The colonization history of British water vole (Arvicola amphibius (Linnaeus, 1758)): origins and development of the Celtic fringe

Selina Brace1, Mark Ruddy2, Rebecca Miller3, Danielle C. Schreve4, John R. Stewart5,† and Ian Barnes1,†

1Department of Earth Sciences, Natural History Museum, London SW7 5BD, UK
2Historic England, Cambridge CB2 8BU, UK
3Service of Prehistory, University of Liège, Liège 4000, Belgium
4Department of Geography, Royal Holloway University of London, Egham, Surrey TW20 0EX, UK
5School of Applied Sciences, Bournemouth University, Poole, Dorset BH12 5BB, UK

The terminal Pleistocene and Early Holocene, a period from 15 000 to 18 000 Before Present (BP), was critical in establishing the current Holarctic fauna, with temperate-climate species largely replacing cold-adapted ones at mid-latitudes. However, the timing and nature of this process remain unclear for many taxa, a point that impacts on current and future management strategies. Here, we use an ancient DNA dataset to test more directly postglacial histories of the water vole (Arvicola amphibius, formerly A. terrestris), a species that is both a conservation priority and a pest in different parts of its range. We specifically examine colonization of Britain, where a complex genetic structure can be observed today. Although we focus on population history at the limits of the species’ range, the inclusion of additional European samples allows insights into European postglacial colonization events and provides a molecular perspective on water vole taxonomy.

1. Introduction

The end of the last (Weichselian/Devensian) glaciation ca 14 700 Before Present (BP) until the Mid-Holocene 8200 BP [1] was a period of climatic and environmental change, including the presence of two minor temperate-climate episodes (the Bølling and Allerød interstadials) and a full cycle of glacial re-advance and retreat, known as the Younger Dryas (YD), 12 800–11 500 BP (collectively referred to as the Lateglacial), followed by the rapid climatic amelioration and ensuing vegetation change that accompanied the start of Holocene interglacial warming. The Lateglacial was a period of rapid vegetational change and widespread faunal extinction and translocation across Eurasia [2,3]. It may be viewed as the most recent example of the dramatic climatic fluctuations associated with the Pleistocene, where over the last 2.6 million years (Myr) ice sheets periodically spread down from the north, leaving Northern Europe almost fully glaciated and permafrost extending throughout Central Europe with only the southernmost peninsulae remaining ice and permafrost free [4]. During the last full glacial cycle, the maximum extent of glaciation (the last glacial maximum or LGM; ca 22 500 BP) extended across Scandinavia and the British Isles, with large parts of Europe becoming too cold for many mammal species to survive [2,5], but see [6]. Hewitt [7–9] made significant progress in the development of a model of organismal response to Holarctic climate change, proposing that in Europe temperate populations survived periods of climatic deterioration in three refugial peninsulae (Iberia, Italy and the Balkans), before recolonizing northwards as glaciers retreated. While this model has been central to our understanding of the population histories of the European biota, subsequent work suggests a more complex pattern of
recolonization that can vary in relation to particular taxa, regions and time points [10]. One example is the proposal by Searle et al. [11], who found that a general pattern for the recolonization of Britain can be inferred from studies of multiple small mammal species. Based on a range of genetic markers, they identified pairs of population groups in five different species, and in each case the two populations formed either ‘core’ (roughly England, sometimes excluding the south coast) or ‘peripheral’ (Scotland, Wales and sometimes the south coast of England) populations. This pattern has been referred to as the ‘Celtic fringe’, since it bears a strong resemblance to the cultural and linguistic distinctions that today separate Scottish, Irish, Welsh, Manx and Cornish peoples from those in central and eastern England [11]. While it has been proposed to stem from multiple colonization events from different populations, the exact timing and nature of this process remains unclear.

The application of ancient DNA in reconstructing species history has played a significant role in identifying source populations and postglacial recolonization events (e.g. [2,12,13]). Here, we focus on one of the ‘Celtic fringe’ species, the northern water vole, Arvicola amphibius (also referred to as A. terrestris), a widely distributed species found across Europe (excluding Ireland and central and southern Spain), east through Siberia to the Lena River Basin, and from the Arctic Sea south to Lake Baikal and Northwest China through north west Iran, Iraq, north Israel, the Caucasus and Turkey [14].

We identified northern water vole as the most suitable organism to explore the formation and origins of the ‘Celtic fringe’, as it (i) exhibits a very clear spatial pattern of mitochondrial DNA (mtDNA) differentiation across the present-day Scottish–English border [15]; (ii) is the largest of the small mammal fauna with a proposed ‘Celtic fringe’ distribution [11], thereby increasing the volume of bone available for each analysis and (iii) both English, and especially Scottish, water vole populations are the focus of considerable conservation efforts, and an improved understanding of the origins of these populations could therefore assist in targeting resources. Water voles are clearly sufficiently polymorphic in both ecology and morphology to present a long-standing taxonomic problem. Membership of the water vole genus has fluctuated, ranging from one all-encompassing species (terrestris; [16]), more commonly two (sapidus and terrestris; [17]), but also four (amphibius, sapidus, scherman and terrestris; [18]) and at its peak seven (amphibius, illyricus, italicus, musignani, sapidus, scherman and terrestris; [19]). Current taxonomic determinations recognize three species: A. amphibius (northern water vole, distributed across Eurasia), A. sapidus (Portugal, Spain and France) and A. scherman (European mountains: Alps, Carpathians, Cantabrian Mountains, Massif Central and Pyrenees) [14].

Several studies have sought to resolve the taxonomy and evolution of water vole lineages through molecular analyses [15,20]. Piertney et al. [15] specifically targeted water vole from across Britain. The resulting phylogeny identified the presence of two distinct clades, one with haplotypes from England/Wales and the second with haplotypes from Scotland. A geographical and genetic division of this nature suggests that two colonization events occurred in Britain. Evidence from the within-clade association of five representative European samples highlighted that the Scottish population was derived from an Iberian population, and the English/Welsh population from Eastern Europe. However, owing to the limitations of an exclusively modern DNA-based dataset, it was impossible to discern whether the two colonization events were separated geographically but occurred at the same time, or, whether events were temporally distinct, the second colonizers replacing the first, in one or other of the geographical regions.

The application of an ancient DNA approach therefore provides an ideal mechanism by which to explore the vole colonization of Britain. Through the analysis of Pleistocene, Early Holocene and additional modern water vole samples, we have tested some of the proposals arising from the Celtic fringe hypothesis of postglacial colonization of Britain [11], namely that:

1. There was an initial re-occupation of the mammal fauna after the LGM, in a temperate-climate period dating sometime within the interval 19 000–12 900 BP. Within Britain, this pre-YD population is inferred to have been small and dispersed.
2. The climatic deterioration of the YD would have played an important role in the subsequent replacement of these lineages, a process that would have taken place prior to the formation of the English Channel (and severance from continental Europe) at 8200–8000 BP (AD 1950; [21]).
3. The replacement populations came westwards, via Doggerland, presumably from source populations located in either the Balkans or European Russia.
4. The post-YD population would have been prone to replacement by incoming populations from Continental Europe during the Holocene, which would have expanded quickly in size due to more favourable climatic conditions. Thus, the displacement of mitochondrial clades is due to drift, rather than any ecotypic advantage for life in lowland environments.

Furthermore, and although not the focus of this study, the use of cross-species samples from a range of locations across Europe provides an opportunity to include a molecular perspective on water vole taxonomy and systematics, and in particular to examine the extent to which mtDNA data are congruent with the currently proposed three species taxonomy.

2. Material and methods
(a) Sample collection
A total of 82 water vole samples were collected from across Europe (figure 1; electronic supplementary material, table S1). Sample choice was restricted by availability of material for destructive purposes but was designed to source material from the Late Pleistocene through to the present day. Britain and surrounding areas were of highest priority, but sampling, particularly for modern materials, extended throughout Europe to allow a wider comparison with extant European haplotypes. Modern samples were obtained from archived museum sources, collected within the last 100 years. Mandibles were used throughout, with species-level identifications conducted by the source museums (electronic supplementary material, table S1).

(b) DNA extraction and sequencing
All DNA extractions were conducted in a dedicated ancient DNA laboratory, physically separated from the post-PCR laboratory. Mandibles were ground into a fine powder and DNA was
extracted using silica spin columns based on Yang et al. [22], with the inclusion of 1 M urea in the extraction buffer. mtDNA was amplified using overlapping fragments spanning 643 base pairs of the control region. Six primer pairs were designed specifically for this study (electronic supplementary material, table S2), each pair amplifying short (150–200 base pair) overlapping fragments. PCR reactions, amplicon purification and sequencing were performed as described in [2] with PCR primer-specific annealing temperatures ranging from 50°C to 52°C. Standard ancient DNA protocols [23] were followed throughout these extraction procedures to prevent contamination, with repeated PCR amplification and sequencing of fragments to ensure DNA authenticity and the absence of miscoding lesions.

(c) Phylogenetic analyses
DNA sequences obtained from this study were aligned with additional sequence data, 27 unique modern haplotypes from the Piertney et al. [15] dataset. Phylogenetic relationships were estimated using Bayesian analysis. The DNA substitution model selected with ModelTest3.7 [24] under Akaike information criterion was general time reversible (GTR) with proportion of invariable sites (I) set to 0.6802 and gamma distribution (G) shape parameter 0.8091. Bayesian trees were generated and compared at posterior probability of 0.95 to 0.50. Portions of the control region were 30% homologous between the pairings involved. The southern water vole (A. sapidus; sample 208) from Portugal was employed as the outgroup in analyses.

Sequence data were partitioned into haplogroups and southern water vole species to establish sequence divergence between haplogroups and the southern water vole. These were calculated in ARLEQUIN v. 3.11 [26], using pairwise estimates of corrected average population sequence divergence.

(d) Radiocarbon dating
Where there was a sufficient mass of sample material, water vole mandibles extracted in this study were also sent for accelerator mass spectrometry (AMS) dating at the Oxford radiocarbon accelerator unit (16 samples). Dates were received as uncalibrated radiocarbon years BP, calibrated calendar ages were generated using Oxcal v4.1 [27] with IntCal09 calibration curve [28].

3. Results

(a) DNA recovery
Water vole mtDNA was successfully amplified from a total of 70 specimens. From the modern samples, 36 gave amplifiable DNA (97%), the 14 samples dating to the Holocene returned a 100% success rate and a total of 20 samples from the Pleistocene (65%) also amplified water vole mtDNA. Three of these samples generated insufficient coverage (less than 200 base pairs) and were therefore excluded from analyses. Of the remaining 67 samples, 62 amplified the entire 643 base pair region of interest; a further five successfully amplified all but one PCR fragment (electronic supplementary material, table S1). To test whether the un-amplified regions contained informative data, phylogenies were generated using both the entire region of interest and with the un-amplified regions omitted. Sequences produced were identical; the five partially amplified samples were therefore included in all further analyses.

(b) Phylogenetic analyses
The Bayesian analyses (figure 1) supports three clades of water voles, exhibiting division between samples from the

Figure 1. Sampling locations for water vole used in this study. British numbered locations key: England: 1, Somerset (187); 2, Surrey; 3, Hertfordshire; 4, Cambridgehire; 5, Norfolk; 6, Shropshire; 7, Staffordshire; 8, Derbyshire (104,106,107); 9, Lincolnshire Read’s Island (195); 10, Yorkshire; 11, Northumberland (200). Wales: 12, Gwent (158,159). Scotland: 13, Fife; 14, Morayshire. European sampling locations are denoted by country name. Bracketed numbers refer to the individual sampling numbers of water vole that are directly referred to in the text.
Table 1. Population average pairwise estimates of sequence divergence with Kimura-2 parameter between clades 1 and 3 as defined by the phylogeny (figure 2) and the southern water vole (*A. sapidus*).

<table>
<thead>
<tr>
<th>clade</th>
<th>all</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>1</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>2.12</td>
</tr>
<tr>
<td>3</td>
<td>2.42</td>
</tr>
<tr>
<td>all</td>
<td>3.9</td>
</tr>
</tbody>
</table>

Pleistocene and older Holocene samples. Further to clade identification, sequence data were partitioned to assess the percentage of sequence divergence between the haplogroups identified and the sister species, the southern water vole (table 1).

4. Discussion

(a) Confirmation of a Celtic fringe distribution in *Arvicola*

A primary aim of this study was to test predictions made about the so-called ‘Celtic fringe’ distribution of small mammal haplotypes in Britain, using ancient DNA, with water vole as a model system. We also incorporated original, recent samples from across Britain, to enable exploration of a major (molecular) division between modern water vole in England and those in Scotland [15]. The resulting phylogeny (figure 2) places all modern Scottish samples as part of a single clade, identified here as clade 1, and nearly all recent English samples in an additional clade, denoted clade 2. There are two exceptions to this pattern: sample 195 (Read’s Island) and 200 (Northumberland) are both from English locations but phylogenetically are placed within clade 1. The Northumberland sample is directly adjacent to the Scottish border and that close proximity likely accounts for its association with the ‘Scottish’ clade. Read’s Island, however, is more of an anomaly. Today, an RSPB reserve situated on the Humber Estuary in East Yorkshire, it is considerably (more than 200 km) south of the Scottish border, and this sample could therefore represent a recent translocation of individuals from Scotland to the Humber. Despite this, the overall trend of a major population division between water vole in the north and the south of Britain persists, with strong support for the nodes that define lineage separation into haplogroup clusters.

(b) Post-last glacial maximum recolonization of Britain

Confirmation of the genetic division observed by Piertney *et al.* [15] can be achieved through modern sampling efforts, but to better establish the timing and mode of colonization, ancient Pleistocene and Holocene water vole samples were included in the analysis. Our AMS dating of these samples shows water vole presence in southern England immediately prior to the LGM (e.g. sample 157; median calibrated date 27 955 BP, see the electronic supplementary material, table S1). While it was not possible to establish a haplotype for sample 157, ancient samples that were successfully haplotyped are indicated in the phylogeny (figure 1) by geographical location and coloured blue (Pleistocene) and pink (Holocene). The phylogenetic placement of these samples is highly informative; Pleistocene samples from England share the same haplogroup as those currently restricted to Scotland, clade 1. In England, only Holocene samples post-dating the YD cluster within the modern English water vole clade (clade 2). This supports the two-phase colonization proposal [11], with members of clade 1 colonizing Britain and subsequently distributed throughout England pre-LGM. Following the end of the Pleistocene, this group was displaced by a second wave of colonizers that remained in England throughout the Holocene to the current day.

The combination of radiocarbon dating and phylogenetic inference provides a clear indication that water vole colonized Britain on (at least) two separate occasions, with a resulting population structure that can be attributed to temporally, rather than spatially, distinct colonization events. This reconstruction also suggests a further potential explanation for the anomalous sample at Read’s Island, as it could represent a relict population of the initial colonizers, a now isolated remnant of the original colonizers prior to their displacement to Scotland.

(c) Timing of the post-Younger Dryas colonization

The timing of the second colonization event can be inferred from the direct dating of samples used in this study (electronic supplementary material, table S1 and figure 1). Only three samples from Britain could be directly dated: two of these (187 and 158) were clade 1 individuals, with dates prior to the end of the YD (median calibrated dates); 14 621 and 12 081 BP, respectively. A third sample, 159, dates to 2900 BP and is associated with the second wave of colonizers, clade 2. Three further clade 2 samples (104, 106, 107) are undated, but were excavated from a Bronze Age barrow, which would provide a maximum date boundary at 4500 BP. Thus, the second colonization event occurred after 12 081 BP, possibly before 4500 BP, and definitely before 2900 years BP. These dates are thus compatible with natural colonization (rather than human translocation as water vole are neither domesticates or commensals) via the landbridge between England and continental Europe that is estimated to have been inundated ca 8000 BP [29].

(d) The process of replacement

The most plausible timing for the second colonization would be before the loss of the landbridge with continental Europe, between 12 and 8 kyr BP. This was a period of oscillating climate, spanning the end of the last glaciation, the intermediate climatic transitions of the Lateglacial and finally, the start of the Holocene interglacial. Population displacements are most commonly associated with one population outcompeting another, through some ecotypic advantage. Water voles from Scotland are generally considered smaller and typically darker in colour than those from England [30], but morphological studies have shown there to be a continuum across Britain and the variability insufficient to support species or subspecies level differences [17,31]. However, in the early 1900s, water vole in Scotland were considered a subspecies, *A. terrestris reta* [32] based on their darker melanic pelage. Additionally, a study by Turk [33] found Bronze Age water vole skulls from Derbyshire, England to be more akin to the Scottish *reta* subspecies than to the English subspecies,
A. t. amphibius. Turk postulated that either the Early Bronze Age population differentiated into two subspecies or that the retra population was once common in England, but was replaced by a second population (amphibius) after the Bronze Age. This led Van den Brink [34] to allocate two species of British water vole, suggesting that the northerly species had been driven back to the Scottish Highlands by the species from the south. Both authors were derided for their claims;

**Figure 2.** Phylogenetic relationships of water vole samples: Bayesian tree constructed in MrBayes. Nodal support is shown through approximate Bayesian probabilities (only values above 50% shown). AMS dates are given as the median calibrated year BP. Southern water vole used as the outgroup (not shown). Sample location written entirely in uppercase = data taken from Piertney et al. [15]; all other data are from this study and 1st number = unique sample identifier, followed by sample location and time period, where blue = Pleistocene (more than 10 kyr BP, oldest dated sample 27 kyr BP), pink = archaeological Holocene (10 kyr BP to 100 yr BP), black = recent/museum samples (100 yr to present). Green star denotes that the sample is directly referred to in the text.
Montgomery [31] asserted that the insufficient sample size of Turk [33] had led to the erroneous identification of two species, and that only one species of water vole had been present in central England during the past 12,000 years. However, in light of the findings from this study, Turk and Van den Brink appear to have been correct with regards to a displacement event due to a second colonization, even if the proposed timings were inaccurate and the two populations are not sufficiently genetically diversified (table 1) as to warrant recognition as separate species.

The evidence from this study indicates that a second wave of colonizers did indeed displace the first population of water vole in England. As there are no reports of discernable ecological or physiological differences between the two, an alternative supposition, also suggested for other ‘Celtic fringe’ species [11], is that the colder climate of the YD heralded a severe reduction in water vole numbers. As the climate warmed, the second colonizers arrived to a region virtually devoid of water voles, resulting in complete genetic replacement in the south. However, as the climate warmed, remnant populations could recover and repopulate, meaning that sufficient numbers were in place to avoid genetic replacement in the north.

(e) Sources of the two British populations and the possible existence of cryptic northern refugia for water voles

Taking a broader geographical outlook, the Pleistocene/Holocene division can be observed across continental Europe (figure 1). Pleistocene samples from Slovakia, Belgium and Germany are found in haplogroup 1, whereas Holocene samples from Belgium and Germany fall within haplogroup 2. This suggests that population replacement in these regions occurred after the YD. Belgium is an interesting case, as the modern sample reverts to a clade 1 haplotype. This could indicate that both haplotypes remain in the region, highlighting the possibility of a Belgian suture zone. As only one modern Belgian sample successfully yielded DNA, there is insufficient support to test this line of investigation, but with greater sampling effort in the region, this question might be resolved. Germany also warrants additional discussion, as dated Pleistocene samples from the same site (Fuchsloch im Krockstein) can be found in both clades 1 and 2. Sample 99, a clade 1 haplotype, dates to immediately prior to the YD, while sample 111, a clade 2 haplotype, has a median calibrated date of 11,538 BP; the very cusp of the YD/Holocene boundary. The temporal interval between these two samples could therefore represent a more accurate estimate of the timing of the second colonization event. An alternative interpretation is that Germany also forms part of a suture zone with both haplotypes present, a question as with Belgium that could be resolved with further sampling efforts.

Although not exhaustive, the inclusion of additional modern samples illustrates a clear lineage division between Eastern and Western Europe. This is in agreement with the previous proposals [15] that major divergent lineages in Europe are derived from Iberian, and eastern refugia.

Expansion from an eastern, rather than a peninsular refugia, is in contrast to the standard ‘Hewitt’ model of post-glacial recolonization, and provides further support for the presence of refugia north of the European continental divide [35,36]. It also underlines the importance of examining a wide spectrum of species, including small mammals, as water vole appear to exhibit an unusual pattern of genetic diversity. In contrast to almost all other species so far studied [37], the Pleistocene–Holocene transition did not result in a major decline in genetic diversity in this species.

(f) Taxonomy and conservation status of European Arvicola

Our results indicate clear confirmation of the species-level status of the southern water vole (A. sapidus), through the monophyly of samples and the high percentage of sequence divergence between southern and northern water vole (3.66%; table 1). However, on the basis of mtDNA, we find no basis for elevating the montane water vole (A. scherman) to species level. Three modern samples (213, 221, 222) identified as montane water vole failed to form a cohesive monophyletic association—in fact, they were designated to different haplogroups (213 and 221 (France) were clade 1, while 222 (Slovenia) was clade 2). A far clearer lineage separation was apparent in samples from Italy (203 and 215) and southern Switzerland (219). Together, these recent samples form a monophyletic clade, clade 3, with relatively high sequence divergence with respect to either clade 1 (2.42%) or clade 2 (2.49%), and with a divergence not much lower than the species-level difference between northern and southern water vole (3.66%; table 1). From a mtDNA perspective, the geographical distribution of this haplogroup, coupled with sequence divergence from other groups, provides a more convincing argument that it should be considered as a separate taxonomic unit distinct from the montane water vole. However, an analysis of the nuclear genome will be required to fully understand the history and taxonomic status of these lineages, as well as the relationship between the lineages currently found in Britain (e.g. [38]).

In summary, this study highlights the benefits and increased depth of knowledge that can be obtained by incorporating an ancient DNA approach into studies of population history. While modern data were sufficient to identify molecular distinction between English and Scottish water voles, this study was able to reveal a more comprehensive explanation of lineage separation observed in British water vole today, through phylogenetic analysis of ancient and modern DNA. We also suggest an order and timing of the colonization of Britain, defining two temporally distinct events, with a second lineage of water vole replacing the first in England ca 12–8 kyr BP, leaving the initial British colonizers restricted to Scotland.

Data accessibility. DNA sequences have been deposited in GenBank (http://www.ncbi.nlm.nih.gov/). Additional supporting information may be found in the electronic supplementary material of this article.

Authors’ contributions. I.B., S.B. and J.R.S. designed the project. D.C.S., J.R.S., R.M. and M.R. provided samples and contextual information. S.B. performed the research. S.B. and I.B. analysed the data and wrote the manuscript, with contributions from the other authors.

Competing interests. We declare we have no competing interests.

Funding. Funding for this study was provided by SYNTHESYS2 (SYNthesis of SYStematic resources), made available by the European Community Research Infrastructure under FP7 (Synthesis of Systematic Resources, 226506-CP-CSA-Infra and BE-TAF-4725), the Natural Environment Research Council Doctoral Training Grant NER/S/A/2006/14031, and the EU FP6 ERA-NET project CLIMIGRATE (Integrating Ancient DNA and Ecological Modeling to Quantify the Impact of Climate Change on Biodiversity).