Received: 19 July 2023 | Revised: 25 October 2023 | Accepted: 16 November 2023

DOI: 10.1111/mec.17225

ORIGINAL ARTICLE

MOLECULAR ECOLOGY WILEY

The effect of divergent and parallel selection on the genomic landscape of divergence

Hisham A. A. Ali[1] | Tim Coulson[1] | Sonya M. Clegg[1] | Claudio S. Quilodrán[1,2]

[1]Department of Biology, Edward Grey Institute of Field Ornithology, University of Oxford, Oxford, UK

[2]Department of Genetics and Evolution, University of Geneva, Geneva, Switzerland

Correspondence
Hisham A. A. Ali, Department of Biology, University of Oxford, Oxford OX1 3SZ, UK.
Email: hisham.ali@biology.ox.ac.uk

Funding information
Schweizerischer Nationalfonds zur Förderung der Wissenschaftlichen Forschung, Grant/Award Number: P5R5PB_203169

Handling Editor: Jeremy B Yoder

Abstract

While the role of selection in divergence along the speciation continuum is theoretically well understood, defining specific signatures of selection in the genomic landscape of divergence is empirically challenging. Modelling approaches can provide insight into the potential role of selection on the emergence of a heterogenous genomic landscape of divergence. Here, we extend and apply an individual-based approach that simulates the phenotypic and genotypic distributions of two populations under a variety of selection regimes, genotype–phenotype maps, modes of migration, and genotype-environment interactions. We show that genomic islands of high differentiation and genomic valleys of similarity may respectively form under divergent and parallel selection between populations. For both types of between-population selection, negative and positive frequency-dependent selection within populations generated genomic islands of higher magnitude and genomic valleys of similarity, respectively. Divergence rates decreased under strong dominance with divergent selection, as well as in models including genotype-environment interactions under parallel selection. For both divergent and parallel selection models, divergence rate was higher under an intermittent migration regime between populations, in contrast to a constant level of migration across generations, despite an equal number of total migrants. We highlight that interpreting a particular evolutionary history from an observed genomic pattern must be done cautiously, as similar patterns may be obtained from different combinations of evolutionary processes. Modelling approaches such as ours provide an opportunity to narrow the potential routes that generate the genomic patterns of specific evolutionary histories.

KEYWORDS
frequency-dependent selection, genomic landscape, genotype–phenotype map, individual-based model, natural selection, speciation

Sonya M. Clegg and Claudio S. Quilodrán are joint senior authors.

1 | INTRODUCTION

New species can form in a variety of ways (Coyne & Orr, 2004), but gradual (rather than abrupt) divergence between lineages, culminating in reproductively isolated forms described by the speciation continuum is thought to be frequent (Nosil et al., 2009). Multiple microevolutionary processes (i.e., drift, mutation, migration and different types of selection) can interact to influence progression along this continuum. While natural selection is considered a primary driver of speciation (Schluter, 2009), reconstructing the roles of, and interactions among, microevolutionary processes as divergence proceeds remains a key goal of evolutionary biology. Genomic signatures of these processes may be observed in the topography of the genomic landscape of divergence: the heterogeneous divergence profile along the genomes of two diverging forms (Gavrilets, 2014; Michel et al., 2010; Wolf & Ellegren, 2017). However, definitive conclusions are difficult because multiple combinations of processes may yield similar genomic profiles (Campbell et al., 2018; Quilodrán et al., 2020; Ravinet et al., 2017; Wolf & Ellegren, 2017). Previous approaches have provided valuable insights into how microevolutionary processes shape genomic landscapes of divergence (e.g., Andrello & Manel, 2015; Feder, Gejji, et al., 2012; Quilodrán et al., 2020; Sedghifar et al., 2016). However, the genomic signatures of combinations of evolutionary histories and processes, interacting with varying genomic architectures, require further investigation.

Two main features of genomic landscapes of divergence that have received particular attention are genomic islands and valleys (Feder, Egan, & Nosil, 2012; Michel et al., 2010; Wolf & Ellegren, 2017). Genomic islands describe regions of the genome with very high levels of divergence, and genomic valleys regions with low levels of divergence, relative to background levels (Ravinet et al., 2017; Roesti et al., 2014). Different forms of natural selection can result in formation of islands and valleys (Seehausen et al., 2014; Shafer & Wolf, 2013). During the early stages of the speciation continuum, genomic islands are typically formed under divergent selection, where different phenotypes, and hence different alleles, are favoured in each population (Feder et al., 2013; Seehausen et al., 2014). However, the formation of genomic islands is not restricted to the action of divergent selection, and they can form stochastically at neutral loci (Bay & Ruegg, 2017; Nosil & Feder, 2012; Quilodrán et al., 2020). A genomic valley may form under parallel selection, when the same phenotype is favoured in both populations, and that phenotype is produced by the same allele(s) (Roesti et al., 2014). However, genomic valleys may also form when similar deleterious mutations are removed in both populations (Cvijović et al., 2018). As the speciation continuum progresses, a heterogenous genomic landscape of divergence may thus emerge between populations influenced by parallel and divergent selection on different aspects of the phenotype (Ravinet et al., 2017; Seehausen et al., 2014).

Frequency-dependent selection (FDS) can occur concurrently with parallel and divergent selection during divergence. FDS describes how the frequency of a phenotype influences its fitness within a population, while parallel and divergent selection influence the phenotypic variation between them (Bolnick & Stutz, 2017). By influencing allele frequencies that underlie these phenotypes, FDS can also impact genetic variation between populations. FDS can be positive (PFDS), where the fitness of a phenotype increases as its frequency increases (Endler, 1988; Thompson, 1984), or negative (NFDS) where the fitness of a phenotype increases as its frequency decreases (Clarke & O'Donald, 1962). Both forms have been empirically documented. For instance, PFDS has been observed for hymenopteran black and yellow warning colouration to signal prey unpalatability (Endler, 1988), as well as phlox flower colouration to attract pollinators (Smithson, 2001). NFDS has been observed for predation-driven apostatic selection in the grove snail (*Cepaea nemoralis*) (Allen, 1988); and reproductive mating strategies in the side blotched lizard (*Uta stansburiana*) (Sinervo & Lively, 1996), and ruff (*Philomachus pugnax*) (Küpper et al., 2015). While NFDS acts to maintain genetic diversity (Clarke, 1979; Doebeli & Ispolatov, 2010), PFDS reduces genetic diversity due to selection against rare phenotypes (Langham, 2004; Mallet & Barton, 1989). However, the potential impact of the interaction of these types of selection with other microevolutionary processes on the genomic landscape of divergence is not well understood (Brisson, 2018; Svensson et al., 2018).

Divergence patterns across the genomic landscape may also be influenced by migration between populations, the genotype–phenotype map and genotype-environment (G×E) interactions. Migration, the movement of alleles between populations, generally acts to hinder genetic divergence (Futuyma, 1987; Morjan & Rieseberg, 2004), though can promote divergence through adaptive introgression, as has been observed in some radiation events (Lamichhaney et al., 2015; Mallet, 2007; Mavárez et al., 2006). The nature of the genotype–phenotype map, and G×E interactions can shape the genomic landscape via effects on genetic variation. The purely genetic contribution to the phenotype can be partitioned into additive, dominance and epistatic variation (Falconer & Mackay, 1996). Where the genotype–phenotype map includes dominance, that is, alleles at a locus interact to produce a phenotype that differs to that expected from the additive effect of those alleles (Falconer & Mackay, 1996), genetic diversity can be maintained when the same phenotype is produced by heterozygous and homozygous dominant loci (Kojima, 1959). G×E interactions, which describe how the phenotype is jointly influenced by the genotype and the environment (Falconer & Mackay, 1996), can also maintain genetic diversity in heterogenous environments (Etges et al., 2007; Mitchell-Olds, 1995).

Several modelling approaches have been developed to simulate the genomic signatures that different microevolutionary processes may produce during divergence. Many of these models are limited to simulating either a single or a few bi-allelic loci, with selection often defined as a single parameter (e.g., Charlesworth et al., 1997; Feder, Gejji, et al., 2012; Feder & Nosil, 2009, 2010; Sedghifar et al., 2016). While considering selection as a single parameter provides valuable information, it represents a specific stage after a given divergence level is reached (Feder, Gejji, et al., 2012; Feder & Nosil, 2009, 2010). Coalescence-based models have proved useful, particularly when

studying neutral evolution (Lohse, 2017). Coalescence-based models also have the advantage of being able to incorporate large amounts of data, potentially simulating entire chromosomes for large populations, and include selection and migration (Haller et al., 2019; Haller & Messer, 2019; Kelleher et al., 2016). Yet, they are typically less informative for small population sizes affected by selection and non-random reproduction (Currat et al., 2015).

Individual-based models (IBMs), or agent-based models, provide population level insights from the collective behaviour of individuals (DeAngelis & Grimm, 2014; Grimm et al., 2006). The simulated individuals are defined by state variables (e.g., physiological traits or genetic attributes), with fixed population behaviours (e.g., average offspring number and migration rates). IBMs use a forward-in-time approach in contrast to backward-in-time coalescence-based models (Hoban et al., 2012), allowing the exploration of more complex demographic and genetic dynamics during divergence. In addition, due to the exponential increase in computational power and the wealth of genomic data gathered in recent decades, IBMs have successfully simulated the evolution of populations using large genomic fragments (Currat et al., 2015; Peng & Kimmel, 2005).

A recent application of the IBM approach introduced a flexible fitness function for simulating evolutionary dynamics (Quilodrán et al., 2020). This function allows the linking of evolutionary, ecological, and population demographic parameters in a form that is not possible when selection is summarized as a single parameter. In addition, the framework facilitates simulation of various scenarios of migration, neutral and non-neutral evolution, using non-uniformly distributed loci with varying recombination and mutation rates. We further developed the IBM of Quilodrán et al. (2020) to model the speciation continuum under both divergent and parallel selection. We introduced four modifications: (i) negative and positive frequency-dependent selection within populations; (ii) genetic dominance rather than a purely additive genotype–phenotype map; (iii) genotype-environment interactions; and (iv) a modified version of the migration parameter allowing a variable level of migration instead of fixed values across all generations. Our aim is to better understand the formation of heterogenous genomic landscapes of divergence, including the development of genomic islands and valleys, as well as to provide a more flexible framework for exploring the genomic patterns that may emerge after particular evolutionary histories.

2 | METHODS

2.1 | Description of the model

The simulations are based on the Genomic Landscape of Divergence Simulations ('*glads*') IBM developed by Quilodrán et al. (2020). The '*glads*' framework can be used to explore the role of different micro-evolutionary processes acting as populations diverge, and via parametrization with empirical data, can also be used to answer questions about specific systems (Sendell-Price et al., 2020) (Figure 1).

Three hierarchical levels are considered in this framework: genotype, phenotype and demographic rate. Individuals are assigned a sex and genetic identity. A genotype–phenotype map defines the expected phenotype from the simulated genotypes, which along with the environment, defines the realized phenotype of individuals. Through the implementation of a fitness function, fitness is computed for each realized phenotype.

Individuals are diploid and are genetically defined by a two-dimensional array that represents homologous chromosomes. Each position in the array represents a locus. With multiple individuals being defined in this manner, we can observe genomic variation at a population level at each locus. For an additive genotype–phenotype map, two integers at each locus influencing a given phenotype determine the locus's breeding values (b_v) (i.e., the homologous alleles in a diploid system), each of which can take the form of any of the simulated alleles, for example, alleles 1 to 20. An individual's phenotype (z) is computed as the sum of all the breeding values contributing to that phenotype (n_a), along with a stochastic environmental component (ε_{env}), with a defined variance (σ_{env}) and mean (0) (Equation 1).

$$z = \sum_{v=1}^{n_a} b_v + \varepsilon_{\text{env}}\left(0, \sigma_{\text{env}}\right) \tag{1}$$

The fitness (ω) of an individual is a modified gaussian function of the phenotype (Equation 2). The following population level parameters are introduced in the first part of Equation 2: maximum number of offspring (b_0); phenotypic optimum (b_1); and variance of the gaussian distribution of the phenotype (b_2). b_0 represents the maximum fitness of individuals in the population, while b_1 represents the value of z that corresponds with b_0. In the second part of Equation 2, a density-dependent component is introduced to prevent exponential growth (b_3) of the population size (N). In the third part of this equation, a stochastic demographic component (ε_{dem}) with a predefined variance (σ_{dem}) and mean (0) is introduced.

$$\omega = b_0 e^{-\frac{1}{2}\left(\frac{z-b_1 n_a}{b_2 n_a}\right)^2} - b_3 N + \varepsilon_{\text{dem}}\left(0, \sigma_{\text{dem}}\right) \tag{2}$$

The input of each simulation run is composed of different populations, containing multiple genetically and sexually defined individuals, in which population level parameters are applied. Divergent selection can be simulated with different phenotypic optima (b_1) between populations, while parallel selection can be simulated with equal b_1 values. The simulation progresses by iterating discrete steps, representing generations. In each generation, individuals mate and generate offspring, where the probability of a male siring offspring, and the number of offspring produced by a female depends on their respective fitness. The offspring inherit half of the genome from each parent, but differ from them by specific rules defining recombination and mutation rates. The recombination parameter gives the probability of a crossover occurring; and is implemented as either an average Poisson probability over all loci or as a fixed probability between adjacent loci that considers a recombination map (see Quilodrán et al., 2020). The mutation parameter is defined by a binomial distribution

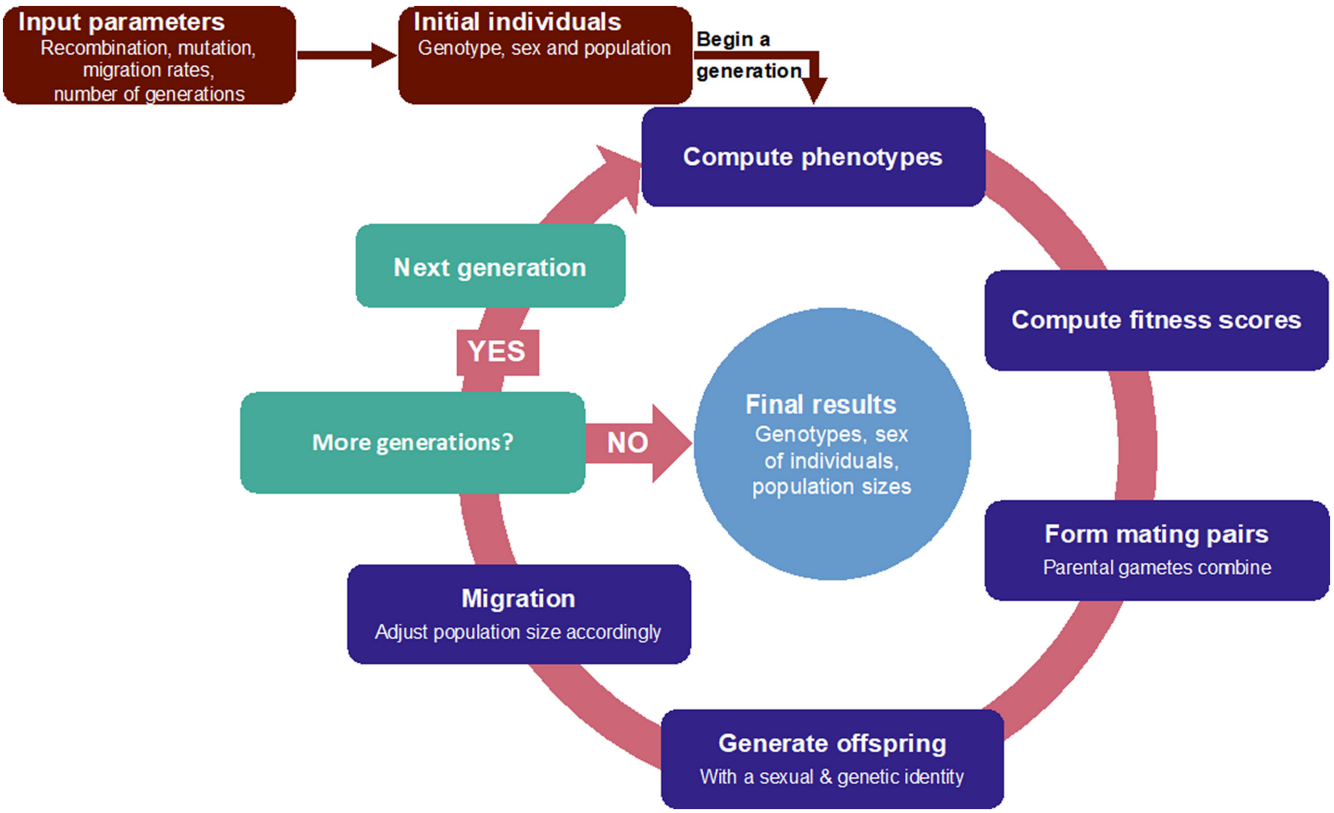

FIGURE 1 Summary of the '*glads*' framework showing each of the main steps in the modelling approach. The boxes in brown represent initialization of the model. The cycle represents a given time-step iteration, with the boxes indicating the different computing steps for each generation. The box 'More generations?' represents a condition variable stating either the running of another generation ('Next generation' box) or the end of simulations (sky-blue circle, representing the output from which the genomic landscape is computed). Adapted from Quilodrán et al. (2020).

computing the expected number of mutations per site, per generation. This parameter is restricted to biallelic genetic structures, and mutations at selected loci can affect the additive phenotype; consequently, shifting the phenotype towards or away from the phenotypic optimum (adaptive and maladaptive respectively). The offspring reach sexual maturity and have the potential to reproduce in the next generation. Finally, the migration rate parameter defines the probability of sexually mature individuals moving between populations. By applying these parameters over the course of iterative simulations, a genomic landscape of divergence can be observed between populations. By changing the parameters, the effects of different evolutionary mechanisms, such as selection and drift, can be ascertained. This model is implemented in *R* using the '*glads*' package (Quilodrán et al., 2020).

2.2 | Frequency-dependant selection

The first modification to the model fitness function (Equations 1 and 2) is a framework for frequency-dependant selection (FDS). FDS was applied to the phenotype derived in Equation 1. The range of phenotypes (z) for each population were partitioned to a pre-defined number of bins of equal size (K). Individuals (i) were subsequently assigned to these bins, based on their phenotypic trait value (Equation 3).

$$zK_i = \left\lfloor \frac{z_i - z_{\min}}{z_{\max} - z_{\min}} K \right\rfloor \quad (3)$$

FDS was applied to the frequency of individuals within each bin (fzK_i), which was normalized by the bin with the highest frequency ($fzK_{\max}$). This facilitates simulation of both positive and negative FDS. Under positive FDS, fitness increases as the frequency of individuals within the bin increases. Conversely, under negative FDS, fitness decreases as the bin frequency increases. The strength of FDS was implemented by the parameter (F) that denotes the maximum increase in offspring number due to FDS. The product of parameter F and the normalized bin frequencies ($\frac{fzK_i}{fzK_{\max}}$) represents the fitness change due to FDS (ω_i'), for an individual belonging to a given phenotype bin. In positive FDS, the product is added to the fitness, from Equation 2, for each individual (Equation 4).

$$\omega_i' = \omega_i + \left(\frac{fzK_i}{fzK_{\max}} F \right) \quad (4)$$

In negative FDS, an inverse relationship between fitness and the bin frequency is computed. This is calculated by subtracting

the product of F and the normalized bin frequency ($\frac{fzK_i}{fzK_{max}}$) from F, as shown in Equation 5.

$$\omega_i' = \omega_i + \left(F - \left(\frac{fzK_i}{fzK_{max}} F \right) \right) \quad (5)$$

2.3 | Genetic dominance

The second modification made to the model was the addition of dominance to the genotype–phenotype map across all loci. Our genotype–phenotype map considers k alleles per locus influencing a given phenotype, denoted in the IBM array of breeding values (b_v) as 1 to k. We provide an option to define the pattern of the dominance hierarchy in alleles. This is simulated by providing a list of numerical values of length k, with each value representing a different allele. These values indicate the position in the dominance hierarchy for the respective allele. Note an equivalent position, or co-dominance, between alleles can be achieved by assigning the alleles equal values in the hierarchy list. Co-dominance was used in the original application of '*glads*' (Quilodrán et al., 2020), and was an additive only model. We simulated three different strengths of dominance: strong, medium and an additive model (co-dominance). Under the additive model, each allele in a heterozygote contributed equally to the phenotype, and had an equal chance of being inherited. Therefore, under this model the hierarchy position for all k alleles are equivalent (Figure S1a). Under the other two models, the chance of inheritance for each allele in a heterozygote was also equal. However, under the strong dominance model, only a single 'dominant' allele would contribute to the phenotype, with a contribution the same as the homozygote for that dominant allele. In Equation 1, this corresponds to only a single allele contributing to the b_v for each locus (i.e., the allele with the highest value). In our simulation 20 alleles were used, and the hierarchy for their dominance were the values 1 to 20, ensuring a dominant allele would be present in any heterozygous combination (Figure S1c). The medium dominance model was the same, except that some heterozygote combinations influenced the phenotype additively instead. Therefore, for Equation 1 under this model, either a single or both alleles will contribute b_v for the locus. To achieve this medium strength dominance, our dominance hierarchy followed a bimodal pattern (Figure S1b).

The three dominance strengths we simulated are found in natural systems (Billiard et al., 2021). For example, in Brassicaceae and the distantly related *Ipomoea* (Convolvulaceae) and *Senecio* (Asteraceae), linear dominance hierarchies, analogous to our model of strong dominance, have been found for the SP11/SCR gene that controls self-incompatibility (Brennan et al., 2002; Fujii & Takayama, 2018; Kowyama et al., 1994). Non-linear hierarchies, analogous to our model of medium dominance, include lactose tolerance and sex specific dominance (Billiard et al., 2021; Fry, 2010; Tishkoff et al., 2007). Examples of co-dominance (additive system) include A and B blood type antigens; haemoglobin producing HBB gene, mutations of which cause beta thalassemia and sickle-cell disease; and gametophytic self-incompatibility in the S gene of a number of *Citrus* species (Kim et al., 2011; Xia, 2013).

2.4 | Genotype-environment interaction

Genotype-environment (G×E) interactions were introduced to the model's genotype–phenotype map function by modifying Equation 1. The first and second components of this linear equation represent the genetic and the environmental effects, respectively. A third component was included in this equation to compute the product of the genotype and the environmental effects, in which parameter (G) represents the strength of this interaction. This product of both components allows us to compute the realized phenotype of each individual by considering the interaction between the genotype and the environment, as shown in Equation 6.

$$z = \sum_{v=1}^{n_a} bv + \varepsilon_{env}(0, \sigma_{env}) + \left(G \times \sum_{v=1}^{n_a} bv \times \varepsilon_{env}(0, \sigma_{env}) \right) \quad (6)$$

2.5 | Migration

We introduced a parameter that allows the specification of changing migration rates across generation. We explored two scenarios of migration: (i) intermittent migration, in which the migration rate (m_{ij}) oscillates between 0 and 0.01 in incremental steps of 0.001; and (ii) constant migration rate (m_{ij}) through time, which represents the original approach included in '*glads*' (see Quilodrán et al., 2020) (see Figure S2). Note that we only consider migration as a homogenizer that hinders divergence, due to gene flow, and do not consider its potential role in promoting divergence through adaptive introgression.

2.6 | Initialization of the model

We parameterized the model following Quilodrán et al. (2020) to allow comparisons with the original implementation (Table 1). These simulations considered two diploid populations of 400 individuals, with genetic identities composed of 300 loci and 20 alleles per locus. A fraction of the simulated genome was assumed to influence the phenotype additively (i.e., an additive genotype–phenotype map composed of 50 loci). The remaining loci, that were not influencing the phenotype, were considered to be neutral. Because the original implementation focused on patterns at early stages of divergence, our models did not include the effect of new mutations. A recombination map was included, that defined the recombination probability between adjacent loci (ρ), to explore the effect of unlinked ($\rho = 0.5$) and strongly-linked loci ($\rho = 0.0001$). Henceforth, we refer to this original implementation as the base model.

Whilst our approach is a general model building on the parameters used by Quilodrán et al. (2020), our simulations are inspired by avian systems. Most landbird species (i.e., bird species that live primarily on or over land) produce two to three eggs per clutch, with relatively few producing 7 or more per clutch (Jetz et al., 2008). While our maximum fitness parameter (b_0) of six (Table 1) could be high for a landbird, it is representative of the range of probable values. The size of the FDS percentage corresponds with the increase

TABLE 1 List of parameters.

Parameter	Definition	Value
N	Population size	400
n_L	Number of simulated loci	300
n_a	Number of additive loci	50
b_0	Maximum offspring number	6
b_1	Phenotypic optima	Pop. 1 = 0.25; Pop. 2 = 0.75 (divergent selection models only)
b_2	Variance of fitness curve	0.5
b_3	Density dependence	Pop.1 = 0.01; Pop.2 = 0.005 (divergent selection models only)
σ_{env}	Stochastic environmental component	0.01
σ_{dem}	Stochastic demographic component	1
k	Number of alleles per locus (breeding values)	20
	Sex ratio	0.5
μ	Mutation rate	0
	Linked loci position	60:69, 150:159, 230:239
ρ	Linkage	0.0001
t	Number of generations	2000
m_{ij}	Gene flow (Migration rate)[a] Figure S2	0, 0.005999 (constant migration rate) 0 to 0.01 (intermittent migration rate)
F[b] K[b]	Frequency-dependent selection strength Number of bins	2% 10
G[c]	Genotype-environment interaction (G × E)	5, 10, 15, 20, 25

[a]Migration is only present for the migration specific models.

[b]F and K in Equations 3–5.

[c]Value of G in Equation 6.

in fitness of individuals. We therefore set the FDS percentage to 2% FD (Table 1) to mitigate exacerbating an already high maximum clutch size, but also at a percentage that appreciably influences the population. FDS acts upon a phenotype, and due to the additive model containing many loci, a large number of different phenotypes (z) could be generated in the population. We captured this variability by defining 10 phenotypic bins (K) (Table 1). The model is flexible and when modelling natural systems, bin number and strength can be parametrized accordingly.

For our simulations of G × E interaction, we used values from 5 to 25 for the strength parameter (G) (Table 1). This range of values was derived using the partitioning of zebra finch phenotypic variance from Woodgate et al. (2014). The genetic (CV_A), environmental (CV_{EC}) and G × E (CV_{GE}) coefficients of variance were used in the equation $G = CV_{GE}/CV_A * CV_{EC}$ to derive G for Equation 6. In the case of migration, we simulated two patterns: intermittent and constant migration. Under the intermittent model, migration levels varied in a bimodal pattern with a period of cessation of migration in between the peaks of migration (Figure S2). Under the constant migration model, a fixed migration rate $m_{ij} = 0.006$ was used, which was the average migrate rate of the intermittent model. This value was used to facilitate comparisons of the two models, by simulating an equivalent number of migrants at the end of each simulation. Finally, divergent selection was simulated by assigning the populations distinct phenotypic optima, while parallel selection was simulated by assigning the populations the same phenotypic optima (b_0, Table 1). Divergent and parallel selection were parametrized in the same way as implemented in Quilodrán et al. (2020).

2.7 | Evaluation of divergence

Genetic divergence between simulated populations was quantified with the fixation index (F_{ST}), the ratio of pairwise differences between the genotypes of two populations, relative to the variation within populations (Weir & Cockerham, 1984). F_{ST} ranges from 0 to 1, where a value of 1 indicates fixation of alternative alleles in each population. F_{ST} was computed every 100 generations across the total of 2000 simulated generations, using the '*pegas*' package

(Paradis, 2010). Simulations and analyses were carried out using R (version 3.61; R Core Team, 2019).

A hidden Markov model (HMM) was applied to the F_{ST} output to differentiate among effects of different types of selection included in our simulations (i.e., parallel or divergent selection with frequency-dependent selection). We simulated four different types of loci (i.e., selected-linked, selected-unlinked, neutral-linked, and neutral-unlinked), selecting an equal number of randomly chosen loci from each type in order to standardize the HMM. The HMM was used to identify three hidden states: (i) genomic islands; (ii) genomic valleys; and (iii) background levels of differentiation. The HMM analysis was performed following Marques et al. (2016) using the 'HiddenMarkov' R package (Harte, 2017). The Baum-Welch algorithm was applied using 1000 randomly chosen initial parameter values to find the parameter estimates with the highest likelihood. The Viterbi algorithm used these parameters to find the most likely sequence of states across the simulated loci. The significance of the assigned states was assessed by randomly permuting islands and valleys 10,000 times across the simulated loci (Quilodrán et al., 2020).

2.8 | Simulations

Our simulations were designed to elucidate how microevolutionary processes may influence the emergence of a heterogeneous genomic landscape of divergence. The parameters used for the simulations are listed in Table 1. The simulations explored how (1) frequency-dependent selection, (2) dominance, (3) G × E interaction, and (4) migration can influence the genomic landscape, alongside the effect of divergent selection, parallel selection, and drift (Table 2). While we analyse each effect separately, the concurrent effect of these processes may be simulated using the *'glads'* model. The code for the modified glads functions can be found in Appendix S1.

All simulations were run for a maximum of 2000 generations. The fixation index (F_{ST}) was measured every 100 generations to characterize the formation of the genomic landscape of divergence. We also measured the phenotypic space, measured as the range of a quantitative phenotype occupied by a population. When the effect of the microevolutionary process being investigated was not easily interpretable, the phenotypic space provided additional data on how a particular genomic landscape formed. Each simulation was repeated 100 times, with the individuals for each iteration varying in genetic identity at the beginning (generation 0). This was implemented by using a different random number generator in *R* for each simulation (set. seed function). These simulations were performed at the University of Geneva on the 'Yggdrasil' HPC cluster and on the Oxford ARC clusters.

3 | RESULTS

3.1 | Frequency-dependent selection

Our first set of simulations explored the effect of frequency-dependent selection on the genomic landscape of divergence between two populations that have evolved independently over 1000 generations (i.e., without migration). In the base model (a purely additive genotype–phenotype map without FDS), the genomic landscape under divergent selection shows the formation of a genomic island at selected loci (Figure 2a) while parallel selection does not produce any feature that can be distinguished from neutral background divergence (Figure 2b). Implementation of PFDS produces a genomic island under both divergent and parallel selection regimes, while NFDS produces a genomic valley under both regimes (Figure 2a,b). The genomic island formed under PFDS is not influenced by linkage to any great extent. The entire selected region approach fixation regardless of position of linked or unlinked loci (Figure 2a,b), with little variation in this outcome over all simulations (Figure 3a,b). This contrasts with both the baseline model and NFDS where the island or valley signature is amplified at linked loci (Figures 2a,b and 3a,b). In general, a high degree of variation in F_{ST} was seen across the 100 simulations for different combinations of selected/neutral and linked/unlinked loci (Figure 3a,b).

While islands or valleys were frequently observed on selected linked loci, significant islands and valleys (as identified by the hidden Markov model) were also observed elsewhere (Figure 4). For instance, the majority of valleys observed for both divergent and parallel selection under PFDS occurred at neutral loci (both linked and unlinked), and the majority of islands formed under NFDS also occurred at neutral loci (Figure 4). Under parallel selection in the baseline additive model, the number of islands and valleys for different loci types (selected/neutral and linked/unlinked loci combinations)

TABLE 2 Scenarios simulated using the glads model.

Selection		
G-P map[a]	Divergent selection	Parallel selection
Additive	Positive frequency-dependent selection	Positive frequency-dependent selection
Additive	Negative frequency-dependent selection	Negative frequency-dependent selection
Non-additive	Dominance (medium & strong)	Dominance (medium & strong)
Additive	Genotype-environment interaction	Genotype-environment interaction
Additive	Gene flow (constant vs. intermittent)	Gene flow (constant vs. intermittent)

[a]Genotype–Phenotype map.

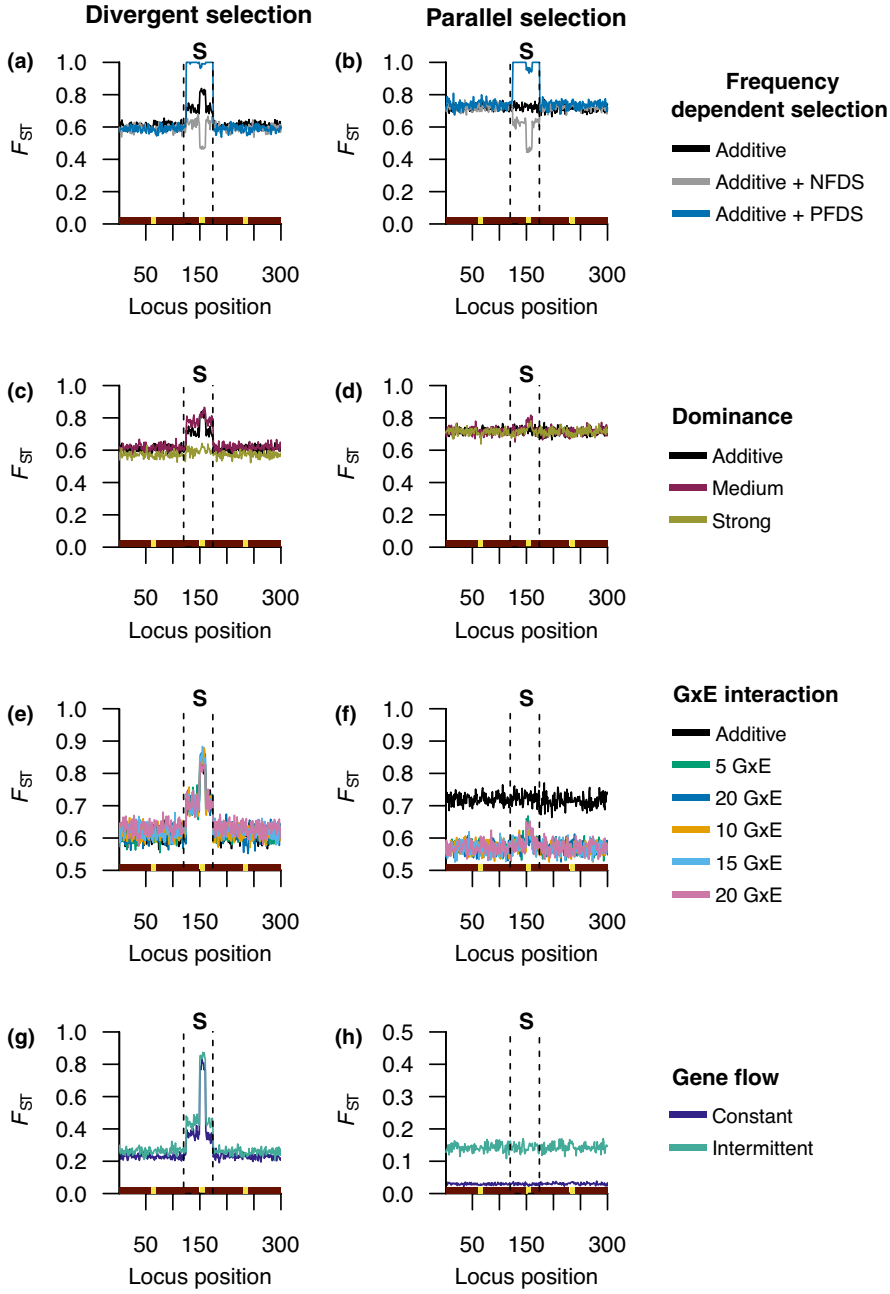

FIGURE 2 The genomic landscape of divergence between two populations. The output is summarized as the average F_{ST} from 100 simulations of 1000 generations each. Locus position represented on the *x*-axis is shaded red for unlinked loci and yellow for linked loci, with loci under selection (S) that contribute to phenotypic trait values indicated by region between the vertical dashed lines. The left-hand panels show divergent selection, and the right-hand panels, parallel selection. For each of these selection regimes, graphs show the addition of: (a, b) frequency-dependent selection; (c, d) varying dominance strength; (e, f) varying genotype-environmental interaction strength (*G* in Equation 6); and (g, h) varying modes of gene flow under an additive model. In (a–f) black lines denote divergence levels under divergent or parallel selection and a purely additive genotype–phenotype map, for comparison with modelled additions.

were similar (Figure 4d). However, under divergent selection, the majority of the islands and a smaller proportion of valleys were found at selected loci (Figure 4a). As expected, under PFDS for both types of selection, a higher number of islands on selected unlinked loci were observed compared to the base additive model (Figure 4c,f). The expected contrast was observed with NFDS, in which an excess of valleys and very few islands were observed on linked selected loci for both types of selection (Figure 4b,e).

A comparison of the genomic landscape of divergence at 1000 generations (above results) with that at 100 generations and 2000 generations showed the dynamics of genomic landscape features develop over time (Figure 5). At 100 generations, incipient island formation was already evident, and for the base model under divergent selection, and PFDS under both divergent and parallel selection, these islands persisted at 2000 generations, albeit at a smaller amplitude. Valley formation occurred later in the divergence process when NFDS was applied, with an island rather than a valley evident in early-stage divergence (100 generations) (Figure 5). The depth of valleys increased between 1000 and 2000 generations as surrounding neutral divergence increased.

The distribution of phenotypes realized by individuals at a range of generation time points under divergent and parallel selection, and with and without FDS is shown in Figure 6. At early stages of divergence (<100 generations), the distribution of phenotypes was similar for all models. The implementation of divergent selection with the base model, or with PFDS, resulted in two disjunct phenotypic groups by 500 generations, each occupying a relatively narrow range of values. Under divergent selection with NFDS, the two populations overlapped in phenotypic space after 2000 generations despite the divergence of the phenotypic space of the populations. Under parallel

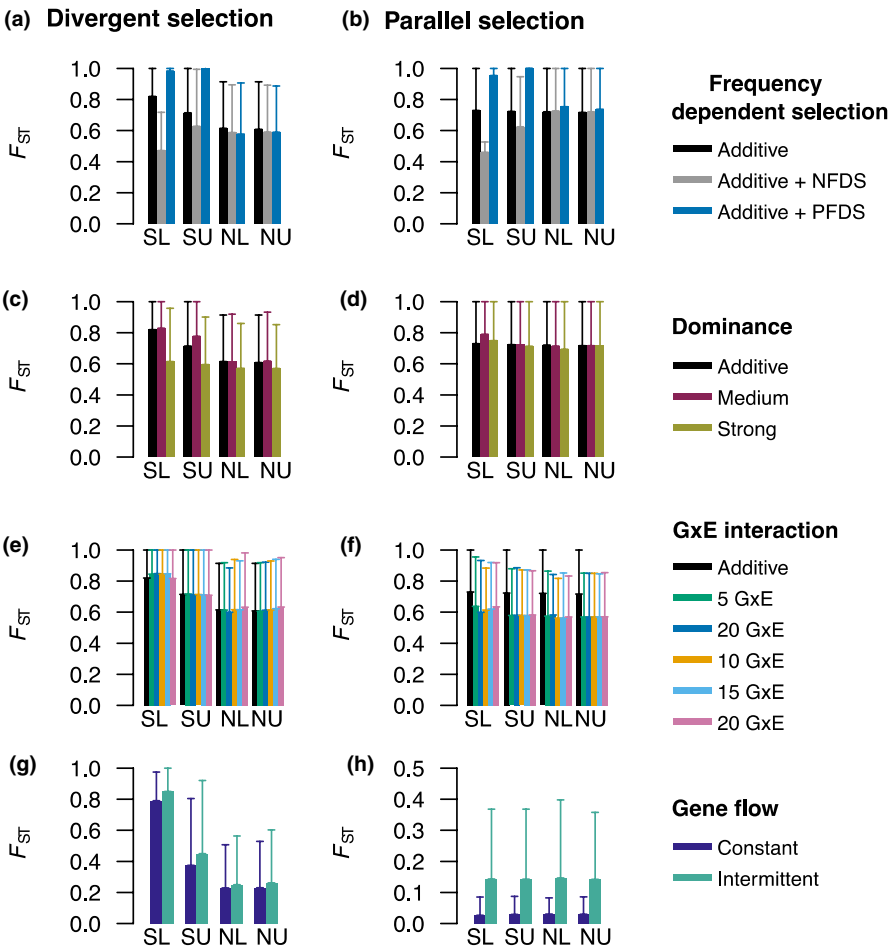

FIGURE 3 Mean F_{ST} for different loci types. Mean F_{ST} is given for selected & linked (SL), selected & unlinked (SU), neutral & linked (NL) and neutral & unlinked (NU) loci calculated from 100 simulations of 1000 generations, with error bars denoting the 95% quantile. The left-hand panels show divergent selection, and the right-hand panels, parallel selection. For each of these selection regimes, graphs show the addition of: (a, b) frequency-dependent selection; (c, d) varying dominance strength; (e, f) varying genotype-environmental interaction strength (G in Equation 6); and (g, h) gene flow. In (a–f) the black bars denote divergence levels under divergent or parallel selection and a purely additive genotype–phenotype map, as a basis for comparison with modelled additions. The bottom panels include different types of gene flow to the same additive model.

selection, phenotypes in the two populations did not diverge, but the range of phenotypes is narrowed when applied to the base model, or with PFDS, compared to a wider range for NFDS. In short, the models showed the expected changes in between and within-population variation under different population selection regimes.

These results highlight that resulting levels of divergence were highly variable. For each independent simulation, the starting genomic composition varied, and under these varying conditions, with the same parameters, similar values of divergence could be obtained across loci irrespective of linkage and selection. This is shown by the HMM, where the formation of islands and valleys potentially occur across all loci types (Figure 4).

3.2 | Dominance

The next set of models modified the genotype–phenotype map to include the effects of dominance rather than alleles acting in a purely additive manner. After 1000 generations of independent evolution, and application of a divergent selection scenario, an island of divergence was evident when moderate dominance was applied (similar to, though slightly higher in amplitude, a purely additive model), however application of strong dominance inhibited island formation altogether (Figure 2c). Under parallel selection, pronounced islands of divergence were not seen for either moderate or strong dominance, as was the case for the base additive model (Figure 3d).

The development of the genomic landscape of divergence over generations is shown in Figure S3. When dominance was strong, divergence levels were comparable for selected and neutral, and linked and unlinked loci from the earliest stages of divergence (100 generations) through to the latest stages (2000 generations), with only a slightly elevated selected region under divergent selection. The genomic island that formed under moderate dominance level and divergent selection was observed from the earliest stages. A transient, low-amplitude genomic island developed in the early stages under parallel selection and moderate dominance.

3.3 | Genotype-Environment interaction

The addition of G × E interactions of varying strengths ($G = 5$ to 25, see Equation 6) did not change the shape of the landscape of divergence to any large degree under divergent selection (Figure 2e). However, inclusion of G × E interactions depressed divergence across the genome compared to the base additive model when parallel selection was applied (Figure 2f). This difference from the additive model was maintained from the early to latest stage of divergence (Figure S4). Despite the strength of the applied G × E interaction, under varying initial genetic diversity, similar values of divergence could be obtained across loci irrespective of linkage and selection (Figure 3e,f).

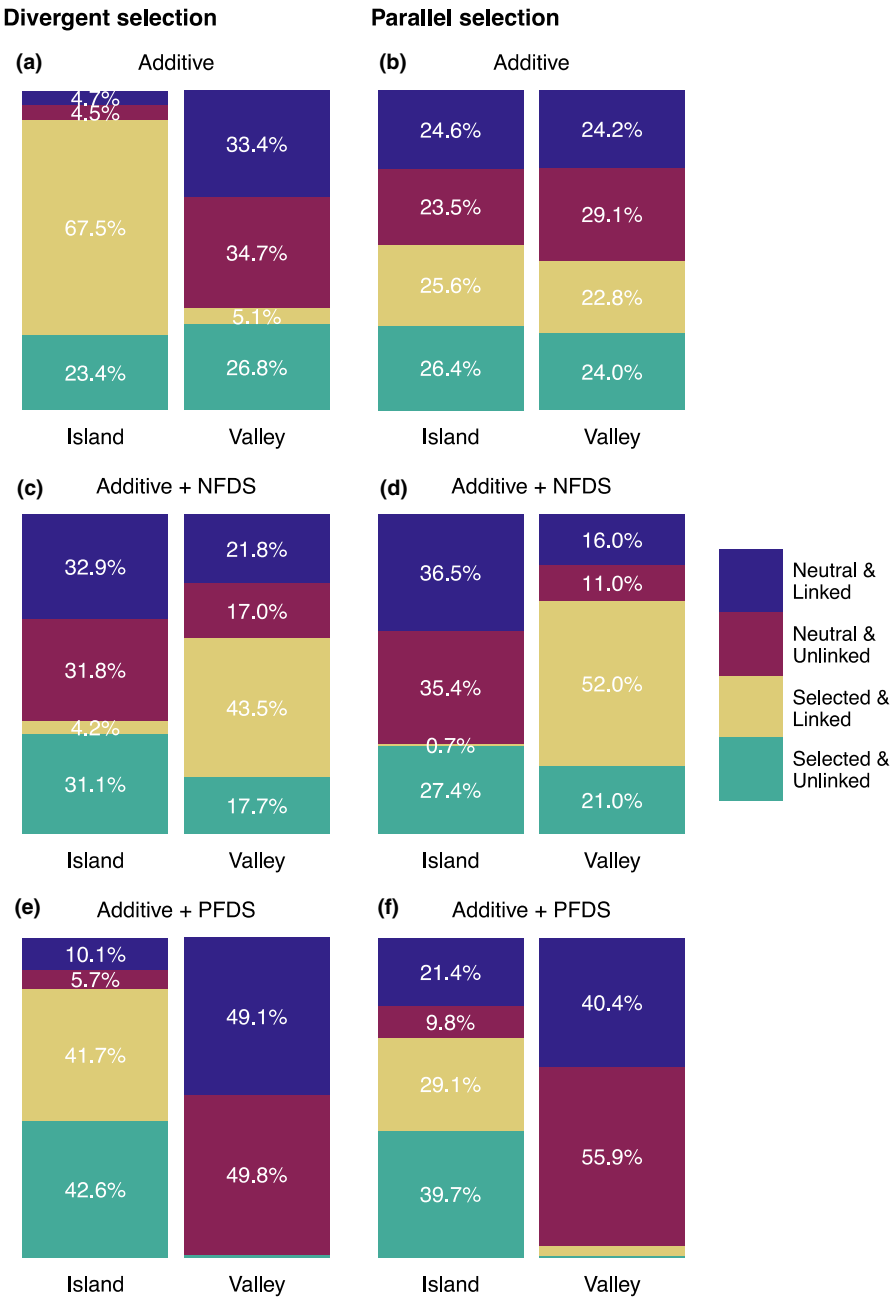

FIGURE 4 Proportions of islands and valleys observed on the genomic landscape of divergence under different type of selection, identified by using a hidden Markov model (100 independent simulations of 1000 generations each). The effect of divergent (left column) and parallel selection (right column) was analysed with: (a, b) an additive genotype–phenotype map; (c, d) with addition of negative frequency-dependent selection (NFDS); and (e, f) with addition of positive frequency-dependent selection (PFDS). The proportions of islands and valleys were computed for neutral and selected loci under linked and unlinked conditions. Under models with PFDS for selected linked and unlinked loci, the proportion of islands and valleys were 0.3% and 0.7% under divergent selection (panel e), and 3.3% and 0.4% under parallel selection, respectively (panel f).

3.4 | Migration

The addition of migration, whether acting in a constant or intermittent manner, slowed divergence overall compared to a model with no migration, but the main features of the profile of the genomic landscape were maintained: the formation of an island under divergent selection, which was more prominent at linked loci; and divergence levels equivalent to neutral loci under parallel selection (Figure 2g,h). However, delivery of the same amount of migration intermittently rather than constantly produced higher levels of divergence. This was more pronounced under a parallel selection regime, where the difference emerged early in the process (100 generations) and was maintained until the end of the simulations (2000 generations) (Figure S5). The variation around the average values obtained from the 100 independent simulations after 1000 generations showed that, under divergent selection for both migration models, selected loci had higher levels of divergence than neutral loci (Figure 3g,h).

4 | DISCUSSION

Our individual-based modelling approach has revealed the shapes of the genomic landscape of divergence when varying (1) the type of selection acting within and between populations, and (2) the form of the genotype–phenotype map. Genomic islands or valleys may form irrespective of whether between-population selection is divergent or parallel when positive and negative frequency-dependent selection also operate within each population, and can also form in the absence of selection. Furthermore, we show that a

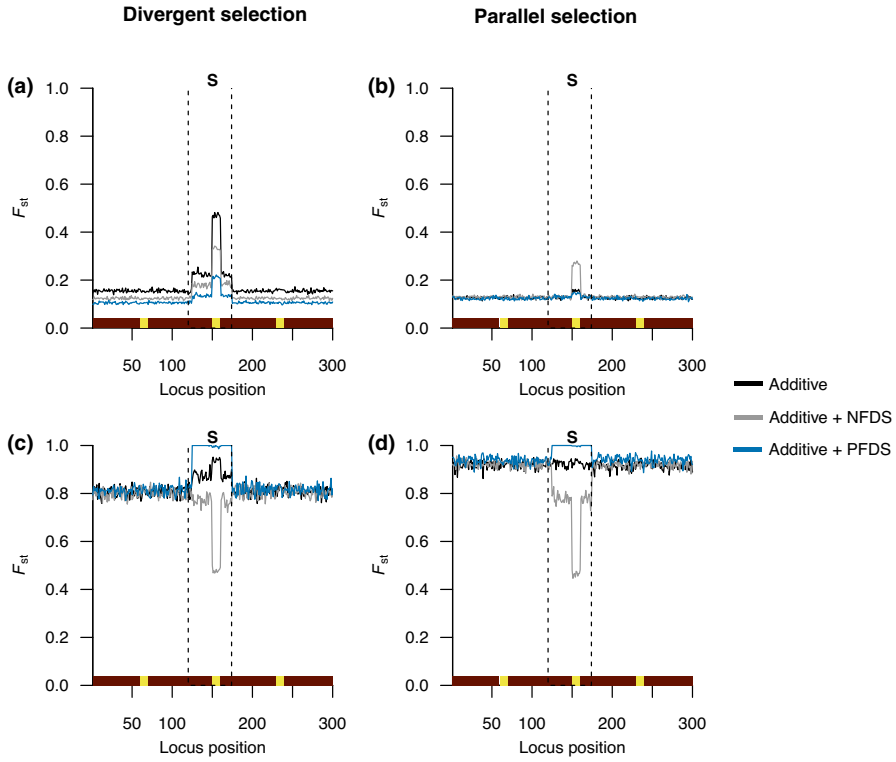

FIGURE 5 The genomic landscape of divergence with frequency-dependent selection summarized as the average F_{ST} from 100 simulations after: (a, b) 100 generations and (c, d) 2000 generations. The left-hand panels show divergent selection, and the right-hand panels, parallel selection. Locus position represented on the *x*-axis is shaded red for unlinked loci and yellow for linked loci, with loci under selection (S) that contribute to phenotypic trait values indicated by region between the vertical dashed lines. The black lines denote divergence levels under divergent or parallel selection without frequency-dependent selection, for comparison.

consequence of dominance in the genotype–phenotype map is the erosion of genomic islands under divergent selection. The level of divergence at both neutral and selected loci were highly variable in our simulations, and the formation of islands or valleys was not always guaranteed. This variable output is known to be influenced by a range of factors (Ravinet et al., 2017; Yeaman et al., 2016), including the initial genetic diversity of the ancestral population and drift (Quilodrán et al., 2020; Yeaman et al., 2016). Our findings confirm that the genotype–phenotype map, gene flow pattern, and selection regime affect the development of a heterogenous genomic landscape of divergence (Quilodrán et al., 2020). We also show they affect the maintenance of genetic diversity under certain conditions: negative frequency-dependent selection and dominance.

4.1 | A heterogenous genomic landscape of divergence under divergent and parallel selection is influenced by within-population frequency-dependent selection

Frequency-dependent selection within populations (PFDS and NFDS) influenced the genomic landscape over time, irrespective of the type of selection between populations (i.e., divergent and parallel). High linkage influenced the shape of the genomic landscape, with selection acting effectively on entire haplotype blocks that then rapidly increase in frequency. This corroborates empirical results where strong linkage typically associated with genomic inversions or proximity to centromeres, is often associated with heterogenous genomic landscapes (Huang et al., 2020; Puig Giribets et al., 2019; Sodeland et al., 2016; Tepolt & Palumbi, 2020).

PFDS resulted in a genomic island that was not evident under parallel selection alone, and was also taller under the effects of divergent selection. PFDS is known to accelerate the fixation of the most frequently occurring allele at a locus under selection (Gordon et al., 2015), which results in a rapid reduction of within population phenotypic variation in our simulations. While this increased the difference between quantitative phenotypes of populations under divergent selection, it decreased the difference under parallel selection, with both populations reaching similar phenotypes. Because there are multiple genetic combinations that result in a given quantitative phenotype, small differences in frequency of the most common allele in each population could be quickly amplified by PFDS, resulting in the development of genomic islands over time. Experimental and empirical investigations have shown the proliferation of monomorphism in species influenced by PFDS (Lindström et al., 2001; Mallet & Barton, 1989; Nokelainen et al., 2014), but also polymorphism (Chouteau et al., 2016; Ogilvie et al., 2021; Rönkä et al., 2020). Both are compatible with our expectation of PFDS under different types of selective pressure between populations.

NFDS as a type of balancing selection is expected to maintain ancestral polymorphism (Brisson, 2018). Genetic diversity within populations negatively correlates with divergence between populations, which may in turn exacerbate valley formation (Wolf & Ellegren, 2017). In our simulations, this is illustrated by the formation of genomic valleys irrespective of the form of selection between populations. The influence of NFDS maintained the within population phenotypic variation and consequently decreased the distance between the phenotypic space of the populations. This finding is in line with empirical observations made under allopatric conditions of morphological overlap between populations under NFDS (Franks & Oxford, 2017). However, under NFDS, the genomic landscape

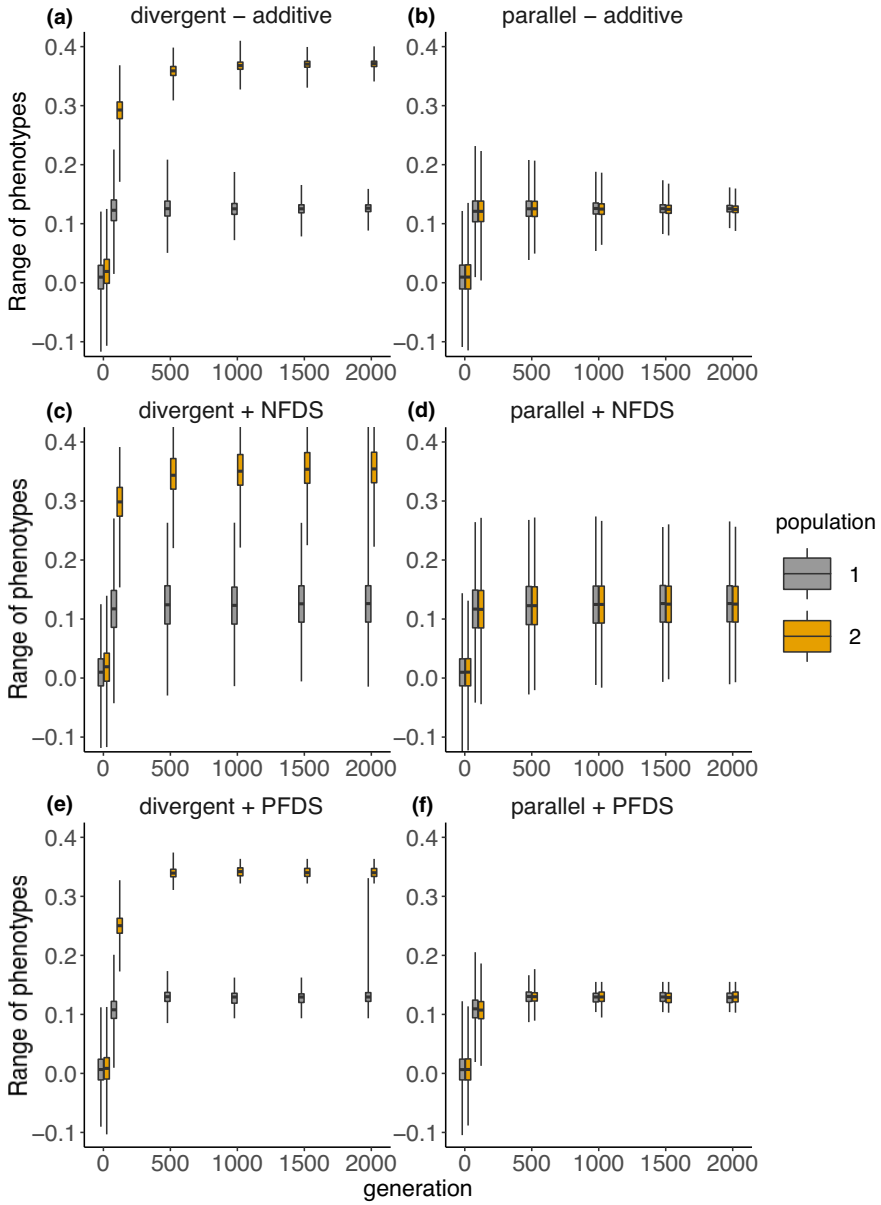

FIGURE 6 The phenotypic range occupied by individuals from two populations across 2000 generations. The phenotypes are standardized by dividing by the number of additive loci (n_a). The effect of divergent (left column) and parallel selection (right column) are shown with: (a, b) an additive genotype–phenotype map; (c, d) with addition of negative frequency-dependent selection (NFDS); and (e, f) with addition of positive frequency-dependent selection (PFDS). The whiskers represent the minimum and maximum phenotype, and the bars 25%, 50% and 75% quantiles.

feature for linked loci often transitioned from a genomic island seen at early stages of divergence (100 generations), to a valley by the later stages (1000 generations). This can be explained by the stochastic generation of different additive routes towards a common phenotype, forming idiosyncratic differences between the populations, resulting in genomic islands in some early-stage simulations (Goldberg et al., 2020). However, as divergence progresses, the cumulative effect of NFDS generates large within-population genotypic and phenotypic variation, resulting in genomic valleys at later stages for both parallel and divergent selection between populations.

4.2 | Maintenance of polymorphism under divergent selection limits island formation

The incorporation of polymorphism-maintaining processes such as dominance and NFDS (Brisson, 2018; Moulherat et al., 2017) shape the genomic landscape of divergence. Under strong dominance and divergent selection, genomic island formation was inhibited as recessive alleles were maintained in the heterozygous form. Other mechanisms not explored here by which dominance could maintain polymorphism include situations where dominance relationships among alleles reverse, for example where dominance is sex-dependent or non-linear relationships exist between fitness and gene activity (Connallon & Chenoweth, 2019; Nabutanyi & Wittmann, 2021). When the average strength of dominance decreases (under the medium dominance model), dominant-recessive combinations of alleles form more rarely, polymorphism is therefore maintained more rarely, allowing formation of islands of divergence as seen in our simulations. NFDS is even more effective at maintaining polymorphism than dominance because the latter is not a selective process (Asmussen et al., 2004), resulting in valley rather than island formation at selected loci. Overall, the incorporation of polymorphism maintaining processes could help reconcile differences between empirical and simulated levels of divergence. For example, in the silvereyes species complex (*Zosterops lateralis*) simulated

levels of divergence exceed empirical observations (Sendell-Price et al., 2020).

4.3 | Genotype-Environment interaction reduces divergence of selected and neutral loci under parallel selection

In heterogenous environments, G×E interactions can favour different genotypes in different environments, maintaining allelic diversity (Côté & Simons, 2020; Falconer & Mackay, 1996; Plesnar-Bielak et al., 2018). In our simulations, the phenotype was an additive combination of both the stochastic environmental component (varied between individuals in a population) and the additive genotypic component. As a result, different additive genotypes complemented different environmental components within a population (e.g., a given additive genotype may form a fitter phenotype with a particular environmental value and a less fit phenotype with a different environmental condition) (Falconer & Mackay, 1996). Furthermore, since the genotype – phenotype function in Equation 6 additively includes the product of the genotype and the environment, higher values of G×E also increase the phenotypic variance of the population. Our simulations reflect the increased phenotypic variance, as a consequence of the implementation of G×E in Equation 6. Under parallel selection, a consequence of the increased genetic variance was that more often common allelic combinations persisted in the simulations, decreasing F_{ST}. However, under divergent selection, different phenotypes were selectively favoured between both populations, in which alternate allelic combinations (generating alternate phenotypes) were favoured. In this last condition, the frequency of common allelic combinations between the two populations was not affected by the increase in phenotypic variance (as a consequence of G×E interaction), and so F_{ST} did not decrease.

4.4 | Intermittent gene flow slows the loss of genomic islands over time

Genomic islands of divergence are expected to disappear over time with the accumulation of genome-wide differences between populations (Nosil & Feder, 2012; Wu & Ting, 2004). However, the temporal dynamic of this process can be moderated by gene flow through migration, in which some gene flow slows the process of genomic island erosion, while a high level can rapidly erase them (Quilodrán et al., 2020). We show that the gene flow pattern (intermittent or continuous) influences the overall level of divergence, with intermittent gene flow resulting in a higher level of divergence despite an equal number of migrants, via a slowed loss of genomic islands over time. This emphasizes the importance of considering different patterns of gene flow when aiming to depict the historical divergence of any biological system (Kirkpatrick & Ravigné, 2002; Rundle & Nosil, 2005).

5 | CONCLUSION

The interpretation of past evolutionary histories based on the observation of given genomic patterns have to be made with caution (e.g., Funk et al., 2021; Ottenburghs et al., 2017), as the emergence of specific patterns is not ensured by a unique set of circumstances. While divergent selection with ongoing gene flow was a prominent explanation for genomic islands of divergence (Feder et al., 2013), more recent studies have shown that they may emerge from a variety of evolutionary routes in the absence of gene flow (Ravinet et al., 2017). Even under divergent selection and strong linkage, two elements that often show patterns of genomic islands when concurrently influencing genomic landscapes (Nam et al., 2020), we have shown that islands could disappear or even turn out to be valleys of similarity when adding other layers of complexity to the system, such as strong dominance and frequency-dependent selection. Computational approaches that concurrently analyse theoretical expectations with summary statistics based on empirical data (e.g., Approximate Bayesian computational methods, machine learning) provide an opportunity to narrow the range of potential drivers that produce a particular genomic landscape of divergence under defined evolutionary histories. The modelling tool used in our approach will allow further exploration of genomic landscape dynamics under the nuanced and complex conditions that characterize biological systems. Empirical data can be fitted to this flexible model to simulate the evolution of natural populations (e.g., Sendell-Price et al., 2020). In doing so, our individual-based approach may help to provide powerful insights into the effects of different microevolutionary processes on particular natural systems.

AUTHOR CONTRIBUTIONS

H.A.A.A., T.C., S.M.C., and C.S.Q. conceived the original idea for the study. H.A.A.A. coded and performed the simulations in R. H.A.A.A. wrote the first draft of the manuscript and all authors provided comments and edits on manuscript versions.

ACKNOWLEDGEMENTS

The authors would like to acknowledge the use of the University of Oxford Advanced Research Computing (ARC) facility (https://doi.org/10.5281/zenodo.22558) and the University of Geneva 'Yggdrasil' HPC cluster. CSQ acknowledges support from the Swiss National Science Foundation (P5R5PB_203169).

CONFLICT OF INTEREST STATEMENT

The authors declare they have no conflict of interest.

DATA AVAILABILITY STATEMENT

All information needed to repeat our simulations are described in Table 1. They can be reproduced by using the R package 'glads', which is available on GitHub and deposited in Zenodo: https: //doi.org/10.5281/zenodo.3516224. The custom functions used in the manuscript are available in Appendix S1.

ORCID

Hisham A. A. Ali https://orcid.org/0000-0001-7467-5199

Claudio S. Quilodrán https://orcid.org/0000-0001-7197-9154

REFERENCES

Allen, J. A. (1988). Frequency-dependent selection by predators. *Philosophical Transactions of the Royal Society of London. B, Biological Sciences*, *319*(1196), 485–503. https://doi.org/10.1098/rstb.1988.0061

Andrello, M., & Manel, S. (2015). MetaPopGen: An R package to simulate population genetics in large size metapopulations. *Molecular Ecology Resources*, *15*(5), 1153–1162. https://doi.org/10.1111/1755-0998.12371

Asmussen, M. A., Cartwright, R. A., & Spencer, H. G. (2004). Frequency-dependent selection with dominance: A window onto the behavior of the mean fitness. *Genetics*, *167*(1), 499–512. https://doi.org/10.1534/genetics.167.1.499

Bay, R. A., & Ruegg, K. (2017). Genomic islands of divergence or opportunities for introgression? *Proceedings of the Royal Society B: Biological Sciences*, *284*(1850), 20162414. https://doi.org/10.1098/rspb.2016.2414

Billiard, S., Castric, V., & Llaurens, V. (2021). The integrative biology of genetic dominance. *Biological Reviews*, *96*(6), 2925–2942. https://doi.org/10.1111/brv.12786

Bolnick, D. I., & Stutz, W. E. (2017). Frequency dependence limits divergent evolution by favouring rare immigrants over residents. *Nature*, *546*(7657), 285–288. https://doi.org/10.1038/nature22351

Brennan, A. C., Harris, S. A., Tabah, D. A., & Hiscock, S. J. (2002). The population genetics of sporophytic self-incompatibility in *Senecio squalidus* L. (Asteraceae) I: S allele diversity in a natural population. *Heredity*, *89*(6), 430–438. https://doi.org/10.1038/sj.hdy.6800159

Brisson, D. (2018). Negative frequency-dependent selection is frequently confounding. *Frontiers in Ecology and Evolution*, *6*, 1–9. https://doi.org/10.3389/fevo.2018.00010

Campbell, C. R., Poelstra, J. W., & Yoder, A. D. (2018). What is speciation genomics? The roles of ecology, gene flow, and genomic architecture in the formation of species. *Biological Journal of the Linnean Society*, *124*(4), 561–583. https://doi.org/10.1093/biolinnean/bly063

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

Chouteau, M., Arias, M., & Joron, M. (2016). Warning signals are under positive frequency dependent selection in nature. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(8), 2164–2169. https://doi.org/10.1073/pnas.1519216113

Clarke, B. C. (1979). The evolution of genetic diversity. *Proceedings of the Royal Society B: Biological Sciences*, *205*(1161), 453–474. https://doi.org/10.1098/rspb.1979.0079

Clarke, B. C., & O'Donald, P. (1962). Frequency-dependent selection. *Heredity*, *19*(2), 201–206. https://doi.org/10.1038/hdy.1964.25

Connallon, T., & Chenoweth, S. F. (2019). Dominance reversals and the maintenance of genetic variation for fitness. *PLoS Biology*, *17*(1), 1–11. https://doi.org/10.1371/journal.pbio.3000118

Côté, K., & Simons, A. M. (2020). Genotype-environment interaction and the maintenance of genetic variation: An empirical study of *Lobelia inflata* (Campanulaceae). *Royal Society Open Science*, *7*(3), 191720. https://doi.org/10.1098/rsos.191720

Coyne, J. A., & Orr, H. A. (2004). Speciation. *Sinauer*, *328*, 494–497. https://doi.org/10.1126/science.1187468

Currat, M., Gerbault, P., Di, D., Nunes, J. M., & Sanchez-Mazas, A. (2015). Forward-in-time, spatially explicit modeling software to simulate genetic lineages under selection. *Evolutionary Bioinformatics*, *11*, 27–39. https://doi.org/10.4137/EBO.S33488

Cvijović, I., Good, B., & Desai, M. (2018). The effect of strong purifying selection on genetic diversity. *Genetics*, *209*, 1235–1278. https://doi.org/10.1534/genetics.118.301058

DeAngelis, D. L., & Grimm, V. (2014). Individual-based models in ecology after four decades. *F1000Prime Reports*, *6*, 39. https://doi.org/10.12703/P6-39

Doebeli, M., & Ispolatov, I. (2010). Complexity and diversity. *Science*, *328*(5977), 494–497. https://doi.org/10.1126/science.1187468

Endler, J. A. (1988). Frequency-dependent predation, crypsis and aposematic coloration. *Philosophical Transactions of the Royal Society of London. B, Biological Sciences*, *319*(1196), 505–523. https://doi.org/10.1098/rstb.1988.0062

Etges, W. J., De Oliveira, C. C., Gragg, E., Ortíz-Barrientos, D., Noor, M. A. F., & Ritchie, M. G. (2007). Genetics of incipient speciation in *Drosophila mojavensis*. I. Male courtship song, mating success, and genotype X environment interactions. *Evolution*, *61*(5), 1106–1119. https://doi.org/10.1111/j.1558-5646.2007.00104.x

Falconer, D. S., & Mackay, T. F. C. (1996). *Introduction to quantitative genetics* (4th ed.). Longman.

Feder, J. L., Egan, S. P., & Nosil, P. (2012). The genomics of speciation-with-gene-flow. *Trends in Genetics*, *28*(7), 342–350. https://doi.org/10.1016/j.tig.2012.03.009

Feder, J. L., Flaxman, S. M., Egan, S. P., Comeault, A. A., & Nosil, P. (2013). Geographic mode of speciation and genomic divergence. *Annual Review of Ecology, Evolution, and Systematics*, *44*, 73–97. https://doi.org/10.1146/annurev-ecolsys-110512-135825

Feder, J. L., Gejji, R., Yeaman, S., & Nosil, P. (2012). Establishment of new mutations under divergence and genome hitchhiking. *Philosophical Transactions of the Royal Society, B: Biological Sciences*, *367*(1587), 461–474. https://doi.org/10.1098/rstb.2011.0256

Feder, J. L., & Nosil, P. (2009). Chromosomal inversions and species differences: When are genes affecting adaptive divergence and reproductive isolation expected to reside within inversions? *Evolution*, *63*(12), 3061–3075. https://doi.org/10.1111/j.1558-5646.2009.00786.x

Feder, J. L., & Nosil, P. (2010). The efficacy of divergence hitchhiking in generating genomic islands during ecological speciation. *Evolution*, *64*(6), 1729–1747. https://doi.org/10.1111/j.1558-5646.2009.00943.x

Franks, D. W., & Oxford, G. S. (2017). The co-evolution of anti-predator polymorphisms in sympatric populations. *Biological Journal of the Linnean Society*, *122*(4), 729–737. https://doi.org/10.1093/biolinnean/blx111

Fry, J. D. (2010). The genomic location of sexually antagonistic variation: Some cautionary comments. *Evolution*, *64*(5), 1510–1516. https://doi.org/10.1111/j.1558-5646.2009.00898.x

Fujii, S., & Takayama, S. (2018). Multilayered dominance hierarchy in plant self-incompatibility. *Plant Reproduction*, *31*(1), 15–19. https://doi.org/10.1007/s00497-017-0319-9

Funk, E. R., Spellman, G. M., Winker, K., Withrow, J. J., Ruegg, K. C., Zavaleta, E., & Taylor, S. A. (2021). Phylogenomic data reveal widespread introgression across the range of an alpine and arctic specialist. *Systematic Biology*, *70*(3), 527–541. https://doi.org/10.1093/sysbio/syaa071

Futuyma, D. J. (1987). On the role of species in anagenesis. *American Naturalist*, *130*(3), 465–473. https://doi.org/10.1086/284724

Gavrilets, S. (2014). Models of speciation: Where are we now? *Journal of Heredity*, *105*(S1), 743–755. https://doi.org/10.1093/jhered/esu045

Goldberg, J. K., Lively, C. M., Sternlieb, S. R., Pintel, G., Hare, J. D., Morrissey, M. B., & Delph, L. F. (2020). Herbivore-mediated negative frequency-dependent selection underlies a trichome dimorphism in nature. *Evolution Letters*, *4*(1), 83–90. https://doi.org/10.1002/evl3.157

Gordon, S. P., Kokko, H., Rojas, B., Nokelainen, O., & Mappes, J. (2015). Colour polymorphism torn apart by opposing positive

frequency-dependent selection, yet maintained in space. *Journal of Animal Ecology*, *84*(6), 1555–1564. https://doi.org/10.1111/1365-2656.12416

Grimm, V., Berger, U., Bastiansen, F., Eliassen, S., Ginot, V., Giske, J., Goss-Custard, J., Grand, T., Heinz, S. K., Huse, G., Huth, A., Jepsen, J. U., Jørgensen, C., Mooij, W. M., Müller, B., Pe'er, G., Piou, C., Railsback, S. F., Robbins, A. M., … DeAngelis, D. L. (2006). A standard protocol for describing individual-based and agent-based models. *Ecological Modelling*, *198*(1–2), 115–126. https://doi.org/10.1016/j.ecolmodel.2006.04.023

Haller, B. C., Galloway, J., Kelleher, J., Messer, P. W., & Ralph, P. L. (2019). Tree-sequence recording in SLiM opens new horizons for forward-time simulation of whole genomes. *Molecular Ecology Resources*, *19*(2), 552–566. https://doi.org/10.1111/1755-0998.12968

Haller, B. C., & Messer, P. W. (2019). SLiM 3: Forward genetic simulations beyond the Wright-fisher model. *Molecular Biology and Evolution*, *36*(3), 632–637. https://doi.org/10.1093/molbev/msy228

Harte, D. (2017). *HiddenMarkov: Hidden Markov Models (R package version 1.8-11)*. Statistics Research Associates. http://www.statsresearch.co.nz/dsh/sslib/

Hoban, S., Bertorelle, G., & Gaggiotti, O. E. (2012). Computer simulations: Tools for population and evolutionary genetics. *Nature Reviews Genetics*, *13*(2), 110–122. https://doi.org/10.1038/nrg3130

Huang, K., Andrew, R. L., Owens, G. L., Ostevik, K. L., & Rieseberg, L. H. (2020). Multiple chromosomal inversions contribute to adaptive divergence of a dune sunflower ecotype. *Molecular Ecology*, *29*(14), 2535–2549. https://doi.org/10.1111/mec.15428

Jetz, W., Sekercioglu, C. H., & Böhning-Gaese, K. (2008). The worldwide variation in avian clutch size across species and space. *PLoS Biology*, *6*(12), 2650–2657. https://doi.org/10.1371/journal.pbio.0060303

Kelleher, J., Etheridge, A. M., & McVean, G. (2016). Efficient coalescent simulation and genealogical analysis for large sample sizes. *PLoS Computational Biology*, *12*(5), 1–22. https://doi.org/10.1371/journal.pcbi.1004842

Kim, J. H., Mori, T., Wakana, A., Ngo, B. X., Sakai, K., & Kajiwara, K. (2011). Determination of self-incompatible citrus cultivars with S1 and/or S2 alleles by pollination with homozygous S1 seedlings (S1S1 or S2S2) of "Banpeiyu" pummelo. *Journal of the Japanese Society for Horticultural Science*, *80*(4), 404–413. https://doi.org/10.2503/jjshs1.80.404

Kirkpatrick, M., & Ravigné, V. (2002). Speciation by natural and sexual selection: Models and experiments. *American Naturalist*, *159*(3 Suppl), S22–S35. https://doi.org/10.2307/3078919

Kojima, K. (1959). Stable equilibria for the optimum model. *Proceedings of the National Academy of Sciences of the United States of America*, *45*(7), 989–993. https://doi.org/10.1073/pnas.45.7.989

Kowyama, Y., Takahasi, H., Muraoka, K., Tani, T., Hara, K., & Shiotani, I. (1994). Number, frequency & dominance relationships of S-alleles in diploid Ipomoea trifida. *Heredity*, *73*(3), 275–283. https://doi.org/10.1038/hdy.1994.134

Küpper, C., Stocks, M., Risse, J. E., Dos Remedios, N., Farrell, L. L., McRae, S. B., Morgan, T. C., Karlionova, N., Pinchuk, P., Verkuil, Y. I., Kitaysky, A. S., Wingfield, J. C., Piersma, T., Zeng, K., Slate, J., Blaxter, M., Lank, D. B., & Burke, T. (2015). A supergene determines highly divergent male reproductive morphs in the ruff. *Nature Genetics*, *48*(1), 79–83. https://doi.org/10.1038/ng.3443

Lamichhaney, S., Berglund, J., Almén, M. S., Maqbool, K., Grabherr, M., Martinez-Barrio, A., Promerová, M., Rubin, C. J., Wang, C., Zamani, N., Grant, B. R., Grant, P. R., Webster, M. T., & Andersson, L. (2015). Evolution of Darwin's finches and their beaks revealed by genome sequencing. *Nature*, *518*(7539), 371–375. https://doi.org/10.1038/nature14181

Langham, G. M. (2004). Specialized avian predators repeatedly attack novel color morphs of Heliconius butterflies. *Evolution*, *58*(12), 2783–2787. https://doi.org/10.1111/j.0014-3820.2004.tb01629.x

Lindström, L., Alatalo, R. V., Lyytinen, A., & Mappes, J. (2001). Strong antiapostatic selection against novel rare aposematic prey. *Proceedings of the National Academy of Sciences of the United States of America*, *98*(16), 9181–9184. https://doi.org/10.1073/pnas.161071598

Lohse, K. (2017). Come on feel the noise – From metaphors to null models. *Journal of Evolutionary Biology*, *30*(8), 1506–1508. https://doi.org/10.1111/jeb.13109

Mallet, J. (2007). Hybrid speciation. *Nature*, *446*(7133), 279–283. https://doi.org/10.1038/nature05706

Mallet, J., & Barton, N. H. (1989). Strong natural selection in a warning-color hybrid zone. *Evolution*, *43*(2), 421–431. https://doi.org/10.1111/j.1558-5646.1989.tb04237.x

Marques, D. A., Lucek, K., Meier, J. I., Mwaiko, S., Wagner, C. E., Excoffier, L., & Seehausen, O. (2016). Genomics of rapid incipient speciation in sympatric Threespine stickleback. *PLoS Genetics*, *12*(2), 1–34. https://doi.org/10.1371/journal.pgen.1005887

Mavárez, J., Salazar, C. A., Bermingham, E., Salcedo, C., Jiggins, C. D., & Linares, M. (2006). Speciation by hybridization in Heliconius butterflies. *Nature*, *441*(7095), 868–871. https://doi.org/10.1038/nature04738

Michel, A. P., Sim, S., Powell, T. H. Q., Taylor, M. S., Nosil, P., & Feder, J. L. (2010). Widespread genomic divergence during sympatric speciation. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(21), 9724–9729. https://doi.org/10.1073/pnas.1000939107

Mitchell-Olds, T. (1995). The molecular basis of quantitative genetic variation in natural populations. *Trends in Ecology & Evolution*, *10*(8), 324–328. https://doi.org/10.1016/S0169-5347(00)89119-3

Morjan, C. L., & Rieseberg, L. H. (2004). How species evolve collectively: Implications of gene flow and selection for the spread of advantageous alleles. *Molecular Ecology*, *13*(6), 1341–1356. https://doi.org/10.1111/j.1365-294X.2004.02164.x

Moulherat, S., Chaine, A., Mangin, A., Aubret, F., Sinervo, B., & Clobert, J. (2017). The roles of plasticity versus dominance in maintaining polymorphism in mating strategies. *Scientific Reports*, *7*(1), 1–10. https://doi.org/10.1038/s41598-017-15078-1

Nabutanyi, P., & Wittmann, M. J. (2021). Models for eco-evolutionary extinction vortices under balancing selection. *American Naturalist*, *197*(3), 336–350. https://doi.org/10.1086/712805

Nam, K., Nhim, S., Robin, S., Bretaudeau, A., Nègre, N., & d'Alençon, E. (2020). Positive selection alone is sufficient for whole genome differentiation at the early stage of speciation process in the fall armyworm. *BMC Evolutionary Biology*, *20*(1), 1–16. https://doi.org/10.1186/s12862-020-01715-3

Nokelainen, O., Valkonen, J., Lindstedt, C., & Mappes, J. (2014). Changes in predator community structure shifts the efficacy of two warning signals in Arctiid moths. *Journal of Animal Ecology*, *83*(3), 598–605. https://doi.org/10.1111/1365-2656.12169

Nosil, P., & Feder, J. L. (2012). Genomic divergence during speciation: Causes and consequences. *Philosophical Transactions of the Royal Society, B: Biological Sciences*, *367*(1587), 332–342. https://doi.org/10.1098/rstb.2011.0263

Nosil, P., Harmon, L. J., & Seehausen, O. (2009). Ecological explanations for (incomplete) speciation. *Trends in Ecology & Evolution*, *24*(3), 145–156. https://doi.org/10.1016/j.tree.2008.10.011

Ogilvie, J. G., Van Belleghem, S., Range, R., Papa, R., McMillan, O. W., Chouteau, M., & Counterman, B. A. (2021). Balanced polymorphisms and their divergence in a Heliconius butterfly. *Ecology and Evolution*, *11*(24), 18319–18330. https://doi.org/10.1002/ece3.8423

Ottenburghs, J., Megens, H. J., Kraus, R. H. S., Van Hooft, P., Van Wieren, S. E., Crooijmans, R. P. M. A., Ydenberg, R. C., Groenen, M. A. M., & Prins, H. H. T. (2017). A history of hybrids? Genomic patterns of introgression in the True Geese. *BMC Evolutionary Biology*, *17*(1), 1–14. https://doi.org/10.1186/s12862-017-1048-2

Paradis, E. (2010). Pegas: An R package for population genetics with an integrated-modular approach. *Bioinformatics*, *26*(3), 419–420. https://doi.org/10.1093/bioinformatics/btp696

Peng, B., & Kimmel, M. (2005). simuPOP: A forward-time population genetics simulation environment. *Bioinformatics*, *21*(18), 3686–3687. https://doi.org/10.1093/bioinformatics/bti584

Plesnar-Bielak, A., Skwierzyńska, A. M., Hlebowicz, K., & Radwan, J. (2018). Relative costs and benefits of alternative reproductive phenotypes at different temperatures—Genotype-by-environment interactions in a sexually selected trait. *BMC Evolutionary Biology*, *18*(1), 1–10. https://doi.org/10.1186/s12862-018-1226-x

Puig Giribets, M., García Guerreiro, M. P., Santos, M., Ayala, F. J., Tarrío, R., & Rodríguez-Trelles, F. (2019). Chromosomal inversions promote genomic islands of concerted evolution of Hsp70 genes in the Drosophila subobscura species subgroup. *Molecular Ecology*, *28*(6), 1316–1332. https://doi.org/10.1111/mec.14511

Quilodrán, C. S., Ruegg, K., Sendell-Price, A. T., Anderson, E. C., Coulson, T., & Clegg, S. M. (2020). The multiple population genetic and demographic routes to islands of genomic divergence. *Methods in Ecology and Evolution*, *11*(1), 6–21. https://doi.org/10.1111/2041-210X.13324

R Core Team. (2019). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.r-project.org/

Ravinet, M., Faria, R., Butlin, R. K., Galindo, J., Bierne, N., Rafajlović, M., Noor, M. A. F., Mehlig, B., & Westram, A. M. (2017). Interpreting the genomic landscape of speciation: A road map for finding barriers to gene flow. *Journal of Evolutionary Biology*, *30*(8), 1450–1477. https://doi.org/10.1111/jeb.13047

Roesti, M., Gavrilets, S., Hendry, A. P., Salzburger, W., & Berner, D. (2014). The genomic signature of parallel adaptation from shared genetic variation. *Molecular Ecology*, *23*(16), 3944–3956. https://doi.org/10.1111/mec.12720

Rönkä, K., Valkonen, J. K., Nokelainen, O., Rojas, B., Gordon, S., Burdfield-Steel, E., & Mappes, J. (2020). Geographic mosaic of selection by avian predators on hindwing warning colour in a polymorphic aposematic moth. *Ecology Letters*, *23*(11), 1654–1663. https://doi.org/10.1111/ele.13597

Rundle, H. D., & Nosil, P. (2005). Ecological speciation. *Ecology Letters*, *8*(3), 336–352. https://doi.org/10.1111/j.1461-0248.2004.00715.x

Schluter, D. (2009). Evidence for ecological speciation and its alternative. *Science*, *323*(5915), 737–741. https://doi.org/10.1126/science.1160006

Sedghifar, A., Brandvain, Y., & Ralph, P. (2016). Beyond clines: Lineages and haplotype blocks in hybrid zones. *Molecular Ecology*, *25*(11), 2559–2576. https://doi.org/10.1111/mec.13677

Seehausen, O., Butlin, R. K., Keller, I., Wagner, C. E., Boughman, J. W., Hohenlohe, P. A., Peichel, C. L., Saetre, G. P., Bank, C., Brännström, Å., Brelsford, A., Clarkson, C. S., Eroukhmanoff, F., Feder, J. L., Fischer, M. C., Foote, A. D., Foote, A. D., Franchini, P., Jiggins, C. D., … Widmer, A. (2014). Genomics and the origin of species. *Nature Reviews Genetics*, *15*(3), 176–192. https://doi.org/10.1038/nrg3644

Sendell-Price, A. T., Ruegg, K. C., Anderson, E. C., Quilodrán, C. S., Van Doren, B. M., Underwood, V. L., Coulson, T., & Clegg, S. M. (2020). The genomic landscape of divergence across the speciation continuum in Island-colonising silvereyes (*Zosterops lateralis*). *G3: Genes, Genomes, Genetics*, *10*(9), 3147–3163. https://doi.org/10.1534/g3.120.401352

Shafer, A. B. A., & Wolf, J. B. W. (2013). Widespread evidence for incipient ecological speciation: A meta-analysis of isolation-by-ecology. *Ecology Letters*, *16*(7), 940–950. https://doi.org/10.1111/ele.12120

Sinervo, B., & Lively, C. M. (1996). The rock–paper–scissors game and the evolution of alternative male strategies. *Nature*, *380*(6571), 240–243. https://doi.org/10.1038/380240a0

Smithson, A. (2001). Pollinator preference, frequency dependence, and floral evolution. In J. D. Thomson & L. Chittka (Eds.), *Cognitive ecology of pollination: Animal behaviour and floral evolution* (pp. 237–258). Cambridge University Press. https://doi.org/10.1017/CBO9780511542268.013

Sodeland, M., Jorde, P. E., Lien, S., Jentoft, S., Berg, P. R., Grove, H., Kent, M. P., Arnyasi, M., Olsen, E. M., & Knutsen, H. (2016). "Islands of Divergence" in the Atlantic cod genome represent polymorphic chromosomal rearrangements. *Genome Biology and Evolution*, *8*(4), 1012–1022. https://doi.org/10.1093/gbe/evw057

Svensson, E. I., Gómez-Llano, M. A., Torres, A. R., & Bensch, H. M. (2018). Frequency dependence and ecological drift shape coexistence of species with similar niches. *American Naturalist*, *191*(6), 691–703. https://doi.org/10.1086/697201

Tepolt, C. K., & Palumbi, S. R. (2020). Rapid adaptation to temperature via a potential genomic Island of divergence in the invasive green crab, *Carcinus maenas*. *Frontiers in Ecology and Evolution*, *8*, 1–14. https://doi.org/10.3389/fevo.2020.580701

Thompson, V. (1984). Polymorphism under apostatic and aposematic selection. *Heredity*, *53*(3), 677–686. https://doi.org/10.1038/hdy.1984.126

Tishkoff, S. A., Reed, F. A., Ranciaro, A., Voight, B. F., Babbitt, C. C., Silverman, J. S., Powell, K., Mortensen, H. M., Hirbo, J. B., & Osman, M. (2007). Convergent adaptation of human lactase persistence in Africa and Europe. *Nature Genetics*, *39*(1), 31–40. https://doi.org/10.1038/ng1946

Weir, B. S., & Cockerham, C. C. (1984). Estimating F-statistics for the analysis of population structure. *Evolution*, *38*(6), 1358–1370. https://doi.org/10.2307/2408641

Wolf, J. B. W., & Ellegren, H. (2017). Making sense of genomic islands of differentiation in light of speciation. *Nature Reviews Genetics*, *18*(2), 87–100. https://doi.org/10.1038/nrg.2016.133

Woodgate, J. L., Buchanan, K. L., Bennett, A. T. D., Catchpole, C. K., Brighton, R., & Leitner, S. (2014). Environmental and genetic control of brain and song structure in the zebra finch. *Evolution*, *68*(1), 230–240. https://doi.org/10.1111/evo.12261

Wu, C. I., & Ting, C. T. (2004). Genes and speciation. *Nature Reviews Genetics*, *5*(2), 114–122. https://doi.org/10.1038/nrg1269

Xia, X. (2013). Codominance. In S. Maloy & K. Hughes (Eds.), *Brenner's encyclopedia of genetics* (2nd ed., pp. 63–64). Academic Press. https://doi.org/10.1016/B978-0-12-374984-0.00278-3

Yeaman, S., Aeschbacher, S., & Bürger, R. (2016). The evolution of genomic islands by increased establishment probability of linked alleles. *Molecular Ecology*, *25*(11), 2542–2558. https://doi.org/10.1111/mec.13611

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Ali, H. A. A., Coulson, T., Clegg, S. M., & Quilodrán, C. S. (2023). The effect of divergent and parallel selection on the genomic landscape of divergence. *Molecular Ecology*, *00*, 1–16. https://doi.org/10.1111/mec.17225