TITLE

The genomic imprint of chromosomal inversions and demographic history in island populations of deer mice

AUTHORS AND AFFILIATIONS

Emma K. Howell[1*], Felix Baier[2,3], Hopi E. Hoekstra[2], Bret A. Payseur[1*]

1. Laboratory of Genetics, University of Wisconsin-Madison, Madison, WI, USA

2. Department of Molecular and Cellular Biology; Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, MA, USA

3. Present address: Max Planck Institute for Brain Research, Frankfurt, Germany

*correspondence: ekhowell@wisc.edu; bret.payseur@wisc.edu

KEYWORDS

Peromyscus maniculatus, inversions, demographic inference, island evolution, balancing selection, whole genome sequencing

ABSTRACT

Populations that colonize islands experience novel selective pressures, fluctuations in size, and changes to their connectivity. Owing to their unique geographic setting, islands can function as natural laboratories in which to examine the interactions between demographic history and natural selection replicated across isolated populations. We used whole genome sequences of wild-caught deer mice (*Peromyscus maniculatus*) from two islands (Saturna and Pender) and one mainland location (Maple Ridge) in the Gulf Islands region of coastal British Columbia to investigate two primary determinants of genome-wide diversity: chromosomal inversions and non-equilibrium demographic history. We found that segregating inversions produce characteristic, large-scale distortions in allele frequencies and linkage disequilibrium that make it possible to identify and characterize them from short-read sequence data. Patterns of variation within and between karyotypes indicate that six inversion polymorphisms have been maintained by a shared history of balancing selection in both island and mainland populations. Whereas the estimated timing of contemporary population splits is consistent with the isolation of island populations from each other following the Last Glacial Maximum, ancestral island and mainland lineages are inferred to have diverged much earlier. These aspects of demographic history suggest that shared inversions existed long ago in a common ancestor or spread via limited gene flow between ancestral island and mainland lineages. Our results raise the possibility that inversions segregating among Gulf Islands populations are on similar evolutionary trajectories, providing a contrast to previous findings in mainland *P. maniculatus* and contributing to the emerging portrait of inversion evolution in this species.

INTRODUCTION

Island colonization alters the evolutionary trajectory of a population in multiple ways. Ecological differences between island and mainland environments in variables such as species composition, habitat diversity, resource availability, and climate impart novel selective challenges on colonizing populations (Losos and Ricklefs 2009). Coincident with these ecological shifts, founder effect bottlenecks amplify the effects of genetic drift and reductions in gene flow driven by island isolation increase differentiation between populations (Losos and Ricklefs 2009). Together, these unique selective and demographic changes that accompany evolution on islands shape the fates of standing variants and new mutations alike.

One source of genetic variation that may be particularly sensitive to these features of island evolution is structural changes to the genome. Compared to single nucleotide mutations, structural variants can be orders of magnitude larger in their genomic scope and change the collinearity of chromosome copies. Chromosomal inversions, which occur when a DNA segment is reinserted in the reverse orientation, can disrupt genes that stretch across breakpoints and alter the proximity of genetic elements within and outside of breakpoints, achieving substantial shifts in genome function and architecture in a single mutational step (reviewed in Berdan et al. 2023). Since Alfred Sturtevant's pioneering discovery of inversions in *Drosophila* (Sturtevant 1921), this class of structural variation has been the focus of intensive evolutionary study aimed at understanding the origin, spread, and maintenance of inversions in populations (Hoffmann and Rieseberg 2008).

Central to the evolutionary significance of inversions is their effect on recombination. Crossing over that occurs within inversion breakpoints can produce unbalanced gametes in individuals that carry both the inverted and standard orientation (McClintock 1931). This underdominant behavior in heterokaryotypes, in addition to potentially deleterious positional and breakpoint effects, suggests that inversions may be more likely to spread in populations experiencing increased genetic drift (Lande 1979; Hedrick 1981; Bush et al. 1977), such as those inhabiting islands. Indeed, accelerated rates of inversion-driven karyotypic evolution in gibbons (Carbone et al. 2014) and annual sunflowers (Burke et al. 2004) have been attributed to dramatic fluctuations in population size. At the same time, inversions can promote adaptive evolution

through their suppression of recombination. By preserving associations between alleles with epistatic fitness effects, inversions can maintain co-adapted variation (Dobzhansky 1948; Dobzhansky 1949; Dobzhansky 1950). Even in the absence of such epistatic interactions, inversions that capture locally adaptive haplotypes can spread when there is gene flow between populations experiencing divergent selective regimes (Kirkpatrick and Barton 2006). Empirical evidence from diverse taxa suggests that inversions are an important source of adaptive divergence between populations (reviewed in Wellenreuther and Bernatchez 2018).

Peromyscus provides a powerful system for examining chromosomal evolution on islands. As one of the most abundant, widespread rodents in North America (Osgood 1909; Bedford and Hoekstra 2015), the range of *Peromyscus* extends to a wide array of coastal islands that line the continent (Redfield 1976; Gill 1980; Adler and Tamarin 1983; Adler et al. 1986; Lucid and Cook 2004; Álvarez-Castañeda et al. 2010; Nolfo-Clements et al. 2017). Although all species share a 2n = 48 karyotype, there is substantial interspecific and intraspecific variation among chromosomes driven by pericentric inversions and heterochromatic additions (Ohno et al 1966; Robbins and Baker 1981; Stangl and Baker 1984; Greenbaum et al. 1994). In the Pacific Northwest, this karyotypic variation has helped to clarify taxonomic relationships between the deer mouse, *Peromyscus maniculatus*, and Keen's mouse, *P. keeni*, where cytogenetic analyses have revealed differences both within and between species in the fundamental number of chromosome arms (Gunn and Greenbaum 1986; Allard and Greenbaum 1988; Gunn 1987; Hogan et al. 1993). Our understanding of such variability is being refined and expanded in the genomics era. Using a combination of long-read sequencing and population genomic data, Harringmeyer and Hoekstra (2022) provide a detailed characterization of 21 large-scale inversions in *P. maniculatus*, including evidence for substantial geographic variability among mainland populations in both the presence and frequencies of inversion polymorphisms. This genomic profile of inversions provides a useful framework for investigating how this important source of variation is partitioned among island and mainland populations.

Deer mice inhabiting the Gulf Islands of British Columbia offer compelling subjects for studying the interaction between demographic history and inversion evolution in island populations. Pleistocene-era glacial activity played an important role in shaping the contemporary

distributions of species occupying higher latitudes (Hewitt 2000; Lessa et al. 2003). Glacial cover of the Gulf Islands during this period (Clague 1980; Clague and James 2002; Hutchinson et al. 2004; Fedje et al. 2009; James et al. 2009) suggests that island colonization could have been accompanied by population size fluctuations as mice dispersed from unglaciated regions or glacial refugia. For low vagility species like *P. maniculatus*, subsequent isolation of the islands from each other and from nearby mainland areas could impose formidable barriers to gene flow (Redfield 1976) that may enable inversions to evolve independently among populations.

Using whole genome sequences from two island populations and one mainland population, we show that segregating inversions have dramatic effects on patterns of variation that span a large portion of the genome. We leverage these genomic signatures to detect polymorphic inversions in short-read population genomic data and to draw inferences about their evolutionary history. By reconstructing the joint demographic history of island and mainland populations, we provide context for our evolutionary interpretations. Together, our genomic characterization of the Gulf Islands populations enriches and expands our current understanding of inversion evolution and population history in *P. maniculatus*.

RESULTS

Population sampling and structure

The Gulf Islands are situated at the southern end of the Strait of Georgia between Vancouver Island (to the west) and mainland British Columbia (to the east). Our sampling of *Peromyscus maniculatus* spans two of the southern Gulf Islands, Saturna (31 km^2) and Pender (34 km^2), and one nearby mainland site at Maple Ridge (Figure 1). Subsequent genomic analyses are based on a subset of unrelated individuals identified among mice sampled from each location (n=20 for Saturna Island, n=10 for Pender Island, and n=17 for mainland Maple Ridge; see Materials and Methods).

To characterize genetic relationships within and among island and mainland sites, we conducted principal component analysis (PCA) and global ancestry inference using genome-wide, unlinked single nucleotide polymorphisms (SNPs). PCA revealed three distinct clusters along the first and second principal component axes that separated mice according to sampling location (Figure

123 2A). Within this combined sample, cross-validation analyses performed with ADMIXTURE
124 (Alexander et al. 2009) recovered three ancestral components (k=3) that almost entirely
125 delineate individuals by sampling location (Figure 2F). In contrast, cross-validation analyses
126 performed separately for each site supported a single ancestral component (k=1) (Figure 2B-D).
127 Together, these findings provide evidence for structure between, but not within, each of the
128 sampled locations.

129 *Inversions drive large-scale distortions in allele frequencies and linkage disequilibrium*

130 Leveraging existing knowledge about chromosomal rearrangements in *P. maniculatus*, we
131 investigated how segregating inversions affect patterns of variation. Due to the suppression of
132 recombination in heterokaryotypes, variation accumulated after the emergence of the inversion
133 may remain private to either class, driving divergence between karyotypes in a manner
134 analogous to local population subdivision. In Gulf Islands *P. maniculatus*, we found that this
135 effect is sufficient to distort the site frequency spectrum (SFS) at the genome-wide level. Within
136 each population, the folded SFS constructed from genome-wide variant calls exhibits subtle
137 elevations in SNP counts at intermediate minor allele frequencies (Figure 3A-C; Supplementary
138 Figure S1). Partitioning the SFS by chromosome, we found that these genome-wide irregularities
139 are driven by extreme distortions in the chromosome-wide SFS for five autosomes: 7, 14, 15, 21,
140 and 22 (Figure 3A-C; Supplementary Figure S1). In each case, the chromosome-wide SFS exhibits
141 an unusually high density of SNPs at a specific minor allele frequency (MAF) bin, consistent with
142 an excess of fixed differences between inverted and standard karyotypes.

143 These same chromosomes also exhibit unusual patterns of linkage disequilibrium (LD). We
144 found that pairwise r^2 across these five chromosomes is elevated over intervals as large as 40
145 Mbp (Figure 3D; Supplementary Figure S1). Such distortions in LD decay could result from the
146 mis-orientation of sampled sequences during their alignment to the reference assembly: when
147 short reads derived from within an inversion are mapped to an inversion-free reference, the re-
148 assignment of genomic coordinates causes nearby SNPs to appear farther apart (and vice versa).
149 In addition to this artificial elevation in LD, the observed patterns of LD decay may also reflect
150 true reductions in recombination across regions harboring polymorphic inversions.

151 SNPs that contribute to the aberrant MAF bins in each chromosome-wide SFS localize to
152 discrete genomic intervals that span the breakpoints identified by Harringmeyer and Hoekstra
153 (2022) for inversions inv7.2, inv14.0, inv15.0, inv21.0, and inv22.0 (Figure 4A-E; Supplementary
154 Figures S2-S3). These MAF-partitioned SNPs exhibit extreme genotype continuity across the
155 entirety of each interval, such that each mouse almost exclusively carries heterozygous or
156 homozygous genotypes at this subset of variants (Figure 4A-E; Supplementary Figures S2-S3).
157 Together, these findings illustrate how the unique evolutionary dynamics induced by non-
158 collinearity can cause patterns of variation spanning polymorphic inversions to depart
159 substantially from genome-wide expectations.

160 We leveraged this diagnostic pattern to screen our short-read sequence data for additional
161 segregating inversions by partitioning autosomal genotypes in each population by all possible
162 minor allele frequencies. This approach revealed that a sixth locus on chromosome 3, inv3.0
163 (Harringmeyer and Hoekstra 2022), is also polymorphic (Figure 4F; Supplementary Figures S2-
164 S3). The shorter span of this inversion is likely insufficient to drive the chromosome-wide
165 distortions in the SFS and LD decay that were observed for the loci on chromosomes 7, 14, 15,
166 21, and 22. In total, our inversion screen indicates that the same six inversion polymorphisms
167 are segregating in all three Gulf Islands populations.

168 Local PCA conducted using SNPs spanning the breakpoints of inv3.0, inv7.2, inv14.0, inv15.0,
169 inv21.0, and inv22.0 yielded three distinct clusters along the first principal component axis
170 (Figure 5). Whereas genome-wide SNPs group individuals by sampling location (Figure 5A), the
171 clustering pattern observed at these loci instead reflects the three possible inversion genotypes
172 (homozygous inversion, heterozygous, or homozygous standard; Figure 5B-G). We find little
173 evidence that additional polymorphisms beyond the six described above are segregating in the
174 Gulf Islands populations. With the possible exception of inv9.0 in the mainland Maple Ridge
175 population, PCA conducted using SNPs within the 21 intervals defined by Harringmeyer and
176 Hoekstra (2022) failed to produce clustering patterns consistent with distinct inversion
177 genotypes (Supplementary Figure S4).

178 *Polarizing inversion polymorphisms*

179 In order to draw inferences about the evolutionary dynamics of segregating inversions, we
180 reconstructed the ancestral orientation at each locus by conducting pairwise whole genome
181 alignments between the *P. maniculatus* reference assembly and genomes from four *Peromyscus*
182 species: *P. polionotus*, *P. leucopus*, *P. eremicus*, and *P. californicus* (see Materials and Methods).
183 Comparisons between *P. maniculatus* and these additional species revealed an abundance of
184 large-scale inversions, duplications, and translocations (Supplementary Figures S5-S8). In some
185 cases, these inter-species rearrangements span over 50 Mbp (Supplementary Figures S5-S8). By
186 assessing collinearity across inversion intervals, we determined that the *P. maniculatus*
187 reference genome carries the derived (i.e., inverted) orientation at inv14.0 and the ancestral
188 (i.e., standard) orientation at inv3.0, inv7.2, inv15.0, inv21.0, and inv22.0 (Supplementary
189 Figures S5-S8). For each locus, we assigned individuals to one of three genotypic classes with
190 respect to the *P. maniculatus* reference genome (homozygous ALT, heterozygous, and
191 homozygous REF) based on their proportion of ALT alleles at (Supplementary Figure S9) or
192 approaching (Supplementary Figures S10-S12) the MAF of the inversion (see Materials and
193 Methods). Based on these polarizations, Table 1 summarizes the frequencies of the inversions at
194 each polymorphic locus in the Gulf Islands populations.

195 Surprisingly, we also uncovered evidence that certain segregating inversions are shared across
196 *Peromyscus* species. Collinearity between the *P. maniculatus* and *P. polionotus* genomes along
197 the majority of the inv14.0 interval suggests that this rearrangement is shared between these
198 sister species (Supplementary Figure S5). In addition, although the *P. maniculatus* assembly
199 carries the standard orientation at inv22.0, evidence for this inversion within *P. leucopus*
200 suggests that this rearrangement is shared across greater phylogenetic distances
201 (Supplementary Figure S6).

202 *High divergence between karyotypes and variable levels of diversity within inversions*

203 To gain insight into how the evolutionary histories of segregating inversions depart from
204 colinear regions of the genome, we compared levels of nucleotide diversity measured at
205 putatively neutral, non-inverted loci (π_{genome}), to three summaries of diversity spanning
206 inversion loci: divergence between standard and inverted karyotypes ($d_{XY}^{karyotype}$), diversity
207 within inverted karyotypes (π_{inv}), and diversity within the standard karyotypes (π_{std}). Table 1

summarizes the means of these distributions calculated in 5 kbp windows across each segregating locus in each population.

Each locus exhibits high levels of divergence between karyotypes ($d_{XY}^{karyotype}$). At all loci, mean $d_{XY}^{karyotype}$ measured in each population far exceeds corresponding levels of diversity measured within either karyotypic class (mean π_{inv} and mean π_{std}) (Table 1). Divergence between karyotypes is also high relative to background levels of diversity. Across all populations and inversion loci, mean $d_{XY}^{karyotype}$ exceeds mean nucleotide diversity measured genome-wide at putatively neutral, non-inverted loci (π_{genome}) (Figure 6). This departure is greatest for inv15.0, which yields a mean $d_{XY}^{karyotype}$ that falls within the right tail of the π_{genome} distribution, and lowest for inv21.0 (Figure 6).

Relative levels of nucleotide diversity within inverted (π_{inv}) and standard (π_{std}) karyotypes exhibit interesting differences across loci. For inv3.0, inv7.2, inv15.0, and inv22.0, inverted karyotypes exhibit reduced nucleotide diversity compared to their corresponding standard karyotypes (i.e., mean π_{inv} < mean π_{std}) (Table 1). These reductions are most extreme for inversions at loci inv7.2 and inv15.0, where mean π_{inv} is less than 14% of mean π_{std}. Both of these inversions feature broad-scale reductions in nucleotide diversity that span nearly the entirety of each locus (Figure 7). In contrast, we observed the opposite trend for inv14.0 and inv21.0, where the inverted karyotype exhibits higher nucleotide diversity (i.e., mean π_{inv} > mean π_{std}) (Table 1). Despite this heterogeneity among inversion loci, all three Gulf Islands populations show remarkable similarity in relative levels of $d_{XY}^{karyotype}$, π_{std}, and π_{inv} at any given locus.

Demographic history

To provide context for the similar patterns of variation we observed at shared polymorphic inversions, we reconstructed the joint demographic history of the Saturna Island, Pender Island, and mainland Maple Ridge populations. Our aim was to estimate population split times, the timing and magnitude of population size changes, and migration rates between populations. Using *Moments* (Jouganous et al. 2017), we fit a variety of simplified demographic models to joint SFS constructed from putatively neutral SNPs falling outside the boundaries of known

236 inversion polymorphisms. Figure 8 depicts the best-fitting three-population demographic
237 model. Maximum likelihood parameter estimates from this model suggest that ancestral island
238 and mainland populations diverged 2,565,592 generations ago and were connected by a low
239 rate of migration (m = 6.75e-6) until 5,090 generations ago (Figure 8). At this time, the
240 contemporary Saturna Island and Pender Island populations split from their shared ancestor,
241 and the mainland Maple Ridge population underwent an over eight-fold size reduction (Figure
242 8). Although this best-fit model provides evidence for migration between the contemporary
243 Saturna Island and Pender Island populations (m = 1.63e-7), it does not support migration
244 between the islands and mainland during this recent epoch (Figure 8). Effective population size
245 (N_e) estimates from this model suggest a large N_e for both the ancestral island (N_e = 255,958)
246 and ancestral mainland (N_e = 579,644) populations and smaller N_e for contemporary
247 populations (N_e = 18,607 for Saturna Island, 14,984 for Pender Island, and 69,313 for mainland
248 Maple Ridge) (Figure 8).

249 Coalescent simulations conducted under this candidate model yielded a close fit to the means
250 of empirical distributions of π, Tajima's *D*, and F_{ST} (Supplementary Figure S13). Variance among
251 loci in these summary statistic estimates is greater for the empirical data than for the simulated
252 data, which could reflect heterogeneity in the recombination rate and mutation rate not
253 captured by simulations. In contrast to the close agreement between empirical data and
254 simulated data for diversity-related summaries, our best-fit demographic model underestimates
255 levels of LD. In each population, average pairwise r^2 decays slower than predicted by simulated
256 data (Supplementary Figure S13). One explanation for this disagreement is that recent
257 population size histories are more dynamic than suggested by our SFS-derived demographic
258 model. To investigate this possibility, we used levels of LD to estimate contemporary N_e,
259 following the approach implemented in *currentNe* (Santiago et al. 2024). Our results indicated
260 that contemporary N_e in each population is more than an order of magnitude smaller than
261 suggested by our SFS-based model (N_e = 679 for Saturna Island, 219 for Pender Island, and 278
262 for mainland Maple Ridge) (Supplementary Table S1).

263 *Patterns of divergence at inversion loci mirror demographic history*

264 Leveraging the shared history of the three Gulf Islands populations, we used patterns of
265 divergence between populations to shed further light on the evolutionary relationships among
266 inverted and standard karyotypes. We extended our comparisons of divergence between
267 karyotypes sampled within the same population ($d_{XY}^{karyotype}$) to include three additional
268 contrasts: inverted karyotypes sampled from different populations (d_{XY}^{inv}), standard karyotypes
269 sampled from different populations (d_{XY}^{std}), and opposite karyotypes sampled from different
270 populations ($d_{XY}^{karyotype\ (between)}$). We used these pairwise summaries of divergence as the basis
271 for distance-based comparisons of inverted and standard karyotypes from all three populations.
272 Figure 9 depicts the resulting trees constructed for each locus using the neighbor-joining
273 algorithm of Saitou and Nei (1987).

274 Surprisingly, despite the old split times we inferred between island and mainland lineages,
275 karyotypes at all polymorphic loci cluster predominantly by orientation rather than by
276 population (Figure 9). This suggests that inversions share a more recent common ancestor with
277 inverted karyotypes in other populations than with their corresponding standard karyotype.
278 Looking within either karyotypic class, branching patterns between populations mirror our best-
279 fit demographic model (i.e., indicating a more recent split between the Saturna Island and
280 Pender Island populations) (Figure 8 and Figure 9). Interestingly, there are major differences
281 among loci in the relative branch lengths subtending inverted versus standard karyotypes.
282 Whereas the inversions at inv3.0, inv15.0, inv22.0, and inv7.2 exhibit lower divergence among
283 populations than their corresponding standard karyotype (Figure 9A-D), we find that the
284 opposite pattern is true for inv21.0 and inv14.0, where inversions exhibit longer branch lengths
285 than standard karyotypes (Figure 9E-F).

286 *Factors shaping divergence between karyotypes*

287 Observed levels of divergence between standard and inverted karyotypes reflect the interplay
288 of two opposing forces. Whereas the gradual accumulation of mutations increases divergence
289 between them, rare "genetic flux" due to gene conversion and double crossover events within
290 heterokaryotypic individuals erodes these differences (Navarro et al. 1997; Guerrero et al.
291 2012). In the absence of genomic heterogeneity in recombination, rates of genetic flux can
292 depend on the length of the inversion and its historical frequency trajectory, both of which

293 determine the propensity for exchange between karyotypes to occur. Longer inversions are
294 more likely to experience the double crossover and gene conversion events that facilitate
295 exchange between karyotypes. Similarly, inversions that have achieved higher frequencies will
296 have spent a greater amount of time in the heterokaryotypic state, increasing the opportunity
297 for genetic flux. Despite these expectations, $d_{XY}^{\text{karyotype}}$ is not strongly associated with inversion
298 length or inversion frequency (Supplementary Figure S14; Supplementary Table S2).

299 In the absence of selection targeting regions located within inverted intervals, rates of genetic
300 flux are expected to be greatest in the center of the inverted locus and lowest near the
301 breakpoints, predicting a negative correlation between $d_{XY}^{\text{karyotype}}$ and distance from breakpoints
302 (Navarro et al. 1997). Treating each inversion locus individually, we find little support for this
303 trend. With the exception of a weak negative correlation at inv21.0 in the Saturna Island sample
304 (Pearson correlation coefficient r = -0.15; p-value = 3.24e-3), all other significant correlations
305 between $d_{XY}^{\text{karyotype}}$ and distance from breakpoints are weakly positive (Supplementary Figure
306 S15; Supplementary Table S3).

307 **DISCUSSION**

308 Our genomic characterization of *P. maniculatus* from the Gulf Islands revealed widespread
309 effects of both polymorphic inversions and demographic history on patterns of variation within
310 and between populations. In the following sections, we discuss how these results enrich our
311 understanding of inversion evolution and the evolution of island populations in this unique
312 system.

313 *A shared history for inversions segregating in island and mainland populations*

314 Our profile of inversions in the Gulf Islands revealed surprising similarities between populations.
315 Despite our estimates that ancestral island and mainland lineages diverged over 2.5 million
316 generations ago, all three Gulf Islands populations segregate the same six inversion
317 polymorphisms. These congruencies stand in stark contrast to the geographic variability among
318 polymorphic inversions observed across mainland populations of *P. maniculatus* (Harringmeyer
319 and Hoekstra 2022).

320 Inversion polymorphisms could be shared among populations because they were inherited from
321 a common ancestor, dispersed by gene flow, or generated by distinct mutations at the same loci.
322 Recent analyses of breakpoint content using long-read sequences have revealed an enrichment
323 of segmental duplications and a lineage-specific centromeric satellite sequence flanking large-
324 scale inversions in *P. maniculatus* (Gozashti et al. 2025). This observation, combined with
325 evidence for breakpoint re-use in *P. maniculatus* (Gozashti et al. 2025), indicates that certain
326 genomic regions are more susceptible to rearrangement than others, as has been observed in
327 diverse species (Cáceres et al. 1999; Pevzner and Tesler 2003; Corbett-Detig et al. 2019;
328 Maggiolini et al. 2020). Although the shared inversion polymorphisms in the Gulf Islands
329 populations could have originated independently at such "fragile" sites, between-population
330 comparisons revealed that inversions sampled from different populations have a lower time to
331 most recent common ancestor (TMRCA) than opposite karyotypes sampled from the same
332 population, indicating that they instead derive from single shared origins. Based on the
333 demographic history we inferred, these inversions either arose in the ultimate island-mainland
334 ancestor or were spread among lineages during the long period of connectivity between
335 ancestral island and ancestral mainland populations. In addition, although we observed distinct
336 trends across inversion loci, levels of within-karyotype diversity and between-karyotype
337 divergence are broadly similar across all three populations within each locus, providing further
338 support for a shared history.

339 *Differences among inversion loci*

340 The distinct patterns of diversity and divergence we observed across inversion loci could reflect
341 differences in the selective regime acting on inversions (discussed in the next section),
342 differences in inversion age, or a combination of both properties. Although levels of divergence
343 between karyotypes can provide useful insights about the age of an inversion (Hasson and
344 Eanes 1996; Corbett-Detig and Hartl 2012), direct estimation of inversion age is complicated by
345 genetic flux between karyotypes and departures from demographic equilibrium (Hasson and
346 Eanes 1996; Corbett-Detig and Hartl 2012; Charlesworth 2023). Surprisingly, we found that
347 inversions at loci with the highest levels of between-karyotype divergence, inv15.0 and inv7.2
348 (Figure 6), also exhibit the lowest levels of divergence across Gulf Islands populations, as

349 illustrated by the shorter branch lengths in Figure 9. In contrast, inv21.0 exhibits the lowest
350 levels of between-karyotype divergence (Figure 6), yet inversions at this locus show high levels
351 of divergence across populations (Figure 9). Though these within-population and between-
352 population patterns appear to paint contrasting portraits of inversion age, they likely reflect the
353 complex interaction between rates of genetic flux, inversion age, and selection. Future efforts
354 aimed at jointly inferring the strength of selection, rates of genetic flux, and inversion age under
355 explicit demographic scenarios will enable clearer insights about inversion dynamics.

356 *Evidence for selection on inversions*

357 Multiple findings suggest that the inversions segregating in Gulf Islands populations have been
358 maintained longer than expected under neutrality. First, all six inversion polymorphisms are
359 shared by the three Gulf Islands populations. Second, whole genome alignments of diverse
360 *Peromyscus* reference assemblies suggest at least two inversions, inv14.0 and inv22.0, may pre-
361 date species splits. Finally, levels of divergence between karyotypes far exceed genome-wide
362 levels of diversity measured at putatively neutral, non-inverted loci, consistent with an elevated
363 TMRCA.

364 Models of long-term balancing selection predict an unusually old TMRCA at the targeted locus
365 and a genealogy characterized by elongated internal branches (Takahata 1990; Hudson 1990;
366 Charlesworth et al. 1997; Bitarello et al. 2023). These features lead to increased neutral
367 variation at the locus, elevated linkage disequilibrium between neutral variants and the selected
368 site, and shifts in the site frequency spectrum toward intermediate-frequency alleles (Bitarello
369 et al 2023; Charlesworth 2023). In most genomic regions, recombination narrows the span of
370 these signals, challenging efforts to detect balancing selection. In regions harboring
371 polymorphic inversions, by contrast, recombination suppression in heterokaryotypes extends
372 the signatures of balancing selection over much greater distances (Charlesworth 2023). At the
373 same time, genetic flux between karyotypes can cause patterns of neutral diversity to vary with
374 distance from breakpoints and selected sites (Hasson and Eanes 1996; Navarro et al. 1997;
375 Navarro et al. 2000; Corbett-Detig and Hartl 2012; Guerrero et al. 2012; Charlesworth 2023;
376 Kapun et al. 2023).

377 Population genetic models that account for some of these complexities have provided insights
378 into the effects of balancing selection on inversion loci. Navarro et al. (2000), Guerrero et al.
379 (2012), and Charlesworth (2023) studied the behavior of coalescence times within and between
380 karyotypes under different models of balancing selection and population structure. Their
381 findings indicated that expected patterns of diversity and divergence strongly depend on
382 whether the system has reached equilibrium and the equilibrium frequency at which the
383 inversion is maintained (Navarro et al. 2000; Guerrero et al. 2012; Charlesworth 2023). After a
384 selectively favored inversion arises, its approach to equilibrium occurs in two phases (Navarro et
385 al. 2000; Guerrero et al. 2012; Charlesworth 2023). First, the inversion quickly rises to its
386 balanced frequency, eliminating variation within the inverted class in a manner analogous to a
387 selective sweep (Maynard Smith and Haigh 1974). Once this balanced frequency is reached,
388 variation within the inverted class recovers through the accumulation of new mutations and
389 genetic flux with the standard karyotype. In a symmetric fashion, variation within the standard
390 class is lost because of its reduced frequency in the population. After equilibrium is established,
391 relative levels of diversity within each karyotypic class should reflect the frequencies at which
392 they are maintained.

393 We observed surprising variability among loci in relative levels of nucleotide diversity within
394 inverted karyotypes. Whereas inversions at most loci (inv3.0, inv7.2, inv15.0, and inv22.0)
395 exhibit reductions in diversity, inversions at two loci (inv14.0 and inv21.0) exhibit higher
396 nucleotide diversity than the standard karyotype. In keeping with theoretical expectations,
397 these contrasting levels of relative nucleotide diversity could reflect the long-term equilibrium
398 frequencies at which each inversion has been maintained. Given the dynamic population size
399 histories we infer, the equilibrium status of these loci among the Gulf Islands populations is
400 uncertain. However, looking across loci, relative levels of π_{inv} do not obviously match the
401 population-specific or continent-wide frequencies of inverted karyotypes (Table 1).

402 Alternatively, reductions in relative π_{inv} could reflect recent selective sweeps of inversions to
403 their balanced frequencies. We found that inv15.0 and inv7.2 exhibit marked reductions in
404 diversity within the inverted karyotype that span nearly the entire locus (Figure 7). Though
405 striking, the extent to which these patterns reflect selective sweeps is unclear. Such broad

406 reductions in diversity would require extremely strong selection, such that there has been little
407 opportunity for recombination within inversion homokaryotypes to narrow the sweep
408 signature. Increases in relative π_{inv} for inv14.0 and inv21.0 also seem inconsistent with neutral
409 evolution, but are difficult to reconcile with existing theory. Current models assume that
410 selection on inversions is sufficiently strong that their equilibrium frequency is quickly achieved,
411 and that the selective regime remains constant through time (Navarro et al. 2000; Guerrero et
412 al. 2012; Charlesworth 2023). In reality, the dynamics could depend on both the mechanism of
413 balancing selection (Fijarczyk and Babik 2015; Bitarello et al. 2023) and its targets (inversion
414 breakpoints or alleles captured by the rearrangement) (Said et al. 2018; Kapun et al. 2023),
415 features that may differ from inversion to inversion. Alternatively, the trends we observed at
416 these two loci could reflect errors in our classification of homozygous genotypes or in our
417 determination of ancestral states. The accuracy of these assignments is influenced by factors
418 including the age of the inversion, the quality of reference assemblies, and the probability of
419 incomplete lineage sorting.

420 More broadly, our findings suggest that there could be important regional differences in the role
421 these inversions play in local adaptation. The same inversions we examined show strong
422 evidence for selection in mainland populations of *P. maniculatus* from the Cascades region,
423 where they exhibit clines across a forest-prairie gradient and associations with phenotypic
424 differences between forest and prairie ecotypes (Hager et al. 2022; Harringmeyer and Hoekstra
425 2022). In contrast to their role in ecotype differentiation among Cascades populations, none of
426 the inversion loci exhibit allele frequency differentiation among Gulf Islands populations, nor do
427 their frequencies in the Gulf Islands mirror continent-wide averages measured by Harringmeyer
428 and Hoekstra (2022) (Table 1). Furthermore, the joint minor allele frequencies of the six
429 inversion polymorphisms across populations (computed by treating the loci as a group) are not
430 unusual with respect to the joint frequencies of neutral, non-inverted SNPs (Supplementary
431 Figure S16).

432 Important differences between the Gulf Islands and Cascades could lead to contrasting
433 evolutionary roles for the same inversions in these two systems. Although island and mainland
434 environments may present divergent selective regimes, the selective gradient responsible for

435 establishing inversion clines among Cascades populations might not exist in the Gulf Islands. In
436 addition, whereas forest and prairie populations in the Cascades region exhibit high rates of
437 migration across divergent environments (Hager et al. 2022), we found little support for
438 migration between island and mainland populations during the contemporary epoch of our
439 demographic model. By linking together adaptive allele combinations, inversions may facilitate
440 local adaptation in the face of high gene flow (Kirkpatrick and Barton 2005; Yeaman and Otto
441 2011; Yeaman and Whitlock 2011). Though this phenomenon may position inversions as an
442 important adaptive mechanism in the Cascades region, our demographic inferences indicate
443 that gene flow is not an impediment to adaptive divergence between island and mainland
444 populations in the Gulf Islands. Hager et al. (2022) suggest that inv15.0 pre-dates formation of
445 the current forest and prairie habitat distribution, raising the possibility that its role in local
446 adaptation has resulted from subsequent accumulation of adaptive alleles (Said et al. 2018;
447 Kapun et al. 2023). Future comparisons of the genetic variation carried by these inversions in
448 the Gulf Islands, Cascades, and other populations of *P. maniculatus* could help clarify their roles
449 in local adaptation.

450 *Prospects for inversion characterization*

451 Our work highlights the valuable insights that can be gained by interpreting patterns of variation
452 spanning inversion polymorphisms within the context of an explicit demographic model. At the
453 same time, we illustrate how the unique evolutionary dynamics experienced by inversions
454 generate characteristic departures in allele frequencies and LD that could have severe
455 consequences for reconstructing demographic history. Indeed, recent demographic analyses
456 using genomic data from the inversion-rich spruce bark beetle demonstrated that segregating
457 inversions affect inferences derived from both SFS and sequentially Markovian coalescent
458 (SMC)-based frameworks depending on the relative contributions of co-linear and non-colinear
459 sequences to the data used for analysis (Zieliński et al. 2025). The six inversions we observed
460 segregating in Gulf Islands *P. maniculatus* span a large portion of the autosomal genome (~6%),
461 have accumulated substantial divergence between karyotypes, and are shared across
462 populations, positioning them in a zone of parameter space with great potential to bias
463 demographic inference. Fortunately, these same properties that are likely to mislead

464 demographic inferences also make it straightforward to detect segregating inversions based on
465 the distortions they produce in the SFS and patterns of LD. Our findings demonstrate that
466 partitioning SNP genotypes by their minor allele frequency provides a simple framework for
467 detecting and genotyping inversions in the absence of long-read sequencing data.

468 Despite recent progress in the characterization of *P. maniculatus* inversions (Harringmeyer and
469 Hoekstra 2022; Gozashti et al. 2025), much remains to be uncovered about inversion evolution
470 in this system. Knowledge of inversion polymorphisms in *P. maniculatus* and the broader
471 karyotypic diversity of the genus pre-dates the genomic era (Ohno et al 1966; Robbins and
472 Baker 1981; Stangl and Baker 1984; Greenbaum et al. 1994), yet important biological properties
473 of these loci are unknown. Uncertainty about local mutation rates and rates of genetic flux
474 between karyotypes preclude more detailed model-based inferences about inversion age and
475 selection. Estimating these key parameters is an important next step for extracting deeper
476 insights about inversion dynamics in *P. maniculatus*.

477 *Insights into the evolution of island populations*

478 The demographic history we reconstructed from genome-wide patterns of variation provides
479 new insights into the population dynamics of *P. maniculatus*. Assuming 1-2 generations per year
480 (Wolff 1985), the estimated split time of 5,090 generations between Pender Island and Saturna
481 Island is broadly consistent with the post-glacial establishment of these populations. During the
482 last glacial period, the southern Strait of Georgia and Gulf Islands region was covered by the
483 Cordilleran ice sheet (Clague 1980; Clague and James 2002; Hutchinson et al. 2004; Fedje et al.
484 2009; James et al. 2009). Following its peak extent 14,500 years ago, the retreat of glacial ice
485 drove rapid uplift of underlying land areas, with relative sea level in the southern Strait of
486 Georgia reaching its lowest point around 11,500 years ago (James et al. 2009). Subsequent rises
487 in global sea level driven by glacial melt established the present-day distribution of land area in
488 the region by 6,000 years ago (James et al. 2009). There is conflicting evidence about the extent
489 of land exposure between southeastern Vancouver Island and mainland British Columbia during
490 this period of sea-level adjustment (Gowan 2005; James et al. 2009), leaving unclear whether
491 mainland *P. maniculatus* could have dispersed to the islands via emergent land bridges. In the
492 absence of such overland routes, it is possible that *P. maniculatus* may have dispersed from

493 mainland areas by rafting on floating ice or debris (Steffan 2016). Notably, our island divergence
494 time estimate also post-dates human inhabitation of coastal British Columbia (Mackie et al.
495 2011; Mackie et al. 2018), raising the possibility that island populations could have instead been
496 founded by inadvertent human introductions.

497 Given its closer proximity and greater connectivity to the Gulf Islands during the post-glacial sea
498 level low stand, Vancouver Island may be a more likely source for the Saturna Island and Pender
499 Island populations than mainland British Columbia. Radiocarbon dating of *Peromyscus* fossils
500 from northern Vancouver Island suggest that *P. maniculatus* could have been present as early as
501 14,000 years ago (Steffan 2016). As this estimate pre-dates deglaciation of the region (Clague
502 and James 2002; Hutchinson et al. 2004; Fedje et al. 2009; James et al. 2009), it indicates that
503 Vancouver Island *Peromyscus* fossils might reflect dispersals from nearby glacial refugia (Steffan
504 2016). Population genetic analysis of *P. maniculatus* from Vancouver Island and other potential
505 source populations could increase the resolution of demographic inferences for mice from
506 Saturna Island and Pender Island.

507 The ancient divergence time between ancestral island and ancestral mainland populations we
508 recovered pre-dates the last inter-glacial period substantially (Clague 1980), leaving its
509 connection to the geological history of the region unclear. Though there is phylogeographic
510 evidence for deep divergence among certain geographic clades of *P. maniculatus* (Sawyer et al.
511 2017), our results are surprising given the proximity of the island and mainland sites we
512 sampled. Demographic analysis of *P. maniculatus* from across the Cascade mountain range
513 produced similar divergence time estimates (2.2 million generations ago) between western
514 populations of the forest ecotype and eastern prairie-ecotype populations (Hager et al. 2022).
515 The similarity in divergence times and the broad proximity of the Gulf Islands and Cascades
516 regions suggest that island-mainland divergence and forest-prairie divergence in these two
517 systems were driven by shared geographic features. Indeed, the Fraser River, which bisects
518 south central British Columbia and northern Washington along its eastern margin, could have
519 shaped population structure among *P. maniculatus* in the Pacific Northwest (Zheng et al. 2003).

520 Our results also speak to demographic processes operating more recently. Analyses of LD in the
521 Gulf Islands populations indicated that values of contemporary N_e are 1-2 orders of magnitude

522 lower than historical estimates based on the SFS. The causes of these recent, severe declines in
523 population size are unknown, but they suggest a reduced capacity for selection to remove
524 deleterious mutations and to spread beneficial mutations in these populations. In addition, our
525 findings serve as a reminder that different aspects of genomic variation respond to evolutionary
526 processes happening on disparate timescales (Beichman et al. 2018; Nadachowska-Brzyska
527 2022). We found a similar discordance between N_e estimates based on allele frequencies and
528 patterns of linkage disequilibrium in island populations of the white-footed mouse, *P. leucopus*
529 (Howell et al. 2025). Commonalities between these systems, such as the fragmentation and
530 urbanization of habitat areas (Eis and Craigdallie 1980; Richburg and Patterson 2005), might
531 suggest that fluctuations in effective population size on contemporary timescales are driven by
532 shared processes. Similar comparisons in natural populations of other species could reveal the
533 extent to which these disparities are a universal reflection of human activity.

534 Finally, our inferred demographic history provides context for interpreting patterns of
535 phenotypic evolution on islands. Mice on Saturna Island exhibit multiple hallmarks of the
536 "island syndrome" (Adler and Levins 1994), including larger body sizes, higher population
537 densities, lower dispersal, and reduced intraspecific aggression (Redfield 1976; Sullivan 1977;
538 Halpin and Sullivan 1978). Saturna Island mice are 35% heavier than mainland mice when raised
539 in a common laboratory environment, indicating a heritable difference in body size (Baier and
540 Hoekstra 2019). The ancient split time between island and mainland lineages substantially
541 extends the temporal window during which the mice that now inhabit Saturna Island and
542 Pender Island could have evolved island syndrome phenotypes. Comparisons of molar sizes
543 between ancient and modern specimens suggest that Pleistocene-era *P. maniculatus* from
544 Vancouver Island were significantly larger than their modern counterparts (Steffan 2016). This
545 inference raises the intriguing possibility that the large body sizes of Saturna Island and Pender
546 Island mice (Redfield 1976; Sullivan 1977; Baier and Hoekstra 2019) are relics of this ancient
547 characteristic.

548 **MATERIALS AND METHODS**

549 *Fieldwork and sample collection*

550 We sampled deer mice (n=120) in the Gulf Island National Park Reserve on Saturna Island and
551 Pender Island (*P. m. saturatus*) and in the Malcolm Knapp Research Forest (MKRF) on the
552 mainland (*P. m. austerus and P. keeni*) in October 2014. Of these, we collected 101 individuals,
553 and either sacrificed and froze them in the field or transported them alive to Harvard University.
554 Specimens were accessioned in the Harvard Museum of Comparative Zoology (MCZ 69713–
555 69813) and liver samples were stored in ethanol. We verified species identity of specimens
556 using a fragment of the cytochrome b gene following Zheng et al. (2003) (see electronic
557 supplementary material, Figure S1, in Baier and Hoekstra (2019) for details). The samples in the
558 present study were previously examined by Baier and Hoekstra (2019). We sequenced 79 of the
559 101 mice for which liver tissue was collected (n=31 for Saturna Island, n=16 for Pender Island,
560 and n=32 for mainland Maple Ridge). Fieldwork was approved by the Harvard University Faculty
561 of Arts and Sciences Institutional Animal Care and Use Committee (protocol #14-08-211).
562 Fieldwork was conducted with permits from the Gulf Island National Park Reserve (GINP-2014-
563 17276), the Pender Island Parks and Recreation Commission, the Malcolm Knapp Research
564 Forest at the University of British Columbia and the Ministry of Forests, Lands and Natural
565 Resource Operations (NASU14-154228, NASU-154230).

566 *Population sampling and sequencing*

567 DNA was extracted from liver samples using Gentra Puregene Tissue Kits (Qiagen, Germantown,
568 MD, USA). Sample purity was assessed by measuring 260/280 absorbance ratios using a
569 NanoDrop Spectrophotometer (Thermo Fisher Scientific, Waltham, MA, USA). DNA
570 concentrations were measured using the Qubit dsDNA HS Assay (Life Technologies, Carlsbad,
571 CA, USA). We used agarose gel electrophoresis to assess DNA sample quality and confirm the
572 presence of high-molecular weight DNA prior to sequencing. Library preparation and
573 sequencing were conducted by the University of Wisconsin-Madison Biotechnology Center. The
574 Illumina DNA Library Prep Kit was used to construct sequencing libraries. Purified and pooled
575 libraries were sequenced on the Illumina NovaSeq X Plus platform to generate paired-end 150
576 bp reads. Samples were sequenced to a targeted coverage of 30X per individual. As a quality
577 control measure, we re-sequenced four mice: one male each from Saturna (FSB85), Pender
578 (FSB114), and Maple Ridge (FSB42) as well as a single female from Maple Ridge (FSB6). These

579 technical replicates were sequenced using unique libraries, allowing us to capture the effects of
580 sequencing/genotyping error introduced at all laboratory and bioinformatic steps downstream
581 of PCR-based library preparation.

582 *Sequence data processing, variant calling, and variant filtering*

583 Prior to alignment and variant calling, raw sequencing reads were pre-processed to remove
584 adapter sequences and low-quality base calls. BBDuk v38.96
585 (https://sourceforge.net/projects/bbmap/) was used to remove Illumina universal adapter
586 sequences and Trimmomatic v0.39 (Bolger et al. 2014) was used to trim low-quality base calls.
587 Processed sequencing reads were then aligned to the *P. maniculatus* NCBI RefSeq assembly
588 (HU_Pman_2.1.3; GCF_003704035.1) using BWA-MEM v0.7.17 (Li and Durbin 2009).

589 We based our variant calling approach on the Genome Analysis Toolkit (GATK) Best Practices
590 recommendations for germline short variant discovery (McKenna et al. 2010; Van der Auwera
591 and O'Connor 2020). Prior to variant identification, we used GATK v4.2.0.0 MarkDuplicates to
592 identify and exclude duplicate reads from downstream variant calling. We then computed
593 genotype likelihoods on a per-individual basis using GATK v4.2.0.0 HaplotypeCaller with the
594 Reference Confidence Model set to “GVCF” mode (Poplin et al. 2017). The resulting GVCF call
595 sets were combined using GATK v4.2.0.0 CombineGVCFs. GATK v4.2.0.0 GenotypeGVCFs was
596 then used to perform joint variant calling separately for the Saturna (n=32), Pender (n=17), and
597 Maple Ridge (n=34) cohorts (Poplin et al. 2017). This approach identified a total of 110,458,672
598 SNPs in the Saturna cohort, 100,102,299 SNPs in the Pender cohort, and 151,071,903 SNPs in
599 the Maple Ridge cohort.

600 We filtered our variant calls with GATK v4.2.0.0 VariantFiltration using site-level annotation
601 thresholds based on GATK Best Practices recommendations for hard filtering (“FisherStrand” >
602 60.0, “StrandOddsRatio” > 3.0, “RMSMappingQuality” < 40.0, “MappingQualityRankSumTest” <
603 -12.5, and “ReadPosRankSumTest” < -8.0; Van der Auwera and O'Connor 2020) with
604 modifications (“QualByDepth” < 5.0). We then used GATK v4.2.0.0 SelectVariants to remove
605 filtered sites and restrict each call set to autosomal, biallelic, single-nucleotide variants. We

606 retained 78,214,290 SNPs in the Saturna cohort, 72,836,785 SNPs in Pender, and 106,361,479
607 SNPs in Maple Ridge following these basic variant filtering and selection procedures.

608 The population-specific joint calling approach we employed can yield incongruent call sets when
609 there is sufficient allele frequency differentiation between populations because variants private
610 to one cohort will not be represented among the variant calls of other cohorts. Variants that are
611 "missing" between call sets could be monomorphic for the reference allele within their cohort,
612 or they could be sites for which high-quality genotype calls could not be obtained. To avoid this
613 ambiguity and ensure that joint allele frequencies were properly calibrated for downstream
614 analyses, we "rescued" genotypes at these missing sites. Briefly, we used the bcftools v1.8
615 (Danecek et al. 2021) "isec" function to extract the locations of private variants observed in
616 each call set. We then used the "-all-sites" mode of GATK v4.2.0.0 GenotypeGVCFs with the "-
617 stand-call-conf" threshold set to 0 to emit monomorphic reference calls at these private variant
618 sites in each of the other cohorts. We used GATK v4.2.0.0 VariantFiltration and SelectVariants to
619 apply the same site-level filtering thresholds to these "rescued" calls (excluding
620 "QualByDepth"). The resulting high-quality invariant calls were then combined with the
621 polymorphic calls of their respective cohort and resulting monomorphic + polymorphic call sets
622 were merged across cohorts, yielding a single multi-population call set that represented the
623 union of variant sites observed across the three populations.

624 *Identification and exclusion of close relatives*

625 Using KING v2.2.8 (Manichaikul et al. 2010), we estimated pairwise kinship coefficients among
626 individuals within each population cohort. Prior to kinship analysis, each of the filtered single-
627 population autosomal call sets described above was pruned for LD by removing SNPs with
628 pairwise $r^2 > 0.1$ in 50 bp sliding windows using the "--indep-pairwise 50 10 0.1" option in plink
629 v1.90 (Chang et al. 2015). This LD-pruning left a total of 2,853,858 SNPs in the Saturna cohort,
630 2,079,267 in Pender, and 5,503,047 in Maple Ridge. We applied the KING-Robust algorithm to
631 each LD-pruned call set and determined relationships from the resulting kinship coefficient
632 estimates using the bounds suggested by the KING documentation (Manichaikul et al. 2010).
633 We identified a large number of closely related individuals within all three cohorts, with 4.44%
634 of Saturna Island pairs, 10.29% of Pender Island pairs, and 6.95% of mainland Maple Ridge pairs

635 exceeding the lower bound for 2nd degree relatives. We did not find close relatives among
636 individuals sampled from different locations. To mitigate the potential biases that close relatives
637 can introduced to population genetic analyses (Wang 2017), we excluded related individuals
638 from downstream analyses. To do so, we used the "relatednessFilter" function of the R package
639 plinkQC v0.3.4 to identify the maximum set of unrelated individuals (defined here as pairs $\geq$ 3rd
640 degree relatives) within each sample. This left us with unrelated sample sizes of n=20 for
641 Saturna, n=10 for Pender, and n=17 for Maple Ridge.

642 *Population structure analysis*

643 We evaluated population structure within and between sampling locations by assessing PCA
644 clustering patterns and conducting model-based inference of global ancestry proportions with
645 ADMIXTURE (Alexander et al. 2009). Since both approaches require unlinked SNPs, we
646 conducted the same LD-pruning described above on the filtered multi-population autosomal
647 call set, restricting our analyses to the unrelated subset of individuals identified in each cohort.
648 PCA was performed on this combined call set using the "--pca" option in plink v1.90 (Chang et
649 al. 2015). Ancestry proportions were inferred either for each population sample separately or
650 for the combined cohort with ADMIXTURE v1.3.0 (Alexander et al. 2009). Five-fold cross-
651 validation was performed for a range of ancestral components (k=1 to k=5) using the "--cv"
652 option. Individual ancestry proportions were then inferred for the k value that produced the
653 lowest cross-validation error.

654 *Inversion identification*

655 We used GATK v4.2.0.0 VariantsToTable to extract information about position, allele count,
656 allele number, allelic states, and individual genotypes at each SNP in the filtered autosomal call
657 sets described above. The resulting genotype tables were used to inspect properties of the
658 genome-wide folded SFS within each population. To investigate the source of "bumps" observed
659 in each population's SFS, we partitioned genotype tables by chromosome in order inspect
660 chromosome-wide allele frequencies. To assess LD decay along each chromosome, we used the
661 "--r2" option in plink v1.90 (Chang et al. 2015) to measure pairwise r^2 among SNPs sampled
662 from each chromosome (set by the "--chr" argument) in the filtered VCF call sets. Singleton

663 variants and sites with missing data were excluded from these LD calculations by setting the "--
664 mac" argument to 1 and the "--geno" argument to 0.

665 In order to "screen" each population sample for evidence of segregating inversions, we looked
666 for intervals that exhibited a high density of either homozygous or heterozygous SNPs in each
667 individual at a specific MAF, consistent with an excess of fixed differences between inverted and
668 standard karyotypes. To perform this screen, we partitioned each population's genotype table
669 by chromosome and by every possible MAF. Coloring SNPs based on their genotype in each
670 individual, we then constructed maps of these MAF-partitioned SNPs and assessed whether
671 regions exhibiting high SNP density and genotype continuity overlapped the boundaries of
672 previously characterized inversion polymorphisms (Harringmeyer and Hoekstra 2022). MAF-
673 partitioned genotype maps were constructed and visualized using ggplot v3.5.1 (Wickham 2016)
674 in R v4.2.1 (R Core Team 2022).

675 To confirm the presence of the inversion polymorphisms identified by our screen and to
676 determine whether additional loci may be segregating in any of the Gulf Islands populations, we
677 conducted PCA within each of the 21 inversion loci characterized by Harringmeyer and Hoekstra
678 (2022). Starting with the filtered multi-population call set, we used the "--extract" function in
679 plink v1.90 (Chang et al. 2015) to extract SNPs falling within each inversion's breakpoints, then
680 performed PCA on each subset of SNPs using the "--pca" function. Because segregating loci
681 carry an abundance of fixed differences that distinguish standard and inverted karyotypes, we
682 did not prune SNPs within inversion intervals for LD prior to analysis. For comparison, we also
683 conducted PCA on genome-wide SNPs falling outside of known inversion intervals. This subset
684 of SNPs was pruned for LD using the same parameters described previously. For each locus, we
685 evaluated whether PCA clustering patterns were consistent with three genotypic classes:
686 homozygous ALT, heterozygous, and homozygous REF. Though inv7.3 exhibits patterns
687 consistent with a segregating inversion (Supplementary Figure S4), we note that this is likely
688 driven by polymorphism at the larger inv7.2 locus in which it is nested, and as such do not treat
689 it as a separate polymorphism.

690 For each polymorphic locus, we assigned individuals to one of three genotypic classes
691 (homozygous ALT, heterozygous, and homozygous REF) by inspecting each mouse's genotype

692 proportions for SNPs segregating at the MAF of the rearrangement. We reasoned that mice with
693 an enrichment of homozygous reference (REF/REF) genotypes carry the same orientation as the
694 reference assembly while those enriched for homozygous alternate (ALT/ALT) genotypes carry
695 the opposite orientation. For inv3.0, inv7.2, inv14.0, inv15.0, and inv22.0, the homozygous
696 classes are clearly distinguished by their proportion of ALT/ALT genotypes (Supplementary
697 Figure S9), and these assignments are visually apparent among MAF-partitioned genotypes
698 (Figure 4 and Supplementary Figures S2-S3). In contrast, neither homozygous class shows a
699 consistent enrichment of ALT/ALT genotypes across all three populations at the inv21.0 locus
700 (Supplementary Figure S9). Instead, one of the homozygous classes (hereby assumed to carry
701 the alternate orientation) shows an enrichment of ALT alleles at MAFs approaching that of the
702 rearrangement (Supplementary Figures S10-S12). This pattern might indicate that although
703 private variants have achieved high frequencies within either karyotypic subpopulation, the
704 inversion is sufficiently young that few of these private differences have fixed. Alternatively, this
705 lack of fixed differences could indicate that the inversion is sufficiently old such that substantial
706 genetic exchange has occurred between karyotypes, impeding the fixation of private mutations.
707 Though we find little evidence for a major role of genetic flux in shaping patterns of divergence
708 between karyotypes (see Results), distinguishing between these explanations remains difficult
709 given our lack of knowledge about rates of exchange.

710 *Inversion polarization*

711 In order to connect the genotype assignments described above (homozygous REF, heterozygous,
712 and homozygous ALT) to rearrangement karyotype in each individual (inv/inv, inv/std, and
713 std/std), we polarized segregating inversion loci to determine whether the *P. maniculatus*
714 assembly carries the inverted (i.e., derived) or standard (i.e., ancestral) orientation at each
715 locus. To assess evolutionary relationships among existing *Peromyscus* reference genomes, we
716 ran Mashtree v 1.4.6 (Katz et al. 2019) on all of the existing *Peromyscus* NCBI RefSeq assemblies:
717 *P. maniculatus* (HU_Pman_2.1.3; GCF_003704035.1), *P. polionotus* (HU_Ppol_1.3.3;
718 GCA_003704135.2), *P. leucopus* (UCI_PerLeu_2.1; GCF_004664715.2), *P. californicus*
719 (ASM782708v3; GCF_007827085.1), *P. eremicus* (PerEre_H2_v1; GCF_949786415.1), *P. nudipes*
720 (Pnud_10x_v1; GCA_902168325.1), *P. melanophrys* (Pmel_10x_v1; GCA_902168415.1), *P.*

721 *attwateri* (Patt_10x_v1; GCA_902168425.1), and *P. aztecus* (Pazt_10x_v1; GCA_902168405.1).
722 The resulting distance-based tree (Supplementary Figure S17) recovers previously characterized
723 phylogenetic relationships by Bradley et al. (2007) and Platt et al. (2015). We focused our
724 polarization analysis on four of these species: *P. polionotus*, *P. leucopus*, *P. californicus*, and *P.*
725 *eremicus*, which collectively span 1.5 to 5 million years of divergence from *P. maniculatus*
726 (Bradley et al. 2007; Platt et al. 2015).

727 We used the nucmer module of the MUMmer v4.0.0 alignment package (Marçais et al. 2018) to
728 conduct pairwise, whole genome alignments between *P. maniculatus* and each of the four focal
729 *Peromyscus* species. Each pair of reference assemblies was aligned using the following
730 parameters: "--maxmatch", "-c 500", "-b 500", and "-l 100". The resulting alignments were
731 filtered using the delta-filter module with parameters "-m", "-i 90", and "-l 100". We then used
732 the show-coords module with parameters "-THrd" to summarize the filtered alignments.

733 Supplementary Figures S18-S21 depict the resulting dotplots from our initial whole genome
734 alignments. Though the closely related *P. polionotus* assembly exhibits one-to-one homology at
735 the chromosome level (Supplementary Figure S18), we identified major inconsistencies
736 between chromosome name and order between *P. maniculatus* and the more diverged *P.*
737 *leucopus* (Supplementary Figure S19) and *P. eremicus* assemblies (Supplementary Figure S20).
738 Both the *P. leucopus* and *P. eremicus* assemblies support the fusion of chromosomes 16 and 21
739 into a single scaffold (chr16_21) and the splitting of chromosome 8 into two separate scaffolds
740 (chr8a and chr8b). According to Long et al. (2019), chromosome assignments for the *P. leucopus*
741 reference were based on the mapping of markers from the *P. maniculatus* linkage map (Kenney-
742 Hunt et al. 2014; Brown et al. 2018) and based on the results of their Hi-C contact analysis,
743 which supported two separate chromosome 8 scaffolds and a fused chromosome 16 and
744 chromosome 21 scaffold. Neither of these fusions or splits are reflected in the *P. maniculatus*
745 assembly. Instead, we found that chr16_21 in both *P. eremicus* and *P. leucopus* is homologous
746 with chr21 in *P. maniculatus* (Supplementary Figures S20 and S19). For the split chromosome 8,
747 we find that chr8a and chr8b in *P. eremicus* align with chr8 and chr16 in *P. maniculatus*,
748 respectively (Supplementary Figure S20). In contrast, chr8a and chr8b in *P. leucopus* align with
749 chr16 and chr8 in *P. maniculatus* (Supplementary Figure S19). We also find that the assignments

750 of chr11 and chr15 in both *P. eremicus* and *P. leucopus* are switched in *P. maniculatus*
751 (Supplementary Figures S20 and S19). We used the "replace" function in seqkit v2.9.0 (Shen et
752 al. 2016) to re-assign chromosome names in both *P. eremicus* and *P. leucopus* reference FASTAs
753 to reflect their homology with *P. maniculatus*. For the scaffold-level *P. californicus* assembly
754 FASTA, we assigned names to the 24 largest scaffolds to reflect their alignment to *P. maniculatus*
755 chromosomes (Supplementary Figure S21).

756 In addition to these chromosome naming inconsistencies, we found that pairwise alignments
757 with *P. leucopus*, *P. eremicus*, and *P. californicus* all exhibit reverse-strand errors in which
758 homologous chromosomes in each pair are represented by the forward strand in one species
759 and the reverse strand in the other (Supplementary Figures S19-S21). These errors manifest as
760 whole-chromosome inversions and can confound the identification of intra-chromosomal
761 rearrangements. To correct these errors, we used the "seq -r -p" option in seqkit v2.9.0 (Shen et
762 al. 2016) to reverse-complement chr1, chr2, chr5, chr8, chr9, chr13, chr16, chr17, chr22, and
763 chrX in the alignment with *P. californicus* and chr1, chr5, chr9, chr11, chr12, chr17, chr18, chr20,
764 and chrX in the alignments with both *P. eremicus* and *P. leucopus*.

765 Following these operations on the reference FASTAs, we re-ran each alignment with the same
766 MUMmer parameters described above. We used SyRI v1.7.1 (Goel et al. 2019) to identify large-
767 scale structural rearrangements between species from these revised alignments and visualized
768 the resulting synteny plots using plotsr v1.1.1 (Goel and Schneeberger 2022). Using established
769 phylogenetic relationships among these species (Bradley et al. 2007; Platt et al. 2015), we then
770 assessed collinearity across inversion intervals to determine the orientation (i.e., standard
771 versus inverted) of each inversion polymorphism the *P. maniculatus* reference genome.

772 *Population genomic analyses of inversions*

773 Taken together, the genotype assignments described above, which reflect the status of each
774 arrangement relative to the *P. maniculatus* reference assembly (i.e., homozygous ALT,
775 heterozygous, or homozygous REF), and our polarization results, which establish the orientation
776 of each locus in the reference, enabled us to determine the inversion genotype carried by each
777 mouse (i.e., inv/inv, inv/std, or std/std). These designations are summarized in Supplementary

Table S4. To avoid switch errors that could be introduced by the phasing of inverted and standard haplotypes in heterozygous individuals, we restricted our analyses of diversity and divergence to the two homozygous classes of individuals (inv/inv and std/std).

Using these homozygous subsets, we computed nucleotide diversity π (Tajima 1983) within each karyotype (π_{inv} and π_{std}) and divergence d_{XY} (Nei and Li 1979) measured either between opposite karyotypes sampled from the same population ($d_{XY}^{karyotype}$), between populations within each karyotypic class (d_{XY}^{inv} and d_{XY}^{std}), or between opposite karyotypes sampled from different populations ($d_{XY}^{karyotype\ (between)}$). We used scikit-allel v1.3.6 (https://github.com/cggh/scikit-allel) to compute each of these summaries in non-overlapping 5 kbp physical windows along each of the six segregating inversion loci: inv3.0, inv7.2, inv14.0, inv15.0, inv21.0, inv22.0. Because the Maple Ridge sample did not include any inversion homozygotes for inv14.0, we excluded within- and between-karyotype summary statistic comparisons involving this population at this locus. To examine spatial patterns of diversity and divergence along each locus, we additionally filtered variant calls to exclude sites exhibiting excess heterozygosity in order to mitigate the effects of repetitive DNA. The GATK commands used to conduct this filtering are described in the next section.

We used the means of the pairwise divergence measures described above to construct distance matrices relating inverted and standard karyotypes sampled from all three Gulf Islands populations at each polymorphic inversion locus (Supplementary Table S5). We then used the algorithm of Saitou and Nei (1987), implemented in the ape v5.7-1 R package (Paradis and Schliep 2019), to construct neighbor-joining trees for each locus.

Obtaining a high-confidence, putatively neutral SNP call set

To obtain a high-confidence SNP call set for demographic analyses, we leveraged the technical replicates of re-sequenced mice to adjust SNP filtering thresholds in a manner that achieved low genotyping error rates while maintaining high SNP retention. In addition to the *site*-level filtering described previously, we focused our more rigorous quality filtering on two *individual*-level annotations: coverage depth (DP), which measures individual read count at a given SNP, and genotype quality score (GQ), which measures the difference between the second lowest

806 and the lowest phred-scaled genotype likelihood. Whereas a low coverage depth can indicate
807 poor support for an individual's genotype call, artificially inflated coverage could reflect copy
808 number variation that is poorly resolved in the reference assembly (Li 2014; Pfeifer 2017). By
809 measuring pairwise genotype concordance between technical replicates, we used the rate of
810 genotype mismatch within each pair of duplicate samples as a proxy for genotyping/sequencing
811 error. To achieve an optimum balance of sensitivity and specificity, we tested a range of
812 maximum DP (41X, 46.5X, 52X, and 57.5X), minimum DP (2.5X, 8X, 13.5X, and 19X), and
813 minimum GQ (0, 25, and 50) thresholds. The impact of these filtering thresholds on SNP
814 retention and error rates is given in Supplementary Table S6. Based on these findings, we
815 imposed the following thresholds using GATK v4.2.0.0 VariantFiltration: "DP" > 8, "DP" < 52, and
816 "GQ" > 25. These individual-level filters left a total of 46,287,217 SNPs in Saturna, 50,060,722
817 SNPs in Pender, and 57,424,563 SNPs in Maple Ridge while reducing the error rates in each
818 sample (measured as one-half the rate of genotype mismatch within each pair of duplicate
819 samples) to 0.06%, 0.05%, and 0.04%, respectively.

820 We also restricted downstream analyses to variants falling outside of annotated repetitive
821 elements, which can cause errors in read mapping and variant calling (Pfeifer 2017). We used
822 the union of the RepeatMasker (hub_2100979_repeatMasker; https://repeatmasker.org),
823 WindowMasker (hub_2100979_windowMasker; Morgulis et al. 2006), TandemRepeatFinder
824 (hub_2100979_simpleRepeat; Benson 1999), and Tandem Duplicates
825 (hub_2100979_tandemDups) annotation tracks of the UCSC Genome Browser
826 (http://genome.ucsc.edu; Karolchik et al 2004) to exclude SNPs within these regions. Repetitive
827 DNA that is poorly resolved in the reference genome can also manifest as regions of high
828 heterozygosity when reads derived from diverged repeats map to the same position. To mitigate
829 this issue, we used GATK v4.2.0.0 VariantFiltration to remove sites with "ExcessHet" > 15.0,
830 corresponding to a p-value less than 0.05 for the Hardy-Weinberg exact test of excess
831 heterozygosity (Wigginton et al. 2005).

832 To reduce the impact of direct and linked selection, we restricted our analysis to putatively
833 neutral SNPs by masking annotated genes in the *P. maniculatus* reference assembly. Using the
834 NCBI RefSeq track (hub_2100979_ncbiRefSeq) of the UCSC Genome Browser

835 (http://genome.ucsc.edu; Karolchik et al 2004), we excluded SNPs falling within 1 kbp of
836 predicted and curated genes. Because inversions may experience distinct recombination and
837 selective landscapes, SNPs falling within these intervals also violate the assumptions of SFS-
838 based demographic inference frameworks. Using the breakpoints defined by Harringmeyer and
839 Hoekstra (2022), we excluded sites within each of the 21 inversion polymorphisms they
840 characterized from downstream analyses. Following these additional masking and filtering
841 procedures, we were left with a total of 2,470,844 SNPs in Saturna, 2,315,952 SNPs in Pender,
842 and 3,455,232 SNPs in Maple Ridge to use for demographic inference.

843 *Demographic inference*

844 We employed the SFS-based inference approach of *Moments* (Jouganous et al. 2017) to
845 reconstruct the joint demographic history of the Gulf Islands populations. Using the high-
846 confidence, putatively neutral variant call set described above, we constructed 2D joint SFS
847 (jSFS) for each pair of populations and a 3D jSFS for the combined set of populations. jSFS were
848 folded to avoid errors in ancestral state identification and all sites containing missing genotype
849 information for one or more individuals were excluded from analysis. To estimate demographic
850 parameters from these jSFS, we used the maximum likelihood framework implemented in
851 *Moments* v1.1.15.

852 We began by fitting simplified two-population demographic histories to the 2D jSFS for each pair
853 of populations (Supplementary Figures S22-S24). Each model featured the divergence of
854 contemporary populations from a shared ancestor, followed by either constant population sizes,
855 discrete changes in size, continuous changes in size, or a combination of discrete and
856 continuous changes in size (Supplementary Figure S22). On top of these basic population size
857 histories, we fit additional parameters that allowed for either symmetric (Supplementary Figure
858 S23) or asymmetric (Supplementary Figure S24) migration between populations following their
859 split.

860 To fit each of the tested two-population models, we used the BFGS optimization routine in
861 *Moments* v1.1.15 to find the parameter values that maximized the likelihood of each observed
862 2D jSFS. To ensure thorough exploration of the parameter space, starting parameter values were

863 permuted across 10,000 independent searches. For each tested model, we retained the top
864 three combinations of parameter values that yielded the highest likelihood. After fitting each
865 model, we inspected parameter convergence, model likelihoods, and residual fit to select
866 between contrasting models. First, we eliminated models where the coefficient of variation for
867 one or more parameters exceeded 0.2 across the top three highest likelihood estimates. Next,
868 we compared model likelihoods either directly (for models with equivalent numbers of
869 parameters), or through adjusted likelihood ratio tests (for models with nested parameters;
870 Coffman et al. 2016). For remaining models, we inspected their fit to both the joint and
871 marginal SFS using Anscombe residuals and selected the simplest model that yielded the best fit
872 to all partitions of the jSFS.

873 Our two-population model fitting and selection procedure indicated that the island-island
874 history was best captured by a model involving recent population divergence (Supplementary
875 Figure S25). In contrast, both island-mainland comparisons yielded best-fit models that involved
876 ancient population divergence followed by recent discrete population size reductions
877 (Supplementary Figures S26 and S27). All three comparisons supported a history of symmetric
878 gene flow between populations (Supplementary Figures S25-S27).

879 To reconcile the extreme differences in divergence times suggested by the island-island and
880 island-mainland comparisons, we used these candidate two-population histories to construct a
881 model that reflects the shared history of all three Gulf Islands populations. This three-
882 population model featured an older epoch in which the ancestral island and mainland lineages
883 diverge from each other, followed by a more recent epoch reflecting the establishment of the
884 contemporary island and mainland populations (Supplementary Figure S28). Given the evidence
885 for gene flow, we fit all possible parameterizations of symmetric migration among ancestral and
886 contemporary populations (Supplementary Figure S28) to the 3D jSFS. We used the same model
887 fitting and model selection approach described above to obtain demographic parameter
888 estimates and select among contrasting models. Residual differences between the joint and
889 marginal SFS observed for each population/population-pair and the expected SFS under the
890 best-fitting three-population demographic model are illustrated in Supplementary Figure S29.

For both two- and three-population candidate models, the ancestral effective population sizes were calculated as $N_{Anc} = \theta/(4*\mu*L)$ using an assumed mutation rate, μ, of 5e-9 per-site, per-generation and an effective sequence length, L, that represents the total length of "callable" sequence. To approximate L, we multiplied the total length of the *P. maniculatus* autosomal genome (2,302,074,966 bp) by the average percentage of SNPs (2.34%) that were retained following both the site- and individual-level filtering as well as the masks we applied to repetitive, genic, and inverted DNA. This yielded an approximate L of 53,870,000 bp. Using the formula above, estimates of population size and time parameters were scaled by N_{Anc} and $2*N_{Anc}$ to convert them into units of individuals and generations, respectively. Parameter uncertainties were estimated using the Godambe Information Matrix methods described in Coffman et al. (2016). We used uncertainty propagation techniques (Ku 1966) to obtain the standard deviations for parameters that are functions of multiple estimated quantities.

We conducted predictive simulations with *msprime* v1.2.0 (Baumdicker et al. 2022) to evaluate the fit of the inferred three-population demographic history. Using the maximum likelihood parameter estimates obtained for the best-fit three-population model, we conducted 2,500 independent simulations of a 1 Mbp genomic element for each component population assuming uniform recombination and mutation rates of 5e-9 per-site, per-generation. Simulated sample sizes were set to match those of the empirical data for each population. We computed nucleotide diversity π (Tajima 1983), Tajima's *D* (Tajima 1989), and F_{ST} (Hudson et al. 1992) in non-overlapping 5 kbp physical windows using scikit-allel v1.3.6 (https://github.com/cggh/scikit-allel) for both the simulated and empirical data. Together, the quality-based filtering we conducted on our empirical SNP call set and the masks we applied to repetitive, genic, and inverted DNA introduced variability in the number of sites "observed" among the 5 kbp physical windows that we used to compute each summary statistic. To account for this variability, we used the "is_accessible" argument in scikit-allel to compute per-site statistics based on the total number of observable sites. To measure the decay of LD in both simulated and empirical data, we used the "--r2" option in plink v1.90 (Chang et al. 2015) to compute pairwise r^2. We restricted this analysis to SNPs separated by at most 1 Mbp by setting the "--ld-window" and "--

919 ld-window-kb" arguments to 1,000,000 and 1,000, respectively. Singleton variants were
920 excluded from these calculations by setting the "--mac" argument to 1.

921 Using the same jSFS employed for demographic inference, we assessed whether the joint minor
922 allele frequencies of the six segregating inversion loci were unusual compared to neutral, non-
923 inverted loci. We summarized joint frequencies by computing the Euclidean distance between
924 each population's minor allele frequencies at the six inversion loci. To obtain the null
925 distributions of this distance measure, we constructed 1,000 random samples of six SNPs from
926 the jSFS and computed the Euclidean distance between minor allele frequencies for each
927 population pair using the dist() function in R v4.2.1 (R Core Team 2022).

928 *Contemporary effective population size estimation*

929 We used the LD-based approach of *currentNe* (Santiago et al. 2024) to estimate contemporary
930 effective population sizes among Gulf Islands populations. Excluding close relatives in a random
931 sample of individuals can artificially inflate estimates of contemporary effective population size
932 obtained through LD-based methods (Waples and Anderson 2017; Santiago et al. 2024). Since
933 the excess of close relatives we found in each sample could be driven by small contemporary
934 effective population sizes or nonrandom trapping, we chose to conduct these analyses on both
935 the full sample (excluding technical replicates) and on the subset of unrelated individuals from
936 each population (obtained by excluding relationships $\leq 2^{nd}$ degree). The same SNP call set
937 employed for demographic inference was used as input to these analyses. We used the "-s"
938 argument to restrict each analysis to 50,000 randomly sampled SNPs and restricted LD
939 calculations to pairs of SNPs located on different chromosomes. When conducting analyses on
940 samples that included close relatives, we used the "-k" option to jointly estimate contemporary
941 effective population size and k, which represents the average number of full siblings each
942 individual has in the sample (Santiago et al. 2024). We use the artificial neutral network
943 approach implemented in *currentNe* to obtain 90% confidence intervals for each point estimate
944 of contemporary effective population size (Santiago et al. 2024).

ACKNOWLEDGEMENTS

We thank members of the Payseur lab for their helpful input on this work. This research was funded by National Institutes of Health (NIH) grants R01GM100426 and R35GM139412 (to B.A.P.). E.K.H. was partially supported by the NIH Graduate Training Grant in Genetics at the University of Wisconsin-Madison (T32GM007133). Computational analyses were conducted using resources provided by the University of Wisconsin-Madison's Center for High Throughput Computing. Fieldwork was supported by a Putnam Expedition Grant of the Harvard Museum of Comparative Zoology. F.B. was supported by an HHMI International Student Research Fellowship, a Joan Brockman Williamson Graduate Research Fellowship, and a Herchel Smith Graduate Fellowship. This work was supported by the Howard Hughes Medical Institute, of which H.E.H. was an Investigator.

DATA ACCESSIBILITY

Upon submission, raw sequencing reads will be deposited in the SRA. Code used to conduct the analyses is available on GitHub (https://github.com/PayseurLabUWMadison/gi_demography_inversions).

AUTHOR CONTRIBUTIONS

E.K.H. and B.A.P. designed the study. F.B. and H.E.H 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 F.B. and H.E.H.

REFERENCES

Adler, G. H., & Levins, R. (1994). The Island Syndrome in Rodent Populations. *The Quarterly Review of Biology*, *69*(4), 473–490.

Adler, G. H., & Tamarin, R. H. (1984). Demography and reproduction in island and mainland white-footed mice (Peromyscus leucopus) in southeastern Massachusetts. *Canadian Journal of Zoology*, *62*(1), 58-64.

Adler, G. H., Wilson, M. L., & Derosa, M. J. (1986). Influence of island area and isolation on population characteristics of Peromyscus leucopus. *Journal of Mammalogy*, *67*(2), 406-409.

Alexander, D. H., Novembre, J., & Lange, K. (2009). Fast model-based estimation of ancestry in unrelated individuals. *Genome research*, *19*(9), 1655-1664.

Allard, M. W., & Greenbaum, I. F. (1988). Morphological variation and taxonomy of chromosomally differentiated Peromyscus from the Pacific Northwest. *Canadian Journal of Zoology*, *66*(12), 2734–2739. https://doi.org/10.1139/z88-400

Álvarez-Castañecla, S. T., Cortés-Calva, P., González-Cózatl, F. X., Rojas, D., & Leyva, I. (2010). Comparison of distribution and habitat characteristics between an endemic and a wide-ranging cryptic species of Peromyscus on the Baja California peninsula. *Western North American Naturalist*, *70*(3), 323-333.

Baier, F., & Hoekstra, H. E. (2019). The genetics of morphological and behavioural island traits in deer mice. *Proceedings of the Royal Society B*, *286*(1914), 20191697.

Baumdicker, F., Bisschop, G., Goldstein, D., Gower, G., Ragsdale, A. P., Tsambos, G., ... & Kelleher, J. (2022). Efficient ancestry and mutation simulation with msprime 1.0. *Genetics*, *220*(3), iyab229.

Bedford, N. L., & Hoekstra, H. E. (2015). Peromyscus mice as a model for studying natural variation. *eLife*, *4*, e06813. https://doi.org/10.7554/eLife.06813

Beichman, A. C., Huerta-Sanchez, E., & Lohmueller, K. E. (2018). Using Genomic Data to Infer Historic Population Dynamics of Nonmodel Organisms. *Annual Review of Ecology,*

Evolution, and Systematics, *49*(1), 433–456. https://doi.org/10.1146/annurev-ecolsys-110617-062431

Benson, G. (1999). Tandem repeats finder: a program to analyze DNA sequences. *Nucleic acids research*, *27*(2), 573-580.

Berdan, E. L., Barton, N. H., Butlin, R., Charlesworth, B., Faria, R., Fragata, I., ... & Flatt, T. (2023). How chromosomal inversions reorient the evolutionary process. *Journal of Evolutionary Biology*, *36*(12), 1761-1782.

Bitarello, B. D., Brandt, D. Y. C., Meyer, D., & Andrés, A. M. (2023). Inferring Balancing Selection From Genome-Scale Data. *Genome Biology and Evolution*, *15*(3), evad032. https://doi.org/10.1093/gbe/evad032

Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*, *30*(15), 2114-2120.

Bradley, R. D., Durish, N. D., Rogers, D. S., Miller, J. R., Engstrom, M. D., & Kilpatrick, C. W. (2007). Toward a Molecular Phylogeny for Peromyscus: Evidence from Mitochondrial Cytochrome-b Sequences. *Journal of Mammalogy*, *88*(5), 1146–1159. https://doi.org/10.1644/06-MAMM-A-342R.1

Brown, J., Crivello, J., & O'Neill, R. J. (2018). An updated genetic map of Peromyscus with chromosomal assignment of linkage groups. *Mammalian Genome*, *29*(5–6), 344–352. https://doi.org/10.1007/s00335-018-9754-7

Burke, J. M., Lai, Z., Salmaso, M., Nakazato, T., Tang, S., Heesacker, A., Knapp, S. J., & Rieseberg, L. H. (2004). Comparative Mapping and Rapid Karyotypic Evolution in the Genus Helianthus. *Genetics*, *167*(1), 449–457. https://doi.org/10.1534/genetics.167.1.449

Bush, G. L., Case, S. M., Wilson, A. C., & Patton, J. L. (1977). Rapid speciation and chromosomal evolution in mammals. *Proceedings of the National Academy of Sciences*, *74*(9), 3942–3946. https://doi.org/10.1073/pnas.74.9.3942

Caceres, M., Ranz, J. M., Barbadilla, A., Long, M., & Ruiz, A. (1999). Generation of a widespread Drosophila inversion by a transposable element. *Science*, *285*(5426), 415-418.

1020 Carbone, L., Alan Harris, R., Gnerre, S., Veeramah, K. R., Lorente-Galdos, B., Huddleston, J.,
1021 Meyer, T. J., Herrero, J., Roos, C., Aken, B., Anaclerio, F., Archidiacono, N., Baker, C.,
1022 Barrell, D., Batzer, M. A., Beal, K., Blancher, A., Bohrson, C. L., Brameier, M., … Gibbs, R.
1023 A. (2014). Gibbon genome and the fast karyotype evolution of small apes. *Nature*,
1024 *513*(7517), 195–201. https://doi.org/10.1038/nature13679

1025 Chang, C. C., Chow, C. C., Tellier, L. C., Vattikuti, S., Purcell, S. M., & Lee, J. J. (2015). Second-
1026 generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience*, *4*(1),
1027 s13742-015.

1028 Charlesworth, B. (2023). The effects of inversion polymorphisms on patterns of neutral genetic
1029 diversity. *Genetics*, *224*(4), iyad116.

1030 Charlesworth, B., Nordborg, M., & Charlesworth, D. (1997). The effects of local selection,
1031 balanced polymorphism and background selection on equilibrium patterns of genetic
1032 diversity in subdivided populations. *Genetical Research*, *70*(2), 155–174.
1033 https://doi.org/10.1017/S0016672397002954

1034 Clague, J. J. (1980). *Late Quaternary geology and geochronology of British Columbia* (No. 13).
1035 Energy, Mines and Resources Canada.

1036 Clague, J. J., & James, T. S. (2002). History and isostatic effects of the last ice sheet in southern
1037 British Columbia. *Quaternary Science Reviews*, *21*(1–3), 71–87.
1038 https://doi.org/10.1016/S0277-3791(01)00070-1

1039 Coffman, A. J., Hsieh, P. H., Gravel, S., & Gutenkunst, R. N. (2016). Computationally Efficient
1040 Composite Likelihood Statistics for Demographic Inference. *Molecular Biology and*
1041 *Evolution*, *33*(2), 591–593. https://doi.org/10.1093/molbev/msv255

1042 Corbett-Detig, R. B., & Hartl, D. L. (2012). Population Genomics of Inversion Polymorphisms in
1043 Drosophila melanogaster. *PLOS Genetics*, *8*(12), e1003056.
1044 https://doi.org/10.1371/journal.pgen.1003056

1045 Corbett-Detig, R. B., Said, I., Calzetta, M., Genetti, M., McBroome, J., Maurer, N. W., ... &
1046 Besansky, N. J. (2019). Fine-mapping complex inversion breakpoints and investigating

1047 somatic pairing in the Anopheles gambiae species complex using proximity-ligation
1048 sequencing. *Genetics*, *213*(4), 1495-1511.

1049 Danecek, P., Bonfield, J. K., Liddle, J., Marshall, J., Ohan, V., Pollard, M. O., ... & Li, H. (2021).
1050 Twelve years of SAMtools and BCFtools. *Gigascience*, *10*(2), giab008.

1051 Dobzhansky, T. (1948). Genetics of natural populations. XVIII. Experiments on chromosomes of
1052 Drosophila pseudoobscura from different geographic regions. *Genetics*, *33*(6), 588.

1053 Dobzhansky, T. (1949). Observations and experiments on natural selection in
1054 Drosophila. *Hereditas*, *35*(S1), 210-224.

1055 Dobzhansky, T. (1950). Genetics of natural populations. XIX. Origin of heterosis through natural
1056 selection in populations of Drosophila pseudoobscura. *Genetics*, *35*(3), 288.

1057 Eis, S., & Craigdallie, D. (1980). Gulf Islands of British Columbia. A landscape analysis.

1058 Fedje, D. W., Sumpter, I. D., & Southon, J. R. (2009). Sea-levels and Archaeology in the Gulf
1059 Islands National Park Reserve. *Canadian Journal of Archaeology / Journal Canadien*
1060 *d'Archéologie*, *33*(2), 234–253.

1061 Fijarczyk, A., & Babik, W. (2015). Detecting balancing selection in genomes: limits and
1062 prospects. *Molecular ecology*, *24*(14), 3529-3545.

1063 Gill, A. E. (1980). Evolutionary genetics of California islands Peromyscus. In *The California*
1064 *Islands: proceedings of a multidisciplinary symposium* (pp. 719-743). Santa Barbara, CA:
1065 Santa Barbara Museum of Natural History.

1066 Goel, M., & Schneeberger, K. (2022). plotsr: Visualizing structural similarities and
1067 rearrangements between multiple genomes. *Bioinformatics*, *38*(10), 2922–2926.
1068 https://doi.org/10.1093/bioinformatics/btac196

1069 Goel, M., Sun, H., Jiao, W.-B., & Schneeberger, K. (2019). SyRI: Finding genomic rearrangements
1070 and local sequence differences from whole-genome assemblies. *Genome Biology*, *20*(1),
1071 277. https://doi.org/10.1186/s13059-019-1911-0

Gowan, E. J. (2007). *Glacio-isostatic adjustment modelling of improved relative sea-level observations in southwestern British Columbia, Canada* (Doctoral dissertation).

Gozashti, L., Harringmeyer, O. S., & Hoekstra, H. E. (2025). How repeats rearrange chromosomes: The molecular basis of chromosomal inversions in deer mice. *Cell Reports*, *44*(5).

Greenbaum, I. F., Gunn, S. J., Smith, S. A., McAllister, B. F., Hale, D. W., Baker, R. J., Engstrom, M. D., Hamilton, M. J., Modi, W. S., Robbins, L. W., Rogers, D. S., Ward, O. G., Dawson, W. D., Elder, F. F. B., Lee, M. R., Pathak, S., & Stangl, Jr., F. B. (1994). Cytogenetic nomenclature of deer mice, *Peromyscus* (Rodentia): Revision and review of the standardized karyotype. *Cytogenetic and Genome Research*, *66*(3), 181–195. https://doi.org/10.1159/000133696

Guerrero, R. F., Rousset, F., & Kirkpatrick, M. (2012). Coalescent patterns for chromosomal inversions in divergent populations. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *367*(1587), 430–438. https://doi.org/10.1098/rstb.2011.0246

Gunn, S. J. (1988). Chromosomal variation and differentiation among insular populations of Peromyscus from the Pacific Northwest. *Canadian Journal of Zoology*, *66*(12), 2726–2733. https://doi.org/10.1139/z88-399

Gunn, S. J., & Greenbaum, I. F. (1986). Systematic Implications of Karyotypic and Morphologic Variation in Mainland Peromyscus from the Pacific Northwest. *Journal of Mammalogy*, *67*(2), 294–304. https://doi.org/10.2307/1380882

Hager, E. R., Harringmeyer, O. S., Wooldridge, T. B., Theingi, S., Gable, J. T., McFadden, S., Neugeboren, B., Turner, K. M., Jensen, J. D., & Hoekstra, H. E. (2022). A chromosomal inversion contributes to divergence in multiple traits between deer mouse ecotypes. *Science*, *377*(6604), 399–405. https://doi.org/10.1126/science.abg0718

Halpin, Z. T., & Sullivan, T. P. (1978). Social Interactions in Island and Mainland Populations of the Deer Mouse, Peromyscus maniculatus. *Journal of Mammalogy*, *59*(2), 395–401. https://doi.org/10.2307/1379924

1098 Harringmeyer, O. S., & Hoekstra, H. E. (2022). Chromosomal inversion polymorphisms shape the
1099 genomic landscape of deer mice. *Nature Ecology & Evolution*, *6*(12), Article 12.
1100 https://doi.org/10.1038/s41559-022-01890-0

1101 Hasson, E., & Eanes, W. F. (1996). Contrasting Histories of Three Gene Regions Associated With
1102 *In(3L)Payne* of *Drosophila melanogaster*. *Genetics*, *144*(4), 1565–1575.
1103 https://doi.org/10.1093/genetics/144.4.1565

1104 Hedrick, P. W. (1981). The Establishment of Chromosomal Variants. *Evolution*, *35*(2), 322–332.
1105 https://doi.org/10.2307/2407841

1106 Hewitt, G. (2000). The genetic legacy of the Quaternary ice ages. *Nature*, *405*(6789), 907–913.
1107 https://doi.org/10.1038/35016000

1108 Hoffmann, A. A., & Rieseberg, L. H. (2008). Revisiting the Impact of Inversions in Evolution: From
1109 Population Genetic Markers to Drivers of Adaptive Shifts and Speciation? *Annual Review*
1110 *of Ecology, Evolution, and Systematics*, *39*(1), 21–42.
1111 https://doi.org/10.1146/annurev.ecolsys.39.110707.173532

1112 Hogan, K. M., Hedin, M. C., Koh, H. S., Davis, S. K., & Greenbaum, I. F. (1993). Systematic and
1113 Taxonomic Implications of Karyotypic, Electrophoretic, and Mitochondrial-DNA Variation
1114 in Peromyscus from the Pacific Northwest. *Journal of Mammalogy*, *74*(4), 819–831.
1115 https://doi.org/10.2307/1382420

1116 Howell, E. K., Nolfo-Clements, L. E., & Payseur, B. A. (2025). Population history across timescales
1117 in an urban archipelago. *Genome Biology and Evolution*, evaf048.

1118 Hudson, R. R. (1990). Gene genealogies and the coalescent process. *Oxford surveys in*
1119 *evolutionary biology*, *7*(1), 44.

1120 Hudson, R. R., Slatkin, M., & Maddison, W. P. (1992). Estimation of levels of gene flow from DNA
1121 sequence data. *Genetics*, *132*(2), 583-589.

1122 Hutchinson, I., James, T. S., Clague, J. J., Barrie, J. V., & Conway, K. W. (2004). Reconstruction of
1123 late Quaternary sea-level change in southwestern British Columbia from sediments in

isolation basins. *Boreas*, *33*(3), 183–194. https://doi.org/10.1111/j.1502-3885.2004.tb01140.x

James, T., Gowan, E. J., Hutchinson, I., Clague, J. J., Barrie, J. V., & Conway, K. W. (2009). Sea-level change and paleogeographic reconstructions, southern Vancouver Island, British Columbia, Canada. *Quaternary Science Reviews*, *28*(13), 1200–1216. https://doi.org/10.1016/j.quascirev.2008.12.022

Jouganous, J., Long, W., Ragsdale, A. P., & Gravel, S. (2017a). Inferring the Joint Demographic History of Multiple Populations: Beyond the Diffusion Approximation. *Genetics*, *206*(3), 1549–1567. https://doi.org/10.1534/genetics.117.200493

Kapun, M., Mitchell, E. D., Kawecki, T. J., Schmidt, P., & Flatt, T. (2023). An ancestral balanced inversion polymorphism confers global adaptation. *Molecular biology and evolution*, *40*(6), msad118.

Karolchik, D., Hinrichs, A. S., Furey, T. S., Roskin, K. M., Sugnet, C. W., Haussler, D., & Kent, W. J. (2004). The UCSC Table Browser data retrieval tool. *Nucleic acids research*, *32*(suppl_1), D493-D496.

Katz, L. S., Griswold, T., Morrison, S. S., Caravas, J. A., Zhang, S., den Bakker, H. C., Deng, X., & Carleton, H. A. (2019). Mashtree: A rapid comparison of whole genome sequence files. *Journal of Open Source Software*, *4*(44), 10.21105/joss.01762. https://doi.org/10.21105/joss.01762

Kenney-Hunt, J., Lewandowski, A., Glenn, T. C., Glenn, J. L., Tsyusko, O. V., O'Neill, R. J., Brown, J., Ramsdell, C. M., Nguyen, Q., Phan, T., Shorter, K. R., Dewey, M. J., Szalai, G., Vrana, P. B., & Felder, M. R. (2014). A genetic map of Peromyscus with chromosomal assignment of linkage groups (a Peromyscus genetic map). *Mammalian Genome*, *25*(3–4), 160–179. https://doi.org/10.1007/s00335-014-9500-8

Kirkpatrick, M., & Barton, N. (2006). Chromosome Inversions, Local Adaptation and Speciation. *Genetics*, *173*(1), 419–434. https://doi.org/10.1534/genetics.105.047985

Ku, H. H. (1966). Notes on the use of propagation of error formulas. *Journal of Research of the National Bureau of Standards*, *70*(4).

Lande, R. (1979). Effective Deme Sizes During Long-Term Evolution Estimated from Rates of Chromosomal Rearrangement. *Evolution*, *33*(1), 234–251. https://doi.org/10.2307/2407380

Lessa, E. P., Cook, J. A., & Patton, J. L. (2003). Genetic footprints of demographic expansion in North America, but not Amazonia, during the Late Quaternary. *Proceedings of the National Academy of Sciences*, *100*(18), 10331–10334. https://doi.org/10.1073/pnas.1730921100

Li, H. (2014). Toward better understanding of artifacts in variant calling from high-coverage samples. *Bioinformatics*, *30*(20), 2843-2851.

Li, H., & Durbin, R. (2009). Fast and accurate short read alignment with Burrows–Wheeler transform. *bioinformatics*, *25*(14), 1754-1760.

Long, A. D., Baldwin-Brown, J., Tao, Y., Cook, V. J., Balderrama-Gutierrez, G., Corbett-Detig, R., ... & Barbour, A. G. (2019). The genome of Peromyscus leucopus, natural host for Lyme disease and other emerging infections. *Science advances*, *5*(7), eaaw6441.

Losos, J. B., & Ricklefs, R. E. (2009). Adaptation and diversification on islands. *Nature*, *457*(7231), 830–836. https://doi.org/10.1038/nature07893

Lucid, M. K., & Cook, J. A. (2004). Phylogeography of Keen's mouse (Peromyscus keeni) in a naturally fragmented landscape. *Journal of Mammalogy*, *85*(6), 1149-1159.

Mackie, Q., Fedje, D., & McLaren, D. (2018). Archaeology and Sea Level Change on the British Columbia Coast. *Canadian Journal of Archaeology / Journal Canadien d'Archéologie*, *42*(1), 74–91.

Mackie, Q., Fedje, D., McLaren, D., Smith, N., & McKechnie, I. (2011). Early Environments and Archaeology of Coastal British Columbia. In N. F. Bicho, J. A. Haws, & L. G. Davis (Eds.), *Trekking the Shore* (pp. 51–103). Springer New York. https://doi.org/10.1007/978-1-4419-8219-3_3

1177 Maggiolini, F. A. M., Sanders, A. D., Shew, C. J., Sulovari, A., Mao, Y., Puig, M., ... & Antonacci, F.
1178 (2020). Single-cell strand sequencing of a macaque genome reveals multiple nested
1179 inversions and breakpoint reuse during primate evolution. *Genome research*, *30*(11),
1180 1680-1693.

1181 Manichaikul, A., Mychaleckyj, J. C., Rich, S. S., Daly, K., Sale, M., & Chen, W. M. (2010). Robust
1182 relationship inference in genome-wide association studies. *Bioinformatics*, *26*(22), 2867-
1183 2873.

1184 Marçais, G., Delcher, A. L., Phillippy, A. M., Coston, R., Salzberg, S. L., & Zimin, A. (2018).
1185 MUMmer4: A fast and versatile genome alignment system. *PLOS Computational Biology*,
1186 *14*(1), e1005944. https://doi.org/10.1371/journal.pcbi.1005944

1187 McClintock, B. (1931). *Cytological observations of deficiencies involving known genes,*
1188 *translocations and an inversion in Zea mays*. University of Missouri, College of
1189 Agriculture, Agricultural Experiment Station.

1190 McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., ... & DePristo, M.
1191 A. (2010). The Genome Analysis Toolkit: a MapReduce framework for analyzing next-
1192 generation DNA sequencing data. *Genome research*, *20*(9), 1297-1303.

1193 Morgulis, A., Gertz, E. M., Schäffer, A. A., & Agarwala, R. (2006). WindowMasker: window-based
1194 masker for sequenced genomes. *Bioinformatics*, *22*(2), 134-141.

1195 Nadachowska-Brzyska, K., Konczal, M., & Babik, W. (2022). Navigating the temporal continuum
1196 of effective population size. *Methods in Ecology and Evolution*, *13*(1), 22–41.
1197 https://doi.org/10.1111/2041-210X.13740

1198 Navarro, A., Barbadilla, A., & Ruiz, A. (2000). Effect of Inversion Polymorphism on the Neutral
1199 Nucleotide Variability of Linked Chromosomal Regions in Drosophila. *Genetics*, *155*(2),
1200 685–698. https://doi.org/10.1093/genetics/155.2.685

1201 Navarro, A., Betrán, E., Barbadilla, A., & Ruiz, A. (1997). Recombination and gene flux caused by
1202 gene conversion and crossing over in inversion heterokaryotypes. *Genetics*, *146*(2), 695–
1203 709. https://doi.org/10.1093/genetics/146.2.695

1204 Nei, M., & Li, W. H. (1979). Mathematical model for studying genetic variation in terms of
1205 restriction endonucleases. *Proceedings of the National Academy of Sciences*, *76*(10),
1206 5269–5273. https://doi.org/10.1073/pnas.76.10.5269

1207 Nolfo-Clements, L., Butcher, R., Leite, M., & Clements, M. (2017). Evidence of the island rule and
1208 microevolution in white-footed mice (Peromyscus leucopus) in an urban harbor
1209 archipelago. *Mammal Research*, *62*, 423-430.

1210 Ohno, S., Weiler, C., Poole, J., Christian, L., & Stenius, C. (1966). Autosomal polymorphism due to
1211 pericentric inversions in the deer mouse (peromyscus maniculatus), and some evidence
1212 of somatic segregation. *Chromosoma*, *18*(2), 177–187.
1213 https://doi.org/10.1007/BF00326866

1214 Osgood, W. H. (1909). *Revision of the mice of the American genus Peromyscus* (No. 28). US
1215 Government Printing Office.

1216 Paradis, E., & Schliep, K. (2019). ape 5.0: an environment for modern phylogenetics and
1217 evolutionary analyses in R. *Bioinformatics*, *35*(3), 526-528.

1218 Pevzner, P., & Tesler, G. (2003). Human and mouse genomic sequences reveal extensive
1219 breakpoint reuse in mammalian evolution. *Proceedings of the National Academy of*
1220 *Sciences*, *100*(13), 7672-7677.

1221 Pfeifer, S. P. (2017). From next-generation resequencing reads to a high-quality variant data
1222 set. *Heredity*, *118*(2), 111-124.

1223 Platt, R. N., Amman, B. R., Keith, M. S., Thompson, C. W., & Bradley, R. D. (2015). What Is
1224 *Peromyscus* ? Evidence from nuclear and mitochondrial DNA sequences suggests the
1225 need for a new classification. *Journal of Mammalogy*, *96*(4), 708–719.
1226 https://doi.org/10.1093/jmammal/gyv067

1227 Poplin, R., Ruano-Rubio, V., DePristo, M. A., Fennell, T. J., Carneiro, M. O., Van der Auwera, G. A.,
1228 ... & Banks, E. (2017). Scaling accurate genetic variant discovery to tens of thousands of
1229 samples. *BioRxiv*, 201178.

Redfield, J. A. (1976). Distribution, abundance, size, and genetic variation of *Peromyscus maniculatus* on the Gulf Islands of British Columbia. *Canadian Journal of Zoology*, *54*(4), 463–474. https://doi.org/10.1139/z76-053

Richburg, J. A., & Patterson III, W. A. (2005). Historical Description of the Vegetation of the Boston Harbor Islands: 1600–2000. *Northeastern Naturalist*, *12*(sp3), 13–30. https://doi.org/10.1656/1092-6194(2005)12[13:HDOTVO]2.0.CO;2

Robbins, L. W., & Baker, R. J. (1981). An assessment of the nature of chromosomal rearrangements in 18 species of *Peromyscus* (Rodentia: Cricetidae). *Cytogenetic and Genome Research*, *31*(4), 194–202. https://doi.org/10.1159/000131649

Said, I., Byrne, A., Serrano, V., Cardeno, C., Vollmers, C., & Corbett-Detig, R. (2018). Linked genetic variation and not genome structure causes widespread differential expression associated with chromosomal inversions. *Proceedings of the National Academy of Sciences*, *115*(21), 5492–5497. https://doi.org/10.1073/pnas.1721275115

Saitou, N., & Nei, M. (1987). The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Molecular biology and evolution*, *4*(4), 406-425.

Santiago, E., Caballero, A., Köpke, C., & Novo, I. (2024). Estimation of the contemporary effective population size from SNP data while accounting for mating structure. *Molecular Ecology Resources*, *24*(1), e13890. https://doi.org/10.1111/1755-0998.13890

Sawyer, Y. E., Flamme, M. J., Jung, T. S., MacDonald, S. O., & Cook, J. A. (2017). Diversification of deermice (Rodentia: Genus Peromyscus) at their north-western range limit: genetic consequences of refugial and island isolation. *Journal of Biogeography*, *44*(7), 1572–1585. https://doi.org/10.1111/jbi.12995

Shen, W., Le, S., Li, Y., & Hu, F. (2016). SeqKit: A Cross-Platform and Ultrafast Toolkit for FASTA/Q File Manipulation. *PLOS ONE*, *11*(10), e0163962. https://doi.org/10.1371/journal.pone.0163962

Smith, J. M., & Haigh, J. (1974). The hitch-hiking effect of a favourable gene. *Genetics Research*, *23*(1), 23-35.

Stangl, F. B., & Baker, R. J. (1984). Evolutionary Relationships in Peromyscus: Congruence in Chromosomal, Genic, and Classical Data Sets. *Journal of Mammalogy*, *65*(4), 643–654. https://doi.org/10.2307/1380846

Steffen, M. L. (2016). Body-size trends in Peromyscus (Rodentia: Cricetidae) on Vancouver Island, Canada, with comments on relictual gigantism. *Paleobiology*, *42*(3), 532–546. https://doi.org/10.1017/pab.2016.14

Sturtevant, A. H. (1921). A case of rearrangement of genes in Drosophila. *Proceedings of the National Academy of Sciences*, *7*(8), 235-237.

Sullivan, T. P. (1977). Demography and Dispersal in Island and Mainland Populations of the Deer Mouse, Peromyscus Maniculatus. *Ecology*, *58*(5), 964–978. https://doi.org/10.2307/1936918

Tajima, F. (1983). Evolutionary relationship of DNA sequences in finite populations. *Genetics*, *105*(2), 437-460.

Tajima, F. (1989). Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics*, *123*(3), 585-595.

Takahata, N. (1990). A simple genealogical structure of strongly balanced allelic lines and trans-species evolution of polymorphism. *Proceedings of the National Academy of Sciences*, *87*(7), 2419-2423.

Van der Auwera, G. A., & O'Connor, B. D. (2020). *Genomics in the cloud: using Docker, GATK, and WDL in Terra*. O'Reilly Media.

Wang, J. (2017). Effects of sampling close relatives on some elementary population genetics analyses. *Molecular Ecology Resources*, *18*(1), 41-54.

Waples, R. S., & Anderson, E. C. (2017). Purging putative siblings from population genetic data sets: a cautionary view.

Wellenreuther, M., & Bernatchez, L. (2018). Eco-Evolutionary Genomics of Chromosomal Inversions. *Trends in Ecology & Evolution*, *33*(6), 427–440. https://doi.org/10.1016/j.tree.2018.04.002

Wickham, H. (2016). *Ggplot2: Elegrant graphics for data analysis*. Springer.

Wigginton, J. E., Cutler, D. J., & Abecasis, G. R. (2005). A Note on Exact Tests of Hardy-Weinberg Equilibrium. *The American Journal of Human Genetics*, *76*(5), 887–893. https://doi.org/10.1086/429864

Wolff, J. O. (1985). Comparative population ecology of Peromyscus leucopus and Peromyscus maniculatus. *Canadian Journal of Zoology*, *63*(7), 1548-1555.

Yeaman, S., & Otto, S. P. (2011). Establishment and Maintenance of Adaptive Genetic Divergence Under Migration, Selection, and Drift. *Evolution*, *65*(7), 2123–2129. https://doi.org/10.1111/j.1558-5646.2011.01277.x

Yeaman, S., & Whitlock, M. C. (2011). The Genetic Architecture of Adaptation Under Migration–Selection Balance. *Evolution*, *65*(7), 1897–1911. https://doi.org/10.1111/j.1558-5646.2011.01269.x

Zheng, X., Arbogast, B. S., & Kenagy, G. J. (2003). Historical demography and genetic structure of sister species: Deermice (Peromyscus) in the North American temperate rain forest. *Molecular Ecology*, *12*(3), 711–724. https://doi.org/10.1046/j.1365-294X.2003.01770.x

Zielinski, P., Morales-Garcia, J., Schebeck, M., Duduman, M. L., & Nadachowska-Brzyska, K. (2025). Demographic history inferred from an inversion-rich spruce bark beetle genome. *bioRxiv*, 2025-07.

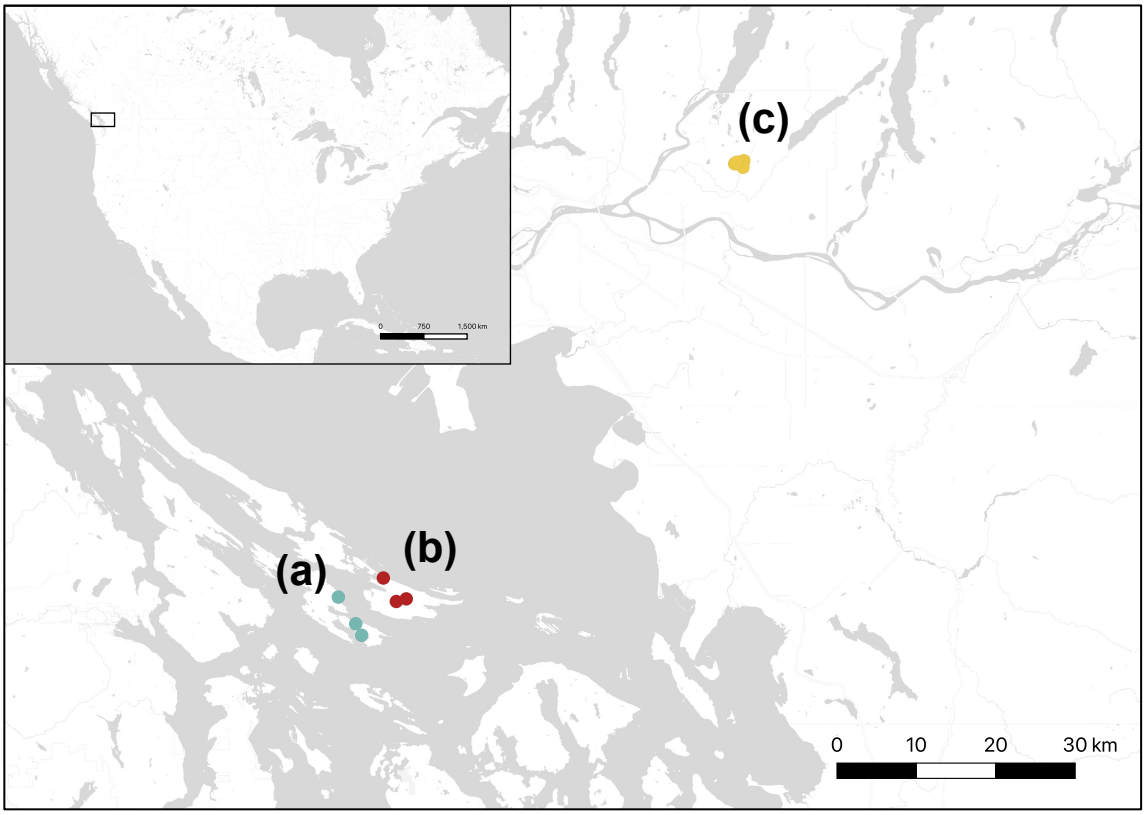

Figure 1. Map of the Gulf Islands in southwestern British Columbia. Land areas are depicted in white and water areas are depicted in gray. Points indicate sampled areas on Pender Island (a; blue), Saturna Island (b; red), and mainland Maple Ridge (c; yellow). Inset depicts location of the Gulf Islands region in the Pacific Northwest.

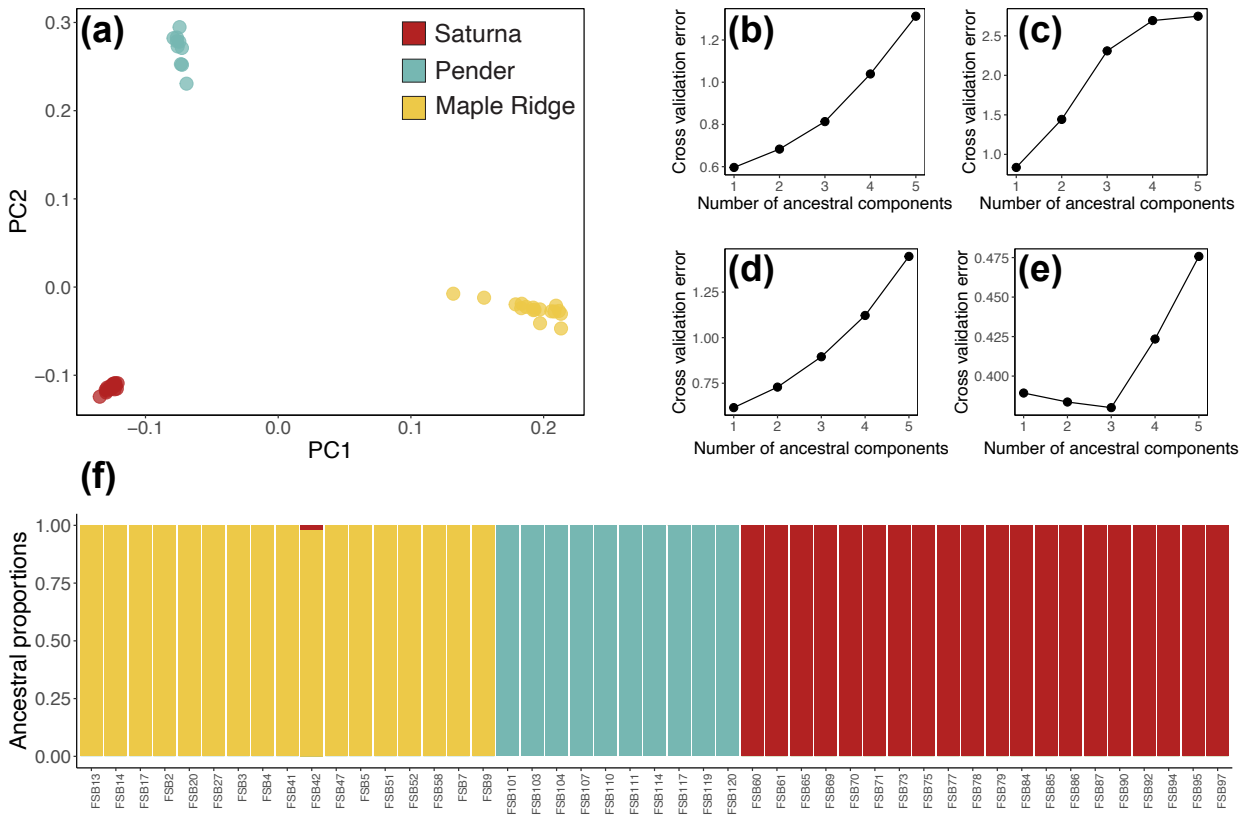

Figure 2. Evidence for structure among island and mainland sampling locations. Panel a depicts the results of PCA conducted using genome-wide unlinked SNPs. Each point represents an individual's position along the first (x-axis) and second (y-axis) principal component. Individuals are colored according to sampling location (yellow for mainland Maple Ridge, blue for Pender Island, and red for Saturna Island). Percentage of variance explained by PC1=7.73% and by PC2=4.65%. Panels b-e present the results of five-fold cross validation analyses performed with ADMIXTURE (Alexander et al. 2009) to identify the best-fit number of ancestral components within the Saturna (b), Pender (c), Maple Ridge (d), and combined (e) samples. Panel f depicts global ancestry proportions estimated for the combined sample, assuming k=3 ancestral components. Each individual is represented by a colored bar where the height of each colored segment within the bar gives the proportion of ancestry an individual derives from each component. Ancestral component colors reflect sampling location (as in Panel a).

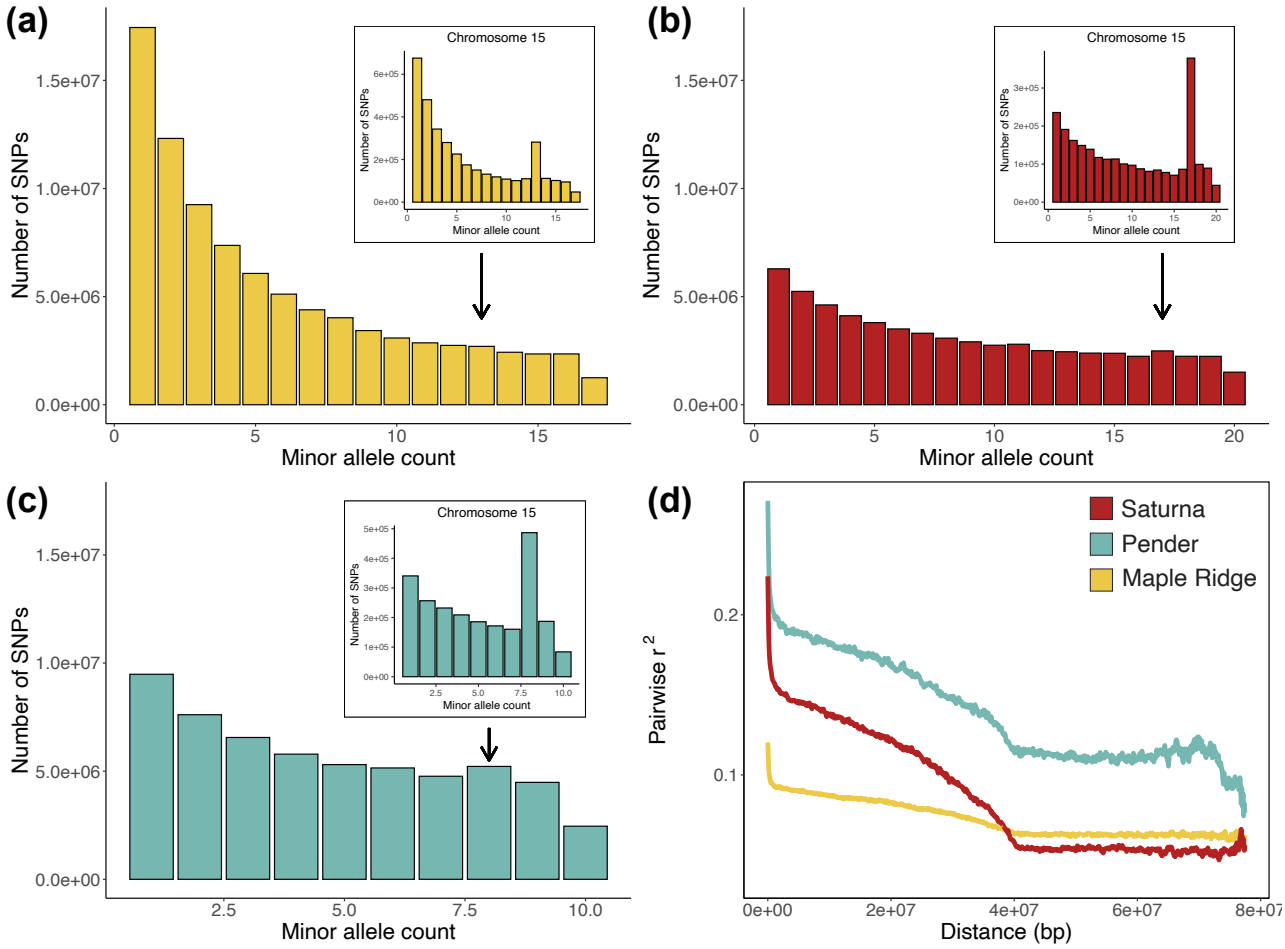

Figure 3. Inversion polymorphisms drive irregularities in the SFS and cause distortions in LD decay. Panels a-c provide an example of how the inversion polymorphism on chromosome 15 (inv15.0) affects both local and genome-wide allele frequencies in the Maple Ridge (a), Saturna Island (b), and Pender Island (c) samples. Main histograms plot the folded SFS constructed from genome-wide SNPs. Arrows indicate minor allele frequency bins in the genome-wide SFS that exhibit elevated SNP counts. Insets illustrate how these minor distortions in the genome-wide SFS correspond to major distortions in the local SFS for chromosome 15 SNPs. Panel d plots average pairwise r^2 (y-axis) between chromosome 15 SNPs separated by increasing distance (x-axis). Each line represents LD decay within a different population sample (yellow for Maple Ridge, red for Saturna, and blue for Pender).

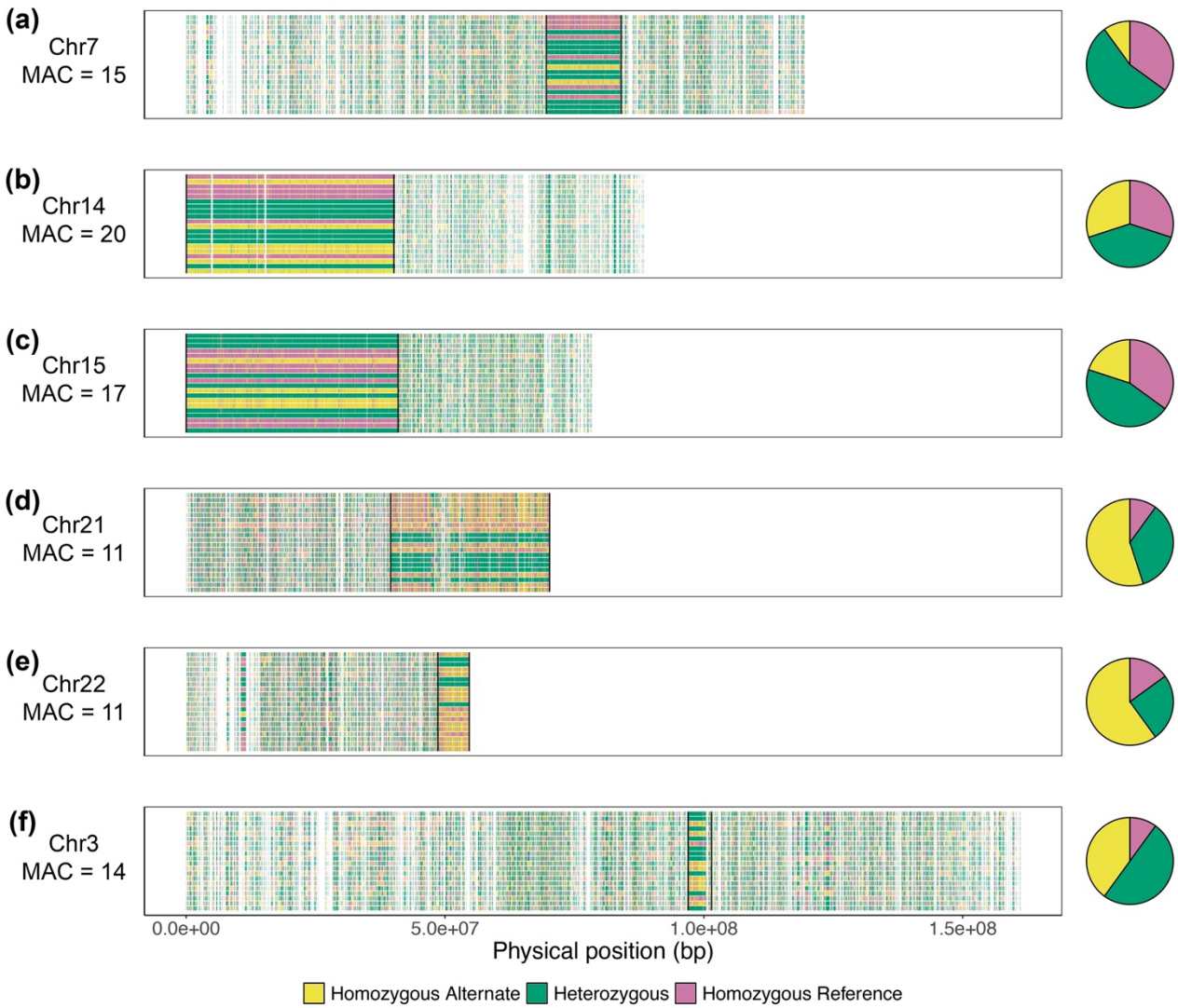

Figure 4. Partitioning SNP genotypes by minor allele frequency enables the detection and genotyping of polymorphic inversions. Plots illustrate the positive results for the inversion screen conducted in the Saturna Island sample. Each horizontal panel represents a different autosome: chr7 (a), chr14 (b), chr15 (c), chr21 (d), chr22 (e), and chr3 (f). Labels beneath each autosome name indicate the corresponding minor allele count of the inversion in the sample. X-axes measure the physical position along each autosome. Within a panel, each row represents a sampled individual, and each vertical tick mark indicates the location of a SNP with the specified minor allele count. SNPs are colored according to their diploid genotype in each individual (yellow for homozygous alternate, green for heterozygous, and pink for homozygous reference). Vertical black lines denote the breakpoints defined by Harringmeyer and Hoekstra (2022) for inv7.2 (a), inv14.0 (b), inv15.0 (c), inv21.0 (d), inv22.0 (e), and inv3.0 (f). Pie charts adjacent to each panel denote the genotype proportions of the arrangement in each sample. See Supplementary Figures S2 and S3 for corresponding plots for the Pender Island and mainland Maple Ridge samples.

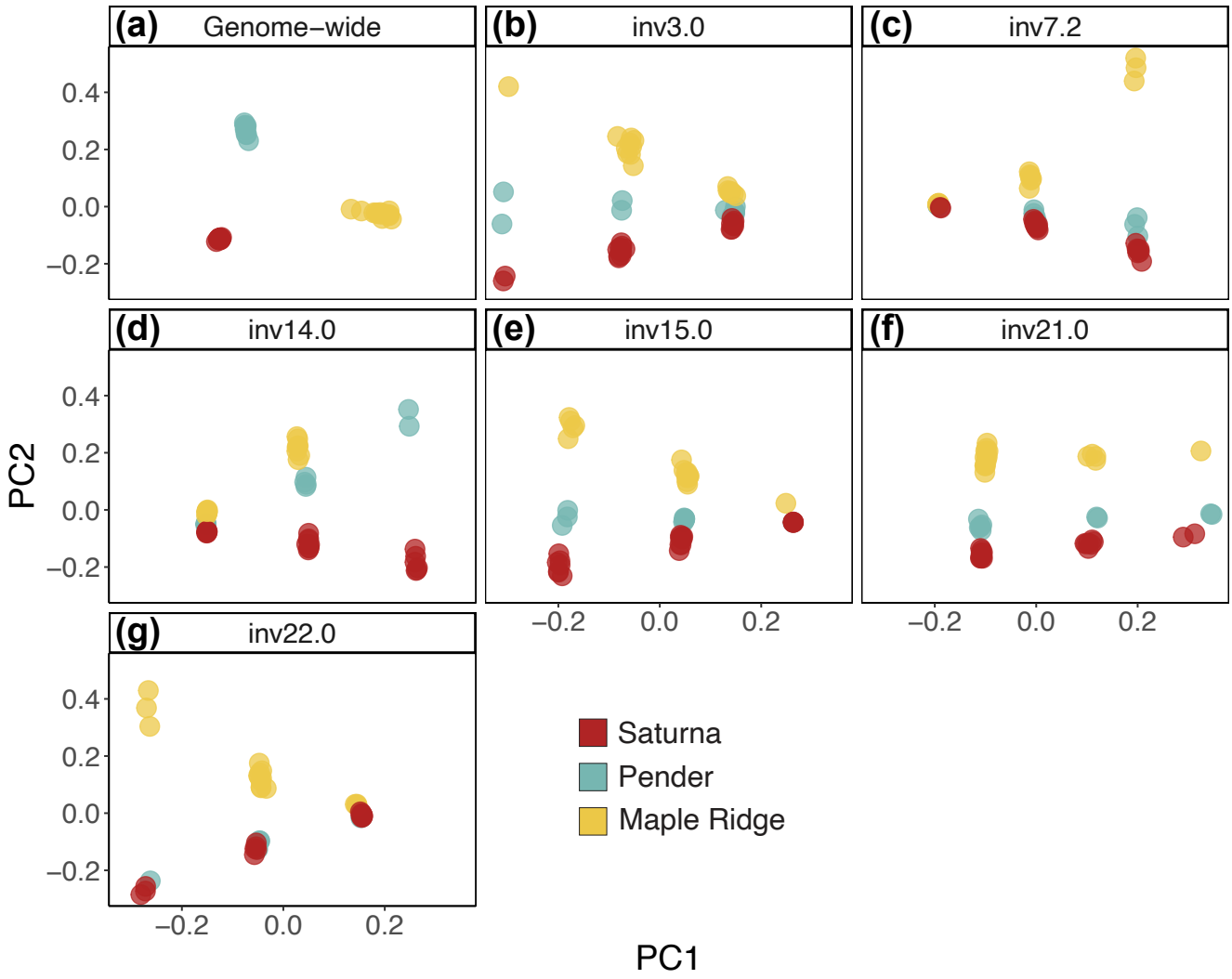

Figure 5. Individuals from distinct populations cluster according to inversion genotype at each polymorphic locus. Panels compare the results of PCA conducted on genome-wide SNPs falling outside of known inversions (a) to PCA performed on SNPs within each polymorphic inversion locus: inv3.0 (b), inv7.2 (c), inv14.0 (d), inv15.0 (e), inv21.0 (f), and inv22.0 (g). Each point denotes an individual's position along the first (x-axis) and second (y-axis) principal component. Points are colored according to population sample (yellow for Maple Ridge, blue for Pender Island, and red for Saturna Island).

Table 1. Polymorphic inversions in the Gulf Islands							
Locus	Population	Inversion frequency	$\bar{\pi}_{\mathrm{std}}$	$\bar{\pi}_{\mathrm{inv}}$	$\bar{d}_{XY}^{karyotype}$	Inversion length	Continent-wide frequency
inv3.0	Saturna	0.65	0.0049	0.0037	0.0150	4.5 Mbp	0.183
	Pender	0.7	0.0078	0.0040	0.0151		
	Maple Ridge	0.676	0.0108	0.0053	0.0154		
inv7.2	Saturna	0.375	0.0084	0.0010	0.0176	14.4 Mbp	0.110
	Pender	0.5	0.0086	0.0006	0.0178		
	Maple Ridge	0.647	0.0102	0.0010	0.0173		
inv14.0	Saturna	0.5	0.0019	0.0061	0.0168	40.1 Mbp	0.862
	Pender	0.4	0.0016	0.0066	0.0169		
	Maple Ridge	0.235	–	–	–		
inv15.0	Saturna	0.425	0.0092	0.0005	0.0192	40.9 Mbp	0.161
	Pender	0.4	0.0085	0.0005	0.0192		
	Maple Ridge	0.382	0.0119	0.0015	0.0186		
inv21.0	Saturna	0.725	0.0041	0.0069	0.0141	30.7 Mbp	0.275
	Pender	0.7	0.0019	0.0065	0.0147		
	Maple Ridge	0.824	0.0028	0.0085	0.0144		
inv22.0	Saturna	0.725	0.0073	0.0027	0.0171	6.1 Mbp	0.571
	Pender	0.6	0.0068	0.0016	0.0171		
	Maple Ridge	0.529	0.0109	0.0025	0.0164		

Table 1. All three Gulf Islands populations are polymorphic at the same inversion loci. Column three summarizes observed frequencies of the inverted karyotype at each polymorphic inversion locus. Individual genotypes at each inversion locus were obtained by partitioning SNP genotypes by minor allele frequency (see Figure 4 and Supplementary Figures S2-S3). The orientation of each arrangement was then determined based on alignments between the *P. maniculatus* reference assembly and those of *P. polionotus*, *P. leucopus*, *P. eremicus*, and *P. californicus*. Columns four through six tabulate mean nucleotide diversity within standard karyotype ($\bar{\pi}_{\mathrm{std}}$), mean nucleotide diversity within inverted karyotype ($\bar{\pi}_{\mathrm{inv}}$) and mean divergence between karyotypes ($\bar{d}_{\mathrm{XY}}^{\mathrm{karyotype}}$) for each locus. Means are computed across 5 kbp windows. Columns eight and nine summarize the physical length and average continent-wide frequency of each inversion, as defined by Harringmeyer and Hoekstra (2022).

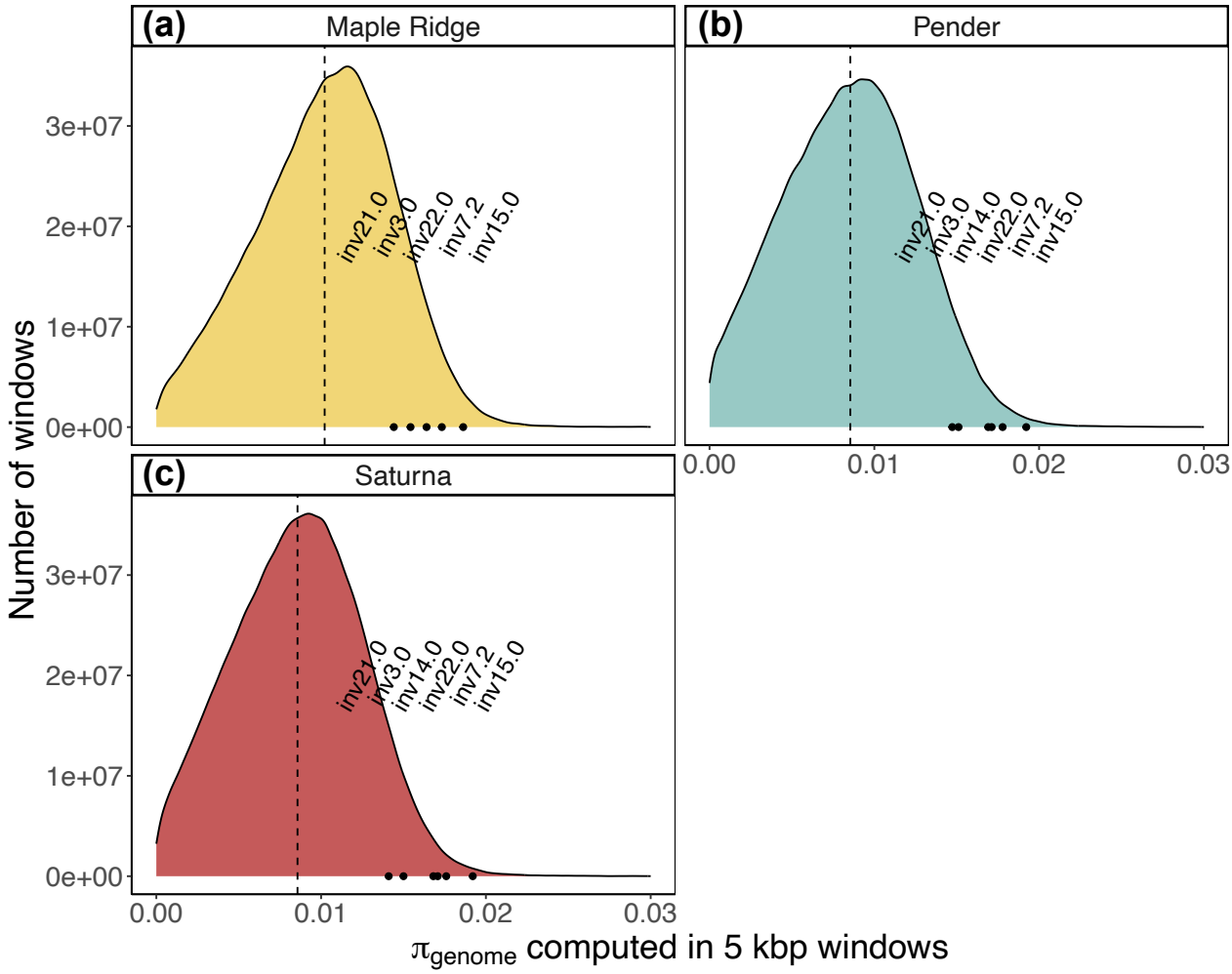

Figure 6. Divergence between karyotypes at polymorphic inversions exceeds mean genome-wide diversity. Density plots illustrate the distribution of nucleotide diversity computed in genome-wide non-overlapping 5 kbp windows outside of known inversion loci (π_{genome}) for the Maple Ridge (a), Pender Island (b), and Saturna Island (c) samples. Dashed black lines denote the mean π_{genome} observed in each population sample. Red points demonstrate where mean divergence between karyotypes ($d_{XY}^{karyotype}$) falls along the distribution of genome-wide nucleotide diversity for each polymorphic inversion.

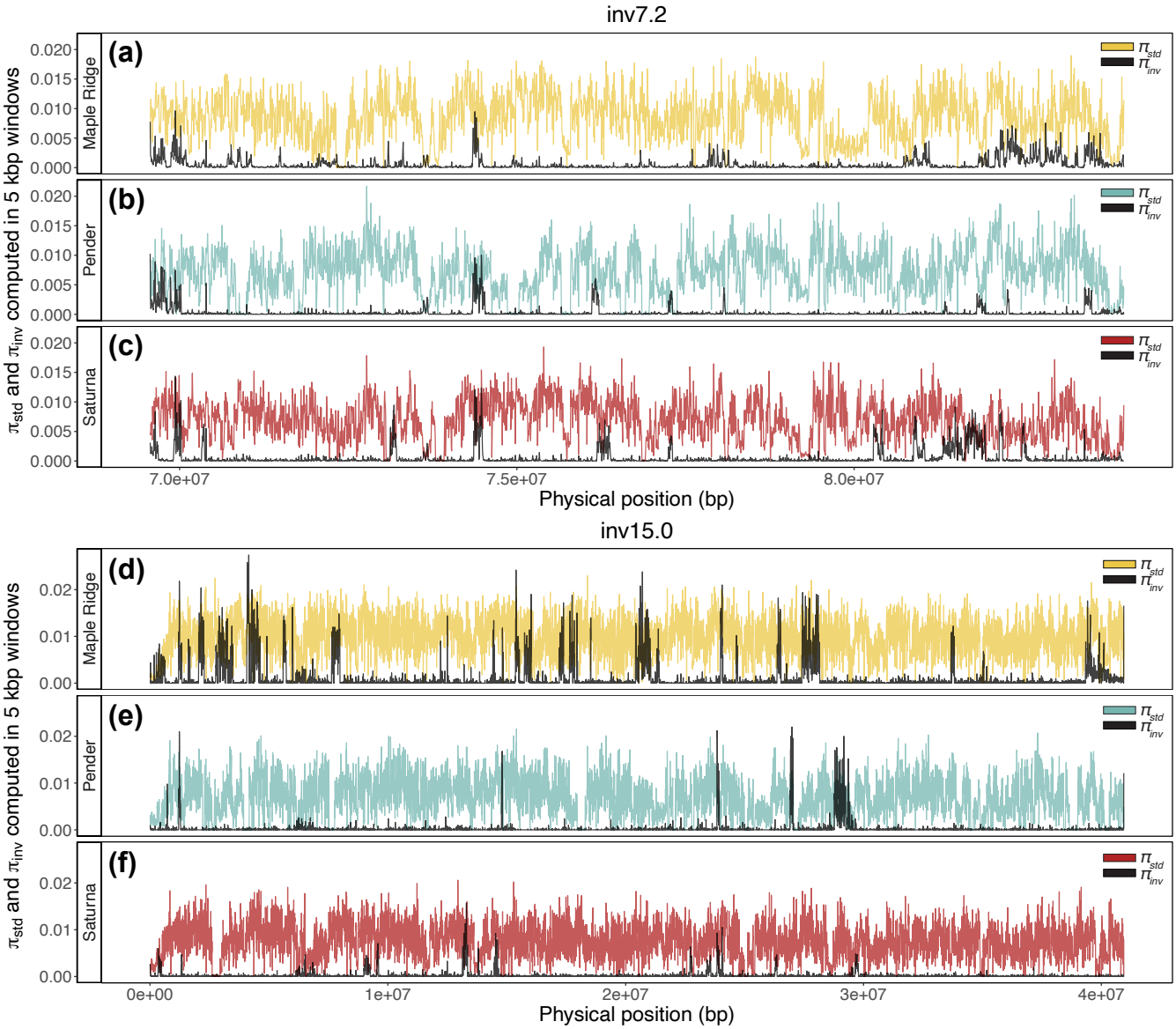

Figure 7. Polymorphic inversions on chromosomes 7 and 15 exhibit extreme reductions in nucleotide diversity. Horizontal panels compare nucleotide diversity among inverted karyotypes (π_{inv}) and nucleotide diversity among standard karyotypes (π_{std}) across the inv7.2 (a-c) and inv15.0 (d-f) loci. Each panel represents a different population: Maple Ridge (a and d), Pender Island (b and e), and Saturna Island (c and f). X-axes represent genomic position along the chromosome, bounded by inversion breakpoints. Y-axes plot each summary of diversity computed in non-overlapping 5 kbp windows. Yellow, blue, and red lines represent nucleotide diversity measured among standard karyotypes (π_{std}) sampled from Maple Ridge, Pender, and Saturna, respectively. Black lines represent nucleotide diversity within the corresponding sample of inverted karyotypes (π_{inv}).

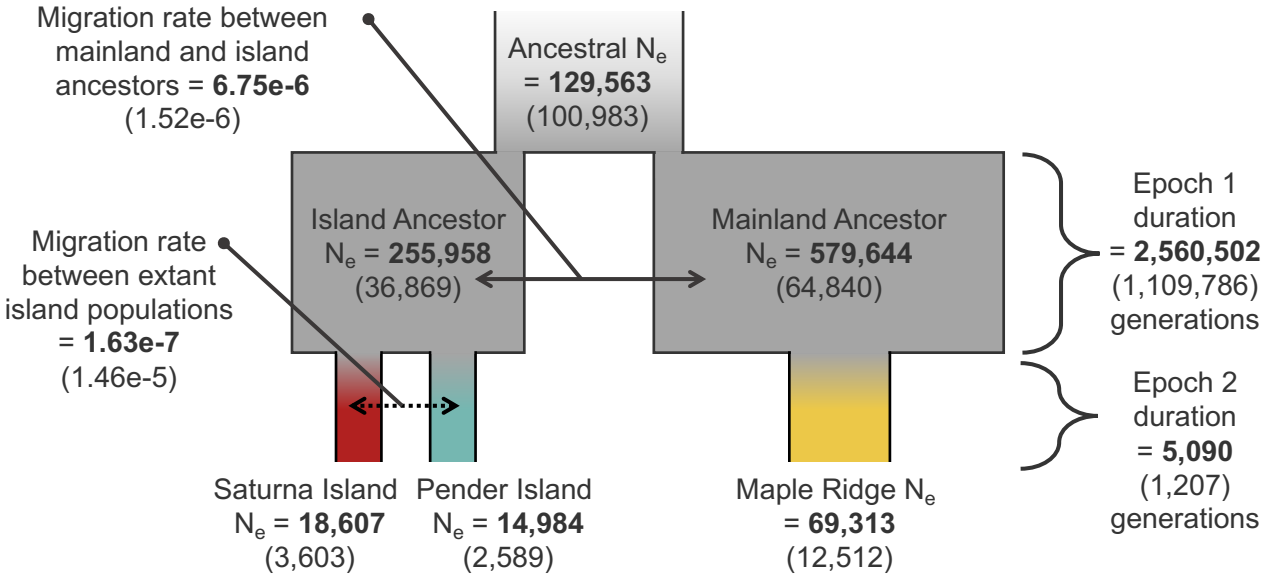

Figure 8. Parameter estimates for the candidate demographic model. Diagram reflects the structure of the best-fit three population demographic model. Boxes reflect population size changes through time (not drawn to scale). Gray shading reflects ancestral populations included in the model. Red, blue, and yellow shading reflects the contemporary Saturna Island, Pender Island, and mainland Maple Ridge populations, respectively. Arrows denote symmetric migration parameters included in the model. Maximum likelihood estimates for each parameter are bolded and standard deviations are given in parentheses. Standard deviations were obtained using the Godambe Information Matrix (Coffman et al. 2016) as an alternative to conventional bootstrapping.

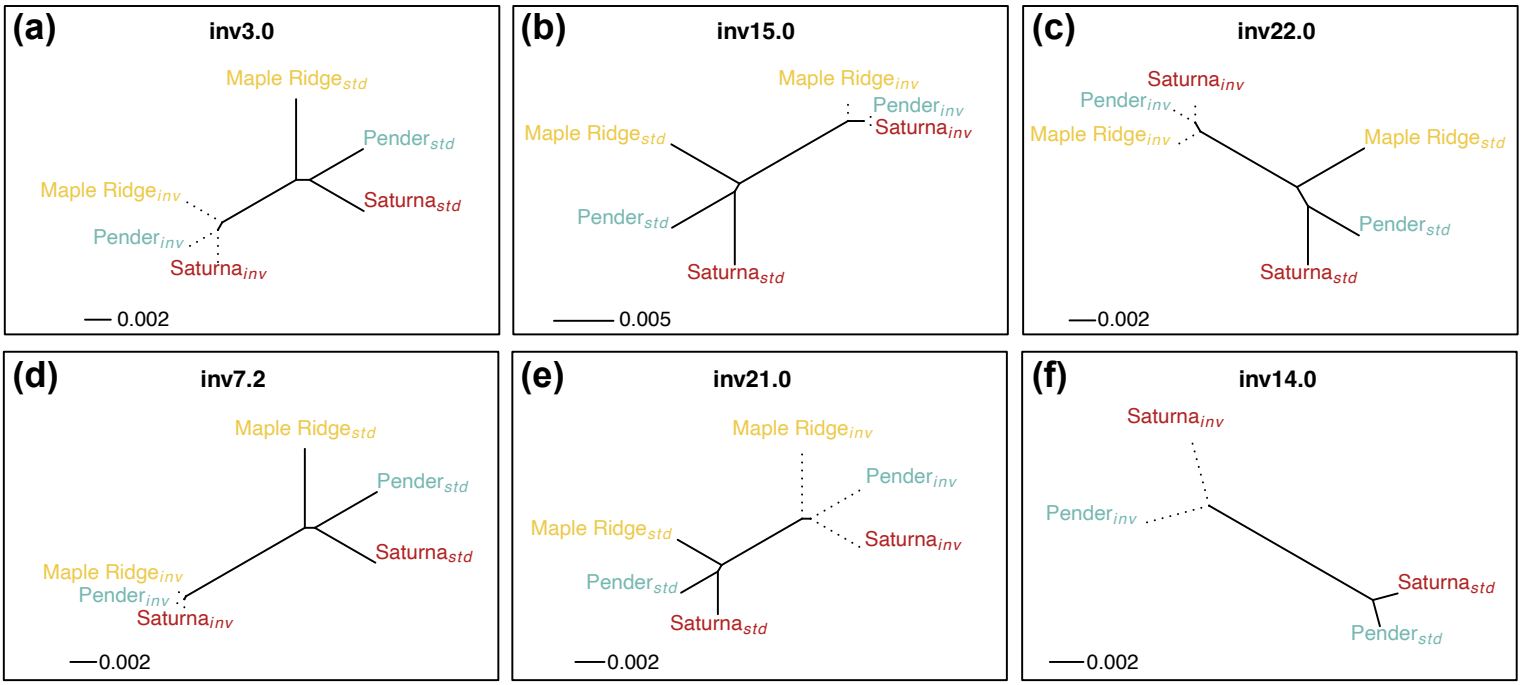

Figure 9. Divergence within and between karyotypes provide insights about the emergence of inversion polymorphisms. Panels depict unrooted, neighbor-joining trees relating standard and inverted karyotypes sampled from each population at each polymorphic locus: inv3.0 (a), inv15.0 (b), inv22.0 (c), inv7.2 (d), inv21.0 (e), and inv14.0 (f). Labels are colored according to population (red for Saturna Island, blue for Pender Island, and yellow for Maple Ridge) and subscripts denote karyotype. Dotted edges represent terminal branches for inverted karyotypes. Distance matrices were constructed using mean divergence (d_{XY}) measured among karyotypic classes sampled within and between populations. Trees were constructed using the neighbor-joining algorithm of Saitou and Nei (1987). Source data is provided in Supplementary Table S5. Maple Ridge karyotypes are excluded from the inv14.0 tree (f) due to a lack of inversion homozygotes.