Of mice and (Viking?) men: phylogeography of British and Irish house mice

Jeremy B. Searle1,8,*†, Catherine S. Jones2,3,†, İslam Gündüz1,4,†, Moira Scascitelli1,5,6,†, Eleanor P. Jones1, Jeremy S. Herman1,7, R. Victor Rambau1,8, Leslie R. Noble2,3, R. J. Berry2, Mabel D. Giménez1 and Fríða Jóhannesdóttir1

1Department of Biology, University of York, PO Box 373, York YO10 5YW, UK
2Department of Biology, University College London, London WC1E 6BT, UK
3School of Biological Sciences, Zoology Building, University of Aberdeen, Tillydrone Avenue, Aberdeen AB24 2TZ, UK
4Department of Biology, Faculty of Arts and Sciences, University of Ondokuz Mayıs, Samsun 55139, Turkey
5Department of Biology, University of Rome ‘Tor Vergata’, 00133 Rome, Italy
6Department of Botany, University of British Columbia, Vancouver, British Columbia, Canada V6T 1Z4,
7Department of Natural Sciences, National Museums of Scotland, Chambers Street, Edinburgh EH1 1JF, UK
8Department of Botany and Zoology, Evolutionary Genomics Group, Stellenbosch University, Private Bag X1, Matieland 7602, Stellenbosch, Republic of South Africa

The west European subspecies of house mouse (Mus musculus domesticus) has gained much of its current widespread distribution through commensalism with humans. This means that the phylogeography of M. m. domesticus should reflect patterns of human movements. We studied restriction fragment length polymorphism (RFLP) and DNA sequence variations in mouse mitochondrial (mt) DNA throughout the British Isles (328 mice from 105 localities, including previously published data). There is a major mtDNA lineage revealed by both RFLP and sequence analyses, which is restricted to the northern and western peripheries of the British Isles, and also occurs in Norway. This distribution of the ‘Orkney’ lineage fits well with the sphere of influence of the Norwegian Vikings and was probably generated through inadvertent transport by them. To form viable populations, house mice would have required large human settlements such as the Norwegian Vikings founded. The other parts of the British Isles (essentially most of mainland Britain) are characterized by house mice with different mtDNA sequences, some of which are also found in Germany, and which probably reflect both Iron Age movements of people and mice and earlier development of large human settlements. MtDNA studies on house mice have the potential to reveal novel aspects of human history.

Keywords: colonization history; D-loop; mitochondrial DNA; Mus musculus domesticus; restriction fragment length polymorphisms; Vikings

1. INTRODUCTION
The colonization history of the house mouse Mus musculus is inextricably linked to human movements; large parts of its cosmopolitan range have been attained through passive transport by humans (Pocock et al. 2005). House mice can exploit the food of humans and their livestock, and are able to live in human dwellings, vehicles and boats. The western subspecies Mus musculus domesticus, in particular, was fortuitously placed, resident in the Fertile Crescent, at the time that humans first formed large settlements (8000 BC) and was able to become commensal with humans in such places (Brothwell 1981; Tchernov 1984). However, M. m. domesticus did not spread widely into Europe from the Middle East until the Iron Age (starting 1000 BC) (Cucchi et al. 2005). At this time, there was apparently extensive migration of the subspecies throughout the Mediterranean and western Europe.

This link between humans and house mice is noteworthy. It means that the colonization history of house mice should reflect human migrations and trading links, and even colonization events. Therefore, a knowledge of the colonization history of this commensal may be a valuable archaeological tool, revealing events in human history that are not well represented in artefacts or written documents (Gündüz et al. 2001; Haynes et al. 2003).

A powerful way to infer the colonization history of a species is through phylogenetic studies of molecular markers in a geographical context. This ‘phylogeographic’ approach (Avise 2000) has been used with much effect, e.g. in tracing the postglacial colonization history of European fauna and flora (Hewitt 2000). Mitochondrial (mt) DNA has been the best-studied genomic region in the phylogeographic studies of mammals, and in the house mouse there have already been extensive studies, particularly by Prager and colleagues on complete D-loop sequences (Prager et al. 1993, 1996, 1998; see also
Given that the phylogeography of house mice around the Atlantic periphery of Europe has yielded such interesting results, it is worthwhile to examine the colonization history of the British Isles by *M. m. domesticus*. The British Isles have had a complex human history over the last 3,000 years including participation by both Danish and Norwegian Vikings. In Britain, the earliest records of house mice date from the Bronze Age (figure 1); though given the finding elsewhere in western Europe that the species did not arrive until the Iron Age, these early records may not be secure (e.g. they could be intrusive). In general, the Iron Age records of house mice come from southern Britain, while there is a substantial amount of Roman material from central Britain. There are no records of house mice from Ireland contemporary with the British Bronze Age, Iron Age or Roman periods.

For the phylogeographic study of the house mouse in the British Isles reported here, we incorporate all relevant mtDNA data available. This includes published mtD-loop sequences as well as our previously unreported RFLP and sequence data collected over the last 20 years. Together, these provide a good coverage of the British Isles, including a particularly detailed representation of Orkney and the neighbouring part of mainland Britain.

2. MATERIAL AND METHODS

(a) Specimens

Using molecular techniques, we typed 310 mice from 96 localities in the British Isles (appendices 2 and 3 in the electronic supplementary material), adding to published data for 18 individuals (nine additional localities). The mice that we typed were either collected by ourselves, generally through live trapping on farms, or provided by colleagues, as listed in appendix 2. Whole specimens or tissue samples were stored frozen (−80°C) or in absolute ethanol.

(b) Molecular methods and analysis

For RFLP analysis, mitochondria were prepared from heart and kidneys using differential centrifugation procedures and mtDNA isolated by a modified phenol extraction procedure (Jones et al. 1988). DNA samples were digested with 13 restriction enzymes (appendix 4 in the electronic supplementary material). Fragments were separated either on 5 per cent polyacrylamide gels run for 3 hours and visualized by silver staining (Tegelström 1986), or run on 0.8 per cent agarose gels in TBE (89 mM Tris–HCl, 89 mM boric acid, 2 mM EDTA (pH 8.0)), and stained with ethidium bromide. The sizes of the fragments were estimated by comparing them with the fragments of known sizes from the published mouse mtDNA sequence (Bibb et al. 1981), run as a standard on each gel and cut with the appropriate enzyme.

Each restriction fragment digestion profile was site mapped using the sequence comparison method (Cann et al. 1982), in which restriction sites required to account for each fragment pattern were determined by reference to Bibb et al. (1981) and the existing maps (Ferris et al. 1983; Sage et al. 1990). A composite genotype was thus generated for each individual based on the presence and absence of each restriction site. These genotypes were incorporated into a data matrix used in an undirected parsimony analysis with PAUP* v. 4.0 (Swofford 2002), generating a 50 per cent majority-rule consensus tree after

Figure 1. Archaeological records of house mice in Britain (see appendix 1 in the electronic supplementary material for further details). Following the chronology by the British and Irish Archaeological Bibliography (http://www.biab.ac.uk/chronology.asp), Bronze Age (black circles) is considered as 2300–700 BC, Iron Age (grey circles; lowland Britain) as 700 BC to AD 43 and Roman (white circles; lowland Britain) as AD 43–450. As most records are from lowland Britain, we retain the chronological framework from lowland Britain when referring to sites elsewhere.
a 1000 replicates of a heuristic search with random stepwise addition of taxa. Nodal support was estimated by bootstrap resampling (n = 1000).

For DNA sequencing of all or part of the mitochondrial D-loop, total genomic DNA was extracted from tissue samples (normally tail tip) by phenol–chloroform or with Qiagen DNeasy tissue kits. The complete mt D-loop and flanking regions were amplified with the PCR primers and the conditions as described by Gündüz et al. (2000), or with the primer pairs L831-5′-ATAGGCTTCTAGCTCAA-3′ and H2228-5′-TTATAAGGCGAGAACAAC-3′. The fragments generated were sequenced in both directions using the same primers. The sequences were aligned and edited using SEQUENCHER v. 4.8 (Gene Codes Corp.). For all specimens, we obtained sequences for the mtDNA region between positions 15 363 and 16 295 (933 bp) of Bibb et al. (1981). These were then truncated to a standard length of 853 bp, between positions 15 424 and 16 276, for comparison with other sequences. Our sequences have been deposited in GenBank (FM211596–FM211631).

All the complete D-loop sequences from the British Isles were combined with equivalent published M. m. domesticus sequence, including those in the companion paper (Searle et al. in press), for phylogenetic analysis. GenBank sequences of the following were used as outgroups, adopting Rajabi-Maham et al. (2008): M. m. gentilulus, AF074545 and AF074544; M. m. castaneus, EF108342, AF088879 and AJ286322; and M. m. musculus, U47504 and U47532. The GTR+I+Γ substitution model was selected using MODELTEST v. 3.7 (Posada & Crandall 1998). Also following Rajabi-Maham et al. (2008), we employed a Bayesian algorithm to analyse our data, using the program MrBayes (Huelsenbeck & Ronquist 2001; Ronquist & Huelsenbeck 2003) to construct a 50 per cent majority-rule consensus tree. Two independent Markov chain Monte Carlo analyses were run, each with one cold chain and four heated chains and the incremental heating parameter set at 0.2. The analyses were terminated after 5 million generations and the first 30 per cent of trees were discarded as burn-in.

For some populations, partial D-loop sequences were obtained as a preliminary to select complete D-loop sequencing (Gündüz 1999). Partial D-loop sequences were generated with primers L15774 and H16498 using the conditions described in Gündüz et al. (2000). This produced a 305 bp sequence between positions 15 363 and 15 667 of Bibb et al. (1981) including the left end of the D-loop (251 bp) and part of the Pro-tRNA (54 bp). Despite the short length of the sequence, partial D-loop sequences capture the majority of nucleotide variation displayed by complete D-loop sequences (see appendix 5 in the electronic supplementary material). Therefore, the partial sequences could be allocated with ease into the major lineages identified by the analysis of the complete D-loop sequences.

Nucleotide (π) and haplotype (h) diversities for the D-loop sequences were calculated according to Nei (1987), using MEGA v. 4 (Tamura et al. 2007) and ARLEQUIN v. 3.0 (Excoffier et al. 2005), respectively. These diversities were estimated on an unbiased dataset of complete D-loop sequences collected from throughout the British Isles (80 individuals). Our new complete D-loop sequences from localities 20–25, 27–29, 37, 38 and 47 were excluded from the analysis owing to the bias introduced by selecting individuals for complete sequencing on the basis of their partial sequence.

### 3. RESULTS

New complete D-loop sequences were generated from 77 house mice from the British Isles adding to 18 published sequences (11 haplotypes). Altogether, the 95 mice are represented by 40 haplotypes (appendices 2 and 5 in the electronic supplementary material). The overall haplotype (h ± s.d.) and nucleotide diversities (π ± s.d.) are 0.955 ± 0.009 and 0.0069 ± 0.0017, respectively. Seven haplotypes were found in five or more individuals within the British Isles. BritIsl.8, 19 and 26 were restricted to the Scottish archipelagos of Shetland, Orkney and the Outer Hebrides, respectively. BritIsl.1, 5, 16 and 31 were wide ranging in the British Isles, and also have been found elsewhere in the world (appendix 6 in the electronic supplementary material). BritIsl.1 was found in northern Scotland (Sutherland), southern Scotland (Dumfries), northern England (York), southern England (e.g. Wiltshire) and the Isle of Man. It has also been found in Germany, New Zealand and Mauritania. BritIsl.5 was found in southern Scotland (East Lothian), southern England (e.g. Somerset), Islay (Scotland) and Anglesey (Wales), and has also been found in Germany, Denmark, Norway, New Zealand and Cambodia; BritIsl.16 was found in Orkney, Ireland and Norway; and BritIsl.31 was found in Orkney, northern Scotland (e.g. Caithness), Ireland, Norway and Portugal. There are two additional haplotypes known from both the British Isles and elsewhere in the world: BritIsl.12 (Germany) and BritIsl.14 (Portugal).

When combined with all other available D-loop sequences of M. m. domesticus in a Bayesian analysis, most of the haplotypes, including all those found to be widely distributed in the British Isles, form two major lineages (figure 2 and appendix 6 in the electronic supplementary material): one which includes BritIsl.1–10 is well supported (posterior probability 0.96) and the other (which includes BritIsl.16–36) is poorly supported (posterior probability 0.65), which is typical of other lineages that have been identified in M. m. domesticus phylogenetic trees (Prager et al. 1993, 1996, 1998; Gündüz et al. 2001, 2005; Rajabi-Maham et al. 2008), due to the recent diversification and high levels of homoplasy associated with the highly variable D-loop region (Rajabi-Maham et al. 2008). The remaining haplotypes (BritIsl.11–15) are located elsewhere in the phylogenetic tree (figure 2).

The lineage consisting of haplotypes BritIsl.16–36 has occurred as a separate lineage in all previous phylogenetic analyses that have included the relevant haplotypes (Prager et al. 1993, 1996, 1998; Nachman et al. 1994; Gündüz et al. 2001, 2005). The lineage has a striking distribution in the British Isles relative to other M. m. domesticus haplotypes (appendix 2 in the electronic supplementary material). It is restricted to the northern and western peripheries (the Outer Hebrides, Shetland and Orkney and nearby mainland, and Ireland). Because its association has been demonstrated particularly clearly with Orkney, we term this the ‘Orkney’ mtDNA lineage. Mus musculus domesticus with haplotypes distinct from this Orkney lineage have been found throughout mainland Britain and in parts of Ireland, the Isle of Man and Shetland.

Further confirmation of the geographical range of the Orkney mtDNA lineage comes from partial D-loop
sequences and RFLP data of the entire mtDNA genome (appendices 2 and 4 in the electronic supplementary material). Again, in the phylogenetic tree generated from the RFLP data, the Orkney lineage appears as a clear entity clustering together individuals from Orkney and its vicinity, the Isle of Man, and Ireland to the exclusion of individuals from England, Wales and southern Scotland. Altogether, there are substantial data to show that Orkney and the nearby mainland (Caithness) are occupied almost solely by the mice of the Orkney mtDNA lineage, which is also predominant in the western parts of the British Isles (figure 3).

4. DISCUSSION

The spread of *M. m. domesticus* throughout much of the world began with populations in the Middle East that first associated with humans ca 8000 BC (Cucchi et al. 2005). The British Isles are clearly at the periphery of the western European expansion of this house mouse subspecies, despite which there is substantial genetic variation here, as reflected, for instance, by comparison of haplotype and nucleotide diversities with Turkey (British Isles: $h=0.96$, $\pi=0.0069$; Turkey: $h=0.98$, $\pi=0.0084$; Gündüz et al. 2005; see also Rajabi-Maham et al. 2008). However, while the variation shows only weak geographical structure in Turkey (Gündüz et al. 2005), it is strongly regionalized in the British Isles. In particular, there is an Orkney mtDNA lineage that is predominant in and restricted to the northern and western peripheries of the British Isles (Orkney, Shetland, Outer Hebrides, Isle of Man, Ireland; figure 3). Given that house mice mtDNA haplotypes are a marker of human movements, it is pertinent to consider how this has come about. It is especially striking that the Orkney mtDNA lineage has a distribution within the British Isles that matches the Norwegian Viking sphere of influence (Logan 1991). Accordingly, the name of the lineage is apposite: Orkney was a key centre within the Norwegian Viking kingdom in the eleventh and twelfth centuries AD (Crawford 1987; Forte et al. 2005).

![Phylogenetic tree for *M. m. domesticus*](image-url)
Further support for a causal linkage between the Orkney lineage and Norwegian Vikings comes from the available mtDNA sequences from Norway, both published and previously unreported (see appendices 2 and 6 in the electronic supplementary material). Eight out of nine mice sampled from two locations on the coast of central Norway were of the Orkney lineage (four haplotypes). Furthermore, two of those Orkney haplotypes found in Norway (BritIsl.16 and 31) have a wide distribution around the area occupied by the Norwegian Vikings (including Orkney, Ireland and Norway), suggesting that the sphere of influence is reflected not just by a lineage but by specific haplotypes. Such haplotypes are likely to have been some of those present at the time of the transport of the house mice; other haplotypes in the lineage may have evolved in situ through mutation.

There is also evidence of distinctiveness of the Orkney, Shetland and Caithness populations of house mice relative to the rest of the British mainland on the basis of morphometrics, allozyme variation and Y chromosome DNA variation (Berry & Peters 1977; Davis 1983; Nash et al. 1983; Jones 1990).

How can the association of the Orkney lineage and Norwegian Vikings be explained? Clearly, house mice were present in Britain long before the first visits to the British Isles by Norwegian Vikings in AD 793 (Logan 1991). Both the analysis of spread of *M. m. domesticus* in the Mediterranean region (Cucchi et al. 2005) and the archaeological records within Britain (figure 1) are consistent with the arrival of house mice in Britain in the Iron Age, i.e. at least 2 kyr ago. Therefore, it is to be expected that at least some of the house mice in the British Isles would have a genetic architecture that reflects pre-Viking origins, and this is indeed seen. The first settlements of any size in Britain are suggested to be Iron Age hill forts in southern England (Collis 1984; Bewley 1994; Cunliffe 2004), while it is with the arrival of the Romans that towns (albeit garrison towns) emerge further north (Harding 2004). This pattern is reflected by the earliest archaeological records of house mice (figure 1), which fits with the dependence of house mice on large settlements for long-term persistence (Cucchi et al. 2005).

Only in large human settlements are house mice able to form viable populations when in competition with the wood mouse *Apodemus sylvaticus* (Cucchi et al. 2005), a species that occurs throughout the British Isles. Apparently, house mice of the mtDNA lineage including BritIsl.1–10 especially exploited these early large settlements and ultimately expanded over most of mainland Britain. Two haplotypes of this lineage are distributed particularly widely in Britain (BritIsl.1 and 5) and are also found in Germany. There were substantial human cultural contacts between middle and northern Europe and Britain in the Iron Age (Cunliffe 2004), and it can be proposed that ancestors of these mice came from Germany at that time.

Considering now the northern and western peripheries of the British Isles currently occupied by the Orkney lineage, there is limited evidence for a pre-Viking presence of house mice in this area. There is a script from Ireland dating to the seventh century AD that refers to house mice (Hayden 2002), but such anecdotal records are contentious owing to the possibility of confusion between house and wood mice. An archaeological investigation by Nicholson et al. (2005) recorded a single house mouse from an apparently secure Iron Age deposit on Shetland, but even such a record needs confirmation with carbon dating of the bones. The lack of records of house mice from pre-Viking archaeological sites on Orkney is telling, because small mammal bones are particularly well represented and studied there (Nicholson et al. 2005).
Whether or not there were occasional pre-Viking house mice in the region now occupied by the Orkney lineage, more important in terms of the current genetic signal would have been the development of large sustained mouse populations associated with the founding of substantial human settlements. For much of the area covered by the Orkney lineage, the Viking era is the time when urban settlements were initiated, particularly so for Ireland. Prior to the Viking era, there were no towns or large settlements in Ireland (Delany 1977); during the Viking period, the first towns were founded at Limerick, Waterford, Wicklow, Wexford, Dublin and near Cork and Lough (Logan 1991; Forte et al. 2005).

It is our contention, therefore, that for the house mouse the Orkney mtDNA lineage represents a marker for Norwegian Viking influence. It should not be assumed, however, that mice with the Orkney lineage originated in Norway and were taken to the British Isles. The transport may have been in the other direction, but still made by Norwegian Vikings. Our results add to the previously described association between Madeiran mice and Danish Vikings (Gündüz et al. 2001), in indicating that house mice are a valuable proxy for Viking movements, as revealed through the studies of mtDNA. The combination in the Viking period of the spread of urbanization in northwestern Europe and the trade facilitated by sophisticated ships capable of travelling substantial distances and carrying large amounts of cargo make the Vikings ideal house mouse vectors. There is clear evidence from Viking-age deposits in Iceland and Greenland that house mice were indeed transported on Viking ships (Nicholson et al. 2005).

An obvious use of the Orkney mouse lineage as a marker for the Norwegian Vikings is that we may be able to help confirm their activities outside the British Isles. Fine-scale genetic analysis of house mice has the potential to reveal precise cultural associations within the former Norwegian Viking kingdom including Faroe, Iceland, Greenland and Newfoundland. We are currently conducting research in these directions (E. P. Jones & J. B. Searle 2008, unpublished data).

The mtDNA variation of house mice in the British Isles is not only relevant to movements of people and their commensals during the Iron Age and Viking times. The occurrence of house mice with British haplotypes in New Zealand is described in the companion paper (Searle et al. in press), and reflects much more recent human movements, in this case from Britain to New Zealand. The clustering of BritIs1.13 with haplotypes from the USA (appendix 6 in electronic supplementary material) may reflect a lineage that colonized North America from Britain.

Thus, on the time scale of the last 3000 years, it is evident that house mice mirror well the movements of human beings, and therefore the study of their phylogeography represents an unusual, underused, but valuable, aid to archaeologists and historians. Here, we have analysed mtDNA variation to useful effect. In the future, there is clearly the exciting opportunity to extend the phylogeographic analysis radically to exploit the nuclear genome resources of the house mouse (Mouse Genome Sequencing Consortium 2002).

We are extremely grateful for all those who provided the specimens used in this analysis (listed in appendix 2 in the electronic supplementary material), including Graham Triggs and Marcus Hewson for their assistance with some island fieldwork. We also thank Keith Dobney for help with archaeological records, Matthew King for aid with the maps, and the editor and referees for their comments on the paper. Funding was provided by the University of York, the Biotechnology and Biological Sciences Research Council (advanced fellowship to C.S.J., 1/AF09056), the Scientific and Technological Research Council of Turkey (TÜBİTAK), the Skye Foundation and the South African National Research Foundation and Programme Allplan (EU). Part of this work was carried out by using the resources of the Computational Biology Service Unit from the Cornell University, which is partially funded by Microsoft Corporation.

REFERENCES


Proc. R. Soc. B