Contingent Convergence: The ability to detect convergent genomic evolution is dependent on demographic similarity

AUTHORS

James R. Whiting[1], Bonnie A. Fraser[1]

[1] *Department of Biosciences, University of Exeter, Geoffrey Pope Building, Exeter, EX4 4QD*

ABSTRACT

Outlier scans, in which the genome is scanned for signatures of selection, have become a prominent tool in studies of local adaptation, and more recently studies of genetic convergence in natural populations. However, these methods have the potential to be confounded by features of demographic history, which are considerably variable across natural populations. In this study, we use forward-simulations to investigate and illustrate how several measures of genetic differentiation commonly used in outlier scans (F_{ST}, D_{XY} and $\Delta\pi$) are influenced by demographic variation across multiple sampling time points. In a factorial design with 32 treatments, we manipulate the strength of bottlenecks (N of founding individuals), contractions (proportional size of derived population) and migration rate between an ancestral population and a derived population. Our results illustrate known constraints of individual measures associated with strong contractions and a lack of gene-flow, but notably we demonstrate how relationships between measures are similarly dependent on demography. We find that false-positive signals of convergent evolution (the same simulated genes marked as outliers in independent treatments) are attainable as a product of similar demographic treatment, and that outliers across different measures (particularly F_{ST} and D_{XY}) can occur with little influence of selection. Taken together, we show how underappreciated, yet quantifiable measures of demographic history can influence commonly employed methods for detecting selection.

KEYWORDS

Outlier scans, demographic history, F_{ST}, D_{XY}, false-positives, convergence

INTRODUCTION

Big-data studies of adaptation and evolution across the genome are increasing in popularity with the availability of modern sequencing technologies. These studies often centre around quantifying patterns of variation in genome-wide SNPs, which can be used to highlight regions or genes having experienced selection relative to the neutral backdrop of the rest of the genome. These analyses, which we refer to here as outlier scans, have become a common tool in population genetics and have seen great success across diverse, natural systems in identifying candidate genes associated with the evolution of a range of adaptive traits (reviewed recently by (Ahrens et al. 2018)). More recently, the method of overlapping outlier scans across independent lineages has been employed to test whether the same regions are involved in independent adaptation events (i.e. genetic convergent evolution). This study seeks to investigate how different outlier scan methods are influenced by demographic variation in natural populations, and how this may lead to overlapping false-positives.

Recent discussions have highlighted the propensity of outlier scans to yield false-positives, given that outliers caused by heterogeneous genomic landscapes are commonplace irrespective of selection (Ellegren and Wolf 2017). Specifically, background selection (BGS), whereby linkage between neutral and deleterious variants reduces local diversity (B. Charlesworth, Morgan, and Charlesworth 1993; Bank et al. 2014; Burri 2017), has been invoked to offer alternative explanations for patterns at first attributed to directional selection (Cruickshank and Hahn 2014). Furthermore, the influence of neutral processes such as genetic drift can similarly produce false-positive signatures of selection (Kimura 1983).

The strength of these processes are dependent on demography. For example, the influence of neutral processes and BGS should be more pronounced in smaller populations (low N_e) (D. Charlesworth, Charlesworth, and Morgan 1995; Cutter and Payseur 2013; B. Charlesworth 2009; Yeaman and Otto 2011); as has been demonstrated in humans (Torres, Szpiech, and Hernandez 2018) and *Drosophila* (Sella et al. 2009; B. Charlesworth 1996). This relationship, however, may not be simply linear, with additional simulation evidence suggesting the effects of BGS are strongest at intermediate-low N_e, with effects becoming weaker again at very low N_e (Zeng 2013).

Lack of connectivity among populations may also elevate measures of genetic differentiation, if a large global population is composed of smaller, isolated, sub-divided populations that each experience the effects of reduced N_e (B. Charlesworth, Nordborg, and Charlesworth 1997; Hoban et al. 2016). In an attempt to mitigate the likelihood of false-positives, some have advocated using multiple measures of population differentiation and divergence to identify regions of the genome likely to be under selection (B. Charlesworth 1998; Cruickshank and Hahn 2014). However, how these measures are correlated with each other and with selection under different demographic scenarios has not been explored.

A diverse array of measures of genetic differentiation and divergence have been employed for outlier scans, but here we focus on two of the most common measures of relative differentiation (F_{ST} and π) and a third measure of absolute divergence (D_{XY}). The most commonly used measure in outlier scans is the fixation-index F_{ST} (Weir and Cockerham 1984; Hudson, Slatkin, and Maddison 1992), which measures the relative importance of within- and between-population variance. This measure is thus maximised when genomic regions exhibit the lowest within- and highest between-population variance. F_{ST} outliers at the right-tail of the distribution are considered candidates for adaptation because they reflect regions with large differences in allele frequencies or high

substitution rate (large between-population variance) and/or low nucleotide variation (low within-population variance) relative to the rest of the genome. Changes in nucleotide diversity (π) are another indicator of adaptation, as selection on a beneficial allele limits variation within a population resulting in selective sweeps (Smith and Haigh 1974). Comparisons of the ratio of π between diverging populations reveal regions under selection as local π is reduced in one population in comparison to the other. Whilst similar to F_{ST}, changes in π ($\Delta\pi$) do not discriminate between which copy of a polymorphism is fixed, such that a substitution between populations is equivalent to a common non-polymorphic site. $\Delta\pi$ outliers therefore represent regions of the genome with reduced π in either population relative to the rest of the genome. As a measure of absolute genetic divergence, D_{XY} (Nei 1987) does not consider the relative frequencies of polymorphisms within populations (B. Charlesworth 1998; Cruickshank and Hahn 2014). D_{XY} can be quantified as the average number of pairwise differences between sequence comparisons between two populations. This measure is therefore influenced by ancestral π and the substitution rate, so outliers highlight regions that are highly variable ancestrally, or in either population (large π), or exhibit many substitutions and thus increased sequence divergence.

Because each measure of differentiation quantifies genetic variation between populations in a slightly different way (Cruickshank and Hahn 2014; Matthey-Doret and Whitlock 2018; Whitlock and Lotterhos 2015), each has a unique relationship with demography. Whilst we can predict how individual measures are influenced by demography, and subsequently neutral processes, we know little about how the relationships between measures are affected. Therefore their combined power to detect selection should also vary with demography, suggesting that combining measures may provide unwarranted confidence in results.

This complex interplay between selection measures and demography may be further exacerbated where studies compare scans from multiple populations to identify convergent genomic evolution.

Here, researchers use measures of population divergence across independent pairs of evolutionary replicates with 'outlier' loci compared across results. This strategy has been employed extensively across diverse taxa, including: birds (Cooper and Uy 2017), fish (Hohenlohe et al. 2010; Jones, Chan, et al. 2012; Rougemont et al. 2017; Fraser et al. 2015; Reid et al. 2016; Meier et al. 2018), insects (Soria-Carrasco et al. 2014; Van Belleghem et al. 2018), mammals (Waterhouse et al. 2018), molluscs (Ravinet et al. 2016; Westram et al. 2014) and plants (Trucchi et al. 2017; Roda et al. 2013). However, different demographic histories within independent scans, and the subsequent effects on measures of population divergence, have rarely been considered in how they affect our ability to detect convergent genetic evolution.

Here, we used forward-simulations to investigate the effects of different demographic histories on the relationships between measures of divergence with selection. We focus on examining SNP variation between ancestral and derived populations, although other methods are employed in the literature (e.g. replicated clines, phylogenomics, single population measures e.g. Tajima's D). We varied demography through manipulating the number of founding individuals (bottlenecks), the population size of the derived population (contractions) and the presence/absence of migration. We simulated the effects of selection on realistic genes designed from features taken from the guppy genome assembly, a prominent model system for studies of convergent evolution (Fraser et al. 2015; Reznick and Endler 1982). Moreover, we measured genetic divergence at four set time points through F_{ST}, D_{XY} and $\Delta\pi$ between an ancestral population and derived population to examine temporal relationships. Our aim was to investigate the significance of demographic history when employing outlier scans and highlight demographic scenarios that are susceptible to false-positives. We also aimed to test when common outliers across measures occur, and whether overlapping outliers are consistently good indicators of selection. We sought to answer the following questions: 1) How do demographic factors influence measures of divergence through time? 2) How well do measures of divergence identify regions of the genome under strong selection through time and

across demographic factors? 3) How are the different measures of divergence related through time and across demographic factors? 4) Where do we detect the strongest signals of convergence (i.e. overlapping outliers using single or multiple measures), and are these consistent with selection?

METHODS

The forward-simulation software SLiM 3.0 (Haller and Messer 2019) was used to simulate population divergence under contrasting demographic treatments in a fully factorial design across an ancestral population of N = 1000 with a mutation rate based on the guppy genome (Künstner et al. 2016) and scaled 100-fold over three raw mutation rates ($4.89e^{-5}$, $4.89e^{-6}$, $4.89e^{-7}$) to ensure robustness of results across different, realistic values of θ ($4N_e\mu$). SLiM employs a classic Wright-Fisher model to simulate populations, in which a population of diploid hermaphrodites proceeds through generations such that an individual's contribution towards the next generation is proportional to its relative fitness.

Demographic treatments included varying degrees of population contraction relative to the ancestral population size (N = 1000) (0.01, 0.1, 0.5, 1.0), migration as a proportion of individuals exchanged between populations (0.0, 0.002) and founding bottlenecks as a number of individuals (2, 20, 100, 1000). For example, a demographic history with bottleneck = 2, contraction = 0.01, and migration = 0.002 would represent the following scenario: 1) 2 individuals are sampled from the ancestral population of 1000 to form the derived population; 2) At the next generation (and for the remainder of the simulation), the derived population's size reflects the contraction treatment, in this case 10 (0.01 of 1000); 3) The ancestral and derived population experience migration in both directions. Thus, each of our treatments can be thought of as a manipulation of the following: 'Bottlenecks' limit the amount of variation within the ancestral population that is transferred to the derived population; 'Contractions' limit the size of the derived population; and 'Migration' dictates the presence of gene flow between the ancestral and derived (Figure 1). Combining these treatment

levels in a fully factorial experimental design generated a total of 32 different independent demographic histories that all genes experienced.

SLiM runs were performed over simulated genes of varying length (in order to ensure robustness across different genomic elements); with exon number (max = 10), exon lengths, and intron lengths drawn at random from the guppy genome gene annotation file (gff). Genes were also given 1kb flanking regions and a divergence selection coefficient (S_σ) drawn at random.

Selection in our model was represented as the fitness consequences incurred through distance from a phenotypic optimum in a one-dimensional landscape. These phenotypic were maintained across the course of simulations; thus, selection was constant throughout. This setup can be considered as analogous to representing two environments with contrasting optimum trait values for a single trait. Intensity of selection was manipulated by modifying the standard deviation (σ) of the normal distribution curve from which the density distribution was calculated through the "dnorm" function in SLiM. Standard deviations for fitness distributions (S_σ) were 0.1, 0.5, 1.0, 5.0, and 10.0, with 0.1 representing the steepest fitness peak and 10.0 the shallowest. All subsequent analyses' standard deviations were 1/x-transformed so that $S_{1/\sigma}$ = 10 represented the strongest selection treatment (Figure 1A). Phenotypes were calculated per individual as the additive phenotypic effects of exonic non-synonymous mutations, which appeared at a rate of 7/3 relative to synonymous mutations (assuming that most mutations in the third base of a codon do not alter the amino acid). Effect sizes for mutations with phenotypic effects were drawn from a Gaussian distribution with mean = 0 and σ = 1. The remaining synonymous exon mutations, and those in introns and flanks, had no effect on fitness. Recombination rate was set at a constant of $1e^{-8}$ across all genomic elements.

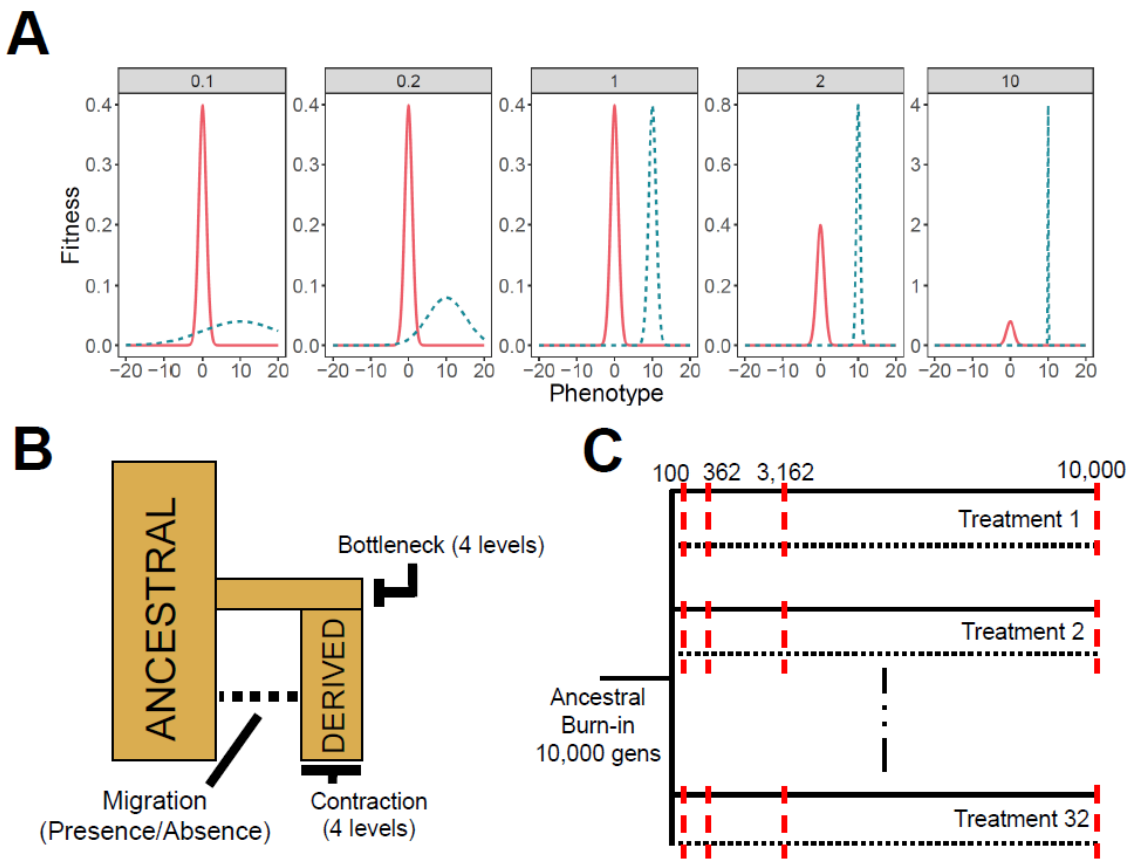

Figure 1: Experimental design for simulations. **A)** Selection treatments assigned to genes, illustrated as the relationship between relative fitness of an individual and its phenotype. Each facet represents a different fitness landscape modified through editing the standard deviation of the normal distribution of fitness consequences in the derived (blue, dashed) population. Facet labels constitute 1/standard deviation applied. The x-axes represent phenotype as calculated through non-synonymous mutations and the y-axes represent relative fitness of individuals. The phenotype of the ancestral population (red, solid) is the same in all treatments (mean = 0, sd = 1), whilst derived populations have a diverged phenotypic optimum (mean = 10, sd = S_σ). **B)** Demographic representation of treatment factors. **C)** Representation of simulation timeline for treatments, illustrating that all treatments share an ancestral burn-in population before splitting into 32 replicated "ancestral" (solid) and "derived" (dotted) population pairs. Red, dashed lines denote sampling generations at which F_{ST}, D_{XY} and $\Delta\pi$ are calculated and averaged across the preceding 50 generations. The purpose of this averaging was to achieve a general sense of population differentiation at four discrete sampling points, such that values represent stable patterns rather than more stochastic generation to generation variation.

For each simulation, populations were seeded with 1000 individuals and allowed to proceed for a burn-in period of 5*2N (10,000) generations to develop stable levels of natural variation under stabilising selection. During this period, burn-in populations evolved towards a phenotypic optimum of 0, defined as a normal distribution with mean = 0 and σ = 1.0 (Figure 1A). Burn-in populations were then subjected to each demographic treatment to simulate the founding of multiple

populations from a shared ancestral state. During this 'divergence period', the ancestral population continued to evolve around a phenotypic optimum of 0, whilst the derived population's phenotypic optimum was centred around 10, with fitness consequences defined according to the selection treatment. Individuals also experienced fitness costs associated with phenotypic proximity to other individuals within the population as a proxy for competition and to ensure a realistic amount of phenotypic variation persisted within populations. Fitness costs due to competitive proximity were drawn from a Gaussian distribution with mean = 1, σ = 0.4 and occurred reciprocally between local individuals with phenotypes with a difference of ≥ 1.2 (3 * 0.4).

Each gene was then run across all demographic treatment levels in parallel. Simulations were sampled at 0.05*2N (100), 0.158*2N (316), 1.581*2N (3162) and 5*2N (10,000) generations (Figure 1C). At each sampling point, F_{ST} (calculated according to Hudson, Slatkin and Maddison (1992)), D_{XY} (calculated according to Nei (1987)), and mean heterozygosity (π) of each population were calculated as the average across the preceding 50 generations. Averaging was performed such that measures would not be dramatically biased by events occurring within individual generations. The change in mean heterozygosity between the ancestral and derived population (Δπ) was calculated at each sampling point as the $\log_{10}$-transformed ratio of ancestral π to derived π, such that reduced diversity in the derived population increases the value of Δπ. Statistics were calculated over all sites of the gene.

$$F_{ST} = 1 - \frac{H_S}{H_T}$$

$$D_{XY} = \sum_{ij} x_i y_j d_{ij}$$

$$\Delta\pi = \log_{10} \frac{\pi_{ancestral}}{\pi_{diverged}}$$

228

229 In total, ~1,000 genes were simulated across 32 demographic treatments across three mutation
230 rates. In addition, a fourth set of simulations was performed in which the phenotypic optimum for
231 both populations was kept at 0 ($Pheno_{Null}$). This dataset was used to assess whether patterns
232 associated with selection were driven by phenotypic divergence or variable stabilising selection
233 within populations. Sets of genes were generated randomly for each mutation rate. Outliers for gene
234 sets were taken as upper quantiles (either the upper 1% or 5% depending on analysis).

235

236 All data analysis was performed in R (3.5) (R Core Team 2016). To assess relationships between
237 divergence measures and selection, we subsetted our data within each treatment level for each
238 demographic factor (e.g. subsetting contraction groups of 0.01, 0.1, 0.5 and 1.0) and calculated
239 Pearson's correlation coefficients across each subset. For example, for each population contraction
240 level, 8 treatment groups were subsetted together and averaged over. To compare the proportions
241 of outliers of different selection treatments from random expectations we applied chi-squared tests.
242 When assessing overlap of outliers across treatment groups, we downsampled the dataset such that
243 we retained a random 10% of $S_{1/\sigma}$ = 10 and a random 50% of $S_{1/\sigma}$ = 2 genes per iteration. This step
244 was taken to try to reduce the proportions of genes under the strongest selection treatment within
245 the dataset from their initial ~20% proportions to a more realistic ~2% and ~10% respectively.
246 Downsampling was bootstrapped 100 times, with the proportional overlap between demographic
247 treatments recorded per iteration and averaged across all runs. Following downsampling, we
248 examined proportional overlap of the two 95% outlier sets with an expected proportional overlap of
249 0.05 under a null hypothesis. To examine how simulated gene features influenced patterns of
250 genetic divergence, we used random forests of 100 trees with the R package *randomForest* and
251 calculated relative importance of each feature. Gene features considered were: within-population π
252 of the derived population, number of exons (Exon N), gene size, the proportion of gene that is

253 coding (Target Proportion). Also included were the selection ($S_{1/\sigma}$) imposed on the gene and the
254 generation at which the optimum phenotype was reached (Pheno Gen).
255

256 **RESULTS**

257 ***Demography modifies measures of divergence***

258 Bottlenecks, simulated by modifying the number of founding individuals, had little effect on F_{ST} and
259 D_{XY}, producing minimal variance between treatments over all sampling points (Figure 2A).
260 Conversely, there was a relatively pronounced effect on Δπ that persisted for several hundred
261 generations. Genes that were simulated with the most extreme bottlenecks displayed inflated
262 measures of divergence in Δπ because of a sharp reduction in π in the derived population.
263

264 Contractions, simulated by modifying the stable number of individuals within the derived
265 population, had pronounced and variable effects on all measures. F_{ST} was negatively associated with
266 contractions, with values increasing with a reduction in derived population size. This effect was
267 generally consistent through time, however early sampling points exhibited the largest variance
268 between the most extreme contraction treatments, whereas later on the strongest variance was
269 observed between weaker treatments (Figure 2B). A similar effect was observed for Δπ; however,
270 this measure was particularly susceptible to inflation under the most extreme contraction scenarios,
271 with substantially more elevated values observed between contraction proportions of 0.01 and 0.1
272 compared with 0.1 and either 0.5 or 1.0. Whilst this pattern was broadly consistent through
273 sampling points, variance among the different contraction levels for Δπ was most pronounced at the
274 initial (100 generations) and final sampling points (10,000 generations). Such observations are
275 unsurprising given that both F_{ST} and Δπ increase with processes that reduce within-population
276 genetic variance. D_{XY} was largely unaffected by contraction treatments up until the third sampling
277 point (3162 generations). Following this, in contrast to F_{ST} and Δπ, D_{XY} decreased when derived
278 population sizes were reduced (Figure 2B).

The inclusion of migration reduced absolute values of all measures of divergence. For F_{ST} and D_{XY}, variance between migration treatments increased across the simulation, however D_{XY} was generally more consistent across migration treatments. Δπ was also reduced in the presence of migration, although the effect of migration on Δπ was generally consistent across all generations and did not increase through time, as was the case for F_{ST} and D_{XY}.

Unsurprisingly, there was a consistent trend regardless of demographic treatment for F_{ST} and D_{XY} to increase generally over time, however this was not the case for Δπ, which was generally constant through time whilst marginally declining. The majority of these patterns were apparent in the $Pheno_{Null}$ simulations (Figure S3), however genetic differentiation for all measures was reduced by the removal of derived phenotypes when gene flow was present.

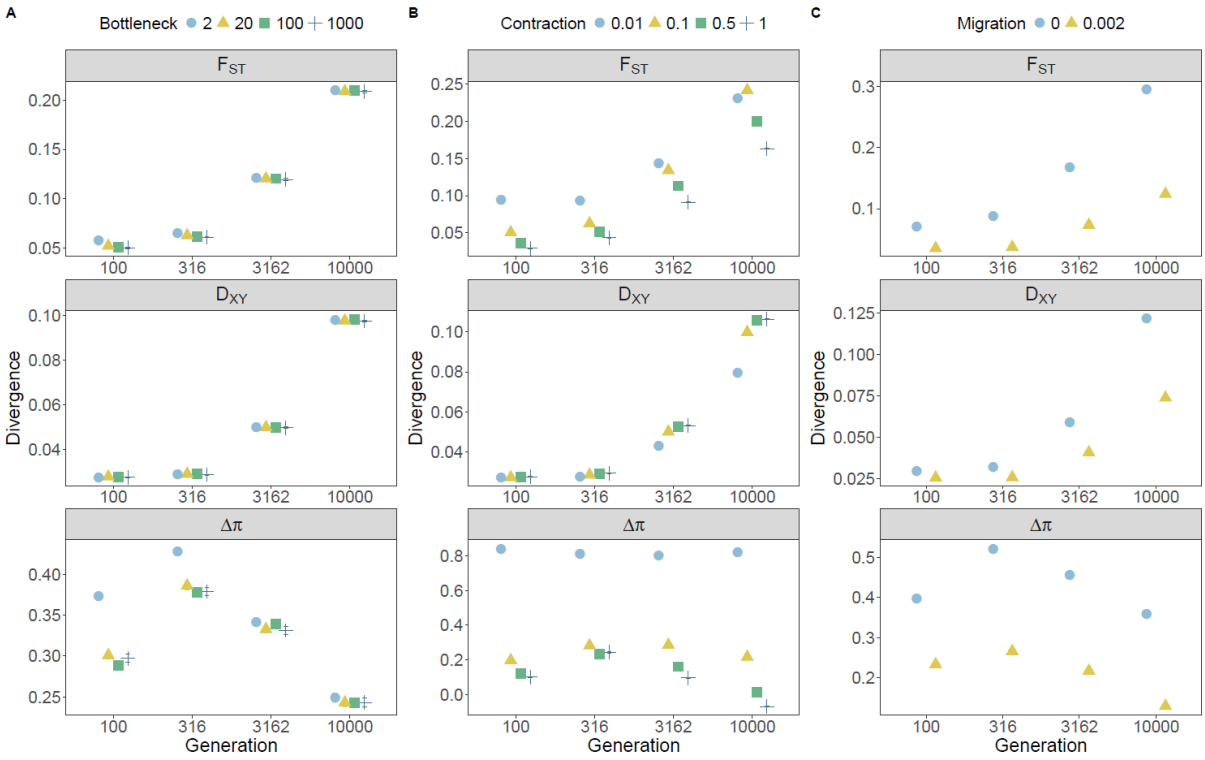

Figure 2: Effects of demographic treatments on measures of genetic divergence across all sampling generations. Point colour and shape denote treatment groups for bottlenecks (**A**), population

contraction (**B**) and migration (**C**). Each point represents values of divergence averaged across all genes, and all other treatment levels beyond those of the treatment group of interest.

Demography moderates the association between measures of divergence and selection

Bottlenecks had a minimal effect on the correlations observed between strength of selection and measures of divergence with minimal variance observed between bottleneck treatments in all comparisons (Figure 3A).

Population contractions had substantial effects on relative (F_{ST} and $\Delta\pi$) measures and marginal effects on absolute (D_{XY}) measures of divergence (Figure 3B). Relative measures in smaller derived populations (0.01, 0.1) were less correlated, and indeed largely uncorrelated with selection than those observed in treatments with minimal (0.5) or no (1.0) contractions (Table 1). For F_{ST} the variance in correlation coefficients between contraction treatment groups generally increased across time whereas variance between treatments for $\Delta\pi$ peaked at 3,162 generations. For D_{XY}, correlations with selection were broadly consistent across contraction treatments up until 10,000 generations, in which associations with selection were reduced by stronger contractions (Figure 3B).

Expectedly, the absence of migration largely precluded the ability of measures of divergence to predict strength of selection, with notable variance observed between no migration (M= 0.0) and minimal migration (M = 0.002) treatments by 3,162 generations for all divergence measures (Figure 3C) (Table 1). Both F_{ST} and D_{XY} variance between migration treatments was greatest at the 10,000 generations sampling point, whereas similar variance was observed between migration treatments for $\Delta\pi$ at both 3,162 and 10,000 generations. This observation again highlights the significance of temporal differences between measures.

$Pheno_{Null}$ simulations resulted in weaker correlations with $S_{1/\sigma}$ for F_{ST}, whilst D_{XY} became negatively correlated with $S_{1/\sigma}$ over time (Figure S6B) rather than positively correlated. $\Delta\pi$, however, became

323 more positively correlated with $S_{1/\sigma}$ over time for all demographic treatments (Figure S6C). These
324 results demonstrate that increases in F_{ST} are partly dependent on divergent selection, increases in
325 D_{XY} are exclusively dependent on divergent selection (and adaptive substitutions), and $\Delta\pi$ increases
326 are less indicative of divergent selection and more of stabilising selection and within-population
327 variance.

Table 1: Correlation coefficients between measures of divergence and selection, separated according to individual treatment levels across all sampling generations. For each measure, the largest correlation coefficient is highlighted in bold.

Treatment	Level	F_{ST}				D_{XY}				$\Delta\pi$			
		100	316	3,162	10,000	100	316	3,162	10,000	100	316	3,162	10,000
Bottleneck	2	-0.006	0.040	0.129	0.169	0.016	0.019	0.033	0.095	0.006	0.083	0.148	0.072
	20	0.013	0.051	0.140	0.165	0.031	0.026	0.043	0.093	0.010	0.097	0.137	0.071
	100	0.010	0.043	0.147	0.177	0.024	0.022	0.050	0.102	0.009	0.084	0.138	0.074
	1000	0.007	0.040	0.133	0.165	0.020	0.021	0.041	0.097	0.019	0.079	0.149	0.070
Contraction	0.01	0.015	0.019	0.033	0.053	0.030	0.034	0.039	0.052	0.004	0.008	0.006	-0.024
	0.1	0.000	-0.001	0.087	0.157	0.015	0.006	0.020	0.110	0.034	0.040	0.189	0.122
	0.5	0.009	0.078	0.269	0.304	0.025	0.027	0.063	0.147	0.013	0.168	0.378	0.214
	1	0.000	0.136	0.306	0.355	0.022	0.020	0.051	0.138	0.031	0.260	**0.406**	0.245
Migration	0	0.003	0.029	0.074	0.062	0.020	0.014	0.006	-0.015	0.018	0.081	0.094	0.007
	0.002	0.012	0.101	0.335	**0.429**	0.027	0.033	0.092	**0.223**	0.003	0.105	0.241	0.191

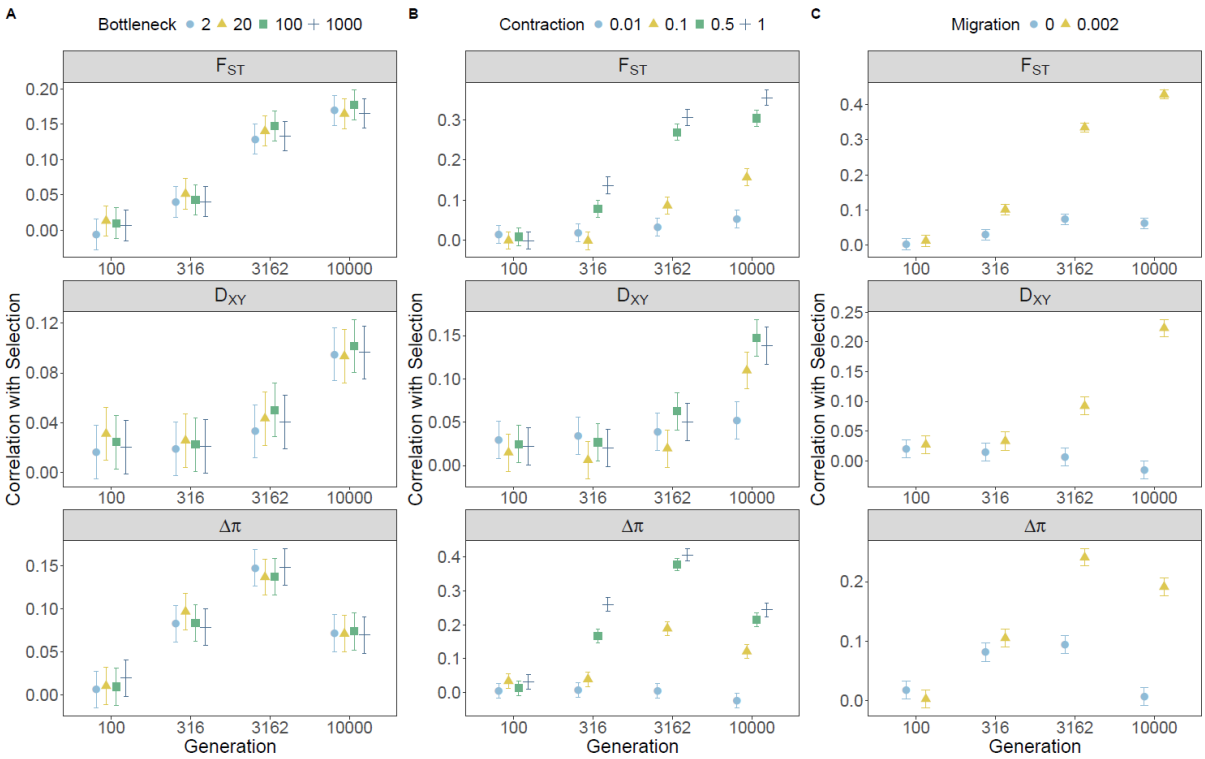

Figure 3: Effects of demographic treatments on the relationship between selection ($S_{1/\sigma}$) and measures of genetic divergence across all sampling generations. Point colour and shape denote treatment groups for bottlenecks (**A**), population contraction (**B**) and migration (**C**). Each point represents correlation coefficients averaged across all genes, and all other treatment levels beyond those of the treatment group of interest.

By examining the correlation coefficients of all 32 unique demographic histories, we can investigate the combined effect of migration and contraction and directly compare effectiveness of individual measures across time (Figure 4). F_{ST} consistently outperforms D_{XY} in terms of associating with $S_{1/\sigma}$ under most demographic treatments, particularly when contractions are weak and migration is present. By 10,000 generations however, the relative dominance of F_{ST} appears to marginally subside, with the trend through time suggesting a relative improvement in D_{XY} in treatments with gene flow (Figure 4). Δπ is similarly more informative than D_{XY} across sampling generations under most demographic treatments. Interestingly, at sampling points 316 and 3,162 contractions produce the biggest bias towards Δπ, however by 10,000 generations zero migration produces the largest deviance away from the y=x relationship (Figure 4). The resilience of Δπ under no-migration

treatments is also apparent in F_{ST} - Δπ comparisons, such that at 316 and 3,162 generations Δπ is more informative than F_{ST} in the absence of migration. By 10,000 generations however F_{ST} outperforms Δπ under most scenarios, indicative of the decaying association between Δπ and $S_{1/\sigma}$ discussed previously.

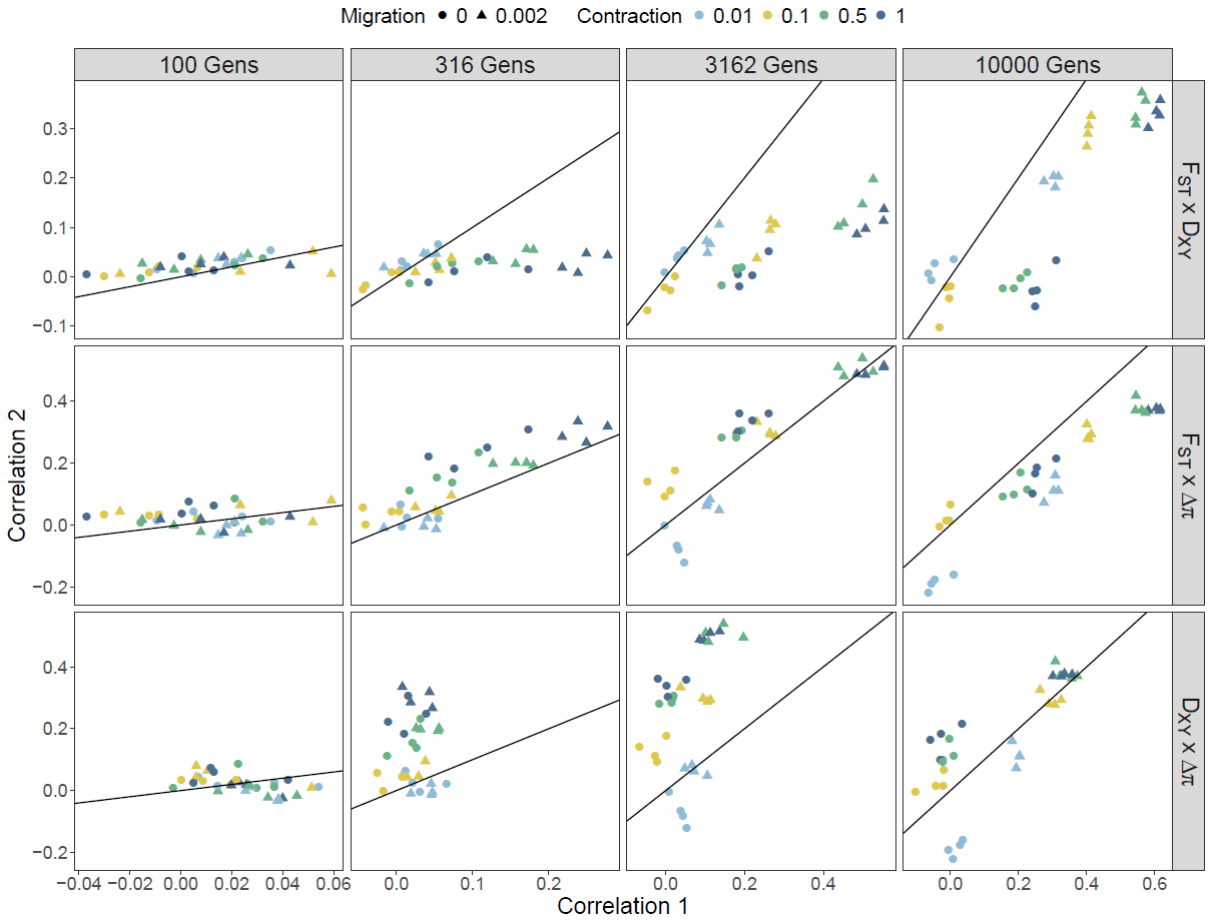

Figure 4: Pairwise comparisons between the correlation coefficients of $S_{1/\sigma}$ with F_{ST}, D_{XY} and Δπ across all sampling points. Each data point represents a unique demographic treatment with points coloured according to contraction level and shaped according to migration level. Correlation 1 refers to the first measure listed in the comparison and Correlation 2 to the second. The y=x line is plotted within each facet to illustrate biases towards one measure. Points below the line are biased towards Correlation 1, whilst points above the line are biased towards Correlation 2.

Demography moderates the tail end of divergence distributions

Ultimately, outlier tests are concerned with the extreme tail end of the distribution, and as such correlation coefficients across the whole dataset may fail to reflect consistent outliers. To counteract

364 this, we asked at each sampling point what the relative proportions of outliers above the 99th
365 quantile were per $S_{1/\sigma}$ treatment, and how these proportions varied according to individual
366 demographic treatments (Figure 5) and collectively (Figure 6) after 10,000 generations (results for
367 other sampling points and mutation rates Figure S10-20). Each simulated 'genome' consisted of
368 ~1000 genes evenly distributed across the 5 treatment levels of $S_{1/\sigma}$, meaning expected null
369 proportions of each $S_{1/\sigma}$ value are 0.2. This analysis assumes that if measures of differentiation are
370 good at highlighting genes under the strongest selection coefficients, genes within the upper 1% of
371 the distribution should be disproportionately associated with stronger selection ($S_{1/\sigma}$ = 2 or 10).
372
373 We found that migration and contraction treatments influenced the relative proportions of $S_{1/\sigma}$
374 treatments in the upper F_{ST} quantile. Under the weakest contraction treatment, the proportion of
375 $S_{1/\sigma}$ = 10 outliers increased from 0.19 (contraction = 0.01) to 0.33 (contraction = 1.0) (null
376 expectation of 0.20 given 5 selection treatments). Migration also dramatically increased the
377 proportion of genes under the strongest selection from 0.19 to 0.63. The extreme effect of migration
378 is evident when combining contraction and migration treatments (Figure 6). Whilst weaker
379 contractions yield a greater proportion of strong selection outliers (moving left to right in the figure),
380 migration is the dominant factor given that the treatment with the most extreme contraction (0.01)
381 and migration present has a stronger selection bias than the treatment with no contraction and
382 migration absent (Figure 6).
383
384 The upper quantile of D_{XY} outliers was unaffected by bottlenecks and contractions but was heavily
385 biased towards $S_{1/\sigma}$ = 10 genes in the presence of migration (Figure 5). Proportions of $S_{1/\sigma}$ = 10
386 outliers doubled in the presence of migration from 0.27 to 0.54. Proportions across the 16 unique
387 contraction/migration patterns produced results analogous to F_{ST}, with migration the dominant
388 factor in biasing proportions towards strong selection genes (Figure 6).
389

390 Except for the most extreme population contraction, Δπ outlier proportions were strongly indicative
391 of genes under strong selection, yielding outlier proportions for $S_{1/\sigma}$ = 10 genes of 0.72 (contraction =
392 0.1), 0.61 (contraction = 0.5) and 0.58 (contraction = 1.0) at sampling generation 10,000. Effects of
393 contractions were visible for Δπ outliers by generation 100 at the most extreme treatments (Figure
394 S10) and clear across all 3 weaker contraction treatments by 3162 generations (Figure S12).
395 Migration had a more modest effect on outlier proportions relative to contractions, increasing the
396 proportion of genes under the strongest selection from 0.16 to 0.41 at sampling generation 10,000.
397 The effect of migration was not as persistent through time however, emerging only by generation
398 3,162 with $S_{1/\sigma}$ =10 proportions of 0.18 (migration = 0) and 0.24 (migration = 0.002). These results
399 are corroborated across the 16 contraction/migration treatments, as proportions acquired from
400 treatments with extreme contractions (contraction = 0.01) are relatively uniform with only a small
401 biasing effect of migration (Figure 6). For all other treatments, outliers are heavily biased towards
402 strong selection genes.
403
404 Taken together, these results highlight that outlier detection using F_{ST} and D_{XY} are heavily dependent
405 on the presence of migration, and to a lesser extent the absence or minimising of population
406 contractions for F_{ST}. Δπ outliers appear rapidly and are generally robust in relation to population
407 contractions up until the most extreme treatment. Whilst migration also influences Δπ, the biasing
408 of outliers towards genes under the strongest selection treatment is less extreme.
409
410 Removing divergent selection ($Pheno_{Null}$) moderated outlier proportions in line with expectations
411 given correlations with $S_{1/\sigma}$ (Figures S21-24). Interestingly, over time both F_{ST} and Δπ outliers were
412 skewed further towards stronger selection treatments when migration was absent (stronger
413 patterns of reduced within-population variance) in comparison to simulations with phenotypic
414 divergence, whilst D_{XY} outliers became biased towards weaker selection treatments in the presence
415 of gene flow.

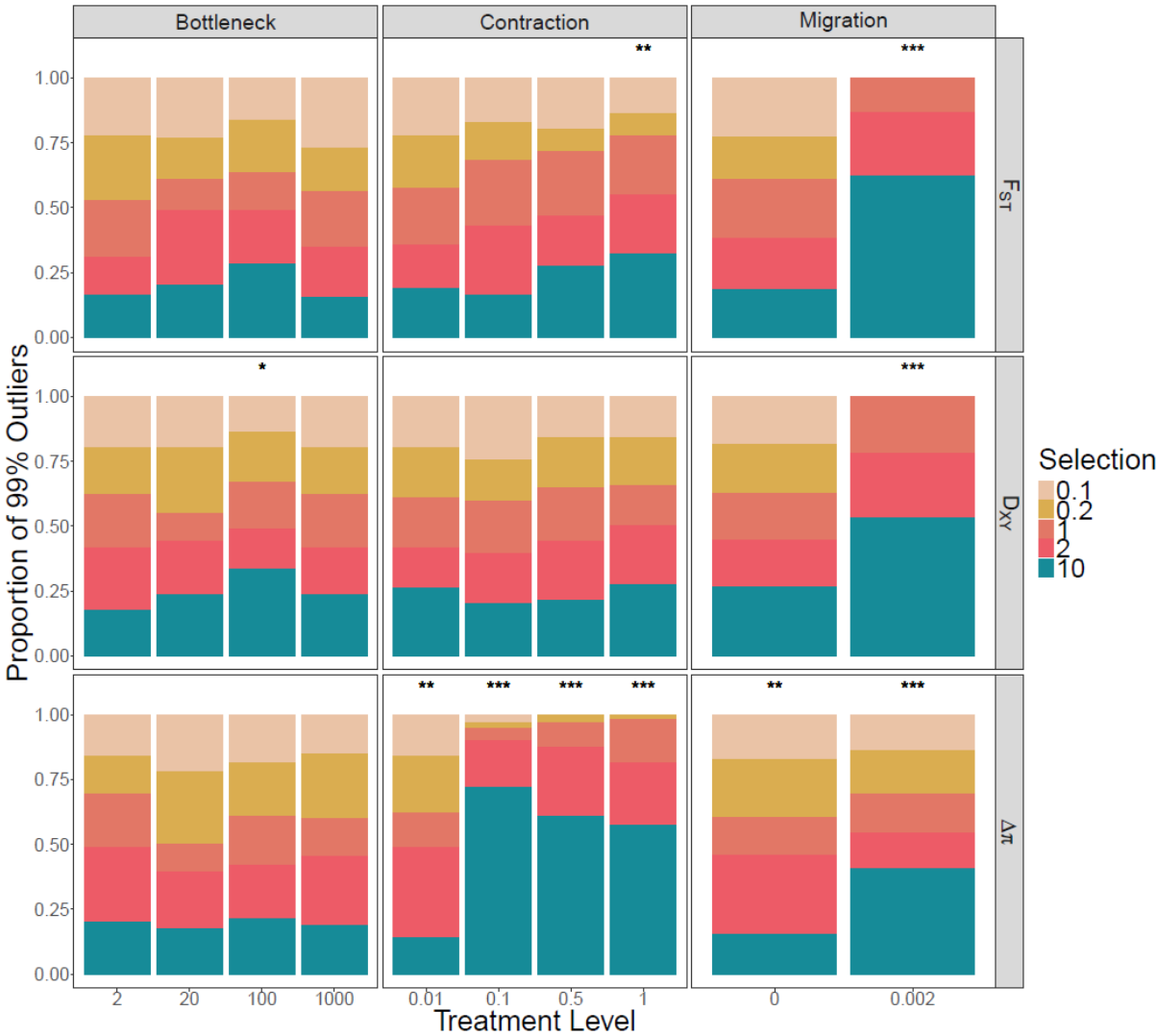

Figure 5: The effect of demography on representation of genes under different selection treatments as a proportion of outliers above the 99% quantile. Colour of bars within stacks denotes different selection treatments. Vertical facets represent different demographic treatments and horizontal facets show effect per measure of genetic divergence. Results shown are from sampling generation 10,000. Results for other sampling points can be found in Figures S10-12.

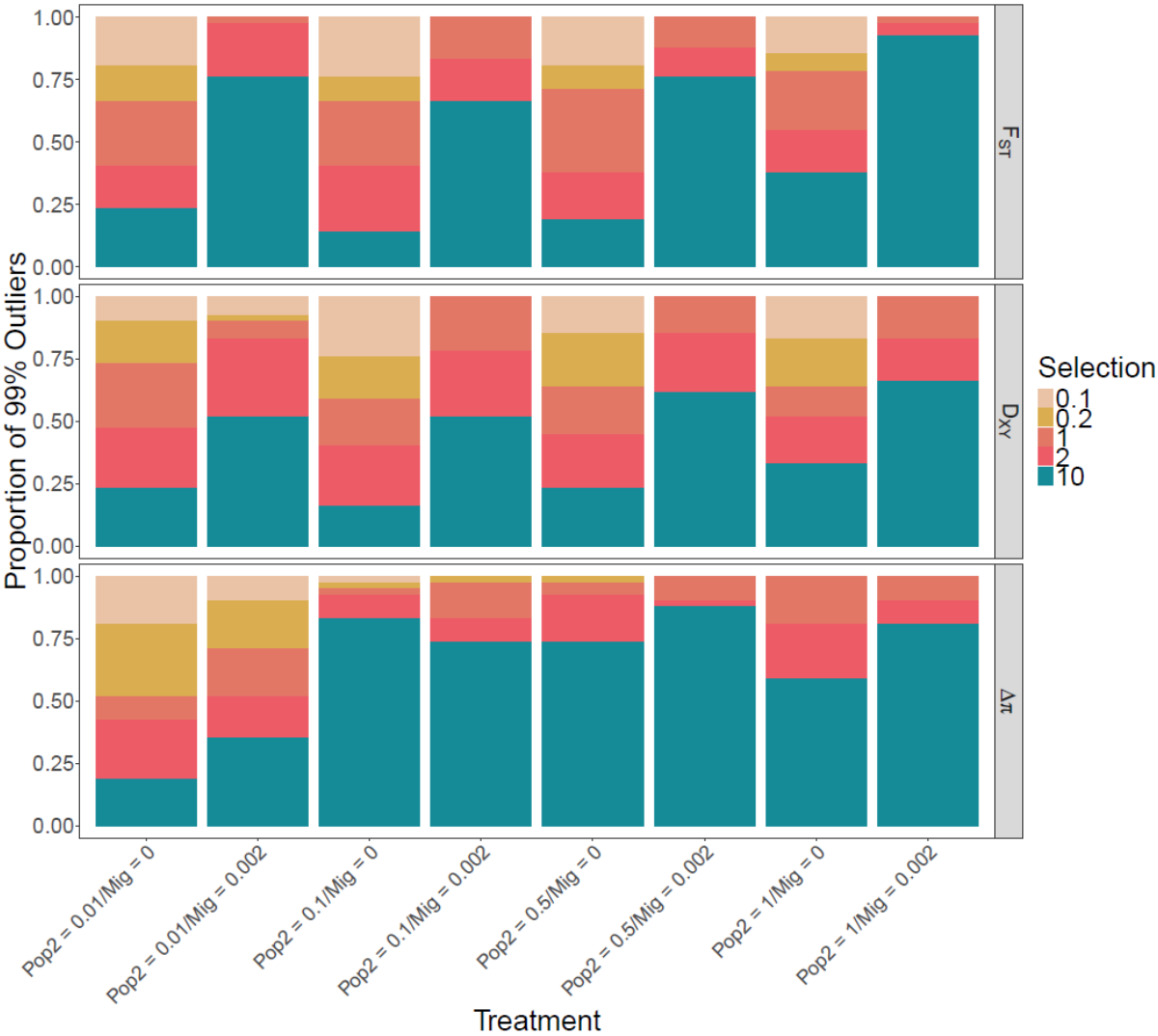

Figure 6: The combined effects of contractions (proportion relative to ancestral population size = 1,000) and migration in 16 unique demographic treatments (all combinations of contraction and migration treatments) on representation of genes under different selection treatments as a proportion of outliers above the 99% quantile. Full information regarding demographic treatment is described on the x-axis. Colour of bars within stacks denotes different selection treatments. Horizontal facets show effect per measure of genetic divergence. Results shown are from sampling generation 10,000.

Demography moderates associations between measures of divergence

Given F_{ST}, D_{XY} and Δπ are all measures of population genetic divergence, there is an assumption that positive associations should exist between them. We employed the same analysis as above for correlations with $S_{1/\sigma}$, but instead examined correlations between individual measures. Bottlenecks had minor effects on the correlations observed between measures of divergence (Figure 7A). Expectedly, effects were strongest at the first sampling point, with the most extreme bottlenecks

increasing the correlation coefficients between F_{ST} & D_{XY} and D_{XY} & $\Delta\pi$ and decreasing the correlation coefficient between F_{ST} & $\Delta\pi$. Variation in correlation coefficients was generally minimal however compared to the effects of other demographic treatments.

F_{ST} & D_{XY} were positively associated irrespective of demography, but smaller population sizes generally increased the association (Figure 7B) with variance between treatments generally decreasing through time. Similarly, F_{ST} & $\Delta\pi$ were typically positively associated, however contractions had a more pronounced effect on the association. Extreme contractions increased the association between F_{ST} & $\Delta\pi$, with contraction-induced variance increasing across sampling points (Figure 7B). At 10,000 generations when variance was largest, F_{ST} & $\Delta\pi$ ranged from being generally well associated under extreme contractions ($R^2 = 0.60$) to unassociated when populations were equally sized ($R^2 = 0.08$). D_{XY} & $\Delta\pi$ correlations were initially weak early in the simulations, however increased generally over time, with populations contractions similarly inflating associations between measures. Interestingly, associations with D_{XY} and both F_{ST} and $\Delta\pi$ appear to peak for weak contraction treatments at 3,162 generations, suggesting a probable temporal factor within the simulations driving positive associations with D_{XY} at some point between 316 and 3,162 generations.

Migration induced substantial variance between correlations of divergence measures, with effects dependent on sampling point (Figure 7C). In the absence of migration, F_{ST} & D_{XY} were strongly associated at 100 and 316 generations, whilst associations were weaker in migration-present treatments. However, by 3,162 generations F_{ST} & D_{XY} associations increased substantially in the presence of migration, and by 10,000 generations the effect of migration had reversed entirely. F_{ST} & $\Delta\pi$ associations were strong initially, but migration weakened the association over time so migration-present associations ranged from $R^2 = 0.67$ (100 gens) to $R^2 = 0.14$ (10,000 gens). In contrast, in the absence of migration positive associations between F_{ST} & $\Delta\pi$ were relatively stable ($R^2 = 0.55 – 0.51$ over the same time period). Interestingly, without migration D_{XY} & $\Delta\pi$ were largely

unassociated, but migration-present treatments induced a negative association between D_{XY} & $\Delta\pi$ over time from R^2 = -0.04 (100 gens) to R^2 = -0.28 (1000 gens).

Contextualised by our previous demonstrations of associations with selection, these results highlight that selection can induce positive associations between measures. Namely, F_{ST} & D_{XY} become strongly associated by generation 10,000 with migration present, which we know is when F_{ST} and D_{XY} are most strongly linked to selection. Crucially, however this relationship only emerges after several thousand generations, before which we observe positive associations in migration-absent treatments when associations with selection are weak for both measures. Extreme contractions also induce positive associations between F_{ST} & D_{XY} that are likely indicative of false-positives given the poor association with selection in these treatments. These results highlight that positive associations between statistics are also achievable under neutrality. It is also interesting to note the decay of associations with $\Delta\pi$ and both F_{ST} and D_{XY} under demographies in which selection should be effective. These observations are most likely driven by the substitution rate. Substitution rates are increased under stronger selection (Lanfear, Kokko, and Eyre-Walker 2014) and drive increased F_{ST} and D_{XY} but not $\Delta\pi$.

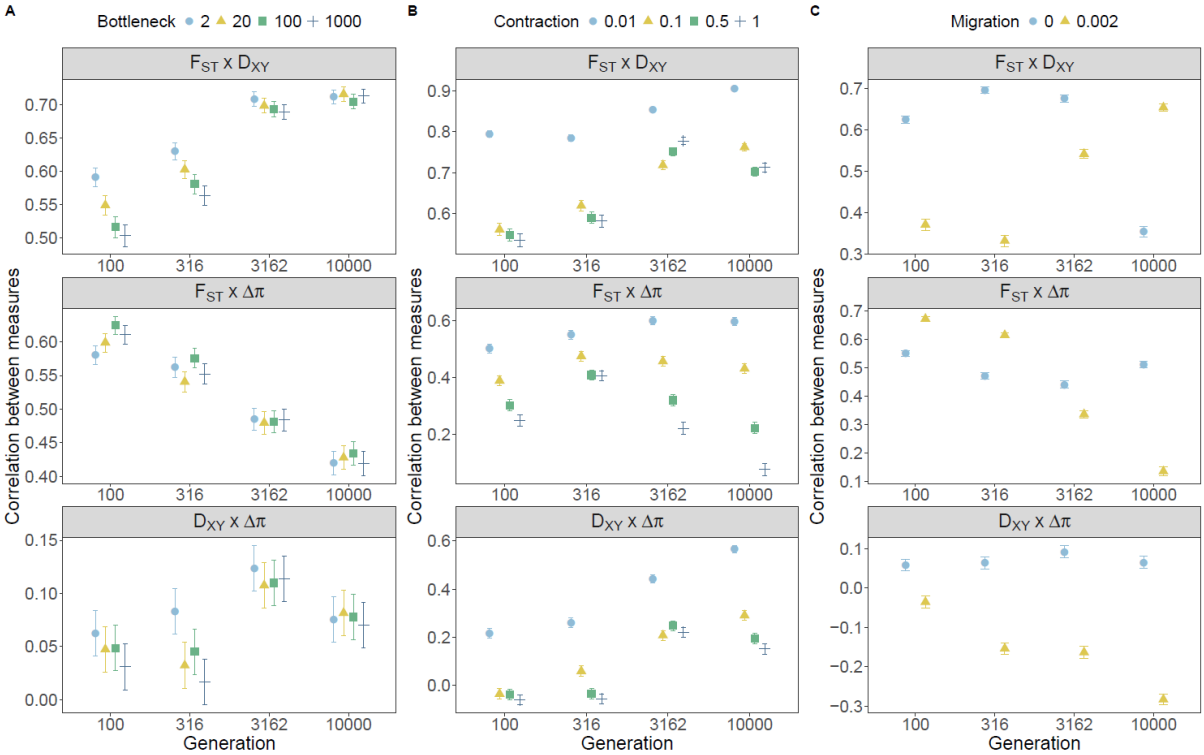

Figure 7: Effects of demographic treatments on the relationship between measures of genetic divergence across all sampling generations. Point colour and shape denote treatment groups for bottlenecks (**A**), population contraction (**B**) and migration (**C**). Each point represents correlation coefficients averaged across all genes, and all other treatment levels beyond those of the treatment group of interest.

In $Pheno_{Null}$ simulations the influence of bottleneck treatments and contraction treatments was broadly unchanged, however the notable emergence of a correlation between measures of F_{ST} and D_{XY} in simulations with gene flow failed to emerge by generation 10,000 (Figure S27). This demonstrates that this positive correlation is driven by divergently adaptive allele frequency shifts and substitutions. In addition, the decay of the relationship between F_{ST} and $\Delta\pi$ in the face of migration did not occur without divergent selection. Rather, this relationship remained strongly positive ($R^2 > 0.5$) up to 10,000 generations regardless of migration treatment (Figure S27C).

To examine the overlap of outliers across measures of divergence, we again sub-divided our treatments into treatment-specific levels and compared the three-way and all pairwise upper 95% quantile gene overlaps across measures of divergence at sampling point 10,000. Three-way

agreement was generally greater than expected by-chance under most conditions, with the largest overlap observed under extreme contractions (Figure 8). Our results demonstrate large overlap between F_{ST} and D_{XY} outliers across all treatments, which was marginally increased by contraction treatments that favour the effectiveness of selection. This association suggests stronger selection may improve the positive association between F_{ST} and D_{XY}, however a causative effect is not clear. Whilst overlap increased by selection-favouring conditions, all treatment levels exhibited above-expected levels of F_{ST} - D_{XY} overlap, suggesting a strong potential for false-positives. Overlap between $\Delta\pi$ and both F_{ST} and D_{XY} was substantially less but had contrary associations with contractions. More extreme contractions tended to increase the overlap between $\Delta\pi$ and F_{ST} outliers, suggesting a shared risk of false-positives. In contrast, $\Delta\pi$ and D_{XY} outlier overlap tended to increase under weaker contractions, suggesting a beneficial effect of combining these measures to remove false-positives. Migration increased outlier overlap between $\Delta\pi$ and both F_{ST} and D_{XY}.

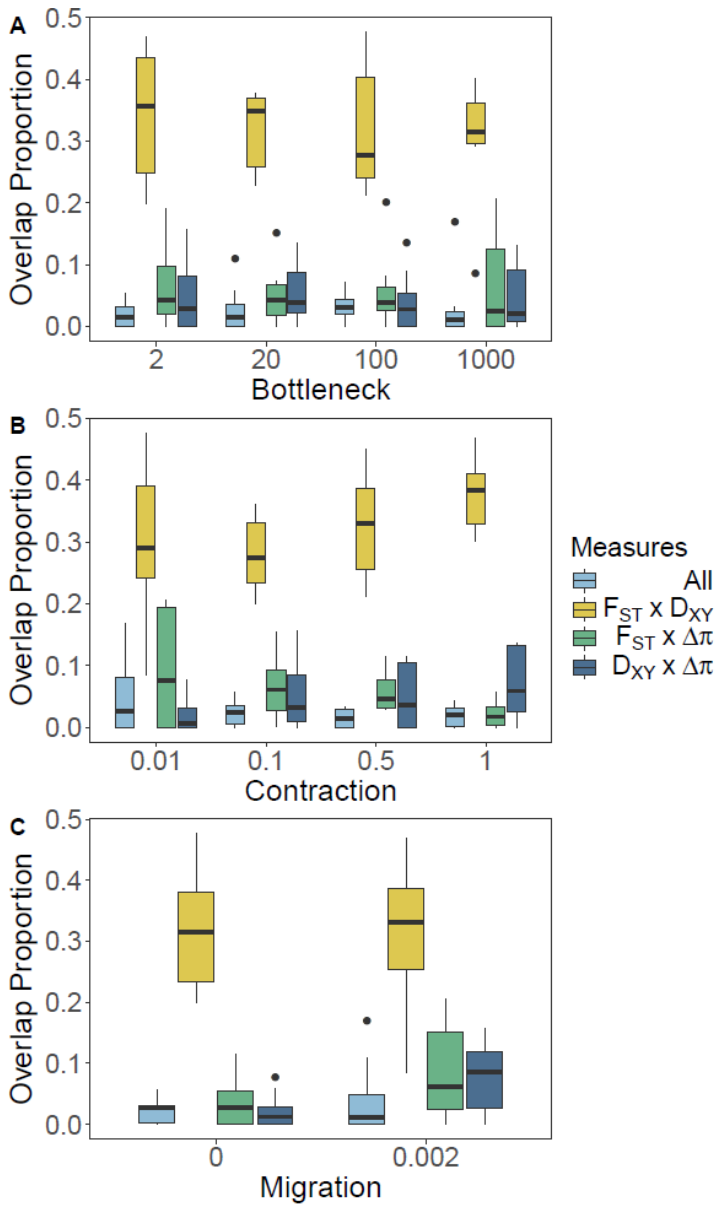

Figure 8: Proportional overlap of 95% outliers across measures of divergence, separated according to treatment levels of bottlenecks (**A**), contractions (**B**) and migration (**C**).

Demography drives signals of convergence irrespective of selection

By sampling generation 10,000, significant overlap in outlier genes was apparent in F_{ST} results across all demographic treatments that had no migration (Figure 9). Within this cluster, the average overlap of downsampled outliers above the 95% ranged between 0.07 and 0.32, suggesting a 1.4- to 6.4-fold enrichment above the null expectation (0.05). These minimum and maximum values were observed between extremes of contraction (minimum) or when contractions were absent (maximum). Given the poor association between F_{ST} and selection in the absence of migration, convergent F_{ST} outliers

522 here are likely driven by neutral processes and would constitute a false-positive signal. Outside of
523 this cluster, there was also reasonable clustering of results across treatments that included
524 migration and weaker contractions (0.1, 0.5 or 1.0). Here, overlap of outliers ranged between 0.08
525 and 0.26, representing a 1.6- to 5.2-fold enrichment above the expected. In the presence of
526 migration and minimal contractions F_{ST} is a good predictor of selection in our dataset, suggesting this
527 convergence may be driven by inferred selection. Beyond these clusters, overlap between
528 treatments of contrasting migration treatment failed to exceed 0.18. Interestingly, the clustering of
529 F_{ST} outliers in the presence of migration was heavily reduced when divergent selection was removed
530 ($Pheno_{Null}$) but clustering of outliers in the absence of migration was still strong (Figure S39).
531

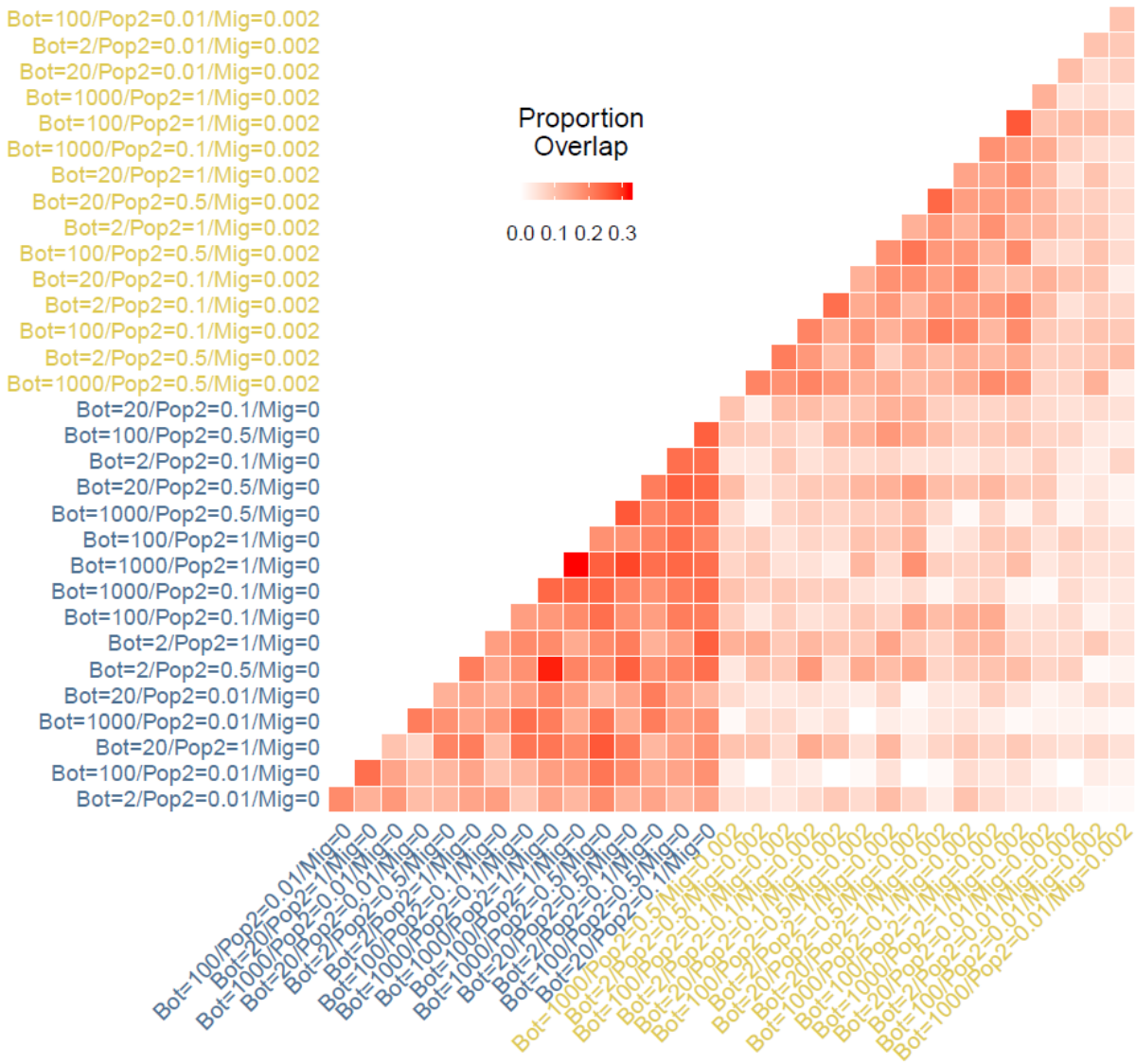

Figure 9: The proportional overlap of outliers above the 95% quantile for F_{ST} values, averaged across 100 randomly downsampled datasets for each of the 32 demographic treatments. Colouring of axis labels denotes migration treatment level. Axis orderings were determined through hierarchical clustering

Clustering of D_{XY} outliers followed a broadly similar pattern to F_{ST} with migration treatments dominating the clustering and inducing proportional overlaps of 0.11-0.34 (2.2- to 6.8-fold enrichment) in the absence of migration (Figure S31). In general, proportional overlap of $\Delta\pi$ outliers was lower than for F_{ST} and D_{XY}, with a maximum overlap observed of 0.21 (fold-enrichment 4.2) (Figure S32). This value was observed in two treatments with minimal contractions (0.5-1.0) and

544 migration present. Generally, however, clustering by demography was weaker in Δπ than for F_{ST} and
545 D_{XY}, corroborating previous results suggesting Δπ may be less susceptible to variation in migration
546 and contraction (with the exception of extreme contraction).
547
548 Because migration was the dominant factor in clustering of treatments with convergent overlaps, we
549 sought to investigate what features of simulated genes drove variance in measures of divergence
550 with treatments separated by migration factor using a Random Forest approach (Figure 10). For F_{ST},
551 we observe that the best predictor of variance in the absence of migration is diversity (π) of the
552 derived population whilst treatments with migration are most strongly influenced by $S_{1/\sigma}$. This is
553 consistent with the notion of F_{ST} relying on within-population genetic variance in the absence of
554 efficient selection. In contrast, D_{XY} in the absence of migration is most strongly driven by gene
555 features, predominantly the proportion of simulated gene that is coding and by the length of the
556 simulated gene. With migration, D_{XY} variance is driven by selection and derived π. Unsurprisingly, Δπ
557 is also strongly influenced by derived π (the denominator in its calculation), but we can see a
558 secondary role for selection in treatments with migration and little influence of gene features.
559
560 Removing divergent selection ($Pheno_{Null}$) did not modify the relative importance of any features of
561 simulations for F_{ST}, but did affect D_{XY} and Δπ (Figure S44). With migration, the relative importance
562 for derived population π in explaining D_{XY} variance increased substantially. For Δπ, selection was as
563 important as derived population π in explaining variance.
564

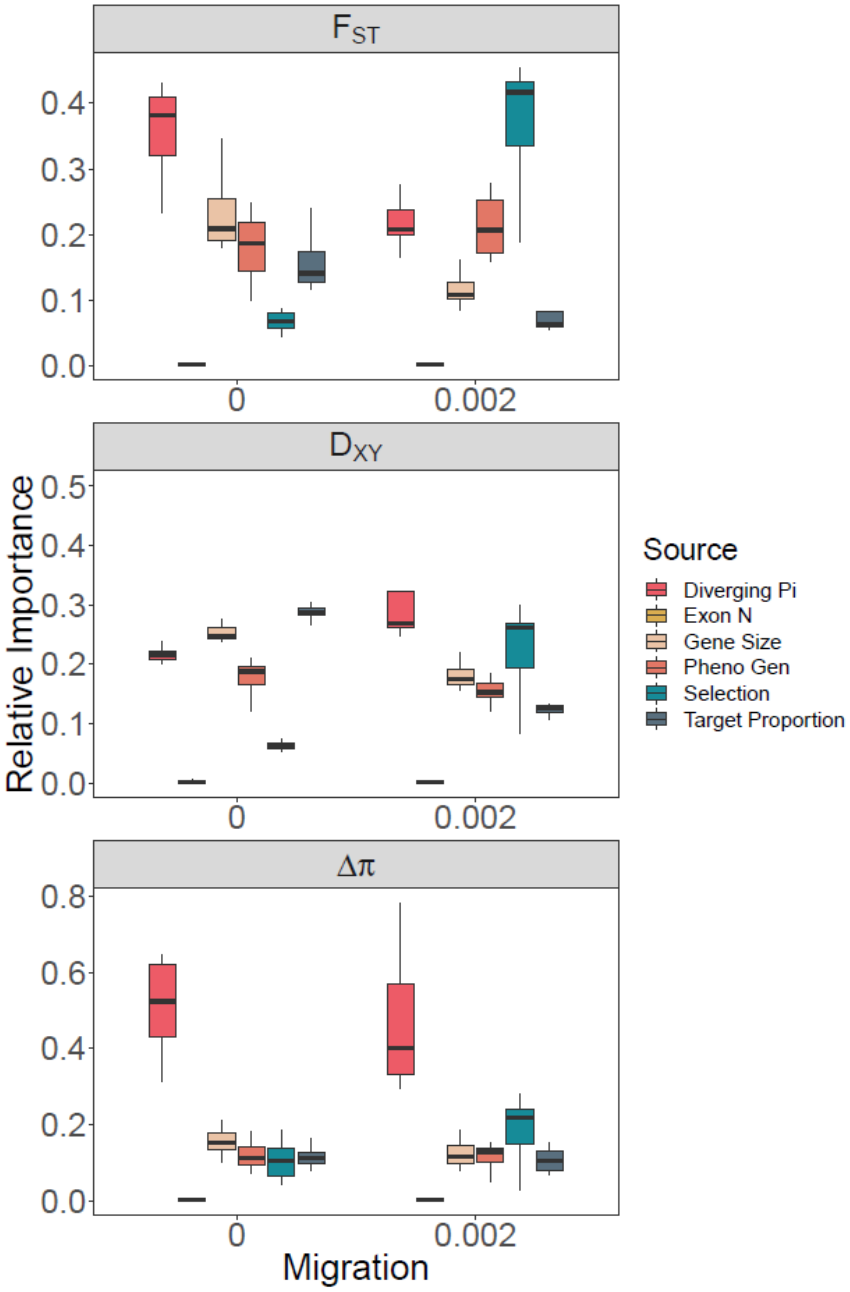

Figure 10: The relative importance of simulated gene features in eliciting variance in measures of genetic variance. Boxplots denote the relative importance of each explanatory variable across 16 individual treatments with/without migration taken at sampling generation 10,000. Variables considered are within-population π of the derived population and gene features including: Number of Exons (Exon N), Gene Size, and the proportion of gene that is coding (Target Proportion). Also included are the selection ($S_{1/\sigma}$) imposed on the gene and the generation at which the optimum phenotype is reached (Pheno Gen).

DISCUSSION

Summary of results

Here, we show that differing demography can have dramatic effects on how measures of population divergence identify regions of the genome under selection. Effects are also strongly time-dependent. Using simulated populations, we have demonstrated the relative influences of founding bottlenecks, population contractions, and migration on three commonly used measures of genetic divergence (F_{ST}, D_{XY}, $\Delta\pi$), whilst demonstrating the relative usefulness of each measure for informing on selection under different demographic conditions.

We find that founding bottlenecks have little effect on population divergence measures, with only minor influence over differences in nucleotide diversity ($\Delta\pi$) shortly after founding. Population contractions (size of diverging population), however, artificially inflate F_{ST} and $\Delta\pi$ but reduce $D_{XY,}$ and can erase the relationship of F_{ST} and D_{XY} with selection under the most extreme contractions. These results can be explained in part due to the relationship between founding population size and relative (F_{ST} and $\Delta\pi$) and absolute (D_{xy}) divergence measures. Relative measures are, in part, driven by intra-population changes in allele frequency, which become exaggerated in smaller populations as a product of drift (B. Charlesworth 2009; Ellegren and Galtier 2016). As a consequence, we observe inflated measures of relative divergence as allele frequencies drift in our smaller derived population.

D_{xy} in contrast increases under weaker population contractions as a product of the relationship between the number of segregating sites and the population-level mutation rate ($4N_e\mu$) (Hartl, Clark, and Clark 1997). D_{xy} is a measure of sequence divergence and is averaged across all sites (although averaging across all segregating sites is possible), which results in higher D_{XY} as segregating sites are introduced into either population at rate of $4N_e\mu$. This relationship with the number of segregating sites can be observed by examining the positive relationships between D_{xy} and π in our derived population (Figure S45). Overall, the relationship between D_{xy} and selection is less affected by population contractions than F_{ST} and $\Delta\pi$, likely due to the lack of allele frequency relevance.

603 Consider for example, two SNPs with minor allele frequencies of 0/0.5 (SNP 1) and 0.5/0.5 (SNP 2) in
604 the ancestral/derived population. Each locus contributes equally towards D_{xy} (SNP 1 = [0 x 0.5] + [1 x
605 0.5] = 0.5; SNP 2 = [0.5 x 0.5] + [0.5 x 0.5] = 0.5), whereas the reduction of within-population
606 variance observed for SNP 1 inflates F_{ST} and $\Delta\pi$.
607
608 Migration, even at the relatively modest rate of 0.2% employed here, reduced absolute values for all
609 measures of divergence substantially. However, whilst absolute values were reduced, their
610 informativeness of selection coefficients increased dramatically (both in terms of their overall
611 relationship with selection and in identifying outliers). Such a result is expected given the role of
612 gene-flow in homogenising neutral loci (reducing measures of divergence), whilst retaining
613 population divergence around adaptive loci (increasing informativeness). The well-known 'genomic
614 islands of divergence' model is often invoked to explain this pattern of heterogenous genomic
615 divergence in studies of speciation-with-gene-flow (Nosil, Funk, and Ortiz-Barrientos 2009; Turner,
616 Hahn, and Nuzhdin 2005).
617
618 Migration also exhibited interesting temporal patterns that are useful for discussing the
619 discrepancies observed between $\Delta\pi$ and both F_{ST} and D_{xy}. Of the measures we considered, $\Delta\pi$
620 uniquely disregards SNP substitutions in its calculation, as fixed substitutions have no influence on
621 intra-population heterozygosity beyond reducing the heterozygosity of proximate SNPs in the
622 aftermath of a hard sweep. This characteristic explains why $\Delta\pi$ responds to selection more rapidly
623 than F_{ST} and D_{xy} and is influenced by migration in a constant manner across time. In addition, the
624 inability of $\Delta\pi$ to account for adaptive substitutions is likely why $\Delta\pi$ informativeness peaked at 3,162
625 generations and decayed by 10,000 generations. In contrast, the contributions of adaptive
626 substitutions to F_{ST} and D_{xy} over time increases their predictive power. This temporal observation
627 has important implications for studies of rapid adaptation on the scale of 10s to 100s of generations,

628 in which our simulations suggest Δπ may be the most informative measure of divergence, whilst D_{xy}
629 was largely uninformative until > 3,162 generations.
630
631 Examining the effects of demographic variation on the associations between measures of divergence
632 themselves is a novel element to this study. The relative weaknesses of individual measures of
633 divergences has prompted a recent movement within the literature to employ multiple measures of
634 divergence to avoid false-positives. Our results provide some support for this strategy, with outlier
635 overlap across measures generally increasing under conditions that favour informativeness of
636 selection, suggesting their combined use may be useful in removing false-positives. Importantly,
637 however, we also find that F_{ST} and D_{xy} (and to a lesser extent F_{ST} and Δπ) have overlapping outliers
638 across demographic treatments with ineffective selection.
639
640 Cruickshank and Hahn (2014) suggested that a disagreement between F_{ST} and D_{xy} outliers in
641 genomic-islands-of-divergence tests highlighted a particular susceptibility of F_{ST} to BGS (although see
642 (Matthey-Doret and Whitlock 2018)). Our results are in line with the notion that D_{xy} is more resistant
643 to false positives due to BGS, given that contraction treatment levels resulted in minimal variance in
644 D_{xy} (assuming treatment levels of population contraction are analogous to different rates of BGS
645 across the genome in exhibiting reduced N_e). We also showed that F_{ST} and D_{XY} are positively
646 correlated under several demographic scenarios, but this strong association only reflects selection
647 when gene-flow is present and only after several thousand generations, consistent with previous
648 simulation work (Ravinet et al. 2017). When migration is absent, and selection ineffective, F_{ST} and D_{XY}
649 are also positively correlated; but this relationship decays over time (Figure 7C), with empirical
650 evidence (see below) suggesting a negative relationship is likely to emerge without gene flow. No-
651 migration treatments are consistent with Isolation-By-Distance demographic histories and, similarly,
652 inter-specific comparisons. Indeed, negative correlations between F_{ST} and D_{XY} within and between
653 clades of birds (Vijay et al. 2017; Irwin et al. 2016), within a radiation of monkeyflowers (Stankowski

654 et al. 2018) and between speciating orca populations (Foote et al. 2016) support a declining
655 relationship between these measures over long periods of time in isolation. This negative
656 association between measures in the absence of effective selection is also predicted, given F_{ST} is
657 negatively associated with within-population π (increases within-population variance), D_{XY} is
658 positively associated with local π (reduces likelihood of common sequence) and π is generally well-
659 conserved in stable populations even across long time periods (Romiguier et al. 2014; Van Doren et
660 al. 2017; Dutoit et al. 2017).
661
662 Similarly, we found F_{ST} – Δπ outliers under most demographic scenarios and overlapping outliers
663 marginally increased under extreme contractions. These results, again, support the notion that using
664 multiple measures may be useful in reducing false-positives under demographic scenarios with
665 effective selection, but there exists a risk of false-positives occurring across populations with
666 ineffective selection.
667
668 In comparing our results to a second dataset that lacked divergent selection ($Pheno_{Null}$), we found
669 consistent support that associations between D_{XY} and selection are strongly dependent on the
670 inclusion of a divergent phenotype; F_{ST} patterns are attainable with stabilising selection (but are
671 weaker than when divergent selection is included); and Δπ patterns are primarily driven by variable
672 stabilising selection and made weaker by the inclusion of phenotypic divergence. These comparisons
673 are useful in highlighting the relative roles of adaptive allele frequency changes and substitutions in
674 driving patterns of genetic differentiation. It is also of interest to note that overlapping outliers are
675 readily attainable across no-migration treatments in $Pheno_{Null}$ simulations, but not for treatments
676 with gene-flow (Figure S39). This confirms that overlapping outliers in no-migration treatments
677 occur due to common neutral processes within genes, whereas overlapping outliers in treatments
678 with migration are driven by the effects of divergent selection. The discrepancies between our

679 $Pheno_{Null}$ and other simulated datasets highlight the necessity in quantifying phenotypic differences
680 or environmental selection pressure when interpreting patterns of variation across the genome.
681
682 ***Detecting genetic convergence***
683 In addition to understanding how outlier detection in individual pairs were affected by demography,
684 we also wanted to explore how studies looking at overlapping outliers in multiple pairs (i.e. genetic
685 convergence) were affected by demography. Our simulation design, in which an ancestral burn-in
686 population is used to found 32 independent pairs of ancestral – derived populations is analogous to
687 replicated ecotype population pairs in model systems, such as various ecotype pairs of the three-
688 spined stickleback, *Gasterosteus aculeatus* (Hohenlohe et al. 2010; Jones, Chan, et al. 2012);
689 high/low predation Trinidadian guppies, *Poecilia reticulata* (Fraser et al. 2015); crab/wave ecotype
690 periwinkles, *Littorina saxitilis* (Ravinet et al. 2016; Kess, Galindo, and Boulding 2018; Westram et al.
691 2014), and alpine/montane ecotypes of *Heliosperma pusillum* (Trucchi et al. 2017). We found
692 overlapping outliers between demographic treatments (proportions above a random null
693 expectation), and therefore that signals of convergent genes are attainable for both F_{ST} and D_{XY}.
694 Clustering of outliers was predominantly driven by the presence or absence of migration (with
695 minimal overlap beyond expectations between clusters).
696
697 Given that within our simulations, selection is expectedly ineffective when migration is absent,
698 overlap of outliers across migration-free treatments suggests a different subset of genes, which are
699 not under strong selection, are repeatedly reporting the strongest signals of genetic divergence for
700 both F_{ST} and D_{XY}. In seeking to explain this, we took a random forest approach to explain how
701 different features of our simulated genes may explain variance in measures of divergence. F_{ST} was
702 heavily influenced (negative association) by within-population genetic diversity (π), and to a lesser
703 extent gene features such as gene length (negative) and the generation at which the new
704 phenotypic optimum was reached ("Pheno Gen", positive), in the absence of migration. This is

705 consistent with previous theoretical and empirical work highlighting the susceptibility of F_{ST} to
706 processes that reduce within-population diversity in the absence of selection, such as heterogeneity
707 in recombination landscape (Cutter and Payseur 2013; Ravinet et al. 2017), structural variation
708 (Kirkpatrick and Barton 2006), and gene density (Enard, Messer, and Petrov 2014; Cutter and
709 Payseur 2013). Without migration, D_{XY} outliers were categorised through features of simulated
710 genes, predominantly length of gene (positive) and proportion of gene that was coding (positive).
711 Such an effect may be explained by reduced neutrality of mutations within coding regions, and
712 therefore genes with increased coding content tend to exhibit reduced neutral polymorphism
713 (across both the ancestral and derived population) and lower D_{XY} as a result across multiple
714 demographic treatments. The secondary effect of gene length is less clear, given that D_{XY} is averaged
715 across all sites, but may stem from potentially lower linkage in larger genes (Ellegren and Galtier
716 2016) or the increased accumulation of segregating sites in larger genes (given mutation occurs with
717 a fixed probability per site). With effective selection in the presence of migration, we observed the
718 expected effect of simulated $S_{1/\sigma}$ in producing overlapping D_{XY} outliers. Interestingly however, we
719 also observed an effect of derived population π that hints towards a susceptibility to BGS even in the
720 presence of effective selection. This dependence on the efficacy of selection in detecting
721 overlapping outliers may explain the discrepancies within the literature regarding the susceptibility
722 to BGS of both F_{ST} (B. Charlesworth, Nordborg, and Charlesworth 1997; Noor and Bennett 2009;
723 Cruickshank and Hahn 2014) and D_{XY} (Matthey-Doret and Whitlock 2018).
724
725 It is perhaps surprising that population contractions (with the exception of the most extreme factor
726 level) did not have a dominant influence in detecting genetic convergence, given the strong effect on
727 population contractions on the relationship between population divergence and selection. However,
728 this is consistent with the increased influence of random drift in populations with reduced N_e,
729 thereby limiting repeatability of outliers.
730

731 ***Limitations of simulations***

732 In our analyses, we grouped genes within 32 unique demographic treatments, assuming the effects
733 of contraction and migration are equivalent for all genes. This may be unrepresentative of genomes
734 sampled from the wild, in which gene flow and effective population size can vary across genomic
735 regions through structural variation, variable recombination rate and BGS (Gossmann, Woolfit, and
736 Eyre-Walker 2011). A prominent example of such a mechanism is the well-characterised
737 consequence of reduced gene-flow within inversions that carry locally adapted alleles. Assuming
738 inversion variants are fixed between populations, migration across the locus is limited by the
739 prevention of recombinant haplotypes and resistance to introgression (Kirkpatrick and Barton 2006;
740 Ravinet et al. 2017). Recent genome scan studies have highlighted convergent outliers within
741 inversions (Jones, Grabherr, et al. 2012; Nishikawa et al. 2015; Morales et al. 2018), confirming
742 theoretical models regarding their role in shielding adaptive haplotypes from introgression during
743 the adaptation process. Our simulations highlight that genes with reduced migration, and reduced
744 N_e as a consequence, will have inflated measures of divergence and differentiation relative to other
745 genomic regions. Therefore, we predict this to lead to their over-representation in genome scan
746 outliers, thus caution should be taken regarding the adaptive significance of these outliers relative to
747 absolutely lower values of genetic divergence attained from regions outside of chromosomal
748 rearrangements.

749

750 By choosing to use a factorial design here, we have increased our understanding in the interplay
751 between individual features of demographic history and multiple measures of population
752 divergence. However, computational limitations constrained absolute population sizes to a
753 maximum of 1,000 individuals and generations to a maximum of 10,000. To mitigate these
754 constraints, we repeated the analysis over multiple mutation rates to illustrate patterns over 100-
755 fold variation of θ. In general, most trends were consistent, suggesting that results should be
756 consistent across taxa of variable effective population size. However certain patterns were

exaggerated or dampened with increased or decreased mutation rate respectively, for instance it is clear that the time lag on informativeness of D_{XY} on selection is shortest when θ is largest, which results in stronger positive associations between F_{ST} and D_{XY} by generation 10,000. It is well-documented that both N_e (Frankham 1995) and mutation rate (μ) (Hodgkinson and Eyre-Walker 2011) are highly variable across taxa, which suggests that applying knowledge of the relationships between measures of population differentiation will vary by system.

Furthermore, temporal variation within our simulations is confounded by the time at which there was a major shift in the derived population's phenotype (Figure S46-49), and by extension when selective sweeps occur. This variation in timing is random with respect to adaptive mutations arising *de novo*, but is also influenced by demography (Figure S50). For example, treatments that are founded with strong bottlenecks are less likely to evolve using variation in the founder, increasing dependence on *de novo* mutations. Additionally, variation in our population contraction parameter modifies the per population number of new mutations per generation. This is also true for features of simulated genes, such as absolute length of coding sequence. Predictable temporal variation in the time at which adaptation is likely to occur is a probable source of variance between measures of differentiation and is particularly clear for Δπ.

Concluding remarks

We have used forward-in-time simulations to perform a factorial experiment in which we explored the relationships between three measures of genetic differentiation, selection, and features of demographic history that are commonly variable in natural populations. In agreement with previous theoretical work, we found the reliance of measures of genetic divergence to indicate regions under selection are dependent on demography and variable through time, with a notable lag in D_{XY}. Furthermore, we provided novel comparisons between measures of genetic differentiation that call into question the use of multiple measures of genetic differentiation to rule out false-positives. We

783 also demonstrated that signals of convergent evolution across independent replicates can be driven
784 by similar demographic histories with minimal influences of selection. Therefore, we strongly advise
785 those using overlapping outlier scans to carefully consider the demographic context of their system
786 to avoid false-positives. It is tempting to assume that replication in study design or analysis in the
787 form of taking multiple measures of genetic divergence can reduce the risk of attaining false-
788 positives. We hope to emphasise in this study that this is not the case, as false-positives (i.e. genome
789 scan outliers that are not associated with regions under natural selection) can be driven by non-
790 random genomic or common demographic features that cannot be bypassed through replication.
791 Moreover, many of the patterns we observe are variable through time, such that the relevant pitfalls
792 of analyses will depend on the age of the populations being considered.
793

794 **ACKNOWLEDGEMENTS**

795 The authors wish to thank all members of the Fraser lab for insightful discussions on this project, as
796 well as members of the University of Exeter's EB theme and the University of Sussex's EBE group.
797 The authors would also like to thank Kai Zeng and Adam Eyre-Walker for feedback on earlier copies
798 of this manuscript.
799

800 **AUTHOR CONTRIBUTIONS**

801 JRW and BAF conceived the project, developed the experimental design and contributed towards
802 the writing of the manuscript. JRW wrote and implemented the pipeline for simulations and analysis.
803

804 **DATA ACCESSIBILITY**

805 A full pipeline including all bash, Eidos and R scripts necessary to repeat this analysis can be
806 downloaded from github (https://github.com/JimWhiting91/Contingent_Convergence_Pipeline).
807 Supplementary figures (S1-50) have been uploaded through the GSA figshare portal. Supplementary
808 figures include results for different mutation rates and $Pheno_{Null}$ data across sampling generations

809 along with additional figures.

810

811 **FUNDING INFORMATION**

812 This work was funded as part of a European Research Council grant (758382 GUPPYCon) awarded to

813 BAF.

814

815 **CONFLICTS OF INTEREST**

816 The authors declare no conflicts of interest

817

818 **REFERENCES**

819 Ahrens, Collin W, Paul D Rymer, Adam Stow, Jason Bragg, Shannon Dillon, Kate D L Umbers, and
820 Rachael Y Dudaniec. 2018. "The Search for Loci under Selection: Trends, Biases and Progress."
821 *Molecular Ecology* 27 (6). Wiley Online Library: 1342–56.
822 Bank, Claudia, Gregory B. Ewing, Anna Ferrer-Admettla, Matthieu Foll, and Jeffrey D. Jensen. 2014.
823 "Thinking Too Positive? Revisiting Current Methods of Population Genetic Selection Inference."
824 *Trends in Genetics* 30 (12): 540–46. doi:10.1016/j.tig.2014.09.010.
825 Belleghem, Steven M Van, Carl Vangestel, Katrien De Wolf, Zoe De Corte, Markus Most, Pasi Rastas,
826 Luc De Meester, and Frederik Hendrickx. 2018. "Evolution at Two Time Frames." *PLoS Genetics*.
827 Burri, Reto. 2017. "Linked Selection, Demography and the Evolution of Correlated Genomic
828 Landscapes in Birds and Beyond." *Molecular Ecology* 26 (15): 3853–56.
829 doi:10.1111/mec.14167.
830 Charlesworth, Brian. 1996. "Background Selection and Patterns of Genetic Diversity in *Drosophila*
831 *Melanogaster*." *Genetics Research* 68 (2). Cambridge University Press: 131–49.
832 ———. 1998. "Measures of Divergence between Populations and the Effect of Forces That Reduce
833 Variability." *Molecular Biology and Evolution* 15 (5): 538–43.
834 ———. 2009. "Effective Population Size and Patterns of Molecular Evolution and Variation." *Nature*
835 *Reviews Genetics* 10 (3). Nature Publishing Group: 195.
836 Charlesworth, Brian, M. T. Morgan, and Deborah Charlesworth. 1993. "The Effect of Deleterious
837 Mutations on Neutral Molecular Variation." *Genetics* 134 (4): 1289–1303. doi:10.1111/j.0014-
838 3820.2002.tb00188.x.
839 Charlesworth, Brian, Magnus Nordborg, and Deborah Charlesworth. 1997. "The Effects of Local
840 Selection, Balanced Polymorphism and Background Selection on Equilibrium Patterns of
841 Genetic Diversity in Subdivided Populations." *Genetics Research* 70 (2). Cambridge University
842 Press: 155–74.
843 Charlesworth, Deborah, Brian Charlesworth, and M. T. Morgan. 1995. "The Pattern of Neutral
844 Molecular Variation under the Background Selection Model." *Genetics* 141 (4): 1619–32.
845 Cooper, Elizabeth A., and J. Albert C. Uy. 2017. "Genomic Evidence for Convergent Evolution of a Key
846 Trait Underlying Divergence in Island Birds." *Molecular Ecology* 26 (14): 3760–74.
847 doi:10.1111/mec.14116.
848 Cruickshank, Tami E., and Matthew W. Hahn. 2014. "Reanalysis Suggests That Genomic Islands of
849 Speciation Are Due to Reduced Diversity, Not Reduced Gene Flow." *Molecular Ecology* 23 (13):

3133–57. doi:10.1111/mec.12796.

Cutter, Asher D., and Bret A. Payseur. 2013. "Genomic Signatures of Selection at Linked Sites: Unifying the Disparity among Species." *Nature Reviews Genetics* 14 (4). Nature Publishing Group: 262–74. doi:10.1007/s10450-013-9579-3.

Doren, Benjamin M. Van, Leonardo Campagna, Barbara Helm, Juan Carlos Illera, Irby J. Lovette, and Miriam Liedvogel. 2017. "Correlated Patterns of Genetic Diversity and Differentiation across an Avian Family." *Molecular Ecology* 26 (15): 3982–97. doi:10.1111/mec.14083.

Dutoit, Ludovic, Nagarjun Vijay, Carina F Mugal, Christen M Bossu, Reto Burri, Jochen Wolf, and Hans Ellegren. 2017. "Covariation in Levels of Nucleotide Diversity in Homologous Regions of the Avian Genome Long after Completion of Lineage Sorting." *Proc. R. Soc. B* 284 (1849). The Royal Society: 20162756.

Ellegren, Hans, and Nicolas Galtier. 2016. "Determinants of Genetic Diversity." *Nature Reviews Genetics* 17 (7). Nature Publishing Group: 422–33. doi:10.1038/nrg.2016.58.

Ellegren, Hans, and J. B. W. Wolf. 2017. "Parallelism in Genomic Landscapes of Differentiation, Conserved Genomic Features and the Role of Linked Selection." *Journal of Evolutionary Biology* 30 (8): 1516–18. doi:10.1111/jeb.13113.

Enard, David, Philipp W Messer, and Dmitri A Petrov. 2014. "Genome-Wide Signals of Positive Selection in Human Evolution." *Genome Research*. Cold Spring Harbor Lab.

Foote, Andrew D, Nagarjun Vijay, María C Ávila-Arcos, Robin W Baird, John W Durban, Matteo Fumagalli, Richard A Gibbs, M Bradley Hanson, Thorfinn S Korneliussen, and Michael D Martin. 2016. "Genome-Culture Coevolution Promotes Rapid Divergence of Killer Whale Ecotypes." *Nature Communications* 7. Nature Publishing Group: 11693.

Frankham, Richard. 1995. "Effective Population Size/Adult Population Size Ratios in Wildlife: A Review." *Genetics Research* 66 (2). Cambridge University Press: 95–107.

Fraser, Bonnie A, Axel Künstner, David N Reznick, Christine Dreyer, and Detlef Weigel. 2015. "Population Genomics of Natural and Experimental Populations of Guppies (*Poecilia Reticulata*)." *Molecular Ecology* 24 (2). Wiley Online Library: 389–408.

Gossmann, Toni I, Megan Woolfit, and Adam Eyre-Walker. 2011. "Quantifying the Variation in the Effective Population Size within a Genome." *Genetics* 189 (4). Genetics Soc America: 1389–1402.

Haller, Benjamin C, and Philipp W Messer. 2019. "SLiM 3: Forward Genetic Simulations beyond the Wright–Fisher Model." *Molecular Biology and Evolution* 36 (3). Oxford University Press: 632–37.

Hartl, Daniel L, Andrew G Clark, and Andrew G Clark. 1997. *Principles of Population Genetics*. Vol. 116. Sinauer associates Sunderland.

Hoban, Sean, Joanna L. Kelley, Katie E. Lotterhos, Michael F. Antolin, Gideon Bradburd, David B. Lowry, Mary L. Poss, Laura K. Reed, Andrew Storfer, and Michael C. Whitlock. 2016. "Finding the Genomic Basis of Local Adaptation: Pitfalls, Practical Solutions, and Future Directions." *The American Naturalist* 188 (4): 379–97. doi:10.1086/688018.

Hodgkinson, Alan, and Adam Eyre-Walker. 2011. "Variation in the Mutation Rate across Mammalian Genomes." *Nature Reviews Genetics* 12 (11). Nature Publishing Group: 756.

Hohenlohe, Paul A, Susan Bassham, Paul D Etter, Nicholas Stiffler, Eric A Johnson, and William A. Cresko. 2010. "Population Genomics of Parallel Adaptation in Threespine Stickleback Using Sequenced RAD Tags." *PLoS Genetics* 6 (2). doi:10.1371/journal.pgen.1000862.

Hudson, Richard R, M Slatkin, and W P Maddison. 1992. "Estimation of Levels of Gene Flow from DNA Sequence Data." *Genetics* 132 (2). Genetics Soc America: 583–89.

Irwin, Darren E., Miguel Alcaide, Kira E. Delmore, Jessica H. Irwin, and Gregory L. Owens. 2016. "Recurrent Selection Explains Parallel Evolution of Genomic Regions of High Relative but Low Absolute Differentiation in a Ring Species." *Molecular Ecology* 25 (18): 4488–4507. doi:10.1111/mec.13792.

Jones, Felicity C., Yingguang Frank Chan, Jeremy Schmutz, Jane Grimwood, Shannon D. Brady,

Audrey M. Southwick, Devin M. Absher, et al. 2012. “A Genome-Wide SNP Genotyping Array Reveals Patterns of Global and Repeated Species-Pair Divergence in Sticklebacks.” *Current Biology* 22 (1): 83–90. doi:10.1016/j.cub.2011.11.045.

Jones, Felicity C., Manfred G. Grabherr, Yingguang Frank Chan, Pamela Russell, Evan Mauceli, Jeremy Johnson, Ross Swofford, et al. 2012. “The Genomic Basis of Adaptive Evolution in Threespine Sticklebacks.” *Nature* 484 (7392): 55–61. doi:10.1038/nature10944.

Kess, Tony, Juan Galindo, and Elizabeth G. Boulding. 2018. “Genomic Divergence between Spanish *Littorina Saxatilis* Ecotypes Unravels Limited Admixture and Extensive Parallelism Associated with Population History.” *International Journal of Business Innovation and Research* 17 (3): 8311–27. doi:10.1002/ece3.4304.

Kimura, Motoo. 1983. *The Neutral Theory of Molecular Evolution*. Cambridge University Press.

Kirkpatrick, Mark, and Nick Barton. 2006. “Chromosome Inversions, Local Adaptation and Speciation.” *Genetics* 173 (1): 419–34. doi:10.1534/genetics.105.047985.

Künstner, Axel, Margarete Hoffmann, Bonnie A. Fraser, Verena A. Kottler, Eshita Sharma, Detlef Weigel, and Christine Dreyer. 2016. “The Genome of the Trinidadian Guppy, Poecilia Reticulata, and Variation in the Guanapo Population.” *PLoS ONE* 11 (12): 1–25. doi:10.1371/journal.pone.0169087.

Lanfear, Robert, Hanna Kokko, and Adam Eyre-Walker. 2014. “Population Size and the Rate of Evolution.” *Trends in Ecology & Evolution* 29 (1). Elsevier: 33–41.

Matthey-Doret, Remi, and Michael C Whitlock. 2018. “Background Selection and the Statistics of Population Differentiation: Consequences for Detecting Local Adaptation.” *Unpublished Data*. Cold Spring Harbor Laboratory, 326256. https://www.biorxiv.org/content/10.1101/326256v1.abstract.

Meier, Joana Isabel, David Alexander Marques, Catherine Elise Wagner, Laurent Excoffier, and Ole Seehausen. 2018. “Genomics of Parallel Ecological Speciation in Lake Victoria Cichlids.” *Molecular Biology and Evolution*, no. April 2018: 1–37. doi:10.1093/molbev/msy051.

Morales, Hernan E, Rui Faria, Kerstin Johannesson, Tomas Larsson, Marina Panova, Anja M Westram, and Roger Butlin. 2018. “Genomic Architecture of Parallel Ecological Divergence: Beyond a Single Environmental Contrast.” *Unpublished Data*. Cold Spring Harbor Laboratory, 447854. https://www.biorxiv.org/content/10.1101/447854v1.abstract.

Nei, Masatoshi. 1987. *Molecular Evolutionary Genetics*. Columbia university press.

Nishikawa, Hideki, Takuro Iijima, Rei Kajitani, Junichi Yamaguchi, Toshiya Ando, Yutaka Suzuki, Sumio Sugano, Asao Fujiyama, Shunichi Kosugi, and Hideki Hirakawa. 2015. “A Genetic Mechanism for Female-Limited Batesian Mimicry in *Papilio* Butterfly.” *Nature Genetics* 47 (4). Nature Publishing Group: 405.

Noor, Mohamed A F, and Sarah M Bennett. 2009. “Islands of Speciation or Mirages in the Desert? Examining the Role of Restricted Recombination in Maintaining Species.” *Heredity* 103 (6). Nature Publishing Group: 439.

Nosil, Patrik, Daniel J Funk, and Daniel Ortiz-Barrientos. 2009. “Divergent Selection and Heterogeneous Genomic Divergence.” *Molecular Ecology* 18 (3). Wiley Online Library: 375–402.

R Core Team. 2016. “R: A Language and Environment for Statistical Computing.” *R Foundation for Statistical Computing Vienna Austria*. doi:ISBN 3-900051-07-0.

Ravinet, Mark, R. Faria, R. K. Butlin, J. Galindo, N. Bierne, M. Rafajlović, M. A.F. Noor, B. Mehlig, and A. M. Westram. 2017. “Interpreting the Genomic Landscape of Speciation: A Road Map for Finding Barriers to Gene Flow.” *Journal of Evolutionary Biology* 30 (8): 1450–77. doi:10.1111/jeb.13047.

Ravinet, Mark, Anja Westram, Kerstin Johannesson, Roger Butlin, Carl André, and Marina Panova. 2016. “Shared and Nonshared Genomic Divergence in Parallel Ecotypes of *Littorina Saxatilis* at a Local Scale.” *Molecular Ecology* 25 (1): 287–305. doi:10.1111/mec.13332.

Reid, Noah M, Dina A Proestou, Bryan W Clark, Wesley C Warren, John K Colbourne, Joseph R Shaw, Sibel I Karchner, Mark E Hahn, Diane Nacci, and Marjorie F Oleksiak. 2016. “The Genomic

Landscape of Rapid Repeated Evolutionary Adaptation to Toxic Pollution in Wild Fish." *Science* 354 (6317). American Association for the Advancement of Science: 1305–8.

Reznick, David, and John A Endler. 1982. "The Impact of Predation on Life History Evolution in Trinidadian Guppies (*Poecilia Reticulata*)." *Evolution* 36 (1). Wiley Online Library: 160–77.

Roda, Federico, Luke Ambrose, Gregory M. Walter, Huanle L. Liu, Andrea Schaul, Andrew Lowe, Pieter B. Pelser, Peter Prentis, Loren H. Rieseberg, and Daniel Ortiz-Barrientos. 2013. "Genomic Evidence for the Parallel Evolution of Coastal Forms in the *Senecio Lautus* Complex." *Molecular Ecology* 22 (11): 2941–52. doi:10.1111/mec.12311.

Romiguier, J, Philippe Gayral, Marion Ballenghien, Arnaud Bernard, Vincent Cahais, A Chenuil, Ylenia Chiari, R Dernat, L Duret, and Nicolas Faivre. 2014. "Comparative Population Genomics in Animals Uncovers the Determinants of Genetic Diversity." *Nature* 515 (7526). Nature Publishing Group: 261.

Rougemont, Quentin, Pierre-Alexandre Gagnaire, Charles Perrier, Clémence Genthon, Anne-Laure Besnard, Sophie Launey, and Guillaume Evanno. 2017. "Inferring the Demographic History Underlying Parallel Genomic Divergence among Pairs of Parasitic and Nonparasitic Lamprey Ecotypes." *Molecular Ecology* 26 (1). Wiley Online Library: 142–62.

Sella, Guy, Dmitri A Petrov, Molly Przeworski, and Peter Andolfatto. 2009. "Pervasive Natural Selection in the *Drosophila* Genome?" *PLoS Genetics* 5 (6). Public Library of Science: e1000495.

Smith, John Maynard, and John Haigh. 1974. "The Hitch-Hiking Effect of a Favourable Gene." *Genetics Research* 23 (1). Cambridge University Press: 23–35.

Soria-Carrasco, Víctor, Zachariah Gompert, Aaron A Comeault, Timothy E Farkas, Thomas L Parchman, J Spencer Johnston, C Alex Buerkle, Jeffrey L Feder, Jens Bast, and Tanja Schwander. 2014. "Stick Insect Genomes Reveal Natural Selection's Role in Parallel Speciation." *Science* 344 (6185). American Association for the Advancement of Science: 738–42.

Stankowski, Sean, Madeline A Chase, Allison M Fuiten, Peter L Ralph, and Matthew A Streisfeld. 2018. "The Tempo of Linked Selection: Rapid Emergence of a Heterogeneous Genomic Landscape during a Radiation of Monkeyflowers." *BioRxiv*, 342352. doi:10.1101/342352.

Torres, Raul, Zachary A. Szpiech, and Ryan D. Hernandez. 2018. "Human Demographic History Has Amplified the Effects of Background Selection across the Genome." *PLoS Genetics* 14 (6): 1–27. doi:10.1371/journal.pgen.1007387.

Trucchi, Emiliano, Božo Frajman, Thomas H.A. Haverkamp, Peter Schönswetter, and Ovidiu Paun. 2017. "Genomic Analyses Suggest Parallel Ecological Divergence in *Heliosperma Pusillum* (Caryophyllaceae)." *New Phytologist* 216 (1): 267–78. doi:10.1111/nph.14722.

Turner, Thomas L, Matthew W Hahn, and Sergey V Nuzhdin. 2005. "Genomic Islands of Speciation in *Anopheles Gambiae*." *PLoS Biology* 3 (9). Public Library of Science: e285.

Vijay, Nagarjun, Matthias Weissensteiner, Reto Burri, Takeshi Kawakami, Hans Ellegren, and Jochen B.W. Wolf. 2017. "Genomewide Patterns of Variation in Genetic Diversity Are Shared among Populations, Species and Higher-Order Taxa." *Molecular Ecology* 26 (16): 4284–95. doi:10.1111/mec.14195.

Waterhouse, Matthew D, Liesl P Erb, Erik A Beever, and Michael A Russello. 2018. "Adaptive Population Divergence and Directional Gene Flow across Steep Elevational Gradients in a Climate-sensitive Mammal." *Molecular Ecology* 27 (11). Wiley Online Library: 2512–28.

Weir, Bruce S, and C. Clark Cockerham. 1984. "Estimating F-statistics for the Analysis of Population Structure." *Evolution* 38 (6): 1358–70.

Westram, A. M., J. Galindo, M. Alm Rosenblad, J. W. Grahame, M. Panova, and R. K. Butlin. 2014. "Do the Same Genes Underlie Parallel Phenotypic Divergence in Different *Littorina Saxatilis* Populations?" *Molecular Ecology* 23 (18): 4603–16. doi:10.1111/mec.12883.

Whitlock, Michael C, and Katie E Lotterhos. 2015. "Reliable Detection of Loci Responsible for Local Adaptation: Inference of a Null Model through Trimming the Distribution of F ST." *The American Naturalist* 186 (S1). University of Chicago Press Chicago, IL: S24–36.

Yeaman, Sam, and Sarah P Otto. 2011. "Establishment and Maintenance of Adaptive Genetic

1003 Divergence under Migration, Selection, and Drift." *Evolution: International Journal of Organic*
1004 *Evolution* 65 (7). Wiley Online Library: 2123–29.
1005 Zeng, K. 2013. "A Coalescent Model of Background Selection with Recombination, Demography and
1006 Variation in Selection Coefficients." *Heredity* 110 (4). Nature Publishing Group: 363.
1007