this type of topology, where most of the haplotypes are derived from the root and differ in only one or two substitutions.

Table 2 Pairwise $F_{ST}$ values (Weir & Cockerham 1984) among Drosophila montana populations (lower triangular matrix) and corresponding $P$ values (upper triangular matrix) for microsatellites (upper line) and mitochondrial DNA (lower line)

<table>
<thead>
<tr>
<th>Pairwise</th>
<th>$F_{ST}$</th>
<th>n</th>
<th>Oulanka</th>
<th>Kemi</th>
<th>Kawasaki</th>
<th>Utah</th>
<th>Colorado</th>
<th>Vancouver</th>
</tr>
</thead>
<tbody>
<tr>
<td>Oulanka</td>
<td></td>
<td>24</td>
<td>—</td>
<td>0.642</td>
<td>0.001***</td>
<td>0.002**</td>
<td>0.001***</td>
<td>0.001***</td>
</tr>
<tr>
<td></td>
<td></td>
<td>45</td>
<td>—</td>
<td>0.662</td>
<td>0.029***</td>
<td>0.001***</td>
<td>0.002**</td>
<td>0.001***</td>
</tr>
<tr>
<td>Kemi</td>
<td></td>
<td>14</td>
<td>—</td>
<td>—</td>
<td>0.006***</td>
<td>0.005</td>
<td>0.001***</td>
<td>0.001***</td>
</tr>
<tr>
<td></td>
<td></td>
<td>5</td>
<td>—</td>
<td>—</td>
<td>0.208</td>
<td>0.026*</td>
<td>0.04**</td>
<td>0.004***</td>
</tr>
<tr>
<td>Kawasaki</td>
<td></td>
<td>4</td>
<td>0.366</td>
<td>0.317</td>
<td>0.057</td>
<td>0.018</td>
<td>0.001***</td>
<td>0.368</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3</td>
<td>0.177</td>
<td>0.178</td>
<td>—</td>
<td>0.048*</td>
<td>0.999</td>
<td>0.577</td>
</tr>
<tr>
<td>Utah</td>
<td></td>
<td>3</td>
<td>0.182</td>
<td>0.128</td>
<td>0.275</td>
<td>—</td>
<td>0.075</td>
<td>0.218</td>
</tr>
<tr>
<td></td>
<td></td>
<td>9</td>
<td>0.159</td>
<td>0.146</td>
<td>0.017</td>
<td>—</td>
<td>0.191</td>
<td>0.024**</td>
</tr>
<tr>
<td>Colorado</td>
<td></td>
<td>5</td>
<td>0.371</td>
<td>0.333</td>
<td>0.404</td>
<td>0.173</td>
<td>—</td>
<td>0.001***</td>
</tr>
<tr>
<td></td>
<td></td>
<td>8</td>
<td>0.149</td>
<td>0.134</td>
<td>0.001</td>
<td>0.014</td>
<td>—</td>
<td>0.001***</td>
</tr>
<tr>
<td>Vancouver</td>
<td></td>
<td>22</td>
<td>0.162</td>
<td>0.135</td>
<td>0.255</td>
<td>—0.007</td>
<td>0.215</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td></td>
<td>23</td>
<td>0.163</td>
<td>0.153</td>
<td>0.042</td>
<td>0.042</td>
<td>0.005</td>
<td>—</td>
</tr>
</tbody>
</table>

Significant $P$ values after Bonferroni correction (B) are marked; n, sample size.
There is one clade corresponding to the lines from the USA, Canada and Japan. It is a large clade with variable branch lengths where most of the haplotypes are unique, differing from each other by a few substitutions. The most common haplotype is represented by eight individuals out of 60. Although lines from Canada and USA did not constitute different clades in the network analysis, separate fluctuate and mismatch distribution analyses were conducted because of the evidence (above) for population structure within North America. However, to increase statistical power, we pooled sequences within regions and added additional lines that could not be included in the AMOVA because they were single representatives of their populations (lines from Lappajarvi, Kemi and Oulu were included). To increase statistical power, we pooled sequences within regions and added additional lines that could not be included in the AMOVA because they were single representatives of their populations (lines from Lappajarvi, Kemi and Oulu were included).

Table 3 Summary statistics of mitochondrial variation (a) and results from the FLUCTUATE and GENIE analyses (b)

<table>
<thead>
<tr>
<th>Region</th>
<th>n</th>
<th>H</th>
<th>θ_s</th>
<th>θ_e</th>
<th>Tajima’s D</th>
<th>Fu’s F</th>
<th>FLUCTUATE results</th>
<th>GENIE results</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>AICc best model</td>
<td>θ_0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(second best model)</td>
<td>(95% CI)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>g (95% CI)</td>
<td>α (95% CI)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>c (95% CI)</td>
<td></td>
</tr>
<tr>
<td>D. montana</td>
<td>23</td>
<td>21</td>
<td>0.0096</td>
<td>0.0075</td>
<td>-0.8429</td>
<td>-9.8345</td>
<td>0.0374</td>
<td>36.451</td>
</tr>
<tr>
<td>USA</td>
<td></td>
<td></td>
<td>(0.0034)</td>
<td>(0.0037)</td>
<td>P = 0.21</td>
<td>P &lt; 0.001</td>
<td>(0.0079)</td>
<td>(90.49)</td>
</tr>
<tr>
<td>D. montana</td>
<td>34</td>
<td>28</td>
<td>0.0065</td>
<td>0.0043</td>
<td>-1.2267</td>
<td>—</td>
<td>0.0986</td>
<td>1231.2</td>
</tr>
<tr>
<td>Canada</td>
<td></td>
<td></td>
<td>(0.0022)</td>
<td>(0.0023)</td>
<td>P &lt; 0.001</td>
<td>(0.00279)</td>
<td>(64.73)</td>
<td></td>
</tr>
<tr>
<td>D. montana</td>
<td>54</td>
<td>25</td>
<td>0.0086</td>
<td>0.0061</td>
<td>-0.9848</td>
<td>-2.4148</td>
<td>0.0103</td>
<td>-22.27</td>
</tr>
<tr>
<td>Finland</td>
<td></td>
<td></td>
<td>(0.0026)</td>
<td>(0.0029)</td>
<td>P = 0.156</td>
<td>P = 0.244</td>
<td>(0.0014)</td>
<td>(42.67)</td>
</tr>
</tbody>
</table>

Estimates (standard errors): n, number of lines surveyed; H, number of haplotypes; θ_s, diversity (Waterson 1975); θ_e, diversity (Tajima 1989); θ_0, current estimate of 2Nµ; g, scaled population growth parameter; α, scaled population size before growth as a proportion of current size; C, shape parameter. GENIE model abbreviations: Exp, exponential growth; Log, logistic growth; Pexp, piecewise expansion; Expan, expansion. Where confidence limits are not given, they were at the default parameter boundaries.
included in the analysis of the Finnish sample, lines from Idaho, Wyoming, Yukon and Nevada were added to Utah and Colorado to form the USA sample and lines from Alaska, Ontario and Quebec were added to the Canadian sample with strains from Vancouver; see Supplementary material). Modeltest favoured the HKY85 model, with a transition/transversion ratio of 7.76 and a shape parameter (alpha) of 0.8434, and this was used in fluctuate. The mismatch distribution for the USA sample is smooth and unimodal (mean number of pairwise differences 8.912, 95% confidence interval from 6.765 to 13.250, Fig. 3a); it does not differ significantly from the expectations of the stepwise expansion model (SSD = 0.0063, \( P = 0.86 \)) and the raggedness index is also very low (\( r = 0.00761, P = 0.97 \)). The exponential model fitted by the fluctuate analysis also suggests growth and model comparison using genie finds the exponential model to be the best fit to the data, with similar estimates for growth rates and current population sizes (Table 3). The skyline plot (Fig. 4a) illustrates the good fit of the exponential expansion model for this population.

Fig. 3 Mismatch distributions among haplotypes of Drosophila montana from the USA (a), Canada (b) and Finland (c). Expectations from the stepwise growth model, fitted in arlequin, are superimposed.

Fig. 4 Generalized skyline plots for D. montana from the USA (a), Canada (b) and Finland (c). Observed values (solid line) and fitted values from the best model (broken line) — see Table 3b. Smoothing parameters (epsilon) were: (a) 0.00102, (b) 0.00041, (c) 0.00036 (maximum likelihood values from option ‘maxepsilon’ in genie).
The mismatch distribution for the Canadian sample (mean number of pairwise differences 5.597, 95% confidence interval from 2.897 to 9.123, Fig. 3b) is more ragged ($r = 0.0134$, $P = 0.91$) but is consistent with a stepwise expansion ($SSD = 0.0106$, $P = 0.73$). A recent stepwise expansion is also indicated by the skyline plot (Fig. 4b), with expansion around 35 000 years ago given a mutation rate of $10^{-8}$. The preferred model in GENIE analysis was the piece-wise expansion model (i.e. exponential expansion from a stable ancestral population) (Table 3). The lines from Finland also have a ragged mismatch distribution ($r = 0.0245$, $P = 0.98$) and are consistent with stepwise expansion ($SSD = 0.0210$, $P = 0.47$) but with much higher maximum divergence (mean number of pairwise differences 9.743, 95% confidence interval from 2.629 to 17.673, Fig. 3c). The skyline plot (Fig. 4c) is very similar to the Canadian sample, except that the ancestral population was somewhat larger and with a longer history. It is not clear why the logistic model has the highest AIC in the GENIE analysis since it does not appear to be a good fit to the data.

All three samples of D. montana yield negative estimates of Tajima’s $D$ and Fu’s $F$, which is consistent with the inferences of population expansion. However, only the $F$ estimates for the Canadian expansion and US samples differ significantly from zero (Table 3).

**Associations between phenotypic and mitochondrial variation**

North American D. montana has been traditionally divided into three subtypes, Giant, Standard and Alaskan–Canadian, according to inversion frequency data, body size and geographic location (see Throckmorton 1982). Finnish D. montana were originally described as Drosophila ovivororum by Lakovaara & Hackman (1973) and Vieira & Hoikkala (2001) showed them to differ genetically from the North-American D. montana populations. The status of the Japanese population is less well known. We divided haplotypes according to this classification (see Supplementary material). Pairwise comparisons between different classes indicated that the Finnish samples were significantly different from all others at the 0.05 level, and that Giant montana differed from all other North American populations but not from lines from Japan (Table 4). However, only two Giant lines were analysed, so this last result should be treated with caution.

Studies on song evolution require confronting the predicted pattern of song changes to a phylogeny as well as tracing selective pressures affecting song traits at the population level. We tested whether divergence between strains for song characters was correlated with genetic distances, using partial Mantel tests, or whether they differed between genetically distinct populations, using analysis of variance. Significant associations are expected if duration or completeness of isolation influences the extent of song differentiation. We obtained song measurements for a subset of the strains: seven lines from Finland, nine from America and three from Japan (see Supplementary material). There were significant Mantel correlations between genetic distance and pulse train length (PTL) or carrier frequency (FRE), although did not remain significant after Bonferroni correction. Only carrier frequency (FRE) varied significantly among geographical regions ($F_{3,32} = 5.43$, $P = 0.0039$) although variation in cycle number was also nearly significant ($F_{3,32} = 2.90$, $P = 0.050$). The variation in FRE, which remains significant after Bonferroni correction, was mainly due to the low frequency of the song of Finnish strains (236.7[16.7]Hz, mean[among strain standard deviation]) relative to the other regions (USA, 261.7[22.6]Hz; Canada, 268.0[90]Hz; Japan, 252.9[17.4]Hz).

**Microsatellite variation**

For the microsatellite analyses, we used the same groupings as for the mtDNA population structure analysis: the sample from Finland was divided into two geographical groupings (Oulanka and Kemi, separated by about 225 km) and the lines from Utah were separated from those collected in Colorado (both USA in the mtDNA analysis) giving six samples in total. Genotype data for 16 polymorphic microsatellite loci showed that D. montana populations are genetically highly differentiated over the geographical area studied (mean $F_{ST}$ across loci = 0.208; 95% confidence interval from 0.175 to 0.261). Significant differentiation was detected for each of the 16 loci (all $P < 0.05$), suggesting that differentiation is a genome-wide phenomenon rather than being concentrated at a few individual loci.

To investigate the influence of geographical separation on genetic differentiation, we calculated a matrix of pairwise $F_{ST}$ values between populations (averaged across loci). Pairwise comparisons showed clear genetic differences between samples originating from different continents, while comparisons within continents showed a much lower

---

**Table 4 Mitochondrial $F_{ST}$ (above the diagonal) and $P$ values (below the diagonal) for morphological subtypes of Drosophila montana. Significant values are shown in bold.**

<table>
<thead>
<tr>
<th></th>
<th>1-Finnish</th>
<th>2-Standard</th>
<th>3-Alaska-Canadian</th>
<th>4-Giant</th>
<th>5-Japanese</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.518</td>
<td>0.517</td>
<td>0.392</td>
<td>0.485</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>0.000</td>
<td>0.032</td>
<td>0.342</td>
<td>0.136</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>0.000</td>
<td>0.186</td>
<td>0.344</td>
<td>0.164</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>0.035</td>
<td>0.043</td>
<td>0.000</td>
<td>—</td>
<td>0.245</td>
</tr>
<tr>
<td>5</td>
<td>0.007</td>
<td>0.057</td>
<td>0.081</td>
<td>0.101</td>
<td>—</td>
</tr>
</tbody>
</table>

---

© 2006 The Authors
Journal compilation © 2006 Blackwell Publishing Ltd
degree of differentiation (Table 2). The impact of large-scale geographical separation on genetic differentiation among *D. montana* populations is best illustrated by the results obtained from AMOVA. Approximately 67% of the among-population genetic variation can be attributed to differentiation among continents (Table 1). Pairwise $F_{ST}$ values indicate that genetic variation within continents mainly resulted from the fact that the Vancouver sample differs from the Colorado sample, while the Finnish populations did not differ genetically (Table 2).

Bayesian analysis supports the inferences of population structure from $F$ statistics, indicating that the six samples originate from four genetically distinct groups: a Finnish group (Kemi, Oulanka), a Japanese group and two American groups. One American group contained the populations from Utah and Vancouver and the second American group contained the population from Colorado. This grouping was supported with very high probability ($P > 0.999$) while the probability for all possible alternative groupings was very low ($P < 0.001$). To further address the role of recent gene flow, especially between the Northern American populations, we performed a Bayesian admixture analysis. The analysis showed that the inferred admixture between individuals belonging to the four distinct genetic clusters is very low (Fig. 5). Most notably, there were no signs of admixture between the Colorado and the Vancouver sample supporting their genetic distinctiveness. Note that one individual of the Finnish group showed a high degree of admixture. Based on the current data set, it is difficult to judge whether this reflects a true admixture event or is due to the contamination of one line in the laboratory. Also, this analysis shows the individuals from Utah to be closer to the individuals from Vancouver than to those in Colorado, which contradicts the pattern obtained based on mtDNA. However, only three individuals were genotyped for microsatellites in the Utah population, and this could be the cause of the result. A table including the number and the frequency of different microsatellite alleles for each locus and population is presented in the Supplementary material.

### Discussion

*Drosophila montana* is an unusual species within the virilis group. It differs from other members of the group in several ways, including mating behaviour, chromosome variability, geographical range, specificity of larval substrate and association with human habitats (Throckmorton 1982). Here, we provide a phylogeographic analysis as a basis for understanding phenotypic variation within the species.

Mitochondrial DNA and microsatellites indicate the presence of at least two distinct populations, one in Eurasia and the other one representing the expansion of the species to the New World. The mtDNA haplotypes found in North America form a clade that nests within the more diverse set of haplotypes present in Finland supporting this direction of colonization. Genetic distances between the two major mtDNA clades range from 0.9 to 1.8%, which, assuming a mitochondrial divergence rate of 2% per million years, implies a separation between them from 450 000 to 900 000 years ago, within the Pleistocene. Päälässahko et al. (2005) obtained congruent divergence times between Finnish and American populations of *D. montana* based on silent substitutions in three X-linked genes, fused, elav and su(s).

Microsatellite data revealed distinct Finnish and North American populations, like mtDNA data, but they further suggested genetic differentiation within North America. Although samples of mtDNA haplotypes from Canada and USA are not phylogenetically distinct, the populations are significantly differentiated as judged by $Φ_{ST}$. They also show evidence for different historical demographic patterns (Fig. 4). This might reflect partial or complete isolation into distinct northern (Beringian) and southern (Rocky Mountains) refugia (Hewitt 2004) during the last glaciation. The possibility of different colonization times could also be considered. Both the Finnish and Canadian samples suggest very rapid population expansion around 35 000 years ago, somewhat older than the end of the last glaciation, while the US sample suggests more gradual expansion starting earlier, which is consistent with the more southerly location of the Rocky Mountains refuge. The mtDNA data do not support a suggestion of recent gene exchange between Finnish and North American populations based on X-linked genes (Päälässahko et al. 2005). However, this suggestion was based on shared variation in short stretches of sequence that might represent remnants of ancestral polymorphisms that have yet to achieve reciprocal monophyly (Päälässahko et al. 2005).

Acoustic signals used by males during courtship may experience various types of selection. Some song traits, important for mate recognition, may be under stabilizing selection and others may be under directional sexual selection while remaining traits have no signalling function and may evolve neutrally. Neutral traits should vary geographically in a way that mirrors divergence at marker
loci, such as mtDNA and microsatellites, while traits under stabilizing or directional selection are expected to be less variable within and among populations. However, a change in the environment may result in rapid divergence between populations in selected traits. It has also been suggested that rapid expansion of a population into a new, vacant habitat might be associated with a reduction in female discrimination among males and so a weakening of selection on male mating signals (Kaneshiro 1980). This might result in greater genetic variability in signals within populations and greater divergence among populations, but the effect may be weakened if the small initial population had low variability.

Our data suggest that D. montana populations in Europe and North America have been isolated for a long period, allowing the opportunity for accumulation of divergence in neutral song traits. They have also expanded in numbers, in some cases rapidly, which might have resulted in relaxed selection or a change in the pattern of selection on key song traits. Our initial analysis of song traits did not detect an association between genetic and phenotypic distance, suggesting that song evolution is not neutral. However, the analysis did detect significant divergence between the Finnish and North American populations, primarily in carrier frequency. In the Finnish D. montana population, females are known to show strong mate preferences based on song, specifically on carrier frequency (Aspi & Hoikkala 1995; Ritchie et al. 1998). It is particularly interesting that the trait that is known to be under sexual selection is also the trait that distinguishes this population from the other regions. Further work will be needed to establish whether this results from relaxation of selection or a change in the direction of selection. However, the fact that carrier frequency is low in the Finnish population, where it experiences directional sexual selection favouring high frequencies and where the effective population size is large, might suggest a relaxation of selective constraints in North America. Future detailed studies on song variation should concentrate on freshly collected strains because the songs, especially the pulse characters of the song, may be liable to change during laboratory maintenance. The Finnish strains used in the song analysis reported here were, on average, younger than strains from other regions but there were not sufficient strains to test for a systematic effect of age.

We have presented here a possible scenario for the biogeographic history of D. montana, which will form the basis for the interpretation of evolution of their mating signals and responses. Currently available information on male song shows divergence among populations that have a long history of separation. Further analyses of these traits will help to show how population demography interacts with selection to generate the divergence in mating behaviour that might ultimately cause speciation.

Aknowledgements

We are grateful to the members of the ‘Co-evolved Traits’ Research Training Network for their valuable input to the work presented here and to the European Commission for funding the network (HPRN-CT-2002-00266). Special thanks are due to Oulanka Biological Station and LAPBIAT project, the Rocky Mountain Biological Laboratory, and Andrew Beckenbach for help in arranging fly-collecting trips in different parts of the species’ distribution area, for all the people who have helped in collecting the flies and establishing of isofemale lines (especially Dominique Mazzi, Susanna Huttunen and Kirsten Klappert). Laboratory strains of the flies were obtained also from Jorge Vieira, Michael Evgeniev and Bowling Green stock centre. Very special thanks to A.F.O. for helping with statistics.

Supplementary material

The supplementary material is available from http://www.blackwellpublishing.com/products/journals/supportmat/MEC/MEC3215/MEC3215sm.htm

Lines of Drosophila montana used in the study, indicating, when it was available, year of collection and coordinates from which the line originates. Morphological subtypes, name of the line and mitochondrial (mt), microsatellite (ms) and song traits data are also indicated

Allele sizes (size) number of alleles (count) and allele frequencies (frq) for 16 microsatellite markers for six Drosophila montana populations

References


