# AF-vapeR: A multivariate genome scan for detecting parallel evolution using allele frequency change vectors

**Authors:**

James R. Whiting[1,2], Josephine R. Paris[1], Mijke J. van der Zee[1], Bonnie A. Fraser[1]

**Corresponding Author:** James R. Whiting, jwhiting2315@gmail.com

**Author Affiliations:**

1) Department of Biosciences, University of Exeter, Stocker Road, Exeter, EX4 4QD, UK
2) Department of Biological Sciences, University of Calgary, Calgary, Canada

**ABSTRACT**

1) The repeatability of evolution at the genetic level has been demonstrated to vary along a continuum from complete parallelism to divergence. In order to better understand why this continuum exists within and among systems, hypotheses must be tested using high-confidence sets of candidate loci for repeatability. Despite this, few methods have been developed to scan SNP data for signatures specifically associated with repeatability, as opposed to local adaptation.
2) Here we present AF-vapeR (Allele Frequency Vector Analysis of Parallel Evolutionary Responses), an approach designed to identify genome regions exhibiting highly correlated allele frequency changes within haplotypes and among replicated allele frequency change vectors. The method divides the genome into windows of an equivalent number of SNPs, and within each window performs eigen decomposition over normalised allele frequency change vectors (AFV), each derived from a replicated pair of populations/species. Properties of the resulting eigenvalue distribution can be used to compare regions of the genome for those exhibiting strong parallelism, and can also be compared against a null distribution derived from randomly permuted AFV.
3) We demonstrate the utility of this approach to detect different modes of parallel evolution using simulations, and also demonstrate a reduction in error rate compared with intersecting $F_{ST}$ outliers. Lastly, we apply AF-vapeR to three previously published datasets (stickleback, guppies, and Galapagos finches) which comprise a range of sampling and sequencing strategies, and lineage ages. We highlight known parallel regions whilst also identifying novel candidates.
4) The main benefits of this approach include a reduced false-negative rate under many conditions, an emphasis on signals associated specifically with repeatable evolution as opposed to local adaptation, and an opportunity to identify different modes of parallel evolution at the first instance.

## INTRODUCTION

Studies of convergent evolution, whereby similar phenotypes evolve independently in separate lineages, provide some of the most compelling examples of the power of natural selection and have been a prominent focus of research over the last decade (Arendt & Reznick, 2008; Blount et al., 2018; Stern, 2013; Storz, 2016). Currently, a large proportion of population genomics studies exploring repeatability among populations tend to follow a similar methodology. This involves sequencing populations in replicated environments, performing genome scans for signatures of selection, and comparing candidate loci across replicates (reviewed in (Fraser & Whiting, 2019)). Once selection candidates have been identified, they can be probed for theories of genetic convergence, for example their mode of convergence (Lee & Coop, 2017), the role of genetic and pleiotropic constraints (Morris et al., 2019; Wollenberg Valero, 2020), the influence of redundancy in the genotype-phenotype map (Láruson et al., 2020; Yeaman et al., 2018), or the effect of gene flow (Fang et al., 2021). It is also of interest whether repeatability occurs through: 1) repeated selection at the same SNP (often termed 'parallel' evolution) (e.g. stickleback adapted to freshwater (Colosimo, 2005) and maize adapted to highlands (L. Wang et al., 2021); 2) nonidentical changes in the same genes or loci ('convergence' (e.g. repeated modification of *Mc1r* during pigmentation evolution (Manceau et al., 2010)); 3) changes in different genes of similar function ('functional convergence' (e.g. development *HOX* genes across aquatic mammals (Nery et al., 2016)); or 4) using no comparable genetic machinery whatsoever ('divergence'). Identifying convergent evolution candidates therefore can help us understand the constraints on evolution and why there exists a continuum of repeatability (as seen in (Stuart et al., 2017)), but it is contingent on our ability to effectively identify the candidates first.

Genome scans are a fundamental tool in identifying regions of the genome involved in local adaptation, but their propensity for false positives (Ahrens et al., 2018; Lotterhos & Whitlock, 2015), biases, and the biological interpretation of outliers (Booker et al., 2021; Cruickshank & Hahn, 2014) has led to increasing scrutiny. This scrutiny has promoted a cautiousness for ensuring minimal false-positives, but if false-positives are randomly distributed, their overlap among replicate populations should be minimal. However, a randomly distributed false-positive rate (FPR) is not necessarily a given, for instance if selection scans are biased by a shared recombination landscape (Booker et al., 2020; Lotterhos, 2019). On the other hand, applying conservative genome scans, in terms of strict cut-offs or structure corrections, and requiring overlap among replicates may have led to an elevated false negative rate (FNR) in studies of convergence. Additional sources of increased FNR when detecting candidate loci for repeatability arise when a low intersection of candidates exists due to excessive genetic drift (Leigh et al., 2021; Perrier et al., 2017), or due to the bias towards detecting hard selective sweeps and excessive/fixed allele frequency change compared with selection on standing

genetic variation (Schlamp et al., 2016; Teshima et al., 2006); the latter thus represents repeatable adaptation that may be more difficult to detect. These issues are particularly concerning given natural populations are demographically complex, and it is expected that most adaptation may occur through minor/moderate allele frequency change (soft sweeps) (Hermisson & Pennings, 2017).

Whilst the exact definitions of convergence and parallelism vary within the literature, different types of convergent evolution can be considered geometrically, where evolutionary change among replicates is compared in multivariate space (De Lisle & Bolnick, 2020; Stuart, 2019). The geometry of convergence/parallelism has been considered in various forms, for instance in assessing the size of, and angles between, phenotypic change vectors (PCV) in multivariate space (Adams & Collyer, 2009). Here, we apply the definitions outlined by De Lisle and Bolnick (2020), and consider convergence as the process by which distinct evolutionary start points converge on a common evolutionary end point. We define 'parallelism' as the process by which distinct evolutionary start points change in the same way along parallel trajectories to arrive at distinct evolutionary end points. The major distinction between convergence and parallelism here, is that lineages evolving in parallel do not evolve to become more similar, but rather change in a repeatable way. At the genetic level, such parallelism would be reflected by changes in allele frequencies in the same direction at the same sites. We can then also define 'anti-parallelism' as changes in allele frequencies in the opposite direction at the same sites, and 'non-parallelism' as changes in allele frequencies at different sites within the same genomic region. We may also find that multiple axes of parallelism exist, where some lineages follow one trajectory and others a different trajectory. We will refer to parallelism along multiple axes as 'multi-parallelism', but this could also be considered 'convergence' in some instances as evolution at different sites affecting the same gene or locus. Multi-parallelism therefore represents special instances of non-parallelism, in which several non-parallel axes exist, each with multiple lineages evolving in parallel on a given axis. These various forms of parallelism in allele frequency change will allow us to identify parallelism, anti-parallelism, and multi-parallelism, within genomes.

We present a novel population genomics scan to specifically identify regions associated with genetic parallelism. To do this, we bypass the requirement to identify regions of local adaptation in multiple datasets, but assess genomic data from all replicates at the same time. The method first scans the genomes for regions in which changes in population pairs' allele frequencies exhibit highly parallel trajectories through multivariate space. We then compare the underlying eigenvalue distributions for extreme positive skew with respect to a permuted null distribution based on allele frequency change among randomly assigned populations. We first outline the background of this approach in detail, before demonstrating its usage using

simulated data comparing it to commonly-applied $F_{ST}$ approaches, and finally apply it to a range of empirical datasets of different sampling design, sequencing resolution, and biological questions. We have released this method as an R package (https://github.com/JimWhiting91/afvaper), which we expect to become an important tool for researchers interested in identifying genomic regions associated with genetic parallelism.

## MATERIALS AND METHODS

### Allele Frequency Vector Analysis of Parallel Evolutionary Responses (AF-vapeR)

Our method is heavily inspired by the work of De Lisle and Bolnick (2020), who presented a framework with which to explore the parallel evolution of phenotypes by examining replicate PCV in multivariate space. We extend this approach to explore and compare many eigenvalue distributions calculated across the genome, as opposed to a single phenotypic distribution, and present a permutation method to assess the significance of genomic regions with respect to potential biases within the SNP dataset, such as linkage disequilibrium and population structure.

AF-vapeR first splits the SNP data into non-overlapping windows (Figure 1A) of equal user-defined SNP count (as opposed to windows based on physical distance). Within each window, we then calculate the change in allele frequency for each SNP between multiple population pairs. Population pairs are defined by the user based on *a-priori* knowledge of the system, such that each pair represents a vector of change from ecotype/phenotype/habitat A to ecotype/phenotype/habitat B. For example, where each A and B pair are unique 'replicate pairs'; or where a single A 'common ancestor' population is compared to multiple B populations; or 'multi-comparisons' in which multiple vectors are set up between a pool of A and a pool of B populations with replacement. These setups reflect sampling designs common to studies of genetic convergence (Fraser & Whiting, 2019).

The first step of the analysis produces a series of *m* x *n* matrices, where *m* represents the number of population pairs and *n* represents the number of SNPs within each window (Figure 1B). Each cell therefore describes the change in allele frequency (ΔAF) for a single SNP from population $P_a$ to $P_b$ (equation 1). The vector $P_a \rightarrow P_b$ describes ΔAF for all SNPs in the window for a given population pair and is termed the allele frequencies vector (**AFV**) (equation 2). The final matrix **X** includes the normalised **AFV** for all population pairs 1 to *m*, where normalisation involves dividing each vector by its length to a unit length of one (equation 3). The **AFV** are normalised in this way such that all population pairs exhibit equivalent variance, and as such, the absolute extent of allele frequency change is only relevant between SNPs within a window, within a population pair (i.e. within each **AFV**).

$$\Delta AF_x = q_{xa} - q_{xb}$$

$$\overrightarrow{P_a P_b} = \begin{bmatrix} \Delta AF_x \\ \Delta AF_{x+1} \\ \vdots \\ \Delta AF_n \end{bmatrix} = \mathrm{AFV} \quad (1)$$

$$\mathrm{X} = \begin{bmatrix} \hat{\mathrm{AFV}}_1 \\ \hat{\mathrm{AFV}}_2 \\ \vdots \\ \hat{\mathrm{AFV}}_m \end{bmatrix} \quad (2)$$

(3)

For each **X** matrix, we then calculate matrix **C,** which is an *m* x *m* matrix describing the correlations between normalised **AFV** within **X** (equation 4), and subsequently perform eigen decomposition of **C** to examine parallelism within the focal genomic region (equation 5). Here, Q is the matrix of eigenvectors and V is the diagonal matrix of eigenvalues, as described in De Lisle and Bolnick (2020).

$$\mathrm{C} = \mathrm{XX}^{\mathrm{T}} \quad (4)$$

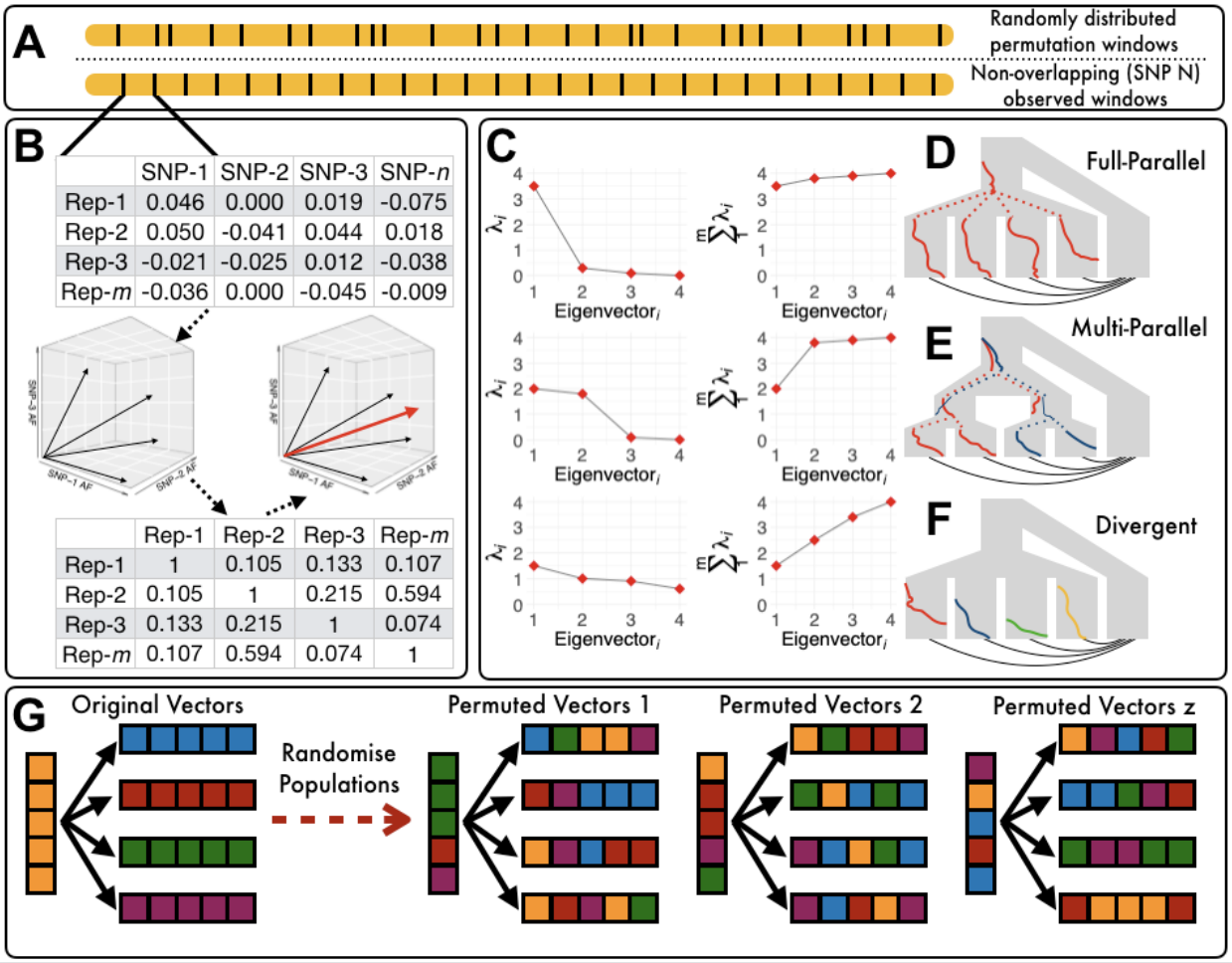
$$\mathrm{C} = \mathrm{QVQ}^{-1} \quad (5)$$

The sum of eigenvalues for each genomic interval sums to *m*, and therefore we are interested in identifying regions of the genome in which the eigenvalue distribution is positively skewed, and so weighted to variance captured by only the first (or first few) eigenvalues (Figure 1C). The sum of eigenvalues is equivalent for all genomic intervals because all AFV are normalised to unit length, and therefore, the raw eigenvalues associated with the *m*-1 eigenvectors can be compared and visualised across the genome. The eigenvectors of **C** can also be related back to the allele frequency change at individual SNPs according to equation 6.

$$\mathrm{A} = \mathrm{X}^{\mathrm{T}}\mathrm{Q}^{\mathrm{T}\ -1} \quad (6)$$

To develop a null expectation for extreme eigenvalues, we repeatedly reassign individuals to random populations, maintaining the original vector design strategy (Figure 1G), and

recalculate random AFV across the genome. This approach captures potential biasing of eigenvalues across regions of the genome and bias associated with vector sampling design that would otherwise fail to be accounted for by a one-size-fits-all null model such as the Wishart distribution (De Lisle & Bolnick, 2020).




| | SNP-1 | SNP-2 | SNP-3 | SNP-*n* |
|---|---|---|---|---|
| Rep-1 | 0.046 | 0.000 | 0.019 | -0.075 |
| Rep-2 | 0.050 | -0.041 | 0.044 | 0.018 |
| Rep-3 | -0.021 | -0.025 | 0.012 | -0.038 |
| Rep-*m* | -0.036 | 0.000 | -0.045 | -0.009 |

| | Rep-1 | Rep-2 | Rep-3 | Rep-*m* |
|---|---|---|---|---|
| Rep-1 | 1 | 0.105 | 0.133 | 0.107 |
| Rep-2 | 0.105 | 1 | 0.215 | 0.594 |
| Rep-3 | 0.133 | 0.215 | 1 | 0.074 |
| Rep-*m* | 0.107 | 0.594 | 0.074 | 1 |

***Figure 1:*** *Methodology of AF-vapeR. Chromosomes at the top (**A**) illustrate that windows for null permutations are randomly placed, but observed windows are non-overlapping and of an equal SNP number. Within a given window (**B**), normalised AFV are calculated for each replicate population pair before calculating the correlation matrix and performing eigen decomposition to yield eigenvectors (red arrow as example). The eigenvalue (λ) distributions within a window, as shown in scree plots (**C**) are informative of constraint and indicative of parallelism. Eigenvalues are summed along consecutive eigenvectors (1..m), with large summed values highlighting that a majority of variance is constrained to a few eigenvectors. Large eigenvalues associated with different eigenvectors are indicative of different evolutionary processes (**D-F**), with large primary eigenvectors highlighting full parallelism, a large second eigenvalue*

*highlighting two axes of parallelism, and no large eigenvalues suggesting divergence/non-parallelism. For the different evolutionary processes, adaptive mutations are acquired at some point backwards in time within the lineage, and change in frequency (left = 0, right = 1) within lineages through time, with dotted lines signifying inheritance during lineage splits. Different mutations are shown as different coloured lines and black lines connecting lineages illustrate which lineages AFV are calculated between. The permutation process is illustrated at the bottom (**G**), where fixed population assignment, denoted as block colour, is repeatedly shuffled and AFV are calculated under the same vector design for a total of* z *permutations. Each square represents an individual in the popmap, and each column/row represents a population, with representative AFV denoted as black arrows between the vertical population to all four horizontal populations in a common-ancestor design.*

**Simulation methods**

To demonstrate the general functionality and utility of AF-vapeR, and explore its effectiveness over demographic parameter space, we simulated data using a Wright-Fisher model with the forward-simulation software SLiM (v3.4 (Haller & Messer, 2019)). The goal of these simulations was to: 1) Demonstrate the power to detect parallel evolution under challenging demographic contexts common to natural populations; 2) Demonstrate the ability to detect different evolutionary scenarios (Figure 1D-F); and 3) Compare *AF-vapeR* to the commonly applied approach of overlapping $F_{ST}$ outlier windows among replicates.

For all simulations, we assembled a metapopulation with an ancestral founding population and four daughter populations (Figure S1). Each daughter population exchanges migrants with the ancestral founding population, such that each daughter population receives a proportion of *M* individuals each generation from the founding population, and the founding population receives a proportion of *M*/4 individuals from each of the four daughters (for a total proportion of *M* new migrants). We then expose the metapopulation to three different scenarios in which evolution occurred at a focal mutation: 1) 'Full Parallel' - All daughters receive the same mutation during their founding and experience positive selection on this common variant; 2) 'Multi Parallel' - Two mutations arise in the ancestral population, one of which is inherited by one pair of daughters and the other is inherited by the remaining pair, and all four daughters experience selection on their respective mutation; 3) Divergent: Four mutations arise in the founding population and each daughter receives and experiences selection on its own mutation. In all cases, selection takes place with a selection coefficient (*s*) of either 0.01, 0.05 or 0.1 that is positive in daughters and negative in the founding population to promote divergence. Migration (*M*) occurs at a rate of 0, 0.0025 or 0.01 between the founding population and daughters. Finally, the population size of daughters (*DP*) was manipulated to vary between 400, 2,500 and 10,000 individuals; founding populations always comprised

10,000 individuals. These three varying parameters (*s*, *DP*, *M*) were combined factorially to yield 27 parameter sets, each of which was simulated with 100 iterations, and across all three evolutionary scenarios. Following the founding of daughter populations, simulations proceeded for 200 generations to allow for divergence to occur. In the event of mutations being lost in any of the four daughter populations, simulations were returned to the point at which daughters had been founded and restarted with a new seed. All parameters are summarised in Table 1.

**Table 1**: Simulation parameters used to simulate Full Parallelism, Multi-Parallelism, and Divergence

| Variable | Value | Description |
|---|---|---|
| *s* | 0.01, 0.05, 0.1 | Relative fitness benefit of selected mutation |
| *M* | 0, 0.025, 0.1 | Migration rate (proportion of migrant individuals per generation) |
| DP | 400, 2500, 10000 | Population size of daughter populations (number of individuals) |
| AP | 10000 | Ancestral population size |
| *r* | 1e-8 | Recombination rate (per bp) |
| μ | 5e-8 | Mutation rate |
| Gen | 200 | Generations post founding prior to sampling |

These 27 scenarios were selected to explore where popular genome scans and summary statistics are predicted to struggle in detecting genetic parallelism/convergence, particularly with respect to migration-selection balance. Under low *DP* values, we expect daughters to experience more drift at neutral loci which may obscure scans based on differentiation outliers. This effect is exacerbated at low or zero *M* where drifting loci cannot be shared between daughters and their founding population. At high rates of *M*, adaptation may be constrained to only weak or subtle shifts in allele frequencies, an effect which is strongest when selection is weak. Indeed, the choice of a maximum *M* of 0.01 and minimum *s* of 0.01 is to explicitly simulate the point at which migration-selection balance is equal and adaptation should be severely limited.

Simulations took place on a 20 Mb chromosome, with the focal site at 10.0 Mb, and a constant recombination rate of $1e^{-8}$. We also explored the effects of variable recombination, in which recombination increased 100-fold linearly from a low of $1e^{-9}$ to a high of $1e^{-7}$ along a 22 Mb chromosome, with recombination increases every 2 Mb. With variable recombination, the focal site for selected variants occurred at 1 Mb ($r = 1e^{-9}$), 11 Mb ($r = 1e^{-8}$) and 21 Mb ($r = 1e^{-7}$), and

we ran 25 iterations at fixed parameters of *s* = 0.05, *m* = 0.0025 and *DP* = 10,000 at low, medium, and high recombination rates (total of 75 iterations).

Following an additional 200 generations of selection, tree sequences (Haller et al., 2019) were exported from SLiM, recapitated, and converted to VCF files of 20 randomly selected individuals per population (total of 100 per VCF) using *pyslim*. For each VCF, we then calculated $F_{ST}$ between each of the daughter populations and the founding population in 10kb windows with vcftools (v0.1.16 (Danecek et al., 2011)). We then used AF-vapeR with window sizes of 50, 200, and 500 SNPs to scan each iteration for signals of parallel evolution at the focal site. For each parameter set, we calculated the false positive (FPR) and false negative rate (FNR). For each iteration, we calculated the FPR as the number of detected signals of parallel evolution by AF-vapeR (compared against null quantiles of α=0.95, α=0.99 and α=0.999), or overlapping $F_{ST}$ outlier windows (overlapping in all four, any three, or any two, outlier sets above the same quantile cut-offs) at neutral loci (beyond an exclusion zone of 0.1 cM around the focal site). We then averaged FPR across all 100 iterations to estimate FPR of each parameter set. For FNR, we counted the number of iterations in which parallel evolution or overlapping outlier $F_{ST}$ windows were not detected within a region of 0.01 cM around the focal site at each null quantile, and divided this by the number of iterations (100).

**Empirical dataset demonstrations**

We chose three publicly available datasets to demonstrate different utilities of AF-vapeR. These include restriction-site associated digestion (RAD) sequencing data from 18 freshwater-adapted stickleback populations British Columbia (BC) and an ancestral marine population (Magalhaes et al., 2016, 2020); whole-genome sequencing (WGS) of five pairs of high-predation (HP) and low-predation (LP) adapted guppies (Whiting et al., 2021); and WGS data from 12 species of Galapagos finches, used previously to explore small vs large, and pointed vs blunt, beak morphologies within the adaptive radiation (Han et al., 2017; Lamichhaney et al., 2015, 2016). These datasets therefore include different sequencing approaches (RAD-seq vs WGS), different sampling approaches (replicate pairs vs common ancestor vs multiple comparisons; see (Fraser & Whiting, 2019)), and a mix of within- and across-species comparisons.

## RESULTS

**Simulation results**

AF-vapeR was successfully able to distinguish between the three different simulation scenarios: Full Parallelism (Figure 2A), Multi-Parallelism (Figure 2B), and Divergent (Figure 2C). These distinctions can be seen as clear peaks arising on the first (Full) or second (Multi) eigenvector, or not at all (Divergent) at the selected site at 10Mb.

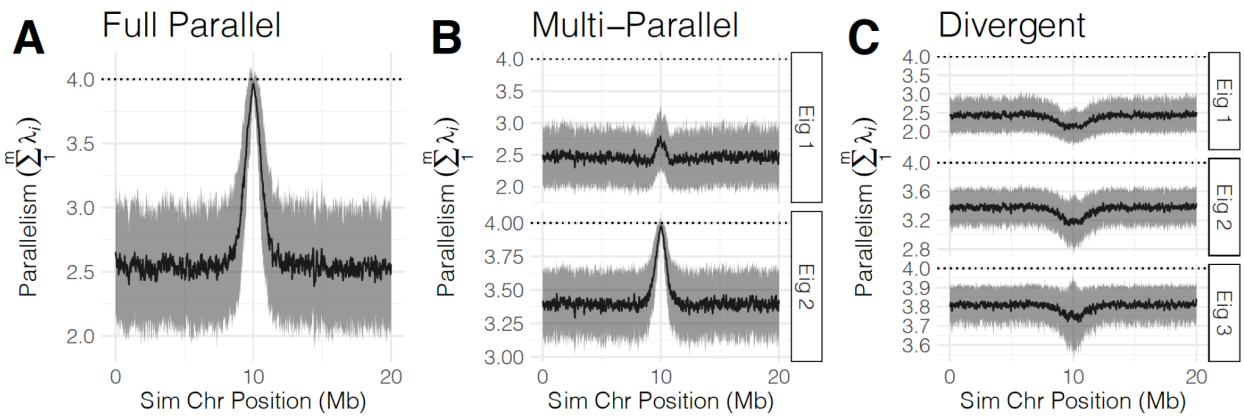



***Figure 2:*** *Simulated eigenvalue peaks detected by AF-vapeR under three different evolutionary scenarios: full parallelism (**A**), multi-parallelism (**B**), and divergence (**C**). Each panel shows the mean summed eigenvalue (black line) across the simulated 20Mb chromosome, averaged within 10kb windows across 100 iterations, along with the standard deviation (grey ribbon). The maximum possible summed eigenvalue, corresponding to m replicates (in this case four), is marked as a dotted line. For panel A, only eigenvalue 1 is shown, for panels B and C, the summed eigenvalues for increasing eigenvectors are shown as facet rows up until a peak is observed. No facet is shown for eigenvector 4 as all windows sum to four by this point, producing a flat line. The simulation parameters shown are for* s *= 0.05, M = 0.025, and DP = 10,000.*

We then looked at error rates associated with detecting full parallelism. Across parameter space, at α = 0.01, FPR were low for all AF-vapeR window sizes (50 SNPs, max FPR = 0.057; 200 SNPs, max FPR = 0.072; 500 SNPs, max FPR = 0.085) and all overlapping $F_{ST}$ thresholds (all 4 daughters, max FPR = 1.1e-4; any 3, max FPR = 7.5e-4; any 2, max FPR = 4.1e-3) (Figure S2). For $F_{ST}$ approaches this is not surprising, whilst for single scans of local adaptation there exists a risk of false positives, the likelihood of these overlapping is minimal if false positives are randomly distributed. AF-vapeR FPR was positively associated with daughter population sizes and negatively associated with migration rate (Figure S2).

In studies of parallelism/convergence, false negatives are of more concern. Across parameter space, FNR for all AF-vapeR window analyses consistently out-performed the most strict and most-likely applied $F_{ST}$ analysis (outlier in all 4 daughters), and often out-performed all $F_{ST}$ overlap cut-offs (Figure 3). Expectedly, strong bottlenecks (*DP* = 400, first column of panels in Figure 3), and strong migration (*M* = 0.01, first row of panels in Figure 3) led to elevated FNR for $F_{ST}$ analyses, but AF-vapeR was generally more robust to these parameters.

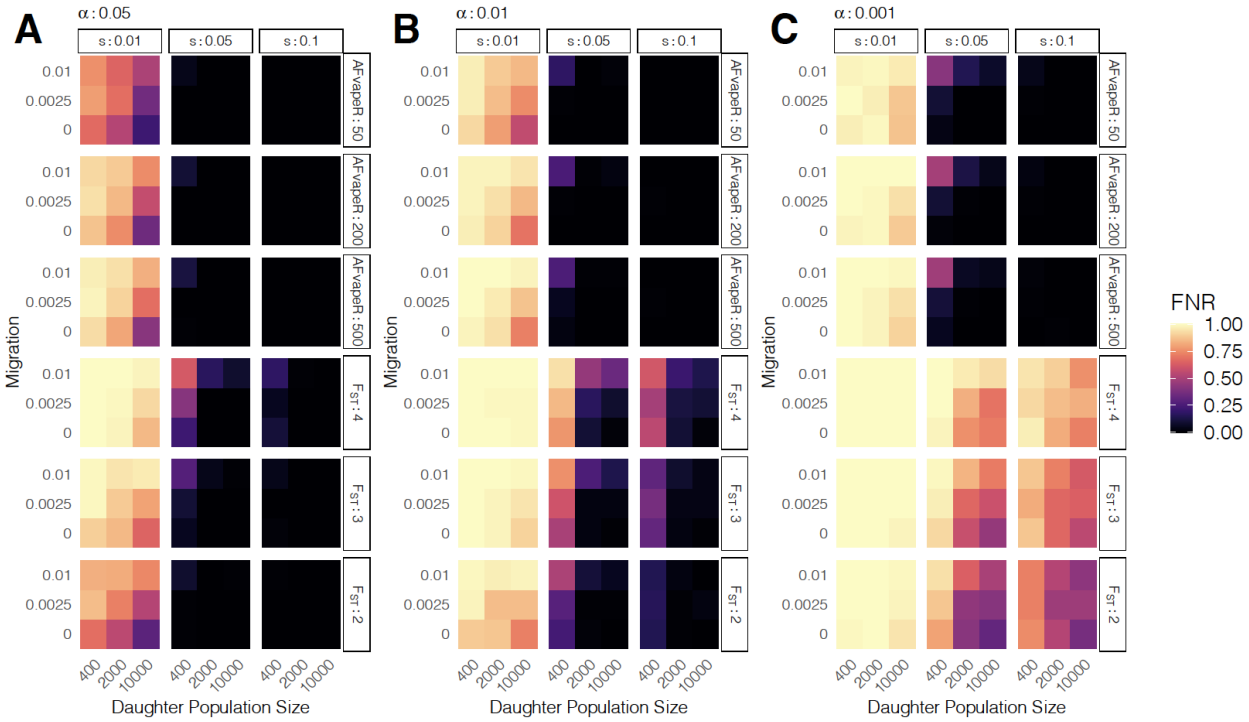



***Figure 3:*** *False negative rates (FNR) associated with AF-vapeR and $F_{ST}$ outliers tests. AF-vapeR facet labels correspond to the number of SNPs per window used in the analysis (50, 200, and 500 SNPs). $F_{ST}$ facet labels correspond to the number of overlapping populations required to call a region as overlapping, (i.e. 2 = an outlier in any 2 of the 4 populations, 3 = an outlier in any 3, 4 = outlier detected in all four). Panels A, B and C show FNR calculated based on taking outliers at quantile cutoffs of 95%, 99%, and 99.9% respectively. Within facet rows, each grid space represents the FNR averaged over 100 iterations for one of the 27 factorial parameter combinations, where facet columns = strength of selection on beneficial mutation, y = migration rate, and x = daughter population size.*

For all $F_{ST}$ overlapping cut-offs, weak selection ($s$ = 0.01) often prevented identification of full parallelism all together unless the least stringent alpha cut-off was applied, but even then, only 15% of simulations detected full parallelism (α = 0.05, $s$ = 0.01, $F_{ST}$ outlier in all 4; min FNR = 0.85, max FNR = 1.0). In contrast, under the same conditions AF-vapeR recovered full parallelism in at least 21% and at most 78% under all nine parameter combinations of daughter population size, migration and $s$ = 0.01 (α = 0.05, $s$ = 0.01, AF-vapeR window size = 50; min FNR = 0.22, max FNR = 0.79). Allowing for the identification of parallelism candidates as $F_{ST}$ outliers in any two daughters improved FNR under weak selection (α = 0.05, $s$ = 0.01, $F_{ST}$ outlier in any 2; min FNR = 0.3, max FNR = 0.86). This approach however still does not perform as well as AF-vapeR and fails to capture the simulation conditions of parallel selection in all four daughters.

343 Under stronger selection, AF-vapeR exhibits consistently low FNR even as cut-off stringency
344 increases from α = 0.05 (Figure 3A) to α = 0.01 (Figure 3B) and α = 0.001 (Figure 3C).
345 Conversely, increasing the stringency of outliers dramatically increases $F_{ST}$-based FNR across all
346 parameter space. This behaviour of AF-vapeR is important because it indicates that the slightly
347 elevated FPR of AF-vapeR, compared with $F_{ST}$ approaches, under lower outlier stringency can be
348 alleviated by increasing the quantile stringency with a minimal impact of FNR. The exception
349 here is under a combination of high migration rates and small population sizes, where an
350 increase in FNR can be seen as outlier stringency increases. Researchers should therefore
351 consider the role of migration and population size when applying the method under such
352 scenarios.
353
354 A comparison of the relationship between allele frequency differentiation (AFD) and FNR
355 (Figure S3) demonstrates that AF-vapeR is more sensitive to changes in AFD, as FNR decreases
356 more rapidly as AFD increases compared with $F_{ST}$-based approaches. This relationship with AFD
357 is most likely what causes AF-vapeR to exhibit a reduced FNR when migration is high and
358 selection is weak. The reduced FNR seen when drift is strongest (DP = 400) is better explained
359 by random drift at neutral sites reducing the genome-wide background signal of parallelism and
360 making peaks at selected sites more distinct.
361
362 Overall, these results show that AF-vapeR has comparably low FPR to overlapping multiple sets
363 of $F_{ST}$ outlier sets, and has a generally reduced FNR across parameter space. AF-vapeR exhibits a
364 considerably reduced FNR under circumstances that are expected to obscure $F_{ST}$ scans, such as
365 scenarios including bottlenecks, weak selection, and strong migration. AF-vapeR consistently
366 outperformed $F_{ST}$ scans when $F_{ST}$ outliers had to be observed in all four daughter populations,
367 therefore it performed better at capturing the simulation conditions of all daughters
368 experiencing parallel selection. FNR were typically reduced with reduced window size for AF-
369 vapeR with only small increases in FPR, suggesting whilst a balance exists between maximising
370 noise and signal, the effect of noise is marginal.
371
372 Comparable results were observed when running the simulation under a multi-parallelism
373 scenario. Here we used AF-vapeR to detect multi-parallelism by looking for outliers at the
374 selected site on the second eigenvector. In these simulations, as above, FPR was generally low
375 (Figure S4), migration rate was positively associated with FNR, and selection coefficient was
376 negatively associated with FNR (Figure S5). Cut-off stringency also behaved similarly, largely
377 obscuring $F_{ST}$-based analyses at higher stringencies. These results demonstrate that AF-vapeR is
378 able to robustly identify parallelism candidates across parameter space whilst also inferring that
379 multiple axes of parallelism exist.
380

381 Introducing a variable recombination landscape to simulations affected eigenvalue peak shape
382 at sites under parallel selection and variance at neutral sites. The width of eigenvalue peaks was
383 inversely associated with recombination rate. Increased recombination also marginally reduced
384 the height of peaks (Figure S6A). The interaction between peak height and width was not
385 affected by the number of SNPs used to calculate eigenvalues. Neutral regions exhibited a
386 greater variance of primary eigenvalues with reduced recombination; however, on average
387 there was a negligible difference in neutral sites across the 100-fold increase in recombination
388 from 1e-9 to 1e-7 cM/Mb (Figure S6B). Peaks of full parallelism were still clearly observable
389 relative to the neutral backdrop regardless of whether they occurred in a region of low,
390 medium, or high recombination rate.
391
392 **Empirical Results: Marine-Freshwater stickleback**
393 The motivation for using this dataset was to demonstrate the ability to detect *a-priori* loci
394 associated with parallel evolution from reduced representation sequencing (RAD-seq). We
395 applied AF-vapeR to a dataset of 18 freshwater-adapted populations and a single marine
396 population in a 'common-ancestor' sampling strategy (Figure 4A) sampled from the northeast
397 Pacific coast in British Columbia (Magalhaes et al., 2020). Eighteen AFVs were therefore
398 calculated between the marine population and each freshwater population. We applied SNP
399 window sizes of 10, 25 and 50 SNPs, corresponding to median physical window sizes of 116.0,
400 336.3, and 714.3 kbp respectively. There were minimal signals of parallelism at window sizes of
401 25 and 50 SNP windows, likely due to the combination of sparse SNP density and linkage decay
402 across most regions at these extended physical sizes. With 10 SNP windows (N=2376), the
403 genomic landscape of full parallelism, defined as the per-window primary eigenvalues (Figure
404 4B), closely resembled the genomic landscape of marine-freshwater divergence commonly
405 observed in this system (Fang et al., 2021; Hohenlohe et al., 2010; Roberts Kingman et al.,
406 2021). In particular, chromosome 4 exhibited a clear chromosome-level increase in parallelism
407 (Figure 4C-D). Chromosomes 21, 12, 20 and 7 exhibited the next highest median primary
408 eigenvalues (Figure 4D). With the exception of chr12, each of these are enriched for QTL
409 associated with marine-freshwater phenotypes (Peichel & Marques, 2017).
410

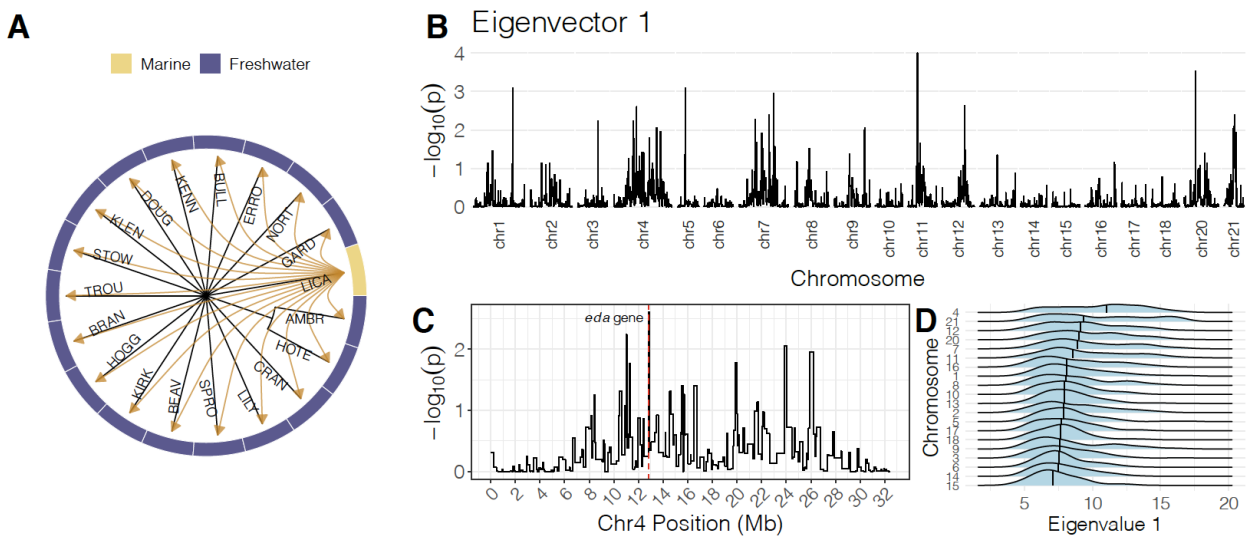



***Figure 4:*** *AF-vapeR results using a RAD-seq dataset for marine-freshwater stickleback from British Columbia. Sampling design is described based on the tree (**A**), with all vectors calculated from the marine population (LICA) to all freshwater populations. The genome-wide distribution of eigenvalue 1 (from 10 SNP windows) is plotted in panel **B**, based on empirical p-values compared to ~10,000 null permutations. Chromosome 4 is highlighted in panel **C**, with the eda gene represented by a red dashed line. The per-chromosome distributions of eigenvalue 1 results are plotted as ridges in panel **D**, with the median denoted in each. Chromosomes are ordered vertically in terms of median strength of eigenvalue 1, with the largest median (chr4) at the top.*

Above the permuted 99.9% cutoff, we detected six windows exhibiting full parallelism on chromosomes 1, 5, 7, 11, and 20, with the window on chromosome 11 exceeding all null permutations. Each window included previously identified candidates for freshwater adaptation: chromosome 1 (21721603-21849857) included the *atp1a1* gene known to be important for freshwater adaptation (Nelson & Cresko, 2018); chromosome 5 (4273793-4332432) is between several QTL associated with marine-freshwater feeding (Erickson et al., 2016; Miller et al., 2014); chromosome 7 (19280164-19396323) harboured SNPs covering the claudin genes *cldnc* and *cldnh*, the latter of which is differentially expressed between marine and freshwater stickleback exposed to salinity gradients (Gibbons et al., 2017); chromosome 11 (5577181-5794574) included feeding (Glazer et al., 2015) and sensory (Wark et al., 2012) QTL and has been detected in other parallel evolution studies in British Columbia (Hohenlohe et al., 2010); and chromosome 20 (6190327-6247151) includes the genes *ddc* and *grb10b*, the former of which is suspected to be involved in parallel evolution through copy number evolution (Hirase et al., 2014). Above the 99% cutoff, we recovered an additional 17 windows, including the window adjacent to the well-known *eda* gene associated with parallel marine-freshwater

armour haplotypes. The window containing *eda* exhibited a weaker signal of full parallelism, likely because three of the 18 freshwater populations exhibit the full armour plating associated with marine fish (Magalhaes et al., 2020).

On the second eigenvector, we detected no windows above a 99.9% cutoff, but 10 windows exceeded the 99% cutoff. For these windows, eigenvalue 1 made up the majority of the summed total, with eigenvalue 2 explaining 15.32% or less of the total variance among AFV. This indicates that these windows represent cases in which most, but not all, freshwater populations are diverging in parallel on eigenvector 1, but additional axes exist to explain non-parallel allele frequency change in a smaller number of replicates. All significant windows on eigenvectors 1 and 2 are summarised in Tables S1-2.

**Empirical Results: High- (HP) and Low-Predation (LP) Guppies**

The goal of using this dataset was to detect known regions associated with HP-LP adaptation in guppies identified in (Whiting et al., 2021), but also to identify novel candidates given hitherto limited evidence of convergence in this system. AF-vapeR was run over WGS data from five HP-LP population pairs (Figure 5A), each from a separate river, using window sizes of 25, 50, 100, 200 SNPs. Variable window sizes were used to evaluate how robust outliers were to changes in this parameter. Median physical window sizes were 4.8, 10.3, 21.4, and 43.9 kb respectively. We looked for regions of the genome that exhibited greater summed eigenvalues on the first and second eigenvectors than expected across two or more window sizes. Taking outliers as those present in two or more window sizes was done to increase confidence in candidate regions, and to ensure spurious signals linked to window size choice were removed. For window sizes of 25 and 50 SNPs we used a quantile cut-off of 99.9%, but relaxed this to 99% for 100 and 200 SNP windows due to the lower overall number of windows and reduced genome-wide variance (Figure 5B-C).

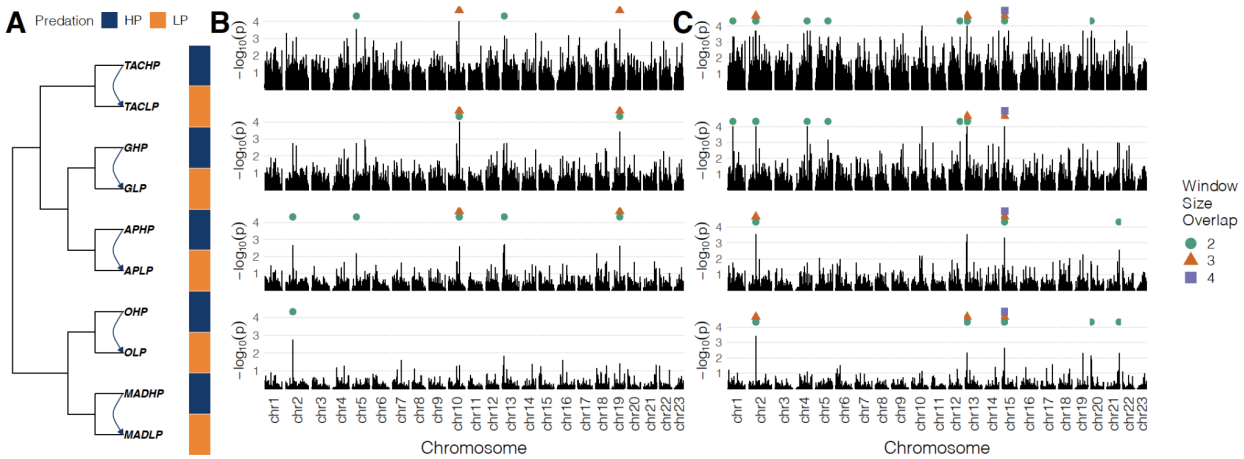



***Figure 5:*** *AF-vapeR results for WGS data from HP-LP guppies. Each AFV was calculated between a distinct HP and its corresponding LP population within five rivers in a "replicate pairs'' sampling design as illustrated in* ***A****. Each row in panels* ***B*** *and* ***C*** *shows genome-wide empirical p-values calculated based on 10,000 null permutations and window sizes of 25 (top), 50, 100 and 200 (bottom) SNPs for primary (****B****) and secondary (****C****) eigenvalues. Agreement among window sizes is shown by points above peaks, with point shape and colour indicative of overlap among window sizes.*

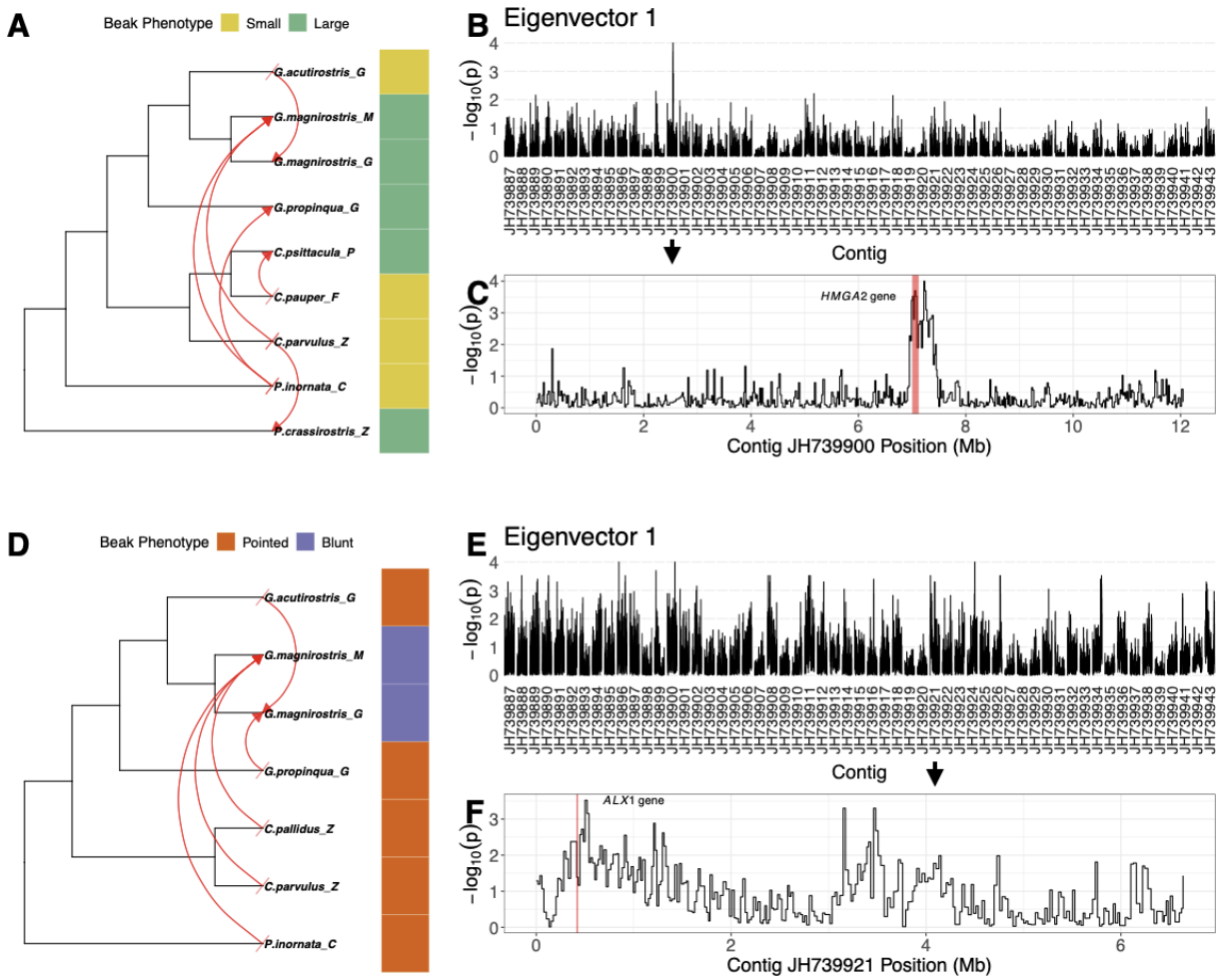
This analysis highlighted six windows across five regions of the genome with significant primary eigenvalues on chromosomes 2, 5, 10, 13 and 19 (Table S3). However, all of these regions represented a combination of parallel and antiparallel vectors (Table S3), suggesting that whilst all population pairs exhibit allele frequency differentiation along a common axis, the direction of change was not always consistent among replicates. The regions with the most parallel changes (i.e. only a single anti-parallel replicate to four parallel replicates) were chr5:8127850-8134640 (Aripo antiparallel), and chr10:21021152-21034438 (Tacarigua antiparallel). The results from the primary eigenvalue thus agree with the previous analyses of these populations, in that full parallelism is limited within this system (Whiting et al., 2021).

We then examined results from the secondary eigenvector (sum of eigenvalues 1 and 2). Here we identified 13 significant windows over nine regions (Table S4, Figure 5C), when taking outliers from two or more window sizes. Large summed eigenvalues can be caused by either a large eigenvalue 1 plus a small eigenvalue 2, or by a more balanced eigenvalue 1 and 2. The latter are better indicators of multiple parallel axes of allele frequency change. Windows on chr1, chr15, chr20 and chr21 exhibited second eigenvalues accounting for at least 30% of total variance (Table S4). Of these, the window on chr21 was the only window where one population pair exhibited no parallelism with any of the other population pairs. Examination of the population loadings onto eigenvectors 1 and 2 (Table S4) highlighted that two of these windows

494 (chr1:9791276-9798829 and chr20:1248797-1296686) exhibited multi-parallelism following a
495 similar pattern, with Caroni drainage rivers (GHP/GLP, TACHP/TACLP, APHP/APLP) parallel along
496 eigenvector 1 and MADHP/MADLP (Northern drainage) and OHP/OLP (Oropouche drainage)
497 parallel along eigenvector 2. These regions thus represent particularly interesting candidates in
498 which divergent haplotypes at the same genomic regions are exhibiting parallel change from HP
499 to LP within phylogenetic lineages seen in Figure 5A.
500
501 The window chr20:1248797-1296686 intersects with the major candidate previously identified
502 as evolving convergently in the three Caroni rivers (Whiting et al., 2021). Here however, we find
503 evidence of multi-parallelism extending selection at this region to all rivers. Plotting the allele
504 frequencies of this region highlights the clear Caroni vs non-Caroni haplotype structure (Figure
505 S7B), and also reveals AF change in the same direction for HP > LP within these two groups.
506 Examining the scores of individual SNPs within this window highlights that both eigenvectors
507 share a common high-loading SNP (denoted by an arrow in Figure S7). This SNP intersects the
508 only predicted gene in this region (ENSPREG00000013524), and indicates that both haplotypes
509 may have independently acquired the same mutation, which is under parallel selection in all
510 rivers. Interestingly, a protein sequence blast against NCBI's nr database highlights that this
511 gene exhibits sequence similarity with various teleost syncytin genes. This family of proteins are
512 essential for placental development in mammals (Dupressoir et al., 2011), and are expressed in
513 the maternal follicle transcriptome of other live-bearing poeciliids (Guernsey et al., 2020).
514 Additional information regarding candidates at chr1:9791276-9798829, and an assessment of
515 multi-parallelism at chr15:5028361-5066375, can be found in the supplementary materials.
516
517 In conclusion, these results confirm previous analyses that full parallelism is likely limited in this
518 system, whilst also identifying a number of novel candidates evolving under a multi-parallel
519 scenario.
520
521 **Empirical Results: Galapagos Finch beak morphology**
522 The goal here was to explore an inter-species dataset to examine limitations associated with
523 divergence over time; the common ancestor of the Galapagos finch species group arrived at the
524 islands ~1.5 million years ago (Lamichhaney et al., 2015). The original study used WGS to
525 investigate beak morphology among Galapagos finches (Han et al., 2017) and included 12
526 between-species comparisons in a multi-comparison sampling design. The focal beak traits of
527 interest were size (small [S] vs large [L]) and shape (pointed [P] vs blunt [B]). These 12
528 comparisons included all combinations of beak phenotypes, e.g S-S, S-L, L-L, P-P, P-B, B-B (Table
529 2 in (Han et al., 2017)). We first used these 12 comparisons with AF-vapeR to fully replicate the
530 original study, but expectedly given the range of phenotype comparisons being made, we
531 observed no signals of parallelism. This emphasises the need to carefully consider the input

532 vectors prior to analysis. We instead looked at targeted phenotype comparisons, i.e. S-L (six
533 comparisons, nine species) and P-B (five comparisons, seven species) (Figures 6A and 6D). For
534 each set of vectors, we again assessed the detection of outliers across four window sizes: 50,
535 100, 200, 500 SNPs. These window sizes were selected due to greater SNP density, to be
536 expected across species, reflecting median physical sizes of 5.4, 11.1, 22.6 and 57.2 kb.
537 Signatures of parallelism and eigenvalue variance were generally greater in this dataset than in
538 the guppy dataset, so we used stricter outlier criteria to examine only the candidate regions
539 with the strongest evidence (>99.9% of all permuted null values across all window sizes).
540
541 For S-L comparisons, we observed a single strong signal of full parallelism clustered around the
542 *HMGA2* locus on contig JH739900 (chromosome 1A of the zebra finch genome) (Figure 6B-C)
543 (Table S5). Interestingly, the strongest peak (JH739900:7234666-7291390, eigenvector 1 =
544 87.58% variance) overlapped the adjacent *MSRB3*, rather than *HMGA2*. There were no signals
545 of multi-parallelism on the second eigenvector for S-L vectors. This strong signal of full
546 parallelism at the *HMGA2* region implies a single axis describing the change from small to large
547 beak phenotypes, indicative of a single S and a single L ancestral haplotype. Indeed, clustered
548 genotypes over the two windows overlapping *HMGA2* illustrate cross-species clustering within
549 S and L haplotypes (Figure S8). These results confirm that the evolution of beak size within the
550 galapagos finch adaptive radiation has involved the segregation of commonly inherited S and L
551 haplotypes at the *HMGA2* gene among species (Lamichhaney et al., 2016), and no other regions
552 of the genome showed this pattern as clearly.
553
554 In contrast, for P-B comparisons we observed signals of full parallelism on fourteen contigs
555 (based on the geoFor1 genome) (Table S6). The *ALX1* gene has been identified as the primary
556 candidate of interest in the evolution of blunt beak shape (Lamichhaney et al., 2015). Our full-
557 parallelism outliers included a peak adjacent to, but not including, the *ALX1* locus on contig
558 JH739921 (Figure 6E-F), but still within the 130kb downstream region identified in Lamichhaney
559 et al. (2015). A lack of parallelism over the *ALX1* gene specifically reflected strong haplotype
560 structure among pointed beak species; genotype clustering suggests four monophyletic P-
561 haplotype groups over the *ALX1* gene (Figure S9). At this locus, divergent "start points" prevent
562 the identification of parallel P-B axes. This result highlights that signatures of full parallelism are
563 likely eroded over long-time scales, which may limit inferences when applying our method to
564 haplotypes with ancient origins. The region adjacent to *ALX1* that exhibited a strong signal of
565 full parallelism did not include any genes, so may be regulatory for *ALX1* and conserved.
566
567 Beyond the *ALX1* locus, we recovered signals of full parallelism at genes previously identified in
568 genome scans of Galapagos finch beaks (*STK3*, *VPS13B*, *NT5C1B*, *RDH14*, *ADK*) (Chaves et al.,
569 2016; Lamichhaney et al., 2015; Lawson & Petren, 2017). We also identified windows with

570 novel candidate genes with probable roles in craniofacial development (*LRIG3*, *IGF1*, *PMCH*,
571 *PARPBP*, *NUP37*, *GDF6*, *PTDSS1*, *OSR2*, *SAMD12*, *COLEC10*, *ENPP2*, *AHSA1*, *AP3M1*, *VCL*, *WNT6*,
572 *WNT10A*, *FEV*, *TP63*) (Table S7). Of particular note is a signal of full-parallelism detected at *IGF1*,
573 which has been linked to bill morphology in black-bellied seedcrackers (*Pyrenestes ostrinus*)
574 (vonHoldt et al., 2018). In the seedcracker study, no evidence was found to link this gene to
575 beak morphology in *Geospiza*, but our analysis suggests this gene may be important for both
576 taxa. In total, 16/31 windows (500 SNPs, ± 50kb) with strong signals of full-parallelism
577 contained previously identified or novel candidate genes for beak morphology in this system.
578
579




580
581 ***Figure 6:*** *AF-vapeR results for WGS data from Galapagos finches. Each AFV was calculated*
582 *between species with contrasting beak phenotypes (small to large or pointed to blunt), with*
583 *species used multiple times in a multi-comparison design (**A**,**D**). Empirical p-values calculated*
584 *based on 10,000 null permutations and a window size of 500 SNPs for primary eigenvectors (**B**)*
585 *highlighting a single strong signature of full parallelism on contig JH739900 associated with the*
586 *HMGA2 locus for beak size (**C**). The comparable figure for beak shape (**E**) highlights several loci*

587 *exhibiting full parallelism, including a window close to, but not including, the candidate ALX1*
588 *gene (marked in red) on contig JH739921 (**F**).*
589
590
591 **DISCUSSION**
592 **Summary of approach**
593 We present AF-vapeR, a novel genome scan method to explore SNP data for signals of parallel
594 evolution based on multivariate allele frequency vectors. Using forward simulations of different
595 evolutionary scenarios, our method performs as well, in terms of FPR, and better, in terms of
596 FNR than commonly used $F_{ST}$ outlier overlap methods. At the same time, we are able to
597 distinguish among different modes of evolution. Then, using empirical datasets we were able to
598 identify previously known candidates of parallel evolution as well as uncovering additional
599 candidates. The reduction in Type-II error is largely derived from re-ordering the workflow used
600 to identify candidates for parallel evolution by explicitly scanning the whole genome for
601 signatures associated with parallelism as opposed to local adaptation. Identified candidates for
602 parallel evolution can later be assessed for the relative strength of selection. This
603 reorganisation of workflow is important in alleviating issues associated with conservative
604 identification of genomic regions under selection within individual replicates.
605
606 AF-vapeR can be considered analogous to approaches that seek to identify regions of the
607 genomes that are associated with or cluster according to ecotype/habitat/phenotype compared
608 with the underlying phylogeny, such as cluster separation score (CSS) (as applied by (Jones et
609 al., 2012; Morales et al., 2019; Roberts Kingman et al., 2021)), Cochran–Mantel–Haenszel
610 (CMH) tests (as applied by (Barghi et al., 2019; Orozco-terWengel et al., 2012)), and, more
611 recently, the C2 statistic (Olazcuaga et al., 2020) (an extension of BayPass' covariate model to
612 binary groups). CSS is a distance measure that can be applied within genomic windows,
613 comparing within-group distances with between-group distances after dimensionality
614 reduction. CMH tests similarly can be used to assess consistency among many replicates for
615 allele frequency change of individual SNPs, and compare consistency against a null expectation
616 using odds ratio tests. As such, each of these assumes a single possible axis for parallelism. A
617 major advantage of AF-vapeR is the ability to highlight multiple axes of parallelism that reflect
618 different trajectories of allele frequency change among replicates. Only under full-parallelism
619 scenarios are haplotypes from common ecotype/habitat/phenotype groups expected to cluster
620 together. If multiple axes of parallelism exist, haplotypes from common
621 ecotype/habitat/phenotype groups would be expected to form multiple distinct clusters. By
622 identifying multiple axes of parallelism, and providing information on eigenvector loadings, our
623 approach can highlight and describe modes of evolution including and beyond full-parallelism.
624 Further, because our permutations take sampling design into account, our approach is flexible

to sampling designs involving biased group membership (such as a single common ancestor vs many diverged replicates) that may affect other clustering approaches.

Our AF-vapeR method, however, is not an explicit test for selection, but for parallel allele frequency differences. We assume that regions exhibiting highly parallel allele frequency change, above a given expectation, are of themselves an indicator of selection (but see (Bailey et al., 2015)), but there may be cases where these could arise beyond selection. Parallel allele frequency change could be due to parallel evolution independently sorting shared standing genetic variation, as is likely the case for stickleback, but could also be due to a single historical evolutionary event that is then simply inherited in all focal replicates, as is likely the case within blunt-beaked finches given these are monophyletic in the sampling design (Figure 6D). This sampling design for blunt-beaked finches is also why the baseline eigenvalue 1 scores are generally high, as there is a genome-wide signal of shared inheritance within the lineage. To compensate for this, we used stricter cutoffs in the finches compared with stickleback and guppies, but it is important to keep the underlying phylogeny in mind when interpreting results. Another possible source of parallel allele frequency change beyond selection may be hybridisation from an unstudied lineage, given heterogeneity in the success of introgression across the genome, as suggested for flycatchers (Ellegren et al., 2012) and anchovies (Le Moan et al., 2016). This should not, however, be confused with hybridisation providing adaptive alleles to multiple lineages that are subsequently selected independently. This latter process, observed for example in *Heliconius* (Edelman et al., 2019), is a genuine representation of parallel evolution given increases in the frequency of introgressed variants occur independently (Lee & Coop, 2019).

Windows identified as exhibiting parallelism may not be those experiencing the strongest selection, but rather may reflect other processes that promote repeatability. These could include heterogeneity in the mutation landscape (Bailey et al., 2018), differential rates of ancient haplotype maintenance (Nelson & Cresko, 2018; B. Wang et al., 2019), or gene/QTL density landscape (Roberts Kingman et al., 2021). The relative influences of these various processes on parallel evolution are not yet well-characterised. Given our approach is specifically sensitive to parallelism, we expect it could be applied to elucidate these effects and how they interact with one another across diverse systems.

The recombination rate landscape is known to obscure many types of genome scan (Booker et al., 2020; Lotterhos, 2019). Reduced recombination inflates variance in divergence, producing longer tails in low-recombination regions which may be identified as false-positives (Booker et al., 2020), or due to increased background selection in low recombination regions (Burri, 2017). Our analysis similarly exhibits a minor association between reduced recombination rate and

increased variance of eigenvalues at neutral regions (Figure S6). Our approach, however, may be more resilient in terms of false positives linked to recombination. Firstly, by using windows of SNP count, instead of physical size, variable linkage and independence among SNPs of varying physical distance is somewhat accounted for. Put another way, because genetic variation is associated with recombination landscape (Burri et al., 2015; Stankowski et al., 2019), we might expect larger, sparser, physical windows in low-recombination regions, and smaller, denser windows in high-recombination regions. This balances the risk of oversampling and having many non-independent windows in tight-linkage in low-recombination regions, and undersampling with non-independence within windows in regions of high-recombination. Secondly, our analysis involves directionality, as we can distinguish between parallel and antiparallel allele frequency change. If low recombination inflates eigenvalue variance randomly, it is unlikely that all allele frequency change will occur in the same direction, and further addition of independent lineages will increase analysis robustness.

**Recommendations**

In terms of datasets best-suited to this approach, we recommend four replicates as a minimum requirement to explore patterns of full and multi-parallelism. Less than this and the likelihood of false-positives increases as large primary eigenvalues are minimally distinct relative to genome-wide values. This is because eigenvalues are constrained between 0 and *m*, where *m* is the number of AFV. As such, the distinctiveness of peaks should increase as the number of independent replicates, and subsequently the maximum possible eigenvalue, increases. Due to its sensitivity to moderate parallel allele frequency shifts, this approach may be well-suited to evolve-and-resequence experimental setups (e.g. (Barghi et al., 2019; Rudman et al., 2021; Tusso et al., 2021)), where replication is common and allele frequency change may be recorded after only a few generations. As observed at the *ALX1* locus in galapagos finches, patterns indicative of parallelism can break down over time, so our approach is best-suited to intra-specific comparisons, or those among relatively young species groups. AF-vapeR can also be used with reduced-representation sequencing data (e.g. RAD-seq shown here), and also pool-sequencing data, given the input is population allele frequencies. Whilst the permutation approach accounts for sampling design, some sampling designs may bias eigenvalues more than others, particularly repeatedly sampling the same population when calculating AFV. Given potential upward biasing of eigenvalues when resampling, a greater stringency can be applied when identifying outliers, as demonstrated here when analysing the guppy dataset compared with the finch dataset. Datasets used here were filtered with a minor allele frequency (MAF) filter of 5%, but 1% may be sufficient. MAF-filtering is important so that SNP windows are not overwhelmed by many SNPs carrying minimal information.

700 We also recommend that candidates for parallelism be interpreted with respect to parallelism
701 and antiparallelism, as well as the underlying tree structure of individual genotypes. As
702 demonstrated here, visualising trees and genotypes at candidate regions can be informative in
703 understanding how replicates load onto eigenvectors and how populations cluster within
704 parallel or multi-parallel clusters (e.g. Figure S7-10). Further, per-SNP scores available in each
705 window's projected A matrix can be used to identify SNPs whose allele frequencies are driving
706 variance (e.g. Figure S7). Examining individual SNP scores is akin to examining the loading of
707 variables during principal component analysis (PCA) to assess which variables are best reflected
708 in each principal component. AF-vapeR is sensitive to geometrically parallel changes in allele
709 frequencies at the same base positions, and thus nonparallel changes in sites affecting the
710 same genes or loci will not be detected unless they have occurred in parallel within multiple,
711 separate replicates (multi-parallelism). If all replicates were to modify the same gene in
712 different ways (molecular convergence), this would not be expected to produce a signal of
713 parallelism, as seen in Figure 2C.
714
715 The study of repeatable evolution in the genome is increasingly popular in natural systems and
716 in experimental evolution. Progression of this field is essential for better understanding the
717 determinism, predictability, and contingencies linked with adaptation, with significant
718 implications for issues such as rapid adaptation to changing environments (Reid et al., 2016)
719 and the evolution of antibiotic resistance (A. C. Palmer & Kishony, 2013). There are also other
720 biological questions where the goal is similarly to identify shared haplotypes among groups,
721 such as identifying regions linked to Y/W chromosomes among members of the heterogametic
722 sex (D. H. Palmer et al., 2019). Despite this, there remain limited approaches tailored towards
723 detecting signatures of repeatability from genomic data. Our method requires only a standard
724 VCF file, a population map, and a description of vectors to calculate. The method should be
725 intuitive to any researcher familiar with common multivariate analyses such as PCA or principal
726 coordinate analysis (PCoA). We therefore expect AF-vapeR to become an important tool for
727 detecting convergence alongside current available approaches.
728
729 **ACKNOWLEDGEMENTS**
730 The authors would like to thank Matthew Webster (Uppsala University), and Andrew MacColl
731 (University of Nottingham) and Isabel Santos Magalhaes (University of Roehampton) for
732 insightful discussions regarding the finch and stickleback datasets respectively. The authors also
733 thank members of the University of Exeter's BS-popgen group and Sam Yeaman (University of
734 Calgary) and members of the Yeaman lab for providing feedback on the manuscript and project.
735 HPC infrastructure support was provided by The University of Exeter's High-Performance
736 Computing (HPC) facility (ISCA).
737

738 **DATA AVAILABILITY**
739 AF-vapeR has been made available for download at github
740 (https://github.com/JimWhiting91/afvaper). The code available for simulations and analyses of
741 simulations and empirical datasets is available at
742 (https://github.com/JimWhiting91/afvaper_analyses_ms). Sequencing reads associated with
743 the stickleback datasets are available at the European Nucleotide Archive (ENA) (accession:
744 PRJEB20851). Sequencing reads associated with the guppy dataset are available at the ENA
745 (accessions: PRJEB43917 and PRJEB10680) and the VCF used is available from FigShare
746 (doi:10.6084/m9.figshare.14315771). The VCF for the galapagos finch data was acquired from
747 Dryad (doi: 10.5061/dryad.7155d).
748
749 **AUTHOR CONTRIBUTIONS**
750 JRW and MJvZ conceived the project. JRW developed the software, and performed simulations,
751 analyses and figure production. JRP assisted software development through testing and
752 simulation development. JRW, JRP and BAF were involved in writing the manuscript. BAF
753 supervised the project.
754
755 **CONFLICTS OF INTEREST**
756 The authors declare no conflicts of interest.
757
758 **REFERENCES**

759 Adams, D. C., & Collyer, M. L. (2009). A general framework for the analysis of phenotypic
760 trajectories in evolutionary studies. *Evolution; International Journal of Organic Evolution*,
761 *63*(5), 1143–1154.

762 Ahrens, C. W., Rymer, P. D., Stow, A., Bragg, J., Dillon, S., Umbers, K. D. L., & Dudaniec, R. Y.
763 (2018). The search for loci under selection: trends, biases and progress. *Molecular Ecology*,
764 *27*(6), 1342–1356.

765 Arendt, J., & Reznick, D. (2008). Convergence and parallelism reconsidered: what have we
766 learned about the genetics of adaptation? *Trends in Ecology & Evolution*, *23*(1), 26–32.

767 Bailey, S. F., Guo, Q., & Bataillon, T. (2018). Identifying Drivers of Parallel Evolution: A

Regression Model Approach. *Genome Biology and Evolution*, *10*(10), 2801–2812.

Bailey, S. F., Rodrigue, N., & Kassen, R. (2015). The effect of selection environment on the probability of parallel evolution. *Molecular Biology and Evolution*, *32*(6), 1436–1448.

Barghi, N., Tobler, R., Nolte, V., Jakšić, A. M., Mallard, F., Otte, K. A., Dolezal, M., Taus, T., Kofler, R., & Schlötterer, C. (2019). Genetic redundancy fuels polygenic adaptation in Drosophila. *PLoS Biology*, *17*(2), e3000128.

Blount, Z. D., Lenski, R. E., & Losos, J. B. (2018). Contingency and determinism in evolution: Replaying life's tape. *Science*, *362*(6415). https://doi.org/10.1126/science.aam5979

Booker, T. R., Yeaman, S., & Whitlock, M. C. (2020). Variation in recombination rate affects detection of outliers in genome scans under neutrality. *Molecular Ecology*, *29*(22), 4274–4279.

Booker, T. R., Yeaman, S., & Whitlock, M. C. (2021). Global adaptation complicates the interpretation of genome scans for local adaptation. *Evolution Letters*, *5*(1), 4–15.

Burri, R. (2017). Interpreting differentiation landscapes in the light of long-term linked selection. *Evolution Letters*, *1*(3), 118–131.

Burri, R., Nater, A., Kawakami, T., Mugal, C. F., Olason, P. I., Smeds, L., Suh, A., Dutoit, L., Bureš, S., Garamszegi, L. Z., Hogner, S., Moreno, J., Qvarnström, A., Ružić, M., Sæther, S.-A., Sætre, G.-P., Török, J., & Ellegren, H. (2015). Linked selection and recombination rate variation drive the evolution of the genomic landscape of differentiation across the speciation continuum of Ficedula flycatchers. *Genome Research*, *25*(11), 1656–1665.

Chaves, J. A., Cooper, E. A., Hendry, A. P., Podos, J., De León, L. F., Raeymaekers, J. A. M., MacMillan, W. O., & Uy, J. A. C. (2016). Genomic variation at the tips of the adaptive

790 radiation of Darwin's finches. *Molecular Ecology*, *25*(21), 5282–5295.

791 Colosimo, P. F. (2005). Widespread Parallel Evolution in Sticklebacks by Repeated Fixation of

792 Ectodysplasin Alleles. *Science*, *307*(5717), 1928–1933.

793 Cruickshank, T. E., & Hahn, M. W. (2014). Reanalysis suggests that genomic islands of speciation

794 are due to reduced diversity, not reduced gene flow. *Molecular Ecology*, *23*(13), 3133–

795 3157.

796 Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., Handsaker, R. E.,

797 Lunter, G., Marth, G. T., Sherry, S. T., McVean, G., Durbin, R., & 1000 Genomes Project

798 Analysis Group. (2011). The variant call format and VCFtools. *Bioinformatics* , *27*(15),

799 2156–2158.

800 De Lisle, S. P., & Bolnick, D. I. (2020). A multivariate view of parallel evolution. *Evolution;*

801 *International Journal of Organic Evolution*, *74*(7), 1466–1481.

802 Dupressoir, A., Vernochet, C., Harper, F., Guégan, J., Dessen, P., Pierron, G., & Heidmann, T.

803 (2011). A pair of co-opted retroviral envelope syncytin genes is required for formation of

804 the two-layered murine placental syncytiotrophoblast. *Proceedings of the National*

805 *Academy of Sciences of the United States of America*, *108*(46), E1164–E1173.

806 Edelman, N. B., Frandsen, P. B., Miyagi, M., Clavijo, B., Davey, J., Dikow, R. B., García-Accinelli,

807 G., Van Belleghem, S. M., Patterson, N., Neafsey, D. E., Challis, R., Kumar, S., Moreira, G. R.

808 P., Salazar, C., Chouteau, M., Counterman, B. A., Papa, R., Blaxter, M., Reed, R. D., …

809 Mallet, J. (2019). Genomic architecture and introgression shape a butterfly radiation.

810 *Science*, *366*(6465), 594–599.

811 Ellegren, H., Smeds, L., Burri, R., Olason, P. I., Backström, N., Kawakami, T., Künstner, A.,

812 Mäkinen, H., Nadachowska-Brzyska, K., Qvarnström, A., Uebbing, S., & Wolf, J. B. W.
813 (2012). The genomic landscape of species divergence in Ficedula flycatchers. *Nature*,
814 *491*(7426), 756–760.
815 Erickson, P. A., Glazer, A. M., Killingbeck, E. E., Agoglia, R. M., Baek, J., Carsanaro, S. M., Lee, A.
816 M., Cleves, P. A., Schluter, D., & Miller, C. T. (2016). Partially repeatable genetic basis of
817 benthic adaptation in threespine sticklebacks. *Evolution; International Journal of Organic*
818 *Evolution*, *70*(4), 887–902.
819 Fang, B., Kemppainen, P., Momigliano, P., & Merilä, J. (2021). Population structure limits
820 parallel evolution in sticklebacks. *Molecular Biology and Evolution*.
821 https://doi.org/10.1093/molbev/msab144
822 Fraser, B. A., & Whiting, J. R. (2019). What can be learned by scanning the genome for
823 molecular convergence in wild populations? *Annals of the New York Academy of Sciences*.
824 https://doi.org/10.1111/nyas.14177
825 Gibbons, T. C., Metzger, D. C. H., Healy, T. M., & Schulte, P. M. (2017). Gene expression
826 plasticity in response to salinity acclimation in threespine stickleback ecotypes from
827 different salinity habitats. *Molecular Ecology*, *26*(10), 2711–2725.
828 Glazer, A. M., Killingbeck, E. E., Mitros, T., Rokhsar, D. S., & Miller, C. T. (2015). Genome
829 Assembly Improvement and Mapping Convergently Evolved Skeletal Traits in Sticklebacks
830 with Genotyping-by-Sequencing. *G3: Genes, Genomes, Genetics*, *5*(7), 1463–1472.
831 Guernsey, M. W., van Kruistum, H., Reznick, D. N., Pollux, B. J. A., & Baker, J. C. (2020).
832 Molecular Signatures of Placentation and Secretion Uncovered in Poeciliopsis Maternal
833 Follicles. *Molecular Biology and Evolution*, *37*(9), 2679–2690.

834 Haller, B. C., Galloway, J., Kelleher, J., Messer, P. W., & Ralph, P. L. (2019). Tree-sequence
835 recording in SLiM opens new horizons for forward-time simulation of whole genomes.
836 *Molecular Ecology Resources*, *19*(2), 552–566.
837 Haller, B. C., & Messer, P. W. (2019). SLiM 3: Forward Genetic Simulations Beyond the Wright–
838 Fisher Model. *Molecular Biology and Evolution*, *36*(3), 632–637.
839 Han, F., Lamichhaney, S., Grant, B. R., Grant, P. R., Andersson, L., & Webster, M. T. (2017). Gene
840 flow, ancient polymorphism, and ecological adaptation shape the genomic landscape of
841 divergence among Darwin's finches. *Genome Research*, *27*(6), 1004–1015.
842 Hermisson, J., & Pennings, P. S. (2017). Soft sweeps and beyond: understanding the patterns
843 and probabilities of selection footprints under rapid adaptation. *Methods in Ecology and*
844 *Evolution / British Ecological Society*, *8*(6), 700–716.
845 Hirase, S., Ozaki, H., & Iwasaki, W. (2014). Parallel selection on gene copy number variations
846 through evolution of three-spined stickleback genomes. *BMC Genomics*, *15*, 735.
847 Hohenlohe, P. A., Bassham, S., Etter, P. D., Stiffler, N., Johnson, E. A., & Cresko, W. A. (2010).
848 Population genomics of parallel adaptation in threespine stickleback using sequenced RAD
849 tags. *PLoS Genetics*, *6*(2). https://doi.org/10.1371/journal.pgen.1000862
850 Jones, F. C., Grabherr, M. G., Chan, Y. F., Russell, P., Mauceli, E., Johnson, J., Swofford, R., Pirun,
851 M., Zody, M. C., White, S., Birney, E., Searle, S., Schmutz, J., Grimwood, J., Dickson, M. C.,
852 Myers, R. M., Miller, C. T., Summers, B. R., Knecht, A. K., … Kingsley, D. M. (2012). The
853 genomic basis of adaptive evolution in threespine sticklebacks. *Nature*, *484*(7392), 55–61.
854 Lamichhaney, S., Berglund, J., Almén, M. S., Maqbool, K., Grabherr, M., Martinez-Barrio, A.,
855 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.

Lamichhaney, S., Han, F., Berglund, J., Wang, C., Almén, M. S., Webster, M. T., Grant, B. R., Grant, P. R., & Andersson, L. (2016). A beak size locus in Darwin's finches facilitated character displacement during a drought. *Science*, *352*(6284), 470–474.

Láruson, Á. J., Yeaman, S., & Lotterhos, K. E. (2020). The Importance of Genetic Redundancy in Evolution. *Trends in Ecology & Evolution*, *35*(9), 809–822.

Lawson, L. P., & Petren, K. (2017). The adaptive genomic landscape of beak morphology in Darwin's finches. *Molecular Ecology*, *26*(19), 4978–4989.

Lee, K. M., & Coop, G. (2017). Distinguishing Among Modes of Convergent Adaptation Using Population Genomic Data. *Genetics*, *207*(4), 1591–1619.

Lee, K. M., & Coop, G. (2019). Population genomics perspectives on convergent adaptation. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *374*(1777), 20180236.

Leigh, D. M., Lischer, H. E. L., Guillaume, F., Grossen, C., & Günther, T. (2021). Disentangling adaptation from drift in bottlenecked and reintroduced populations of Alpine ibex. *Molecular Ecology Resources*. https://doi.org/10.1111/1755-0998.13442

Le Moan, A., Gagnaire, P.-A., & Bonhomme, F. (2016). Parallel genetic divergence among coastal-marine ecotype pairs of European anchovy explained by differential introgression after secondary contact. *Molecular Ecology*, *25*(13), 3187–3202.

Lotterhos, K. E. (2019). The Effect of Neutral Recombination Variation on Genome Scans for Selection. *G3* , *9*(6), 1851–1867.

878 Lotterhos, K. E., & Whitlock, M. C. (2015). The relative power of genome scans to detect local
879 adaptation depends on sampling design and statistical method. *Molecular Ecology*, *24*(5),
880 1031–1046.
881 Magalhaes, I. S., D'Agostino, D., Hohenlohe, P. A., & MacColl, A. D. C. (2016). The ecology of an
882 adaptive radiation of three-spined stickleback from North Uist, Scotland. *Molecular*
883 *Ecology*, *25*(17), 4319–4336.
884 Magalhaes, I. S., Whiting, J. R., D'Agostino, D., Hohenlohe, P. A., Mahmud, M., Bell, M. A.,
885 Skúlason, S., & MacColl, A. D. C. (2020). Intercontinental genomic parallelism in multiple
886 three-spined stickleback adaptive radiations. *Nature Ecology & Evolution*.
887 https://doi.org/10.1038/s41559-020-01341-8
888 Manceau, M., Domingues, V. S., Linnen, C. R., Rosenblum, E. B., & Hoekstra, H. E. (2010).
889 Convergence in pigmentation at multiple levels: mutations, genes and function.
890 *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*,
891 *365*(1552), 2439–2450.
892 Miller, C. T., Glazer, A. M., Summers, B. R., Blackman, B. K., Norman, A. R., Shapiro, M. D., Cole,
893 B. L., Peichel, C. L., Schluter, D., & Kingsley, D. M. (2014). Modular Skeletal Evolution in
894 Sticklebacks Is Controlled by Additive and Clustered Quantitative Trait Loci. *Genetics*,
895 *197*(1), 405–420.
896 Morales, H. E., Faria, R., Johannesson, K., Larsson, T., Panova, M., Westram, A. M., & Butlin, R.
897 K. (2019). Genomic architecture of parallel ecological divergence: Beyond a single
898 environmental contrast. *Science Advances*, *5*(12), eaav9963.
899 Morris, J., Navarro, N., Rastas, P., Rawlins, L. D., Sammy, J., Mallet, J., & Dasmahapatra, K. K.

(2019). The genetic architecture of adaptation: convergence and pleiotropy in Heliconius wing pattern evolution. *Heredity*, *123*(2), 138–152.

Nelson, T. C., & Cresko, W. A. (2018). Ancient genomic variation underlies repeated ecological adaptation in young stickleback populations. *Evolution Letters*, *2*(1), 9–21.

Nery, M. F., Borges, B., Dragalzew, A. C., & Kohlsdorf, T. (2016). Selection on different genes with equivalent functions: the convergence story told by Hox genes along the evolution of aquatic mammalian lineages. *BMC Evolutionary Biology*, *16*(1), 113.

Olazcuaga, L., Loiseau, A., Parrinello, H., Paris, M., Fraimout, A., Guedot, C., Diepenbrock, L. M., Kenis, M., Zhang, J., Chen, X., Borowiec, N., Facon, B., Vogt, H., Price, D. K., Vogel, H., Prud'homme, B., Estoup, A., & Gautier, M. (2020). A Whole-Genome Scan for Association with Invasion Success in the Fruit Fly Drosophila suzukii Using Contrasts of Allele Frequencies Corrected for Population Structure. *Molecular Biology and Evolution*, *37*(8), 2369–2385.

Orozco-terWengel, P., Kapun, M., Nolte, V., Kofler, R., Flatt, T., & Schlötterer, C. (2012). Adaptation of Drosophila to a novel laboratory environment reveals temporally heterogeneous trajectories of selected alleles. *Molecular Ecology*, *21*(20), 4931–4941.

Palmer, A. C., & Kishony, R. (2013). Understanding, predicting and manipulating the genotypic evolution of antibiotic resistance. *Nature Reviews. Genetics*, *14*(4), 243–248.

Palmer, D. H., Rogers, T. F., Dean, R., & Wright, A. E. (2019). How to identify sex chromosomes and their turnover. *Molecular Ecology*, *28*(21), 4709–4724.

Peichel, C. L., & Marques, D. A. (2017). The genetic and molecular architecture of phenotypic diversity in sticklebacks. *Philosophical Transactions of the Royal Society of London. Series B,*

922 *Biological Sciences*, *372*(1713). https://doi.org/10.1098/rstb.2015.0486

923 Perrier, C., Ferchaud, A.-L., Sirois, P., Thibault, I., & Bernatchez, L. (2017). Do genetic drift and
924 accumulation of deleterious mutations preclude adaptation? Empirical investigation using
925 RADseq in a northern lacustrine fish. *Molecular Ecology*, *26*(22), 6317–6335.

926 Reid, N. M., Proestou, D. A., Clark, B. W., Warren, W. C., Colbourne, J. K., Shaw, J. R., Karchner,
927 S. I., Hahn, M. E., Nacci, D., Oleksiak, M. F., Crawford, D. L., & Whitehead, A. (2016). The
928 genomic landscape of rapid repeated evolutionary adaptation to toxic pollution in wild
929 fish. *Science*, *354*(6317), 1305–1308.

930 Roberts Kingman, G. A., Vyas, D. N., Jones, F. C., Brady, S. D., Chen, H. I., Reid, K., Milhaven, M.,
931 Bertino, T. S., Aguirre, W. E., Heins, D. C., von Hippel, F. A., Park, P. J., Kirch, M., Absher, D.
932 M., Myers, R. M., Di Palma, F., Bell, M. A., Kingsley, D. M., & Veeramah, K. R. (2021).
933 Predicting future from past: The genomic basis of recurrent and rapid stickleback
934 evolution. *Science Advances*, *7*(25). https://doi.org/10.1126/sciadv.abg5285

935 Rudman, S. M., Greenblum, S. I., Rajpurohit, S., Betancourt, N. J., Hanna, J., Tilk, S., Yokoyama,
936 T., Petrov, D. A., & Schmidt, P. (2021). Direct observation of adaptive tracking on ecological
937 timescales in Drosophila. In *bioRxiv* (p. 2021.04.27.441526).
938 https://doi.org/10.1101/2021.04.27.441526

939 Schlamp, F., van der Made, J., Stambler, R., Chesebrough, L., Boyko, A. R., & Messer, P. W.
940 (2016). Evaluating the performance of selection scans to detect selective sweeps in
941 domestic dogs. *Molecular Ecology*, *25*(1), 342–356.

942 Stankowski, S., Chase, M. A., Fuiten, A. M., Rodrigues, M. F., Ralph, P. L., & Streisfeld, M. A.
943 (2019). Widespread selection and gene flow shape the genomic landscape during a

944 radiation of monkeyflowers. *PLoS Biology*, *17*(7), e3000391.

945 Stern, D. L. (2013). The genetic causes of convergent evolution. *Nature Reviews. Genetics*,
946 *14*(11), 751–764.

947 Storz, J. F. (2016). Causes of molecular convergence and parallelism in protein evolution. *Nature*
948 *Reviews. Genetics*, *17*(4), 239–250.

949 Stuart, Y. E. (2019). Divergent Uses of "Parallel Evolution" during the History of The American
950 Naturalist. *The American Naturalist*, *193*(1), 11–19.

951 Stuart, Y. E., Veen, T., Weber, J. N., Hanson, D., Ravinet, M., Lohman, B. K., Thompson, C. J.,
952 Tasneem, T., Doggett, A., Izen, R., Ahmed, N., Barrett, R. D. H., Hendry, A. P., Peichel, C. L.,
953 & Bolnick, D. I. (2017). Contrasting effects of environment and genetics generate a
954 continuum of parallel evolution. *Nature Ecology & Evolution*, *1*(6), 158.

955 Teshima, K. M., Coop, G., & Przeworski, M. (2006). How reliable are empirical genomic scans for
956 selective sweeps? *Genome Research*, *16*(6), 702–712.

957 Tusso, S., Nieuwenhuis, B. P. S., Weissensteiner, B., Immler, S., & Wolf, J. B. W. (2021).
958 Experimental evolution of adaptive divergence under varying degrees of gene flow. *Nature*
959 *Ecology & Evolution*, *5*(3), 338–349.

960 vonHoldt, B. M., Kartzinel, R. Y., Huber, C. D., Le Underwood, V., Zhen, Y., Ruegg, K., Lohmueller,
961 K. E., & Smith, T. B. (2018). Growth factor gene IGF1 is associated with bill size in the black-
962 bellied seedcracker Pyrenestes ostrinus. *Nature Communications*, *9*(1), 4855.

963 Wang, B., Mojica, J. P., Perera, N., Lee, C.-R., Lovell, J. T., Sharma, A., Adam, C., Lipzen, A., Barry,
964 K., Rokhsar, D. S., Schmutz, J., & Mitchell-Olds, T. (2019). Ancient polymorphisms
965 contribute to genome-wide variation by long-term balancing selection and divergent

966 sorting in Boechera stricta. *Genome Biology*, *20*(1), 126.

967 Wang, L., Josephs, E. B., Lee, K. M., Roberts, L. M., Rellán-Álvarez, R., Ross-Ibarra, J., & Hufford,
968 M. B. (2021). Molecular Parallelism Underlies Convergent Highland Adaptation of Maize
969 Landraces. *Molecular Biology and Evolution*. https://doi.org/10.1093/molbev/msab119

970 Wark, A. R., Mills, M. G., Dang, L.-H., Chan, Y. F., Jones, F. C., Brady, S. D., Absher, D. M.,
971 Grimwood, J., Schmutz, J., Myers, R. M., Kingsley, D. M., & Peichel, C. L. (2012). Genetic
972 architecture of variation in the lateral line sensory system of threespine sticklebacks. *G3* ,
973 *2*(9), 1047–1056.

974 Whiting, J. R., Paris, J. R., van der Zee, M. J., Parsons, P. J., Weigel, D., & Fraser, B. A. (2021).
975 Drainage-structuring of ancestral variation and a common functional pathway shape
976 limited genomic convergence in natural high- and low-predation guppies. *PLoS Genetics*,
977 *17*(5), e1009566.

978 Wollenberg Valero, K. C. (2020). Aligning functional network constraint to evolutionary
979 outcomes. *BMC Evolutionary Biology*, *20*(1), 58.

980 Yeaman, S., Gerstein, A. C., Hodgins, K. A., & Whitlock, M. C. (2018). Quantifying how
981 constraints limit the diversity of viable routes to adaptation. *PLoS Genetics*, *14*(10),
982 e1007717.

983