










# Population Genomics of Giant Mice from the Faroe Islands: Hybridization, Colonization, and a Novel Challenge to Identifying Genomic Targets of Selection

Bret A. Payseur <sup>1,\*</sup>, Peicheng Jing <sup>1</sup>, Emma K. Howell <sup>1</sup>, Megan E. Frayer <sup>1</sup>, Eleanor P. Jones <sup>2,3</sup>, Eyðfinn Magnussen <sup>4</sup>, Jens-Kjeld Jensen <sup>5</sup>, Yingguang Frank Chan <sup>6,7</sup>, Jeremy B. Searle <sup>8</sup>

<sup>1</sup>Laboratory of Genetics, University of Wisconsin, Madison, WI 53706, USA

<sup>2</sup>Fera Science, The National Agri-Food Innovation Campus, York YO41 1LZ, UK

<sup>3</sup>School of Natural and Environmental Sciences, University of Newcastle, Newcastle NE1 7RU, UK

<sup>4</sup>Faculty of Science and Technology, University of the Faroe Islands, Tórshavn, Faroe Islands

<sup>5</sup>Naturalist, Nólsoy FO-270, Faroe Islands

<sup>6</sup>Friedrich Miescher Laboratory of the Max Planck Society, Tübingen 72076, Germany

<sup>7</sup>Groningen Institute for Evolutionary Life Sciences (GELIFES), University of Groningen, Groningen 9747AG, The Netherlands

<sup>8</sup>Department of Ecology and Evolutionary Biology, Cornell University, Ithaca, NY 14853, USA

\*Corresponding author: Email: [bret.payseur@wisc.edu](mailto:bret.payseur@wisc.edu).

Accepted: July 09, 2025

## Abstract

Populations that colonize islands provide unique insights into demography, adaptation, and the spread of invasive species. House mice on the Faroe Islands evolved exceptionally large bodies after colonization, generating longstanding interest from biologists. To reconstruct the evolutionary history of these mice, we sequenced genomes of population samples from three Faroe Islands (Sandoy, Nólsoy, and Mykines) and Norway as a mainland comparison. Mice from the Faroe Islands are hybrids between the subspecies *Mus musculus domesticus* and *M. m. musculus*, with ancestry alternating along the genome. Analyses based on the site frequency spectrum of single nucleotide polymorphisms and the ancestral recombination graph (ARG) indicate that mice arrived on the Faroe Islands on a timescale consistent with transport by Norwegian Vikings, with colonization of Sandoy likely preceding colonization of Nólsoy. Substantial reductions in nucleotide diversity and effective population size associated with colonization suggest that mice on the Faroe Islands evolved large body size during periods of heightened genetic drift. Genomic scans for positive selection uncover windows with unusual site frequency spectra, but this pattern is mostly generated by clusters of singletons in individual mice. Three genomic regions show evidence for selection on islands based on the ARG, including variants located in transcription factor binding sites. Our findings reveal a dynamic evolutionary history for the enigmatic mice from Faroe Island and emphasize the challenges that accompany population genomic inferences in island populations.

**Key words:** island evolution, hybridization, demographic inference, positive selection, house mouse.

© The Author(s) 2025. Published by Oxford University Press on behalf of Society for Molecular Biology and Evolution.

This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial License (<https://creativecommons.org/licenses/by-nc/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact [reprints@oup.com](mailto:reprints@oup.com) for reprints and translation rights for reprints. All other permissions can be obtained through our RightsLink service via the Permissions link on the article page on our site—for further information please contact [journals.permissions@oup.com](mailto:journals.permissions@oup.com).

## Significance

Populations that colonize islands are expected to have unusual histories compared to their mainland counterparts. Using population genomic data, we conclude that giant mice living on the Faroe Islands originated from hybrids, invaded the islands on a timescale consistent with transport by Vikings, and persisted despite drastic reductions in population size. We also discover a novel challenge to scanning genomes for genes involved in adaptation.

## Introduction

The colonization of islands is a perilous task. To persist, populations that colonize islands must quickly adapt to novel environments in the face of drastic changes in population size and migration. As a result, island colonizers provide rich systems for studying the interplay of complex demography and natural selection (Losos and Ricklefs 2009). Evolutionary inferences from island populations also reveal conditions that favor the spread of invasive species, thereby informing conservation strategies (Estoup and Guillemaud 2010).

House mice (*Mus musculus*) are a model system for understanding island colonization on recent evolutionary timescales (Berry 1996). Among mammalian species, house mice are second only to humans in their global distribution (Angel et al. 2009). Human commensalism and a capacity for rapid adaptation has enabled house mice to invade islands around the world that encompass a wide variety of environments (Berry 1964; Berry and Peters 1975; Berry et al. 1978a, 1978b, 1979; Rowe-Rowe and Crafford 1992; Berry 1996; Berry 2009; Gabriel et al. 2010). Patterns of genetic variation in house mice can even be used to reconstruct human migration to islands (Jones et al. 2012, 2013).

One island radiation of house mice has long captivated biologists. The Faroe Islands comprise a small archipelago located in the North Atlantic approximately halfway between Norway and Iceland. The archipelago includes 18 islands, six of which are inhabited by house mice (Wilches et al. 2021). Mice on the Faroe Islands were cited as early examples of rapid adaptation and speciation by Degerbøl (1942) and Huxley (1942). Based on the unusually large size of five specimens from Nólsoy, Clarke (1904) assigned mice from the Faroe Islands to a new subspecies (*Mus musculus faroensis*). Miller (1912) subsequently elevated the distinction of the Faroe Island mice to the level of species. Later measurements confirmed the Faroe Island mice to be among the largest wild-living house mice in the world (Berry et al. 1978b). Due to cold temperatures on the Faroe Islands, mice there are suspected to be among the most climatically stressed in western Europe (Berry and Jakobson 1975), “probably approaching the physiological limit of their outdoor range” (Berry and Peters 1977). Mice from the Faroe Islands constitute one of several examples of house mice evolving extreme body size on islands

(Berry 1964; Berry and Peters 1975; Berry et al. 1978a, 1978b; Rowe-Rowe and Crafford 1992). Genetic mapping revealed a complex genetic basis to size evolution, with 111 quantitative trait loci (QTL) that contribute to differences in body weight, body proportions, and growth hormone concentrations between mice from the Faroe Island of Mykines and a laboratory mouse strain selected for small body size (Wilches et al. 2021).

The first historical record of house mice inhabiting the Faroe Islands is in 1676 (Debes 1676). History (Edwards 2005), biogeography (Berry et al. 1978b), and mitochondrial sequence similarity (Jones et al. 2011) suggest that Norwegian Vikings (the Norse) were responsible for passively introducing mice to the Faroe Islands around 800 A.D., with potential later introductions of mice by the British, Dutch, German, and Danish (Jones et al. 2011). Based on analyses at a few molecular markers (Jones et al. 2011) and suspected origins in Norway, where many hybrids are found (Jones et al. 2010), it has been proposed that mice from the Faroe Islands derive ancestry from house mouse subspecies from western and eastern Europe (*M. musculus domesticus* and *M. m. musculus*, respectively) (Jones et al. 2011).

Key questions remain unanswered about the evolutionary history of the unusual mice living on the Faroe Islands. Are house mice on the Faroe Islands descended from hybrids? If so, how is the history of hybridization recorded in their genomes? When did house mice colonize the Faroe Islands? What population size changes accompanied island colonization? Did mice on the Faroe Islands continue to exchange genes with mice on the mainland after the island populations were established? Is there evidence for positive selection associated with island adaptation? The answers to these questions will reveal the conditions under which mice evolved extreme body sizes on the Faroe Islands.

Here, we take advantage of the unique perspective and expansive toolkit provided by population genomics to revisit the evolutionary history of the enigmatic house mice on the Faroe Islands, making comparisons with house mice from Norway as the colonization source. We reveal the genomic footprint of hybridization between subspecies and characterize the dynamics of demographic history associated with island colonization. We report genome-wide scans for positive selection that reveal a novel challenge with population genomic characterization of adaptation.

Overall, we provide an increasingly rich evolutionary portrait of a fascinating, very recent island radiation.

## Results

### Genetic Differentiation Between Populations

To visualize population structure among mice from the Faroe Islands, we conducted principal components analysis on the covariance matrix of genotypes (Price et al. 2006). Mice from Nólsoy, Sandoy, Mykines, and Norway form visually distinct groups based on their scores along principal components 1 and 2 (Fig. 1). As a result, we treat mice from these geographic locations as separate populations in subsequent analyses. There also is evidence for population differentiation among mice within Norway, which we consider in the next section.

### Hybrid Ancestry

We used a probabilistic approach to infer the subspecific ancestries of Faroe Island mice (Corbett-Detig and Nielsen 2017). Ancestry analyses indicate that Faroe Island mice are hybrids between *M. m. domesticus* and *M. m. musculus* (*domesticus* and *musculus*, hereafter). Mice are inferred to be descended primarily from *domesticus* (Fig. 2), with the genomic proportion of *domesticus* ancestry spanning 0.949 to 0.955 across Nólsoy mice, 0.906 to 0.920 across Sandoy mice, and 0.941 in one Mykines mouse (Table S1). Mice from Norway vary widely, with the genomic proportion of *domesticus* ancestry ranging from 0.095 to 0.925. Four mice from Norway with high proportions (>0.9) of *domesticus* ancestry were used for analyses of island colonization history (see next section). Heterogeneity, the proportion of the genome heterozygous for ancestry, is 0.006 to 0.140 among mice from Nólsoy, 0.031 to 0.046 in mice from Sandoy, and 0.002 in one Mykines mouse. Heterogeneity values are higher in mice from Norway (0.024 to 0.259).

Moving along a chromosome, subspecific ancestry alternates between *domesticus*, *musculus*, and heterogenic (ancestry from both subspecies) (Fig. 2). Individual mice from Nólsoy, Sandoy, and Mykines harbor an average of 10.2, 10.3, and 10.6 junctions (switches) in ancestry per Mb, respectively. Junction numbers are higher in mice from Norway. There is variation in the spatial distribution of junctions within and among chromosomes (Fig. S1). Junction patterns are more similar among island mice than among mainland mice from Norway.

Junctions represent historical recombination events that occurred in heterogenic regions of the genome (Baird 2006). To determine whether the rate of ancestry switching increases with recombination, we asked whether junction number is correlated with local recombination rate (cM/Mb; Cox et al. 2009) using comparisons across 10 Mb windows in mice from island populations. Spearman's rank

correlations between recombination rate and the average number of junctions across mice are  $\rho = 0.47$  ( $P = 7.23 \times 10^{-13}$ ) in Nólsoy,  $\rho = 0.47$  ( $P = 6.78 \times 10^{-13}$ ) in Sandoy, and  $\rho = 0.37$  ( $P = 4.82 \times 10^{-8}$ ) in Mykines.

The hybrid nature of the Faroe Island mice raises the possibility that body size evolution was partly shaped by subspecies ancestry. To test this hypothesis, we evaluated whether quantitative trait loci (QTL) for body size identified by crossing mice from Mykines with a laboratory strain (Wilches et al. 2021) are enriched for *domesticus* or *musculus* ancestry in 2 Mb windows in island populations. QTL intervals are enriched for *domesticus* ancestry in Nólsoy (Welch's *t*-test;  $t_{45} = 4.55$ ;  $P = 4.05 \times 10^{-5}$ ), but not in Sandoy ( $t_{25} = -0.39$ ;  $P = 0.7$ ) nor in Mykines ( $t_{25} = -0.47$ ;  $P = 0.64$ ).

The *domesticus* and *musculus* subspecies show evidence of partial reproductive isolation in the form of reduced male fertility in  $F_1$  (Forejt and Ivanyi 1974; Good et al. 2008),  $F_2$  (White et al. 2011), and wild-caught hybrids (Turner et al. 2012). Selection against hybrids is expected to remove ancestry from the minor subspecies disproportionately in regions of low recombination, where neutral variants are more likely to be linked to incompatible alleles (Schumer et al. 2018; Moran et al. 2021). To test this prediction, we compared the average proportion of *musculus* ancestry to the recombination rate in 2 Mb windows along the genome. The proportion of *musculus* ancestry is positively correlated with recombination rate in mice from Nólsoy (Spearman's  $\rho = 0.09$ ;  $P = 0.003$ ), Sandoy ( $\rho = 0.06$ ;  $P = 0.04$ ), and Norway (four mice with majority *domesticus* ancestry used in demographic analyses below;  $\rho = 0.16$ ;  $P = 5 \times 10^{-8}$ ), consistent with selection against hybrid incompatibilities.

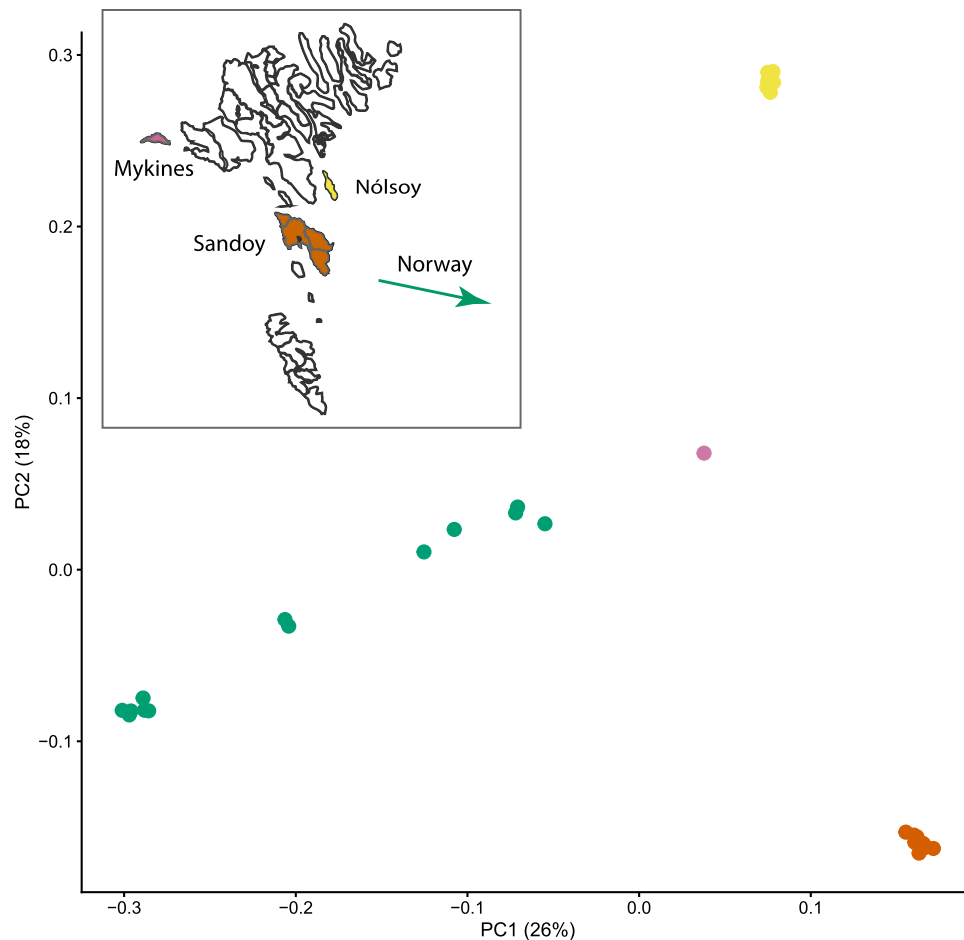
### Island Colonization History

#### Levels of Variation

Focusing on all genomic regions with shared *domesticus* ancestry, mice on the Faroe Islands show substantial reductions in nucleotide diversity (the average number of pairwise differences; Tajima 1983). Average nucleotide diversity across windows is 0.066% in mice from Nólsoy, 0.078% in mice from Sandoy, and 0.199% in mice from Norway. Mice on islands harbor significantly less diversity than mice in Norway (paired *t*-tests; Nólsoy vs. Norway:  $t_{322,778} = -205.49$ ,  $P < 2.2 \times 10^{-16}$ ; Sandoy vs. Norway:  $t_{289,431} = -186.8$ ,  $2.2 \times 10^{-16}$ ). The percentage of windows with zero diversity is 43.0% in mice from Nólsoy and 33.1% in mice from Sandoy, compared to 10.1% in mice from Norway. These low levels of genetic variation suggest that the history of Faroe Island mice involved reductions in population size.

#### The Site Frequency Spectrum

To prepare for demographic analyses, we chose sequences with shared *domesticus* ancestry filtered to minimize



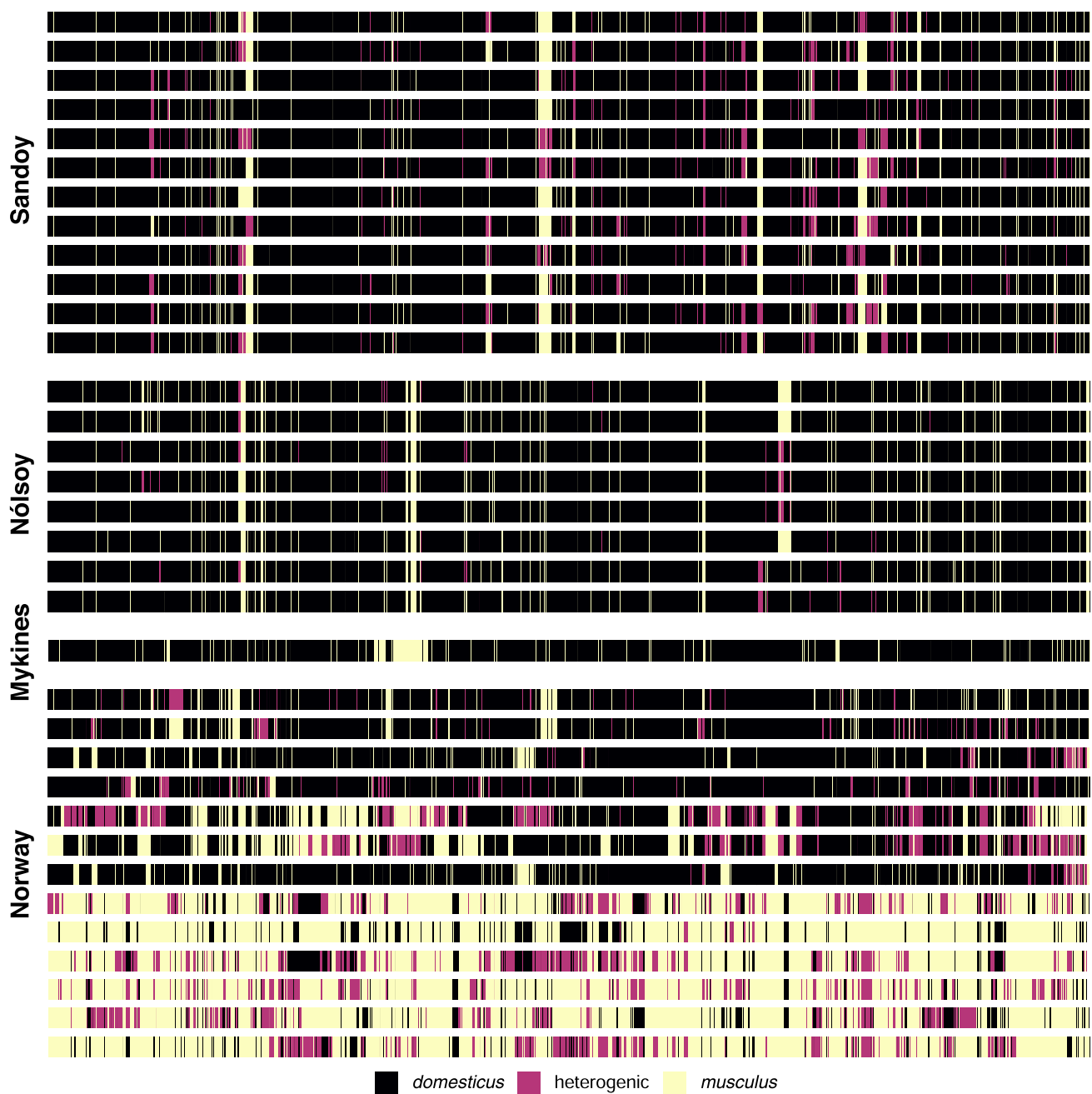
**Fig. 1.** Genetic differentiation between mice from Nólsoy, Sandoy, Mykines, and Norway, and a map of the Faroe Islands. Each dot represents one mouse, plotted using its scores along principal components 1 and 2 inferred from principal components analysis (PCA) of the genotype covariance matrix. Map of the Faroe Islands is inset, with an arrow marking the direction toward Norway. Color codes on the PCA plot match color codes on the map.

selection at linked sites. We focused on sequences with *domesticus* ancestry to reduce effects of hybridization on the site frequency spectrum (SFS) (Blanckaert and Payseur 2021). Folded SFS of single nucleotide polymorphisms (SNPs) residing within these regions differ among populations (Fig. 3). Compared to SFS for mainland Norway, SFS for island populations (Nólsoy and Sandoy) display more evenness in the proportions of SNPs at common frequencies. Site frequency spectra for island populations also show bigger disparities between proportions of rare versus common SNPs. These characteristics of the island population SFS are consistent with nonequilibrium demographic histories, including changes in population size over time. Site frequency spectra also differ depending on whether they were estimated from genotype likelihoods (with ANGSD) or reconstructed from called genotypes (with GATK). For mice from Nólsoy, ANGSD estimates a much higher proportion of singleton SNPs than GATK. For mice from Sandoy, ANGSD estimates higher proportions of singletons and doubletons than GATK.

### Model-based Inferences

To reconstruct demographic history associated with island colonization, we used two complementary approaches. In the first approach, we fit a variety of models to the folded, 2D SFS (Gutenkunst et al. 2009) for three pairs of populations: Norway–Sandoy, Norway–Nólsoy, and Sandoy–Nólsoy. Considering differences between the SFS recovered with GATK and ANGSD, we conducted separate sets of demographic analyses on the 2D SFS inferred by each method.

In the second approach, we estimated demographic parameters from ancestral recombination graphs for each pair of populations (ARG) (Speidel et al. 2019). In contrast to SFS-based analyses, ARG-based analyses incorporated variation throughout the genome, regardless of subspecies ancestry, used only GATK-derived genotypes, and focused on derived alleles inferred using two outgroup species (*M. caroli* and *M. spretus*). Here, we highlight comparisons between estimates of the same demographic parameters from the two approaches.

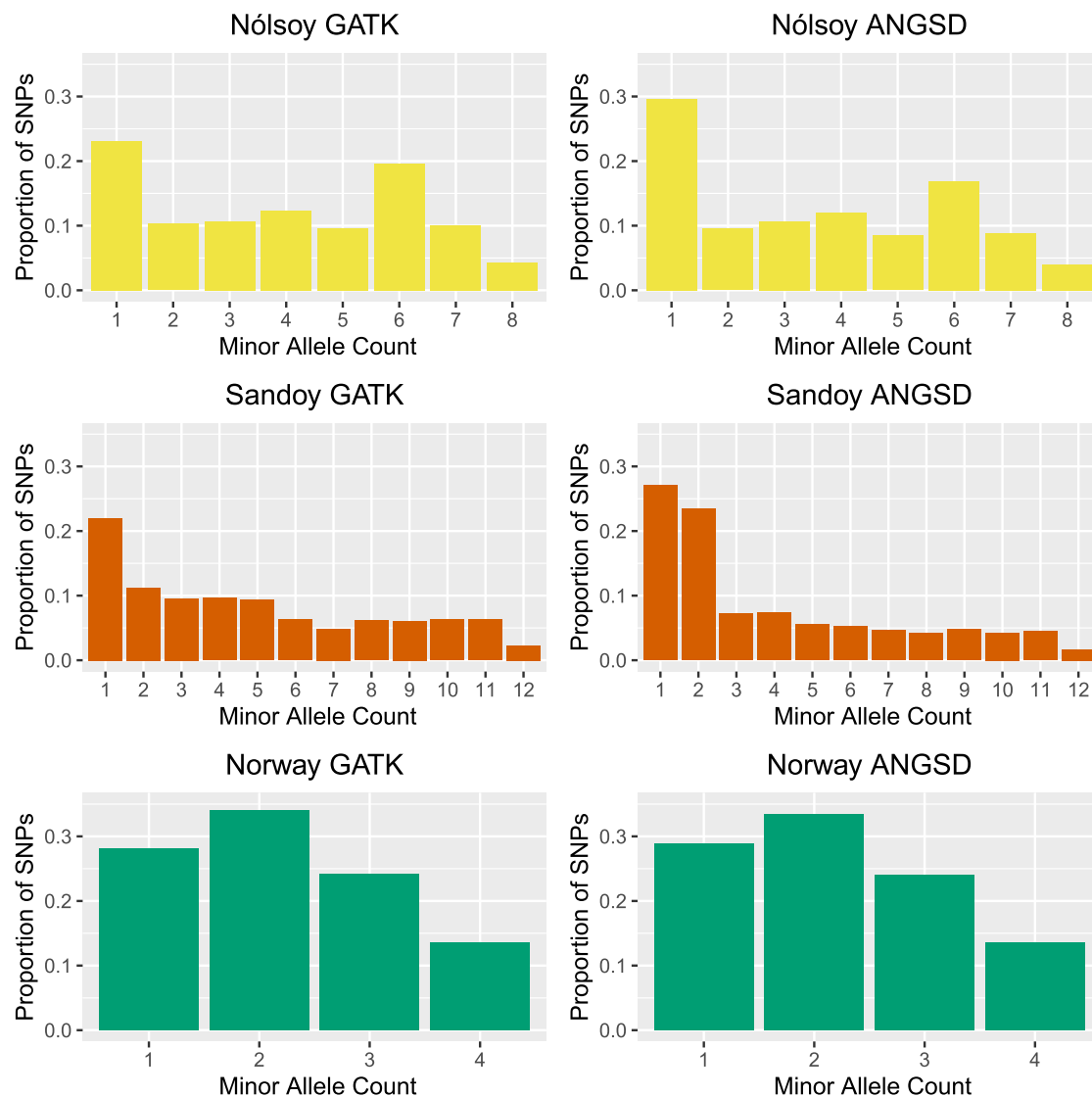


**Fig. 2.** Subspecies ancestries along chromosome 3 for mice from Nólsoy, Sandoy, Mykines, and Norway. Each row is one mouse. Ancestries are color-coded. Ancestries for other chromosomes are shown in Fig. S1.

### Norway–Sandoy

The best-fitting demographic model for the Norway–Sandoy GATK-derived 2D SFS includes a population size reduction in Norway followed by a severe bottleneck in Sandoy after its split from Norway (Fig. 4). The best-fitting model for the Norway–Sandoy ANGSD-derived 2D SFS includes these features plus one-way migration from Norway to Sandoy (Fig. S2). For some parameter estimates,

the disparities are large between GATK-derived SFS and ANGSD-derived SFS (Table S2). Effective population size ( $N_e$ ) estimates for Sandoy immediately following the split from mainland Norway are 760 (GATK) and 7 (ANGSD). The timing of the population split ranges from 3,176 (GATK) to 5,411 (ANGSD) generations ago. The Sandoy bottleneck is estimated to end 1,286 (GATK) or 5,396 (ANGSD) generations ago. Both best-fitting models include historical inbreeding (Table S2).



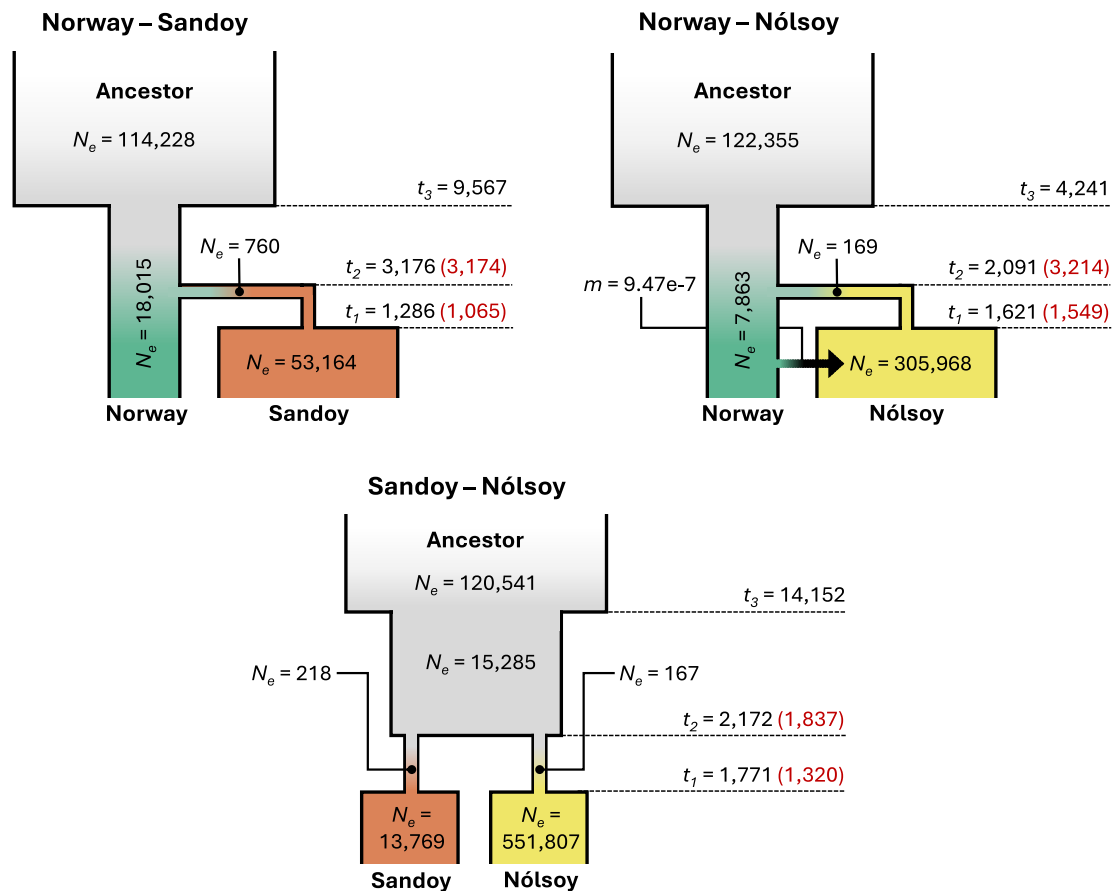
**Fig. 3.** Site frequency spectra for Nólsoy, Sandoy, and Norway. Site frequency spectra were inferred from putatively neutral SNPs in genomic regions with shared *M. m. domesticus* ancestry, separately using GATK and ANGSD. Color-coding matches Figs. 1 and 4.

The split time estimate for Norway–Sandoy from the ARG is 3,174 generations ago, very close to the corresponding parameter estimate from the GATK-derived SFS of 3,176 generations ago (Fig. 4). For Sandoy, the ARG yields a minimum  $N_e$  at 1,065 generations ago (Fig. S3), which is close to the end of the bottleneck from the GATK-derived SFS at 1,286 generations ago. The ARG estimate for lowest Sandoy  $N_e$  is 1,178, compared to the GATK-derived SFS  $N_e$  of 760 immediately after the split from mainland Norway.

#### Norway–Nólsoy

Similar to the Norway–Sandoy comparison, the best-fitting demographic model for the Norway–Nólsoy GATK-derived

2D SFS includes a population size reduction in Norway followed by a bottleneck in Nólsoy that accompanies the population split from the mainland (Fig. 4). The best-fitting GATK-derived model additionally includes one-way migration from Norway to Nólsoy after their split (Fig. 4). The pre-split mainland size reduction is absent from the best-fitting model for the Norway–Nólsoy ANGSD-derived 2D SFS (Fig. S2). Disparities in parameter estimates between GATK-derived SFS and ANGSD-derived SFS are smaller for Norway–Nólsoy than for Norway–Sandoy. For both models,  $N_e$  of Nólsoy is small (169 vs. 9) immediately following its split from mainland Norway and  $N_e$  recovers soon afterward, suggesting rapid expansion of island mice. The split time between Norway and Nólsoy is estimated to be 2,091 (GATK-derived SFS) or 3,532 (ANGSD-derived SFS)



**Fig. 4.** Best-fitting demographic models for mice from Nólsoy and Sandoy, treating mice from Norway as a mainland reference population. Parameter estimates come from analyses of 2D site frequency spectra for pairs of populations inferred from GATK-derived genotypes for putatively neutral SNPs in genomic regions with shared *M. m. domesticus* ancestry. Estimates for comparable parameters from analyses of ancestral recombination graphs are in parentheses.  $N_e$ , effective population size;  $t$ , timing in generations of population splits or changes in  $N_e$ , ordered beginning with the most recent event;  $m$ , migration rate.

generations ago. Both models support historical inbreeding and a low rate of migration from Norway to Nólsoy (Table S2). Modest standard deviations accompany most parameter estimates (Table S2).

The Nólsoy–Norway split time is estimated from the ARG to be 3,214 generations ago, which can be compared to the time of 2,091 generations ago estimated from the GATK-derived SFS (Fig. 4). For Nólsoy, the minimum  $N_e$  reconstructed from the ARG occurs 1,549 generations ago, which is close to the time of 1,621 generations ago at which the bottleneck ends, as estimated from the GATK-derived SFS. The ARG estimate for the lowest Nólsoy  $N_e$  is 1,183 (Fig. S3), compared to 169 from the GATK-derived SFS.

#### Sandoy–Nólsoy

The best-fitting model for the Sandoy–Nólsoy GATK-derived 2D SFS includes a reduction in ancestral

population size followed by a bottleneck on both islands (Fig. 4). The best-fitting model for the Sandoy–Nólsoy ANGSD-derived 2D SFS excludes the ancestral size reduction and includes migration from Nólsoy to Sandoy following their split (Fig. S2).  $N_e$  estimates are similar for Sandoy and Nólsoy for both GATK-derived SFS and ANGSD-derived SFS (Table S2). Population split time is 2,172 (GATK) or 3,774 (ANGSD) generations ago. Both models include historical inbreeding (Table S2).

The split time for the two island populations based on the ARG (1,837 generations ago) is close to the split time inferred from the GATK-derived SFS (2,172 generations ago) (Fig. 4). The times with the lowest  $N_e$  are similar for Sandoy (1,320 generations ago) and Nólsoy (1,349 generations ago) (Fig. S3), and a few hundred generations more recent than the bottleneck end estimated from the GATK-derived SFS (1,771). The ARG estimate for lowest island  $N_e$  is 955, compared to  $N_e$  following the population split based on the GATK-derived SFS (Sandoy: 218; Nólsoy: 167).

## Genomic Scans for Selection Connected to Island Adaptation

### Selection Scans Based on Summary Statistics

To generate patterns of variation expected under neutrality, we conducted coalescent simulations assuming parameter estimates from models without inbreeding that showed the best fit to the GATK-derived SFS. Distributions of nucleotide diversity, Tajima's  $D$  (a measure of skew in the SFS; Tajima 1989), and  $F_{ST}$  (a measure of population differentiation; Hudson et al. 1992) computed from windows in datasets simulated under these models are close to distributions from the real datasets including all windows with shared *domesticus* ancestry (Fig. S4), suggesting they provide reasonable null models for genomic scans for selection.

Our first genomic scan for selection considered metrics that summarize the SFS (Tajima's  $D$ ) and population differentiation ( $F_{ST}$ ) between island and mainland. Across windows with shared *domesticus* ancestry, 84 windows in Nólsoy and 57 windows in Sandoy have Tajima's  $D$  values that are unexpectedly negative for neutrality (i.e. in the direction expected under a simple selective sweep; Braverman et al. 1995) (Table S3), given simulated demographic histories and strict corrections for multiple testing (genome-wide Bonferroni  $P < 0.05$ ). Tajima's  $D$  values in these windows range from  $-2.52$  to  $-2.43$  in Nólsoy and range from  $-2.75$  to  $-2.62$  in Sandoy.

As expected, large percentages of the variants in these windows with unusual Tajima's  $D$  values are singletons (96.6% to 100% in Nólsoy; 96.1% to 100% in Sandoy). Surprisingly, all or most of the singletons in these windows are found in a single mouse, with the identity of the mouse varying among windows (Table S3). This observation comes from a nontrivial number of singletons in these outlier windows (Nólsoy average: 59.2 singletons; Sandoy average: 64.5 singletons), so it is not a simple consequence of low levels of variation. This pattern seems incompatible with positive selection as the cause of the unusual values of Tajima's  $D$  in these windows, so we do not pursue them further. No windows show unusual values of  $F_{ST}$  once demographic history and multiple testing are accounted for.

### Selection Scans Based on the Full Site Frequency Spectrum

Our second genomic scan for selection searched for windows with SFS that depart from the genome-wide SFS (DeGiorgio et al. 2016) in shared *domesticus* regions. Two thousand seven hundred and forty-six windows in Nólsoy and 154 windows in Sandoy have likelihood ratio test statistics that are unusual for neutrality, given the demographic history and strict corrections for multiple testing (genome-wide Bonferroni  $P < 0.05$ ) (Table S4). None of these windows are the same in Nólsoy and Sandoy. These

windows with unusual SFS form several clusters in the genome.

Although this approach detects more windows with potential evidence of positive selection than scans based on Tajima's  $D$  (above), most of these windows are afflicted by the same issue: an excess of singletons derived mostly or exclusively from a single mouse, with the identity of this mouse varying among windows (Table S4). Since this pattern seems unexpected under models of positive selection, we do not pursue these windows further.

### Selection Scans Based on the Ancestral Recombination Graph

Our third approach to scanning the genome for selection searched for derived SNPs with unusually high frequencies given their genealogies, as inferred using the ARG (Speidel et al. 2019). In contrast to scans based on the SFS, this approach incorporated regions throughout the genome, irrespective of subspecies ancestry. Although ARG analyses identify SNPs with evidence for positive selection, none of the tests conducted separately in Nólsoy and Sandoy survive corrections for multiple testing based on either the number of SNPs or the estimated number of genealogies across the genome. A total of 670 SNPs have uncorrected  $P < 0.05$  in both Nólsoy and Sandoy, but not in Norway (Table S5). These SNPs are clustered into a modest number of genomic regions. Genes in these intervals are significantly enriched for biological processes and cellular components connected to neurological features and functions, including synapses between neurons and synaptic vesicle fusion (Table S5).

To increase our power to identify island-specific positive selection, we repeated the ARG selection tests, treating mice from Nólsoy and Sandoy as a single population. In these analyses, nine SNPs have  $P$ -values that survive corrections for multiple testing and show differences in allele frequency of at least 0.5 when compared to Norway (Table S6). Seven of these SNPs lie in a single 6,241 bp window on chromosome 4 that contains L1 long interspersed nuclear elements (LINEs). The frequencies of the 7-SNP haplotype containing all derived alleles are 22/24 in Sandoy and 7/16 in Nólsoy. Although the four mice we used as a reference population for Norway do not carry the haplotype, another mouse from Norway (with 75% *domesticus* ancestry genome-wide) is homozygous for it, suggesting that the haplotype arose prior to island colonization. Two of the SNPs in the 7-SNP haplotype are found in sequences with evidence of binding transcription factors (CEBPB and KLF4), as curated by ReMap (Hammal et al. 2022). The CEBPB transcription factor regulates a variety of functions (Ramji and Foka 2002), including fasting responses in the liver (Goldstein et al. 2017). The KLF4 transcription factor is involved in the epithelial–mesenchymal

transition in mammary epithelial cells (Tiwari et al. 2013). The SNPs in sequences bound by CEBPB and KLF4 are located 3,064 bp and 2,674 bp upstream of predicted gene Gm12817.

## Discussion

Mice living on the Faroe Islands have complex evolutionary histories. As suspected based on previous findings (Jones et al. 2011), they are products of hybridization between *domesticus* and *musculus*. Faroe Island mice are primarily *domesticus* in origin, probably reflecting a history of repeated backcrossing to *domesticus* following initial hybridization. The skew toward *domesticus* ancestry matches the morphological resemblance of Faroe Island mice to this subspecies (Jones et al. 2011). The hybrid nature of Faroe Island mice is likely connected to ancestors in Norway, where our genomic ancestry inferences support earlier claims of a hybridization gradient (Jones et al. 2010). Although firm conclusions await the application of methods that consider a range of hybridization histories, the overall similarity in junction patterns among island mice, as well as between island mice and some Norway mice, suggests that the hybridization visible in Faroe Island mouse genomes occurred prior to island colonization. Distributions of ancestry also show evidence of selection against hybrids and contributions to body size evolution. To reconstruct colonization histories, we identified best-fit models based on separate analyses of the ANGSD-derived 2D SFS in genomic regions with *domesticus* ancestry, the GATK-derived 2D SFS in genomic regions with *domesticus* ancestry, and the ancestral recombination graph across the genome. Assuming three generations per year (Berry et al. 2008), the ranges of point estimates across the three approaches indicate that mice colonized Nólsoy 697 to 1,071 years ago and Sandoy 1,058 to 1,804 years ago. The split time between Nólsoy and Sandoy is estimated to be 724 to 1,258 years ago. Collectively, these results suggest that mice colonized Sandoy first and Nólsoy later. Overall, the timing of colonization is consistent with passive transport of mice to the Faroe Islands by Norwegian Vikings (Edwards 2005; Jones et al. 2011). Mice on these islands likely had several hundred to a few thousand generations following colonization to evolve unusually large bodies.

Mice on the Faroe Islands appear to have persisted through substantial declines in effective population size ( $N_e$ ). Estimates of  $N_e$  associated with island colonization are as low as 2 and the island populations show low levels of nucleotide diversity. In populations with temporal variation in number,  $N_e$  is expected to be influenced disproportionately by intervals with small numbers. Consequently, low  $N_e$  values associated with colonization raise the prospect of heightened genetic drift and reduced selective efficacy during a substantial proportion of the time mice have

occupied the Faroe Islands. In turn, these inferences should motivate careful evaluation of the relative roles of natural selection and genetic drift in the evolution of extreme body size and other phenotypes in these mice.

We uncovered evidence for limited migration following island colonization. Best-fit models to the 2D SFS feature low rates of migration from Norway to Nólsoy (both ANGSD-derived SFS and GATK-derived SFS) and from Norway to Sandoy (ANGSD-derived SFS). Restrictions to gene flow may have facilitated the evolution of genetic and phenotypic differences between these groups (Berry et al. 1978 a, 1978b; Davis 1983; Jones et al. 2011). In addition, our results suggest that mice on the Faroe Islands and in Norway have a history of inbreeding. This inference matches the conclusion from a wider survey of house mice based on microarray genotypes, which uncovered widespread and variable evidence of inbreeding across populations (Morgan et al. 2022).

Our study underscores challenges that accompany evolutionary inferences from genome sequences in populations with complex demographic histories. The characterization of demographic history can depend on which methods are used to reconstruct the SFS. In our study, discordance between best-fit models from ANGSD-derived SFS versus GATK-derived SFS persisted despite decent sequencing coverage and the application of best practices for these popular approaches. Although explaining this discordance is beyond the scope of our study, we suggest a few potential causes. In both Nólsoy and Sandoy, the GATK-derived SFS contains a lower proportion of singletons than the ANGSD-derived SFS. This pattern is consistent with the under-calling of rare variants by GATK's joint genotyping framework, an effect detected in simulations of short-read sequences (at coverages lower than those in our study; Han et al. 2014). In Sandoy, the ANGSD-derived SFS contains a higher proportion of doubletons than the GATK-derived SFS. To investigate this discrepancy, we conducted leave-one-out analyses, removing one mouse at a time and then reconstructing the Sandoy SFS ( $n=11$  mice). These analyses revealed that a single mouse (FSa16) is largely responsible for the excess doubletons in the ANGSD-derived SFS (Fig. S5). Removing this mouse has no effect on the GATK-derived SFS, which exhibits greater stability across subsamples (Fig. S6). Although it is difficult to determine the cause of this finding (FSa16 has a sequencing coverage that is typical for Sandoy mice), it suggests that the Sandoy ANGSD-derived SFS is more sensitive to sample composition. ANGSD also assumes independence among sites in its estimation of the SFS, an assumption that could be violated by elevated linkage disequilibrium in island populations.

Disparities between GATK-derived SFS and ANGSD-derived SFS lead to differences in inferred demographic histories. To better understand this connection, we partitioned

2D SFS into the percentages of shared and private variants contributed by each pair of populations (Table S7). In both island–mainland comparisons (Nólsoy–Norway and Sandoy–Norway), there are higher percentages of private variants in ANGSD-derived 2D SFS than in GATK-derived 2D SFS, contributing to the inference of larger island  $N_e$  values and older split times with Norway. ANGSD-derived models also differ by the inclusion of one-way migration into Sandoy, which is consistent with the higher percentages of shared variants in both ANGSD-derived 2D SFS involving Sandoy. Although estimates of key parameters are broadly similar between the two methods, our results provide a reminder that uncertainty in the data should be incorporated during population genomic analyses (Johri et al. 2022).

Consideration of the model-based population genetic methods we used to reconstruct colonization histories also emphasizes the role of uncertainty. Some parameter estimates of best-fit models from analyses of the SFS feature nontrivial confidence intervals, and the ARG method implemented in *Relate* currently yields only point estimates. Furthermore, samples of wild populations can contain high incidences of close relatives. The modest sample sizes that remained after we excluded close relatives represent additional sources of error. Based on our experience, we recommend that population geneticists examine relatedness before proceeding to reconstruct demographic history.

Another lesson of general interest from our investigation concerns the goal of identifying genomic windows that contain beneficial mutations involved in adaptation. Scanning the local SFS in *domesticus* regions pinpointed many windows showing conventional evidence of positive selection even after accounting for demographic history and multiple testing. Interestingly, most of these windows harbor an excess of singletons contributed by a single mouse (with the identity of the mouse varying among windows). This pattern contrasts with expectations for a selective sweep, in which singletons should accumulate randomly across individuals following the loss of neutral variation caused by fixation of a linked beneficial variant. We confirmed that mice harboring clusters of singletons in these windows have high-confidence ancestry calls, arguing against the explanation that most of the clusters are caused by mis-inference of heterogenic mice as carrying homozygous *domesticus* ancestry. Whatever explains this constellation of singletons, it could be a source of false positives in selection scans that is currently unrecognized in the field and deserves further examination. We recommend that researchers using the SFS to search for positive selection inspect the partitioning of variants among individuals. Overall, our results underscore the barriers to localizing targets of adaptive evolution in the genomes of recently bottlenecked populations (Barton 2000; Poh et al. 2014), including populations that colonize islands.

The ARG offers strategies for characterizing positive selection that complement those based on traditional population genetic approaches. SNPs with low  $P$ -values in both Nólsoy and Sandoy given their local genealogies may reflect parallel adaptation of mice on these two islands. The enrichment of genes with neurological functions we detected raises the prospect that mice evolved new behaviors following their colonization of the Faroe Islands. Mice from Gough Island evolved increased exploration and boldness (Stratton et al. 2021, 2023), and genes in genomic regions showing evidence of positive selection in these mice are enriched for neurological functions (Payseur and Jing 2021). Still, gene ontology results from Faroe Island mice should be viewed with caution since they come from applying relaxed significance thresholds. By combining mice from Nólsoy and Sandoy, we found SNPs in three genomic regions that show strong statistical evidence for selection. The frequencies of the chromosome 4 7-SNP haplotype in island populations, along with the presence of the haplotype in Norway, suggests that it marks a partial selective sweep originating from variation that was transported to the islands.

## Materials and Methods

### Mice

Mice were collected on Sandoy (from five localities) during 2005 to 2006, and on the small islands of Mykines (October 2009) and Nólsoy (between 1984 and 2003). Because of the small sizes of the villages on Nólsoy and Mykines, sampling locations were less than 1 km apart and treated as a single locality. DNA from island mice was extracted from tail or ear clips. Tissues were lysed with lysis buffer with Tris-EDTA-SDS, SDS, and protein were precipitated using potassium acetate, and DNA was purified using magnetic beads. In Norway, mice were collected from five geographical districts from 2007 August 31 to October 13: Sogn og Fjordane, Hordaland, Sogn, Vestfold, and Telemark. DNA from Norway mice was extracted from muscle using the Qiagen Puregene kit.

### Genome Sequencing, Alignment, Variant Calling, and Filtering

We sequenced the genomes of 15 mice from Nólsoy, 20 mice from Sandoy, seven mice from Mykines, and 15 mice from Norway (seven from Telemark, two from Sogn og Fjordane, two from Hordaland, two from Sogn, and two from Vestfold). Sequencing was completed by the Biotechnology Center at the University of Wisconsin-Madison. Genomic DNA concentrations were verified using the Qubit dsDNA HS Assay Kit (Life Technologies, Grand Island, NY). Samples were prepared according to the Celero PCR Workflow with Enzymatic Fragmentation (Tecan Genomics, Redwood City,

CA). Quality and quantity of the finished libraries were assessed using an Agilent TapeStation (Agilent, Santa Clara, CA) and Qubit dsDNA HS Assay Kit, respectively. Paired end, 150 bp sequences were generated using the Illumina NovaSeq6000 (Illumina, San Diego, CA). Sequencing reads from genome sequences for eight mice from Germany, eight mice from France, eight mice from Czech Republic, and eight mice from Kazakhstan (Harr et al. 2016) were downloaded from: <https://www.ebi.ac.uk/ena/browser/view/ERA441366>; <https://www.ebi.ac.uk/ena/browser/view/ERA532416>.

Sequencing reads were aligned to the mouse mm10 genome reference sequence downloaded from the UCSC Genome Browser (Kent et al. 2002), using bwa-mem (Li and Durbin 2009) with default settings. BAM files were sorted and merged using Samtools (Danecek et al. 2021). Duplicates were marked and removed with Picard tools (<https://broadinstitute.github.io/picard/>). Indel realignment was performed using IndelRealigner in GATK. Variants were identified and genotyped using GATK (McKenna et al. 2010). Average realized sequencing coverage was close to the target of 20x (Table S8). GVCF files generated by HaplotypeCaller (GATK) were merged into a joint VCF file using GenotypeGVCFs (GATK). SNPs were processed with VariantFiltration using filtering criteria provided by GATK guidelines for best practices, including the following cutoffs: QualByDepth (QD)  $\geq 2.0$ ; FisherStrand (FS)  $\leq 60.0$ ; StrandOddsRatio (SOR)  $\leq 3.0$ ; RMSMappingQuality (MQ)  $\geq 40.0$ ; MappingQualityRankSumTest (MQRankSum)  $\geq -12.5$ ; ReadPosRankSumTest (ReadPosRankSum)  $\geq -8.0$ . We expected that combining individuals from populations with divergent allele frequencies into a single cohort could lead to poorly calibrated QUAL scores during genotyping (Altinkaya et al. 2025). To mitigate this issue, genotypes were called separately for each population.

Close relatives were identified by pairwise identity-by-state comparisons with KING (Manichaikul et al. 2010) using the “kinship” option on all SNPs that passed GATK filters (results using SNPs pruned for linkage disequilibrium were similar). Prior to downstream analyses, mice inferred to be close relatives were removed. For each population, we chose unrelated mice by starting with the set selected using the KING “unrelated” option (which removed first-degree and second-degree relatives), adjusting it to leave no third-degree relatives and to preferentially remove the mouse with lower sequencing coverage from each relative pair. This approach left eight mice from Nólsoy, 12 mice from Sandoy, and 13 mice from Norway (seven from Telemark, two from Hordaland, two from Sogn, one from Sogn og Fjordane, and one from Vestfold). Following the same procedure left unrelated samples of eight mice from Germany, four from France, seven from Czech Republic, and four from Kazakhstan. All seven mice from Mykines were inferred to be first-degree relatives. As a result, we did not attempt to reconstruct colonization history for

Mykines, though we inferred the subspecific ancestry of a single mouse from Mykines. Subsequent analyses only considered these samples of unrelated mice; genotypes were recalled after closely related mice were removed.

Genotypes were merged across populations with BCFtools (Danecek et al. 2021). Positions with missing genotypes were identified in each population and filled in using the “all-sites” option in GATK. SNPs within repetitive sequences and segmental duplications annotated for the reference genome were removed with Bedtools (Quinlan and Hall 2010) “subtract” using positions from “RepeatMasker” and “Segmental Dups” tracks downloaded from the UCSC Genome Browser.

All analyses focused on SNPs located on the autosomes.

### Inference of Genetic Differentiation and Ancestry

Genetic differentiation between mice from different locations was visually examined using principal component analysis (PCA) of the genotype covariance matrix, as implemented in SmartPCA (Price et al. 2006). Prior to PCA, SNPs with minor allele frequency less than 0.05, greater than 50% missing genotypes, and/or pairwise  $R^2$  higher than 0.2 were removed.

We used AncestryHMM (Corbett-Detig and Nielsen 2017) to probabilistically assign subspecific ancestry (homozygous *domesticus*, homozygous *musculus*, and heterozygous *domesticus/musculus*) along the genome of individual mice from Nólsoy, Sandoy, Mykines, and Norway. We computed recombination distances between pairs of adjacent SNPs by interpolation from their physical distances and local recombination rates taken from the mouse genetic map constructed from the heterogeneous stock (Cox et al. 2009; downloaded at <https://github.com/kbroman/CoxMapV3>). We treated mice from France and Germany as *domesticus* reference populations and treated mice from Czech Republic and Kazakhstan as *musculus* reference populations. We ran separate sets of AncestryHMM analyses, treating different intersubspecific pairs of populations as references (France–Czech, France–Kazakhstan, Germany–Czech, and Germany–Kazakhstan). To improve the quality of inferred ancestry calls, we used two conservative approaches. First, we only considered SNPs for which one of the three possible ancestries was assigned a posterior probability  $> 0.95$ . Second, we required that ancestry calls agreed across analyses using each of the four pairs of reference populations.

We created plots to visualize ancestry along each chromosome. We also used a few metrics to summarize inferred patterns of ancestry. For each mouse, we computed the proportion of each chromosome with *domesticus* ancestry. Heterogenicity was calculated as the proportion of each chromosome in each mouse inferred to harbor ancestry from both *domesticus* and *musculus* (i.e. heterozygosity

for ancestry). We counted “junctions” (Fisher 1949; Baird 2006; Janzen et al. 2018) as inferred transitions from one ancestry to another along a chromosome. With unphased data, we considered the locations of junctions in diploid space, but we could still count junctions in a pseudohaploid manner. A location where diploid ancestry changes from homozygous (*domesticus* or *musculus*) to heterogenic requires the presence of one junction on one of the two chromosomes. A location where diploid ancestry changes from homozygous *domesticus* to homozygous *musculus* (or the reverse) requires two junctions, one on each chromosome. To be conservative, we required that junctions be identified across analyses using all four pairs of reference populations and we ignored junctions supported by ancestry at single SNPs.

## Reconstruction of Island Colonization

### The Two-Population Site Frequency Spectrum

Our first approach to reconstructing island colonization focused on analyses of the folded, 2D site frequency spectrum (SFS) for each pair of populations. We did not pursue three-population models because the 3D SFS was too sparse with our sample sizes after close relatives were removed. We separately analyzed the 2D SFS for Norway–Sandoy, Norway–Nólsoy, and Sandoy–Nólsoy. To mitigate the effects of hybridization on the SFS (Blanckaert and Payseur 2021), we focused on genomic regions with *domesticus* ancestry and treated four mice from Norway with high proportions (>0.9) of *domesticus* ancestry as a mainland reference population. We considered putatively neutral SNPs found in regions with the following characteristics: (i) shared *domesticus* ancestry across the three populations, (ii) at least 25 kb away from genes, and (iii) recombination rates higher than the genomic average (>0.5 cM/Mb; Cox et al. 2009). Windows were chosen to be at least 5 kb from each other. SNPs were taken from windows ranging from 500 bp to 10 kb in size. The summed genomic length considered was 15,312,841 bp. For each population pair, the 2D SFS was separately generated twice, using two independent approaches: either (i) GATK-called genotypes or (ii) ANGSD (Korneliussen et al. 2014) with options “-minMapQ 30 -minQ 30 -baq 1”.

To estimate demographic parameters from the 2D SFS, we used the maximum likelihood framework implemented in Diffusion Approximation for Demographic Inference (*dadi*; Gutenkunst et al. 2009) version 2.2.0. This approach employs a diffusion approximation to compute the expected SFS, assuming a specified model of population history. We considered a series of two-population models motivated by island colonization. For mainland–island analyses (Norway–Sandoy and Norway–Nólsoy), potential model parameters included an ancestral effective population size ( $N_e$ ), an instantaneous change in the ancestral  $N_e$ , divergence of the

island population from the mainland accompanied by an instantaneous change in island  $N_e$ , a constant mainland  $N_e$ , an additional instantaneous change in island  $N_e$  following its split for the mainland, and migration between the island and the mainland. Population splits and size changes were each associated with an additional parameter that reflects the timing of the event.

For the island–island analyses (Sandoy–Nólsoy), the framework was similar, except both populations split directly from an ancestral population and could undergo an additional size change following their split. For all models, we allowed for historical inbreeding in both island and mainland populations (Blischak et al. 2020).

To measure model fit and estimate parameters, we ran many replicate analyses for each model/dataset combination. The set of starting parameter values for each replicate was drawn randomly from log-uniform prior distributions with the following ranges (in units of the *dadi* reference population size, which defaults to 1; times are given in units of  $2N_{e,ref}$  generations, whereas migration rates are in units of  $2N_{e,ref}m_{ij}$ ):  $10^{-4}$  to 10 for  $N_e$ ;  $10^{-5}$  to 10 for event times;  $10^{-6}$  to 20 for migration rates, and  $10^{-7}$  to 1 for inbreeding coefficients. We initially ran 1,000 replicates using parameter bounds recommended by the *dadi* manual. Resulting likelihoods suggested that some bounds should be expanded. We subsequently ran 10,000 replicates using the following bounds:  $10^{-5}$  to 20 for population sizes;  $10^{-6}$  to 10 for event times; 0 to 100 for migration rates; and 0 to 1 for inbreeding coefficients. If any of the parameter estimates for the top five replicates with highest likelihoods were not in the same order of magnitude, more replicates were run until all parameter estimates converged in order of magnitude. Inferences were based on results from these 10,000 or more replicates, with the best-fit set of parameter values taken to be those in the replicate with the highest log-likelihood. Parameter estimates were rescaled to units of interest (e.g. time in generations) following recommendations in the *dadi* manual. To identify best-fit models, we compared log-likelihoods among nested models through adjusted likelihood ratio tests (Coffman et al. 2016). Likelihood ratio tests comparing all models evaluated are available in Table S9. We also visually compared plots of residuals between observed and expected 2D SFS among all models. Parameter uncertainty for best-fit models was quantified using the Godambe Information Matrix (Coffman et al. 2016) with 200 bootstrap replicates.

Recognizing the potential for uncertainty in reconstructing the SFS, we performed separate demographic analyses that used either GATK-derived 2D SFS or ANGSD-derived 2D SFS as input datasets.

### The Ancestral Recombination Graph

Our second approach to reconstructing island colonization focused on the ancestral recombination graph (ARG).

We used the ARG-based method *Relate* v1.2.2 (Speidel et al. 2019) to infer historical changes in  $N_e$  and to estimate the split times between populations. *Relate* infers topological features of local genealogies based on position-specific distance matrices between haplotypes. Following the reconstruction of local tree topology, *Relate* estimates branch lengths using a Markov chain Monte Carlo algorithm, initially assuming constant  $N_e$ . Estimated branch lengths are then used to re-estimate the historical trajectory of  $N_e$  as the inverse of the average coalescence rate during each epoch, enabling the re-estimation of branch lengths under the inferred demography. This process is iteratively repeated to coestimate branch lengths and the trajectory of  $N_e$ .

We conducted this set of analyses with the same collection of unrelated mice from Nólsoy, Sandoy, and Norway used for SFS-based demographic inference. We included all autosomal sequence information, regardless of inferred subspecific ancestry. To make inferences from the ARG more directly comparable to inferences from the SFS, we conducted a separate ARG analysis for each pair of populations: Norway–Sandoy, Norway–Nólsoy, and Sandoy–Nólsoy. *Relate* requires that SNP alleles be polarized as ancestral or derived and haplotypes be phased. To polarize GATK-filtered SNPs, we examined genotype calls from genome sequences of two outgroups, *Mus caroli* (Thybert et al. 2018) and *Mus spretus* (Keane et al. 2011), aligned to the mouse reference genome sequence. For each species, we downloaded conversion files between genome assemblies from the UCSC Genome Browser and used UCSC “LiftOver” to find the positions of variants. To be conservative, we required that these two outgroups harbored the same allele, which we inferred to be the ancestral allele. SNPs that did not meet this criterion were eliminated. Haplotypes were phased separately for each population by passing polarized SNP calls to *Beagle* v5.4 (Browning et al. 2021). We assigned SNPs positions on the genetic map using linear interpolation from the mouse reference genetic map (Cox et al. 2009). Phased VCF files were merged and converted to the hap/sample format using *RelateFileFormats* –mode ConvertFromVcf. The haps, sample, and genetic map files were passed to *Relate* v1.2.2 (Speidel et al. 2019) to perform genome-wide genealogy estimation assuming a mutation rate of  $5 \times 10^{-9}$  (per site, per generation) and an initial  $N_e$  of 20,000. Historical  $N_e$  values were calculated from estimated coalescence rates as  $\frac{1}{2}$  divided by the coalescence rate. Coalescence rates were calculated from the genealogy results using the “EstimatePopulationSize” module in *Relate* with “year\_per\_generation” = 0.5. Population split times were inferred as the times when the inverse of the cross-population coalescence rates were the smallest (Kingman 1982). One hundred replicates initiated with different seed numbers were collected to assess convergence.

### Genomic Scans for Selection

To locate genomic regions with evidence for positive directional selection in Faroe Island mice, we used three strategies. First, we searched for genomic regions showing a skew toward rare alleles in the single-population SFS or an elevation in differentiation between mainland–island pairs of populations that were unusual under the inferred demographic history. To match assumptions of the neutral demographic models we reconstructed, we restricted these analyses to genomic regions with shared *domesticus* ancestry (again excluding those masked for repeats and segmental duplications). We separated 5 kb windows into four bins based on the number of accessible (unfiltered) sites: 1–2 kb, 2–3 kb, 3–4 kb, and 4–5 kb. Windows smaller than 1 kb were discarded. We used *scikit-allel* v1.3.6 (<https://github.com/cggh/scikit-allel>) to calculate summary statistics of variation. We computed nucleotide diversity (Tajima 1983) and Tajima’s D (Tajima 1989) separately for Nólsoy, Sandoy, and Norway in these windows. We computed  $F_{ST}$  (Hudson et al. 1992) for Norway–Sandoy and Norway–Nólsoy in these windows. For each bin and summary statistic, we counted the number of windows with non-NA values and noninfinite values. To reconstruct the null distribution of Tajima’s D and  $F_{ST}$  expected under neutrality, we used *msprime* (Baumdicker et al. 2022) to conduct neutral simulations under the demographic models that provided the best fits to the GATK-derived SFS (from *dadl*) without inbreeding. We assumed the mutation rate and recombination rate were  $5 \times 10^{-9}$  per site per generation. We simulated ten times the number of windows, separately for each bin and summary statistic. We treated the Tajima’s D values and  $F_{ST}$  values resulting from these simulations as null distributions for genomic scans for selection. For each window in the observed dataset, we computed a *P*-value for Tajima’s D as the proportion of simulated windows with values more negative than the observed value. For each window in the observed dataset, we computed a *P*-value for  $F_{ST}$  as the proportion of simulated windows with values higher than the observed value. Windows with *P*-values less than 0.05 divided by the number of tested windows were declared significant after a Bonferroni correction. This conservative approach accounts for the characteristics of the dataset, demographic history, and the performance of many tests along the genome.

In a second strategy to search the genome for positive selection, we compared the local SFS to the genomic SFS, using the method implemented in *Sweepfinder2* (DeGiorgio et al. 2016). For this analysis, we considered SNPs in regions of shared *domesticus* ancestry (again excluding those masked for repeats and segmental duplications). Nólsoy, Sandoy, and Norway were analyzed separately using the same grids with 20 kb spacing to enable direct comparisons. The background SFS was computed using all autosomal SNPs for each

population. Recombination rates were again estimated by linear extrapolation from the mouse reference genetic map (Cox et al. 2009). We used the following *SweepFinder2* command: “SweepFinder2 -lru grids\_chrNum.txt popName\_chrNum\_sf2\_alleleFreq.txt popName\_SpectFile popName\_chrNum\_sf2\_recMap.txt sf2\_grids\_output\_popName\_chrNum.txt”. For each window, *SweepFinder2* computes a likelihood ratio test that compares the fit of a model allowing the local SFS to differ from the background SFS to the fit of a model assuming the local SFS is drawn from the background SFS. To reconstruct null distributions of the likelihood ratio test statistic, we used *msprime* to separately simulate neutral datasets for each pair of mainland and island populations (Norway–Sandoy and Norway–Nólsoy) following the demographic models that provided the best fits to the GATK-derived SFS without inbreeding. We simulated 20,000 replicates of a 1 Mb sequence with mutation rate and recombination rate set to  $5 \times 10^{-9}$  per site per generation. The background SFS was reconstructed from all 20,000 replicates. We conducted a *SweepFinder2* scan on each of the 20,000 simulated datasets, using a grid size of 20 kb with windows centered at the grid points and treating the full SFS across all simulations as the background SFS. For each 20 kb window in the observed dataset, we compared the likelihood ratio test statistic to the simulated distribution of this statistic to compute a *P*-value under the null hypothesis of neutrality. Windows with *P*-values less than 0.05 divided by the number of tested windows were declared significant after a Bonferroni correction. This conservative approach again accounts for the characteristics of the dataset, demographic history, and the performance of many tests along the genome.

Our third tactic for scanning the genome for positive selection focused on the ARG. For this analysis, we separately considered Nólsoy, Sandoy, and Norway, using the set of polarized, GATK-filtered SNPs that were used in the demographic reconstruction by *Relate*. The set of variants included in this ARG analysis was much larger than the sets of variants included in the other two genomic scan approaches, since it did not require that we filter for subspecific ancestry. *Relate* evaluates the evidence for selection at a SNP by estimating how quickly the lineage carrying the mutation spreads through the population relative to other lineages (Speidel et al. 2019). The program treats the number of present-day carriers of the mutation as a test statistic, conditioning on the number of lineages present when the mutation first enters the population. The null distribution of the test statistic is computed analytically and is claimed to be robust to changes in population size (Speidel et al. 2019). Our selection scans followed recommendations in the *Relate* user’s manual. Population size history was estimated using the setting `–threshold 0`, so that the branch lengths of all trees were updated for the estimated population size history. SNPs with low derived allele frequencies were filtered out

using `RelateFileFormats –mode GenerateSNPAnnotations` to annotate .mut files. *Relate* provides a *P*-value for each SNP under the null hypothesis of neutrality. We used a Bonferroni adjustment to account for multiple testing, considering the number of SNPs or the estimated number of trees as the number of tests in separate corrections. To explore evidence for parallel evolution across island populations, we identified the subset of SNPs with uncorrected  $P < 0.05$  in both Nólsoy and Sandoy. To search for enrichment of functions among genes in which these SNPs were located, we conducted a gene ontology analysis using g:Profiler (<https://biit.cs.ut.ee/gprofiler/gost>). To increase our power to identify island-specific positive selection, we repeated the selection tests, treating mice from Nólsoy and Sandoy as a single population.

## Supplementary Material

Supplementary material is available at *Genome Biology and Evolution* online.

## Acknowledgments

We thank the following individuals for assistance with field collections: Emilie Hardouin and the Johannessens on Mykines, Noomi Gregersen, and Heidi S. Hansen in the Faroe Islands, and Jeroen van der Kooij in Norway. We thank Mark Nolte for assistance with DNA extractions. Members of the Payseur lab provided helpful input during the project.

## Author Contributions

B.A.P. designed the study, with assistance from P.J., M.E.F., and E.K.H. E.P.J., E.M., J.-K.M., Y.F.C., and J.B.S. acquired samples used for the study. P.J. conducted the research, with contributions from M.E.F. and E.K.H. and supervision from B.A.P. B.A.P. wrote the manuscript, with substantive input from E.K.H., P.J., J.B.S., and M.E.F.

## Funding

This research was supported by National Institutes of Health grants R35GM139412 and R01GM100426 to B.A.P. M.E.F. and E.K.H. were partially supported by the National Institutes of Health Graduate Training Grant in Genetics at the University of Wisconsin-Madison (T32GM007133). Computational analyses were conducted using resources provided by the University of Wisconsin-Madison’s Center for High Throughput Computing.

## Data Availability

Raw sequencing reads are available from NCBI’s Sequence Read Archive under the BioProject accession

PRJNA1233310. Variant and summary statistics files are available on Dryad (<https://doi.org/10.5061/dryad.5qftdz4>). Scripts are provided on GitHub ([https://github.com/PayseurLabUWMadison/Faroe\\_Demography/tree/main](https://github.com/PayseurLabUWMadison/Faroe_Demography/tree/main)).

## Literature Cited

- Altinkaya I, Nielsen R, Korneliusen TS. vcfGL: a flexible genotype likelihood simulator for VCF/BCF files. *Bioinformatics*. 2025;41:btaf098. <https://doi.org/10.1093/bioinformatics/btaf098>.
- Angel A, Wanless RM, Cooper J. Review of impacts of the introduced house mouse on islands in the Southern Ocean: are mice equivalent to rats? *Biol Invasions*. 2009;11:1743–1754. <https://doi.org/10.1007/s10530-008-9401-4>.
- Baird SJ. Phylogenetics: Fisher's markers of admixture. *Heredity* (Edinb). 2006;97:81–83. <https://doi.org/10.1038/sj.hdy.6800850>.
- Barton NH. Genetic hitchhiking. *Philos Trans R Soc Lond B Biol Sci*. 2000;355:1553–1562. <https://doi.org/10.1098/rstb.2000.0716>.
- Baumdicker F, et al. Efficient ancestry and mutation simulation with msprime 1.0. *Genetics*. 2022;220:iyab229. <https://doi.org/10.1093/genetics/iwab229>.
- Berry R, Tattersall F, Hurst J. House mouse *Mus domesticus*. In: Harris S, Yalden D, editors. *Mammals of the British Isles*. Handbook. 4th ed. The Mammal Society; 2008. p. 141–149.
- Berry RJ. The evolution of an island population of house mouse. *Evolution*. 1964;18:468–483. <https://doi.org/10.2307/2406357>.
- Berry RJ. Small mammal differentiation on islands. *Philos Trans R Soc Lond B Biol Sci*. 1996;351:753–764. <https://doi.org/10.1098/rstb.1996.0070>.
- Berry RJ. Evolution rampant: house mice on Madeira. *Mol Ecol*. 2009;18:4344–4346. <https://doi.org/10.1111/j.1365-294X.2009.04345.x>.
- Berry RJ, Bonner WN, Peters J. Natural selection in house mice *Mus musculus* from South Georgia (South Atlantic Ocean). *J Zool*. 1979;185:385–398. <https://doi.org/10.1111/j.1469-7998.1979.tb03970.x>.
- Berry RJ, Jakobson ME. Adaptation and adaptability in wild-living house mice. *J Zool Lond*. 1975;176:391–402. <https://doi.org/10.1111/j.1469-7998.1975.tb03210.x>.
- Berry RJ, Jakobson ME, Peters J. The house mice of the Faroe Islands: a study in microdifferentiation. *J Zool*. 1978b;185:73–92. <https://doi.org/10.1111/j.1469-7998.1978.tb03314.x>.
- Berry RJ, Peters J. Macquarie island house mice: a genetical isolate on a sub-Antarctic island. *J Zool*. 1975;176:375–389. <https://doi.org/10.1111/j.1469-7998.1975.tb03209.x>.
- Berry RJ, Peters J. Heterogeneous heterozygosities in *Mus musculus* populations. *Proc R Soc B*. 1977;197:485–503. <https://doi.org/10.1098/rspb.1977.0082>.
- Berry RJ, Peters J, Aarde RVJ. Sub-Antarctic house mice: colonization, survival and selection. *J Zool*. 1978a;184:127–141. <https://doi.org/10.1111/j.1469-7998.1978.tb03270.x>.
- Blanckaert A, Payseur BA. Finding hybrid incompatibilities using genome sequences from hybrid populations. *Mol Biol Evol*. 2021;38:4616–4627. <https://doi.org/10.1093/molbev/msab168>.
- Blischak PD, Barker MS, Gutenkunst RN. Inferring the demographic history of inbred species from genome-wide SNP frequency data. *Mol Biol Evol*. 2020;37:2124–2136. <https://doi.org/10.1093/molbev/msaa042>.
- Braverman JM, Hudson RR, Kaplan NL, Langley CH, Stephan W. The hitchhiking effect on the site frequency spectrum of DNA polymorphisms. *Genetics*. 1995;140:783–796. <https://doi.org/10.1093/genetics/140.2.783>.
- Browning BL, Tian X, Zhou Y, Browning SR. Fast two-stage phasing of large-scale sequence data. *Am J Hum Genet*. 2021;108:1880–1890. <https://doi.org/10.1016/j.ajhg.2021.08.005>.
- Clarke WE. On some forms of *Mus musculus*, Linn., with description of a new subspecies from the Faroe Islands. *Proc R Phys Soc Edin*. 1904;15:160–167.
- Coffman AJ, Hsieh PH, Gravel S, Gutenkunst RN. Computationally efficient composite likelihood statistics for demographic inference. *Mol Biol Evol*. 2016;33:591–593. <https://doi.org/10.1093/molbev/msv255>.
- Corbett-Detig R, Nielsen R. A hidden Markov model approach for simultaneously estimating local ancestry and admixture time using next generation sequence data in samples of arbitrary ploidy. *PLoS Genet*. 2017;13:e1006529. <https://doi.org/10.1371/journal.pgen.1006529>.
- Cox A, et al. A new standard genetic map for the laboratory mouse. *Genetics*. 2009;182:1335–1344. <https://doi.org/10.1534/genetics.109.105486>.
- Danecek P, et al. Twelve years of SAMtools and BCFtools. *Gigascience*. 2021;10:giab008. <https://doi.org/10.1093/gigascience/giab008>.
- Davis SJM. Morphometric variation of populations of house mice *Mus domesticus* in Britain and Faroe. *J Zool*. 1983;199:521–534. <https://doi.org/10.1111/j.1469-7998.1983.tb05103.x>.
- Debes LJ. Faeroae, and Faeroae Referata: that is, a description of the islands and inhabitants of Faeroe. William Hes; 1676. English translation.
- Degerbøl M. Mammalia. In: Jensen AS, et al., editors. *Zoology of the Faroes*. Horst; 1942. p. 1–133.
- DeGiorgio M, Huber CD, Hubisz MJ, Hellmann I, Nielsen R. SweepFinder2: increased sensitivity, robustness and flexibility. *Bioinformatics*. 2016;32:1895–1897. <https://doi.org/10.1093/bioinformatics/btw051>.
- Edwards KJ. 'On the windy edge of nothing': a historical human ecology of the Faroe Islands. *Hum Ecol*. 2005;33:585–596. <https://doi.org/10.1007/s10745-005-7678-9>.
- Estoup A, Guillemaud T. Reconstructing routes of invasion using genetic data: why, how and so what? *Mol Ecol*. 2010;19:4113–4130. <https://doi.org/10.1111/j.1365-294X.2010.04773.x>.
- Fisher RA. The theory of inbreeding. Oliver and Boyd; 1949.
- Forejt J, Ivanyi P. Genetic studies on male sterility of hybrids between laboratory and wild mice (*Mus musculus* L.). *Genet Res*. 1974;24:189–206. <https://doi.org/10.1017/S0016672300015214>.
- Gabriel SI, Jóhannesdóttir F, Jones EP, Searle JB. Colonization, mouse-style. *BMC Biol*. 2010;8:131. <https://doi.org/10.1186/1741-7007-8-131>.
- Goldstein I, et al. Transcription factor assisted loading and enhancer dynamics dictate the hepatic fasting response. *Genome Res*. 2017;27:427–439. <https://doi.org/10.1101/gr.212175.116>.
- Good JM, Handel MA, Nachman MW. Asymmetry and polymorphism of hybrid male sterility during the early stages of speciation in house mice. *Evolution*. 2008;62:50–65. <https://doi.org/10.1111/j.1558-5646.2007.00257.x>.
- Gutenkunst RN, Hernandez RD, Williamson SH, Bustamante CD. Inferring the joint demographic history of multiple populations from multidimensional SNP frequency data. *PLoS Genet*. 2009;5:e1000695. <https://doi.org/10.1371/journal.pgen.1000695>.
- Hammal F, de Langen P, Bergon A, Lopez F, Ballester B. Remap 2022: a database of human, mouse, Drosophila and Arabidopsis regulatory regions from an integrative analysis of DNA-binding sequencing experiments. *Nucleic Acids Res*. 2022;50:D316–D325. <https://doi.org/10.1093/nar/gkab996>.
- Han E, Sinsheimer JS, Novembre J. Characterizing bias in population genetic inferences from low-coverage sequencing data. *Mol Biol Evol*. 2014;31:723–735. <https://doi.org/10.1093/molbev/mst229>.

- Harr B, et al. Genomic resources for wild populations of the house mouse, *Mus musculus* and its close relative *Mus spretus*. *Sci Data*. 2016;3:160075. <https://doi.org/10.1038/sdata.2016.75>.
- Hudson RR, Slatkin M, Maddison WP. Estimation of levels of gene flow from DNA sequence data. *Genetics*. 1992;132:583–589. <https://doi.org/10.1093/genetics/132.2.583>.
- Huxley JS. *Evolution, the modern synthesis*. Allen and Unwin; 1942.
- Janzen T, Nolte AW, Traulsen A. The breakdown of genomic ancestry blocks in hybrid lineages given a finite number of recombination sites. *Evolution*. 2018;72:735–750. <https://doi.org/10.1111/evo.13436>.
- Johri P, et al. Recommendations for improving statistical inference in population genomics. *PLoS Biol*. 2022;20:e3001669. <https://doi.org/10.1371/journal.pbio.3001669>.
- Jones EP, et al. A molecular characterization of the charismatic Faroe house mouse. *Biol J Linn Soc*. 2011;102:471–482. <https://doi.org/10.1111/j.1095-8312.2010.01597.x>.
- Jones EP, et al. Fellow travellers: a concordance of colonization patterns between mice and men in the North Atlantic region. *BMC Evol Biol*. 2012;12:35. <https://doi.org/10.1186/1471-2148-12-35>.
- Jones EP, Eager HM, Gabriel SI, Jóhannesdóttir F, Searle JB. Genetic tracking of mice and other bioproxies to infer human history. *Trends Genet*. 2013;29:298–308. <https://doi.org/10.1016/j.tig.2012.11.011>.
- Jones EP, Van Der Kooij J, Solheim R, Searle JB. Norwegian house mice (*Mus musculus musculus/domesticus*): distributions, routes of colonization and patterns of hybridization. *Mol Ecol*. 2010;19:5252–5264. <https://doi.org/10.1111/j.1365-294X.2010.04874.x>.
- Keane TM, et al. Mouse genomic variation and its effect on phenotypes and gene regulation. *Nature*. 2011;477:289–294. <https://doi.org/10.1038/nature10413>.
- Kent WJ, et al. The human genome browser at UCSC. *Genome Res*. 2002;12:996–1006. <https://doi.org/10.1101/gr.229102>.
- Kingman JFC. On the genealogy of large populations. *J Appl Prob*. 1982;19:27–43. <https://doi.org/10.2307/3213548>.
- Korneliusson TS, Albrechtsen A, Nielsen R. ANGSD: analysis of next generation sequencing data. *BMC Bioinformatics*. 2014;15:356. <https://doi.org/10.1186/s12859-014-0356-4>.
- Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*. 2009;25:1754–1760. <https://doi.org/10.1093/bioinformatics/btp324>.
- Losos JB, Ricklefs RE. Adaptation and diversification on islands. *Nature*. 2009;457:830–836. <https://doi.org/10.1038/nature07893>.
- Manichaikul A, et al. Robust relationship inference in genome-wide association studies. *Bioinformatics*. 2010;26:2867–2873. <https://doi.org/10.1093/bioinformatics/btq559>.
- McKenna A, et al. The genome analysis toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res*. 2010;20:1297–1303. <https://doi.org/10.1101/gr.107524.110>.
- Miller GS. *Catalogue of the mammals of Western Europe*. British Museum (Natural History); 1912.
- Moran BM, et al. The genomic consequences of hybridization. *Elife*. 2021;10:e69016. <https://doi.org/10.7554/eLife.69016>.
- Morgan AP, et al. Population structure and inbreeding in wild house mice (*Mus musculus*) at different geographic scales. *Heredity (Edinb)*. 2022;129:183–194. <https://doi.org/10.1038/s41437-022-00551-z>.
- Payseur BA, Jing P. Genomic targets of positive selection in giant mice from Gough Island. *Mol Biol Evol*. 2021;38:911–926. <https://doi.org/10.1093/molbev/msaa255>.
- Poh YP, Domingues VS, Hoekstra HE, Jensen JD. On the prospect of identifying adaptive loci in recently bottlenecked populations. *PLoS One*. 2014;9:e110579. <https://doi.org/10.1371/journal.pone.0110579>.
- Price AL, et al. Principal components analysis corrects for stratification in genome-wide association studies. *Nat Genet*. 2006;38:904–909. <https://doi.org/10.1038/ng1847>.
- Quinlan AR, Hall IM. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics*. 2010;26:841–842. <https://doi.org/10.1093/bioinformatics/btq033>.
- Ramji DP, Foka P. CCAAT/enhancer-binding proteins: structure, function and regulation. *Biochem J*. 2002;365:561–575. <https://doi.org/10.1042/bj20020508>.
- Rowe-Rowe DT, Crafford JE. Density, body size, and reproduction of feral house mice on Gough Island. *S Afr J Zoology*. 1992;27:1–5. <https://doi.org/10.1080/02541858.1992.11448252>.
- Schumer M, et al. Natural selection interacts with recombination to shape the evolution of hybrid genomes. *Science*. 2018;360:656–660. <https://doi.org/10.1126/science.aar3684>.
- Speidel L, Forest M, Shi S, Myers SR. A method for genome-wide genealogy estimation for thousands of samples. *Nat Genet*. 2019;51:1321–1329. <https://doi.org/10.1038/s41588-019-0484-x>.
- Stratton JA, Nolte MJ, Payseur BA. Evolution of boldness and exploratory behavior in giant mice from Gough Island. *Behav Ecol Sociobiol*. 2021;75:65. <https://doi.org/10.1007/s00265-021-03003-6>.
- Stratton JA, Nolte MJ, Payseur BA. Genetics of behavioural evolution in giant mice from Gough Island. *Proc Biol Sci*. 2023;290:20222603. <https://doi.org/10.1098/rspb.2022.2603>.
- Tajima F. Evolutionary relationship of DNA sequences in finite populations. *Genetics*. 1983;105:437–460. <https://doi.org/10.1093/genetics/105.2.437>.
- Tajima F. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics*. 1989;123:585–595. <https://doi.org/10.1093/genetics/123.3.585>.
- Thybert D, et al. Repeat associated mechanisms of genome evolution and function revealed by the *Mus caroli* and *Mus pahari* genomes. *Genome Res*. 2018;28:448–459. <https://doi.org/10.1101/gr.234096.117>.
- Tiwari N, et al. Klf4 is a transcriptional regulator of genes critical for EMT, including Jnk1 (Mapk8). *PLoS One*. 2013;8:e57329. <https://doi.org/10.1371/journal.pone.0057329>.
- Turner LM, Schwahn DJ, Harr B. Reduced male fertility is common but highly variable in form and severity in a natural house mouse hybrid zone. *Evolution*. 2012;66:443–458. <https://doi.org/10.1111/j.1558-5646.2011.01445.x>.
- White MA, Steffy B, Wiltshire T, Payseur BA. Genetic dissection of a key reproductive barrier between nascent species of house mice. *Genetics*. 2011;189:289–304. <https://doi.org/10.1534/genetics.111.129171>.
- Wilches R, Beluch WH, McConnell E, Tautz D, Chan YF. Independent evolution toward larger body size in the distinctive Faroe Island mice. *G3 (Bethesda)*. 2021;11:jkaa051. <https://doi.org/10.1093/g3journal/jkaa051>.

Associate editor: Aida Andres