# 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

**Word Count: 8146**

## 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. Further, the shape of the eigenvalue distribution can reveal multiple axes of parallelism within datasets.

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 detect 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.

**Keywords:**

Parallel evolution, parallelism, genome scan, repeatability, convergence, allele frequency

34 **INTRODUCTION**

35 Studies of convergence, where separate lineages evolve similar phenotypes, provide compelling
36 evidence of natural selection and have been a prominent focus of research in the last decade
37 (Stern 2013; Storz 2016; Blount, Lenski, and Losos 2018; Arendt and Reznick 2008). Common
38 practice in population genomics studies of repeatability involves sequencing populations in
39 replicated environments, performing selection scans, and comparing candidate loci across
40 replicates (the overlapping outlier approach reviewed in (Fraser and Whiting 2019)). Identified
41 selection candidates are then probed for theories of genetic convergence; for example their
42 mode of convergence (Lee and Coop 2017), the role of genetic and pleiotropic constraints
43 (Morris et al. 2019; Wollenberg Valero 2020), the influence of redundancy in the genotype-
44 phenotype map (Yeaman et al. 2018; Láruson, Yeaman, and Lotterhos 2020), or the effect of
45 gene flow (Fang et al. 2021). It is also of interest whether repeatability occurs through: 1)
46 selection at the same SNP (often termed 'parallel' evolution), as seen in stickleback (Colosimo
47 2005) and maize (L. Wang et al. 2021); 2) nonidentical changes in the same genes or loci
48 ('convergence') e.g. repeated modification of *Mc1r* during pigmentation evolution (Manceau et
49 al. 2010); 3) changes in different genes of similar function ('functional convergence') e.g.
50 developmental HOX genes across aquatic mammals (Nery et al. 2016); or 4) using no
51 comparable genetic machinery whatsoever ('divergence').

52

Genome scans are a fundamental tool in identifying locally-adapted genomic regions, but scrutiny has increased in response to concerns of false-positives, biases, and interpretation of outliers (Booker et al., 2021; Cruickshank & Hahn, 2014; Lotterhos & Whitlock, 2015). False-positives may be non-randomly associated with shared recombination landscapes between lineages (Booker et al., 2020; Lotterhos, 2019), but random false-positives could be due to drift (Leigh et al., 2021), or a tendency towards detecting hard-sweeps over soft-sweeps (Teshima et al., 2006). Generally, the overlapping outlier test assumes false-positives are randomly distributed and they should not overlap among replicate lineages. However, over-conservative cut-offs or structure corrections subsequently may mask overlapping true positives and inflate the false-negative rate (FNR) in convergence studies. This implies that some mechanisms of repeatable adaptation may be more difficult to detect, which is concerning given most adaptation may occur through minor/moderate allele frequency change (soft sweeps) (Hermisson & Pennings, 2017).

Many definitions of convergence/parallelism exist (Stuart, 2019), including those which consider these as geometric processes in multivariate space (Adams & Collyer, 2009; De Lisle & Bolnick, 2020). Applying these here, ‘convergence’ is a process where distinct evolutionary starting points converge on a common endpoint. ‘Parallelism’ is a process where distinct evolutionary start points change in the same way along parallel trajectories to arrive at distinct

endpoints. Parallel lineages, therefore, may not evolve to become more similar, contrary to convergent lineages. At the genetic level, parallelism can 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 regions. 'Multi-parallelism' may also exist, where multiple parallel trajectories, each with 'parallel' lineages, are observed in the same genomic region. Multi-parallelism can therefore be analogous to definitions of 'convergence' involving 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 (summarised in Figure S1) 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, bypassing the requirement to first identify local adaptation candidates in replicate datasets. We first outline the background of this approach in detail, before demonstrating its usage and comparing it to $F_{ST}$-based overlapping outlier approaches using simulations, 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), and extends their framework for deriving a single phenotypic eigenvalue distribution to many allele frequency-derived eigenvalue distributions across the genome. These are then compared against randomly permuted distributions across the genome.

SNP data is first split into non-overlapping windows (Figure 1A) of equal user-defined SNP count, within which we calculate per-SNP allele frequency change between multiple population pairs. Population pairs are defined 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. These could be unique 'replicate pairs' (each A-B population is unique); a single A to many B ('common' ancestor); or vectors between a pool of A and B populations sampled with replacement ('multi-comparisons'). These setups reflect sampling designs common to studies of genetic convergence (Fraser & Whiting, 2019).

The process 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; termed the allele frequencies vector (**AFV**) (equation 2). The final matrix **X** includes the normalised **AFV** for all population pairs 1..*m*, where normalisation involves dividing each vector by its length to a unit length of one (equation 3), such that all population pairs exhibit equivalent variance. 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**).

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

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

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

(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 window (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).

$$\mathrm{C} = \mathrm{X}\mathrm{X}^{\mathrm{T}} \tag{4}$$

$$\mathrm{C} = \mathrm{Q}\mathrm{V}\mathrm{Q}^{-1} \tag{5}$$

Because the **AFV** are normalised to unit length, all eigenvalues in a window sum to *m*, so we are interested in identifying regions of the genome in which the eigenvalue distribution is positively skewed and weighted to variance captured by only the first (or first few) eigenvalues (Figure 1C). The raw eigenvalues associated with the *m*-1 eigenvectors can thus 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.

$$\mathrm{A} = \mathrm{X}^{\mathrm{T}}\mathrm{Q}^{\mathrm{T}\ -1} \tag{6}$$

142 To develop a null expectation for extreme eigenvalues, we repeatedly reassign individuals to
143 random populations, maintaining the original vector design strategy (Figure 1G), and
144 recalculate random **AFV** across the genome. This approach captures potential biasing of
145 eigenvalues across regions of the genome, for example due to recombination and linkage
146 variation, and bias associated with vector sampling design that would otherwise fail to be
147 accounted for by a one-size-fits-all null model such as the Wishart distribution (De Lisle &
148 Bolnick, 2020).

149

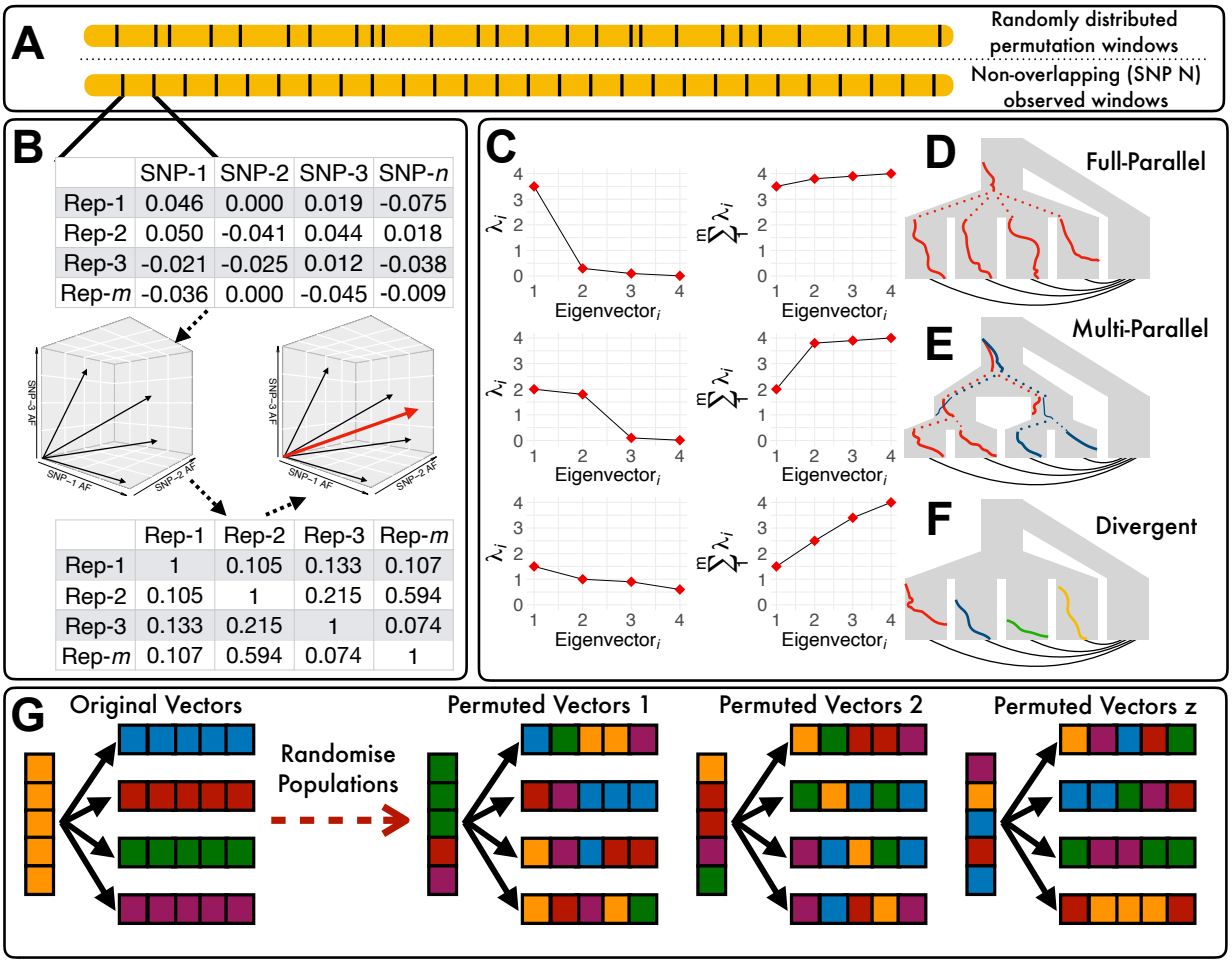



***Figure 1:*** *Methodology of AF-vapeR. Chromosomes are windowed by SNP count (**A**). Normalised AFV are calculated for all population pairs within a window (**B**) prior to calculating the correlation matrix and eigen decomposition to yield eigenvectors (red arrow as example). Eigenvalues (λ) are summed along consecutive eigenvectors (1..m) (**C**), with large summed values highlighting that a majority of variance is constrained to a few eigenvectors; indicative of parallelism. Distribution shape is indicative of different evolutionary processes (**D-F**), with large primary eigenvectors highlighting full parallelism, large second eigenvalues highlighting two*

158 *axes of parallelism, and no large eigenvalues suggesting divergence/non-parallelism. Different*
159 *evolutionary processes are shown as increasing in frequency from left to right within lineages.*
160 *Dotted lines show inheritance through lineage branching. Colours denote shared mutations, and*
161 *black lines connecting lineages show AFV pairs. The permutation process is illustrated at the*
162 *bottom (**G**), where fixed population assignment, denoted as block colour, is repeatedly shuffled*
163 *and AFV are calculated under the same vector design for a total of* z *permutations. Individuals*
164 *are squares and each column/row represents a population, with AFV shown as black arrows*
165 *between the vertical population to all four horizontal populations in a common-ancestor design.*

166

167 **Simulation methods**

168 Simulation methods are described in full in the supplementary materials. Briefly, to
169 demonstrate functionality and explore the effectiveness of AF-vapeR over parameter space, we
170 used a Wright-Fisher model with the forward-simulation software SLiM (v3.4 (Haller & Messer,
171 2019)). The goals of these simulations were to: 1) Demonstrate the power to detect parallelism
172 under different demographies and selection contexts; 2) Demonstrate the ability to detect
173 different evolutionary scenarios (Figure 1D-F); and 3) Compare *AF-vapeR* to the commonly
174 applied approach of overlapping $F_{ST}$ outlier windows among replicates.

175

176 We simulated a five-population metapopulation with an ancestral founding population
177 (N=10,000) and four daughters (N=400, 250, or 10,000). After founding, each daughter
178 experienced 200 generations of selection on an ancestrally-neutral mutation inherited at low-
179 frequency from the ancestral population. Either: 1) all daughters inherited and swept the same
180 mutation (full-parallelism); 2) daughters were paired, and each pair swept their pair-specific
181 mutations (multi-parallelism); or 3) all daughters swept their own mutation (non-
182 parallelism/divergence) (Figure S2). We manipulated the metapopulation by modifying: the size
183 of all daughter populations (to encourage drift), the reciprocal migration rate between the
184 ancestral and all four daughters ($M$ = 0, 0.0025, 0.01), and the selection coefficient applied to
185 the focal mutation in all daughters ($s$ = 0.01, 0.05, 0.1). Factorially combining these parameters
186 yielded 27 treatment groups. Separately, we explored how positioning the focal mutation in
187 regions of low, medium, and high recombination, affected results.

188

189 AF-vapeR (window sizes of 50, 200, 500 SNPs) was compared to $F_{ST}$ scans (10kb windows) in
190 false-positive rate (FPR) and false-negative rate (FNR) with variable cutoffs ($\alpha$ = 0.95, 0.99,
191 0.999). FPR reflected the proportion of falsely-identified windows beyond the selected
192 mutation averaged across all 100 iterations. FNR reflected the number of iterations that failed
193 to identify parallelism at the focal mutation. All simulation parameters are in Table S1.

194

195 **Empirical dataset demonstrations**

196 We chose three publicly available datasets to demonstrate different utilities of AF-vapeR. These
197 include restriction-site associated digestion (RAD) sequencing data from 18 freshwater-adapted
198 three-spined stickleback populations British Columbia (BC) and an ancestral marine population
199 (Magalhaes et al., 2016, 2020); whole-genome sequencing (WGS) of five pairs of high-predation
200 (HP) and low-predation (LP) adapted guppies (Whiting et al., 2021); and WGS data from 12
201 species of Galapagos finches, used previously to explore small vs large, and pointed vs blunt,
202 beak morphologies within the adaptive radiation (Han et al., 2017; Lamichhaney et al., 2015,
203 2016). These datasets therefore include different sequencing approaches (RAD-seq vs WGS),
204 different sampling approaches (replicate pairs vs common ancestor vs multiple comparisons;
205 see (Fraser & Whiting, 2019)), and a mix of within- and across-species comparisons.

206

207 **RESULTS**

208 **Simulation results**

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

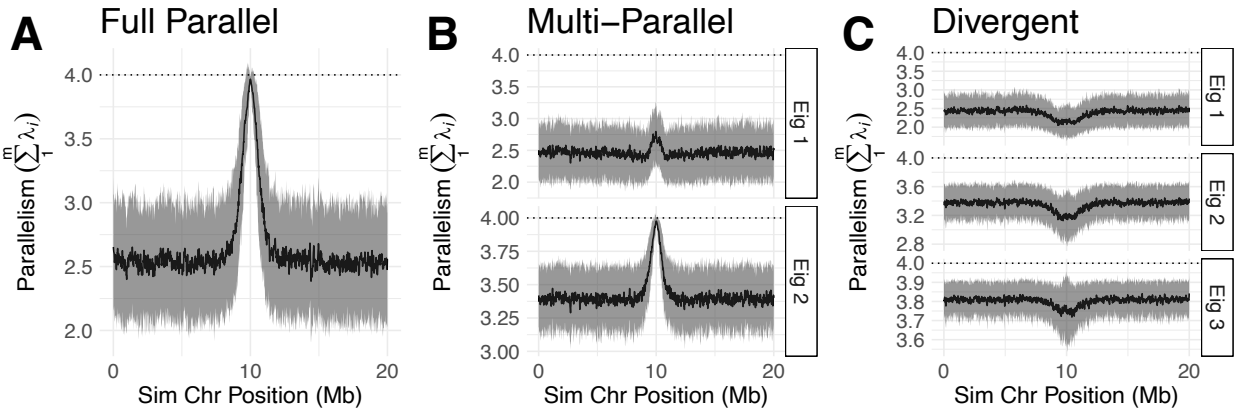



***Figure 2:*** *Simulated eigenvalue peaks (with standard deviation) detected by AF-vapeR under three different evolutionary scenarios: full parallelism (**A**), multi-parallelism (**B**), and divergence (**C**). The maximum eigenvalue is 4. For each scenario we show the eigenvalues up until a peak is observed. The simulation parameters shown are for* s *= 0.05, M = 0.025, and DP = 10,000.*

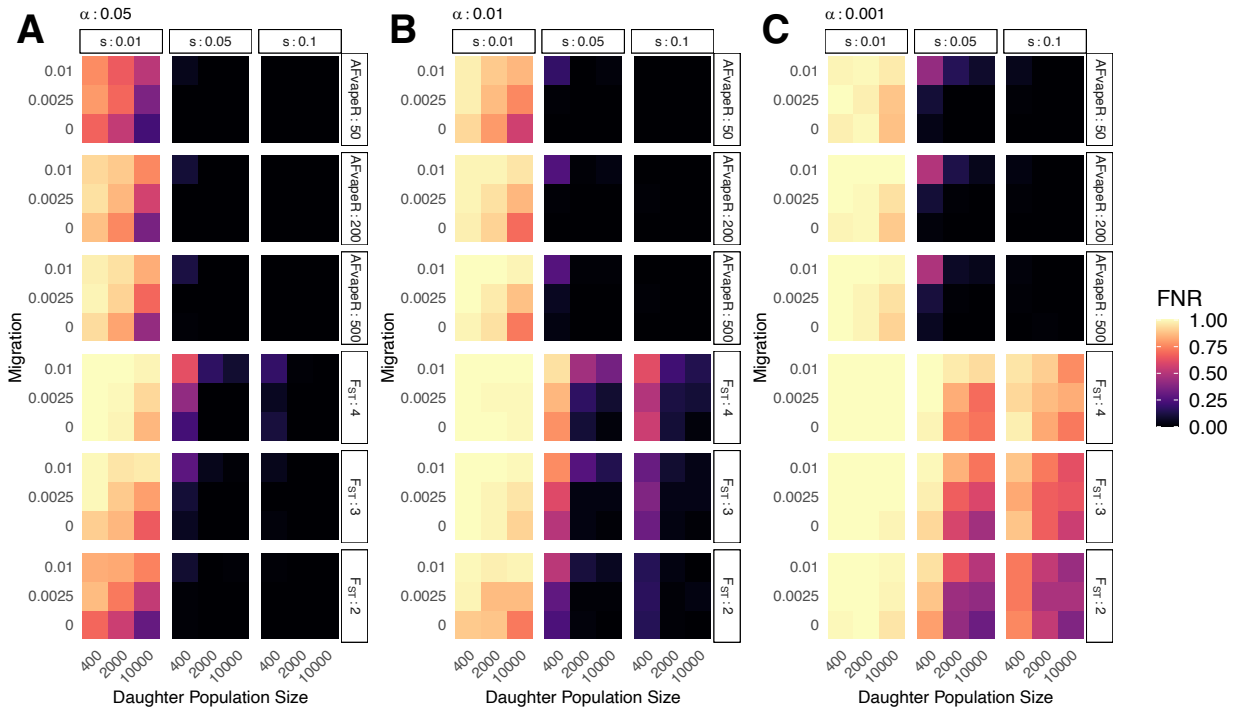
We then looked at error rates associated with detecting full parallelism. Across parameter space, at α = 0.01, FPR were low for all AF-vapeR window sizes (50 SNPs, max FPR = 0.057; 200 SNPs, max FPR = 0.072; 500 SNPs, max FPR = 0.085) and all overlapping $F_{ST}$ thresholds (all 4 daughters, max FPR = 1.1e-4; any 3, max FPR = 7.5e-4; any 2, max FPR = 4.1e-3) (Figure S3). Low $F_{ST}$ FPR suggests false-positives are randomly distributed within daughters. AF-vapeR FPR was positively associated with daughter population sizes and negatively associated with migration rate (Figure S3).

227

228 Across parameter space, FNR for all AF-vapeR window analyses consistently out-performed the
229 strictest and most applicable $F_{ST}$ analysis (outlier in all 4 daughters), and often out-performed
230 all $F_{ST}$ overlap cut-offs (Figure 3). Expectedly, strong bottlenecks (*DP* = 400, first column of
231 panels in Figure 3), and strong migration (*M* = 0.01, first row of panels in Figure 3) led to
232 elevated FNR for $F_{ST}$ analyses, but AF-vapeR was generally more robust to these parameters.

233

234




235

236 ***Figure 3:*** *FNR associated with AF-vapeR and $F_{ST}$ outliers tests. Facet labels correspond to AF-*
237 *vapeR window size (50, 200, and 500 SNPs) or number of overlapping populations for $F_{ST}$ (i.e. 2 =*
238 *an outlier in any 2 of the 4 populations, 3 = an outlier in any 3, 4 = outlier in all four). Panels* ***A, B***
239 *and* ***C*** *show FNR calculated using quantile cutoffs of 95%, 99%, and 99.9% respectively. FNR was*
240 *averaged over 100 iterations within cells. Facet columns = strength of selection on beneficial*
241 *mutation, y = migration rate, and x = daughter population size.*

242

243 For all $F_{ST}$ overlapping cut-offs, weak selection (*s* = 0.01) often prevented identification of full-
244 parallelism. Even using the least stringent cut-off, only 15% of simulations detected full-
245 parallelism (α = 0.05, *s* = 0.01, $F_{ST}$ outlier in all 4; min FNR = 0.85, max FNR = 1.0). In contrast,
246 under the same conditions AF-vapeR recovered full parallelism in at least 21% and at most 78%
247 of iterations under all nine parameter combinations of daughter population size, migration and
248 *s* = 0.01 (α = 0.05, *s* = 0.01, AF-vapeR window size = 50; min FNR = 0.22, max FNR = 0.79).
249 Detecting parallelism candidates as $F_{ST}$ outliers in any two daughters improved FNR under weak
250 selection (α = 0.05, *s* = 0.01, $F_{ST}$ outlier in any 2; min FNR = 0.3, max FNR = 0.86). This approach,
251 however, still does not perform as well as AF-vapeR and fails to capture the simulation
252 conditions of parallel selection in all four daughters.

253

254 Under stronger selection (*s* = 0.05, 0.1), AF-vapeR exhibits consistently low FNR even as cut-off
255 stringency increases (Figure 3A-C). Conversely, increasing cut-off stringency dramatically
256 increases $F_{ST}$-based FNR across all parameter space. This behaviour of AF-vapeR is important
257 because it indicates that FPR at weak cut-offs can be ameliorated with stricter cut-offs at a
258 minimal cost to false-negatives. Only with high migration and small population sizes (top left of
259 parameter grids) did FNR increase with outlier stringency. Researchers should therefore
260 consider the role of migration and population size when applying the method under such
261 scenarios.

262

263 AF-vapeR was found to be more sensitive to lower allele frequency differentiation (AFD), as FNR
264 reduced more rapidly with increasing AFD (Figure S4), which may explain the observed
265 reduction in FNR when migration is high and selection is weak (low AFD). The reduced FNR seen
266 when drift is strongest (DP = 400) is better explained by random drift at neutral sites reducing
267 the background signal of parallelism and making peaks at selected sites more distinct.

268

269 Overall, these results show that AF-vapeR has comparably low FPR to overlapping multiple $F_{ST}$
270 outlier sets, and has a generally reduced FNR across parameter space. AF-vapeR exhibits a
271 considerably reduced FNR under drift, weak selection and high migration. AF-vapeR
272 consistently outperformed $F_{ST}$ when $F_{ST}$ outliers must be observed in all four daughters, and

273 thus performs better at capturing the simulation conditions of all daughters experiencing
274 parallel selection. FNR were typically reduced with reduced window size for AF-vapeR with only
275 small increases in FPR, suggesting the effect of noise is marginal when balancing signal and
276 noise.

277

278 Comparable results were observed when running the simulation under a multi-parallelism
279 scenario when looking for outliers at the selected site on the second eigenvector. In these
280 simulations, as above, FPR was generally low (Figure S5), migration rate was positively
281 associated with FNR, and selection coefficient was negatively associated with FNR (Figure S6).
282 Cut-off stringency also behaved similarly, obscuring $F_{ST}$-based analyses at higher stringencies.
283 These results demonstrate that AF-vapeR is able to robustly identify parallelism candidates
284 across parameter space whilst also inferring that multiple axes of parallelism exist.

285

286 Increasing recombination from 1e-9 to 1e-7 cM/bp increased the width of eigenvalue peaks and
287 marginally reduced peak-height (Figure S7A). This was not affected by window SNP count.
288 Eigenvalue variance at neutral sites increased by approximately 50% with a 100-fold reduction
289 in recombination (Figure S7B-C), however peaks of full-parallelism were still clearly observable
290 relative to the neutral backdrop regardless of local recombination rate.

291

292 **Empirical Results: Marine-Freshwater stickleback**

293 The goal for this dataset was to detect *a-priori* loci associated with parallel evolution from RAD-
294 sequencing data. We applied AF-vapeR to a dataset of 18 freshwater-adapted populations and
295 a marine population in a 'common-ancestor' sampling strategy (Figure 4A) sampled from the
296 northeast Pacific coast in British Columbia (Magalhaes et al., 2020) (Eighteen marine-
297 freshwater AFV). Due to low SNP density (total = 26,658), we used windows of 10 SNPs
298 (N=2376) with a median physical window size of 116.0 kb. At this resolution, the genomic
299 landscape of full-parallelism, defined by the per-window primary eigenvalues (Figure 4B),
300 closely resembled the genomic landscape of marine-freshwater divergence commonly observed
301 in this system (Fang et al., 2021; Hohenlohe et al., 2010; Roberts Kingman et al., 2021). In
302 particular, chromosome 4 exhibited a clear chromosome-level increase in parallelism (Figure
303 4C-D). Chromosomes 21, 12, 20 and 7 exhibited the next highest median primary eigenvalues
304 (Figure 4D). With the exception of chr12, each of these are enriched for QTL associated with
305 marine-freshwater phenotypes (Peichel & Marques, 2017).

306

307

308

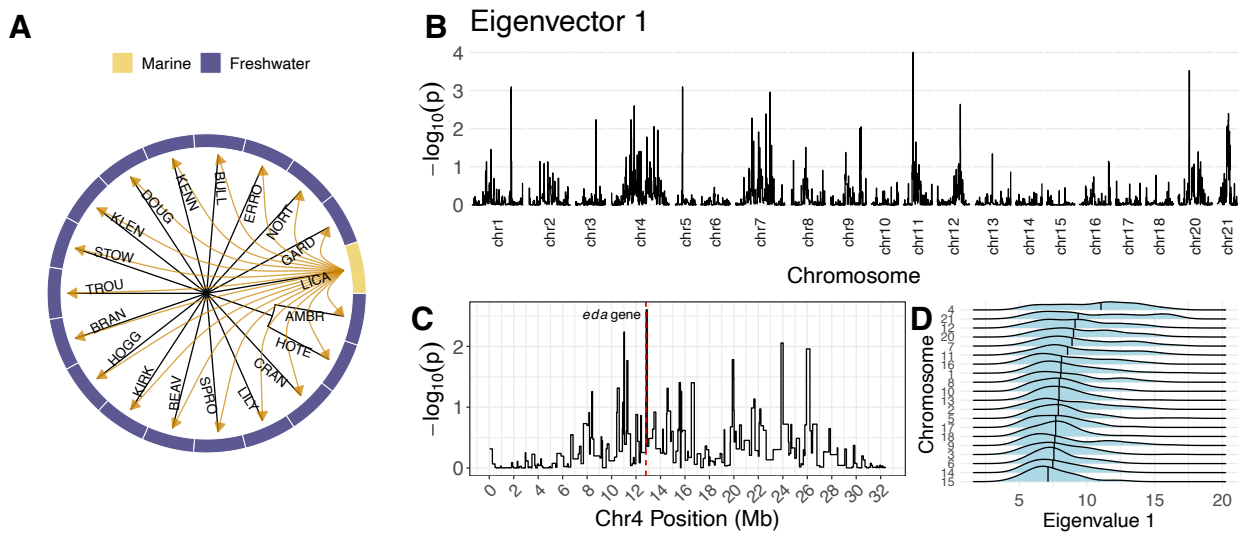



***Figure 4:*** *AF-vapeR results using a RAD-seq dataset for marine-freshwater stickleback from British Columbia. Sampling design is described based on the tree (**A**), with all vectors calculated from the marine population (LICA) to all freshwater populations. The genome-wide distribution of eigenvalue 1 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. 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.*

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. All windows included previously identified candidates for freshwater adaptation

322 (discussed in full in the supplementary results). Above the 99% cutoff, we recovered an
323 additional 17 windows, including the window adjacent to the well-known *eda* gene associated
324 with parallel marine-freshwater armour haplotypes. The window containing *eda* exhibited a
325 weaker signal of full parallelism, likely because three of the 18 freshwater populations exhibit
326 full-armour plating associated with marine fish (Magalhaes et al., 2020). Ten windows exceeded
327 the 99% cutoff on eigenvector 2; none were above 99.9%. For these windows, eigenvalue 1
328 made up the majority of the summed total, with eigenvalue 2 explaining 15.32% or less of the
329 total variance among AFV. This indicates that these windows represent cases in which most, but
330 not all, freshwater populations are diverging in parallel on eigenvector 1, but additional axes
331 exist to explain non-parallel allele frequency change in a smaller number of replicates. All
332 significant windows on eigenvectors 1 and 2 are summarised in Tables S2-3.

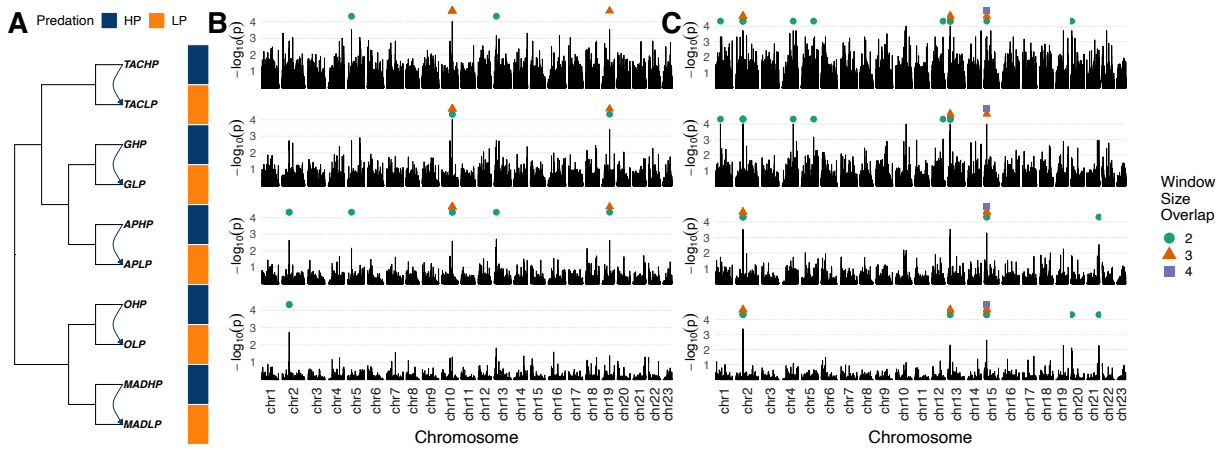
333

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

335 The motivation for using this dataset was to detect previously identified (Whiting et al., 2021)
336 and novel regions associated with repeated HP-LP adaptation in guppies. AF-vapeR was run
337 over WGS data from five HP-LP population pairs (Figure 5A), each from a separate river, using
338 window sizes of 25, 50, 100, 200 SNPs. Variable window sizes were used to evaluate how robust
339 outliers were to changes in this parameter. Median physical window sizes were 4.8, 10.3, 21.4,
340 and 43.9 kb respectively. We took windows that exceeded null cut-offs for at least two window

341 sizes to increase confidence in candidates, and to ensure spurious signals linked to window size
342 choice were removed. For window sizes of 25 and 50 SNPs we used a quantile cut-off of 99.9%,
343 but relaxed this to 99% for 100 and 200 SNP windows due to the lower overall number of
344 windows and reduced genome-wide variance (Figure 5B-C).

345




346

347 ***Figure 5:*** *AF-vapeR results for WGS data from HP-LP guppies. Each AFV was calculated between*
348 *a distinct HP and its corresponding LP population within five rivers in a "replicate pairs''*
349 *sampling design as illustrated in* ***A****. Each row in panels* ***B*** *and* ***C*** *shows genome-wide empirical p-*
350 *values calculated based on 10,000 null permutations and window sizes of 25 (top), 50, 100 and*
351 *200 (bottom) SNPs for primary (****B****) and secondary (****C****) eigenvalues. Agreement among window*
352 *sizes is shown by points above peaks, with point shape and colour indicative of overlap among*
353 *window sizes.*

354

355 This analysis highlighted six windows across five regions of the genome with significant primary
356 eigenvalues on chromosomes 2, 5, 10, 13 and 19 (Table S4). However, all of these regions
357 represented a combination of parallel and antiparallel vectors (Table S4), suggesting that whilst
358 all population pairs exhibit allele frequency differentiation along a common axis, the direction
359 of change was not always consistent among replicates. The regions with the most parallel
360 changes (i.e. only a single anti-parallel replicate to four parallel replicates) were chr5:8127850-
361 8134640 (APHP-APLP antiparallel), and chr10:21021152-21034438 (TACHP-TACLP antiparallel).
362 The results from the primary eigenvalue thus agree with the previous analyses of these
363 populations, in that full-parallelism is limited within this system (Whiting et al., 2021).

364

365 On the second eigenvector we identified 13 significant windows over nine regions (Table S5,
366 Figure 5C). Large summed eigenvalues can be caused by either a large eigenvalue 1 plus a small
367 eigenvalue 2, or by a more balanced eigenvalue 1 and 2. The latter are better indicators of
368 multiple parallel axes of allele frequency change. Windows on chr1, chr15, chr20 and chr21
369 exhibited second eigenvalues accounting for at least 30% of total variance (Table S5). Of these,
370 the window on chr21 was the only window where one population pair exhibited no parallelism
371 with any of the other population pairs. Population loadings onto eigenvectors 1 and 2 (Table S4)
372 highlighted that two of these windows (chr1:9791276-9798829 and chr20:1248797-1296686)

373 show particularly interesting distinct phylogenetic axes of parallelism for the TAC-G-AP and O-
374 MAD lineages (Figure 5A).

375

376 The window chr20:1248797-1296686 (Figure S8) intersects with the major candidate previously
377 identified as evolving convergently in the three Caroni rivers (Whiting et al., 2021), and the
378 window chr15:5028361-5066375 (Figure S9) has been demonstrated to evolve in parallel
379 following experimental translocations of HP guppies to LP environments (van der Zee et al.,
380 2021). Neither of these windows exhibited evidence of parallelism from all five rivers, as shown
381 here, and in our previous analysis (Whiting et al., 2021). An in-depth assessment of candidate
382 windows is in the supplementary results.

383

384 In conclusion, these results confirm previous analyses that full parallelism is likely limited in this
385 system, whilst also identifying novel candidates and extending previous candidates to more
386 populations that are evolving under a multi-parallel scenario.

387

388 **Empirical Results: Galapagos Finch beak morphology**

389 The goal here was to explore an inter-species dataset to examine limitations associated with
390 divergence over time; the common ancestor of the Galapagos finch species group arrived at the

391 islands ~1.5 million years ago (Lamichhaney et al., 2015). The original study used WGS to
392 investigate beak morphology among Galapagos finches (Han et al., 2017) and included 12
393 between-species comparisons in a multi-comparison sampling design. The focal beak traits of
394 interest were size (small [S] vs large [L]) and shape (pointed [P] vs blunt [B]). These 12
395 comparisons included all combinations of beak phenotypes, e.g S-S, S-L, L-L, P-P, P-B, B-B (Table
396 2 in (Han et al., 2017)). We first used all 12 comparisons with AF-vapeR to fully replicate the
397 original study, but expectedly given the range of phenotype comparisons being made, we
398 observed no signals of parallelism. This emphasises the need to carefully consider the input
399 vectors prior to analysis. We next looked at targeted phenotype comparisons, i.e. S-L (six
400 comparisons, nine species) and P-B (five comparisons, seven species) (Figures 6A and 6D). We
401 applied four window sizes (50, 100, 200, 500 SNPs), that were selected due to greater SNP
402 density in this cross-species dataset. Windows reflected median physical sizes of 5.4, 11.1, 22.6
403 and 57.2 kb. Signatures of parallelism and eigenvalue variance were generally greater in this
404 dataset than in the guppy dataset, so we used stricter outlier criteria to examine only the
405 candidate regions with the strongest evidence (>99.9% of all permuted null values across all
406 window sizes).

407

408 A single clear peak of full-parallelism was observed for S-L comparisons, clustered around the
409 *HMGA2* locus and neighbouring genes on contig JH739900 (chromosome 1A) (Figure 6B-C)

410 (Table S6). There were no signals of multi-parallelism on the second eigenvector for S-L vectors.
411 This strong signal of full-parallelism at the *HMGA2* region implies a single axis describing allele
412 frequency change from small to large beak phenotypes; indicative of single S and L haplotype
413 groups. Indeed, clustered genotypes over the two windows overlapping *HMGA2* illustrate cross-
414 species clustering within S and L haplotypes (Figure S10). These results confirm that the
415 evolution of beak size within the galapagos finch radiation has involved the segregation of
416 commonly inherited S and L haplotypes at the *HMGA2* gene among species (Lamichhaney et al.,
417 2016), and no other regions of the genome show this process.

418

419 In contrast, for P-B comparisons we observed signals of full-parallelism on fourteen contigs
420 (Table S7-8). The *ALX1* gene has been identified as the primary candidate of interest in the
421 evolution of blunt beak shape (Lamichhaney et al., 2015). Our full-parallelism outliers included
422 a peak adjacent to, but not including, the *ALX1* locus on contig JH739921 (Figure 6E-F), but
423 within the 130kb downstream region identified by Lamichhaney et al. (2015). A lack of
424 parallelism over the *ALX1* gene specifically reflected strong haplotype structure among pointed
425 beak species; genotype clustering suggests four monophyletic P-haplotype groups over the
426 *ALX1* gene (Figure S11). This suggests divergent starting points prevent the identification of
427 parallel P-B axes, and highlight that signatures of full-parallelism erode over time. Inferences
428 may therefore be limited when applying our method to haplotypes with ancient origins.

429

430 In total, 16/31 windows (500 SNPs, ± 50kb) with strong signals of full-parallelism contained
431 previously identified and novel candidate genes for beak morphology in this system (discussed
432 fully in supplementary results and Table S9). Of particular note was a signal of full-parallelism
433 detected at *IGF1*, which has been linked to bill morphology in black-bellied seedcrackers
434 (*Pyrenestes ostrinus*) (vonHoldt et al., 2018). In the seedcracker study, no evidence was found
435 to link this gene to beak morphology in *Geospiza*, but our analysis suggests this gene may be
436 important for both taxa.

437

438

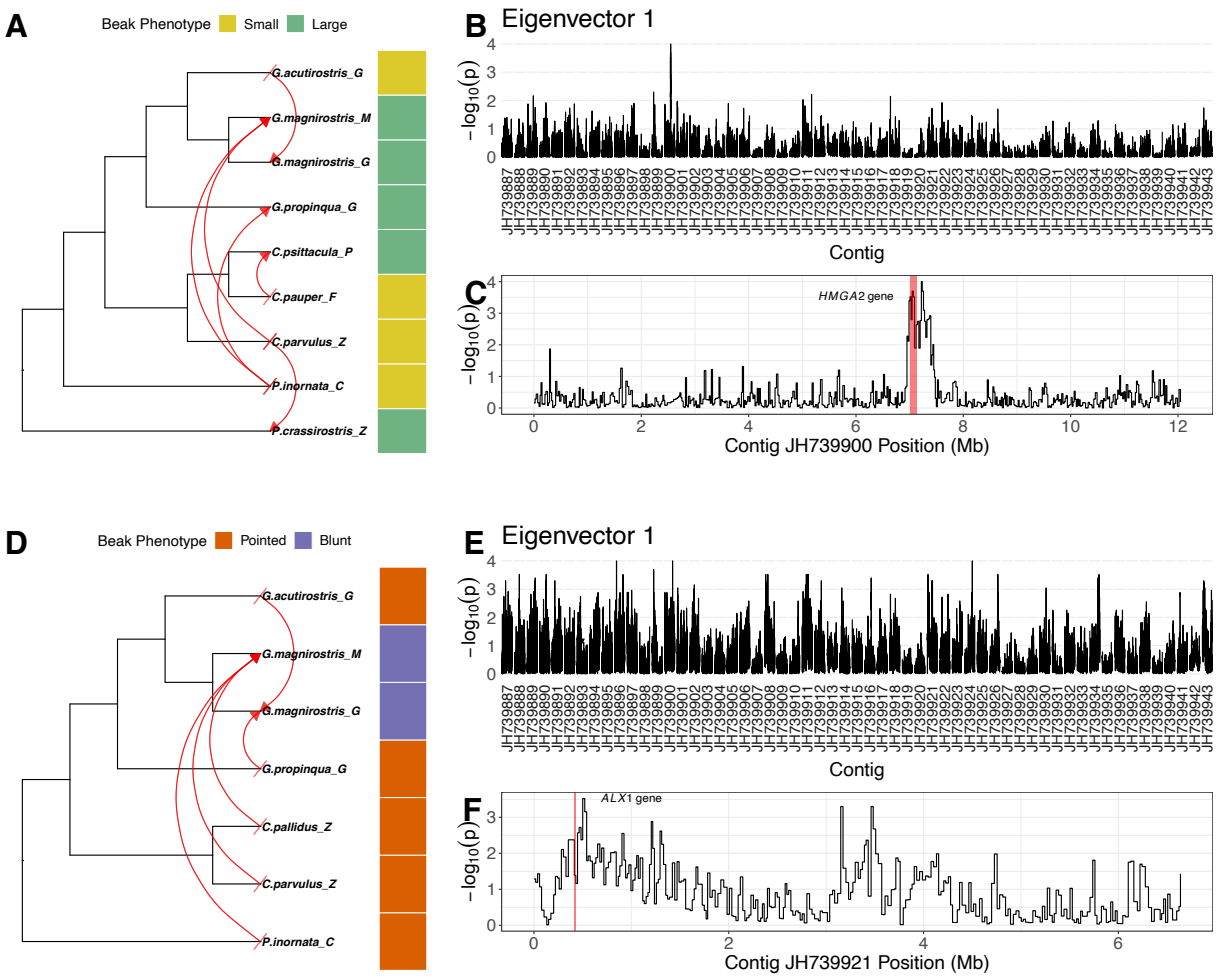



***Figure 6:*** *AF-vapeR results for WGS data from Galapagos finches. AFV were calculated between species with contrasting beak phenotypes (small to large or pointed to blunt), with species used multiple times in a multi-comparison design (**A**,**D**). Empirical p-values calculated based on 10,000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

446 *exhibiting full parallelism, including a window close to, but not including, the candidate ALX1*
447 *gene (marked in red) on contig JH739921 (**F**).*

448

449 **DISCUSSION**

450 **Summary of approach**

451 We present AF-vapeR, a novel genome scan method to explore SNP data for signals of parallel
452 evolution based on multivariate allele frequency vectors. Simulating different evolutionary
453 scenarios demonstrates low FPR and considerably reduced FNR relative to common $F_{ST}$ outlier
454 overlap approaches, whilst also distinguishing among different modes of evolution. Empirical
455 analyses demonstrate the ability to identify previously-known candidates of parallelism and to
456 uncover additional candidates. The reduction in Type-II error is derived from re-ordering the
457 workflow to scan for signatures of parallelism explicitly rather than of local adaptation.
458 Identified candidates for parallel evolution can later be assessed for the relative strength of
459 selection or differentiation if desired.

460

461 AF-vapeR is analogous to clustering approaches such as cluster separation score (CSS) (as
462 applied by (Jones et al., 2012; Morales et al., 2019; Roberts Kingman et al., 2021)), Cochran–
463 Mantel–Haenszel (CMH) tests (as applied by (Barghi et al., 2019; Orozco-terWengel et al.,

464 2012)), and the C2 statistic (Olazcuaga et al., 2020) (an extension of BayPass' covariate model to
465 binary groups). Importantly, each of these assumes a single possible axis for parallelism,
466 whereas AF-vapeR can detect multiple axes of parallelism reflecting different trajectories of
467 allele frequency change among replicates. In addition, CSS tests for low within-
468 ecotype/habitat/phenotype distances relative to between, which is true of some cases of full-
469 parallelism, but not all cases as defined here and detected by AF-vapeR. Specifically, AF-vapeR
470 detects similarity in direction of trajectory, as opposed to clustering of starting and endpoints,
471 so both clustered and unclustered cases of full-parallelism (Figure S1) can be detected. Further,
472 our permutation approach to deriving the null distribution incorporates sampling design, and
473 therefore is flexible to sampling designs involving biased group membership (such as a single
474 common ancestor vs many diverged replicates) that may affect other clustering approaches.

475

476 AF-vapeR, however, is not an explicit test for selection, but for parallel allele frequency
477 differences. This assumes highly parallel allele frequency change is an indicator of selection
478 through repeated, independent sorting of alleles (but see (Bailey et al., 2015)), but parallelism
479 could arise by other means. For instance, regions linked to blunt beaks in finches
480 parsimoniously evolved once in the common ancestor and were inherited by both
481 monophyletic blunt-beaked species. This sampling design is also likely why baseline P-B
482 eigenvalue 1 scores are high, as there is a genome-wide signal of shared inheritance within the

483 blunt-beaked lineage. To compensate for this, we used stricter cutoffs in the finches compared
484 with stickleback and guppies. Generally, we advise to keep the underlying phylogeny in mind
485 when interpreting results. Hybridisation from an unstudied lineage may also produce parallel
486 allele frequency change, given heterogeneity in the success of introgression across the genome,
487 as suggested for flycatchers (Ellegren et al., 2012) and anchovies (Le Moan et al., 2016). This
488 should not, however, be confused with hybridisation providing adaptive alleles to multiple
489 lineages that are subsequently selected independently. This latter process, observed for
490 example in *Heliconius* (Edelman et al., 2019), is a genuine representation of parallel evolution
491 given increases in the frequency of introgressed variants occur independently (Lee & Coop,
492 2019).

493

494 Recombination rate landscapes commonly obscure genome scans (Lotterhos, 2019), either by
495 inflating variance in divergence, producing longer tails in low-recombination regions that may
496 be falsely identified as signals of selection (Booker et al., 2020), or through inverse associations
497 with background selection (Burri, 2017). Increased variance with reduced recombination is also
498 observed with AF-vapeR (Figure S7). Some resilience to recombination is expected by using SNP
499 count, rather than physical size, windows, given SNP density is linked to recombination
500 landscape (Burri et al., 2015; Stankowski et al., 2019). We therefore expect larger, sparser
501 physical windows in low-recombination regions, and smaller, denser windows in high-

502 recombination regions. This balances the risk of oversampling and having many non-
503 independent windows in tight-linkage in low-recombination regions, and vice-versa. Secondly,
504 because our analysis involves directionality, we can rule out windows with high eigenvalues
505 with parallel and antiparallel vectors. If low recombination inflates variance randomly, it is not
506 expected to produce allele frequency change in the same direction, and additional independent
507 lineages will increase analysis robustness.

508

509 Importantly, regions showing strong parallelism may not be under the strongest selection, but
510 may reflect processes governing repeatability specifically, including: heterogeneity in the
511 mutation landscape (Bailey et al., 2018); differential rates of ancient haplotype maintenance
512 (Nelson & Cresko, 2018; Wang et al., 2019); or gene/QTL density landscape (Roberts Kingman et
513 al., 2021). Identifying these regions will help elucidate the influences of these processes on
514 repeatability.

515

516 **Recommendations**

517 We recommend a minimum of four replicate vectors to explore full- and multi-parallelism.
518 Because maximum eigenvalues are constrained to *m* (for *m* AFV), increasing *m* increases the
519 distinctiveness of eigenvalue peaks and reduces the likelihood of false-positives. Beyond studies
520 of repeatability, this approach may be well-suited to evolve-and-resequence experiments (e.g.

521 (Barghi et al., 2019; Rudman et al., 2021; Tusso et al., 2021)), where replication is common and
522 allele frequency change may be moderate and recorded after only a few generations. Because
523 patterns of parallelism break down over time, seen here with *ALX1* haplotypes in finches, our
524 approach is best-suited to intra-specific comparisons, or those among relatively young clades.
525 Whole-genome data is desirable, but not a prerequisite (e.g. RAD-seq shown here), and given
526 allele frequencies are the focus pool-sequencing data could also be used. Whilst the
527 permutation approach accounts for sampling design, some sampling designs may bias
528 eigenvalues more than others, particularly repeatedly sampling the same population when
529 calculating AFV. In such cases, greater stringency can be applied when calling outliers, as
530 demonstrated here when analysing the guppy dataset compared with the finch dataset.
531 Datasets used here were filtered with a minor allele frequency (MAF) filter of 5%, but 1% may
532 be sufficient. MAF-filtering is important so that SNP windows are not overpopulated by many
533 SNPs carrying minimal information.

534

535 Visualising trees and genotypes at candidate regions can be informative in understanding how
536 replicates load onto eigenvectors and how populations cluster within parallel or multi-parallel
537 clusters (e.g. Figure S8-11). Further, per-SNP scores available in each window's projected **A**
538 matrix can be used to identify SNPs whose allele frequencies are driving variance (e.g. Figure
539 S8). Examining individual SNP scores is akin to examining the loading of variables during

principal component analysis (PCA) to assess which variables are best reflected in each principal component. We reiterate that AF-vapeR cannot detect cases where all replicates divergently modify the same gene or locus (molecular convergence; Figure 2C), as this produces geometric non-parallelism. Such regions may still be of interest, however, and combining AF-vapeR with selection scans may highlight such regions.

The study of repeatable evolution in the genome is increasingly popular in natural systems and in experimental evolution. Progression of this field is essential for better understanding the determinism, predictability, and contingencies linked with adaptation, with implications for issues including rapid adaptation to changing environments (Reid et al., 2016) and the evolution of antibiotic resistance (A. C. Palmer & Kishony, 2013). There are also other biological questions where the goal is similarly to identify shared haplotypes among groups, such as sex-linked loci in the heterogametic sex (D. H. Palmer et al., 2019). Despite this, there remain limited approaches tailored towards detecting signatures of repeatability from genomic data. Our method requires only a standard VCF file, a population map, and a description of vectors to calculate, and should be accessible to researchers familiar with common multivariate analyses such as PCA or principal coordinate analysis (PCoA). We therefore expect AF-vapeR to become an important tool for detecting convergence alongside currently available approaches.

## ACKNOWLEDGEMENTS

The authors would like to thank Matthew Webster (Uppsala University), and Andrew MacColl (University of Nottingham) and Isabel Santos Magalhaes (University of Roehampton) for insightful discussions regarding the finch and stickleback datasets respectively. The authors also thank members of the University of Exeter's BS-popgen group and Sam Yeaman (University of Calgary) and members of the Yeaman lab for providing feedback on the manuscript and project. HPC infrastructure support was provided by The University of Exeter's High-Performance Computing (HPC) facility (ISCA).

## DATA AVAILABILITY

AF-vapeR has been made available for download at github (https://github.com/JimWhiting91/afvaper). The code available for simulations and analyses of simulations and empirical datasets is available at (https://github.com/JimWhiting91/afvaper_analyses_ms). Sequencing reads associated with the stickleback datasets are available at the European Nucleotide Archive (ENA) (accession: PRJEB20851). Sequencing reads associated with the guppy dataset are available at the ENA (accessions: PRJEB43917 and PRJEB10680) and the VCF used is available from FigShare (doi:10.6084/m9.figshare.14315771). The VCF for the galapagos finch data was acquired from Dryad (doi: 10.5061/dryad.7155d).

578

579 **AUTHOR CONTRIBUTIONS**

580 JRW and MJvZ conceived the project. JRW developed the software, and performed simulations,
581 analyses and figure production. JRP assisted software development through testing and
582 simulation development. JRW, JRP and BAF were involved in writing the manuscript. BAF
583 supervised the project.

584

585 **CONFLICTS OF INTEREST**

586 The authors declare no conflicts of interest.

587

588 **REFERENCES**

589 Adams, D. C., & Collyer, M. L. (2009). A general framework for the analysis of phenotypic
590 trajectories in evolutionary studies. *Evolution; International Journal of Organic Evolution*,
591 *63*(5), 1143–1154.
592 Bailey, S. F., Guo, Q., & Bataillon, T. (2018). Identifying Drivers of Parallel Evolution: A
593 Regression Model Approach. *Genome Biology and Evolution*, *10*(10), 2801–2812.
594 Bailey, S. F., Rodrigue, N., & Kassen, R. (2015). The effect of selection environment on the
595 probability of parallel evolution. *Molecular Biology and Evolution*, *32*(6), 1436–1448.

596 Barghi, N., Tobler, R., Nolte, V., Jakšić, A. M., Mallard, F., Otte, K. A., Dolezal, M., Taus, T., Kofler,
597 R., & Schlötterer, C. (2019). Genetic redundancy fuels polygenic adaptation in Drosophila.
598 *PLoS Biology*, *17*(2), e3000128.
599 Booker, T. R., Yeaman, S., & Whitlock, M. C. (2020). Variation in recombination rate affects
600 detection of outliers in genome scans under neutrality. *Molecular Ecology*, *29*(22), 4274–
601 4279.
602 Booker, T. R., Yeaman, S., & Whitlock, M. C. (2021). Global adaptation complicates the
603 interpretation of genome scans for local adaptation. *Evolution Letters*, *5*(1), 4–15.
604 Burri, R. (2017). Interpreting differentiation landscapes in the light of long-term linked
605 selection. *Evolution Letters*, *1*(3), 118–131.
606 Burri, R., Nater, A., Kawakami, T., Mugal, C. F., Olason, P. I., Smeds, L., Suh, A., Dutoit, L., Bureš,
607 S., Garamszegi, L. Z., Hogner, S., Moreno, J., Qvarnström, A., Ružić, M., Sæther, S.-A.,
608 Sætre, G.-P., Török, J., & Ellegren, H. (2015). Linked selection and recombination rate
609 variation drive the evolution of the genomic landscape of differentiation across the
610 speciation continuum of Ficedula flycatchers. *Genome Research*, *25*(11), 1656–1665.
611 Cruickshank, T. E., & Hahn, M. W. (2014). Reanalysis suggests that genomic islands of speciation
612 are due to reduced diversity, not reduced gene flow. *Molecular Ecology*, *23*(13), 3133–
613 3157.
614 De Lisle, S. P., & Bolnick, D. I. (2020). A multivariate view of parallel evolution. *Evolution;*
615 *International Journal of Organic Evolution*, *74*(7), 1466–1481.

616 Edelman, N. B., Frandsen, P. B., Miyagi, M., Clavijo, B., Davey, J., Dikow, R. B., García-Accinelli,
617 G., Van Belleghem, S. M., Patterson, N., Neafsey, D. E., Challis, R., Kumar, S., Moreira, G. R.
618 P., Salazar, C., Chouteau, M., Counterman, B. A., Papa, R., Blaxter, M., Reed, R. D., …
619 Mallet, J. (2019). Genomic architecture and introgression shape a butterfly radiation.
620 *Science*, *366*(6465), 594–599.
621 Ellegren, H., Smeds, L., Burri, R., Olason, P. I., Backström, N., Kawakami, T., Künstner, A.,
622 Mäkinen, H., Nadachowska-Brzyska, K., Qvarnström, A., Uebbing, S., & Wolf, J. B. W.
623 (2012). The genomic landscape of species divergence in Ficedula flycatchers. *Nature*,
624 *491*(7426), 756–760.
625 Fang, B., Kemppainen, P., Momigliano, P., & Merilä, J. (2021). Population structure limits
626 parallel evolution in sticklebacks. *Molecular Biology and Evolution*.
627 https://doi.org/10.1093/molbev/msab144
628 Fraser, B. A., & Whiting, J. R. (2019). What can be learned by scanning the genome for
629 molecular convergence in wild populations? *Annals of the New York Academy of Sciences*.
630 https://doi.org/10.1111/nyas.14177
631 Haller, B. C., & Messer, P. W. (2019). SLiM 3: Forward Genetic Simulations Beyond the Wright–
632 Fisher Model. *Molecular Biology and Evolution*, *36*(3), 632–637.
633 Han, F., Lamichhaney, S., Grant, B. R., Grant, P. R., Andersson, L., & Webster, M. T. (2017). Gene
634 flow, ancient polymorphism, and ecological adaptation shape the genomic landscape of
635 divergence among Darwin's finches. *Genome Research*, *27*(6), 1004–1015.

636 Hermisson, J., & Pennings, P. S. (2017). Soft sweeps and beyond: understanding the patterns
637 and probabilities of selection footprints under rapid adaptation. *Methods in Ecology and*
638 *Evolution / British Ecological Society*, *8*(6), 700–716.
639 Hohenlohe, P. A., Bassham, S., Etter, P. D., Stiffler, N., Johnson, E. A., & Cresko, W. A. (2010).
640 Population genomics of parallel adaptation in threespine stickleback using sequenced RAD
641 tags. *PLoS Genetics*, *6*(2). https://doi.org/10.1371/journal.pgen.1000862
642 Jones, F. C., Grabherr, M. G., Chan, Y. F., Russell, P., Mauceli, E., Johnson, J., Swofford, R., Pirun,
643 M., Zody, M. C., White, S., Birney, E., Searle, S., Schmutz, J., Grimwood, J., Dickson, M. C.,
644 Myers, R. M., Miller, C. T., Summers, B. R., Knecht, A. K., … Kingsley, D. M. (2012). The
645 genomic basis of adaptive evolution in threespine sticklebacks. *Nature*, *484*(7392), 55–61.
646 Lamichhaney, S., Berglund, J., Almén, M. S., Maqbool, K., Grabherr, M., Martinez-Barrio, A.,
647 Promerová, M., Rubin, C.-J., Wang, C., Zamani, N., Grant, B. R., Grant, P. R., Webster, M. T.,
648 & Andersson, L. (2015). Evolution of Darwin's finches and their beaks revealed by genome
649 sequencing. *Nature*, *518*(7539), 371–375.
650 Lamichhaney, S., Han, F., Berglund, J., Wang, C., Almén, M. S., Webster, M. T., Grant, B. R.,
651 Grant, P. R., & Andersson, L. (2016). A beak size locus in Darwin's finches facilitated
652 character displacement during a drought. *Science*, *352*(6284), 470–474.
653 Lee, K. M., & Coop, G. (2019). Population genomics perspectives on convergent adaptation.
654 *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*,
655 *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.

Lotterhos, K. E., & Whitlock, M. C. (2015). The relative power of genome scans to detect local adaptation depends on sampling design and statistical method. *Molecular Ecology*, *24*(5), 1031–1046.

Magalhaes, I. S., D'Agostino, D., Hohenlohe, P. A., & MacColl, A. D. C. (2016). The ecology of an adaptive radiation of three-spined stickleback from North Uist, Scotland. *Molecular Ecology*, *25*(17), 4319–4336.

Magalhaes, I. S., Whiting, J. R., D'Agostino, D., Hohenlohe, P. A., Mahmud, M., Bell, M. A., Skúlason, S., & MacColl, A. D. C. (2020). Intercontinental genomic parallelism in multiple three-spined stickleback adaptive radiations. *Nature Ecology & Evolution*. https://doi.org/10.1038/s41559-020-01341-8

Morales, H. E., Faria, R., Johannesson, K., Larsson, T., Panova, M., Westram, A. M., & Butlin, R. K. (2019). Genomic architecture of parallel ecological divergence: Beyond a single

environmental contrast. *Science Advances*, *5*(12), eaav9963.

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

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

Reid, N. M., Proestou, D. A., Clark, B. W., Warren, W. C., Colbourne, J. K., Shaw, J. R., Karchner,

696 S. I., Hahn, M. E., Nacci, D., Oleksiak, M. F., Crawford, D. L., & Whitehead, A. (2016). The
697 genomic landscape of rapid repeated evolutionary adaptation to toxic pollution in wild
698 fish. *Science*, *354*(6317), 1305–1308.
699 Roberts Kingman, G. A., Vyas, D. N., Jones, F. C., Brady, S. D., Chen, H. I., Reid, K., Milhaven, M.,
700 Bertino, T. S., Aguirre, W. E., Heins, D. C., von Hippel, F. A., Park, P. J., Kirch, M., Absher, D.
701 M., Myers, R. M., Di Palma, F., Bell, M. A., Kingsley, D. M., & Veeramah, K. R. (2021).
702 Predicting future from past: The genomic basis of recurrent and rapid stickleback
703 evolution. *Science Advances*, *7*(25). https://doi.org/10.1126/sciadv.abg5285
704 Rudman, S. M., Greenblum, S. I., Rajpurohit, S., Betancourt, N. J., Hanna, J., Tilk, S., Yokoyama,
705 T., Petrov, D. A., & Schmidt, P. (2021). Direct observation of adaptive tracking on ecological
706 timescales in Drosophila. In *bioRxiv* (p. 2021.04.27.441526).
707 https://doi.org/10.1101/2021.04.27.441526
708 Stankowski, S., Chase, M. A., Fuiten, A. M., Rodrigues, M. F., Ralph, P. L., & Streisfeld, M. A.
709 (2019). Widespread selection and gene flow shape the genomic landscape during a
710 radiation of monkeyflowers. *PLoS Biology*, *17*(7), e3000391.
711 Stuart, Y. E. (2019). Divergent Uses of “Parallel Evolution” during the History of The American
712 Naturalist. *The American Naturalist*, *193*(1), 11–19.
713 Teshima, K. M., Coop, G., & Przeworski, M. (2006). How reliable are empirical genomic scans for
714 selective sweeps? *Genome Research*, *16*(6), 702–712.
715 Tusso, S., Nieuwenhuis, B. P. S., Weissensteiner, B., Immler, S., & Wolf, J. B. W. (2021).

716 Experimental evolution of adaptive divergence under varying degrees of gene flow. *Nature*
717 *Ecology & Evolution*, *5*(3), 338–349.
718 van der Zee, M. J., Whiting, J. R., Paris, J. R., Bassar, R. D., Travis, J., Weigel, D., Reznick, D. N., &
719 Fraser, B. A. (2021). Rapid genomic convergent evolution in experimental populations of
720 Trinidadian guppies (Poecilia reticulata). In *bioRxiv* (p. 2021.02.10.430609).
721 https://doi.org/10.1101/2021.02.10.430609
722 vonHoldt, B. M., Kartzinel, R. Y., Huber, C. D., Le Underwood, V., Zhen, Y., Ruegg, K., Lohmueller,
723 K. E., & Smith, T. B. (2018). Growth factor gene IGF1 is associated with bill size in the black-
724 bellied seedcracker Pyrenestes ostrinus. *Nature Communications*, *9*(1), 4855.
725 Wang, B., Mojica, J. P., Perera, N., Lee, C.-R., Lovell, J. T., Sharma, A., Adam, C., Lipzen, A., Barry,
726 K., Rokhsar, D. S., Schmutz, J., & Mitchell-Olds, T. (2019). Ancient polymorphisms
727 contribute to genome-wide variation by long-term balancing selection and divergent
728 sorting in Boechera stricta. *Genome Biology*, *20*(1), 126.
729 Whiting, J. R., Paris, J. R., van der Zee, M. J., Parsons, P. J., Weigel, D., & Fraser, B. A. (2021).
730 Drainage-structuring of ancestral variation and a common functional pathway shape
731 limited genomic convergence in natural high- and low-predation guppies. *PLoS Genetics*,
732 *17*(5), e1009566.

733