PROCEEDINGS B

royalsocietypublishing.org/journal/rspb

Research

Cite this article: Booker TR, Payseur BA, Tigano A. 2022 Background selection under evolving recombination rates. *Proc. R. Soc. B* **289**: 20220782.
https://doi.org/10.1098/rspb.2022.0782

Received: 23 April 2022
Accepted: 25 May 2022

Subject Category:
Evolution

Subject Areas:
evolution, genetics, genomics

Keywords:
background selection, recombination rate, evolutionary genetics, chromosomal rearrangements, *Mus musculus*

Author for correspondence:
Tom R. Booker
e-mail: booker@zoology.ubc.ca

Electronic supplementary material is available online at https://doi.org/10.6084/m9.figshare.c.6026286.

THE ROYAL SOCIETY PUBLISHING

Background selection under evolving recombination rates

Tom R. Booker[1], Bret A. Payseur[2] and Anna Tigano[3]

[1]Department of Zoology, University of British Columbia, Vancouver Campus, Vancouver, BC, Canada
[2]Laboratory of Genetics, University of Wisconsin - Madison, Madison, WI, USA
[3]Department of Biology, University of British Columbia, Okanagan Campus, Kelowna, BC, Canada

TRB, 0000-0001-8403-6219; AT, 0000-0001-9240-3058

Background selection (BGS), the effect that purifying selection exerts on sites linked to deleterious alleles, is expected to be ubiquitous across eukaryotic genomes. The effects of BGS reflect the interplay of the rates and fitness effects of deleterious mutations with recombination. A fundamental assumption of BGS models is that recombination rates are invariant over time. However, in some lineages, recombination rates evolve rapidly, violating this central assumption. Here, we investigate how recombination rate evolution affects genetic variation under BGS. We show that recombination rate evolution modifies the effects of BGS in a manner similar to a localized change in the effective population size, potentially leading to underestimation or overestimation of the genome-wide effects of selection. Furthermore, we find evidence that recombination rate evolution in the ancestors of modern house mice may have impacted inferences of the genome-wide effects of selection in that species.

1. Introduction

Different modes of selection (e.g. positive, purifying and balancing) all affect genetic variation at sites linked to the actual targets of selection (reviewed in [1]). In the case of purifying selection, the removal of deleterious mutations causes linked neutral variants to be lost along with them through a process referred to as background selection (BGS; [2]). Of the mutations that affect fitness in natural populations, the vast majority are likely deleterious, with a comparatively small proportion of beneficial mutations [3]. For those reasons, it has been proposed that BGS is ubiquitous across eukaryotic genomes and should be incorporated into null models for population genomics [4,5]. Indeed, recent studies have used BGS to set baseline patterns for identifying the locations and effects of positively selected mutations [6,7] and understanding Lewontin's paradox of genetic diversity [8]. Interpreting genome-wide patterns of genetic diversity in terms of BGS, however, requires accurate estimates of population genetic parameters, particularly recombination rates.

In many species, the recombination rate per base pair (r) varies across the genome both between and within chromosomes [9]. For example, in the house mouse (*Mus musculus*), the average r for chromosome 19 (the shortest chromosome) is around 60% higher than for chromosome 1 (the longest chromosome; [10]). The requirement of at least one crossover per chromosome per meiosis in mammals causes shorter chromosomes to recombine at a higher average rate than longer ones [11–13]. Local recombination rates can vary substantially across chromosomes as well, and, in some cases, this variation is predicted by gross features of chromosome architecture such as the locations of centromeres and telomeres [14]. In mice, the majority of crossovers occur in a minority of the genome, in narrow windows (in the order of 1 to 5 kbp) referred to as hotspots [14]. The positions of recombination hotspots in mice and, in some other vertebrates, are determined by the binding of a protein encoded by the *PRDM9* gene to specific DNA motifs [15,16], although hotspots are still observed in *PRDM9* knockout lines and in dogs, which lack a functional copy of *PRDM9* [17,18].

Estimates of r can be obtained empirically by examining the inheritance of genetic markers through controlled crosses or through pedigrees, or by comparing an individual's genome to that of its gametes (e.g. [19]). Both methods reconstruct recombination events over one or a few generations and thus provide estimates of r for contemporary populations. Alternatively, estimates of r can be obtained indirectly by analysing patterns of linkage disequilibrium across the genome (e.g. [20]), in which case estimates reflect both recent and ancestral recombination events. Whether recombination rates are estimated from marker transmission or population genetics, using such estimates when analysing variation across the genome in terms of BGS implicitly assumes that the recombination landscape has not changed over the time in which patterns of diversity have been established. However, recombination rate landscapes can evolve rapidly in some lineages. For example, owing to the relationship between chromosome size and average r, changes in chromosome length (i.e. karyotype evolution) may induce changes in r. The lineage leading to *Mus musculus* ($2n = 40$) has experienced large chromosomal rearrangements since it shared a common ancestor with *Mus pahari* ($2n = 48$) 3–5 million years ago [21]. Moreover, populations of *Mus musculus domesticus* harbouring different karyotypes exhibit different genomic landscapes of recombination [22]. Chromosomal fusions can exhibit meiotic drive [23] so new karyotypes may spread to fixation very rapidly. Even mice with the same karyotype vary in regional recombination rate across substantial proportions of the genome [24,25] and in total number of crossovers [26,27], both within and between sub-species. There is also evidence that *PRDM9*, the gene that encodes the protein that dictates the locations of recombination events, has undergone recurrent bouts of positive selection in mice [28], and wild-derived lines of *M. musculus* spp. possess various *PRDM9* alleles corresponding to different suites of recombination hotspots [29]. Overall, there is clear evidence from mice that recombination rates can evolve rapidly on broad and fine scales.

Changes in the recombination rate over time may influence patterns of genetic variation across the genome [4,30]. For example, chromosomal fusions would decrease recombination rates experienced by individual nucleotides in the fused chromosomes and thus increase the effects of BGS and other processes mediated by recombination. Consistent with this hypothesis, Cicconardi *et al*. [31] found evidence suggesting that chromosomes that underwent fusions in the ancestors of extant *Heliconius* butterfly species now exhibit reduced recombination rates and nucleotide diversity (π), presumably owing to amplified BGS effects. Following the evolution of the recombination rate landscape, there will be a lag period wherein patterns of genetic variation more closely reflect ancestral recombination rates than derived rates. Furthermore, the fixation of recombination modifiers may be associated with demographic effects. For example, the fixation of a new chromosomal variant would resemble a localized population bottleneck. Over time, as new deleterious mutations arise and cause BGS, patterns of genetic variation will come to reflect derived recombination rates. Depending on the extent and pace of recombination rate evolution, effects of BGS may be obscured in lineages that are still within the lag period. In this paper, we examine how patterns of neutral genetic variation under BGS respond to evolution of the recombination rate and describe how this could affect analyses that are used to identify the effects of selection on a genome-wide scale.

2. Results

(a) Background selection under evolving recombination rates

The effects of BGS reflect the interplay of purifying selection and recombination [32], so changes to the recombination rate will influence the effects of BGS. Under strong purifying selection, BGS resembles a localized reduction in the effective population size (N_e). N_0 is used to denote the hypothetical effective population size that would be expected in the absence of all linked selection effects. It is difficult to estimate N_0 from empirical data [33], but in some species N_0 may be approximated by examining genomic regions that are expected to be only weakly affected by BGS [5]. An increase in the recombination rate between neutral sites and sites subject to purifying selection will decrease the effect of BGS (bringing N_e closer to N_0) and *vice versa* for a decrease in the recombination rate. At a neutral locus v, coalescence times under BGS ($T_{BGS,v}$) are shorter than those expected in the absence of linked selection (T_0) [32] and the effect of BGS is often expressed as $B_v = T_{BGS,v}/T_0 = N_{e,v}/N_0$ (e.g. [32]). Consider a population that underwent a change in the recombination rate such that v experiences a BGS effect of B'_v under the derived recombination rate regime. Even with instantaneous changes in the recombination rate, genetic variation at v would not reflect B'_v immediately, as there would be a lag period after recombination rate change wherein coalescence times (and patterns of genetic variation) would more closely reflect B_v.

If BGS resembles a localized reduction in N_e under strong purifying selection, the period of lag after a change in the recombination rate may resemble the change in coalescence times following a change in the population size. If the recombination rate changed at time t in the past (measured in $2N_0$ generations), then BGS under the new recombination rate can be described with

$$B_{v,t} = B_v\left(1 + \left(\frac{B_v}{B'_v} - 1\right)e^{-t}\right). \tag{2.1}$$

We obtained equation (2.1) by modifying an expression that describes coalescence times after an instantaneous change in the population size from Johri *et al*. [5]. Note that Pool & Nielsen [34] provided similar expressions to those given by Johri *et al*. [5].

We modelled deleterious mutations occurring in a single functional element (e.g. a protein-coding exon) and examined π for neutral mutations in and around this region after an instantaneous change in the recombination rate (electronic supplementary material, figure S1). The value of π gradually departed from the expectations based on the ancestral recombination rate over $4N_0$ generations, when it finally aligned to the expectation under the derived recombination rate (figure 1; electronic supplementary material, figure S2). Up to approximately $2N_0$ generations after a change in the recombination rate, π more closely resembled the expectation under the ancestral recombination rate than it did under the derived rate (figure 1; electronic supplementary material, figure S2). After around $4N_0$ generations, coalescence times closely reflected those expected under BGS given the derived

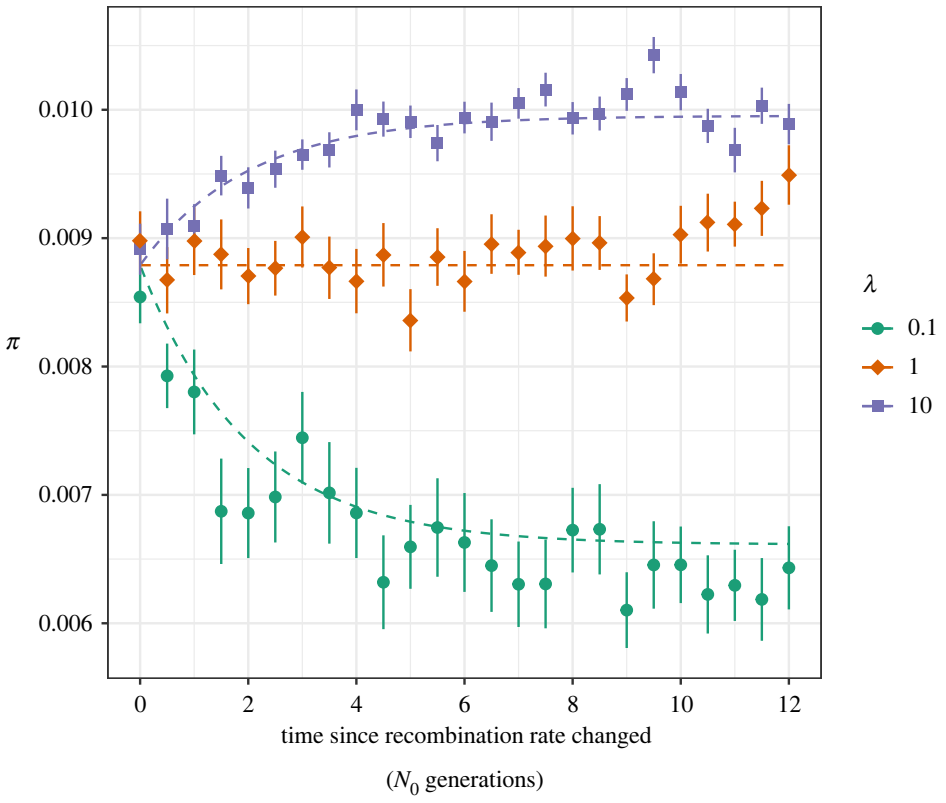

Figure 1. The effect of BGS on nucleotide diversity (π) over time after recombination rates change by a factor λ. The dashed lines were calculated using equation (2.1) and formulae from Nordborg *et al.* [32]. Points indicate the mean from 50 replicate simulations. Nucleotide diversity was calculated for neutral sites 10 000 bp away from sites subject to purifying selection. Points are shown with error bars indicating ± 1 standard error. (Online version in colour.)

recombination rate, as measured by π (figure 1; electronic supplementary material, figure S2). However, it is important to note that when deleterious mutations have nearly neutral deleterious effects, equation (2.1) may not predict changes in nucleotide diversity particularly well because in such cases BGS does not resemble a simple reduction in N_e [35,36].

In the case of a population that has recently undergone shifts in the recombination rate landscape (i.e. fewer than $2N_0$ generations ago), estimates of r would likely reflect contemporary recombination rates regardless of how they were obtained. Estimates of r from patterns of marker inheritance in crosses or pedigrees always reflect contemporary rates, and our results suggest that population genetic estimates (i.e. obtained from patterns of linkage disequilibrium (LD)) may reflect contemporary recombination rates within $0.5N_0$ generations of an increase or decrease in r (electronic supplementary material, figure S3). However, LD-based estimates of the recombination rate are downwardly biased by the effects of BGS (electronic supplementary material, figure S3).

Depending on the direction and magnitude of recombination rate change, population genomic analyses that compare features of genetic variation with estimates of r could underestimate or overestimate the effects of BGS (and other forms of selection) on patterns of genetic variation. For example, LD-based estimates of r (in units $4N_er$) have been compared with estimates of nucleotide diversity, which, for neutral sites, approximates $4N_e\mu$, to investigate the relative role of recombination and mutation in shaping patterns of linked selection (e.g. [37]). Applying such an analysis to lineages that have recently undergone evolution of the recombination map may overestimate or underestimate the effects of BGS, depending on how recombination rates have changed.

(b) Patterns of genetic variation after evolution of the recombination landscape

To understand how population genomic analyses may be affected by changes in r, we simulated two scenarios of BGS under evolving recombination rates. In the first, the broad-scale landscape of r was rearranged (electronic supplementary material, figure S4A). In the second, the locations of recombination hotspots were shifted, as if a new PRDM9 allele had fixed in a population (electronic supplementary material, figure S4B). In both scenarios, deleterious mutations occurred at random across the genome, generating widespread BGS such that there was a positive correlation between π and r at equilibrium (figure 2). For the sake of simplicity, our analyses assumed that recombination rate is invariant among individuals, even as heritable variation in recombination rates has been reported in several species (reviewed in [9]).

A positive correlation between π and r is a hallmark of selection at linked sites across a genome [38], but the evolution of the recombination rate may obscure this pattern. In both the scenarios we simulated, changes in r did not influence the average nucleotide diversity across simulated chromosomes (electronic supplementary material, figure S5), because, under the models of recombination rate evolution we implemented, the average map length was constant over time. However, before the change in the recombination rate, there was a positive correlation between π and r in both scenarios that was detectable when examining 10 kbp,

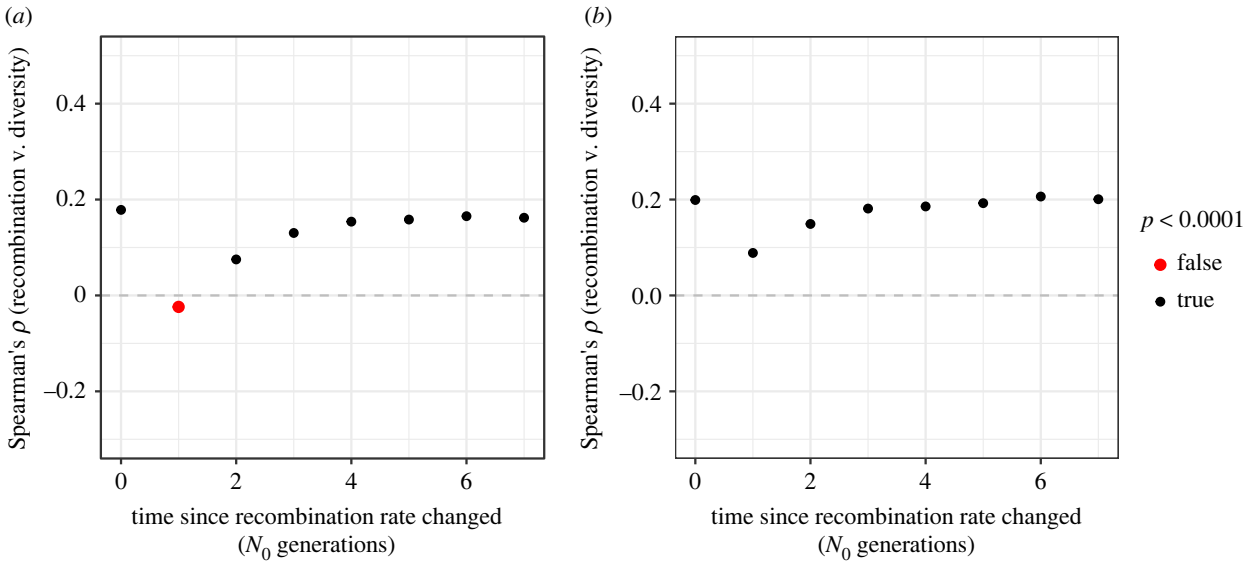

Figure 2. Spearman's correlation between nucleotide diversity (π) and recombination rate (r) over time after recombination rates evolve; (*a*) shows results for a broad-scale shift in the recombination landscape and (*b*) shows results for recombination rate evolution by the movement of hotspots. Results are shown for 10 kbp analysis windows. (Online version in colour.)

100 kbp and 1 Mbp analysis windows (figure 2; electronic supplementary material, figure S6). Following changes in the recombination rate landscape under the model of broad-scale recombination rate variation, the correlation between π and r was either absent or misleading (figure 2*a*; electronic supplementary material, figure S6A). Under the model of recombination hotspot evolution, the correlation between π and r was weakened by change in the landscape of hotspots (figure 2*b*). In both cases we simulated, after about $4N_0$ generations a positive correlation between π and derived r was restored to levels similar to what had been observed before the recombination maps changed (figure 2; electronic supplementary material, figure S6). Figure 2 shows results for 10 000 bp analysis windows; similar results were found when examining larger windows (electronic supplementary material, figure S5).

Both scenarios we simulated are hypothetical, but they demonstrate that recombination rate evolution can readily obscure population genomic analyses aimed at understanding genome-wide processes. In the case of the broad-scale map, we simulated a quite extreme scenario that caused the correlation between π and r to change sign (figure 2*a*). While this scenario may be unrealistic in nature, it serves to demonstrate the effects that recombination rate evolution could exert on patterns of diversity. The simulations modelling hotspot evolution were more realistic, as we based the length, number and intensity of recombination hotspots on results for mice. Given the extensive evidence of recombination hotspot evolution in mice, the simulations we performed demonstrate that evolution of hotspots could indeed weaken genome-wide patterns such as the relationship between π and r (figure 2*b*).

(c) Rapid recombination rate evolution in house mice

Rapid evolution of recombination rates in *Mus musculus* may have influenced our ability to identify the effects of selection across that species' genome. Kartje *et al*. [39] recently demonstrated that natural populations of *M. m. domesticus* exhibit a very weak correlation between π and r (when examining analysis windows of various widths) and concluded that selection at linked sites exerted only modest effects on genetic variation throughout the genome. This is notable because wild mice are thought to have large effective population sizes for mammals [40] and genome-wide effects of selection are thought to be more pronounced in species with large N_e [38]. As discussed in the Introduction, there is evidence that mice have undergone rapid evolution of the recombination rate. For example, around 3–5 Ma, the lineage leading to *M. musculus* experienced a burst of karyotype evolution [21]. If that burst of karyotype evolution affected recombination rates and ancestral mouse populations were very large, then contemporary mice may still be within the lag period described by equation (2.1). Patterns of genetic diversity in mice may still be adjusting to historical changes in the recombination rate, and we may see a stronger correlation between π and r in genomic regions that have not undergone dramatic changes in the recombination rate.

Using an alignment of genomes from closely related species, Thybert *et al*. [21] distinguished chromosomes in the *M. musculus* genome that have undergone dramatic rearrangements in the last 5 million years from those that have not. We re-analysed data from Kartje *et al*. [39] and found that the correlation coefficient between π and r was more positive on chromosomes that have not undergone large-scale rearrangements in the last 3–5 million years (table 1) for *M. m. domesticus* individuals from France and Germany. This pattern holds when looking at analysis windows of 5 kbp and 1 Mbp (table 1). No substantial correlations were found for mice from Gough Island in any comparison. *M. m. domesticus* are believed to have colonized Gough Island in the nineteenth century and to have experienced a severe population bottleneck [41], a demographic history that could have further obscured the correlation between nucleotide diversity and recombination rate in that population.

3. Discussion

The evolution of the recombination rate will influence the effects of selection at linked sites (e.g. BGS and selective

Table 1. The correlation between nucleotide diversity (π) and recombination rate (r) for three populations of house mice (*Mus musculus domesticus*) calculated from all autosomes, conserved chromosomes that exhibit no syntenic breaks between *M. musculus* and *M. pahari*, and chromosomes that experienced large-scale rearrangements as identified by Thybert *et al*. [21]. Correlations with p-values less than 0.00056 (a Dunn–Šidák corrected $\alpha = 0.01$) are highlighted in italics.

		whole genome		conserved chromosomes		rearranged chromosomes	
window size	population	Spearman's ρ	p-value	Spearman's ρ	p-value	Spearman's ρ	p-value
5 kbp	Gough Island	*0.00767 [0.00340–0.0113]*	*4.28×10^{-5}*	0.00880 [0.00208–0.0155]	0.0102	0.00486 [0.000466–0.00925]	0.0302
5 kbp	France	0.00408 [0.000407–0.00775]	0.0295	*0.0403 [0.0336–0.0470]*	*6.10×10^{-32}*	*−0.0107 [−0.0151 – −0.00632]*	*1.76×10^{-6}*
5 kbp	Germany	*0.00752 [0.00385–0.0112]*	*6.05×10^{-5}*	*0.0152 [0.00844–0.0219]*	*9.63×10^{-6}*	0.00386 [—0.00053–0.00825]	0.0849
1 Mbp	Gough Island	0.0536 [0.0131–0.0938]	0.009 46	0.0588 [—0.0162–0.1332]	0.124	0.0437 [—0.00438–0.0916]	0.0748
1 Mbp	France	0.0450 [0.00452–0.0853]	0.0294	*0.135 [0.0606–0.208]*	*0.000 400*	0.00999 [—0.0381–0.0580]	0.684
1 Mbp	Germany	0.0535 [0.0131–0.0938]	0.00953	0.0775 [0.0025–0.0152]	0.0428	0.0426 [—0.00554–0.0905]	0.0828

sweeps) and thus influence patterns of genetic variability. Estimates of the recombination rate made from contemporary populations may not adequately predict genetic variability up to $2N_0$ generations following evolution of the recombination rate landscape (figures 1 and 2). Our re-analysis of the Kartje *et al*. [39] data suggests that mice are still within the lag period after evolution of the recombination rate, such that π in *M. m. domesticus* does not fully reflect contemporary recombination rates in *M. musculus*. By contrast, the ancestors of *Heliconius* butterflies also underwent large-scale karyotype evolution, but gross patterns of π versus chromosome length in those species suggest that patterns of variation have largely re-equilibrated after changes in r [31].

While our re-analysis of the data from Kartje *et al*. [39] suggests that recombination rate evolution in the ancestors of mice obscures the evidence for natural selection across the genome, the overall correlations between π and r were still fairly weak on the conserved chromosomes (table 1). The largest rank correlation coefficient we found was 0.135 for the sample of *M. m. domesticus* from France (1 Mbp windows; table 1). By contrast, Spearman's rank correlation between nucleotide diversity and recombination rate in humans has been reported to be 0.219 for 400 kbp analysis windows [42]. The variance in recombination rates across the *M. musculus* genome is less than half of that reported for humans [43], so perhaps the effects of BGS across the genome are more homogeneous in *M. musculus* than they are in humans, contributing to the weak correlations between π and r shown in table 1. Beyond the pulse of karyotype evolution reported by Thybert *et al*. [21], there is clear evidence of recent and likely ongoing evolution of the recombination rate in *M. musculus* (see Introduction), which may further obscure genome-wide evidence for the effects of natural selection. For example, there is evidence that the landscape of recombination hotspots in the *M. musculus* genome has evolved rapidly among sub-species and populations [29]. Our simulations suggest that even a single change to the locations of hotspots can substantially weaken the correlation between π and r (figure 2; electronic supplementary material, figure S6). Of course, there are reasons why species may not exhibit a strong positive correlation between π and r that have nothing to do with recombination rate evolution [38]. For example, wild rice and domesticated rice (*Oryza* spp.) exhibit negative correlations between π and r, but in those species, there is a strong positive correlation between the density of functional sites (i.e. sites subject to purifying selection) and the recombination rate [44]. In such a case, the effects of BGS are primarily occurring in regions of high recombination.

In the model we used in this study, we assumed that recombination rate modifiers fix instantaneously. While instantaneous change is unlikely, there is reason to expect that recombination rate evolution may occur rapidly. For example, in the case of karyotype evolution, derived chromosomal variants may exhibit meiotic drive [23] and rapidly spread to fixation. In addition, the spread of some kinds of recombination modifiers may mimic demographic effects. For example, the fixation of a derived chromosomal variant would generate the effect of a localized population size bottleneck followed by an expansion, changes that can interact with BGS [4,5,45,46]. BGS acting during the course of a bottleneck + expansion scenario leads to levels of nucleotide diversity that deviate from standard BGS predictions [46]. The interaction between recombination rate evolution and demography in the context of BGS is an area for further research.

This paper should add to the growing appreciation of recombination as an evolutionarily labile trait. As pointed out by Smukowski Heil *et al*. [47], Comeron [4], and Pettie *et al*. [30], information on recombination rates in outgroup species is an important covariate when performing population genomic analyses. In some lineages, recombination rates may evolve slowly. Birds, for example, have highly conserved karyotypes and in some cases highly conserved recombination landscapes [48,49]. Evolution of the recombination rate is one of many possible reasons why one might not be able to adequately identify the effects of BGS (or natural selection more broadly) from population genomic data (see reviews in [38] and [4]), but conservation of recombination landscapes will likely make comparative population genomics more straightforward [30].

4. Methods

(a) Model

BGS has been modelled as the reduction in effective population size (N_e) at a neutral site due to the removal of linked deleterious variants. The effects of BGS are often expressed as $B = N_e/N_0$, where N_e is the effective population size and N_0 is the expected population size under strict neutrality. In a non-recombining genome, B is proportional to the ratio of the deleterious mutation rate to the strength of selection acting on harmful mutations [2]. For a neutral site present on a recombining chromosome, the effects of BGS depend on the density of functional sites (i.e. those that can mutate to deleterious alleles), the strength of selection at functional sites, the mutation rate at functional sites and the recombination rate between the neutral site and the functional sites [32,50,51]. For a neutral locus v linked to x functional sites, the reduction in N_e has been described with the following equation:

$$B_v = \frac{N_e}{N_0} = \exp\left[-\sum_x \frac{u_x}{t(1+(1-t)r_{x,v}/t)^2}\right],$$

where u_x is the deleterious mutation rate at functional site x, t is the heterozygous fitness effect of a deleterious mutation (i.e. $0.5s$ in the case of semi-dominance) and $r_{x,v}$ is the recombination map distance between the neutral locus and functional site x. In the above equation, deleterious mutations have fixed effects, but it is straightforward to incorporate a distribution of fitness effects [32]. The above equation holds when selection is sufficiently strong that random drift does not overwhelm selection ($N_e s > 1$) [35].

(b) Simulations

We simulated BGS under recombination rate evolution using two types of simulations in *SLiM* v. 3.2 [52]. We simulated diploid populations of $N_e = 5000$ individuals. In all cases, we scaled mutation, recombination and the strength of selection to approximate evolution in a large population.

The first set of simulations was designed to examine how long it takes for patterns of neutral diversity under BGS to equilibrate after the recombination rate evolves. In these simulations, the genome was 25 kbp long with a 5 kbp functional element in the centre. Mutations occurred in the functional element at rate $\mu = 5 \times 10^{-7}$ and had semi-dominant fitness effects with a fixed selection coefficient of $s = -0.01$. Recombination occurred at a uniform rate of $r = 5 \times 10^{-7}$ across the chromosome. After 80 000 generations ($16N_e$ generations), we simulated an

instantaneous change in the recombination rate, multiplying r by λ, giving $r = 5 \times 10^{-7}\lambda$. We simulated cases with $\lambda = 0.1$, 1.0 and 10.0. Simulated populations were sampled every 2500 generations after the recombination rate changed and we performed 200 replicates for each set of parameters tested. Note that these simulations were not designed to be particularly realistic, but to provide clear-cut patterns to test the theoretical predictions.

The second set of simulations was designed to examine how patterns of π versus r varied over time when recombination rates evolved at fine and/or broad scales. For these simulations, we modelled chromosomes that were 10 Mbp long. Neutral mutations occurred at random across the length of the sequence at a rate of 5×10^{-7} (such that expected nucleotide diversity was 0.01). Deleterious mutations occurred at random across the length of the sequence at a rate of 5×10^{-8} with semi-dominant fitness effects drawn from a gamma distribution with a mean ($\bar{s}$) of −0.1 and a shape parameter of 0.1. The deleterious mutation rate was chosen so that 10% of the genome was subject to purifying selection. Populations evolved under BGS for 80 000 generations (i.e. $16N_e$ generation). In generation 80 000, there was instantaneous evolution of the recombination landscape, after which we recorded the tree-sequence of the population every 5000 generations for a further 40 000 generations. We incorporated two models of recombination rate variation and evolution of the recombination map as follows.

(1) We modelled recombination rate evolution at broad scales by rearrangement of the recombination landscape. Recombination rates vary across the genome [9]. For example, recombination rates vary by a factor of 3 across chromosome 1 in mice. In these simulations, recombination varied from $r = 2.08 \times 10^{-7}$ to $r = 6.24 \times 10^{-7}$ across the simulated chromosome (electronic supplementary material, figure S4A). When the recombination landscape evolved, we reversed the order of recombination rates across the genome (electronic supplementary material, figure S4A).

(2) We modelled evolution of the recombination map by the movement of hotspots. Recombination occurred at a uniform rate of $r = 6 \times 10^{-8}$, except in 5 kbp hotspots, where it occurred at a rate of $r = 6 \times 10^{-6}$. At the beginning of a simulation, a Poisson number of hotspots was sampled with an expectation of 120. Hotspots were placed at random across the simulated chromosome. When the recombination landscape evolved, we resampled the locations of hotspots (electronic supplementary material, figure S4B).

In both cases, rates were chosen such that the total map length was similar to one that recombined at a constant rate of $4N_e r = 0.008$, the value reported for wild mice [53]. For both models of recombination rate map evolution, we performed 20 simulation replicates, giving a total of 200 Mbp-worth of simulated data, similar to the length of chromosome 1 in mice.

For all simulations, we used the tree-sequence recording option in *SLiM* and neutral mutations were added to the resulting tree-sequences at a rate of 5×10^{-7} using *PySLiM* and *msprime* [54,55]. Nucleotide diversity (π) was calculated in windows of varying size using *sci-kit-allel*. We used the program *PyRho* [20] to estimate recombination rates from samples of 10 diploid individuals from 20 replicate simulations. Spearman's ρ between π and r was calculated using R [56] . All figures were made using *ggplot2*. All simulation scripts and analysis and plotting scripts are deposited at https://github.com/TBooker/BGS_RecombinationRateEvolution.

Data accessibility. All simulation scripts and analysis and plotting scripts are deposited at https://github.com/TBooker/BGS_RecombinationRateEvolution.

Authors' contributions. T.R.B.: conceptualization, formal analysis, investigation, methodology, writing—original draft, and writing—review and editing; B.A.P.: conceptualization, resources, and writing—review and editing; A.T.: conceptualization, and writing—review and editing.

All authors gave final approval for publication and agreed to be held accountable for the work performed herein.

Conflict of interest declaration. We declare we have no competing interests.

Funding. T.R.B. was funded by a Bioinformatics Fellowship awarded by the Biodiversity Research Centre at the University of British Columbia. B.A.P. was supported by NIH R35 GM139412.

Acknowledgements. Thanks to Nadia Singh and Judith Mank for the invitation to present this work at vSMBE 2021. Thanks to Michael Whitlock, Brian Charlesworth and Mikey Kartje for helpful discussions.

References

1. Charlesworth B. 2009 Fundamental concepts in genetics: effective population size and patterns of molecular evolution and variation. *Nat. Rev. Genet.* **10**, 195–205. (doi:10.1038/nrg2526)
2. Charlesworth B, Morgan MT, Charlesworth D. 1993 The effect of deleterious mutations on neutral molecular variation. *Genetics* **134**, 1289–1303. (doi:10.1093/genetics/134.4.1289)
3. Eyre-Walker A, Keightley PD. 2007 The distribution of fitness effects of new mutations. *Nat. Rev. Genet.* **8**, 610–618. (doi:10.1038/nrg2146)
4. Comeron JM. 2017 Background selection as null hypothesis in population genomics: insights and challenges from *Drosophila* studies. *Phil. Trans. R. Soc. Lond. B* **372**, 1736. (doi:10.1098/rstb.2016.0471)
5. Johri P, Charlesworth B, Jensen JD. 2020 Toward an evolutionarily appropriate null model: jointly inferring demography and purifying selection. *Genetics* **215**, 173–192. (doi:10.1534/genetics.119.303002)
6. DeGiorgio M, Huber CD, Hubisz MJ, Hellmann I, Nielsen R. 2016 SweepFinder 2: increased sensitivity, robustness and flexibility. *Bioinformatics* **32**, 1895–1897. (doi:10.1093/bioinformatics/btw051)
7. Campos JL, Zhao L, Charlesworth B. 2017 Estimating the parameters of background selection and selective sweeps in *Drosophila* in the presence of gene conversion. *Proc. Natl Acad. Sci. USA* **114**, E4762–E4771. (doi:10.1073/pnas.1619434114)
8. Buffalo V. 2021 Quantifying the relationship between genetic diversity and population size suggests natural selection cannot explain Lewontin's paradox. *eLife* **10**, e67509. (doi:10.7554/eLife.67509)
9. Stapley J, Feulner PGD, Johnston SE, Santure AW, Smadja CM. 2017 Variation in recombination frequency and distribution across eukaryotes: patterns and processes. *Phil. Trans. R. Soc. B* **372**, 20160455. (doi:10.1098/rstb.2016.0455)
10. Cox A *et al.* 2009 A new standard genetic map for the laboratory mouse. *Genetics* **182**, 1335–1344. (doi:10.1534/genetics.109.105486)
11. Pardo-Manuel de Villena F, Sapienza C. 2001 Recombination is proportional to the number of chromosome arms in mammals. *Mamm. Genome* **12**, 318–322. (doi:10.1007/s003350020005)
12. Segura J *et al.* 2013 Evolution of recombination in eutherian mammals: insights into mechanisms that affect recombination rates and crossover interference. *Proc. R. Soc. B* **280**, 20131945. (doi:10.1098/rspb.2013.1945)
13. Dumont BL. 2017 Variation and evolution of the meiotic requirement for crossing over in mammals. *Genetics* **205**, 155–168. (doi:10.1534/genetics.116.192690)
14. Paigen K, Szatkiewicz JP, Sawyer K, Leahy N, Parvanov ED, Ng SH, Graber JH, Broman KW, Petkov PM. 2008 The recombinational anatomy of a mouse chromosome. *PLoS Genet.* **4**, e1000119. (doi:10.1371/journal.pgen.1000119)
15. Baudat F, Buard J, Grey C, Fledel-Alon A, Ober C, Przeworski M, Coop G, de Massy B. 2010 *PRDM9* is a major determinant of meiotic recombination hotspots in humans and mice. *Science* **327**, 836–840. (doi:10.1126/science.1183439)
16. Baker Z, Schumer M, Haba Y, Bashkirova L, Holland C, Rosenthal GG, Przeworski M. 2017 Repeated

losses of *PRDM9*-directed recombination despite the conservation of PRDM9 across vertebrates. *eLife* **6**, e24133. (doi:10.7554/eLife.24133)
17. Brick K *et al.* 2012 Genetic recombination is directed away from functional genomic elements in mice. *Nature* **485**, 642–645. (doi:10.1038/nature11089)
18. Auton A *et al.* 2013 Genetic recombination is targeted towards gene promoter regions in dogs. *PLoS Genet.* **9**, e1003984. (doi:10.1371/journal.pgen.1003984)
19. Sun H, Rowan BA, Flood PJ, Brandt R, Fuss J, Hancock AM, Michelmore RW, Huettel B, Schneeberger K. 2019 Linked-read sequencing of gametes allows efficient genome-wide analysis of meiotic recombination. *Nat. Commun.* **10**, 4310. (doi:10.1038/s41467-019-12209-2)
20. Spence JP, Song YS. 2019 Inference and analysis of population-specific fine-scale recombination maps across 26 diverse human populations. *Sci. Adv.* **5**, 10. (doi:10.1126/sciadv.aaw9206)
21. Thybert D *et al.* 2018 Repeat associated mechanisms of genome evolution and function revealed by the *Mus caroli* and *Mus pahari* genomes. *Genome Res.* **28**, 448–459. (doi:10.1101/gr.234096.117)
22. Vara C *et al.* 2021 The impact of chromosomal fusions on 3D genome folding and recombination in the germ line. *Nat. Commun.* **12**, 2981. (doi:10.1038/s41467-021-23270-1)
23. Chmátal L, Gabriel SI, Mitsainas GP, Martínez-Vargas J, Ventura J, Searle JB, Schultz RM, Lampson MA. 2014 Centromere strength provides the cell biological basis for meiotic drive and karyotype evolution in mice. *Curr. Biol.* **24**, 2295–2300. (doi:10.1016/j.cub.2014.08.017)
24. Dumont BL *et al.* 2011 Extensive recombination rate variation in the house mouse species complex inferred from genetic linkage maps. *Genome Res.* **21**, 114–125. (doi:10.1101/gr.111252.110)
25. Wang RJ *et al.* 2017 Recombination rate variation in mice from an isolated island. *Mol. Ecol.* **26**, 457–470. (doi:10.1111/mec.13932)
26. Dumont BL, Payseur BA. 2011 Evolution of the genomic recombination rate in murid rodents. *Genetics* **187**, 643–657. (doi:10.1534/genetics.110.123851)
27. Peterson AL, Payseur BA. 2021 Sex-specific variation in the genome-wide recombination rate. *Genetics* **217**, iyaa019. (doi:10.1093/genetics/iyaa019)
28. Oliver PL *et al.* 2009 Accelerated evolution of the *Prdm9* speciation gene across diverse metazoan taxa. *PLoS Genet.* **5**, e1000753. (doi:10.1371/journal.pgen.1000753)
29. Smagulova F, Brick K, Yongmei P, Camerini-Otero RD, Petukhova GV. 2016 The evolutionary turnover of recombination hotspots contributes to speciation in mice. *Genes Dev.* **30**, 277–280. (doi:10.1101/gad.270009.115)
30. Pettie N, Llopart A, Comeron JM. 2022 Meiotic, genomic and evolutionary properties of crossover distribution in *Drosophila yakuba*. *PLoS Genet.* **18**, e1010087. (doi:10.1371/journal.pgen.1010087)
31. Cicconardi F, Lewis JJ, Martin SH, Reed RD, Danko CG, Montgomery SH. 2021 Chromosome fusion affects genetic diversity and evolutionary turnover of functional loci but consistently depends on chromosome size. *Mol. Biol. Evol.* **38**, 4449–4462. (doi:10.1093/molbev/msab185)
32. Nordborg M, Charlesworth B, Charlesworth D. 1996 The effect of recombination on background selection. *Genet. Res.* **67**, 159–174. (doi:10.1017/S0016672300033619)
33. Kern AD, Hahn MW. 2018 The neutral theory in light of natural selection. *Mol. Biol. Evol.* **35**, 1366–1371. (doi:10.1093/molbev/msy092)
34. Pool JE, Nielsen R. 2007 Population size changes reshape genomic patterns of diversity. *Evolution* **61**, 3001–3006. (doi:10.1111/j.1558-5646.2007.00238.x)
35. Good BH, Walczak AM, Neher RA, Desai MM. 2014 Genetic diversity in the interference selection limit. *PLoS Genet.* **10**, e1004222. (doi:10.1371/journal.pgen.1004222)
36. Cvijović I, Good BH, Desai MM. 2018 The effect of strong purifying selection on genetic diversity. *Genetics* **209**, 1235–1278. (doi:10.1534/genetics.118.301058)
37. Wang J *et al.* 2016 Natural selection and recombination rate variation shape nucleotide polymorphism across the genomes of three related *Populus* species. *Genetics* **202**, 1185–1200. (doi:10.1534/genetics.115.183152)
38. Cutter AD, Payseur BA. 2013 Genomic signatures of selection at linked sites: unifying the disparity among species. *Nat. Rev. Genet.* **14**, 262–274. (doi:10.1038/nrg3425)
39. Kartje ME, Jing P, Payseur BA. 2020 Weak correlation between nucleotide variation and recombination rate across the house mouse genome. *Genome Biol. Evol.* **12**, 293–299. (doi:10.1093/gbe/evaa045)
40. Leffler EM, Bullaughey K, Matute DR, Meyer WK, Segurel L, Venkat A, Andolfatto P, Przeworski M. 2012 Revisiting an old riddle: what determines genetic diversity levels within species? *PLoS Biol.* **10**, e1001388. (doi:10.1371/journal.pbio.1001388)
41. Gray MM, Wegmann D, Haasl RJ, White MA, Gabriel SI, Searle JB, Cuthbert RJ, Ryan PG, Payseur BA. 2014 Demographic history of a recent invasion of house mice on the isolated island of Gough. *Mol. Ecol.* **23**, 1923–1939. (doi:10.1111/mec.12715)
42. Cai JJ *et al.* 2009 Pervasive hitchhiking at coding and regulatory sites in humans. *PLoS Genet.* **5**, e1000336. (doi:10.1371/journal.pgen.1000336)
43. Jensen-Seaman MI *et al.* 2004 Comparative recombination rates in the rat, mouse, and human genomes. *Genome Res.* **14**, 528–538. (doi:10.1101/gr.1970304)
44. Flowers JM, Molina J, Rubinstein S, Huang P, Schaal BA, Purugganan MD. 2011 Natural selection in gene-dense regions shapes the genomic pattern of polymorphism in wild and domesticated rice. *Mol. Biol. Evol.* **29**, 675–687. (doi:10.1093/molbev/msr225)
45. Beissinger TM, Wang L, Crosby K, Durvasula A, Hufford MB, Ross-Ibarra J. 2016 Recent demography drives changes in linked selection across the maize genome. *Nat. Plants* **2**, 16084. (doi:10.1038/nplants.2016.84)
46. Torres R, Stetter MG, Hernandez RD, Ross-Ibarra J. 2020 The temporal dynamics of background selection in nonequilibrium populations. *Genetics* **214**, 1019–1030. (doi:10.1534/genetics.119.302892)
47. Smukowski Heil CS, Ellison CS, Dubin M, Noor MA. 2015 Recombining without hotspots: a comprehensive evolutionary portrait of recombination in two closely related species of *Drosophila*. *Genome Biol. Evol.* **7**, 2829–2842. (doi:10.1093/gbe/evv182)
48. Damas J, Kim J, Farré M, Griffin DK, Larkin DM. 2018 Reconstruction of avian ancestral karyotypes reveals differences in the evolutionary history of macro- and microchromosomes. *Genome Biol.* **19**, 155. (doi:10.1186/s13059-018-1544-8)
49. Singhal S *et al.* 2015 Stable recombination hotspots in birds. *Science* **350**, 928–932. (doi:10.1126/science.aad0843)
50. Hudson RR, Kaplan NL. 1995 Deleterious background selection with recombination. *Genetics* **141**, 1605–1617. (doi:10.1093/genetics/141.4.1605)
51. Nordborg M. 1997 Structured coalescent processes on different time scales. *Genetics* **146**, 1501–1514. (doi:10.1093/genetics/146.4.1501)
52. Haller BC, Messer PW. 2019 SLiM 3: forward genetic simulations beyond the Wright–Fisher model. *Mol. Biol. Evol.* **36**, 632–637. (doi:10.1093/molbev/msy228)
53. Booker TR, Ness RW, Keightley PD. 2017 The recombination landscape in wild house mice inferred using population genomic data. *Genetics* **207**, 297–309. (doi:10.1534/genetics.117.300063)
54. Haller BC, Galloway J, Kelleher J, Messer PW, Ralph PL. 2019 Tree-sequence recording in SLiM opens new horizons for forward-time simulation of whole genomes. *Mol. Ecol. Resour.* **19**, 552–566. (doi:10.1111/1755-0998.12968)
55. Kelleher J, Etheridge AM, McVean G. 2016 Efficient coalescent simulation and genealogical analysis for large sample sizes. *PLoS Comput. Biol.* **12**, e1004842. (doi:10.1371/journal.pcbi.1004842)
56. R Core Team. 2021 *R: a language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. See https://www.R-project.org/.