

ORIGINAL ARTICLE

Origin of the chromosomal radiation of Madeiran house mice: a microsatellite analysis of metacentric chromosomes

DW Förster^{1,2}, ML Mathias³, J Britton-Davidian⁴ and JB Searle^{2,5}

Chromosome races of *Mus musculus domesticus* are characterised by particular sets of metacentric chromosomes formed by Robertsonian fusions and whole-arm reciprocal translocations. The Atlantic island of Madeira is inhabited by six chromosome races of house mice with 6–9 pairs of metacentric chromosomes. Three of these races are characterised by the metacentric 3.8 also found elsewhere in the distribution of *M. m. domesticus*, including Denmark and Spain. We investigated the possibility that metacentric 3.8 was introduced to Madeira during the initial colonisation, as this could have ‘seeded’ the cascade of chromosomal mutation that is the basis of the extraordinary chromosomal radiation observed on the island. Variation at 24 microsatellite loci mapping to three different chromosomal regions (proximal, interstitial and distal) of mouse chromosomes 3 and 8 was investigated in 179 mice from Madeira, Denmark, Portugal, Spain, Italy and Scotland. Analyses of microsatellite loci closely linked to the centromeres of these chromosomes (‘proximal loci’) do not support a common evolutionary origin of metacentric 3.8 among Madeiran, Danish and Spanish mouse populations. Our results suggest that Madeiran mice are genetically more similar to standard karyotype mice from Portugal than to metacentric mice from elsewhere. There is expected to be an interruption to gene flow between hybridising metacentric races on Madeira, particularly in the chromosomal regions close to the rearrangement breakpoints. Consistent with this, relating to differentiation involving chromosomes 3 and 8 on Madeira, we found greater genetic structure among races for proximal than interstitial or distal loci. *Heredity* (2013) **110**, 380–388; doi:10.1038/hdy.2012.107; published online 12 December 2012

Keywords: chromosomal rearrangements; metacentric; house mouse; island colonisation; *Mus musculus domesticus*; microsatellite

INTRODUCTION

The Robertsonian (Rb) fusion chromosomal rearrangement (the joining of two acrocentrics to form a metacentric) is commonly involved in karyotypic differences between and within animal species (King, 1993; Searle, 1993). Given the potential of chromosomal rearrangements such as this to cause an interruption of gene flow, and therefore speciation (King, 1993; Butlin, 2005; Faria and Navarro, 2010), it is important to investigate their origin and impact in natural systems. The western house mouse (*Mus musculus domesticus*) is very variable chromosomally (Piálek *et al.*, 2005): while the standard karyotype of the house mouse consists of 40 acrocentric chromosomes, there are chromosomal forms with different combinations of autosomal metacentrics, which represent the products of Rb fusions or derivatives of these (that is, whole-arm reciprocal translocations: the exchange of chromosome arms between metacentrics or between metacentrics and acrocentrics). Where metacentrics are fixed within a population, they define a metacentric race. Because each metacentric involves a reduction in chromosome number, the diploid number that characterises such races can be as low as $2n = 22$ (nine pairs of metacentrics).

Here we focus on the notable chromosomal radiation in the house mice on the island of Madeira (Figure 1), where there are six metacentric races with diploid chromosome numbers ranging from 22 to 28 (Britton-Davidian *et al.*, 2000). Of the 20 metacentric

chromosomes, 7 found on Madeira have been documented elsewhere in the distribution of *M. m. domesticus* (Piálek *et al.*, 2005). This includes metacentric 3.8, which phylogenetic studies suggest is likely to be one of the first Rb fusions to have arisen on the island (Britton-Davidian *et al.*, 2005; White *et al.*, 2010). As the proposed north European source of the Madeiran mice (based on mitochondrial (mt) DNA analysis: Gündüz *et al.* (2001a) and Förster *et al.* (2009)) includes a region of Denmark inhabited by races fixed for metacentric 3.8 (Piálek *et al.* (2005), and references therein), it is possible that this metacentric was introduced to Madeira from Denmark during colonisation. Such (human-mediated) long-distance dispersal of metacentric-bearing mice has already been confirmed for some European populations (Riginos and Nachman, 1999). In the case of Madeira, such a scenario may help to explain the extensive chromosomal diversification observed. Further, it has been reported that the presence of an Rb fusion in mouse laboratory stocks appears to predispose such stocks to further Rb fusions (Nachman and Searle (1995), and references therein). This presents the intriguing possibility that a metacentric introduced during colonisation could have led to a fusion ‘cascade’ on Madeira.

In this study, we use mapped microsatellite loci to investigate the Madeiran metacentric races. Our first aim was to investigate the possibility that metacentric 3.8 was introduced to Madeira. For this purpose, microsatellite loci closely linked to this Rb fusion (that is,

¹Research Group Evolutionary Genetics, Leibniz-Institute for Zoo and Wildlife Research, Berlin, Germany; ²Department of Biology, University of York, Wentworth Way, York, UK; ³Departamento de Biologia Animal, Faculdade de Ciências, Universidade de Lisboa, Lisboa, Portugal; ⁴Université Montpellier II, CNRS Institut des Sciences de l'Évolution, Montpellier, France and ⁵Department of Ecology and Evolutionary Biology, Cornell University, Ithaca, NY, USA
Correspondence: Dr DW Förster, Research Group Evolutionary Genetics, Leibniz-Institute for Zoo and Wildlife Research, Alfred-Kowalke-Strasse 17, Berlin 10315, Germany.
E-mail: dwgfoerster@gmail.com

Received 26 June 2012; revised 25 October 2012; accepted 5 November 2012; published online 12 December 2012

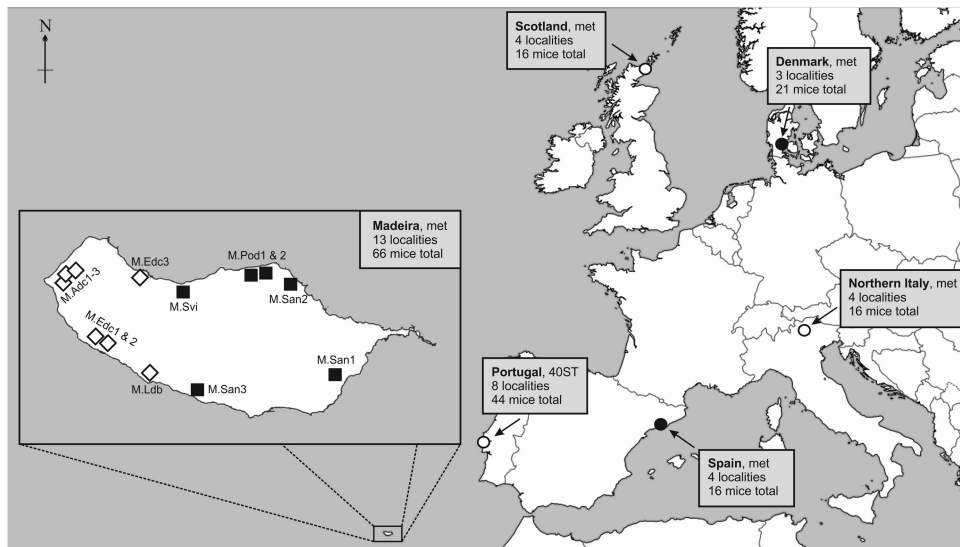


Figure 1 Map indicating the sampling localities used in this study; the island of Madeira is enlarged (left). Localities in continental Europe are indicated by circles (○), filled circles (●) represent localities where metacentric 3.8 is present. Text in boxes gives the chromosome status of *M. m. domesticus* at these localities (met = metacentric races; 40ST = standard karyotype (all-acrocentric) mice; see Table 1 for details); the number of sampling localities and total number of mice per geographic region is also indicated. Madeiran sampling localities are shown in the inset; see Table 1 for details about races. (■) Races with metacentric 3.8; and (◇) races with metacentrics 3.14 and 8.11.

loci mapping to the centromeres of the constituent chromosomes 3 and 8) were genotyped in mice from Madeira and elsewhere in the distribution of *M. m. domesticus* in Europe, including chromosome races with metacentric 3.8. Such loci have previously proven effective in demonstrating a shared evolutionary origin of a metacentric in the house mouse (Riginos and Nachman, 1999). Races with metacentric 3.8 from Denmark represent a potential source from a north European origin, if the metacentric was introduced during colonisation. An alternative source of metacentric 3.8, a race in Spain, may have contributed this metacentric subsequent to colonisation.

Additionally, we aimed to investigate the genetic input from mainland Portugal to Madeira. The island is a Portuguese dependency and passive transport (by humans) of mice from mainland Portugal to Madeira is expected to have occurred since the island's settlement in the 15th century. In spite of a virtual absence of Portuguese mtDNA sequences on Madeira (Gündüz *et al.*, 2001a; Förster *et al.*, 2009), there is high genic similarity between most Madeiran metacentric races and Portuguese house mice at nuclear allozyme markers (Britton-Davidian *et al.*, 2007). It has been proposed that mice introduced from Portugal likely represent a secondary wave of colonisers after an initial colonisation from northern Europe (Britton-Davidian *et al.*, 2007; Förster *et al.*, 2009). Several Portuguese mouse populations that could have contributed to this second wave of colonisers were included in the study (that is, from cities along the Atlantic coast with active harbours since the settlement period of Madeira from Portugal).

Finally, chromosomal speciation theory predicts that karyotypic divergence facilitates genetic divergence (for example, King, 1993), particularly in those chromosome regions involved in the karyotypic rearrangement(s) (for example, Faria and Navarro, 2010). We aimed to investigate this in the metacentric races of Madeira. As our efforts were focused on chromosomes 3 and 8, the metacentric races on Madeira were considered as two karyotypic groups: those with metacentric 3.8 and those with metacentrics 3.14 and 8.11 (see Figure 1). Gene flow between the metacentric races was investigated using microsatellite loci mapping to three chromosome regions (proximal, interstitial and

distal) of chromosomes 3 and 8. As the formation of metacentric chromosomes by Rb fusions or whole-arm reciprocal translocations involves the centromere, it was predicted that loci closely linked to the centromere (that is, proximal loci) would show an interruption to gene flow between the karyotypic groups. It was not expected that gene flow at loci located in the other two chromosome regions (that is, interstitial and distal loci) would be affected by the chromosome disposition (that is, 3.8 versus 3.14 and 8.11).

MATERIALS AND METHODS

Specimens

The 179 mice used in this study were collected at 36 localities in six geographical regions: Madeira, Portugal, Spain, Italy, Denmark and Scotland (Figure 1). Collection dates and karyotype information are shown in Table 1.

Karyotypes of mice were determined in previous studies, as follows: Madeira and Portugal, Förster *et al.* (2009); Denmark, Smadja *et al.* (2004); Spain, Gündüz *et al.* (2001b); northern Italy, Hauffe and Searle (1993) and Panithanarak *et al.* (2004); and Scotland, Searle (1991), Searle *et al.* (1993) and Palmer (1997).

Microsatellite typing

In all, 24 loci were chosen based on their genetic distance from the centromere (see Supplementary Table S1 for details). *Proximal loci*: D3mit60, D3mit117, D3mit221, D3mit62, D8mit155, D8mit58, D8mit124 and D8mit171. *Interstitial loci*: D3mit72, D3mit156, D3mit283, D3mit28, D8mit54, D8mit193, D8mit130 and D8mit178. *Distal loci*: D3mit87, D3mit116, D3mit163, D3mit19, D8mit52, D8mit93, D8mit56 and D8mit92.

Tail tips, spleens and/or toe clippings were stored in 100% ethanol and maintained at 4 °C. DNA was extracted using the Qiagen DNeasy Tissue Kit (Qiagen, Hilden, Germany) following the manufacturers' instructions. The polymerase chain reaction volume was 5 µl; polymerase chain reactions were carried out using the Qiagen Multiplex PCR Kit (Qiagen) following the instructions for multiplexing microsatellite loci. Reverse primers were labelled with one of the fluorescent dyes: FAM, HEX, VIC, NED or PET. Amplifications were carried out following the 'microsatellite cycling protocol' of the Qiagen Multiplex PCR Kit.

When microsatellites were labelled with FAM or HEX, alleles were separated using an Applied Biosystems 377a sequencer (Applied Biosystems, Foster City,

Table 1 Specimen details

Geographic area	Locality	Code	N	Metacentrics	2n	Trap date
Madeira	Camacha	M.San1	5	2.19 3.8 4.16 5.14 6.7 9.10 11.12 13.17 15.18	22	2004
	Santana	M.San2	5	2.19 3.8 4.16 5.14 6.7 9.10 11.12 13.17 15.18	22	2004
	Campanario	M.San3	5	2.19 3.8 4.16 5.14 6.7 9.10 11.12 13.17 15.18	22	2004
	Ponta Delgada no. 1	M.Pod1	5	3.8 4.5 10.16 11.12 9.18 14.17	28	2004
	Ponta Delgada no. 2	M.Pod2	5	3.8 4.5 10.16 11.12 9.18 14.17	28	2004
	São Vicente	M.Svi	4	2.4 3.8 5.18 6.7 9.12 10.16 11.19 13.17	24	2004
	Lugar de Baixo	M.Ldb	7	2.4 3.14 6.7 8.11 9.12 10.16 13.17 15.18	24	2004
	Estreito da Calheta no. 1	M.Edc1	5	2.4 3.14 5.18 6.7 8.11 9.12 10.16 13.17	24	2004
	Estreito da Calheta no. 2	M.Edc2	5	2.4 3.14 5.18 6.7 8.11 9.12 10.16 13.17	24	2004
	Seixal	M.Edc3	5	2.4 3.14 5.18 6.7 8.11 9.12 10.16 13.17	24	2004
	Ponta do Pargo	M.Adc1	5	2.4 3.14 5.18 7.15 8.11 9.12 10.16 13.17	24	2004
	Sra. do Amparo no. 1	M.Adc2	5	2.4 3.14 5.18 7.15 8.11 9.12 10.16 13.17	24	2004
	Sra. do Amparo no. 2	M.Adc3	5	2.4 3.14 5.18 7.15 8.11 9.12 10.16 13.17	24	2004
Portugal	Lisbon	P.Lis	6	—	40	2005
	Setubal	P.Set	6	—	40	2005
	Vila Franca de Xira	P.Vfx	6	—	40	2005
	Vianna de Castelo	P.Vdc	5	—	40	2005
	Porto	P.Por	4	—	40	2005
	Sines	P.Sin	6	—	40	2005
	Lagos	P.Lag	5	—	40	2005
	Tavira	P.Tav	6	—	40	2005
Denmark	Lunderskov	DK.Lu	6	2.5 3.8 6.9	34	1992
	Sommersted	DK.Sm	9	2.5 3.8 6.9	34	1992
	Sonder Bjert	DK.Sn	6	3.8 6.9	36	1992
Spain	Avinyonet del Penedes	E.Avi	4	4.14 5.15 9.11 12.13	32	1996
	Viladecans	E.Vil	4	4.14 5.15 6.10 9.11 12.13	30	1997
	Sant Marti de Sarroca	E.Sms	4	4.14 5.15 9.11 12.13	32	1997
	Garraf	E.Gar	4	3.8 4.14 5.15 6.10 9.11 12.13	28	1997
Scotland	West Canisbay	UK.Wc	5	4.10 6.13 9.12 11.14	32	1994
	Mains of Ollrig	UK.Mo	5	4.10 6.13 9.12	34	1987
	Dunnet	UK.Du	2	4.10 6.13 9.12 11.14	32	1994
	John O'Groats	UK.Jo	4	4.10 6.13 9.12 11.14	32	1992
Northern Italy	Sondalo	I.Son	4	1.3 2.8 4.6 5.15 9.14 10.12 11.13 16.17	24	2001
	Migiondo	I.Mig	4	1.3 4.6 5.15 8.12 9.14 11.13 16.17	26	1990
	Sernio	I.Ser	4	1.3 2.8 4.6 5.15 9.14 10.12 11.13 16.17	24	1990
	Pedrengo	I.Ped	4	1.6 2.8 3.4 5.15 7.18 9.14 10.12 11.13 16.17	22	2002

Metacentric chromosomes with constituent chromosomes 3 and/or 8 are highlighted in bold.

CA, USA); when they were labelled with FAM, VIC, NED or PET, they could be separated using an Applied Biosystems 3130xl sequencer (Applied Biosystems). GENEMAPPER v.4.0 (Applied Biosystems) was used to determine allele sizes. Alleles were scored two times; peak morphologies and sizes were assessed by eye. Alleles were allocated using histograms of raw allele sizes given to two decimal places of a base pair.

Analysis

MICROCHECKER v.2.2.3 (Van Oosterhout *et al.*, 2004) was used to test for null alleles, large allele drop-out and scoring errors. One locus (D8mit92) was removed from the data set because of the presence of null alleles. Two further loci were excluded because they either did not amplify consistently (D3mit62) or had peak profiles with extensive stuttering (D3mit28), making it impossible to score with confidence. All further analyses were carried out with the remaining 21 loci (Supplementary Table S1).

GENEPOP v.3.4 (Raymond and Rousset, 1995) was used to test deviations from Hardy–Weinberg proportions. Observed heterozygosity (H_O) and unbiased

expected heterozygosity (H_E) were estimated following Nei (1978) in POPGENE v.1.32 (Yeh *et al.*, 1997).

F-statistic (F_{ST}) values were calculated as θ (Weir and Cockerham, 1984) in FSTAT v.2.9.3.2 (Goudet, 2002). FSTAT was also used to estimate allelic frequencies. R-statistic (R_{ST}) values were determined in RSTCALC v.2.2 (Goodman, 1997). R_{ST} values were calculated by first averaging variance components over loci, following Slatkin (1995). For multilocus estimates, data were 'globally standardised' with the *standard.exe* application of RSTCALC, so that loci with low variances may contribute to the final estimates of R_{ST} . For the two summary statistics, P -values were corrected for multiple comparisons when appropriate (Bonferroni correction).

Several analyses were carried out in the statistical programming environment R (<http://www.cran.r-project.org/>). We calculated Spearman's correlation coefficients (and their significance levels) to examine if levels of genetic diversity (H_O , H_E , mean number of alleles) were affected by differences in sampling. A principal component analysis (PCA) was carried out using the R package adegenet v.1.3-3 (Jombart, 2008), to examine the genetic relationship

among localities (at two scales: by locality and by geographic region). Three further PCAs were carried out to determine the genetic differentiation among the six Madeiran races (13 localities) by chromosome region (proximal, interstitial and distal). To examine to what degree chromosome disposition (3.8 versus 3.14 and 8.11) and chromosome race affected population genetic structuring, we carried out an analysis of molecular variance with four hierarchical levels using the R package Hierfstat v.0.04-6 (Goudet, 2005).

RESULTS

Microsatellite variation

The microsatellite loci were highly polymorphic. Considering 179 mice from 36 localities in six geographic regions (Table 1), 331 alleles were found for the 21 microsatellite loci. The mean number of alleles per locus varied between localities, with a range of 1.2–4.7 (overall mean = 3.5; Table 2). This was significantly correlated with sample size ($r_s = 0.534$; $P < 0.01$), as expected with small but varying sample sizes. H_E 's are also significantly correlated with sample size ($r_s = 0.38$; $P = 0.02$). They show values between 0.069 (I.Ser) and 0.672 (M.San3), similar to the range of H_O values: from 0.119 (I.Ser) to 0.771 (P.Lag). Partitioning of genetic diversity was significant under both the infinite-allele model ($F_{ST} = 0.285$, $P < 0.001$) and the stepwise-mutation model ($R_{ST} = 0.319$, $P < 0.001$). These values compare to previous microsatellite studies in *M. m. domesticus* (Dallas *et al.*, 1998; Panithanarak *et al.*, 2004; Franchini *et al.*, 2008, 2010).

Genetic differentiation between geographic regions

Summary statistics (F_{ST} and R_{ST}) for pairwise comparisons between geographic regions are shown in Table 3. Localities within geographic regions were pooled with the exception of Madeira, for which specimens were pooled into two groups: mice with metacentric 3.8 versus those without (that is, all specimens with 3.14 and 8.11). This was carried out because of interest in the amount of differentiation between these two karyotypic groups on Madeira (see below). Differentiation among geographic regions measured under the infinite-allele model and stepwise-mutation model was highly significant (Table 3); R_{ST} values tended to be higher in this analysis (F_{ST} : 0.067–0.289; R_{ST} : 0.101–0.478). Although the extent of differentiation was sometimes inconsistent between F_{ST} and R_{ST} , both summary statistics show that the pairwise comparison between Madeiran samples with and without metacentric 3.8 have the lowest values, followed by pairwise comparisons involving either of the Madeiran samples and the Portugal sample (Table 3).

Relationship of Madeiran metacentric populations and other localities

Selectively neutral loci linked to the centromere of a metacentric (that is, proximal loci) are the most likely to reflect accurately the evolutionary history of the chromosome involved (Riginos and Nachman, 1999). Regarding the colonisation of Madeira, use of such loci on mouse chromosomes 3 and 8 enabled us to examine if metacentric 3.8 was introduced to Madeira during colonisation of the island.

Table 4 shows pairwise F_{ST} values (proximal loci only) between Madeiran localities characterised by metacentric 3.8 and the remaining 30 localities. This summary statistic does not indicate a close genetic relationship between Madeiran mice with metacentric 3.8 and mice carrying this Rb fusion from Denmark (DK.Lu, DK.Sm and DK.Sn) or Spain (E.Gar). Rather, the data on proximal loci recapitulates the pattern seen when examining all 21 microsatellites (Table 3): the lowest pairwise values are observed among the two

Table 2 Summary of multilocus (21 microsatellite loci) population genetic analyses

Region	Code ^a	N	A (s.d.)	H_O (s.d.)	H_E (s.d.)
<i>Madeira</i>					
	M.San1	5	3.5 (1.2)	0.597 (0.333)	0.589 (0.182)
	M.San2	5	3.6 (1.0)	0.595 (0.238)	0.543 (0.168)
	M.San3	5	4.6 (1.6)	0.695 (0.242)	0.672 (0.162)
	M.Pod1	5	4.4 (1.6)	0.657 (0.229)	0.636 (0.159)
	M.Pod2	5	4.0 (1.1)	0.648 (0.268)	0.625 (0.163)
	M.Svi	4	3.6 (1.2)	0.635 (0.252)	0.589 (0.187)
	M.Ldb	7	5.1 (1.9)	0.603 (0.221)	0.669 (0.154)
	M.Edc1	5	4.6 (1.7)	0.657 (0.269)	0.671 (0.209)
	M.Edc2	5	4.7 (1.6)	0.657 (0.211)	0.650 (0.172)
	M.Edc3	5	4.0 (1.1)	0.686 (0.273)	0.630 (0.163)
	M.Adc1	5	4.5 (1.5)	0.681 (0.260)	0.647 (0.189)
	M.Adc2	5	3.8 (1.5)	0.612 (0.303)	0.545 (0.239)
	M.Adc3	5	3.9 (1.3)	0.657 (0.317)	0.581 (0.193)
<i>Portugal</i>					
	P.Lis	6	3.5 (1.0)	0.643 (0.232)	0.572 (0.140)
	P.Set	6	3.4 (1.3)	0.587 (0.227)	0.547 (0.195)
	P.Vfx	6	4.1 (1.1)	0.437 (0.227)	0.569 (0.146)
	P.Vdc	5	2.9 (1.0)	0.533 (0.292)	0.483 (0.205)
	P.Por	4	3.2 (1.1)	0.619 (0.245)	0.557 (0.150)
	P.Sin	6	4.6 (1.7)	0.620 (0.172)	0.667 (0.137)
	P.Lag	5	3.0 (0.9)	0.771 (0.312)	0.539 (0.182)
	P.Tav	6	2.9 (0.9)	0.605 (0.192)	0.529 (0.170)
<i>Denmark</i>					
	DK.Lu	6	4.3 (1.7)	0.519 (0.256)	0.622 (0.214)
	DK.Sm	9	4.0 (1.5)	0.466 (0.246)	0.580 (0.176)
	DK.Sn	6	3.0 (1.0)	0.598 (0.329)	0.522 (0.191)
<i>Spain</i>					
	E.Avi	4	2.7 (1.3)	0.726 (0.378)	0.503 (0.238)
	E.Vil	4	3.9 (1.4)	0.714 (0.266)	0.631 (0.136)
	E.Sms	4	3.2 (0.9)	0.726 (0.295)	0.580 (0.134)
	E.Gar	4	2.6 (0.9)	0.571 (0.327)	0.432 (0.205)
<i>Scotland</i>					
	UK.Wc	5	2.2 (0.9)	0.524 (0.355)	0.392 (0.238)
	UK.Mo	5	3.0 (1.1)	0.524 (0.300)	0.501 (0.233)
	UK.Du	2	2.5 (1.0)	0.595 (0.407)	0.476 (0.261)
	UK.Jo	4	2.6 (1.2)	0.536 (0.356)	0.418 (0.269)
<i>Northern Italy</i>					
	I.Son	4	3.4 (1.0)	0.619 (0.302)	0.573 (0.181)
	I.Mig	4	2.0 (1.2)	0.250 (0.274)	0.275 (0.275)
	I.Ser	4	1.2 (0.4)	0.119 (0.302)	0.069 (0.158)
	I.Ped	4	2.0 (0.6)	0.345 (0.331)	0.331 (0.213)
<i>Total</i>	179			0.584 (0.096)	0.857 (0.045)

Abbreviations: A, mean number of alleles per locus; H_E , expected heterozygosities; H_O , observed heterozygosities; N, number of typed specimens; s.d., standard deviation.
^aFollowing Table 1.

karyotypic groups on Madeira (F_{ST} : 0.125–0.405), followed by comparisons involving Portuguese localities (F_{ST} : 0.206–0.413).

A PCA based on only proximal loci (Figure 2) reflects the genetic affinities discerned from the summary statistics: the two Madeiran samples (mice with and without metacentric 3.8) show genetic similarity, and are clearly separated from all but the Portuguese mice

Table 3 Estimates of pairwise genetic differentiation (F_{ST} and R_{ST}) between geographic regions based on 21 microsatellite loci

	Madeira		Portugal	Denmark	Spain	Scotland	Italy
	3.8	w/o 3.8					
Madeira 3.8		0.067	0.086	0.193	0.154	0.204	0.203
Madeira w/o 3.8	0.101		0.104	0.229	0.175	0.231	0.230
Portugal	0.142	0.153		0.143	0.088	0.169	0.157
Denmark	0.399	0.363	0.306		0.217	0.225	0.289
Spain	0.226	0.279	0.211	0.478		0.196	0.140
Scotland	0.375	0.354	0.258	0.298	0.376		0.279
Italy	0.399	0.415	0.295	0.423	0.295	0.364	

F-statistic (F_{ST}) values above diagonal; R-statistic (R_{ST}) values below diagonal. All values are highly significant ($P < 0.001$) after correction for multiple comparisons. Note that the Madeiran samples have been divided into specimens homozygous for metacentric 3.8 ('Madeira 3.8') and those that carry its monobrachial homologues ('Madeira w/o 3.8'); see Table 1 and text for details.

along the first principal component (which explains 17.28% of variation). The same analysis conducted at the scale of locality (Supplementary Figure S1) also shows no evidence of genetic similarity between Madeiran, Danish or Spanish localities characterised by metacentric 3.8. In the latter PCA, three Portuguese localities (Setubal, Sines and Tavira) appear genetically most similar to the Madeiran populations characterised by metacentric 3.8. However, these three Portuguese localities do not stand out based on F_{ST} values (proximal loci only; Table 4), in comparison to other Portuguese localities, or Madeiran populations characterised by metacentrics 3.14 and 8.11. It is worth noting that the populations on Madeira characterised by metacentric 3.8 occupy the eastern portion of the island (Figure 1), where the main port and airport are located; this is the area most likely to receive mouse immigrants from Portugal, and mice here show within-race polymorphism involving metacentrics and their acrocentric homologues (Britton-Davidian *et al.*, 2005, 2007). However, it has not been conclusively determined if this chromosomal polymorphism is due to immigration of standard karyotype (all-acrocentric) mice, and no mice with what could be considered Portuguese mtDNA haplotypes have been detected in this area (Gündüz *et al.*, 2001a; Förster *et al.*, 2009).

Relationship among Madeiran karyotypic groups

Overall differentiation on the island was highly significant when considering all loci and all individuals ($F_{ST} = 0.129$; $R_{ST} = 0.168$; $P < 0.001$ in both cases).

To examine the impact of chromosome disposition (metacentric 3.8 versus 3.14 and 8.11) and chromosome race on genetic structure, we conducted an analysis of molecular variance at four hierarchical levels: (1) among karyotypic groups (3.8 versus 3.14 and 8.11); (2) among chromosome races within karyotypic groups; (3) among populations within races; and (4) within populations. The results are shown in Table 5. For this analysis, we examined the three chromosome regions separately (proximal, interstitial and distal). The proximal region shows a significant proportion of variation partitioned among karyotypic groups (14.05%), whereas the interstitial and distal regions do not (Table 5). At the lower hierarchical levels, all three chromosome regions show significant genetic structuring—with the 'within-population' level harbouring by far the greatest proportion of variation. Of note is that the proportion of variation at proximal loci is greater between karyotypic groups (14.05%) than between races within these groups (4.09%) or among populations

Table 4 Pairwise F_{ST} values between Madeiran localities characterised by metacentric 3.8 and the remaining 30 localities

Geographic region	Locality	Differentiation at seven proximal loci (F_{ST})					
		M.San1	M.San2	M.San3	M.Pod1	M.Pod2	M.Svi
Madeira, met	M.Ldb	0.164	0.249	0.125	0.228	0.154	0.163
	M.Edc1	0.201	0.251	0.156	0.313	0.160	0.177
	M.Edc2	0.221	0.304	0.255	0.350	0.236	0.252
	M.Edc3	0.167	0.263	0.231	0.280	0.200	0.204
	M.Adc1	0.244	0.285	0.275	0.353	0.220	0.251
	M.Adc2	0.291	0.389	0.290	0.405	0.303	0.325
	M.Adc3	0.147	0.265	0.184	0.355	0.175	0.223
Portugal, 40ST	P.Lis	0.265	0.353	0.212	0.274	0.273	0.267
	P.Set	0.345	0.376	0.313	0.401	0.291	0.319
	P.Vfx	0.278	0.345	0.227	0.281	0.251	0.208
	P.Vdc	0.339	0.395	0.310	0.387	0.283	0.300
	P.Por	0.288	0.336	0.231	0.311	0.206	0.256
	P.Sin	0.273	0.316	0.238	0.310	0.231	0.260
	P.Lag	0.340	0.404	0.325	0.335	0.286	0.333
	P.Tav	0.412	0.413	0.382	0.403	0.355	0.370
	Denmark, met	DK.Lu ^a	0.295	0.374	0.273	0.296	0.280
DK.Sm ^a		0.339	0.436	0.300	0.412	0.336	0.326
DK.Sn ^a		0.284	0.381	0.251	0.341	0.274	0.270
Spain, met	E.Avi	0.288 ^b	0.365 ^b	0.266 ^b	0.272 ^b	0.252 ^b	0.235 ^b
	E.Vil	0.267	0.349	0.243	0.288	0.223	0.249
	E.Sms	0.339	0.386	0.297	0.328	0.288	0.276
	E.Gar ^a	0.465	0.551	0.465	0.496	0.462	0.444
Scotland, met	UK.Wc	0.376	0.449	0.356	0.369	0.362	0.343
	UK.Mo	0.292	0.407	0.283	0.304	0.302	0.288
	UK.Du	0.260	0.395	0.268	0.317	0.260	0.234
	UK.Jo	0.388	0.493	0.380	0.378	0.402	0.380
Northern Italy, met	I.Son	0.333	0.439	0.360	0.365	0.332	0.325
	I.Mig	0.580	0.675	0.587	0.638	0.564	0.568
	I.Ser	0.533	0.635	0.547	0.589	0.522	0.532
	I.Ped	0.386	0.497	0.427	0.488	0.434	0.435

Pairwise comparisons are significant at $P < 0.01$, except those in italics that are significant at $P < 0.05$.

^aMice at these four localities have metacentric 3.8 (see Table 1).

^bToo much missing data to calculate significance (included for completeness sake).

within races (5.78%). At the latter two hierarchical levels, the proportion of variation at proximal loci lays in the range of the other two chromosome regions (Table 5).

Three PCAs were carried out on the Madeiran data set, each corresponding to one of the three chromosome regions investigated (Figure 3). The plot for proximal loci (Figure 3a) shows a clear separation of specimens homozygous for metacentric 3.8 and its monobrachial homologues on the first principal component (which explains 19.43% of variation). The PCA plots for the interstitial and distal regions (Figures 3b and c) do not show a similar separation of the two karyotypic groups.

The high differentiation in the proximal region between the karyotypic groups appears to be more a reflection of the loci on chromosome 3 (three loci, $F_{ST} = 0.237$, $P < 0.001$) than those on chromosome 8 (four loci, $F_{ST} = 0.081$, $P < 0.001$). Our outlier analysis

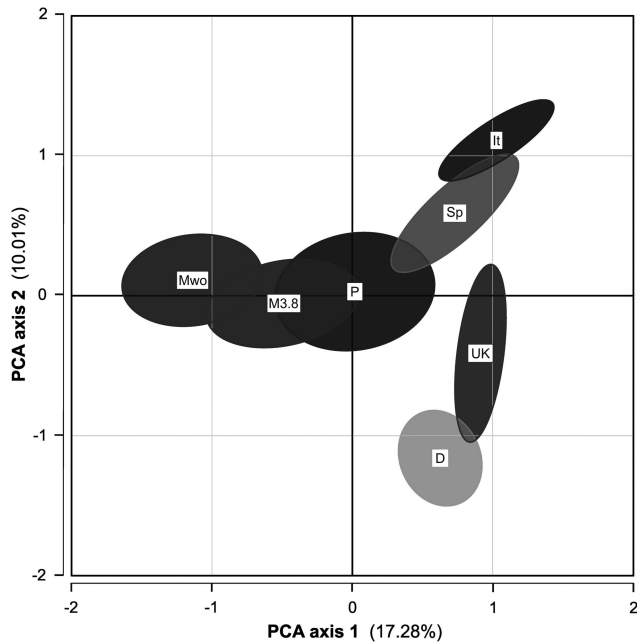


Figure 2 PCA plot showing relationship among *M. m. domesticus* in this study by geographic region (seven proximal loci; see Supplementary Table S1) on first two principal components. Ellipses capture 95% variation among mice from a given geographic region. M3.8, Madeiran specimens with metacentric 3.8; Mwo, Madeiran specimens without metacentric 3.8; P, Portugal; Sp, Spain; It, Northern Italy; UK, Scotland; and D, Denmark.

identified two proximal loci on chromosome 3 as having significantly higher F_{ST} values than expected given the level of genetic diversity and differentiation among Madeiran specimens, but tests of neutrality do not suggest that this is the result of selection (Supplementary Figure S2). Supplementary Table S2 shows the frequencies of alleles at three proximal loci that appear exclusive or nearly exclusive to one of the karyotypic groups; those on chromosome 3 show the strongest allelic association with karyotype.

DISCUSSION

Common versus multiple origin of Rb metacentric 3.8

For metacentrics generated by Rb fusion, selectively neutral loci linked to the centromere are most likely to reflect the evolutionary history of the chromosome concerned (Riginos and Nachman, 1999). Neither summary statistics nor principal component analyses were able to demonstrate strong genetic similarity at such loci between any of the mainland populations of house mice characterised by metacentric 3.8 and those on Madeira with this chromosome. Instead, our microsatellite results and the allozyme results of Britton-Davidian *et al.* (2007) demonstrate an affinity between Madeiran mice and those on the Portuguese mainland.

The chromosomal divergence on Madeira is remarkable, particularly given the timeframe within which this must have arisen (Förster *et al.*, 2009). The introduction of a metacentric from elsewhere could potentially have contributed to the karyotypic diversity observed. We chose to investigate metacentric 3.8 in this context, because it was likely one of the earliest metacentrics to arise on Madeira (Britton-Davidian *et al.*, 2005; White *et al.*, 2010) and its presence in the proposed colonisation source (Gündüz *et al.*, 2001a; Förster *et al.*, 2009). Our results do not support the introduction

Table 5 AMOVA among Madeiran mice, by chromosome region and by locus

	Per cent of variance			
	Among groups	Among races within groups	Among populations within races	Within populations
<i>Proximal loci</i>	14.05*	4.09**	5.78*	76.08***
D3mit60	44.22	-0.75 ^a	5.23	51.31
D3mit117	-3.06 ^a	3.62	10.64	88.80
D3mit221	20.42	5.53	6.21	67.84
D8mit155	12.39	5.51	0.10	82.00
D8mit58	7.77	10.07	4.45	77.71
D8mit124	3.81	-1.47 ^a	7.62	90.04
D8mit171	-2.50 ^a	6.99	8.08	87.43
<i>Interstitial loci</i>	-1.32^a	7.09***	4.39***	89.84***
D3mit72	-4.02 ^a	13.71	-2.81 ^a	93.13
D3mit156	-6.13 ^a	17.34	8.35	80.43
D3mit283	0.21	1.68	5.21	92.91
D8mit54	1.30	0.36	10.76	87.57
D8mit193	2.32	2.84	1.07	93.77
D8mit178	-1.65 ^a	7.72	0.98	92.95
D8mit130	-1.48 ^a	6.44	6.55	88.49
<i>Distal loci</i>	0.44	2.38*	8.89***	88.30***
D3mit87	2.32	4.88	5.32	87.48
D3mit116	-3.49 ^a	6.36	12.79	84.34
D3mit163	0.46	1.35	8.23	89.97
D3mit19	4.26	1.37	9.23	85.13
D8mit52	0.30	1.44	16.33	81.92
D8mit93	-0.23 ^a	1.14	4.07	95.02
D8mit56	-1.14 ^a	0.35	7.60	93.19

Abbreviation: AMOVA, analysis of molecular variance.

To examine the impact of chromosome disposition (metacentric 3.8 versus 3.14 and 8.11) and chromosome race on genetic structure, we conducted an AMOVA at four hierarchical levels: (1) among karyotypic groups (3.8 vs 3.14 and 8.11); (2) among chromosome races within karyotypic groups; (3) among populations within races; and (4) within populations. Significance of values for chromosome regions (in bold) is indicated: * $P < 0.05$, ** $P < 0.01$ and *** $P < 0.001$.

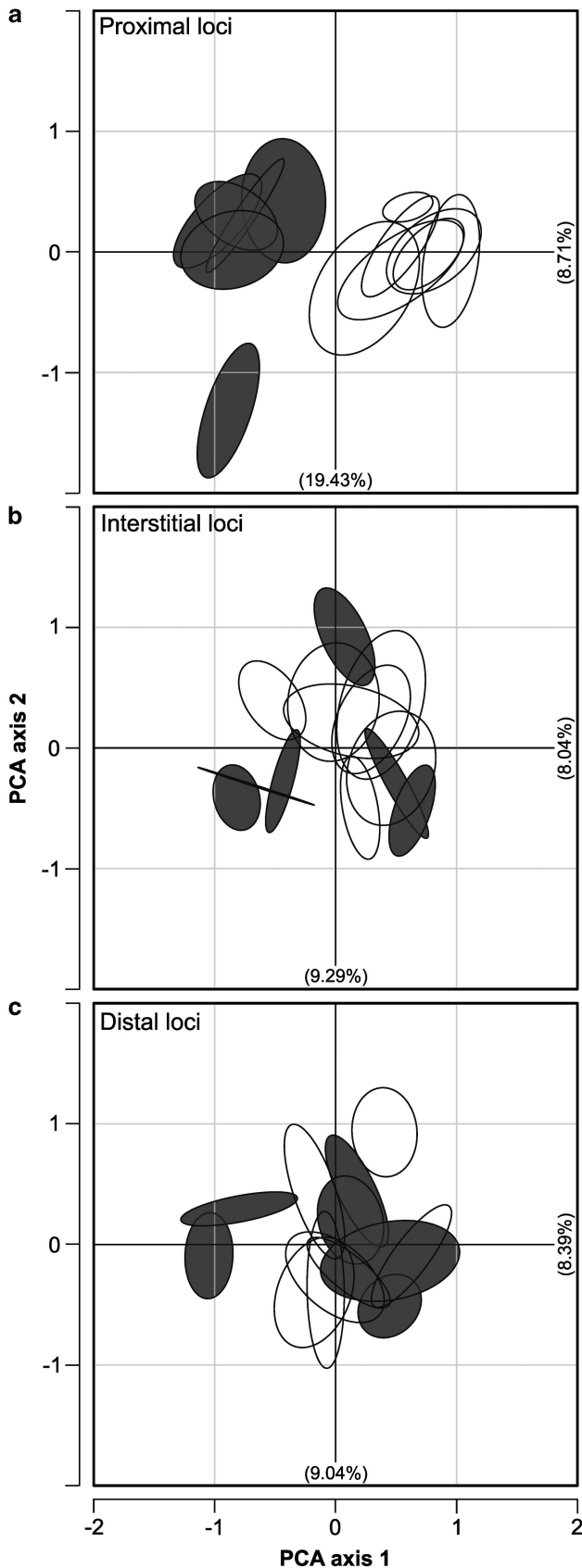
^aNegative values may occur when the true value is zero (Schneider *et al.*, 1997).

of metacentric 3.8 during colonisation from northern Europe (Denmark), or thereafter from Spain.

Relationship of Madeiran metacentric mice with those of elsewhere

Island populations are expected to show genetic signatures of their source area, particularly when oceanic barriers greatly reduce post-colonisation gene flow (Avice, 2000). The large distance of Madeira to the closest continental landmass (600 km) supports a human-mediated colonisation event linked to the arrival of seafarers on the island. Phylogenetic analyses with mtDNA provide very strong evidence for a north European source of the first mouse colonists (Gündüz *et al.*, 2001a; Förster *et al.*, 2009). This result was unexpected because of the absence of historical documentation for a north European presence on the archipelago, whereas the human association with Portugal from the fifteenth century onward is well documented (Albuquerque and Vieira, 1988). Our analysis of microsatellite variation confirms results of previous studies indicating gene flow from mainland Portugal to Madeira (Britton-Davidian *et al.*, 2007; Förster *et al.*, 2009).

The high level of genetic diversity observed by us and Britton-Davidian *et al.* (2007) for two types of nuclear marker



(microsatellites, allozymes) is unexpected considering (i) the process of chromosomal divergence and (ii) the fragmentation of habitat on Madeira suitable for house mouse occupation. Britton-Davidian *et al.* (2007) suggested that a significant influx of mice from Portugal might have taken place during the rapid expansion of farming communities on Madeira, thus potentially introducing variation to a genetically impoverished Madeiran mouse population. This interpretation is consistent with the genetic similarity between Madeiran mouse populations and those of mainland Portugal (this study; Britton-Davidian *et al.*, 2007), as well as the high levels of variation detected at nuclear markers. A similar situation has been reported for mouse populations on Mediterranean islands, which show no reduction in genetic variability compared with mainland populations, presumably as a result of multiple (human-mediated) introductions of house mice (Navarro and Britton-Davidian, 1989).

However, a high expansion rate, which is consistent with the mtDNA results (Förster *et al.*, 2009), could also explain the high levels of microsatellite and allozyme variability detected (Zenger *et al.*, 2003). In addition, microsatellite studies investigating population bottlenecks (for example, Keller *et al.*, 2001; Busch *et al.*, 2007) have demonstrated that very low levels of immigration (on the level of a few individuals) can recover levels of variability to pre-bottleneck levels in a small number of generations, even when the incidence of inbreeding is high. It is thus difficult to ascertain conclusively to what extent Portuguese mouse immigrants have contributed to the current high levels of variability, although the genetic similarity at nuclear loci between Madeiran and Portuguese mice does provide compelling evidence for an influx of individuals. Of note is that the microsatellite data suggest that the Madeiran populations most likely to come into contact with mouse immigrants (that is, closest to the port city of Funchal) are those that show signs of gene flow from mainland Portugal.

Highly divergent chromosome races on Madeira show low-to-moderate genetic differentiation

While genetic differentiation among the Madeiran metacentric mice is highly significant ($F_{ST} = 0.129$; $R_{ST} = 0.168$; $P < 0.001$ in both cases), the degree of differentiation observed is lower than may be expected considering the extensive chromosomal divergence on the island, which involves 20 or more chromosomal rearrangements (Britton-Davidian *et al.*, 2005; White *et al.*, 2010). Previous studies in house mice involving chromosomally less divergent races typically show higher levels of genetic differentiation at microsatellite loci ($F_{ST} = 0.39$, 13 microsatellite loci (Panithanarak *et al.*, 2004); $F_{ST} = 0.187$, 8 microsatellite loci (Franchini *et al.*, 2008); $F_{ST} = 0.133$, 12 microsatellite loci (Franchini *et al.*, 2010)). More pronounced differentiation has also been recorded in a study involving metacentric and standard karyotype (all-acrocentric) mice ($F_{ST} = 0.29$; eight microsatellite loci (Dallas *et al.*, 1998)), as well as between standard karyotype mice, even in localities only 300 m distant from one another ($F_{ST} = 0.18$; six microsatellite loci (Dallas *et al.*, 1995)). The effects of individual microsatellite loci may be strong and studies using fewer loci may show greater differentiation. A study of Madeiran metacentric races using 39 microsatellite loci spread over

Figure 3 PCA plot showing relationship among Madeiran metacentric races of *M. m. domesticus*. Ellipses capture 95% variation among mice from a given population. Specimens homozygous for metacentric 3.8 are represented by filled ellipses (six populations), whereas those homozygous for its monobrachial homologues 3.14 and 8.11 are represented by open ellipses (seven populations). PCAs are shown by chromosome region: (a) proximal loci; (b) interstitial loci; and (c) distal loci.

six chromosomes (Förster, 2007) indicates that the degree of differentiation reported here is not unique to chromosomes 3 and 8. As three of the Italian samples in this study (I.Son, I.Mig and I.Ser; Table 1) correspond to sampling localities analysed by Panithanarak *et al.* (2004), we were able to examine if our results were affected by the particular loci chosen. Results for this analysis of Italian mice with our set of loci ($F_{ST}=0.404$; 95% confidence interval (CI): 0.308–0.487) are clearly similar to those obtained by Panithanarak *et al.* (2004) and confirm that the degree of differentiation observed for the Madeiran metacentric mice is not because of some property of the particular microsatellites used in this study. This is further supported by the observed degree of differentiation between localities in Portugal ($F_{ST}=0.287$; 95% CI: 0.249–0.326), Scotland ($F_{ST}=0.225$; 95% CI: 0.158–0.301) and Spain ($F_{ST}=0.149$; 95% CI: 0.106–0.189), which is in the same range as previous studies (Dallas *et al.*, 1995, 1998; Panithanarak *et al.*, 2004; Franchini *et al.*, 2008, 2010).

As there are several lines of evidence supporting a single origin of the Madeiran mice (Gündüz *et al.*, 2001a; Britton-Davidian *et al.*, 2005; Förster *et al.*, 2009), it is not surprising that these races exhibit genetic similarities. The low degree of differentiation we found among Danish mice ($F_{ST}=0.082$; 95% CI: 0.027–0.161), which share metacentric 3.8 by descent (Piálek *et al.* (2005), and references therein), appears consistent with such an interpretation. As the number of fixed karyotypic differences between races increases, the strength of the chromosomal barrier to gene flow is expected to increase (for example, Hauffe and Searle, 1998). The low degree of differentiation observed on Madeira for microsatellites (this study), allozymes (Britton-Davidian *et al.*, 2007) and mt sequences (Förster *et al.*, 2009) may thus reflect that karyotypic divergence on Madeira is recent and involved individuals from one source.

Differential gene flow between chromosome regions

Karyotypic differences are expected to act as barriers to gene flow (for example, King, 1993; Faria and Navarro, 2010). Some chromosomal speciation models propose that it is *reduced recombination* in those parts of the genome involved in chromosomal rearrangements that facilitates divergence (see Butlin (2005) and Faria and Navarro (2010)), whereas earlier models emphasise the reduction of *reproductive fitness* of karyotypic heterozygotes (that is, post-zygotic isolation, see King (1993) and Coyne and Orr (2004)). In both types of model it is the portion of the genome affected by chromosomal rearrangements that experiences a reduction in gene flow, promoting divergence. In the case of Rb fusions and whole-arm reciprocal translocations, the prediction is that loci in the centromeric region (referred to by us as 'proximal loci') should experience diminished gene flow, because the centromere is involved in these rearrangements.

There is evidence that the processes described by both types of chromosomal speciation model may apply in the house mouse. It has been demonstrated that house mice heterozygous for metacentrics have impaired reproductive fitness (for example, Wallace *et al.*, 1992, 2002; Saïd *et al.*, 1993; Hauffe and Searle, 1998), and that there are recombination effects in the vicinity of the centromere in such heterozygotes (for example, Davisson and Akeson, 1993; Bidau *et al.*, 2001; Merico *et al.*, 2003). Thus, one or both of these processes could affect gene flow in the centromeric region of the metacentric chromosomes investigated by us.

This potential chromosomal barrier to gene flow was investigated using microsatellite loci mapping to three chromosome regions (proximal, interstitial and distal) of mouse chromosomes 3 and 8. We found that the extent of genetic differentiation between karyotypic groups on Madeira varied according to the

chromosome region examined. As predicted, higher genetic structure was observed near the centromere (proximal microsatellite loci) than in the other two chromosome regions examined (interstitial and distal microsatellite loci). This resulted from the association of alleles at proximal loci with either metacentric 3.8 or with metacentrics 3.14 and 8.11. A similar situation was reported by Panithanarak *et al.* (2004), who examined two chromosome regions (proximal and distal) in a system of house mouse metacentric races using mapped microsatellite loci.

It should be noted that the centromeric region is expected to show reduced recombination in general (for example, Carneiro *et al.*, 2009; Geraldès *et al.*, 2011; Nachman and Payseur, 2012), and that this has been documented for metacentric chromosomes in the house mouse (Dumas and Britton-Davidian, 2002). As the extent of genetic differentiation in the three chromosome regions was essentially the same among races sharing the same disposition of chromosomes (3.8 or 3.14 and 8.11), as well as within races, it seems unlikely that solely a 'centromeric effect' is responsible for the significant divergence observed at proximal loci between karyotypic groups (that is, 3.8 versus 3.14 and 8.11).

Not all proximal loci were characterised by alleles with a strong association with karyotype. Although this may reflect the short timeframe of the chromosomal radiation on Madeira (Britton-Davidian *et al.*, 2007; Förster *et al.*, 2009), the following factors may have contributed:

- (1) *Recombination*: It has been suggested (Panithanarak *et al.*, 2004) that association with a particular karyotype is not absolute, and that recombination breakpoints may occur between a proximal locus and the centromere, even when its genetic distance to the centromere is reported to be 0 cM. Not all of the proximal microsatellite loci investigated had zero recombination distance from the centromere (Supplementary Table S1).
- (2) *Random sampling of alleles*: The random sampling of alleles during the fixation process in founding populations may involve common or uncommon alleles. Loci initially characterised by uncommon alleles may subsequently show a strong association with karyotype (like those in Supplementary Table S2), whereas loci initially characterised by common alleles may not.

Overall, our results are in agreement with chromosomal speciation models that predict higher divergence (or less introgression) in the vicinity of chromosomal rearrangements despite hybridisation and gene flow (for example, Butlin, 2005; Faria and Navarro, 2010). However, studies on more chromosomes are desirable. Thus, Panithanarak *et al.* (2004) compared chromosomes that did and did not differ between chromosomal races in Northern Italy and showed that centromeric regions of rearranged chromosomes showed less introgression than non-rearranged chromosomes. Such a study would be worthwhile on the Madeiran mice, to rule out the possibility that the reduced gene flow associated with proximal markers of metacentric 3.8 is a purely centromeric effect, rather than an impact of the chromosomal rearrangement (Carneiro *et al.*, 2009; Geraldès *et al.*, 2011; Nachman and Payseur, 2012). Such work is currently being carried out, using a suite of centromeric microsatellites that includes three loci per chromosome (total 60 loci: three for each of the 19 autosomes, plus three on the X chromosome).

DATA ARCHIVING

Genotype data have been submitted to Dryad: doi:10.5061/dryad.jv2hd.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ACKNOWLEDGEMENTS

We are grateful to C Bidau, S Gabriel, G Ganem, M Giménez, İ Gündüz, H Hauffe, M López-Fuster, A Nunes, C Palmer, T Panithanarak, C Smadja and J Ventura for contributing house mouse samples to this study. This project benefited from the financial support of the Fundacao para a Ciencia e a Tecnologia (PRAXIS/PCNA/C/BIA/135/96) and SAPIENS (POCTI/BSE/47019/02), the University of York (UK) and the Natural Environment Research Council (NERC).

Albuquerque L, Vieira A (1988). *The Arquipelago of Madeira in the XV-century*. Coleção Atlântica. Secretaria Regional do Turismo e Cultural/Centro de Estudos de História do Atlântico, Região Autónoma da Madeira.

Avice JC (2000). *Phylogeography: The History and Formation of Species*. Harvard University Press: Boston, MA.

Bidau CJ, Giménez MD, Palmer CL, Searle JB (2001). The effects of Robertsonian fusions on chiasma frequency and distribution in the house mouse (*Mus musculus domesticus*) from a hybrid zone in northern Scotland. *Heredity* **87**: 305–313.

Britton-Davidian J, Catalan J, Ramalinho G, Auffray J-C, Nunes AC, Gazave E *et al.* (2005). Chromosomal phylogeny of Robertsonian races of the house mouse on the island of Madeira: testing between alternative mutational processes. *Genet Res* **86**: 171–183.

Britton-Davidian J, Catalan J, Lopez J, Ganem G, Nunes AC, Ramalinho MG *et al.* (2007). Patterns of genic diversity and structure in a species undergoing rapid chromosomal radiation: an allozyme analysis of house mice from the Madeira archipelago. *Heredity* **99**: 432–442.

Britton-Davidian J, Catalan J, Ramalinho M, Ganem G, Auffray J-C, Capela R *et al.* (2000). Rapid chromosomal evolution in island mice. *Nature* **403**: 158.

Busch JD, Waser PM, DeWoody JA (2007). Recent demographic bottlenecks are not accompanied by a genetic signature in banner-tailed kangaroo rats (*Dipodomys spectabilis*). *Mol Ecol* **16**: 2450–2462.

Butlin RK (2005). Recombination and speciation. *Mol Ecol* **14**: 2621–2635.

Carneiro M, Ferrand N, Nachman MW (2009). Recombination and speciation: loci near centromeres are more differentiated than loci near telomeres between subspecies of the European rabbit (*Oryctolagus cuniculus*). *Genetics* **181**: 593–606.

Coyne JA, Orr HA (2004). *Speciation*. Sinauer Associates: Sunderland, MA.

Dallas JF, Bonhomme F, Boursot P, Britton-Davidian J, Bauchau V (1998). Population genetic structure in a Robertsonian race of house mice: evidence from microsatellite polymorphism. *Heredity* **80**: 70–77.

Dallas JF, Dod B, Boursot P, Prager EM, Bonhomme F (1995). Population subdivision and gene flow in Danish house mice. *Mol Ecol* **4**: 311–320.

Davison MT, Akeson EC (1993). Recombination suppression by heterozygous Robertsonian chromosomes in the mouse. *Genetics* **133**: 649–667.

Dumas D, Britton-Davidian J (2002). Chromosomal rearrangements and evolution of recombination: comparison of chiasma distribution patterns in standard and Robertsonian populations of the house mouse. *Genetics* **162**: 1355–1366.

Faria R, Navarro A (2010). Chromosomal speciation revisited: rearranging theory with pieces of evidence. *Trends Ecol Evol* **25**: 660–669.

Förster DW (2007). A molecular study on the evolution of island races of house mice on Madeira. DPhil thesis, University of York.

Förster DW, Gündüz I, Nunes AC, Gabriel S, Ramalinho MG, Mathias ML *et al.* (2009). Molecular insights into the colonization and chromosomal diversification of Madeiran house mice. *Mol Ecol* **18**: 4477–4494.

Franchini P, Castiglia R, Capanna E (2008). Reproductive isolation between chromosomal races of the house mouse *Mus musculus domesticus* in a parapatric contact area revealed by an analysis of multiple unlinked loci. *J Evol Biol* **21**: 502–513.

Franchini P, Colangelo P, Solano E, Capanna E, Verheyen E, Castiglia R (2010). Reduced gene flow at pericentromeric loci in a hybrid zone involving chromosomal races of the house mouse *Mus musculus domesticus*. *Evolution* **64**: 2020–2032.

Geraldes A, Basset P, Smith KL, Nachman MW (2011). Higher differentiation among subspecies of the house mouse (*Mus musculus*) in genomic regions with low recombination. *Mol Ecol* **20**: 4722–4736.

Goodman SJ (1997). *rst calc*: a collection of computer programs for calculating unbiased estimates of genetic differentiation and gene flow from microsatellite data and determining their significance. *Mol Ecol* **6**: 881–885.

Goudet J (2002). FSTAT 2.9.3.2. Lausanne: Switzerland. Available at: <http://www2.unil.ch/popgen/softwares/fstat.htm>.

Goudet J (2005). Hierstat, a package for R to compute and test hierarchical F-statistics. *Mol Ecol Notes* **5**: 184–186.

Gündüz I, Auffray J-C, Britton-Davidian J, Catalan J, Ganem G, Ramalinho MG *et al.* (2001a). Molecular studies on the colonization of the Madeiran archipelago by house mice. *Mol Ecol* **10**: 2023–2029.

Gündüz I, López-Fuster MJ, Ventura J, Searle JB (2001b). Clinal analysis of a chromosomal hybrid zone in the house mouse. *Genet Res* **77**: 41–51.

Hauffe HC, Searle JB (1993). Extreme karyotypic variation in a *Mus musculus domesticus* hybrid zone: the tobacco mouse story revisited. *Evolution* **47**: 1374–1395.

Hauffe HC, Searle JB (1998). Chromosomal heterozygosity and fertility in house mice (*Mus musculus domesticus*) from Northern Italy. *Genetics* **150**: 1143–1154.

Jombart T (2008). adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics* **24**: 1403–1405.

Keller LF, Jeffery KJ, Arcese P, Beaumont MA, Hochachka WM, Smith JNM *et al.* (2001). Immigration and the ephemerality of a natural population bottleneck: evidence from molecular markers. *Proc R Soc Lond Ser B* **268**: 1387–1394.

King M (1993). *Species Evolution. The Role of Chromosome Change*. Cambridge University Press: Cambridge.

Merico V, Pigozzi MI, Esposito A, Merani MS, Garagna S (2003). Meiotic recombination and spermatogenic impairment in *Mus musculus domesticus* carrying multiple simple Robertsonian translocations. *Cytogenet Genome Res* **103**: 321–329.

Nachman MW, Payseur BA (2012). Recombination rate variation and speciation: theoretical predictions and empirical results from rabbits and mice. *Philos Trans R Soc Ser B* **367**: 409–421.

Nachman MW, Searle JB (1995). Why is the mouse karyotype so variable? *Trends Ecol Evol* **10**: 397–402.

Navarro YN, Britton-Davidian J (1989). Genetic structure of insular Mediterranean populations of the house mouse. *Biol J Linn Soc* **36**: 377–390.

Nei M (1978). Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics* **89**: 583–590.

Palmer CL (1997). A chromosomal hybrid zone of the house mouse (*Mus musculus domesticus*) in northern Scotland. DPhil thesis, University of York.

Panithanarak T, Hauffe HC, Dallas JF, Glover A, Ward RG, Searle JB (2004). Linkage-dependent gene flow in a house mouse chromosomal hybrid zone. *Evolution* **58**: 184–192.

Piálék J, Hauffe HC, Searle JB (2005). Chromosomal variation in the house mouse. *Biol J Linn Soc* **84**: 535–563.

Raymond M, Rousset F (1995). GENEPOP, Version 3.1d: population genetics software for exact tests and ecumenism. *J Hered* **86**: 248–249.

Riginos C, Nachman MW (1999). The origin of a Robertsonian chromosomal translocation in house mice inferred from linked microsatellite markers. *Mol Biol Evol* **16**: 1763–1773.

Said K, Saad A, Auffray J-C, Britton-Davidian J (1993). Fertility estimates in the Tunisian all-acrocentric and Robertsonian populations of the house mouse and their chromosomal hybrids. *Heredity* **71**: 532–538.

Schneider S, Kueffer JM, Roessli D, Excoffier L (1997). Arlequin, version 1.1: a software for population genetic data analysis. Genetics and Biometry Laboratory, University of Geneva.

Searle JB (1991). A hybrid zone comprising staggered chromosomal clines in the house mouse (*Mus musculus domesticus*). *Proc R Soc Lond Ser B* **246**: 47–52.

Searle JB (1993). Chromosomal hybrid zones in eutherian mammals. In: Harrison RG (ed.). *Hybrid Zones and the Evolutionary Process*. Oxford University Press: New York, pp 507–532.

Searle JB, Navarro YN, Ganem G (1993). Further studies of a staggered hybrid zone in *Mus musculus domesticus* (the house mouse). *Heredity* **71**: 523–531.

Slatkin M (1995). A measure of population subdivision based on microsatellite allele frequencies. *Genetics* **139**: 457–462.

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

Van Oosterhout C, Hutchinson WF, Wills DPM, Shipley P (2004). Micro-checker: software for identifying and correcting genotyping errors in microsatellite data. *Mol Ecol Notes* **4**: 535–538.

Wallace BMN, Searle JB, Everett CA (1992). Studies on meiosis and gametogenesis in wild house mice (*Mus musculus domesticus*): Male 'simple' heterozygotes and homozygotes from a chromosomal hybrid zone. *Cytogenet Cell Genet* **61**: 211–220.

Wallace BMN, Searle JB, Everett CA (2002). The effect of multiple simple Robertsonian heterozygosity on chromosome pairing and fertility of wild-stock house mice (*Mus musculus domesticus*). *Cytogenet Genome Res* **96**: 276–286.

Weir BS, Cockerham CC (1984). Estimating F-statistics for the analysis of population structure. *Evolution* **38**: 1358–1370.

White TA, Bordewich M, Searle JB (2010). A network approach to study karyotypic evolution: the chromosomal races of the common shrew (*Sorex araneus*) and house mouse (*Mus musculus*) as model systems. *Syst Biol* **59**: 262–276.

Yeh FC, Yang RC, Boyle TBJ, Ye ZH, Mao JX (1997). *popgene, the User-Friendly Shareware for Population Genetic Analysis*. Molecular Biology and Biotechnology Centre, University of Alberta: Canada.

Zenger KR, Richardson BJ, Vachot-Griffin AM (2003). A rapid population expansion retains genetic diversity within European rabbits in Australia. *Mol Ecol* **12**: 789–794.

Supplementary Information accompanies the paper on Heredity website (<http://www.nature.com/hdy>)