MOLECULAR ECOLOGY

Molecular Ecology (2011) **20**, 2403–2424 doi: 10.1111/j.1365-294X.2011.05106.x

Reinforcement selection acting on the European house mouse hybrid zone

BARBORA VOŠLAJEROVÁ BÍMOVÁ,*† MILOŠ MACHOLÁN,† STUART J. E. BAIRD,‡ PAVEL MUNCLINGER,§ PETRA DUFKOVÁ,*¶ CHRISTINA M. LAUKAITIS,** ROBERT C. KARN,** KENNETH LUZYNSKI,†† PRISCILLA K. TUCKER,†† and JAROSLAV PIÁLEK*‡‡

*Department of Population Biology, Institute of Vertebrate Biology, Academy of Sciences of the Czech Republic, Brno, Czech Republic, †Laboratory of Mammalian Evolutionary Genetics, Institute of Animal Physiology and Genetics, Academy of Sciences of the Czech Republic, Veveří 97, CZ-60200 Brno, Czech Republic, ‡CIBIO, Centro de Investigação em Biodiversidade e Recursos Genéticos, Campus Agrário de Vairão, Vairão, Portugal, §Biodiversity Research Group, Department of Zoology, Faculty of Science, Charles University in Prague, Prague, Czech Republic, ¶Department of Genetics, Faculty of Science, University of South Bohemia, České Budějovice, Czech Republic, **College of Medicine, University of Arizona, Tucson, AZ, USA, ††Department of Ecology and Evolutionary Biology and Museum of Zoology, University of Michigan, Ann Arbor, MI, USA

Abstract

Behavioural isolation may lead to complete speciation when partial postzygotic isolation acts in the presence of divergent-specific mate-recognition systems. These conditions exist where *Mus musculus musculus* and *M. m. domesticus* come into contact and hybridize. We studied two mate-recognition signal systems, based on urinary and salivary proteins, across a Central European portion of the mouse hybrid zone. Introgression of the genomic regions responsible for these signals: the major urinary proteins (MUPs) and androgen binding proteins (ABPs), respectively, was compared to introgression at loci assumed to be nearly neutral and those under selection against hybridization. The preference of individuals taken from across the zone regarding these signals was measured in Y mazes, and we develop a model for the analysis of the transition of such traits under reinforcement selection. The strongest assortative preferences were found in males for urine and females for ABP. Clinal analyses confirm nearly neutral introgression of an *Abp* locus and two loci closely linked to the *Abp* gene cluster, whereas two markers flanking the *Mup* gene region reveal unexpected introgression. Geographic change in the preference traits matches our reinforcement selection model significantly better than standard cline models. Our study confirms that behavioural barriers are important components of reproductive isolation between the house mouse subspecies.

Keywords: androgen binding protein, assortative mating, behavioural cline analysis, major urinary protein, reproductive isolation, salivary cues, speciation, urinary cues

Received 22 June 2010; revision received 21 February 2011; accepted 7 March 2011

Introduction

According to the Biological Species Concept (Dobzhansky 1937; Mayr 1942), speciation, i.e. a process leading to evolution of separate species, requires the development of reproductive barriers between gene pools. Assortative mating is a potentially efficient prezygotic reproductive barrier and may thus prevent loss of genes into unfit hybrids (Lande 1981; Butlin 1995; Wells & Henry 1998; Panhuis *et al.* 2001; Turelli *et al.* 2001; Ptacek 2002; Coyne & Orr 2004; Ritchie 2007). When partial postzygotic isolation acts in the presence of divergent-specific mate-recognition systems, selection for increased mating specificity may lead to complete speciation (Dobzhansky 1940; Howard 1993; Servedio &

Correspondence: Miloš Macholán, Fax: +420 541212988; E-mail: macholan@iach.cz

‡‡Present address: Department of Population Biology, IVB, ASCR, Studenec, Czech Republic.

Noor 2003; Coyne & Orr 2004). This phenomenon can best be studied in closely related or recently divergent taxa where limited hybridization still occurs and speciation may be incipient. Here, selection may act to reinforce prezygotic isolation in regions of secondary contact, leading to avoidance of disadvantageous mating between genetically diverged taxa (Dobzhansky 1940; Butlin 1987, 1995; Howard 1993; Kirkpatrick & Ravigné 2002; Marshall *et al.* 2002; Servedio & Noor 2003; Servedio 2004; Hoskin *et al.* 2005; Smadja & Butlin 2006).

Two subspecies of the house mouse, *Mus musculus musculus* and *M. m. domesticus*, meet and hybridize along a contact front more than 2500 km long, running across southern Danish Jutland and from the Baltic Sea to the Black Sea coast through Central Europe (Boursot *et al.* 1993; Sage *et al.* 1993; Macholán *et al.* 2003). The width of this hybrid zone may vary for different markers (see Teeter *et al.* 2010; Macholán *et al.* 2011; and references therein), but usually it is more than a factor of ten wider than the per generation dispersal of individual mice. Virtually no F_1 hybrids are found within the zone, and the locality samples in the centre consist of a complex mixture of late-generation hybrids and backcrosses. Both sexes contribute equally to hybridization, and for most markers, zone structure is unimodal with predominating intermediate genotypes in the centre, characterized by the lowest fitness (Raufaste *et al.* 2005; Macholán *et al.* 2007).

Two lines of indirect evidence suggest that selection is acting against hybrids: (i) hybrid male sterility and partial female sterility have been described in different crosses of laboratory or wild populations (Forejt & Iványi 1974; Forejt 1996; Oka *et al.* 2004; Storchová *et al.* 2004; Britton-Davidian *et al.* 2005; Vyskočilová *et al.* 2005, 2009; Good *et al.* 2008; Mihola *et al.* 2009) and (ii) limited introgression of sex chromosome markers as compared with autosomes has been shown across four studied transects (Vanlerberghe *et al.* 1986; Tucker *et al.* 1992; Dod *et al.* 1993, 2005; Raufaste *et al.* 2005; Macholán *et al.* 2007). Recent genome-wide mapping studies have highlighted several candidate speciation genes, some of which are associated with olfaction, pheromone response and other behavioural aspects of reproduction (Harr 2006; Teeter *et al.* 2008), highlighting the potential importance of behavioural isolation between the two mouse subspecies. As olfactory cues represent the primary means of communication in nocturnal animals such as the house mouse (Beauchamp & Yamazaki 2003; Brennan & Kendrick 2006), diverged chemosignals and their receptors should be ideal candidates for premating isolation barriers.

Two candidate-specific mate-recognition systems have been proposed as behavioural barriers acting in the house mouse hybrid zone: one based on urinary cues (Smadja & Ganem 2002, 2005; Smadja *et al.* 2004; Ganem *et al.* 2008), including possibly the major urinary proteins (MUPs; Cheetham *et al.* 2007; Stopková *et al.* 2007; Thom *et al.* 2008), the other based on saliva cues encoded by genes of the androgen-binding protein (*Abp*) family (Karn & Dlouhy 1991; Laukaitis *et al.* 1997; Talley *et al.* 2001). A rigorous analysis of the potential roles of these systems in modulating house mouse interactions requires consideration of both parts of each mate-recognition system (the signal and the associated preference) across the geographic context of the hybrid zone, a situation requiring some modification of existing clinal models.

Haldane (1948) introduced the theory of a cline, i.e. a smooth change in the expectation of traits with geographic distance, to the neo-Darwinian synthesis. Clines of similar smooth sigmoid shape can be maintained by environmental (extrinsic) selection for different optima or intrinsic selection against admixture (Kruuk *et al.* 1999). This shape is modified when selection acts on multiple loci because dispersal leads to strong associations across loci where clinal change is rapid and these associations mean the effective selection on each locus is greater than if they were independent. The result is clines with steepened central portions or 'stepped' clines (Szymura & Barton 1986, 1991; Barton & Gale 1993; Baird 1995; Kruuk *et al.* 1999). These models of clinal change apply equally to frequencies of alleles and measures of quantitative traits (Barton & Gale 1993), and a number of clines in quantitative traits have been analysed, e.g. in toads (Nürnberger *et al.* 1995), grasshoppers (Bridle *et al.* 2001; Bridle & Butlin 2002), ground beetles (Takami & Suzuki 2005), butterflies (Blum 2008), sea gulls (Gay *et al.* 2008), and house mice (O. Mikula & M. Macholán, unpublished data).

Mate preference might similarly be expected to change as a quantitative trait cline across a contact zone. If two taxa prefer cues originating from their own rather than the other taxon, then a simple expectation is that cue preference will change from one extreme to the other across the zone, passing through a point of no-preference where hybridization has produced individuals that identify with neither pure taxon (e.g. grasshoppers: Butlin & Hewitt 1985a,b; Butlin & Ritchie 1991; or mice: Ganem *et al.* 2008). However, the phenomenon of reinforcement (i.e. the strengthening of postzygotic isolation as a result of emergence of prezygotic barriers; Coyne & Orr 2004) would modify this simple cline expectation, and in a manner quite different from the stepped cline effect arising from multilocus associations. This is because we expect reinforcement to modify behaviour most where the consequences of choice can lead to unfit descendants. If hybrids are less fit, rein-

forcement should then amplify consubspecific preference most close to a contact zone resulting in increased prezygotic isolation in sympatry relative to allopatry, a phenomenon called reproductive character displacement (Butlin 1995; Lemmon *et al.* 2004).

This will not only make the transition in preference at the centre of the zone steeper, it will qualitatively change the shape of the observed preference trait cline from a monotonic sigmoid, to one resembling a 'soliton' (a self-reinforcing solitary wave or pulse that maintains its shape; Bullough & Caudrey 1980; Lakshmanan 1988; Barton & Hewitt 1989). Butlin & Ritchie (1991) noted such a modification, and as well as fitting a sigmoid cline to their data they fitted a polynomial curve in order to better capture this unusual shape. Here, we develop an explicit model of the effects of reinforcement on a preference trait cline. An explicit model has the advantage that we can directly compare the likelihood of our observations when different levels of reinforcement are allowed. Moreover, testing for the effects of reinforcement using comparison of nested explicit hypotheses is much more straightforward to interpret in a biologically meaningful way than comparing clines and polynomial curves, because clines are described in terms of their centre and width, while these are not explicitly part of a polynomial curve parameterization.

In this study, our goal was to understand the contributions of mate-choice preference based on both salivary and urinary signals to the maintenance of the European mouse hybrid zone. We analysed both parts of these recognition systems, i.e. diverged signals and associated preferences, in two candidate subspecies-specific systems, in contrast to other studies that evaluated either only preference (Smadja & Ganem 2002, 2005; Smadja *et al.* 2004; Bímová *et al.* 2005; Ganem *et al.* 2008) or only signal genotypes (Dod *et al.* 2005). We asked how behavioural isolation contributes to the dynamics of the mouse hybrid zone. Specifically, we predicted that: (i) If the signals contribute to isolation, they must be diverged between the two subspecies and genes encoding the signals should have more limited introgression across the hybrid zone than neutral genes and (ii) If the associated preferences are to contribute to isolation then we should find assortative mating on both sides of the hybrid zone and we would expect to see reproductive character displacement in the hybrid zone, in terms of enhanced preferences, as predicted by the theory of reinforcement. We analysed clines for sexual preferences measured using simple two-way choice tests for urinary signals and androgen binding protein (ABP) signals. We also compared the introgression of molecular markers on chromosome 4 (*Mup* region) and chromosome 7 (*Abp* region) with loci assumed to be under strong selection and those assumed to be selectively neutral or nearly neutral with respect to the hybrid zone.

Materials and methods

Sampling

In total, 2408 mice were trapped at 128 sites scattered across an area 145 km long and 50 km wide, stretching from north-eastern Bavaria (Germany) to western Bohemia (Czech Republic). The sampling sites with numbers of tested individuals and scored markers are listed in Table S1 (Supporting information), and their position in the field area is shown in Fig. 1. Mice were live-trapped and either euthanized and dissected directly in the field or transported to the breeding facility of the Institute of Vertebrate Biology, Studenec (Czech Republic), for behavioural experiments. After sacrifice and dissection, samples of kidney and muscle were frozen and kept at −80 °C for subsequent electrophoretic analyses while a piece of spleen and/or tail was put in alcohol for DNA extraction. DNA was extracted using the DNeasy® 96 Tissue Kit (Qiagen), following the manufacturer's instructions.

Behavioural tests for ABP signals were carried out on 644 mice collected from 26 (females) and 28 (males) sites, respectively, while 490 individuals from 29 (females) and 26 (males) sites, respectively, were used for urinary preference experiments.

Molecular analyses

To analyse the introgression pattern of genes encoding signals involved in assortative preferences, we used several subspecies-specific markers either mapping to the corresponding region or located close to it. Fixation of each marker for alternative alleles was tested on a panel of 20 *M. m. musculus* and 20 *M. m. domesticus* individuals sampled from 40 allopatric populations well outside the hybrid zone, scattered well across the European ranges of the two subspecies. Only a single individual was taken from each population sample. Besides these mice, the testing panel also included individuals from wild-derived inbred strains representing the two taxa (Piálek *et al.* 2008). The only exception to this design was marker *4.057*, in which the number of sampled localities was increased to 83 with a total of 156 tested wild mice (see Table S2, Supporting Information).

Mouse ABPs are dimers composed of an alpha subunit disulfide bridged to a beta-gamma subunit (Dlouhy & Karn 1983; Dlouhy *et al.* 1987; see Emes *et al.* 2004 and Laukaitis *et al.* 2008 for revised nomenclature). The whole *Abp* gene region encompasses 3 Mb on the proximal end of Chromosome 7 and consists of 30 *Abpa*

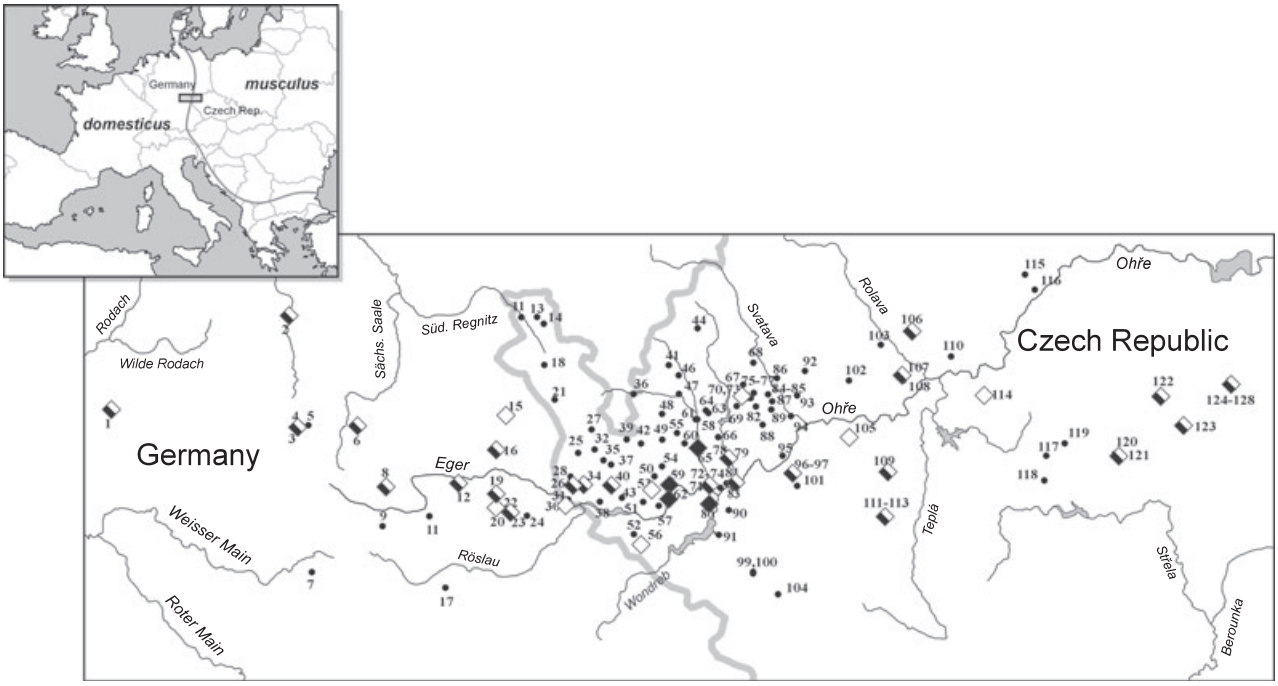

Fig. 1 The Czech-Bavarian transect across the *Mus musculus musculus/M. m. domesticus* hybrid zone in Europe with positions of collecting sites depicted. Black dots indicate sampling sites from which only molecular data are available whereas diamonds show localities sampled also for behavioural analyses; dark and light colours refer to individuals subjected to tests using ABP and urinary cues, respectively. The locality numbers are the same as in Table S1 (Supporting information). On the upper left panel, the position of the zone in a global scale with the study area indicated with shaded rectangle.

genes encoding alpha subunits and 34 *Abpbg* genes encoding beta–gamma subunits (Emes *et al.* 2004; Laukaitis *et al.* 2008). We analysed one of the *Abpa* genes, *Abpa27* (hereinafter referred to as *a27*) which maps near the distal end of the region (N = 2044 mice) and has a different allele fixed in each of the three *Mus musculus* subspecies ($Abpa^a$ in *M. m. domesticus*, $Abpb^b$ in *M. m. musculus* and $Abpa^c$ in *M. m. castaneus*) (Karn & Dlouhy 1991; Karn *et al.* 2002). We used a PCR-based method modified from Dod *et al.* (2005) and checked the results by sequencing 280 mice sampled from the most polymorphic localities. Sequencing revealed differences in 18 (6.4%) of tested individuals. In further analyses, only corrected genotypes were included.

Major urinary proteins are encoded by a gene cluster composed of approximately 40 genes and pseudogenes classified into two subgroups spanning a 1.9-Mb region located in the central part of Chromosome 4 (Bishop *et al.* 1982; Clissold & Bishop 1982; Logan *et al.* 2008; Mudge *et al.* 2008). No *Mup* has been evidenced to be diagnostic of the *M. musculus* subspecies. As a proxy for these genes, we analysed two SNPs flanking the *Mup* region on Chromosome 4: *4.59941702* (hereinafter abbreviated as *4.060*; N = 1477), which maps 137 640 bp proximally of the *Mup* gene cluster, and *4.62782738* (*4.063*; N = 1482), which is located 1 286 471 bp distal to it. Both markers were selected from a high-density genotyping array (Yang *et al.* 2009) and show fixed differences between *M. m. musculus* and *M. m. domesticus* mice sampled from across their ranges (L. Wang *et al.*, in preparation). For both markers, a subset of mice were genotyped using the Illumina® Goldengate® Assay on an Illumina® Beadstation 500 at the University of Michigan Genotyping Core and the remaining samples were genotyped using the TaqMan protocol. In addition, we analysed diagnostic loci on the same chromosomes, located at various distances from the *Abp* and *Mup* regions. On Chromosome 7, two B2 SINE retroelements and one U2 element were scored. Both B2 markers (*7B2-3720*, hereinafter referred to as *3720*; N = 1769; and *7B2-3746*, hereinafter *3746*; N = 1798) map very close to the proximal side of the *Abp* region, whereas the U2 element (*7U2_318M16*, hereinafter referred to as *318M16*; N = 1853) maps about halfway between the centromere and the *Abp* region (Fig. 2a). On Chromosome 4, we scored one SNP, *4.057* (N = 1426; Teeter *et al.* 2008), 2.4 Mb proximal to the *Mup* gene cluster, one B2 SINE element, *4B2_141I14* (*141I14*; N = 2313), and one allozyme locus, hexose-6-phosphate dehydrogenase (*H6pd*, E.C. 1.1.1.47; N = 2007), both located distally to the *Mup* genes (Fig. 2b). A brief description of each marker, its position, the primer sequences and PCR conditions are given in Table S3 (Supporting information). Genetic map positions were retrieved from the current mouse

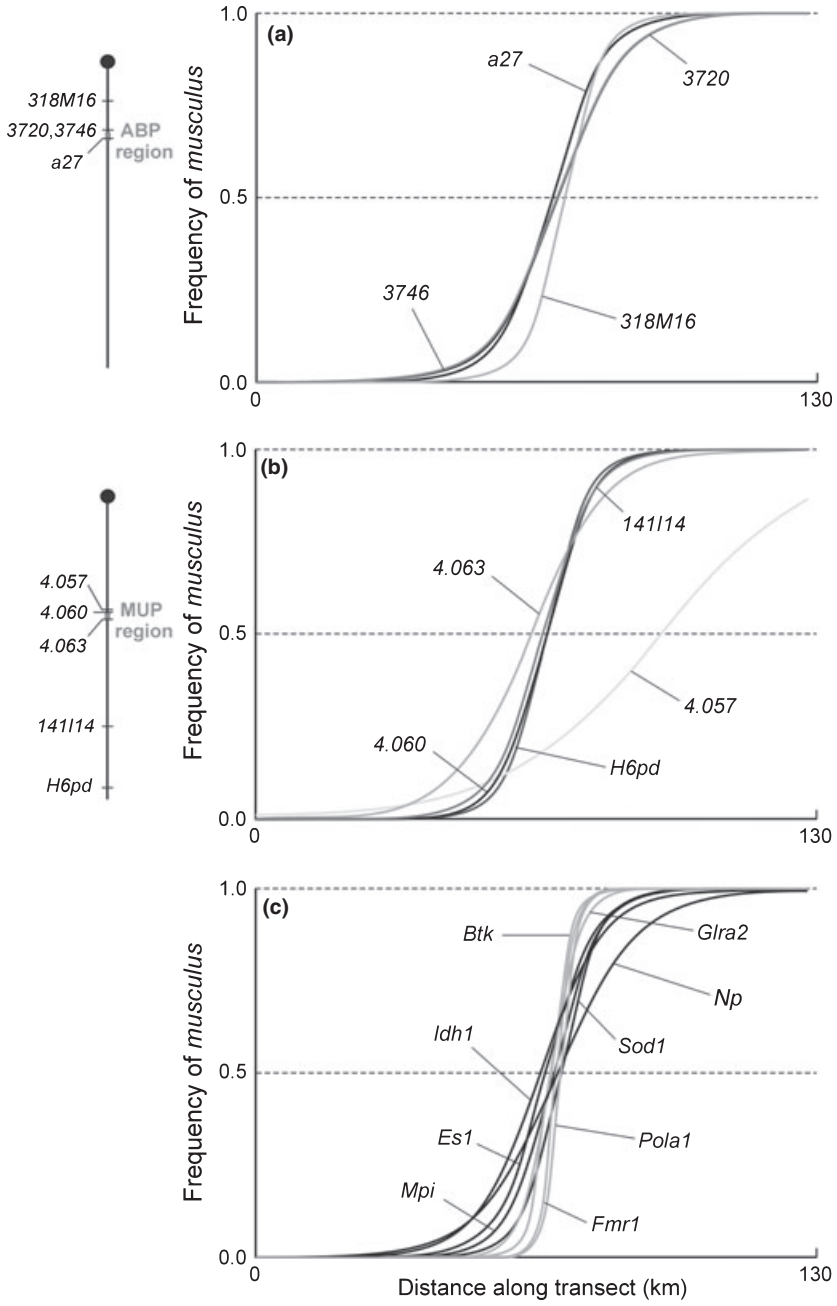

Fig. 2 Sigmoid clines for molecular markers on Chromosome 7 (a), Chromosome 4 (b), and loci on the X chromosome, assumed to be under selection (*Fmr1*, *Pola1*, *Btk*, *Glra2*), and various autosomes (*Es1*, *Idh1*, *Mpi*, *Np*, *Sod1*) considered neutral or nearly neutral markers (c). The positions of the Chromosome 4 and Chromosome 7 markers are indicated on the left side. Despite the large number of loci displayed in panel c, the clines at the *X*-linked loci (grey) are clearly steeper than those at autosomal loci (black).

genome build 37.1 (NCBI database; http://www.ncbi.nlm.nih.gov).

For comparison, we scored four *X*-linked and five autosomal markers. The X-chromosome loci were proposed because these are assumed to be under selection against admixture or linked to a selected locus (Payseur *et al*. 2004; Macholán *et al*. 2007, 2011), and owing to the map distance between them, we assumed linkage disequilibrium to be negligible. Three of these loci were SNPs, *Fmr1* (N = 1630; 65.95 Mb; 24 cM); *Pola1* (N = 1520; 90.88 Mb; 27.3 cM); and *Glra2* (N = 2285; 161.69 Mb; 57.9 cM). They were originally described in Payseur *et al*. (2004) and analysed as described in Dufková *et al*. (2011). The fourth *X*-linked marker was a B1 SINE mapping to the *Btk* gene (N = 2397; 131.08; 43.7 cM; see Munclinger *et al*. 2002, 2003, for details).

The autosomal markers were diagnostic allozyme loci assumed to be neutral or nearly neutral with respect to subspecies admixture (Macholán *et al.* 2007) and located on different chromosomes than the *Abp* and *Mup* regions: Isocitrate dehydrogenase 1 (*Idh1*, E.C.1.1.1.42; $N = 2138$; Chromosome 1); Superoxide dismutase-1 (*Sod1*, E.C. 1.15.1.1; $N = 2098$; Chr. 16); Nucleoside phosphorylase (*Np*, E.C.2.4.2.1; $N = 2209$; Chr. 14); Esterase 1 (*Es1*, E.C.3.1.1.1; $N = 2033$; Chr. 8); and Mannose phosphate isomerase (*Mpi*, E.C. 5.3.1.8; $N = 2093$). All allozymes were scored with standard horizontal starch gel electrophoresis (Harris & Hopkinson 1976; Pasteur *et al.* 1988) using samples of the C57BL/6J inbred strain as standards (Munclinger *et al.* 2002).

Behavioural experiments

After transportation to the breeding facility, all mice were cleaned of both ectoparasites (Arpalit® Neo applied onto the cage walls) and endoparasites (1% solution of Biomectin® administered peroraly). Males were housed singly and females either singly or in pairs captured at the same sampling site, under standard laboratory conditions (plastic cages $30 \times 15 \times 15$ cm, food and water available *ad libitum*, a 14:10 photoperiod, light on between 06.30 and 20.30 h). Prior to testing, the animals were allowed to habituate to the laboratory conditions for at least 1 month after capture.

The preferences for urinary and ABP signals were tested with a simple two-way choice test using a Y-maze (Talley *et al.* 2001; Bímová *et al.* 2005). The experimental apparatus consisted of a habituating box ($35 \times 25 \times 13$ cm) connected to the stem of a Y-maze (diameter: 5 cm; stem length: 35 cm; side arms length: 23 cm) and an electric pump insuring one-way air circulation in the apparatus (for the apparatus design and method details, see Bímová *et al.* 2005, 2009). Each tested individual was allowed to choose between *musculus* and *domesticus* (urinary or ABP) cues of the opposite sex.

Urinary signals were collected from wild animals at two localities, Straas (Germany), *c.* 46 km west of the hybrid zone centre (50°11′N, 11°46′E), representing pure *M. m. domesticus*, and Buškovice (Czech Republic), *c.* 57 km east of the zone centre (50°13′N, 13°21′E), representing pure *M. m. musculus* (for more details, see Piálek *et al.* 2008). Salivary ABP signals were collected using the isoproterenol-stimulated salivation method (Karn 1981) from two strains with a genetic background differing only in their *Abpa* allele (Laukaitis *et al.* 1997; Bímová *et al.* 2005). One of the strains possessed the $a27^a$ allele carried by the C3H/HeJ strain (purchased from ANLAB, Prague), whereas the $a27^b$ allele was carried by the $Abpa^b$-congenic strain established from DBA mice backcrossed to C3H/HeJ [provided by two of the authors (RCK & CML)]. Signal samples were pooled from at least eight individuals from the same sex and locality/strain and stored at −80 °C. Prior to each experiment, 10-μL aliquots of the signal were defrosted and spotted in the middle of a sterile 1.5×20 cm strip of filter paper and left to dry at room temperature for 30 min; then each sample was positioned at the bottom of one of the maze side arms.

At the beginning of each experiment, the tested individual was placed in the habituating box for at least 15 min. All experiments were carried out during the light phase of the diurnal cycle. Diurnal experiments were shown to yield results comparable to tests carried out during the dark period (i.e. the phase of increased activity) in a pilot study (B. Vošlajerová Bímová, unpublished results). Moreover, this experimental design allowed a direct comparison with previously published results (Laukaitis *et al.* 1997; Talley *et al.* 2001; Smadja & Ganem 2002; Bímová *et al.* 2005, 2009; Ganem *et al.* 2008).

After habituation, the door leading to the Y-maze was opened and the animal's behaviour was recorded for 5 min starting immediately after it left the habituating box. The Observer software (Noldus Technologies, Noldus *et al.* 2000) was used to aid analysis. If the same individual was to be involved in more than one test, at least 21 days intervened between tests. Female sexual receptivity was checked using vaginal smears after each test. No significant effect of oestrous cycle phase on sexual preference was found for either of the cue signals (ANOVA, urinary preferences: $F_{1,199} = 0.656$; $P = 0.419$; ABP preferences: $F_{1,244} = 1.578$; $P = 0.210$; see also Bímová *et al.* 2005).

At the end of the study, the mice were euthanized and samples of tissues were taken for molecular analyses. The whole study followed the experimental protocol (No. 5/05) approved by the Institutional Committee and Czech Academy of Sciences Committee for animal welfare. The breeding facility has been licensed (3245/2003–1020) for keeping small mammals according to Czech law since 2000, and the first author holds a license (V/1/2005/03) for experimental work on vertebrates in accordance with Czech law.

Data analyses

Molecular data. Because of the demic structure of mouse populations, with a single dominant male siring most of the offspring, estimates of allele frequency across localities were weighted by taking into account potential non-independence of observations of alleles because of relatedness and deviation from Hardy–Weinberg equilibrium. This weighting is expressed as the effective

number of alleles sampled (N_e) following the procedure described in detail in Macholán *et al.* (2008). The likelihood of observing a particular allele frequency at a locality was calculated using a modified binomial distribution (Edwards 1992) parameterized by N_e.

Geographic coordinates of sampling sites were gnomonically projected onto the plane using a routine written in Mathematica (Wolfram 1992) by one of the authors (SJEB). Subsequently, the two-dimensional transect was collapsed onto a line parallel to the most likely direction of change in allele frequencies, estimated in Macholán *et al.* (2008); the position of each locality was then defined as a distance along this one-dimensional transect line from the most distant locality on the *domesticus* side (Kübelhof).

Three cline models were fitted to the molecular data as described in Macholán *et al.* (2007): a simple sigmoid or *Tanh* model (Haldane 1948; Bazykin 1969), hereinafter referred to as the 'Sig' model, and two 'stepped' models (Szymura & Barton 1986), one symmetrical and one asymmetrical, hereinafter referred to as the 'Sstep' and 'Astep' model, respectively (Raufaste *et al.* 2005; Macholán *et al.* 2007). Two-unit support limits of the maximum likelihood estimate (MLE) of each parameter approximate frequentist 95% confidence intervals (Edwards 1992). For each model and each marker, the likelihood profile was constructed as described in Phillips *et al.* (2004) and Macholán *et al.* (2007) and when appropriate, these profiles were used for testing cline coincidence and/or concordance, employing ANALYSE 1.3 (Barton & Baird 1995).

Behavioural data. For each individual, cue preference in the Y-maze was assessed according to Smadja & Ganem (2002; see also Bímová *et al.* 2005), as the time spent sniffing one of the cues (T_{mus}, T_{dom}). The H_0 hypothesis that $T_{mus} = T_{dom}$ was tested using the Wilcoxon-matched pairs test with type I error set to $\alpha = 0.05$ adjusted with the Bonferroni correction. We also tested the difference in the total time spent by sniffing ($T_{mus} + T_{dom}$) between males and females for each cue using the Mann–Whitney *U*-test.

When treating genetic information, the amount of independent information sampled for alleles at each locality (N_e) was estimated. Similarly, consideration was given to how much information supports each estimate of the preference of individual *i*:

$$Y_i = \frac{(T_{mus} - T_{dom})}{(T_{mus} + T_{dom})}.$$

Previous approaches have calculated a locality preference estimated as the equal weight arithmetic mean over individual preference estimates (Smadja & Ganem 2002; Bímová *et al.* 2005); however, some of those individual estimates will be poorly supported, and others more strongly. If weakly supported estimates have random polarity they will tend to cancel each other out (i.e. their mean will tend to zero). If all individual estimates are given the same weight, this weakly supported noise will disproportionately obscure any strongly supported preferences. This can be taken into account by using a weighted arithmetic mean, with weights corresponding to the amount of information supporting each individual estimate, i.e. the total time an individual spends expressing interest as the measure to which its preference estimate is supported. The weighted mean estimate of preference at locality *L* is:

$$Y_L = \frac{\Sigma_L(T_{mus} - T_{dom})}{\Sigma_L(T_{mus} + T_{dom})}.$$

These estimates (Y_i, Y_L) range from −1 to +1 so that $Y < 0$ indicates preference for the *domesticus*-type signal, $Y > 0$ indicates preference for the *musculus* signal, and $Y = 0$ indicates the absence of any preference. The same weighting scheme is applied between localities. Some locality estimates will be better supported than others and should have more influence when fitting explanatory models. The information supporting the estimate as a weighting, in this case the weight for a locality estimate, is the total time all individuals spent expressing interest:

$$T_L = \sum_L (T_{mus} + T_{dom}).$$

Model of reinforced cline shape

To model clinal expectations for traits at localities with different distances across a hybrid zone, consider a locality distance x from the centre c of a cline width w: a natural simplification of this description is to work in terms of the locality's displacement from the centre measured in units of the width: $y = (x - c)/w$. First consider expectations for the hybrid index h, scaled from 0 (pure individuals of one taxon) to 1 (pure individuals of the other taxon). We follow a common wisdom that the house mouse hybrid zone is maintained by the balance between dispersal and selection against hybrids, i.e. it is a 'tension zone' (Key 1968). A simple tension zone expectation is that h will change as the sigmoid function $h(y) = [Tanh(2y) + 1]/2$ (Bazykin 1969; Nagylaki 1975; Slatkin 1975; Barton & Gale 1993). The gradient of change is then $h'(y) = Sech(2y)^2$, i.e. it is steepest at the centre, achieving a maximum value of unit width (equivalently, the width of a cline is defined as the gradient at the centre). Now consider the current study's preference measure (see behaviour data above). In the

case of complete assortative mating of pure individuals, this trait will change from −1 at one extreme to +1 at the other. Between these extremes, if preference for a cue is proportional to an individual's similarity to individuals at the cue's source (where similarity is measured using hybrid indices), then the expectation for the trait will follow the sigmoid cline running from −1 to +1, $Tanh(y)$. More generally, individuals may not show such extreme (−1,+1) levels of preference. If preference instead varies between p_0 and p_1, the trait will follow the cline $p(y, p_0, p_1) = p_0 + Tanh(y\Delta p^{-1})\ \Delta p$, where Δp is the difference in preference levels across the zone, $p_1 - p_0$.

So far, traits have been assumed to change according to a standard tension zone model. Our model of a *reinforced* preference cline p^* has two aspects: (i) *The selection driving reinforcement should be strongest where the most heterogeneous genetic backgrounds come into contact* (and so produce the most unfit hybrids). A measure of the local diversity of genetic backgrounds is the gradient in the hybrid index, and so the strength of reinforcing selection can be modelled as $s \propto h'(y)$. (ii) *Reinforcement amplifies existing mate preferences* (then, where no preference is expected, amplification will have no effect). For small s, p_0, p_1, we model reinforced preference as $p^*(y, p_0, p_1) = p(y, p_0, p_1)\ (1 + s)$. Substituting the model of selection (1) into the model of reinforcement (2):

$$p^*(y, p_0, p_1) = p(y, p_0, p_1)[1 + (R - 1)h'(y)],$$

where $(R - 1)$ is a constant of proportionality governing how much the local diversity in genetic backgrounds (producing unfit hybrids) has led to amplification of existing preferences. The gradient of the reinforced cline at $y = 0$ is R, and so R has a natural interpretation: it is the factor by which reinforcement increases the steepness of clines in preference traits in comparison with the cline in hybrid index.

The model is approximate in two senses: First, amplification of strong preferences may lead to a preference cline with values exceeding the (−1,+1) bounds of the preference measure. Second, amplification of asymmetric preferences (when preferences in the extremes p_0, p_1 are of very different magnitude) shifts the centre of the preference cline away from $y = 0$ (there is a portion of the cline with steeper gradient than at $y = 0$). However, both these effects are negligible for the cases we consider: preferences never exceed (−1,+1) and the maximum shift in $y = 0$ is <5% of the underlying cline width. It should be emphasized that this model is an oversimplification of reality: if preference is amplified and heritable, the genes encoding the preference will themselves diffuse, changing the expected form of the soliton. Those moving across the hybrid zone will be counter selected. Those moving away from the hybrid zone will arrive in areas where the consequences of poor choosiness are reduced. While these considerations are worthy of further exploration, for current purposes we believe the simple soliton-like expectation developed earlier is sufficient as an alternative model to be compared with standard cline expectations.

Results

Modelling clines of genes and genetic markers

Geographic cline analyses show that when the influence of outliers is reduced through estimating the effective number of alleles (N_e) only symmetrical (Sig, Sstep) models are statistically justified (Table 1). Moreover, for some loci both the symmetrical stepped model (Sstep) and the asymmetrical stepped model (Astep) converged at nonsensical combinations of parameters, and thus, only the Sig model was used for subsequent coincidence and concordance tests.

The locus assumed to be associated with subspecies recognition, *a27*, revealed a cline concordant (i.e. of the same width) and coincident (i.e. of the same position) with all autosomal loci except two markers flanking the *Mup* gene cluster (see below). Any differences in cline width between *a27* and the other three Chromosome 7 loci were nonsignificant, including *318M16* (Table 1, Fig. 2). Of the three SNP markers flanking the *Mup* region, *4.060* (0.14 Mb proximal) reveals a cline similar to *a27* whereas the sigmoid cline for *4.063* (1.3 Mb distal) is non-significantly broader and shifted westwards relative to both the loci (Table 1). Though there is considerable introgression on both sides for *4.063*, introgression of *musculus* alleles into the *domesticus* territory is stronger than the opposite process (Fig. 3c), making the cline almost significantly asymmetric (Table 1). The third locus, *4.057*, mapping 2.4 Mb away from the *Mup* region, shows a steep transition on the *domesticus* side, whereas there is no apparent decline of the *domesticus* allele frequency on the *musculus* side (Fig. 3a). This causes both the sigmoid and symmetrical stepped cline models to have extremely high width estimates with centre estimates substantially shifted towards the *musculus* range (Table 1, Fig. 2b). The proximal cause of these effects is confirmed by comparison with estimates allowing for asymmetry: the Astep model produces 'normal' width and centre estimates with poor explicative power on the *musculus* side. Because of the problems presented by their clines, the *4.057* and *4.063* loci were excluded from tests of coincidence and concordance. The remaining two loci (*141I14* and *H6pd*) on this

Table 1 Maximum likelihood estimates (MLEs) of cline shape parameters for 10 analysed loci: *c* = centre; *w* = width; *LL* = log-likelihood. Two-unit support limits of MLE for each parameter are given in parentheses. The parameters were estimated for three models: sigmoid (Sig), symmetrical stepped (Sstep) and asymmetrical stepped (Astep), and the best-fit model is indicated with asterisk

Locus	Model	*LL*	*w* (km)	*c* (km)
4.057	Sig*	−23.433	71.29 (47.4–122.0)	92.56 (83.7–108.4)
	Sstep	−23.433	71.30 (47.4–122.0)	92.57 (83.7–108.4)
	Astep	−20.660	12.60 (7.5–20.4)	70.29 (68.8–70.3)
4.060	Sig*	−20.339	20.80 (14.4–32.1)	65.78 (63.1–68.2)
	Sstep	−18.742	6.9 (0.0–32.1)	66.80 (64.8–68.7)
	Astep	−18.245	9.79 (3.2–22.1)	67.06 (62.2–73.3)
4.063	Sig	−26.320	34.84 (23.4–56.9)	62.67 (57.9–66.3)
	Sstep*	−22.592	4.00 (0.0–15.1)	66.72 (65.5–68.1)
	Astep	−19.889	9.92 (34.1–18.5)	66.85 (64.1–72.2)
141/14	Sig	−35.087	22.95 (16.9–32.6)	65.9 (63.6–68.1)
	Sstep*	−28.758	8.11 (0.0–16.1)	67.08 (65.7–68.6)
	Astep	−26.48	8.40 (0.0–14.2)	67.45 (65.9–73.4)
H6pd	Sig	−30.988	18.71 (13.8–26.5)	66.68 (64.7–68.6)
	Sstep*	−25.635	4.90 (0.0–13.6)	66.71 (65.7–68.1)
	Astep	−25.622	4.98 (0.0–14.9)	66.74 (64.2–68.5)
a27	Sig*	−22.213	24.14 (17.6–34.6)	67.62 (65.4–70.0)
	Sstep	−21.157	15.76 (0.0–30.9)	68.03 (65.8–69.9)
	Astep	−21.136	7.09 (0.0–29.2)	69.50 (66.0–78.3)
3720	Sig	−25.491	29.83 (21.5–43.6)	68.86 (66.2–71.8)
	Sstep*	−22.09	12.37 (0.0–26.9)	69.32 (67.0–71.2)
	Astep	−21.873	10.89 (0.0–25.8)	69.24 (62.2–74.9)
3746	Sig	−25.916	29.54 (21.3–43.1)	68.98 (66.3–71.9)
	Sstep*	−22.137	8.27 (0.0–24.4)	68.81 (67.1–71.2)
	Astep	−21.974	10.40 (0.0–24.4)	69.29 (62.4–72.1)
318M16	Sig*	−20.157	17.94 (13.0–25.6)	70.67 (68.8–72.8)
	Sstep	−19.196	4.66 (0.0–23.4)	70.7 (69.3–72.4)
	Astep	−18.668	4.40 (0.0–22)	68.13 (63.3–72.2)
Idh1	Sig*	−29.080	28.67 (20.6–42.0)	64.91 (61.8–67.8)
	Sstep	−26.610	10.5 (0.0–27.4)	66.69 (64.4–68.8)
	Astep	−24.196	9.19 (0.0–20.8)	67.42 (65.3–70.9)
Es1	Sig	−22.982	21.97 (15.8–31.8)	66.18 (63.7–68.6)
	Sstep*	−16.820	7.89 (0.0–17.0)	67.13 (65.5–68.6)
	Astep	−16.029	8.92 (0.0–15.2)	67.26 (65.7–70.5)
Mpi	Sig*	−27.631	20.98 (15.4–29.7)	66.98 (65.9–70.1)
	Sstep	−24.732	5.80 (0.0–20.6)	67.68 (66.6–70.3)
	Astep	−24.615	11.75 (7.6–21.7)	68.26 (62.2–76.1)
Np	Sig	−36.940	37.53 (26.7–56.3)	68.75 (65.2–72.7)
	Sstep*	−32.759	8.20 (0.0–25.2)	69.27 (67.3–71.3)
	Astep	−32.512	9.34 (0.0–24.6)	69.38 (67.3–73.6)
Sod1	Sig*	−28.231	17.99 (13.2–25.5)	69.45 (67.7–71.4)
	Sstep	−28.231	17.99 (13.2–25.5)	69.45 (67.7–71.4)
	Astep	−28.231	17.99 (13.2–25.5)	69.45 (67.7–71.4)
Fmr1	Sig*	−16.167	7.97 (5.5–11.8)	68.99 (67.7–70.4)
	Sstep	−13.729	6.16 (0.0–9.5)	69.12 (67.9–70.3)
	Astep	−13.429	6.22 (2.1–9.5)	69.12 (67.9–70.4)
Pola1	Sig*	−14.960	7.78 (5.2–11.7)	69.51 (68.2–70.9)
	Sstep	−14.484	7.21 (4.9–11.2)	69.53 (68.3–70.8)
	Astep	−11.966	2.6 (0.0–6.7)	67.94 (64.6–69.5)
Btk	Sig	−20.784	9.54 (7.0–13.3)	67.41 (66.2–68.5)
	Sstep*	−14.174	5.97 (0.8–8.7)	67.67 (66.7–68.6)
	Astep	−13.035	5.33 (1.9–8.4)	67.75 (67.0–70.0)
Glra2	Sig*	−24.105	13.41 (9.9–20.4)	67.43 (65.7–69.0)
	Sstep	−21.464	3.22 (0.0–15.7)	67.50 (66.6–69.0)
	Astep	−20.843	3.50 (0.0–12.5)	67.74 (66.7–72.7)

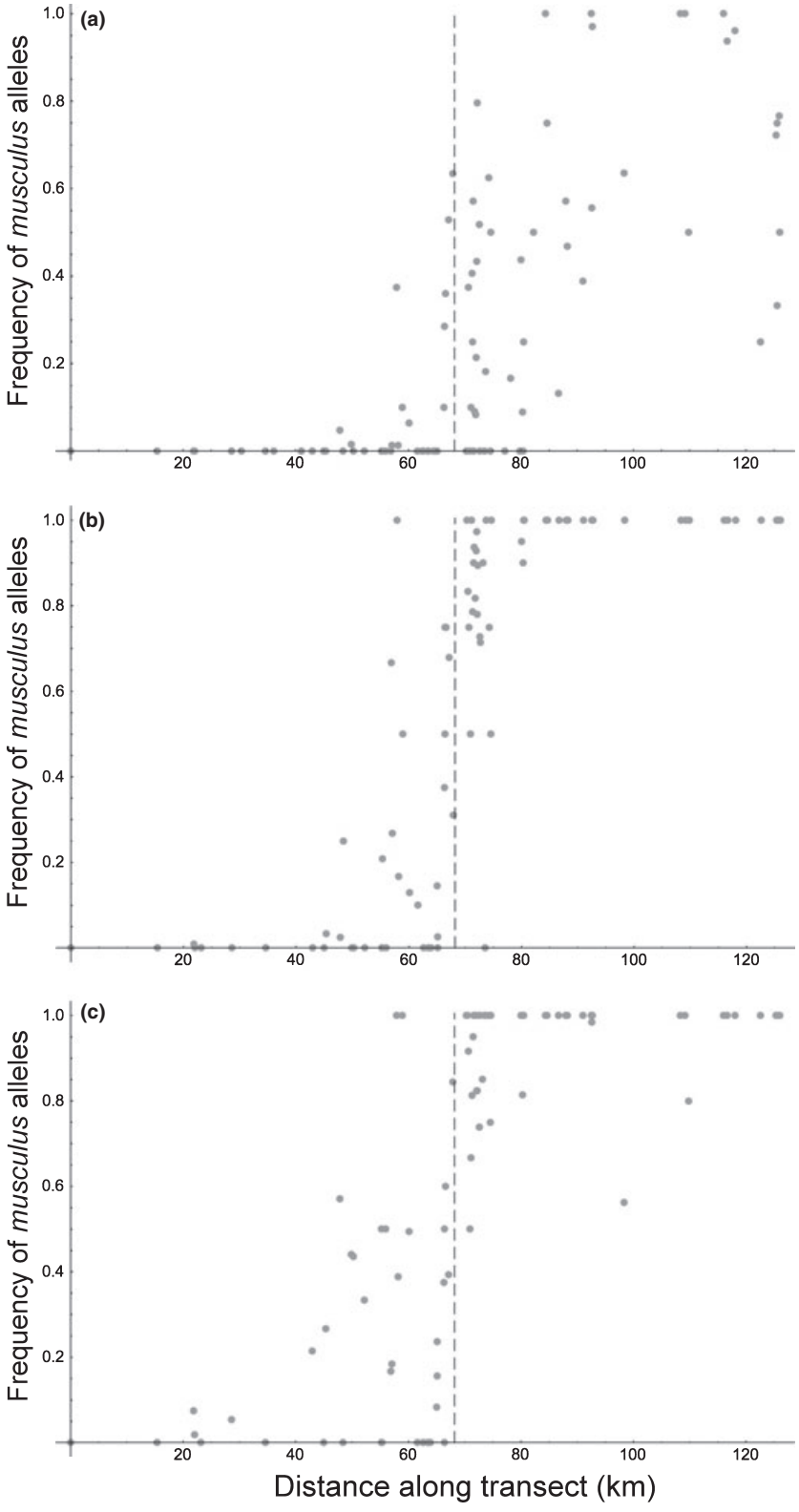

Fig. 3 Frequencies of *musculus* alleles at three Chromosome 4 loci plotted against geographic distance along the one-dimensional transect across the hybrid zone. (a) *4.057*; (b) *4.060*; (c) *4.063*. Vertical dashed lines depict the position of a consensus centre estimated from seven autosomal loci (Macholán *et al*. 2008).

chromosome did not yield unusual cline parameter estimates (Table 1, Fig. 2b). By contrast to the clines for *a27*, *4.060* and the others, clines for the X-chromosome loci are significantly narrower than all autosomal loci tested.

Behavioural data

The results of the behavioural tests are summarized in the Appendix separately for each locality, cue and sex. The data indicate that males show a higher level of preference than females for urinary cues, while the preferences of males and females for ABP are more similar to each other (see Appendix for details). Across populations, the direction of preference was varied (i.e. mice preferred individuals of either the same or the other subspecies), so we tested whether there was any consistent and significant trend towards either assortative or disassortative mating. Our null hypothesis that the direction of preference varies randomly within the same sex and cue, i.e. that the underlying ratio of consubspecific preferences (*domesticus*-oriented west of the zone centre and *musculus*-oriented east of the zone centre) to heterosubspecific preferences is parity. We tested this null hypothesis with a Chi-square test. In all cues and sexes, the overall preferences were consistently skewed towards assortative preferences, although the results were significant only in males sniffing urine ($\chi^2 = 18.615$; $P < 0.001$) and females sniffing the ABP signal ($\chi^2 = 4.840$; $P = 0.028$), others being only marginally (ABP males: $\chi^2 = 3.000$; $P = 0.083$; urine female: $\chi^2 = 2.793$; $P = 0.095$). As far as the total time of sniffing any cue is concerned, males spent significantly more time sniffing urine (Mann–Whitney $U = 251.1$, $Z = -2.116$, $P = 0.034$), whereas no difference between the sexes was found for ABP ($U = 359.0$, $Z = 0.087$, $P = 0.931$).

Testing for the presence of reinforcing selection

We considered a hierarchy of nested hypotheses designed to summarise the key features of the preference clines, accepting additional parameters only when they significantly increased the likelihood of the data conditioned on the model. The simplest hypothesis (H_0) assumes symmetric preference trait clines (i.e. with $|p_0| = |p_1|$), sharing the same level of reinforcement R, all with the same cline centre c and width w. We constrained the preference clines centre to be the best estimate of the hybrid zone centre MLE ($c = 68.2$ km). H_0 therefore had six free parameters: four for preference levels (male and female extreme preferences for urine and ABP), width w, and reinforcement level R; MLE (H_0) = −63.812.

Hypothesis H_1 allowed for preference asymmetry. The extreme preference magnitudes for one trait were allowed to be asymmetric ($|p_0| \neq |p_1|$), i.e. there was one additional free parameter in comparison with H_0; MLE (H_1) = −61.425. Only the cline for male preference for consubspecific urine appeared significantly asymmetric (H_1 vs. H_0: $2\Delta LL = 4.774$, 1 d.f., $P = 0.0289$), whereas the likelihood increase for the remaining three signals was nonsignificant (female urine: $2\Delta LL = 0.004$, $P = 0.9496$; male ABP: $2\Delta LL = 1.228$, $P = 0.2678$; female ABP: $2\Delta LL = 0.548$, $P = 0.4591$).

Hypothesis H_2 allowed for shifts in cline centre. Here, symmetric preferences were assumed (apart from that for male urine), but one preference cline at a time was allowed to be displaced from the zone centre (8 free parameters: 5 for preference levels and 1 each for w, R, and the displaced centre c; MLE (H_2) = −57.488). All but one cline was coincident (H_2 vs. H_1: male urine: $2\Delta LL = 1.008$, $P = 0.3154$; female preference for urine: $2\Delta LL = 2.022$, $P = 0.1550$; female preference for ABP: $2\Delta LL = 0.002$, $P = 0.9643$), the noncoincident cline being male preference for ABP ($2\Delta LL = 7.874$, $P = 0.0050$).

Hypothesis H_3 explored heterogeneity of reinforcement levels. The previously demonstrated asymmetry in male preferences for consubspecific urine and male ABP preference cline displacement were taken into account, but now one trait was allowed a different level of reinforcement (9 free parameters: 5 preference levels, w, R, c[ABP], and the reinforcement outlier value R_1). The test revealed no significant difference between the two hypotheses (H_3 vs. H_2: male urine: $2\Delta LL = 2.494$, $P = 0.1143$; female urine: $2\Delta LL = 2.692$, $P = 0.1009$; male ABP: $2\Delta LL = 1.008$, $P = 0.3846$; female ABP: $2\Delta LL = 0.034$, $P = 0.8537$). That is, there was insufficient evidence to accept a complicated hypothesis of multiple levels of reinforcement over the simpler single-level alternative.

Finally, we contrasted the ML values for H_0, H_1 and H_2 with no reinforcement (i.e. fixing R at 1) and allowing R to vary. In all cases, reinforcement yielded significantly higher likelihoods (Fig. 4; see Table 2 for the H_2 parameters).

We conclude that (i) In contrast to other traits, male preference levels for female urine cues were asymmetric: *domesticus* males preferred consubspecific female urine signals roughly twice as much as *musculus* males. This asymmetry was robust across details of the hybrid zone model. (ii) The change in male preferences for ABP was displaced *c.* 10 km west from the MLE zone centre, no matter whether reinforcement was invoked or not. (iii) There was no significant evidence in the data of different levels of reinforcement acting on the traits (H_3 vs. H_2), i.e. the cost associated with being a hybrid seemed to act equally on all traits. If there is no reinforcement (R con-

Table 2 Parameters of hypothesis H_2 with reinforcement; w and wR are in kilometres, c is expressed as westward displacement in kilometres of the male ABP preference cline from the consensus cline (see text for details); in the last row, the asymmetry of the male urine preference cline is given as the absolute value of the ratio p_0/p_1; in parentheses are 2-unit MLE support bounds

Parameter	MLE
w	24.9 (16–41)
R	2.59 (1.7–3.1)
Rw	64.4 (36–124)
c (male ABP)	9.30 (11.2–6.2)
$\left\|\frac{p_0}{p_1}\right\|$(male urine)	1.63 (1.1–2.1)

strained at 1), cline width estimates were not robust to details of the model (this may be seen as an indication that the nonreinforcement models in general were poor), while reinforcement ($R > 1$) was strongly supported over all scenarios considered. The MLE strength of reinforcement, R, and width of preference trait clines were robust to model details: reinforcement steepens preference clines by a factor of 2–3, making them between 25 and 30 km wide. The scale of change in underlying hybrid index on which reinforcement is proposed to have acted is also robust to model details.

Discussion

Introgression of signal markers

Both theoretical models and empirical studies on patterns of variation in secondary sexual traits and mating preferences suggest that behavioural premating isolation may play an important role in speciation (Lande 1981; West-Eberhard 1983; Butlin & Ritchie 1991; Ryan & Rand 1993; Butlin 1995; Ptacek 2000, 2002; Tregenza *et al.* 2000; Panhuis *et al.* 2001; Turelli *et al.* 2001; Coyne & Orr 2004; Ritchie 2007). In the presence of hybridization, premating isolation may be maintained either by the direct action of sexual selection or through reinforcement driven by natural selection against hybridization (see Panhuis *et al.* 2001; Turelli *et al.* 2001; Coyne & Orr 2004; and references therein). In this study, we assessed the contribution of mate choice preferences based on salivary and urinary signals to the dynamics of the house mouse hybrid zone. More specifically, we analysed the pattern of transition of both parts of sexual mate recognition signalling, i.e. molecular markers mapping at various distances from candidate loci encoding olfactory cues likely to be signals, and the reception of those signals as expressed in terms of preferences, across the Central-European portion of the zone.

Notwithstanding whether a behavioural barrier has come about through drift, sexual or natural selection, we should expect the transition of both the signal and preference loci across the zone to be steeper than the transition of selectively neutral or nearly neutral traits. However, molecular markers mapping or flanking the *Abp* and *Mup* regions revealed clines contradicting this expectation. The cline for *a27*, mapping an *Abpa* paralogue at the distal end of the *Abp* region and the cline for *4.060*, the marker most closely linked and proximal to the *Mup* region, were similar to those for allozyme loci (Table 1; Fig. 2), while cline for the X-chromosome loci were significantly narrower than the autosomal loci tested.

In the case of *a27*, this result seems to corroborate previous studies from the Danish transect (Dod *et al.* 2005) suggesting that linkage of *a27* with a selected locus in close proximity, rather than selection acting on *a27 per se*, is responsible for the pattern observed. The two markers mapping the proximal end of the *Abp* region and located more than 3 Mb from *a27* (*3720*, *3746*) revealed clines wider by more than 5 km (Table 1, Fig. 2) than *a27*. Although this difference was not significant, it is consistent with the notion that they map further from the region where selection is acting. On the other hand, the clines for the *318M16* U2 element, mapping *c*. 15 Mb from the *Abp* region, were substantially narrower than the remaining loci on Chromosome 7, probably reflecting increased incompatibilities of pericentromeric chromosomal regions within the hybrid zone, as suggested by Fel-Clair *et al.* (1996) based on data from Denmark (see also results of Tucker *et al.* 1992; from southern Germany). However, neither of these markers revealed clines concordant with *Pola1*, i.e. the X-chromosome locus that has been shown to be under strong selection against hybridization (Payseur *et al.* 2004; Teeter *et al.* 2010; Dufková *et al.* 2011; Macholán *et al.* 2011), nor with another two *X*-linked loci analysed, *Fmr1* and *Btk*.

The results of a large-scale study of Karn *et al.* (2002) showing limited secondary admixture between house mouse subspecies suggested that the evolution of *Abpa* was more complex than previously thought (Karn & Dlouhy 1991; Laukaitis *et al.* 1997). In fact, it is possible that the positive selection acting on *a27* reported by Karn & Nachman (1999) and Karn *et al.* (2002) was actually affecting the closely linked *Abpbg26* and *Abpbg27* genes, because they also have sites that differ in representatives of the three subspecies of *M. m. musculus* (Karn & Laukaitis 2003). Indeed, Emes *et al.* (2004) showed that the *Abpbg* subunits have more sites under selection than the *Abpa* subunits. Androgen-binding proteins function as dimers consisting of an alpha and beta (now beta-gamma) subunit, encoded by one of 30

Abpa paralogues and one of 34 *Abpbg* paralogues, respectively, found in a rapidly evolving region of the mouse genome (Laukaitis *et al.* 2008; Karn & Laukaitis 2009) and their relative roles as olfactory signals are yet to be unravelled. In any event, all the evidence available strongly suggests that it is ABP that is acting as the pheromonal signal, as there are amino acid residues under selection in both subunits, and in the dimeric combination, these sites all map to one face of the dimer (Emes *et al.* 2004). Consequently, we suggest that the analysis of introgression patterns across the European hybrid zone should be extended to other *Abp* loci.

The situation around the *Mup* cluster appears even more complex. The marker closest to this region, *4.060* (0.14 Mb proximal), revealed a cline similar to most autosomal loci, whereas the other flanking locus, *4.063* (1.29 Mb distal), appeared slightly more introgressed (Fig. 3c) to the *domesticus* side, rendering the sigmoid cline almost 35 km wide and shifted about 5 km westwards. Interestingly, this introgression appears to be localised to the area of massive introgression of *musculus* Y chromosome and some *X*-linked loci into the *domesticus* territory described by Macholán *et al.* (2008, 2011); both stepped models rendered clines similar to *4.060*.

The third SNP relatively close to the *Mup* region, *4.057* (2.4 Mb proximal), displayed an unusual pattern with massive introgression of *domesticus* alleles into the *musculus* territory (Fig. 3a). This pattern was first described by Teeter *et al.* (2008) from southern Germany who suggested it was because of strong positive selection acting on *Mup* genes. If so, this introgression would span hundreds of kilometres beyond the zone because *domesticus* alleles are also found in considerable frequencies in southern Moravia (Czech Republic), Poland, Hungary, and as far to the east as in Kharkov in eastern Ukraine, i.e. more than 1700 km from the zone centre (Table S2, in Supporting information). However, there are alternative explanations, namely ancestral polymorphism in *M. m. musculus* because of incomplete lineage sorting or balancing selection, or a westward movement of the zone, the latter scenario being corroborated by introgression patterns at other loci (Macholán *et al.* 2011).

Sexual preferences

The Y-maze tests revealed differences in odour preferences both between sexes and analysed stimuli (Appendix, Fig. 4). Despite this heterogeneity, preference clines were in general symmetrical and coincident, the exceptions being the 'male-ABP' cline (i.e. the cline for male preference of the ABP signal) which was found to be significantly shifted into the *domesticus* territory and the 'male-urine' cline (i.e. the cline for male preference of urine) with a higher level of consubspecific preference in *M. m. domesticus* males (Fig. 4). If the westward shift of the 'male-ABP' cline is a genuine reflection of the state of the hybrid zone, this pattern suggests a stronger preference of hybrids for the *musculus* signal in the hybrid zone where both *a27* alleles are present (Bímová *et al.* 2005). Similar *musculus*-biased preferences were revealed for urinary stimuli in F_1 hybrids by Christophe & Baudoin (1998); however, we did not find any significant prevalence of preferences for *musculus* urinary signals on the *domesticus* side of the zone. If species-specific preference is a result of a self-referencing system (Todrank *et al.* 2005), we should expect a similar shift also in associated signals. However, we found no significant cline shift of *a27* relative to all analysed allozyme and X-chromosome markers.

Salivary ABPs probably play an important role as signals transmitted during the close contact of interacting individuals (Laukaitis *et al.* 1997; Luo *et al.* 2003; Bímová *et al.* 2009) and thus their importance may be masked in Y-maze tests using dried saliva as suggested by results of tests based on a direct contact of the test subjects and the donor individuals when compared with those based on saliva spots only (Laukaitis *et al.* 1997; Talley *et al.* 2001; Bímová *et al.* 2009). Thus, even though during experiments using Y-maze and saliva spots tested animals are also in close contact with the signal, the tests may be too conservative and their results should be considered to be a lower bound to 'true' preferences. In contrast, preferences for urinary stimuli are less likely to be affected by the test design, as these cues are naturally used as long-lasting scent marks (Beynon *et al.* 2001; Hurst & Beynon 2004; Bímová *et al.* 2009).

Each adult mouse expresses a unique fixed pattern of 8–14 different MUP isoforms corresponding to its genotype (Beynon *et al.* 2002), and this has been likened to a protein 'bar code' (Beynon & Hurst 2003; Cheetham *et al.* 2007; Logan *et al.* 2008). However, for any protein to constitute a subspecies recognisable and discrimination mechanism, it must possess a molecule, or a combination of molecules consistently similar among members of either subspecies but significantly different between the two subspecies to be recognisable. In the case of MUPs, there has been speculation that their subspecies specificity can be attributed to expression and concentration differences rather than to differences in individual specific isoforms (Robertson *et al.* 2007; Stopka *et al.* 2007; Stopková *et al.* 2007; Hurst 2009; Janotová & Stopka 2009). Recent genetic studies suggest that high structural heterogeneity of *Mup* genes may reflect functional divergence within the family (Logan *et al.* 2008; Mudge *et al.* 2008) with a potential for subspecies rec-

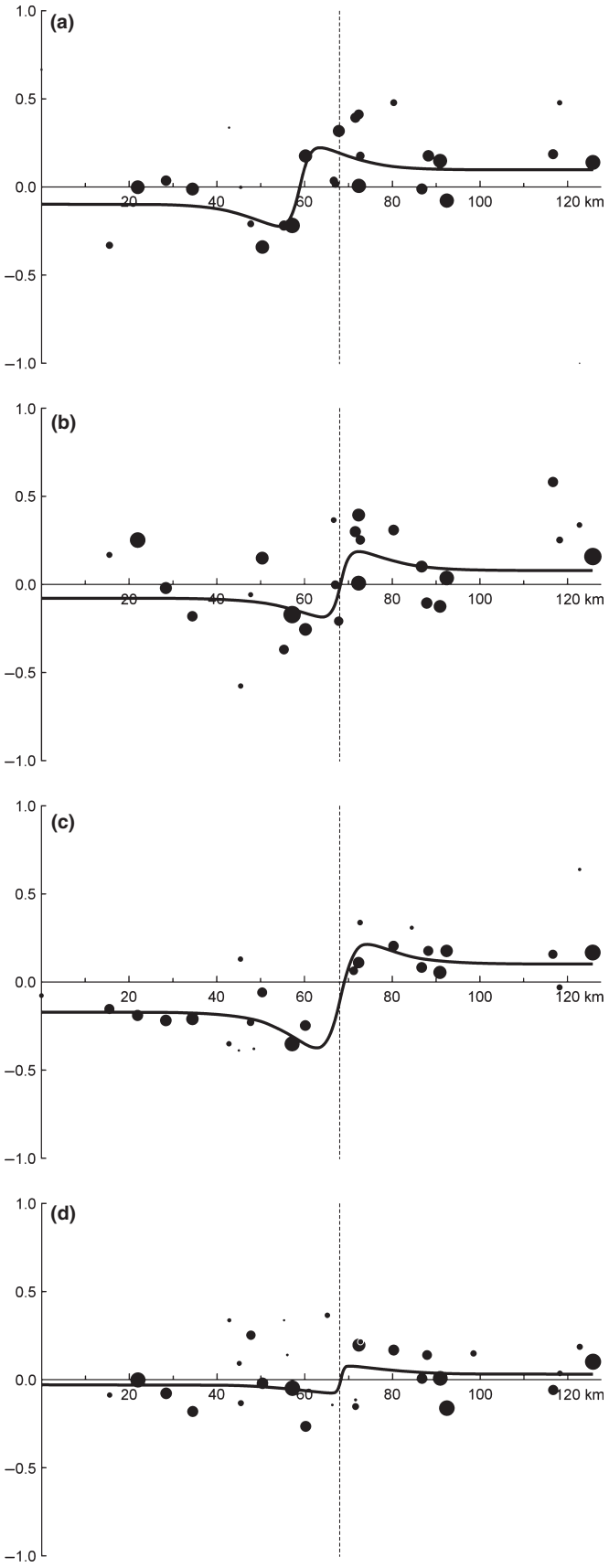

Fig. 4 Behavioural clines along the transect (abscissa) fitted with the reinforcement model separately for each sex and olfactory cue. (a) male, preference for either consubspecific female ABP; (b) female, preference for male ABP; (c) male, preference for female urine; (d) female, preference for male urine. Black bubbles represent estimates of Y_L for each locality; the area of each bubble is proportional to the weight for each sample (see text for explanation and Appendix for exact values): $Y_L < 0$ indicates preference for *domesticus* signals, $Y_L > 0$ preference for *musculus* signals, and $Y_L = 0$ mean null preference. Dashed vertical lines indicate approximate position of the consensus molecular cline centre.

ognition (Janotová & Stopka 2009) indicated by detectable differences in the native PAGE banding pattern (K. Janotová, personal communication). Nevertheless, direct evidence for this capacity is still lacking, and the results of molecular analyses presented in this study seem to contradict this hypothesis.

Even though MUPs constitute major protein components of mouse urine, there are also other involatile peptides in mouse scent detected with the vomeronasal organ (VNO) epithelium and coded for by multigene families: MHC peptides and exocrine-gland secreting peptides (ESPs) may play additional roles in sexual assessment (Cotton 2007; Hurst 2009), although Hurst (2009) has pointed out problems with the proposed role of MHC peptides. Recently, other urine constituents capable of firing VNO receptors, in particular sulfated steroids (Hsu *et al.* 2008; Nodari *et al.* 2008) and (methylthio) methanethiol or MTMT (Dulac & Wagner 2006) have been identified. Mouse chemical communication is thus likely to be mediated through a complex mixture of low-molecular-weight components with complementary roles, some of which are probably unknown, and hence, the results of urine-based olfactory preference tests should not be simply extrapolated to MUPs and vice versa.

Unlike the urine targets, the saliva targets used in this study are much more specific for the pheromonal signal they contain. The salivas we used were from mice congenic for the C3H strain background and selected to differ only in a 12 ± 8 cM region centred on the *a27* gene (Laukaitis *et al.* 1997). Therefore, the signal must be a salivary protein from a gene in that region. Putative pheromone genes that map to other chromosomes, such as *Mup*s (Chromosome 4) and *Esp*s (Chromosome 17), can be ruled out because they are identical in the congenic strains from which the saliva targets were collected. Furthermore, of all the genes mapping on Chromosome 7 in the 12 ± 8 cM region transferred onto the C3H background, only the *a27*, *bg26* and *bg27* gene products, originally described as the alpha, beta and gamma subunits secreted into saliva (Dlouhy & Karn 1983; Karn & Laukaitis 2003), were found in a recent saliva proteome study. In that work, triplicate saliva samples from both males and females were analysed by multi-dimensional protein identification technology (R. C. Karn and C. M. Laukaitis, unpublished data), and the findings are consistent with *Abp* gene expression studies that reported *a27*, *bg26* and *bg27* transcripts in the parotid, submandibular and submaxillary glands of house mice (Laukaitis *et al.* 2005; paralogs renumbered as per Laukaitis *et al.* 2008).

Regardless of the type of signal, males consistently displayed relatively strong preferences (even stronger than females in the case of urine), in agreement with previous studies (Bímová *et al.* 2005, 2009; Ganem *et al.* 2008). This seems to be at odds with the notion that females are the sex with higher reproductive costs and hence choosier than males (Darwin 1871; Fisher 1930; Trivers 1972; Andersson 1994). The weaker female preferences could be explained by variation in sexual receptivity as suggested by Ganem *et al.* (2008); however, we have not observed any differences in preference between receptive and nonreceptive females during our experiments (see also Bímová *et al.* 2005, 2009). An alternative hypothesis may point to differences in mating strategy between the sexes. Mice live in relatively closed demes so that females mate almost exclusively with a single dominant male during the most fertile period of the oestrous cycle and more than 70% of pups are thus sired by him (Bronson 1979; Drickamer *et al.* 2000; Dean *et al.* 2006). Females probably do not have enough opportunities to engage in mate choice, and thus, selection on assortative mating acts more strongly on males as the more dispersing sex (Gerlach 1996, 1998). On the other hand, absence of assortative female preference based on urine stimuli may not mean that females do not choose whatsoever. Rather, females may only engage in mate-quality recognition. The value of urine scents for species recognition is less useful, as the probability of meeting a hetero(sub)specific male may be low even in cases where the deme structure is disrupted. A factor favouring male choosiness is the need to save sperm, given the polygynous mating system of the species. The sperm-allocation strategy (Dewsbury 1982; Parker 1984) may be further strengthened from the danger of wasting valuable sperm in hetero(sub)specific matings. The slightly stronger female preferences based on salivary ABPs relative to males found in this study could then be ascribed to the need for assessing both individual quality and (sub)specific status of subordinate males. As leaving extensive scent marks would be detrimental for these 'sneaking' males, ABPs, being effective in close inter-individual contacts, could be an efficient cue for the females.

Is there reinforcement of behavioural isolation in the mouse hybrid zone?

It appears that when discussing reinforcement, two issues must be clearly separated: the presence of the action of reinforcement, and the likelihood that it could lead to speciation. While the circumstances where reinforcement alone is expected to produce biological species are very restricted (Howard 1993; Butlin 1987, 1995), recent theoretical models have relaxed these restrictions to some extent (e.g. Kirkpatrick & Ravigné 2002; Servedio 2004) and empirical studies (Hoskin *et al.* 2005; Smadja & Butlin 2006; see also Marshall

et al. 2002; Servedio & Noor 2003; Coyne & Orr 2004; for reviews) have provided some potential examples of its occurrence in nature. But even if reinforcement does not lead to speciation, we feel it is incorrect to assume that the presence of the reinforcement process itself is uninteresting.

Those who study barriers to gene flow are well aware that all types of barriers attract each other: tension zones move towards environmental barriers (whether abrupt dispersal blockades or simple density troughs) and clines at multiple loci involved with postzygotic barriers tend to coincide (Barton & Bengtsson 1986). Not only do all types of barrier attract each other, they interact to steepen each other, producing stepped clines (Barton & Bengtsson 1986; Barton & Gale 1993). It seems clear then that examining the potential action of reinforcement in increasing (prezygotic) barrier strength is an integral part of any understanding of a secondary contact. For us, the interest in exploring whether reinforcement is acting in the mouse hybrid zone is not about whether it will lead to speciation. Rather, we are interested in what proportion of the sharp and consistent barrier between mouse gene pools might come about through amplification of consubspecific preferences. We would argue that the interaction of barriers means any attempt to interpret, for example, postzygotic barrier strength independently of other potential barriers is flawed.

For both tested signals and sexes, models of behavioural clines including a reinforcement parameter showed significantly better fits than sigmoid cline models. The reinforcement parameter allows us to explore whether there is amplification of consubspecific preferences in areas where different genetic backgrounds meet. If hybrids are unfit, such amplification is predicted by reinforcement theory (Dobzhansky 1937; Howard 1993; Butlin 1995; Servedio & Noor 2003). We present direct evidence that amplification is a better explanation than no amplification when describing preferences in the central portion of our hybrid zone transect. Smadja & Ganem (2005, 2008) previously demonstrated that both urinary cues and associated preferences are more diverged in populations closer to the contact zone than in allopatric populations, suggesting a pattern similar to reproductive character displacement, and a similar observation has been made with respect to ABP (Bímová *et al.* 2005). This pattern is commonly considered as an evolutionary signature of reinforcement shaping premating isolation (Coyne & Orr 2004); however, it has been pointed out that other processes can give rise to similar observations and that reproductive character displacement may not be used as sole proof of reinforcement (Butlin 1995; Lemmon *et al.* 2004). For example, Albert & Schluter (2004) show that direct selection may be more effective than reinforcement in establishing reproductive barrier between limnetic and benthic sticklebacks, suggesting necessity of controlling for the effects of ecological character displacement and adaptation to different niches on mate preferences. However, we believe that the direct nature of our measurements of preference amplification *in situ* in the hybrid zone severely reduces the number of alternative processes than might explain the observations. Moreover, house mice are human commensals, and at least in Central Europe, there are no differences in ecology between the two mouse taxa.

Based on our results, we conclude that recognition between the two house mouse subspecies is a complex system involving several signals and associated preferences, where the latter are shaped by reinforcing selection. Our model of behavioural clines involving reinforcement provides an efficient tool for testing reinforcing selection in a unimodal hybrid zone, and our data provide the first direct proof of reinforcement acting on the barrier to gene flow between mammalian species. We thus conclude that behavioural barriers are an important component of a mosaic of reproductive isolation between house mouse taxa.

Acknowledgements

The immense mouse sampling would not be possible without many colleagues and local farmers, who are acknowledged for their continuous help in the field. We are grateful to E. Božíková (Charles University in Prague) who provided a part of the data on urinary preferences and to four anonymous referees for their comments on an earlier version of the manuscript. This work was supported with the Czech Science Foundation grant 206/08/0640 (to JP, MM, PM), the Grant Agency of the Academy of Sciences grant IAA600930506 (to JP), the grant of the Ministry of Education, Youth and Sports 0021620828 (to PM), the Grant Agency of the Charles University grant 167/2005 (to BVB) and US National Science Foundation grant DEB0746560 (to PKT). RCK was supported by a Senior Postdoctoral Fellowship (5F33HD055016-02) from the National Institute of Child Health and Human Development (NICHD). CML was supported by NIH/NCI University of Arizona SPORE in GI Cancer (P50 CA95060).

References

Albert AYK, Schluter D (2004) Reproductive character displacement of amle stickleback mate preference: reinforcement or direct selection? *Evolution*, **58**, 1099–1107.

Andersson M (1994) *Sexual Selection*. Princeton University Press, Princeton, NJ.

Baird SJE (1995) A simulation study of multilocus clines. *Evolution*, **49**, 1038–1045.

Barton NH, Baird SJE (1995) *Analyse – An Application for Analyzing Hybrid Zones*. Freeware, Edinburgh, UK.

Barton NH, Bengtsson BO (1986) The barrier to genetic exchange between hybridising populations. *Heredity*, **56**, 357–376.

Barton NH, Gale KS (1993) Genetic analysis of hybrid zones. In:*Hybrid Zones and the Evolutionary Process* (ed. Harrison RG), pp. 13–45, Oxford University Press, Oxford, UK.

Barton NH, Hewitt GM (1989) Adaptation, speciation and hybrid zones. *Nature*, **341**, 497–503.

Bazykin AD (1969) Hypothetical mechanism of speciation. *Evolution*, **23**, 685–687.

Beauchamp GK, Yamazaki K (2003) Chemical signalling in mice. *Biochemical Society Transactions*, **31**, 147–151.

Beynon RJ, Hurst JL (2003) Multiple roles of major urinary proteins in the house mouse. *Mus domesticus. Biochemical Society Transactions*, **31**, 142–146.

Beynon RJ, Hurst JL, Gaskell SJ *et al.* (2001) Mice, MUPs and myths – structure-function relationships of the major urinary proteins. In:*Chemical Signals in Vertebrates* (eds Marchlewska-Koj A, Muller-Schwarze D, Lepri J), pp. 149–156, Plenum Press, New York.

Beynon RJ, Veggerby C, Payne CE *et al.* (2002) Polymorphism in major urinary proteins: molecular heterogeneity in a wild mouse population. *Journal of Chemical Ecology*, **28**, 1429–1446.

Bímová B, Karn RC, Piálek J (2005) The role of salivary androgen-binding protein in reproductive isolation between two subspecies of house mouse: *Mus musculus musculus* and *Mus musculus domesticus*. *Biological Journal of the Linnean Society*, **84**, 349–361.

Bímová B, Albrecht T, Macholán M, Piálek J (2009) Signalling components of mate recognition system in the house mouse. *Behavioural Processes*, **80**, 20–27.

Bishop JO, Clark AJ, Clissold PM, Hainey S, Francke U (1982) Two main groups of mouse major urinary protein genes, both largely located on chromosome 4. *EMBO Journal*, **1**, 615–620.

Blum MJ (2008) Ecological and genetic associations across a *Heliconius* hybrid zone. *Journal of Evolutionary Biology*, **21**, 330–341.

Boursot P, Auffray J-C, Britton-Davidian J, Bonhomme F (1993) The evolution of house mice. *Annual Review of Ecology and Systematics*, **24**, 119–152.

Brennan PA, Kendrick KM (2006) Mammalian social odours: attraction and individual recognition. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **361**, 2061–2078.

Bridle JR, Butlin RK (2002) Mating signal variation and bimodality in a mosaic hybrid zone between *Chorthippus* grasshopper species. *Evolution*, **56**, 1184–1198.

Bridle JR, Baird SJE, Butlin RK (2001) Spatial structure and habitat variation in a grasshopper hybrid zone. *Evolution*, **55**, 1832–1843.

Britton-Davidian J, Fel-Clair F, Lopez J, Alibert P, Boursot P (2005) Postzygotic isolation between the two European subspecies of the house mouse: estimates from fertility patterns in wild and laboratory-bred hybrids. *Biological Journal of the Linnean Society*, **84**, 379–393.

Bronson FH (1979) The reproductive ecology of the house mouse. *The Quarterly Review of Biology*, **54**, 265–299.

Bullough RK, Caudrey PJ (eds) (1980) *Solitons*. Springer-Verlag, Berlin.

Butlin RK (1987) Speciation by reinforcement. *Trends in Ecology and Evolution*, **2**, 8–13.

Butlin RK (1995) Reinforcement: an idea evolving. *Trends in Ecology and Evolution*, **10**, 432–434.

Butlin RK, Hewitt GM (1985a) A hybrid zone between *Chorthippus parallelus parallelus* and *Chorthippus parallelus erythropus* (Orthoptera: Acrididae): behavioural characters. *Biological Journal of the Linnean Society*, **26**, 287–299.

Butlin RK, Hewitt GM (1985b) A hybrid zone between *Chorthippus parallelus parullelus* and *Chorthippus parallelus erythropus* (Orthoptera: Acrididae): morphological and electrophoretic characters. *Biological Journal of the Linnean Society*, **26**, 269–285.

Butlin RK, Ritchie MG (1991) Variation in female mate preference across a grasshopper hybrid zone. *Journal of Evolutionary Biology*, **4**, 227–240.

Cheetham SA, Thom MD, Jury F, Ollier WER, Beynon RJ, Hurst JL (2007) The genetic basis of individual-recognition signals in the mouse. *Current Biology*, **17**, 1771–1777.

Christophe N, Baudoin C (1998) Olfactory preferences in two subspecies of mice *Mus musculus musculus* and *Mus musculus domesticus* and their hybrids. *Animal Behaviour*, **56**, 365–369.

Clissold PM, Bishop JO (1982) Variation in mouse major urinary protein (MUP) genes and the MUP gene products within and between inbred lines. *Gene*, **18**, 211–220.

Cotton S (2007) Individual recognition: mice, MUPs and the MHC. *Current Biology*, **17**, R971–R973.

Coyne JA, Orr HA (2004) *Speciation*. Sinauer Associates, Inc. Sunderland, Massachusetts.

Darwin C (1871) *The Descent of Man, and Selection in Relation to Sex*. John Murray, London.

Dean MD, Ardlie KG, Nachman MW (2006) The frequency of multiple paternity suggests that sperm competition is common in house mice (*Mus domesticus*). *Molecular Ecology*, **15**, 4141–4151.

Dewsbury DA (1982) Ejaculate cost and male choice. *The American Naturalist*, **119**, 601–610.

Dlouhy SR, Karn RC (1983) The tissue source and cellular control of the apparent size of androgen binding protein (Abp), a mouse salivary protein whose electrophoretic mobility is under the control of sex-limited saliva pattern (Ssp). *Biochemical Genetics*, **21**, 1057–1070.

Dlouhy SR, Taylor BA, Karn RC (1987) The genes for mouse salivary androgen-binding protein (ABP) subunits alpha and gamma are located on chromosome 7. *Genetics*, **115**, 535–543.

Dobzhansky T (1937) *Genetics and the Origin of Species*. Columbia University Press, New York.

Dobzhansky T (1940) Speciation as a stage in evolutionary divergence. *The American Naturalist*, **74**, 312–321.

Dod B, Jermiin LS, Boursot P, Chapman VH, Nielsen JT, Bonhomme F (1993) Counterselection on sex chromosomes in the *Mus musculus* European hybrid zone. *Journal of Evolutionary Biology*, **6**, 529–546.

Dod B, Smadja C, Karn RC, Boursot P (2005) Testing for selection on the androgen-binding protein in the Danish mouse hybrid zone. *Biological Journal of the Linnean Society*, **84**, 447–459.

Drickamer LC, Gowaty PA, Holmes CM (2000) Free female mate choice in house mice affects reproductive success and

offspring viability and performance. *Animal Behaviour*, **59**, 371–378.
Dufková P, Macholán M, Piálek J (2011) Inference of selection and stochastic effects in the house mouse hybrid zone. *Evolution*, **65**, 993–1010.
Dulac C, Wagner S (2006) Genetic analysis of brain circuits underlying pheromone signalling. *Annual Review of Genetics*, **40**, 449–467.
Edwards AWF (1992) *Likelihood*. Cambridge University Press, Cambridge, Massachusetts.
Emes RD, Riley MC, Laukaitis CM, Goodstadt L, Karn RC, Ponting CP (2004) Comparative evolutionary genomics of androgen-binding protein genes. *Genome Research*, **14**, 1516–1529.
Fel-Clair F, Lenormand T, Catalan J *et al.* (1996) Genomic incompatibilities in the hybrid zone between house mice in Denmark: evidence from steep and non-coincident chromosomal clines for Robertsonian fusions. *Genetical Research*, **67**, 123–134.
Fisher RA (1930) *The Genetical Theory of Natural Selection. A Complete Variorum Edition*. Oxford University Press, Inc., New York.
Forejt J (1996) Hybrid sterility in the mouse. *Trends in Genetics*, **12**, 412–417.
Forejt J, Iványi P (1974) Genetic studies on male sterility of hybrids between laboratory and wild mice (*Mus musculus* L.). *Genetical Research*, **24**, 189–206.
Ganem G, Litel C, Lenormand T (2008) Variation in mate preference across a house mouse hybrid zone. *Heredity*, **100**, 594–601.
Gay L, Crochet P-A, Bell DA, Lenormand T (2008) Comparing clines on molecular and phenotypic traits in hybrid zones: a window on tension zone models. *Evolution*, **62**, 2789–2806.
Gerlach G (1996) Emigration mechanisms in feral house mice – a laboratory investigation of the influence of social structure, population density, and aggression. *Behavioral Ecology and Sociobiology*, **39**, 159–170.
Gerlach G (1998) Impact of social ties on dispersal, reproduction and dominance in feral house mice (*Mus musculus domesticus*). *Ethology*, **104**, 487–499.
Good JM, Handel MA, Nachman MW (2008) Asymmetry and polymorphism of hybrid male sterility during the early stages of speciation in house mice. *Evolution*, **62**, 50–65.
Haldane JBS (1948) The theory of a cline. *Journal of Genetics*, **48**, 277–284.
Harr B (2006) Genomic islands of differentiation between house mouse subspecies. *Genome Research*, **16**, 730–737.
Harris H, Hopkinson DA (1976) *Handbook of Enzyme Electrophoresis in Human Genetics*. North-Holland Publishing Company, Oxford, UK.
Hoskin CJ, Higgie M, McDonald KR, Moritz C (2005) Reinforcement drives rapid allopatric speciation. *Nature*, **437**, 1353–1356.
Howard DJ (1993) Reinforcement: the origin, dynamics, and fate of an evolutionary hypothesis. In:*Hybrid Zones and the Evolutionary Process* (ed. Harrison RG), pp. 46–69, Oxford University Press, New York.
Hsu F-F, Nodari F, Kao L-F *et al.* (2008) Structural characterization of sulfated steroids that activate mouse pheromone-sensing neurons. *Biochemistry*, **47**, 14009–14019.
Hurst JL (2009) Female recognition and assessment of males through scent. *Behavioural Brain Research*, **200**, 295–303.
Hurst JL, Beynon RJ (2004) Scent wars: the chemobiology of competitive signalling in mice. *Bioessays*, **26**, 1288–1298.
Janotová K, Stopka P (2009) Mechanisms of chemical communication: the role of Major Urinary Proteins. *Folia Zoologica*, **58**, 43–59.
Karn RC (1981) Sex-limited effects of the expression of the db gene in mice during puberty. *Biochemical Genetics*, **19**, 355–371.
Karn RC, Dlouhy SR (1991) Salivary androgen-binding protein variation in *Mus* and other rodents. *Journal of Heredity*, **82**, 453–458.
Karn RC, Laukaitis CM (2003) Characterization of two forms of mouse salivary Androgen-Binding Protein (ABP): implications for evolutionary relationships and ligand-binding function. *Biochemistry*, **42**, 7162–7170.
Karn RC, Laukaitis CM (2009) The mechanism of expansion and the volatility it created in three pheromone gene clusters in the mouse (*Mus musculus*) genome. *Genome Biology and Evolution*, **1**, 494–503.
Karn RC, Nachman MW (1999) Reduced nucleotide variability at an Androgen-Binding Protein locus (*Abpa*) in house mice: evidence for positive natural selection. *Molecular Biology and Evolution*, **16**, 1192–1197.
Karn RC, Orth A, Bonhomme F, Boursot P (2002) The complex history of a gene proposed to participate in a sexual isolation mechanism in house mice. *Molecular Biology and Evolution*, **19**, 462–471.
Key KH (1968) The concept of stasipatric speciation. *Systematic Zoology*, **17**, 14–22.
Kirkpatrick M, Ravigné V (2002) Speciation by natural and sexual selection: models and experiments. *The American Naturalist*, **159**, S22–S35.
Kruuk LEB, Baird SJE, Gale KS, Barton NH (1999) A comparison of multilocus clines maintained by environmental adaptation or by selection against hybrids. *Genetics*, **153**, 1959–1971.
Lakshmanan M (ed.) (1988) *Solitons*. Springer-Verlag, New York.
Lande R (1981) Models of speciation by sexual selection on polygenic traits. *Proceedings of the National Academy of Sciences, USA*, **78**, 3721–3725.
Laukaitis CM, Critser ES, Karn RC (1997) Salivary androgen-binding protein (ABP) mediates sexual isolation in *Mus musculus*. *Evolution*, **51**, 2000–2005.
Laukaitis CM, Dlouhy SR, Emes RD, Ponting CP, Karn RC (2005) Diverse spatial, temporal, and sexual expression of recently duplicated androgen-binding protein genes in *Mus musculus*. *BMC Evolutionary Biology*, **5**, 40.
Laukaitis CM, Heger A, Blakley TD, Munclinger P, Ponting CP, Karn RC (2008) Rapid bursts of androgen-binding protein (Abp) gene duplication occurred independently in diverse mammals. *BMC Evolutionary Biology*, **8**, 46.
Lemmon AR, Smadja C, Kirkpatrick M (2004) Reproductive character displacement is not the only possible outcome of reinforcement. *Journal of Evolutionary Biology*, **17**, 177–183.
Logan DW, Marton TF, Stowers L (2008) Species specificity in major urinary proteins by parallel evolution. *PLoS ONE*, **3**, e3280.

Luo MM, Fee MS, Katz LC (2003) Encoding pheromonal signals in the accessory olfactory bulb of behaving mice. *Science*, **299**, 1196–1201.

Macholán M, Kryštufek B, Vohralík V (2003) The location of the *Mus musculus/M. domesticus* hybrid zone in the Balkans: clues from morphology. *Acta Theriologica*, **48**, 177–188.

Macholán M, Munclinger P, Šugerková M *et al.* (2007) Genetic analysis of autosomal and X-linked markers across a mouse hybrid zone. *Evolution*, **61**, 746–771.

Macholán M, Baird SJE, Munclinger P, Dufková P, Bímová B, Piálek J (2008) Genetic conflict outweighs heterogametic incompatibility in the mouse hybrid zone? *BMC Evolutionary Biology*, **8**, 271–284.

Macholán M, Baird SJE, Dufková P, Munclinger P, Vošlajerová Bímová B, Piálek J (2011) Assessing multilocus introgression patterns: a case study on the mouse X chromosome in Central Europe. *Evolution*, Early View, DOI: 10.1111/j.1558-5646.2011.01228.x.

Marshall JL, Arnold ML, Howard DJ (2002) Reinforcement: the road not taken. *Trends in Ecology and Evolution*, **17**, 558–563.

Mayr E (1942) *Systematics and the Origin of Species*. Columbia University Press, New York.

Mihola O, Trachtulec Z, Vlček C, Schimenti JC, Forejt J (2009) A mouse speciation gene encodes a meiotic histone H3 methyltransferase. *Science*, **323**, 373–375.

Mudge JM, Armstrong SD, McLaren K *et al.* (2008) Dynamic instability of the major urinary protein gene family revealed by genomic and phenotypic comparisons between C57 and 129 strain mice. *Genome Biology*, **9**, R91.

Munclinger P, Božíková E, Šugerková M, Piálek J, Macholán M (2002) Genetic variation in house mice (*Mus*, Muridae, Rodentia) from the Czech and Slovak Republics. *Folia Zoologica*, **51**, 81–92.

Munclinger P, Boursot P, Dod B (2003) B1 insertions as easy markers for mouse population studies. *Mammalian Genome*, **14**, 359–366.

Nagylaki T (1975) Conditions for the existence of clines. *Genetics*, **80**, 595–615.

Nodari F, Hsu F-F, Fu X *et al.* (2008) Sulfated steroids as natural ligands of mouse pheromone-sensing neurons. *The Journal of Neuroscience*, **18**, 6407–6418.

Noldus LPJJ, Trienes RJH, Hendriksen AHM, Jansen H, Jansen RG (2000) The Observer Video-Pro: new software for the collection, management, and presentation of time-structured data from videotapes and digital media files. *Behavior Research Methods, Instruments, and Computers*, **32**, 197–206.

Nürnberger B, Barton N, MacCallum C, Gilchrist J, Appleby M (1995) Natural selection on quantitative traits in the *Bombina* hybrid zone. *Evolution*, **49**, 1224–1238.

Oka A, Mita A, Sakurai-Yamatani N *et al.* (2004) Hybrid breakdown caused by substitution of the X chromosome between two mouse subspecies. *Genetics*, **166**, 913–924.

Panhuis TM, Butlin RK, Zuk M, Tregenza T (2001) Sexual selection and speciation. *Trends in Ecology and Evolution*, **16**, 364–371.

Parker GA (1984) Sperm competition and the evolution of animal mating strategies. In:*Sperm Competition and the Evolution of Animal Mating Systems* (ed. Smith RL), pp. 2–60, Academic Press, London.

Pasteur N, Pasteur G, Bonhomme F, Catalan J, Britton-Davidian J (1988) *Practical Isozyme Genetics*. Ellis Horwood, Ltd., Chichester, UK.

Payseur BA, Krenz JG, Nachman MW (2004) Differential patterns of introgression across the X chromosome in a hybrid zone between two species of house mice. *Evolution*, **58**, 2064–2078.

Phillips BL, Baird SJE, Moritz C (2004) When vicars meet: a narrow contact zone between morphologically cryptic phylogeographic lineages of the rainforest skink, *Carlia rubrigularis*. *Evolution*, **58**, 1536–1548.

Piálek J, Vyskočilová M, Bímová B *et al.* (2008) Development of unique house mouse resources suitable for evolutionary studies of speciation. *Journal of Heredity*, **99**, 34–44.

Ptacek MB (2000) The role of mating preferences in shaping interspecific divergence in mating signals in vertebrates. *Behavioural Processes*, **51**, 111–134.

Ptacek MB (2002) Patterns of inheritance of mating signals in interspecific hybrids between sailfin and shortfin mollies (Poeciliidae: *Poecilia: Mollienesia*). *Genetica*, **116**, 329–342.

Raufaste N, Orth A, Belkhir K *et al.* (2005) Inference of selection and migration in the Danish house mouse hybrid zone. *Biological Journal of the Linnean Society*, **84**, 593–616.

Ritchie MG (2007) Sexual selection and speciation. *Annual Review of Ecology, Evolution, and Systematics*, **38**, 79–102.

Robertson DHL, Hurst JL, Searle JB, Gündüz I, Beynon RJ (2007) Characterization and comparison of major urinary proteins from the house mouse, *Mus musculus domesticus*, and the aboriginal mouse, *Mus macedonicus*. *Journal of Chemical Ecology*, **33**, 613–630.

Ryan MJ, Rand AS (1993) Species recognition and sexual selection as a unitary problem in animal communication. *Evolution*, **47**, 647–657.

Sage RD, Atchley WR, Capanna E (1993) House mice as models in systematic biology. *Systematic Biology*, **42**, 523–561.

Servedio MR (2004) The what and why of research on reinforcement. *PLoS Biology*, **2**, e420.

Servedio MR, Noor MAF (2003) The role of reinforcement in speciation: theory and data. *Annual Review of Ecology and Systematics*, **34**, 339–364.

Slatkin M (1975) Gene flow and selection on a two-locus system. *Genetics*, **81**, 787–802.

Smadja C, Butlin RK (2006) Speciation: a new role for reinforcement. *Heredity*, **96**, 422–423.

Smadja C, Ganem G (2002) Subspecies recognition in the house mouse: a study of two populations from the border of a hybrid zone. *Behavioral Ecology*, **13**, 312–320.

Smadja C, Ganem G (2005) Asymmetrical reproductive character displacement in the house mouse. *Journal of Evolutionary Biology*, **18**, 1485–1493.

Smadja C, Ganem G (2008) Divergence of odorant signals within and between the two European subspecies of the house mouse. *Behavioral Ecology*, **19**, 223–230.

Smadja C, Catalan J, Ganem G (2004) Strong premating divergence in a unimodal hybrid zone between two subspecies in the house mouse. *Journal of Evolutionary Biology*, **17**, 165–176.

Stopka P, Janotová K, Heyrovský D (2007) The advertisement role of major urinary proteins in mice. *Physiology and Behavior*, **91**, 667–670.

Stopková R, Stopka P, Janotová K, Jedelský P (2007) Species-specific expression of major urinary proteins in the house mice (*Mus musculus musculus* and *Mus musculus domesticus*). *Journal of Chemical Ecology*, **33**, 861–869.

Storchová R, Gregorová S, Buckiová D, Kyselová D, Divina P, Forejt J (2004) Genetic analysis of X-linked hybrid sterility in the house mouse. *Mammalian Genome*, **15**, 515–524.

Szymura JM, Barton NH (1986) Genetic analysis of a hybrid zone between the fire-bellied toads, *Bombina bombina* and *B. variegata*, near Cracow in southern Poland. *Evolution*, **40**, 1141–1159.

Szymura JM, Barton NH (1991) The genetic structure of the hybrid zone between the firebellied toads *Bombina bombina* and *B. variegata*: comparisons between transects and between loci. *Evolution*, **45**, 237–261.

Takami Y, Suzuki H (2005) Morphological, genetic and behavioural analyses of a hybrid zone between the ground beetles *Carabus lewisianus* and *C. albrechti* (Coleoptera, Carabidae): asymmetrical introgression caused by movement of the zone? *Biological Journal of the Linnean Society*, **86**, 79–94.

Talley HM, Laukaitis CM, Karn RC (2001) Female preference for male saliva: implications for sexual isolation of *Mus musculus* subspecies. *Evolution*, **55**, 631–634.

Teeter KC, Payseur BA, Harris LW *et al.* (2008) Genome-wide patterns of gene flow across a house mouse hybrid zone. *Genetical Research*, **18**, 67–76.

Teeter KC, Thibodeau LM, Gompert Z, Buerkle CA, Nachman MW, Tucker PK (2010) The variable genomic architecture of isolation between hybridizing species of house mice. *Evolution*, **64**, 472–485.

Thom MD, Stockley P, Jury F, Ollier WER, Beynon RJ, Hurst JL (2008) The direct assessment of genetic heterozygosity through scent in the mouse. *Current Biology*, **18**, 619–623.

Todrank J, Busquet N, Baudoin C, Heth G (2005) Preferences of newborn mice for odours indicating closer genetic relatedness: is experience necessary? *Proceedings of the Royal Society B: Biological Sciences*, **272**, 2083–2088.

Tregenza T, Pritchard VL, Butlin RK (2000) The origins of premating reproductive isolation: testing hypotheses in the grasshopper *Chorthippus parallelus*. *Evolution*, **54**, 1687–1698.

Trivers RL (1972) Parental investment and sexual selection. In:*Sexual Selection and the Descent of Man, 1871–1971* (ed. Campbell B), pp. 136–179, Heinemann, London.

Tucker PK, Sage RD, Warner J, Wilson AC, Eicher EM (1992) Abrupt cline for sex chromosomes in a hybrid zone between two species of mice. *Evolution*, **46**, 1146–1163.

Turelli M, Barton NH, Coyne JA (2001) Theory and speciation. *Trends in Ecology and Evolution*, **16**, 330–343.

Vanlerberghe F, Dod B, Boursot P, Bellis M, Bonhomme M (1986) Absence of Y-chromosome introgression across the hybrid zone between *Mus musculus domesticus* and *Mus musculus musculus*. *Genetical Research*, **48**, 191–197.

Vyskočilová M, Trachtulec Z, Forejt J, Piálek J (2005) Does geography matter in hybrid sterility in house mice? *Biological Journal of the Linnean Society*, **84**, 663–674.

Vyskočilová M, Pražanová G, Piálek J (2009) Polymorphism in hybrid male sterility in wild-derived Mus musculus musculus strains on proximal chromosome 17. *Mammalian Genome*, **20**, 83–91.

Wells MM, Henry CS (1998) Songs, reproductive isolation and speciation in cryptic species of insects: a case study using green lacewings. In:*Endless Forms: Species and Speciation* (ed. Howard E), pp. 217–233, Oxford University Press, New York.

West-Eberhard M (1983) Sexual selection, social competition, and speciation. *The Quarterly Review of Biology*, **58**, 155–183.

Wolfram S (1992) *Mathematica: A System for Doing Mathematics by Computer*. Addison-Wesley Publishing Company, Inc., Redwood City.

Yang H, Ding Y, Hutchins LM *et al.* (2009) A customized and versatile high-density array for the mouse. *Nature Methods*, **6**, 663–666.

B.V.B., M.M., S.J.E.B., P.M., P.D., and J.P. are involved in a long-term survey of the central-European portion of the house mouse hybrid zone; B.V.B. is engaged in behavioural genetics; M.M. is interested in population genetics, systematics and evolution of the genus *Mus* and other small mammals; S.J.E.B. in modelling outcomes of admixture on secondary contacts and in extending the tools of population genetics to allow sound inference about the evolutionary processes that give rise to data collected in the field; P.M. in population genetics and evolution of mammals and birds; and J.P. in population genetics, hybridisation and speciation. R.C.K. and C.M.L. are engaged in the evolution and genomics of secrotoglobins and their role in animal communication; K.L. in population genetics, hybrid zone introgression, banana agroecology and bat interactions; and P.K.T. in molecular evolution and systematics of mammals, and in speciation. P.D. is currently working on borreliosis and transmission of Borrelia from ticks to mice and from mice to ticks.

Data accessibility

Data deposited at Dryad: http://dx.doi.org/10.5061/dryad.8807.

Appendix

The time spent sniffing one of the cues ($\sum T_{mus}$, $\sum T_{dom}$) in the preference tests assessed for each population, sex, and signal. Numbers of tested individuals with results of the Wilcoxon matched pair test ($\mu_{\mathrm{Tmus}} = \mu_{\mathrm{Tdom}}$; Z and T statistics, P-value) are given. Populations are ordered along the transect from the westernmost site (*M. m. domesticus* territory) to the *M. m. musculus* territory, with the zone centre indicated by the dashed line. Significant values are given in bold. Alpha values ($\alpha = 0.05$) were adjusted using Bonferroni correction to $\alpha = 0.00046$.

			ABP						Urine					
Map No	Code	Sex	*N*	Σ*Tmus*	Σ*Tdom*	*T*	*Z*	*P*	*N*	Σ*Tmus*	Σ*Tdom*	*T*	*Z*	*P*
1	KUBL	F	–	–	–	–	–	–	–	–	–	–	–	–
		M	1	5	1	0.0	–	–	2	17	20	0.0	–	–
2	WEI1	F	4	14	10	3.5	0.548	0.584	3	11	13	0.0	–	–
		M	5	14	28	6.0	0.405	0.686	8	80	109	8.0	1.400	0.161
3–4	STR1–2	F	41	185	110	267.0	1.501	0.133	39	171	172	262.5	0.322	0.748
		M	27	114	115	139.5	0.914	0.361	16	104	152	14.5	2.386	0.017
6	BENK	F	13	61	64	32.0	0.549	0.583	11	65	76	14.5	0.948	0.343
		M	12	53	49	22.5	0.510	0.610	12	110	171	7.5	2.267	0.023
8	LEHS	F	10	34	49	15.0	0.889	0.374	11	50	72	21.5	1.022	0.307
		M	12	93	95	27.0	0.051	0.959	16	139	211	1.5	3.323	0.001
12	HEBA	F	–	–	–	–	–	–	1	10	5	0.0	–	–
		M	1	4	2	0,0	–	–	2	16	33	0.0	–	–
15	PLOS	F	–	–	–	–	–	–	2	12	10	0.0	–	–
		M	–	–	–	–	–	–	1	4	9	0.0	–	–
16	UNWE	F	2	4	15	0.0	–	–	2	13	17	0.0	–	–
		M	2	5	5	0.0	–	–	3	27	21	2.0	0.535	0.593
19	THIE	F	3	8	9	2.0	0.535	0.593	10	50	30	14.0	1.007	0.314
		M	5	15	23	4.5	0.809	0.418	5	36	57	0.0	1.826	0.068
20	HOCH	F	–	–	–	–	–	–	–	–	–	–	–	–
		M	–	–	–	–	–	–	1	5	11	0.0	–	–
22–23	NEUE–8	F	23	88	65	114.0	0.052	0.958	8	67	70	7.0	0.135	0.893
		M	21	73	148	46.0	2.203	0.028	6	86	96	7.0	0.734	0.463
26	LIB2	F	11	22	48	9.5	1.835	0.067	1	4	2	0.0	–	–
		M	9	38	59	9.0	0.845	0.398	–	–	–	–	–	–
30	HAM2	F	–	–	–	–	–	–	1	4	3	0.0	–	–
		M	–	–	–	–	–	–	–	–	–	–	–	–
34	HUR1	F	21	191	271	76.5	0.392	0.695	23	170	189	82.5	1.147	0.251
		M	15	136	211	41.0	0.314	0.753	25	214	442	8.5	4.144	0.000
40	LUZN	F	22	55	93	77.5	1.027	0.305	7	44	76	1.0	2.197	0.028
		M	14	118	82	27.0	1.601	0.109	8	88	145	3.5	1.775	0.076
53	DLMO	F	–	–	–	–	–	–	1	17	8	0.0	–	–
		M	–	–	–	–	–	–	–	–	–	–	–	–
56	SVKR	F	–	–	–	–	–	–	1	3	4	0.0	–	–
		M	–	–	–	–	–	–	–	–	–	–	–	–
59	DOU3	F	7	15	7	5.0	1.521	0.128	–	–	–	–	–	–
		M	12	30	28	25.5	0.204	0.838	–	–	–	–	–	–
62	JIND	F	7	27	27	11.0	0.507	0.612	–	–	–	–	–	–
		M	6	25	24	7.5	0.000	1.000	–	–	–	–	–	–
65	MIL1	F	8	25	38	12.5	0.770	0.441	–	–	–	–	–	–
		M	9	101	52	1.5	2.488	0.013	–	–	–	–	–	–
71–72	NEB2–3	F	8	66	36	6.0	1.352	0.176	2	17	23	0.0		
		M	8	67	29	6.0	1.680	0.093	4	71	62	1.0	1.069	0.285
73	KRA4	F	–	–	–	–	–	–	2	4	5	0.0		
		M	–	–	–	–	–	–	–	–	–	–	–	–
79	KAC2	F	20	126	124	67.0	0.052	0.959	12	119	80	1.0	2.845	0.004
		M	17	134	133	40.0	1.448	0.148	11	154	124	16.0	1.172	0.241
80	OBIL	F	12	107	47	18.5	1.289	0.197	–	–	–	–	–	–
		M	12	62	26	17.5	1.019	0.308	–	–	–	–	–	–
81	MOST	F	11	40	24	8.0	1.014	0.310	2	23	15	0.0	–	–
		M	6	36	25	4.5	0.809	0.418	1	36	18	0.0	–	–
96–98	RUD1–2	F	11	59	31	13.0	1.478	0.139	10	68	49	4.0	1.960	0.050
		M	8	28	10	7.0	1.540	0.123	9	131	86	8.5	1.659	0.097
105	SSED	F	–	–	–	–	–	–	–	–	–	–	–	–
		M	–	–	–	–	–	–	1	19	10	0.0	–	–

Appendix (Continued).

			ABP						Urine					
Map No	Code	Sex	*N*	Σ*Tmus*	Σ*Tdom*	*T*	*Z*	*P*	*N*	Σ*Tmus*	Σ*Tdom*	*T*	*Z*	*P*
106	DEPO	F	8	67	55	14.0	0.560	0.575	10	56	55	25.0	0.255	0.799
		M	9	58	59	13.5	0.085	0.933	9	123	105	10.5	1.422	0.155
107–108	POC1–2	F	13	46	57	35.0	0.314	0.754	8	50	38	10.5	1.050	0.294
		M	12	78	55	25.5	1.059	0.290	9	112	78	4.5	2.132	0.033
109	HOSL	F	11	66	85	16.5	0.711	0.477	31	155	154	204.5	0.281	0.779
		M	8	144	107	13.0	0.700	0.484	13	220	199	23.5	0.845	0.398
111–113	NVES	F	30	120	112	169.5	0.152	0.879	26	160	220	31.5	3.386	0.001
		M	20	124	144	42.0	1.022	0.307	13	209	145	18.0	1.922	0.055
114	SEDL	F	–	–	–	–	–	–	3	19	14	1.5	0.802	0.423
		M	–	–	–	–	–	–	–	–	–	–	–	–
120–121	TYN1–2	F	8	65	17	6.5	1.268	0.205	8	47	53	14.0	0.560	0.575
		M	7	55	38	9.0	0.845	0.398	7	87	63	5.0	1.521	0.128
122	PROH	F	6	20	12	4.0	0.944	0.345	3	14	13	0.0	–	–
		M	3	17	6	0.0	1.604	0.109	3	29	31	0.0	–	–
123	VRBI	F	5	14	7	1.5	1.618	0.106	3	19	13	1.0	1.069	0.285
		M	1	0	2	0.0	–	–	1	18	4	0.0	–	–
124–128	BUSK	F	35	263	191	192.0	1.581	0.114	32	233	190	191.0	1.365	0.172
		M	33	175	133	142.0	1.861	0.063	31	467	334	116.5	2.386	0.017

Supporting information

Additional supporting information may be found in the online version of this article.

Table S1 Frequencies of *M. m. musculus* alleles for each locus and sampling site. Distance along the transect and effective number of alleles (N_e) are also given; h = hybrid index expressed as the frequency of *musculus* alleles averaged across the six diagnostic allozyme loci (see Macholán *et al.* 2007 and text for details).

Table S2 A panel of allopatric populations and wild-derived inbred strains representing both taxa used for usability of each locus as a diagnostic marker (see Piálek *et al.* 2008 for more details on the inbred strains). Geographic coordinates are given for each site along with numbers of analysed individuals and frequencies of *M. m. musculus* alleles for marker *4.057*.

Table S3 List of molecular markers used in this study. For each locus, the marker type, chromosome, position according to the NCBI m37 mouse assembly, and when appropriate, primer sequences and annealing temperature is given.

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