

# Population History Across Timescales in an Urban Archipelago

Emma K. Howell <sup>1,\*</sup>, Lauren E. Nolfo-Clements <sup>2</sup>, Bret A. Payseur <sup>1,\*</sup>

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

<sup>2</sup>Framingham State University, Framingham, MA, USA

\*Corresponding authors: E-mail: ekhowell@wisc.edu; bret.payseur@wisc.edu.

Accepted: March 09, 2025

## Abstract

Contemporary patterns of genetic variation reflect the cumulative history of a population. Population splitting, migration, and changes in population size leave genomic signals that enable their characterization. Existing methods aimed at reconstructing these features of demographic history are often restricted in their temporal resolution, leaving gaps about how basic evolutionary parameters change over time. To illustrate the prospects for extracting insights about dynamic population histories, we turn to a system that has undergone dramatic changes on both geological and contemporary timescales—an urbanized, near-shore archipelago. Using whole genome sequences, we employed both common and novel summaries of variation to infer the demographic history of three populations of endemic white-footed mice (*Peromyscus leucopus*) in Massachusetts' Boston Harbor. We find informative contrasts among the inferences drawn from these distinct patterns of diversity. While demographic models that fit the joint site frequency spectrum (jSFS) coincide with the known geological history of the Boston Harbor, patterns of linkage disequilibrium reveal collapses in population size on contemporary timescales that are not recovered by our jSFS-derived models. Historical migration between populations is also absent from best-fitting models for the jSFS, but rare variants show unusual clustering along the genome within individual mice, a novel pattern that is reproduced by simulations of recent migration. Together, our findings indicate that these urban archipelago populations have been shaped by both ancient geological processes and recent human influence. More broadly, our study demonstrates that the temporal resolution of demographic history can be extended by examining multiple facets of genomic variation.

**Key words:** demographic inference, timescales, island evolution, whole genome sequencing.

## Significance

Detailed information about a population's history can be obtained by studying patterns of genomic variation, but these insights can be limited in their temporal scope. Our investigation of white-footed mice in the urban Boston Harbor archipelago demonstrates how combining multiple summaries of genomic variation enables a more complete reconstruction of population history over both the recent and distant past.

## Introduction

The evolutionary processes and events that shape the history of a population leave signatures in patterns of DNA sequence variation. With advances in genome sequencing, statistical methods, and computational approaches, it is

now feasible to reconstruct population sizes, migration rates, and split times in any set of populations for which samples can be obtained (Beichman et al. 2018). In addition to quantifying these important evolutionary parameters, demographic inference is a key step toward the goal of

© 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 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.

using genomic data to characterize natural selection (Pavlidis et al. 2010; Crisci et al. 2013; Lotterhos and Whitlock 2014; Johri et al. 2022).

Existing approaches for demographic inference can be distinguished by the genomic summary used to draw inferences, the statistical framework they employ, or the assumptions they make (reviewed in Beichman et al. 2018). An important distinction concerns the timescale over which available strategies are best suited to estimate demographic parameters of interest (Nadachowska-Brzyska et al. 2022). This temporal scope is determined by how rapidly a given genomic signal responds to relevant demographic processes and how quickly the signal decays. For example, tracts of identity-by-descent (IBD) are immediately established by recent common ancestry and readily broken up by recombination during subsequent meioses (Browning and Browning 2012; Thompson 2013). As a consequence, long-distance genotype correlations that reflect this coancestry form sensitive indicators of recent history (Palamara et al. 2012; Harris and Nielsen 2013). In contrast, mutation serves as the record keeper of demographic change for analyses based on nucleotide diversity or allele frequencies at unlinked sites. Here, inferences about the recent past are limited by the low rate at which new mutations occur and by the capacity to observe them in a sample (Adams and Hudson 2004; Keinan and Clark 2012; Robinson et al. 2014). Such properties of the sample can also constrain the temporal resolution of approaches based on the sequentially Markovian coalescent (SMC), where the times to most recent common ancestry (TMRCA) observed for small numbers of genomes reflect population size history over more ancient timescales (Li and Durbin 2011; Schiffels and Durbin 2014).

Recognition of these limitations (Nadachowska-Brzyska et al. 2022) has spurred efforts to broaden the temporal resolution captured by demographic inference, either by extending existing approaches to larger samples (Schiffels and Durbin 2014), combining multiple informative summaries (Boitard et al. 2016; Terhorst et al. 2017), or leveraging new summaries (Harris and Nielsen 2013). Despite this progress, the temporal limits of any given method are rarely investigated in an empirical context. As demographic inference is extended to a wider range of species encompassing diverse histories, it is increasingly important to consider the timescales to which these genomic analyses provide access.

Continental islands are useful settings for investigating how the interplay between historical and contemporary processes shape genomic patterns of variation. Formed by the marine transgressions that followed Pleistocene glacial cycles, the young geological ages of continental islands and their geographic proximity to the mainland offer multiple routes for the establishment of island populations (Martínková et al. 2013). Colonization may be achieved

by short-distance dispersal from nearby mainland areas or by vicariance events if species habitation predates island formation. For less vagile species, geological history may constrain the demographic history of island populations. Contemporary anthropogenic influence adds a layer of complexity to the population dynamics underlying these island systems, as human transport can overcome otherwise impermeable barriers to dispersal.

The Boston Harbor archipelago presents a prime opportunity to elucidate the contributions of historical and contemporary demography to genomic diversity in island populations. Ranging in size from 0.4 ha to 105 ha, the islands were formed by the glacial retreat and subsequent sea level rise that followed the Last Glacial Maximum (Olmstead Center for Landscape Preservation 2017). Owing to their abundance of natural resources and proximity to the mainland, the islands in Boston Harbor have served a prominent role in the ecological and cultural history of the region (Luedtke and Rosen 1993; Luedtke 1996; Richburg and Patterson 2005). Prior to their incorporation into the Boston Harbor Islands National Recreation Area in 1996, the islands supported diverse public and private endeavors including farms, schools, military forts, and waste management facilities (Olmstead Center for Landscape Preservation 2017). Although there have been efforts to revitalize the natural landscapes of the Boston Harbor (Richburg and Patterson 2005; Olmstead Center for Landscape Preservation 2017), the lasting impact of human development is evident in the high burden of exotic plant species surveyed among the islands (Elliman 2005).

Despite its dynamic history, multiple islands within this “urban archipelago” support stable populations of the white-footed mouse, *Peromyscus leucopus* (Nolfo-Clements and Clements 2015; Nolfo-Clements 2018), a North American endemic that has colonized a wide range of environments (Osgood 1909; Bedford and Hoekstra 2015). *Peromyscus leucopus* is an emerging model system for native species that persist in the face of human disturbance (Munshi-South 2012; Harris and Munshi-South 2017; Yu et al. 2017). For example, in the highly urbanized landscape of New York City, *P. leucopus* are ubiquitous in forested fragments and demonstrate an impressive ability to disperse along even the most marginal “green” corridors formed by roadside vegetation, residential areas, and cemeteries (Munshi-South 2012).

Reconstructing the historical and contemporary demography of island systems also lays the foundation for understanding the phenotypic departures that are characteristic of island evolution. Comparisons of body size between island and mainland populations spanning diverse taxa have demonstrated broad adherence to the “island rule,” a pattern in which large-bodied species become smaller on islands and small-bodied species become larger (Foster

1964; Van Valen 1973; Lomolino 1985; Benítez-López et al. 2021). Mice surveyed on both Bumpkin and Peddocks are 40% to 50% heavier than mainland *P. leucopus* (Nolfo-Clements et al. 2017), suggesting that island populations within the Boston Harbor follow the island rule. While meta-analyses have suggested that differences in predation, competition, resource availability, and dispersal ability between island and mainland populations may be drivers of this phenomenon (Foster 1964; Case 1978; Lomolino 1985; Adler and Levins 1994; Palkovacs 2003; Lomolino 2005), few studies have elucidated the demographic conditions under which the island rule emerges (Martínková et al. 2013; Gray et al. 2014; Baier and Hoekstra 2019; Wang et al. 2023).

In this article, we take advantage of the unique geological and ecological legacy of the Boston Harbor to elucidate the impact of historical and contemporary demographic processes on genomic variation in populations of *P. leucopus*. Using whole genome sequences from island and mainland populations of mice, we draw inferences from allele frequencies, patterns of linkage disequilibrium (LD), and properties of rare variants. We illustrate how the insights gained from these distinct signals enrich our understanding of demographic history on multiple timescales.

## Results

### Sampling Island and Mainland Sites Within the Boston Harbor

The Boston Harbor archipelago is comprised of 34 small islands and peninsulas just east of the city center. Our study focused on two of the inner harbor islands, Peddocks (74.6 ha) and Bumpkin (12.2 ha), and one mainland site on the World's End peninsula (Fig. 1). Using tissue samples collected from wild-caught *P. leucopus*, we constructed short-read whole genome sequence datasets for mice from each of these three locations ( $n = 37$  for Bumpkin Island,  $n = 21$  for Peddocks Island, and  $n = 19$  for mainland World's End). Samples were sequenced to a target coverage of 30X. Realized coverage ranged from 17X to 34X with a mean of 28X across mice.

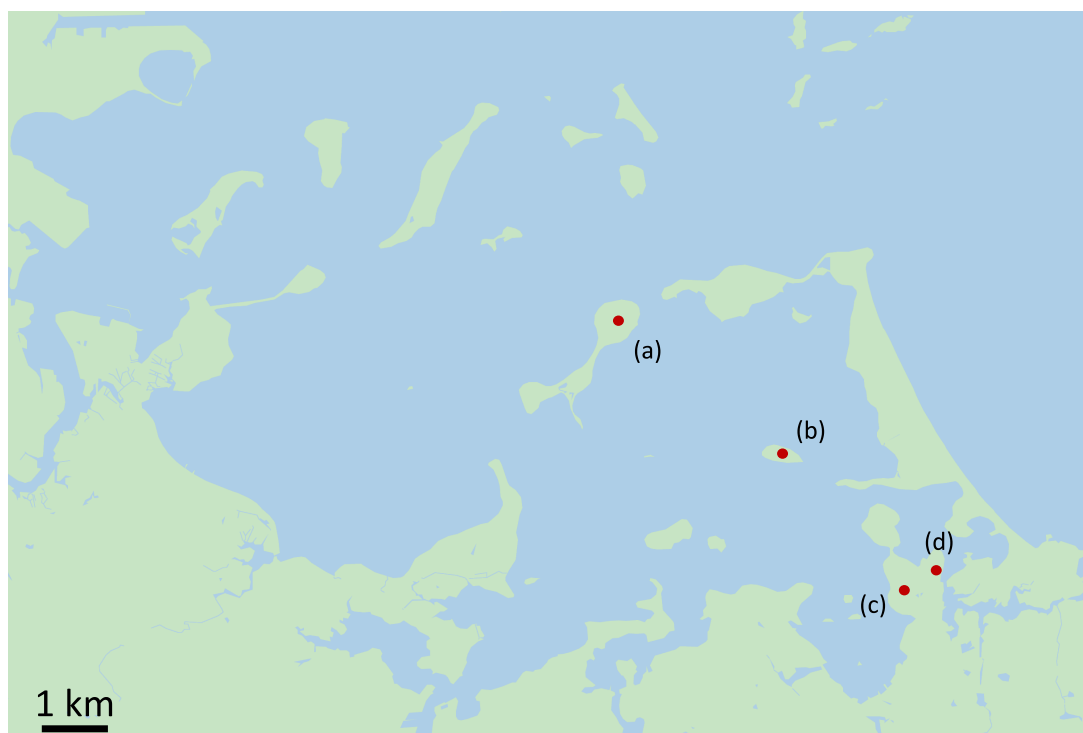
We used a subset of unlinked single nucleotide polymorphisms (SNPs) to estimate pairwise kinship coefficients between sampled mice with KING (Manichaikul et al. 2010). Surprisingly, we inferred a high incidence of close relatives in both island cohorts, with 13.77% of Bumpkin Island pairs and 10.28% of Peddocks Island pairs exceeding the lower bound for 2nd degree relatives (supplementary fig. S1a and b, Supplementary Material online). In contrast, only 3.33% of mainland World's End pairs were inferred as closely related (defined here as pairs  $\leq$  2nd degree relatives) (supplementary fig. S1c, Supplementary Material online).

We did not find any close relatives among mice sampled from different locations.

In addition to this abundance of close relatives, we also identified a single Bumpkin Island mouse (B19) that exhibits substantial genetic divergence from the rest of the Bumpkin sample. Specifically, we found that B19 contributes a disproportionately large number of singleton variants (i.e. those with a minor allele count of 1 in the Bumpkin sample) to the genome-wide folded site frequency spectrum (SFS) (supplementary fig. S2, Supplementary Material online). Many of these singleton variants are shared with the Peddocks and World's End cohorts, reducing the likelihood that they originated from sequencing or genotyping errors (supplementary fig. S3, Supplementary Material online). Principal component analyses (PCA) restricted to Bumpkin mice and PCA involving mice from all three populations identified this individual as an outlier with respect to the rest of the Bumpkin sample (supplementary fig. S4, Supplementary Material online). In addition, ADMIXTURE (Alexander et al. 2009) analyses conducted on the combined dataset suggested that this individual derives ancestry from both World's End and Peddocks (supplementary fig. S5, Supplementary Material online). Given their potential to bias the SFS, we excluded both B19 and the close relatives we identified in each cohort from allele frequency-based analyses, reducing sample sizes to  $n = 13$  for Bumpkin Island,  $n = 13$  for Peddocks Island, and  $n = 17$  for mainland World's End.

### Population Structure Within the Boston Harbor

To characterize genetic relationships between sampled locations, we conducted PCA using genome-wide, unlinked SNPs from each island and mainland cohort (see Materials and Methods). We found that individuals cluster according to sampling location along the first and second principal component axes, indicating genetic divergence between mice inhabiting different sites (Fig. 2a). To further assess the degree of structure both within and among sampled locations, we performed model-based inference of global ancestry proportions with ADMIXTURE (Alexander et al. 2009) using the same collection of genome-wide SNPs. Evidence for structure within the combined sample was strong, with cross-validation analyses identifying three ancestral components (Fig. 2b) that delineated individuals according to location (Fig. 2c). In contrast, a single ancestral component provided the best fit to each of the location-specific samples, suggesting little evidence for structure within each site (Fig. 2b). Given these results, we treated each of the sampled locations (Peddocks Island, Bumpkin Island, and mainland World's End) as representing distinct populations of mice. To orient subsequent demographic analyses, we inspected the SFS of shared and private variants between each pair of populations (supplementary fig. S6,



**Fig. 1.** Map of the Boston Harbor archipelago in Boston, Massachusetts. Land areas are depicted in green and water areas are depicted in blue. Red points indicated sampled areas on Peddocks Island (a), Bumpkin Island (b), and mainland World's End (c and d).

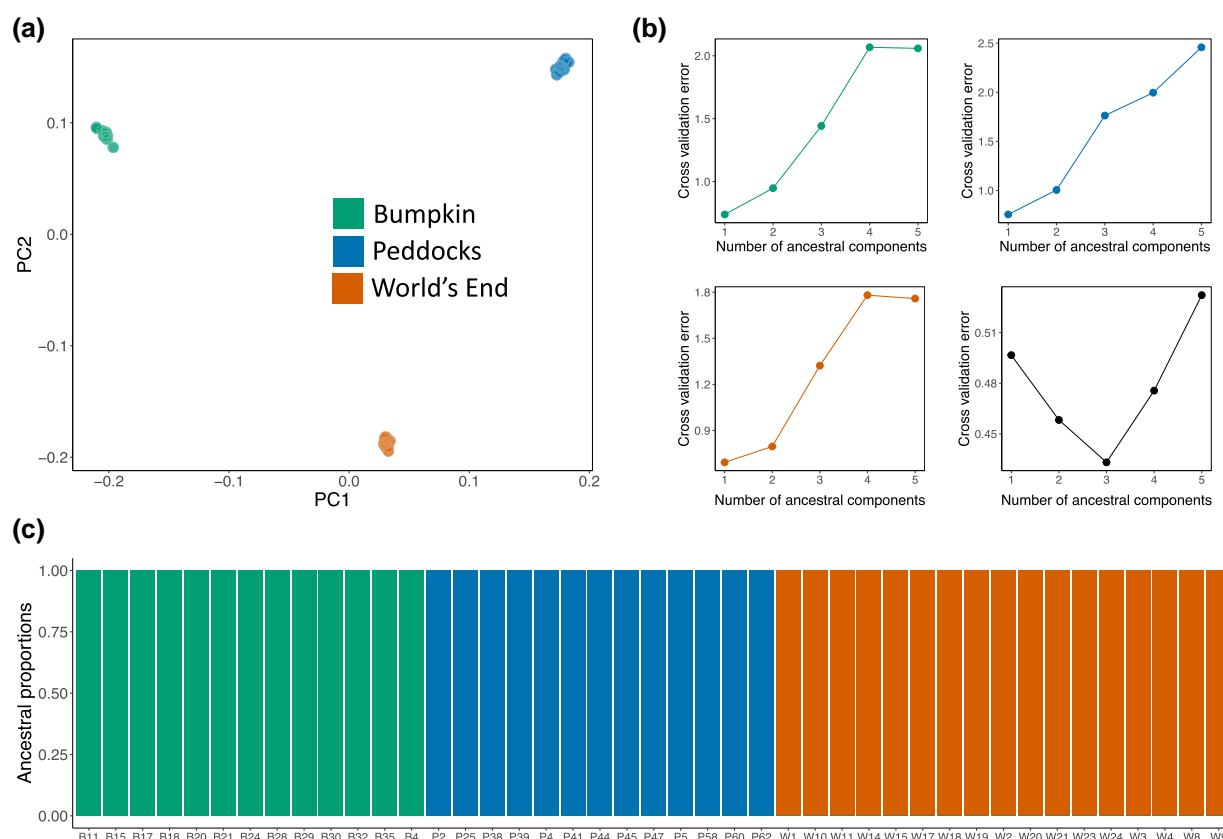
Supplementary Material online). These partitions indicated that mainland World's End retains much of the genetic variation present in the Peddocks Island and Bumpkin Island samples, providing evidence that the three sampled populations are indeed derived from a recent shared ancestral population.

### Historical Insights About Demography

To reconstruct the demographic histories of these island and mainland populations, we employed an approach based on the joint site frequency spectrum (jSFS). Using *Moments* (Jouanous et al. 2017), we fit two-population demographic models to the folded jSFS of putatively neutral variants observed in each pair of population samples (see Materials and Methods). Our aims were to estimate population divergence times, changes in effective population size ( $N_e$ ) through time, and historical levels of gene flow between populations. For each pair of populations, our model selection criteria (see Materials and Methods) favored the simplest model of population divergence. This model involves a split between sampled populations from a shared ancestor, constant island and mainland  $N_e$  following their split, and no migration (supplementary fig. S7a, Supplementary Material online). Maximum likelihood

estimates of  $N_e$  and divergence times yielded congruent values across the three pairwise comparisons (Fig. 3). Parameter estimates across these three inferred histories suggest a large ancestral  $N_e$  (ranging from  $N_e = 413,589$  to  $N_e = 470,513$ ), small postsplit  $N_e$  for both Bumpkin Island (between  $N_e = 7,257$  and  $N_e = 7,501$ ) and Peddocks Island (between  $N_e = 9,465$  and  $N_e = 9,669$ ), and a larger postsplit  $N_e$  for mainland World's End (between  $N_e = 65,483$  and  $N_e = 67,573$ ) (Fig. 3). Although there is overlap among the confidence intervals for all inferred divergence times, their point estimates imply an older split between the Peddocks and World's End populations (7,659 generations ago) and a more recent split between the Bumpkin and World's End populations (6,911 generations ago) (Fig. 3).

To evaluate the ability of the inferred demographic histories to recover summaries of genetic variation beyond the jSFS, we conducted predictive simulations under the best-fit parameter estimates using *msprime* (Baumdicker et al. 2022) (see Materials and Methods). For  $\pi$ , Tajima's  $D$ , and  $F_{ST}$ , we found close agreement between the means of the observed window-based estimates and simulated distributions (Fig. 4). For all three measures, the models underpredicted the variance in the observed summary statistic distributions (Fig. 4), but much of this discrepancy is likely driven by unmodeled heterogeneity in the recombination



**Fig. 2.** Evidence of population structure among Boston Harbor localities. a) Principal component analysis conducted using genome-wide, unlinked SNPs. Points represent individuals sampled from either Bumpkin Island (green), Peddocks Island (blue), or mainland World's End (orange). Percentage of variance explained by PC1 = 10.36% and by PC2 = 8.82%. b) Five-fold cross-validation analysis with ADMIXTURE (Alexander et al. 2009). A single ancestral component ( $k = 1$ ) yields the minimum cross-validation error within each of the sampled localities (top left, top right, and bottom left). In the combined sample (bottom right), three ancestral components ( $k = 3$ ) yield the minimum cross-validation error. c) Global ancestry proportions among sampled individuals, assuming  $k = 3$  ancestral components. Each individual is represented by a colored bar. The height of each colored segment in a bar gives the proportion of ancestry an individual derives from a given component. Ancestral component colors reflect population of origin (as in a).

and the mutation rate as well as differences in SNP density caused by the filtering conducted on the empirical callset. In contrast to the close fit our inferred model provides to allele frequency-related summaries, we found that LD decays more quickly and reaches a lower minimum at 1 Mbp for simulated data (Fig. 4). This discrepancy is largest for the two island populations, Bumpkin and Peddocks, where LD decay curves are unusually high and flat. In both cases,  $r^2$  exceeds 0.25 even for SNPs separated by as much as 1 Mbp (Fig. 4).

Beyond these common summaries of genomic variation, we detected an additional genomic pattern that our simulations failed to predict. In both the Bumpkin Island and Peddocks Island samples, singleton variants form distinct tracts in which a single mouse contributes a disproportionate number of singletons to the local (within-population) SFS (Fig. 5a and c). At the genomic scale, this pattern results in substantial inter-chromosomal variance in individual singleton contributions (supplementary figs. S8 and S9, Supplementary Material online), with multiple Bumpkin

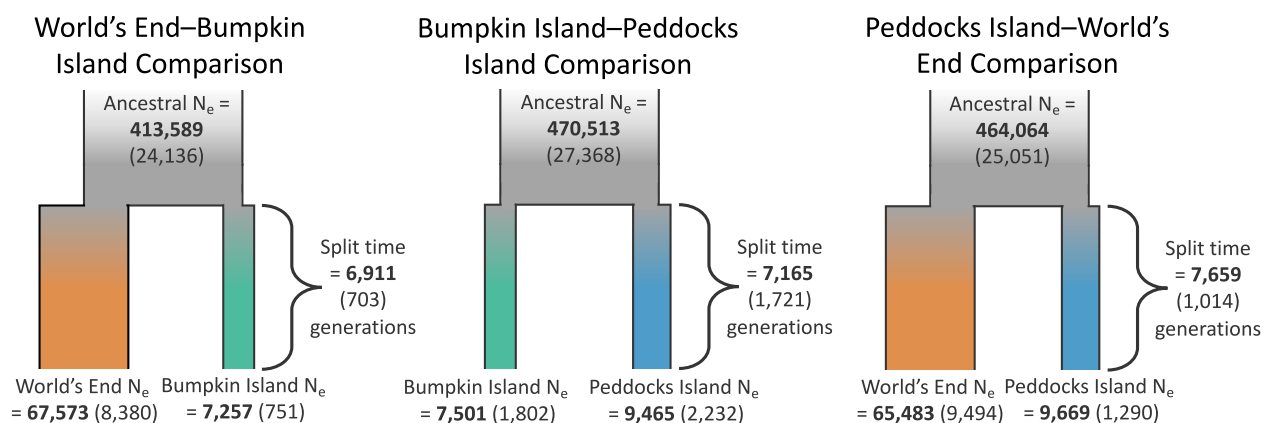
mice deriving more than 50% of their contributed singletons from a single chromosome (supplementary fig. S8, Supplementary Material online). This pattern is less evident in mainland World's End, where singleton variants appear more uniformly distributed across the genome (Fig. 5e) and across mice (supplementary fig. S10, Supplementary Material online). Chromosome-scale predictive simulations demonstrate that the extreme spatial clustering of singletons observed in the island samples is unlikely to be produced solely by the inferred reductions in population size (Fig. 5b, d, and f).

In the following sections, we explore possible causes for these intriguing features of the data that are not fit by our historical models of demography based on the jSFS.

### Inter-archipelago Migration May Explain the Observed Clustering of Rare Variants

The unusual distribution of singleton variants in the island samples may be a more localized manifestation of the





**Fig. 3.** Demographic inferences from the SFS support old divergence times and smaller  $N_e$  for island populations. Model fitting and parameter estimation were performed separately for the jSFS constructed from each pair of populations, yielding three separate histories for the mainland World's End–Bumpkin Island (left), Bumpkin Island–Peddocks Island (middle), and Peddocks Island–mainland World's End (right) comparisons. Maximum likelihood estimates for each parameter are bolded. Standard deviations for each parameter estimate are given in parentheses and were ascertained using the Godambe Information Matrix (Coffman et al. 2016) as an alternative to conventional bootstrapping.

excess singletons we observed in the excluded Bumpkin Island mouse, B19 (supplementary fig. S11, Supplementary Material online). Our finding that B19's singletons are segregating in both the Peddocks and World's End populations (supplementary fig. S3, Supplementary Material online) raises the possibility that this pattern is driven by gene flow within the archipelago. Indeed, the proximity of island and mainland areas within the Boston Harbor, combined with the frequent human traffic between them, leaves multiple routes for the migration of small mammals. Given the substantial genetic divergence we observed between the three populations (Fig. 2), we should expect much of the genomic content exchanged by within-archipelago migration to involve private variation. Consequently, when tracts of migrant ancestry are observed in the recipient population, the private variation they carry will manifest as new polymorphisms. If migration is sufficiently rare or recent such that individuals carry distinct migrant haplotypes, then these tracts of migrant ancestry will contribute disproportionately to rare allele frequency bins.

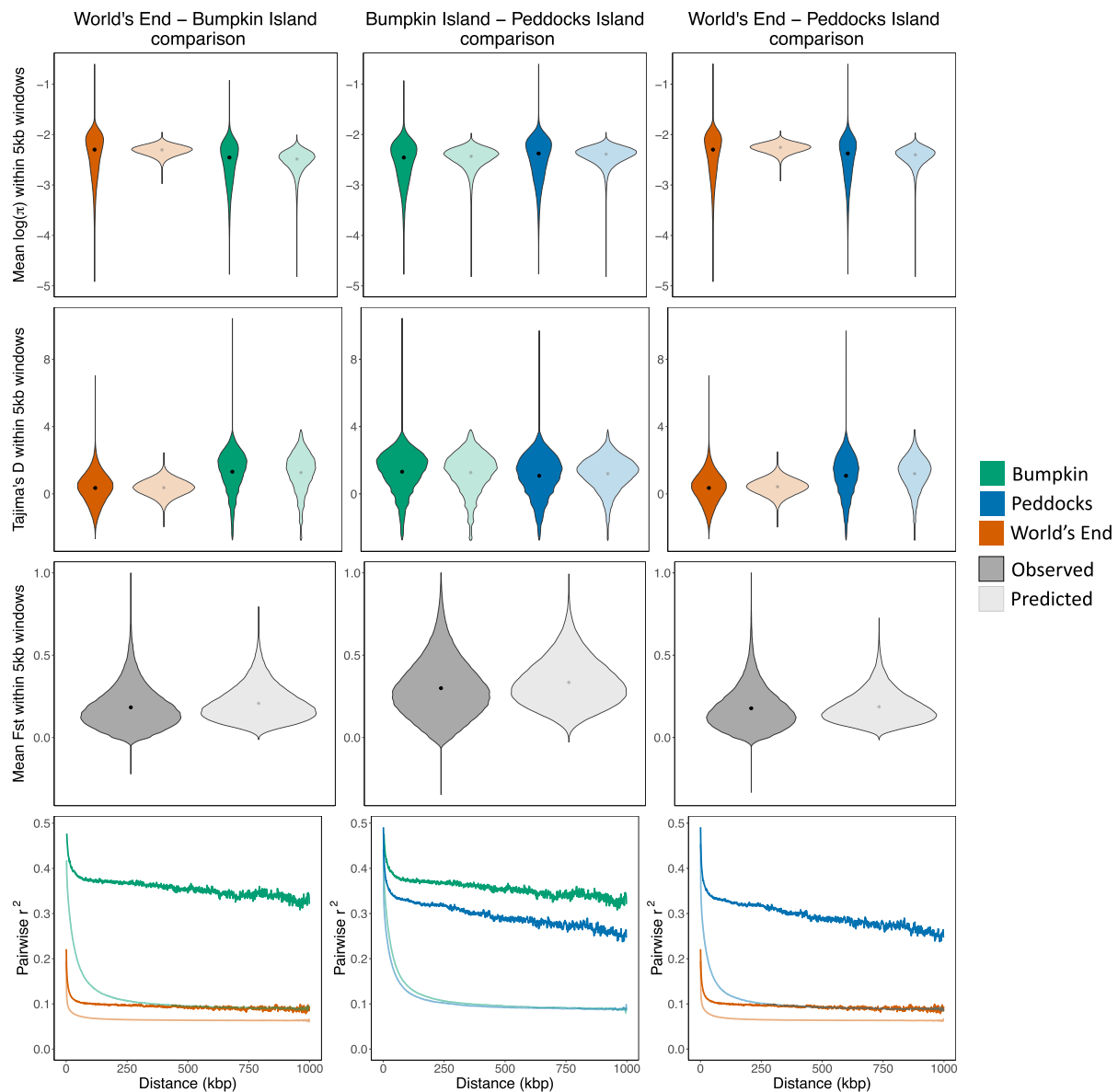
We used coalescent simulations of genome-scale datasets to test this verbal model and explore the possibility that migration (which was absent from our best-fitting demographic models based on the jSFS) could explain the pattern of singleton clustering observed in both island samples. For each pair of populations, we augmented our best-fit jSFS demographic model with migration from an unsampled “ghost” population, varying both the migration rate and the time at which migration begins following population divergence (see Materials and Methods).

By inspecting the spatial distribution of singleton variants across simulated chromosomes, we found qualitative

evidence that migration can generate dense tracts of singletons, such as those observed in both the Bumpkin Island and Peddocks Island samples (supplementary figs. S12 and S13, Supplementary Material online). Specifically, regimes involving either extended periods of weak migration (Fig. 6a and c) or brief periods of strong migration (Fig. 6b and d) yield singleton tracts that are similar in length and abundance to those found in the island datasets (Fig. 5a and c).

The other salient feature of singleton variants in the island samples is their uneven distribution among mice. We found high variance in the number of singletons each mouse contributes to the SFS of a given chromosome (supplementary figs. S8 and S9, Supplementary Material online). To evaluate whether the simulated migration regimes could recover this variability, we compared chi-squared test statistics computed for each empirical and simulated chromosome assuming a null hypothesis of equal singleton contributions among mice. We found that the mean chi-squared statistic measured across empirical Peddocks Island and Bumpkin Island chromosomes falls within the range of values produced by the simulated migration regimes (supplementary fig. S14, Supplementary Material online). Although these empirical chi-squared test statistics are more extreme than those obtained under simulations without migration, even migration-free simulations yield large deviations from the null hypothesis. These results indicate that the population size history represented in our jSFS-based demographic models may itself affect the distribution of singleton variants among individuals.

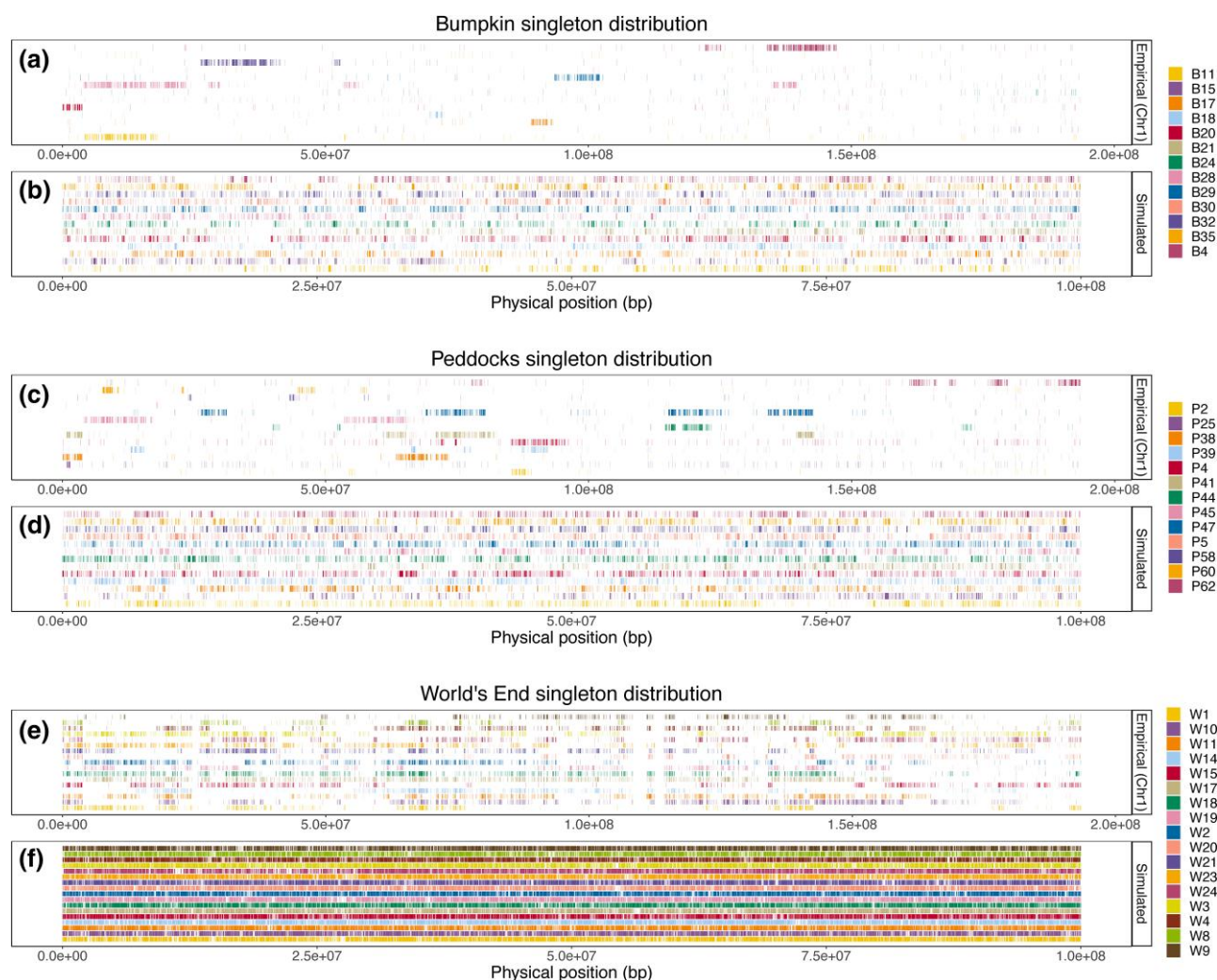
To develop intuition for how population size history and migration interact to shape individual singleton contributions, we leveraged the functionality of the tskit library (Kelleher et al. 2018) to inspect properties of the local genealogies underlying simulated chromosomes (see



**Fig. 4.** SFS-based demographic models provide a close fit to some, but not all, summaries of variation. Panels compare summaries of variation measured from empirical data (dark shading) to those obtained from predictive simulations conducted using best-fit demographic parameter estimates (light shading). Violins plot the distributions of  $\pi$  (log-scaled; first row), Tajima's  $D$  (second row), and  $F_{ST}$  (third row) computed in 5 kbp windows. Black points denote the mean value of each distribution. LD decay curves (fourth row) plot average pairwise  $r^2$  between SNPs separated by increasing physical distances (up to 1 Mbp). For single-population summaries, empirical and simulated distributions/curves are colored according to population (Bumpkin Island, green; Peddocks Island, blue; mainland World's End, orange). Column titles denote the demographic model used for each set of predictive simulations (World's End-Bumpkin Island, left; Bumpkin Island-Peddocks Island, middle; World's End-Peddocks Island, right).

Materials and Methods). Focusing on Bumpkin Island chromosomes simulated under the Recent/Strong migration regime featured in Fig. 6b, we compared various measures of branch lengths between genealogical trees spanning regions with and without migrant ancestry (Fig. 7a and b). As expected, the presence of one or more migrant haplotypes increases the total external branch length ( $E$ ), the mean external branch length within a tree ( $\text{Mean}(E)$ ), and the standard deviation of external

branch lengths within a tree ( $\text{Sd}(E)$ ) (Fig. 7c). Although migration increases the means of these external branch length measurements, we note that there is still substantial variability in these quantities across trees spanning regions with and without migrant ancestry. Interestingly, despite the large distortions it produces along external branches, migrant ancestry appears to have little effect on the total tree height ( $H$ ) or the total branch length ( $T$ ) (Fig. 7c).



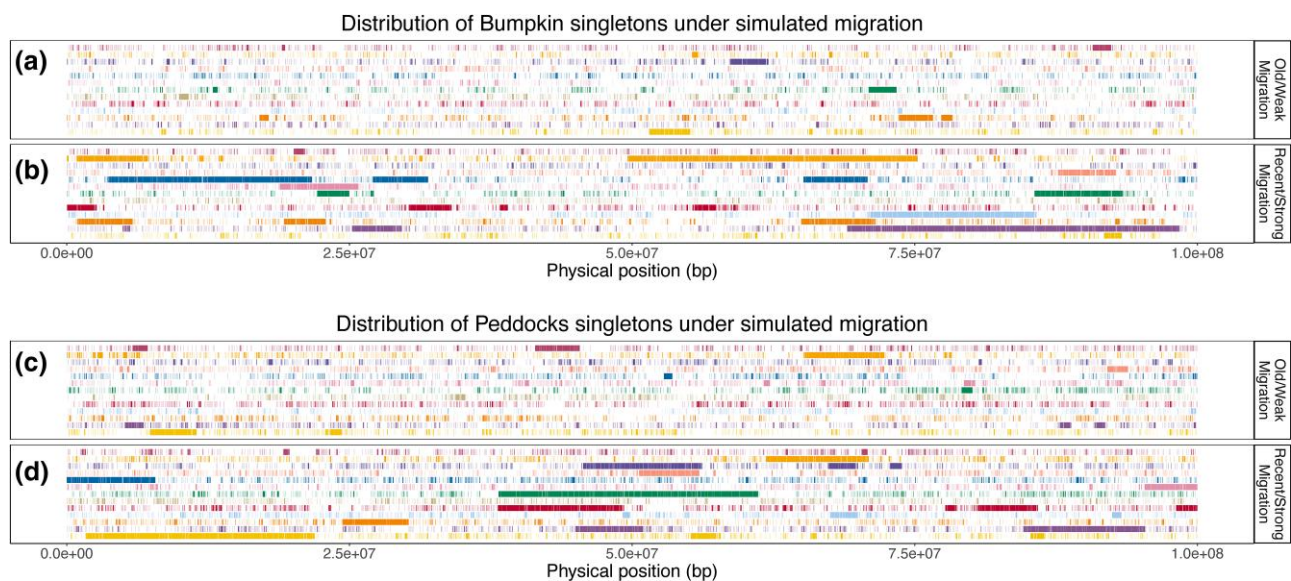
**Fig. 5.** SFS-based demographic models fail to recover the spatial clustering of rare variants observed in island samples. a, c, and e) Distribution of singleton variants observed on chromosome 1 (NC\_051063.1) within the Bumpkin Island (a), Peddocks Island (c), and mainland World's End (e) samples. X-axes give the physical position along chromosome 1. Vertical, colored tick marks denote observed singleton variants. Each row reflects the singleton variants of a distinct individual. b, d, and f) Corresponding singleton distributions predicted by the inferred demographic histories for Bumpkin Island (b), Peddocks Island (d), and mainland World's End (f). X-axes give the location of singletons along a 100 Mbp simulated chromosome and tick marks denote singleton variants observed in each simulated individual. For clarity, we only show the simulated singleton maps derived from the Bumpkin-World's End demographic model (for b and f) and the Peddocks-World's End demographic model (for d), though we note that patterns are similar regardless of which specific demographic model is used for a given population.

### Small Contemporary Effective Population Sizes May Explain Elevated Linkage Disequilibrium

The pattern of LD decay we observed in the Bumpkin Island and Peddocks Island populations yielded curves that are unusually high and flat. Such characteristics may indicate a recent and severe population collapse (Rogers 2014). In addition to this trend of elevated LD, the abundance of close relatives we identified in each sample may signify smaller effective population sizes than suggested by our jSFS-based demographic models (Wang 2009; Waples and Anderson 2017).

Given these findings and the dynamic ecological histories of the islands, we hypothesized that the demographic models we inferred from the jSFS may not capture more recent changes in effective population size. To investigate this possibility, we reconstructed contemporary  $N_e$  using two approaches that leverage different facets of LD (see Materials and Methods). *CurrentNe* uses LD measured between unlinked markers to obtain a point estimate of  $N_e$  that reflects the several most recent generations (Santiago et al. 2024). In contrast, GONE uses LD measured between markers separated by increasing genetic distances





**Fig. 6.** The physical clustering of rare variants is recovered by simulated migration regimes. a and c) Singleton distributions predicted for Bumpkin Island (a) and Peddocks Island (c) samples under a demographic history that includes weak migration (at a rate of  $1e-5$ ) between sampled populations and an unsampled “ghost” population that occurs for 10% of the total divergence time. b and d) Predicted singleton distributions for Bumpkin (b) and Peddocks (d) samples under a demographic history that features stronger migration (at a rate of  $1e-3$ ) occurring for only 1% of the total divergence time. Singleton maps are arranged as in Fig. 5b and d. For clarity, we only show the simulated singleton maps derived using the population size and divergence time estimates from the Bumpkin-World’s End (a and b) and Peddocks-World’s End (c and d) demographic models, though we note that patterns are similar regardless of which specific demographic model is used as the basis for each set of migration simulations.

to reconstruct the trajectory of  $N_e$  in the recent past (Santiago et al. 2020). Using these approaches, we estimated recent  $N_e$  and the recent  $N_e$  trajectory for each population using either the full sample or the subset of unrelated individuals (excluding relationships  $\leq 2$ nd degree).

Regardless of whether the sample included close relatives, point estimates of contemporary  $N_e$  for each population are over an order of magnitude smaller than suggested by best-fit jSFS demographic models (Figs. 8a and 3). For the unrelated samples of mice, we obtained  $N_e$  point estimates of 20, 66, and 87 for the Bumpkin Island, Peddocks Island, and mainland World’s End populations, respectively (Fig. 8a). As expected, datasets that were pruned for close relatives yielded larger  $N_e$  estimates (Fig. 8a). Recent  $N_e$  trajectories inferred with GONE are similar between datasets with and without close relatives, with the exception of the recent bottleneck inferred for Peddocks from the full sample (Fig. 8b). For all populations,  $N_e$  estimates in the immediate past (generation 0) are close to the point estimates obtained with *currentNe* (Fig. 8a and b). These trajectories suggest that each population has experienced dynamic changes in  $N_e$  over the last 150 generations (Fig. 8b).

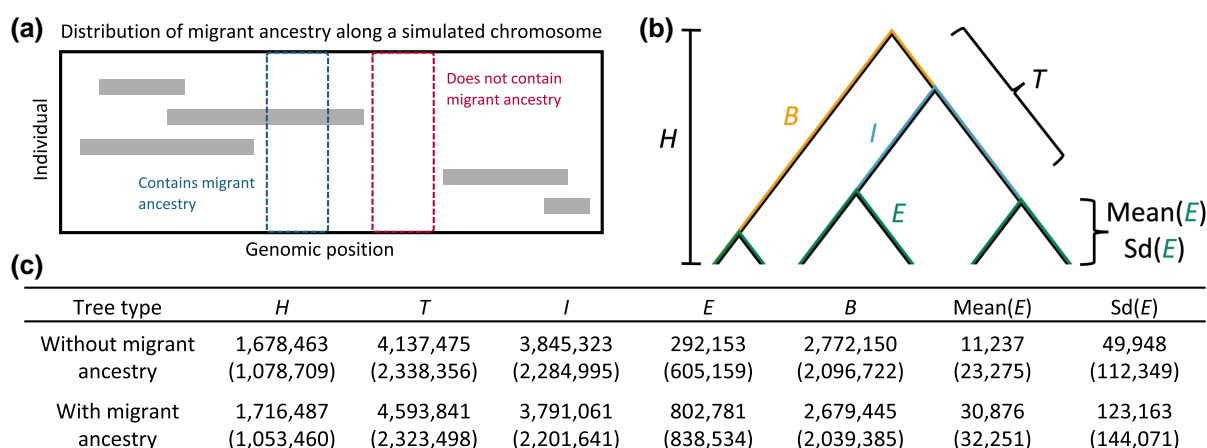
## Discussion

Demographic histories of natural populations are shaped by both historical processes (e.g. glacial cycles, sea-level

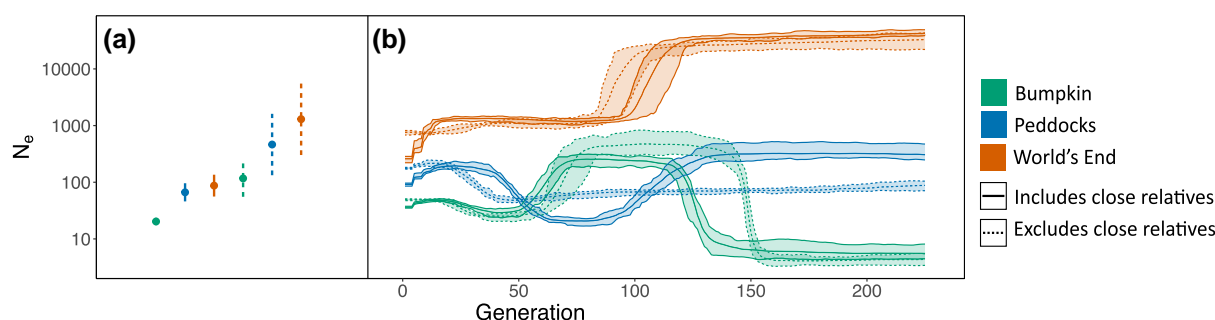
changes, and tectonic activity) and contemporary shifts (e.g. anthropogenic changes in climate, habitat loss, and fragmentation). While fully capturing the dynamic nature of these histories is beyond the capacity of population genetics, increasingly complex insights about population history are being extracted from genomic data. Despite this progress, we lack a unified framework for drawing inferences about demographic history that span the recent and distant past (Nadachowska-Brzyska et al. 2022). Using island and mainland populations of endemic mice as a model system, we illustrate the importance of time-scale in the interpretation of empirical patterns of genomic variation. In the following sections, we discuss both the historical and contemporary insights gained about *P. leucopus* in the Boston Harbor and conclude with general prospects for conducting demographic inference along the temporal continuum.

## Historical Insights

The candidate demographic models we obtained from our SFS-based inference suggest that the current distribution of *P. leucopus* in the Boston Harbor was primarily mediated by the geological formation of the archipelago. The topography of the Boston Harbor was shaped by Pleistocene glacial activity which formed the drumlin field that constitutes islands within the archipelago (Kaye 1982; Aubrey 1996). Rises in sea level and the subsequent inundation of the



**Fig. 7.** Singleton clustering reflects the interaction of migration and population size history on the external branch lengths of local genealogies. We inspected coalescent genealogies for the Bumpkin Island sample simulated under the Recent/Strong migration regime featured in Fig. 6b. a) Diagram illustrates how local genealogies were classified into two types. Each row represents variation observed in a different simulated individual. Gray rectangles reflect migrant haplotypes derived from the “ghost” population. The tskit library (Kelleher et al. 2018) was used to distinguish between local genealogies spanning regions containing one or more migrant haplotypes (blue box) and those spanning regions without migrant ancestry (red box). b) We measured multiple genealogical quantities for each category of trees: the total tree height ( $H$ ), sum of total branch lengths ( $T$ ), sum of basal branch lengths ( $B$ ), sum of internal branch lengths ( $I$ ), sum of external branch lengths ( $E$ ), mean external branch length within a tree (Mean( $E$ )), and standard deviation of external branch lengths within a tree (Sd( $E$ )). c) Table summarizes the mean value of each quantity across all trees observed along a simulated 100 Mbp chromosome. Standard deviations across trees are given in parentheses.



**Figure 8.** LD-based estimates of recent effective population size are an order of magnitude smaller than SFS-based inferences. a) Point estimates of contemporary  $N_e$  obtained with *currentNe* (Santiago et al. 2024). Vertical lines represent 90% confidence intervals. b) Recent  $N_e$  trajectories estimated by GONE (Santiago et al. 2020).  $N_e$  trajectories span the present (generation 0) to 200 generations into the past. Center lines represent  $N_e$  estimates averaged over 15 replicate GONE analyses. Ribbons represent the range of each  $N_e$  estimate across replicates. Point estimates in (a) and  $N_e$  trajectories in (b) are colored according to population (Bumpkin Island, green; Peddocks Island, blue; mainland World's End, orange). Dotted confidence intervals/trajectories reflect estimates made using a subset of the data in which close relatives (defined as relationships closer than or equal to 2nd degree) were excluded. Solid confidence intervals/trajectories reflect estimates made using the full sample.

Boston Harbor basin severed existing land connections and established the present-day distribution of island and mainland areas (Kaye 1982; Aubrey 1996). Historical shoreline reconstructions suggest that the islands were isolated 3,000 to 6,000 years ago (Aubrey 1996). Assuming 1 to 2 generations per year (as observed in natural populations of *P. leucopus* in Michigan (Burt 1940)) our inferred divergence times, ranging from 6,000 to 7,000 generations ago, show broad agreement with the estimated timing of island formation. While overlapping confidence intervals raise

uncertainty about the order of inferred population splits, their point estimates are consistent with the piecemeal changes in connectivity suggested by the geological history of the region, in which outer islands such as Peddocks were isolated earlier from the mainland (and are therefore older) than the inner islands of the archipelago such as Bumpkin (Aubrey 1996). Although such comparisons should be interpreted cautiously in light of uncertainty about generation time and mutation rate in these populations, our inferences about divergence times are largely congruent

with the geological formation of the Boston Harbor archipelago.

Demographic models based on the SFS also indicate that all three populations experienced reductions in  $N_e$  following their divergence from a shared ancestor. These size reductions were most extreme for the populations on Bumpkin (60-fold decrease) and Peddocks (49-fold decrease). The modest  $N_e$  values inferred for the island populations could reflect founder effect bottlenecks that accompanied their isolation from the mainland, or constraints on habitat and resource availability driven by the small land areas of both Bumpkin (12.2 ha) and Peddocks (74.6 ha).

Given the inferences about migration and recent changes in  $N_e$  drawn from orthogonal approaches, the apparent simplicity of these best-fitting jSFS-based models should be interpreted with caution. Although we attempted to fit models that incorporated migration and additional population size changes (see Materials and Methods), they did not improve the fit to the data beyond that obtained for these simpler models of population divergence. The lack of support for more parameter-rich models may be driven by inherent limitations of our population genomic dataset. Excluding close relatives reduced sample sizes for each population, which limits our ability to observe rare variants. In turn, this constraint likely decreased the complexity of demographic models that could be fit and reduced our power to resolve demographic changes that occurred in the more recent past (Robinson et al. 2014; Kim et al. 2017). In addition, certain features of the observed SFS could not be recovered by any of the tested models. In particular, residual differences between the inferred models and observed data were driven largely by the “roughness” of the marginal SFS of both the Bumpkin and Peddocks samples (supplementary fig. S15, Supplementary Material online).

Despite their limitations, these historical insights provide useful context for understanding phenotypic evolution in this system. The vicariance mechanism of island colonization suggested by our inferred histories provides a contrast to classic examples of the island rule in house mice that involve recent, human-mediated introductions (Berry 1964; Berry et al. 1978a, 1978b; Berry et al. 1979; Rowe-Rowe and Crafford 1992). Unlike the rapid evolution of body size observed in these cases, our findings raise the possibility that mice in the Boston Harbor evolved large bodies over a longer period of time.

### Contemporary Insights

The divergence times and historical  $N_e$  values estimated from the SFS provide a useful starting point for reconstructing the demographic history of *P. leucopus* in the Boston Harbor. By employing orthogonal approaches to  $N_e$  estimation and by leveraging the unique empirical pattern we

observed among singleton variants in the island populations, we find strong evidence for a more dynamic recent history than indicated by this candidate model. Our results suggest an important role for recent human activity in shaping multiple demographic parameters.

The stark contrast between the LD decay predicted by our jSFS-based models and the LD decay observed in the empirical data may be explained by the differences in  $N_e$  that we estimated using SFS-based versus LD-based approaches. Although the rank order of  $N_e$  among island and mainland populations is consistent across both historical and contemporary estimates, their magnitudes differ considerably. Multiple studies have found that empirical patterns of LD are particularly difficult to recover with models derived from the SFS (Harris and Nielsen 2013; Garud et al. 2015; Beichman et al. 2017), but the source of this discrepancy is rarely investigated. In *P. leucopus* from the Boston Harbor, patterns of LD are consistent with very small  $N_e$  in the recent past, with variation in the temporal trajectory of  $N_e$  among populations starting at 150 generations ago. These findings could reflect the ecological transformation of the Boston Harbor during the last 400 years. European colonization initiated dramatic shifts in the natural landscape through the clear-cutting of forested islands for wood and the conversion of land into sites for agricultural, military, and public use (Richburg and Patterson 2005; Olmstead Center for Landscape Preservation 2017). This intense human usage has likely decreased habitat and resource availability for native species like *P. leucopus*.

Regardless of the cause for small contemporary  $N_e$ , our findings have important implications for the future of these populations. Small  $N_e$  limits the amount of additive genetic variance that can be maintained, increasing the likelihood of extinction in the face of environmental change (Lande and Barrowclough 1987; Bürger and Lynch 1995; Kardos et al. 2021). Even in the absence of new or changing selective pressures, the continual input of deleterious mutations leave small  $N_e$  populations vulnerable to extinction by “mutational meltdown,” a compounding sequence of events in which the fixation of deleterious mutations leads to declines in population size and subsequent increases in the rate at which additional deleterious mutations are fixed (Lynch et al. 1995).

The recent history of *P. leucopus* populations in the Boston Harbor probably featured changes in migration, in addition to dynamic population sizes. Our simulations of ghost migration indicate that inter-archipelago migration is a likely cause for the unusual partitioning of singleton variants in the island mice. Despite the geographic proximity of island and mainland sites in the Boston Harbor, the ocean could impose a strong barrier to the natural dispersal of *P. leucopus*. However, portions of the harbor occasionally freeze, allowing for mammals to move between island and mainland areas by crossing the ice (Nolfo-Clements

2018). In addition, Bumpkin Island experiences transient connections to the mainland via a land bridge during periods of low tide (Nolfo-Clements and Clements 2015). A more general mechanism for the migration of *P. leucopus* may be their inadvertent transport by humans. The Boston Harbor is a popular recreation area and both public ferries and private watercraft generate substantial island–island and island–mainland traffic.

To better understand why our jSFS-based demographic models failed to support a history with migration, we explored how ghost migration manifests in the jSFS used for inference (see Materials and Methods). Under our simulated scenarios, the signal that ghost migration leaves in the jSFS derives from both the movement of shared ancestral polymorphisms and the introduction of variants that are private to the unsampled population. The strength of this signal depends on the rate of migration and the duration over which it occurs (supplementary figs. S16 and S17, Supplementary Material online). Comparisons of the island–mainland jSFS under models with and without ghost migration reveal that only high intensity regimes (i.e. those involving a high rate of migration over a long period of time) can produce correlations in the island-mainland jSFS that extend to shared intermediate allele frequencies (supplementary figs. S16 and S17, Supplementary Material online). Notably, the types of migration capable of producing the observed singleton clustering (Fig. 6) have either a negligible impact on the jSFS (for the Old/Weak regime), or one that is restricted to private and low allele frequencies (for the Recent/Strong regime) (supplementary figs. S16 and S17, Supplementary Material online). These findings raise the possibility that the migration experienced by the island populations occupies a zone of parameter space that precludes robust estimation by SFS-based approaches but is nonetheless sufficient to create the observed pattern of singleton clustering.

It is unclear how unmodeled migration of this nature might affect demographic inferences drawn from the jSFS and patterns of LD. Because allele frequency correlations between populations are jointly shaped by the timing of population splits, levels of gene flow, and effective population sizes, model misspecification (i.e. the exclusion of one or more pertinent features) can have a cascading impact on other parameter estimates obtained from the jSFS (Gutenkunst et al. 2009). While it is possible that our divergence time estimates are downwardly biased by the absence of migration in our best-fitting jSFS-based models, the lack of signal that tested migration regimes leave in the jSFS indicates that our inferences may instead be robust to the level of migration suggested by the observed singleton clustering. Migration can also increase the magnitude of LD measured among loci within a population, leading to biases in LD-based estimates of contemporary  $N_e$  that depend on the rate of migration and the manner in which

interacting demes are sampled (Waples and England 2011; Novo et al. 2023; Santiago et al. 2024). To generate the observed singleton pattern, however, migration must be rare or recent enough such that the bulk of variation carried on migrant haplotypes manifests as singletons in the focal sample. This logic suggests that the level of migration relevant to the Boston Harbor populations may have negligible effects on LD. Although  $N_e$  estimates for the most immediate generations should be minimally affected by such weak migration (Novo et al. 2023; Santiago et al. 2024), Novo et al. (2023) demonstrate that even low rates of migration can still bias recent  $N_e$  estimates for more distant generations. Changes in  $N_e$  that appear earlier in our LD-based population size histories should therefore be interpreted with caution.

### Prospects for Demographic Inference Along the Temporal Continuum

As demographic inference is increasingly aimed at characterizing the histories of endangered, fragmented, and declining populations (Beichman et al. 2021; Morin et al. 2021; Kyriazis et al. 2024), it is important to jointly consider estimates of  $N_e$  that capture historical levels of variability and those that reflect the immediate past. Approaches for estimating contemporary  $N_e$  are widespread in conservation genetics (Luikart et al. 2010) but have seen limited application in the broader field of population genetics. Yet, these estimates may be particularly relevant for understanding processes that occur over ecological timescales, such as seasonal adaptation in *D. melanogaster* (Bergland et al. 2014; Machado et al. 2021). They may also be important for forecasting the future evolutionary trajectories of populations, especially in the face of anthropogenic change.

In addition to motivating deeper consideration of  $N_e$ , our investigation suggests new avenues for reconstructing contemporary migration histories. We find that the distributions of rare variants along the genome and across sampled individuals serve as sensitive indicators of recent or rare gene flow. In contrast to SFS-based approaches, which treat SNPs in different genomic locations as interchangeable (Sawyer and Hartl 1992; Gutenkunst et al. 2009), the signature of recent migration we discovered leverages the positions of SNPs along chromosomes. Furthermore, detecting this signature requires no knowledge or sampling of potential source populations. Few approaches exist to drawing inferences about migration from an unsampled ghost population. A commonly used statistic,  $S^*$ , relies on the capacity for introgressed ancestry to partition a sample in a congruent manner over long genomic distances (Wall 2000; Plagnol and Wall 2006). The singleton pattern we describe appears suited to detect migration in cases where it is too weak to generate such patterns of



LD. Future work is necessary to characterize the dependence of this signal on levels of ancestral diversity, divergence times between populations, and sample size. More detailed inferences may be achieved by extending existing theory about the distribution of external branch lengths to incorporate gene flow and population size change (Blum and François 2005; Caliebe et al. 2007; Disanto and Wiehe 2020).

Altogether, our study illustrates how combining multiple features of genomic variation can improve the temporal breadth of demographic inference. While substantial progress remains to be made in bridging the gap between the recent and distant past, important insights can still be gained by deeper evaluation of model fit and interrogation of empirical patterns.

## Materials and Methods

### Population Sampling and Sequencing

#### Animal Trapping

Animals were captured using Sherman live traps spaced at approximately 7 m intervals in grids or transect lines depending on the location. Traps were baited with a mixture of peanut butter and oats. Bedding in the form of dried leaves was added to the traps to assist with thermoregulation. We monitored the traps on Bumpkin for a total of 380 trap nights, for 480 trap nights on Peddocks, and 200 trap nights on World's End. Trapping for this study occurred June–August 2019.

#### Animal Handling and Sampling

We checked traps once daily during trapping intervals. Trapped individuals were transferred into a large, unsealed plastic bag to allow for species identification, sexing, maturity evaluation, and weighing with a spring scale. Adult and young adult mice were sampled using a 2 mm tissue punch applied to the right ear. Animals were subsequently released. We chose this sampling method due to its low level of invasiveness and because many captured wild mice have similar injuries to their ears that we presume have occurred naturally. Tissue samples were preserved and transported in 95% ethanol-filled polypropylene microcentrifuge tubes and chilled on ice. The ear punch and forceps were sterilized in between samplings using 95% ethanol to prevent cross contamination. This work was approved by the National Parks Service's Institutional Animal Care and Use Committee (project NER\_BOHA\_Nolfo-Clements\_SmMammals\_2019.A3).

#### Whole Genome Sequencing

DNA was extracted from ear punch tissue samples using Gentra Puregene Tissue Kits (Qiagen, Germantown, MD, USA). The purity of extracted DNA was assessed by

measuring 260/280 absorbance ratios with a NanoDrop Spectrophotometer (Thermo Fisher Scientific, Waltham, MA, USA) and sample concentrations were measured using the Qubit dsDNA HS Assay (Life Technologies, Carlsbad, CA, USA). Prior to library preparation, the presence of high-molecular weight DNA was confirmed using agarose gel electrophoresis. Library preparation and sequencing were conducted by the University of Wisconsin-Madison Biotechnology Center. Sequencing libraries were constructed with the Illumina DNA Library Prep using 100 to 500 ng of input DNA per sample. Purified and pooled libraries were sequenced on the Illumina NovaSeq6000 platform to generate paired-end 150 bp reads. Samples were sequenced to a targeted coverage of 30X per individual.

### Sequence Data Processing and Variant Calling

#### Sequence Alignment

Raw sequencing reads were preprocessed to remove adapter sequences and low-quality base calls. Illumina universal adapter sequences were removed using BBDuk v38.96 (<https://sourceforge.net/projects/bbmap/>) and low-quality base calls were trimmed with Trimmomatic v0.39 (Bolger et al. 2014). Processed sequencing reads were then aligned to the *P. leucopus* reference genome (UCI\_PerLeu\_2.1; GenBank accession: GCA\_004664715.2; Long et al. 2019) using BWA-MEM v0.7.17 (Li and Durbin 2009).

#### Variant Calling and Filtering

We adapted our variant calling approach from the Genome Analysis Toolkit (GATK) Best Practices recommendations for germline short variant discovery (McKenna et al. 2010; Van der Auwera and O'Connor 2020). Duplicate reads were identified from the alignments using GATK v4.2.0.0 MarkDuplicates and excluded from downstream variant calling. The removal of duplicates and other aberrant reads by GATK's default filters yielded a mean realized coverage of 28X (averaged across sites and across individuals). Genotype likelihoods were computed for each individual using GATK v4.2.0.0 HaplotypeCaller with the Reference Confidence Model set to "GVCF" mode (Poplin et al. 2017). To accommodate between-population differences in allele frequencies and ensure well-calibrated measures of variant confidence, individual call sets were combined according to sampling location with GATK v4.2.0.0 CombineGVCFs and joint variant calling was performed separately for the Bumpkin ( $n = 37$ ), Peddocks ( $n = 21$ ), and World's End ( $n = 19$ ) cohorts using GATK v4.2.0.0 GenotypeGVCFs (Poplin et al. 2017). Prior to filtering, our variant calling approach yielded a total of 51,857,593 SNPs in the Bumpkin cohort, 51,225,571 SNPs in the Peddocks cohort, and 74,861,232 SNPs in the World's End cohort.



Due to a lack of available high-confidence call sets for training more sophisticated variant filtering approaches like variant quality score recalibration (VQSR), we elected to filter our variant calls based on a combination of site-level and individual-level annotations. GATK v4.2.0.0 VariantFiltration was used to remove variants based on the Best Practices recommendations for hard filtering (“FisherStrand” > 60.0, “StrandOddsRatio” > 3.0, “RMSMappingQuality” < 40.0, “MappingQualityRankSumTest” < -12.5, and “ReadPosRankSumTest” < -8.0; Van der Auwera and O’Connor 2020) with some minor modifications (“QualByDepth” < 5.0). Given that our downstream inferences about population structure and demographic history rely on accurate measures of sample allele frequencies, we additionally removed sites for which any individual’s read depth fell below 10X to ensure the accuracy of individual genotype calls.

Since low-complexity and repetitive regions of the genome can pose issues for read mapping and variant calling (Pfeifer 2017), we restricted our population genetic analyses to variants falling outside of annotated repetitive elements. We used the union of the RepeatMasker (hub\_2100979\_repeatMasker\*; <https://repeatmasker.org>), TandemRepeatFinder (hub\_2100979\_simpleRepeat; Benson 1999), and WindowMasker (hub\_2100979\_windowMasker; Morgulis et al. 2006) annotation tracks of the UCSC Genome Browser (<http://genome.ucsc.edu>; Karolchik et al. 2004) to exclude variants falling within these regions. Copy number variation that is poorly resolved in the reference assembly could instead manifest as regions of artificially inflated coverage (Li 2014; Pfeifer 2017). To mitigate the impact of unannotated repeats on the accuracy of variant calls, we additionally imposed a maximum coverage filter by removing sites for which any individual’s read depth exceeded 50X. Together, these quality-based and accessibility-based filtering criteria yielded a total biallelic SNP count of 7,964,169 in the Bumpkin Island cohort, 9,573,228 in the Peddocks Island cohort, and 12,990,178 in the mainland World’s End cohort.

### Rescuing Invariant Sites

When there is sufficient allele frequency differentiation between populations, the joint calling approach described above will yield incongruent call sets, as variants that are private to one cohort will not be observed among the variant calls produced by other cohorts. This creates ambiguity about the status of missing variants, which could either represent monomorphic reference sites or sites for which there is insufficient information to obtain high-confidence genotypes. To ensure that our joint allele frequencies were properly calibrated for downstream analyses, we reinspected the genotypes at these missing sites. Briefly, we used the bcftools v1.8 (Danecek et al. 2021) “isec” function to extract the locations of private variants within each

pairwise comparison of call sets. We then produced VCF records for these missing sites within each cohort using the “-all-sites” mode of GATK GenotypeGVCFs with the “-stand-call-conf” threshold set to “0.0” to emit monomorphic reference calls. GATK v4.2.0.0 SelectVariants was used to restrict these “rescued” calls to invariant sites and the same site-level and individual-level filtering thresholds described above (excluding “QualByDepth”) were applied to obtain a collection of high-confidence, monomorphic calls. These rescued invariant calls were then merged with their respective variant calls, yielding a collection of datasets that represent the high-confidence union of variant sites observed across the three cohorts.

### Code Availability

The pipeline used to process raw sequencing reads, perform the alignment, and conduct the variant calling and variant quality filtering is available on GitHub ([https://github.com/PayseurLabUWMadison/boha\\_demography/tree/main/variant\\_calling](https://github.com/PayseurLabUWMadison/boha_demography/tree/main/variant_calling)).

### Exclusion of Close Relatives

To mitigate the potential biases that closely related individuals can introduce into population genetic analyses (Wang 2017), we estimated kinship coefficients between all pairs of individuals within each population sample using KING v2.2.8 (Manichaikul et al. 2010). Prior to kinship analysis, each variant call set was pruned for LD by removing SNPs with pairwise  $r^2 > 0.1$  in 50 bp sliding windows using the “-indep-pairwise 50 10 0.1” option in plink v1.90 (Chang et al. 2015). This LD-pruning yielded a total of 1,987,614 SNPs in the Bumpkin Island cohort, 1,253,128 in the Peddocks Island cohort, and 2,240,885 in the mainland World’s End cohort remaining for kinship analysis. The KING-Robust algorithm was used to estimate pairwise kinship coefficients from the LD-pruned autosomal variant calls and relationships were inferred from these estimates using the bounds provided by KING (Manichaikul et al. 2010). We used the “relatednessFilter” function of the R package plinkQC v0.3.4 to identify the maximum set of unrelated individuals (defined here as pairs  $\geq$  3rd degree relatives) within each sample.

### Analysis of Population Structure

We analyzed population structure both between and within geographic locations by clustering individuals according to genetic similarity and by inferring individual ancestry proportions. Since both analyses assume that different SNPs experience free recombination, we conducted the same LD pruning described above within the Bumpkin, Peddocks, and World’s End cohorts. After excluding close relatives, this LD-pruning left a total of 92,099 SNPs in the Bumpkin cohort, 112,369 in the Peddocks cohort,

and 213,660 in the World's End cohort remaining for population structure analysis. PCA was performed on the combined, LD-pruned, autosomal call sets using the “-pca” option in plink v1.90 (Chang et al. 2015). Ancestry proportions were inferred either for each population sample separately or for the combined cohort with ADMIXTURE v1.3.0 (Alexander et al. 2009) using the same collection of SNPs. Five-fold cross-validation was performed for a range of ancestral components ( $k = 1$  to  $k = 5$ ) using the “-cv” option and individual ancestry proportions were inferred for the  $k$  value that yielded the lowest cross-validation error.

## Demographic Inference

### Model Fitting and Parameter Estimation

We estimated demographic parameters using the maximum likelihood framework implemented in *Moments* v1.1.15 (Jouanous et al. 2017), which employs a recursion approach to compute the expected sample site frequency spectrum under a given model. To simplify the space of possible models and to ensure sufficient information in each shared allele frequency bin, we restricted our analysis to two-population models of divergence from a common ancestor. Tested models included either discrete or continuous changes in  $N_e$  for both island and mainland populations (supplementary fig. S7a, Supplementary Material online). On top of these size changes, we fit two additional classes of models that allowed for either symmetric or asymmetric migration between sampled populations (supplementary fig. S7b and c, Supplementary Material online, respectively). Given our limited sampling of the archipelago, we considered the possibility that observed variation might also reflect contributions made by unsampled populations inhabiting the many islands and peninsulas that comprise the Boston Harbor. To address this possibility, we attempted to fit a third class of migration models that allowed for symmetric migration between sampled populations and an unsampled “ghost” population (supplementary fig. S7d, Supplementary Material online). To reduce the number of parameters in the ghost migration models, we held the ghost population's  $N_e$  constant at the ancestral  $N_e$ , assumed that it diverged from the ancestral population at the same time as the sampled populations, and assumed symmetric migration rates between the ghost population and focal population. For these three-population ghost migration models, we used the “marginalize” function to obtain the expected sample frequency spectrum for the two sampled populations.

To circumvent issues of ancestral state misidentification, we fit these models to the folded jSFS constructed from putatively neutral SNPs for each pair of populations. To increase the accuracy of estimates of sample allele frequencies, all sites containing missing genotype information for one or more individuals were excluded from

demographic analysis. To mitigate the impact of direct and linked selection on observed allele frequencies, we masked annotated genes using the NCBI RefSeq annotations of the *P. leucopus* (UCI\_PerLeu\_2.1; GenBank accession: GCA\_004664715.2; Long et al. 2019) reference assembly. Starting with our high-confidence variant callsets described above, we used the NCBI RefSeq track (hub\_2100979\_ncbiRefSeq; Pruitt et al. 2014) of the UCSC Genome Browser (<http://genome.ucsc.edu>; Karolchik et al. 2004) to exclude SNPs falling within 100 bp of predicted and curated genes. Given the potential for the effects of linked selection to extend farther distances from these functional regions, we tested the impact of more stringent distance-based gene masking on the observed jSFS for each pair of populations (supplementary fig. S18, Supplementary Material online). We found that extending the masked distance beyond 100 bp upstream and downstream of gene annotations to distances of 5 kbp, 50 kbp, and even 100 kbp had little impact on the genome-wide jSFS constructed from the remaining SNPs (supplementary fig. S18, Supplementary Material online). Given this finding, we limited our masking to 100 bp surrounding annotated genes to retain the maximal number of SNPs for inference. This approach left a total of 4,219,108 Bumpkin SNPs, 5,063,432 Peddocks SNPs, and 6,707,809 World's End SNPs remaining for inference.

For each of the tested models, we used the BFGS optimization algorithm implemented in *Moments* to find the parameter values that maximized the likelihood of the observed jSFS. We permuted starting parameter values across 10,000 independent searches and retained the top three parameter combinations that yielded the highest likelihood for each model. Following model fitting, we employed a hierarchical approach to selecting between contrasting demographic models. First, we evaluated parameter convergence and eliminated models for which the coefficient of variation for any parameter exceeded 0.2 across the top three highest likelihood estimates. Next, we compared likelihoods between models, either directly for models with equivalent numbers of parameters, or through adjusted likelihood ratio tests (Coffman et al. 2016) for models with nested parameters. Finally, for remaining models, we inspected their fit to both the jSFS and the marginal SFS for each component population and proceeded with the simplest model that yielded the best fit to these frequency spectrum partitions.

For selected models, we calculated the ancestral effective population size as  $N_{Anc} = \theta / (4 * \mu * L)$ , where we assumed a mutation rate,  $\mu$ , of  $4e-9$  per-site, per-generation and used an effective sequence length,  $L$ , that represents the total length of sequence from which we could have called variants. To obtain an approximate  $L$  of 210,600,000 bp, we multiplied the total length of the autosomal portion of the *P. leucopus* reference assembly (2,336,931,804 bp)

by the average percent of called SNPs that survived our quality-based, accessibility-based, and neutrality-based filtering (9.01%). Population size and time parameters were scaled by  $N_{\text{Anc}}$  and  $2 \times N_{\text{Anc}}$  to obtain these parameter estimates in units of individuals and generations, respectively. LD among SNPs used to construct the jSFS violates the independence assumption of the Poisson Random Field framework used to compute model likelihoods (Sawyer and Hartl 1992) and precludes the use of standard techniques to estimate the uncertainty of model parameters. As an alternative to the computationally intensive approach of conventional bootstrapping, we used the Godambe Information Matrix methods described in Coffman et al. (2016) to estimate parameter uncertainties in the presence of this nonindependence. Since certain population size and time parameters are functions of multiple estimated quantities, we used uncertainty propagation techniques (Ku 1966) to obtain the standard deviations of these parameter estimates.

### Predictive Simulations

We evaluated the fit of the inferred demographic histories to other patterns of neutral variation by conducting predictive simulations with *msprime* v1.2.0 (Baumdicker et al. 2022). For each population pair, we simulated genomic datasets using the maximum likelihood parameter estimates obtained for the best-fit demographic model. These genomic datasets were constructed from 1,000 independent simulations of a 1 Mbp genomic element. Simulated sample sizes were set to match those of the empirical data.

Since we lack estimates of the mutation and recombination rate in *P. leucopus*, our choice of these parameter values was guided by existing estimates for house mice (which have a mutation rate of  $3.9 \times 10^{-9}$  per generation (Lindsay et al. 2019) and a recombination rate of 0.58 cM/Mb (Cox et al. 2009)). Centering our assumed mutation rate on the house mouse estimate, we conducted predictive simulations within a range of rates above and below this value:  $5 \times 10^{-9}$ ,  $4 \times 10^{-9}$ , and  $3 \times 10^{-9}$ . We found that a mutation rate of  $4 \times 10^{-9}$  provided the best fit with respect to diversity-related summary statistics. To achieve a  $\theta/\rho = 1$ , we fixed our assumed recombination rate to match this mutation rate at 0.4 cM/Mb.

For both the simulated and empirical neutral datasets, we computed nucleotide diversity  $\pi$  (Tajima 1983), Tajima's  $D$  (Tajima 1989), and  $F_{\text{ST}}$  based on nucleotide diversity (Hudson et al. 1992) in nonoverlapping 5 kbp physical windows using *scikit-allel* v1.3.6 (<https://github.com/cggh/scikit-allel>). To compare the decay of LD between simulated and observed data, we used the “-r2” option in *plink* v1.90 (Chang et al. 2015) to measure pairwise  $r^2$  between a random subset of SNPs sampled from simulated and empirical chromosomes. We restricted this analysis to SNPs separated by at most 1 Mbp by setting the

“-ld-window” and “-ld-window-kb” arguments to 1,000,000 and 1,000, respectively. Singleton variants were excluded from LD calculations by setting the “-mac” argument to 1. The rigorous quality-based, accessibility-based, and neutrality-based filtering that we performed on the empirical data introduced substantial heterogeneity in the number of “observable” sites that comprise each 5 kbp physical window. To account for this, we computed per-site statistics like  $\pi$  based on the total number of observed sites using the “is\_accessible” argument in *scikit-allel* calculations. For SNP-based statistics like Tajima's  $D$  and  $F_{\text{ST}}$ , we note that this heterogeneity could be a major source of between-window variability in summary statistic estimates from the empirical data.

### Code Availability

The scripts used to conduct demographic inference with *Moments* and to perform the predictive simulations with *msprime* are available on GitHub ([https://github.com/PayseurLabUWMadison/boha\\_demography/tree/main/demographic\\_inference](https://github.com/PayseurLabUWMadison/boha_demography/tree/main/demographic_inference)).

### Singleton Analysis

Inspection of variant positions along chromosomes revealed clustering of singleton variants by mouse and by genomic location in the island samples. To explore the capacity of migration to generate this pattern, we used chromosome-scale simulations conducted with *msprime* v1.2.0 (Baumdicker et al. 2022). Given that best-fit demographic models based on the jSFS did not include migration, we augmented these best-fit demographic histories with gene flow from an unsampled “ghost” population (supplementary fig. S19, Supplementary Material online). Our aim was to capture a range of scenarios that might be relevant to migration that occurs locally within the archipelago. To this end, we varied both the strength of migration between the focal populations and ghost population (spanning rates of  $1 \times 10^{-3}$ ,  $1 \times 10^{-4}$ , and  $1 \times 10^{-5}$ ) as well as time at which migration begins following population divergence (supplementary fig. S19, Supplementary Material online). These intervals of migration are expressed as the most recent proportion of the split time (either 0.1%, 1%, 10%, or 100% of the inferred split time between a given pair of populations). To reduce the number of tested migration models, we made several simplifying assumptions. First, we assumed that the divergence of the ghost population from the ancestral population coincided with the divergence of the sampled populations, and we held the ghost  $N_e$  constant at the ancestral  $N_e$ . Second, we assumed that both sampled populations experienced the same amount of migration from the same ghost population. Third, we assumed that migration between the sampled and ghost population occurred in a continuous and symmetric fashion into the present. For each migration regime, we simulated

25 independent 100 Mbp genomic elements using the parameter estimates from the best-fit demographic model for each pair of populations. For each simulation, we assumed a uniform recombination and mutation rate of  $4e-9$ . Simulated sample sizes were set to match those of the empirical data. For comparison, we generated additional chromosome-scale datasets in the absence of migration.

We used the chi-squared test statistic to quantify the deviations of individual singleton contributions from the null hypothesis that singletons are contributed equally among individuals. For each of the simulated and empirical chromosomes, we used R v4.2.1 (R Core Team 2022) to compute the chi-squared test statistic from the observed and expected individual singleton counts. To evaluate the impact of migration on the jSFS between island and mainland populations, we used scikit-allel v1.3.6 (<https://github.com/cggh/scikit-allel>) to compute joint allele frequencies from datasets simulated with and without ghost migration.

To examine the combined effect of population size history and migration on the genealogies underlying the singleton pattern, we revisited the simulations of Bumpkin Island chromosomes conducted using the best-fit parameter estimates from the Bumpkin-World's End demographic history (Fig. 3) with migration parameters from the Recent/Strong regime featured in Fig. 6b. We prioritized this model because it involves the most extreme reduction in  $N_e$  for the island population (Fig. 3) and provided the closest qualitative fit to the observed singleton tracts (Figs. 5a and 6b). We used tskit v0.5.8 (Kelleher et al. 2018) to inspect the properties of marginal trees spanning a single 100 Mbp Bumpkin Island chromosome. We used the “migrations” function to record the breakpoints of migrant ancestry along the chromosome. Then, we used the “trees” function to iterate over all of the marginal trees in the tree sequence. For each of the 1,010,931 trees, we compared the tree boundaries to the recorded breakpoints of migrant ancestry to determine if the tree overlapped one or more migrant haplotypes (classified as “with migration”) (Fig. 7a). Trees were classified as “no migration” if they did not span migrant tracts (Fig. 7a). We computed the branch length quantities illustrated in Fig. 7b for each tree. We also recorded the TMRCA of each tree along with the population in which the common ancestor originated.

### Code Availability

The scripts used to perform the migration simulations with *msprime* and analyze the simulation output with scikit-allel and tskit are available on GitHub ([https://github.com/PayseurLabUWMadison/boha\\_demography/tree/main/migration\\_analysis](https://github.com/PayseurLabUWMadison/boha_demography/tree/main/migration_analysis))

### Estimating Contemporary Effective Population Sizes

We estimated contemporary effective population sizes in the Boston Harbor using two LD-based approaches,

*currentNe* (Santiago et al. 2024) and GONE (Santiago et al. 2020). Observed levels of LD can be influenced by the structure of relatedness in a sample, and there is evidence that excluding relatives may artificially inflate estimates of contemporary effective population size obtained by these methods (Waples and Anderson 2017; Santiago et al. 2024). Because we cannot know whether the abundance of close relatives we identified (supplementary fig. S1, Supplementary Material online) reflects random sampling or an artifact of trapping procedure, we chose to conduct these analyses on both the complete sample and on the subset of unrelated individuals from each population (obtained by excluding relationships  $\leq$  2nd degree). For all analyses of the Bumpkin Island population, we additionally excluded individual B19 due to its genetic divergence from the rest of the sample (supplementary figs. S2, S4, and S5, Supplementary Material online). We used the same putatively neutral, high-confidence, autosomal SNP callset used for demographic inference with *Moments* as the input to these analyses (see Sequence Data Processing and Variant Calling and Demographic Inference sections). All sites with missing genotype information for one or more individuals were excluded from downstream analyses.

For each dataset, we obtained point estimates of  $N_e$  reflecting the few most recent generations with *currentNe* (Santiago et al. 2024). For computational tractability, we used the “-s” argument to restrict pairwise measurements of LD to 50,000 randomly sampled SNPs.  $N_e$  estimates were based only on pairs of SNPs located on different chromosomes. To account for the impact of close relatives, we used the “-k” option to jointly estimate  $N_e$  and  $k$ , the average number of full siblings each individual has in the sample (Santiago et al. 2024). Consistent with our KING analyses (supplementary fig. S1, Supplementary Material online), this parameter yielded estimates of 0.53, 0.43, and 0.19 for the complete Bumpkin Island, Peddocks Island, and mainland World's End samples, respectively. As expected, we obtained  $k$  values of 0 for each sample of unrelated individuals. 90% confidence intervals for each point estimate were obtained using the artificial neural network (ANN) implemented in the *currentNe* software (Santiago et al. 2024).

To reconstruct the recent trajectory of  $N_e$  in each population, we analyzed each dataset with GONE (Santiago et al. 2020). We used plink v1.90 (Chang et al. 2015) to convert input SNP callsets from VCF to plink text files (.ped and .map) with the “-recode” option. Singleton variants were excluded from each analysis by setting the “-mac” argument to 1. For all GONE analyses, we used the default settings for unphased genotypes and specified a uniform recombination rate of 0.4 cM/Mb to match the rate assumed for predictive simulations (see Demographic Inference section). We conducted 15 replicate analyses for each dataset (with and without close relatives) using



disjoint subsets of SNPs. For each replicate analysis, we report the geometric mean  $N_e$  estimate for each time interval obtained across 40 independent optimizations. In contrast to *currentNe*, GONE does not explicitly account for the presence of close relatives in the dataset (Santiago et al. 2024). Despite this, the impact of close relatives is expected to be negligible, since they exert greater influence on LD between unlinked markers than between the closely linked markers used by GONE to estimate the recent  $N_e$  trajectory (Santiago et al. 2024).

## Supplementary Material

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

## Acknowledgments

We thank members of the Payseur lab for their helpful input on this work. We thank the National Park Service for assistance with boat transportation, mouse capture, and sample collection. Computational analyses were conducted using resources provided by the University of Wisconsin-Madison's Center for High Throughput Computing.

## Author Contributions

E.K.H., B.A.P., and L.E.N. designed the study. L.E.N. acquired the samples used for the study. E.K.H. conducted the research with supervision from B.A.P. E.K.H. and B.A.P. wrote the manuscript with input from L.E.N.

## Funding

This research was funded by National Institutes of Health grants R01GM100426 and R35GM139412 (to B.A.P.). E.K.H. was partially supported by the National Institutes of Health Graduate Training Grant in Genetics at the University of Wisconsin-Madison (T32GM007133).

## Data Availability

The code used to conduct all analyses is provided on GitHub ([https://github.com/PayseurLabUWMadison/boha\\_demography](https://github.com/PayseurLabUWMadison/boha_demography)). Raw sequencing reads are available from NCBI's Sequence Read Archive under the BioProject accession PRJNA1231036. A static snapshot of the code is also available from Dryad (<https://doi.org/10.5061/dryad.pk0p2nh08>).

## Literature Cited

Adams AM, Hudson RR. Maximum-likelihood estimation of demographic parameters using the frequency spectrum of unlinked single-nucleotide polymorphisms. *Genetics*. 2004;168(3):1699–1712. <https://doi.org/10.1534/genetics.104.030171>.

- Adler GH, Levins R. The Island Syndrome in Rodent Populations. *The Quarterly Review of Biology*. 1994;69(4):473–490. <https://doi.org/10.1086/418744>.
- Alexander DH, Novembre J, Lange K. Fast model-based estimation of ancestry in unrelated individuals. *Genome Res*. 2009;19(9):1655–1664. <https://doi.org/10.1101/gr.094052.109>.
- Aubrey DG. Geomorphological investigation of spectacle island. Appendix A. In G. McHargue editor. *The spectacle island site: middle to late woodland adaptations in Boston harbor, Suffolk county, Massachusetts, central artery/tunnel project, Boston Massachusetts*, Vol. 2. Timelines, Inc. Massachusetts Highway Department. Report on file (25-1935, Vol. 2 of 2). Boston, MA: Massachusetts Historical Commission; 1996.
- Baier F, Hoekstra HE. The genetics of morphological and behavioural island traits in deer mice. *Proceedings of the Royal Society B: Biological Sciences*. 2019;286(1914):20191697. <https://doi.org/10.1098/rspb.2019.1697>.
- Baumdicker F, Bisschop G, Goldstein D, Gower G, Ragsdale AP, Tsambos G, Zhu S, Eldon B, Ellerman EC, Galloway JG, et al. Efficient ancestry and mutation simulation with msprime 1.0. *Genetics*. 2022;220(3):iyab229. <https://doi.org/10.1093/genetics/iyab229>.
- Bedford NL, Hoekstra HE. *Peromyscus* mice as a model for studying natural variation. *eLife*. 2015;4:e06813. <https://doi.org/10.7554/eLife.06813>.
- Beichman AC, Huerta-Sanchez E, Lohmueller KE. Using genomic data to infer historic population dynamics of nonmodel organisms. *Annu Rev Ecol Evol Syst*. 2018;49(1):433–456. <https://doi.org/10.1146/annurev-ecolsys-110617-062431>.
- Beichman AC, Kalhori P, Kyriazis CC, DeVries AA, Nigenda-Morales S, Heckel G, Schramm Y, Moreno-Estrada A, Kennett DJ, Hylkema M, et al. Genomic analyses reveal range-wide devastation of sea otter populations. *Mol Ecol*. 2021;32(2):281–298. <https://doi.org/10.1111/mec.16334>.
- Beichman AC, Phung TN, Lohmueller KE. Comparison of single genome and allele frequency data reveals discordant demographic histories. *G3 (Bethesda)*. 2017;7(11):3605–3620. <https://doi.org/10.1534/g3.117.300259>.
- Benítez-López A, Santini L, Gallego-Zamorano J, Milá B, Walkden P, Huijbregts MA, Tobias JA. The island rule explains consistent patterns of body size evolution in terrestrial vertebrates. *Nat Ecol Evol*. 2021;5(6):768–786. <https://doi.org/10.1038/s41559-021-01426-y>.
- Benson G. Tandem repeats finder: a program to analyze DNA sequences. *Nucleic Acids Res*. 1999;27(2):573–580. <https://doi.org/10.1093/nar/27.2.573>.
- Bergland AO, Behrman EL, O'Brien KR, Schmidt PS, Petrov DA. Genomic evidence of rapid and stable adaptive oscillations over seasonal time scales in *Drosophila*. *PLoS Genet*. 2014;10(11):e1004775. <https://doi.org/10.1371/journal.pgen.1004775>.
- Berry RJ. The evolution of an island population of the house mouse. *Evolution*. 1964;18(3):468–483. <https://doi.org/10.2307/2406357>.
- Berry RJ, Bonner WN, Peters J. Natural selection in house mice (*Mus musculus*) from South Georgia (South Atlantic Ocean). *J Zool*. 1979;189(3):385–398. <https://doi.org/10.1111/j.1469-7998.1979.tb03970.x>.
- Berry RJ, Jakobson ME, Peters J. The house mice of the Faroe Islands: a study in microdifferentiation. *J Zool*. 1978a;185(1):73–92. <https://doi.org/10.1111/j.1469-7998.1978.tb03314.x>.
- Berry RJ, Peters J, Aarde RV. Sub-Antarctic house mice: colonization, survival and selection. *J Zool*. 1978b;184(1):127–141. <https://doi.org/10.1111/j.1469-7998.1978.tb03270.x>.
- Blum MG, François O. Minimal clade size and external branch length under the neutral coalescent. *Adv Appl Probab*. 2005;37(3):647–662. <https://doi.org/10.1239/aap/1127483740>.



- Boitard S, Rodríguez W, Jay F, Mona S, Austerlitz F. Inferring population size history from large samples of genome-wide molecular data—an approximate Bayesian computation approach. *PLoS Genet.* 2016;12(3):e1005877. <https://doi.org/10.1371/journal.pgen.1005877>.
- Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics.* 2014;30(15):2114–2120. <https://doi.org/10.1093/bioinformatics/btu170>.
- Browning SR, Browning BL. Identity by descent between distant relatives: detection and applications. *Annu Rev Genet.* 2012;46(1):617–633. <https://doi.org/10.1146/annurev-genet-110711-155534>.
- Bürger R, Lynch M. Evolution and extinction in a changing environment: a quantitative-genetic analysis. *Evolution.* 1995;49(1):151–163. <https://doi.org/10.1111/j.1558-5646.1995.tb05967.x>.
- Burt WH. Territorial behavior and populations of some small mammals in southern Michigan. Vol. 45. Ann Arbor, MI: Miscellaneous Publications of the University of Michigan Museum of Zoology; 1940. p. 16–41.
- Caliebe A, Neininger R, Krawczak M, Rösler U. On the length distribution of external branches in coalescence trees: genetic diversity within species. *Theor Popul Biol.* 2007;72(2):245–252. <https://doi.org/10.1016/j.tpb.2007.05.003>.
- Case TJ. A general explanation for insular body size trends in terrestrial vertebrates. *Ecology.* 1978;59(1):1–18. <https://doi.org/10.2307/1936628>.
- Chang CC, Chow CC, Tellier LC, Vattikuti S, Purcell SM, Lee JJ. Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience.* 2015;4(1):s13742–s13015. <https://doi.org/10.1186/s13742-015-0047-8>.
- Coffman AJ, Hsieh PH, Gravel S, Gutenkunst RN. Computationally efficient composite likelihood statistics for demographic inference. *Mol Biol Evol.* 2016;33(2):591–593. <https://doi.org/10.1093/molbev/msv255>.
- Cox A, Ackert-Bicknell CL, Dumont BL, Ding Y, Bell JT, Brockmann GA, Wergedal JE, Bult C, Paigen B, Flint J, et al. A new standard genetic map for the laboratory mouse. *Genetics.* 2009;182(4):1335–1344. <https://doi.org/10.1534/genetics.109.105486>.
- Crisci JL, Poh YP, Mahajan S, Jensen JD. The impact of equilibrium assumptions on tests of selection. *Front Genet.* 2013;4:235. <https://doi.org/10.3389/fgene.2013.00235>.
- Danecek P, Bonfield JK, Liddle J, Marshall J, Ohan V, Pollard MO, Whitwham A, Keane T, McCarthy SA, Davies RM, et al. Twelve years of SAMtools and BCFtools. *Gigascience.* 2021;10(2):giab008. <https://doi.org/10.1093/gigascience/giab008>.
- Disanto F, Wiehe T. Measuring the external branches of a Kingman tree: a discrete approach. *Theor Popul Biol.* 2020;134:92–105. <https://doi.org/10.1016/j.tpb.2020.05.002>.
- Elliman T. Vascular flora and plant communities of the Boston harbor islands. *Northeast Nat (Steuben).* 2005;12(sp3):49–74. [https://doi.org/10.1656/1092-6194\(2005\)12\[49:VFAPCO\]2.0.CO;2](https://doi.org/10.1656/1092-6194(2005)12[49:VFAPCO]2.0.CO;2).
- Foster JB. Evolution of mammals on islands. *Nature.* 1964;202(4929):234–235. <https://doi.org/10.1038/202234a0>.
- Garud NR, Messer PW, Buzbas EO, Petrov DA. Recent selective sweeps in North American *Drosophila melanogaster* show signatures of soft sweeps. *PLoS Genet.* 2015;11(2):e1005004. <https://doi.org/10.1371/journal.pgen.1005004>.
- Gray MM, Wegmann D, Haas RJ, White MA, Gabriel SI, Searle JB, Cuthbert RJ, Ryan PG, Payseur BA. Demographic history of a recent invasion of house mice on the isolated Island of Gough. *Mol Ecol.* 2014;23(8):1923–1939. <https://doi.org/10.1111/mec.12715>.
- 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(10):e1000695. <https://doi.org/10.1371/journal.pgen.1000695>.
- Harris K, Nielsen R. Inferring demographic history from a spectrum of shared haplotype lengths. *PLoS Genet.* 2013;9(6):e1003521. <https://doi.org/10.1371/journal.pgen.1003521>.
- Harris SE, Munshi-South J. Signatures of positive selection and local adaptation to urbanization in white-footed mice (*Peromyscus leucopus*). *Mol Ecol.* 2017;26(22):6336–6350. <https://doi.org/10.1111/mec.14369>.
- Hudson RR, Slatkin M, Maddison WP. Estimation of levels of gene flow from DNA sequence data. *Genetics.* 1992;132(2):583–589. <https://doi.org/10.1093/genetics/132.2.583>.
- Johri P, Aquadro CF, Beaumont M, Charlesworth B, Excoffier L, Eyre-Walker A, Keightley PD, Lynch M, McVean G, Payseur BA, et al. Recommendations for improving statistical inference in population genomics. *PLoS Biol.* 2022;20(5):e3001669. <https://doi.org/10.1371/journal.pbio.3001669>.
- Jouganous J, Long W, Ragsdale AP, Gravel S. Inferring the joint demographic history of multiple populations: beyond the diffusion approximation. *Genetics.* 2017;206(3):1549–1567. <https://doi.org/10.1534/genetics.117.200493>.
- Kardos M, Armstrong EE, Fitzpatrick SW, Hauser S, Hedrick PW, Miller JM, Tallmon DA, Funk WC. The crucial role of genome-wide genetic variation in conservation. *Proc Natl Acad Sci U S A.* 2021;118(48):e2104642118. <https://doi.org/10.1073/pnas.2104642118>.
- Karolchik D, Hinrichs AS, Furey TS, Roskin KM, Sugnet CW, Haussler D, Kent WJ. The UCSC table browser data retrieval tool. *Nucleic Acids Res.* 2004;32(suppl\_1):D493–D496. <https://doi.org/10.1093/nar/gkh103>.
- Kaye CA. Bedrock and quaternary geology of the Boston area, Massachusetts. In: Leggett RF, editors. *Geology Under Cities*. Vol. 5. Boulder, CO: Geological Society of America Reviews in Engineering Geology; 1982. p. 25–40.
- Keinan A, Clark AG. Recent explosive human population growth has resulted in an excess of rare genetic variants. *Science.* 2012;336(6082):740–743. <https://doi.org/10.1126/science.1217283>.
- Kelleher J, Thornton KR, Ashander J, Ralph PL. Efficient pedigree recording for fast population genetics simulation. *PLoS Comput Biol.* 2018;14(11):e1006581. <https://doi.org/10.1371/journal.pcbi.1006581>.
- Kim BY, Huber CD, Lohmueller KE. Inference of the distribution of selection coefficients for new nonsynonymous mutations using large samples. *Genetics.* 2017;206(1):345–361. <https://doi.org/10.1534/genetics.116.197145>.
- Ku HH. Notes on the use of propagation of error formulas. *J Res Natl Bur Stand.* 1966;70C(4):161–187. <https://doi.org/10.6028/jres.070c.025>.
- Kyriazis CC, Serieys LE, Bishop JM, Drouilly M, Viljoen S, Wayne RK, Lohmueller KE. The influence of gene flow on population viability in an isolated urban caracal population. *Mol Ecol.* 2024;33(9):e17346. <https://doi.org/10.1111/mec.17346>.
- Lande R, Barrowclough GF. Effective population size, genetic variation, and their use in population management. In: Soulé ME, editors. *Viable populations for conservation*. Vol. 87. New York, NY; 1987. p. 87–124. <https://doi.org/10.1017/CBO9780511623400.007>.
- Li H. Toward better understanding of artifacts in variant calling from high-coverage samples. *Bioinformatics.* 2014;30(20):2843–2851. <https://doi.org/10.1093/bioinformatics/btu356>.
- Li H, Durbin R. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics.* 2009;25(14):1754–1760. <https://doi.org/10.1093/bioinformatics/btp324>.
- Li H, Durbin R. Inference of human population history from individual whole-genome sequences. *Nature.* 2011;475(7357):493–496. <https://doi.org/10.1038/nature10231>.
- Lindsay SJ, Rahbari R, Kaplanis J, Keane T, Hurles ME. Similarities and differences in patterns of germline mutation between mice and

- humans. *Nat Commun.* 2019;10(1):4053. <https://doi.org/10.1038/s41467-019-12023-w>.
- Lomolino MV. Body size of mammals on islands: the island rule reexamined. *Am Nat.* 1985;125(2):310–316. <https://doi.org/10.1086/284343>.
- Lomolino MV. Body size evolution in insular vertebrates: generality of the island rule. *J Biogeography.* 2005;32(10):1683–1699. <https://doi.org/10.1111/j.1365-2699.2005.01314.x>.
- Long AD, Baldwin-Brown J, Tao Y, Cook VJ, Balderrama-Gutierrez G, Corbett-Detig R, Mortazavi A, Barbour AG. The genome of *Peromyscus leucopus*, natural host for Lyme disease and other emerging infections. *Sci Adv.* 2019;5(7):eaaw6441. <https://doi.org/10.1126/sciadv.aaw6441>.
- Lotterhos KE, Whitlock MC. Evaluation of demographic history and neutral parameterization on the performance of FST outlier tests. *Mol Ecol.* 2014;23(9):2178–2192. <https://doi.org/10.1111/mec.12725>.
- Luedtke BE. The archaeology of Thompson island. Boston, MA: University of Massachusetts Boston Anthropology Faculty Publication Series; 1996. p.14.
- Luedtke BE, Rosen PS. Archaeological geology on long island, Boston harbor. In: Field Trip Guidebook for the Northeastern United States, Vol. 2. Geological Society of America, Boston, MA; 1993.
- Luikart G, Ryman N, Tallmon DA, Schwartz MK, Allendorf FW. Estimation of census and effective population sizes: the increasing usefulness of DNA-based approaches. *Conserv Genet.* 2010;11(2):355–373. <https://doi.org/10.1007/s10592-010-0050-7>.
- Lynch M, Conery J, Bürger R. Mutation accumulation and the extinction of small populations. *Am Nat.* 1995;146(4):489–518. <https://doi.org/10.1086/285812>.
- Machado HE, Bergland AO, Taylor R, Tilk S, Behrman E, Dyer K, Fabian DK, Flatt T, González J, Karasov TL, et al. Broad geographic sampling reveals the shared basis and environmental correlates of seasonal adaptation in *Drosophila*. *Elife.* 2021;10:e67577. <https://doi.org/10.7554/eLife.67577>.
- Manichaikul A, Mychaleckyj JC, Rich SS, Daly K, Sale M, Chen WM. Robust relationship inference in genome-wide association studies. *Bioinformatics.* 2010;26(22):2867–2873. <https://doi.org/10.1093/bioinformatics/btq559>.
- Martínková N, Barnett R, Cucchi T, Struchen R, Pascal M, Pascal M, Fischer MC, Higham T, Brace S, Ho SYW. Divergent evolutionary processes associated with colonization of offshore islands. *Mol Ecol.* 2013;22(20):5205–5220. <https://doi.org/10.1111/mec.12462>.
- McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernysky A, Garimella K, Altshuler D, Gabriel S, Daly M, et al. The genome analysis toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* 2010;20(9):1297–1303. <https://doi.org/10.1101/gr.107524.110>.
- Morgulis A, Gertz EM, Schäffer AA, Agarwala R. WindowMasker: window-based masker for sequenced genomes. *Bioinformatics.* 2006;22(2):134–141. <https://doi.org/10.1093/bioinformatics/bti774>.
- Morin PA, Archer FI, Avila CD, Balacco JR, Bukhman YV, Chow W, Fedrigo O, Formenti G, Fronczek JA, Functammasan A, et al. Reference genome and demographic history of the most endangered marine mammal, the vaquita. *Mol Ecol Resour.* 2021;21(4):1008–1020. <https://doi.org/10.1111/1755-0998.13284>.
- Munshi-South J. Urban landscape genetics: canopy cover predicts gene flow between white-footed mouse (*Peromyscus leucopus*) populations in New York City. *Mol Ecol.* 2012;21(6):1360–1378. <https://doi.org/10.1111/j.1365-294X.2012.05476.x>.
- Nadachowska-Brzyska K, Konczal M, Babik W. Navigating the temporal continuum of effective population size. *Methods Ecol Evol.* 2022;13(1):22–41. <https://doi.org/10.1111/2041-210X.13740>.
- Nolfo-Clements L. Mammals of the Boston harbor islands: permanent and ephemeral residents. *Northeast Nat (Steuben).* 2018;25(sp9):77–89. <https://doi.org/10.1656/045.025.s901>.
- Nolfo-Clements L, Butcher R, Leite M, Clements M. Evidence of the island rule and microevolution in white-footed mice (*Peromyscus leucopus*) in an urban harbor archipelago. *Mamm Res.* 2017;62(4):423–430. <https://doi.org/10.1007/s13364-017-0324-2>.
- Nolfo-Clements L, Clements M. Small-mammal population dynamics and habitat use on bumpkin island in the Boston harbor. *Northeast Nat (Steuben).* 2015;22(1):NENHC-14–NENHC-25. <https://doi.org/10.1656/045.022.0105>.
- Novo I, Ordás P, Moraga N, Santiago E, Quesada H, Caballero A. Impact of population structure in the estimation of recent historical effective population size by the software GONE. *Genet Sel Evol.* 2023;55(1):86. <https://doi.org/10.1186/s12711-023-00859-2>.
- Olmstead Center for Landscape Preservation, National Park Service. Cultural landscape report for the Boston harbor islands; 2017.
- Osgood WH. Revision of the mice of the American genus *Peromyscus*. *North American Fauna.* 1909;28:1–285. <https://doi.org/10.3996/nafa.28.0001>.
- Palamara PF, Lencz T, Darvasi A, Pe'er I. Length distributions of identity by descent reveal fine-scale demographic history. *Am J Hum Genet.* 2012;91(5):809–822. <https://doi.org/10.1016/j.ajhg.2012.08.030>.
- Palkovacs EP. Explaining adaptive shifts in body size on islands: a life history approach. *Oikos.* 2003;103(1):37–44. <https://doi.org/10.1034/j.1600-0706.2003.12502.x>.
- Pavlidis P, Jensen JD, Stephan W. Searching for footprints of positive selection in whole-genome SNP data from nonequilibrium populations. *Genetics.* 2010;185(3):907–922. <https://doi.org/10.1534/genetics.110.116459>.
- Pfeifer SP. From next-generation resequencing reads to a high-quality variant data set. *Heredity (Edinb).* 2017;118(2):111–124. <https://doi.org/10.1038/hdy.2016.102>.
- Plagnol V, Wall JD. Possible ancestral structure in human populations. *PLoS Genet.* 2006;2(7):e105. <https://doi.org/10.1371/journal.pgen.0020105>.
- Poplin R, Ruano-Rubio V, DePristo MA, Fennell TJ, Carneiro MO, Van der Auwera GA, Kling DE, Gauthier LD, Levy-Moonshine A, Roazen D, et al. Scaling accurate genetic variant discovery to tens of thousands of samples. *bioRxiv.* 2017:201178. <https://doi.org/10.1101/201178>, 24 July 2018, preprint: not peer reviewed.
- Pruitt KD, Brown GR, Hiatt SM, Thibaud-Nissen F, Astashyn A, Ermolaeva O, Farrell CM, Hart J, Landrum MJ, McGarvey KM, et al. RefSeq: an update on mammalian reference sequences. *Nucleic Acids Res.* 2014;42(D1):D756–D763. <https://doi.org/10.1093/nar/gkt1114>.
- Richburg JA, Patterson WA III. Historical description of the vegetation of the Boston harbor islands: 1600–2000. *Northeast Nat (Steuben).* 2005;12(sp3):13–30. [https://doi.org/10.1656/1092-6194\(2005\)12\[13:HDOTVO\]2.0.CO;2](https://doi.org/10.1656/1092-6194(2005)12[13:HDOTVO]2.0.CO;2).
- Robinson JD, Coffman AJ, Hickerson MJ, Gutenkunst RN. Sampling strategies for frequency spectrum-based population genomic inference. *BMC Evol Biol.* 2014;14(1):254. <https://doi.org/10.1186/s12862-014-0254-4>.
- Rogers AR. How population growth affects linkage disequilibrium. *Genetics.* 2014;197(4):1329–1341. <https://doi.org/10.1534/genetics.114.166454>.
- Rowe-Rowe DT, Crafford JE. Density, body size, and reproduction of feral house mice on Gough island. *South Afr J Zool.* 1992;27(1):1–5. <https://doi.org/10.1080/02541858.1992.11448252>.
- Santiago E, Caballero A, Köpke C, Novo I. Estimation of the contemporary effective population size from SNP data while accounting

- for mating structure. *Mol Ecol Resour.* 2024;24(1):e13890. <https://doi.org/10.1111/1755-0998.13890>.
- Santiago E, Novo I, Pardiñas AF, Saura M, Wang J, Caballero A. Recent demographic history inferred by high-resolution analysis of linkage disequilibrium. *Mol Biol Evol.* 2020;37(12):3642–3653. <https://doi.org/10.1093/molbev/msaa169>.
- Sawyer SA, Hartl DL. Population genetics of polymorphism and divergence. *Genetics.* 1992;132(4):1161–1176. <https://doi.org/10.1093/genetics/132.4.1161>.
- Schiffels S, Durbin R. Inferring human population size and separation history from multiple genome sequences. *Nat Genet.* 2014;46(8):919–925. <https://doi.org/10.1038/ng.3015>.
- Tajima F. Evolutionary relationship of DNA sequences in finite populations. *Genetics.* 1983;105(2):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(3):585–595. <https://doi.org/10.1093/genetics/123.3.585>.
- Terhorst J, Kamm JA, Song YS. Robust and scalable inference of population history from hundreds of unphased whole genomes. *Nat Genet.* 2017;49(2):303–309. <https://doi.org/10.1038/ng.3748>.
- Thompson EA. Identity by descent: variation in meiosis, across genomes, and in populations. *Genetics.* 2013;194(2):301–326. <https://doi.org/10.1534/genetics.112.148825>.
- Van der Auwera GA, O'Connor BD. Genomics in the cloud: using Docker, GATK, and WDL in terra. Sebastopol, CA: O'Reilly Media; 2020.
- Van Valen L. Patterns and the balance of nature. *Evol Theory.* 1973;1: 31–49.
- Wall JD. Detecting ancient admixture in humans using sequence polymorphism data. *Genetics.* 2000;154(3):1271–1279. <https://doi.org/10.1093/genetics/154.3.1271>.
- Wang J. A new method for estimating effective population sizes from a single sample of multilocus genotypes. *Mol Ecol.* 2009;18(10): 2148–2164. <https://doi.org/10.1111/j.1365-294X.2009.04175.x>.
- Wang J. Effects of sampling close relatives on some elementary population genetics analyses. *Mol Ecol Resour.* 2017;18(1):41–54. <https://doi.org/10.1111/1755-0998.12708>.
- Wang X, Peischl S, Heckel G. Demographic history and genomic consequences of 10,000 generations of isolation in a wild mammal. *Curr Biol.* 2023;33(10):2051–2062.e4. <https://doi.org/10.1016/j.cub.2023.04.042>.
- Waples RS, Anderson EC. Purging putative siblings from population genetic data sets: a cautionary view. *Mol. Ecol.* 2017;26(5): 1211–1224. <https://doi.org/10.1111/mec.14022>.
- Waples RS, England PR. Estimating contemporary effective population size on the basis of linkage disequilibrium in the face of migration. *Genetics.* 2011;189(2):633–644. <https://doi.org/10.1534/genetics.111.132233>.
- Yu A, Munshi-South J, Sargis EJ. Morphological differentiation in white-footed mouse (*Mammalia: Rodentia: Cricetidae: Peromyscus leucopus*) populations from the New York city metropolitan area. *Bull Peabody Mus Nat Hist.* 2017;58(1):3–16. <https://doi.org/10.3374/014.058.0102>.

Associate editor: Kirk Lohmueller