Received: 25 February 2020 | Revised: 26 May 2020 | Accepted: 8 June 2020

DOI: 10.1111/jbi.13932

METHODS AND TOOLS

Journal of Biogeography | WILEY

Efficient summary statistics for detecting lineage fusion from phylogeographic datasets

Ryan C. Garrick | Chaz Hyseni | Ísis C. Arantes

Department of Biology, University of Mississippi, Oxford, MS, USA

Correspondence
Ryan Garrick, Department of Biology, University of Mississippi, Oxford, MS 38677, USA.
Email: rgarrick@olemiss.edu

Funding information
National Science Foundation, Grant/Award Number: 1738817; University of Mississippi

Handling Editor: Luis Valente

Abstract

Aim: Lineage fusion (merging of two or more populations of a species resulting in a single panmictic group) is a special case of secondary contact. It has the potential to counteract diversification and speciation, or to facilitate it through creation of novel genotypes. Understanding the prevalence of lineage fusion in nature requires reliable detection of it, such that efficient summary statistics are needed. Here, we report on simulations that characterized the initial intensity and subsequent decay of signatures of past fusion for 17 summary statistics applicable to DNA sequence haplotype data.

Location: Global.

Taxa: Diploid out-crossing species.

Methods: We considered a range of scenarios that could reveal the impacts of different combinations of read length versus number of loci (arrangement of DNA sequence data), and whether or not pre-fusion populations experienced bottlenecks coinciding with their divergence (historical context of fusion). Post-fusion gene pools were sampled along 10 successive time points representing increasing lag times following merging of sister populations, and summary statistic values were recalculated at each.

Results: Many summary statistics were able to detect signatures of complete merging of populations after a sampling lag time of 1.5 N_e generations, but the most informative ones included two neutrality tests and four diversity metrics, with Z_{nS} (a linkage disequilibrium-based neutrality test) being particularly powerful. Correlation was relatively low among the two neutrality tests and two of the diversity metrics. There were clear benefits of many short (200-bp × 200) loci over a handful of long (4-kb × 10) loci. Also, only the latter genetic dataset type showed impacts of bottlenecks during divergence upon the number of informative summary statistics.

Main conclusions: This work contributes to identifying cases of lineage fusion, and advances phylogeography by enabling more nuanced reconstructions of how individual species, or multiple members of an ecological community, responded to past environmental change.

KEYWORDS

demographic history, ephemeral lineage, fission-fusion, lineage fusion, neutrality test, nucleotide diversity, phylogeography, population merging, secondary contact

1 | INTRODUCTION

Pleistocene climatic cycles have had profound impacts on how intraspecific genetic diversity is structured across landscapes (Hewitt, 2004). Despite these events being quite recent, they coincide with speciation in a broad suite of taxa (e.g., Carstens & Knowles, 2007; Dorsey, Gregory, Sass, & Specht, 2018; Ho, Saarma, Barnett, Haile, & Shapiro, 2008). Globally, alternations between cool dry glacials and warm humid interglacials have occurred on ~100 thousand-year (Kyr) rotations over the past ~800 Kyr (Bennett, 1990). The local effects of these Milankovitch oscillations have caused recurrent contraction, fragmentation and isolation, followed by expansion and reconnection of populations over time (Emerson & Hewitt, 2005; Jesus, Wilkins, Solferini, & Wakeley, 2006). Differentiated gene pools that come into secondary contact are expected to merge, unless postzygotic isolating mechanisms maintain their integrity (Burton, Pereira, & Barreto, 2013; Orr & Presgraves, 2000). Thus, Pleistocene climatic cycles repeatedly acted as crucibles for speciation.

Lineage fusion (merging of two or more populations of a species resulting in a single panmictic group) is a special case of the broader phenomenon of secondary contact (Garrick, Banusiewicz, Burgess, Hyseni, & Symula, 2019). While the term has been in use for over a decade (Campbell et al., 2008), the broader concept (i.e., including merging of reproductively isolated lineages) was described much earlier. For example, Lewis and Bloom (1972) described a case of reverse speciation driven by new environmental conditions leading to range overlap. In the context of phylogenetic theory, Savage (1983) argued for inclusion of lineage fusion and reverse speciation in the branching structure of trees to provide more realistic models of diversification. Indeed, visual representations of reticulations were presented by Jansson and Dynesius (2002) to illustrate impacts of glacial–interglacial cycles upon formation, merging, and persistence of population gene pools and intraspecific clades (also see Dynesius & Jansson, 2014; Rosenblum et al., 2012). However, demographic histories involving intraspecific lineage fusion have been largely overlooked in the empirical phylogeographic literature (but see Garrick et al., 2014; Kearns et al., 2018; Webb, Marzluff, & Omland, 2011).

The ability to detect past lineage fusion is important because upon population merging, an abrupt, pronounced, and spatially localized increase in genetic diversity occurs, creating opportunities for rapid evolutionary change via reshuffling of variation (Alcala & Vuilleumier, 2014; Bennett, 1990). Indeed, some adaptive radiations were preceded by admixture events leading to recombination among divergent alleles (Marques, Meier, & Seehausen, 2019). Rather than simply causing a net reduction in extant lineages, fission-fusion dynamics may create new opportunities for selection, and subsequent diversification. Furthermore, if unrecognized, lineage fusion can mask the extent of similarity in which co-distributed species responded to past environmental change (Garrick et al., 2019). By extension, lineage fusion can affect conclusions about the importance of biotic versus abiotic influences on phylogeographic structure (e.g., Garrick, Nason, Fernández-Manjarrés, & Dyer, 2013; Satler & Carstens, 2017).

Several analytical frameworks can reconstruct complex demographic histories, including lineage fusion. For example, approximate Bayesian computation (ABC; Beaumont, Zhang, & Balding, 2002) is quite flexible (Bertorelle, Benazzo, & Mona, 2010), and has shown promise for distinguishing between lineage fusion versus competing demographic models, using multi-locus DNA haplotype datasets for non-model organisms (Garrick et al., 2019). However, for complex fission-fusion scenarios such as those with short divergence between parental populations, asymmetrical mixing during merging, and/or long lag times between fusion and gene pool sampling, considerable power may be needed. The performance of ABC and other frameworks relevant to phylogeography (e.g., Smith et al., 2017) can be optimized through informed choice of summary statistics used to characterize and compare empirical versus simulated datasets (Hickerson, Dolman, & Moritz, 2006). Indeed, a handful of information-rich summary statistics can avoid the 'curse of dimensionality' – where many summary statistics not only impose a larger computational burden, but also lead to an increase in variability and reduction in overall quality of parameter estimates, such that costs outweigh the benefits (Beaumont et al., 2002).

In this study, we used simulations to characterize intensity and subsequent decay of signatures of past fusion for a suite of summary statistics applicable to DNA haplotype datasets. We constructed simple lineage fusion models and corresponding 'baseline' non-fusion scenarios, and then used successive temporal sampling of population gene pools to quantify the magnitude of changes over time in summary statistic values. Within this framework, we examined the impact of the: (a) arrangement of DNA sequence data (different combinations of number of loci versus read length, while holding total sequence length constant); (b) lag time between fusion and gene pool sampling (measured as number of generations, scaled by effective population size); and (c) historical context of a fusion event (whether or not pre-fusion populations experienced a bottleneck coinciding with their divergence). To provide insights into how reduction in dimensionality may be achieved, we identified relatively weakly correlated subsets of informative summary statistics (see Blum, Nunes, Prangle, & Sisson, 2013 for a review of dimensionality reduction methods for ABC). Ultimately, this work provides useful guidelines for empirical phylogeographic and comparative phylogeographic studies.

2 | MATERIALS AND METHODS

2.1 | Overview of simulations

Garrick et al. (2019) used simulations to assess detectability of lineage fusion within an ABC framework, under simplifying assumptions of drift-induced divergence followed by instantaneous merging. That study compared fusion scenarios with different durations of drift-induced divergence among parental populations against long-term

panmixia or population expansion-contraction to understand the circumstances under which a history of fusion can be correctly inferred. Our goal here was to characterize the signature of lineage fusion in a suite of summary statistics, and to identify the most informative statistics for use within ABC or other frameworks. Simulations focused on one scenario that reflected the timing of paleoclimatic events frequently cited in the phylogeographic literature (e.g., Emerson & Hewitt, 2005; Hewitt, 2004; Jesus et al., 2006), and successively sampled the same population gene pool at 10 intervals after fusion. In this model, divergence between sister lineages occurred during the Penultimate Glacial Maximum (PGM; t_{div} = 140,000 generations ago, assuming a 1-year generation time) and merging occurred soon after the end of the Last Glacial Maximum (LGM; t_{fuse} = 15,000 generations ago). Other fixed parameters included effective population size (N_e = 10,000 diploid individuals along all branches of the population tree), and mixing = 0.5 (each sister population contributed equally to fusion). The population gene pool was initially sampled at three non-uniform intervals in relatively quick succession soon after the merging event (i.e., 10, 100, and 1,000 generations following fusion), and subsequently, at seven uniform intervals of 2,000 generations through to the present day (t_0). For comparative purposes, a 'baseline' non-fusion model was constructed and sampled at the same time intervals (Figure 1).

To explore impacts of the historical context of a fusion event upon the responses of summary statistics, two pairs of lineage fusion versus baseline models were considered. The first ($Fusion_{bottleneck}$) represented a case where the amount of suitable habitat did not vary over time, such that the cumulative effective population size (N_e) of the species remained constant. In this case, PGM divergence was coupled with bottlenecks that reduced each sister population to half the size of the ancestor, followed by recovery to the original size upon fusion. The corresponding baseline scenario simply modelled a single population with constant N_e over time (Figure 1, left panels). The second scenario ($Fusion_{no\ bottleneck}$) emulated contraction and expansion of suitable habitat in concert with glacial-interglacial cycles. Here, N_e of each sister lineage was the same as the ancestor and remained constant over time, but because the number of populations temporarily doubled, there was a transient increase in cumulative N_e of the species (i.e., following PGM divergence, but prior to fusion). Thus, the corresponding baseline scenario modelled a single population that doubled in size at t_{div} and then returned to its original size at t_{fuse} (Figure 1, right panels).

2.2 | Genetic datasets

Within the constraints of each demographic scenario, nuclear autosomal DNA sequence haplotypes (i.e., loci with alleles comprising of a set of linked nucleotide polymorphisms) were simulated using the HKY model of nucleotide evolution (Hasegawa, Kishino, & Yano, 1985) with proportion of invariant sites = 10%, a gamma model for rate heterogeneity across sites (Yang, 1994) with two discrete categories, and a mutation rate (μ) of 1×10^{-7} substitutions per site per generation, assuming a 1-year generation time. In all cases, non-recombining independent loci with phase-known alleles were generated, and 10 diploid individuals were sampled at each of the 10 successive time intervals. To understand the influence of the arrangement of DNA sequence data, three different combinations of read length versus number of loci were examined, holding the total amount of DNA sequence data per multi-locus alignment constant: 200-bp alignments × 200 loci, 400-bp × 100 loci, and 4-kb × 10 loci (i.e., 80-kb of phase-known DNA sequence per diploid individual × 10 individuals in each case). In practice, such datasets are attainable using short-read next-generation sequencing of reduced representation libraries for shallow-scale phylogenomics (e.g., Andrews, Good, Miller, Luikart, & Hohenlohe, 2016; Harvey, Smith, Glenn, Faircloth, & Brumfield, 2016; Lemmon, Emme, & Lemmon, 2012; Rochette, Rivera-Colón, & Catchen, 2019). For each demographic scenario,

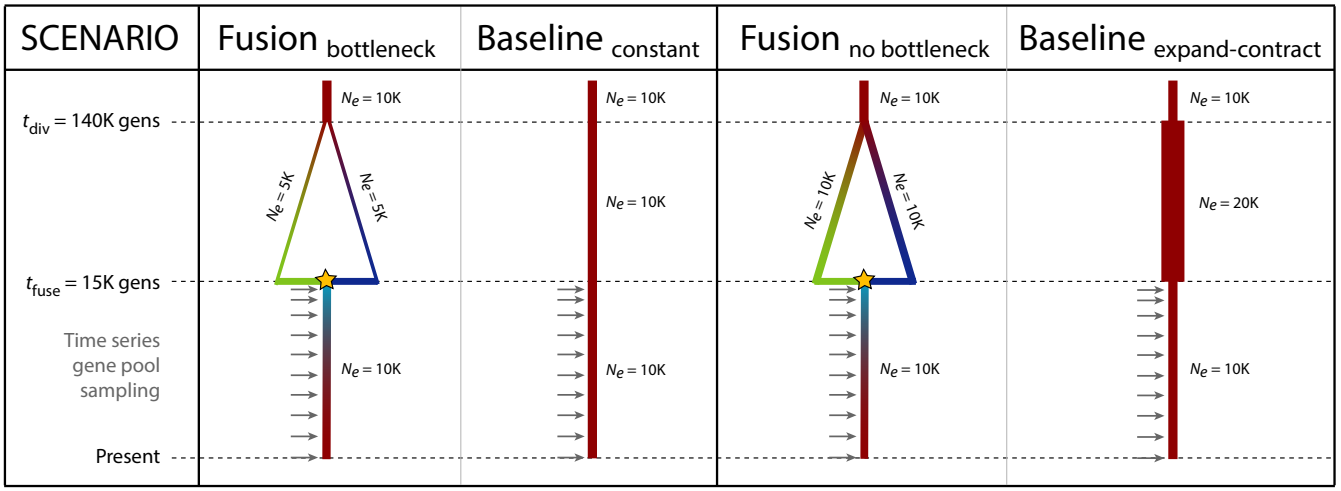

FIGURE 1 Demographic models used to simulate DNA sequence haplotype datasets. For the two fusion scenarios, colour transitions indicate accumulated effects of genetic drift on isolated gene pools derived from a common ancestral population, beginning with the initial divergence of two sister lineages, followed by their complete fusion (yellow stars). Grey arrows mark 10 successive time intervals (not to scale) of gene pool resampling. Population size is indicated by N_e values annotated on each model

TABLE 1 Summary statistics evaluated for their ability to detect signatures of past lineage fusion from DNA sequence haplotype datasets. Given that multi-locus datasets were simulated, point estimates were averaged across loci

General class	Notation	Summary statistic name	Reference	Basic description
Neutrality test	D	Tajima's D	Tajima (1989)	Comparison of the mean number of pairwise nucleotide differences (K) versus the number of segregating sites (S), per locus
	D^*	Fu and Li's D^*	Fu and Li (1993)	Comparison of the number of singletons (η_S) versus the number of mutations (i.e., accounting for sites segregating for >2 nucleotides; η), per locus
	F^*	Fu and Li's F^*	Fu and Li (1993)	Comparison of the number of singletons (η_S) versus the mean number of pairwise nucleotide differences (K), per locus
	F_S	Fu's F_S	Fu (1997)	Comparison of the observed number of haplotypes (H_N) versus the expected number of haplotypes given the estimated value of θ_π, per locus
	R_2	Ramos-Onsins and Rozas' R_2	Ramos-Onsins and Rozas (2002)	Comparison of the number of singletons (η_S) versus the mean number of pairwise nucleotide differences (K), divided by the number of segregating sites (S), per locus
	Y^*	Achaz's Y^*	Achaz (2008)	Comparison of the mean number of pairwise nucleotide differences ignoring singletons ($K_{-\eta_S}$) versus the number of segregating sites ignoring singletons ($S_{-\eta_S}$), per locus
	Z_{nS}	Kelly's Z_{nS}	Kelly (1997)	A linkage disequilibrium-based measure of allele frequency equivalency among pairs of segregating sites (S), within non-recombining DNA regions
Diversity	θ_W	Watterson's estimate of theta (per nucleotide)	Watterson (1975)	The observed S in sample size n, divided by the expected S given a neutral model of evolution calculated per nucleotide
	θ_{Kvar}	Variance of Tajima's theta (per locus)	Tajima (1983)	Variance of the mean number of pairwise nucleotide differences (K)
	π	Nucleotide diversity	Nei (1987)	Mean per-site number of nucleotide differences between pairs of randomly selected haplotypes, per locus
	S	Segregating sites	Watterson (1975)	Number of polymorphic sites, excluding insertion-deletion mutations, per locus
	Hd	Haplotype diversity	Nei (1987)	Mean proportion of pairs of randomly selected haplotypes that are different from one another, per locus
	Hd_{var}	Variance of Hd	Nei (1987)	Sampling variance of haplotype diversity (Hd) across loci
	$H(p)$	Number of heterozygous sites	Kimura (1969)	Mean number of observed segregating sites (S) that are heterozygous, per diploid individual, per locus
	H_N	Number of haplotypes	Nei (1987)	Number of distinct haplotypes per locus
	MNS	Mean numbers of rarest nucleotide at S	Cornuet et al. (2015)	Mean of the number of the rarest nucleotide at segregating sites, averaged across loci
	VNS	Variance of MNS	Cornuet et al. (2015)	Variance of the mean of the numbers of the rarest nucleotide at segregating site (MNS), averaged across loci

10 replicate datasets were simulated, giving a total of 1,200 single population gene pool samples (i.e., 2 fusion scenarios + 2 baseline scenarios × 3 dataset types × 10 time points × 10 replicates). All simulations were performed in DIY-ABC v.2.1.0 (Cornuet et al., 2014), which implements a backward-in-time coalescent process.

2.3 | Summary statistics

For DNA sequence haplotype data, summary statistics can be broadly categorized as tests of neutrality (or population size change), or measures of nucleotide diversity. We calculated 17 statistics that represent both classes using either or DIY-ABC for three statistics (θ_{Kvar}, MNS and VNS) or DNASP v.6.12.03 (Rozas et al., 2017) for all others (Table 1). Although not exhaustive, the chosen statistics are useful for phylogeographic inference from next-generation sequencing datasets because they can be calculated quickly. Furthermore, these statistics should be applicable to haplotype data from non-model organisms as there is no requirement for an outgroup, or knowledge of physical distances among loci.

2.4 | Informativeness and signal intensity, and patterns of decay over time

To identify which summary statistics best captured the signatures of past lineage fusion, we first assessed whether there were significant differences in mean values (across 10 replicates) between lineage fusion versus its associated baseline scenario, using two-tailed *t*-tests. As these tests were repeated for 10 successive sampling events from the same post-fusion population gene pool, we considered it prudent to use a stringent significance threshold (herein, 'informative' summary statistics are those for which *t*-test $p \leq 0.001$). To efficiently summarize these outcomes for each of the three DNA sequence haplotype dataset types, we focused on two time points of gene pool sampling (out of the aforementioned set of 10): a 'short' and 'long' lag time after fusion (i.e., 0.5 and 1.5 N_e generations, respectively). This gave a total of six parameter combinations per fusion scenario (i.e., $Fusion_{bottleneck}$ and $Fusion_{no\ bottleneck}$). For each informative summary statistic, we quantified its initial signal intensity and subsequent decay. Thus, we calculated the proportional difference (PD) between the mean values of the summary statistic on fusion versus baseline scenarios at each of the nine most recent sampling time points, relative to the initial difference at the oldest time point, which was 10 generations after fusion and defined PD = 1.0 [i.e., $PD_{generation\ N} = (Fusion_{generation\ N} - Baseline_{generation\ N})/(Fusion_{generation\ 10} - Baseline_{generation\ 10})$]. Note that in principle, PD should be upper limited at one and approach zero as the lag time between fusion and gene pool sampling increases (but stochasticity may lead to some departures from these expectations). The PD metric was used to establish a rank-ordering of informative summary statistics for all six parameter combinations. We also constructed signal decay plots to explore how the arrangement of DNA sequence data, sampling lag time, and the historical context of lineage fusion, affected summary statistics.

2.5 | Independence

To determine which subsets of summary statistics are semi-independent, we calculated the Pearson correlation coefficient (*r*) between all pairs of informative summary statistics for a given lineage fusion scenario (corresponding baseline scenarios had much narrower ranges of values and were considered uninformative for this purpose). Correlation matrices were converted to dendrograms, and a threshold of $r \leq 0.85$ demarcated natural clusters. This cut-off is commonly used for an analogous redundancy reduction problem in the ecological niche modelling literature (e.g., Duan, Kong, Huang, Fan, & Wang, 2014 and references therein), so it was adopted here. From each cluster, PD scores were used to select the most informative summary statistic for inclusion in the final subset. As above, this was done for all six parameter combinations.

3 | RESULTS

3.1 | Genetic datasets

An overview of diversity within simulated DNA sequence haplotype datasets is provided in Table S1 (all raw sequence data and summary statistic values are available from Dryad). For datasets that comprised of 200-bp × 200 loci, the mean number of unique haplotypes (H_N) ranged from 3.16 to 5.16, consistent with values reported in some empirical studies (e.g., Baetscher, Clemento, Ng, Anderson, & Garza, 2018). H_N ranged from 4.57 to 7.29 for 400-bp × 100 loci, which approximates polymorphism in an intraspecific ultraconserved element dataset generated by Harvey et al. (2016). As expected, H_N continued to increase with longer read lengths, with values ranging from 12.90 to 15.88 for 4-kb × 10 loci. Overall, our simulated data seem representative of empirical datasets, such that conclusions should be of practical value.

3.2 | Informativeness and signal intensity, and patterns of decay over time

Across replicate datasets, summary statistic values appeared to be normally distributed (e.g., >99% of datasets had skewness values, which measure degree of asymmetry around the mean, between −2 and 2), supporting the utility of *t*-tests. Outcomes from *t*-tests showed that a larger number of summary statistics were informative ($p \leq 0.001$) when more loci were used (Table 2). For example, across both scenarios ($Fusion_{bottleneck}$ and $Fusion_{no\ bottleneck}$) and the two focal lag times ('short' and 'long'), an average of 16 summary statistics (range: 14–17) were informative for datasets consisting of 200-bp × 200 loci. However, this dropped to 14 (range: 12–15) for

TABLE 2 Rank-ordering of summary statistics, from highest (1) to lowest (17) sensitivity to past fusion, based on proportional difference scores. For each type of DNA sequence dataset and fusion scenario, two time points of gene pool sampling were examined: a 'short' and 'long' lag time (0.5 and 1.5 N_e generations, for diploid autosomal loci, respectively). Uninformative summary statistics are indicated by '–'. Colour coding highlights the top six summary statistics, where warm to cool colours correspond with high to low sensitivity

Dataset	200-bp × 200 loci				400-bp × 100 loci				4-kb × 10 loci			
Fusion scenario	$Fusion_{bottleneck}$		$Fusion_{no\ bottleneck}$		$Fusion_{bottleneck}$		$Fusion_{no\ bottleneck}$		$Fusion_{bottleneck}$		$Fusion_{no\ bottleneck}$	
Lag time (N_e gens)	Short (0.5)	Long (1.5)	Short (0.5)	Long (1.5)	Short (0.5)	Long (1.5)	Short (0.5)	Long (1.5)	Short (0.5)	Long (1.5)	Short (0.5)	Long (1.5)
D	12	12	12	12	11	11	11	11	11	–	–	–
*D**	9	8	10	8	7	8	9	7	6	–	8	–
*F**	10	10	11	11	9	9	10	9	9	–	–	–
F_S	7	6	6	6	6	6	6	3	8	6	6	–
R_2	14	13	14	14	10	10	12	10	10	–	–	–
*Y**	15	15	15	–	15	13	14	–	13	–	–	–
Z_{nS}	1	1	1	1	1	4	1	1	1	5	4	1
θ_W	3	3	3	3	3	2	3	5	3	4	2	3
θ_{Kvar}	5	5	5	5	5	5	5	2	5	1	5	4
π	8	7	8	7	8	7	7	8	7	7	7	6
S	2	2	2	2	2	1	2	4	2	3	1	2
Hd	11	11	9	10	13	–	8	–	–	–	–	–
Hd_{var}	–	–	16	–	14	–	–	–	–	–	–	–
H(*p*)	4	4	4	4	4	3	4	6	4	2	3	5
H_N	6	9	7	9	–	–	–	–	–	–	–	–
MNS	13	14	13	13	12	12	13	12	12	–	–	–
VNS	16	–	17	–	–	–	–	–	–	–	–	–

400-bp × 100 loci, and then down to nine (range: 6–13) when using 4-kb × 10 loci. For the latter genetic dataset type, the number of informative summary statistics was generally greater when the historical context of lineage fusion involved a fixed amount of habitat over time (i.e., $Fusion_{bottleneck}$) compared to when this fluctuated (i.e., $Fusion_{no\ bottleneck}$). Conversely, little or no influence of the historical context of lineage fusion upon the number of informative summary statistics was seen for 200-bp × 200 loci or 400-bp × 100 loci. Not surprisingly, more summary statistics were informative when there was a short lag between fusion and gene pool sampling (Table 2).

Informative summary statistics varied widely in their signal intensity (e.g., PD ranged from 0.202–0.990 for the 'long' lag time between fusion and gene pool sampling). Furthermore, the historical context of lineage fusion was important for some summary statistics. For instance, for the $Fusion_{no\ bottleneck}$ scenario, PD values > 0.95 were retained after 1.5 N_e generations by Z_{nS} for datasets consisting of 200-bp × 200 loci and 400-bp × 100 loci, yet for the $Fusion_{bottleneck}$ scenario, PD values were much lower (0.70 and 0.67, respectively). Similar but less pronounced impacts on Z_{nS} were seen when using 4-kb × 10 loci (i.e., PD = 0.81 for $Fusion_{no\ bottleneck}$ vs. 0.63 for $Fusion_{bottleneck}$). Considering all informative summary statistics and genetic dataset types together, there was a trend towards higher PD values for $Fusion_{no\ bottleneck}$ when there was a 'long' lag time between fusion and gene pool sampling, whereas the historical context of lineage fusion had little impact on signal retention when lag time was 'short'.

Across all parameter combinations, there was considerable variability in the rank-ordering of summary statistics (Table 2) based on their informativeness, and signal intensity (from *t*-tests and PD, respectively). To facilitate interpretation, we focused on the six most informative statistics. For datasets with 200-bp × 200 loci, outcomes were very consistent. Thus, the historical context of lineage fusion and sampling lag times had little impact on choice of the most informative summary statistics in this case. For 400-bp × 100 loci, outcomes matched one another (and those seen for the previous genetic dataset type) when sampling lag time was 'short', but some rank-orderings changed when lag time was 'long'. Although there were some lag time effects and fusion scenario effects, these were rather minor. Conversely, when using 4-kb × 10 loci, all rank-ordering outcomes were different, indicating stronger impacts of sampling lag time and the historical context of lineage fusion. Over all parameter combinations, the six most informative summary statistics included four diversity metrics [θ_W, θ_{Kvar}, *H*(*p*) and *S*] and two neutrality tests (F_S and Z_{nS}; Table 2; Figure 2).

Signal decay plots (Figure 2) indicated that arrangement of DNA sequence data had an impact on variance across replicates. Variance was lowest for datasets with 200-bp × 200 loci, and highest for

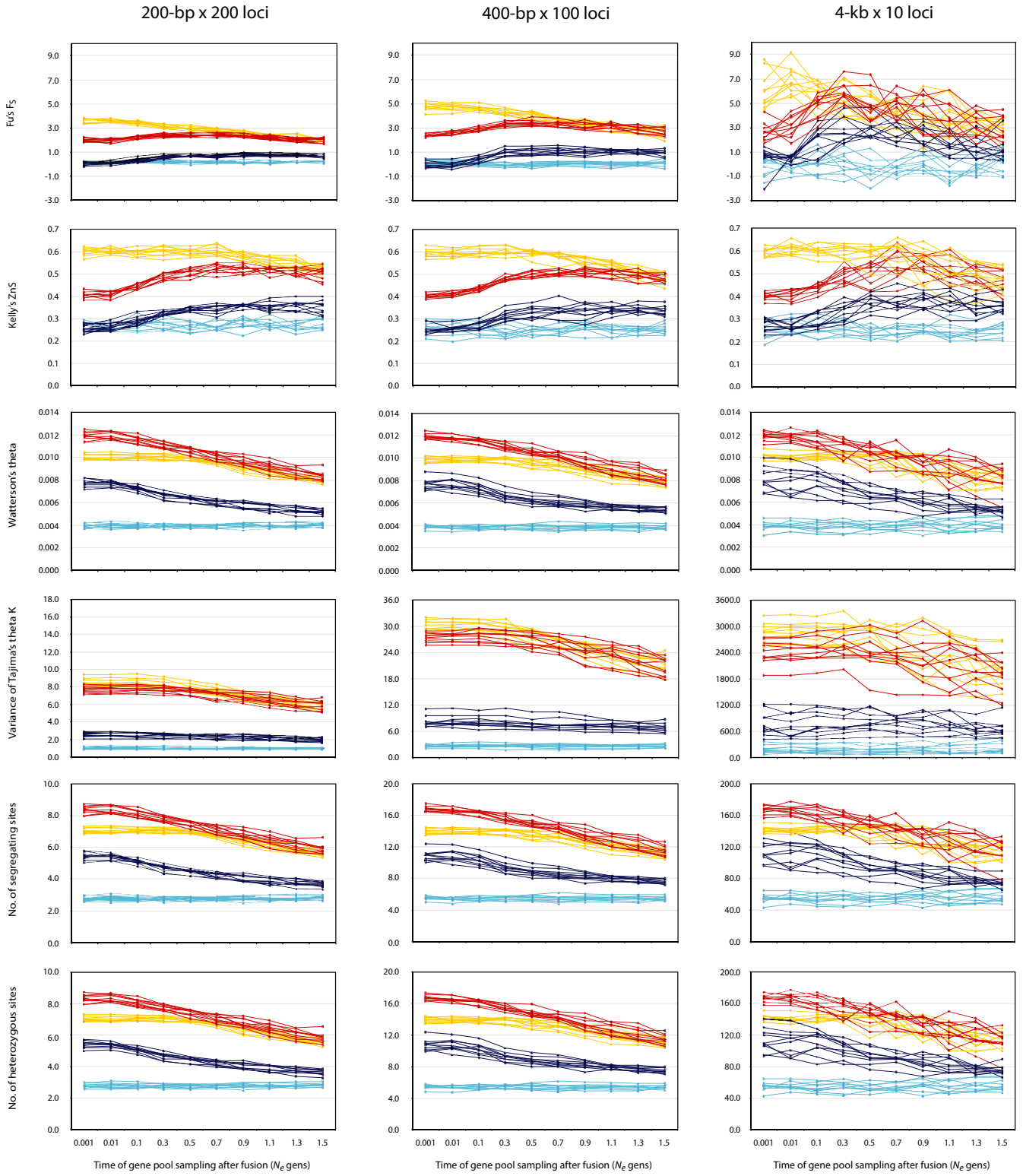

FIGURE 2 Signal decay plots showing change in values over time for six of the most sensitive summary statistics for detecting past lineage fusion (also see Table 1). The pairs of lineage fusion versus baseline models are coloured as follows: $Fusion_{bottleneck}$ (yellow) versus $Baseline_{constant}$ (pale blue) and $Fusion_{no\ bottleneck}$ (red) versus $Baseline_{expand\text{-}contract}$ (dark blue). The *x*-axis measures time as N_e generations for diploid autosomal loci

4-kb × 10 loci. When considering absolute values of those summary statistics expected to be independent of read length (i.e., per-site diversity metrics and neutrality tests), there was consistency across the three dataset types (except for F_S). Signal decay plots also showed that overall, most summary statistics were sensitive to bottlenecks during divergence. However, a few were mostly insensitive

(e.g., π and θ_{Kvar}). Many summary statistics showed a progressive decrease in their values following fusion, but interestingly, for the $Fusion_{no\ bottleneck}$ scenario, some statistics (e.g., Z_{nS}, and F_S) showed a temporary increase. However, the latter was matched by the associated expansion-contraction baseline scenario. Ultimately, all signal decay plots showed that, over time, signatures of fusion are increasingly difficult to detect.

3.3 | Independence

Across all six parameter combinations, the number of informative semi-independent summary statistics ranged between two and four, indicating high overall correlation among the 17 statistics (Table 3). There were no clear effects of the type of DNA sequence haplotype data, sampling lag time, or the historical context of lineage fusion. However, some statistics were repeatedly included in each final subset. Z_{nS} was the most predominant, followed by θ_{Kvar} and S. These three were the most broadly useful for capturing strong signal of lineage fusion while simultaneously reducing dimensionally. F_S was also a prominent inclusion when sampling lag times were 'long'.

4 | DISCUSSION

Here, we built on previous work that highlighted the importance of considering lineage fusion when making inferences about a species' demographic history or assessing concordance of species' responses to past environmental change (Garrick et al., 2019). Our goal was to characterize the intensity and decay of signatures of past fusion for a suite of summary statistics. Overall, the most informative summary statistics were two neutrality tests, Kelly's (1997) Z_{nS} and Fu's (1997) F_S, and four diversity metrics: Watterson's (1975) theta (θ_W), Tajima's (1983) variance of theta (θ_{Kvar}), Kimura's (1969) number of heterozygous sites [$H(p)$] and Watterson's (1975) number of segregating sites (S). Furthermore, these neutrality tests and two of the diversity metrics (i.e., θ_{Kvar}, and S) were semi-independent ($r < 0.85$). Thus, their inclusion in analytical frameworks for reconstructing demographic histories that may include lineage fusion is recommended. Notably, Z_{nS} was particularly powerful (Table 2; Figure 2), which may be because this statistic provides a unique linkage disequilibrium-based assessment of the pattern and extent of associations among mutations at different polymorphic sites (however, the impact of assuming no intralocus recombination warrants further investigation). Below,

TABLE 3 Subsets of semi-independent summary statistics that are informative about past lineage fusion, for different genetic dataset types, fusion scenarios and sampling lag times. When two or more informative statistics were highly correlated ($r > 0.85$), their proportional difference scores were used to select the best representative for inclusion (✓) (cf. omission, ✗). Uninformative summary statistics are indicated by '–'

Dataset	200-bp × 200 loci				400-bp × 100 loci				4-kb × 10 loci			
Fusion scenario	$Fusion_{bottleneck}$		$Fusion_{no\ bottleneck}$		$Fusion_{bottleneck}$		Fusion scenario		$Fusion_{bottleneck}$		$Fusion_{no\ bottleneck}$	
Lag time (N_e gens)	Short (0.5)	Long (1.5)	Short (0.5)	Lag time (N_e gens)	Short (0.5)	Long (1.5)	Short (0.5)	Lag time (N_e gens)	Short (0.5)	Long (1.5)	Short (0.5)	Lag time (N_e gens)
D	✗	✗	✗	✗	✗	✗	✓	✗	✗	–	–	–
D^*	✗	✗	✗	✗	✗	✗	✓	✓	✓	–	✗	–
F^*	✗	✗	✗	✗	✗	✗	✗	✗	✗	–	–	–
F_S	✗	✓	✗	✓	✗	✓	✗	✓	✗	✓	✓	–
R_2	✗	✗	✗	✗	✗	✗	✗	✗	✗	–	–	–
Y^*	✗	✗	✗	–	✗	✗	✗	–	✗	–	–	–
Z_{nS}	✓	✓	✓	✓	✓	✗	✓	✓	✓	✗	✗	✓
θ_W	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗
θ_{Kvar}	✗	✗	✗	✓	✓	✗	✓	✓	✓	✓	✓	✗
π	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗
S	✗	✗	✓	✓	✓	✓	✗	✗	✓	✗	✓	✓
Hd	✗	✗	✗	✗	✗	–	✗	–	–	–	–	–
Hd_{var}	–	–	✗	–	✗	–	–	–	–	–	–	–
$H(p)$	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗
H_N	✗	✗	✗	✗	–	–	–	–	–	–	–	–
MNS	✗	✗	✗	✗	✗	✗	✗	✗	✗	–	–	–
VNS	✓	–	✗	–	–	–	–	–	–	–	–	–

we consider outcomes along three axes of parameter space, and then comment on the importance of lineage fusion in evolution.

4.1 | The arrangement of DNA sequence data

Many simulation studies have considered how the number of loci included in a phylogeographic dataset impacts historical inferences and/or demographic parameter estimates (e.g., Carling & Brumfield, 2007; Huang, Takebayashi, Qi, & Hickerson, 2011; Jackson, Morales, Carstens, & O'Meara, 2017; Robinson, Bunnefeld, Hearn, Stone, & Hickerson, 2014). However, few have explicitly considered how the arrangement of DNA sequence polymorphisms within and among loci influences outcomes. That said, Felsenstein (2006) explored the balance between haplotype length versus number of loci, given a fixed amount of sequencing resources, when estimating theta from a single isolated population, and concluded it is better to add loci rather than extend read length. This is because although long sequences are superior in their ability to estimate a coalescent tree, any one of these trees is just a stochastic representation of population history (Felsenstein, 2006). While we focused only on comparison of estimated summary statistic values (cf. the accuracy *per se*), we also saw clear benefits of many short (200-bp × 200) loci over a handful of long (4-kb × 10) loci. These advantages included a larger number of informative summary statistics, as well as a higher consistency in their rank-ordering with respect to their signal intensity (Table 2). Thus, when polymorphisms are distributed across many short loci, summary statistics are less sensitive to the underlying historical context of lineage fusion and sampling lag time, which are not known *a priori* in most empirical studies. Accordingly, sequencing approaches that yield microhaplotype data (i.e., loci with two or more single nucleotide polymorphisms within 200-bp that do not show complete linkage, resulting in >2 alleles; Kidd et al., 2013) are valuable.

4.2 | Lag time between lineage fusion and gene pool sampling

As the lag time between fusion and gene pool sampling was extended, signal of the initial excess of genetic diversity diminished. However, how long the signal persists before it becomes indistinguishable from non-fusion scenarios depends on many factors, including the duration of isolation prior to fusion, N_e, mutation rate and mixing ratio, among others (Alcala, Jensen, Telenti, & Vuilleumier, 2016; Alcala & Vuilleumier, 2014; Garrick et al., 2019). Here, we examined signal decay based on a series of 10 sampling lag time intervals along only one axis of parameter space (i.e., with t_{div}, N_e, μ and mixing fixed). While it is encouraging that many summary statistics were still able to detect signatures of past fusion after 1.5 N_e generations (Table 2; Figure 2), our simulations modelled 12.5 N_e generations of drift-induced divergence between ancestral sister populations (i.e., a divergence vs. lag time ratio of ~8.3). Although the prevalence of lineage fusion in nature is unknown, its apparent rarity in the literature may be attributable to observation bias, since these events are ephemeral (Emerson & Faria, 2014; Garrick et al., 2014). Given that phylogeographic divergence scales with reproductive isolation (Singhal & Moritz, 2013), the most commonly occurring cases of past lineage fusion (i.e., collapse of young lineages) may also be most difficult to detect owing to their low divergence versus lag time ratio (Garrick et al., 2019). Recently, Kearns et al. (2018) provided compelling evidence for ancient lineage fusion in the Common Raven (*Corvus corax*), with initial divergence estimated at ~1.5 Mya and subsequent merging on secondary contact likely occurring at least 140–440 Kya, resulting in what is now a single group of randomly mating individuals with mosaic genomes. This suggests a long lag time between fusion and gene pool sampling is not insurmountable, provided that dense genomic and geographic sampling are conducted.

4.3 | Historical context of lineage fusion

Whether or not population size reductions occurred during the divergence of lineages that subsequently collapsed together can impact the detectability of past lineage fusion (Garrick et al., 2019). Indeed, many summary statistics that we examined were sensitive to bottlenecks (Table 2). In empirical studies, non-genetic data could inform priors about changes in the distribution and abundance of habitat over time, and thus, the likelihood of bottlenecks. For instance, comparisons of present-day versus hindcast (i.e., LGM or earlier) projections of ecological niche models provide insights into the number and locations of habitat refugia (Waltari et al., 2007) and enable inferences about distributional shifts (Hyseni & Garrick, 2019). We found that $\text{Fusion}_{\text{bottleneck}}$ scenarios were generally associated with a larger number of informative summary statistics when DNA haplotypes are long (Table 2). This not only indicates that the historical context of lineage fusion can be important, but also suggests that founder effects, habitat fragmentation and/or range contractions may facilitate detection, if DNA sequence read length is sufficient. Although few empirical studies have considered lineage fusion in its strict sense (i.e., complete merging of two populations resulting in a single randomly mating group), a few cases of 'incipient fusion' have been reported. For example, following a ~200 Kyr separation, two genetically distinct groups of *Chelonoidis becki* Galápagos giant tortoises are currently merging (Garrick et al., 2014). Although founder effects during their initial divergence are likely, genetic evidence for or against this was equivocal. However, bottlenecks were almost certainly coupled with divergence of Western Carpathian grey wolf (*Canis lupus*) populations that have subsequently fused (Hulva et al., 2018). More complex pre-fusion scenarios, such as contraction (or expansion) affecting just one parental lineage, warrant investigation.

4.4 | Importance of lineage fusion in evolution

The immediate outcome of lineage fusion is an abrupt increase in locally co-occurring allelic and genotypic diversity that may persist for many generations, providing opportunities for rapid evolution (Alcala & Vuilleumier, 2014; Kearns et al., 2018; Marques et al., 2019). Indeed, it has long been recognized that hybridization at or below the species-level can be a constructive process (e.g., Barton & Hewitt, 1989; Lewontin & Birch, 1966). However, long-standing questions remain unanswered, such as under which demographic and ecological circumstances would merging lead to loss of previously distinct evolutionary lineages, rather than formation of a stable hybrid zone (Rhymer & Simberloff, 1996)? Likewise, how common is lineage fusion, and in which taxonomic groups and landscape settings is it most prevalent (Garrick et al., 2019)? To address this, we need efficient summary statistics for detecting such events, and the present work is an early attempt. Follow-up work should explore the utility of gene tree-based summary statistics (e.g., metrics for shape, Fu & Li, 1993; Pybus & Harvey, 2000, or incongruence, Woodhams, Lockhart, & Holland, 2016), scenarios involving merging of non-sister populations (Garrick et al., 2014; Kearns et al., 2018) or repeated fission-fusion cycles (Alcala et al., 2016; Alcala & Vuilleumier, 2014), and assess the impact of simplifying assumptions (e.g., instantaneous fusion, recombination-free neutral independent loci, or equivalence of cumulative N_e between fusion and associated baseline scenarios). Notwithstanding remaining knowledge gaps, the capacity of phylogeographers to explicitly consider lineage fusion, owing to the field's transformation from data-limited to data-rich (Edwards, Shultz, & Campbell-Staton, 2015; Garrick et al., 2015), represents a meaningful advancement.

ACKNOWLEDGEMENTS

RCG was supported by start-up funds and resources from the University of Mississippi, and a National Science Foundation EPSCoR RII Track-4 research fellowship (award #1738817). We benefited from feedback from members of the Carstens lab at Ohio State University, Marlène Chiarello, Mark McCauley, and two anonymous reviewers.

DATA AVAILABILITY STATEMENT

All simulated DNA sequence haplotype datasets, and raw summary statistic values, are available via Dryad Repository entry https://doi.org/10.5061/dryad.7sqv9s4pq.

ORCID

Ryan C. Garrick https://orcid.org/0000-0002-4057-7061
Chaz Hyseni https://orcid.org/0000-0003-2567-8013

REFERENCES

Achaz, G. (2008). Testing for neutrality in samples with sequencing errors. *Genetics*, *179*, 1409–1424. https://doi.org/10.1534/genetics.107.082198

Alcala, N., Jensen, J. D., Telenti, A., & Vuilleumier, S. (2016). The genomic signature of population reconnection following isolation: From theory to HIV. *G3: Genes, Genomes, Genetics*, *6*, 107–120. https://doi.org/10.1534/g3.115.024208

Alcala, N., & Vuilleumier, S. (2014). Turnover and accumulation of genetic diversity across large time-scale cycles of isolation and connection of populations. *Proceedings of the Royal Society B: Biological Sciences*, *281*, 20141369.

Andrews, K. R., Good, J. M., Miller, M. R., Luikart, G., & Hohenlohe, P. A. (2016). Harnessing the power of RADseq for ecological and evolutionary genomics. *Nature Reviews Genetics*, *17*, 81–92. https://doi.org/10.1038/nrg.2015.28

Baetscher, D. S., Clemento, A. J., Ng, T. C., Anderson, E. C., & Garza, J. C. (2018). Microhaplotypes provide increased power from short-read DNA sequences for relationship inference. *Molecular Ecology Resources*, *18*, 296–305. https://doi.org/10.1111/1755-0998.12737

Barton, N. H., & Hewitt, G. M. (1989). Adaptation, speciation and hybrid zones. *Nature*, *341*, 497–503. https://doi.org/10.1038/341497a0

Beaumont, M. A., Zhang, W., & Balding, D. J. (2002). Approximate Bayesian computation in population genetics. *Genetics*, *162*, 2025–2035.

Bennett, K. D. (1990). Milankovitch cycles and their effects on species in ecological and evolutionary time. *Paleobiology*, *16*, 11–21. https://doi.org/10.1017/S0094837300009684

Bertorelle, G., Benazzo, A., & Mona, S. (2010). ABC as a flexible framework to estimate demography over space and time: Some cons, many pros. *Molecular Ecology*, *19*, 2609–2625. https://doi.org/10.1111/j.1365-294X.2010.04690.x

Blum, M. G. B., Nunes, M. A., Prangle, D., & Sisson, S. A. (2013). A comparative review of dimension reduction methods in approximate Bayesian computation. *Statistical Science*, *28*, 189–208. https://doi.org/10.1214/12-STS406

Burton, R. S., Pereira, R. J., & Barreto, F. S. (2013). Cytonuclear genomic interactions and hybrid breakdown. *Annual Review of Ecology, Evolution, and Systematics*, *44*, 281–302. https://doi.org/10.1146/annurev-ecolsys-110512-135758

Campbell, N. A., Reece, J. B., Urry, L. A., Cain, M. L., Wasserman, S. A., Minorsky, P. V., & Jackson, R. B. (2008). *Biology* (8th ed.). San Francisco, CA: Pearson Benjamin Cummings.

Carling, M. D., & Brumfield, R. T. (2007). Gene sampling strategies for multilocus population estimates of genetic diversity (θ). *PLoS One*, *1*, e160. https://doi.org/10.1371/journal.pone.0000160

Carstens, B. C., & Knowles, L. L. (2007). Shifting distributions and speciation: Species divergence during rapid climate change. *Molecular Ecology*, *16*, 619–627. https://doi.org/10.1111/j.1365-294X.2006.03167.x

Cornuet, J.-M., Pudlo, P., Veyssier, J., Dehne-Garcia, A., Gautier, M., Leblois, R., … Estoup, A. (2014). DIYABC v2.0: A software to make approximate Bayesian computation inferences about population history using single nucleotide polymorphism, DNA sequence and microsatellite data. *Bioinformatics*, *30*, 1187–1189. https://doi.org/10.1093/bioinformatics/btt763

Cornuet, J.-M., Pudlo, P., Veyssier, J., Loire, E., Santos, F., Dehne-Garcia, A., & Estoup, A. (2015). *DIYABC version 2.1: A user-friendly software for inferring population history through Approximate Bayesian Computations using microsatellite, DNA sequence and SNP data*. Montferrier-sur-Lez cedex, France: Centre de Biologie et de Gestion des Populations, Institut National de la Recherche Agronomique. Retrieved from http://www1.montpellier.inra.fr/CBGP/diyabc/

Dorsey, B. L., Gregory, T. J., Sass, C., & Specht, C. D. (2018). Pleistocene diversification in an ancient lineage: A role for glacial cycles in the evolutionary history of *Dioon* Lindl. (Zamiaceae). *American Journal of Botany*, *105*, 1512–1530.

Duan, R.-Y., Kong, X.-Q., Huang, M.-Y., Fan, W.-Y., & Wang, Z.-G. (2014). The predictive performance and stability of six species distribution

models. *PLoS One*, *9*, e112764. https://doi.org/10.1371/journal.pone.0112764

Dynesius, M., & Jansson, R. (2014). Persistence of within-species lineages: A neglected control of speciation rates. *Evolution*, *68*, 923–934. https://doi.org/10.1111/evo.12316

Edwards, S. V., Shultz, A., & Campbell-Staton, S. C. (2015). Next-generation sequencing and the expanding domain of phylogeography. *Folia Zoologica*, *64*, 187–206. https://doi.org/10.25225/fozo.v64.i3.a2.2015

Emerson, B. C., & Faria, C. M. A. (2014). Fission and fusion in island taxa – Serendipity, or something to be expected? *Molecular Ecology*, *23*, 5132–5134. https://doi.org/10.1111/mec.12951

Emerson, B. C., & Hewitt, G. M. (2005). Phylogeography. *Current Biology*, *15*, R367–371. https://doi.org/10.1016/j.cub.2005.05.016

Felsenstein, J. (2006). Accuracy of coalescent likelihood estimates: Do we need more sites, more sequences, or more loci? *Molecular Biology and Evolution*, *23*, 691–700. https://doi.org/10.1093/molbev/msj079

Fu, Y.-X. (1997). Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics*, *147*, 915–925.

Fu, Y.-X., & Li, W.-H. (1993). Statistical tests of neutrality of mutations. *Genetics*, *133*, 693–709.

Garrick, R. C., Banusiewicz, J. D., Burgess, S., Hyseni, S., & Symula, R. E. (2019). Extending phylogeography to account for lineage fusion. *Journal of Biogeography*, *46*, 268–278. https://doi.org/10.1111/jbi.13503

Garrick, R. C., Benavides, E., Russello, M. A., Hyseni, C., Edwards, D. L., Gibbs, J. P., … Caccone, A. (2014). Lineage fusion in Galápagos giant tortoises. *Molecular Ecology*, *23*, 5276–5290. https://doi.org/10.1111/mec.12919

Garrick, R. C., Bonatelli, I. A. S., Hyseni, C., Morales, A., Pelletier, T. A., Perez, M. F., … Carstens, B. C. (2015). The evolution of phylogeographic data sets. *Molecular Ecology*, *24*, 1164–1171. https://doi.org/10.1111/mec.13108

Garrick, R. C., Nason, J. D., Fernández-Manjarrés, J. F., & Dyer, R. J. (2013). Ecological coassociations influence species' responses to past climatic change: An example from a Sonoran Desert bark beetle. *Molecular Ecology*, *22*, 3345–3361. https://doi.org/10.1111/mec.12318

Harvey, M. G., Smith, B. T., Glenn, T. C., Faircloth, B. C., & Brumfield, R. T. (2016). Sequence capture versus restriction site associated DNA sequencing for shallow systematics. *Systematic Biology*, *65*, 910–924. https://doi.org/10.1093/sysbio/syw036

Hasegawa, M., Kishino, K., & Yano, T. (1985). Dating the human-ape splitting by a molecular clock of mitochondrial DNA. *Journal of Molecular Evolution*, *22*, 160–174.

Hewitt, G. M. (2004). The structure of biodiversity – Insights from molecular phylogeography. *Frontiers in Zoology*, *1*, 4.

Hickerson, M. J., Dolman, G., & Moritz, C. (2006). Comparative phylogeographic summary statistics for testing simultaneous vicariance. *Molecular Ecology*, *15*, 209–223. https://doi.org/10.1111/j.1365-294X.2005.02718.x

Ho, S. Y. W., Saarma, U., Barnett, R., Haile, J., & Shapiro, B. (2008). The effect of inappropriate calibration: Three case studies in molecular ecology. *PLoS One*, *3*, e1615. https://doi.org/10.1371/journal.pone.0001615

Huang, W., Takebayashi, N., Qi, Y., & Hickerson, M. J. (2011). MTML-msBayes: Approximate Bayesian comparative phylogeographic inference from multiple taxa and multiple loci with rate heterogeneity. *BMC Bioinformatics*, *12*, 81. https://doi.org/10.1186/1471-2105-12-1

Hulva, P., Černá Bolfíková, B., Woznicová, V., Jindřichová, M., Benešová, M., Mysłajek, R. W., … Antal, V. (2018). Wolves at the crossroad: Fission–fusion range biogeography in the Western Carpathians and Central Europe. *Diversity and Distributions*, *24*, 179–192. https://doi.org/10.1111/ddi.12676

Hyseni, C., & Garrick, R. C. (2019). The role of glacial-interglacial climate change in shaping the genetic structure of eastern subterranean termites in the southern Appalachian Mountains, USA. *Ecology and Evolution*, *9*, 4621–4636. https://doi.org/10.1002/ece3.5065

Jackson, N. D., Morales, A. E., Carstens, B. C., & O'Meara, B. C. (2017). PHRAPL: Phylogeographic inference using approximate likelihoods. *Systematic Biology*, *66*, 1045–1053. https://doi.org/10.1093/sysbio/syx001

Jansson, R., & Dynesius, M. (2002). The fate of clades in a world of recurrent climatic change: Milankovitch oscillations and evolution. *Annual Review of Ecology and Systematics*, *33*, 741–777. https://doi.org/10.1146/annurev.ecolsys.33.010802.150520

Jesus, F. F., Wilkins, J. F., Solferini, V. N., & Wakeley, J. (2006). Expected coalescence times and segregating sites in a model of glacial cycles. *Genetics and Molecular Research*, *5*, 466–474.

Kearns, A. M., Restani, M., Szabo, I., Schrøder-Nielsen, A., Kim, J. A., Richardson, H. M., … Omland, K. E. (2018). Genomic evidence of speciation reversal in ravens. *Nature Communications*, *9*, 906. https://doi.org/10.1038/s41467-018-03294-w

Kelly, J. K. (1997). A test of neutrality based on interlocus associations. *Genetics*, *146*, 1197–1206.

Kidd, K. K., Pakstis, A. J., Speed, W. C., Lagace, R., Chang, J., Wootton, S., & Ihuegbu, N. (2013). Microhaplotype loci are a powerful new type of forensic marker. *Forensic Science International: Genetics Supplement Series*, *4*, e123–e124.

Kimura, M. (1969). The number of heterozygous nucleotide sites maintained in a finite population due to steady flux of mutations. *Genetics*, *61*, 893–903.

Lemmon, A. R., Emme, S. A., & Lemmon, E. M. (2012). Anchored hybrid enrichment for massively high-throughput phylogenomics. *Systematic Biology*, *61*, 727–744. https://doi.org/10.1093/sysbio/sys049

Lewis, H., & Bloom, W. L. (1972). The loss of a species through breakdown of a chromosomal barrier. *Symposia Biologica Hungarica*, *12*, 61–64.

Lewontin, R. C., & Birch, L. C. (1966). Hybridization as a source of variation for adaptation to new environments. *Evolution*, *20*, 315–336. https://doi.org/10.1111/j.1558-5646.1966.tb03369.x

Marques, D. A., Meier, J. I., & Seehausen, O. (2019). A combinatorial view on speciation and adaptive radiation. *Trends in Ecology and Evolution*, *34*, 531–544. https://doi.org/10.1016/j.tree.2019.02.008

Nei, M. (1987). *Molecular evolutionary genetics*. New York, NY: Columbia University Press.

Orr, H. A., & Presgraves, D. C. (2000). Speciation by postzygotic isolation: Forces, genes and molecules. *BioEssays*, *22*, 1085–1094. https://doi.org/10.1002/1521-1878(200012)22:12<1085:AID-BIES6>3.0.CO;2-G

Pybus, O. G., & Harvey, P. H. (2000). Testing macro-evolutionary models using incomplete molecular phylogenies. *Proceedings of the Royal Society B: Biological Sciences*, *267*, 2267–2272. https://doi.org/10.1098/rspb.2000.1278

Ramos-Onsins, S. E., & Rozas, J. (2002). Statistical properties of new neutrality tests against population growth. *Molecular Biology and Evolution*, *19*, 2092–2100. https://doi.org/10.1093/oxfordjournals.molbev.a004034

Rhymer, J. M., & Simberloff, D. (1996). Extinction by hybridization and introgression. *Annual Review of Ecology and Systematics*, *27*, 83–109. https://doi.org/10.1146/annurev.ecolsys.27.1.83

Robinson, J. D., Bunnefeld, L., Hearn, J., Stone, G. N., & Hickerson, M. J. (2014). ABC inference of multi-population divergence with admixture from unphased population genomic data. *Molecular Ecology*, *23*, 4458–4471. https://doi.org/10.1111/mec.12881

Rochette, N. C., Rivera-Colón, A. G., & Catchen, J. M. (2019). Stacks 2: Analytical methods for paired-end sequencing improve RADseq-based population genomics. *Molecular Ecology*, *28*, 4737–4754. https://doi.org/10.1111/mec.15253

Rosenblum, E. B., Sarver, B. A. J., Brown, J. W., Des Roches, S., Hardwick, K. M., Hether, T. D., … Harmon, L. J. (2012). Goldilocks meets Santa Rosalia: An ephemeral speciation model explains patterns of diversification across time scales. *Evolutionary Biology*, *39*, 255–261. https://doi.org/10.1007/s11692-012-9171-x

Rozas, J., Ferrer-Mata, A., Sánchez-DelBarrio, J. C., Guirao-Rico, S., Librado, P., Ramos-Onsins, S. E., & Sánchez-Gracia, A. (2017). DnaSP 6: DNA sequence polymorphism analysis of large data sets. *Molecular Biology and Evolution*, *34*, 3299–3302. https://doi.org/10.1093/molbev/msx248

Satler, J. D., & Carstens, B. C. (2017). Do ecological communities disperse across biogeographic barriers as a unit? *Molecular Ecology*, *26*, 3533–3545. https://doi.org/10.1111/mec.14137

Savage, H. M. (1983). The shape of evolution: Systematic tree topology. *Biological Journal of the Linnean Society*, *20*, 225–244. https://doi.org/10.1111/j.1095-8312.1983.tb01874.x

Singhal, S., & Moritz, C. (2013). Reproductive isolation between phylogeographic lineages scales with divergence. *Proceedings of the Royal Society B: Biological Sciences*, *280*, 20132246.

Smith, M. L., Ruffley, M., Espíndola, A., Tank, D. C., Sullivan, J., & Carstens, B. C. (2017). Demographic model selection using random forests and the site frequency spectrum. *Molecular Ecology*, *26*, 4562–4573. https://doi.org/10.1111/mec.14223

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

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

Waltari, E., Hijmans, R. J., Peterson, A. T., Nyari, A. S., Perkins, S. L., & Guralnick, R. P. (2007). Locating Pleistocene refugia: Comparing phylogeographic and ecological niche model predictions. *PLoS One*, *2*, e563. https://doi.org/10.1371/journal.pone.0000563

Watterson, G. A. (1975). On the number of segregating sites in genetical models without recombination. *Theoretical Population Biology*, *7*, 256–276. https://doi.org/10.1016/0040-5809(75)90020-9

Webb, W. C., Marzluff, J. M., & Omland, K. E. (2011). Random interbreeding between cryptic lineages of the Common Raven: Evidence for speciation in reverse. *Molecular Ecology*, *20*, 2390–2402. https://doi.org/10.1111/j.1365-294X.2011.05095.x

Woodhams, M. D., Lockhart, P. J., & Holland, B. R. (2016). Simulating and summarizing sources of gene tree incongruence. *Genome Biology and Evolution*, *8*, 1299–1315. https://doi.org/10.1093/gbe/evw065

Yang, Z. (1994). Maximum likelihood phylogenetic estimation from DNA sequences with variable rates over sites: Approximate methods. *Journal of Molecular Evolution*, *39*, 306–314. https://doi.org/10.1007/BF00160154

BIOSKETCH

Ryan C. Garrick's lab focuses on understanding processes that generate and maintain biodiversity within and among species, with an emphasis on montane forest biota.

Author contributions: R.C.G. conceived the study, simulated and analyzed data; R.C.G., C.H. and I.D.A developed the ideas and contributed to drafting and revising the manuscript.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

How to cite this article: Garrick RC, Hyseni C, Arantes ÍC. Efficient summary statistics for detecting lineage fusion from phylogeographic datasets. *J Biogeogr.* 2020;00:1–12. https://doi.org/10.1111/jbi.13932