First asserted record of the house mouse in Morocco: application of a multidisciplinary approach to the site of Rirha (5th – 1st c. BC)

José Utge, Camille Larrue, Ángel Domínguez-García, Abel Moclán, Mohamed Kbiri Alaoui, Elsa Rocca, Charlotte Carrato, Laurent Callegarin, Claire-Anne de Chazelle, Tarek Oueslati, et al.

Archaeological and Anthropological Sciences (2024) 16:93
https://doi.org/10.1007/s12520-024-02002-8

First asserted record of the house mouse in Morocco: application of a multidisciplinary approach to the site of Rirha (5th – 1st c. BC)

Ángel C. Domínguez-García [1,2], José Utge [3], Camille Larrue [3], Abel Moclán [4], Mohamed Kbiri Alaoui [5], Elsa Rocca [6], Charlotte Carrato [6,7], Laurent Callegarin [8], Claire-Anne De Chazelle [6], Tarek Oueslati [9], Emmanuelle Stoetzel [10]

[1] Departamento de Ciencias de la Tierra, Facultad de Ciencias, Aragosaurus-IUCA, Universidad de Zaragoza, Zaragoza, Spain
[2] Departamento de Geodinámica, Estratigrafía y Paleontología, Facultad de Ciencias Geológicas, Universidad Complutense de Madrid, Madrid, Spain
[3] Eco-Anthropologie (EA) - UMR 7206, MNHN/CNRS/ Université Paris Diderot, Paris, France
[4] Institute of Evolution in Africa (IDEA), University of Alcalá de Henares, Madrid, Spain
[5] Institut National des Sciences de L'archéologie et du Patrimoine (INSAP), Rabat, Morocco
[6] Archéologie des Sociétés Méditerranéennes (ASM) - UMR 5140, CNRS/Université Paul Valéry, Montpellier, France
[7] Mosaïques Archéologie, Cournonterral, France
[8] Université de Pau et des Pays de l'Adour, UAR 3155 IRAA, Pau, France
[9] Histoire Archéologie Littérature des Mondes Anciens (HALMA) - UMR 8164, CNRS/Université de Lille, Lille, France
[10] Histoire Naturelle de l'Homme Préhistorique (HNHP) - UMR 7194, CNRS/MNHN/UPVD, Paris, France
*corresponding author : Emmanuelle Stoetzel (emmanuelle.stoetzel@mnhn.fr)

Abstract
The biogeographic and demographic history of the house mouse, *Mus musculus domesticus*, has been closely linked to human history since pre-Neolithic times – this species first appeared in the Near East and subsequently dispersed westward with humans into regions with local *Mus species*. However, if this issue is now well known for the eastern Mediterranean, data are still lacking for the western Mediterranean, and especially for north-western Africa. In the present study, we combine morphological, geometric morphometric and genetic analyses to *Mus* remains from the Moroccan site of Rirha, dating to the Mauretanian period (5th to 1st century BC). Thanks to this multidisciplinary approach, we were able to confirm the first asserted record of the house mouse in a well-established archaeological context in the Maghreb. The morphometric and genetic results are largely congruent, with a single discrepancy observed for one specimen, which may be linked to possible hybridization between *M. m. domesticus* and *M. spretus* species. Further analyses are required on material from other North African sites dating to earlier chrono-cultural periods to better document the oldest appearance and the dispersal of the house mouse in the south-western Mediterranean through time.

Keywords : *Mus musculus domesticus,* Molar morphology, Geometric morphometrics, Palaeogenetics, Morocco, Mauretanian period.

INTRODUCTION

The biogeographic and demographic history of certain animal species can be closely linked to human history. This is the case, for example, of anthropophilic and commensal species, which have been able to adapt to new ecological niches offered by humans, and then expand geographically and demographically thanks to humans (e.g., Dobson 1998; Grayson 2001; Hulme-Beaman et al. 2016).

Probably one of the best-known examples among vertebrates is the house mouse (*Mus musculus domesticus*), which is thought to have begun developing commensal behaviour around 15,000 years ago in the Levant, shortly before the development of agriculture in the region (Weissbrod et al. 2017). The emergence of commensal behaviour was probably linked to the sedentarization of human populations and the development of villages, bringing several advantages for these small animals, such as access to waste and food stocks and some protection from predators and weather (Weissbrod et al. 2017). *Mus musculus domesticus* would then have dispersed from the Levant in the eastern Mediterranean region, spreading to Cyprus and Anatolia between 11,000 and 9,000 cal BP and towards Transcaucasia between 5,000 and 4,000 cal BP (Cucchi et al. 2005, 2020). For the western Mediterranean region, both the fossil record and genetic data of extant populations indicate that it expanded after 3,000 years ago (~ 1000 BC), favoured by the intensification of trade and navigation along the Mediterranean by Phoenicians, as well as human population expansion (Bonhomme et al. 2011; Cucchi et al. 2005, 2020; Jones et al. 2013; Domínguez García et al. 2019). However, recent studies suggest that the house mouse could have arrived even earlier (4,400–4,100 cal BP) to northern Iberia, following an alternative route through inland Europe (Álvarez-Vena et al. 2023). The dispersal to north-western Africa could then have occurred either by sea from Iberia, or through parallel dispersal events on both sides of the Mediterranean. While these dispersal processes are already well known for the eastern Mediterranean (Cucchi et al. 2020), this is not the case for the western Mediterranean, in particular for the southern shore (North Africa), for which palaeontological data are scarce.

In north-western Africa, another mouse species occurs today: the Algerian mouse or short-tailed mouse, *Mus spretus*. The biogeographic history of this species is also related to the dispersion of human groups. Originating from north-western Africa, and known in this region since the Middle Pleistocene (Stoetzel 2013), *Mus spretus* would have colonized south-western Europe during the Middle Holocene, probably translocated by maritime transport through the Strait of Gibraltar (e.g., Domínguez García et al. 2019; Lalis et al. 2019). In some areas, prior to the arrival of the house mouse, some wild species, such as the Algerian mouse (*Mus spretus*) or wood mice (*Apodemus* spp.), are thought to have developed anthropophilic behaviours before being supplanted by the much more competitive house mouse (Cucchi et al. 2005; Weissbrod et al. 2017).

Concerning the palaeontological record of north-western Africa, studies published prior to the 2000s have mistakenly attributed the remains of the Algerian mouse (*Mus spretus*) to *Mus musculus*, contributing to the absence of ascertained identifications of the house mouse from securely dated archaeological sites in this region (Stoetzel 2013; Stoetzel et al. 2013). Moreover, the morphological proximity of the two species can sometimes complicate identifications, when based on conventional morphology alone, meaning further analyses are required, notably geometric morphometric methods (Cucchi et al. 2006, 2011; Michaux et al. 2007; Valenzuela-Lamas et al. 2011; Stoetzel et al. 2013; Moclán et al. 2023). In addition to the morphometric methods, palaeogenetic analyses have become a powerful tool for archaeological and palaeontological research in the last decades (e.g., Utge et al. 2020), particularly for small mammal records, and numerous studies have provided valuable insights concerning their taxonomy and phylogeography (e.g., Baca et al. 2019, 2020; Cucchi et al. 2020; García-Rodríguez et al. 2021; Alfaro-Ibañez et al. 2023).

A recent paper focusing on excavated layers dating to the 1st century BC from the site of Rirha in northern Morocco mentions the presence of both *M. spretus* and *M. m. domesticus* on the basis of conventional morphological characters (Oueslati et al. 2020). This would represent the first asserted record of the house mouse in a well-established archaeological context in the Maghreb, except that the identifications remain uncertain because of the method used.

The present study aims to validate the identification of *Mus* spp. remains from Rirha by applying a multidisciplinary approach using geometric morphometric (GMM) analyses of the first lower molar (m1) outline shape combined with Machine Learning (ML) models (Moclán et al. 2023). In addition, palaeogenetic analyses were performed to verify the morphometric-based taxonomic identification of the material, with the final objective being to better document the dispersal of the house mouse in south-western Mediterranean and its co-occurrence with the Algerian mouse.

MATERIALS AND METHODS

The site of Rirha

The site of Rirha (34°18'37"N 5°55'52"W) lies on an artificial mound, created by successive layers of human occupations through the centuries, in a meander of the wadi Beht in the Gharb plain (Fig. 1). Ongoing excavations have been undertaken every year since 2005 (Callegarin et al. 2016). The earliest Mauretanian levels have been dated by associated ceramics to the 5th century BC, and the occupation continued without interruption until the 1st century BC. Radiocarbon dating has provided ages of between ~ 410 and ~ 110 cal BC (Table 1). Excavation has revealed that the structures are directly overlaid by Roman walls.

The Mauretanian-era levels consist of a group of buildings constructed exclusively from unfired brick corresponding to a residential block. This sector has benefited from detailed excavation, systematic sampling for palaeoenvironmental studies and soil analysis through micro-morphology. Currently, levels older than the 2nd century BC are yet to be excavated; therefore, we can expect new data to appear from an older time span during successive campaigns.

Study material

The studied material comes from the Mauretanian Tell at Rirha. Small murid assemblages were sampled from well-stratified strata and the sediments were sieved using 0.5 mm mesh. The material was often associated with living floors or building and construction levels with characteristic mud brick architecture containing well-dated ceramics.

Although all the cranio-dental elements were counted, for this study we only focused on the molars – in archaeological and palaeontological contexts whole skulls and mandibles are rare or absent and cannot be used for reliable identification to species level. The first attempt at species attribution was made on the basis of molar morphological characters according to Darviche and Orsini (1982) and Darviche et al. (2006). The main characteristics to distinguish *M. spretus* and *M. m. domesticus* can be summarized as follows: the major axis of the t1 of the M1 is rather oblique in *M. m. domesticus*, while it is much less oblique and backwardly oriented in *M. spretus*; the anterior part of the m1 is trilobate in *M. m. domesticus* (tE reduced compared with tF) and tetralobate in *M. spretus* (tE well individualized); the presence of a cingulate margin with an accessory tubercule (c1) on the labial side of the m1 is frequent in *M. spretus* and rare in *M. m. domesticus*; on the m2, a cingulum (vestigial tE) is sometimes observed on the antero-labial base of the crown in *M. m. domesticus*. This preliminary identification work was published in Oueslati et al. (2020) and is here compared to our geometric morphometric and genetic data.

A modern reference sample was built, including 375 first lower molars (m1) of *M. m. domesticus* and *M. spretus* (Table 2) – m1 was chosen as it has been proved to be the most useful tooth for identifying mice taxa of the genus *Mus* (Cucchi et al. 2006; Darviche et al. 2006). We photographed specimens of different populations from south-western Europe (France, Iberian Spain, Canary Islands, Mallorca and Switzerland) and North Africa (Algeria, Morocco and Tunisia), stored in the collections of various institutions: Estación Biológica de Doñana (EBD - Sevilla, Spain), Institut des Sciences de l'Evolution de Montpellier (ISEM – Montpellier, France), Museo Nacional de Ciencias Naturales (MNCN – Madrid, Spain) and the Muséum national d'Histoire naturelle (MNHN – Paris, France), as well as those

captured during field campaigns in Morocco by the French ANR project "Modern Human installation in Morocco, influence on the small terrestrial vertebrate biodiversity and evolution" (MOHMIE), stored at the MNHN. Some of these specimens are genotyped (ISEM and MOHMIE), thus providing precise taxonomic identification. Classification of the reference material was carefully reviewed following the cranial and external criteria commonly used to discriminate between both species (Britton et al. 1976; Darviche et al. 2006; Darviche and Orsini 1982; Gerasimov et al. 1990). See supplementary material (Table S1) for detailed catalogue information for each specimen.

The archaeological material studied in this work consisted of 38 first lower molars (m1) from the site of Rirha (Table 2). Only complete and unaltered molars were selected. Broken teeth or those showing advanced tooth wear were excluded from the GMM analysis. Nine m1s were selected for palaeogenetic analyses.

Geometric morphometrics

To obtain reliable taxonomic identifications of the archaeological material assigned to *Mus* spp. from the site of Rirha, Geometric Morphometric (GMM) analysis was performed. To minimize wear-related variation, GMM techniques were applied to 2D images of the occlusal view to obtain the outline of each m1 included in this study (Renaud et al. 1996). To ensure consistency in the shape of outlines and the positioning of the landmark and semi-landmarks, particular attention was given to maintaining the occlusal surface of the tooth in a horizontal orientation during the photographic process (Fox et al. 2020). Following the procedure described by Cucchi et al. (2013, 2020) and Moclán et al. (2023), we digitized one landmark located at the furthermost point in the anterior axe and 63 sliding semi-landmarks equally spaced along the crown's external outline of the molar using tpsUtil v.1.76 (Rohlf 2018) and tpsDig v.2.32 (Rohlf 2010). Generalized Procrustes Analysis (GPA) was then carried out using the minimum bending-energy criterion for sliding semi-landmarks (Green 1996; Bookstein 1997) using tpsRelw v. 1.70 (Rohlf 2019). To explore the relationship between shape and taxonomy, Principal Component Analysis (PCA) was conducted on the Procrustes coordinates of each specimen, resulting in the PC scores being employed as molar shape variables in the subsequent analyses.

As previously shown (Courtenay et al. 2019; Bellin et al. 2021), ML algorithms are a powerful system of data classification when combined with GMM techniques, particularly in *Mus* classification tasks (Moclán et al. 2023). Here, we combined GMM data with ML techniques following the proposal of Moclán et al. (2023) to classify the archaeological material from Rirha. A complete description of the methods is available in the original paper of Moclán et al. (2023). The supplementary code was developed in R (R Core Team 2022) with 'caret' (Kuhn et al. 2020) and 'caretEnsemble' (Deane-Mayer and Knowles 2019; Deane-Mayer 2019) libraries.

A series of 11 algorithms were first trained with 82 PC scores obtained from the PCA – the PC scores were selected based on the number that did not exhibit collinearity. The algorithms used were Neural Networks (NNET), Linear Support Vector Machines (SVMl), Radial Support Vector Machines (SVMr), *k*-nearest neighbour (*k*NN), Logistic Regression (LG), Decision Trees using the 5.0 algorithm (DTC5.0), Random Forest (RF), Gradient Boosting (GB), Naïve Bayes (NB), Linear Discriminant Analysis (LDA) and Partial Least Squares (PLS). Ensemble learning (EL) techniques were then employed to train new meta-learners from the 'caret' models, as these techniques usually produce a better performance than "simple" base learners (Opitz and Maclin 1999; Dietterich 2000; Rokach 2010; Sagi and Rokach 2018). Thus, we trained a Generalized Linear Model (Ensemble: GLM) and three stacked models: a Neural Network, a Random Forest and a Gradient Boosting algorithm.

The training process was developed using a typical sequence – dividing the sample into training (70%) and testing (30%) subsamples and using cross-validation as the control method (base learners = 10-folds, 10 repeats; meta-learners = 11-folds, 10 repeats). Accuracy and kappa values were considered in combination with sensitivity, specificity and balanced accuracy to evaluate the algorithm performance. Note that as we have improved the modern sample of *Mus* in comparison with Moclán et al. (2023), we have also retrained the models using the 'tunelengh = 20' function of 'caret'.

The final classification of *Mus* remains from Rirha was achieved using a combination of the results provided by the ensemble/stacked models following the criteria of Moclán et al. (2023). As a further precaution, we classified the difference between those samples with a posterior probability value higher than 0.9 (Cucchi et al. 2013).

Palaeogenetic analyses

The methodology used for the taxonomic identification of archaeological remains was based on the use of real-time Polymerase Chain reactions (PCRs) and TaqMan MG probes (Kutyavin et al. 2000), which allowed an amplification of the specific and diagnostic genomic regions for the determination or comparison between the study species (Elalouf 2022; Utge et al. 2020). Compared to conventional PCRs based on the use of primers targeting the amplified DNA region, the TaqMan method also includes a probe that hybridizes to the recopied DNA fragment. The DNA polymerase, which copies the DNA from the primer, has a 5'–3' exonuclease activity that cleaves the probe. Once this cleavage has taken place, the suppressor in the probe is far enough away from the fluorophore for the latter to emit a signal when agitated. The advantage of this method is the increased specificity compared with a simple PCR. While oligonucleotides allow the amplification of fragments other than the one required, these undesirable fragments will not hybridize the probe and therefore no spurious signal will be produced. In principle, this approach should detect only the desired fragment, thus avoiding the need for subsequent manipulations (gel purification, cloning, sequencing).

Primers and probes for the TaqMan MG assays (Table 3) were designed with the help of Geneious Primes 2021 software. In total, 95 reference sequences of the cytochrome B (cytB) for the genus *Mus*, within different haplotypes of *M. spretus* and *M. m. domesticus*, were aligned and used as a reference to obtain the TaqMan assays (Supplement material, Table S2).

Validation of the assays was conducted by qPCR with the DNA extracted from the tissue samples, preserved in alcohol, of contemporaneous specimens from the MNHN collection. The tissue sample of *M. m. domesticus* (ID: SPOT1308) came from a specimen collected in Marcq-en- Bareuil (France), and the tissue sample of *M. spretus* (ID: 21,001) came from a specimen collected in Montbazin (France). The DNA of both samples were extracted using the NucleoSpin Tissue kit (Macherey Nagel), following the steps suggested by the manufacturer and with a final elution of 100 µl. The concentrations of DNA obtained were measured with a Nanodrop 2000 (ThermoScientific) (Table 4).

The specificity and efficiency of the probe were then measured. The specificity of the probes was tested in a real-time PCR using the DNA of both samples with the two TaqMan assays. The efficiency was tested in a real-time PCR using, for each TaqMan probe, a DNA dilution series of 1:10 covering five dilution points for each sample. The real-time PCR was carried out in a 20-µl reaction volume containing 10 µl of 2X TaqMan fast advanced master mix (Thermo Fisher Scientific), 900 nM of forward and reverse primers, 250 nM of TaqMan probe, and water (PCR blank), with the DNA diluted extracts. Amplification was performed in a BioRad CFX96 Touch Real-Time PCR Detection System (Bio-Rad Laboratoires) in 50 °C, 2 min; 95 °C, 20 s; then 40 PCR cycles (95 °C, 3 s; 60 °C, 30 s). Data were analysed using CFX Maestro Software (Bio-Rad Laboratories) set with default parameters to determine the cycle threshold (*C*T), i.e., the number of PCR cycles required for the fluorescent signal to exceed the background level. Only samples with a CT value below or equal to 37 cycles were considered positive for DNA content.

Once the assays were verified, DNA extraction and the pre-PCR steps using the nine archaeological rodent remains were performed in a laboratory dedicated to ancient DNA studies – the Plateau de Paléogénomique et de Génétique Moléculaire at the Musée de l'Homme (P2GM platform). Ancient DNA (aDNA) samples were extracted from rodent hemi-mandibles. The total quantity of material per specimen ranged from 10 to 40 mg. Lysis was carried out by digesting the extracted complete teeth (without prior grinding), placing them in 1 ml of DNA extraction buffer (0.25 M EDTA, 10 mM Tris-EDTA (pH 8), 0.2% N-lauryl-sarcosyl, 200 µg/ml proteinase K) and then incubating for 16 h at 42 °C under constant agitation. The tubes were then centrifuged (5,000 g; 5 min) in an Eppendorf microcentrifuge, and the supernatant was transferred into an Amicon ultra- 0.5 ml 30 kDa centrifugal filter unit

(Millipore; Burlington, MA, USA). The filter unit was centrifuged (14,000 g; 7 min), rinsed 4 times by adding 500 µl of ultrapure DNase/ RNase-free distilled water (Thermo Fisher Scientific) and centrifuged again (14,000 g; 7 min). The DNA extract was recovered as a ~ 50-µl sample volume by centrifugation (1,000 g; 1 min) of the filter unit in the reverse position. The extract was further purified using a Qiagen minelute PCR purification kit (Qiagen; Venlo, Netherlands), according to the manufacturer's instructions, and eluted in 60 µl of 10 mM Tris, pH 8. Quantification of the extracted DNA was performed using a Qubit HS kit (Invitrogen) (Table 5).

Real-time PCR was carried out in a 20-µl reaction volume, as before, using the TaqMan fast advanced master mix, which contains dUTP (instead of dTTP) and uracil- N-glycosylase to prevent carryover contamination from one experiment to another. For each sample, 3 to 4 serial dilutions corresponding to between 2 and 0.03 µl of the DNA extract were analyzed for each TaqMan assay. Only samples for which a CT value below or equal to 37 cycles were considered positive and the taxonomic identification was confirmed.

RESULTS

Geometric morphometrics and machine learning

The PCA applied to the shape data of modern samples and the archaeological material from Rirha shows that the first two principal components (PC1-PC2) captured 59.1% of the shape variation, providing a separation between both species with some overlapping (Fig. 2). The major shape difference in this plot is the relative development of t1 in the anterior part of the tooth. *M*. *m. domesticus* specimens are mainly located at the right part of the plot, showing positive values in PC1, while *M. spretus* clusters at the left part of the *x*-axis with mainly negative values. The archaeological specimens from Rirha show a non-clear distribution overlapping with both modern subsamples.

The results obtained from the ML algorithm training (Table 6) show high-classification ratios (Accuracy > 0.9; kappa = > 0.8) for the modern testing subsample. The highest accuracy value is shown by LDA (0.973), while the lowest is presented by SVMl (0.90) and DTC5.0, GB and NB (0.91). In general, the values of sensitivity (0.852–0.963) and specificity (0.912–1) are also very high. The final hyperparameter configuration of each algorithm is shown in Table 7.

When the correlation among algorithmic classification was tested (Table 8), the results showed that a general correlation between classifications is a rarity. Those correlations identified by the test were found between SVMl and LG (r = 0.77), SVMl and LDA (r = 0.58), LDA and LG (r = 0.61) and RF and GB (r = 0.78). These results allow us to consider ensemble and stacking techniques as extremely useful for classifying the archaeological samples.

The training process of the meta-learners provided extremely similar results to those delivered by the base learners. The lowest accuracy value is shown in the stacked RF (accuracy = 0.946; kappa = 0.892), while the highest value was provided by the ensembled GLM (accuracy = 0.964; kappa = 0.928). All the results provided perfect specificity values, high sensitivity values (0.88–0.926) and balanced accuracy values (0.944–0.963). In this sense, the highest accuracy value provided by a base learner had to be discarded (LDA accuracy = 0.973) due to the possible presence of overfitting in that specific model; therefore, the ensemble/stacked models were used to classify the archaeological samples.

The taxonomic assignment of the m1s from Rirha provided similar results among the different ensemble/stacked models (Table 9). The ensembled GLM indicated the presence of 16 *M. spretus* specimens against 22 *M. m. domesticus* ones, out of the 38 analysed specimens. On the other hand, the stacked models highlighted the presence of 17 *M. spretus* and 21 *M. m. domesticus*. In all models, the vast majority of molars analysed were classified to one or another species with posterior probabilities (p) above the 0.9 threshold (GLM = 97.37%; NNET and GB = 100%; RF = 92.11%). Even those specimens classified with $p < 0.9$ by some models were assigned by at least two to the same species with $p > 0.9$. Discrepancies between models were found in only one specimen (RHA06 US109 2), which was classified as *M*. *m*. *domesticus* by the ensembled GLM (p = 1), whereas stacking models assigned it to *M*. *spretus* (p = 0.79–0.99); it should therefore be considered as indeterminate.

To summarize the results of the ML classification, a total of 21 specimens were reliably assigned to *M. m. domesticus*, and 16 were classified as *M. spretus*, while just one specimen should be considered indeterminate. Considering the laterality of the analysed specimens, these belong to a minimum number of individuals (MNI) of 12 *M*. *m*. *domesticus* and 9 *M*. *spretus*.

Palaeogenetic analyses

DNA obtained from each of the nine archaeological samples was used in a real-time PCR where taxonomic identification was performed using the two proposed TaqMan assays. The results obtained from the qPCR (Table 10) show a strong correlation between GMM + ML and genetic determination. Four specimens had the same taxonomic assignment, while one specimen had a conflicting assignment. Three samples were not taxonomically determined using the proposed assays, as there was no qPCR amplification of the chosen citB gene region, possibly due to poor DNA preservation.

DISCUSSION AND CONCLUSION

This study is the first attempt to approach combining GMM + ML and palaeogenetics for the taxonomic identification of the genus *Mus* in the western Mediterranean region, following the model developed by Cucchi et al. (2020) for the Near East. It has provided a new case supporting the usefulness of this procedure in studying fossil small mammals.

The presence of both species, *M. spretus* and *M. m. domesticus*, at Rirha is clearly confirmed by GMM combined with ML, as well as palaeogenetics. GMM + ML analyses provided high-accuracy values for discrimination between both species (0.946–0.955, Table 6). These results are slightly lower than those achieved by Moclán et al. (2023) who reported a perfect classification of the testing/modern samples with a stacked NNET. However, this decrease in accuracy seems to be caused by the increase in the number of modern mice m1s employed in this study compared to the previous one. An effort has been made to add 72 new *Mus* spp. remains, including new populations such as *Mus spretus* from Algeria and Tunisia, and *Mus m*. *domesticus* from Switzerland. At the same time, the sample size of populations with low specimen numbers (*M*. *spretus* and *M*. *m*. *domesticus* from France) was increased. This approach, therefore, provided a greater morphometric variability, allowing us to improve the validity/applicability of the models. Nonetheless, the high accuracy values obtained are similar or even greater than those obtained by previous studies applying elliptic Fourier analysis to the outline shape and cross-validated linear discriminant analysis for classification (Michaux et al. 2007; Valenzuela-Lamas et al. 2011; Stoetzel et al. 2013), employing lower sample sizes for the modern referential dataset in some cases (Valenzuela-Lamas et al. 2011; Stoetzel et al. 2013).

Palaeogenetic analysis provided a taxonomic identification for one specimen that was excluded from the GMM + ML analysis (RHA05-US0029-5, Table 10), since the m1 was partially broken and displayed signs of digestion. However, another excluded specimen (RHA06- US104-3), which also showed signs of light digestion, was not identified by palaeogenetics. The remaining specimens that did not provide positive results by aDNA analysis (RHA06-US104-2, RHA06-US109-1) were classified as *Mus m*. *domesticus* by GMM + ML according to their m1 shape, as those molars were complete and displayed a good preservation status.

Overall, the two approaches concur, evidencing the high reliability of these methods for taxonomic identification. Among the analysed teeth (some examples are shown in Fig. 3), only two are problematic: RHA06 US109-2 (Fig. 3e), which was classified as indeterminate by GMM + ML and for which no palaeogenetic analysis was performed; and RHA06 US1106-1 (Fig. 3f), which showed discrepancies between the GMM + ML and palaeogenetics assignments (Table 10). This indicates that the classification achieved by GMM + ML is not perfect. Despite the high accuracy of ensemble/stacking models employed, low misclassification percentages (3.6–5.4%) were also obtained due to the elevated intraspecific variability of the two species (Darviche and Orsini 1982; Cucchi et al. 2002). If we analyse the morphology of this m1 (RHA06 US1106-1) in a traditional way, it displays a

moderately reduced t1 and a cingular margin showing an incipient c1 (Fig. 3f), which is rare in *M. m. domesticus* populations (Darviche et al. 2006). However, the modern reference database includes similar morphotypes, meaning that traditional morphological criteria must be used with caution. In addition, it is striking that all models provided a posterior probability of 1 for assignment as *M. spretus* for RHA06 US1106-1 (Table 9). Another methodological issue that can explain this discrepancy is related to the incorrect positioning of the tooth when taking the picture. Although special care was taken in this regard, since changes in tooth orientation may lead to shape alterations (Fox et al. 2020), this tooth seems to have partially shifted to its anterior region. However, although we considered that this small shift in tooth orientation had not caused any significant changes in the shape of the tooth outline, this possibility cannot be ruled out. In addition, the possible hybridization between *M. spretus* and *M. domesticus* (Orth et al. 2002; Banker et al. 2022), a phenomenon with poorly known effects on the phenotype (notably the dental morphology), may also explain the discrepancy between morphometric and genetic results in this specimen. This possibility could represent a new research direction. In conclusion, this single case of discrepancy in the classification observed between GMM + ML and palaeogenetics represents a challenging result, reinforcing the complexity in discriminating between sibling species characterized by overlapping morphometric traits. These results also highlight the importance of some methodological considerations such as the low misclassification rate or accurate tooth orientation, as well as the potential influence of the hybridization process between both species.

The site of Rirha indicates the existence of a strongly anthropized context, since it comprises a succession of consecutive settlements from the Mauretanian period to Roman times, and two medieval phases (Oueslati et al. 2020). The house mouse would have already benefited from the presence of a commensal niche at Rirha since 390–200 cal BC, providing a food supply and protection against climatic variations and predation. This is also consistent with the presence at the site of other commensal rodents such as *Rattus rattus* and the domestic exploitation of chickens since the first half of the 1st century BC (Oueslati et al. 2020). In addition, the accumulation of mouse remains and other small vertebrates including *Lemniscomys barbarus*, shrews and reptiles is linked to predation, probably by domestic cats (Oueslati et al. 2020). These wild taxa are represented in very low abundance in the assemblage, which also shows a lower diversity compared to the modern natural community of Morocco (Mediani et al. 2015; Aulagnier 2017). Following a comparison of both classification approaches, our sample shows a higher abundance of the house mouse (MNI = 13; %MNI = 61.9%) compared to the Algerian mouse (MNI = 8; %MNI = 39.1%). This abundance pattern can be attributed to their distinct ecological behaviours: *M. m. domesticus* is a typically commensal taxon, inhabiting areas near or within human settlements; conversely, *M. spretus*, while still benefitting from anthropogenic environments, tends to avoid direct proximity to human settlements (Denys et al. 2017). This ecological distinction appears to result in a gradual displacement of the Algerian mouse from the human-made habitats in the Rirha site and its immediate surroundings.

The multidisciplinary approach adopted by this study allowed us to confirm the first asserted record of the house mouse in a well-established archaeological context in the Maghreb. However, European genetic and archaeological data assume that the arrival of *M. m. domesticus* in the western Mediterranean should be dated to at least the 1st millennium BC (Bonhomme et al. 2011; Cucchi et al. 2005, 2020), much earlier than the Rirha occupations (1st − 4th century BC). We must therefore continue our investigations by applying a similar multidisciplinary approach to the older levels of Rirha, as well as other sites dated to earlier chrono-cultural periods (from the Phoenician up to the Neolithic), with well-established archaeological, stratigraphic and chronological contexts, to verify this hypothesis in north-western Africa.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s12520- 024-02002-8.

Acknowledgements The Rirha research project, directed since 2017 by Mohamed Kbiri Alaoui, Elsa Rocca and Charlotte Carrato, is supported by INSAP Rabat, Morocco, French Ministry of Foreign Affairs, Casa de Velázquez, Archéologie des Sociétés Méditerranéennes UMR5140 French laboratory, Université Paul Valéry Montpellier 3 and LabEx ARCHIMEDE (from "Investissement d'Avenir" program ANR-11-LABX-0032-01). We would like to thank the other members of the "SOuMed" project for their scientific support: Thomas Cucchi, Xavier Gallet, Loïc Lebreton, Matthieu Lebon, Violaine Nicolas-Colin, Olivier Tombret and Antoine Zazzo. We would also like to thank Thomas Cucchi for sharing pictures of *Mus* molars, and the curators of the collections who gave us access to modern mice material from different institutions, including the Estación Biológica de Doñana (Seville, Spain), the Museo Nacional de Ciencias Naturales (Madrid, Spain), the Muséum national d'Histoire naturelle (Paris, France) and the Institut des Sciences de l'Évolution de Montpellier (France). We thank Jill Cucchi (Editing and Translation Services) for copy-editing the English (financed by the PAST team – UMR 7194 HNHP). Finally, we would like to thank the two anonymous reviewers for their careful reading of our manuscript and their insightful comments.

Author contributions A.C.D.G., J.U. and E.S. designed, conceptualized the research, and wrote the original draft. A.C.D.G., J.U., C.L. and A.M. performed the analyses. T.O., L.C., C.A.D.C. and E.R. carefully proofread the article and helped to write certain sections. T.O. conducted the zooarchaeological identification of micromammals and provided the photos for the GMM analysis. M.K.A., L.C., C.A.D.C. and E.R. are the directors of the excavation team at the Rirha site. E.S. supervised and coordinated the project. All authors approved the final version of the manuscript and agree to be responsible for its content.

Funding ACD-G is supported by a Postdoctoral Margarita Salas Contract (CT18/22) funded by the European Union "NextGenerationEU/ PRTR". This study was carried out within the framework of the following projects: "SOuMed - Approche pluridisciplinaire de la diffusion des souris commensales et sauvages dans l'Ouest de la Méditerranée" (ES dir.) and "ModelPal - Modélisation de la conservation de molécules organiques dans les restes archéologiques du paléolithique" (JU dir.), both founded by the Département Homme et Environnement of the Muséum national d'Histoire naturelle, Paris, France, as well as the project "Documenting the introduction and dispersal process of mice in the Western Mediterranean Region" (SYNTHESY + FR-TAF_ Call4_007, ACD-G dir.).

Data availability All data supporting our results are given in the manuscript and in the supplementary material.

Declarations
Competing interests The authors declare no competing interests.

REFERENCES

Alfaro-Ibáñez MP, Cuenca-Bescós G, Bover P et al (2023) Implications of population changes among the Arvicolinae (Rodentia, Mammalia) in El Mirón Cave (Cantabria, Spain) for the climate of the last c. 50,000 years. Quat Sci Rev 315:108234. https://doi. org/10.1016/j.quascirev.2023.108234

Álvarez-Vena A, Marín-Arroyo AB, Álvarez-Lao DJ et al (2023) Mammalian turnover as an indicator of climatic and anthropogenic landscape modification: a new Meghalayan record (late holocene) in northern Iberia. Palaeogeogr Palaeoclimatol Palaeoecol 616:111476. https://doi.org/10.1016/j.palaeo.2023.111476

Aulagnier S, Cuzin F, Thévenot M (2017) Mammifères sauvages du Maroc: peuplement, repartition, ecologie. Société Française pour l'Etude et la Protection des Mammifères (SFEPM), Toulouse, France

Baca M, Popović D, Lemanik A et al (2019) Highly divergent lineage of narrow-headed Vole from the late Pleistocene Europe. Sci Rep 9:17799. https://doi.org/10.1038/s41598-019-53937-1

Baca M, Popović D, Baca K et al (2020) Diverse responses of common Vole (*Microtus arvalis*) populations to late glacial and early Holocene climate changes – evidence from ancient DNA. Quat Sci Rev 233:106239. https://doi.org/10.1016/j.quascirev.2020.106239

Banker SE, Bonhomme F, Nachman MW (2022) Bidirectional introgression between *Mus musculus domesticus* and *Mus spretus*. Genome Biol Evol 14(1):evab288

Bellin N, Calzolari M, Callegari E et al (2021) Geometric morphometrics and machine learning as tools for the identification of sibling mosquito species of the Maculipennis complex (*Anopheles*). Infect Genet Evol 95:105034. https://doi.org/10.1016/j. meegid.2021.105034

Bonhomme F, Orth A, Cucchi T et al (2011) Genetic differentiation of the house mouse around the Mediterranean basin: matrilineal footprints of early and late colonization. Proc R Soc B: Biol Sci 278:1034–1043. https://doi.org/10.1098/rspb.2010.1228

Bookstein FL (1997) Landmark methods for forms without landmarks: morphometrics of group differences in outline shape. Med Image Anal 1:225–243

Britton J, Pasteur N, Thaler L (1976) Les Souris Du Midi De La France: caractérisation génetique de deux groupes de populations sympatriques. C R Acad Sci 258:515–518

Bronk Ramsey C Bayesian analysis of radiocarbon dates. Radiocarbon, 51(1): 337–360., Alaoui L, Ichkhakh MK, Roux AA (2009) JC (2016). Rirha: site antique et médiéval du Maroc. I Cadre historique et géographique général, Collection de la Casa de Velázquez no 150, Madrid

R Core Team (2022) R: A language and environment for statistical computing

Courtenay LA, Yravedra J, Huguet R et al (2019) Combining machine learning algorithms and geometric morphometrics: a study of Carnivore tooth marks. Palaeogeogr Palaeoclimatol Palaeoecol 522:28–39. https://doi.org/10.1016/j.palaeo.2019.03.007

Cucchi T, Vigne J-D, Auffray J-C et al (2002) Introduction involontaire de la souris domestique (*Mus musculus domesticus*) à Chypre dès Le Néolithique précéramique ancien (fin IXe et VIIIe millénaires av. J.-C). C R Palevol 1:235–241. https://doi.org/10.1016/ S1631-0683(02)00033-7

Cucchi T, Vigne J-D, Auffray J-C (2005) First occurrence of the house mouse (*Mus musculus domesticus* Schwarz & Schwarz, 1943) in the western Mediterranean: a zooarchaeological revision of subfossil occurrences. Biol J Linn Soc 84:429–445. https://doi. org/10.1111/j.1095-8312.2005.00445.x

Cucchi T, Orth A, Auffray J-C et al (2006) A new endemic species of the subgenus *Mus* (Rodentia, Mammalia) on the island of Cyprus. Zootaxa 1241:1–36. https://doi.org/10.11646/zootaxa.1241.1.1

Cucchi T, Bălăşescu A, Bem C et al (2011) New insights into the invasive process of the eastern house mouse (*Mus musculus musculus*): evidence from the burnt houses of Chalcolithic Romania. Holocene 21:1195–1202. https://doi.org/10.1177/0959683611405233

Cucchi T, Kovács ZE, Berthon R et al (2013) On the trail of neolithic mice and men towards Transcaucasia: zooarchaeological clues from Nakhchivan (Azerbaijan). Biol J Linn Soc 108:917–928. https://doi.org/10.1111/bij.12004

Cucchi T, Papayianni K, Cersoy S et al (2020) Tracking the Near Eastern origins and European dispersal of the western house mouse. Sci Rep 10:8276. https://doi.org/10.1038/s41598-020-64939-9

Darviche D, Orsini P (1982) Critères de différenciation morphologique et biométrique de deux espèces de souris sympatriques: *Mus spretus* Et *Mus musculus domesticus*. Mammalia 46:205–218. https://doi.org/10.1515/mamm.1982.46.2.205

Darviche D, Orth A, Michaux J (2006) Mus spretus Et M. musculus (Rodentia, Mammalia) en zone méditerranéenne: différenciation biométrique et morphologique: application à Des fossiles marocains pléistocènes. Mammalia 70:90–97. https://doi. org/10.1515/MAMM.2006.010

Deane-Mayer ZA (2019) A Brief Introduction to caretEnsemble. In: CRAN. https://cran.r-project.org/web/packages/caretEnsemble/ vignettes/caretEnsemble-intro.html. Accessed 12 Nov 2021

Deane-Mayer ZA, Knowles JE (2019) Package 'caretEnsemble&#8217

Denys C, Taylor P, Aplin K (2017) Family Muridae (true mice and rats, gerbils and relatives). In: Wilson DE, Lacher TE, Mittermeier RA (eds) Handbook of the mammals of the World, vol 7. Rodents II. Lynx Edicionas, Barcelona, pp 204–535

Dietterich TG (2000) Ensemble methods in machine learning. Multiple Classifier systems. Springer, Berlin, Heidelberg, pp 1–15

Dobson M (1998) Mammal distributions in the western Mediterranean: the role of human intervention. Mamm Rev 28:77–88

Domínguez-García ÁC, Laplana C, Sevilla P et al (2019) New data on the introduction and dispersal process of small mammals in southwestern Europe during the Holocene: Castillejo Del Bonete site (southeastern Spain). Quat Sci Rev 225:106008. https://doi.org/10.1016/j.quascirev.2019.106008

Elalouf J-M, Palacio P, Bon C, Berthonaud V, Maksud F, Stafford TW, Hitte C (2022) The genome and diet of a 35,000-year-old *Canis lupus* specimen from the palaeolithic painted cave, Chauvet- Pont d'Arc, France. Ecol Evol 12:e9238. https://doi.org/10.1002/ece3.9238

Fox NS, Veneracion JJ, Blois JL (2020) Are geometric morphometric analyses replicable? Evaluating landmark measurement error and its impact on extant and fossil *Microtus* classification. Ecol Evol 10:3260–3275. https://doi.org/10.1002/ece3.6063

García-Rodríguez O, Hardouin EA, Hambleton E et al (2021) Ancient mitochondrial DNA connects house mice in the British Isles to trade across Europe over three millennia. BMC Ecol Evol 21(1):9. https://doi.org/10.1186/s12862-021-01746-4

Gerasimov S, Nikolov H, Mihailova V et al (1990) Morphometric stepwise discriminant analysis of the five genetically determined European taxa of the Genus *Mus*. Biol J Linn Soc 41:47–64. https://doi.org/10.1111/j.1095-8312.1990.tb00820.x

Grayson DK (2001) The archaeological record of human impacts on animal populations. J World Prehist 15(1):1–68. https://doi.org/1 0.1023/A:1011165119141

Green WDK (1996) The thin-plate spline and images with curving features. In: Mardia KV, Gill CA, Dryden IL (eds) image fusion and shape variability. University of Leeds Press, Leeds, pp 79–87

Hulme-Beaman A, Dobney K, Cucchi T, Searle JB (2016) An Ecological and Evolutionary Framework for Commensalism in Anthropogenic environments. Trends Ecol Evol 31(8):633–645

Jones EP, Eager HM, Gabriel SI et al (2013) Genetic tracking of mice and other bioproxies to infer human history. Trends Genet 29(5):298–308. https://doi.org/10.1016/j.tig.2012.11.011

Kuhn M, Wing J, Weston S et al (2020) caret: Classification and Regression Training

Kutyavin IV, Afonina IA, Mills A, Gorn VV, Lukhtanov EA, Belousov ES, Singer MJ, Walburger DK, Lokhov SG, Gall AA, Dempcy R, Reed MW, Meyer RB, Hedgpeth J (2000) 3'-minor groove binder-DNA probes increase sequence specificity at PCR extension temperatures. Nucleic Acids Res 28(2):655–661. https://doi.org/10.1093/nar/28.2.655PMID: 10606668; PMCID: PMC102528

Lalis A, Mona S, Stoetzel E et al (2019) Out of Africa: demographic and colonization history of the Algerian mouse (*Mus spretus* Lataste). Heredity 122:150–171. https://doi.org/10.1038/s41437-018-0089-7

Mediani M, Brito JC, Fahd S (2015) Atlas of the amphibians and reptiles of northern Morocco: updated distribution and patterns of habitat selection. Basic Appl Herpeto 29:81–107

Michaux J, Cucchi T, Renaud S et al (2007) Evolution of an invasive rodent on an archipelago as revealed by molar shape analysis: the house mouse in the Canary Islands. J Biogeogr 34:1412–1425. https://doi.org/10.1111/j.1365-2699.2007.01701.x

Moclán A, Domínguez-García ÁC, Stoetzel E et al (2023) Machine learning interspecific identification of mouse first lower molars (Genus *Mus* Linnaeus, 1758) and application to fossil remains from the Estrecho Cave (Spain). Quat Sci Rev 299:107877. https://doi.org/10.1016/j.quascirev.2022.107877

Opitz D, Maclin R (1999) Popular Ensemble methods: an empirical study. J Artif Intell Res 11:169–198. https://doi.org/10.1613/jair.614

Orth A, Belkhir K, Britton-Davidian J et al (2002) Hybridation naturelle entre deux espèces sympatriques de souris *Mus musculus domesticus* L. Et *Mus spretus* Lataste. C R Biologies 325:89– 97. https://doi.org/10.1016/S1631-0691(02)01413-0

Oueslati T, Kbiri Alaoui M, Ichkhakh A et al (2020) 1st century BCE occurrence of chicken, house mouse and black rat in Morocco: Socio-economic changes around the reign of Juba II on the site of Rirha. J Archaeol Sci Rep 29:102162. https://doi.org/10.1016/j.jasrep.2019.102162

Reimer P, Austin W, Bard E et al (2020) The IntCal20 Northern Hemisphere radiocarbon age calibration curve (0–55 cal kBP). Radiocarbon 62(4):725–757. https://doi.org/10.1017/RDC.2020.41

Renaud S, Michaux J, Jaeger J-J, Auffray J-C (1996) Fourier analysis applied to *Stephanomys* (Rodentia, Muridae) molars: nonprogressive evolutionary pattern in a gradual lineage. Paleobiology 22:255–265

Rohlf FJ (2010) tpsDig

Rohlf FJ (2018) tpsUtil

Rohlf FJ (2019) tps Relative Warps

Rokach L (2010) Ensemble-based classifiers. Artif Intell Rev 33:1–39. https://doi.org/10.1007/s10462-009-9124-7

Sagi O, Rokach L (2018) Ensemble learning: a survey. WIREs Data Min Knowl Discov 8:e1249. https://doi.org/10.1002/widm.1249

Stoetzel E (2013) Late cenozoic micromammal biochronology of northwestern Africa. Palaeogeogr Palaeoclimatol Palaeoecol 392:359–381

Stoetzel E, Denys C, Michaux J, Renaud S (2013) *Mus* in Morocco: a quaternary sequence of intraspecific evolution. Biol J Linn Soc 109:599–621. https://doi.org/10.1111/bij.12065

Utge J, Sévêque N, Lartigot-Campin AS, Testu A, Moigne AM et al (2020) A mobile laboratory for ancient DNA analysis. PLoS ONE 15(3):e0230496. https://doi.org/10.1371/journal.pone.0230496

Valenzuela-Lamas S, Baylac M, Cucchi T, Vigne J-D (2011) House mouse dispersal in Iron Age Spain: a geometric morphometrics appraisal. Biol J Linn Soc 102:483–497. https://doi.org/10.1111/j.1095-8312.2010.01603.x

Weissbrod L, Marshall FB, Valla FR et al (2017) Origins of house mice in ecological niches created by settled hunter-gatherers in the Levant 15,000 y ago. PNAS 114:4099–4104. https://doi.org/10.1073/pnas.1619137114

FIGURES

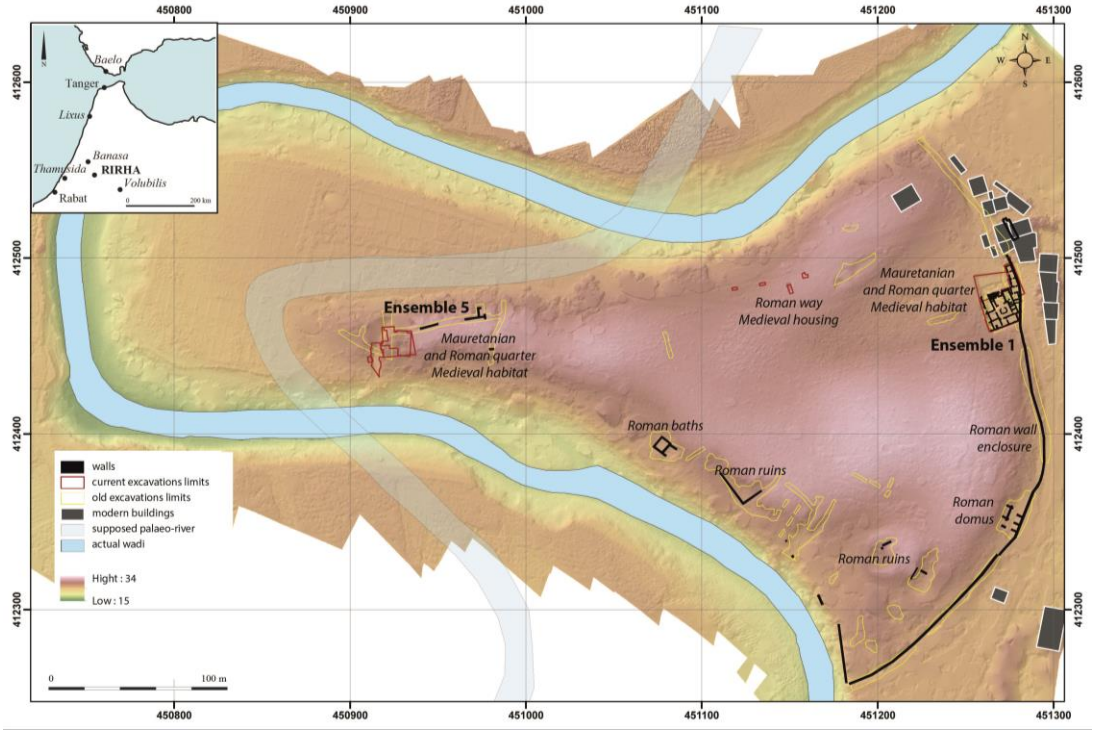

Figure 1. Location of the Rirha site in northern Morocco (top left) and plan of the remains with location of the main Mauretanian quarter 'Ensemble 5' (© Rirha Project).

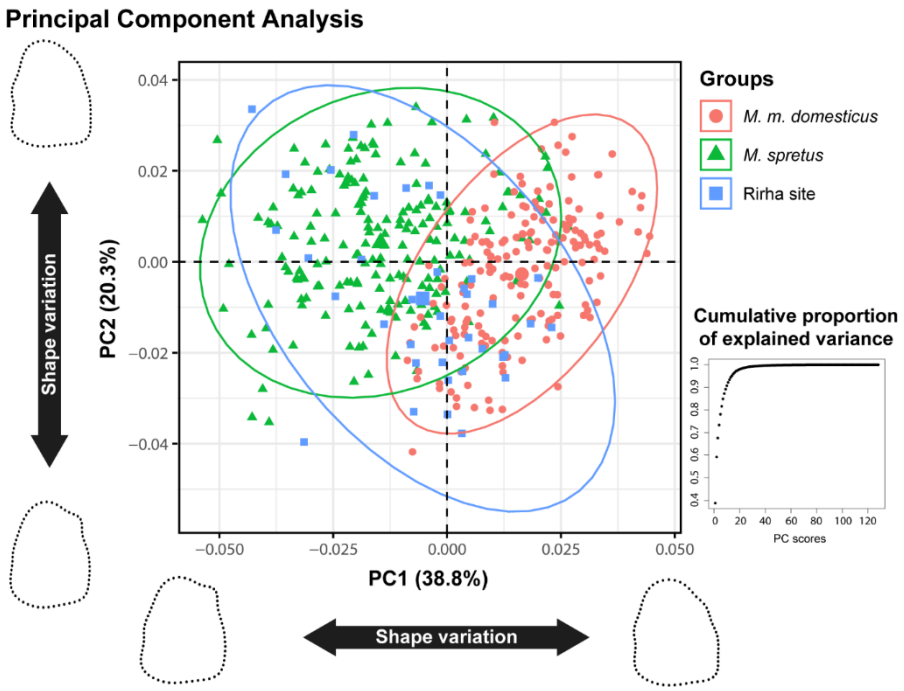

Figure 2. Biplot of the first two Principal Components of the modern samples and the specimens from Rirha.

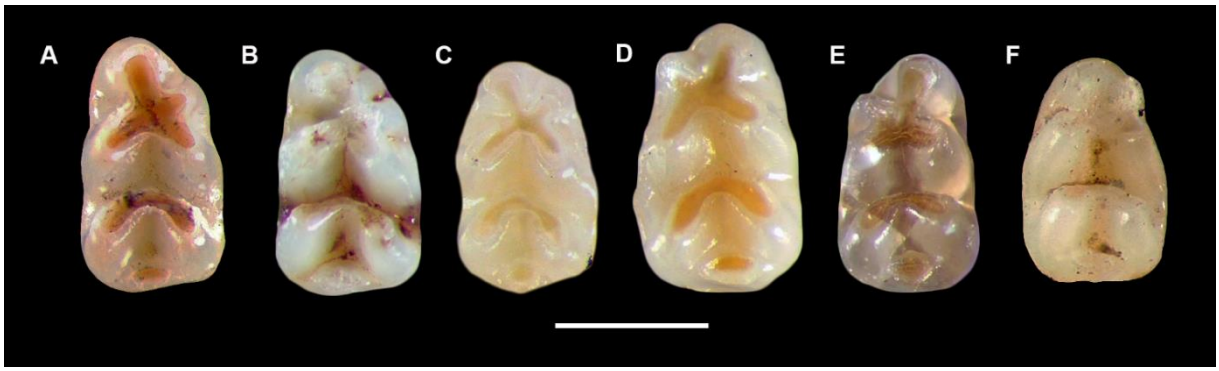

Figure 3. Selected *Mus* spp. m1 from Rirha in occlusal view. *Mus musculus domesticus*: a) RHA07-US5143-16, b) RHA07 US5143-6, f) RHA06-US1106-1; *Mus spretus*: c) RHA06-US5044-3, d) RHA06 US5044-1; indeterminate (GMM+ML, palaeogenetic analysis not performed): e) RHA06 US109-2; coincident classification between GMM+ML and palaeogenetic: a) and c); palaeogenetic analyses not performed: b) and d); discrepancy between GMM+ML and palaeogenetic identifications: f). Scale bar = 1 mm.

TABLES

Samples	Non calibrated ^{14}C ages	Calibrated ^{14}C ages	
		68.3% probability	95.4% probability
RHA 06 US 109	2240±30BP	380 (19.3%) 352 cal BC	390 (26.1%) 343 cal BC
		290 (42.7%) 227 cal BC	320 (69.4%) 202 cal BC
		221 (6.2%) 210 cal BC	
RHA 08 US 184	2295±30BP	400 (60%) 361 cal BC	407 (65.3%) 352 cal BC
		272 (4.7%) 266 cal BC	289 (28.6%) 227 cal BC
		241 (3.5%) 236 cal BC	221 (1.6%) 210 cal BC
RHA 08 US 5182	2175±30BP	351 (41.6%) 290 cal BC	361 (92.0%) 147 cal BC
		209 (26.7%) 171 cal BC	138 (3.5%) 110 cal BC

Table 1. Radiocarbon ages obtained for Rirha (Poznań Radiocarbon Laboratory). Here, intervals of calendar age are provided, where the true ages of the samples encompass the probability of ca. 68% and ca. 95%. Calibration was made using the Intcal 20 curve of atmospheric data (Reimer et al. 2020) and OxCal v4.4 (Bronk Ramsey 2009).

Population	Collection	*n*
Modern *Mus musculus domesticus*		
Gran Canaria	EBD, MNCN	14
Lanzarote	EBD, MNCN	11
Iberian Peninsula	EBD, MNCN	61
France	ISEM, MNHN	28
Algeria	ISEM	17
Morocco	ISEM	36
Switzerland	MNHN	16
TOTAL		183
Modern *Mus spretus*		
Mallorca	MNCN	10
Iberian Peninsula	EBD, MNCN	57
France	ISEM, MNHN	23
Algeria	MNHN	8
Morocco	ISEM, MOHMIE	65
Tunisia	MNHN	29
TOTAL		192
Archaeological material		
Rirha site		38

Table 2. Modern reference and archaeological samples used in this study, according to species, populations (location), collections and number of specimens studied (*n*).

	Oligonucleotide sequence (5'-3')	Target, size (bp)
TaqMan assay - *Mus spretus*		
Primer Forward	CGA TTC TTC GCY TTC CAC TTC	CYTB, 74
Primer Reverse	TGG AGR AAR AGG AGG TGW ACG	
Probe	[FAM] TGC TAG GGC TGC GAT GAT GAA TGG CA [MGBEQ]	
TaqMan assay - *Mus musculus domesticus*		
Primer Forward	CGA TTC TTC GCY TTC CAC TTC	CYTB, 74
Primer Reverse	TGG AGG AAR AGG AGG TGA ACG	
Probe	[FAM] TGC TAG GGC CGC GAT AAT AAA TGG TAA G [MGBEQ]	

Table 3. TaqMan assays designed for the taxonomic identification of *Mus spretus* and *Mus musculus domesticus* archaeological remains.

Identification	Collection Number	Origin	DNA (ng/µl)
Mus musculus domesticus	SPOT1308	France, Marcq-en-Bareuil	6
Mus spretus	21001	France, Montbazin	13.6

Table 4. MNHN collection tissue samples used for DNA extraction.

Specimen ID	morphotype	Weight lysate sample (mg)	DNA (ng/µl)
RHA05-US0029-5	*M. m. d.*	10	>0.02
RHA06-US104-2	*M. m. d.*	40	0.08
RHA06-US109-1	*M. m. d.*	10	>0.02
RHA06-US1106-1	*M. m. d.*	10	0.07
RHA07-US5143-16	*M. m. d.*	10	>0.02
RHA06-US104-3	*M. s.*	10	>0.02
RHA06-US5044-3	*M. s.*	20	0.07
RHA06-US5044-4	*M. s.*	30	0.29
RHA11-US1436-1	*M. s.*	30	0.14

Table 5. DNA extraction from archaeological remains. *M. m. d. = Mus musculus domesticus, M. s. = Mus spretus*.

	Accuracy	Kappa	Acc. Lower	Acc. Upper	Sensitivity	Specificity	Balanced Acc.
NNET	0.955	0.910	0.898	0.985	0.907	1	0.954
SVMl	0.901	0.801	0.830	0.950	0.870	0.930	0.900
SVMr	0.955	0.910	0.898	0.985	0.963	0.947	0.955
***k*NN**	0.964	0.928	0.910	0.990	0.944	0.983	0.964
LG	0.919	0.838	0.852	0.962	0.889	0.947	0.918
DTC5.0	0.910	0.819	0.841	0.956	0.870	0.947	0.909
RF	0.919	0.837	0.852	0.962	0.852	0.983	0.917
GB	0.910	0.819	0.841	0.956	0.852	0.965	0.908
NB	0.910	0.820	0.841	0.956	0.907	0.912	0.910
LDA	0.973	0.946	0.923	0.994	0.963	0.983	0.973
PLS	0.964	0.928	0.910	0.990	0.944	0.983	0.964
Ensemble: GLM	0.964	0.928	0.910	0.990	0.926	1	0.963
Stacking: NNET	0.955	0.910	0.898	0.985	0.907	1	0.954
Stacking: RF	0.946	0.892	0.886	0.980	0.889	1	0.944
Stacking: GB	0.955	0.910	0.898	0.985	0.907	1	0.954

Table 6. Results provided by ML and EL algorithms with the modern samples of *M. m. domesticus* and *M. spretus*.

NNET	
Size	11
Decay	0.00464159
SVMl	
C	1
SVMr	
Sigma	0.00677087
C	32768
kNN	
k	17
RF	
mtry	14
GB	
n.trees	700
interaction.depth	2
shrinkage	0.1
n.minobsinnode	10
NB	
fl	0
usekernel	False
adjust	1
PLS	
ncomp	4
Stacking: NNET	
Size	11
Decay	0.03162278
Stacking: RF	
mtry	4
Stacking: GB	
n.trees	350
interaction.depth	16
shrinkage	0.1
n.minobsinnode	10

Table 7. Hyperparameter configuration of the best tunned ML algorithms.

	NNET	**SVMl**	**SVMr**	**kNN**	**LG**	**DTC5.0**	**RF**	**GB**	**NB**	**LDA**
SVMl	0.31	1	-	-	-	-	-	-	-	-
SVMr	0.23	0.47	1	-	-	-	-	-	-	-
***k*NN**	0.45	0.19	0.15	1	-	-	-	-	-	-
LG	0.26	0.77*	0.50	0.11	1	-	-	-	-	-
DTC5.0	0.15	0.13	-0.03	0.35	0.10	1	-	-	-	-
RF	0.11	0.14	-0.07	0.38	0.10	0.31	1	-	-	-
GB	0.05	0.14	-0.07	0.26	0.07	0.28	0.78*	1	-	-
NB	0.05	0.01	0.21	0.03	0.16	0.16	-0.11	-0.02	1	-
LDA	0.06	0.58*	0.41	-0.01	0.61*	0.01	-0.05	-0.14	0.12	1
PLS	0.43	0.24	-0.08	0.15	0.23	0.10	0.09	-0.10	0.00	0.21

Table 8. Results obtained after calculating the pairwise correlation of the different algorithms applied to the modern sample. * = presence of correlation between algorithms.

Specimen ID	Ensemble: GLM	*M. m. d.*	*M. s.*	Stacking: NNET	*M. m. d.*	*M. s.*	Stacking: RF	*M. m. d.*	*M. s.*	Stacking: GB	*M. m. d.*	*M. s.*	ML class.
RHA05 D042 1	*M. m. d.*	1	0	*M. m. d.*	1	0	*M. m. d.*	1	0	*M. m. d.*	1	0	*M. m. d.*
RHA05 US0029 1	*M. m. d.*	1	0	*M. m. d.*	1	0	*M. m. d.*	1	0	*M. m. d.*	1	0	*M. m. d.*
RHA05 US0029 2	*M. m. d.*	1	0	*M. m. d.*	1	0	*M. m. d.*	1	0	*M. m. d.*	1	0	*M. m. d.*
RHA05 US0029 6	*M. m. d.*	1	0	*M. m. d.*	1	0	*M. m. d.*	0.98	0.02	*M. m. d.*	1	0	*M. m. d.*
RHA05 US0045 1	*M. m. d.*	1	0	*M. m. d.*	1	0	*M. m. d.*	1	0	*M. m. d.*	1	0	*M. m. d.*
RHA05 US0045 4	*M. m. d.*	1	0	*M. m. d.*	1	0	*M. m. d.*	1	0	*M. m. d.*	1	0	*M. m. d.*
RHA06 US104 2	*M. m. d.*	1	0	*M. m. d.*	1	0	*M. m. d.*	1	0	*M. m. d.*	1	0	*M. m. d.*
RHA06 US109 1	*M. m. d.*	1	0	*M. m. d.*	1	0	*M. m. d.*	0.984	0.016	*M. m. d.*	1	0	*M. m. d.*
RHA06 US109 2	*M. m. d.*	1	0	*M. s.*	0.081	0.919	*M. s.*	0.21	0.79	*M. s.*	0.01	0.99	Indeterminate
RHA06 US109 5	*M. m. d.*	1	0	*M. m. d.*	1	0	*M. m. d.*	1	0	*M. m. d.*	1	0	M. m. d.
RHA06 US1106 1	*M. s.*	0	1	*M. s.*	0	1	*M. s.*	0	1	*M. s.*	0	1	*M. s.*
RHA06 US1130 bis 1	*M. s.*	0	1	*M. s.*	0	1	*M. s.*	0	1	*M. s.*	0	1	*M. s.*
RHA06 US1130 1	*M. s.*	0.145	0.855	*M. s.*	0.001	0.999	*M. s.*	0.308	0.692	*M. s.*	0.004	0.996	*M. s.*
RHA06 US5044 1	*M. s.*	0	1	*M. s.*	0	1	*M. s.*	0	1	*M. s.*	0	1	*M. s.*
RHA06 US5044 3	*M. s.*	0	1	*M. s.*	0	1	*M. s.*	0	1	*M. s.*	0	1	*M. s.*
RHA06 US5044 4	*M. s.*	0	1	*M. s.*	0	1	*M. s.*	0	1	*M. s.*	0	1	*M. s.*
RHA07 US1209 2	*M. m. d.*	1	0	*M. m. d.*	1	0	*M. m. d.*	1	0	*M. m. d.*	1	0	M. m. d.
RHA07 US1248 1	*M. m. d.*	1	0	*M. m. d.*	1	0	*M. m. d.*	1	0	*M. m. d.*	1	0	M. m. d.
RHA07 US5136 2	*M. s.*	0.001	0.999	*M. s.*	0	1	*M. s.*	0.24	0.76	*M. s.*	0	1	*M. s.*
RHA07 US5136 5	*M. s.*	0	1	*M. s.*	0	1	*M. s.*	0	1	*M. s.*	0	1	*M. s.*
RHA07 US5143 2	*M. m. d.*	1	0	*M. m. d.*	1	0	*M. m. d.*	1	0	*M. m. d.*	1	0	*M. m. d.*
RHA07 US5143 3	*M. m. d.*	1	0	*M. m. d.*	1	0	*M. m. d.*	1	0	*M. m. d.*	1	0	*M. m. d.*
RHA07 US5143 4	*M. m. d.*	1	0	*M. m. d.*	1	0	*M. m. d.*	1	0	*M. m. d.*	1	0	*M. m. d.*
RHA07 US5143 6	*M. m. d.*	1	0	*M. m. d.*	1	0	*M. m. d.*	1	0	*M. m. d.*	1	0	*M. m. d.*
RHA07 US5143 7	*M. m. d.*	1	0	*M. m. d.*	1	0	*M. m. d.*	1	0	*M. m. d.*	1	0	*M. m. d.*
RHA07 US5143 12	*M. m. d.*	1	0	*M. m. d.*	1	0	*M. m. d.*	1	0	*M. m. d.*	1	0	*M. m. d.*
RHA07 US5143 16	*M. m. d.*	1	0	*M. m. d.*	1	0	*M. m. d.*	0.988	0.012	*M. m. d.*	1	0	*M. m. d.*
RHA11 US1436 1	*M. s.*	0	1	*M. s.*	0	1	*M. s.*	0	1	*M. s.*	0	1	*M. s.*
RHA11 US1436 2	*M. s.*	0	1	*M. s.*	0	1	*M. s.*	0	1	*M. s.*	0	1	*M. s.*

RHA12 US5412 1	*M. m. d.*	1	0	*M. m. d.*	1	0	*M. m. d.*	1	0	*M. m. d.*	1	0	*M. m. d.*
RHA14 US5502 1	*M. s.*	0	1	*M. s.*	0	1	*M. s.*	0.008	0.992	*M. s.*	0	1	*M. s.*
RHA15 US5546 1	*M. m. d.*	1	0	*M. m. d.*	1	0	*M. m. d.*	1	0	*M. m. d.*	1	0	*M. m. d.*
RHA15 US5568 1	*M. m. d.*	1	0	*M. m. d.*	1	0	*M. m. d.*	1	0	*M. m. d.*	1	0	*M. m. d.*
RHA16 US5680 2	*M. s.*	0	1	*M. s.*	0	1	*M. s.*	0	1	*M. s.*	0	1	*M. s.*
RHA16 US5680 3	*M. s.*	0	1	*M. s.*	0	1	*M. s.*	0	1	*M. s.*	0	1	*M. s.*
RHA16 US5680 5	*M. s.*	0	1	*M. s.*	0	1	*M. s.*	0	1	*M. s.*	0	1	*M. s.*
RHA16 US5680 6	*M. s.*	0	1	*M. s.*	0	1	*M. s.*	0	1	*M. s.*	0	1	*M. s.*
RHA16 US5680 7	*M. s.*	0	1	*M. s.*	0	1	*M. s.*	0	1	*M. s.*	0	1	*M. s.*

Table 9. Classification results of the different *Mus* individuals from Rirha using each of the ensemble/stacking models. *M. s.* = *Mus spretus*; *M. m. d.* = *Mus musculus domesticus.* Specimens showing discrepancies in their taxonomic attribution are highlighted in grey.

Specimen ID	**GMM+ML classification**	assay TaqMan *Mus musculus domesticus* (identification/Ct)	assay TaqMan *Mus spretus* (identification/Ct)
RHA05-US0029-5	Excluded *Morphotype *M. m. d.*	*M. m. d.* / 34	- / 0
RHA06-US104-2	*M. m. d.*	- / 0	- / 0
RHA06-US109-1	*M. m. d.*	- / 0	- / 0
RHA07-US5143-16	*M. m. d.*	*M. m. d.* / 32	- / 0
RHA06-US1106-1	*M. s.*	*M. m. d.* / 30	- / 0
RHA06-US104-3	Excluded *Morphotype *M. s.*	- / 0	- / 0
RHA06-US5044-3	*M. s.*	- / 0	*M. s.* / 36
RHA06-US5044-4	*M. s.*	- / 0	*M. s.* / 35
RHA11-US1436-1	*M. s.*	- / 0	*M. s.* / 35

Table 10. The qPCR results for the nine archaeological remains using the TaqMan assays and comparison with GMM+ML classification results. *M. s.* = *Mus spretus*; *M. m. d.* = *Mus musculus domesticus.*, - = indeterminate, * = morphotype according to traditional morphology (Darviche & Orsini, 1982; Darviche et al. 2006) of specimen excluded from GMM+ML analyses. The specimen showing a discrepancy in its taxonomic attribution is highlighted in grey.

SUPPLEMENTARY MATERIAL

Table S1. Specimen ID number and data of modern *Mus musculus domesticus* and *Mus spretus* material used in this study – EBD: Estación Biológica de Doñana, Sevilla; ISEM: Institut des Sciences de l'Evolution de Montpellier; MNCN: Museo Nacional de Ciencias Naturales, Madrid; MOHMIE: project MOdern Human installation in Morocco; MNHN: Museum National d'Histoire Naturelle.

Specimen ID	Species	Population	Collection
EBD5459M	*Mus m. domesticus*	Gran Canaria	EBD
EBD5460M	*Mus m. domesticus*	Gran Canaria	EBD
EBD5462M	*Mus m. domesticus*	Gran Canaria	EBD
EBD5466M	*Mus m. domesticus*	Gran Canaria	EBD
EBD5469M	*Mus m. domesticus*	Gran Canaria	EBD
EBD5470M	*Mus m. domesticus*	Gran Canaria	EBD
EBD5472M	*Mus m. domesticus*	Gran Canaria	EBD
EBD5506M	*Mus m. domesticus*	Gran Canaria	EBD
EBD5507M	*Mus m. domesticus*	Gran Canaria	EBD
EBD5508M	*Mus m. domesticus*	Gran Canaria	EBD
MNCN_9686	*Mus m. domesticus*	Gran Canaria	MNCN
MNCN_9692	*Mus m. domesticus*	Gran Canaria	MNCN
MNCN_9695	*Mus m. domesticus*	Gran Canaria	MNCN
MNCN_9697	*Mus m. domesticus*	Gran Canaria	MNCN
EBD5473M	*Mus m. domesticus*	Lanzarote	EBD
EBD5477M	*Mus m. domesticus*	Lanzarote	EBD
EBD5479M	*Mus m. domesticus*	Lanzarote	EBD
EBD5480M	*Mus m. domesticus*	Lanzarote	EBD
EBD5481M	*Mus m. domesticus*	Lanzarote	EBD
MNCN_9687	*Mus m. domesticus*	Lanzarote	MNCN
MNCN_9688	*Mus m. domesticus*	Lanzarote	MNCN
MNCN_9689	*Mus m. domesticus*	Lanzarote	MNCN
MNCN_9690	*Mus m. domesticus*	Lanzarote	MNCN
MNCN_9693	*Mus m. domesticus*	Lanzarote	MNCN
MNCN_9699	*Mus m. domesticus*	Lanzarote	MNCN
EBD2454M	*Mus m. domesticus*	Iberian Peninsula	EBD
EBD2459M	*Mus m. domesticus*	Iberian Peninsula	EBD
EBD2466M	*Mus m. domesticus*	Iberian Peninsula	EBD
EBD2467M	*Mus m. domesticus*	Iberian Peninsula	EBD
EBD2476M	*Mus m. domesticus*	Iberian Peninsula	EBD
EBD2614M	*Mus m. domesticus*	Iberian Peninsula	EBD
EBD2616M	*Mus m. domesticus*	Iberian Peninsula	EBD
EBD2617M	*Mus m. domesticus*	Iberian Peninsula	EBD
EBD3540M	*Mus m. domesticus*	Iberian Peninsula	EBD
EBD3942M	*Mus m. domesticus*	Iberian Peninsula	EBD
EBD4049M	*Mus m. domesticus*	Iberian Peninsula	EBD
EBD4050M	*Mus m. domesticus*	Iberian Peninsula	EBD
EBD4052M	*Mus m. domesticus*	Iberian Peninsula	EBD
EBD4053M	*Mus m. domesticus*	Iberian Peninsula	EBD

EBD6343M	*Mus m. domesticus*	Iberian Peninsula	EBD
EBD6344M	*Mus m. domesticus*	Iberian Peninsula	EBD
EBD6361M	*Mus m. domesticus*	Iberian Peninsula	EBD
EBD6623M	*Mus m. domesticus*	Iberian Peninsula	EBD
EBD6646M	*Mus m. domesticus*	Iberian Peninsula	EBD
EBD14830M	*Mus m. domesticus*	Iberian Peninsula	EBD
EBD14837M	*Mus m. domesticus*	Iberian Peninsula	EBD
EBD14839M	*Mus m. domesticus*	Iberian Peninsula	EBD
EBD17170M	*Mus m. domesticus*	Iberian Peninsula	EBD
EBD17171M	*Mus m. domesticus*	Iberian Peninsula	EBD
EBD17285M	*Mus m. domesticus*	Iberian Peninsula	EBD
EBD19863M	*Mus m. domesticus*	Iberian Peninsula	EBD
EBD19864M	*Mus m. domesticus*	Iberian Peninsula	EBD
EBD22537M	*Mus m. domesticus*	Iberian Peninsula	EBD
EBD22538M	*Mus m. domesticus*	Iberian Peninsula	EBD
EBD22541M	*Mus m. domesticus*	Iberian Peninsula	EBD
EBD22557M	*Mus m. domesticus*	Iberian Peninsula	EBD
EBD22558M	*Mus m. domesticus*	Iberian Peninsula	EBD
EBD22564M	*Mus m. domesticus*	Iberian Peninsula	EBD
EBD22565M	*Mus m. domesticus*	Iberian Peninsula	EBD
EBD22567M	*Mus m. domesticus*	Iberian Peninsula	EBD
EBD22569M	*Mus m. domesticus*	Iberian Peninsula	EBD
EBD22570M	*Mus m. domesticus*	Iberian Peninsula	EBD
EBD22571M	*Mus m. domesticus*	Iberian Peninsula	EBD
EBD22572M	*Mus m. domesticus*	Iberian Peninsula	EBD
EBD22580M	*Mus m. domesticus*	Iberian Peninsula	EBD
MNCN_3467	*Mus m. domesticus*	Iberian Peninsula	MNCN
MNCN_3468	*Mus m. domesticus*	Iberian Peninsula	MNCN
MNCN_3472	*Mus m. domesticus*	Iberian Peninsula	MNCN
MNCN_3475	*Mus m. domesticus*	Iberian Peninsula	MNCN
MNCN_3476	*Mus m. domesticus*	Iberian Peninsula	MNCN
MNCN_3477	*Mus m. domesticus*	Iberian Peninsula	MNCN
MNCN_3478	*Mus m. domesticus*	Iberian Peninsula	MNCN
MNCN_3480	*Mus m. domesticus*	Iberian Peninsula	MNCN
MNCN_3481	*Mus m. domesticus*	Iberian Peninsula	MNCN
MNCN_3482	*Mus m. domesticus*	Iberian Peninsula	MNCN
MNCN_3483	*Mus m. domesticus*	Iberian Peninsula	MNCN
MNCN_3484	*Mus m. domesticus*	Iberian Peninsula	MNCN
MNCN_3486	*Mus m. domesticus*	Iberian Peninsula	MNCN
MNCN_3487	*Mus m. domesticus*	Iberian Peninsula	MNCN
MNCN_3488	*Mus m. domesticus*	Iberian Peninsula	MNCN
MNCN_3490	*Mus m. domesticus*	Iberian Peninsula	MNCN
MNCN_3492	*Mus m. domesticus*	Iberian Peninsula	MNCN
MNCN_3493	*Mus m. domesticus*	Iberian Peninsula	MNCN
MNCN_3499	*Mus m. domesticus*	Iberian Peninsula	MNCN
MNCN_18919	*Mus m. domesticus*	Iberian Peninsula	MNCN

MNCN_18920	*Mus m. domesticus*	Iberian Peninsula	MNCN
ISEM_145	*Mus m. domesticus*	France	ISEM
ISEM_376	*Mus m. domesticus*	France	ISEM
ISEM_377	*Mus m. domesticus*	France	ISEM
ISEM_408	*Mus m. domesticus*	France	ISEM
ISEM_409	*Mus m. domesticus*	France	ISEM
ISEM_416	*Mus m. domesticus*	France	ISEM
ISEM_838	*Mus m. domesticus*	France	ISEM
ISEM_839	*Mus m. domesticus*	France	ISEM
ISEM_840	*Mus m. domesticus*	France	ISEM
ISEM_842	*Mus m. domesticus*	France	ISEM
ISEM_843	*Mus m. domesticus*	France	ISEM
ISEM_844	*Mus m. domesticus*	France	ISEM
ISEM_845	*Mus m. domesticus*	France	ISEM
ISEM_846	*Mus m. domesticus*	France	ISEM
ISEM_847	*Mus m. domesticus*	France	ISEM
ISEM_848	*Mus m. domesticus*	France	ISEM
ISEM_849	*Mus m. domesticus*	France	ISEM
ISEM_850	*Mus m. domesticus*	France	ISEM
MNHN-ZM-MO-1942-328	*Mus m. domesticus*	France	MNHN
MNHN-ZM-MO-1942-336	*Mus m. domesticus*	France	MNHN
MNHN-ZM-MO-1942-338	*Mus m. domesticus*	France	MNHN
MNHN-ZM-MO-1942-339	*Mus m. domesticus*	France	MNHN
MNHN-ZM-MO-1942-378	*Mus m. domesticus*	France	MNHN
MNHN-ZM-MO-1942-389	*Mus m. domesticus*	France	MNHN
MNHN-ZM-MO-1949-144	*Mus m. domesticus*	France	MNHN
MNHN-ZM-MO-1949-145	*Mus m. domesticus*	France	MNHN
MNHN-ZM-MO-1949-146	*Mus m. domesticus*	France	MNHN
MNHN-ZM-MO-1956-1164	*Mus m. domesticus*	France	MNHN
ISEM_9936	*Mus m. domesticus*	Morocco	ISEM
ISEM_9939	*Mus m. domesticus*	Morocco	ISEM
ISEM_9942	*Mus m. domesticus*	Morocco	ISEM
ISEM_9960	*Mus m. domesticus*	Morocco	ISEM
ISEM_9978	*Mus m. domesticus*	Morocco	ISEM
ISEM_9984	*Mus m. domesticus*	Morocco	ISEM
ISEM_9986	*Mus m. domesticus*	Morocco	ISEM
ISEM_9987	*Mus m. domesticus*	Morocco	ISEM
ISEM_9991	*Mus m. domesticus*	Morocco	ISEM
ISEM_9994	*Mus m. domesticus*	Morocco	ISEM
ISEM_9995	*Mus m. domesticus*	Morocco	ISEM
ISEM_9997	*Mus m. domesticus*	Morocco	ISEM
ISEM_10000	*Mus m. domesticus*	Morocco	ISEM
ISEM_10002	*Mus m. domesticus*	Morocco	ISEM
ISEM_10005	*Mus m. domesticus*	Morocco	ISEM
ISEM_10006	*Mus m. domesticus*	Morocco	ISEM
ISEM_10007	*Mus m. domesticus*	Morocco	ISEM

ISEM_10009	*Mus m. domesticus*	Morocco	ISEM
ISEM_10010	*Mus m. domesticus*	Morocco	ISEM
ISEM_10012	*Mus m. domesticus*	Morocco	ISEM
ISEM_10013	*Mus m. domesticus*	Morocco	ISEM
ISEM_10015	*Mus m. domesticus*	Morocco	ISEM
ISEM_10016	*Mus m. domesticus*	Morocco	ISEM
ISEM_10017	*Mus m. domesticus*	Morocco	ISEM
ISEM_10018	*Mus m. domesticus*	Morocco	ISEM
ISEM_10019	*Mus m. domesticus*	Morocco	ISEM
ISEM_10021	*Mus m. domesticus*	Morocco	ISEM
ISEM_10022	*Mus m. domesticus*	Morocco	ISEM
ISEM_10023	*Mus m. domesticus*	Morocco	ISEM
ISEM_10024	*Mus m. domesticus*	Morocco	ISEM
ISEM_10026	*Mus m. domesticus*	Morocco	ISEM
ISEM_10027	*Mus m. domesticus*	Morocco	ISEM
ISEM_10045	*Mus m. domesticus*	Morocco	ISEM
ISEM_10058	*Mus m. domesticus*	Morocco	ISEM
ISEM_10060	*Mus m. domesticus*	Morocco	ISEM
ISEM_17219	*Mus m. domesticus*	Morocco	ISEM
ALG_1	*Mus m. domesticus*	Algeria	ISEM
ALG_2	*Mus m. domesticus*	Algeria	ISEM
ALG_3	*Mus m. domesticus*	Algeria	ISEM
ALG_4	*Mus m. domesticus*	Algeria	ISEM
ALG_5	*Mus m. domesticus*	Algeria	ISEM
ALG_6	*Mus m. domesticus*	Algeria	ISEM
ALG_7	*Mus m. domesticus*	Algeria	ISEM
ALG_8	*Mus m. domesticus*	Algeria	ISEM
ALG_12	*Mus m. domesticus*	Algeria	ISEM
ALG_13	*Mus m. domesticus*	Algeria	ISEM
ALG_14	*Mus m. domesticus*	Algeria	ISEM
ALG_15	*Mus m. domesticus*	Algeria	ISEM
ALG_16	*Mus m. domesticus*	Algeria	ISEM
ALG_17	*Mus m. domesticus*	Algeria	ISEM
ALG_18	*Mus m. domesticus*	Algeria	ISEM
ALG_20	*Mus m. domesticus*	Algeria	ISEM
ALG_21	*Mus m. domesticus*	Algeria	ISEM
MNHN-ZM-MO-1994-2401	*Mus m. domesticus*	Swizterland	MNHN
MNHN-ZM-MO-1994-2402	*Mus m. domesticus*	Swizterland	MNHN
MNHN-ZM-MO-1994-2403	*Mus m. domesticus*	Swizterland	MNHN
MNHN-ZM-MO-1994-2404	*Mus m. domesticus*	Swizterland	MNHN
MNHN-ZM-MO-1994-2405	*Mus m. domesticus*	Swizterland	MNHN
MNHN-ZM-MO-1994-2406	*Mus m. domesticus*	Swizterland	MNHN
MNHN-ZM-MO-1994-2407	*Mus m. domesticus*	Swizterland	MNHN
MNHN-ZM-MO-1994-2408	*Mus m. domesticus*	Swizterland	MNHN
MNHN-ZM-MO-1994-2409	*Mus m. domesticus*	Swizterland	MNHN
MNHN-ZM-MO-1994-2430	*Mus m. domesticus*	Swizterland	MNHN

MNHN-ZM-MO-1994-2431	*Mus m. domesticus*	Swizterland	MNHN
MNHN-ZM-MO-1994-2432	*Mus m. domesticus*	Swizterland	MNHN
MNHN-ZM-MO-1994-2433	*Mus m. domesticus*	Swizterland	MNHN
MNHN-ZM-MO-1994-2434	*Mus m. domesticus*	Swizterland	MNHN
MNHN-ZM-MO-1994-2435	*Mus m. domesticus*	Swizterland	MNHN
MNHN-ZM-MO-1994-2437	*Mus m. domesticus*	Swizterland	MNHN
MNCN_9978	*Mus spretus*	Mallorca	MNCN
MNCN_9979	*Mus spretus*	Mallorca	MNCN
MNCN_9980	*Mus spretus*	Mallorca	MNCN
MNCN_9982	*Mus spretus*	Mallorca	MNCN
MNCN_9983	*Mus spretus*	Mallorca	MNCN
MNCN_9984	*Mus spretus*	Mallorca	MNCN
MNCN_9985	*Mus spretus*	Mallorca	MNCN
MNCN_9986	*Mus spretus*	Mallorca	MNCN
MNCN_9987	*Mus spretus*	Mallorca	MNCN
MNCN_9988	*Mus spretus*	Mallorca	MNCN
2019.013.002.01	*Mus spretus*	Iberian Peninsula	EBD
2019.013.002.02	*Mus spretus*	Iberian Peninsula	EBD
2019.013.002.03	*Mus spretus*	Iberian Peninsula	EBD
2019.013.002.04	*Mus spretus*	Iberian Peninsula	EBD
2019.013.002.05	*Mus spretus*	Iberian Peninsula	EBD
2019.013.002.06	*Mus spretus*	Iberian Peninsula	EBD
2019.013.002.07	*Mus spretus*	Iberian Peninsula	EBD
2019.013.002.08	*Mus spretus*	Iberian Peninsula	EBD
2019.013.002.09	*Mus spretus*	Iberian Peninsula	EBD
2019.013.002.10	*Mus spretus*	Iberian Peninsula	EBD
2019.013.007.01	*Mus spretus*	Iberian Peninsula	EBD
2019.013.007.02-1	*Mus spretus*	Iberian Peninsula	EBD
2019.013.007.02-2	*Mus spretus*	Iberian Peninsula	EBD
2019.013.007.02-3	*Mus spretus*	Iberian Peninsula	EBD
2019.013.007.02-4	*Mus spretus*	Iberian Peninsula	EBD
2019.013.007.03	*Mus spretus*	Iberian Peninsula	EBD
2019.013.007.04	*Mus spretus*	Iberian Peninsula	EBD
2019.013.028.01	*Mus spretus*	Iberian Peninsula	EBD
2019.013.028.02	*Mus spretus*	Iberian Peninsula	EBD
2019.013.028.03	*Mus spretus*	Iberian Peninsula	EBD
2019.013.026.01	*Mus spretus*	Iberian Peninsula	EBD
2019.013.026.02	*Mus spretus*	Iberian Peninsula	EBD
EBD1985M	*Mus spretus*	Iberian Peninsula	EBD
EBD1986M	*Mus spretus*	Iberian Peninsula	EBD
EBD1994M	*Mus spretus*	Iberian Peninsula	EBD
EBD3923M	*Mus spretus*	Iberian Peninsula	EBD
EBD4034M	*Mus spretus*	Iberian Peninsula	EBD
EBD4324M	*Mus spretus*	Iberian Peninsula	EBD
EBD4326M	*Mus spretus*	Iberian Peninsula	EBD
EBD5819M	*Mus spretus*	Iberian Peninsula	EBD

EBD5829M	*Mus spretus*	Iberian Peninsula	EBD
EBD5830M	*Mus spretus*	Iberian Peninsula	EBD
EBD5870M	*Mus spretus*	Iberian Peninsula	EBD
EBD5884M	*Mus spretus*	Iberian Peninsula	EBD
EBD5885M	*Mus spretus*	Iberian Peninsula	EBD
EBD5887M	*Mus spretus*	Iberian Peninsula	EBD
EBD6696M	*Mus spretus*	Iberian Peninsula	EBD
EBD19457M	*Mus spretus*	Iberian Peninsula	EBD
EBD19458M	*Mus spretus*	Iberian Peninsula	EBD
EBD19459M	*Mus spretus*	Iberian Peninsula	EBD
EBD19460M	*Mus spretus*	Iberian Peninsula	EBD
EBD19477M	*Mus spretus*	Iberian Peninsula	EBD
EBD19481M	*Mus spretus*	Iberian Peninsula	EBD
EBD19482M	*Mus spretus*	Iberian Peninsula	EBD
EBD21702M	*Mus spretus*	Iberian Peninsula	EBD
EBD25999M	*Mus spretus*	Iberian Peninsula	EBD
EBD26092M	*Mus spretus*	Iberian Peninsula	EBD
EBD26093M	*Mus spretus*	Iberian Peninsula	EBD
EBD30911M	*Mus spretus*	Iberian Peninsula	EBD
EBD30922M	*Mus spretus*	Iberian Peninsula	EBD
EBD30925M	*Mus spretus*	Iberian Peninsula	EBD
MNCN_3517	*Mus spretus*	Iberian Peninsula	MNCN
MNCN_3518	*Mus spretus*	Iberian Peninsula	MNCN
MNCN_3519	*Mus spretus*	Iberian Peninsula	MNCN
MNCN_3520	*Mus spretus*	Iberian Peninsula	MNCN
MNCN_3521	*Mus spretus*	Iberian Peninsula	MNCN
MNCN_9994	*Mus spretus*	Iberian Peninsula	MNCN
ISEM_717	*Mus spretus*	France	ISEM
ISEM_725	*Mus spretus*	France	ISEM
ISEM_726	*Mus spretus*	France	ISEM
ISEM_728	*Mus spretus*	France	ISEM
ISEM_729	*Mus spretus*	France	ISEM
ISEM_731	*Mus spretus*	France	ISEM
ISEM_732	*Mus spretus*	France	ISEM
ISEM_734	*Mus spretus*	France	ISEM
ISEM_735	*Mus spretus*	France	ISEM
ISEM_736	*Mus spretus*	France	ISEM
ISEM_737	*Mus spretus*	France	ISEM
ISEM_739	*Mus spretus*	France	ISEM
ISEM_740	*Mus spretus*	France	ISEM
ISEM_741	*Mus spretus*	France	ISEM
MNHN-ZM-MO-1933-1836	*Mus spretus*	France	MNHN
MNHN-ZM-MO-1933-1851	*Mus spretus*	France	MNHN
MNHN-ZM-MO-1933-1854	*Mus spretus*	France	MNHN
MNHN-ZM-MO-1933-1856	*Mus spretus*	France	MNHN
MNHN-ZM-MO-1980-417	*Mus spretus*	France	MNHN

MNHN-ZM-MO-1993-1847	*Mus spretus*	France	MNHN
MNHN-ZM-MO-1993-2724	*Mus spretus*	France	MNHN
MNHN-ZM-MO-1993-2752	*Mus spretus*	France	MNHN
MNHN-ZM-MO-1993-2772	*Mus spretus*	France	MNHN
ISEM_742	*Mus spretus*	Morocco	ISEM
ISEM_743	*Mus spretus*	Morocco	ISEM
ISEM_744	*Mus spretus*	Morocco	ISEM
ISEM_745	*Mus spretus*	Morocco	ISEM
ISEM_746	*Mus spretus*	Morocco	ISEM
ISEM_747	*Mus spretus*	Morocco	ISEM
ISEM_750	*Mus spretus*	Morocco	ISEM
ISEM_751	*Mus spretus*	Morocco	ISEM
ISEM_752	*Mus spretus*	Morocco	ISEM
ISEM_753	*Mus spretus*	Morocco	ISEM
ISEM_754	*Mus spretus*	Morocco	ISEM
ISEM_755	*Mus spretus*	Morocco	ISEM
ISEM_756	*Mus spretus*	Morocco	ISEM
ISEM_757	*Mus spretus*	Morocco	ISEM
ISEM_759	*Mus spretus*	Morocco	ISEM
ISEM_760	*Mus spretus*	Morocco	ISEM
ISEM_761	*Mus spretus*	Morocco	ISEM
ISEM_762	*Mus spretus*	Morocco	ISEM
ISEM_763	*Mus spretus*	Morocco	ISEM
ISEM_764	*Mus spretus*	Morocco	ISEM
ISEM_765	*Mus spretus*	Morocco	ISEM
ISEM_766	*Mus spretus*	Morocco	ISEM
ISEM_767	*Mus spretus*	Morocco	ISEM
ISEM_768	*Mus spretus*	Morocco	ISEM
ISEM_769	*Mus spretus*	Morocco	ISEM
ISEM_772	*Mus spretus*	Morocco	ISEM
ISEM_777	*Mus spretus*	Morocco	ISEM
ISEM_7560	*Mus spretus*	Morocco	ISEM
ISEM_7561	*Mus spretus*	Morocco	ISEM
ISEM_9852	*Mus spretus*	Morocco	ISEM
ISEM_9853	*Mus spretus*	Morocco	ISEM
ISEM_9856	*Mus spretus*	Morocco	ISEM
ISEM_9857	*Mus spretus*	Morocco	ISEM
ISEM_9862	*Mus spretus*	Morocco	ISEM
ISEM_9863	*Mus spretus*	Morocco	ISEM
ISEM_9864	*Mus spretus*	Morocco	ISEM
ISEM_9865	*Mus spretus*	Morocco	ISEM
ISEM_9868	*Mus spretus*	Morocco	ISEM
ISEM_9869	*Mus spretus*	Morocco	ISEM
ISEM_9870	*Mus spretus*	Morocco	ISEM
ISEM_9891	*Mus spretus*	Morocco	ISEM
ISEM_9899	*Mus spretus*	Morocco	ISEM

ISEM_9907	*Mus spretus*	Morocco	ISEM
ISEM_9908	*Mus spretus*	Morocco	ISEM
ISEM_9912	*Mus spretus*	Morocco	ISEM
ISEM_9913	*Mus spretus*	Morocco	ISEM
ISEM_9914	*Mus spretus*	Morocco	ISEM
ISEM_9915	*Mus spretus*	Morocco	ISEM
ISEM_9916	*Mus spretus*	Morocco	ISEM
ISEM_9921	*Mus spretus*	Morocco	ISEM
MOHMIE_SB1	*Mus spretus*	Morocco	MOHMIE
MOHMIE_SB18	*Mus spretus*	Morocco	MOHMIE
MOHMIE_SB2	*Mus spretus*	Morocco	MOHMIE
MOHMIE_SB3	*Mus spretus*	Morocco	MOHMIE
MOHMIE_SB35	*Mus spretus*	Morocco	MOHMIE
MOHMIE_SB4	*Mus spretus*	Morocco	MOHMIE
MOHMIE_SB40	*Mus spretus*	Morocco	MOHMIE
MOHMIE_SB41	*Mus spretus*	Morocco	MOHMIE
MOHMIE_SB43	*Mus spretus*	Morocco	MOHMIE
MOHMIE_SB44	*Mus spretus*	Morocco	MOHMIE
MOHMIE_SB48	*Mus spretus*	Morocco	MOHMIE
MOHMIE_SB5	*Mus spretus*	Morocco	MOHMIE
MOHMIE_SB50	*Mus spretus*	Morocco	MOHMIE
MOHMIE_SB6	*Mus spretus*	Morocco	MOHMIE
MOHMIE_SB7	*Mus spretus*	Morocco	MOHMIE
MNHN-ZM-MO-1955-79	*Mus spretus*	Algeria	MNHN
MNHN-ZM-MO-1980-353	*Mus spretus*	Algeria	MNHN
MNHN-ZM-MO-1980-354	*Mus spretus*	Algeria	MNHN
MNHN-ZM-MO-1980-355	*Mus spretus*	Algeria	MNHN
MNHN-ZM-MO-1980-356	*Mus spretus*	Algeria	MNHN
MNHN-ZM-MO-1980-357	*Mus spretus*	Algeria	MNHN
MNHN-ZM-MO-1980-358	*Mus spretus*	Algeria	MNHN
MNHN-ZM-MO-1980-359	*Mus spretus*	Algeria	MNHN
MNHN-ZM-MO-1980-209	*Mus spretus*	Tunisia	MNHN
MNHN-ZM-MO-1980-210	*Mus spretus*	Tunisia	MNHN
MNHN-ZM-MO-1980-212	*Mus spretus*	Tunisia	MNHN
MNHN-ZM-MO-1980-213	*Mus spretus*	Tunisia	MNHN
MNHN-ZM-MO-1980-214	*Mus spretus*	Tunisia	MNHN
MNHN-ZM-MO-1980-215	*Mus spretus*	Tunisia	MNHN
MNHN-ZM-MO-1981-635	*Mus spretus*	Tunisia	MNHN
MNHN-ZM-MO-1981-636	*Mus spretus*	Tunisia	MNHN
MNHN-ZM-MO-1981-637	*Mus spretus*	Tunisia	MNHN
MNHN-ZM-MO-1981-648	*Mus spretus*	Tunisia	MNHN
MNHN-ZM-MO-1991-1167	*Mus spretus*	Tunisia	MNHN
MNHN-ZM-MO-1991-1168	*Mus spretus*	Tunisia	MNHN
MNHN-ZM-MO-1991-1169	*Mus spretus*	Tunisia	MNHN
MNHN-ZM-MO-1991-1171	*Mus spretus*	Tunisia	MNHN
MNHN-ZM-MO-1991-1172	*Mus spretus*	Tunisia	MNHN

MNHN-ZM-MO-1991-1173	*Mus spretus*	Tunisia	MNHN
MNHN-ZM-MO-1991-1174	*Mus spretus*	Tunisia	MNHN
MNHN-ZM-MO-1991-1175	*Mus spretus*	Tunisia	MNHN
MNHN-ZM-MO-1991-1176	*Mus spretus*	Tunisia	MNHN
MNHN-ZM-MO-1991-1177	*Mus spretus*	Tunisia	MNHN
MNHN-ZM-MO-1991-1178	*Mus spretus*	Tunisia	MNHN
MNHN-ZM-MO-1991-1179	*Mus spretus*	Tunisia	MNHN
MNHN-ZM-MO-1991-1180	*Mus spretus*	Tunisia	MNHN
MNHN-ZM-MO-1991-1181	*Mus spretus*	Tunisia	MNHN
MNHN-ZM-MO-1991-1183	*Mus spretus*	Tunisia	MNHN
MNHN-ZM-MO-1991-1184	*Mus spretus*	Tunisia	MNHN
MNHN-ZM-MO-1991-1186	*Mus spretus*	Tunisia	MNHN
MNHN-ZM-MO-1991-1187	*Mus spretus*	Tunisia	MNHN
MNHN-ZM-MO-1991-1188	*Mus spretus*	Tunisia	MNHN

Table S2. Ninety-five cytB reference sequences of *Mus* genus used to design the TaqMan probes of *Mus musculus domesticus* and *Mus spretus*.

Taxonomy	GenBank Ascension number
Mus booduga	AB125760
Mus booduga	AB125761
Mus bufo	KJ935781
Mus bufo	KJ935782
Mus caroli	AB033698
Mus caroli	AB109792
Mus caroli	AB109794
Mus caroli	AB109795
Mus cervicolor	AB125764
Mus cervicolor	AB125765
Mus cf. callewaerti	KJ935752
Mus cf. callewaerti	KJ935753
Mus cf. gerbillus	KJ935825
Mus cf. gerbillus	KJ935826
Mus cf. gratus	KJ935815
Mus cf. gratus	KJ935816
Mus cf. kasaicus	KJ935811
Mus cf. kasaicus	KJ935812
Mus cf. proconodon	KJ935768
Mus cf. proconodon	KJ935769
Mus cf. proconodon	KJ935770
Mus cf. saxicola	AJ698879
Mus cf. setulosus	KJ935778
Mus cf. setulosus	KJ935779
Mus cf. setulosus	KJ935780
Mus cf. tenellus	KJ935833
Mus cf. tenellus	KJ935834
Mus cookii	AB125767
Mus cookii	AB125768
Mus crociduroides	AJ698878
Mus cypriacus	FR751074
Mus famulus	AJ698872
Mus fragilicauda	AB125779

Mus haussa	AJ698877
Mus haussa	AJ875071
Mus indutus	AJ698874
Mus indutus	AJ875070
Mus lepidoides	AB262414
Mus lepidoides	AB262415
Mus macedonicus	AB125770
Mus macedonicus	AY057808
Mus mahomet	KJ935786
Mus mahomet	KJ935787
Mus mattheyi	AB125781
Mus mattheyi	AJ698876
Mus minutoides	AJ875076
Mus minutoides	AJ875077
Mus musculoides	AJ698875
Mus musculoides	AJ875075
Mus musculoides	KJ935835
Mus musculus	AB819907
Mus musculus	AB819908
Mus musculus bactrianus	KY418170
Mus musculus bactrianus	KY418171
Mus musculus castaneus	AB125771
Mus musculus castaneus	AB125772
Mus musculus castaneus	AB125773
Mus musculus domesticus	AB125774
Mus musculus domesticus	AB649455
Mus musculus domesticus	AB649456
Mus musculus homourus	AB649506
Mus musculus molossinus	AB205281
Mus musculus molossinus	AB205282
Mus musculus molossinus	AB205283
Mus musculus musculus	AB033699
Mus musculus musculus	AB205273
Mus nitidulus	AB262422
Mus nitidulus	AB262423
Mus pahari	AB096839

Mus pahari	AY057814
Mus pahari	EU349767
Mus platythrix	AB125782
Mus platythrix	AJ698880
Mus saxicola	AY057815
Mus setulosus	KJ935775
Mus setulosus	KJ935776
Mus setulosus	KJ935777
Mus sp. 2000082	AJ875085
Mus sp. Dakawa	KJ935803
Mus sp. Harenna	KJ935742
Mus sp. Harenna	KJ935743
Mus sp. Kiwit	KJ935751
Mus spicilegus	AB125775
Mus spicilegus	AY057809
Mus spicilegus	KU375154
Mus spicilegus	KY754048
Mus spretus	AB033700
Mus spretus	AY057810
Mus spretus	AY224678
Mus terricolor	AB125776
Mus terricolor	AB125777
Mus terricolor	AB125778
Mus triton	KJ935746
Mus triton	KJ935747
Mus triton	KJ935750