1 **Coalescent models at small effective population sizes and population declines**

2 **are positively misleading**

3 Corresponding author:

4 M. Elise Lauterbur[1]

5 Department of Ecology and Evolution

6 Stony Brook University

7 Stony Brook, NY 11794

1 1 Address as of August 2019: Department of Ecology and Evolutionary Biology, University of Arizona,
2 Tucson, AZ 85721

8   Running title: **Coalescent models of small populations**

9   Corresponding author:

10   M. Elise Lauterbur

11   650 Life Sciences Building

12   Department of Ecology and Evolution

13   Stony Brook University

14   Stony Brook, NY 11794

15   (217) 649-2566

16   lauterbur@gmail.com

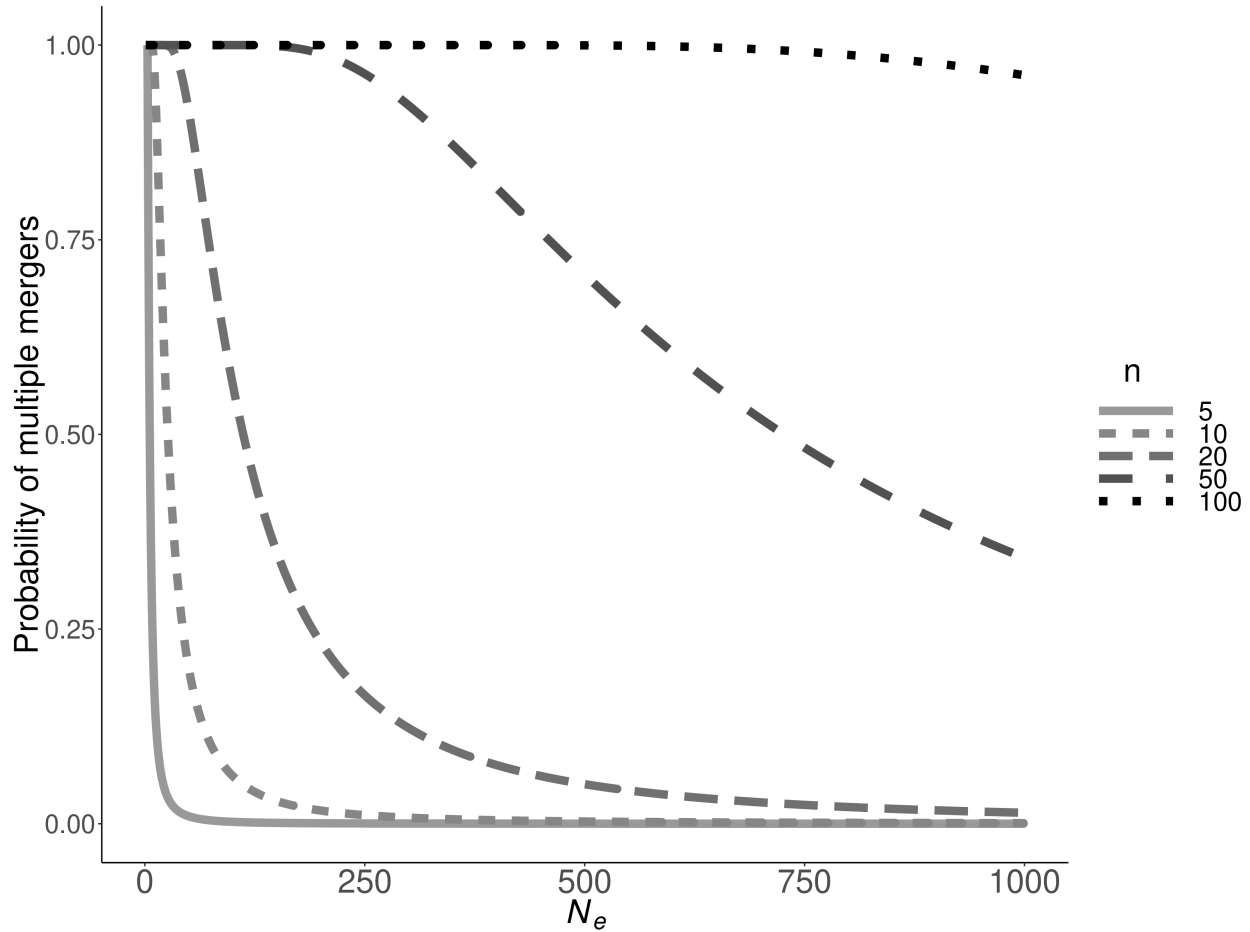17   **Keywords**: Coalescent, effective population size, conservation, bottleneck

18 **Abstract**

19 Population genetics employs two major models for conceptualizing genetic relationships

20 among individuals – outcome-driven (coalescent) and process-driven (forward). These

21 models are complementary, but the basic Kingman coalescent and its extensions make

22 fundamental assumptions to allow analytical approximations: a constant effective

23 population size much larger than the sample size. These make the probability of

24 multiple coalescent events per generation negligible. Although these assumptions are

25 often violated in species of conservation concern, conservation genetics often uses

26 coalescent models of effective population sizes and trajectories in endangered species.

27 Despite this, the effect of very small effective population sizes, and their interaction with

28 bottlenecks and sample sizes, on such analyses of genetic diversity remains

29 unexplored. Here, I use simulations to analyze the influence of small effective

30 population size, population decline, and their relationship with sample size, on

31 coalescent-based estimates of genetic diversity. Compared to forward process-based

32 estimates, coalescent models significantly overestimate genetic diversity in

33 oversampled populations with very small effective sizes. When sampled soon after a

34 decline, coalescent models overestimate genetic diversity in small populations

35 regardless of sample size. Such overestimates artificially inflate estimates of both

36 bottleneck and population split times. For conservation applications with small effective

37 population sizes, forward simulations that do not make population size assumptions are

38 computationally tractable and should be considered instead of coalescent-based

39 models. These findings underscore the importance of the theoretical basis of analytical

40 techniques as applied to conservation questions.

41 **Introduction**

42 Inferring how a population evolves or has evolved under circumstances of interest is a

43 key goal of population genetics. To understand the dynamics of real-world populations,

44 we often turn to analytically and computationally tractable coalescent models of Wright-

45 Fisher population genetics, with the standard being the Kingman coalescent (1982a,

46 1982b). These models allow the estimation and simulation of current and past

47 population genetic parameters of interest, perhaps foremost among them, genetic

48 diversity or $\theta = 4N_e\mu$ (Fu and Li 1993; Fu and Li 1999).

49 Coalescent theory in its standard form (Kingman 1982a, 1982b) both makes certain

50 simplifying assumptions that allow its analytical elegance, and remains the typical model

51 used to interpret genetic variation. These include a stable effective population size large

52 enough to be effectively infinite, that is much larger than the sample size. This is

53 because of the fundamental simplifying assumption that only two lineages coalesce per

54 generation. However when the effective population size is very small ($N_e < 1000$),

55 and/or the sample size approaches or exceeds the effective population size, the

56 probability of more than two lineages coalescing in a single generation (multiple

57 mergers) increases exponentially (figure 1) (Wakeley 2016).

6   4

58 Figure 1. Effective population size and multiple merger probability. The probability that

59 more than two lineages coalesce per generation (multiple mergers) as a function of

60 sample size (n) and effective population size from N_e = 0 to 1000.

61 In species of conservation concern, as well as widely-sampled species such as

62 humans, one or more of these assumptions (large N_e, $N_e \gg n$, stable N_e) are often

63 unmet. For example, in a sample of 138 studies of species with genetically estimated

64 effective population sizes, Palstra and Ruzzante (2008) found 120 species to have

65 estimated effective population sizes of less than 1000, and 27 to have effective

7   5

66 population sizes of less than 100. Included among these are a population North Sea

67 cod (*Gadus morhua*) off the Yorkshire coast of England, with an effective population

68 size of 70 - 120 depending on sampling period (Hutchinson et al. 2003); Brazilian water

69 hyacinth (*Eichhornia paniculata*) in northeastern Brazil, with an effective population size

70 of < 5 - 70 depending on the population sampled (Husband and Barrett 1992); and the

71 Siberian tiger (*Panthera tigris altaica*) with an effective population size of approximately

72 14 (Alasaad et al. 2011).

73 Given these small N_e, some species and populations may be oversampled relative to

74 effective population size. For example, the hihi (*Notiomysts cincta*) of New Zealand was

75 estimated to have an effective population size of 10 and a census size of 30 after

76 reintroduction to Mokoia island (Castro et al. 2004). In this case, because of banding

77 and monitoring, it was possible to sample the entire population, and thus have a sample

78 size three times that of the effective population size. Estimates of human effective

79 population sizes vary depending on population(s) and method of estimation, but are

80 generally in the range of a few thousand to ~15,000 (Tenesa et al. 2007; Mezzavilla

81 2015; The 1000 Genomes Project Consortium 2015). This means that current human

82 research involving samples of tens or even hundreds of thousands of individuals (eg.

83 Mazet et al. 2016; Martin et al. 2017) is also likely to be oversampled relative to

84 effective population size.

85 As in humans, small effective population sizes are often the result of population

86 bottlenecks. Bottlenecks are characteristic of populations of conservation concern as

8   6

87 well as the founding of new populations. Immediately after a bottleneck, genetic

88 diversity, as measured by descriptive statistics involving the number of alleles, such as

89 Watterson's theta (θ_W), drops precipitously (Maruyama and Fuerst 1985; Leberg 1992).

90 This is because of the immediate loss of rare alleles, particularly singletons. In contrast,

91 heterozygosity is reduced less immediately after a bottleneck, but continues to fall if the

92 bottleneck is prolonged (Allendorf 1986).

93 The robustness of coalescent theory to violations of theoretical assumptions has been

94 relatively well-studied with respect to sample size, migration, population size

95 fluctuations, and variation in reproductive success, among others. For example,

96 Wakeley and Takahashi (2003) and Bhaskar et al. (2014) found that large sample sizes,

97 relative to the effective population size, increase the proportion of singleton

98 polymorphisms relative to standard coalescent expectations. Bhaskar et al. (2014) were

99 able to attribute this directly to multiple mergers in oversampled genealogies. In

100 addition, variance in reproductive success can also result in multiple coalescent events

101 per generation, and increases the complexity of the relationship between effective

102 population size and the per-locus neutral mutation rate (Eldon and Wakeley 2005).

103 Migration and recolonization in metapopulations also causes skew in the site frequency

104 spectrum (Wakeley and Aliacar 2001). In these cases, deviations are attributable to

105 multiple mergers.

106 Intuition might suggest that very small effective population sizes could have the same

107 effects on parameter inference as do other violations resulting in multiple mergers, but

9   7

108   the exact nature of the consequences and how they interact with population bottlenecks

109   remains unexplored. As previously suggested, application of multiple merger coalescent

110   models could resolve erroneous estimates of population genetic parameters in

111   populations that violate assumptions of the standard coalescent (Schweinsberg 2000;

112   Mohle and Sagitov 2003; Eldon and Wakeley 2005; Birkner and Blath 2008; Tellier and

113   Lemaire 2014; Koskela, Jenkins, and Spano 2015; Montano 2016). However, if and how

114   this applies to populations with very small effective population sizes (N_e < 1000) and

115   bottlenecks has so far been neglected.

116   Here, I use simulations to examine the effect of small effective population sizes and

117   population declines, and their interaction with sample size, on coalescent estimates of

118   genetic diversity. I compare coalescent estimates of genetic diversity to genetic diversity

119   generated by forward simulations for a range of constant effective population sizes

120   between N_e = 10 and 1000, and a range of sample sizes between n = 2 and 1000. This

121   includes scenarios of oversampling: parameter spaces in which the sample size

122   approaches and even exceeds the effective population size ($n > N_e$). Because small

123   effective population sizes are often the result of more or less recent bottlenecks, I

124   extend the simulations to include bottlenecks of severity ranging from 0% to 99% and

125   sampling times (without recovery) between T = 1 and 1000 generations post-bottleneck.

126   These comparisons show that coalescent models tend to incorrectly estimate genetic

127   diversity when effective population sizes are very small and when sample size exceeds

128   effective population size. These effects are especially pronounced after a recent

129   bottleneck.

Materials and Methods

130

131 To ascertain the reliability of coalescent-based estimates of genetic diversity at low

132 effective population sizes (N_e < 1000) and sample sizes (n) approaching N_e, both with

133 and without bottleneck at varying times (T), I used forward, process-based simulations

134 to ground-truth coalescent simulations (table 1). Unlike coalescent simulations, which

135 probabilistically model the history of a sample backward in time, Wright-Fisher forward

136 simulations follow the entire population of individuals generation by generation from the

137 beginning of the simulation until it ends and individuals are sampled (eg. Hernandez

138 2008; Haller and Messer 2017, 2019).

139 Table 1. Explored parameter space of N_e and n, with approximate relationship between

140 N_e and n in the cells.

$n \mid N_e$	10	20	50	100	200	500	1000
2	$n < N_e$	$n \ll N_e$	$n \ll N_e$	$n \ll N_e$	$n \ll N_e$	$n \ll N_e$	$n \ll N_e$
5	$n < N_e$	$n < N_e$	$n \ll N_e$	$n \ll N_e$	$n \ll N_e$	$n \ll N_e$	$n \ll N_e$
10	$n = N_e$	$n < N_e$	$n < N_e$	$n \ll N_e$	$n \ll N_e$	$n \ll N_e$	$n \ll N_e$
20	$n > N_e$	$n = N_e$	$n < N_e$	$n < N_e$	$n \ll N_e$	$n \ll N_e$	$n \ll N_e$
50	$n > N_e$	$n > N_e$	$n = N_e$	$n < N_e$	$n < N_e$	$n \ll N_e$	$n \ll N_e$
100	$n \gg N_e$	$n > N_e$	$n > N_e$	$n = N_e$	$n < N_e$	$n < N_e$	$n \ll N_e$
200	$n \gg N_e$	$n \gg N_e$	$n > N_e$	$n > N_e$	$n = N_e$	$n < N_e$	$n < N_e$
500	$n \gg N_e$	$n \gg N_e$	$n \gg N_e$	$n > N_e$	$n > N_e$	$n = N_e$	$n < N_e$
1000	$n \gg N_e$	$n \gg N_e$	$n \gg N_e$	$n \gg N_e$	$n > N_e$	$n > N_e$	$n = N_e$

11

141 *Simulations*

142 Forward simulations were performed using SFS_CODE (Hernandez 2008) and SLiM v2

143 (Haller and Messer 2017), and coalescent simulations were performed with Hudson's

144 ms (Hudson 2002). I performed 10,000 simulations of each parameter combination {N_e,

145 n, T} with SFS_CODE and Hudson's ms. SLiM v2 was used to verify the SFS_CODE

146 results by performing 10,000 simulations of each parameter combination {N_e, n} to

147 mutation-drift equilibrium. Effective population size N_e could take the values {10, 20, 50,

148 100, 200, 500, 1000}, and sample size n the values {2, 5, 10, 20, 50, 100, 200, 500,

149 1000}, chosen to reflect plausible effective population sizes of populations of

150 conservation concern. For the forward simulations, when sample size exceeded

151 effective population size, individuals were resampled until the sample size was reached.

152 The non-bottleneck simulations were run with constant N_e = {10, 20, 50, 100, 200, 500,

153 1000}, and forward simulations were run with a burn-in of 10 * 2N_e.

154 For the bottleneck simulations, pre-bottleneck N_e = 1000, since preliminary simulations

155 showed no significant differences between forward and coalescent diversity estimates

156 at N_e = 1000, and because forward simulations were computationally impractical at N_e >

157 1000. The bottleneck reduced the population to one of the above set of possible N_e

158 values at T = {1, 10, 100, 1000} generations, resulting in sampling times ranging from 1

159 to 1000 generations after the bottleneck event. There was no post-bottleneck recovery

160 of N_e. For both non-bottleneck and bottleneck scenarios, a 5,000-base-pair locus was

161 simulated with mutation rate = 4N_e * 10^{-6} and recombination rate = 4N_e * 10^{-8} per bp.

162 See table 1 for {N_e, n} parameter space.

163   Since rare variants are lost first during a bottleneck, which reduces the number of

164   segregating sites, Watterson's θ (θ_W) is an appropriate and sensitive summary statistic

165   for comparing genetic diversity estimated by each type of simulation (Allendorf 1986). It

166   is also commonly used in studies of genetic diversity both within and among

167   populations. I compared θ_W estimated by each type of simulation at each parameter

168   combination {N_e, n, T}.

169   *Statistical comparisons*

170   Each set of 10,000 simulations per parameter combination {N_e, n, T} and simulation

171   type (forward, coalescent) created a distribution of estimated θ_W values. At small N_e, the

172   distributions of the number of segregating sites, thus θ_W, is zero-inflated, and parametric

173   tests are inappropriate. To compare mean θ_W among coalescent and both forward

174   simulators for each set of simulations with the same parameter combination (without a

175   bottleneck), I used Kruskal-Wallis tests followed by a post-hoc Dunn's test for each

176   significant case. For the bottleneck scenarios, I used Kolmogorov-Smirnov tests to

177   compare the distributions, and Mann-Whitney U tests to determine if the coalescent and

178   forward samples could have come from a distribution with the same mean θ_W, between

179   pairs of forward-coalescent distributions created with the same parameter combination.

180   All p-values were converted to false discovery rate (FDR) with a threshold of 0.001 to

181   control for multiple comparisons.

182   *Data Availability*

183   Supplementary tables have been uploaded to figshare, and code and data for analyses have

13   11

184   been uploaded to https://github.com/lauterbur/Ne.

14   12

185 **Results**

186 *Constant small effective population size*

187 There were few differences between genetic diversity (as measured by Watterson's θ)

188 calculated from both forward simulators, SFS_CODE and SLiM v2. The post-hoc

189 Dunn's test showed false discovery rates (FDR) under the 0.001 threshold only

190 occurred for 4 out of 63 comparisons (6%). Thus, only comparisons between

191 SFS_CODE (forward) and Hudson's ms (coalescent) are shown and used for

192 subsequent bottleneck analyses (Table S1).

193 For those scenarios in which significant differences between θ_W calculated from

194 SFS_CODE and SLiM simulations were found, these differences were 16-30% the

195 magnitude of differences between the coalescent estimate and either forward θ_W. All

196 such scenarios had sample sizes approximately double the effective population size

197 ($\{N_e = 100, n = 200\}$, $\{N_e = 200, n = 500\}$, $\{N_e = 500, n = 1000\}$). Hence, the discrepancy

198 could be an effect of the resampling of individuals performed when sample size exceeds

199 effective population size.

200 At constant effective population size, statistically significant differences between θ_W

201 estimated by coalescent models and calculated from forward models were most

202 pronounced when sample size was much different from effective population size

203 (Kruskal Wallis FDR < 0.001 and Dunn's FDR < 0.001). When $N_e \geq n$, the coalescent

204 overestimated genetic diversity, and when $N_e < n$, the coalescent underestimated

15

204   genetic diversity (figure 3).
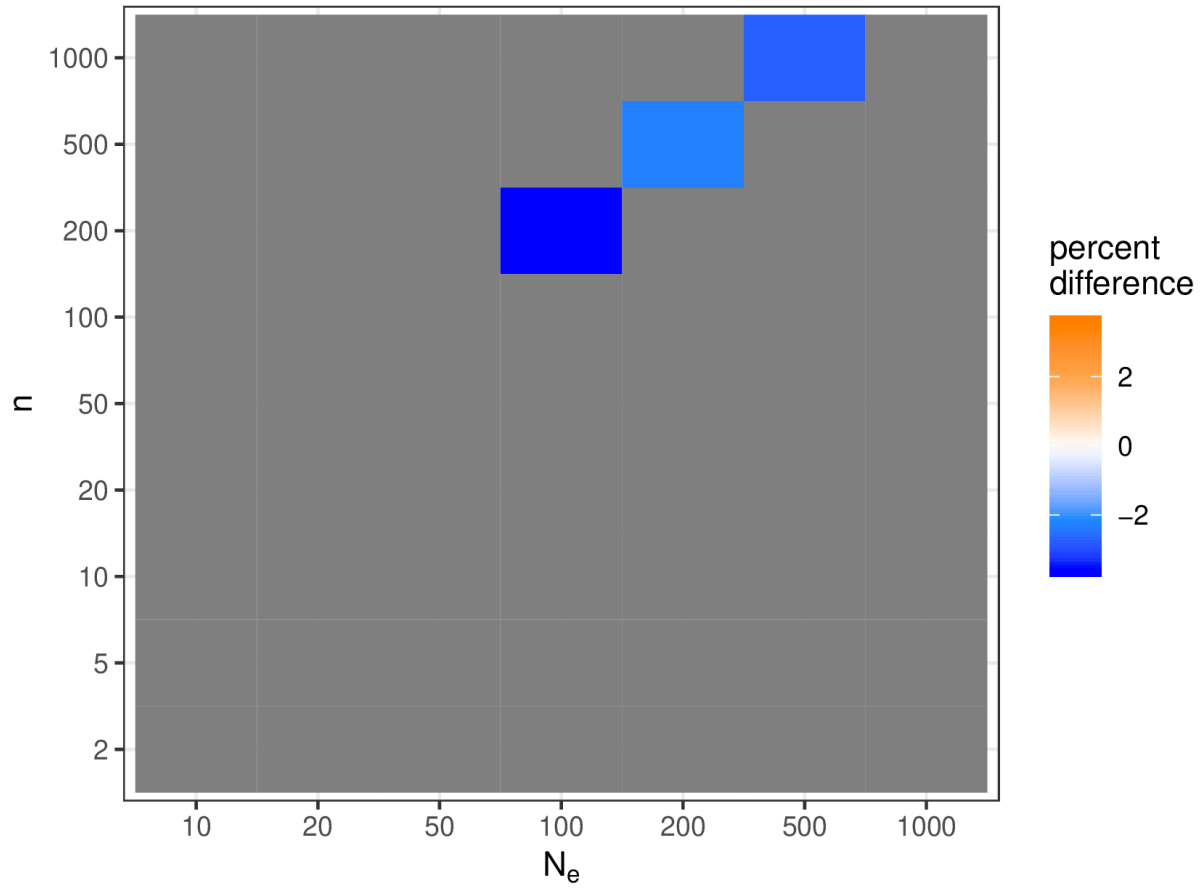
205   Figure 2. Significant differences between θ_W values for SLiM and SFS_CODE. Grey

206   indicates no significant difference, colors indicate significant differences with darker

207   oranges associated with larger mean percent differences.
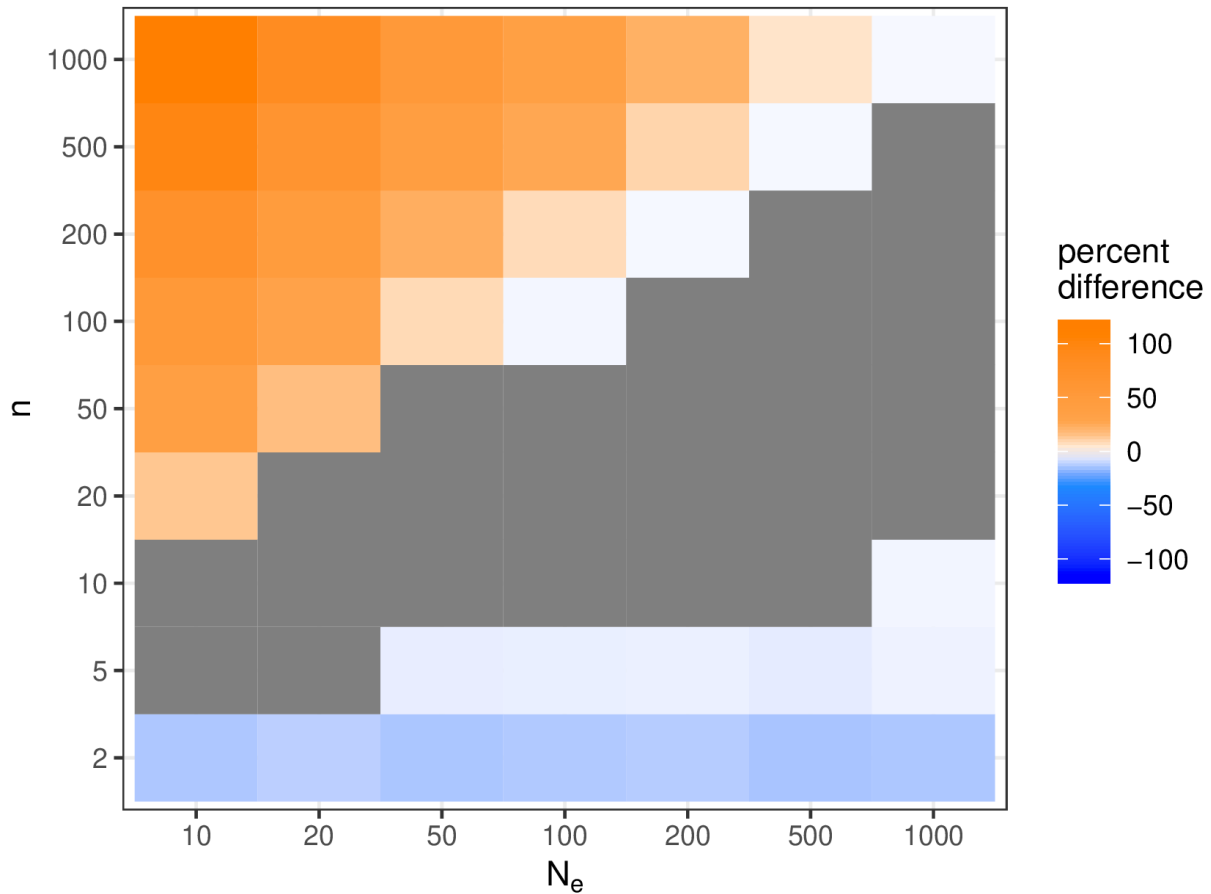
Figure 3. Significant differences between θ_W values for forward (SFS_CODE) and coalescent (ms) simulations. Grey indicates no significant difference, colors indicate significant differences with darker oranges associated with larger mean positive percent differences and darker blues associated with larger mean negative percent differences.

The relationship between sample size (n) and effective population size (N_e) influenced the result of forward to coalescent comparisons of genetic diversity. At small n/N_e ratios (≤ 0.1), coalescent simulations were more likely to underestimate θ_W by up to 13%, especially at small N_e, while at large n/N_e ratios (≥ 2.0) coalescent

216   simulations were more likely to overestimate θ_W by up to 60% (figure 4)
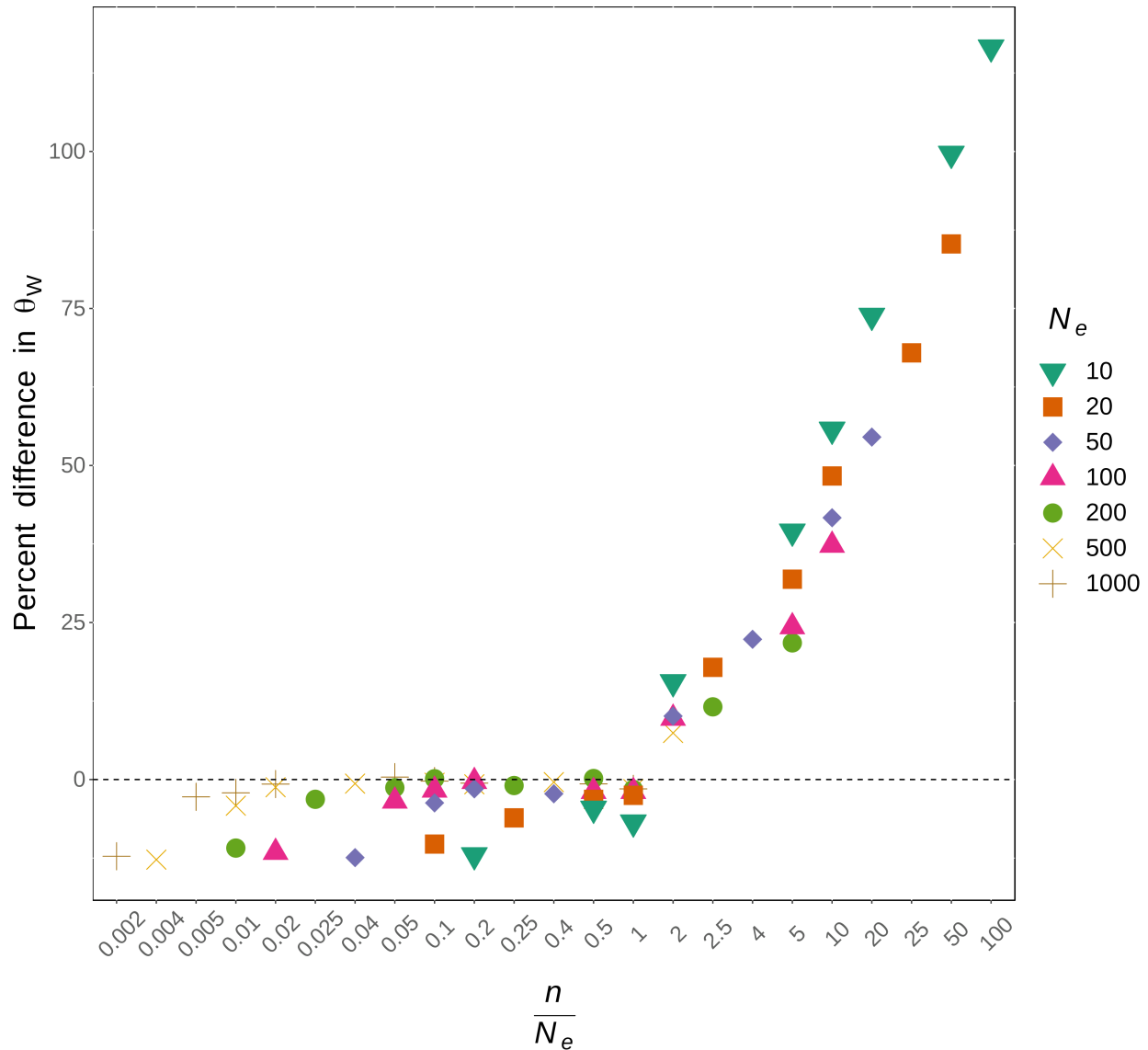

217   Figure 4. The effect of the relationship between sample size and effective population

218   size (n/N_e) on coalescent estimates of θ_W relative to forward calculations of θ_W. When

219   the percent difference is positive, the coalescent estimate is larger and thus

220   overestimating θ_W, and when the percent difference is negative, the coalescent estimate

221   is smaller and thus underestimating θ_W.

18   16

222   *Bottlenecks to small effective population size*

223   The time between sampling and a population bottleneck in small populations had a

224   greater effect on the accuracy of coalescent estimates of θ_W than effective population

225   size and its relationship with sample size in populations of constant size (Table S2).

226   Immediately after a bottleneck (T = {1, 10}), coalescent estimates of θ_W overestimate

227   genetic diversity at all effective population sizes except N_e = 1000, though not all

228   comparisons at N_e = 200 and N_e = 500 are significant. At N_e = 1000, coalescent

229   estimates underestimate genetic diversity (figure 5, figure 6). There was no effect of the

230   relationship between sample size and effective population size at T = 1. At T = 10 there

231   was an effect only at the smallest effective population sizes (N_e = {10, 20}, figure 7a, b).

232   As the time since the bottleneck increases, coalescent estimates of θ_W approach

233   forward calculations of θ_W. Despite this, there are still statistically significant differences

234   between distributions and means of coalescent and forward θ_Ws (Kruskal Wallis and

235   Mann-Whitney U tests FDR for most scenarios < 0.001, figure 5). For most scenarios at

236   T = 100 and T = 1000, coalescent estimates overestimate θ_W (figure 5, figure 8). At

237   these longer times since the bottleneck, the relationship between sample size and

238   effective population size becomes important (figure 7c, d). Since all bottlenecks were

239   from a starting N_e = 1000, these comparisons do not distinguish between percent

240   population size reduction during the bottleneck and final effective population size.
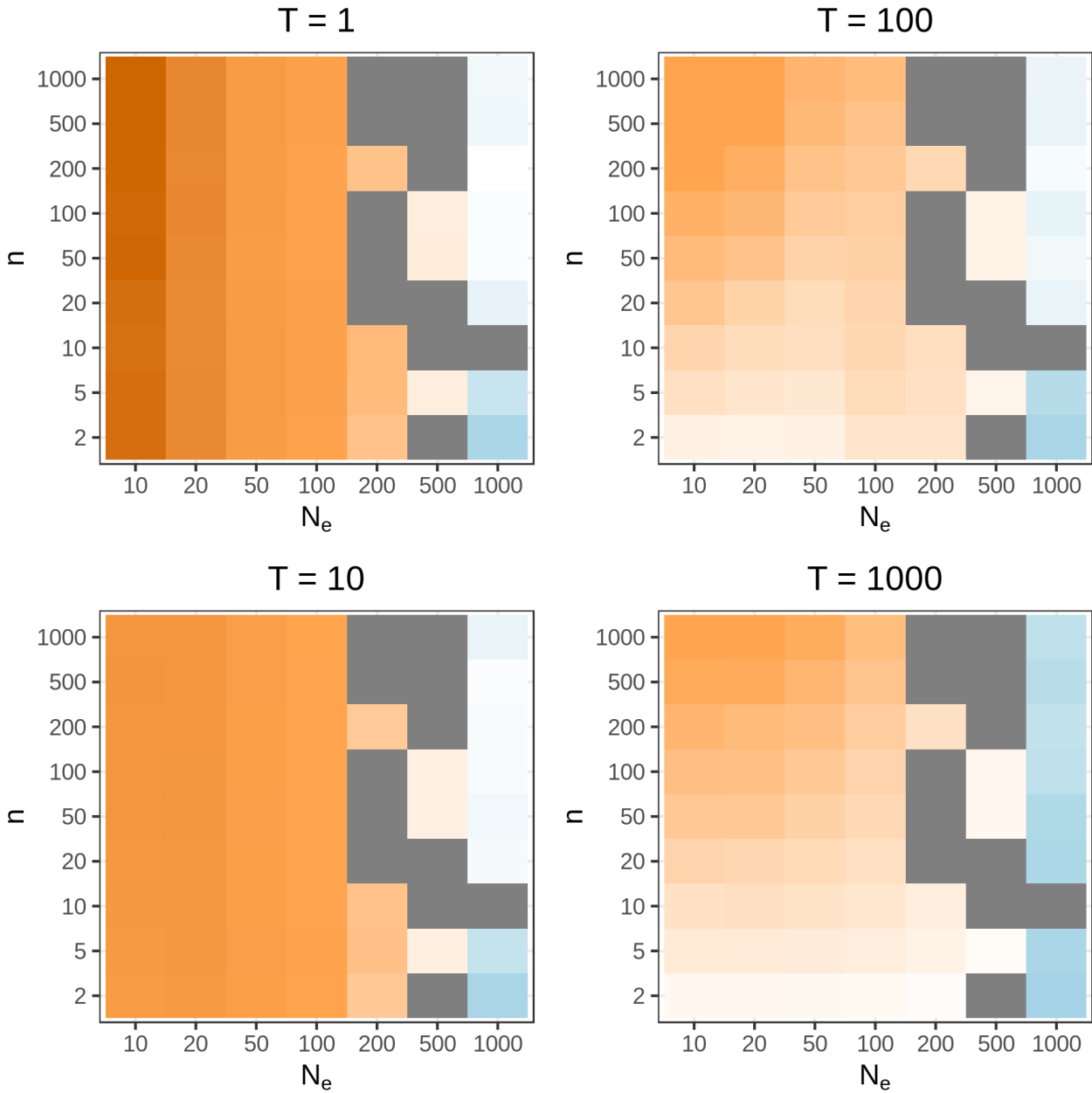
Figure 5. Significant differences between θ_W values for forward (SFS_CODE) and coalescent (ms) simulations, by time since bottleneck (T). Grey indicates no significant difference, colors indicate significant differences with darker oranges associated with larger mean positive percent differences and darker blues associated with larger mean negative percent differences.
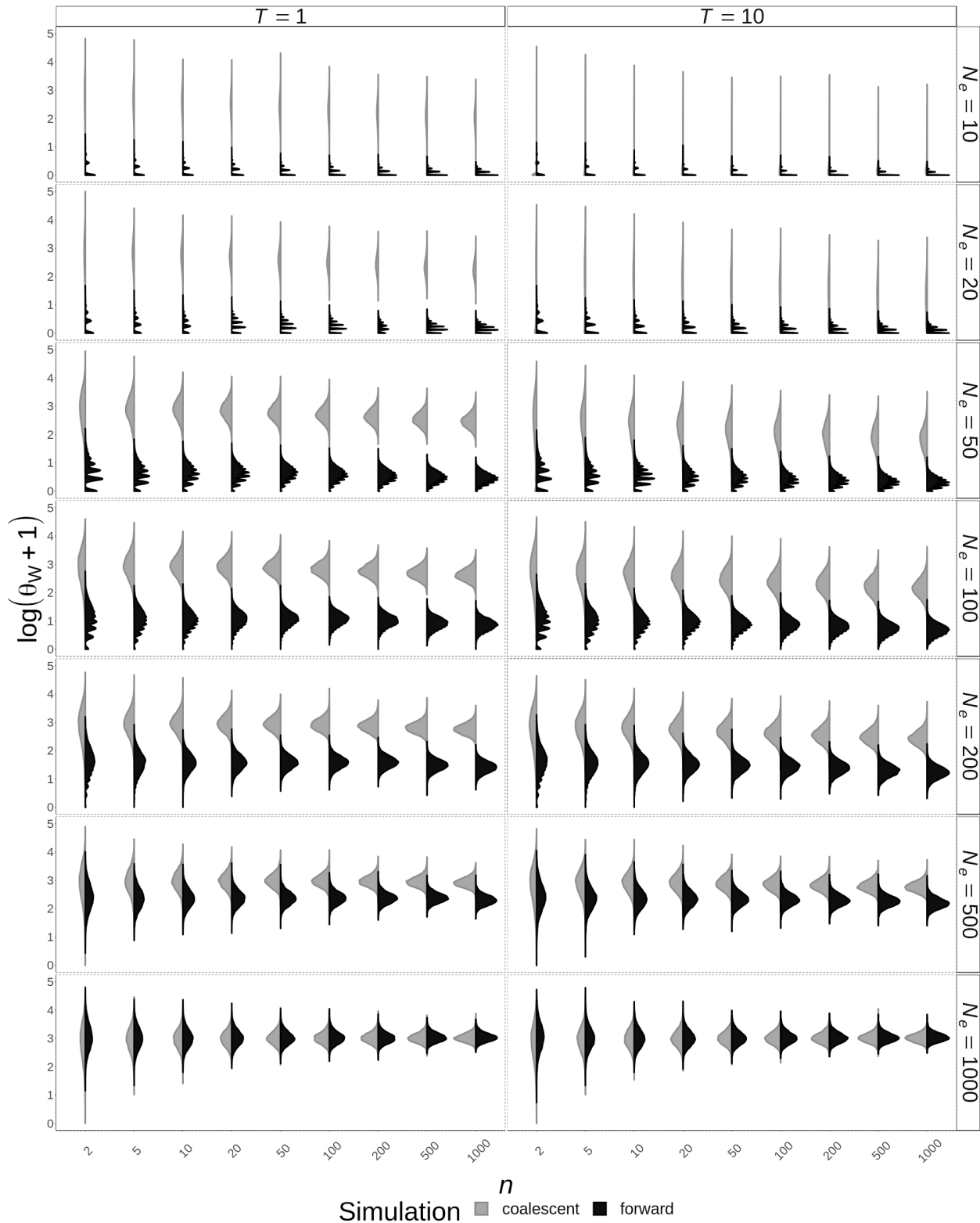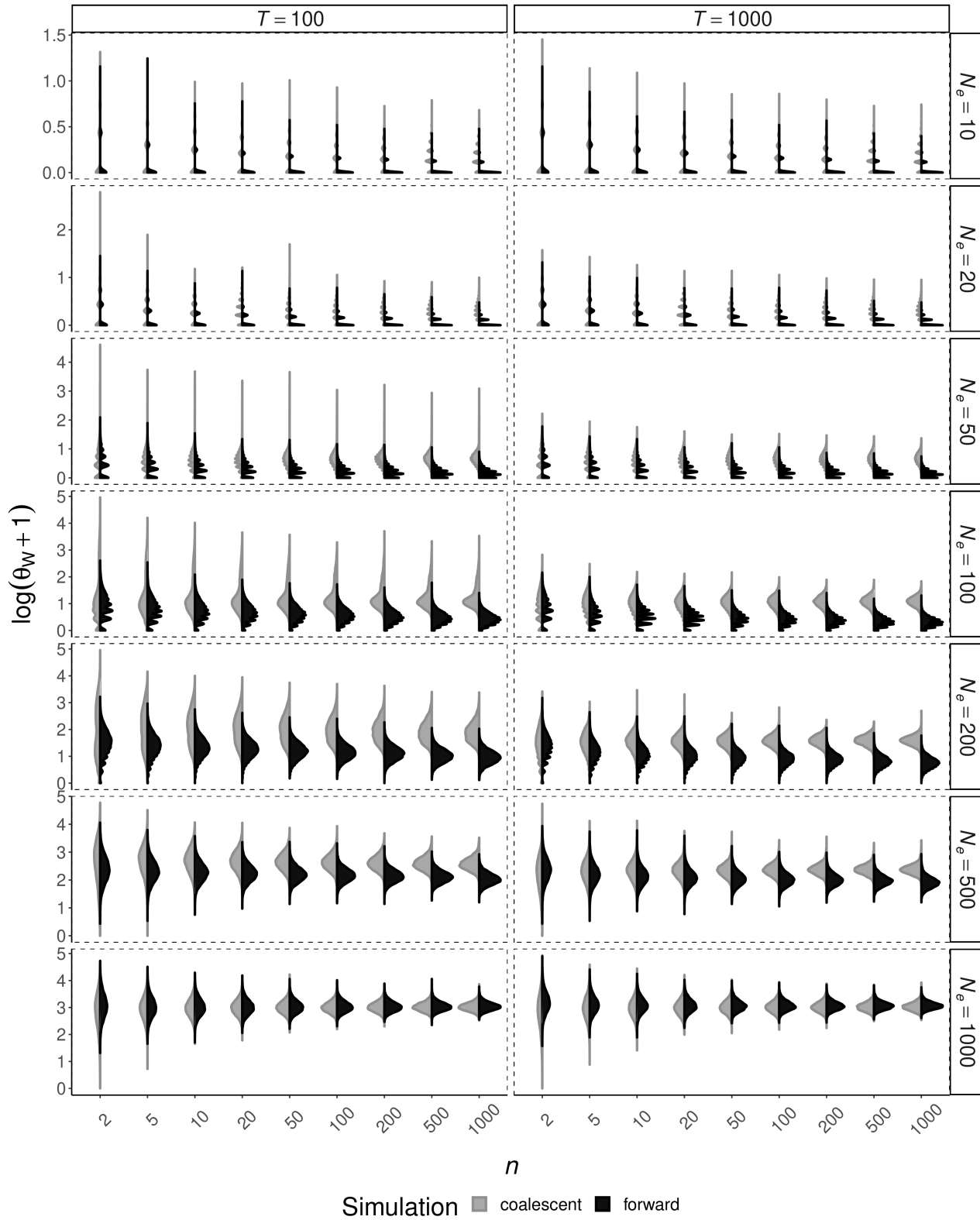
Figure 6. Distributions of θ_W (log transformed for visibility) estimated from coalescent

simulations (grey) compared to θ_W calculated from forward simulations (black), across

21   19

247   sample size (x axis) and effective population size (vertical facets) for $T = 1$ (left) and $T =$

248   10 (right).
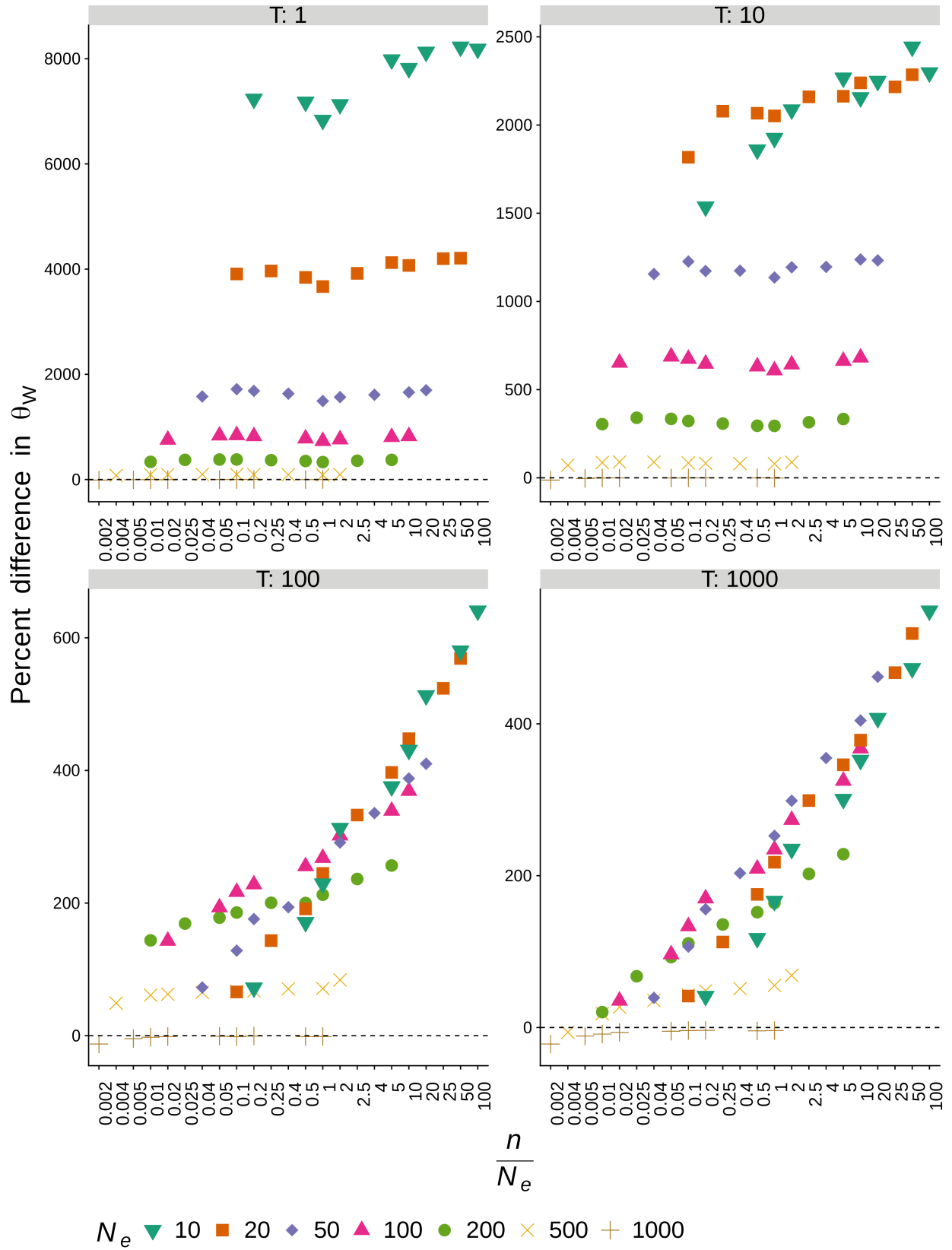
250   Figure 7. The effect of the relationship between sample size and effective population

251   size (n/N_e) on coalescent estimates of θ_W relative to forward calculations of θ_W at

252   different sampling times after a bottleneck (T = {1, 10, 100, 1000)}. When the percent

253   difference is positive, the coalescent estimate is larger and thus overestimating θ_W, and

254   when the percent difference is negative, the coalescent estimate is smaller and thus

255   underestimating θ_W. Note the scales of the y-axes of each panel are different for

256   readability.

257   Figure 8. Distributions of θ_W (log transformed for visibility) estimated from coalescent

258   simulations (grey) compared to θ_W calculated from forward simulations (black), across

259   sample size (x axis) and effective population size (vertical facets) for $T = 100$ (left) and

260    $T = 1000$ (right). Note the scales of the y-axes of each panel are different for readability.

261 **Discussion**

262 To determine if and how standard coalescent models influence estimates of genetic

263 diversity in populations with small effective population sizes, this analysis compares

264 genetic diversity estimates based on coalescent models to genetic diversity calculated

265 in forward simulation ground-truths at a range of effective population sizes, sample

266 sizes, and sampling times since a bottleneck. Coalescent models give unreliable

267 estimates of genetic diversity, as measured by Watterson's θ, regardless of the

268 relationship between sample size and effective population size. This occurs particularly

269 when the population is oversampled with respect to effective population size (sample

270 size exceeds effective population size), and when sampled soon after a bottleneck.

271 The profound differences between coalescent estimates of θ_w and forward calculations

272 of θ_w soon after a bottleneck show that applying standard coalescent models to

273 bottlenecked populations can give misleading results. Overestimates would have the

274 effect of fitting null coalescent models with incorrectly long times to the most recent

275 common ancestor to estimates of θ_w from real data, thus artificially extending estimates

276 of both bottleneck times and population split times into the past. Underestimates of θ_w

277 that occur after prolonged small effective population sizes would artificially shorten

278 estimates of bottleneck times or population split times.

279 Bottlenecks are very common, both in species of conservation concern and others,

280 including humans (Hey and Harris 1999). There have been suggestions that standard

281 coalescent models would be inappropriate for bottlenecked populations because a

27 25

282   population bottleneck does not merely shorten branches (thus necessitating the

283   rescaling of coalescent time), it increases the likelihood of multiple mergers (Wakeley

284   and Aliacar 2001; Wakeley and Takahashi 2003). However, previous analyses have

285   been restricted to metapopulations, which include migration and

286   extinction/recolonization dynamics (Wakeley and Aliacar 2001), and the results herein

287   show that coalescent models are inappropriate even for the simplest of bottlenecked

288   populations.

289   The result that coalescent models overestimate θ_W when $n > N_e$ is corroborated by

290   previous observations that when $n > N_e$ the proportion of singleton polymorphisms

291   increases relative to standard coalescent expectations (Wakeley and Takahashi 2003;

292   Bhaskar, Clark, and Song 2014). Wakeley and Takahashi attribute this excess of

293   singletons to mutations in the generation immediately previous to sampling, at a rate (θ

294   * n/N_e)/2. These mutations account for most of the singletons expected in the entire tree

295   (E[S] = θ) (Watterson 1975).

296   Multiple merger coalescent models, which are more general than the Kingman

297   coalescent, include the Ξ-, Ψ-, and Λ-coalescent (Irwin et al. 2016). These have been

298   developed since the early 2000s, and vary in how they handle multiple merger events

299   as well as the number of descendants per lineage (Árnason and Halldórsdóttir 2015).

300   Though most of the development has so far been theoretical, potential applications to

301   real-world multiple merger scenarios are clear (Schweinsberg 2000; Mohle and Sagitov

302   2003; Eldon and Wakeley 2005; Birkner and Blath 2008; Tellier and Lemaire 2014;

303   Koskela, Jenkins, and Spano 2015; Montano 2016). At least one software package,

304   Hybrid-Lambda, has been developed to allow the simulation of multiple merger trees,

305   including conversion of the resulting branch lengths to coalescent estimates (Zhu et al.

306   2015).

307   When applied to the limpet *Cellana ornata*, which has high reproductive skew, Zhu et al.

308   (2015) showed that multiple merger coalescent models using the Hybrid-Lambda

309   package estimated a population split time of approximately 9,000 generations with a

310   multiple merger model vs. approximately 48,000 generations with the standard Kingman

311   coalescent. Similarly, Árnasson and Halldórsdóttir (2015) applied Ξ- and Λ-coalescent

312   null models (Eldon et al. 2015) to another species, Atlantic cod (*Gadus morhua*), whose

313   high reproductive skew may have contributed to its small effective population size

314   (Hutchinson et al. 2003). They found that previous population divergence time estimates

315   may be too high because of inappropriate applications of the standard coalescent, and

316   that multiple merger models may provide better null hypotheses for testing both split

317   times and natural selection.

318   Despite the importance of exploring and applying multiple merger coalescent models for

319   species with small effective population sizes and those subject to bottlenecks, a more

320   expedient alternative when simulating null expectations is to use forward instead of

321   coalescent simulations. Forward simulators require more computational resources than

322   coalescent simulators, but effective population sizes of $N_e = 1000$ and even larger can

323   be manageable, even for tens of thousands of iterations. There are many forward

324   simulation options to choose from, including SFS_CODE (Hernandez 2008), SLiM

325   (Haller and Messer 2017, 2019), and fwdpp (Thornton 2014). SLiM v3 in particular

326   supports non-Wright-Fisher models, which has the potential to extend our

327   understanding of the dynamics of natural populations even further.

328   Loci from natural populations are typically sampled with ascertainment bias toward

329   polymorphic sites. This would reduce or eliminate sites without polymorphisms, which

330   account for most sites at small effective population sizes, and also accounts for some of

331   the differences in mean θ_W between coalescent and forward simulations. The resulting

332   distributions of both coalescent and forward θ_W would be biased upward, with reduced

333   right skew. This has the potential to reduce apparent differences between coalescent

334   and forward θ_W. In addition, θ_W is expected to be more sensitive to the immediate loss of

335   rare variants (and thus segregating sites) caused by a bottleneck than some other

336   commonly used summary statistics such as heterozygosity (Allendorf 1986). Like loci

337   chosen with ascertainment bias, heterozygosity might also reduce the apparent

338   differences between coalescent and forward θ_W.

339   Since natural selection has the effect of reducing effective population sizes at (and

340   near) loci under selection, these results have implications for tests for selection using

341   coalescent-based null models. A similar effect has recently been suggested of the

342   influence of sample size on tests for selection under exponential population growth

343   (Subramanian 2016). The effect of small effective population size and bottlenecks could

344   be easily tested using similar methods as were applied in this study.

345 This study shows limits in the accuracy of coalescent models with respect to small

346 effective population sizes and bottlenecked populations, including in combination with

347 oversampling. Under these population conditions, standard coalescent models have the

348 potential to produce misleading results and generate incorrect conclusions about

349 population history and structure. In applications to empirical data, it is thus imperative to

350 take into account how robust coalescent models may or may not be to violations of their

351 assumptions.

352 Thus, standard coalescent models are inappropriate for developing null expectations of

353 genetic diversity for many scenarios with small effective population sizes, especially

354 soon after a bottleneck. This is particularly relevant to studies of species of conservation

355 concern, in which small effective population sizes and bottlenecks are common

356 (Allendorf 1986; Palstra and Ruzzante 2008), but may also be relevant to other species

357 subject to oversampling or bottlenecks. This has implications for studies of species of

358 conservation concern, because coalescent models that profoundly mischaracterize the

359 timings of bottlenecks, population splits, and other demographic events could

360 detrimentally influence management actions. In addition, these mischaracterizations are

361 potentially pernicious in studies of disease population genetics, because repeated

362 bottlenecks and intense selection are classic aspects of disease agent ecology.

32

References

370   Alasaad, Samer, Ramón C. Soriguer, Galina Chelomina, Yury Petrovich Sushitsky, and

371   Joerns Fickel. 2011. "Siberian Tiger's Recent Population Bottleneck in the

372   Russian Far East Revealed by Microsatellite Markers." *Mammalian Biology* 76

373   (6): 722–26. https://doi.org/10.1016/j.mambio.2011.01.011.

374   Allendorf, Fred W. 1986. "Genetic Drift and the Loss of Alleles versus Heterozygosity."

375   *Zoo Biology* 5 (2): 181–90. https://doi.org/10.1002/zoo.1430050212.

376   Árnason, Einar, and Katrín Halldórsdóttir. 2015. "Nucleotide Variation and Balancing

377   Selection at the *Ckma* Gene in Atlantic Cod: Analysis with Multiple Merger

378   Coalescent Models." *PeerJ* 3 (February): e786. https://doi.org/10.7717/peerj.786.

379   Bhaskar, Anand, Andrew G. Clark, and Yun S. Song. 2014. "Distortion of Genealogical

380   Properties When the Sample Is Very Large." *Proceedings of the National

381   Academy of Sciences* 111 (6): 2385–90.

382   https://doi.org/10.1073/pnas.1322709111.

383   Birkner, Matthias, and Jochen Blath. 2008. "Computing Likelihoods for Coalescents with

384   Multiple Collisions in the Infinitely Many Sites Model." *Journal of Mathematical

385   Biology* 57 (3): 435–65. https://doi.org/10.1007/s00285-008-0170-6.

386   Castro, Isabel, Karen M. Mason, Doug P. Armstrong, and David M. Lambert. 2004.

387   "Effect of Extra-Pair Paternity on Effective Population Size in a Reintroduced

388   Population of the Endangered Hihi, and Potential for Behavioural Management."

389   *Conservation Genetics* 5 (3): 381–93.

390   https://doi.org/10.1023/B:COGE.0000031146.51681.b0.

391   Eldon, Bjarki, Matthias Birkner, Jochen Blath, and Fabian Freund. 2015. "Can the Site-

392 Frequency Spectrum Distinguish Exponential Population Growth from Multiple-

393 Merger Coalescents?" *Genetics* 199: 841–56.

394 Eldon, Bjarki, and John Wakeley. 2005. "Coalescent Processes When the Distribution of

395 Offspring Number Among Individuals Is Highly Skewed." *Genetics* 172 (4): 2621–

396 33. https://doi.org/10.1534/genetics.105.052175.

397 Fu, Yun-Xin, and Wen-Hsiung Li. 1993. "Maximum Likelihood Estimation of Population

398 Parameters." *Genetics* 134: 1261–70.

399 Fu, Yun-Xun, and Wen-Hsiung Li. 1999. "Coalescing into the 21st Century: An

400 Overview and Prospects of Coalescent Theory." *Theoretical Population Biology*

401 56 (1): 1–10. https://doi.org/10.1006/tpbi.1999.1421.

402 Haller, Benjamin C., and Philipp W. Messer. 2017. "SLiM 2: Flexible, Interactive

403 Forward Genetic Simulations." *Molecular Biology and Evolution* 34 (1): 230–40.

404 https://doi.org/10.1093/molbev/msw211.

405 Haller, Benjamin C, and Philipp W Messer. 2019. "SLiM 3: Forward Genetic Simulations

406 Beyond the Wright–Fisher Model." Edited by Ryan Hernandez. *Molecular Biology*

407 *and Evolution* 36 (3): 632–37. https://doi.org/10.1093/molbev/msy228.

408 Hernandez, R. D. 2008. "A Flexible Forward Simulator for Populations Subject to

409 Selection and Demography." *Bioinformatics* 24 (23): 2786–87.

410 https://doi.org/10.1093/bioinformatics/btn522.

411 Hey, J., and E. Harris. 1999. "Population Bottlenecks and Patterns of Human

412 Polymorphism." *Molecular Biology and Evolution* 16 (10): 1423–26.

413 https://doi.org/10.1093/oxfordjournals.molbev.a026054.

414 Hudson, R. R. 2002. "Generating Samples under a Wright-Fisher Neutral Model of

415   Genetic Variation." *Bioinformatics* 18 (2): 337–38.

416   https://doi.org/10.1093/bioinformatics/18.2.337.

417 Husband, Brian C., and Spencer C. H. Barrett. 1992. "Effective Population Size and

418   Genetic Drift in Tristylous *Eichhornia Paniculata* (Pontederiaceae)." *Evolution* 46

419   (6): 1875–90. https://doi.org/10.1111/j.1558-5646.1992.tb01175.x.

420 Hutchinson, William F., Cock van Oosterhout, Stuart I. Rogers, and Gary R. Carvalho.

421   2003. "Temporal Analysis of Archived Samples Indicates Marked Genetic

422   Changes in Declining North Sea Cod (*Gadus Morhua*)." *Proceedings of the*

423   *Royal Society of London. Series B: Biological Sciences* 270 (1529): 2125–32.

424   https://doi.org/10.1098/rspb.2003.2493.

425 Irwin, K K, S Laurent, S Matuszewski, S Vuilleumier, L Ormond, H Shim, C Bank, and J

426   D Jensen. 2016. "On the Importance of Skewed Offspring Distributions and

427   Background Selection in Virus Population Genetics." *Heredity* 117 (6): 393–99.

428   https://doi.org/10.1038/hdy.2016.58.

429 Kingman, J. F. C. 1982a. "On the Genealogy of Large Populations." *Journal of Applied*

430   *Probability* 19: 27. https://doi.org/10.2307/3213548.

431 ———. 1982b. "The Coalescent." *Stochastic Processes and Their Applications* 13:

432   235–48. https://doi.org/10.1016/0304-4149(82)90011-4.

433 Koskela, Jere, Paul Jenkins, and Dario Spano. 2015. "Computational Inference beyond

434   Kingman's Coalescent." *Journal of Applied Probability* 52 (2): 519–37.

435 Leberg, Paul L. 1992. "Effects of Population Bottlenecks on Genetic Diversity as

436   Measured by Allozyme Electrophoresis." *Evolution* 46 (2): 477–94. https://doi.org/

437   10.1111/j.1558-5646.1992.tb02053.x.

438   Martin, Alicia R., Christopher R. Gignoux, Raymond K. Walters, Genevieve L. Wojcik,

439   Benjamin M. Neale, Simon Gravel, Mark J. Daly, Carlos D. Bustamante, and

440   Eimear E. Kenny. 2017. "Human Demographic History Impacts Genetic Risk

441   Prediction across Diverse Populations." *The American Journal of Human*

442   *Genetics* 100 (4): 635–49. https://doi.org/10.1016/j.ajhg.2017.03.004.

443   Maruyama, Takeo, and Paul A. Fuerst. 1985. "Population Bottlenecks and

444   Nonequilibrium Models in Population Genetics. II. Number of Alleles in a Small

445   Population That Was Formed by a Recent Bottleneck." *Genetics* 111: 675–89.

446   Mazet, O, W Rodríguez, S Grusea, S Boitard, and L Chikhi. 2016. "On the Importance

447   of Being Structured: Instantaneous Coalescence Rates and Human Evolution—

448   Lessons for Ancestral Population Size Inference?" *Heredity* 116 (4): 362–71.

449   https://doi.org/10.1038/hdy.2015.104.

450   Mezzavilla, Massimo. 2015. "Neon: An R Package to Estimate Human Effective

451   Population Size and Divergence Time from Patterns of Linkage Disequilibrium

452   between SNPS." *Journal of Computer Science & Systems Biology* 8 (1).

453   https://doi.org/10.4172/jcsb.1000168.

454   Mohle, Martin, and Serik Sagitov. 2003. "Coalescent Patterns in Diploid Exchangeable

455   Population Models." *Journal of Mathematical Biology* 47 (4): 337–52.

456   https://doi.org/10.1007/s00285-003-0218-6.

457   Montano, Valeria. 2016. "Coalescent Inferences in Conservation Genetics: Should the

458   Exception Become the Rule?" *Biology Letters* 12 (6): 20160211.

459   https://doi.org/10.1098/rsbl.2016.0211.

460   Palstra, Friso P., and Daniel E. Ruzzante. 2008. "Genetic Estimates of Contemporary

461   Effective Population Size: What Can They Tell Us about the Importance of

462   Genetic Stochasticity for Wild Population Persistence?" *Molecular Ecology* 17

463   (15): 3428–47. https://doi.org/10.1111/j.1365-294X.2008.03842.x.

464 Schweinsberg, Jason. 2000. "Coalescents with Simultaneous Multiple Collisions."

465   *Electronic Journal of Probability* 5 (0). https://doi.org/10.1214/EJP.v5-68.

466 Subramanian, Sankar. 2016. "The Effects of Sample Size on Population Genomic

467   Analyses – Implications for the Tests of Neutrality." *BMC Genomics* 17 (1).

468   https://doi.org/10.1186/s12864-016-2441-8.

469 Tellier, Aurélien, and Christophe Lemaire. 2014. "Coalescence 2.0: A Multiple

470   Branching of Recent Theoretical Developments and Their Applications."

471   *Molecular Ecology* 23 (11): 2637–52. https://doi.org/10.1111/mec.12755.

472 Tenesa, A., P. Navarro, B. J. Hayes, D. L. Duffy, G. M. Clarke, M. E. Goddard, and P.

473   M. Visscher. 2007. "Recent Human Effective Population Size Estimated from

474   Linkage Disequilibrium." *Genome Research* 17 (4): 520–26.

475   https://doi.org/10.1101/gr.6023607.

476 The 1000 Genomes Project Consortium, Richard A. Gibbs, Eric Boerwinkle, Harsha

477   Doddapaneni, Yi Han, Viktoriya Korchina, Christie Kovar, et al. 2015. "A Global

478   Reference for Human Genetic Variation." *Nature* 526 (7571): 68–74.

479   https://doi.org/10.1038/nature15393.

480 Thornton, Kevin R. 2014. "A C++ Template Library for Efficient Forward-Time

481   Population Genetic Simulation of Large Populations." *Genetics* 198 (1): 157–66.

482   https://doi.org/10.1534/genetics.114.165019.

483 Wakeley, John. 2016. *Coalescent Theory: An Introduction*. Macmillan Learning.

484   Wakeley, John, and Nicolas Aliacar. 2001. "Gene Genealogies in a Metapopulation."

485   *Genetics* 159: 893–905.

486   Wakeley, John, and Tsoyoshi Takahashi. 2003. "Gene Genealogies When the Sample

487   Size Exceeds the Effective Size of the Population." *Molecular Biology and*

488   *Evolution* 20 (2): 208–13. https://doi.org/10.1093/molbev/msg024.

489   Watterson, G.A. 1975. "On the Number of Segregating Sites in Genetical Models

490   without Recombination." *Theoretical Population Biology* 7 (2): 256–76.

491   https://doi.org/10.1016/0040-5809(75)90020-9.

492   Zhu, Sha, James H. Degnan, Sharyn J. Goldstien, and Bjarki Eldon. 2015. "Hybrid-

493   Lambda: Simulation of Multiple Merger and Kingman Gene Genealogies in

494   Species Networks and Species Trees." *BMC Bioinformatics* 16 (1).

495   https://doi.org/10.1186/s12859-015-0721-y.