MOLECULAR ECOLOGY

Detecting natural selection across environmental gradients: A performance evaluation of statistical genetic methods using simulations

1   **Title**: Detecting natural selection across environmental gradients: A performance evaluation of

2   statistical genetic methods using simulations

3   **Authors**: Gretchen H. Roffler[1, 2], Erin L. Landguth[3], Michael K. Schwartz[4], Gordon Luikart[5, 6]

4

5   [1]Alaska Department of Fish and Game, Division of Wildlife Conservation, 802 3rd Street,

6   Douglas, AK 99824, USA

7   [2]United States Geological Survey, Alaska Science Center, 4210 University Drive, Anchorage,

8   AK 99508, USA

9   [3]Computational Ecology Laboratory, Division of Biological Sciences, University of Montana,

10   Missoula, MT, 59812, USA

11   [4]US Forest Service Rocky Mountain Research Station, National Genomics Center for Wildlife

12   and Fish Conservation, 800 E. Beckwith, Missoula, MT 59801, USA

13   [5]Fish and Wildlife Genomics Group, Division of Biological Sciences, University of Montana,

14   Missoula, MT 59812, USA

15   [6]Flathead Lake Biological Station, Fish and Wildlife Genomics Group, Division of Biological

16   Sciences, University of Montana, Polson, MT 59860, USA

17   **Corresponding author:** Gretchen H. Roffler, Alaska Department of Fish and Game, Division of

18   Wildlife Conservation, 802 3rd Street, Douglas, AK 99824, fax: (907) 465-4272, email:

19   gretchen.roffler@alaska.gov

20   **Key words:** adaptive differentiation, directional selection, environmental gradient, landscape

21   genomics, local adaptation, outlier tests, simulations

22   **Running title:** Methods to detect natural selection

Abstract

Identifying loci under selection is important for understanding how populations adapt to different environments. Most research on natural selection, gene flow, and effective population size interactions has focused on idealized theoretical populations without accounting for spatial complexities including selective gradients or discrete environments. Using spatially-explicit, individual-based simulations, we incorporated realistic demographic parameters and landscapes to address the following questions: (1) How strong must selection be for distributions of adaptive versus neutral alleles to diverge given different strengths of genetic drift (population size), gene flow (dispersal distance), and spatial patterns of selection, and (2) What is the relative power and error rate of population- and individual-based statistical tests for detecting loci under selection? The development of adaptive differentiation was most influenced by low dispersal ($N_m = 1$–2), and low drift ($N_e \leq 475$). Type I error rates (false positives) were highest with individual-based (e.g., Samβada and GLMs) and lowest with population-based methods (LOSITAN and BayeScan). Alternatively, population- and individual-based methods produced low type II error rates (false negatives) given strong regional differences in selection, but only the individual-based methods consistently performed well when selection was weak. Further, in a selection gradient, only individual-based methods consistently detected the locus under selection. Our results suggest that individual-based methods will most often correctly identify genotype-environment interactions, especially with large populations and low dispersal (type II error rate $0.0 - 0.3$, in comparison to population-based methods $0.7 - 1.0$). We suggest using

44   <mark>both population- and individual-based approaches for conservative detection of false positives</mark>

45   <mark>and negatives</mark>.

46

47   **Introduction**

48   Understanding how populations adapt to spatial and temporal heterogeneity of their

49   environments is a key topic in ecology, evolution, and conservation biology (Schoville *et al*.

50   2010; Hoffmann & Sgrò 201; Savolainen *et al*. 2013). Integral to this line of research is

51   identification of genes underlying adaptation and signatures of natural selection (Storz 2005;

52   Neale & Kremer 2011; Allendorf *et al*. 2010). A proven and increasingly common way to test

53   for signatures of selection in natural populations involves detection of the geographical

54   distribution of candidate adaptive alleles using <mark>statistical outlier tests and related methods</mark>

55   (Luikart *et al*. 2003; Narum *et al*. 2011). Detection of adaptive alleles and correlating their

56   spatial distributions with key landscape and environmental variables can help in understanding

57   and verifying patterns of selection and local adaptation (Foll & Gaggiotti 2006; Joost *et al*.

58   2007). Moreover, quantifying the geographic distribution of adaptive alleles can help identify

59   adaptively differentiated populations and predict population persistence and shifts in species

60   distributions due to climate fluctuations, or anthropogenic changes of landcover (Hancock *et al*.

61   2008; Manel *et al*. 2010; Kovach et al. 2015).

62   Although much theory exists, and several model systems have been examined

63   (Hohenlohe *et al*. 2010; Fournier-Level *et al*. 2011; Hancock *et al*. 2011), little is currently

64   known concerning the pattern or rate of the spread of adaptive alleles across populations in

65 natural landscapes (Holderegger & Wagner 2008; Schoville *et al*. 2012). Advancement of

66 research in detecting distinct signatures of adaptive allele distributions (e.g., outliers) across

67 realistic landscapes is currently limited by poor understanding of the accuracy and precision of

68 statistical tests and empirical approaches (Yang *et al*. 2012), and failure of methods to account

69 for complications of spatial autocorrelation (Meirmans 2012; Neel *et al*. 2013). Genetic adaptive

70 differentiation among populations often evolves if distinct environmental gradients are present or

71 strong variation exists between regions, and if the selection is not countered by gene flow or

72 genetic drift (Hohenlohe *et al*. 2010; Narum *et al*. 2010; Zardi *et al*. 2011). Given strong enough

73 selection, frequencies of adaptive alleles (influenced by differences in individual fitness) will

74 diverge between populations (Lewotin & Krakauer 1973; Beaumont 2005; Savolainen *et al*.

75 2013), and are predicted to have a different spatial distribution than neutral alleles (influenced by

76 genetic drift, migration and mutation; Schwartz *et al*. 2010).

77   There is debate over whether neutral and adaptive markers will be spatially correlated

78 (McKay & Latta 2002; Coop *et al*. 2009), or have different spatial distributions, as adaptive

79 alleles can respond to environmental differences among populations (Hohenlohe *et al*. 2010;

80 Limborg *et al*. 2012). Population genetics theory predicts large populations with limited gene

81 flow will develop significant adaptive genetic differences among them (Crow & Kimura 1970;

82 Waples 1998; Lenormand 2002; Hedrick 2005), due to the efficiency of natural selection

83 differentiating adaptive allele frequencies, but not neutral alleles (Kohn *et al*. 2003; Franks *et al*.

84 2007, Lowe & Allendorf 2010). Conversely, gene flow may play a role in spreading

85 advantageous mutations, whereas balancing selection could homogenize allele frequencies at

86 certain adaptive loci (Rieseberg & Burke 2001; Charbonnel & Pemberton 2005; Holderegger &

87   Wagner 2008). However, population genetics theory has not been extensively tested or extended

88   to complex landscapes or tested for using individual-based approaches. Advances have been

89   made in identifying adaptive differentiation at broad geographic scales including globally

90   (Prugnolle *et al*. 2005; Hanncock *et al*. 2008; Coop *et al*. 2009), continent or ocean-wide (Joost

91   *et al*. 2007, Limborg *et al*. 2012), and over multiple ecotypes (Matala *et al*. 2014;Cheviron *et al*.

92   2014), but less work has been accomplished at smaller, landscape level scales.

93   To understand the roles and interaction between gene flow, population size, and natural

94   selection, it is necessary to answer these questions**:** how does gene flow distribute adaptive genes

95   across a heterogeneous landscape given different selection regimes and population sizes, and

96   how does adaptive gene flow differ from neutral gene flow patterns? Comparing loci that are

97   outliers to what would be expected under neutral distributions is necessary to identify locally

98   adapted populations, and prevent biases associated with using loci under selection when testing

99   for neutral genetic structure (Luikart *et al*. 2003). Once they are identified, adaptive loci can be

100   used to facilitate the study of interactions between environmental selective pressures and genes

101   by examining patterns of adaptive divergence (Hand *et al*. 2015). In addition, adaptive loci can

102   be used in combination with neutral loci to elucidate the effects of population genetic parameters

103   such as gene flow and effective population size (N_e) on the spatial scale of adaptive

104   differentiation.

105   Until recently, the majority of research on the role and interaction between natural

106   selection, gene flow, and effective population size has focused on idealized theoretical

107   populations without accounting for spatial complexities such as selective gradients or spatially

108   distinct environments (Epperson *et al*. 2010). Using discrete, population-based models under

109   ideal Wright-Fisher assumptions limits their broad application, and makes interpretation of

110   results in complex environments equivocal. Incorporating biologically realistic demographic and

111   life history parameters into spatially explicit, individual-based models will increase our

112   understanding of how population genetic structure is influenced by environmental variability

113   (Landguth *et al*. 2010). This is now possible with the development of landscape genetics

114   approaches and modeling tools that specifically seek to reveal how spatial complexity of the

115   environment affects population genetic processes (Landguth & Cushman 2010).

116   Simulation studies are useful for understanding how complex systems function (Hoban *et*

117   *al*. 2011), how population genetic theory may be informed by approaches that incorporate

118   realistic landscape and environmental information (Hedrick 2005; Epperson *et al*. 2010;

119   Landguth *et al*. 2010), how sampling impacts results (Schwartz & McKelvey 2009), and

120   furthermore may be combined with empirical data to improve inference (Cushman & Landguth

121   2010). Recent simulation studies incorporating spatially-explicit data evaluated the performance

122   of population-based (Narum & Hess 2011; Villas *et al*. 2012; Lotterhos & Whitlock 2015) and

123   correlation-based (Jones *et al*. 2013; De Mitta *et al*. 2013; Forester *et al*. 2015; Lotterhos &

124   Whitlock 2015) methods to detect loci under selection. However, additional quantification of the

125   strengths and weaknesses of these methods is necessary, particularly in in a spatially-explicit

126   modeling scenario. Whereas some studies have examined how environmental gradients affect

127   selection (Jones *et al*. 2013; Forester *et al*. 2015), very few have used scenarios with >1

128   population (De Mitta *et al*. 2013). Finally, in addition to examining the effect of gene flow by

129   varying dispersal distance, it is also necessary to vary population sizes to assess the effects of

130   genetic drift on the development of adaptive differentiation.

131   We used a spatially-explicit, individual-based simulation approach and incorporated

132   biologically realistic demographic parameters and heterogeneous landscapes to evaluate the

The main focus of the study

133   power and sensitivity of multiple widely-used methods and software programs to identify

134   candidate loci under selection given varying strengths of natural selection, population sizes, and

135   dispersal distances. Using this approach, we addressed the following questions: (1) How strong

136   must selection be for spatial distributions of adaptive alleles to differ significantly from neutral

137   patterns given different strengths of genetic drift (population size), gene flow (approximated

138   through dispersal distance), and spatial pattern of selection, and (2) What is the relative power

139   and error rate (type I and II) of population-based and individual-based statistical tests for

140   detecting loci under selection? We predicted that adaptive and neutral alleles would become

141   more differentiated with increasing population size and strength of selection, and decreasing

142   dispersal distances, and these differences would become more distinct over time (generations).

143   We further predicted that signatures of selection would be stronger in large populations with low

144   dispersal, and therefore have lower type I and II error rates for the outlier loci detection methods.

145

146   **Methods**

147   *Simulation of spatially-explicit selection*

148   We used the spatially-explicit, individual-based, landscape genetics program CDPOP v1.0

149   (Landguth *et al*. 2012) to simulate the effects of varying selection strength, population size, and

150   dispersal distances across a hypothetical landscape of size 200 x 300 km. CDPOP is a

151   simulation program that has been demonstrated to accurately reproduce natural selection

152   (Landguth *et al*. 2012), movement, and mating processes that affect spatial population genetic

153   substructure (Landguth *et al*. 2010). Therefore, ecological processes that affect genetic patterns

154   can be tested by varying model parameters and evaluating the results. We conducted a series of

155   simulation modeling experiments with the following factors: (1) Population size (small = 1,000,

156   large = 10,000), (2) Dispersal distance (5 % and 50 % of maximum dispersal distance in study

157   area; 9 and 93 km respectively), (3) Strengths of selection (strong [selection coefficient (*s*) =

158   0.5] and weak [*s* = 0.01], absent [*s* = 0.00]; 50 %, 10 %, and 0 % reduced survival, respectively),

159   and (4) Spatial patterns of selection: selection present in the northern half and absent in the

160   southern half, a continuous selection gradient, and selection absent (Fig. 1). We subsequently

161   ran 10 Monte Carlo replicates for each simulation scenario across 500 non-overlapping

162   generations which resulted in 2,640 simulated datasets (Table 1), a computationally challenging

163   and time-consuming endeavor.

164   The landscape was initially populated with random locations representing individuals in 2

165   subpopulations with an isolation-by-distance (IBD) genetic pattern (Fig. 1). We limited

166   migration between the 2 subpopulations by maintaining a barrier to dispersal (Landguth *et al*.

167   2010, Landguth & Balkenhol 2012) for 50 generations to achieve desired levels of population

168   differentiation (F_{ST} = 0.02). We then used the genotypes of these 2 populations as the starting

169   point for subsequent simulations. We simulated genetic exchange as functions of individual-

170   based movement, mating, and dispersal acting on offspring at a given location. The

171   subpopulations had an equal ratio of males to females, a polygynous mating system, and an

172   inverse square probability function for dispersal. Individual genotypes were composed of 100

173   bi-allelic loci (99 neutral and 1 under selection), with a *k*-allele mutation rate of 2.5×10^{-8} per

174   nucleotide per generation, based on previous estimates of SNP (single nucleotide polymorphism)

175   mutation rate (Nachman & Crowell 2000; Kondrashov 2002, and references therein) included in

176   the genetic model.

177   We then generated differences in selection by portraying individual fitness as a function

178   of the number of surviving offspring, which was determined by the interaction of the offspring's

179   genotype and spatial patterns of selection. We created 3 simulated landscape surfaces based on

180   different spatial configuration of directional genetic selection: (a) two adjacent discrete regions

181   with selection present in one and absent in the other to represent spatially distinct environments,

182   (b) a gradient with the strength of selection continuously distributed from north to south to

183   represent a clinal environment, and (c) a uniform surface with no selection (Fig. 1). For the first

184   2 surfaces, we simulated both strong and weak selection. In surface "a", we represented strong

185   selection by assigning a relative fitness value of 1.0 to the genotype *AA* in the northern region

186   (selection "absent"), and a value of 0.5 in the southern (selection "present") region. Weak

187   selection was simulated by assigning a relative fitness value of 1.0 for the *AA* genotype in the

188   north, and 0.9 in the south. The genotypes *Aa* and aa had a uniform fitness value of 1.0. In the

189   selection surface "b" (selection gradient), the relative fitness value of the *AA* genotype was 1.0

190   (for both strong and weak selection) at the northern end, and either 0.5 (strong) or 0.9 (weak) at

191   the southern end, with intermediate values in between. The *Aa* and *aa* genotypes had a uniform

192   fitness value of 1.0 for both strong selection weak selection throughout the study area (Fig. 1).

193   Pixel size on the selection surface was 100 m^2.

194   We estimated N_e using the bias-corrected version of the linkage disequilibrium method

195   with 95 % confidence intervals generated with a jack-knife approach (Waples & Do 2008; Do *et*

196   *al.* 2014). We determined pairwise population F_{ST} values (Weir and Cockerham 1984) using

197   GENEPOP, and calculated number of migrants per generation using the flowing equation: $N_m =$

198    $(1- F_{ST}) / 4 F_{ST}$.

199

200   *Genetic differentiation of adaptive vs. neutral markers*

201   To measure the degree of divergence of adaptive and neutral alleles, we estimated genetic

202   differentiation at each simulated generation using both a population-based and an individual-

203   based approach. For the population-based approach, we used the estimator D_{est} (Jost 2008) to

204   quantify genetic differentiation over time using only the 99 neutral loci, the 1 loci under

205   selection, and all 100 loci. Although several genetic differentiation metrics are available, we

206   selected D_{est} because is not constrained by measures of within-population diversity as much as

207   F_{ST}, it scales linearly, and the low mutation rates common in SNPs would have a negligible

208   effect (Meirmans & Hedrick 2011).

209   For the landscape genetics approach, we used Mantel tests (Mantel 1967) to compare

210   pairwise genetic distance of all individuals to geographic distance using only neutral loci, only

211   adaptive loci, and all loci combined. Genetic distance among individuals was calculated using

212   the proportion of shared alleles (Dps; Bowcock *et al.* 1994). For each scenario, selection

213   strength, and selection surface we evaluated the10 independent Monte Carlo replicates to assess

214   variability in results. To calculate changes in genetic differentiation across time, we plotted the

215   average D_{est} and average Mantel correlation coefficient r across 500 generations in 4 scenarios

216   portraying combinations of genetic drift (population size) and dispersal distance (Table 2). We

217   compared resulting test statistics and assessed differences using the neutral r values as the null

218   distribution.

219

220   *Identifying loci under selection*

221   We computed the type I and II error rates of population- and individual-based methods to

222   correctly detect loci under selection in the simulated datasets. We used the simulated allele

223   frequencies at generations 50 and 500 under 4 different scenarios (Table 2), with both strong and

224   weak selection and using 3 selection surfaces (present/absent, latitudinal gradient, selection

225   absent), averaged across the 10 Monte Carlo replicates, resulting in 400 tests of each population-

226   and individual-based method. We enumerated how often each method falsely identified loci

227   under selection in each of the surfaces (type I error). We also summarized the number of times

228   outlier programs failed to detect the 1 locus under selection in each run and under each scenario

229   (type II error).

230   *Population-based methods*. – We tested the ability of the frequentist D_{FDIST} F_{ST} outlier approach

231   (Beaumont & Nichols 1996) implemented in to detect loci under selection in the simulated

232   dataset. Using LOSITAN (Antão *et al*. 2008), distributions of F_{ST} values were simulated from

233   thousands of loci under neutral expectations and candidate loci were identified as falling outside

234   of confidence intervals from a plot of average locus heterozygosity versus F_{ST} (Beaumont &

235   Nichols 1996; Beaumont & Balding 2004). We conducted pairwise population comparisons

236   using 500,000 replicates for the 100 loci under an infinite alleles mutation model. We tested for

237   outliers at both the 95 % and 99 % confidence interval levels and used a false discovery rate

238   (FDR) of 0.1.

239   We compared results to those produced using BayeScan v.2.1 (Foll & Gaggiotti 2008),

240   which incorporates a Bayesian approach and directly estimates the posterior probability of a

241   locus being under selection by comparing global and population-specific allele frequencies

242   derived from F_{ST} coefficients. The probabilities of models including and excluding selection are

243   estimated separately through a reversible-jump Markov chain Monte Carlo algorithm. BayeScan

244   is based on a multinomial-Dirichlet distribution, and incorporates estimates of gene flow based

245   on detection of linkage disequilibrium among loci pairs. BayeScan is considered to be more

246   realistic in describing ecological scenarios such as migration in comparison to the Beaumont &

247   Nichols (1996) method, which applies more simplified demographic models (Beaumont 2005;

248   Foll & Gaggiotti 2008).

249   To identify loci under selection using BayeScan, we first performed 20 pilot runs of

250   50,000 iterations, then used 100,000 iterations on a sample size of 5,000 and thinning interval of

251   10, setting the prior odds value to10. We used posterior probabilities establishing Log10 values

252   of the posterior odds > 1and 2 ("strong" and "decisive" support [Jefferys 1961]) that the model

253   included loci under selection. This probability is not directly comparable to P values (Foll &

254   Gaggiotti, 2008), but does allow for using threshold values obtained from the FDR to calculate q

255   values, which are analogous to P values. We identified outlier loci as those with FDR q values <

256   0.05 and 0.01, meaning outlier loci are identified by values lower than the threshold, and are

257   expected to have 5 % and 1 % false positive rate, respectively.

258   *Individual-based methods*. – We tested the ability of the individual-based spatial analysis method

259   Samβada (Joost *et al*. 2007) to estimate the probability of the presence of an allele given the

260   sampling site environmental characteristics. Samβada tests for associations between all possible

261   pairs of each environmental variable and each allele frequency at spatially referenced sample

262   locations using logistic regressions.   In our simulations, the environmental variable used as the

263   response was the latitude at each individual's location, corresponding to the north-south selection

264   gradient, and selection surfaces with selection present in the north and absent in the south (Fig.

265   1). Models including and excluding the latitude selection surface were compared through

266   examination of the significance of regression coefficients. Models were considered significant at

267   the Bonferroni-corrected alpha levels of 0.05 and 0.01 if both likelihood ratio (G) and Wald tests

268   indicated latitude was more informative than a model with a constant only (Joost *et al*. 2007).

269   We also used generalized linear models (GLMs) to detect loci under selection identified as

270   alleles correlated with environmental characteristics (latitude). Alleles for each locus were

271   defined as "present" or absent", and used as the response variable in the models. We identified

272   loci as under selection when they were significant at 95% and 99% confidence levels (CIs) after

273   a Bonferonni correction.

274

275   **Results**

276   *Genetic differentiation of adaptive vs. neutral markers*

277   Linkage disequilibrium-based estimates of N_e in local subpopulation ranged from approximately

278   10–1,741, and were higher in the scenarios with longer dispersal distances (Table 3). F_{ST}

279   estimates were highest in the low N_m, high drift (small N_e) scenarios, and lowest in the high N_m,

280   low drift (large N_e) scenarios, as expected. Values of N_m calculated from F_{ST} were accordingly

281   lowest in the small populations with highest genetic differentiation between the two populations.

282   Results from the population-based genetic distance estimator D_{est} revealed development

283   of substantial differences between adaptive and neutral alleles over time in the selection

284   present/absent surface, but little difference in the selection gradient (Fig. 2 and Appendix S2,

285   Supporting Information). Differences were especially clear with strong selection present/absent

286   (Fig. 1a) in the scenarios with short dispersal distances (5 % of maximum dispersal distance in

287   the study area [9 km]; Fig. 2ai and iii). Differences between adaptive and neutral genetic

288   structure developed rapidly with the adaptive locus reaching the maximum D_{est} value (0.88 for

289   adaptive loci versus only 0.06 and 0.01for neutral loci in the populations with both low and high

290   genetic drift, respectively) by 200 generations, and remaining at this level throughout the rest of

291   the 500 simulated generations.

292   The greatest differentiation occurred in the low dispersal, low drift scenario, as neutral

293   structure did not develop over time (Fig. 2aiii), whereas in the low dispersal, high drift scenario

294   neutral D_{est} values gradually increased to 0.1 by 500 generations (Fig. 2ai). In the scenarios with

295   strong selection and high dispersal, D_{est} values increased at a similar rate initially to the low

296   dispersal scenarios, but then declined slightly and plateaued after approximately 100 generations

297   (Fig. 2aii and iv). Adaptive and neutral alleles also showed substantial differentiation where

298   present/absent selection was weak and dispersal low, but developed at a slower rate reaching D_{est}

299   = 0.62 and 0.57 at 500 generations for populations with high and low drift, respectively (Fig. 2bi

300   and iii).

301   The adaptive locus and neutral loci both showed minimal differentiation in selection

302   gradients as measured by the D_{est} estimator ($D_{est} \leq 0.25$). With weak selection, high drift, and

303   low dispersal, D_{est} based on the adaptive locus rose more rapidly than D_{est} based on neutral loci

304   and reached its highest genetic differentiation between marker types at approximately 300

305   generations (Appendix S2, Supporting Information). In scenarios with relatively low drift,

306   slightly more genetic structure developed for both marker types within 100 generations, but both

307   followed the same pattern. D_{est} values of the combined neutral and adaptive loci followed the

308   same trend as the neutral loci in all selection surfaces and scenarios. In contrast, when selection

309   was absent genetic structure did not develop except for in the low dispersal, high drift scenario

310   reaching $D_{est} = 0.12$ after 500 generations (Appendix S3. Supporting Information).

311   Results from the landscape genetics approach measured by Mantel r indicated substantial

312   development of neutral and adaptive genetic structure over time in both the selection

313   present/absent and selection gradient surfaces, whereas the population-based approach only

314   detected this pattern in the selection present/absent surface. Genetic structure was strongest and

315   developed the quickest in scenarios with low dispersal in both selection surfaces. Adaptive

316   differentiation between neutral and adaptive loci was most pronounced in the selection gradient

317   surfaces with low dispersal, for both weak and strong selection. In this case, the genetic

318   structure based on neutral and combined loci increased to $r = 0.7$ with both strong and weak

319   selection after 500 generations, whereas the genetic structure based on the adaptive locus

320   remained r < 0.2 for the same duration of time (Fig. 3 bi and biii).

321

322 *Identifying loci under selection*

323 Our results revealed that high dispersal ($N_m > 19$) and genetic drift ($N_e < 121$) can reduce the

324 ability of both population- and individual-based methods to detect loci under selection. All 4

325 methods had low type II error rates when selection was strong in present/absent scenarios, but

326 individual-based methods performed best in selection gradients, and in the present/absent surface

327 when selection was weak. Overall population-based methods had the lowest type I error rates,

328 and GLMs had the highest.

329 *Population-based methods*. – With strong selection in the present/absent surface, both

330 population-based methods (LOSITAN and BayeScan) rarely failed to reject the false null

331 hypothesis (failed to detect the adaptive locus, type II error; Table 4). With weak selection

332 however, LOSITAN failed to detect the adaptive locus in ≥90% of the Monte Carlo replicates at

333 both confidence levels in all scenarios after 50 generations (Appendix S4, Supporting

334 Information). After 500 generations Type II error rates declined to 0% in the scenarios with

335 large populations, but remained high for the small populations (high drift) with both high and

336 low dispersal (Table 4). Bayescan failed to detect the adaptive locus in 60–100 % of the Monte

337 Carlo replicates in all scenarios after 50 generations. BayeScan also performed better after 500

338 generations in the large populations, but still failed to reject the false null hypothesis in ≥ 70% of

339 the replicates in the scenarios with high drift (Table 4). Type I error rates were low (≤ 0.1) for

340 both methods at the 99 % CI levels, but high at the 95 % CI for the low dispersal, high drift

341 scenario with LOSITAN (Table 5).

342   Both population-based methods had high type II error rates in both the weak and strong

343   selection gradient after 50 and 500 generations (Table 4, Appendix S4, Supporting Information).

344   The exception was BayeScan, which detected the locus under selection in 100 % of the replicates

345   after 50 generations when selection was strong, the population large, and dispersal low. Neutral

346   markers were incorrectly identified as outliers with both population-based methods in all

347   scenarios (Table 5). In particular LOSITAN identified between 1–5 false outliers in 10–60 % of

348   the replicates at 95 % CI with selection both weak and strong and both the present/absent and

349   environmental gradient selection surfaces. Type II error rate decrease when using 99 % CIs. The

350   low dispersal, high drift scenario had the highest frequency of neutral loci identified as outliers

351   by LOSITAN. When selection was absent, population-based methods rarely falsely detected loci

352   under selection, except for LOSITAN when dispersal was low and drift high (error rate = 0.8;

353   Appendix S4, Supporting Information).

354   *Individual-based methods*. – Both methods (GLMs and Samβada) consistently detected the

355   adaptive locus in the present/absent surface with selection both weak and strong and at 50 and

356   500 generations (Table 4, Appendix S4, Supporting Information), with one exception. In the

357   high dispersal, high drift scenario, both methods failed to reject the false null hypothesis (that an

358   adaptive locus was present), but still type II error rates were low (0.1–0.2). When a selection

359   gradient was present, error rate increased, especially in small populations with high dispersal,

360   after 500 generations. Type II error rate also increased over time when the selection gradient was

361   strong. Whereas both Samβada and GLMs had low type II error after 50 generations, both

362   methods failed to detect the adaptive locus in any of the replicates with small populations after

363   500 generations (adaptive locus was not significantly correlated with latitude; Table 4). When

364 the selection gradient was weak, Samβada detected the adaptive locus more often than GLMs.

365 Type I error rates (neutral markers were significantly correlated with latitude = false positive)

366 were quite high for both classes of methods (Table 6, Appendix S1, Supporting Information).

367 Type I error rates were high for both individual-based methods in all scenarios, but the highest

368 when dispersal levels were low, and the lowest in the selection gradient when dispersal was high.

369 When selection was absent, individual-based methods also performed poorly, especially GLMs

370 which falsely identified neutral loci as under selection in 100 % of replicates in all scenarios

371 (Appendix S1, Supporting Information). Overall, type I error rate increased over time,

372 demonstrated by an escalation of false positives between 50 and 500 generations (Table 6,

373 Appendix S1, Supporting Information). This was especially the case in scenarios with low

374 dispersal distances regardless of selection surface or strength.

375

376 **Discussion**

377 We used simulations to gain understanding of the relationship between complex environments,

378 demographic parameters, including population size and dispersal distances, and the development

379 of adaptive genetic differentiation over time. Specifically, we wished to refine understanding of

380 our capacity to quantify and avoid large type I and II error in outlier loci detection methods and

381 correlative approaches between genetic and environmental variables.

382   We found that individual-based methods performed best in selection gradients and more

383 often correctly identified gene-environment interactions, but that type I error rates were quite

384 high, especially when dispersal levels were low and drift high. Our results suggest caution is

385   required in applying both population- and individual- based methods to detect molecular signals

386   of local adaptation, and careful consideration of N_m, genetic drift, and patterns of IBD in the

387   study population is recommended when interpreting results. A combination of methods and

388   using corroborating results is the most conservative approach to identifying loci under selection;

389   an even better approach is to additionally conduct simulations and combine with empirical data

390   to improve inference.

391   *Influence of dispersal, genetic drift and selection on adaptive differentiation*

392   Our research expanded upon recent simulation studies by introducing a spatial

393   component (Narum & Hess 2011; Villas *et al*. 2012; Lotterhos & Whitlock 2014; Forester *et al*.

394   2015; Lotterhos & Whitlock 2015) and by examining temporal trends in adaptive differentiation

395   by simulating genetic exchange over 500 generations (DeMita *et al*. 2013; Jones *et al*. 2013).

396   Our study is unique in varying population size to assess the effect of genetic drift on adaptive

397   differentiation.

398   Our results revealed that the strength and rate of development of different genetic

399   structure between neutral and adaptive loci was most highly influenced by low dispersal (e.g., 5

400   % of maximum dispersal distance in the study area [9 km]; $N_m = 1$–2), and low drift ($N_e \leq 475$).

401   As predicted, we found different geographic distributions of adaptive versus neutral alleles in

402   landscapes where natural selection results in reduced survival for particular genotypes, and these

403   differences became more distinct over time. Furthermore, the differences in adaptive and neutral

404   allele distribution was more distinct with strong selection, suggesting adaptive differentiation can

405   both occur over fine spatial scales and be detected given strong enough differences in individual

406   survival.

407   Our simulation results confirm the importance of distinguishing between adaptive and

408   neutral loci, and measuring genetic differentiation separately in studies of population structure.

409   When measured together, the combined neutral and adaptive loci generally followed the same

410   trend of genetic structure over time as the neutral loci assessed independently (Figs. 2, 3).

411   Further, when neutral and adaptive loci were combined, there was little difference in the rate or

412   strength of IBD development among the same scenarios comparing strong and weak selection, or

413   the selection present/absent or gradient surfaces. Without testing for loci under selection and

414   assessing genetic structure of adaptive and neutral loci separately, these patterns of genetic

415   variation resulting from adaptation to different environmental conditions would not be detected.

416   Significantly, identifying and removing adaptive loci improves inference of widely applied

417   population genetics parameters such as F_{ST} and N_m by removing bias of loci that do not behave

418   according to neutral expectations.

419   We also found differences in the ability of the population-based (D_{est}) and individual-

420   based (Mantel r) tests to measure development in genetic structure given different selection

421   surfaces. D_{est} was effective at detecting development of adaptive differentiation only in the

422   selection present/absent surface, whereas Mantel r detected this pattern in both the selection

423   present/absent and selection gradient surfaces. These results suggest that the individual-based

424   methods for detecting structure commonly used in landscape genetic studies are most appropriate

425   when a selection gradient is present, or when populations have a clinal genetic pattern. As

426   demonstrated with simulation modeling (Landguth & Schwartz 2014), matching the metric used

427   to detect differentiation given continuous or clustered genetic patterns is a key consideration.

428   *Power and limitations of methods*

429   We predicted that selection would be stronger in large populations with shorter simulated

430   dispersal distances, and therefore adaptive loci would be more readily detectable. We further

431   expected that outlier loci detection methods would have corresponding lower type I and II error

432   rates under these scenarios. We found that the performance of both population- and individual-

433   based methods to detect loci under selection (type II error) was in fact improved with low

434   dispersal and low genetic drift (large populations), but that type I error (falsely identifying

435   neutral loci as under selection) was highest using individual-based methods under all scenarios,

436   in particular low dispersal and high drift.

437   The efficacy of natural selection is predicted as the product of the effective population

438   size and the strength of selection ($N_e * s$), and selection is expected to be less effective than

439   genetic drift in changing alleles frequencies when the value of this equation is < 1 (Li 1978).

440   The low dispersal, high drift scenario with weak selection ($s = 0.01$) produced a value < 1 (0.15

441   in the present/absent selection surface, and 0.25 in the selection gradient surface); all the other

442   scenarios were > 1. We would therefore expect that the high type I error rates produced by the

443   individual-based methods in this scenario are a result of drift overpowering selection in small

444   populations. These results suggest that caution is required when interpreting results of outlier

445   and correlative tests for adaptive differentiation, especially in small and isolated populations

446   where genetic drift exerts a greater influence than natural selection.

447   We were able to directly compare population- and individual-based methods in our

448   analyses by examining 2 simulated populations separated by a partial barrier.   In addition to

449   dispersal and drift, differences in rates of type II error (false negatives) between the methods

450   were influenced by the strength of selection and the selection surfaces. When environments

451   were distinct (present versus absent selection), and selection was strong, both classes of methods

452   performed well, whereas with weak selection, only the individual-based methods consistently

453   detected the locus under selection. In contrast, type II error was very high (0.7–1.0) in the

454   selection gradient surface when using population-based methods but generally low for

455   individual-based methods (with the exception of in populations with high dispersal and high

456   drift). DeMita *et al*. (2013) also compared individual- and population-based methods and found

457   that the former was more powerful in detecting adaptive loci in a selection gradient. These

458   results suggest that in selection gradients, individual-based methods most often correctly identify

459   gene-environment interactions, especially with large populations and low dispersal (type II error

460   rate 0–0.3) in comparison to population-based methods (type II error rate 0.7–1.0; Table 4).

461   Jones *et al*. (2013) found that individual-based methods produced few type I errors (0.0–

462   0.2) within a selection gradient surface when selection varied from weak to strong (e.g., *s* =

463   0.01–0.5). Low type I error rates were attributed to high dispersal in all of the simulated

464   populations (equivalent to 25% maximum dispersal distance of their simulated study area) which

465   resulted in minimal development of IBD patterns. Our analyses varied dispersal distances and

466   thus expanded upon quantification of error rates using individual-based methods, and

467   demonstrated generally high type I error rates (0.3–1, Appendix S1, Supporting Information,

468   Table 6). Samβada performed better than GLMs for detecting selection in all scenarios, but still

469   produced high type I error rates between 0.7–1.0 except for the scenario with high dispersal

470   (50% of landscape) and high drift (Appendix S1, Supporting Information). The high dispersal,

471   high drift error rate (0.3) fell within the range of values of error rates reported by Jones *et al*.

472   (0.1–0.4) reflecting similar dispersal distances relative to the scale of the simulated landscape.

473   These results corroborate that individual-based approaches may have a relatively low

474   type I error rate (falsely identifying neutral loci as significantly correlated with environment)

475   given comparatively high dispersal. However, our results support the conclusion of DeMita *et*

476   *al*. (2013) that population-based methods have lower type I error rates in comparison to

477   individual-based approaches and are particularly effective in detecting adaptive variation when

478   selection is strong. Furthermore, our investigation into the effects of varying dispersal distances

479   revealed limitations to the application of individual-based methods given low dispersal.

480   Our results indicate that dispersal influenced development of an IBD pattern in neutral

481   and adaptive loci over time. Over the 500 simulated generations, IBD developed most quickly in

482   the scenarios with the shortest dispersal distances, and had a diminishing rate of increase and

483   lower resulting D_{est} and Mantel r values at generation 500 with longer dispersal distances.

484   Spatial autocorrelation of genetic variation can arise in continuously distributed populations from

485   limited dispersal, mating patterns, and other demographic processes. These spatially structured

486   processes can be difficult to distinguish from spatial patterns of natural selection, and

487   demonstrates a drawback of population genetics statistical methods that do not account for

488   spatial autocorrelation (Neel *et al*. 2013).

489   Dispersal patterns leading to IBD can generate spurious correlations of selection

490   gradients and allele frequencies, resulting in high type I error (De Mita *et al*. 2013). Our work

491   supports this conclusion, as well as evidence from other research that a strong pattern of IBD can

492   reduce the effectiveness of landscape genetic methods to detect loci affected by selection

493   (Meirmans 2012; Jones *et al*. 2013; Lotterhos & Whitlock 2014; Forester *et al*. 2015) and

494   produce high levels of false positives. We concur that these methods may be most applicable to

495   situations where strong population structure is not present, such as in populations with clinal

496   genetic patterns. However, in our analysis overall type I error rates were lower using population-

497   based methods than with individual-based methods particularly when populations were large and

498   dispersal distances short. Thus, we suggest choice of methods based on assessment of N_m,

499   genetic drift, and patterns of IBD in the study population. In most situations it would be most

500   appropriate to use both approaches for conservative detection of both false positives and false

501   negatives (De Mitta *et al*. 2013; Jones *et al*. 2013; Forester *et al*. 2015).

502

503   *Conclusions*

504   Our research provides insights into molecular adaptation and detection of natural

505   selection. We performed our analyses in a spatially-explicit modelling scenario using realistic

506   population structure, which provided richer insights into the influences of spatial genetic

507   structure on adaptive divergence than simplistic ideal models (e.g., Wright-Fisher). Thus, this

508   work advances knowledge about evolutionary mechanisms influencing adaptive differentiation.

509   In this context we found individual-based methods performed well in selection gradients

510   suggesting these methods are more appropriate for continuously distributed populations.

511   However, considering the effects of spatial autocorrelation of genetic variation, our results

512   suggest some limitations in the capacity of these increasingly applied statistical methods to

513   detect molecular signatures of local adaptation, and therefore should be used with a great deal of

514   caution and in combination with other methods. Identification of correlations between candidate

515   loci and environmental variables does not necessarily indicate causation, as these patterns may

516   be spurious and moreover do not elucidate the actual mechanisms of selection. However,

517   identification of gene-environment correlations can inform further research into genes associated

518   with specific traits that may influence fitness (Hansen *et al*. 2012).

519   By improving our ability to detect and model the distribution and presence of adaptive

520   alleles across complex landscapes, this work should help transform population genetics by

521   facilitating discovery and use of adaptive genomic information (Allendorf *et al*. 2010), or

522   accurately exclude adaptive loci when only neutral loci should be used, e.g. for estimation of

523   parameters like N_e, N_m, and F_{ST} (Luikart *et al*. 2003). Our results and modeling provide guidance

524   for designing and conducting both genomic research and species monitoring or conservation

525   programs. For example, identification of patterns of adaptation is critical in designing predictive

526   models of a species' ability to persist, shift its geographic distribution (Allendorf *et al*. 2010;

527   Manel *et al*. 2010; Hoffmann & Sgrò 2011), and to maintain connectivity among populations

528   (Fortin & Dale 2005; Foll & Gaggiotti 2006) despite changes in climate patterns (Kovach *et al*.

529   2015).

530   The use of adaptive genetic markers is a promising approach for defining management

531   units and evolutionary significant units (Allendorf *et al*. 1997; Fraser & Bernatchez 2001; Funk

532 *et al*. 2012), and would provide useful insight to the appropriate scale for current and long-term

533 population management and monitoring strategies. Further, discovery of locally-adapted

534 populations may inform management actions and prioritize policy decisions so that the

535 evolutionary processes themselves may be targeted for conservation (Funk *et al*. 2012; Limborg

536 *et al*. 2012). However, we caution that verification that a genetic marker is adaptive (or linked to

537 adaptive loci) can require detection of the same outlier behavior in multiple independent

538 populations and/or experiments (Allendorf et al. 2013). We hope this work sparks additional

539 research and understanding of natural selection in spatially heterogeneous environments by

540 providing insights into factors that influence adaptive vs. neutral population genetic structure and

541 appropriate methods to study them.

542

550 **References**

551 Allendorf FW, Bayles D, Bottom DL, Currens KP, Frissell CA, Hankin D, Lichatowich JA,
552   Nehlsen W, Trotter PC, Williams TH (1997) Prioritizing Pacific salmon stocks for
553   conservation. *Conservation Biology*, **11**, 140–152.

554 Allendorf FW, Hohenlohe PA., Luikart G (2010) Genomics and the future of conservation
555   genetics. *Nature Reviews Genetics*, **11**, 697–709.

556 Allendorf FW, Luikart G, Aitken SN (2013) Conservation and the genetics of populations. 2nd
557   ed. Oxford (UK): Wiley-Blackwell.

558 Antão T, Lopes A, Lopes RJ, Beja-Pereira A, Luikart G (2008) LOSITAN: a workbench to
559   detect molecular adaptation based on a Fst-outlier method. *BMC Bioinformatics*, **9**, 323.

560 Beaumont MA, Nichols RA (1996) Evaluating loci for use in the genetic analysis of population
561   structure. *Proceedings of the Royal Society B: Biological Sciences*, **263**, 1619–1626.

562 Beaumont MA., Balding DJ (2004) Identifying adaptive genetic divergence among populations
563   from genome scans. *Molecular Ecology*, **13**, 969–980.

564 Beaumont MA (2005) Adaptation and speciation: what can F(st) tell us? *Trends in Ecology &*
565   *Evolution*, **20**, 435–40.

566 Bowcock AM, Ruiz-Linares A, Tomfohrde J *et al.* (1994) High resolution of human evolutionary
567   trees with polymorphic microsatellites. *Letters to Nature*, **368**, 455–457.

568 Charbonnel N, Pemberton J (2005) A long-term genetic survey of an ungulate population reveals
569   balancing selection acting on MHC through spatial and temporal fluctuations in selection.
570   *Heredity*, **95**, 377–88.

571 Cheviron ZA, Natarajan C, Projecto-Garcia J *et al.* (2014) Integrating Evolutionary and
572   Functional Tests of Adaptive Hypotheses : A Case Study of Altitudinal Differentiation in
573   Hemoglobin Function in an Andean Sparrow, Zonotrichia capensis. *Molecular Biology*
574   *and Evolution*, **31**, 2948–2962.

575 Coop G, Pickrell JK, Novembre J *et al.* (2009) The role of geography in human adaptation. *PLoS*
576   *Genetics*, **5**, e1000500.

577 Crow JF, Kimura M (1970) An introduction to population genetics theory. New York: Harper
578   and Row.

579 De Mita S, Thuillet A-C, Gay L *et al.* (2013) Detecting selection along environmental gradients:
580   analysis of eight methods and their effectiveness for outbreeding and selfing populations.
581   *Molecular Ecology*, **22**, 1383–1399.

582 Do C, Waples RS, Peel D *et al.* (2014) NeEstimator v2: Re-implementation of software for the
583   estimation of contemporary effective population size (N_e) from genetic data. *Molecular*
584   *Ecology Resources*, **14**, 209–214.

585 Epperson BK, McRae BH, Scribner K *et al.* (2010) Utility of computer simulations in landscape
586   genetics. *Molecular Ecology*, **19**, 3549–3564.

587   Foll M, Gaggiotti O (2006) Identifying the environmental factors that determine the genetic
588   structure of populations. *Genetics*, **174**, 875–91.

589   Foll M, Gaggiotti O (2008) A genome-scan method to identify selected loci appropriate for both
590   dominant and codominant markers: a Bayesian perspective. *Genetics*, **180**, 977–93.

591   Forester BR, Jones MR, Joost S, Landguth EL, Lasky JR (2015) Detecting spatial genetic
592   signatures of local adaptation in heterogeneous landscapes. *Molecular Ecology*.

593   Fortin, M-J, Dale, MRT (2005) Spatial Analysis. Cambridge University Press, Cambridge.

594   Franks SJ, Sim S, Weis AE (2007) Rapid evolution of flowering time by an annual plant in
595   response to a climate fluctuation. *Proceedings of the National Academy of Sciences*, **104**,
596   1278–1282.

597   Fraser DJ, Bernatchez L (2001) Adaptive evolutionary conservation: towards a unified concept
598   for defining conservation units. *Molecular Ecology*, **10**, 2741–52.

599   Fournier-Level A, Korte A, Cooper MD *et al.* (2011) A Map of local adaptation in *Arabidopsis
600   thaliana*. *Science*, 86–89.

601   Funk WC, McKay JK, Hohenlohe PA, Allendorf FW (2012) Harnessing genomics for
602   delineating conservation units. *Trends in Ecology & Evolution*, **27**, 489–496.

603   Hancock AM, Witonsky DB, Gordon AS *et al.* (2008) Adaptations to climate in candidate genes
604   for common metabolic disorders. *PLoS Genetics*, **4**, e32.

605   Hancock AM, Jarymowycz LB, Sperone FG *et al.* (2011) Adaptation to climate across the
606   *Arabidopsis thaliana* genome. *Science*, **334**, 83–86.

607   Hand BK, Lowe WH, Kovach RP, Muhlfeld CC, Luikart G (2015) Landscape community
608   genomics : understanding eco-evolutionary processes in complex environments. *Trends
609   in Ecology & Evolution*, **30**, 161–168.

610   Hansen MM, Olivieri I, Waller DM, Nielsen EE (2012) Monitoring adaptive genetic responses to
611   environmental change. *Molecular Ecology*, **21**, 1311–1329.

612   Hedrick PW (2005) A standardized genetic differentiation measure. *Evolution*, **59**, 1633–8.

613   Hoban S, Bertorelle G, Gaggiotti OE (2011) Computer simulations: tools for population and
614   evolutionary genetics. *Nature reviews. Genetics*, **13**, 110–22.

615   Hohenlohe PA, Bassham S, Etter PD *et al.* (2010) Population genomics of parallel adaptation in
616   threespine stickleback using sequenced RAD tags. *PLoS genetics*, **6**, e1000862.

617   Hoffmann AA, Sgrò CM (2011) Climate change and evolutionary adaptation. *Nature*, **470**, 479–
618   85.

619   Holderegger R, Wagner HH (2008) Landscape Genetics. *BioScience*, **58**, 199–207.

620   Jeffreys H (1961) Theory of Probability, 3rd edn. Oxford University Press, London, p. 432.

621   Jones MR, Forester BR, Teufel AI *et al.* (2013) Integrating landscape genomics and spatially
622   explicit approaches to detect loci under selection in clinal populations. *Evolution*, **67**,
623   3455–68.

624   Joost S, Bonin A, Bruford MW *et al.* (2007) A spatial analysis method (SAM) to detect
625   candidate loci for selection: towards a landscape genomics approach to adaptation.
626   *Molecular Ecology*, **16**, 3955–69.

627   Jost L (2008) GST and its relatives do not measure differentiation. *Molecular Ecology*, **17**,
628   4015–4026.

629   Kohn MH, Pelz H-J, Wayne RK (2003) Locus-specific genetic differentiation at Rw among
630   warfarin-resistant rat (*Rattus norvegicus*) populations. *Genetics*, **164**, 1055–70.

631   Kovach RP, Muhlfeld CC, Wade AA, Hand BK, Whited DC, DeHaan PW, Al-Chokhachy R,
632   Luikart G (2015) Genetic diversity is related to climatic variation and vulnerability in
633   threatened bull trout. *Global Change Biology*, **21**, 2510–2524.

634   Landguth EL, Cushman SA (2010) cdpop: A spatially explicit cost distance population genetics
635   program. *Molecular Ecology Resources*, **10**, 156–61.

636   Landguth EL, Cushman SA, Murphy MA, Luikart G (2010) Relationships between migration
637   rates and landscape resistance assessed using individual-based simulations. *Molecular
638   Ecology Resources*, **10**, 854–62.

639   Landguth EL, Cushman SA, Johnson NA (2012) Simulating natural selection in landscape
640   genetics. *Molecular Ecology Resources*, **12**, 363-368.

641   Landguth EL, Balkenhol N (2012) Relative sensitivity of neutral versus adaptive genetic data for
642   assessing population differentiation. *Conservation Genetics*, **13**, 1421–1426.

643   Landguth EL, Schwartz MK (2014) Evaluating sample allocation and effort in detecting
644   population differentiation for discrete and continuously distributed individuals.
645   *Conservation Genetics*, **15**, 981–992.

646   Lenormand T (2002) Gene flow and the limits to natural selection. *Trends in Ecology &
647   Evolution*, **17**, 183–189.

648   Lewontin RC, Krakauer J (1973) Distribution of gene frequency as a test of the theory of the
649   selective neutrality of polymorphisms. *Genetics*, **74**, 175–195.

650   Li, W.H. (1978) Maintenance of genetic variability under joint effect of mutation, selection and
651   random drift. *Genetics*, **90**, 349–382.

Limborg MT, Helyar SJ, De Bruyn M *et al.* (2012) Environmental selection on transcriptome-derived SNPs in a high gene flow marine fish, the Atlantic herring (*Clupea harengus*). *Molecular Ecology*, **21**, 3686–703.

Lotterhos KE, Whitlock MC (2014) Evaluation of demographic history and neutral parameterization on the performance of FST outlier tests. *Molecular Ecology*, **23**, 2178–2192.

Lotterhos KE, Whitlock MC (2015) The relative power of genome scans to detect local adaptation depends on sampling design and statistical method. *Molecular Ecology*, **24**, 1031–1046.

Lowe WH, Allendorf FW (2010) What can genetics tell us about population connectivity? *Molecular Ecology*, **19**, 3038–3051.

Luikart G, England PR, Tallmon D, Jordan S, Taberlet P (2003) The power and promise of population genomics: from genotyping to genome typing. *Nature Reviews.Genetics*, **4**, 981–994.

Kondrashov AS (2003) Direct estimates of human per nucleotide mutation rates at 20 loci causing Mendelian diseases. *Human Mutation*, **21**, 12–27.

Kovach RP, Muhlfeld CC, Wade AA, Hand BK, Whited DC, DeHaan PW, Al-Chokhachy R, Luikart G (2015) Genetic diversity is related to climatic variation and vulnerability in threatened bull trout. *Global Change Biology*. doi: 10.1111/gcb.12850

McKay JK, Latta RG (2002) Adaptive population divergence: markers, QTL and traits. *Trends in Ecology and Evolution,* **17**, 285–291.

Manel S, Joost S, Epperson BK *et al.* (2010) Perspectives on the use of landscape genetics to detect genetic adaptive variation in the field. *Molecular Ecology*, **19**, 3760–72.

Mantel N (1967) The detection of disease clustering and a generalized regression approach. *Cancer Research*, **27**, 209– 220.

Matala AP, Ackerman MW, Campbell MR, Narum SR (2014) Relative contributions of neutral and non-neutral genetic differentiation to inform conservation of steelhead trout across highly variable landscapes. *Evolutionary Applications*, **7**, 682–701.

Meirmans PG, Hedrick PW (2011) Assessing population structure: F_{ST} and related measures. *Molecular Ecology Resources*, **11**, 5–18.

Meirmans PG (2012) The trouble with isolation by distance. Molecular Ecology, **21**, 2839–46.

Nachman MW, Crowell SL (2000) Estimate of the mutation rate per nucleotide in humans. *Genetics*, **156**, 297–304.

685   Narum SR, Campbell NR, Kozfkay CC, Meyer K a (2010) Adaptation of redband trout in desert
686   and montane environments. *Molecular Ecology*, **19**, 4622–37.

687   Narum SR, Hess JE (2011) Comparison of F_{ST} outlier tests for SNP loci under selection.
688   *Molecular Ecology Resources*, **11**(s.1), 184-194.

689   Neale DB, Kremer A (2011) Forest tree genomics: growing resources and applications. *Nature
690   reviews. Genetics*, **12**, 111–122.

691   Neel MC, McKelvey K, Ryman N *et al.* (2013) Estimation of effective population size in
692   continuously distributed populations: there goes the neighborhood. *Heredity*, **111**, 189–
693   99.

694   Prugnolle F, Manica A, Balloux F (2005) Geography predicts neutral genetic diversity of human
695   populations. *Current Biology: CB*, **15**, R159–60.

696   Rieseberg, L.H., Burke, J.M. 2001. The biological reality of species: gene flow, selection, and
697   collective evolution. *Taxon,* **50**, 47–67.

698   Savolainen O, Lascoux M, Merilä J (2013) Ecological genomics of local adaptation. *Nature
699   Reviews Genetics*, **14**, 807–820.

700   Schoville SD, Bonin A, François O *et al.* (2012) Adaptive genetic variation on the landscape:
701   methods and cases. *Annual Review of Ecology, Evolution, and Systematics*, **43**, 23–43.

702   Schwartz MK, McKelvey KS (2009) Why sampling scheme matters: the effect of sampling
703   scheme on landscape genetic results. *Conservation Genetics*, **10**, 441–452.

704   Storz JF (2005) Using genome scans of DNA polymorphism to infer adaptive population
705   divergence. *Molecular Ecology*, **14**, 671–688.

706   Thibert-Plante X, Hendry AP (2010) When can ecological speciation be detected with neutral
707   loci? *Molecular Ecology*, **19**, 2301–14.

708   Vilas A, Pérez-Figueroa A, Caballero A (2012) A simulation study on the performance of
709   differentiation-based methods to detect selected loci using linked neutral markers.
710   *Journal of Evolutionary Biology*, **25**, 1364–76.

711   Waples RS (1998) Separating the wheat from the chaff: Patterns of genetic differentiation in high
712   gene flow species. *Journal of Heredity*, **89**, 438–450.

713   Waples RS, Do C (2008) Ldne: a Program for Estimating Effective Population Size From Data
714   on Linkage Disequilibrium. *Molecular Ecology Resources*, **8**, 753–756.

715   Weir BS, Cockerham CC (1984) Estimating F-statistics for the analysis of population structure.
716   *Evolution* **38**:1358–1370.

717   Yang W-Y, Novembre J, Eskin E, Halperin E (2012) A model-based approach for analysis of
718   spatial structure in genetic data. *Nature Genetics*, **44**, 725–31.

719 Zardi GI, Nicastro KR, Canovas F *et al.* (2011) Adaptive traits are maintained on steep selective
720   gradients despite gene flow and hybridization in the intertidal zone. *PloS One*, **6**, e19402.

721

Author Contributions

723

724 G.H.R., E.L.L., M.K.S. and G.L. designed the study. G.H.R. and E.L.L. ran the simulations.
725 G.H.R. performed the analyses, interpreted the data, and wrote the manuscript with input from
726 all authors.

727

Data accessibility

729

730 Simulation data, landscape surfaces, and output files are available at Dryad doi:XXX. CDPOP
731 software and user manual are available at http://cel.dbs.umt.edu/cms/CDPOP.
732

733   Table 1. Summary of factors we varied in simulated populations. For each population size (used
734   to approximate genetic drift), different combinations of dispersal distance, and strength of
735   selection, and selection surface were simulated.

Population size	Dispersal distance[*]	Strength of selection	Selection surface
1,000	5 %	Strong (s = 0.5)[†]	Present/absent[§]
10,000	50 %	Weak (s = 0.1) [‡]	Latitudinal gradient
		Absent	Uniform[¶]

736   [*] Percentage of maximum possible dispersal distance (i.e., length the entire study area).
737   [†] 0.5 selection coefficient ($s = 1$ - relative fitness genotype *AA*). *Aa* and *aa* = uniform fitness
738   value of 1.0.
739   [‡] 0.1 selection coefficient ($s = 1$ - relative fitness genotype *AA*). *Aa* and aa genotypes = uniform
740   fitness value of 1.0.
741   [§]Two adjacent discrete regions with selection present in one and absent in the other.
742   [¶]A uniform surface with no selection to test for type I statistical errors.

743   Table 2. Scenarios used to quantify type I and II error rates for population- and individual-based
744   tests to detect loci under selection. Scenarios were tested with both strong and weak selection
745   (Table 1) and using 3 selection surfaces (present/absent, latitudinal gradient, selection absent
746   uniformly).

	low N_m[*], high drift	low N_m, low drift	high N_m, high drift	high N_m, low drift
population size	1,000	10,000	1,000	10,000
maximum % dispersal distance[†]	5 %	5 %	50 %	50 %

747   [*] N_m = effective population size (N_e) * rate of dispersal among populations; estimated through
748   dispersal distance parameter (Table 1).
749   [†]Percent of the length of the entire study area.
750

751 Table 3. Estimates of effective population size (N_e) using the linkage disequilibrium method,
752 population differentiation (F_{ST}), and number of migrants per generation (N_m) for each of the 4
753 simulation scenarios, for strong and weak directional selection (s = 0.5 and 0.05, respectively), in
754 the selection present/absent and gradient surfaces.
755

a.)	Present/Absent				
Selection strength	Estimated Metric	low N_m, high drift	low N_m, low drift	high N_m, high drift	high N_m, low drift
weak	N_e	15.25	475.20	126.90	1572.40
	F_{ST}	0.108	0.027	0.013	0.001
	N_m	2.06	8.26	19.45	249.00
strong	N_e	10.25	407.15	127.50	1549.60
	F_{ST}	0.162	0.012	0.014	0.001
	N_m	1.29	20.58	17.61	249.70
b.)	Gradient				
Selection strength	Metric	low N_m, high drift	low N_m, low drift	high N_m, high drift	high N_m, low drift
weak	N_e	25.20	456.00	136.85	1741.85
	F_{ST}	0.143	0.015	0.110	0.002
	N_m	1.50	15.67	22.42	124.00
strong	N_e	13.00	394.20	121.20	1490.20
	F_{ST}	0.131	0.014	0.013	0.001
	N_m	1.66	17.61	18.88	249.75

Table 4. Type II error rates of population- and individual-based methods to detect loci under selection after 500 generations, and the proportion of replicates where this occurred at the 95% and 99% confidence interval levels (LOSITAN, GLM, Samβada) and with FDR q value thresholds < 0.05 and 0.01(BayeScan), which are analogues of *P*-values (displayed under 95% and 99% columns). Type II error rates (false negatives) are shown for a.) selection present/absent surfaces, and b.) gradient surfaces, for both weak and strong selection scenarios.

a.)

Selection strength	Method	Present/Absent							
		low N_m, high drift		low N_m, low drift		high N_m, high drift		high N_m, low drift	
	CIs	95%	99%	95%	99%	95%	99%	95%	99%
weak	LOSITAN	0.2	0.4	0	0	1	1	0	0
	BayeScan	0.7	0.6	0.1	0.1	1	1	0.3	0.3
	Samβada	0	0	0	0	0.2	0.2	0	0
	GLM	0	0	0	0	0.1	0.1	0	0
strong	LOSITAN	0	0	0	0	0	0	0	0
	BayeScan	0	0	0.2	0.2	0	0	0	0
	Samβada	0	0	0	0	0.2	0.2	0	0
	GLM	0	0	0	0	0	0	0	0

b.)

Selection strength	Method	Gradient							
		low N_m, high drift		low N_m, low drift		high N_m, high drift		high N_m, low drift	
	CIs	95%	99%	95%	99%	95%	99%	95%	99%
weak	LOSITAN	0.8	0.9	1	1	0.8	0.8	1	1
	BayeScan	1	1	1	1	1	1	1	1
	Samβada	0	0	0	0	0.5	0.5	0	0
	GLM	0.2	0.2	0	0	0.7	0.7	0	0
strong	LOSITAN	0.8	0.7	1	1	0.9	0.9	1	1
	BayeScan	1	1	1	1	1	1	1	1
	Samβada	0.3	0.3	0.3	0.3	1	1	0	0
	GLM	1	1	0	0	1	1	0.4	0.4

Table 5. Type I error rates of population-based methods defined as the number of neutral loci (1 - 6) identified as outliers using LOSTIAN and BayeScan, and the proportion of replicates where this occurred at the 95% and 99% confidence interval levels (LOSITAN) and with FDR q value thresholds < 0.05 and 0.01(BayeScan), which are analogues of P-values (displayed under 95% and 99% columns). The expected number of outlier loci for 95% and 99% is 5 and 1, respectively (with 100 simulated loci).

		Present/absent									Gradient							
	Number of loci	low N_m, high drift		low N_m, low drift		high N_m, high drift		high N_m, low drift		Number of loci	low N_m, high drift		low N_m, low drift		high N_m, high drift		high N_m, low drift	
		95%	99%	95%	99%	95%	99%	95%	99%		95%	99%	95%	99%	95%	99%	95%	99%
weak selection	*LOSITAN*									*LOSITAN*								
	1	0.6	0.0	0.3	0.0	0.0	0.0	0.0	0.0	1	0.3	0.3	0.3	0.0	0.2	0.0	0.0	0.0
	2	0.0	0.1	0.0	0.0	0.0	0.0	0.0	0.0	2	0.3	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	3	0.2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	>4	0.1*	0.0	0.0	0.0	0.0	0.0	0.0	0.0	>4	0.3§	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	BayeScan									*BayeScan*								
	1	0.0	0.0	0.1	0.1	0.0	0.0	0.1	0.1	1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	2	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
strong selection	*LOSITAN*									*LOSITAN*								
	1	0.3	0.3	0.1	0.0	0.2	0.1	0.0	0.0	1	0.2	0.1	0.2	0.0	0.1	0.1	0.0	0.0
	2	0.2	0.2	0.0	0.0	0.1	0.0	0.0	0.0	2	0.2	0.1	0.0	0.0	0.0	0.0	0.0	0.0
	3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	3	0.2	0.0	0.0	0.0	0.1	0.0	0.0	0.0
	>4	0.3†	0.0	0.0	0.0	0.0	0.0	0.0	0.0	4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	BayeScan									*BayeScan*								
	1	0.2	0.0	0.0	0.0	0.1	0.0	0.0	0.0	1	0.0	0.0	0.1	0.1	0.0	0.0	0.0	0.0
	2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

*One Monte Carlo replicate detected 5 false positive loci.

† Two Monte Carlo replicates detected 4, and one detected 6 false positive loci.

§ Two Monte Carlo replicates detected 4, and one detected 5 false positive loci.

Table 6. Type I error rates (false positives) of individual-based methods (Samβada and GLMs) defined as the mean number of neutral loci significantly correlated with the selection gradient (latitude) at the 95% and 99% confidence interval levels after 500 generations. The expected number of significantly correlated loci for 95% and 99% is 5 and 1, respectively (with 100 simulated loci). Type I error rates are shown for a.) selection present/absent surfaces, and b.) gradient surfaces, for both weak and strong selection scenarios (where one locus was under selection).

a.) Present/absent

Selection strength	Method	low N_m, high drift		low N_m, low drift		high N_m, high drift		high N_m, low drift	
	CIs	95%	99%	95%	99%	95%	99%	95%	99%
weak	Samβada	56	52	48	45	19	14	28	22
	GLM	57	54	49	47	18	14	30	25
strong	Samβada	57	54	50	47	16	13	21	17
	GLM	59	56	52	49	18	15	22	19

b.) Gradient

Selection strength	Method	low N_m, high drift		low N_m, low drift		high N_m, high drift		high N_m, low drift	
	CIs	95%	99%	95%	99%	95%	99%	95%	99%
weak	Samβada	43	40	49	47	16	13	28	24
	GLM	58	54	50	48	18	15	30	26
strong	Samβada	56	53	50	47	14	10	26	23
	GLM	58	54	51	49	15	12	29	25

Figure 1. Simulated 200 x 300 km landscape with 1,000 individuals (black filled dots). In the selection present/absent scenario (a), individuals with the AA genotype have a relative fitness value of 1.0 in the light grey area, and 0.5 (strong selection), or 0.1 (weaker selection) in the dark grey area, and all other genotypes have a relative fitness value of 1.0. In the gradient scenario (b) individuals with the AA genotype have a relative fitness value of 1.0 on the northern (upper) end of the landscape (white), and linearly decreasing values to 0.5 (strong) or 0.1 (weak) at the lower end (dark grey). Scenario c.) represents a uniform surface with no selection. The Aa and aa genotypes have a uniform relative fitness value of 1.0 for both strong selection weak selection in scenarios a.), b.), and c.).
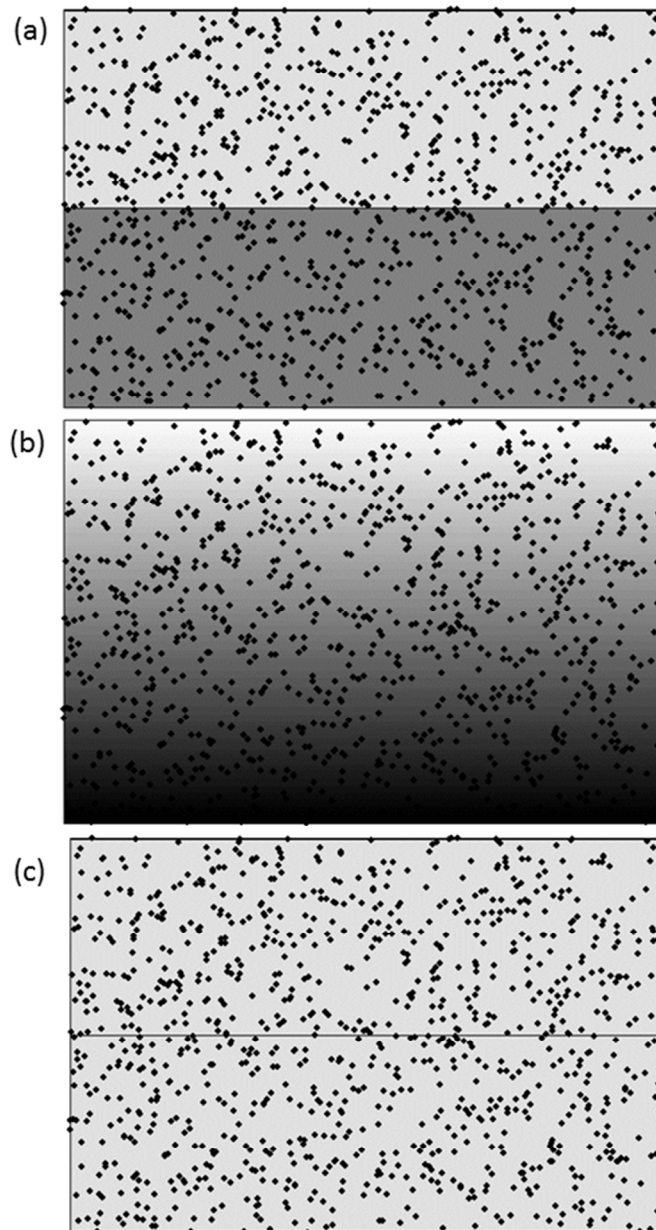
Figure 2. D_{est} results depicting genetic differentiation for adaptive (selection-driven), neutral and combined loci. The present/absent selection surface is shown with selection strong (a) and weak (b), for the i) low N_m, high drift, ii) high N_m, high drift, iii) low N_m, low drift, iv) high N_m, low drift scenarios. See Tables 5.1 and 5.2 for descriptions.
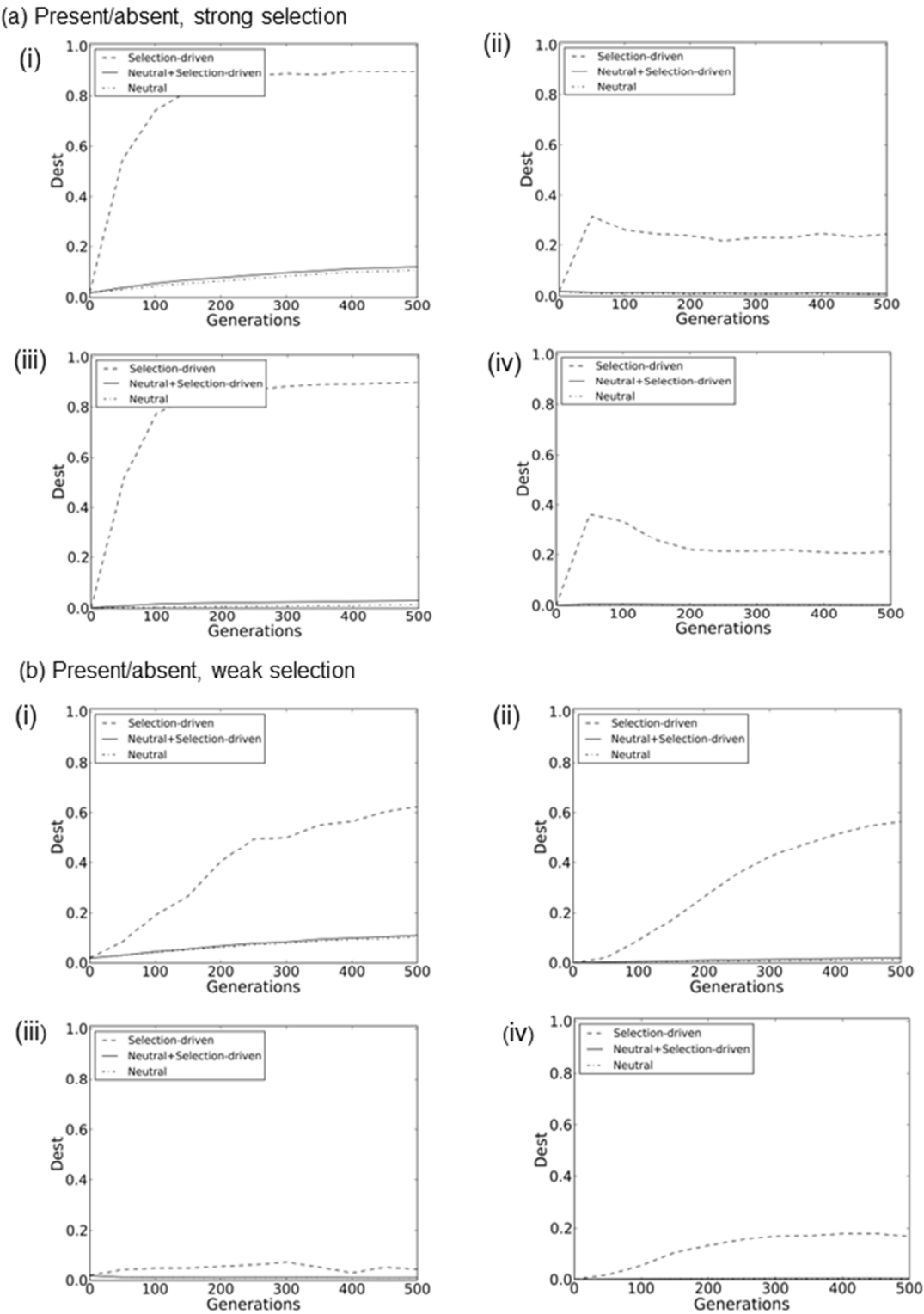

(a) Present/absent, strong selection

(b) Present/absent, weak selection

Figure 3. Mantel *r* results depicting genetic differentiation for adaptive (selection-driven), neutral and combined loci. The present/absent selection (a) and selection gradient (b) surfaces are shown with selection strong for the i) low N_m, high drift, ii) high N_m, high drift scenarios, and selection weak for the iii) low N_m, low drift, iv) high N_m, low drift scenarios. See Tables 5.1 and 5.2 for descriptions.
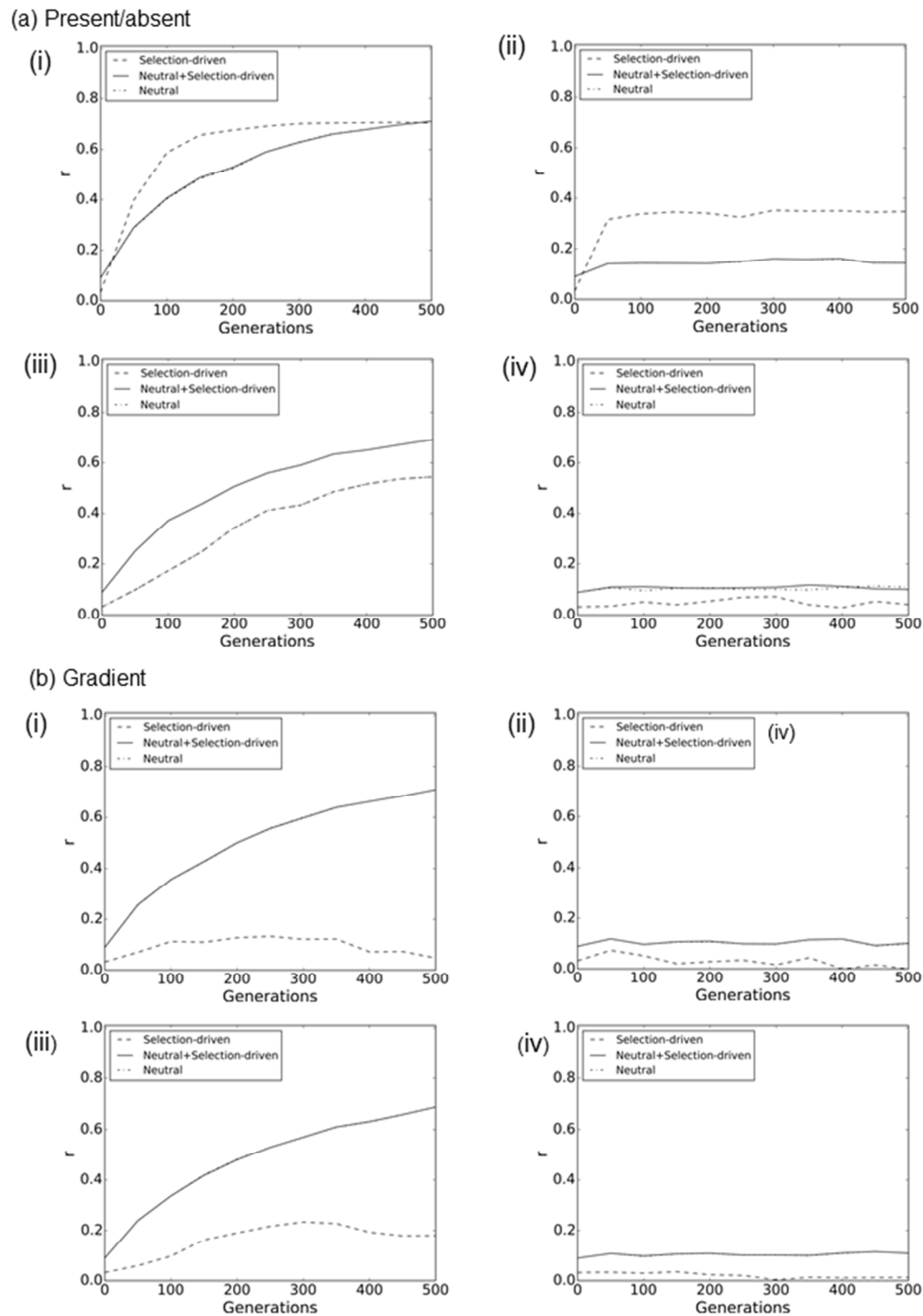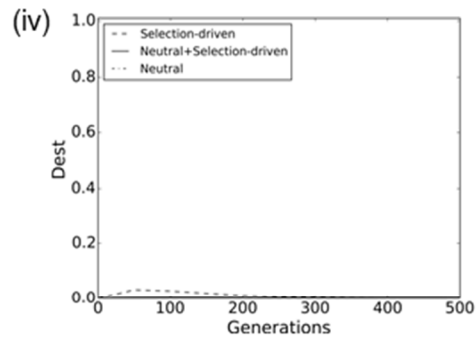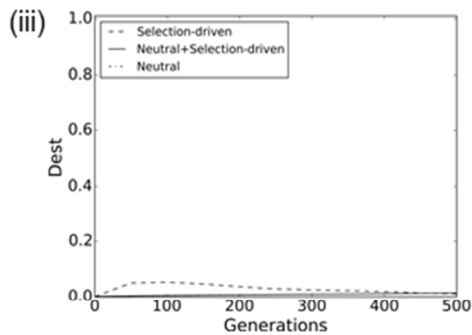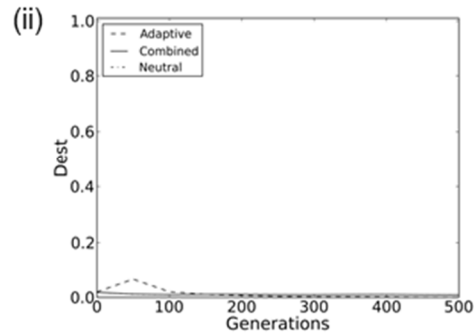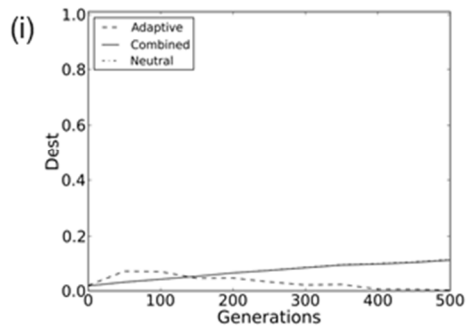
APPENDIX S1. Type I error rates of population- and individual-based methods when using the uniform selection surface (selection absent). Type I error rates are reported at 50 and 500 generations, and refer to the proportion of replicates where neutral loci were identified as under selection at the 95% and 99% confidence interval levels (LOSITAN, GLM, Samβada) and with FDR q value thresholds < 0.05 and 0.01(BayeScan), which are analogues of *P*-values (displayed under 95% and 99% columns).
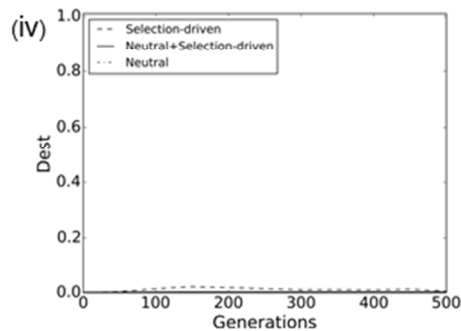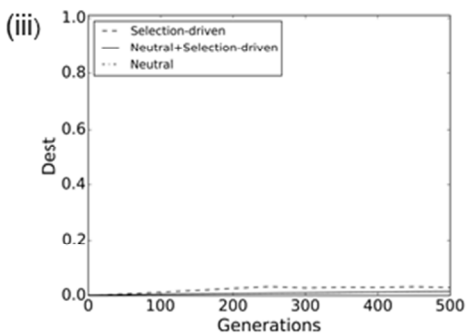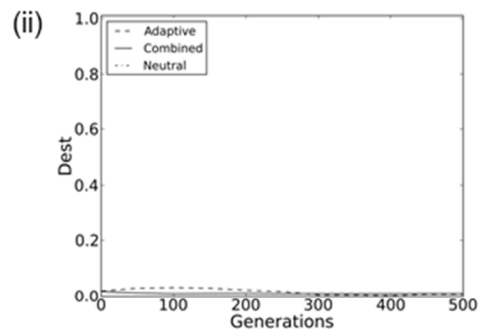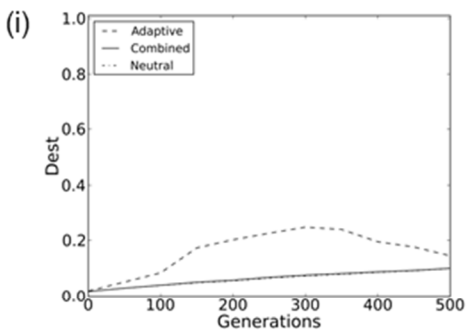
Selection surface	low N_m, high drift		low N_m, low drift		high N_m, high drift		high N_m, low drift	
Absent	95%	99%	95%	99%	95%	99%	95%	99%
LOSITAN								
50	0.3	0.1	0	0	0.1	0	0	0
500	0.8	0.1	0.1	0	0.2	0.1	0	0
BayeScan								
50	0.1	0.1	0.1	0.1	0	0	0	0
500	0.1	0.1	0.1	0.1	0.1	0.1	0	0
Samβada								
50	1	1	0.7	0.7	0.8	0.8	0.8	0.8
500	1	1	0.7	0.7	0.3	0.3	0.7	0.7
GLM								
50	1	1	1	1	1	1	1	1
500	1	1	1	1	1	1	1	1

APPENDIX S2. D_{est} results depicting genetic differentiation for adaptive (selection-driven), neutral and combined loci. The selection gradient surface is shown with (a) selection strong ($s = 0.5$) and (b) weak ($s = 0.1$), for the i) low N_m, high drift, ii) high N_m, high drift, iii) low N_m, low drift, iv) high N_m, low drift scenarios. See Tables 1 and 2 for descriptions.

(a) Gradient, strong selection

(b) Gradient, weak selection

**APPENDIX S3**. D_{est} (a) and Mantel r (b) results depicting genetic differentiation for adaptive (selection-driven), neutral and combined loci in the uniform selection surface (selection absent) for the i) low Nm, high drift, ii) high Nm, high drift, iii) low Nm, low drift, iv) high Nm, low scenarios. See Tables 1 and 2 for descriptions.

(a) D_{est}, uniform selection (absent)
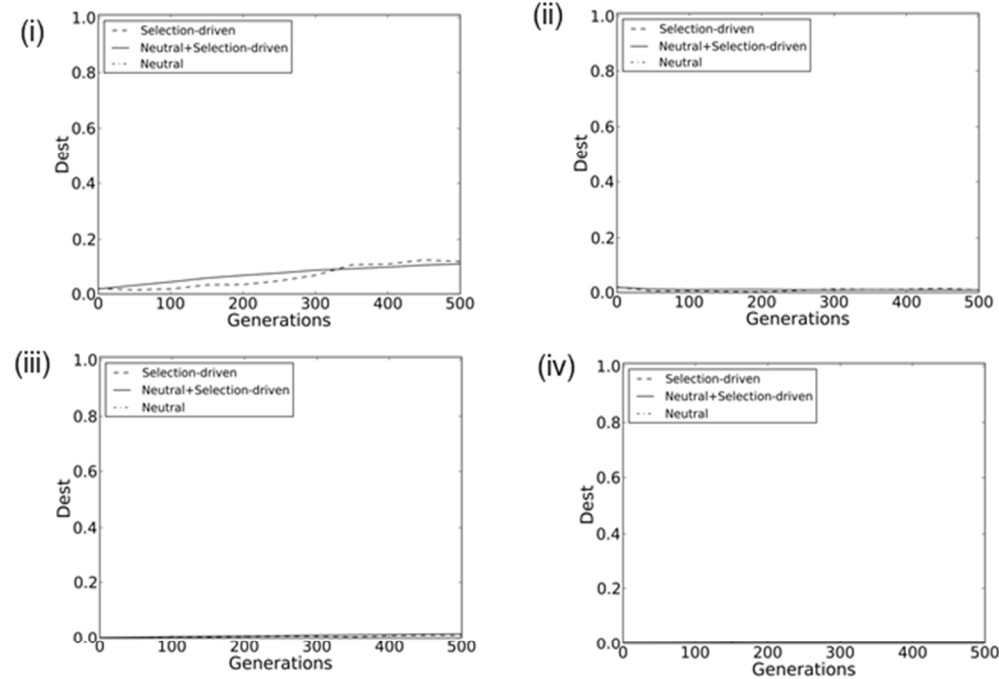
(b) Mantel r, uniform selection (absent)
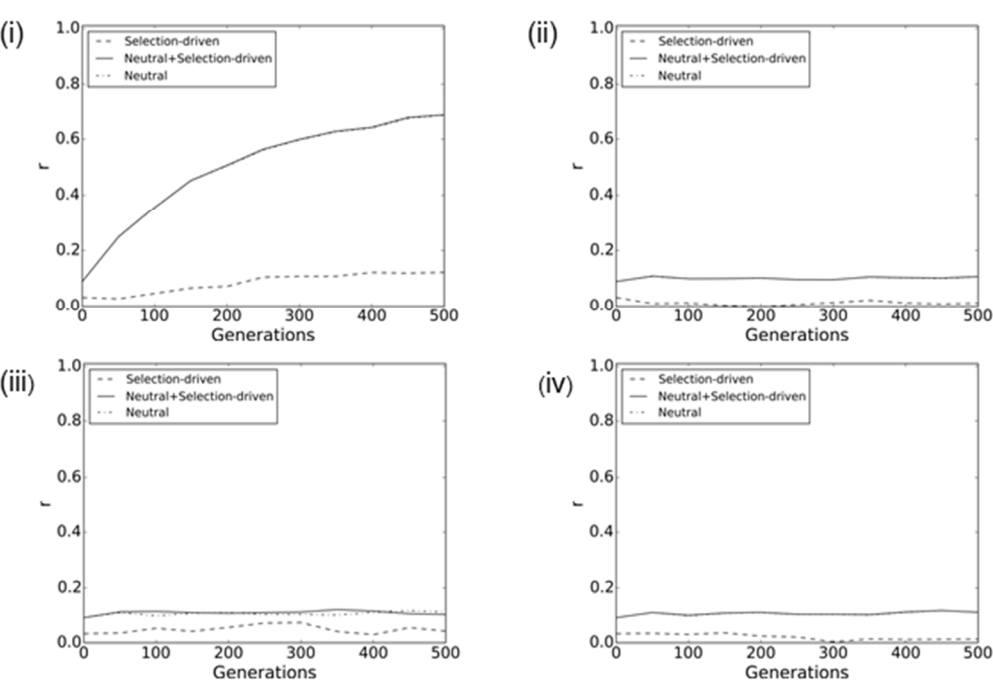
APPENDIX S4. Type II error rates of population- and individual-based methods to detect loci under selection after 50 generations, and the proportion of replicates where this occurred at the 95% and 99% confidence interval levels (LOSITAN, GLM, Samβada) and with FDR q value thresholds < 0.05 and 0.01(BayeScan), which are analogues of *P*-values (displayed under 95% and 99% columns). Type II error rates are shown for a.) selection present/absent surfaces, and b.) gradient surfaces, for both weak and strong selection scenarios.

a.)

Selection strength	Method	Present/Absent							
		low N_m, high drift		low N_m, low drift		high N_m, high drift		high N_m, low drift	
	CIs	95%	99%	95%	99%	95%	99%	95%	99%
weak	LOSITAN	1	1	1	1	0.9	1	1	1
	BayeScan	1	1	0.6	0.6	1	1	0.7	0.7
	Samβada	0	0	0	0	0.2	0.2	0	0
	GLM	0	0	0	0	0.2	0.2	0	0
strong	LOSITAN	0	0	0	0	0	0	0	0
	BayeScan	0	0	0	0	0	0	0	0
	Samβada	0	0	0	0	0.2	0.2	0	0
	GLM	0	0	0	0	0	0	0	0

b.)

Selection strength	Method	Gradient							
		low N_m, high drift		low N_m, low drift		high N N_m, high drift		high N_m, low drift	
	CIs	95%	99%	95%	99%	95%	99%	95%	99%
weak	LOSITAN	1	1	1	1	0.9	1	1	1
	BayeScan	1	1	1	1	1	1	1	1
	Samβada	0.2	0.2	0.2	0.2	0.3	0.3	0	0
	GLM	0.3	0.3	0.1	0.1	0.3	0.3	0	0
strong	LOSITAN	0.9	1	0.7	1	0.7	0.9	1	1
	BayeScan	1	1	0	0	0.9	1	0.1	0.1
	Samβada	0.2	0.2	0	0	0	0	0	0
	GLM	0	0	0	0	0.1	0.1	0	0

APPENDIX S5. Type I error rates of individual-based methods (Samβada and GLMs) defined as the mean number of neutral loci significantly correlated with the selection gradient (latitude) at the 95% and 99% confidence interval levels after 50 generations. The expected number of significantly correlated loci for 95% and 99% is 5 and 1, respectively (with 100 simulated loci).

a.)		Present/absent							
Selection strength	Method	low N_m, high drift		low N_m, low drift		high N_m, high drift		high N_m, low drift	
	CIs	95%	99%	95%	99%	95%	99%	95%	99%
weak	Samβada	27	24	30	21	16	13	21	17
	GLM	28	26	25	22	18	14	22	19
strong	Samβada	29	25	25	21	16	13	20	17
	GLM	31	27	26	23	16	14	22	19
b.)		Gradient							
Selection strength	Method	low N_m, high drift		low N_m, low drift		high N_m, high drift		high N_m, low drift	
	CIs	95%	99%	95%	99%	95%	99%	95%	99%
weak	Samβada	26	22	27	22	15	12	22	18
	GLM	27	24	29	25	17	14	23	20
strong	Samβada	27	23	26	22	15	12	23	19
	GLM	35	29	27	25	17	14	25	22