



Water voles of Lake Hula: assessing their past, present, and future

Nimrod Marom¹ · Adva Olga Peretz² · Ignacio A. Lazagabaster^{3,4} · Meirav Meiri⁵ · Shai Meiri^{2,5}

Received: 5 October 2023 / Revised: 18 February 2024 / Accepted: 26 February 2024
© The Author(s) 2024

Abstract

The southernmost population of Eurasian water vole (*Arvicola amphibius*) inhabited Lake Hula in the upper Jordan Valley until the lake was drained in the 1950s. Considering the continuous conservation and restoration initiatives in the Hula Valley, we set out to verify the extinction of the Hula water vole population using trap surveys, field sign surveys, and owl pellets' content. Having confirmed its recent extirpation, we used museum and archaeological specimens to study the morphological and genetic similarity of the extirpated Hula water voles to both modern conspecifics in Eurasia and to local Pleistocene specimens. Our results suggest that the Hula population represented an admixture of extinct local Pleistocene and extant, probably European, ancestors. The recent anthropogenic extirpation of this unique population could justify its reintroduction. Species distribution modelling, however, suggests future deterioration of habitat suitability over the coming decades. This calls for careful consideration of how sustainable a reintroduction would be.

Keywords Water vole · *Arvicola* · Paleobiology · Geometric morphometrics · Ancient DNA · Species distribution modelling

Introduction

Lake Hula, in northeastern Israel, was drained in the 1950s, severely damaging local plant and animal communities (Hambright and Zohary 1999) (Fig. 1) and driving the probable extirpation of the southernmost water vole (*Arvicola amphibius* s.l.) population (Mendelsshon and Yom-Tov 1999). The recent, human-caused, extirpation suggests the possibility of restoring water voles to the Hula Valley as part of the ongoing re-flooding and conservation efforts in the region (Hambright and Zohary 1998; Tsipris and Meron 1998; Cohen-Shacham et al. 2011).

Water voles (*Arvicola* spp.) are widely distributed, from Britain to the Lena River, and from Scandinavia to the Middle East (Lawton and Woodroffe 1991; Batsaikhan et al. 2021). *Arvicola* includes three species: *A. sapidus* in Iberia and France, *A. amphibius* in much of northwestern Eurasian (Brace et al. 2016), and *A. persicus* in Iran (Mahmoudi et al. 2020). Genetic studies suggest allopatric population diversification in Glacial Maximum refugia and consequent spread and recolonisation during interglacials (Centeno-cuadros and Delibes 2009; Brace et al. 2016). Water voles have both fossorial and semi-aquatic forms, which differ in body size and molar shape (Piras et al. 2012), but do not match phylogenetic structure (Brace et al. 2016).

The earliest known *Arvicola* specimens in the Southern Levant date to the early Pleistocene of Ubeidiya in the Jordan Valley (Tchernov 1975). The earliest Hula Valley fossils are from ~65 kya in Nahal Mahanayim Outlet. These have been ascribed to a new species, *Arvicola nahalensis*, based on the first lower molar (m1) antecoid morphology (Maul et al. 2021). Maul et al. (2021) suggested that the recent Hula water voles, known from the owl pellets collected in the 1940s by Dor (Mendelsshon and Yom-Tov 1999), were similar in their enamel thickness to final Pleistocene and early Holocene specimens from Eynan (ca. 14 kya, Hula Valley) and Netiv Hagdud (ca. 11 kya, near Jericho), and that they were probably

✉ Nimrod Marom
nmarom2@univ.haifa.ac.il

¹ School of Archaeology and Maritime Cultures, University of Haifa, Haifa, Israel

² School of Zoology, Tel Aviv University, Tel Aviv, Israel

³ National Research Center On Human Evolution, Burgos, Spain

⁴ Department of Ecology, Evolution & Behaviour, University of Liverpool, Liverpool, UK

⁵ Steinhardt Museum of Natural History, Tel Aviv University, Tel Aviv, Israel

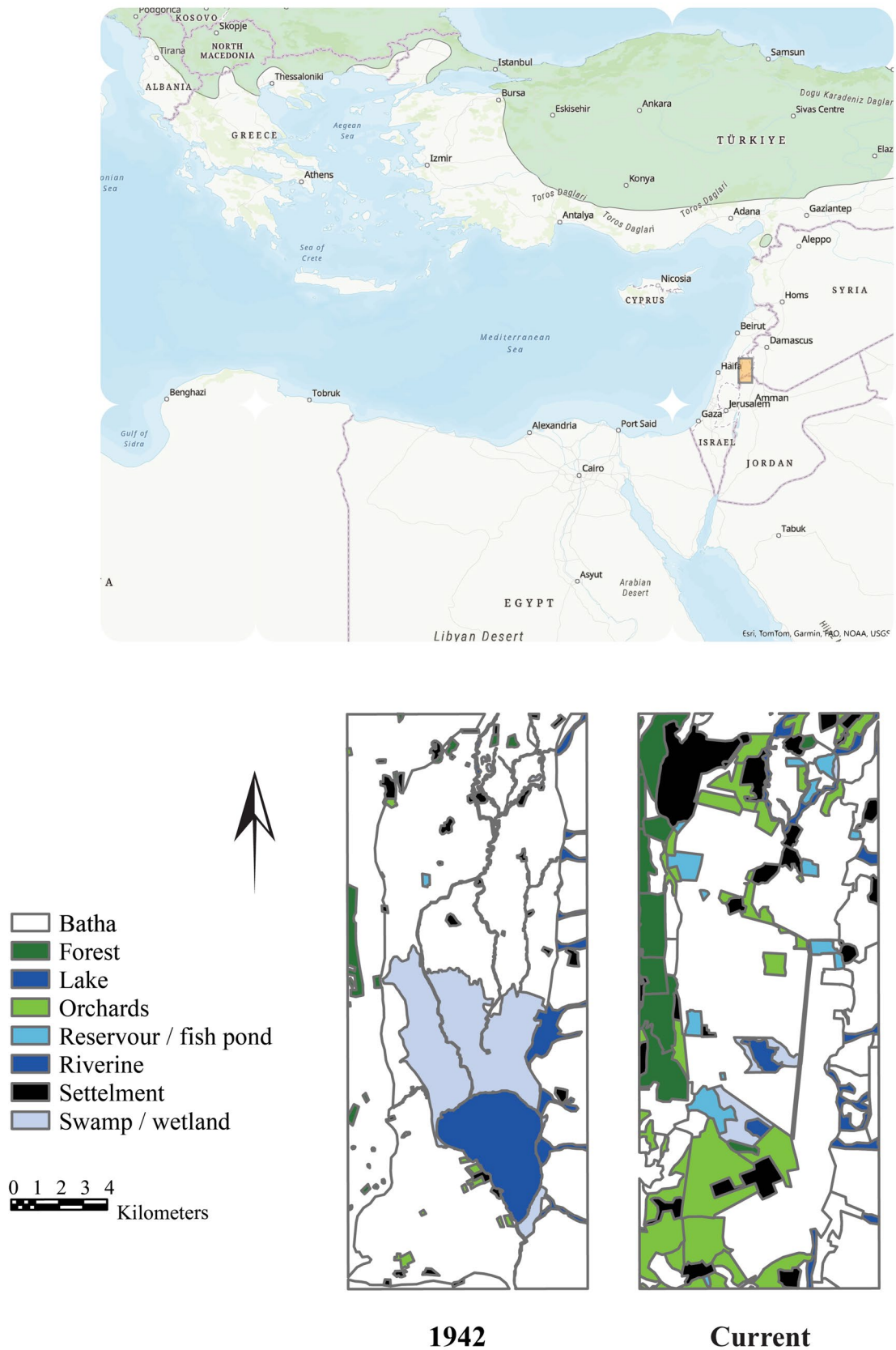


Fig. 1 The study region (yellow rectangle) against the current southern part of *A. amphibius* range (top). The Hula Lake in 1942 and today (bottom; Map courtesy of Uri Roll, habitat definitions follow Renan et al. 2017)

part of, or closely related to, the extant Iranian lineage, *A. persicus*. This contradicts the traditional assignment of the recent Hula water voles to *A. amphibius* (Mendelsshon and Yom-Tov 1999), although the Iranian population has only been elevated to the species level in 2020 (Mahmoudi et al. 2020). Maul et al. (2021) suggested that, during the late Pleistocene, the endemic *A. nahalensis* was replaced by *A. persicus*, which persisted in the Jordan Valley until recently. By the twentieth century CE, southern Levant *Arvicola* were restricted to the northernmost part of the Jordan Valley in the Hula region. The 1950s drainage of the lake and the adjacent swamp are thought to have driven this last population to extinction (Mendelsshon and Yom-Tov 1999; Meiri et al. 2019).

The water voles of Lake Hula are known from 21 crania found in owl pellets recovered during the 1940s from Yesud Hamaala on the western bank of the Hula, where Menahem Dor had also spotted water vole droppings and field marks (Dor 1947). Recent pellets of several owl species from the Galilee did not yield *Arvicola* remains (Comay et al. 2022). Some doubts, however, remain due to a report on a water vole south of the Sea of Galilee (32.649N 35.562E) in 1995, and another potential sighting at Tel Dan (33.249N, 35.651E) (Horvitz 2015).

We therefore aimed

1. To test whether water voles really went extinct in Israel. This was accomplished by trap and sign surveys along water courses in the Hula and Dan Nature reserves, and studying rodent remains from owl pellets.
2. To assess which extant water vole population is most similar to the extinct Hula voles using genetic and morphometric tools.
3. To evaluate current and future habitat suitability of the Hula region to water voles using species distribution modelling.

Assessing the present condition of the Hula water vole population, comparing it to existing populations, and

evaluating the regional suitability, can inform decisions regarding the potential reintroduction of this rodent to Israel.

Materials and methods

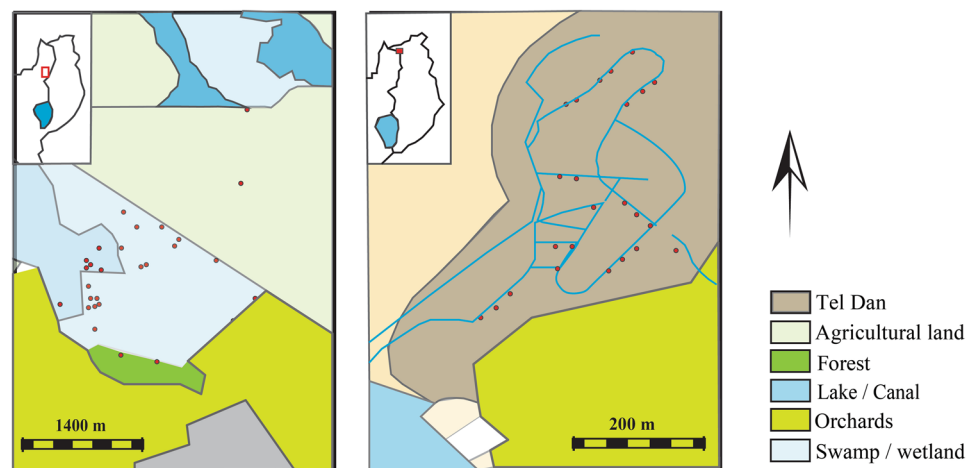
Sign and trap surveys

During March 2018, we carried out a field sign survey in the Dan and Hula Nature Reserves. We surveyed stretches of lake shore that were adjacent to water at least 25 cm deep, had > 70% vegetation coverage, and had a non-composite soil profile that is raised at least 20 cm above the lake level. The sign survey focused on scats, holes, eating platforms, tracks, and nests. *Arvicola*-specific traps were purchased from Wildcare Nationwide Ecology Supplies (UK) and deployed in the surveyed regions during the 2nd–6th (Tel Dan) and 9th–12th (Hula) of June 2018, in regions where the pre-drainage outlines of the water bodies remained relatively unaltered. We used peanut butter snacks as a lure. Twenty-five traps were laid each night (Fig. 2).

Owl pellet recovery

Owl pellet contents are important means for reconstructing regional micro-faunal sequences (Dodson and Wexlar 1979; Kusmer 1990). Owls were effective ‘collectors’ of water voles in the Hula (Dor 1947). We thus attempted to recover owl pellets that span the time between the drainage (1951–1957) and the present. We examined a standing building from the ruins of the village of Tulayl (33.110N, 35.595E WGS84), on the banks of the Hula Lake, which was abandoned in the 1948 war (Supplement 1, Fig. S1-1). The roofed building was used by nocturnal raptors, and microfaunal remains and nesting materials were concentrated below the window and door lintels. We opened five 1 × 1 m excavation squares below window/door arches and excavated the

Fig. 2 Location of trap deployment in the Hula (left) and Dan (right) nature reserves. Map courtesy of Uri Roll



20–25 cm of deposits overlying the floor (Fig. S1-2). These deposits cannot be dated accurately, but their horizontal stratification (Fig. S1-3, Table S1-1) suggests gradual accumulation over decades. The sediments were sieved through 1-mm aperture mesh, and the rodent remains collected and identified to taxon based on craniodental morphology (Table S1-3).

Ancient DNA

DNA extraction

One mandible, collected in the 1940s from Yesod Hamaala, Upper Galilee, was used for genetic analysis. DNA was extracted in a dedicated ancient DNA laboratory that is physically isolated from post-PCR laboratory. We followed the protocol described in Dabney et al. (2013).

Library preparation

The DNA extract and negative controls were converted into genomic libraries following a modified single-tube BEST library protocol (Carøe et al. 2018): 32 µl DNA sample was mixed with 8 µl of a master mix consisting of 0.4 µl T4 DNA polymerase (NEB, cat#M0203S, 3 U/µl), 1 µl T4 Polynucleotide Kinase (NEB, cat#M0201S, 10 U/µl), 4 µl 10×T4 DNA Ligase Reaction Buffer (NEB), and 0.4 µl dNTP (25 mM). The end-repair reaction was incubated for 30 min at 20 °C followed by 30 min at 65 °C. For adapter-ligation, 2 µl of blunt-end adapter BEDC3 (20 µM) was added to the same reaction tube and mixed by pipetting, followed by 8 µl of master mix consisting of 6 µl PEG 4000 (Sigma-Aldrich, 50%), 1 µl T4 DNA Ligase Reaction Buffer (10×), and 1 µl T4 DNA ligase (NEB, cat#M0202S, 400 U/µl). The ligation reaction was incubated for 30 min at 20 °C followed by 10 min at 65 °C. The fill-in step was performed by adding 10 µl of master mix consisting of 0.8 µl dNTP (25 mM), 2 µl Isothermal Amplification Buffer (10×) (NEB), 5.6 µl molecular biology grade water, and 1.6 µl Bst 2.0 Warmstart polymerase (NEB, cat#M0538S, 8 U/µl). The fill-in reaction was incubated at 65 °C for 15 min, followed by 15 min at 80 °C. Following library preparation, the reactions were purified with a MinElute column following the manufacturer's instructions and eluted twice in 18 µl EB after 12 min incubation time at 42 °C.

Quantitative PCR

Quantitative PCR (qPCR) was performed on purified libraries, using 2 µl template of a 10× dilution of the libraries and a 5 µl of Power SYBR Green PCR master mix 2× (Applied Biosystems), 0.2 µM forward and reverse primer IS7 and IS8 (Meyer and Kircher 2010), and 2.6 µl molecular biology grade water to a final reaction volume of 10 µl. qPCR was

performed on StepOnePlus Real-Time PCR System instrument (Applied Biosystems). Cycling conditions included the following: 95 °C for 10 min, followed by 40 cycles of 95 °C for 15 s, and 60 °C for 60 s.

Indexing, PCR amplification, and sequencing

The two libraries were indexed and amplified for sequencing, using conventional full-length P7 and P5 Illumina primers with indexes (Carøe et al. 2018). PCR was performed in 25 µl reactions using 4 µl template, 2.5 µl 10×AmpliTaQ Gold buffer, 2.5 µl 25 mM MgCl₂, 1 µl 1 µg/µl Rabbit Serum Albumin (RSA), 0.2 µl 25 mM dNTP, 0.5 µl 10 µM forward and reverse indexed primers (specific for each sample), 0.3 µl 5 U/µl AmpliTaQ Gold polymerase and 4.5 µl molecular biology grade water to final volume of 25 µl. Libraries were amplified in the Applied Biosystems ProFlex PCR System using the following conditions: 95 °C for 10 min, followed by 20 cycles of 95 °C for 30 s, 60 °C for 60 s, and 72 °C for 45 min, followed by 5 min at 72 °C. Quantification and size estimation was performed with an Agilent 4200 TapeStation system. Sequencing was performed on an Illumina NextSeq™ 500 system for 75 cycles in single read mode.

Data processing

Raw reads were processed using AdapterRemoval (Schubert et al. 2016), trimming the adapters and low-quality bases, and discarding reads shorter than 30 bp. Reads were mapped against water vole mitogenome from GenBank MT381921 using BWA (Li and Durbin 2009) with BWA-backtrack logarithm and disabled seed. Reads mapping to the reference were sorted and duplicates removed using Picard (<http://broadinstitute.github.io/picard>), then realign to minimise the number of mismatches around the indels using GATK (<https://gatk.broadinstitute.org/hc/en-us>). MapDamage2.0 (Jónsson et al. 2013) was used to analyse the damage pattern in the fragments. We annotated the mitogenome using MITOS2 (beta version) (Donath et al. 2019), followed by manual validation of the coding regions using the NCBI ORF Finder (<http://www.ncbi.nlm.nih.gov/projects/gorf/>).

Phylogenetic analyses

We first built a dataset with complete mitogenomes of the Arvicolinae based on the published data in Abramson et al. (2021). The dataset included 27 complete voles and lemming mitogenomes and two water voles mitogenomes from Denmark (from GenBank, Accession numbers: MN12283 and MN122828). We used the concatenated alignment of 13 genes using MAFFT version 7.222 (Katoh et al. 2002). There are only three published complete mitogenomes of water voles. We therefore assembled a second dataset of cytochrome b (CYTB)

gene sequences using published data from Chevret et al. (2020). This dataset includes four *A. amphibius* lineages, *Arvicola sapidus*, and *A. persicus*. The total dataset included 268 individuals and 911 bp with common vole (*Microtus arvalis*), bank vole (*Myodes glareolus*), Père David's vole (*Eothenomys melanogast*), and Zaisan mole (*Ellobius tancrei*) as outgroups (Table S2-1; Chevret et al. 2020).

For both datasets, phylogenetic trees were reconstructed using maximum likelihood (ML) and Bayesian inference (BI). We ran ML and BI analyses with W-IQ-TREE (Trifinopoulos et al. 2016) and MrBayes v3.2 (Ronquist et al. 2012), respectively. The analyses were run according to the results of PartitionFinder (Lanfear et al. 2017), using W-IQ-TREE (Trifinopoulos et al. 2016). Robustness of the nodes was estimated with 1000 ultrafast bootstrap replicates (Hoang et al. 2018) and posterior probability with MrBayes. Markov chain Monte Carlo analyses were run independently for 3,000,000 generations with sampling every 100 generations, discarding the first 10% as burn-in. For the complete mitogenomes dataset, the analysis was carried out with genes partition only, while for the CYTB dataset, the analysis was carried out with codon partition. Trees were visualised and edited in FigTree v1.4.4 (Rambaut 2012). DNA sequences were deposited in GenBank (BankIt2793521 Seq1 PP291566).

Geometric morphometrics

We ran geometric morphometric analysis on 102 lower first molars to determine the similarity of the extirpated Hula water voles to recent Eurasian populations and to local archaeological specimens from Eynan (~ 14 kya) (Valla et al. 2017) and Nahal Mahanayim Outlet (NMO, ~ 65 kya) (Kalbe et al. 2015; Maul et al. 2021). Teeth were photographed at the Natural History Museum, London, the Steinhardt Museum of Natural History (Tel Aviv University), and the Natural History Collections of the Hebrew University (Table S3-1). The full database including landmark files and the script used in the geometric morphometric analysis are deposited in <https://doi.org/https://doi.org/10.5281/zenodo.8189164>.

Specimens were ascribed to a southern European group (EuS; $N=11$ from Macedonia, Italy, and Kosovo), a north European group (EuN; $N=50$, UK, Estonia, Finland, France, Germany, Switzerland, and Russia), a central Asian group (AsiaC; $N=13$ from Iran, Kazakhstan, and Turkey), and a local group (Nahal Mahanayim Outlet (IL-MP; $N=11$), Eynan (IL=Epi; $N=4$), and recent Hula (IL-Mod; $N=14$)), based on the specimens geographic origins.

All teeth were photographed by AOP using a Nikon D7500 camera with a Nikkor 40-mm DX micro lens. The photographs were taken perpendicular to the occlusal plane. Digitisation followed the protocol of Piras et al. (2012) for water vole m1, using TPSdig2 (Rohlf 2017). Procrustes transformation of the

landmarks and semi-sliding landmarks (*geomorph::gpgen*) was followed by principal components analysis (*geomorph::gm.prcomp*) and canonical variates analysis (CVA, *Morpho::CVA* with jackknife cross-validation) of the first ten principal components. Shape disparity between groups was calculated based on Procrustes distances (*geomorph::morphol.disparity*). Data analysis was carried out on R 4.2.1 using the libraries 'geomorph' (Adams and Otárola-Castillo 2013; Collyer and Adams 2018; Adams et al. 2022) and 'Morpho' (Schlager 2017). Data visualisation also used 'ggplot2' (Wickham et al. 2016) and 'ggsci' (Xiao 2018).

Species distribution modelling

Environmental data and study area

Climate data for current and future climatic scenarios were obtained from CHELSA (<https://chelsa-climate.org/>, accessed May 2023) at a 1-km (30 arc-sec) resolution for the period 1981–2010. We selected three global future climate models from CMIP6 (GFDL-ESM4, PSL-CM6A-LR, and MRI-ESM2-0) for the periods 2041–2070 and 2071–2100, under two shared socioeconomic pathways (ssp) scenarios (ssp126 and ssp585) which represent two contrasting scenarios of future greenhouse gas emissions, with ssp126 assuming a rapid and sustained reduction in emissions and ssp585 assuming a continued increase in emissions throughout the twenty-first century. We downloaded data for 19 bioclimatic variables, plus the mean monthly climate moisture index (CMI), the growing degree days heat sum above 0 °C (gdd0), the mean monthly near-surface relative humidity (hurs), the net primary productivity (NPP), and the mean monthly potential evapotranspiration (PET).

We also generated four topographic variables, elevation, aspect, slope, and roughness from a 1-km resolution from NASA Shuttle Radar Topography Mission (SRTM) Version 3.0 (SRTM Plus) digital elevation model using the function *terra::terrain* in R (Wilson et al. 2007). We further used distance to water (DW) by calculating the distance to water bodies using the function *rgeos::gDistance* in R (Bivand 2018). The water bodies' vectorial data, at 10-m resolution, were derived from World Data Bank 2 by smoothing and adjusting the river and lake outlines to fit a shaded relief generated. The study area was selected to encompass the current and historical geographical range of *A. amphibius*. Thus, all climatic, topographic, and environmental rasters were cropped to the extent 20°W to 140°E longitude and 20°S to 70°N latitude. Raster resolution was downscaled to a 5-km grid.

Occurrence data

Occurrence data for *A. amphibius* were downloaded from iNaturalist (only dated 'research-grade' occurrences with coordinate

uncertainty < 10,000 m, that were verified by ≥ 2 users and backed by photographs). We also used the *rgbif::occ_search* function in R to obtain occurrence data from GBIF. The query was limited to 10,000 records and filtered to select only human observations (museum specimens were discarded) from the year 1800 to the present and with a coordinate uncertainty < 10,000 m (Fig. S4-1). We cleaned the data (e.g. for coordinate uncertainty, taxonomic issues, country of origin) with the *CoordinateCleaner::clean_coordinates* function in R. The filtered and cleaned GBIF dataset included 9132 records. The GBIF database also contained some (but not all) of the iNaturalist data. We combined the datasets, removed duplicates, and adjusted the presence data to the climate raster resolution (so that there was only one presence point per pixel). The final dataset contained 1072 presences (Fig. S4-2).

Observations were heavily spatially clustered. Thus, we further generated pseudo-presence data from the IUCN Red List of Threatened Species distribution range which represents the current distribution of water voles. After adjusting the presence data to the climate raster resolution, the IUCN dataset contained 1129 pseudo-presences.

Absence data

We generated pseudo-absence data for the presence (iNat-GBIF) and pseudo-presence (IUCN) datasets. In the case of the iNat-GBIF dataset, we drew 100-km buffers around each presence point and took a random sample of pseudo-absences outside of these buffers throughout the study area. We generated 10 different sets of random pseudo-absences. For the IUCN dataset, the pseudo-absence points were taken outside of the polygon. The number of pseudo-absences equaled the number of presences and pseudo-presences, following Barbet-Massin et al. (2012), who recommended weighted presences and absences for tree-classifier algorithms (see sections below).

Variable importance and selection

For each presence-absence dataset, we removed variables with Variance Inflation Factors (VIF) > 10 using the *flexsdm::correct_colinvar* function in R (Velazco et al. 2022). The remaining variables were put through the automated stepwise variable set reduction procedure following the Bayesian additive regression trees (BART) method provided in the *embarcadero::var.step* function (Carlson 2020). In this method, the least informative variable is eliminated across all model iterations ($n = 200$) with 20 trees. The models are rerun (also with iterations = 200) without the least informative variable, recording the root mean square error (RMSE). The previous steps are repeated until there are only three covariates left. The number of variables to be dropped is decided based on the model with the lowest average RMSE.

Variable importance was examined using the *embarcadero::varimp.diag* function in R, set up to 200 iterations (Carlson 2020). We selected only the model runs with 20 trees (Chipman et al. 2010). Variable importance in BART results from counting the number of times a given variable is used by a tree split across the full posterior draw of trees (Cutler et al. 2007; Chipman et al. 2010; Carlson 2020). This method weighs the relative importance of each variable across all decision trees and thus spreads variable contribution across all covariates. This approach guards against the elimination of variables that could be good predictors of the response and may be ecologically important (Cutler et al. 2007).

Model implementation and evaluation

We modelled the water vole potential geographic distribution running BART models with the default parameters as implemented in *embarcadero::dbarts* in R (Carlson 2020). BART models were run using 200 trees and 1000 back-fitting Markov Chain Monte Carlo (MCMC) iterations, discarding 20% as burn-in. Maps with environmental suitability projections were generated for each of the future climatic projections and environmental suitability estimates were extracted at the Hula (35.599°N, 33.077°E) and Tel Dan (35.651°N, 33.248°E) nature reserves. In addition, a threshold that maximised the true skill statistic (TSS) was calculated for each BART model, which was then used to generate binary suitable/non-suitable maps.

Model performance was evaluated using a 5 k-fold cross-validation approach and model performance metrics: the correct classification rate (CCR), sensitivity, specificity, precision, recall, kappa, the operating characteristic curve (AUC), and true skill statistics (TSS) (Allouche et al. 2006; Liu et al. 2009; Márcia Barbosa et al. 2013). To facilitate comparisons across metrics, AUC and TSS were scaled to the interval 0 to 1 (Barbosa 2015), with values closer to unity indicating optimal prediction performance and values < 0.5 indicating that the model performs no better than random chance. We generated partial dependence plots for each of the variables used in the models to show the response curves of individual variables in relation to habitat suitability estimations. The response curves depict the average for each posterior draw of sum-of-trees models and true Bayesian credible intervals (Carlson 2020).

Results

Sign survey, trap survey, and owl pellets

No scats, tracks, holes, feeding platform, or other field signs of water vole have been found in the field survey. All

the traps deployed were recovered empty except two traps deployed at Tel Dan which caught a stone marten (*Martes foina*) and an Eastern broad-toothed field mouse (*Apodemus mystacinus*), which were released unharmed at the point of capture. No water voles were trapped in either locality.

The excavation in a-Tulayl yielded an assemblage yielded ~150 rodent remains including 27 craniodental fragments (NISP). These were identified as *Crocidura* ($N=2$), *Microtus* ($N=22$), *Rattus* ($N=2$), and (probably) *Mus* ($N=1$) (Table S1-2). No *Arvicola* remains were found.

Ancient DNA

The specimen from Yesod Hamaala was well preserved, and the mitochondrial genome was sequenced to high coverage (>50 , Table S2-1). The substitution model which best fit the 13 mitochondrial genes was GTR+I+G, while for the CYTB gene alone the TPM24+G+1 model was the best fit. For both datasets, BI and ML analyses yielded similar topologies. The phylogenetic tree shows that the Hula specimen (labelled 'MM1000_Israel') occupies a basal split within the *Arvicola* (100% ML bootstrap support and 1.0 Bayesian posterior probabilities) (Fig. S2-1).

In the CYTB tree, the Hula specimen falls with the *Arvicola amphibious* clade (99% ML bootstrap support and 1.0 Bayesian posterior probabilities), and groups with Italian samples with high support (95% ML bootstrap support and 0.91 Bayesian posterior probabilities) (Fig. 3; see supplementary Fig. S2-2). Overall, tree topologies are consistent with the outcomes of other molecular studies on Arvicolinae and on *Arvicola* (Chevret et al. 2020; Abramson et al. 2021).

Geometric morphometrics

The effects of regional grouping, size, and their interaction, on m1 shape (executed on the Procrustes transformed landmarks; Fig. 4a) are statistically significant (Table S3-2). The effect of group association ($R^2=0.24$, $P=0.001$) is by far the strongest (size, $R^2=0.01$, $P=0.02$; group/size interaction, $R^2=0.06$, $P=0.04$). Centroid size estimates (Fig. 4b) suggest that the recent Hula voles had a similar m1 size range as water voles from all other recent groups but were smaller, on average, and more variable in size than the Pleistocene specimens from the same region.

Shape PCA results (Fig. 4c) suggest that the main difference between specimens is the robusticity of the anterior cap and neck (see also Piras et al. 2012). When group structure is considered, the Pleistocene specimens from Hula have reduced anterior caps. Recent specimens (including Eurasian and Persian water voles) have a more robust anterior caps and necks. Most (77%, $N=10$) recent Hula specimens have positive PC1 values, like their Pleistocene predecessors, but three (23%) show strong negative values indicating a robust anterior cap. Such bimodality is not apparent in any of the other groups.

CVA and the Procrustes distance were used to examine the difference between groups. The CVA (Fig. 5a, see Table S3-3 for classification statistics) clearly separates Middle Palaeolithic Hula water voles from the other, including Hula Epipalaeolithic, groups. Cluster analysis based on Mahalanobis distance also places the terminal Pleistocene

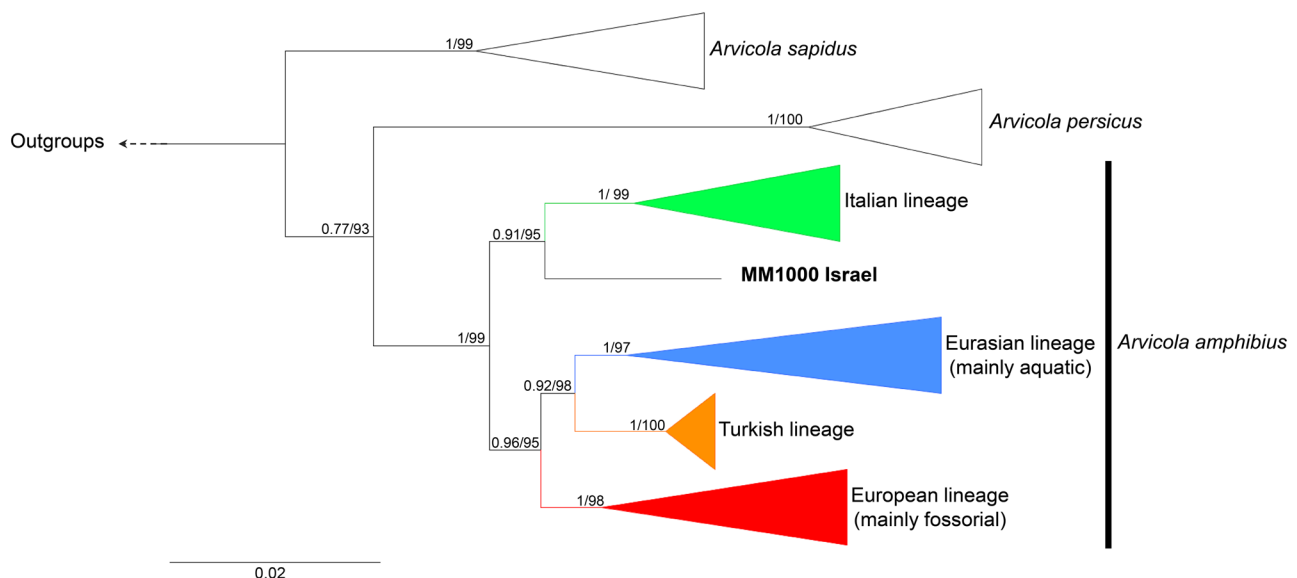


Fig. 3 A simplified Bayesian phylogeny of *Arvicola* water voles, with 268 individuals from Europe and Asia and based on a 911 bp of CYTB. Node support is presented by Bayesian posterior probabilities

and bootstrap supports respectively. Based on the data from Chevret et al. (2020). For a full tree see Fig. S2-2

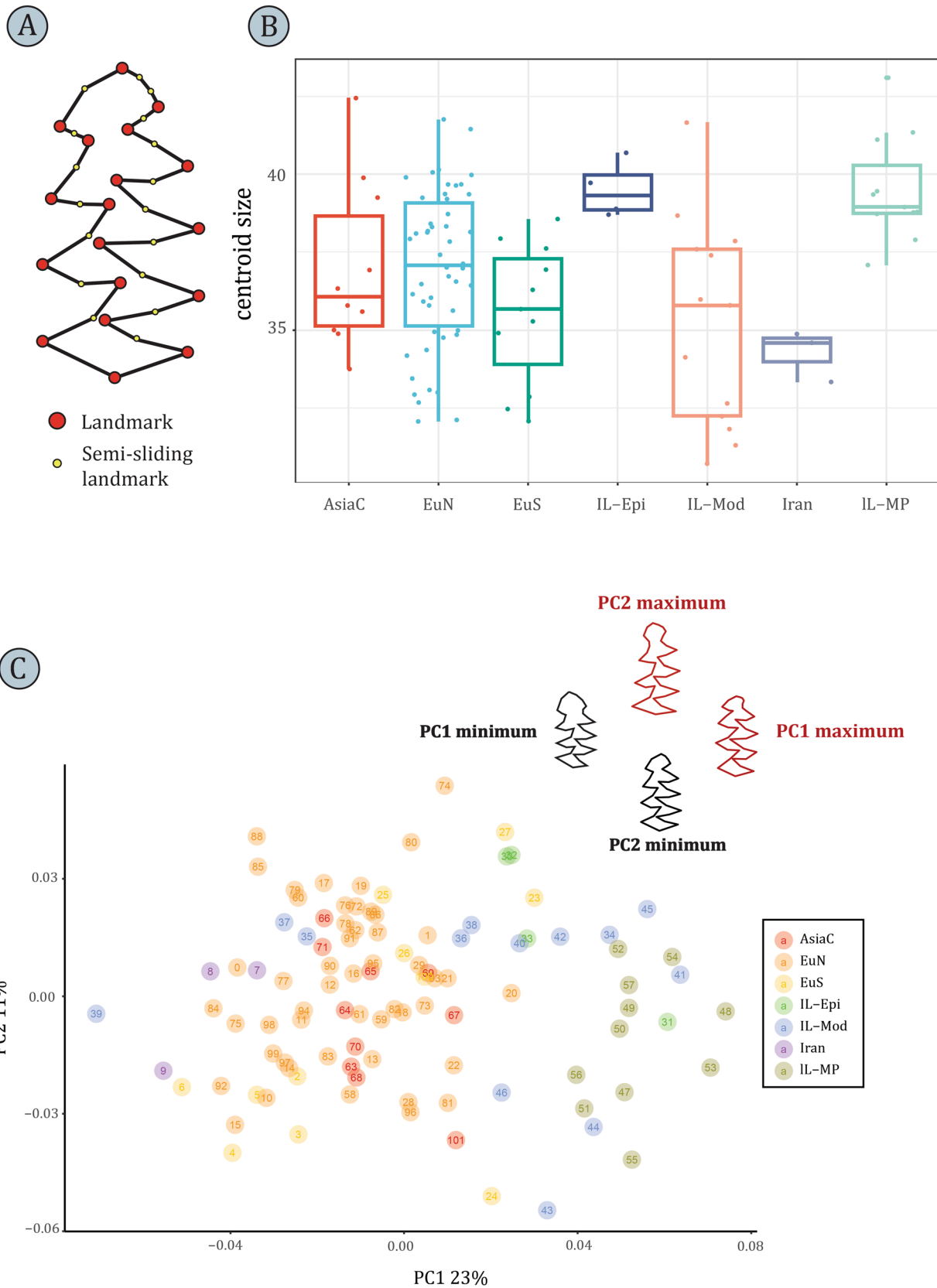
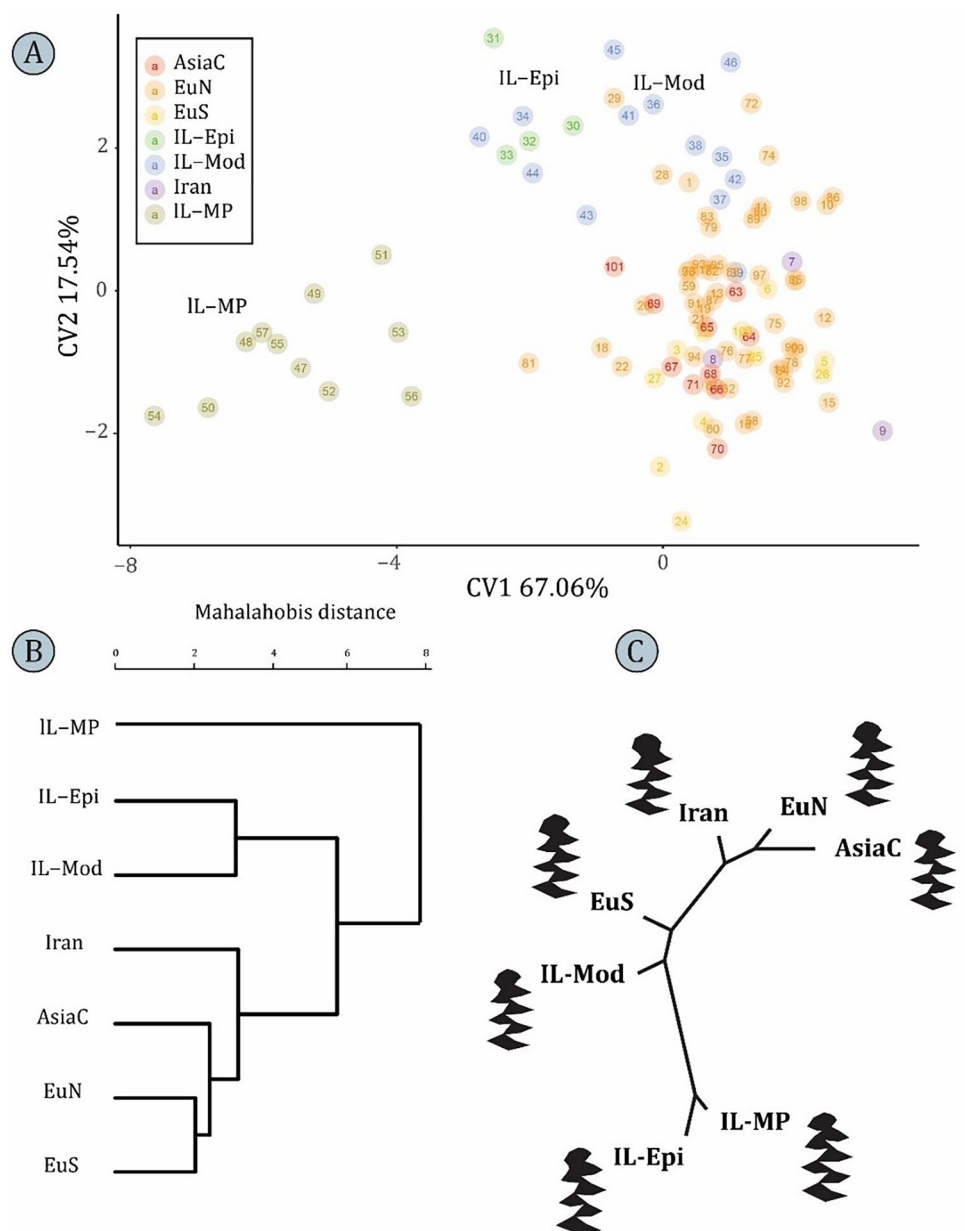


Fig. 4 **A** Landmark and semi-sliding landmark configuration. **B** Centroid size by group. **C** Shape PCA: colours denote groups and numbers inside circles are specimen index numbers (<https://doi.org/10.5281/zenodo.8189164>). Specimens EuS, southern European group; EuN, northern European group; AsiaC, central Asian group; IL-MP, Nahal Mahanayim Outlet; IL-Epi, Eynan; IL-Mod, recent Hula

and recent Hula water voles as separated from, but closer to, all other modern groups (Fig. 5b).

Procrustes distances (Fig. 5c, Table S3-2, Table S3-4) suggest close affinity between the late and terminal Pleistocene Hula groups. The recent Hula voles take an intermediate position with the southern European group between the archaic and the Eurasian voles.

Fig. 5 **A** Visualisation of the first two canonical variate scores colour-coded by group. Numbers are specimen identifications (S3-2). **B** Mahalanobis distance between groups based on the CV scores. **C** Neighbour-joining tree based on the inter-group Procrustes distance



Species distribution modelling

The BART selection procedure resulted in ten variables. These are, in order of importance, temperature seasonality (bio4), net primary productivity (npp), mean daily maximum air temperature in the warmest month (bio5), precipitation seasonality (bio15), mean monthly precipitation amount of the coldest quarter (bio19), elevation, mean temperature of driest quarter (bio8), mean diurnal air temperature range (bio2), mean monthly precipitation amount of the warmest quarter (bio18), and slope (Fig. S4-3).

The partial dependence plots (Fig. S4-4, Figure S4-5) indicate that water voles prefer productive habitats ($NPP > 1000 \text{ g C m}^{-2} \text{ year}^{-1}$), low temperature and precipitation seasonality,

maximum temperatures in the warmest month not surpassing 30 °C, mean diurnal temperature range < 0.2 °C, elevation < 1200 m, fewer than 500-mm rain in summer, and 300- and 700-mm rain in winter. The average AUC for the generated models was higher than 0.9 and the TSS was higher than 0.8, suggesting that the models performed adequately (Table S4-1, Fig. S4-6, Fig. S4-7). The mean environmental suitability estimates with BART for all future climate scenarios (Fig. 6) suggest low environmental suitability in the Hula and Dan nature reserves. None of the environmental suitability estimates were higher than the thresholds calculated for the models. Furthermore, there is a continuous trend towards less suitable habitats in the Levant in 2040, 2070, and 2100 as the projected range of the species shifts north,

regardless of the presence-absence dataset, climatic model, or carbon emissions scenario used (Fig. 7).

Discussion

We find unequivocal support for the notion that the water vole went extinct from the Hula Valley. The traps captured no water voles, and there were neither field signs nor remains in owl pellets, despite historical evidence of their common occurrence in the 1940s. Furthermore, a search for Arvicoline DNA in the Environmental DNA database recovered from Lake Hula (Perl et al. 2022) yielded negative results (Bina Perl, pers. comm.). Of 24 vertebrate species

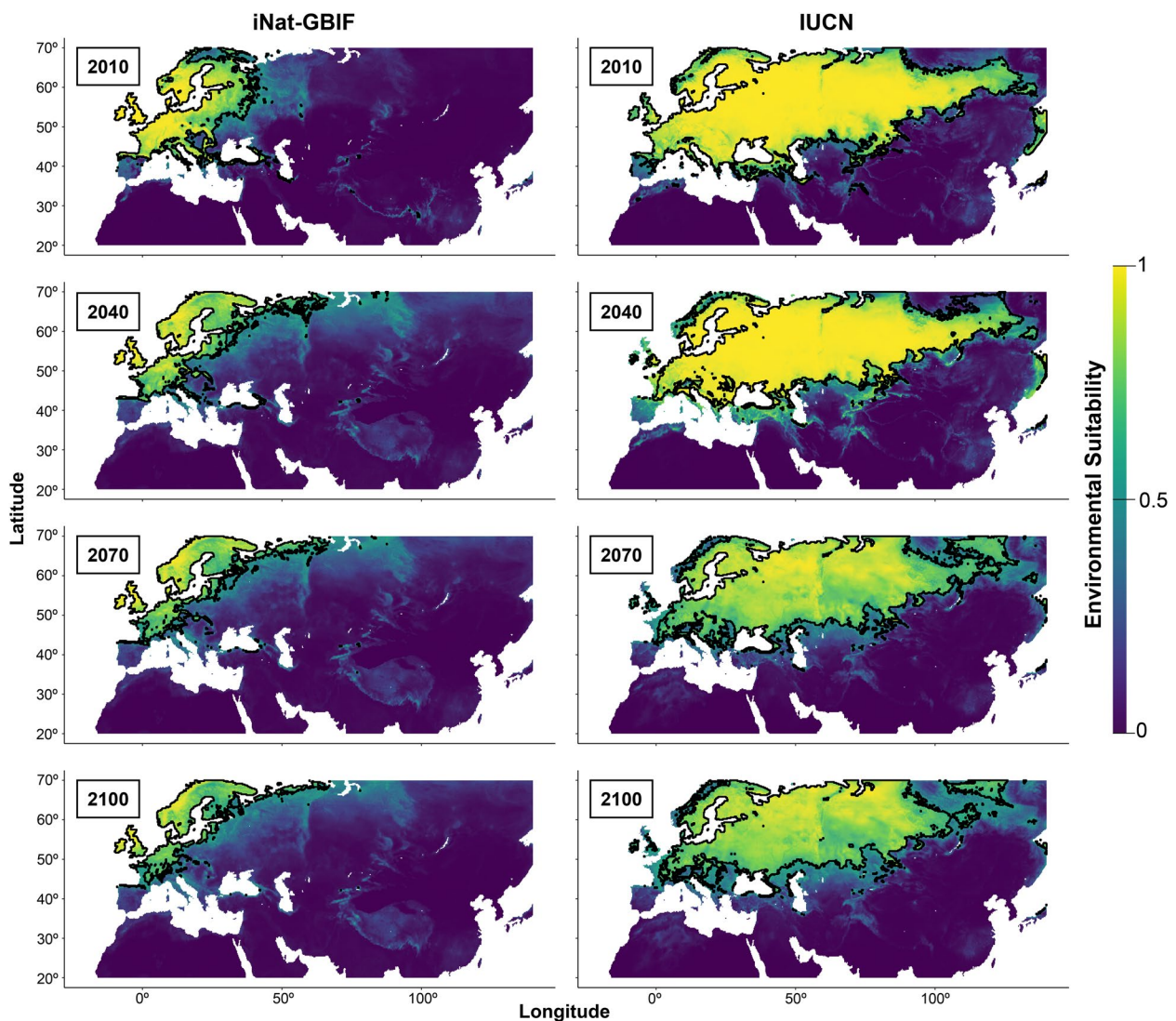


Fig. 6 Forecasted distribution of water vole suitable environments across Eurasia averaged for all climatic models and emission scenarios. The results obtained with the iNat-GBIF occurrence dataset

are shown in the left panels and the results obtained using the IUCN occurrence dataset are shown in the right panel

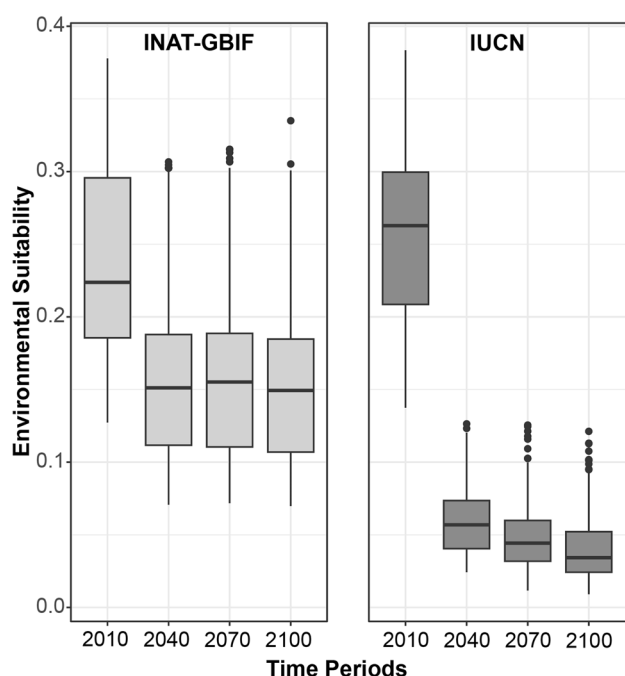


Fig. 7 Boxplots showing the mean environmental suitability estimations for water voles in the Hula Valley in both the INAT-GBIF and IUCN datasets. All environmental suitability estimations are under the average thresholds generated for the models (0.58 and 0.53 for INAT-GBIF and IUCN datasets, respectively)

sequences, only one has been identified as a rodent, and a BLAST result gave 100% conformance with a rat (*Rattus* sp. indet.). These results strongly support the widely acknowledged assumption that the local water vole population is extirpated, most likely because of the drainage of the lake in the 1950s.

A single mtDNA sequence of a twentieth-century water vole from the Hula clusters with recent southern European sequences of *A. amphibius*. While we are certain of this specific result, the question remains how well it represents the population from which it is derived. M1 geometric morphometrics suggests that shape variation is mainly in the anterior cap, which was suggested to relate to aquatic (gracile) and fossorial (derived, robust) ecomorphs (Piras et al. 2012). Most recent Hula specimens cluster with the Palaeolithic and Epipalaeolithic period specimens from the same region, but a few have a robust morphology—a bimodal distribution of this trait not observed in other populations. Specimen centroid sizes of recent Hula voles are likewise more variable than archaic populations. This Holocene diversity in size and shape may reflect post-glacial demographic expansion and admixture.

Discriminatory canonical variates analysis suggests that the Middle Palaeolithic voles form a distinct cluster, in accordance with the observations of Maul et al. (2021). Both canonical variates analysis and Procrustes disparity suggest closer relations of later Pleistocene Hula water voles to modern Eurasian *A. amphibius*. The recent Hula vole shapes are similar to southern European populations mirroring the results obtained using aDNA.

The species distribution probabilities indicate that the Hula Valley currently represents the southernmost boundary of potentially suitable water vole habitats in Asia. The Hula Valley population of the twentieth century was located at the edges of the water vole distribution. Although it is possible that we may have overlooked certain environmental factors influencing water vole distribution, the modelling results consistently indicate future decline in favourable climatic conditions in the Levant region and across Eurasia. This decline is associated with increasing aridity and temperatures in the southern regions, leading to a northward restriction of their potential range. A reintroduction of water voles in the local conditions of the Hula Valley will likely face substantial challenges from climate change, necessitating meticulous conservation planning and management.

In conclusion, we verified the extirpation of the species from the study region, produced evidence that the closest extant population is the south European one, and determined that the environmental condition in the coming decades would pose serious challenges to water vole reintroduction. Future decisions regarding water vole restoration and conservation in the Hula region should take these findings into consideration.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s10344-024-01781-8>.

Acknowledgements We wish to thank Bina Perl and Eli Geffen for their help with eDNA from the Hula region; to the National Parks Authority for facilitating the field surveys; to Lior Weissbrod, Rivka Rabinovitch, and Gonen Sharon for allowing us access to archaeological specimens used in this study; and to the British Museum of Natural History for access to their collections.

Author contributions Conceptualization: SM, NM; data collection: AOP, MM, IAL, NM; data analysis: IAL, MM, NM; writing: NM, IAL, MM, SM. All authors reviewed the manuscript.

Funding Open access funding provided by University of Haifa.

Data availability Genetic data can be found in GenBank [BankIt2793521 Seq1 PP291566]. Geometric morphometric source files and code are available in <https://doi.org/https://doi.org/10.5281/zenodo.8189164>.

Declarations

Competing interests The authors declare no competing interests.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing,

adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Abramson NI, Bodrov SY, Bondareva OV, Genelt-Yanovskiy EA, Petrova TV (2021) A mitochondrial genome phylogeny of voles and lemmings (Rodentia: Arvicolinae): evolutionary and taxonomic implications. *PLoS ONE* 16:e0248198. <https://doi.org/10.1371/journal.pone.0248198>
- Adams DC, Collyer ML, Kaliontzopoulou A, Baken EK (2022) Geomorph: software for geometric morphometric analyses. <https://cran.r-project.org/package=geomorph>. Accessed on 02 June 2023
- Adams DC, Otárola-Castillo E (2013) Geomorph: an R package for the collection and analysis of geometric morphometric shape data. *Methods in Ecology and Evolution* / British Ecological Society 4:393–399. <https://doi.org/10.1111/2041-210X.12035>
- Allouche O, Tsoar A, Kadmon R (2006) Assessing the accuracy of species distribution models: prevalence, kappa and the true skill statistic (TSS). *J Appl Ecol* 43:1223–1232. <https://doi.org/10.1111/j.1365-2664.2006.01214.x>
- Barbet-Massin M, Jiguet F, Albert CH, Thuiller W (2012) Selecting pseudo-absences for species distribution models: how, where and how many? *Methods in Ecology and Evolution* / British Ecological Society 3:327–338. <https://doi.org/10.1111/j.2041-210X.2011.00172.x>
- Barbosa AM (2015) fuzzySim: applying fuzzy logic to binary similarity indices in ecology. *Methods in Ecology and Evolution* / British Ecological Society. <https://doi.org/10.1111/2041-210X.12372>
- Batsaikhan N, Henttonen H, Meinig H, Shenbrot G, Bukhnikashvili A, Hutterer R, Kryštufek B, Yigit N, Mitsainas G, Palomo L, Batsaikhan N, Tsytsulina K, Formozov N, Sheftel B (2021) *Arvicola amphibius* (amended version of 2016 assessment). The IUCN Red List of Threatened Species 2021: e.T2149A Posted: 2008. Accessed on 05 May 2023
- Bivand MR (2018) Package “rgeos.” [cran.r-hub.io. https://cran.r-hub.io/web/packages/rgeos/rgeos.pdf](https://cran.r-hub.io/web/packages/rgeos/rgeos.pdf). Accessed on 25 Jul 2023
- Brace S, Ruddy M, Miller R, Schreve DC, Stewart JR, Barnes I (2016) The colonization history of British water vole (*Arvicola amphibius* (Linnaeus, 1758)): origins and development of the Celtic fringe. *Proceedings. Biological sciences / The Royal Society* 283. <https://doi.org/10.1098/rspb.2016.0130>
- Carlson CJ (2020) embarcadero: species distribution modelling with Bayesian additive regression trees in r. *Methods in Ecology and Evolution* / British Ecological Society 11:850–858
- Carøe C, Gopalakrishnan S, Vinner L, Mak SS, Sinding MHS, Samaniego JA, Wales N, Sicheritz-Pontén T, Gilbert MTP (2018) Single-tube library preparation for degraded DNA. *Methods in Ecology and Evolution* / British Ecological Society 9:410–419. <https://doi.org/10.1111/2041-210X.12871>
- Centeno-cuadros A, Delibes M (2009) Phylogeography of southern water vole (*Arvicola sapidus*): evidence for refugia within the Iberian glacial refugium? *Molecular. Wiley Online Library. https://doi.org/10.1111/j.1365-294X.2009.04297.x*
- Chevret P, Renaud S, Helvacı Z, Ulrich RG, Quéré J, Michaux JR (2020) Genetic structure, ecological versatility, and skull shape differentiation in *Arvicola* water voles (Rodentia, Cricetidae). *J Zool Syst Evol Res* 24:7. <https://doi.org/10.1111/jzs.12384>
- Chipman HA, George EI, McCulloch RE (2010) BART: Bayesian additive regression trees. *Ann Appl Stat* 4:266–298
- Cohen-Shacham E, Dayan T, Feitelson E, de Groot RS (2011) Ecosystem service trade-offs in wetland management: drainage and rehabilitation of the Hula, Israel. *Hydrol Sci J* 56:1582–1601. <https://doi.org/10.1080/02626667.2011.631013>
- Collyer ML, Adams DC (2018) RRPP: an r package for fitting linear models to high-dimensional data using residual randomization. *Methods in Ecology and Evolution* / British Ecological Society 9:1772–1779. <https://doi.org/10.1111/2041-210X.13029>
- Comay O, Ezov E, Yom-Tov Y, Dayan T (2022) In its southern edge of distribution, the Tawny Owl (*Strix aluco*) is more sensitive to extreme temperatures than to rural development. *Animals: an open access journal from MDPI* 12. <https://doi.org/10.3390/ani12050641>
- Cutler DR, Edwards TC Jr, Beard KH, Cutler A, Hess KT, Gibson J, Lawler JJ (2007) Random forests for classification in ecology. *Ecology* 88:2783–2792. <https://doi.org/10.1890/07-0539.1>
- Dabney J et al (2013) Complete mitochondrial genome sequence of a Middle Pleistocene cave bear reconstructed from ultrashort DNA fragments. *Proc Natl Acad Sci* 110:15758–15763. <https://doi.org/10.1073/pnas.1314445110>
- Dodson P, Wexlar D (1979) Taphonomic investigations of owl pellets. *Paleobiology* 5:275–284. <https://doi.org/10.1017/S0094837300006564>
- Donath A, Jühling F, Al-Arab M, Bernhart SH, Reinhardt F, Stadler PF, Middendorf M, Bernt M (2019) Improved annotation of protein-coding genes boundaries in metazoan mitochondrial genomes. *Nucleic Acids Res* 47:10543–10552. <https://doi.org/10.1093/nar/gkz833>
- Dor M (1947) Observations sur les micromammifères trouvés dans les pelotes de la chouette effraye (*Tyto alba*) en Palestine. 11:50–54. <https://doi.org/10.1515/mamm.1947.11.1.50>
- Hambright KD, Zohary T (1999) The Hula Valley (Northern Israel) Wetlands rehabilitation project. In: Streever W (ed) *An international perspective on wetland rehabilitation*. Springer, Netherlands, Dordrecht, pp 173–180. https://doi.org/10.1007/978-94-011-4683-8_18
- Hambright KD, Zohary T (1998) Lakes Hula and Agmon: destruction and creation of wetland ecosystems in northern Israel. *Wetlands Ecol Manage* 6:83–89. <https://doi.org/10.1023/A:1008441015990>
- Hoang DT, Chernomor O, von Haeseler A, Minh BQ, Vinh LS (2018) UFBoot2: improving the ultrafast bootstrap approximation. *Mol Biol Evol* 35:518–522. <https://doi.org/10.1093/molbev/msx281>
- Horvitz N (2015) BioGIS - Mammals - SPNI. Israel Nature and Parks Authority. Occurrence dataset, catalogue number 1167; serial number 386653. <https://doi.org/10.15468/foby5h>. <https://www.gbif.org/occurrence/1094465249>. Accessed on 15 Mar 2023
- Jónsson H, Ginolhac A, Schubert M, Johnson PLF, Orlando L (2013) mapDamage2.0: fast approximate Bayesian estimates of ancient DNA damage parameters. *Bioinformatics* 29:1682–1684. <https://doi.org/10.1093/bioinformatics/btt193>
- Kalbe J, Mischke S, Dulski P, Sharon G (2015) The Middle Palaeolithic Nahal Mahanayeem Outlet site, Israel: reconstructing the environment of Late Pleistocene wetlands in the eastern Mediterranean from ostracods. *J Archaeol Sci* 54:385–395. <https://doi.org/10.1016/j.jas.2014.04.018>
- Katoh K, Misawa K, Kuma K-I, Miyata T (2002) MAFFT: a novel method for rapid multiple sequence alignment based on fast

- Fourier transform. *Nucleic Acids Res* 30:3059–3066. <https://doi.org/10.1093/nar/gkf436>
- Kusner KD (1990) Taphonomy of owl pellet deposition. *J Paleontol* 64:629–637. <https://doi.org/10.1017/S0022336000042669>. Cambridge University Press
- Landfear R, Frandsen PB, Wright AM, Senfeld T, Calcott B (2017) PartitionFinder 2: new methods for selecting partitioned models of evolution for molecular and morphological phylogenetic analyses. *Mol Biol Evol* 34:772–773. <https://doi.org/10.1093/molbev/msw260>
- Lawton JH, Woodroffe GL (1991) Habitat and the distribution of water voles: why are there gaps in a species' range? *J Anim Ecol* 60:79–91. <https://doi.org/10.2307/5446>
- Li H, Durbin R (2009) Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 25:1754–1760. <https://doi.org/10.1093/bioinformatics/btp324>
- Liu C, White M, Newell G (2009) Measuring the accuracy of species distribution models: a review. *Proceedings 18th World IMACs/MODSIM Congress*. Cairns, p 4247
- Mahmoudi A, Maul LC, Khoshyar M, Darvish J, Aliabadian M, Kryštufek B (2020) Evolutionary history of water voles revisited: confronting a new phylogenetic model from molecular data with the fossil record. *Mammalia* 84:171–184. <https://doi.org/10.1515/mammalia-2018-0178>
- Márcia Barbosa A, Real R, Muñoz A-R, Brown JA (2013) New measures for assessing model equilibrium and prediction mismatch in species distribution models. *Divers Distrib* 19:1333–1338. <https://doi.org/10.1111/ddi.12100>
- Maul LC, Rabinovich R, Biton R (2021) At the southern fringe: extant and fossil water voles of the genus *Arvicola* (Rodentia, Cricetidae, Arvicolinae) from Israel, with the description of a new species. *Hist Biol* 33:2273–2793. <https://doi.org/10.1080/08912963.2020.1827240>
- Meiri S, Belmaker A, Berkowicz D, Kazes K, Maza E, Bar-Oz G, Dor R (2019) A checklist of Israeli land vertebrates. *Isr J Ecol Evol* 65:43–70. <https://doi.org/10.1163/22244662-20191047>
- Mendelsshon H, Yom-Tov Y (1999) Fauna Palaestina: mammalia of Israel. The Israel Academy of Sciences and Humanities
- Meyer M, Kircher M (2010) Illumina sequencing library preparation for highly multiplexed target capture and sequencing. *Cold Spring Harb Protoc* 2010:db.prot5448. <https://doi.org/10.1101/pdb.prot5448>
- Perl RGB, Avidor E, Roll U, Malka Y, Geffen E, Gafny S (2022) Using eDNA presence/non-detection data to characterize the abiotic and biotic habitat requirements of a rare, elusive amphibian. *Environmental DNA* 4:642–653. <https://doi.org/10.1002/edn3.276>
- Piras P, Marcolini F, Claude J, Ventura J, Kotsakis T, Cubo J (2012) Ecological and functional correlates of molar shape variation in European populations of *Arvicola* (Arvicolinae, Rodentia). *Zoologischer Anzeiger - A Journal of Comparative Zoology* 251:335–343. <https://doi.org/10.1016/j.jcz.2011.12.002>. Elsevier
- Rambaut A (2012) FigTree v1. 4. Molecular evolution, phylogenetics and epidemiology. Edinburgh: University of Edinburgh, Institute of Evolutionary Biology
- Renan S, Gafny S, Perl RGB, Roll U, Malka Y, Vences M, Geffen E (2017) Living quarters of a living fossil-Uncovering the current distribution pattern of the rediscovered Hula painted frog (*Latonina nigriventer*) using environmental DNA. *Mol Ecol* 26:6801–6812. <https://doi.org/10.1111/mec.14420>
- Rohlf JF (2017) TPSDig. Digitize landmarks and outlines. Stony Brook, NY: Department of Ecology and Evolution, State University of New York
- Ronquist F, Teslenko M, van der Mark P, Ayres DL, Darling A, Höhna S, Larget B, Liu L, Suchard MA, Huelsenbeck JP (2012) MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Syst Biol* 61:539–542. <https://doi.org/10.1093/sysbio/sys029>
- Schlager S (2017) Morpho and Rvcg – shape analysis in R: R-packages for geometric morphometrics, shape analysis and surface manipulations. In: Zheng G, Li S, Székely G (eds) *Statistical shape and deformation analysis*. Academic Press, London, pp 217–256. <https://doi.org/10.1016/B978-0-12-810493-4.00011-0>
- Schubert M, Lindgreen S, Orlando L (2016) AdapterRemoval v2: rapid adapter trimming, identification, and read merging. *BMC Res Notes* 9:88. <https://doi.org/10.1186/s13104-016-1900-2>
- Tchernov E (1975) Rodent faunas and environmental changes in the Pleistocene of Israel. In: Prakash I, Ghosh PK (eds) *Rodents in desert environments*. Monographiae Biologicae, 28. Springer, Dordrecht, pp 331–362
- Trifinopoulos J, Nguyen L-T, von Haeseler A, Minh BQ (2016) W-IQ-TREE: a fast online phylogenetic tool for maximum likelihood analysis. *Nucleic Acids Res* 44:W232–W235. <https://doi.org/10.1093/nar/gkw256>
- Tsipris J, Meron M (1998) Climatic and hydrological aspects of the Hula restoration project. *Wetlands Ecol Manage* 6:91. <https://doi.org/10.1023/A:1008459816898>
- Valla FR, Khalaily H, Samuelian N, Bocquentin F, Bridault A, Rabinovich R (2017) Eynan (Ain Mallaha). In: Enzel Y, Bar-Yosef O (eds) *Quaternary of the levant: environments, climate change, and humans*. Cambridge University Press, pp 295–302. <https://doi.org/10.1017/9781316106754.034>
- Velazco SJE, Rose MB, de Andrade AFA, Minoli I, Franklin J (2022) flexsdm: an R package for supporting a comprehensive and flexible species distribution modelling workflow. *Methods in Ecology and Evolution / British Ecological Society* 13:1661–1669. <https://doi.org/10.1111/2041-210X.13874>
- Wickham H, Chang W, Wickham MH (2016) Package ‘ggplot2’. Create elegant data visualisations using the grammar of graphics Version 2 1:189
- Wilson MFJ, O'Connell B, Brown C, Guinan JC, Grehan AJ (2007) Multiscale terrain analysis of multibeam bathymetry data for habitat mapping on the continental slope. *Mar Geodesy* 30:3–35. <https://doi.org/10.1080/01490410701295962>
- Xiao N (2018) ggsci: Scientific Journal and Sci-Fi Themed Color Palettes for “ggplot2.” <https://CRAN.R-project.org/package=ggsci>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.