The extent of variation in male song, wing and genital characters among allopatric *Drosophila montana* populations

J. ROUTTU,* D. MAZZI,*[1] K. VAN DER LINDE,† P. MIROL,‡[2] R. K. BUTLIN‡ & A. HOIKKALA*

*Department of Biological and Environmental Sciences, University of Jyväskylä, Jyväskylä, Finland
†Department of Biological Science, Florida State University, Tallahassee, FL, USA
‡Animal and Plant Sciences, University of Sheffield, Sheffield, UK

Abstract

Drosophila montana, a species of the *Drosophila virilis* group, has distributed around the northern hemisphere. Phylogeographic analyses of two North American and one Eurasian population of this species offer a good background for the studies on the extent of variation in phenotypic traits between populations as well as for tracing the selection pressures likely to play a role in character divergence. In the present paper, we studied variation in the male courtship song, wing and genital characters among flies from Colorado (USA), Vancouver (Canada) and Oulanka (Finland) populations. The phenotypic divergence among populations did not coincide with the extent of their genetic divergence, suggesting that the characters are not evolving neutrally. Divergence in phenotypic traits was especially high between the Colorado and Vancouver populations, which are closer to each other in terms of their mtDNA genotypes than they are to the Oulanka population. The males of the Colorado population showed high divergence especially in song traits and the males of the Vancouver population in wing characters. Among the male song traits, two characters known to be under sexual selection and a trait important in species recognition differed clearly between populations, implying a history of directional and/or diversifying rather than balancing selection. The population divergence in wing characters is likely to have been enhanced by natural selection associated with environmental factors, whereas the male genitalia traits may have been influenced by sexual selection and/or sexual conflict.

Introduction

Populations spreading into new environments may diverge from the ancestral population both genetically and phenotypically. Phenotypic divergence may reflect genetic divergence or it can be enhanced by natural selection or by stochastic effects of mutation and genetic drift. The role of selection is of special importance in populations adapting to novel environmental conditions (Fisher, 1930; Wright, 1931; Dobzhansky, 1951). Phylogenetic or phylogeographic analyses based on neutral molecular markers provide a good background for population level studies on phenotypic divergence but, as Grandcolas & D'Haese (2003) have emphasized, the selective value of a character can best be considered when examining patterns of phenotypic variation among wild populations. Phenotypic studies should, if possible, be targeted on several behavioural and morphological traits at the same time to distinguish neutrally evolving traits from those whose evolution has been affected by different kinds of selective pressures.

During the courtship rituals, *Drosophila* males emit various kinds of species-specific stimuli, e.g. male courtship song. The songs may play an important role in sexual selection and/or in species recognition, and they

Correspondence: J. Routtu, Department of Biological and Environmental Sciences, 40014 University of Jyväskylä, Jyväskylä, Finland.
Tel.: +358 142 604239; fax: +358 142 602321;
e-mail: Jarkko.Routtu@bytl.jyu.fi
[1]*Present address:* Institute of Plant Sciences, Applied Entomology, ETH (Swiss Federal Institute of Technology), CH-8092 Zurich, Switzerland.
[2]*Present address:* Museo Argentino de Ciencias Naturales, C1405DJR, Buenos Aires, Argentina.

may simultaneously be affected by directional, diversifying and balancing selection. However, signals important in species recognition may not vary too much if they are to retain species specificity (Lambert & Henderson, 1986). Lande (1982) has shown that the evolution of directional female mating preferences for male secondary sexual characters can greatly amplify large-scale geographic variation in male characters. This coevolutionary process can be enhanced by variation in the strength of direct or indirect selection on female preferences through the species distribution area (see Houde, 1993). Diversifying selection (where the direction of selection varies between populations) may speed up the evolution of species-specific songs and might increase the effectiveness of prezygotic sexual isolation between sympatric species (Etges *et al.*, 2006).

Among morphological traits, most variation (> 90%) in wing shape of *Drosophila* flies is correlated with the phylogenetic history of the species (K. van der Linde & D. Houle, unpublished data), although shape differences between closely related species can vary considerably (Houle *et al.*, 2003; K. van der Linde & D. Houle, unpublished data). Variation in wing shape can lead to functionally identical outcomes with various internal structural rearrangements, which suggests a large influence of genetic drift and, therefore, a correlation with phylogenetic history. For example, the convergence of clinal variation in wing size of *Drosophila subobscura* on different continents has been achieved through analogous, not homologous, changes in the relative lengths of different parts of the wing (Huey *et al.*, 2000). Additionally, wing traits have been found to evolve quite rapidly in response to geographic clines, e.g. in *D. subobscura* (Gilchrist *et al.*, 2000; Huey *et al.*, 2000), and they respond well to artificial selection (Houle *et al.*, 2003; Kennington *et al.*, 2003). These rapid changes are not surprising because of the abundant genetic variation related to wing shape that is available (Mezey & Houle, 2005; Weber *et al.*, 2005). The flies use their wings also for producing courtship songs, and so the wing can potentially be influenced by sexual selection. To date, little is known of the influence of wing morphology on male courtship song traits.

The size and shape of male genitalia are rapidly evolving species-specific characters and they are often used for species identification, e.g. in *Drosophila* species (Grimaldi, 1990). Genitalia have been suggested to evolve via lock–key mechanics (Dufour, 1844), pleiotropy (Mayr, 1963) or cryptic female choice (Eberhard, 1985). The striking morphological diversity of genitalia in the species of the *Drosophila virilis* group (Kulikov *et al.*, 2004) may, also, have arisen through sperm competition (Parker, 1970) or sexual conflict (Hosken & Stockley, 2004; Arnqvist & Rowe, 2005).

Our study species, *D. montana*, has a wide circumpolar distribution in temperate forests around the northern hemisphere (Throckmorton, 1982). The physical barrier of the Bering Sea between Eurasia and North America isolates populations on the two continents. Nevertheless, a study of nuclear sequences of three X chromosomal genes showed evidence for recurrent gene flow among North American populations and possibly also from North America to Finland (Päällysaho *et al.*, 2005). A recent population genetic study (Mirol *et al.*, 2007) showed that the North American populations from Vancouver and Colorado do not have distinct mitochondrial haplotypes, whereas the haplotypes of the Finnish Oulanka population differ from those of the North American populations and are more diverse. In the same study, the microsatellites differentiated all of the study populations. Mirol *et al.* (2007) estimated population divergence time for the Finnish and North American populations to be 0.45–0.90 Myr, with rapid population expansion of Oulanka and Vancouver populations about 35 000 years ago and a more gradual expansion of the Colorado population starting earlier. These estimates are in concordance with the expected genetic effects of likely range changes during the Pleistocene glaciations (Hewitt, 2004).

The objective of the present study was to analyse the phenotypic divergence of *D. montana* populations in male courtship song and in wing and genital size and shape in two North American and one Finnish population. The data on the genetic divergence of these populations (Mirol *et al.*, 2007) is used as a reference to compare the levels of genotypic and phenotypic divergence between populations. We hypothesize that adaptation of *D. montana* populations to different physical and biotic conditions (including different *Drosophila* communities) and the restricted gene flow between populations has led to the divergence of populations at several behavioural and morphological characters. The main issues we address here are the extent of the phenotypic divergence of populations in the male song, wing and genital traits when compared with their genetic divergence, and whether the evolution of behavioural and morphological characters has occurred in concert. Moreover, we attempted to pinpoint the selective pressures likely to be responsible for enhancing phenotypic divergence in the characters studied.

Materials and methods

Isofemale strains

Drosophila montana flies were collected from three locations: Oulanka (Finland; 66°27′N, 29°00′E, near sea level), Colorado (USA; 38°52′N, 106°59′W, 2700 m altitude) and Vancouver (Canada; 49°15′N, 123°06′W, near sea level). Wild-caught females were transferred individually into malt-food vials to lay eggs. They were moved into new vials after a few days to control for larval density and food availability. The F1 generation flies were sexed within 2 days of their emergence under light

CO_2 anaesthesia and the males and the females were kept in separate vials for 3 weeks, i.e. until they were sexually mature. We established seven isofemale lines for the Oulanka, six isofemale lines for the Colorado and 20 isofemale lines for the Vancouver population in 2003. Phenotypic characters (male courtship song, male wing size and shape, and male genitalia size and shape) were measured for three F1 males per isofemale line to allow the estimation of variation in these characters within and between populations. The same study lines were used for the mtDNA analysis and most of them also for the microsatellite analysis in the phylogeographic analysis of *D. montana* populations (Mirol *et al.*, 2007).

Recording and analyses of the male courtship songs

Drosophila montana males produce courtship song by wing vibration. To record this song we transferred individual males with a virgin, sexually mature female of the same strain into a Petri dish with a moistened filter paper on the bottom and a nylon net roof. The courting flies walked upside down on the roof of the chamber allowing song to be recorded by holding the microphone (AKG C 1000 S) directly above the roof of the chamber. Male courtship songs were recorded on a Marantz Professional (model no. 74PDM 502/02B) tape recorder at the temperature of 20 ± 1 °C and analysed with Signal 4.0 (Engineering Design, Belmont, MA, USA) sound analysis system. The analysed song traits were the number of pulses in a pulse train (PN), the length of a pulse train (PTL), the length of a sound pulse (PL), the interpulse interval (IPI, the length of the time from the beginning of one pulse to the beginning of the next one), the number of cycles in a sound pulse (CN) and the carrier frequency of the song (FRE). The first five song traits were measured manually in the oscillograms, PN and PTL for the whole pulse trains and PL, IPI and CN for the fourth pulse of each train (see Fig. 1). The carrier frequency of the song (FRE) was measured from the frequency spectrum of the studied pulse trains. For statistical analyses, we calculated the mean values of the song traits over three pulse trains for each male. Data for the male song frequencies have been used also in the study by K. Klappert *et al.* (unpublished data).

Analysis of wing morphology

The original procedure for imaging the wings and for estimating the vein locations and the corresponding landmarks has been described in detail by Houle *et al.* (2003). Recently, the extraction procedure of the landmark data has been modified, such that we can localize allometric variation more precisely (K. van der Linde & D. Houle, unpublished data). The wings of the flies, stored in 70% ethanol after the song recording, were cut with a sharp scalpel and attached to a microscope slide with transparent tape. A randomly chosen left or right wing of an individual was used for analysis. After a digital image of the wing was obtained, the operator marked two start points at the basal end of the wing. The position of the veins was described as clamped quadratic B-splines and extracted from the images using the splining procedure *FindWing* (Lu & Houle, 1995–1997). This programme is embedded in a wrapper programme *Wings* (van der Linde & Houle, 2004–2006), which automates its input and output. *Wings* also includes modules for outlier detection (e.g. faulty splined wings) using Minimum Volume Ellipsoids (MVE) (Rousseeuw & Leroy, 1987; Rousseeuw & van Zomeren, 1990; K. van der Linde & D. Houle, unpublished data) and subsequent correction of the control points of those incorrectly positioned splines.

The wing landmark data were aligned using the Generalized Procrustes Analysis (GPA). GPA seeks to remove nonshape variation by centring, scaling and rotating the landmark data to minimize the least-squares deviations among them. Centroid size was retained as the scaling variable so that size-dependent changes in shape

Fig. 1. Oscillogram of *Drosophila montana* courtship song. The traits measured from the oscillogram are: the number of pulses in a pulse train (PN), the length of a pulse train (PTL), the length of a sound pulse (PL), the interpulse interval (IPI) and the number of cycles in a sound pulse (CN). Song frequency (FRE) was measured from the Fourier spectra of the pulse trains.
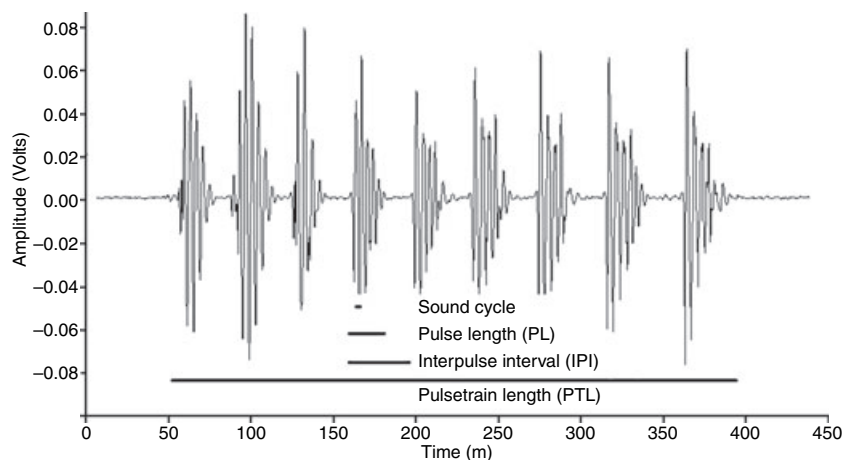
can be explored. Three degrees of freedom are lost in the estimation of the nuisance variables centring and rotation, and another degree of freedom is used to estimate centroid size. Two types of data were extracted from the wings: 15 pseudo-landmarks describing the outlines of the wing and 12 landmarks describing the junctions between wing veins and between the veins and the outline of the wing. The outline data contain considerably less allometric variation, and as such, fulfil better the assumption of isotropy that is required for the GPA (Dryden & Mardia, 1998). The two data sets were aligned simultaneously (van der Linde, 2005–2006). The outline data were used to determine the nuisance parameters and the nuisance parameters estimated for the outlines were used to determine the position of the landmarks. The outline shape data have 26 degrees of freedom (15 pseudo-landmarks, each with two dimensions, minus the four degrees of freedom lost as described above), whereas the landmark data have 24 degrees of freedom (12 landmarks times two dimensions, no loss of degrees of freedom as the nuisance parameters were estimated for the outline data). The repeatability of wing landmarks with this measurement technique is as high as 93% (Houle *et al.*, 2003).

Genital morphology

The genitalia were removed from the male bodies, which had been preserved in 70% ethanol after the song recordings had been completed. They were transferred singly into Eppendorf tubes with 0.1 M NaOH and kept at 95 °C for 8 min. Thereafter, NaOH was replaced with water. The distiphallus, the distal part of the aedeagus, was removed from the rest of the genital apparatus in water with fine forceps and placed on its side on a glass slide. After removing the excess water, the distiphallus was covered with EUKITT (Kindler, Freiburg, Germany). Slides were photographed with a SPOT Insight colour digital camera installed on a Zeiss Axioscop 40 microscope and stored as bitmap files. Files were analysed with the SHAPE 1.2 (Iwata & Ukai, 2002) programme, which is based on principal component analysis (PCA) performed on elliptical Fourier descriptors (EFD) of an enclosed contour (Kuhl & Giardina, 1982; see Fig. 3). The PCA was used on EFDs to describe morphological variation in the distiphalli, which, owing to their irregular shape, lack reliable landmarks. The resulting normalised Principal Component (PC) scores can be used as measured trait values that include only the allometric variation. The size component (area) was measured separately by the programme. The distiphallus will be referred to as the 'genitalia' for simplicity in the rest of the paper. Genital hook length was included in the analysis as an additional measure because it has frequently been used in species identification (Grimaldi, 1990). The hook length from the endpoint of the hook to the base of the hook at the head of the distiphallus

(Fig. 3) was measured twice for 10 individuals from each of the three populations giving a repeatability of 98.6% averaged over all repeats. Pictures used in the PCA were edited to black and white to enable automated measurements with program SHAPE 1.2. Repeatability for the PCA was measured for similar set of individuals as for the hook length and it varied from 100% to 94%, the last PCs being the least accurate ones.

Statistical analyses

All the statistical analyses were performed with SPSS 12.0.1 for Windows (SPSS Inc., Chicago, Illinois, USA), except for the PCA on genitalia shape described above. The data for the male courtship songs, wings and genitalia consisted of the measurements for three males of each isofemale line. Normal distribution and homogeneity of variance were tested with Kolmogorov–Smirnov and Levene's tests respectively. Log_{10} transformation was performed when necessary.

Principal component analysis was carried out to reduce the number of dependent variables to be entered in subsequent analyses for male genitalia. Only the traits describing the shape were included in the PCA; genitalia size (area) and genital hook length were used in the subsequent analysis without modification. The PCA was based on variance–covariance matrices to reduce the effects of elliptical harmonics describing minor aspects of variation that are not likely to have biological importance. The PCs that explained more than 1% of the shape variation were included in the variance–covariance-based PCA.

Variation between allopatric *D. montana* populations in male song, wing and genitalia traits was studied using the F1 progenies of the females collected in Oulanka, Vancouver and Colorado populations. The structure of the data (three males per isofemale strain) allowed us to analyse the magnitude of variance components within and between populations in the measured behavioural and morphological traits. The proportions of variation at different hierarchical levels were calculated with nested ANOVAS for unequal sample sizes (Sokal & Rohlf, 1997). The *F*-values and their significance for different hierarchical levels (between and within populations and within strains) were tested for each trait group (courtship song, wings and genitalia), using a sequential Bonferroni (Holm, 1979; Rice, 1989) corrected significance value to avoid type I error. Percentages of variation at different levels were calculated on the basis of the mean squares (MS) from the nested ANOVA. Significantly different traits between populations in nested ANOVAS were used in a discriminant analysis (DA) to test for the influence of overall interactions of the combined dependent variables and variance levels between populations.

The extent of overall phenotypic divergence between populations was studied using DA. The DA was carried out for the mean values calculated over the three males

from each isofemale strain to avoid pseudoreplication. The analysis defines axes that maximally separate the populations and yields correlations of the trait values to the separating axes. The reliability of the classification obtained was confirmed with a cross-validation test. Wilks' λ was used, in connection with the DA, to test the significance of the multivariate patterns and to determine the proportions of variance that the multivariate method explained.

MtDNA data for COI and II genes from Mirol *et al.* (2007) for the populations used in this study (seven strains from Oulanka, 20 strains from Vancouver and six strains from Colorado) were plotted using multi-dimensional scaling (MDS) analysis with the 'metric' option (performed in GenStat 8.0; VSN International Ltd, Oxford, UK). We also performed AMOVA for the same data set. We did not have microsatellite data for the same set of strains. Consequently, microsatellite and phenotypic data for the Oulanka, Vancouver and Colorado populations have been compared only in the discussion.

The courtship songs are produced by male wing vibration. Possible covariance between the wing and song traits was studied using ANCOVA on the mean values of these traits in isofemale strains.

Results

Variation in male song, wing and genitalia traits between populations

Male courtship songs

Nested ANOVAs for the male song traits showed that the CN, PL and IPI, describing the quality of sound pulses and the PN, giving the number of pulses in a pulse train, vary significantly between populations after sequential Bonferroni correction (see Table S1 in Supplementary material). CN and PL had the highest between population variance components (66.3% and 58.1% respectively) and IPI the lowest one (24.1%). Song frequency did not vary significantly among populations ($F_{2,30} = 4.74$, $P = 0.016$) after sequential Bonferroni correction ($p_{crit} = 0.0125$) and so it was not included in the subsequent DA.

In DA, the first and the second discriminant axes for the four song traits varying between populations accounted for 84.4% (eigenvalue 3.21) and 15.6% (eigenvalue 0.59) of the total variation between populations respectively. Wilks' λ of 0.15 for the two discriminant axes confirmed significant divergence in song among populations ($P < 0.001$). Divergence remained significant after the removal of the influence of the first discriminant axis (Wilks' $\lambda = 0.63$; $P = 0.004$). The first discriminant axis was mainly affected by CN and PL and the second one by IPI and PN (Table 1). The first axis separated the Colorado population from the Oulanka and Vancouver populations, whereas the second axis had little explanatory power

Table 1 The discriminant analysis structure matrices of the traits, highest correlation of a trait with a discriminant axis is marked with an asterisk (*). The abbreviations are: CN (the number of cycles in a sound pulse), PL (the length of a sound pulse), IPI (an interpulse interval), PN (the number of pulses in a pulse train), LM (a landmark), OL (an outline (pseudo)landmark), and PC (a principal component).

Songs			Wings			Genitalia		
	Discriminant axis			Discriminant axis			Discriminant axis	
Trait	1	2	Trait	1	2	Trait	1	2
CN	0.844*	0.395	SIZE	0.295*	0.107	PC4	0.564*	−0.008
PL	0.769*	0.363	LM Y1	−0.253*	−0.167	PC7	−0.181	0.912*
IPI	0.249	0.717*	LM X9	−0.245*	−0.017	PC3	0.523	0.527*
PN	−0.437	0.474*	LM X10	−0.239*	−0.036			
			LM X7	−0.230*	0.124			
			OL Y14	0.178	0.620*			
			OL Y13	0.086	0.488*			
			LM Y3	−0.037	−0.430*			
			OL Y9	−0.095	−0.401*			
			LM Y8	−0.308	0.395*			
			OL X10	−0.128	−0.395*			
			OL X11	−0.139	−0.372*			
			OL Y8	−0.095	−0.366*			
			OL X2	0.134	0.362*			
			OL X12	−0.135	−0.359*			
			OL X13	−0.132	−0.348*			
			OL Y15	0.119	0.341*			
			OL X14	−0.125	−0.310*			
			OL X5	0.133	0.301*			
			OL X3	0.151	0.299*			
			OL X4	0.157	0.292*			
			LM X4	−0.243	0.286*			
			LM Y7	−0.191	−0.286*			
			LM X1	−0.181	0.211*			

(Fig. 4). Cross-validated classification of the strains into their correct population of origin was high for the strains from the Colorado and Vancouver populations (83.3% and 95% respectively), whereas 71.4% of the strains from the Oulanka population were (mis)classified to the Vancouver population.

Male wing size and shape

The traits describing variation in male wing size and shape showed, on average, less variation between populations than the male song traits. Nested ANOVAs resulted in 24 traits, from the original 55 traits, varying between populations after sequential Bonferroni correction ($p_{crit} = 0.002$). The among population variance levels for these traits were between 59.9% and 16.5%. Variation within the strains was quite high, whereas the variation among the strains of the same population was generally low.

The first and the second discriminant axes for the wing measures explained 83.7% (eigenvalue 17.64) and
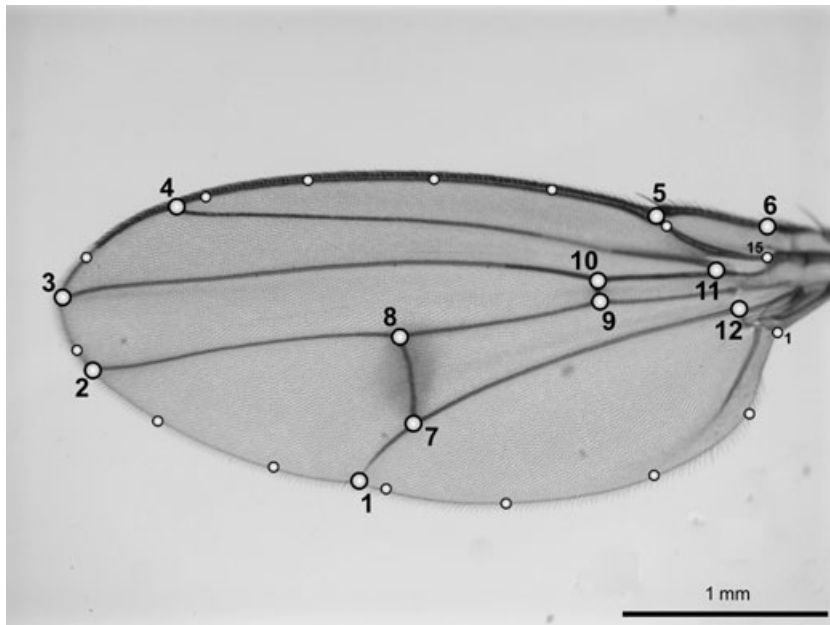
Fig. 2 Wing landmarks. Each landmark has been portioned into an *x* and *y* variable. Landmarks (LM) (large numbers) and outline (pseudo)landmarks (OL) (small numbers) are shown in the picture. Only the first and the last number, for the outline (pseudo)landmarks, are illustrated.
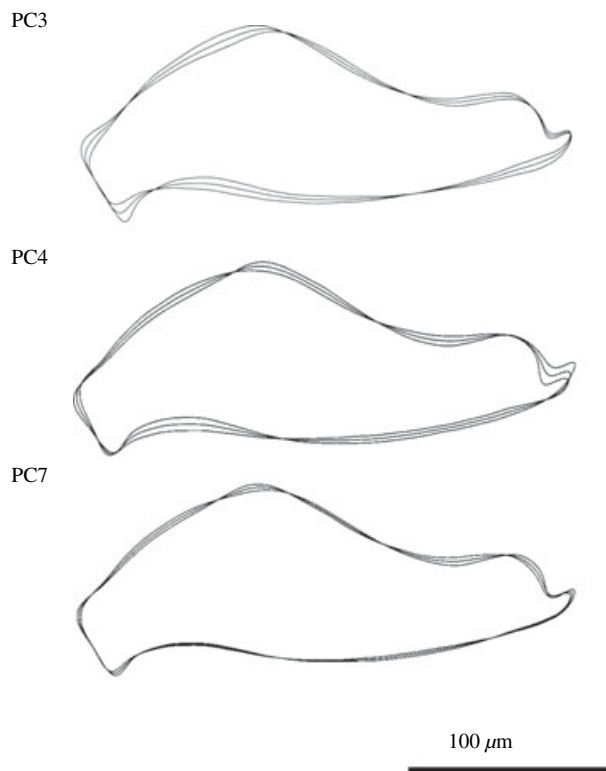

PC3

PC4

PC7

100 µm

Fig. 3 Genitalia outlines. Shapes corresponding to the mean and two times the standard deviation in both directions are shown for the three axes that contribute to significant variation among populations. A long distance between the lines indicates the highest variance. Shapes are normalized, but the approximate scale is shown.

16.3% (eigenvalue 3.44) of total variation. Wilks' λ had a combined effect of 0.012 for the two axes ($P < 0.001$) and of 0.23 ($P = 0.045$) after the removal of the first discriminant axis. The first discriminant axis was significantly affected by the centroid size of the wing and the landmarks (LM) 1, 7, 9 and 10, whereas the second discriminant axis was affected by several traits (Table 1, Fig. 2). The first discriminant axis separated the Vancouver population from the Oulanka and Colorado populations, whereas the second axis separated all populations (Fig. 4). Cross-validated classification of the strains into the correct population of origin was 57.1% for the Oulanka strains, 83.3% for Colorado strains and 95% for Vancouver strains. Oulanka strains were misclassified mainly to Colorado (28.6%) and vice versa (16.7%).

Male genitalia size and shape
The analysis of male genitalia shape gave a complicated pattern of correlations between EFDs, resulting in 14 PCs. Altogether 30 harmonics were used for the extraction of EFDs, the normalization method being based on the first harmonic. Variance–covariance matrices of the resulting 117 EFDs (*X* and *Y* dimensions, sine and cosine components for each harmonic) were analysed with PCA embedded in the SHAPE 1.2 software. The PCA with 14 axes explained 93.3% of the total variation. The genital hook length and the genital size (area) were treated separately.

The between populations variance level for the male genitalia traits was not as high as for the song and wing traits. Nested ANOVAs resulted to only three traits, out of the original 16 traits, varying between populations after sequential Bonferroni correction ($p_{crit} = 0.025$). PC7 showed the highest among population variance (18%),
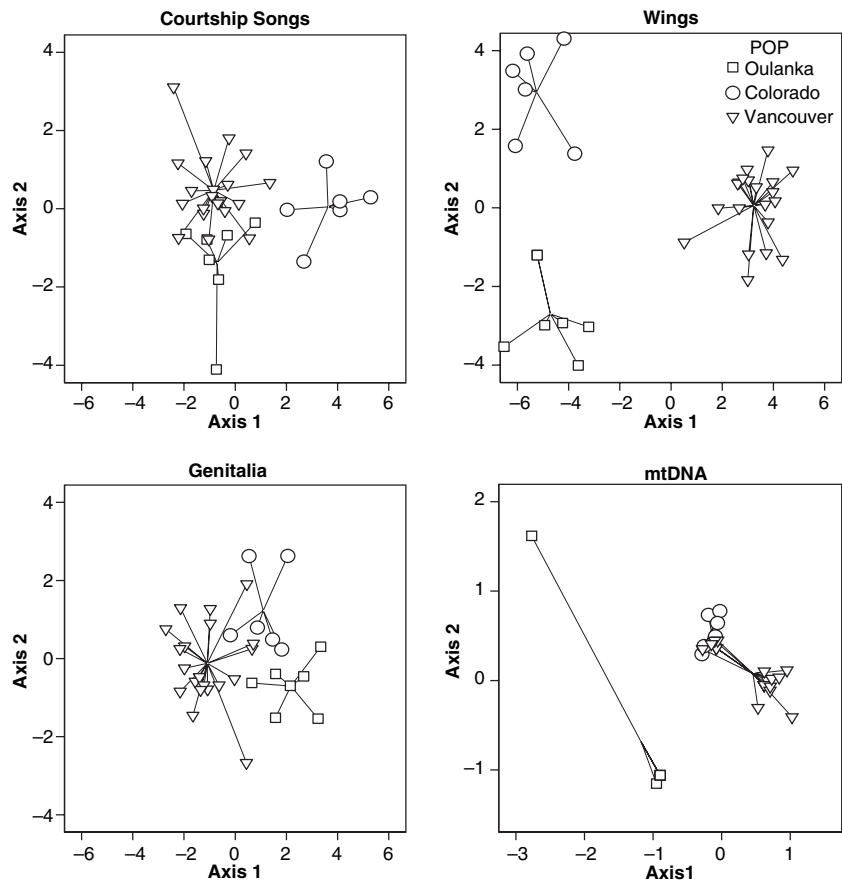
Fig. 4 Scatterplots of the discriminant axes and multi-dimensional scaling plot for mitochondrial DNA. Each point is the mean value for an isofemale strain. The Colorado population (six strains) is denoted by circles, the Oulanka population (seven strains) by squares and the Vancouver population (20 strains) by triangles. Note that fife Oulanka strains have coincident positions in the mtDNA plot around position (−1, −1).

followed by PC3 (13.3%) and PC4 (13.1%). PC3 describes variation in overall thickness of distiphallus, PC4 mainly in the angle and extension of the genital hook and the basal curvature of the distiphallus and PC7 in a specific curvature just above the hook (Fig. 3.). The hook length ($F_{2,30} = 3.61$, $P = 0.039$) did not vary significantly among populations after sequential Bonferroni correction.

The first and the second discriminant axis for the male genitalia traits accounted for 83.3% (eigenvalue 2.13) and 16.7% (eigenvalue 0.43) of among-population variance respectively. Wilks' λ for the two discriminant axes had a combined effect of 0.22 ($P < 0.001$) and an effect of 0.70 ($P = 0.006$) after the removal of the influence of the first discriminant axis. The first discriminant axis was mainly affected by PC4 and the second one by PC7 and PC3 (Table 1). The first axis separated all three populations from each other and the second one separated Colorado from the rest of the populations (Fig. 4). The cross-validated classification into correct population was 85.7% for the Oulanka strains, 50% for the Colorado strains and 80% for the Vancouver strains. A total of 33.3% of the strains from Colorado was misclassified to the Oulanka population.

Genetic vs. phenotypic divergence of the strains from different populations

Molecular data concerning the strains used in this study were obtained from Mirol *et al.* (2007). Mitochondrial (mt)DNA sequences (COI and II, 1358 bp) from seven Oulanka, six Colorado and 20 Vancouver strains were plotted in a MDS figure, based on F84 distances. The North American strains form two overlapping clusters with the mtDNA haplotypes mixed between the two populations. The Oulanka strains are clearly differentiated from the North American strains, this population showing high within population variation in mtDNA haplotypes. AMOVA results for the mtDNA were $F_{st} = 0.187$ ($P < 0.0001$), 18.7% variation among populations.

Covariation between wing and song traits

Covariation between the sexually selected pulse characters of male song (PL, CN and FRE; Ritchie *et al.*, 2001) and the wing traits was studied with ANCOVA to find out whether variation in wing morphology could have an effect on the quality of male song. Significant covariation was detected in ANCOVA for the wing landmark character LM X4 as the dependent variable, the populations as

the independent variable and the song characters PL ($F_{1,27} = 10.92$, $P = 0.003$), CN ($F_{1,27} = 9.05$, $P = 0.006$) and FRE ($F_{1,27} = 11.32$, $P = 0.002$) as covariates. The landmark character LM X4 pinpoints the endpoint of the second long vein on the outline of a wing (Houle *et al.*, 2003), i.e. the shorter the vein, the longer the PL and the higher the CN and the FRE.

Discussion

We analysed the extent of population divergence in phenotypic characters and compared it with the genetic data to trace the effects of selection in the phenotypic characters. We discovered strong divergence in the song and wing characters and milder divergence in the genital characters. Population divergence in these traits did not coincide with the extent of their genetic divergence, suggesting that the characters are not evolving neutrally.

The three *D. montana* populations, two from North America and one from Finland, showed significant geographic variation both in behavioural and morphological characters. As all the flies used in this analysis were reared under standardized laboratory conditions, this variation must have a genetic basis even though the effects of maternal and environmental factors cannot be fully excluded. The divergence between populations in the studied characters did not coincide with the degree of genetic divergence of the same populations (Mirol *et al.*, 2007), suggesting that the phenotypic characters are not evolving neutrally. If the phenotypic traits had evolved in line with the neutral genetic divergence between populations, the Finnish Oulanka population should have formed a separate cluster from the North American Colorado and Vancouver populations. At the phenotypic level, the divergence between the North American populations in song and wing data was, however, as high as or even higher than their divergence from the Oulanka population.

Low concordance between the genetic and phenotypic divergence of populations has earlier been found in several studies comparing the genotypic and phenotypic divergence between conspecific populations (see McKay & Latta, 2002) or subspecies (e.g. Chan & Arcese, 2003), the divergence at the phenotypic level usually exceeding the divergence at the genetic level. We were not able to use the F_{ST} and Q_{ST} statistics (e.g. Merilä & Crnokrak, 2001; McKay & Latta, 2002) for our study populations, mainly because of the low number of populations. In the present study, variation between the Colorado, Vancouver and Oulanka populations in phenotypic traits was 66.3% for the most diverged song trait and 59.9% and 18% for the most diverged wing and genitalia traits respectively. In comparison, the mtDNA divergence level between the three populations for the same set of strains as used in the present study was about 18.7% and the microsatellite divergence level 8.78% (same populations,

but partly different set of strains; data in Mirol *et al.*, 2007). Even though the data on the phenotypic and genetic divergence of populations are not strictly comparable (analysis of variance comprising different levels and partly different strains for microsatellites), the magnitude of the population divergence in phenotypic characters is clearly higher than their genetic divergence.

The most important selection pressures liable to have an effect on population divergence in male song characters are directional sexual selection exercised by females on male traits (e.g. Heisler *et al.*, 1987), diversifying selection towards the increase of divergence among sympatric species (Etges *et al.*, 2006) and balancing selection for maintaining species specificity (Lambert & Henderson, 1986). The song traits may be affected simultaneously by different kinds of selective pressures and they may also change through alterations in other genetically correlated song traits. In *D. montana*, the song is an essential element of male courtship rituals and a requirement for mating (Liimatainen *et al.*, 1992). The females of this species (Finnish population) have been found to exercise choice on male song traits, both in the wild (Aspi & Hoikkala, 1995) and in playback experiments with synthetic songs (Ritchie *et al.*, 2001). In these studies, the females have shown a preference for songs with a short PL and high CN and FRE, opening a possibility for Fisherian runaway or good-genes selection (Hoikkala *et al.*, 1998).

When studying variation in the male courtship songs of the laboratory strains of *D. montana* from a wide geographic area, Mirol *et al.* (2007) found the largest divergence to be accounted for by the pulse characters of the song, mainly the song frequency (FRE). In our study, we detected high variation between *D. montana* populations in the pulse characters CN and PL, variation in FRE becoming nonsignificant after Bonferroni correction. The fact that the progenies of wild-caught females showed no clear divergence in FRE in our study was mainly because of high within-strain variation. Our study strains had not been maintained for multiple generations in laboratory conditions, which can decrease within strain and increase between strain variation in FRE (S. Huttunen, J. Aspi, A. Hoikkala, J. Routtu, C. Schlötterer, unpublished). The song frequency is quite sensitive to changing environmental factors (Hoikkala & Suvanto, 1999) and it also has a low heritability (Aspi & Hoikkala, 1993; Suvanto *et al.*, 1998), and thus one would expect any change under selection in this character to be slow.

In addition to PL and CN, also IPI (the interpulse interval) varied significantly among populations. This trait has been found to play an important role in species recognition in *D. montana* (Saarikettu *et al.*, 2005) as well as in several other *Drosophila* species (e.g. Cowling & Burnet, 1981). The songs of all species of the montana phylad of the *D. virilis* group have a species-specific IPI (Hoikkala & Lumme, 1987). Variation between *D. montana* populations in male courtship song could

have been enhanced by character displacement, if the flies of different populations interact with different species in the wild. In Finland, interspecific courtships are quite common between sympatric species (Liimatainen & Hoikkala, 1998). In Colorado, *D. montana* occurs sympatrically with *Drosophila borealis* and *Drosophila flavomontana* (A. Hoikkala & D. Mazzi, personal observation) and in Oulanka with *Drosophila ezoana, Drosophila littoralis* and *Drosophila lummei* (Aspi *et al.*, 1993), whereas in Vancouver it is probably the only representative of the *D. virilis* group (K. Klappert & L. Orsini, unpublished observation). It is not possible to trace the historical patterns of sympatry between different species, but it is intriguing that a species-specific trait shows such high variation between conspecific populations. On the other hand, it is important for the species-recognition signals to differ from those of other sympatric species to effectively prevent hybridization.

The wing size and shape characters separated the two North-American *D. montana* populations from each other, the differentiation between the Colorado and Oulanka populations being of lower level. *Drosophila* species are known to evolve latitudinal wing morphology clines when introduced into novel environments (Gilchrist *et al.*, 2000; Santos *et al.*, 2004), reflecting increasing body size in colder climate (Huey *et al.*, 2000). Similarly, altitude has been show to have an effect on wing morphology by increasing wing load because of colonization or the scarcity of food resources (Norry *et al.*, 2001). In the present study, the morphological divergence in wing traits was most pronounced in the Vancouver population, which was separated from the Oulanka and Colorado populations by changes influencing mainly the size and the internal landmarks of the wings. Each population differed from the other two populations by changes mainly on the tip and the front and back edges of the wings. The differentiation, likely to have aerodynamic consequences, could be because of natural selection causing the wing form to adjust to the aerodynamic optimum of a local climate. Also, a wing landmark pinpointing the endpoint of the second long vein on the outline of the wing showed covariation with the combined effect of three pulse characters of the song. However, further experiments are required to find out what is the precise role of the wing morphology on song production.

The genitalia shape and size showed divergence between the Vancouver and Oulanka and Vancouver and Colorado populations in the DA, but the cross-validated classification test revealed an overlap in these traits between the Colorado and Oulanka populations. Nested ANOVAs (see Table S1) revealed strong within strain variance for all the genitalia shape traits suggesting them to be sensitive to environmental factors or to effects of sample preparation. *D. montana* females mate repeatedly in the wild, with the last male siring most of the offspring (Aspi & Lankinen, 1991), and so selective

pressure affecting these traits could be caused by cryptic female choice or by variation in the ability of males to remove a previous male's sperm with their genitalia. Also sexual conflict over the duration of copulation (K. Klappert & D. Mazzi, unpublished data) may influence the curvature of the genitalia and the genital hook angle to increase the males' ability to prolong copulation (Arnqvist & Rowe, 2005).

Adaptation of *D. montana* populations to different environmental conditions together with restricted gene flow between populations has led to the divergence of populations for several behavioural and morphological characters, as predicted. The phenotypic divergence of allopatric populations in the male courtship songs, wings and genitalia has clearly been enhanced by selection. It remains to be seen whether this divergence results in any reproductive isolation between populations.

Acknowledgments

References

Arnqvist, G. & Rowe, L. 2005. *Sexual Conflict*. Princeton University Press, Princeton and Oxford.

Aspi, J. & Hoikkala, A. 1993. Laboratory and natural heritabilities of male courtship song characters in *Drosophila montana* and *D. littoralis*. *Heredity* **70**: 400–406.

Aspi, J. & Hoikkala, A. 1995. Male mating success and survival in the field with respect to size and courtship song characters in *Drosophila littoralis* and *D. montana* (Diptera: Drosophilidae). *J. Insect Behav.* **8**: 67–87.

Aspi, J. & Lankinen, P. 1991. Frequency of multiple insemination in a natural population of *Drosophila montana*. *Hereditas* **117**: 169–177.

Aspi, J., Lumme, J., Hoikkala, A. & Heikkinen, E. 1993. Reproductive ecology of the boreal riparian guild of *Drosophila*. *Ecography* **16**: 65–72.

Chan, Y.L. & Arcese, P. 2003. Morphological and Microsatellite differentiation in *Melospiza melodia* (Aves) at a Microgeographic Scale. *J. Evol. Biol.* **16**: 936–947.

Cowling, D.E. & Burnet, B. 1981. Courtship songs and genetic control of their acoustic characteristics in sibling species of

the *Drosophila melanogaster* subgroup. *Anim. Behav.* **29**: 924–935.

Dobzhansky, T. 1951. *Genetics and the Origin of Species*. Columbia University Press, New York.

Dryden, I.L. & Mardia, K.V. 1998. *Statistical Shape Analysis*. John Wiley and Sons, Chichester, UK.

Dufour, L. 1844. Anatomie generale des dipteres. *Ann. Sci. Nat.* **1**: 244–264.

Eberhard, W.G. 1985. *Sexual Selection and Animal Genitalia*. Harvard University Press, Cambridge.

Etges, W.J., Over, K.F., Cardoso de Oliveira, C. & Ritchie, M.G. 2006. Inheritance of courtship song variation among geographically isolated populations of *Drosophila mojavensis*. *Anim. Behav.* **71**: 1205–1214.

Fisher, R.A. 1930. *The Genetical Theory of Natural Selection*. Clarendon Press, Oxford.

Gilchrist, G.W., Huey, R.B. & Serra, L. 2000. Rapid evolution of wing size clines in *Drosophila subobscura*. *Genetica* **112–113**: 273–286.

Grandcolas, P. & D'Haese, C. 2003. Testing adaptation with phylogeny: how to account for phylogenetic pattern and selective value together? *Zoologica Scripta* **32**: 483–490.

Grimaldi, D.A. 1990. A phylogenetic, revised classification of genera in the *Drosophilidae* (*Diptera*). *Bull. Am. Mus. Nat. Hist.* **197**: 1–139.

Heisler, L., Andersson, M.B., Arnold, S.J., Boak, C.R., Borgia, G., Hausfater, G., Kirkpatrik, M., Lande, R., Maynard Smith, J., O'Donald, P., Thornhill, A.R. & Weissing, F.J. 1987. The evolution of mating preferences and sexually selected traits. In: *Sexual Selection: Testing the Alternatives* (J. W. Bradbury & M. B. Andersson, eds), pp. 97–118. John Wiley and Sons, New York.

Hewitt, G.M. 2004. Genetic consequences of climatic oscillations in the Quaternary. *Philos. Trans. R. Soc. Lond. B.* **359**: 183–195.

Hoikkala, A. & Lumme, J. 1987. The genetic basis of evolution of the male courtship sounds in the *Drosophila virilis* group. *Evolution* **41**: 827–845.

Hoikkala, A. & Suranto, L. 1999. Male courtship song frequency as an indicator of male mating success in *Drosophila montana*. *J. Insect Behav.* **12**: 599–609.

Hoikkala, A., Aspi, J. & Suvanto, L. 1998. Male courtship song frequency as an indicator of male genetic quality in an insect species, *Drosophila montana*. *Proc. R. Soc. Lond. B.* **265**: 503–508.

Holm, S. 1979. A simple sequentially rejective multiple test procedure. *Scand. J. Stat.* **6**: 65–70.

Hosken, D.J. & Stockley, P. 2004. Sexual selection and genital evolution. *Trends Ecol. Evol.* **19**: 87–93.

Houde, A.E. 1993. Evolution by sexual selection: what can population comparisons tell us?. *Am. Nat.* **141**: 796–803.

Houle, D., Mezey, J.G., Galpern, P. & Carter, A. 2003. Automated measurement of *Drosophila* wings. *BMC Evol. Biol.* **3**: 25.

Huey, R.B., Gilchrist, G.W., Carlson, M.L., Berrigan, D. & Serra, L. 2000. Rapid evolution of a geographic cline in size in an introduced fly. *Science* **287**: 308–309.

Iwata, H. & Ukai, Y. 2002. SHAPE: a computer program package for quantitative evaluation of biological shapes based on elliptic Fourier descriptors. *J. Hered.* **93**: 384–385.

Kennington, W.J., Killeen, J.R., Goldstein, D.B. & Partridge, L. 2003. Rapid laboratory evolution of adult wing area in *Drosophila melanogaster* in response to humidity. *Evolution* **57**: 932–936.

Kuhl, F.P. & Giardina, C.R. 1982. Elliptic Fourier features of a closed contour. *Comp. Graph. Imag. Proc.* **18**: 236–258.

Kulikov, A.M., Melnikov, A.I., Gornostaev, N.G., Lazebny, O.E. & Mitrofanov, V.G. 2004. Morphological analysis of male mating organ in the *Drosophila virilis* species group: a multivariate approach. *J. Zool. Syst. Evol. Res.* **42**: 135–144.

Lambert, D.M. & Henderson, N.R. 1986. The stability of the specific-mate recognition system of *Drosophila melanogaster*. *Behav. Genet.* **16**: 369–373.

Lande, R. 1982. Rapid origin of sexual isolation and character divergence in a cline. *Evolution* **36**: 213–223.

Liimatainen, J. & Hoikkala, A. 1998. Interactions of the males and females of three sympatric *Drosophila virilis* group species, *D. montana*, *D. littoralis* and *D. lummei*, in intra- and interspecific courtships in the wild and in the laboratory. *J. Insect Behav.* **11**: 399–417.

Liimatainen, J., Hoikkala, A., Aspi, J. & Welbergen, Ph. 1992. Courtship in *Drosophila montana*: the effects of male auditory signals on the behaviour of the flies. *Anim. Behav.* **43**: 35–48.

van der Linde, K. 2005–2006. *Geometrics Package*. http://www.kimvdlinde.com/professional/geometrics.html

van der Linde, K. & Houle, D. 2004–2006. *Wings*. http://www.kimvdlinde.com/professional/wings.html

van der Linde, K. & Houle, D. 2006. *Applied Usage of the Minimum-Volume Ellipsoid*. [www document]. URL http://www.kimvdlinde.com/professional/mvepaper.html.

Lu, F. & Houle, D. 1995–1997. *FindWing*. http://bio.fsu.edu/dhoule/Software/

Mayr, E. 1963. *Animal Species and Evolution*. Harvard University Press, Cambridge.

McKay, J.K. & Latta, R.G. 2002. Adaptive population divergence: markers, QTL and traits. *Trends Ecol. Evol.* **17**: 285–291.

Merilä, J. & Crnokrak, P. 2001. Comparison of genetic differentiation at marker loci and quantitative traits. *J. Evol. Biol.* **14**: 892–903.

Mezey, J.G. & Houle, D. 2005. The dimensionality of genetic variation for wing shape in *Drosophila melanogaster*. *Evolution* **59**: 1027–1038.

Mirol, P., Schäfer, M., Orsini, L., Routtu, J., Schlötterer, C., Butlin, R.K. & Hoikkala, A. 2007. Phylogeographic patterns in *Drosophila montana*. *Mol. Ecol.* **16**: 1085–1097.

Moorhead, P.S. 1954. Chromosome variation in giant forms of *Drosophila montana*. *Univ. Texas Publ.* **5422**: 106–129.

Norry, F.M., Bubliy, O.A. & Loeschcke, V. 2001. Developmental time, body size and wing loading in *Drosophila buzzatii* from lowland and highland populations in Argentina. *Hereditas* **135**: 35–40.

Päällysaho, S., Vieira, C.P., Hoikkala, A. & Vieira, J. 2005. Evidence for introgression in differentiated North-American and Finnish *Drosophila montana* populations. *Genetica* **123**: 285–293.

Parker, G.A. 1970. Sperm competition and its evolutionary consequences in insects. *Biol. Rev. Camb. Philos. Soc.* **45**: 525–567.

Rice, W.R. 1989. Analyzing tables of statistical tests. *Evolution* **43**: 223–225.

Ritchie, M., Saarikettu, M., Livingstone, S. & Hoikkala, A. 2001. Characterization of female preference functions for *Drosophila montana* courtship song and a test of the temperature coupling hypothesis. *Evolution* **55**: 721–727.

Rousseeuw, P.J. & Leroy, A.M. 1987. *Robust Regression and Outlier Detection*. Wiley-Interscience, New York.

Rousseeuw, P.J. & van Zomeren, B.C. 1990. Unmasking multivariate outliers and leverage points. *J. Am. Stat. Assoc.* **85**: 633–639.

Saarikettu, M., Liimatainen, J.O. & Hoikkala, A. 2005. The role of male courtship song in species recognition in *Drosophila montana*. *Behav. Genet.* **35**: 257–263.

Santos, M., Iriarte, P.F., Céspedes, W., Balanyà, J., Fontdevila, A. & Serra, L. 2004. Swift laboratory thermal evolution of wing shape (but not size) in *Drosophila subobscura* and its relationship with chromosomal inversion polymorphism. *J. Evol. Biol.* **17**: 841–855.

Sokal, R.R. & Rohlf, F.J. 1997. *Biometry. The Principles and Practice of Statistics in Biological Research*, 3rd edn. Freeman, New York.

Suvanto, L., Liimatainen, J. & Hoikkala, A. 1998. Variability and evolvability of male song characters in *Drosophila montana* populations. *Hereditas* **130**: 13–18.

Throckmorton, L.H. 1982. *The Genetics and Biology of Drosophila*, Vol. 3b. Academic Press, New York.

Weber, K., Johnson, N., Champlin, D. & Patty, A. 2005. Many P-element insertions affect wing shape in *Drosophila melanogaster*. *Genetics* **169**: 1461–1475.

Wright, S. 1931. Evolution in Mendelian populations. *Genetics* **16**: 97–159.

Supplementary material

The following supplementary material is available for this article:

Table S1 Nested ANOVAS.

This material is available as part of the online article from: http://www.blackwell-synergy.com/doi/abs/10.1111/j.1420-9101.2007.01323.x

Please note: Blackwell Publishing are not responsible for the content or functionality of any supplementary materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.