The LINNEAN SOCIETY of London BIOLOGICAL Journal of the Linnean Society

Biological Journal of the Linnean Society, 2011, **102**, 471–482. With 5 figures

A molecular characterization of the charismatic Faroe house mouse

ELEANOR P. JONES[1,2]*, JENS-KJELD JENSEN[3], EYÐFINN MAGNUSSEN[4], NOOMI GREGERSEN[4], HEIDI S. HANSEN[4] and JEREMY B. SEARLE[1,5]

[1]*Department of Biology (Area 2), University of York, PO Box 373, York YO10 5YW, UK*
[2]*Population Biology and Conservation Biology, Evolutionary Biology Centre, Uppsala University, Norbyvägen 18 D, SE-752 36 Uppsala, Sweden*
[3]*Í Geilini 37, FO-270 Nólsoy, Faroe Islands*
[4]*University of the Faroe Islands, Faculty of Science and Technology, Noatun 3, FO-100 Tórshavn, Faroe Islands*
[5]*Department of Ecology and Evolutionary Biology, Cornell University, Corson Hall, Ithaca, NY 14853-2701, USA*

Received 25 August 2010; revised 29 September 2010; accepted for publication 29 September 2010

Faroe house mice are a 'classic' system of rapid and dramatic morphological divergence highlighted by J. S. Huxley during the development of the Modern Synthesis. In the present study, we characterize these charismatic mice using modern molecular techniques, examining specimens from all Faroe islands occupied by mice. The aims were to classify the mice within the modern house mouse taxonomy (i.e. as either *Mus musculus domesticus* or *Mus musculus musculus*) using four molecular markers and a morphological feature, and to examine the genetic diversity and possible routes of colonization using mitochondrial (mt) control region DNA sequences and microsatellite data (15 loci). Mice on the most remote islands were characterized as *M. m. domesticus* and exhibited exceptionally low genetic diversity, whereas those on better connected islands were more genetically diverse and had both *M. m. musculus* and *M. m. domesticus* genetic elements, including one population which was morphologically *M. m. musculus*-like. The mtDNA data indicate that the majority of the mice had their origins in south-western Norway (or possibly southern Denmark/northern Germany), and probably arrived with the Vikings, earlier than suggested by Huxley. The *M. m. musculus* genetic component appears to derive from recent mouse immigration from Denmark. © 2011 The Linnean Society of London, Biological Journal of the Linnean Society, 2011, **102**, 471–482.

ADDITIONAL KEYWORDS: hybridization – human commensal – *Mus musculus domesticus* – *Mus musculus musculus* – phylogeography – introgression.

INTRODUCTION

Faroe is an archipelago of 18 islands in the North Atlantic with three introduced terrestrial mammalian species (Bloch, 1982), of which the house mouse, *Mus musculus*, was the first colonist. These inconspicuous animals from a remote archipelago have been the subject of much scientific interest, and were cited by Darwin (1869) as an example of adaptability to extreme environments, and by others as an example of rapid diversification. An early study by Clarke (1904) described the Faroe mice as a distinct subspecies on the basis of their great size, pelage and 'robust habit'. Evans & Vevers (1938) and Degerbøl (1942) reinforced the interest in the mice, as did Huxley (1942), who popularized the Faroe house mice as an example of very rapid speciation (within 250 years) in his contribution to the Modern Synthesis. The studies by Evans & Vevers (1938) and Degerbøl (1942) documented the morphological distinctiveness of various island populations of Faroe mice, and hypothesized

*Corresponding author. E-mail: eleanor.jones@ebc.uu.se

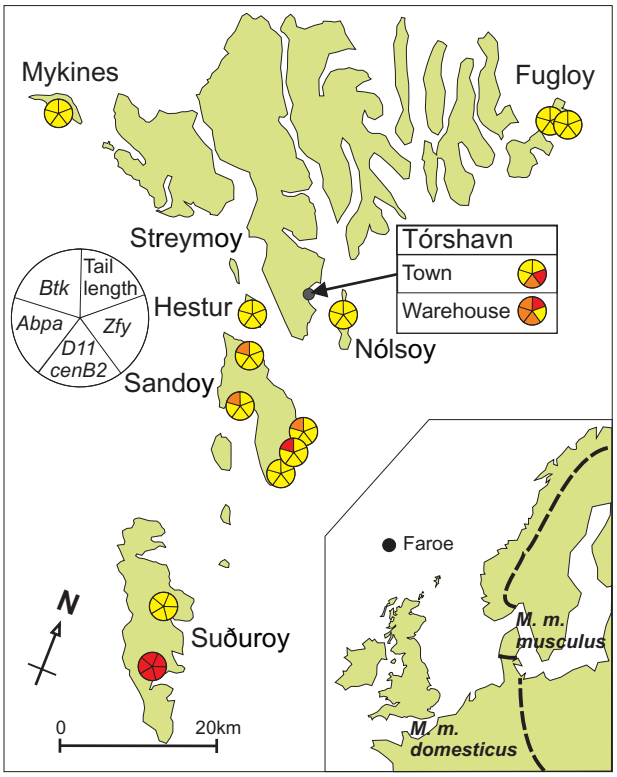

Figure 1. Sample localities and subspecies identification of house mice on Faroe. The diagnostic markers are shown as pies; yellow is *domesticus*-like, red *musculus*-like, orange polymorphic. Inset map shows the distribution of *Mus musculus musculus* and *Mus musculus domesticus* in Northern Europe (*sensu* Božíková *et al*., 2005; Dod, Smadja & Karn, 2005; Jones *et al*., 2010b); the location of Faroe is shown as a black dot.

that their specific morphological traits were adaptive responses to the local environment. More recent studies have more systematically described the differentiation between the mice from the different islands using morphological measurements and allozymes (Berry & Peters, 1977; Berry, Jakobson & Peters, 1978; Davis, 1983), although these differences were attributed to stochastic effects such as founder events rather than to environmental adaptation.

Given the suggested subspecific or specific status of Faroe house mice, we use modern genetic markers to re-examine their taxonomy, in particular considering them in the context of the two currently recognized subspecies in northern Europe, *M. m. musculus* and *M. m. domesticus* (Fig. 1, inset map).

Associated with the determination of the taxonomy of the Faroe house mice is the investigation of their origin. There has been a significant body of work on the phylogeography and colonization history of the house mouse, in part a result of their association with humans. The house mouse global range expansion appears to track aspects of human history such as urbanization and intensity of cultural linkages (Cucchi, Vigne & Auffray, 2005) and human expansion and colonization (Searle *et al*., 2009a,b; Nunome *et al*., 2010). In Faroe, the colonization history of the mice has already been the subject of investigation. Both Berry *et al*. (1978) and Davis (1983) noted morphological similarities between the Faroe and Scottish island mice, suggesting a shared ancestry for these populations. In the present study, we examine the origin of the Faroe mice using the distributions of mitochondrial (mt)DNA control region haplotypes from Faroe in the context of previously published sequences, and use genetic markers to infer the between-island colonization history.

The Faroe mice were viewed as special by Huxley and others because of the pattern of their evolution after colonization, including the emergence of great morphological divergence between the populations on the different islands. In the present study, we investigate the genetic variation among these populations using microsatellite loci, to determine whether the morphological divergence is matched by genetic divergence. Small populations on islands are well known to exhibit low, sometimes very low, levels of heterozygosity (Frankham, 1997). From this, we expect the mice on the smaller, more isolated islands on Faroe to exhibit low genetic diversity.

Therefore, with these various molecular analyses, we are able to provide a new characterization of the Faroe house mouse, providing a framework in which to consider the rapid morphological evolution highlighted so forcefully by Huxley and others.

MATERIAL AND METHODS

Sample collection

House mice were collected from all islands where there were stable populations: Hestur (single site), Streymoy (two sites), Nólsoy, Fugloy (two sites), Mykines (single site), and Sandoy (five sites) (Fig. 1, Table 1). Two mice were obtained from Suðuroy, an island which does not appear to have a stable population, with the mice instead arriving recently with ships (one was caught as it disembarked). The mice from Streymoy came from two locations in a single town (Tórshavn), from houses within the town or from a warehouse that contains imports of animal feed from Jutland in Denmark (Fig. 1, Table 1).

Where whole adult mice were obtained, their pelage was examined, and they were identified to subspecies on the basis of tail length. Adult mice with tail length of 74 mm or above were classified as *M. m. domesticus*; those below 73 mm as *M. m. musculus*; and intermediates were left unclassified (*sensu* Prager

Table 1. Sample characteristics for house mice from Faroe

Location	Island	Year	*Btk**	*Zfy2**	*Abpa**	*D11 cenB2**	Tail†	mtDNA‡
Hestur	Hestur	2006	dom (10)	dom (2)	dom (10)	dom (10)	dom 80.8 (10)	U47455 (9)
Hattarvík	Fugloy	2001–2007	dom (4)	dom (2)	dom (4)	dom (4)	dom 75.0 (3)	U47455 (4)
Kirkja	Fugloy	2001–2007	dom (6)	dom (2)	dom (6)	dom (6)	dom 77.3 (2)	U47455 (5)
Dalur	Sandoy	2006	dom (1)	dom (1)	dom (1)	dom (1)	dom 83.0 (1)	FSa2 (1)
Húsavík	Sandoy	2006	musc (1)	dom (1)	dom (1)	dom (1)	dom 74.1 (1)	FSa2 (1)
Skálavík	Sandoy	2006	musc (1), het (1), dom (3)	dom (3)	dom (5)	dom (5)	dom 74.6 (4)	FSa1 (3), FSa2 (2)
Skopun	Sandoy	2006	musc (1), dom (3)	dom (4)	dom (4)	dom (4)	dom 85.3 (8)	FSa1 (1), FSa2 (3)
Sandur	Sandoy	2005–2006	musc (1), het (2), dom (2)	dom (2)	dom (5)	dom (5)	dom 84.2 (5)	FSa1 (3), FSa2 (2)
Nólsoy	Nólsoy	1984–2003	dom (11)	dom (8)	dom (11)	dom (11)	dom 82.9 (17)	U47455 (10)
Tórshavn town	Streymoy	2000–2005	dom (8)	musc (5)	dom (8)	dom (7), het (1)	dom 74.6 (4)	U47455 (7)
Tórshavn warehouse	Streymoy	2006	musc (5), het (2), dom (1)	dom (5)	musc (2), het (3), dom (3)	het (2), dom (6)	musc 72.9 (10)	AM182650 (7)
Mykines	Mykines	1999	dom (10)	dom (5)	dom (10)	dom (10)	dom 84.8 (9)	Made27 (2), U47455 (8)
Tvøroyri	Suðuroy	2005–2006	musc (1), dom (1)	musc (1), dom (1)	musc (1), dom (1)	musc (1), dom (1)	NA	U47456 (1), FSu1 (1)

*The *Btk*, *Zfy2*, *Abpa*, and *D11 cenB2* markers are scored as 'musc': *Mus musculus musculus*-like, 'dom': *Mus musculus domesticus*-like or 'het': heterozygous for the marker. The number of individuals that score for each are given in brackets.

†The mean tail length for the population is assigned either as 'musc': *M. m. musculus*-like or 'dom': *M. m. domesticus*-like. Number of individuals scored is given in brackets.

‡'mtDNA' lists the haplotypes found in each locality with the number of individuals that share the haplotype given in parentheses.

© 2011 The Linnean Society of London, *Biological Journal of the Linnean Society*, 2011, **102**, 471–482

et al., 1993). This measurement has been reliably used in north Germany and Scandinavia (Prager *et al*., 1993; Jones *et al*., 2010b). The shape of the fossa mesopterygoidea, as described in Degerbøl (1942), was examined on all available adult skulls.

Molecular methods

DNA was extracted from tissue using Qiagen Blood and Tissue kits. The mtDNA control region and adjacent tRNA sequences were amplified by polymerase chain reaction (PCR) between positions 15 193 and 16 492 of Bibb *et al*. (1981) for selected samples using primers H2228 and L15774 (Searle *et al*., 2009b). Sequences were shortened to positions 15 424 and 16 276 and aligned with previously published house mouse sequences using BIOEDIT, version 7.0.9 (Hall, 1999). Sequences have been deposited in GenBank under accession numbers HQ326585–HQ326587. Fifty percent majority-rule consensus trees were generated using MrBayes (Ronquist & Huelsenbeck, 2003), which estimates the prior probability of trees using Bayesian methods and Markov chain Monte Carlo simulations. Run conditions were six million generations of two independent runs of five chains (four heated, one cold) using an incremental heating parameter of 0.01; the burn-in was determined from the convergence of the two runs using the PSRF statistics. The analysis was run with gamma-distributed rates across sites and variable transition and transversion rates, as determined by JMODELTEST, version 0.1.1 (Posada, 2008). Nucleotide diversity (π) and haplotype diversity (H_D) (Nei, 1987) were calculated for the Faroe mtDNA sequences (excluding those from Suðuroy) using DNAsp, version 4.20.2 (Rozas *et al*., 2003).

Individuals were typed for four genetic markers which reliably differentiate between *M. m. musculus* and *M. m. domesticus*: *Btk* on the X chromosome, *Zfy2* on the Y chromosome, *Abpa* on chromosome 7, and *D11 cenB2* on chromosome 11 (Searle *et al*., 2009a).

Fifteen centromeric microsatellite loci on different chromosomes were amplified by PCR using the Qiagen Multiplex PCR Kit: D1Mit64, D2Mit1, D3Mit117, D4Mit103, D5Mit145, D8Mit58, D9Mit218, D10Mit188, D12Mit145, D13Mit153, D15Mit12, D16Mit2, D17Mit19, D18Mit116, D19Mit150 (primer details available at: http://www.ncbi.nlm.nih.gov/). The fluorescent-labelled fragments were scored on an ABI 3130 (Applied Biosystems) and alleles assigned using GENEMAPPER, version 3.7.

The mean number of microsatellite alleles per locus and observed (Levene, 1949) and expected heterozygosities (Nei, 1978) were calculated per population using the POPGENE, version 1.32 (Yeh *et al*., 1997). Deviations from the Hardy–Weinburg equilibrium per locus and per population were assessed in GENEPOP (Raymond & Rousset, 1995); run parameters were 5000 batches of 20 000 iterations. The robustness of the population clustering (defined by geographic origin) was assessed using the BAPS (Corander *et al*., 2008) and STRUCTURE, version 2.2 (Pritchard, Stephens & Donnelly, 2000) software. BAPS was run with between 0 and 20 populations, STRUCTURE with ten independent runs for $K = 1$ to 12 with 10^6 MCMC iterations, using a burn-in of 10^5 and an admixture and correlated allele frequencies model. The STRUCTURE average log-likelihoods and standard deviation were calculated for each value of K. Among population differentiation was assessed using pairwise F_{ST} and R_{ST} values, using Weir and Cockerham's θ (1984; calculated in FSTAT, version 2.9.3.2, Goudet, 2001) and ρ (RST CALC; Goodman, 1997). P-values were estimated from 10 000 permutations. Holm's correction was used for multiple tests (Aickin & Gensler, 1996). A tree based on the genetic divergence between populations was created using Cavalli-Sforza & Edwards' (1967) chord distance, D_C, calculated in the software POPULATIONS (O. Langella; available at http://www.bioinformatics.org/project/?group_id=84). GENETIX, version 4.05 (Belkhir *et al*., 1999) was used to visualize a factorial correspondence analysis plot of the multilocus relationship of individuals of different populations.

RESULTS

The mice from Nólsoy, Fugloy, Sandoy, Hestur, Mykines, and Tórshavn town resembled *M. m. domesticus* morphologically, with a dull brown pelage and no obvious line of demarcation between the dorsal and ventral fur (Marshall & Sage, 1981) and tail lengths of 74 mm and above. The mice from the warehouse in Tórshavn were more *M. m. musculus*-like, with paler bellies, an obvious line of demarcation between dorsal and ventral fur, and tail lengths of below 73 mm (Table 1). Three of the seven mice examined from Fugloy had a white belly spot, as recorded in previous studies (Evans & Vevers, 1938; Degerbøl, 1942). The fossa mesopterygoidea of mice from Mykines narrowed to a point (five skulls examined), as noted by Degerbøl (1942), a feature that we did not find on any other island (79 skulls examined).

The mice from Fugloy, Hestur, Mykines, and Nólsoy were *M. m. domesticus*-like for four diagnostic nuclear markers (*Abpa*, *D11 cenB2*, *Btk*, and *Zfy2*; Fig. 1, Table 1). The mice from Sandoy were *M. m. domesticus*-like for *Abpa*, *D11 cenB2*, and *Zfy2* but had both *M. m. domesticus* and *M. m. musculus* alleles for the X chromosome marker, *Btk,* in all heterozygous and homozygous combinations; both alleles were widespread, appearing in three of the five settlements sampled. The samples from Tórshavn

town were *M. m. domesticus*-like except for the Y chromosome marker, *Zfy2*, whereas the mice from Tórshavn warehouse had both *M. m. musculus* and *M. m. domesticus* alleles for *Abpa*, *D11 cenB2* and, *Btk*, and only *M. m. domesticus* alleles for *Zfy2*. Of the two mice caught on Suðuroy, one was *M. m. musculus* for all markers, whereas the other was *M. m. domesticus* for all markers.

A total of 70 mtDNA control region sequences were obtained, distributed among seven haplotypes (Table 1). Haplotype diversity ($H_D \pm SD$) was 0.599 ± 0.0002, and nucleotide diversity ($\pi \pm SD$) 0.00434 ± 0.00066. Inspection of the sequences showed that they were *M. m. domesticus*-like and the Bayesian analysis therefore utilized published *M. m. domesticus* sequences (a list of references for the sequences used is available in Jones *et al*., 2010a, plus additional sequences from Jones *et al*., 2010b) and two *M. m. musculus* sequences (GenBank accession numbers U47504 and U47532) and three *M. m. castaneus* sequences (ED108342, AJ286322, and AF088879) as out-groups. As is usual for phylogenetic trees of house mouse control region sequences, the support on the tree was not high (Fig. 3; Rajabi-Maham, Orth & Bonhomme, 2008). The sequences from Fugloy, Mykines, Nólsoy, Streymoy, Hestur, and Suðuroy belonged to a single clade, referred to as clade D1 (Jones *et al*., 2010a), represented by five haplotypes; those from Fugloy, Nólsoy, Tórshavn town, and the majority from Mykines belong to a single haplotype within that clade, U47455 (Fig. 2). The Sandoy sequences were two haplotypes in clade E, FSa1 and 2 (Figs 2, 3).

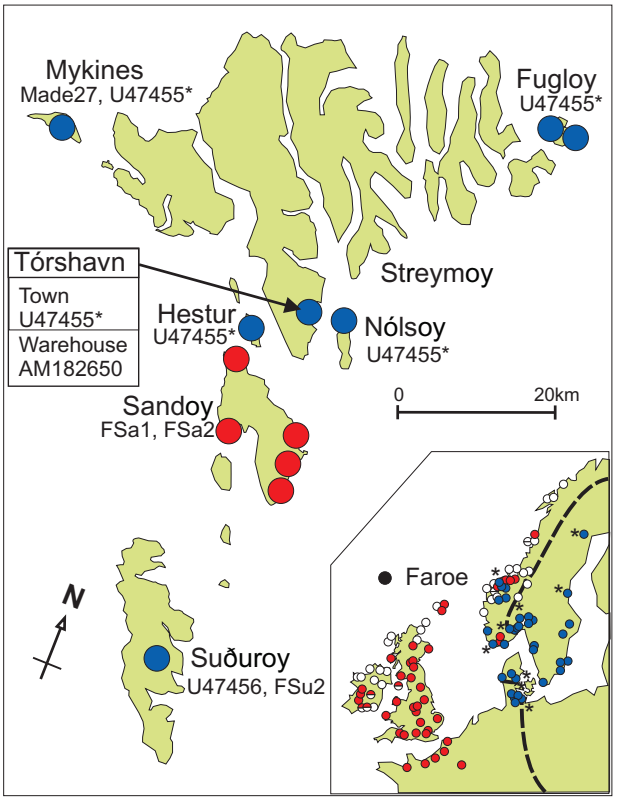

Figure 2. Distribution of the different haplotypes and mitochondrial DNA lineages of house mouse found on Faroe (colours are the same as the phylogenetic tree in Fig. 3). Inset map shows the distribution of the featured lineages in Northern Europe; locations where haplotype U47455 is found are highlighted with an asterisk (*).

Microsatellites

A total of 79 individuals were scored for fifteen microsatellite loci from seven populations in Faroe (see Supporting information, Table S1). Two *M. m. musculus* populations (from Burg-auf-Fehmarn in Germany and Berg in south-east Norway, Prager *et al*., 1993; Jones *et al*., 2010b) were included in the analysis as outgroups. Two of the fifteen microsatellites (D5Mit145 and D16Mit2) failed to amplify reliably and were excluded from further analysis. The assignment of individuals to genetic populations was assessed in BAPS and STRUCTURE and gave very similar results to the geographic populations. BAPS allocated the individuals to ten populations, identical to the geographic populations except that one individual from Berg in Norway was assigned its own population. The optimum value for *K* from the STRUCTURE output, estimated from the average log-likelihood scores and visual output (see Supporting information, Fig. S1) was nine or ten populations. As with the BAPS output, the populations as defined by STRUCTURE were the same as the geographic populations. For both analyses, the house mice from Tórshavn were assigned to two populations: Tórshavn town and Tórshavn warehouse.

The number of alleles per locus over all individuals ranged from five (D13mit153) to seventeen (D18mit116). Many of the populations sampled from Faroe had very little variation: Fugloy and Mykines, the most geographically remote islands with small human populations, had house mouse populations which were highly monomorphic for all microsatellite loci tested (the Mykines mouse population was monomorphic for all fifteen markers scored, including the two markers excluded from subsequent analysis, the Fugloy population monomorphic for fourteen of the fifteen markers), whereas the island of Hestur, which again has a very small human population, had a house mouse population which was monomorphic for all but two microsatellite loci scored (see Supporting information, Table S1). Other diversity measures reflect this lack of diversity (Table 2). The populations from the islands of

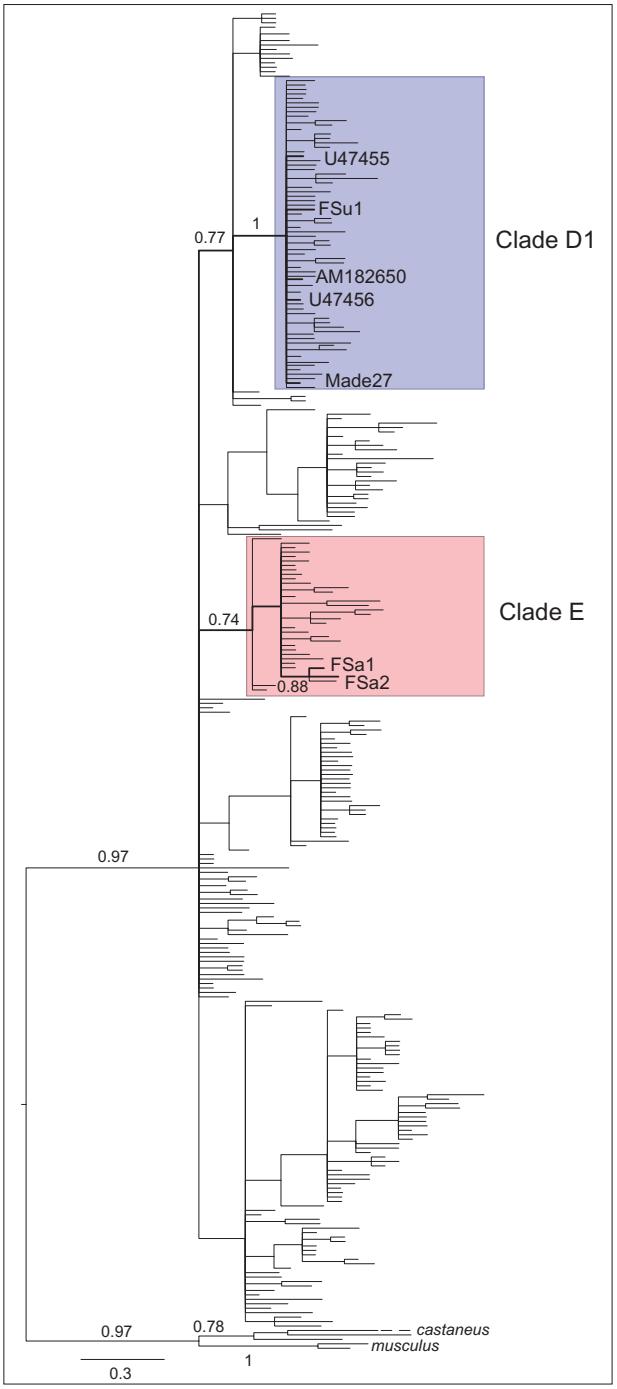

Figure 3. Bayesian inference phylogenetic tree of published mitochondrial DNA D-loop sequences and the Faroe sequences obtained during the study (named). Clade colours and highlighted taxa as in Fig. 2.

Nólsoy and Sandoy had higher levels of genetic diversity, whereas the two populations in the town of Tórshavn were the most diverse of all. The number of inhabitants per island/town is given in Table 2 as a measure of the house mouse habitat availability, and shows a significant correlation with the genetic diversity measures (A: $r = 0.939$, $P < 0.002$, H_E: $r = 0.936$, $P < 0.002$ and H_O: $r = 0.942$, $P < 0.002$). Divergence between the populations was also very high because even the near-monomorphic populations did not share the same alleles; the values for the pairwise divergence measures R_{ST} and F_{ST} are given in Table 3.

The tree based on Cavalli-Sforza & Edwards' (1967) chord distance, D_C, is shown in Fig. 4. There is a clear division between the *M. m. musculus* populations included as outgroups (Berg in Norway, and Burg-auf-Fehmarn in Germany) and the *M. m. domesticus* Faroe populations, with the exception of the population from the warehouse in Tórshavn, which branches between the *M. m. musculus* and *M. m. domesticus* populations. There is little bootstrap support elsewhere on the tree. Factorial correspondence analysis showed that the greatest axis of variation was a result of the difference between the two subspecies, with the *M. m. musculus* mice at one end of the axis, *M. m. domesticus* at the other, and the population of mice from Tórshavn warehouse lying between the two (Fig. 5). The second, third, and fourth axes of variation largely reflected differentiation of the Faroe mice.

DISCUSSION

TAXONOMY

The Faroe mice are predominantly *M. m. domesticus*, as determined by the mtDNA sequences, the four diagnostic nuclear markers, the microsatellites, and the tail length data. However, the mice on the islands of Suðuroy, Sandoy, and in the town of Tórshavn had genetic traces of *M. m. musculus*. The two mice from Suðuroy, one pure *M. m. musculus* and one pure *M. m. domesticus*, were likely very recent immigrants. The presence of *M. m. musculus* X chromosome alleles in otherwise *M. m. domesticus* mice on Sandoy is more surprising, given the suggested incompatibility between the X chromosome and the genomic background of the alternate subspecies (Geraldes *et al.*, 2008). This incompatibility is most clearly shown at hybrid zones, where the X chromosome introgresses far less than the autosomes (Dod *et al.*, 1993) and is believed to be under greater selective pressure (Macholán *et al.*, 2007).

For the Tórshavn mice, the town mice were *M. m. domesticus*-like morphologically and for most markers but were *M. m. musculus*-like for the Y chromosome; one individual was also heterozygous for *D11 cenB2*. The mice from the warehouse (which imports

Table 2. Per population genetic diversity measures for the microsatellite data

	N	*A*	SD	H_O	SD	H_E	SD	Human population
Fugloy	12	1.1538	*0.5547*	0.0192	*0.0693*	0.0222	*0.0799*	45
Hestur	11	1.2308	*0.5991*	0.0979	*0.2390*	0.0750	*0.1844*	39
Mykines	15	1.0000	*0.0000*	0.0000	*0.0000*	0.0000	*0.0000*	22
Nólsoy	11	2.0000	*1.2910*	0.2049	*0.2209*	0.2377	*0.2386*	256
Sandoy	14	1.6154	*0.7679*	0.1795	*0.2312*	0.1872	*0.2482*	1 317
Tórshavn town	8	3.2308	*0.9268*	0.4615	*0.2904*	0.4946	*0.1930*	12 600
Tórshavn warehouse	8	3.0000	*0.5774*	0.5673	*0.1739*	0.5625	*0.1092*	
Berg, Norway	8	2.4615	*1.0500*	0.3027	*0.2777*	0.3334	*0.2568*	
Burg-auf-Fehmarn, Germany	15	4.0769	*1.8467*	0.5413	*0.3044*	0.4899	*0.2612*	

N, number of individuals scored; *A*, mean number of alleles per locus; H_O, the observed heterozygosity; H_E, unbiased expected heterozygosity.

Table 3. Pairwise F_{ST} (above the diagonal) and R_{ST} (below the diagonal) values for between population comparisons among Faroe populations and populations from Berg in Norway and Burg-auf-Fehmarn in Germany

	Fugloy	Hestur	Mykines	Nólsoy	Sandoy	Tórs. town	Tórs. wareh.	Berg	Burg-auf-F.
Fugloy	0	0.9154	0.9849	0.7660	0.8417	0.6647	0.6806	0.8319	0.6715
Hestur	0.9946	0	0.9644	0.7063	0.8181	0.4624	0.5885	0.8073	0.6817
Mykines	0.9986	0.9967	0	0.8681	0.8789	0.7662	0.7384	0.8634	0.7027
Nólsoy	0.9064	0.9468	0.9362	0	0.6983	0.3920	0.4971	0.6889	0.5891
Sandoy	0.9858	0.9831	0.9909	0.9494	0	0.5184	0.5957	0.6714	0.5949
Tórshavn town	0.5742	0.6704	0.7445	0.6583	0.6795	0	0.2487	0.5266	0.4392
Tórshavn warehouse	0.4917	0.6869	0.5423	0.6496	0.6362	0.2995	0	0.4044	0.3221
Berg	0.8472	0.8887	0.8634	0.8602	0.7916	0.6696	0.5037	0	0.2864
Burg-auf-Fehmarn, Germany	0.7752	0.8280	0.7464	0.7798	0.7956	0.5937	0.4047	0.4488	0

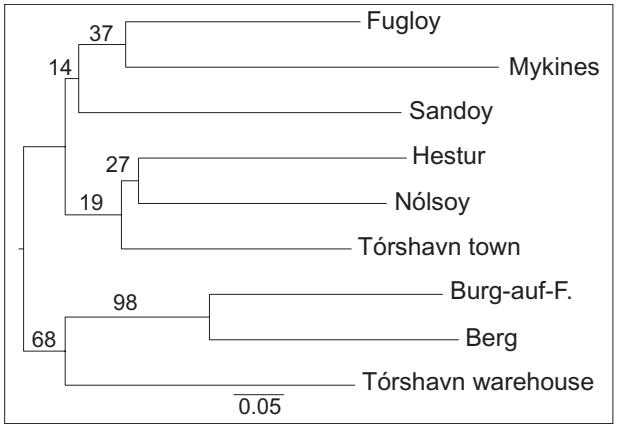

Figure 4. Microsatellite tree based on Cavalli-Sforza & Edwards' (1967) chord distance.

agricultural foodstuffs from Denmark) were hybrids of the two subspecies to a considerable degree, being *M. m. musculus*-like morphologically and polymorphic for all the nuclear markers except *Zfy2*, for which they were *M. m. domesticus*-like. The warehouse mice also had *M. m. domesticus*-like mtDNA haplotypes, a feature that they share with Danish *M. m. musculus* mice (Prager *et al.*, 1993).

The widespread presence of genetic material from *M. m. musculus* in otherwise *M. m. domesticus* house mice on some Faroe islands is of interest because the *M. m. musculus*/ *M. m. domesticus* hybrid system is one of the best studied (Božíková *et al.*, 2005; Raufaste *et al.*, 2005; Macholán *et al.*, 2007; Teeter *et al.*, 2010) and has provided great insights into the interaction of distinct genomes within a species of mammal. The apparently stable hybrid forms on Sandoy, where the *M. m. musculus* X chromosome is widespread in *M. m. domesticus* mice, and the (presumably) more dynamic hybridization occurring in Tórshavn, add more information to this well-studied system. Studies have shown that cryptic hybridization can occur between the *M. musculus* subspecies, particularly in relatively recently derived populations (Orth *et al.*, 1998; Searle *et al.*, 2009a; Jones *et al.*, 2010b).

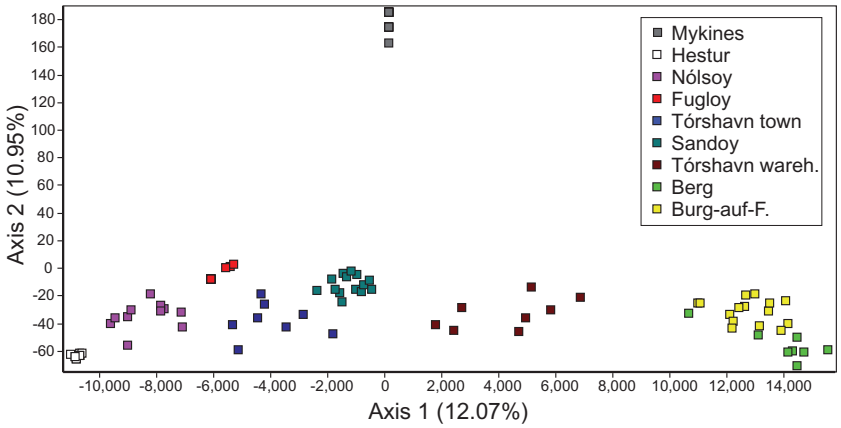

Figure 5. Factorial correspondence analysis of microsatellite data.

Genetic diversity

The house mice in Faroe have low levels of genetic diversity, both for mtDNA sequences and microsatellites. For the mtDNA, there were only five haplotypes from 68 sequences (excluding the two Suðuroy mice), many monomorphic populations and a low haplotype diversity ($H_D = 0.599$) compared to house mice elsewhere: 0.896 for Norway (Jones *et al*., 2010b), 0.955 for Great Britain and Ireland (Searle *et al*., 2009b), and 0.819 for Bulgaria (Vanlerberghe *et al*., 1988). Nucleotide diversity, $\pi = 0.0043$, was low compared to 0.0082 for Norway (Jones *et al*., 2010b), 0.0069 for Great Britain and Ireland (Searle *et al*., 2009b), and 0.0066 for Bulgaria (Vanlerberghe *et al*., 1988), although the Madeiran archipelago (0.0015; Förster *et al*., 2009) had a lower value. This low mtDNA diversity mirrors that of the Faroe human population (Als *et al*., 2006).

There was also little microsatellite variation as reflected by mean numbers of alleles per locus and heterozygosity values (Table 2). The most remote islands with the smallest human populations had exceptionally low levels of genetic diversity in mice, with one (Mykines) being monomorphic for all loci scored. This lack of genetic diversity matches that found for allozymes by Berry & Peters (1977), with entirely homozygous populations on the island of Fugloy and almost entirely homozygous populations on Mykines. However, considering microsatellites, the genetic diversity values for the Tórshavn populations (observed heterozygosity 0.46 and 0.57) in the present study are similar to those found in other wild mouse populations: 0.44–0.70 in Belgium (Dallas *et al*., 1998), 0.56–0.67 in Madeira, and 0.47–0.61 in Denmark (Förster, 2007)].

Colonization history and population processes are important because the islands that are more remote and have little house mouse habitat (Mykines, Hestur, Fugloy) have genetically less diverse mice, whereas the better connected islands with more habitat have mice with greater genetic diversity. In the context of Faroe, Sandoy is a large island with much of the agricultural land and receives large ferries, whereas Nólsoy is well connected to the town of Tórshavn by a frequent ferry service. Tórshavn itself is a relatively large town that receives overseas shipping, increasing the likelihood of immigrant mice arriving.

Colonization and population dynamics

The earliest introduction of the house mouse to Faroe is likely to date to the main human colonization of Faroe by Norwegian Vikings around 800 AD (Edwards, 2005), coming from two principal areas: first, the region of Norway between Sogn, Hordaland and East Agder and, second, from Northern Scotland, Orkney, Shetland or Ireland (Marcus, 1980; Arge *et al*., 2005). There is also a controversial suggestion that the islands were first colonized by Irish seafarers (Arge *et al*., 2005; Edwards, 2005). The later history of Faroe reflects British, Dutch, German, and Danish interests (Marcus, 1980).

Relating the mtDNA haplotype data to the possible routes of colonization of the house mice, the mice on Fugloy, Mykines, Nólsoy, Hestur, and Tórshavn town all belong to a single haplotype from clade D, U47455 (or to a haplotype a single mutation from it, Made27), which is found elsewhere in *M. m. domesticus* mice in south-western Norway and Northern Germany (also found in *M. m. musculus* mice in Sweden, Finland, and Denmark, Prager *et al*., 1993). In Norway, U47455 is found in Hordaland and East Agder, the suggested region of origin for the human settlers of Faroe. The congruence of the mouse mtDNA data from Faroe and the putative human origin is striking,

although the mtDNA data alone could not confirm that this is where the mice arrived from, particularly because relevant haplotypes may be present in unsampled areas. It is likely that the mice would have been inadvertently transported to the islands with livestock, in accordance with the findings from contemporary Faroe Island cattle, which are found to be most closely related to old breed cattle from Western Norway (Li *et al*., 2005).

The mtDNA clade found on Sandoy (clade E) is widespread in Shetland and the British mainland, and present in parts of Norway, Germany, and Denmark. These mice may not only represent the second main route of human immigration from the British Isles identified by historians, but also reflect more recent arrivals from any of these places. The two mice of different subspecies caught on the island of Suðuroy were clearly recent arrivals from different locations. Partially on the basis of a shared morphological character (the narrowing of the fossa mesopterygoidea), it has been suggested that the St Kilda and Faroe house mice share a common origin (Berry *et al*., 1978); however, they belong to different mtDNA clades (Jones *et al*., 2010a), making this unlikely.

It has been suggested that mice arrived in Faroe around 1670 (Huxley, 1942; Matthews, 1952; Berry & Scriven, 2005) based on the first reference to Faroe house mice (Debes, 1676). No other external data are available to date the arrival of the house mouse in Faroe but, because house mouse fossils have been found in the earliest archaeological deposits in Iceland (from 871–940 AD; McGovern *et al*., 2006), colonized by humans in a similar manner to Faroe, it is probable that the Faroe mice would have arrived at a similarly early date. Because house mouse populations in similar geographic contexts can persist and thrive, even in the absence of human populations (e.g. on the island of Kerguelen in the South Atlantic; Hardouin *et al*., 2010), it is likely that the early founding populations in Faroe will have been able to persist to the modern day, also considering the absence of competing species. As an indication of this, although over a far shorter time frame, the distinctive white belly spot found on the Fugloy mice in the present study and studies dating back to 1938 (Evans & Vevers, 1938) suggests that populations can be stable for at least 70 years.

The mice are highly unlikely to have pre-Viking Irish origin, as has sometimes been suggested (Bloch, 1999), because the mice in Ireland belong to mtDNA clade not found in Faroe (Searle *et al*., 2009b; Jones *et al*., 2010a), with the exception of the Sandoy mice, which belong to a clade found in south western Ireland (Searle *et al*., 2009b; Jones *et al*., 2010a). If the mice were the result of a later, Medieval, colonization event, it would be expected that the mtDNA sequences would be similar to those found around Bergen (which they are not; Jones *et al*., 2010b) because the trade between Norway and Faroe from the fourteenth century AD went via that port. Alternatively, if the mice were of Danish origin, once Faroe became part of Denmark, we would expect the mice to be more *M. m. musculus*-like (the majority of Danish ports lie in the area occupied by *M. m. musculus*). In general they are not, although the mice from Tórshavn (particularly Tórshavn warehouse) do have a strong *M. m. musculus* genetic component, which likely reflects recent mouse arrivals with animal feed from Denmark, as found in the warehouse. If the Tórshavn *M. m. musculus* genetic component is recent, the same is also reasonably argued for the Sandoy mice, although there is no firm confirmation of this.

Previous studies have considered the sequence of colonization among the Faroe islands. The mtDNA data do not provide sufficient resolution to contribute to this, other than that the current population on Sandoy is, at least in part, the product of a colonization process separate from that of the other islands, and that Fugloy, Mykines, Nólsoy, Hestur, and Tórshavn town were likely originally the product of a single colonization event.

Although the microsatellite and morphological diversification of the Faroe house mice from their mainland counterparts remains impressive (Berry & Peters, 1977; Berry *et al*., 1978; Davis, 1983), the time in which this has occurred is probably closer to 1000 years than to the 250 years popularized by Huxley (1942). The mechanisms are likely to be a combination of founder events, genetic drift, inbreeding, and selection, as suggested in earlier studies (Degerbøl, 1942; Berry *et al*., 1978).

In summary, it appears most likely that the original house mice to arrive on Faroe came with Vikings from southern Norway (or possibly from northern Germany or Denmark south of the *M. m. musculus*/ *M. m. domesticus* hybrid zone) and were *M. m. domesticus* mice bearing the haplotype U47455. These house mouse populations remained isolated from each other, with the combination of a low population size and isolation giving rise to extremely low genetic diversity, as measured by the microsatellite and mtDNA data. More recently, stronger links with Denmark led to *M. m. musculus* arriving in the bigger ports at Tórshavn on Streymoy and in Sandoy, leading to a *M. m. musculus* input into the otherwise *M. m. domesticus* mice. This input has not managed to spread to the more remote islands, likely because of the small opportunities for the mice to be transported to and from the islands, and the resistance of established populations of house mice to more recent arrivals (Hardouin *et al*., 2010). Remnants of the original

founding population, which remain exclusively *M. m. domesticus* and carry haplotype U47455, apparently persist in Fugloy, Mykines, Nólsoy, Hestur and, to a lesser extent, Tórshavn town. The island of Sandoy was colonized, at least in part, by a different event, with mice arriving from somewhere in the British Isles, Norway or Denmark, although it is not possible to say when.

The house mice in Faroe remain extraordinary; the remarkable between island morphological divergence recorded in previous studies is matched by extremely high inter-island genetic divergence in the microsatellite data reported in the present study. Although we have updated the study of these mice into the microsatellite and mtDNA era, elucidating the colonization history and documenting the hybridization between subspecies, it is likely that these bizarre huge island mice will again be the focus of attention in the genomic era.

ACKNOWLEDGEMENTS

We are grateful to the various people from Faroe who aided in the collection of specimens. E.P.J. is partially funded by the University of York. Part of this work was carried out at the Computational Biology Service Unit, Cornell University, partially funded by Microsoft Corporation.

REFERENCES

Aickin M, Gensler H. 1996. Adjusting for multiple testing when reporting research results: the Bonferroni vs Holm methods. *American Journal of Public Health* **86:** 726–728.

Als TD, Jorgensen TD, Børglum AD, Petersen PA, Mors O, Wang AG. 2006. Highly discrepant proportions of female and male Scandinavian and British Isles ancestry within the isolated population of the Faroe Islands. *European Journal of Human Genetics* **14:** 497–504.

Arge SV, Sveinbjarnardóttir G, Edwards KJ, Buckland PC. 2005. Viking and Medieval settlement in the Faroes: people, place and environment. *Human Ecology* **33:** 597–620.

Belkhir K, Borsa P, Goudet J, Bonhomme F. 1999. *GENETIX: logiciel sous Windows pour la génétique des populations*. Montpellier: Laboratoire Génome et Populations, CNRS-UPR, Université de Montpellier II.

Berry RJ, Peters J. 1977. Heterogeneous heterozygosities in *Mus musculus* populations. *Proceedings of the Royal Society of London Series B, Biological Sciences* **197:** 485–503.

Berry RJ, Scriven PN. 2005. The house mouse: a model and motor for evolutionary understanding. *Biological Journal of the Linnean Society* **84:** 335–347.

Berry RJ, Jakobson ME, Peters J. 1978. The house mice of the Faroe Islands: a study in microdifferentiation. *Journal of Zoology* **185:** 73–92.

Bibb MJ, Van Etten RA, Wright CT, Walberg MW, Clayton DA. 1981. Sequence and gene organization of mouse mitochondrial DNA. *Cell* **26:** 167–180.

Bloch D. 1982. Animal life on the Faeroe Islands. In: Rutherford GK, ed. *The physical environment of the Faeroe Islands*. The Hague: Dr W. Junk, 53–68.

Bloch D. 1999. *Villini súgdjór í Útnorðri*. Tórshavn: Føroya Skúlabókagrunnur.

Božíková E, Munclinger P, Teeter KC, Tucker PK, Macholán M, Piálek J. 2005. Mitochondrial DNA in the hybrid zone between *Mus musculus musculus* and *Mus musculus domesticus*: a comparison of two transects. *Biological Journal of the Linnean Society* **84:** 363–378.

Cavalli-Sforza L, Edwards AWF. 1967. Phylogenetic analysis: models and estimation procedures. *Evolution* **32:** 550–570.

Clarke WE. 1904. On some forms of *Mus musculus*, Linn., with description of a new subspecies from the Faeroe islands. *Proceedings of the Royal Physical Society, Edinburgh* **15:** 160–167.

Corander J, Marttinen P, Sirén J, Tang J. 2008. Enhanced Bayesian modelling in BAPS software for learning genetic structures of populations. *BMC Bioinformatics* **9:** 539.

Cucchi T, Vigne J-D, Auffray J-C. 2005. First occurrence of the house mouse (*Mus musculus domesticus* Schwarz and Schwarz, 1943) in the Western Mediterranean: a zooarchaeological revision of subfossil occurrences. *Biological Journal of the Linnean Society* **84:** 429–445.

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.

Darwin CR. 1869. *On the origin of species by means of natural selection*, 5th edn. London: John Murray.

Davis SJM. 1983. Morphometric variation of populations of house mice *Mus domesticus* in Britain and Faroe. *Journal of Zoology* **199:** 521–534.

Debes LJ. 1676. *Færoæ, and Færoæ Referata: that is, a description of the islands and inhabitants of Faeroe*. London: William Hes. (English translation).

Degerbøl M. 1942. Mammalia. In: Jensen AS *et al.*, ed. *Zoology of the Faroes*. Copenhagen: Horst, 1–133.

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. 2005. Testing for selection on the androgen-binding protein in the Danish mouse hybrid zone. *Biological Journal of the Linnean Society* **84:** 447–459.

Edwards KJ. 2005. 'On the windy edge of nothing': a historical human ecology of the Faroe Islands. *Human Ecology* **33:** 585–596.

Evans FC, Vevers HG. 1938. Notes on the biology of the Faeroe mouse (*Mus musculus faeroensis*). *Journal of Animal Ecology* **7:** 290–297.

Förster DW, Gündüz İ, Nunes AC, Gabriel S, Ramalhinho MG, Mathias ML, Britton-Davidian J, Searle

JB. 2009. Molecular insights into the colonization and chromosomal diversification of Madeiran house mice. *Molecular Ecology* **18:** 4477–4494.

Förster DWG. 2007. A molecular study on the evolution of island races of house mice on the island of Madeira. PhD thesis, University of York.

Frankham R. 1997. Do island populations have less genetic variation than mainland populations? *Heredity* **78:** 311–327.

Geraldes A, Basset P, Gibson B, Smith KL, Harr B, Yu H-T, Bulatova N, Ziv Y, Nachman MW. 2008. Inferring the history of speciation in house mice from autosomal, X-linked, Y-linked and mitochondrial genes. *Molecular Ecology* **17:** 5349–5363.

Goodman SJ. 1997. RSt Calc: a collection of computer programs for calculating estimates of genetic differentiation from microsatellite data and determining their significance. *Molecular Ecology* **6:** 881–885.

Goudet J. 2001. *FSTAT, a program to estimate and test gene diversities and fixation indices*, Version 2.9.3. Lausanne: University of Lausanne.

Hall TA. 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Service* **41:** 95–98.

Hardouin EA, Chapuis J-L, Stevens MI, van Vuuren BJ, Quillfeldt P, Scavetta RJ, Teschke M, Tautz D. 2010. House mouse colonization patterns on the sub-Antarctic Kerguelen Archipelago suggest singular primary invasions and resilience against re-invasion. *BMC Evolutionary Biology* **10:** 325–339.

Huxley JS. 1942. *Evolution, the modern synthesis*. London: Allen and Unwin.

Jones EP, Jóhannesdóttir F, Gündüz I, Richards MB, Searle JB. 2010a. The expansion of the house mouse into NW Europe. *Journal of Zoology* DOI: 10.1111/j.1469-7998.2010.00767.x.

Jones EP, van der Kooij J, Solheim R, Searle JB. 2010b. Colonisation and interactions of two subspecies of house mouse (*Mus musculus*) in Norway. *Molecular Ecology* **19:** 5252–5264.

Levene H. 1949. On a matching problem arising in genetics. *Annals of Mathematical Statistics* **20:** 91–94.

Li MH, Sternbauer K, Haahr PT, Kantanen J. 2005. Genetic components in contemporary Faroe Islands cattle as revealed by microsatellite analysis. *Journal of Animal Breeding and Genetics* **122:** 309–317.

McGovern TH, Perdikaris P, Harrison R, Smiarowski K, Manigault N. 2006. An interim report of the Viking Age archaeofauna from Hrísheimar, Mývatn District, N Iceland. *NORSEC Laboratory Report 32*. NABO.

Macholán M, Munclinger P, Šugerková M, Dufková P, Bímová B, Božíková E, Zima J, Piálek J. 2007. Genetic analysis of autosomal and X-linked markers across a mouse hybrid zone. *Evolution* **61:** 746–771.

Marcus GJ. 1980. *The conquest of the North Atlantic*. Woodbridge: Boydell Press.

Marshall JT, Sage RD. 1981. Taxonomy of the house mouse. *Symposia of the Zoological Society of London* **47:** 15–25.

Matthews LH. 1952. *British mammals*. London: Collins.

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

Nei M. 1987. *Molecular evolutionary genetics*. New York: Columbia University Press.

Nunome M, Ishimori C, Alpin KP, Tsuchiya K, Yonekawa H, Moriwaki K, Suzuki H. 2010. Detection of recombinant haplotypes in wild mice (*Mus musculus*) provides new insights into the origin of Japanese mice. *Molecular Ecology* **19:** 2474–2489.

Orth A, Adama T, Din W, Bonhomme F. 1998. Hybridation naturelle entre deux sous-espèces de souris domestique, *Mus musculus domesticus* et *Mus musculus castaneus*, près du lac Casitas (Californie). *Genome* **41:** 104–110.

Posada D. 2008. jModelTest: phylogenetic model averaging. *Molecular Biology and Evolution* **25:** 1253–1256.

Prager EM, Sage RD, Gyllensten U, Thomas WK, Hübner R, Jones CS, Noble L, Searle JB, Wilson AC. 1993. Mitochondrial DNA sequence diversity and the colonisation of Scandinavia by house mice from East Holstein. *Biological Journal of the Linnean Society* **30:** 80–122.

Pritchard JK, Stephens M, Donnelly P. 2000. Inference of population structure using multilocus genotype data. *Genetics* **155:** 945–959.

Rajabi-Maham H, Orth A, Bonhomme F. 2008. Phylogeography and postglacial expansion of *Mus musculus domesticus* inferred from mitochondrial DNA coalescent, from Iran to Europe. *Molecular Ecology* **17:** 627–641.

Raufaste N, Orth A, Belkhir K, Senet D, Smadja C, Baird SJE, Bonhomme F, Dod B, Boursot P. 2005. Inferences of selection and migration in the Danish house mouse hybrid zone. *Biological Journal of the Linnean Society* **84:** 593–616.

Raymond M, Rousset F. 1995. GENEPOP, version 1.2: population genetics software for exact tests and ecumenicism. *Journal of Heredity* **86:** 248–249.

Ronquist F, Huelsenbeck JP. 2003. MRBAYES 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* **19:** 1572–1574.

Rozas J, Sánchez-Del Barrio JC, Messeguer X, Rozas R. 2003. DNAsp, DNA polymorphism analyses by the coalescent and other methods. *Bioinformatics* **19:** 2496–2497.

Searle JB, Jamieson PM, Gündüz I, Stevens MI, Jones EP, Gemmill CEC, King CM. 2009a. The diverse origins of New Zealand house mice. *Proceedings of the Royal Society of London Series B, Biological Sciences* **276:** 209–217.

Searle JB, Jones CS, Gündüz I, Scascitelli M, Jones EP, Herman JS, Rambau RV, Noble LR, Berry RJ, Giménez MD, Jóhannesdóttir F. 2009b. Of mice and (Viking?) men: phylogeography of British and Irish house mice. *Proceedings of the Royal Society of London Series B, Biological Sciences* **276:** 201–207.

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.

Vanlerberghe F, Boursot P, Catalan J, Gerasimov S, Bonhomme F, Botev BA, Thaler L. 1988. Analyse génétique de la zone d'hybridation entre les deux sous-espèce de souris *M. m. domesticus* et *M. m. musculus* en Bulgarie. *Genome* **30:** 427–437.

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

Yeh FC, Yang R-C, Boyle T, Ye Z-H, Mao JX. 1997. *POPGENE, the user-friendly shareware for population genetic analysis*. Edmonton: Molecular and Biotechnology Centre, University of Alberta.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article:

Figure S1. Visualization of the STRUCTURE output for $K = 8$–11.

Table S1. Microsatellite genotypes of all individuals in each population. Loci D5Mit145 and D16Mit2 were not included in the analysis (see text).

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