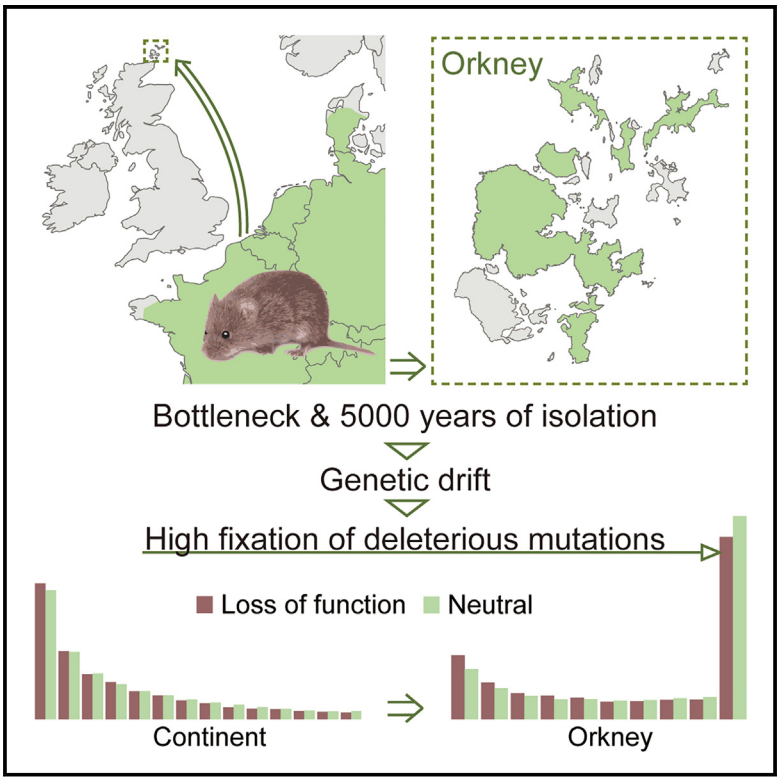
Article

Current Biology

# Demographic history and genomic consequences of 10,000 generations of isolation in a wild mammal

## Graphical abstract




## Highlights

- Common voles on Orkney remained completely isolated for more than 5,000 years
- Genetic drift led to a strong reduction of genetic diversity and population divergence
- Orkney voles have high levels of detrimental mutations, especially on small islands
- Simulations suggest purging of highly deleterious alleles, while mild ones persisted

## Authors

Xuejing Wang, Stephan Peischl, Gerald Heckel

## Correspondence

gerald.heckel@unibe.ch

## In brief

Wang et al. show that common voles on the Orkney archipelago have remained genetically isolated for more than 5,000 years after human introduction. Orkney voles lost most genetic diversity and harbor high levels of inferred strongly deleterious mutations, yet simulations and large current population sizes suggest rather mild effects on fitness.

Wang et al., 2023, Current Biology *33*, 1–16
May 22, 2023 
https://doi.org/10.1016/j.cub.2023.04.042

CellPress

Current Biology

CellPress
OPEN ACCESS

Article

# Demographic history and genomic consequences of 10,000 generations of isolation in a wild mammal

Xuejing Wang,[1] Stephan Peischl,[2,3] and Gerald Heckel[1,3,4,*]
[1]Institute of Ecology and Evolution, University of Bern, Baltzerstrasse 6, 3012 Bern, Switzerland
[2]Interfaculty Bioinformatics Unit, University of Bern, Baltzerstrasse 6, 3012 Bern, Switzerland
[3]Swiss Institute of Bioinformatics, Amphipôle, Quartier UNIL-Sorge, 1015 Lausanne, Switzerland
[4]Lead contact
*Correspondence: gerald.heckel@unibe.ch
https://doi.org/10.1016/j.cub.2023.04.042

## SUMMARY

Increased human activities caused the isolation of populations in many species—often associated with genetic depletion and negative fitness effects. The effects of isolation are predicted by theory, but long-term data from natural populations are scarce. We show, with full genome sequences, that common voles (*Microtus arvalis*) in the Orkney archipelago have remained genetically isolated from conspecifics in continental Europe since their introduction by humans over 5,000 years ago. Modern Orkney vole populations are genetically highly differentiated from continental conspecifics as a result of genetic drift processes. Colonization likely started on the biggest Orkney island and vole populations on smaller islands were gradually split off, without signs of secondary admixture. Despite having large modern population sizes, Orkney voles are genetically depauperate and successive introductions to smaller islands resulted in further reduction of genetic diversity. We detected high levels of fixation of predicted deleterious variation compared with continental populations, particularly on smaller islands, yet the fitness effects realized in nature are unknown. Simulations showed that predominantly mildly deleterious mutations were fixed in populations, while highly deleterious mutations were purged early in the history of the Orkney population. Relaxation of selection overall due to benign environmental conditions on the islands and the effects of soft selection may have contributed to the repeated, successful establishment of Orkney voles despite potential fitness loss. Furthermore, the specific life history of these small mammals, resulting in relatively large population sizes, has probably been important for their long-term persistence in full isolation.

## INTRODUCTION

Isolated populations are of major concern in conservation biology, human and animal health, and evolutionary biology.[1,2] As human activities have affected and will continue to affect ecosystems worldwide, fragmentation of many species into isolated populations is ongoing and may lead to increased risks of local or global extinction.[3,4] Human transfer of organisms to previously unoccupied areas can result in populations that may either vanish or thrive and turn into biological invasions given suitable conditions.[5,6] Even though both scenarios typically involve isolated populations, their demographic trajectories are diametrically opposed and are thought to be influenced by the amount and quality of genetic variation.[7]

Isolation of populations leads to loss of genetic diversity over time, and the rate of the loss is inversely related to the effective population size ($N_e$).[2] In small populations, the loss of genetic diversity is mostly governed by random genetic drift, which affects neutral and non-neutral genetic variation equally, and might leave long-lasting signatures even if a population expanded afterward.[8,9] Deleterious mutations may reduce the fitness of populations, a phenomenon for which the term "mutation load" was coined.[10–13] Deleterious mutations are usually present at low frequency in large populations and mostly express their negative fitness effects when inbreeding occurs (inbreeding load).[14] The loss of diversity in small populations can cause recessive deleterious variants to drift to high frequency and be more likely present in homozygous form. This can lead to an increased mutation load (drift load)[15] of mildly and moderately deleterious mutations but also to efficient purging of strongly deleterious mutations over time by purifying selection.[16–18]

Theory has shown that, at the front of spatially expanding populations initially going through a series of bottlenecks, the proportion of deleterious alleles in a homozygous state may increase while the total number of deleterious alleles remains unchanged.[19,20] Such an increase correlates with a decrease in fitness of derived populations and has thus been called "expansion load." The expansion load can last for thousands of generations, especially in the absence of gene flow from core populations.[21] Empirical studies on the genetic effects of deleterious variation in isolated populations have mainly focused either on model species[19,22,23] in experiments[24,25] or on

Please cite this article in press as: Wang et al., Demographic history and genomic consequences of 10,000 generations of isolation in a wild mammal, Current Biology (2023), https://doi.org/10.1016/j.cub.2023.04.042



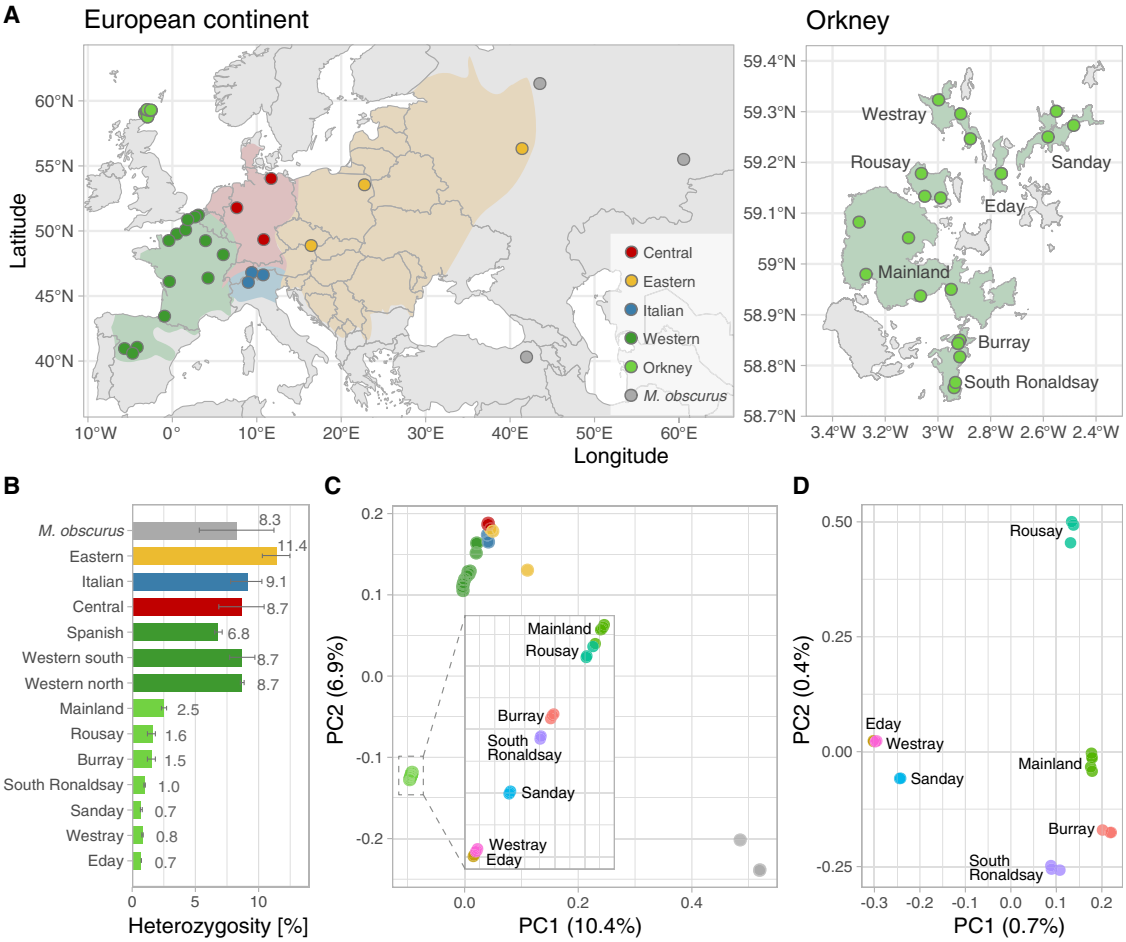



**Figure 1. Distribution, genetic diversity and divergence of common voles**

(A) Sampling locations and distribution range of the common vole *Microtus arvalis* modified after IUCN Red List.[42,43] Major evolutionary lineages in *Microtus arvalis* are coded as: Eastern lineage, yellow; Italian lineage, blue; Central lineage, red; Western lineage, dark green; Orkney, light green. Sampling locations of *M. obscurus* are marked with gray dots. One genome was sequenced for each location except for Eday (n = 2) and one location of Burray (n = 2).

(B) The average percentage of heterozygous sites and the standard deviation marked as the error bar in each group.

(C) Principal component analysis (PCA) of genetic variation based on 57.5 million SNPs without missing data (54.9% of all) with a zoom of the Orkney voles in the insert (x axis: −0.097 to −0.092; y axis: −0.128 to −0.118). Patterns in Orkney populations correlate with heterozygosity in (B) and indicate the varied strength of genetic drift during the history.

(D) PCA of Orkney vole genomes only.

See also Tables S1–S3.

endangered wild populations,[26–30] but knowledge of the effects in the wild exceeding a period of a few hundred generations[29,30] remains scarce.

Islands are natural systems to study the short- and long-term effects of isolation on populations.[2,31,32] The Orkney archipelago in the north of Scotland has served as an iconic model for studying the history of Neolithic human settlements and culture[33] and associated organisms.[34,35] The Orkney vole (*Microtus arvalis orcadensis*)[36] results from one of the oldest-known human introductions of a wild species to an island system[37,38] and provides a unique system to study the long-term consequences of genetic isolation. Common voles on Orkney are geographically separated from their conspecifics in continental Europe (Figure 1) and were once believed to be an independent species based on morphology. Common voles are small herbivorous rodents with a very high capability of reproduction and extensive population size fluctuations in many areas.[39] In Europe, four major evolutionary lineages are found that are genetically and geographically distinct but morphologically cryptic (Figure 1A).[40,41]

Limited mitochondrial and nuclear DNA information has shown that Orkney voles belong to the Western evolutionary lineage within the species and were most likely introduced by Neolithic farmers from coastal Belgium or France more than 5,000 years ago,[37,38] probably as a food item.[44] Orkney voles are present and abundant on seven of the islands, with their census population size estimated to be over one million.[45] Their bones have been found on multiple Orkney islands in archeological sites dating back to the Neolithic, which supports early splits between vole populations[38,44,46,47] at a time when most islands of Orkney were already isolated by the sea.[48,49] However, the last intentional introduction of voles to an unoccupied Orkney island,

Please cite this article in press as: Wang et al., Demographic history and genomic consequences of 10,000 generations of isolation in a wild mammal, Current Biology (2023), https://doi.org/10.1016/j.cub.2023.04.042





**Table 1. Genomic diversity of common voles (*Microtus arvalis*) and the outgroup *M. obscurus***

| Lineage | Group | Number of genomes | Heterozygosity | Deleterious homozygous | Deleterious heterozygous | Deleterious sites total | Deleterious alleles |
|---|---|---|---|---|---|---|---|
| *M. obscurus* | – | 3 | 8.3% | – | – | – | – |
| Eastern | – | 3 | 11.4% | – | – | – | – |
| Italian | – | 3 | 9.1% | – | – | – | – |
| Central | – | 3 | 8.7% | – | – | – | – |
| Western | north | 7 | 8.7% | 1,371 | 2,661 | 4,032 | 5,468 |
| | south | 3 | 8.7% | 1,598 | 2,849 | 4,447 | 6,045 |
| | Spanish | 3 | 6.8% | 1,995 | 2,456 | 4,451 | 6,446 |
| | overall | 13 | 8.2% | 1,585 | 2,657 | 4,242 | 5,827 |
| Orkney | Mainland | 5 | 2.5% | 2,419 | 911 | 3,330 | 5,750 |
| | Rousay | 3 | 1.6% | 2,580 | 597 | 3,177 | 5,757 |
| | Burray | 3 | 1.5% | 2,602 | 597 | 3,199 | 5,801 |
| | South Ronaldsay | 3 | 1.0% | 2,690 | 431 | 3,121 | 5,813 |
| | Sanday | 3 | 0.7% | 2,662 | 334 | 2,996 | 5,659 |
| | Westray | 3 | 0.8% | 2,683 | 383 | 3,066 | 5,749 |
| | Eday | 2 | 0.7% | 2,705 | 331 | 3,036 | 5,741 |
| | overall | 22 | 1.4% | 2,598 | 556 | 3,155 | 5,753 |

Groups in the Western lineage are defined based on geographical origin. Heterozygosity is shown as the average percentage of heterozygous sites. Deleterious genetic variation in Orkney voles and the Western evolutionary lineage is reported as the average numbers of derived deleterious homozygous and heterozygous sites, deleterious sites, and deleterious alleles, rounded up to the whole number.

Eday, occurred only in 1987 (M. Cockram, personal communication), which indicates the potential of human interference with the demographic and genetic history of seemingly isolated island populations.

Here, we use full genomic data to first reconstruct the history of Orkney voles throughout the Neolithic until the modern period relative to their continental conspecifics. On this base, we then assess the genomic consequences of ancient colonization events in Orkney potentially leading to strong genetic bottlenecks, thousands of generations of isolation, and mutation load. This allows us to empirically test some of the theoretical predictions regarding changes in the number, frequency, and heterozygosity of deleterious variation in natural mammal populations after extremely long isolation.

## RESULTS

### Genomic diversity and diversification

We sequenced 45 common vole and three additional *Microtus obscurus* genomes at a mean read depth of 26× (Tables 1 and S1). This included 22 individuals from all seven Orkney islands occupied by *M. arvalis*, 14 individuals chosen to cover the continental distribution range of the Western evolutionary lineage in the species, and three samples from each of the other major evolutionary lineages (Figure 1A). The genetic diversity of Orkney voles, reported as the average percentage of heterozygous sites among all called single-nucleotide polymorphism (SNP, 100.8 million in total) sites of each population, was extremely low compared with continental populations (Figure 1B; Table 1). On the European continent, genetic diversity in the Eastern evolutionary lineage was the highest (11.4%). Within the Western evolutionary lineage, Spanish populations had the lowest heterozygosity (6.8%), while Western-south and Western-north populations had equal levels (8.7%) despite the divergence within the lineage. In Orkney, voles on Mainland Orkney had the highest heterozygosity (2.5%), which is 1.5–3.4 times more than in the other islands (Figure 1B). Orkney voles on islands closer to Mainland Orkney had higher heterozygosity than voles on the islands—Sanday, Westray, and Eday—farther to the north (Figures 1A and 1B). Though the Eday vole population started from only 15 individuals from Westray released in 1987, the population did not show a strong decrease in heterozygosity compared with Westray.

A principal component analysis (PCA) of all genomes without missing data showed Orkney voles as highly divergent from the continental populations (Figure 1C). Most of the variance was explained by differences between *M. arvalis* and *M. obscurus*, but Orkney voles formed a cluster separate from all other individuals spanning the distribution range of the species between Spain and Russia. On PC1, PC2, and PC3, Orkney voles were closest to individuals from the northern part of the Western evolutionary lineage (Table S2). In the detailed view, populations are ordered almost perfectly linear to levels of heterozygosity (Figure 1B) probably due to the effects of genetic drift. Orkney Mainland was closest to the continental populations, which gives support to this largest island as the original founding site of Orkney voles. The pattern in the PCA showed Rousay and Burray as genetically closest to Mainland Orkney, consistent with the geography. South Ronaldsay was beyond Burray, and populations from the most distant northern islands, Sanday and Westray, were farther out (Figure 1A; Table S3). As expected, the Eday population was close to Westray voles that were the source of the introduction. Focusing only on Orkney voles confirmed the genetic distinctness of populations on separate islands and little variation within populations. The three northern islands separated from the rest on PC1, and the

Please cite this article in press as: Wang et al., Demographic history and genomic consequences of 10,000 generations of isolation in a wild mammal, Current Biology (2023), https://doi.org/10.1016/j.cub.2023.04.042



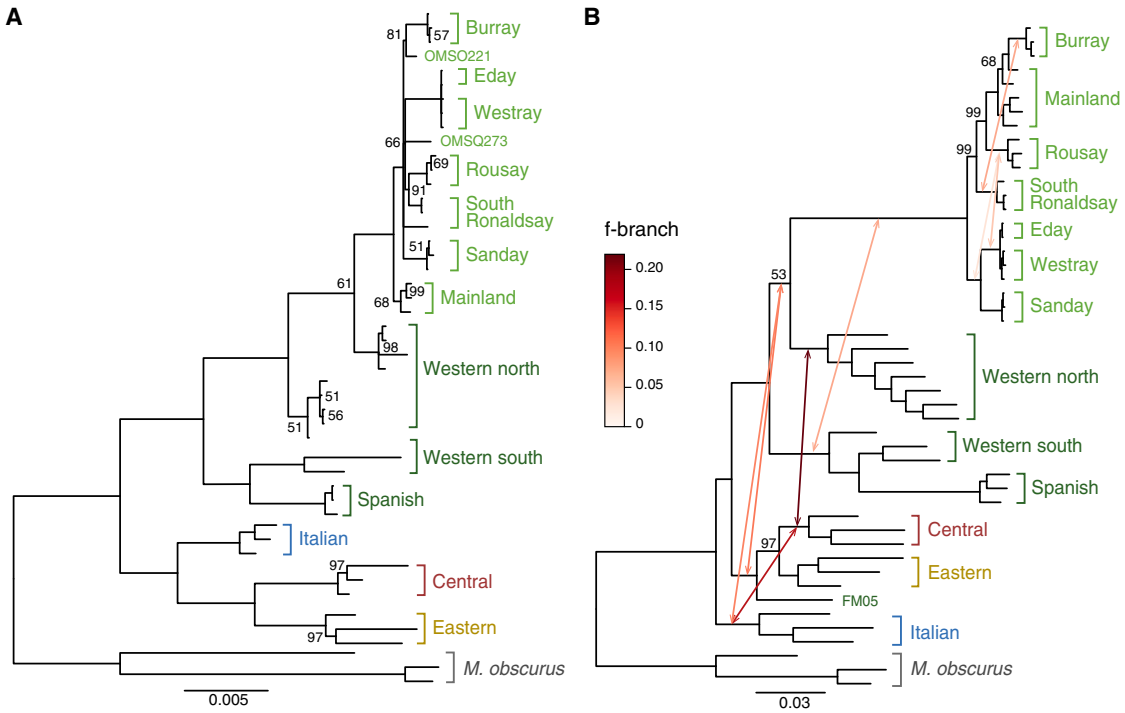



**Figure 2. Phylogenetic relationships and gene flow inference**

(A) Bayesian phylogeny based on complete mitochondrial vole genomes.

(B) ML tree conducted with IQ-TREE based on 5 million random nuclear single-nucleotide polymorphism sites without missing data. Arrows show the *f*-statistics with $p < 0.01$ and the colors representing the values of *f*-statistics. Note that *f*-statistics do not represent the direction of gene flow. All posterior probabilities in (A) and (B) were equal to 1 unless shown as percentage on the tree. Information on evolutionary lineage and grouping of the vole individuals is given in Table 1. Individuals that did not cluster with the geographical groups are marked with the sample name.

See also Figures S1 and S2.

southern islands separated from each other on PC2 (Figure 1D; Table S4).

## Phylogeny and admixture

Phylogenetic analyses of both nuclear and mitochondrial genomes showed Orkney voles as distinct, with the closest affinity to individuals from the northern area of the Western evolutionary lineage (Figure 2). Mainland individuals were either at basal positions or grouped with voles from other islands on the mitochondrial tree, potentially due to preserved ancestral diversity in this biggest Orkney population.[38]

Nuclear genomes strongly supported the notion that Orkney voles have remained isolated from continental populations for a very long period, as no evidence of post-introduction admixture between Orkney populations and the continent was found (Figures 2B, S1A, and S2). Signals of genomic admixture were mostly detected between basal branches on the continent, which is consistent with the sharing of ancestral variation and limited gene flow due to incomplete reproductive isolation between evolutionary lineages.[42,50–52] An individual from northeastern France (FM05) showed strong admixture (Figure S1A) between the Western-north and the Central lineage and was thus not included in further, more specific analyses. *f*-branch statistics suggested also limited ancestral gene flow between continental lineages and Orkney voles (Figures 2B and S2). The signal of admixture between the Western-south branch and ancestral Orkney is likely a consequence of shared ancestral variation that was passed over to the Orkney population at the time of founding (for details, see Malinsky et al.[53]). The results of ADMIXTURE analyses support that Orkney vole genomes are distinct from current continental populations, with only very minor sharing of ancestries (Figure S1).

Within the monophyletic cluster of Orkney samples, individuals from each island clustered separately except for Orkney Mainland (Figures 2 and S1B), suggesting isolation and divergence between the islands. Populations from the northern Orkney islands (Sanday, Westray, and Eday) were diverged from the rest in the phylogeny and in ADMIXTURE analyses (Figure S1B). *f*-branch statistics suggested multiple admixture events between the Orkney islands, mainly at the early stage of divergence (Figures 2B and S2). Treemix or ADMIXTURE detected no signals of admixture between islands, except for the South Ronaldsay and Burray populations (Figure S1B), where recent gene flow between these nearby islands cannot be excluded.

## Demographic history of Orkney voles

We used the Orkney Mainland population to represent Orkney voles in demographic analyses because of its preserved diversity and absence of potential admixture after introduction (Figures 2B and S2). Analyses with SMC++[54] showed a strong drop in the effective population size of Orkney Mainland voles

Please cite this article in press as: Wang et al., Demographic history and genomic consequences of 10,000 generations of isolation in a wild mammal, Current Biology (2023), https://doi.org/10.1016/j.cub.2023.04.042



CellPress
OPEN ACCESS

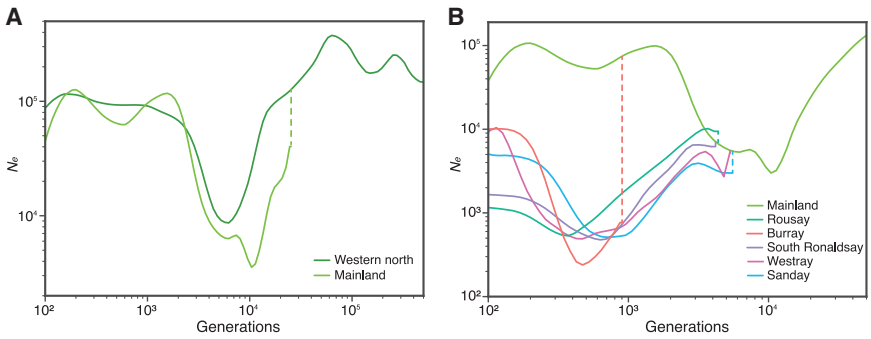



**Figure 3. Demographic history of Orkney voles**

(A) Demographic histories of common voles on Orkney Mainland relative to the genetically and geographically closest Western-north populations on the continent inferred by SMC++. Solid curves show variation in the effective population size $N_e$ over time and the dashed vertical line represents the inferred time of divergence. The generation time is 0.5 year.

(B) Demographic histories of voles on smaller Orkney islands relative to Orkney Mainland. Multiple analyses were combined into one graph for comparison. The Eday population is not shown because of its very recent introduction from Westray in 1987. See also Table S4.

ca. 5,300 years ago (bootstrap median 4,875, 5$^{th}$ percentile 3,014, 95$^{th}$ percentile 5,410), with a minimum $N_e$ of 2,781 (Figure 3B; bootstrap median 3,153, 5$^{th}$ percentile 2,275, 95$^{th}$ percentile 3,771). The time of this bottleneck closely fits the dates from radiocarbon analyses on vole bones (oldest 5,100 years old) and the previously estimated time (about 5,000 years) of the introduction to Orkney.[38] The split between Orkney Mainland and the continental population was estimated to be ~8,400 years ago (Figure 3A; Table S4). This is much older than the bottleneck on Orkney Mainland and probably a consequence of the Western-north individuals not being the descendants of the precise source populations for the founding of Orkney (see discussion).[38] After a relatively slow initial population size increase in Orkney, the vole population grew particularly fast in the last millennium and remained, with some fluctuation, at an $N_e$ of roughly 100,000. Surprisingly, this is the same order of magnitude that the population size in the Western-north group reached after a less severe bottleneck about 2,000–3,000 years ago.

The comparison between Orkney islands showed that the demographic history of Orkney Mainland voles differed from populations on the smaller islands, but there were similar trends among the latter (Figure 3B). With the Mainland population as the potential source, the splits of most island populations (Sanday, Westray, Rousay, and South Ronaldsay) were estimated at between 2,000 and 3,000 years ago (Table S4). The population on Burray split only about 450 years ago, at a time when all Orkney populations, including the Mainland, experienced a phase with the lowest effective population sizes in their history (Figure 3B). Populations on the smaller islands showed partially signs of very strong growth only in the last 200 years (e.g., on Burray from ~250 to 10,000). We attempted also to get insights into the history of the recently introduced voles on Eday through analyses with SMC++. The estimated population size changes are relatively similar to those on Westray, the source for the introduction. The split time of about 3 years ago (Table S4) is clearly an underestimate, but this is not surprising because of the limits of SMC++ to estimate very recent demographic changes. Events occurring less than ~160 generations since introduction are too recent to be estimated reliably.[55] Orkney voles had runs of homozygosity (ROH) longer than 1 Mb, covering a total of 2.1%–5.4% of the genome (mean = 3.6%), which is 8–20 times longer compared with the continental voles (mean = 0.25%; range 0%–0.35%). However, the ROH regions in Orkney voles were short (mean 1.4 Mb, longest 3.4 Mb; mean continental voles 1.6 Mb, longest 4.1 Mb), indicating that the ROHs resulted from the deeper demographic history rather than inbreeding due to recent common ancestors.

## Accumulation of deleterious mutations

Assessing the potential functional relevance of genetic variation showed that Orkney voles accumulated high amounts of homozygous, putatively deleterious mutations in their specific history compared with continental populations (Figures 4A, 4B, and S3A; Table 1). We identified in total 11,562 SNPs and 22,006 indels with defined derived states in Orkney that were classified as highly deleterious (mainly loss-of-function mutations, e.g., frameshift mutations; see STAR Methods), with 9.7 million neutral SNPs and 1.8 million neutral indels as comparison. On Orkney Mainland, over 40% of the deleterious variants were fixed in the population. On smaller islands, 63.6%–89.9% of the deleterious variants were fixed (Figures 4A and S3A). The majority of the fixed variants were shared between Mainland Orkney and the other islands, probably inherited from the ancestral population of Orkney (Figure S4). Gene enrichment analysis suggested that the affected genes were distributed widely in many fundamental pathways without specific enrichment, and we did not detect under-represented gene ontology terms (results not shown).

We used individual-based simulations to investigate whether this large proportion of fixations is a plausible outcome of the estimated demographic history. Using a distribution of fitness effects (gamma distributed) with mean $s = -0.01$ (Ns = 50; N being the effective population size during the bottleneck), we found very good agreement between observed and simulated site frequency spectra (SFS) for both neutral and deleterious alleles (Figure 4A). The mean strength of selection and the shape parameter of the distribution of fitness effects (DFE, alpha = 1) were chosen such that a broad range of selection coefficients are contained in the DFE. Using refined DFEs (Figure S5) that incorporate a negative relationship between the strength of selection and the dominance coefficient yielded very similar SFS,[56] in particular whether the DFE was wide enough to include a substantial proportion of nearly neutral mutations (Ns < 1 during the bottleneck, Figure S5). Re-tracing the fixation events in the simulations showed that they mostly occurred during the ancestral bottleneck more than 5,000 years ago (Figure S5C). Furthermore, virtually all mutations that became fixed fell below

Please cite this article in press as: Wang et al., Demographic history and genomic consequences of 10,000 generations of isolation in a wild mammal, Current Biology (2023), https://doi.org/10.1016/j.cub.2023.04.042



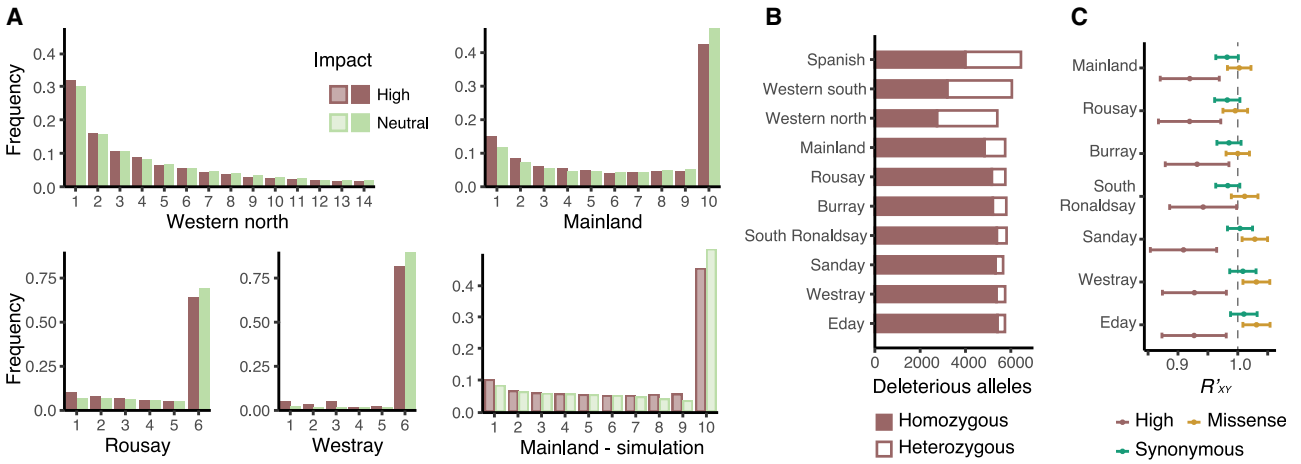



**Figure 4. Accumulation of deleterious genetic variation in Orkney voles**

(A) Unfolded site frequency spectra of derived high-impact (dark red) and neutral sites (green) for real data, including nucleotide polymorphism sites (SNP) and indels, or for simulated data. A very high proportion of deleterious mutations were fixed in Orkney vole populations.

(B) Number of high-impact deleterious alleles per individual for Orkney and continental voles from the Western evolutionary lineage. Deleterious alleles were more often found in homozygous state in Orkney voles.

(C) $R'_{XY}$ of highly deleterious, missense, and synonymous SNPs of Orkney populations. The error bars stand for ± 2 standard errors.

See also Figures S3–S5.

the threshold for nearly neutral mutations during the bottleneck (Figure S5). This shows that only mildly deleterious mutations could fix, whereas more strongly deleterious mutations were successfully purged from the population in the simulations. Including a proportion of lethal and sub-lethal mutations (3% of mutations with s = −0.75 or s = −1) did not affect our simulation results (Figure S3), in line with the observation that genomic patterns of deleterious diversity are composed almost exclusively of small effect mutations across all simulations.

We investigated the accumulation of deleterious alleles in empirical data with the $R_{XY}$ method.[57,58] The ratio $R'_{XY}$ of highly deleterious mutations in Orkney populations was close but significantly lower than 1 when calculated against the Western-north group (Figure 4C). This showed that a number of highly deleterious alleles were removed from Orkney populations due to purging,[58] especially in the northern islands (Sanday, Westray, and Eday). However, $R'_{XY}$ values of missense SNPs exceeding 1 suggest that the removal of weakly deleterious alleles was less efficient in Orkney populations (Figure 4C). We further tested for evidence of pseudo-overdominance (POD) in Orkney voles that could potentially reduce the expression of fitness effects. POD is a form of balancing selection in regions of low recombination that leads to the masking of recessive deleterious variants by maintaining complementary haplotypes at high frequencies.[59,60] However, we did not find signatures of POD, such as an excess of deleterious mutations in low recombination rate regions nor an excess of intermediate frequency variants in the SFS of deleterious variants (Figures 4A and S4). Also, the SFS from our simulations showed no evidence of potential POD affecting Orkney voles (Figure 4A and S5).

The genomic patterns of homozygosity of Orkney populations were consistent with the theoretical predictions for range expansions. The number of sites with derived deleterious alleles in a homozygous state observed per genome was strongly correlated with the number of sites with derived neutral alleles in homozygous states ($R^2$ = 0.992, p < 2.2e−16). The same pattern was found for the number of heterozygous sites ($R^2$ = 0.992, p < 2.2e−16). This correlation is mainly due to genetic drift because the alleles with lower initial frequency are more likely to be lost during the bottleneck.[61] Despite much lower heterozygosity overall, the total number of deleterious alleles per individual in Orkney voles was comparable to their conspecifics on the continent (t test; Orkney versus Western lineage p = 0.96; Figure 4B). Very similar patterns were observed when classifying the mutations according to different functional consequences (frameshift mutations, premature stop codons, start codon lost and missense mutations; Figure S3B).

## DISCUSSION

Our analyses show that the Orkney vole is likely one of the oldest cases of human introduced wild species on islands with complete genetic isolation for more than 5,000 years (Figures 2 and 3A). The early separation between Orkney islands makes these voles an ancient "experiment" with repeated trials on the evolution of isolated populations. The outcome of this exceptionally long-lasting experiment supports population genetics theory on the consequences of bottlenecks and founder events[8,10,62,63] beyond most time frames investigated in natural systems so far.[64,65] Remarkably, the demographic trajectories of the populations are largely decoupled from the depauperate genomic background of Orkney voles. Orkney voles have high modern $N_e$ as indicated by demographic reconstruction (Figure 3; Table S4) and they have experienced ecological success,[45] despite the ubiquitous genomic signs associated with potential negative fitness consequences that still persist after thousands of generations in isolation.

Please cite this article in press as: Wang et al., Demographic history and genomic consequences of 10,000 generations of isolation in a wild mammal, Current Biology (2023), https://doi.org/10.1016/j.cub.2023.04.042





## Multiple replicates of long-lasting isolation induced by humans

Genetic isolation of Orkney voles for over 10,000 generations after human introduction is exceptional among the studied systems of vertebrates. The most similar example is the island fox (*Urocyon littoralis*), which started colonizing the California Channel Islands about 9,000 years ago (9,000 generations for fox), and for which the later close association with humans contributed to their spread and recent gene flow among populations.[64] Other completely isolated vertebrate populations with relatively clear history have been separated from conspecific populations for hundreds of generations (e.g., Italian brown bears[66]) and, rarely, a few thousand generations (e.g., Soay sheep[29]). Except for domestic animals like sheep,[29] there was either no human transfer involved in establishing these completely isolated populations or it remains unknown. Overall, human introductions of vertebrates apart from domestic animals are very numerous but they concern mostly comparatively very recent cases that led to biological invasions.[6,67,68]

The strongest bottleneck in the history of Orkney voles ca. 5,300 years ago probably represents the establishment of founding populations[46] with increasing human occupation at a time when most Orkney islands were already separated by the North Atlantic Ocean.[48,49] This time point is consistent with the oldest vole remains on Mainland dated 4,800 years old, while the earliest archeological evidence of voles on Orkney dated to 5,600 years ago stems from the island Papa Westray in the north of Westray.[38] The estimated split between Orkney Mainland and the continental population ~8,400 years ago coincides with the earliest evidence for Mesolithic humans in Orkney,[69] but this time estimate may have been pushed earlier by population structure within the Western lineage (e.g., shown in Heckel et al.[40]). A genetic replacement event in the Western-north region[38] occurred probably due to land-use changes on the continent within the last two millenia[38]—much after the separation of Orkney populations. The signal of introgression from the Western-south group to the ancestor of the Orkney voles (Figures 2B and S2) is thus likely caused by shared ancestral variation in the Western lineage that is still partially preserved in Orkney populations.

The demographic history of Orkney voles was tightly linked to human activities. Early archeological records from multiple islands[46] show that voles have spread fast across the Orkney archipelago given suitable ecological conditions and likely with the help of Neolithic humans. The signals of gene flow between islands (Figures 2B and S2) probably relate mostly to this relatively early phase of colonization. Our divergence times estimates for extant Orkney populations will thus not necessarily reflect the first waves of vole introduction, but rather the order in which island populations were split off. Human transport of voles around the Orkney system may have continued for centuries after the initial colonization and our divergence time estimates are likely to also reflect such gene flow.[70]

The absence of voles from Papa Westray in modern times demonstrates that some of the initial island populations have vanished. Human actions have affected other Orkney islands like Shapinsay[71] and Eday (M. Cockram, personal communication) where voles went extinct in the early 20$^{th}$ century.[46] It is unclear what the causes of these extinctions were and whether mostly ecological change or also genetically caused fitness decline ("mutational meltdown")[72] was involved. The phase of low $N_e$ on all Orkney islands around 500 to 200 years ago is compatible with negative effects of agricultural practices that reduced vole habitats by using burned kelp and hay as fertilizer.[73] The Orkney vole system provides a unique opportunity to distinguish between the scenarios of repeated ancient introductions versus continued gene flow. In the future, the concrete timing of events can be provided by combining more genomic information from modern vole populations with time series analyses of ancient DNA based on the rich vole bone material recovered during archeological excavations.[38,46,47]

## Extreme accumulation of deleterious mutations

The bottleneck related to the founding event of Orkney voles has left the shared genomic legacy of a large proportion of fixed derived sites that include many potentially highly deleterious alleles. Homozygosity of loss-of-function mutations (Figure 4) has reached an extremely high level in Orkney voles compared with populations of other mammals that have gone through much shorter periods of isolation.[27,28,57] Our individual-based simulations suggest that the fixation of many deleterious mutations in Orkney voles likely happened shortly after the bottleneck started (Figure S5C). Fixation at the founding stage of Orkney populations can also explain the sharing of most fixed deleterious mutations among successively colonized islands (Figure S4). It is noteworthy that these genomic patterns resulting from relatively few consecutive founding events are largely consistent with the patterns predicted by theoretical models of population contraction[74] and of extensive range expansions in continuous or discrete space,[20,21] and also with empirical patterns in human populations.[22] This similarity may stem from the fact that evolutionary forces at the front of expanding populations can have effects that are very similar to that of a single bottleneck.[75] The amount of drift during a spatial expansion comprising serial founder events is determined by the harmonic mean of population sizes at the expansion front.[76] Thus, a single bottleneck and a spatial expansion with the same harmonic mean of population size and the same duration will yield very similar genomic signatures as the spatial expansion. Relatively large population sizes per island may have contributed to maintaining the high frequency of derived deleterious alleles over the extended period in isolation.[77]

The relatively large effective population size of Orkney voles even during the initial bottleneck may also explain why we found no strong signal of reduction in the number of deleterious alleles in contrast to e.g., Mountain gorilla,[57] Bengal tiger,[78] or reintroduced Alpine ibex,[27] large mammals typically with population sizes of dozens to hundreds. Genetic purging in the narrow sense as a consequence of inbreeding is expected to decrease the number of deleterious alleles[79] and a few dozens of generations can lead to detectable signatures in genomes.[27] This process may be subtle in the Orkney vole system given even large effective population sizes per island compared with the mentioned big mammals (Figure 3). Indeed, the amount of genetic drift and inbreeding is determined by the harmonic mean of population sizes over time,[80] and the harmonic mean gives most weight to the lowest population sizes. Populations of voles with high growth rates can rapidly regain large sizes after

Please cite this article in press as: Wang et al., Demographic history and genomic consequences of 10,000 generations of isolation in a wild mammal, Current Biology (2023), https://doi.org/10.1016/j.cub.2023.04.042



introduction given suitable environments, and thus experience far less drift and fixation through bottlenecks than those of large mammals with much smaller growth rates. Furthermore, distant bottlenecks have a smaller impact on current genomic diversity, inbreeding and mutation load as compared with more recent ones (see Table 1 in Robinson et al.[81]). For these demographic reasons, we would not expect strong signals of purging in Orkney vole populations in contrast to the big mammal species that experienced prolonged and recent bottlenecks at much lower population sizes.

However, even the Eday population, which started from only 15 individuals in 1987 and whose ancestors have gone through multiple bottlenecks in the last 5,000 years, did not show a noticeable reduction of the total number of deleterious alleles (Figure 4; Table 1). Systematically collected data on its expansion are not available but the 27 $km^2$ island was already largely occupied by voles 4 years after introduction (M. Cockram, personal communication). This suggests that the current mutation load effectively leads to only minor reduction in the absolute fitness of Orkney voles and that selection has at least in part been soft in the Orkney archipelago. Deleterious variation with the strongest effects on survival or reproduction may have been purged already thousands of generations ago at the initial stages of Orkney colonization[13] as shown in our simulations (Figure S5). However, extinction of a few island populations mentioned above circumstantially suggests that hard selection from environmental or intrinsic genetic pressure may have had a large effect on certain populations.

Several potential explanations exist for the demographic success of Orkney voles despite the relatively large amount of inferred deleterious mutations. Reduced inter-specific competition could contribute to predominantly soft selection on Orkney, such that fitness is mostly affected by intra-specific competition.[82] Furthermore, selection could have been relaxed overall due to generally benign environmental conditions on Orkney (e.g., stable climatic conditions, lower predation pressure, lower pathogen burden) compared with continental Europe,[46,83] resulting in a DFE that is shifted toward more neutral variants. Selective sweeps of beneficial mutations that compensate effects of deleterious mutations might also play a role, but these are very difficult to identify after long bottlenecks.[74,75]

The realized fitness effects of the deleterious alleles in the present-day Orkney populations remain unknown. There is no information on current demographic trajectories for any of the Orkney islands, and published data on survival rates or other fitness relevant parameters in natural populations are absent. Our own field data suggest a lower reproductive rate in Orkney populations (mean number of embryos: 4.4; N = 80 litters) than in the Western lineage (mean: 5.5[84]), but Orkney voles have probably a higher early survival rate and longer life expectancy based on data from captivity.[85,86] Given the large recent effective population size, it is possible that some of the Orkney populations carry a high inbreeding load.[87] In order to improve our knowledge about the realized or masked effects of deleterious variation, it would be interesting to expose and quantify it at the individual level, for example, through crossing experiments with voles with high or low genetic load. Experimental comparisons between Orkney and continental voles could clarify the extent of differences in absolute and relative fitness between these populations. Field experiments in different environments might elucidate particularly components of hard and soft selection affecting populations and individuals differently (e.g., Barrett et al.[88]). Similar experimental approaches with Orkney voles could also shed light on how demographic events, such as prolonged periods at reduced population sizes in isolated populations, can set the stage for the occurrence of heterosis and POD after secondary contact.[60]

## The success of an ancient off-site introduction

The Orkney vole system can be regarded as an ancient off-site introduction and may thus enable a comparative glimpse into the potential future of isolated vertebrate populations of concern. In conservation biology, the need is rising to better understand the relationship between fitness and genome-wide genetic variation.[87,89] Genetic diversity is the hallmark of evolutionary potential of the populations, and reduced diversity and increased homozygosity of deleterious mutations for extended periods of time are expected to lead to increased risk of extinction of isolated populations.[90] However, many Orkney vole populations have prevailed for thousands of generations. One of the beneficial factors may be that the initial effective sizes were not extremely small compared with other systems mentioned above and may have left sufficient functionally relevant diversity to cope with the environment.

An important reason for the persistence of Orkney voles may consist in the life history of the common vole. Unlike the big mammals that have been mostly studied,[26,27,57,91] the common vole is a typical *r*-strategist with short life expectancy and high reproductivity, and extensive fluctuations as part of its normal demography of populations.[92,93] Recent work suggested that species with shorter life spans tend to bear more deleterious mutations,[17] possibly driven by insufficient selection on the genes involved in the late stage of life, or due to genetic drift from repeated bottlenecks.[94] Moreover, organisms with high reproductive rates are less likely affected by severe fitness loss due to inbreeding depression than slowly reproducing species because of more potential for variation among offspring.[95] Yet, more theoretical and empirical studies are necessary to disentangle the evolutionary connection between life history traits and mutation load. Thus, it would be important to consider for off-site introductions, e.g., of species of conservation concern, not only general life history traits but also information about historical levels of deleterious variation in genomes.[96,97] Although being of no conservation concern itself, the Orkney vole offers the opportunity to examine in more detail the relative importance of adaptive versus chance events for the persistence of isolated populations in the future.

## STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS

## SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.cub.2023.04.042.

## ACKNOWLEDGMENTS

We thank Julian Branscombe, John Crossley, Keith Dobney, and Sydney Gauld for logistical support, and the following landowners and institutions for access for fieldwork: C. and N. Bichan, M. and D. Cockram, S. and A. Gauld, A. and D. Thomson, S. and L. Hagan, J. and S. Crossley, B. and L. Hamill, B. Turnbull and the Westray Golf Club, E. Spence, R. and A. Manson, and the Orkney Islands council. We thank Susanne Tellenbach, Melanie Hiltbrunner, Joana Rechsteiner, and Guy Schnidrig for help with fieldwork in Orkney, and S. Braaker, M. Fischer, E. Kindler, B. Gauffre, A. Labutin, N. Bulatova, L. Yalkovskaya, M. Ratkiewicz, and C. Polat for assistance with sample collection. We thank the Next Generation Sequencing Platform of the University of Bern and Pamela Nicholson for excellent sequencing services. Alexandre Thiéry and Alexandre Gouy provided support for bioinformatics. We thank Nina Marchi, Manuel Schweizer, and David Marques for their comments on the manuscript. We also thank Donald Waller, Marty Kardos, and one anonymous reviewer for helpful comments and suggestions. Computation was performed in part on UBELIX (http://www.id.unibe.ch/hpc), the HPC cluster at the University of Bern. X.W. received a scholarship for PhD study from the China Scholarship Council (no. 201706380049). This study was supported by Swiss National Science Foundation grant 31003A_176209 to G.H.

## AUTHOR CONTRIBUTIONS

G.H. conceptualized the study; G.H. designed research; X.W. and G.H. performed research; X.W. analyzed genomic data; S.P. performed evolutionary simulations; X.W., S.P., and G.H. wrote the paper.

## DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: September 28, 2022
Revised: December 20, 2022
Accepted: April 17, 2023
Published: May 12, 2023

## REFERENCES

1. Kruess, A., and Tscharntke, T. (1994). Habitat fragmentation, species loss, and biological control. Science *264*, 1581–1584.
2. Mayr, E. (1963). Animal Species and Evolution*797* (Belknap Press of Harvard University Press).
3. Schipper, J., Chanson, J.S., Chiozza, F., Cox, N.A., Hoffmann, M., Katariya, V., Lamoreux, J., Rodrigues, A.S., Stuart, S.N., and Temple, H.J. (2008). The status of the world's land and marine mammals: diversity, threat, and knowledge. Science *322*, 225–230.
4. Crooks, K.R., Burdett, C.L., Theobald, D.M., King, S.R.B., Di Marco, M., Rondinini, C., and Boitani, L. (2017). Quantification of habitat fragmentation reveals extinction risk in terrestrial mammals. Proc. Natl. Acad. Sci. USA *114*, 7635–7640.
5. Pringle, R.M., Kartzinel, T.R., Palmer, T.M., Thurman, T.J., Fox-Dobbs, K., Xu, C.C.Y., Hutchinson, M.C., Coverdale, T.C., Daskin, J.H., Evangelista, D.A., et al. (2019). Predator-induced collapse of niche structure and species coexistence. Nature *570*, 58–64.
6. Luque, G.M., Bellard, C., Bertelsmeier, C., Bonnaud, E., Genovesi, P., Simberloff, D., and Courchamp, F. (2014). The 100th of the world's worst invasive alien species. Biol. Invas. *16*, 981–985.
7. Colautti, R.I., Alexander, J.M., Dlugosch, K.M., Keller, S.R., and Sultan, S.E. (2017). Invasions and extinctions through the looking glass of evolutionary ecology. Philos. Trans. R. Soc. Lond. B Biol. Sci. *372*, 20160031.
8. Nei, M., Maruyama, T., and Chakraborty, R. (1975). The bottleneck effect and genetic variability in populations. Evolution *29*, 1–10.
9. Simons, Y.B., Turchin, M.C., Pritchard, J.K., and Sella, G. (2014). The deleterious mutation load is insensitive to recent population history. Nat. Genet. *46*, 220–224.
10. Kimura, M., Maruyama, T., and Crow, J.F. (1963). The mutation load in small populations. Genetics *48*, 1303–1312.
11. González-Martínez, S.C., Ridout, K., and Pannell, J.R. (2017). Range expansion compromises adaptive evolution in an outcrossing plant. Curr. Biol. *27*, 2544–2551.e4.
12. Perrier, A., Sánchez-Castro, D., and Willi, Y. (2020). Expressed mutational load increases toward the edge of a species' geographic range. Evolution *74*, 1711–1723.
13. Bertorelle, G., Raffini, F., Bosse, M., Bortoluzzi, C., Iannucci, A., Trucchi, E., Morales, H.E., and Van Oosterhout, C. (2022). Genetic load: genomic estimates and applications in non-model animals. Nat. Rev. Genet. *23*, 1.
14. Charlesworth, D., and Charlesworth, B. (1987). Inbreeding depression and its evolutionary consequences. Annu. Rev. Ecol. Syst. *18*, 237–268.
15. Lynch, M., Conery, J., and Burger, R. (1995). Mutation accumulation and the extinction of small populations. Am. Nat. *146*, 489–518.
16. Crnokrak, P., and Barrett, S.C. (2002). Perspective: purging the genetic load: a review of the experimental evidence. Evolution *56*, 2347–2358.
17. . Preprint at van der Valk, T., de Manuel, M., Marques-Bonet, T., and Guschanski, K. (2021). Estimates of genetic load suggest frequent purging of deleterious alleles in small populations bioRxiv. https://doi.org/10.1101/696831.
18. Kirkpatrick, M., and Jarne, P. (2000). The effects of a bottleneck on inbreeding depression and the genetic load. Am. Nat. *155*, 154–167.
19. Aris-Brosou, S. (2019). Direct evidence of an increasing mutational load in humans. Mol. Biol. Evol. *36*, 2823–2829.
20. Peischl, S., and Excoffier, L. (2015). Expansion load: recessive mutations and the role of standing genetic variation. Mol. Ecol. *24*, 2084–2094.
21. Peischl, S., Dupanloup, I., Bosshard, L., and Excoffier, L. (2016). Genetic surfing in human populations: from genes to genomes. Curr. Opin. Genet. Dev. *41*, 53–61.
22. Peischl, S., Dupanloup, I., Foucal, A., Jomphe, M., Bruat, V., Grenier, J.C., Gouy, A., Gilbert, K.J., Gbeha, E., Bosshard, L., et al. (2018). Relaxed selection during a recent human expansion. Genetics *208*, 763–777.
23. Mallet, M.A., and Chippindale, A.K. (2011). Inbreeding reveals stronger net selection on *Drosophila melanogaster* males: implications for mutation load and the fitness of sexual females. Heredity *106*, 994–1002.
24. Jiménez, J.A., Hughes, K.A., Alaks, G., Graham, L., and Lacy, R.C. (1994). An experimental study of inbreeding depression in a natural habitat. Science *266*, 271–273.
25. Meagher, S., Penn, D.J., and Potts, W.K. (2000). Male–male competition magnifies inbreeding depression in wild house mice. Proc. Natl. Acad. Sci. USA *97*, 3324–3329.

26. Robinson, J.A., Räikkönen, J., Vucetich, L.M., Vucetich, J.A., Peterson, R.O., Lohmueller, K.E., and Wayne, R.K. (2019). Genomic signatures of extensive inbreeding in Isle Royale wolves, a population on the threshold of extinction. Sci. Adv. *5*, eaau0757.

27. Grossen, C., Guillaume, F., Keller, L.F., and Croll, D. (2020). Purging of highly deleterious mutations through severe bottlenecks in Alpine ibex. Nat. Commun. *11*, 1001.

28. Robinson, J.A., Brown, C., Kim, B.Y., Lohmueller, K.E., and Wayne, R.K. (2018). Purging of strongly deleterious mutations explains long-term persistence and absence of inbreeding depression in island foxes. Curr. Biol. *28*, 3487–3494.e4.

29. Stoffel, M.A., Johnston, S.E., Pilkington, J.G., and Pemberton, J.M. (2021). Genetic architecture and lifetime dynamics of inbreeding depression in a wild mammal. Nat. Commun. *12*, 2972.

30. Huisman, J., Kruuk, L.E.B., Ellis, P.A., Clutton-Brock, T., and Pemberton, J.M. (2016). Inbreeding depression across the lifespan in a wild mammal population. Proc. Natl. Acad. Sci. USA *113*, 3585–3590.

31. Whittaker, R.J., and Fernández-Palacios, J.M. (2007). Island Biogeography: Ecology, Evolution, and Conservation (Oxford University Press).

32. Latter, B.D. (1973). The island model of population differentiation: a general solution. Genetics *73*, 147–157.

33. Richards, C. (1996). Monuments as landscape: creating the centre of the world in late Neolithic Orkney. World Archaeol. *28*, 190–208.

34. Stanton, D.W.G., Mulville, J.A., and Bruford, M.W. (2016). Colonization of the Scottish Islands via long-distance Neolithic transport of red deer (*Cervus elaphus*). Proc. Biol. Sci. *283*, 20160095.

35. Chevret, P., Hautier, L., Ganem, G., Herman, J., Agret, S., Auffray, J.C., and Renaud, S. (2021). Genetic structure in Orkney island mice: isolation promotes morphological diversification. Heredity *126*, 266–278.

36. Millais, J. (1904). On a new British vole from the Orkney Islands. Zoologist *8*, 241–246.

37. Haynes, S., Jaarola, M., and Searle, J.B. (2003). Phylogeography of the common vole (*Microtus arvalis*) with particular emphasis on the colonization of the Orkney archipelago. Mol. Ecol. *12*, 951–956.

38. Martínková, N., Barnett, R., Cucchi, T., Struchen, R., Pascal, M., Pascal, M., Fischer, M.C., Higham, T., Brace, S., Ho, S.Y.W., et al. (2013). Divergent evolutionary processes associated with colonization of offshore islands. Mol. Ecol. *22*, 5205–5220.

39. Lambin, X., Bretagnolle, V., and Yoccoz, N.G. (2006). Vole population cycles in northern and southern Europe: is there a need for different explanations for single pattern? J. Anim. Ecol. *75*, 340–349.

40. Heckel, G., Burri, R., Fink, S., Desmet, J.-F., and Excoffier, L. (2005). Genetic structure and colonization processes in European populations of the common vole, *Microtus arvalis*. Evolution *59*, 2231–2242.

41. Lischer, H.E.L., Excoffier, L., and Heckel, G. (2014). Ignoring heterozygous sites biases phylogenomic estimates of divergence times: implications for the evolutionary history of *Microtus* voles. Mol. Biol. Evol. *31*, 817–831.

42. Beysard, M., and Heckel, G. (2014). Structure and dynamics of hybrid zones at different stages of speciation in the common vole (*Microtus arvalis*). Mol. Ecol. *23*, 673–687.

43. Yigit, N., Hutterer, R., Kryštufek, B., and Amori, G.; The IUCN red list threatened species (2016). *Microtus arvalis*. e.T13488A22351133. https://doi.org/10.2305/IUCN.UK.2016-2.RLTS.T13488A22351133.en.

44. Romaniuk, A.A., Shepherd, A.N., Clarke, D.V., Sheridan, A.J., Fraser, S., Bartosiewicz, L., and Herman, J.S. (2016). Rodents: food or pests in Neolithic Orkney. R. Soc. Open Sci. *3*, 160514.

45. Reynolds, P. (1992). The Impact of Changes in Land-Use in Orkney, on the Vole *Microtus arvalis orcadensis* and Its Avian Predators.. PhD thesis (University of Aberdeen).

46. Cucchi, T., Barnett, R., Martínková, N., Renaud, S., Renvoisé, E., Evin, A., Sheridan, A., Mainland, I., Wickham-Jones, C., Tougard, C., et al. (2014). The changing pace of insular life: 5000 years of microevolution in the Orkney vole (*Microtus arvalis orcadensis*). Evolution *68*, 2804–2820.

47. Fraser, S.M. (2015). Mammals in Late Neolithic Orkney (with reference to mammal bone recovered from Links of Noltland, Westray). PhD thesis (The University of Edinburgh).

48. Phillips, T. (2004). Seascapes and landscapes in Orkney and northern Scotland. World Archaeol. *35*, 371–384.

49. Smith, D.E., Barlow, N.L.M., Bradley, S.L., Firth, C.R., Hall, A.M., Jordan, J.T., and Long, D. (2019). Quaternary sea level change in Scotland. Earth Environ. Sci. Trans. R. Soc. Edinb. *110*, 219–256.

50. Braaker, S., and Heckel, G. (2009). Transalpine colonisation and partial phylogeographic erosion by dispersal in the common vole (*Microtus arvalis*). Mol. Ecol. *18*, 2518–2531.

51. Saxenhofer, M., Schmidt, S., Ulrich, R.G., and Heckel, G. (2019). Secondary contact between diverged host lineages entails ecological speciation in a European hantavirus. PLoS Biol. *17*, e3000142.

52. Saxenhofer, M., Labutin, A., White, T.A., and Heckel, G. (2022). Host genetic factors associated with the range limit of a European hantavirus. Mol. Ecol. *31*, 252–265.

53. Malinsky, M., Svardal, H., Tyers, A.M., Miska, E.A., Genner, M.J., Turner, G.F., and Durbin, R. (2018). Whole-genome sequences of Malawi cichlids reveal multiple radiations interconnected by gene flow. Nat. Ecol. Evol. *2*, 1940–1955.

54. Terhorst, J., Kamm, J.A., and Song, Y.S. (2017). Robust and scalable inference of population history from hundreds of unphased whole genomes. Nat. Genet. *49*, 303–309.

55. Patton, A.H., Margres, M.J., Stahlke, A.R., Hendricks, S., Lewallen, K., Hamede, R.K., Ruiz-Aravena, M., Ryder, O., McCallum, H.I., Jones, M.E., et al. (2019). Contemporary demographic reconstruction methods are robust to genome assembly quality: A case study in Tasmanian devils. Mol. Biol. Evol. *36*, 2906–2921.

56. Agrawal, A.F., and Whitlock, M.C. (2011). Inferences about the distribution of dominance drawn from yeast gene knockout data. Genetics *187*, 553–566.

57. Xue, Y., Prado-Martinez, J., Sudmant, P.H., Narasimhan, V., Ayub, Q., Szpak, M., Frandsen, P., Chen, Y., Yngvadottir, B., Cooper, D.N., et al. (2015). Mountain gorilla genomes reveal the impact of long-term population decline and inbreeding. Science *348*, 242–245.

58. Do, R., Balick, D., Li, H., Adzhubei, I., Sunyaev, S., and Reich, D. (2015). No evidence that selection has been less effective at removing deleterious mutations in Europeans than in Africans. Nat. Genet. *47*, 126–131.

59. Gilbert, K.J., Pouyet, F., Excoffier, L., and Peischl, S. (2020). Transition from background selection to associative overdominance promotes diversity in regions of low recombination. Curr. Biol. *30*, 101–107.e3.

60. Abu-Awad, D., and Waller, D. (2023). Conditions for maintaining and eroding pseudo-overdominance and its contribution to inbreeding depression. Peer Community J. *3*, e8.

61. Kimura, M. (1962). On the probability of fixation of mutant genes in a population. Genetics *47*, 713–719.

62. Slatkin, M., and Excoffier, L. (2012). Serial founder effects during range expansion: A spatial analog of genetic drift. Genetics *191*, 171–181.

63. Deshpande, O., Batzoglou, S., Feldman, M.W., and Cavalli-Sforza, L.L. (2009). A serial founder effect model for human settlement out of Africa. Proc. Biol. Sci. *276*, 291–300.

64. Funk, W.C., Lovich, R.E., Hohenlohe, P.A., Hofman, C.A., Morrison, S.A., Sillett, T.S., Ghalambor, C.K., Maldonado, J.E., Rick, T.C., Day, M.D., et al. (2016). Adaptive divergence despite strong genetic drift: genomic analysis of the evolutionary mechanisms causing genetic differentiation in the island fox (*Urocyon littoralis*). Mol. Ecol. *25*, 2176–2194.

65. Martin, C.H., and Höhna, S. (2018). New evidence for the recent divergence of Devil's Hole pupfish and the plausibility of elevated mutation rates in endangered taxa. Mol. Ecol. *27*, 831–838.

66. Benazzo, A., Trucchi, E., Cahill, J.A., Maisano Delser, P.M., Mona, S., Fumagalli, M., Bunnefeld, L., Cornetti, L., Ghirotto, S., and Girardi, M.

(2017). Survival and divergence in a small group: the extraordinary genomic history of the endangered Apennine brown bear stragglers. Proc. Natl. Acad. Sci. USA *114*, E9589–E9597.

67. Iannella, A., Peacock, D., Cassey, P., and Schwensow, N. (2019). Genetic perspectives on the historical introduction of the European rabbit (*Oryctolagus cuniculus*) to Australia. Biol. Invas. *21*, 603–614.
68. Puckett, E.E., Magnussen, E., Khlyap, L.A., Strand, T.M., Lundkvist, Å., and Munshi-South, J. (2020). Genomic analyses reveal three independent introductions of the invasive brown rat (*Rattus norvegicus*) to the Faroe Islands. Heredity *124*, 15–27.
69. Farrell, M., Bunting, M.J., Lee, D.H.J., and Thomas, A. (2014). Neolithic settlement at the woodland's edge: palynological data and timber architecture in Orkney, Scotland. J. Archaeol. Sci. *51*, 225–236.
70. Leaché, A.D., Harris, R.B., Rannala, B., and Yang, Z. (2014). The influence of gene flow on species tree estimation: a simulation study. Syst. Biol. *63*, 17–30.
71. Miller, G.S. (1912). Catalogue of the Mammals of Western Europe (Europe Exclusive of Russia): in the Collection of the British Museum (Trustees of the British Museum).
72. Lynch, M., Bürger, R., Butcher, D., and Gabriel, W. (1993). The mutational meltdown in asexual populations. J. Hered. *84*, 339–344.
73. Fenton, A. (1997). The Northern Isles Orkney and Shetland (Dundurn).
74. Moinet, A., Schlichta, F., Peischl, S., and Excoffier, L. (2022). Strong neutral sweeps occurring during a population contraction. Genetics *220*, iyac021.
75. Schlichta, F., Peischl, S., and Excoffier, L. (2022). The impact of genetic surfing on neutral genomic diversity. Mol. Biol. Evol. *39*, msac249.
76. Wakeley, J. (2009). Coalescent Theory (Roberts & Company).
77. Marsden, C.D., Ortega-Del Vecchyo, D., O'Brien, D.P., Taylor, J.F., Ramirez, O., Vilà, C., Marques-Bonet, T., Schnabel, R.D., Wayne, R.K., and Lohmueller, K.E. (2016). Bottlenecks and selective sweeps during domestication have increased deleterious genetic variation in dogs. Proc. Natl. Acad. Sci. USA *113*, 152–157.
78. Khan, A., Patel, K., Shukla, H., Viswanathan, A., van der Valk, T., Borthakur, U., Nigam, P., Zachariah, A., Jhala, Y.V., Kardos, M., et al. (2021). Genomic evidence for inbreeding depression and purging of deleterious genetic variation in Indian tigers. Proc. Natl. Acad. Sci. USA *118*, e2023018118.
79. García-Dorado, A. (2012). Understanding and predicting the fitness decline of shrunk populations: inbreeding, purging, mutation, and standard selection. Genetics *190*, 1461–1476.
80. Crow, J.F., and Kimura, M. (1970). An Introduction to Population Genetics Theory (Burgess Publishing Company).
81. Robinson, J., Kyriazis, C.C., Yuan, S.C., and Lohmueller, K.E. (2023). Deleterious variation in natural populations and implications for conservation genetics. Annu. Rev. Anim. Biosci. *11*, 93–114.
82. Agrawal, A.F., and Whitlock, M.C. (2012). Mutation load: the fitness of individuals in populations where deleterious alleles are abundant. Annu. Rev. Ecol. Evol. Syst. *43*, 115–135.
83. Adler, G.H., and Levins, R. (1994). The island syndrome in rodent populations. Q. Rev. Biol. *69*, 473–490.
84. Spitz, F. (1974). Démographie du campagnol des champs *Microtus arvalis* en Vendée. Ann. Zool. Écologie Anim. *6*, 259–312.
85. Leslie, P.H., Tener, J.S., Vizoso, M., and Chitty, H. (1955). The longevity and fertility of the Orkney vole, *Microtus orcadensis*, as observed in the laboratory. Proc. Zool. Soc. Lond. *125*, 115–125.
86. Daketse, M.-J., and Martinet, L. (1977). Effect of temperature on the growth and fertility of the field-vole, *Microtus arvalis*, raised in different daylength and feeding conditions. Ann. Biol. anim. Bioch. Biophys. *17*, 713–721.
87. Kardos, M., Armstrong, E.E., Fitzpatrick, S.W., Hauser, S., Hedrick, P.W., Miller, J.M., Tallmon, D.A., and Funk, W.C. (2021). The crucial role of genome-wide genetic variation in conservation. Proc. Natl. Acad. Sci. USA *118*, e2104642118.
88. Barrett, R.D.H., Laurent, S., Mallarino, R., Pfeifer, S.P., Xu, C.C.Y., Foll, M., Wakamatsu, K., Duke-Cohan, J.S., Jensen, J.D., and Hoekstra, H.E. (2019). Linking a mutation to survival in wild mice. Science *363*, 499–504.
89. DeWoody, J.A., Harder, A.M., Mathur, S., and Willoughby, J.R. (2021). The long-standing significance of genetic diversity in conservation. Mol. Ecol. *30*, 4147–4154.
90. Allendorf, F.W., Hohenlohe, P.A., and Luikart, G. (2010). Genomics and the future of conservation genetics. Nat. Rev. Genet. *11*, 697–709.
91. Kleinman-Ruiz, D., Lucena-Perez, M., Villanueva, B., Fernández, J., Saveljev, A.P., Ratkiewicz, M., Schmidt, K., Galtier, N., García-Dorado, A., and Godoy, J.A. (2022). Purging of deleterious burden in the endangered Iberian lynx. Proc. Natl. Acad. Sci. USA *119*, e2110614119.
92. Boyce, C.C.K., and Boyce, J.L., III. (1988). Population biology of *Microtus arvalis*. I. Lifetime reproductive success of solitary and grouped breeding females. J. Anim. Ecol. *57*, 711–722.
93. Tkadlec, E., and Zejda, J. (1995). Precocious breeding in female common voles and its relevance to rodent fluctuations. Oikos *73*, 231–236.
94. Cui, R., Medeiros, T., Willemsen, D., Iasi, L.N.M., Collier, G.E., Graef, M., Reichard, M., and Valenzano, D.R. (2019). Relaxed selection limits lifespan by increasing mutation load. Cell *178*, 385–399.e20.
95. Hedrick, P.W., Hellsten, U., and Grattapaglia, D. (2016). Examining the cause of high inbreeding depression: analysis of whole-genome sequence data in 28 selfed progeny of *Eucalyptus grandis*. New Phytol. *209*, 600–611.
96. Hohenlohe, P.A., Funk, W.C., and Rajora, O.P. (2021). Population genomics for wildlife conservation and management. Mol. Ecol. *30*, 62–82.
97. Robinson, J.A., Kyriazis, C.C., Nigenda-Morales, S.F., Beichman, A.C., Rojas-Bracho, L., Robertson, K.M., Fontaine, M.C., Wayne, R.K., Lohmueller, K.E., Taylor, B.L., et al. (2022). The critically endangered vaquita is not doomed to extinction by inbreeding depression. Science *376*, 635–639.
98. Lunter, G., and Goodson, M. (2011). Stampy: a statistical algorithm for sensitive and fast mapping of Illumina sequence reads. Genome Res. *21*, 936–939.
99. DePristo, M.A., Banks, E., Poplin, R., Garimella, K.V., Maguire, J.R., Hartl, C., Philippakis, A.A., del Angel, G., Rivas, M.A., Hanna, M., et al. (2011). A framework for variation discovery and genotyping using next-generation DNA sequencing data. Nat. Genet. *43*, 491–498.
100. Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows–Wheeler transform. Bioinformatics *25*, 1754–1760.
101. Danecek, P., and McCarthy, S.A. (2017). BCFtools/csq: haplotype-aware variant consequences. Bioinformatics *33*, 2037–2039.
102. Ronquist, F., Teslenko, M., Van Der Mark, P., Ayres, D.L., Darling, A., Höhna, S., Larget, B., Liu, L., Suchard, M.A., and Huelsenbeck, J.P. (2012). MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. Syst. Biol. *61*, 539–542.
103. Darriba, D., Taboada, G.L., Doallo, R., and Posada, D. (2012). JModelTest 2: more models, new heuristics and parallel computing. Nat. Methods *9*, 772.
104. Martin, M., Patterson, M., Garg, S., Fischer, O.S., Pisanti, N., Klau, G.W., Schöenhuth, A., and Marschall, T. (2016). WhatsHap: fast and accurate read-based phasing. Preprint at bioRxiv. https://doi.org/10.1101/085050.
105. Delaneau, O., Zagury, J.-F., Robinson, M.R., Marchini, J.L., and Dermitzakis, E.T. (2019). Accurate, scalable and integrative haplotype estimation. Nat. Commun. *10*, 5436.
106. Hermann, P., Heissl, A., Tiemann-Boege, I., and Futschik, A. (2019). *LDJump*: estimating variable recombination rates from population genetic data. Mol. Ecol. Resour. *19*, 623–638.

107. Zhou, Y., Browning, S.R., and Browning, B.L. (2020). A fast and simple method for detecting identity-by-descent segments in large-scale data. Am. J. Hum. Genet. *106*, 426–437.

108. Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M.A., Bender, D., Maller, J., Sklar, P., De Bakker, P.I., and Daly, M.J. (2007). PLINK: a tool set for whole-genome association and population-based linkage analyses. Am. J. Hum. Genet. *81*, 559–575.

109. Minh, B.Q., Schmidt, H.A., Chernomor, O., Schrempf, D., Woodhams, M.D., von Haeseler, A., and Lanfear, R. (2020). IQ-TREE 2: New models and efficient methods for phylogenetic inference in the genomic era. Mol. Biol. Evol. *37*, 1530–1534.

110. Pickrell, J.K., and Pritchard, J.K. (2012). Inference of population splits and mixtures from genome-wide allele frequency data. PLoS Genet. *8*, e1002967.

111. Malinsky, M., Matschiner, M., and Svardal, H. (2021). Dsuite – fast *D*-statistics and related admixture evidence from VCF files. Mol. Ecol. Resour. *21*, 584–595.

112. Cingolani, P., Platts, A., Wang, L.L., Coon, M., Nguyen, T., Wang, L., Land, S.J., Lu, X., and Ruden, D.M. (2012). A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w1118; iso-2; iso-3. Fly *6*, 80–92.

113. Raudvere, U., Kolberg, L., Kuzmin, I., Arak, T., Adler, P., Peterson, H., and Vilo, J. (2019). g:profiler: a web server for functional enrichment analysis and conversions of gene lists (2019 update). Nucleic Acids Res. *47*, W191–W198.

114. Haller, B.C., and Messer, P.W. (2019). SLiM 3: forward genetic simulations beyond the Wright–Fisher model. Mol. Biol. Evol. *36*, 632–637.

115. Fink, S., Excoffier, L., and Heckel, G. (2004). Mitochondrial gene diversity in the common vole *Microtus arvalis* shaped by historical divergence and local adaptations. Mol. Ecol. *13*, 3501–3514.

116. Baca, M., Popović, D., Lemanik, A., Bañuls-Cardona, S., Conard, N.J., Cuenca-Bescós, G., Desclaux, E., Fewlass, H., Garcia, J.T., Hadravova, T., et al. (2023). Ancient DNA reveals interstadials as a driver of common vole population dynamics during the last glacial period. J. Biogeogr. *50*, 183–196.

117. Poplin, R., Ruano-Rubio, V., DePristo, M.A., Fennell, T.J., Carneiro, M.O., Van der Auwera, G.A., Kling, D.E., Gauthier, L.D., Levy-Moonshine, A., and Roazen, D. (2017). Scaling accurate genetic variant discovery to tens of thousands of samples. Preprint at bioRxiv. https://doi.org/10.1101/201178.

118. Stumpf, M.P.H., and McVean, G.A.T. (2003). Estimating recombination rates from population-genetic data. Nat. Rev. Genet. *4*, 959–968.

119. Chapman, N.H., and Thompson, E.A. (2001). Linkage disequilibrium mapping: the role of population history, size, and structure. Adv. Genet. *42*, 413–437.

120. Jensen-Seaman, M.I., Furey, T.S., Payseur, B.A., Lu, Y., Roskin, K.M., Chen, C.F., Thomas, M.A., Haussler, D., and Jacob, H.J. (2004). Comparative recombination rates in the rat, mouse, and human genomes. Genome Res. *14*, 528–538.

121. Palamara, P.F., Francioli, L.C., Wilton, P.R., Genovese, G., Gusev, A., Finucane, H.K., Sankararaman, S., Genome of the Netherlands Consortium, Sunyaev, S.R., de Bakker, P.I., et al. (2015). Leveraging distant relatedness to quantify human mutation and gene-conversion rates. Am. J. Hum. Genet. *97*, 775–789.

122. Zheng, Z., Wang, X., Li, M., Li, Y., Yang, Z., Wang, X., Pan, X., Gong, M., Zhang, Y., Guo, Y., et al. (2020). The origin of domestication genes in goats. Sci. Adv. *6*, eaaz5216.

123. Boyko, A.R., Williamson, S.H., Indap, A.R., Degenhardt, J.D., Hernandez, R.D., Lohmueller, K.E., Adams, M.D., Schmidt, S., Sninsky, J.J., and Sunyaev, S.R. (2008). Assessing the evolutionary impact of amino acid mutations in the human genome. PLoS Genet. *4*, e1000083.

124. Kim, B.Y., Huber, C.D., and Lohmueller, K.E. (2017). Inference of the distribution of selection coefficients for new nonsynonymous mutations using large samples. Genetics *206*, 345–361.

125. Eyre-Walker, A., and Keightley, P.D. (2007). The distribution of fitness effects of new mutations. Nat. Rev. Genet. *8*, 610–618.

Please cite this article in press as: Wang et al., Demographic history and genomic consequences of 10,000 generations of isolation in a wild mammal, Current Biology (2023), https://doi.org/10.1016/j.cub.2023.04.042





## STAR★METHODS

### KEY RESOURCES TABLE

| REAGENT or RESOURCE | SOURCE | IDENTIFIER |
|---|---|---|
| Deposited data | | |
| Reference genome *Microtus arvalis* | NCBI website (https://www.ncbi.nlm.nih.gov/bioproject/737461) | PRJNA737461 |
| Mitochondrial genome *Microtus arvalis* | NCBI website (https://www.ncbi.nlm.nih.gov/nuccore/MG948434.1) | MG948434.1 |
| Gene annotation *Microtus arvalis* | A. Gouy, G.H., X.W. et al., unpublished data | N/A |
| Raw resequencing data | This study | PRJNA963218 (https://www.ncbi.nlm.nih.gov/sra/PRJNA963218) |
| Software and algorithms | | |
| Stampy 1.0.32 | Lunter and Goodson[98] | https://github.com/uwb-linux/stampy |
| GATK 4.1.6.0 | DePristo et al.[99] | https://gatk.broadinstitute.org |
| BWA 0.7.17 | Li et al.[100] | https://github.com/lh3/bwa |
| BCFtools 1.9 | Danecek and McCarthy[101] | https://samtools.github.io/bcftools |
| MrBayes 3.2.7 | Ronquist et al.[102] | https://nbisweden.github.io/MrBayes |
| JModelTest 2.7 | Darriba et al.[103] | https://github.com/ddarriba/jmodeltest2 |
| WhatsHap 1.0 | Martin et al.[104] | https://whatshap.readthedocs.io/en/latest |
| SHAPEIT4 4.1.2 | Delaneau et al.[105] | https://github.com/odelaneau/shapeit4 |
| R package LDJump | Hermann et al.[106] | https://github.com/PhHermann/LDJump |
| hap-IBD | Zhou et al.[107] | https://github.com/browning-lab/hap-ibd |
| PLINK 1.9 | Purcell et al.[108] | https://www.cog-genomics.org/plink |
| IQ-TREE 2.1.4 | Minh et al.[109] | http://www.iqtree.org |
| Treemix 1.13 | Pickrell et al.[110] | https://web.stanford.edu/group/pritchardlab/software.html |
| R package "optM" | https://cran.r-project.org/web/packages/OptM/index.html | |
| Dsuite 0.5 | Malinsky et al.[111] | https://github.com/millanek/Dsuite |
| SMC++ 1.15.2 | Terhorst et al.[54] | https://github.com/popgenmethods/smcpp |
| SnpEff 5.0 | Cingolani et al.[112] | https://pcingola.github.io/SnpEff/ |
| g:Profiler | Raudvere et al.[113] | https://biit.cs.ut.ee/gprofiler/ |
| SLiM 3.2 | Haller et al.[114] | https://www.slimframework.com/ |

### RESOURCE AVAILABILITY

#### Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Gerald Heckel (gerald.heckel@unibe.ch).

#### Materials availability

This study did not generate new unique reagents.

#### Data and code availability

Resequencing short-read Fastq files generated in this study have been deposited to SRA (BioProject ID: PRJNA963218). Code for data processing and analysis, and simulation used in this study have been deposited to Github (https://github.com/s-peischl/Wang_et_al_2023).

Please cite this article in press as: Wang et al., Demographic history and genomic consequences of 10,000 generations of isolation in a wild mammal, Current Biology (2023), https://doi.org/10.1016/j.cub.2023.04.042



## EXPERIMENTAL MODEL AND SUBJECT DETAILS

Tissue samples were obtained from individuals captured with snap traps and stored in absolute ethanol or as material stored at -20 °C. The samples were collected either for earlier studies Martínková et al.,[38] Beysard and Heckel,[42] Heckel et al.,[40] Fink et al.,[115] Baca et al.,[116] or in 2019.

## METHOD DETAILS

### Sampling and Sequencing

DNA was extracted using the phenol-chloroform method. DNA quality and concentration were checked with 1% agarose gels, Qubit fluorometer (Life Technologies) and NanoDrop 2000 Spectrophotometer (Thermo Fisher Scientific). Sequencing libraries were produced with Illumina TruSeq DNA PCR-Free Library Prep Kit and sequenced on Illumina Hiseq 2000 or Novaseq 6000 by the NGS platform of the University of Bern (2 x 150 cycles).

### Variant calling and filtering

In total, the genomes of 45 common voles covering most of the species' distribution range and three additional *Microtus obscurus* were sequenced (Figure 1A; Tables 1 and S1; see supplemental information for the methods). Raw reads were mapped to the reference genome of *M. arvalis* assembled at chromosome scale (BioProject ID: PRJNA737461, A. Gouy, G.H., X.W., et al., unpublished data) using Stampy 1.0.32[98] with default parameters except "—substitutionrate" set to 0.05. After mapping, duplicated reads were marked and removed using MarkDuplicates tool from GATK 4.1.6.0.[99] Variants were called individually with GATK HaplotypeCaller[117] using the GVCF pipeline.

After calling, SNPs and indels were filtered separately. For each individual, variants with read depth lower than 5 or higher than 150, or genotype quality (GQ) lower than 20 were marked as missing sites. For each SNP, the overall filter parameters were: QD<10.0, SOR>3.0, FS>60.0, MQ<40.0, MQRankSum<-12.5 and ReadPosRankSum<-8.0. The overall filter parameters for indels were: QD<10.0, FS>60.0, and ReadPosRankSum<-8.0. Indel data was only used in the analysis of mutation load. Given the broad distribution of samples from different evolutionary lineages with multiple levels of divergence in the common vole,[40,41] setting a filter of allele frequency would cause a bias towards less heterozygosity (for example up to 39% less when set to 0.05) in samples of *M. obscurus*, from the Italian lineage and some from the Western lineage, so variants with low allele frequency were preserved. With overall high sequencing depth and quality, the accuracy of called singletons was expected to be high. Keeping singletons was also essential for SFS analysis. Indel data was only used in the analysis of mutation load. Only variants from 22 chromosome-scale scaffolds corresponding to the autosomes were used for analyses.

### Mitochondrial genome analyses

Raw reads were mapped to the complete mitochondrial genome of *M. arvalis* (GenBank: MG948434.1) using BWA 0.7.17[100] with default parameters. Duplicated reads were marked and removed with GATK MarkDuplicates. Joint SNP calling was done with BCFtools[101] mpileup (version 1.9) and only sites with QUAL higher than 50 were kept. For each individual, the full mitochondrial sequence was converted from VCF with reference (GenBank: MG948434.1) using BCFtools consensus function, and sites with read depth lower than 3 or GQ lower than 99 were marked as missing sites. Phylogenetic reconstructions based on full mitochondrial sequences were achieved using MrBayes 3.2.7[102] with the GTR+I+G substitution model which was chosen with JModelTest 2.7.[103] Two runs were completed with 4 chains, 2,000,000 iterations and 25% burn-in to make sure the final standard deviation of split frequencies was lower than 0.01. The mitochondrial sequences of *M. obscurus* were used as outgroup for phylogenetic reconstruction.

### Recombination map and mutation rate

To estimate the mutation rate, SNP phasing and a recombination map were needed. Two steps were taken to obtain phased genome sequences. First, for each individual, SNPs were phased into small blocks based on reads covering two SNPs or more using WhatsHap 1.0.[104] Then the pre-phased data of individuals from Orkney and the Western lineage were pooled together and phased with SHAPEIT4 4.1.2[105] taking into account read-based blocks with an error rate of 0.0001. The specific default parameter set designed for sequencing data (–sequencing) was used. Though WhatsHap was originally designed for long-read sequencing, using short-read data with high depth can still effectively reduce the runtime and increase the accuracy for grouped phasing with SHAPEIT4.[105]

Phased SNP data was converted to fasta files using the BCFtools consensus function and missing sites were masked as "N". Considering that population structure could bias the estimation of linkage disequilibrium (LD),[118,119] only five individuals from the northern coastal area of the Western lineage were used for the estimation. The R package LDJump,[106] designed for small sample size, was used to produce the recombination map. Default parameters were used except for the segment length set to 2kb. The population recombination parameter $\rho$ was then converted to cM using the autosomal average recombination rate 0.63 cM/Mb for rat.[120]

The genome wide mutation rate was estimated using the TMRCA (time to the most recent common ancestor) regression method from Palamara et al.[121] which is based on the genetic differences of identical-by-descent (IBD) segments from an inbred population. In short, the average mismatch rate of IBD segments longer than a certain length linearly regresses to the posterior mean age of these segments. Five individuals from Mainland Orkney were used for mutation rate estimation. IBD segments were detected using phased

Please cite this article in press as: Wang et al., Demographic history and genomic consequences of 10,000 generations of isolation in a wild mammal, Current Biology (2023), https://doi.org/10.1016/j.cub.2023.04.042





SNP data and the estimated recombination map with hap-IBD[107] and default parameters. The mismatch rate of each segment was calculated by directly counting mismatch sites. IBD segments longer than 0.5 cM and only MRCA timepoints with more than 100 segments were used for regression.

### Nuclear genomic diversity and divergence

The numbers of SNPs and heterozygous sites were counted using BCFtools 1.9. PCA was done using PLINK 1.9[108] on either SNPs of all individuals or only Orkney vole individuals without any missing data. We used IQ-TREE 2.1.4[109] and Treemix 1.13[110] to obtain a nuclear phylogeny and estimate the number and extent of historical migration events. To avoid excessive run times of the Maximum Likelihood reconstruction of the phylogeny, we subsampled our dataset to 5 million random SNPs without any missing data. The three *M. obscurus* individuals were used as outgroup. Implemented substitution model test, 1,000 ultrafast bootstraps and 1,000 SH-like approximate likelihood ratio tests were performed with IQ-TREE. Independent runs of Treemix were performed with 0 to 8 migration events added to the tree with 10 iterations for each number of migration events. The optimal number of migration events was estimated with the R package "optM" using the "Evanno" method (https://cran.r-project.org/web/packages/OptM/). For the Western evolutionary lineage, three geographical groups of genomes were defined according to Treemix results.

To further examine our dataset for genetic admixture between Orkney and continental voles and between different Orkney islands, *f*-branch statistics were calculated with Dsuite[111] based on the phylogeny obtained with IQ-TREE. This method is based on the $f_4$-ratio and can assign admixture events to specific branches given a phylogenetic tree.[53] Additionally, we performed population clustering with ADMIXTURE for either all *M. arvalis* individuals in the dataset or only for Orkney individuals to help interpret the admixture and gene flow events.

### Demographic history

To investigate the demographic history and the sequence of colonization and divergence of voles on different Orkney islands, we used SMC++[54] to estimate the change of population sizes and divergence times. Population size changes were first estimated for each population from 100 to 500,000 generations ago, with 40 iterations using pchip spline, 30 spline knots, and regularization penalty at 8. We used the mutation rate estimate of $8.7\times10^{-9}$ per generation (SI) and a generation time of 0.5 years.[41] The divergence time was then estimated between Orkney Mainland voles and the Western-north group which is closest to the ancestral populations. We obtained further estimates for the split between Orkney Mainland and the other islands where Orkney voles occur, except for Eday island. Given the known introduction and release of 15 voles from Westray in Eday in 1987 by M. Cockram (personal communication), we attempted to estimate the split time between Eday and Westray.

### Mutation load

In order to estimate the frequency of deleterious mutations in Orkney voles and Western lineage populations, SnpEff 5.0[112] was used to annotate the functional effects of the variants. Gene annotation of *M. arvalis* (BioProject ID: PRJNA737461, A. Gouy, G.H., X.W., et al., unpublished data) was used to locate SNPs and indels in genes. Variants marked with the impact category "high" by SnpEff were considered deleterious mutations. This impact category includes mutations heavily affecting the function of the protein, for example mutations which eliminate start or stop codons, or frameshifting insertions and deletions. The number of variants with different functional consequences was counted per individual. Putatively neutral variants were defined as variants assumed to have no effect on any gene (category "modifier" defined by SnpEff) and at least 10 Kb away from any gene. The genomes of the three *M. obscurus* individuals were used to define derived states of the genotypes. Only those sites were kept where the allele frequencies in *M. obscurus* were zero, and for each site the alternative genotype was considered as the derived allele. Note that the derived states were inferred from a single outgroup species and thus may not be as reliable as when based on multiple species. To avoid the uncertainty of ancestral state, non-biallelic variants were filtered. Variants that were heterozygous in more than 50% of the individuals or had missing data were also filtered. The unfolded allele site frequency spectrum (SFS) of deleterious mutations and neutral mutations was then calculated for each group of the Western evolutionary lineage and the Orkney islands. Statistical enrichment and underrepresentation analyses of Gene Ontology terms were performed on g:Profiler web server[113] using genes with deleterious variants fixed in Orkney populations and the gene list of mouse as reference.

## QUANTIFICATION AND STATISTICAL ANALYSIS

### Demographic history

For each SMC++ analysis, we produced 50 bootstrap replicates for each population with the script from Zheng et al.[122] and performed a SMC++ run for each replicate with the same parameters as above. Due to the limitation of computational time, the size of each bootstrap replicate was reduced to 5 chromosomes of 100 Mb formed by 5 Mb blocks randomly chosen from the original genomes. For the reported population sizes and split times, we then estimated the median, 5$^{th}$ percentile and 95$^{th}$ percentile for each in R.

### Mutation load

To estimate the change of selection efficacy in Orkney voles compared to their continental conspecifics, we calculated the standardized rate of private deleterious allele $R'_{XY}$[58] of Orkney populations relative to the Western-north group. For each Orkney population,

Please cite this article in press as: Wang et al., Demographic history and genomic consequences of 10,000 generations of isolation in a wild mammal, Current Biology (2023), https://doi.org/10.1016/j.cub.2023.04.042



$R_{XY}$ of highly deleterious, missense (SnpEff impact category "moderate" in coding regions), and synonymous (SnpEff impact category "low" in coding regions) SNPs were calculated and divided by $R_{XY}$ of neutral SNPs to acquire the standardized $R'_{XY}$ to eliminate the effect of unequal distance to the outgroup. The standard error was estimated by jackknifing of 100 blocks of equal number of SNPs.

### Individual based simulations

We conducted forward-time, individual-based simulations using SLiM v3.2.[114] We modeled the Mainland population as a single panmictic population with a demographic history that mimics the one inferred using SMC++. We simplified the demographic history in a 3-epoch model with instantaneous changes in population size. We first simulated an ancestral population of 150,000 diploid individuals for a burn-in period of 450,000 generations. After 12,000 generations the population size changed to 5,000 individuals for 8,000 generations. Then the population increased to the current effective population size of 75,000 individuals for another 4,000 generations (see Figure S5C for an illustration of the demographic model). We then sampled 10 individuals from the population to estimate the expected SFS corresponding to the number of genomes from Orkney Mainland. First, we simulated 100,000 sites with a recombination rate of $10^{-6}$ between consecutive sites and the estimated mutation rate (see supplemental information) mentioned above. We simulated both neutral and deleterious SNPs. The distribution of fitness effects for deleterious mutations was modelled as a Gamma distribution with mean -0.01 and shape parameter 0.1. The dominance coefficient of deleterious mutations was set to h = 0.25. To speed up computation times all parameters were rescaled by a factor of 50 (see chapter 5.5 in the SLiM manual). We chose the parameters of the DFE to span the range of plausible values for selection coefficients of deleterious mutations estimated in natural populations.[123–125] Second, to test the robustness of our results and to further investigate the role of the shape of the DFE, we performed an additional set of simulations with a refined DFE combining different dominance coefficients, a lower recombination rate of $10^{-8}$ between consecutive sites, and smaller scaling factors to allow for broader DFEs including more large-effect mutations including sub-lethal and lethal mutations (Figure S5). We further examined the simulation results for evidence of pseudo-overdominance (POD) that could potentially reduce the expression of fitness effects, including low recombination rate regions with an excess of deleterious mutations or an excess of intermediate frequency deleterious variants in the SFS.