MOLECULAR ECOLOGY

Identifying signatures of sexual selection using genome-wide selection components analysis

Journal:	*Molecular Ecology*
Manuscript ID:	MEC-15-0027
Manuscript Type:	Original Article
Date Submitted by the Author:	09-Jan-2015
Complete List of Authors:	Flanagan, Sarah; Texas A&M University, Biology Jones, Adam; Texas A&M University, Biology
Keywords:	simulation model, RAD-seq, population genomics, Adaptation, reproductive success, population genetics

Identifying signatures of sexual selection using genome-wide selection components analysis

Sarah P. Flanagan[1], Adam G. Jones[1]

[1]Biology Department, Texas A&M University, 3258 TAMU, College Station, TX 77843, USA

Key words: simulation model, RAD-seq, population genomics, adaptation, reproductive success, population genetics

Corresponding Author: Sarah P. Flanagan, sflanagan@bio.tamu.edu, fax: 979-845-2891

Running title: Genome-wide selection components analysis

Abstract

Sexual selection must affect the genome for it to have an evolutionary impact, yet signatures of selection remain elusive. Here we use an individual-based model to investigate the potential of genome-wide selection components analysis, which compares allele frequencies of individuals at different life history stages within a single population to detect selection without requiring *a priori* knowledge of traits under selection. We modeled a diploid, sexually reproducing population and introduced strong mate choice to simulate sexual selection. Genome-wide allele frequencies in adults and offspring were compared using weighted F_{ST} values. The average number of outlier peaks (i.e., those with significantly large F_{ST} values) with a quantitative trait locus in close proximity ('real' peaks) represented correct diagnoses of loci under selection, whereas peaks above the F_{ST} significance threshold without a quantitative trait locus reflected spurious peaks. We found that even with moderate sample sizes signatures of strong sexual selection were detectable, but larger sample sizes improved detection rates. The model was better able to detect selection when the genome size was large and distributed across multiple chromosomes. Although environmental variation decreased detection rates, the identification of real peaks nevertheless remained feasible. We also found that detection rates can be improved by sampling multiple populations experiencing similar selection regimes. In short, genome-wide selection components analysis is a challenging but feasible approach for the identification of regions of the genome under selection.

Introduction

One of the most important questions in evolutionary biology is how selection, which by definition acts on phenotypes, causes heritable changes (Nielsen 2005). Recent advances in DNA sequencing technologies have provided many new opportunities to explore how genomes are affected by selection, but no method currently exists to detect the signature of individual episodes of selection within the timeframe of a single generation on a genome-wide scale. Yet, we know that total selection can be decomposed into several components of selection that affect individuals at various stages during the life cycle (Christiansen and Frydenberg 1973) and that these episodes can provide important insights into mating systems (Emlen and Oring 1977), ecological factors (Loehle and Pechmann 1988), and can even help evaluate threats and conservation issues (Stockwell et al. 2003). Additionally, much of the quantitative genetics theory commonly used in empirical studies focuses on individual episodes of selection (Arnold and Wade 1984a; Arnold and Wade 1984b). Therefore, a method to detect the signature of each component of selection within a natural population would be an important addition to an evolutionary biologist's toolkit.

Currently, three principal analytical methods are used to diagnose the effects of selection on the genome. First, quantitative traits can be mapped to specific loci using linkage mapping techniques. Quantitative trait locus mapping is very effective, but requires crossing specific parents, generating numerous offspring, and having a trait of interest to map. Second, genome-wide association studies can be used to correlate a specific trait (often disease-related) with loci that differ between groups with different values of the trait (e.g. a group with diabetes compared to a group without; reviewed in Carlson et al. 2004). Finally, population genomics methods compare summary statistics describing allele frequencies, genetic diversity, and linkage

disequilibrium between multiple populations of the same species to identify loci that lie outside of a specified significance threshold (Hohenlohe et al. 2010b). This method can be very powerful at detecting signatures of positive selection (e.g. Gagnaire et al. 2013; Hess et al. 2013), balancing selection (e.g. Reitzel et al. 2013), local adaptation (e.g. Hohenlohe et al. 2010a; Miller et al. 2012; Catchen et al. 2013a; Vincent et al. 2013), and selective sweeps (e.g. Boitard and Rocha 2013; Clement et al. 2013; Harris et al. 2013; Hubner et al. 2013; Rellstab et al. 2013), but does not facilitate a diagnosis of the type of selection (e.g. sexual selection, viability selection) causing the pattern.

A complementary approach, which has not yet been applied on a whole-genome scale, is to measure the effects of selection at various stages in the life cycle of a population. At least four major types of selection occur during a typical life cycle. These different components can be isolated within a single generation using a cross-sectional study design (e.g. Christiansen et al. 1973; Christiansen and Frydenberg 1973), or by tracking a population over multiple generations in a longitudinal design (Bundgaard and Christiansen 1972; Clark and Feldman 1981; Clark et al. 1981; Anderson et al. 2014). Although a longitudinal design allows researchers to track allele frequencies over multiple generations, it is not a feasible experimental design for many organisms and is difficult to implement in studies of natural populations.

Selection components analysis can be used to decompose total selection into its parts in a variety of ways. For instance, some researchers have compared pre-observation and post-observation components of selection (Prout 1965, 1969, 1971a, b), while others have examined mother-offspring combinations, allowing a subset of the male breeding population to be inferred and compared to a random sample of adult males containing both mated and unmated individuals (e.g. Christiansen and Frydenberg 1973; Nadeau and Baccus 1981). Allele frequencies of

individuals at different life history stages were commonly compared in studies using allozyme markers (e.g. Yamazaki 1971; Christiansen et al. 1973; Christiansen and Frydenberg 1973, 1974; Christiansen and Frydenberg 1976; Allard et al. 1977; Baccus et al. 1977; Clegg et al. 1978; Baccus et al. 1980; Christiansen 1980; Curtsinger and Feldman 1980; Nadeau and Baccus 1981; Østergaard and Christiansen 1981; Nadeau and Baccus 1983; Siegismund and Christiansen 1985; Heath et al. 1988; McDonald 1989), but these studies usually did not target enough loci to detect selection. Selection components analysis was also used to investigate patterns of selection on entire chromosomes (e.g. Anderson 1969; Prout 1971a; Bundgaard and Christiansen 1972; Anderson et al. 1979; Clark and Feldman 1981; Clark et al. 1981; Curtsinger 1984; Barbadilla et al. 1994), but chromosomes were too broad of a target and so only crude estimates of selection were detectable. However, with next-generation sequencing approaches it is now possible to identify large numbers of single nucleotide polymorphisms distributed across the entire genome, opening up the possibility to detect a genome-wide signature of selection components.

The field of population genomics has room for additional approaches aimed at detecting signatures of selection, and selection components analysis may be one solution. In this paper, we present findings from an individual-based simulation model that tests selection components analysis in the context of thousands of loci. We show that this approach holds great promise for detecting genome-wide signatures of strong selection, without needing to target a trait of interest, but that very large studies might be required to identify loci of relatively small effect. Additionally, this model allows us to make predictions about characteristics of populations that might benefit most from a selection components analysis approach.

Methods

Model overview

We modeled a population with a carrying capacity of N individuals, each of which had c chromosomes with m markers (i.e. single-nucleotide polymorphisms) and q quantitative trait loci. The quantitative trait loci additively determined, sometimes with added environmental effects, the phenotype of each individual. In a life cycle, individuals produced gametes, mated, and produced offspring. Females chose mates based on the encountered males' phenotypes, putting sexual selection pressure on the male phenotype only. The male offspring then underwent viability selection on the same trait females used to choose a mate. Finally, the offspring matured into adults and replaced the previous generation.

The life cycle was repeated for a given number of initial generations followed by five experimental generations, during which the population was randomly sampled and summary statistics calculated. Allele frequencies were also compared between adults and offspring (see below, "Sampling the population", for more detail) using weighted F_{ST} values.

Specifics of the model: genetics of the population

The simulated organism was assumed to be diploid, so each individual had two copies of each chromosome. Marker loci were evenly distributed along chromosomes, and each quantitative trait locus was randomly placed in close proximity to one marker. For instance, under the basic parameter combinations, each of the 4 chromosomes had 1000 marker loci and two quantitative trait loci, so that the total genome size was 4000 marker loci and eight quantitative trait loci. The alleles for the quantitative trait loci were drawn from a normal distribution with a mean of zero and a standard deviation of 0.5. Each locus could have up to four alleles, and we started each simulation run with only four chromosome-wide genotypes for each chromosome. In other words, each run started with complete linkage disequilibrium within

particular chromosomes. Linkage disequilibrium then deteriorated during the initial generations due to recombination, which occurred during the production of offspring in the form of crossing-over events at a constant rate per chromosome, $r = 0.2$. Each recombination event was randomly assigned a location between two marker loci. This approach allows the genome-wide level of linkage disequilibrium to be altered by merely changing the number of initial generations. No mutation occurred in the production of gametes, because the simulation runs consisted of so few generations that mutation would not be a major factor affecting allele frequencies. Phenotypes were calculated by summing across all alleles at all quantitative trait loci plus an added value, *e*, a random number from a normal distribution with a mean of 0 and a specified environmental standard deviation.

We tested different numbers of initial generations to see their effect on linkage disequilibrium and on the prospects for reliably detecting quantitative trait loci. For this analysis, we calculated pairwise linkage disequilibrium between 100 randomly chosen loci on each chromosome (i.e. all comparisons were from loci on the same chromosome). The pairwise linkage disequilibrium between randomly chosen locus A and locus B was then calculated as follows. For each allele A_i and B_j, D was calculated as $f_{ij} - p_i q_j$, where f_{ij} is the frequency of alleles i and j occurring together in the population, p_i is the frequency of allele i, and q_j is the frequency of allele j. $D_{\max}$ is the lesser of $p_i q_j$ or $1 - p_i q_j$ when $D_{ij} < 0$ and is the minimum of $(1 - p_i)*q_i$ or $p_i*(1 - q_i)$ when $D > 0$. Finally, D' was evaluated as:

$$D' = \sum_i^m \sum_j^n \frac{p_i q_j |D_{ij}|}{D_{max_{ij}}},$$

where m is the number of alleles at locus A and n is the number of alleles at locus B.

Specifics of the model: mating, production of gametes, and selection

In this model, each female mated with at most one male, and males were capable of mating with multiple females. Females randomly sampled 50 males in the population, and if they could not find an acceptable mate within those 50, they did not mate. In the initial generations, females mated with the first male they encountered, so no sexual selection occurred. Sexual selection was introduced to the model during the experimental generations, after trait values were standardized to a mean of 0 and a standard deviation of 1. When mate choice was implemented, the probability that the female would mate with a given male was determined by a Gaussian-shaped function comparing the male's phenotype, z, to a population-level female preference value, θ,

$$P(z) = e^{\frac{-(z-\theta)^2}{2\omega_S^2}},$$

where ω_S^2 is the width of the selection surface (i.e., it determines the strength of selection). We set θ to an arbitrary value of 4 for all runs of the model. If a random number from a normal distribution was less than $P(z)$, the female mated with that male and they produced four offspring.

When females found a mate and produced offspring, we simulated meiosis in the following way. For each chromosome pair, one of the mother's two chromosomes was randomly chosen to combine with a randomly chosen one of the father's two chromosomes. When these two chromosomes were combined in the zygote, recombination occurred at a rate of 0.2 recombination events per chromosome. If recombination occurred in a given mating event on a given chromosome, a randomly chosen chunk of the maternally-derived chromosome was exchanged with the matching region from the paternally-derived chromosome, while maintaining the total size of each chromosome.

After the zygotes were produced, viability selection acted on the male offspring. This selection was implemented as the following Gaussian fitness surface with a given width, ω_V^2:

$$W(z) = e^{\frac{-(z-\theta)^2}{2\omega_V^2}}$$

where z is an individual's phenotype and θ is the optimum value (zero). Viability selection was implemented during both the initial generations and the experimental generations, although the strength of viability selection was weak ($\omega_{VI}^2 = \omega_{VE}^2 = 500$). If an individual was affected by viability selection, that individual did not survive to the next generation. After the viability selection event, offspring were randomly chosen to survive to adulthood, so that the number of surviving offspring was less than or equal to the carrying capacity.

Sampling the population

Population demographic statistics were calculated for the entire population each generation. Some of the statistics calculated were population size, sex ratios, and mean trait values for males and females. Additive genetic and phenotypic variances were calculated from the distribution of values in all adults. Heritability was calculated as the additive genetic variance divided by the phenotypic variance. 'Long distance' linkage disequilibrium was calculated for randomly selected loci throughout the genome as described above (see "Specifics of the model: genetics of the population"), and the same equations were used to calculate pairwise linkage disequilibrium between neighboring polymorphic loci. Mating differentials were calculated as the covariance between standardized trait values and relative mating success (Jones 2009).

During the experimental generations, the population was randomly sampled after mating occurred and offspring were produced, but before the offspring experienced viability selection. Since females all produced the same number of offspring (4) and meiotic drive was not included

in the model, this sampling strategy captured the effects of sexual selection on allele frequencies. Both parents and offspring were sampled without replacement, and genealogical relationships were assumed to be unknown. Summary statistics, including allele frequencies and observed and expected heterozygosities, were calculated for adults, offspring, and the total population. Expected heterozygosity, H_E, was calculated as $1 - \sum_{i=1}^{a} p_i^2$ for each locus with a alleles. We then compared allele frequencies in adults and offspring using the Weir and Cockerham (1984) F_{ST} calculation:

$$F_{ST} = \left(1 - \frac{H_{E_{progeny}} + H_{E_{adults}}}{2 * H_{E_{total\ population}}}\right).$$

F_{ST} values were weighted using a kernel-smoothing moving average, which incorporates the contribution of nearby values to the F_{ST} for each locus. Specifically, each variable locus, k, was weighted by the F_{ST} values at each nucleotide position, d, within the sliding window region in each direction, using the Gaussian function:

$$F_{ST}{}' = \frac{\sum_{k-\sigma_s}^{k+\sigma_s} F_{ST} * e^{\frac{-(d-k)^2}{2\sigma_s^2}}}{\sum_{k-\sigma_s}^{k+\sigma_s} e^{\frac{-(d-k)^2}{2\sigma_s^2}}}$$

where σ_s is the width of the sliding window region (Hohenlohe et al. 2010).

To determine significance of our summary statistics, we implemented three methods of determining cut-offs. First, we calculated p-values using χ^2 distributions, and then calculated the Benjamini and Hochberg (1995) false discovery rate to establish a cut-off value. The false discovery rate ranks all of the p-values from smallest to largest. For each p-value, its relative rank (its order in the sorted list of p-values divided by the total number of p-values) was multiplied by the significance value, 0.05. The largest p-value that was less than or equal to this

weighted rank was the false discovery rate significance threshold. Second, we implemented the bootstrapping algorithm used by the software package STACKS (Catchen et al. 2011; Catchen et al. 2013b), a common population genomics bioinformatics program, re-sampling the genome 10,000 times. This algorithm re-weights the weighted F_{ST}' values using the kernel smoothing approach described above, but the nucleotide positions (d) are randomly chosen loci from anywhere in the genome, rather than the neighboring nucleotides. Confidence intervals were then calculated from the distribution of the 10,000 bootstrapped F_{ST} values in the same way as described below for the genome-wide confidence intervals. Finally, we determined the genome-wide distribution of F_{ST}' values and calculated confidence intervals. To do this, the mean F_{ST}' and the variance and standard deviation in F_{ST}' values were calculated across all sampled loci on all chromosomes. Genome-wide confidence intervals were then calculated as the mean F_{ST}' value plus the appropriate value from the cumulative normal distribution function multiplied by the standard deviation of F_{ST}' values. For example, the 95% genome-wide confidence interval is calculated as follows:

$$95\%\ \mathrm{CI} = \overline{F_{ST}'} \pm 1.95996 * \sigma_{F_{ST}'}.$$

After determining these various cut-off values, each peak in F_{ST} value was detected and the value compared to the cut-offs. If the peak was above the cut-off value, and if a known quantitative trait locus was within 50 marker loci of the peak, then it was counted as a 'real' peak. If the peak was above the cut-off value but a quantitative trait locus was not within 50 marker loci, then it was counted as a 'spurious' peak. The average and standard errors of both the number of real and spurious peaks were calculated. We also calculated the average ratio of real peaks relative the total number of quantitative trait loci underlying the trait, which allowed us to assess how often real peaks were not detected. In other words, the proportion of real peaks

220 detected reflected the amount of type II error in the analysis. In contrast, the average number of
221 spurious peaks indicated the extent to which type I error occurred (i.e. the frequency of false
222 positives).

223 *Testing parameter combinations*

224 The starting parameters from the model are in Table 1. To address whether genome-wide
225 selection components analysis could be used in empirical studies of natural populations, we
226 focused on the effects of sample size, strength of sexual selection, and environmental variance in
227 the focal trait. We also assessed the effects of genome size (number of marker loci and number
228 of chromosomes), population size, and the number of quantitative trait loci underlying the trait
229 on this type of selection-components analysis. We tested many of these parameters in
230 combination. The pairwise parameter combinations tested included sample size and carrying
231 capacity; the number of markers and the number of chromosomes; the strength of sexual
232 selection and number of quantitative trait loci; and environmental variation and the number of
233 quantitative trait loci.

234 **Results**

235 *Linkage disequilibrium*

236 Linkage disequilibrium when sexual selection was introduced and when sampling
237 occurred was controlled by changing the number of initial generations. The number of initial
238 generations tested varied from 1 to 1000 to examine how linkage disequilibrium affected our
239 ability to detect selection. We found that after 200 initial generations linkage disequilibrium was
240 0.1007 and that detection rates appeared to peak at this level of linkage disequilibrium (Fig. 1).
241 We chose to use 200 generations in the rest of our permutations of the model.

242 *Determining significance*

Our three methods of choosing cut-off values for determining whether a peak was significant showed strikingly different patterns. The false discovery rate was highly unpredictable, such that in some cases nearly every locus was significant, and at other times nearly none of the loci were significant, when the parameters remained constant. This unpredictability is reflected in the standard errors of the mean number of spurious peaks detected and especially in the mean proportion of peaks detected (Table 2). Additionally, the false discovery rate detected very few actual peaks. The bootstrapped confidence intervals were better than the false discovery rate, and detected a high number of real peaks, but consistently detected spurious peaks (Table 2). In contrast, the genome-wide confidence intervals were more conservative in the number of peaks detected, but consistently detected over 30% of the real peaks, and importantly very rarely detected spurious peaks (Table 2). Thus, for the rest of our results, significance was determined by 95% or 99% genome-wide confidence intervals. It is important to note that rarely were all quantitative trait loci detected, largely because quantitative trait loci of small effect tended not to show strong signatures of selection. The best detection rate achieved was 92.5% of the quantitative trait loci, with a spurious detection rate of 0, but more commonly a 'good' detection rate was 50-70%.

The effects of population size and sample size

Population genetics theory predicts that selection will have a stronger effect in larger populations due to a reduction in the effects of drift (Hartl and Clark 2007). Thus, we tested how well our selection components analysis detected selection in populations of varying sizes (1000, 2500, 5000, and 10000) with different sample sizes (100, 250, 500, 1000, 2000, and 4000). Sample size and population size both impacted the detection rates, and they appeared to have an interactive effect. The minimum sample size tested, 100 adults and 100 offspring, had a high

266 average number of spurious peaks detected (but this number was still below 1, suggesting that on
267 average fewer than one spurious peak was detected) and a low proportion of peaks detected (only
268 18.5% at the 99% confidence level, which means that between 1 and 2 of the eight actual
269 quantitative trait loci were detected; Fig. 2). This pattern was consistent across population sizes.
270 In other words, regardless of the actual population size, a sample of 100 adults and 100 offspring
271 was barely adequate to detect any quantitative trait loci. However, increasing the sample size
272 improved detection rates dramatically, especially in larger populations. As the population size
273 increased, the mean number of spurious loci fell below the mean detection rate (Fig. 2). Large
274 sample sizes alone improved detection rates, but the combination of a large sample with a large
275 population led to high detection rates as well as very low numbers of spurious loci detected.
276 *The effects of genome size*
277 In this model, the genome size could be manipulated by changing the number of
278 chromosomes, changing the number of markers per chromosome, or changing both. We
279 investigated the interaction between the total number of markers (1000, 2000, 4000, and 9000)
280 and chromosome number (1, 2, 4, and 8) on the detection of real and spurious peaks, and two
281 major patterns emerged. First, regardless of how many chromosomes among which the loci were
282 distributed, having more marker loci increased the average proportion of real peaks detected, but
283 also slightly increased the number of spurious loci detected. The number of chromosomes
284 seemed to determine the magnitude of the increase in spurious peaks. Small genome sizes (1000
285 and 2000) consistently had low detection rates (Fig. 3). However, with the markers distributed
286 across many chromosomes, the detection rate increased dramatically once there were more than
287 2000 markers. Indeed, with 12,000 markers spread evenly across 8 chromosomes, the
288 quantitative trait locus detection rate was 87.3% at the 99% confidence level and 94.3% at the

95% confidence level, which are some of the highest values we recorded. Essentially no spurious peaks were detected (Fig. 3) under these parameter combinations. This pattern may be due to the fact that the quantitative trait loci were equally distributed among chromosomes, so with 8 chromosomes and 8 quantitative trait loci, there was exactly 1 quantitative trait locus on each chromosome.

The effects of number of quantitative trait loci and strength of selection

The number of quantitative trait loci had a strong effect on the ability to detect real peaks, as we had predicted. The total strength of selection was distributed among all of the quantitative trait loci, so with fewer quantitative trait loci, each locus received a greater portion of the total selection. We tested sexual selection surface widths in the experimental generations (ω_{SE}^2) of 2, 8, 20, 50, 100, and 500 acting on a total of 4, 8, 16, and 32 quantitative trait loci distributed equally on 4 chromosomes. Selection strength was greatest at $\omega_{SE}^2 = 2$ in each case (male $m' \approx$ 1.2). The number of quantitative trait loci had a large effect on the detection of selection: when strong selection was acting on a total of 32 quantitative trait loci, only 6.5% real peaks were detected at the 99% confidence level (~2 of the quantitative trait loci), whereas with only 4 quantitative trait loci, 89.5% were detected on average (Fig. 4). The improvement in detection rates with few quantitative trait loci when selection was strong came with a cost when selection was weak: the number of spurious peaks detected at low selection strengths and few quantitative trait loci was higher than the number of peaks detected at low selection strengths but many quantitative trait loci (Fig. 4). Overall, these results suggest that accurately and reliably detecting selection requires that selection acts strongly on phenotypes that are mainly determined by a few quantitative trait loci of major effect.

The effects of environmental variation

The phenotype of an individual was determined by two components: the genotype derived from the quantitative trait loci and environmental effects. We tested how adding environmental variation to individuals' phenotypes affected the reliability of our genome-wide selection components analysis by changing the environmental variance (0, 0.1, 0.5, 1, 2, 8, 12, and 20) and the number of quantitative trait loci underlying the phenotype (4, 8, 16, and 32). Adding a small amount of environmental variance (0.1 or 0.5) did not have a large effect on the ability to detect quantitative trait loci under selection, and even led to a slight increase in the proportion of real peaks detected (Fig. 5). Once the environmental variance reached values greater than 1, the ability to detect loci under selection declined and the number of spurious loci detected increased (Fig. 5). However, the variance in male trait values without added environmental variance was typically between 0 and 1 in this model, so adding a value of up to ±20 to the phenotype may not be biologically relevant. Adding perhaps more relevant values (0.1, 0.5, 1, and 2) did not substantially alter the ability to detect selection (Fig. 5).

The number of quantitative trait loci underlying the trait also affected the prospects for detecting selection but buffered the effects of environmental variation; when 32 quantitative trait loci affected the trait, the proportion of quantitative trait loci detected was consistently below 20%, but did not decline significantly with added amounts of environmental variation, and the number of spurious loci detected remained near zero (Fig. 5). So although environmental variance added noise to the data, especially when many quantitative trait loci affected the trait of interest, it was still possible to detect selection even with a large amount of environmental variance.

Comparing multiple populations

We ran the model on multiple populations that all had the same quantitative trait loci, but

that experienced no gene flow, to determine if comparing even distant populations of the same species might help improve detection of quantitative trait loci. We found that there was no increase in reliability of detection (within a population, the average number of real and spurious loci remained the same). However, as predicted, the spurious loci differed between populations, allowing peaks that were at consistent loci to be identified as 'real' loci (Fig. 6).

Discussion

With this model, we set out to investigate the prospects for detecting a signature of sexual selection by comparing adults and offspring from one population using the type of data generated by next-generation sequencing approaches. We found that the genetic architecture of the trait was one of the most important determining factors in the ability to detect selection. As expected, more real peaks were detected when fewer quantitative trait loci contributed to variation in the trait. When sexual selection was strong and acted on a phenotype determined by few quantitative trait loci, even small sample sizes (i.e., 100 parent-offspring pairs) could accurately detect some of the real quantitative trait loci without generating prohibitively huge numbers of spurious peaks. Overall, we showed that a genome-wide selection components analysis has the potential to be very powerful in detecting signatures of sexual selection within a single population.

The results from our simulations suggest that current empirical methods for assessing significance may be unreliable. The false discovery rate, in particular, was unpredictable. This finding is sobering, since many studies have used the false discovery rate in analysis of RAD-seq and genome-wide scans (e.g. Hohenlohe et al. 2010a, Heylar et al. 2012, Hohenlohe et al. 2013, Narum et al. 2013). Although the observed unreliability may be a feature of selection components analysis *per se*, we suggest that researchers use caution when applying the false discovery rate. We instead suggest that using the very simple measure of genome-wide

confidence intervals would not only be appropriate but also be an accurate and repeatable method for defining cut-off values, since it best excluded spuriously significant loci while catching the majority of real peaks in our model.

Even though the occurrence of spurious loci was rare in our model, they occurred occasionally. In empirical studies, identifying which significant peaks are real and which are spurious may be challenging, but our results suggest that comparing multiple populations could help differentiate between real and spurious peaks. Although spurious loci occurred occasionally in our model, with multiple populations it was possible to identify the real peaks as those that occur consistently in all populations. This observation is consistent with recent evidence from empirical work that spurious loci should not be repeatable between populations or replicates; Tobler et al. (2014) showed that comparing replicates of lab- reared populations of *Drosophila melanogaster* was a very effective way to filter out false positives when looking for single nucleotide polymorphisms that responded to artificial selection regimes.

Environmental variation is expected to contribute to quantitative traits in real-world settings. We included environmental variation in our model and found intriguing results. Although large amounts of environmental variation dramatically reduced quantitative trait locus detection rates and increased the number of spurious loci, small amounts of environmental variation had very little effect on the detection of selection, even with many quantitative trait loci contributing to the trait. These small amounts of environmental variation (up to an environmental variance of 8) led to average heritability values within the range of 0.1 to 0.8, which is the range reported in studies with animal models (Visscher et al. 2008), so our genome-wide selection components analysis was robust to biologically relevant amounts of environmental variation.

Genome-wide selection components analysis was most effective at detecting only real

peaks and not spurious ones when the sample size was large (>1000 adults and >1000 offspring). Such large samples may be difficult to collect and also would be very costly to genotype. Fortunately, even small sample sizes identified real quantitative trait loci, at least when sexual selection was strong, with little change in the average number of spurious loci detected: in a large population (carrying capacity = 10,000), sampling 100 adults and 100 offspring identified on average 14.5% of the quantitative trait loci (at least one real peak), but only detected 0.228 spurious peaks (less than one spurious peak, on average). If 100 parent-offspring pairs were sampled from each of two distant populations (assuming both populations had the same genetic architecture of the trait and similar selection pressures), then the real peaks could be identified. This plan would not be logistically prohibitive, especially if other population genomics questions could be answered in the comparison of the two populations.

Empirical work suggests that most quantitative traits in several model organisms appear to have many underlying quantitative trait loci of small effect (Flint and Mackay 2009). Our model's ability to detect signatures of sexual selection was negatively impacted by an increase in the number of quantitative trait loci, which suggests that there may be limitations to the applicability of this method. However, we also modeled much smaller genome sizes than are observed in most organisms, and we observed that larger genome sizes (especially having additional chromosomes) increased detection rates of real peaks. Thus, genome-wide selection components analysis may indeed be able to capture signatures of selection on traits that are determined by many loci of small effect.

If researchers want to apply genome-wide selection components analysis to natural populations, we can provide several recommendations. First, this analysis was most effective in populations experiencing strong sexual selection, so it would be best implemented in a species

with clear evidence that sexual selection is occurring. Additionally, although larger sample sizes are always better, small sample sizes had the most spurious loci when the carrying capacity was small. Therefore, if the species of interest is known to have a small population size (as might be the case in some endangered species), investing in more comprehensive sampling may be especially worthwhile. Finally, it is important to note that we are detecting signatures of selection with this model and are not necessarily identifying the exact locus underlying the trait. When we identified peaks as 'real', the quantitative trait locus had to be within 50 loci in either direction from the peak F_{ST}' value. In some cases, the nearest marker locus did not have the greatest F_{ST} value, so additional work will probably be necessary in most cases to identify the actual DNA-level variant affecting fitness.

In summary, we investigated the potential for genome-wide selection components analysis to detect signatures of sexual selection by using an individual-based simulation model, in which allele frequencies in adults and offspring from a single population were compared. We were able to accurately detect some or most of the quantitative trait loci underlying the trait under selection, even when sample sizes were low or the trait was highly polygenic. Implementation of this method in studies of natural populations could provide another tool to identify genomic regions that are affected by sexual selection, leading to a better understanding of how selection affects the phenotype and results in the heritable changes that allow evolutionary change in natural populations.

Acknowledgments

This material is based upon work supported by the National Science Foundation Graduate Research Fellowship under Grant No. DGE-1252521. This work was supported by grant number DEB-1119261 from the National Science foundation to A.G.J.

References

Allard RW, Kahler AL, Clegg MT (1977) Estimation of mating cycle components of selection in plants. In: *Measuring selection in natural populations* (eds Christiansen F, Fenchel T) pp. 1-19. Springer Berlin Heidelberg.

Anderson JT, Lee CR, Mitchell-Olds T (2014) Strong selection genome-wide enhances fitness trade-offs across environments and episodes of selection. *Evolution,* **68**, 16-31.

Anderson WW (1969) Selection in experimental populations I. Lethal genes. *Genetics*, **62,** 653.

Anderson WW, Levine L, Olvera O, Powell JR, de la Rosa ME, Salceda VM, Gaso MI, Guzmán J (1979) Evidence for selection by male mating success in natural populations of *Drosophila pseudoobscura*. *Proceedings of the National Academy of Sciences of the United States of America,* **76,** 1519-1523.

Arnold SJ, Wade MJ (1984a). On the measurement of natural and sexual selection: Applications. *Evolution,* **38**, 720-734.

Arnold SJ, Wade MJ (1984b). On the measurement of natural and sexual selection: Theory. *Evolution*, **38**, 709-719.

Baccus R, Hillestad H, Johns P, Manlove M, Marchinton R, Smith M (1977) Prenatal selection in white-tailed deer. *Proceedings of the Annual Conference of the Southeast Association of Fish and Wildlife Agencies*, **31**, 173-179.

Baccus R, Joule J, Kimberling W (1980) Linkage and selection analysis of biochemical variants in peromyscus maniculatus. *Journal of Mammalogy,* **61**, 423-435.

Barbadilla A, Ruiz A, Santos M, Fontdevila A (1994) Mating pattern and fitness-component analysis associated with inversion polymorphism in a natural population of *Drosophila buzzatii*. *Evolution*, **48**, 767-780.

Benjamini Y, Hochberg Y (1995) Controlling the false discover rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society, Series B,* **57**, 289-300.

Boitard S, Rocha D (2013) Detection of signatures of selective sweeps in the blonde d'aquitaine cattle breed. *Animal Genetics*, **44**, 579-583.

Bundgaard J, Christiansen F (1972) Dynamics of polymorphisms. I. Selection components in an experimental population of *Drosophila melanogaster*. *Genetics,* **71**, 439.

Carlson CS, Eberle MA, Kruglyak L, Nickerson D (2004) Mapping complex disease loci in whole-genome association studies. *Nature,* **429**, 446-452.

Catchen JM, Amores A, Hohenlohe P, Cresko WA, Postlethwait JH (2011) Stacks: Building and genotyping loci de novo from short-read sequences. *G3: Genes, Genomes, Genetics,* **1**, 171-182.

Catchen J, Bassham S, Wilson T, Currey M, O'Brien C, Yeates Q, Cresko WA (2013a). The population structure and recent colonization history of oregon threespine stickleback determined using restriction-site associated DNA-sequencing. *Molecular Ecology,* **22**, 2864-2883.

Catchen J, Hohenlohe PA, Bassham S, Amores A, Cresko WA (2013b) Stacks: An analysis tool set for population genomics. *Molecular Ecology,* **22**, 3124-3140.

Christiansen FB (1980) Studies on selection components in natural populations using population samples of mother-offspring combinations. *Hereditas,* **92,** 199-203.

Christiansen FB, Frydenberg O (1973) Selection component analysis of natural polymorphisms using population samples including mother-offspring combinations. *Theoretical Population Biology*, **4**, 425-445.
Christiansen FB, Frydenberg O (1974) Geographical patterns of four polymorphisms in *Zoarces viviparus* as evidence of selection. *Genetics*, **77**, 765-770.
Christiansen FB, Frydenberg O (1976) Selection component analysis of natural polymorphisms using mother-offspring samples of successive cohorts. In: *Population Genetics and Ecology* (eds Karlin S, Nevo E), pp. 277-301. Academic Press, Inc., New York.
Christiansen FB, Frydenberg O, Simonsen V (1973) Genetics of Zoarces populations IV. Selection component analysis of an esterase polymorphism using population samples of mother-offspring combinations. *Hereditas*, **73**, 291-304.
Clark A, Feldman M (1981) The estimation of epistasis in components of fitness in experimental populations of *Drosophila melanogaster* II. Assessment of meiotic drive, viability, fecundity and sexual selection. *Heredity*, **46**, 347.
Clark A, Feldman M, Christiansen F (1981) The estimation of epistasis in components of fitness in experimental populations of *Drosophila melanogaster* I. A two-stage maximum likelihood model. *Heredity*, **46**, 321.
Clegg M, Kahler A, Allard R (1978) Estimation of life cycle components of selection in an experimental plant population. *Genetics,* **89**, 765-792.
Clement JAJ, Toulza E, Gautier M, Parrinello H, Roquis D, Boissier J, Rognon A, Mone H, Mouahid G, Buard J, Mitta G, Grunau C (2013) Private selective sweeps identified from next-generation pool-sequencing reveal convergent pathways under selection in two inbred *Schistosoma mansoni* strains. *PLOS Neglected Tropical Diseases*, **7**, e2591.
Curtsinger JW (1984) Components of selection in X chromosome lines of *Drosophila melanogaster*: Sex ratio modification by meiotic drive and viability selection. *Genetics,* **108**, 941-952.
Curtsinger JW, Feldman MW (1980) Experimental and theoretical analysis of the "sex-ratio" polymorphism in *Drosophila pseudoobscura*. *Genetics,* **94**, 445-466.
Emlen ST, Oring LW (1977) Ecology, sexual selection, and the evolution of mating systems. *Science,* **197**, 215-223.
FlintJ, Mackay TFC (2009) Genetic architecture of quantitative traits in mice, flies, and humans. *Genome Research*, **19**, 723-733.
Gagnaire PA., Pavey SA, Normandeau E, Bernatchez L (2013) The genetic architecture of reproductive isolation during speciation-with-gene-flow in lake whitefish species pairs assessed by RAD sequencing. *Evolution*, **67**, 2483-2497.
Harris SE, Munshi-South J, Obergfell C, O'Neill R (2013). Signatures of rapid evolution in urban and rural transcriptomes of white-footed mice (*Peromyscus leucopus*) in the New York metropolitan area. *PLoS ONE,* **8**, e74938.
Hartl DL, Clark AG (2007) Principles of population genetics. Sinauer Associates, Sunderland, Massachusetts.
Heath D, Riddoch B, Childs D, Ratford J (1988) Selection component analysis of the *pgi* polymorphism in *Sphaeroma rugicauda*. *Heredity,* **60**, 229-235.
Hess JE, Campbell NR, Close DA, Docker MF, Narum SR (2013) Population genomics of pacific lamprey: adaptive variation in a highly dispersive species. *Molecular Ecology,* **22**, 2898-2916.

Helyar SJ, Limborg MT, Bekkevold D, *et al.* (2012) SNP discovery using next generation transcriptomic sequencing in Atlantic herring (*Clupea harengus*). *PLoS ONE,* **7**, e42089.
Hohenlohe PA, Bassham S, Etter PD, Stiffler N, Johnson EA, Cresko WA (2010a) Population genomics of parallel adaptation in threespine stickleback using sequenced RAD tags. *PLoS Genetics,* **6**, e1000862.
Hohenlohe PA, Day MD, Amish SJ, Miller MR, Kamps-Hughes N, Boyer MC, Muhlfeld CC, Allendorf FW, Johnson EA, Luikart G (2013) Genomic patterns of introgression in rainbow and westslope cutthroat trout illuminated by overlapping paired-end RAD sequencing. *Molecular Ecology*, **22**, 3002–3013.
Hohenlohe PA, Phillips PC, Cresko WA (2010b) Using population genomics to detect selection in natural populations: key concepts and methodological considerations. *International Journal of Plant Sciences*, **171**, 1059-1071.
Hubner S, Rashkovetsky E, Kim YB, Oh JH, Michalak K, Weiner D, Korol AB, Nevo E, Michalak P (2013) Genome differentiation of *Drosophila melanogaster* from a microclimate contrast in Evolution Canyon, Israel. *Proceedings of the National Academy of Sciences of the United States of America*, **110**, 21059-21064.
Jones AG (2009) On the opportunity for sexual selection, the Bateman gradient and the maximum intensity of sexual selection. *Evolution*, **63**, 1673-1684.
Loehle C, Pechmann JH (1988) Evolution: The missing ingredient in systems ecology. *The American Naturalist* 132:884.
McDonald J (1989) Selection component analysis of the *mpi* locus in the amphipod *Platorchestia platensis*. *Heredity,* **62**, 243-249.
Miller MR, Brunelli JP, Wheeler PA, Liu S, Rexroad CE, Palti Y, Doe CQ, Thorgaard GH (2012) A conserved haplotype controls parallel adaptation in geographically distant salmonid populations. *Molecular Ecology,* **21**, 237-249.
Nadeau JH, Baccus R (1981) Selection components of four allozymes in natural populations of *Peromyscus maniculatus*. *Evolution,* **35**, 11-20.
Nadeau JH, Baccus R (1983) Gametic selection and hemoglobin polymorphisms in *Peromyscus maniculatus*: A rejoinder. *Evolution,* **37**, 642-646.
Narum SR, Campbell NR, Meyer KA, Miller MR, Hardy RW (2013) Thermal adaptation and acclimation of ectotherms from differing aquatic climates. *Molecular Ecology,* **22**, 3090–3097.
Nielsen R (2005) Molecular signatures of natural selection. *Annual Review of Genetics*, **44**, 197-218.
Østergaard H, Christiansen FB (1981) Selection component analysis of natural polymorphisms using population samples including mother-offspring combinations, II. *Theoretical Population Biology,* **19,** 378-419.
Prout T (1965) The estimation of fitnesses from genotypic frequencies. *Evolution*, **19**, 546-551.
Prout T (1969) The estimation of fitnesses from population data. *Genetics*, **63**, 949.
Prout T (1971a). The relation between fitness components and population prediction in Drosophila. I: The estimation of fitness components. *Genetics,* **68,** 127.
Prout T (1971b) The relation between fitness components and population prediction in Drosophila. II: Population prediction. *Genetics,* **68,** 151.
Reitzel AM, Herrera S, Layden MJ, Martindale MQ, Shank TM (2013) Going where traditional markers have not gone before: utility of and promise for rad sequencing in marine

invertebrate phylogeography and population genomics. *Molecular Ecology,* **22**, 2953-2970.

Rellstab C, Zoller S, Tedder A, Gugerli F, Fischer MC (2013) Validation of SNP allele frequencies determined by pooled next-generation sequencing in natural populations of a non-model plant species. *PLoS ONE,* **8**, e80422.

Siegismund HR, Christiansen FB (1985) Selection component analysis of natural polymorphisms using population samples including mother-offspring combinations, III. *Theoretical Population Biology,* **27**, 268-297.

Stockwell CA, Hendry AP, Kinnison MT (2003) Contemporary evolution meets conservation biology. *Trends in Ecology & Evolution*, **18**, 94-101.

Tobler R, Franssen SU, Kofler R, Orozco-terWengel P, Nolte V, Hermisson J, Schlötterer C (2014) Massive habitat-specific genomic response in *Drosophila melanogaster* populations during experimental evolution in hot and cold environments. *Molecular Biology and Evolution,* **31**, 364-375.

Vincent B, Dionne M, Kent MP, Lien S, Bernatchez L (2013) Landscape genomics in Atlantic salmon (*Salmo salar*): Searching for gene-environment interactions driving local adaptation. *Evolution*, **67**, 3469-3487.

Visscher PM, Hill WG, Wray NR (2008). Heritability in the genomics era—concepts and misconceptions. *Nature Reviews Genetics,* **9**, 255-266.

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

Yamazaki T (1971) Measurement of fitness at the *esterase-5* locus in *Drosophila pseudoobscura*. *Genetics,* **67**, 579.

Data Accessibility

Data will be archived on Dryad upon acceptance of the manuscript, and the source code for the model will be made available on the web.

Author Contributions

A.G.J. and S.P.F. contributed to study design, writing of the simulation model, data analysis, and writing of the manuscript.

Table 1. The baseline parameters for running the simulation model. Selection variances refer to the Gaussian selection surface variance. "Initial" refers to the variance during the initial generations before sexual selection was imposed, and "Experimental" refers to the variance during the subsequent five generations during which the population was sampled.

Parameter	**Starting Value**
Carrying capacity	5000
Sample size (adults)	4000
Sample size (offspring)	4000
Number of markers per chromosome	1000
Number of QTL per chromosome	2
Number of chromosomes	4
Initial mate choice strength (ω_{SI}^2)	Random mating
Experimental mate choice strength (ω_{SI}^2)	2
Initial viability selection strength (ω_{VE}^2)	500
Experimental viability selection strength (ω_{VE}^2)	500
Environmental variance	0
Number of populations	1

Table 2. The reliability of three methods to determine significance cut-off thresholds for weighted F_{ST} values were compared using the mean proportion of actual peaks detected and the mean number of spurious peaks detected and their standard errors. These means were generated by running the model with its default parameters (Table 1) in10 replicates, each of which had 200 initial generations where no sampling occurred, followed by 5 generations in which the population was sampled and F_{ST} measures calculated between adults and offspring. Allele frequencies were calculated from a sample of 4000 adults and 4000 offspring from a population with a carrying capacity of 5000. Each individual had 1000 marker loci and 2 quantitative trait loci on each of 4 chromosomes. Viability selection was weak during both the initial and the experimental generations ($\omega_{VI}^2 = \omega_{VE}^2 = 500$) and strong sexual selection was introduced at the start of the experimental generations (ω_{SI}^2 = random mating, $\omega_{SE}^2 = 2$).

	Mean Proportion of Actual Peaks Detected	SE	Mean Number of Spurious Peaks Detected	SE
99% Genome-wide CI	0.3475	0.0245	0.1000	0.0428
95% Genome-wide CI	0.4925	0.0318	0.1000	0.0428
99% Bootstrapped CI	0.8000	0.0404	1.1400	0.1874
95% Bootstrapped CI	0.8075	0.0395	1.5600	0.2063
False Discovery Rate	0.1525	0.0455	4.0400	1.5367

Figure 1. The effect of linkage disequilibrium (determined by the number of initial generations) on the percentage of quantitative trait loci accurately detected and the number of spurious loci called as significant by 99% and 95% genome-wide confidence intervals for weighted F_{ST} values. The measure of linkage disequilibrium presented here is D' calculated as a pairwise measure between 100 loci randomly selected from each chromosome and averaged across chromosomes and replicates. Linkage disequilibrium was calculated in the final initial generation, and the number of initial generations was varied to change linkage disequilibrium (initial generation numbers are presented above the points on the graphs). Panel A shows all permutations of the number of initial generations (1, 2, 4, 8, 10, 20, 40, 50, 75, 100, 150, 200, 250, 500, and 1000). Panel B presents a close-up view of generations 40, 50, 75, 100, 150, 200, 250, 500, and 1000 to highlight the changes that occur at low levels of linkage disequilibrium (below 0.3). The model was run with weak viability selection ($\omega_{VI}^2 = 500$) and random mating during the initial generations and weak viability selection ($\omega_{VE}^2 = 500$) and strong sexual selection during the experimental generations ($\omega_{SE}^2 = 2$). There were 5000 individuals in the population. The genome comprised four chromosomes, each of which carried 1000 marker loci and 2 quantitative trait loci. During each of the 5 experimental generations, 4000 adults and 4000 offspring were sampled. Values presented here are averages from 10 replicates, each of which was sampled over five experimental generations (i.e. a total of 50 generations were sampled). Bars are the standard error of the mean.

Figure 2. The effect of population size and sample size on detection rates. For each of 10 replicates, the model ran for 200 initial generations followed by 5 experimental generations. The number of real and spurious peaks detected were averaged across the five experimental generations of all ten replicates. The carrying capacity and sample size were constant throughout all initial and experimental generations, and the same number of adults and offspring were sampled in each of the five experimental generations. Each individual had 1000 marker loci and 2 quantitative trait loci on each of 4 chromosomes. Viability selection was weak during both the initial and the experimental generations ($\omega_{VI}^2 = \omega_{VE}^2 = 500$) and strong sexual selection was introduced at the start of the experimental generations ($\omega_{SI}^2 = 0$, $\omega_{SE}^2 = 2$). Bars show the standard error of the mean.

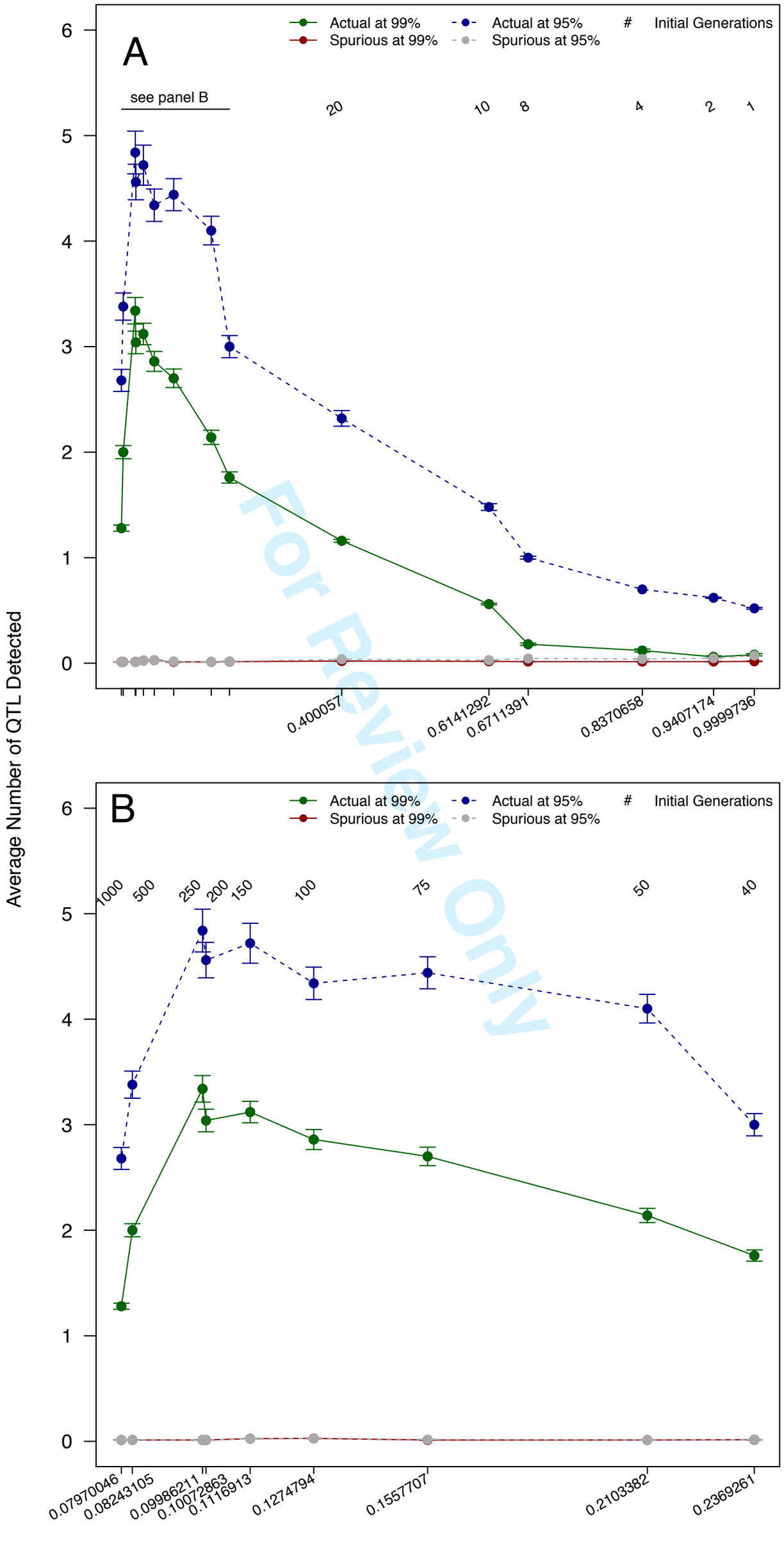
Figure 3. The effect of the number of markers and number of chromosomes on detection rates. These data represent averages from five experimental generations of ten replicates, each of which had a carrying capacity of 5000, adult and offspring sample sizes of 4000 each, and ran for an initial 200 generations followed by 5 experimental generations. A constant number of 8 quantitative trait loci were distributed equally across the chromosomes (so with 4 chromosomes, there were 2 quantitative trait loci on each, but the location of each quantitative trait locus on the chromosome was randomly chosen). Viability selection was weak during both the initial and the experimental generations ($\omega_{VI}^2 = \omega_{VE}^2 = 500$) and strong sexual selection was introduced at the start of the experimental generations (ω_{SI}^2 = random mating, $\omega_{SE}^2 = 2$). Bars indicate the standard error of the mean.

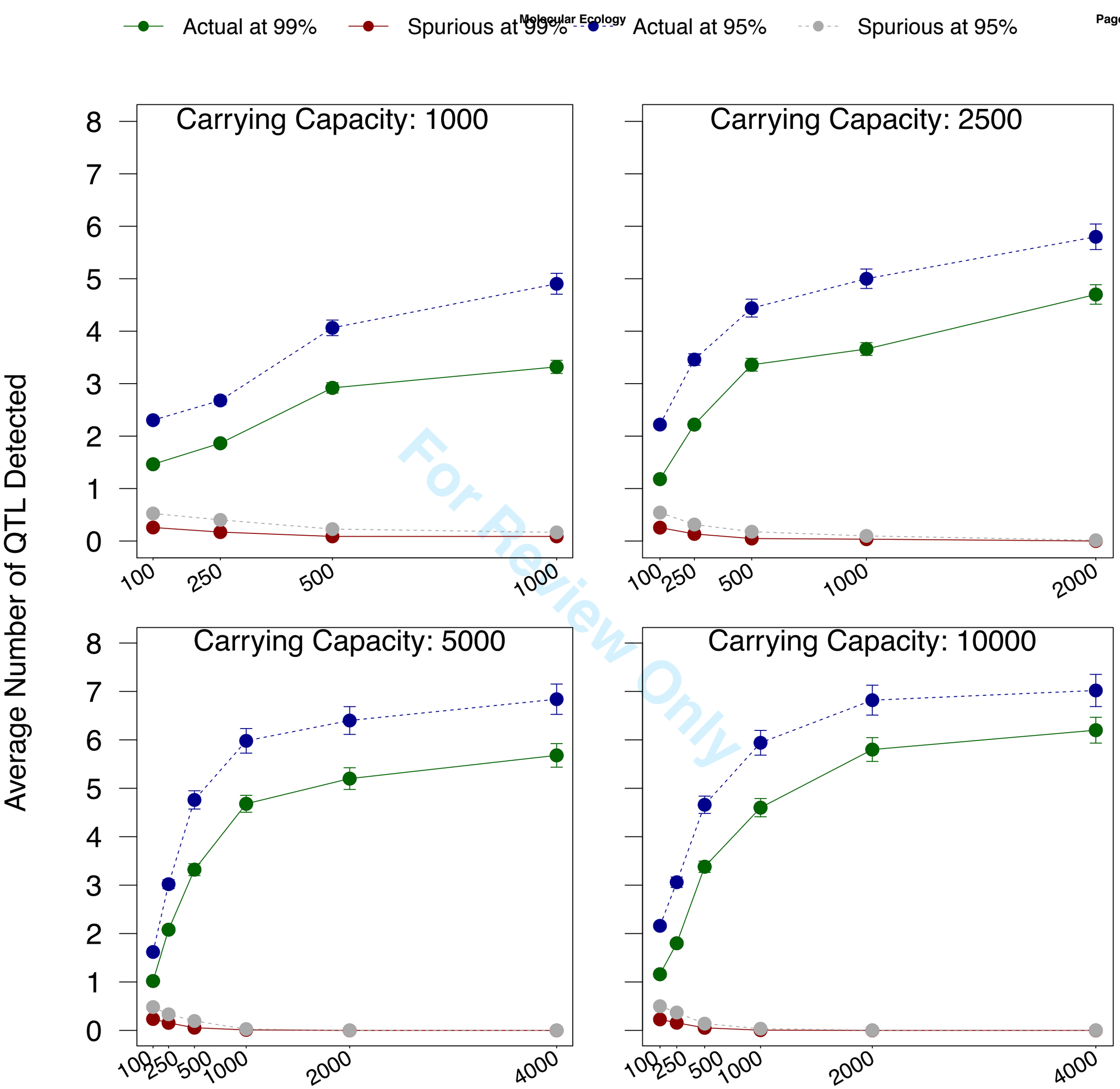
Figure 4. The effect of sexual selection strength and the number of quantitative trait loci on detection rates. We tested selection surface widths (ω_{SI}^2) of 2, 8, 20, 50, and 100 during the experimental generations. There was random mating during the initial 200 generations, and these

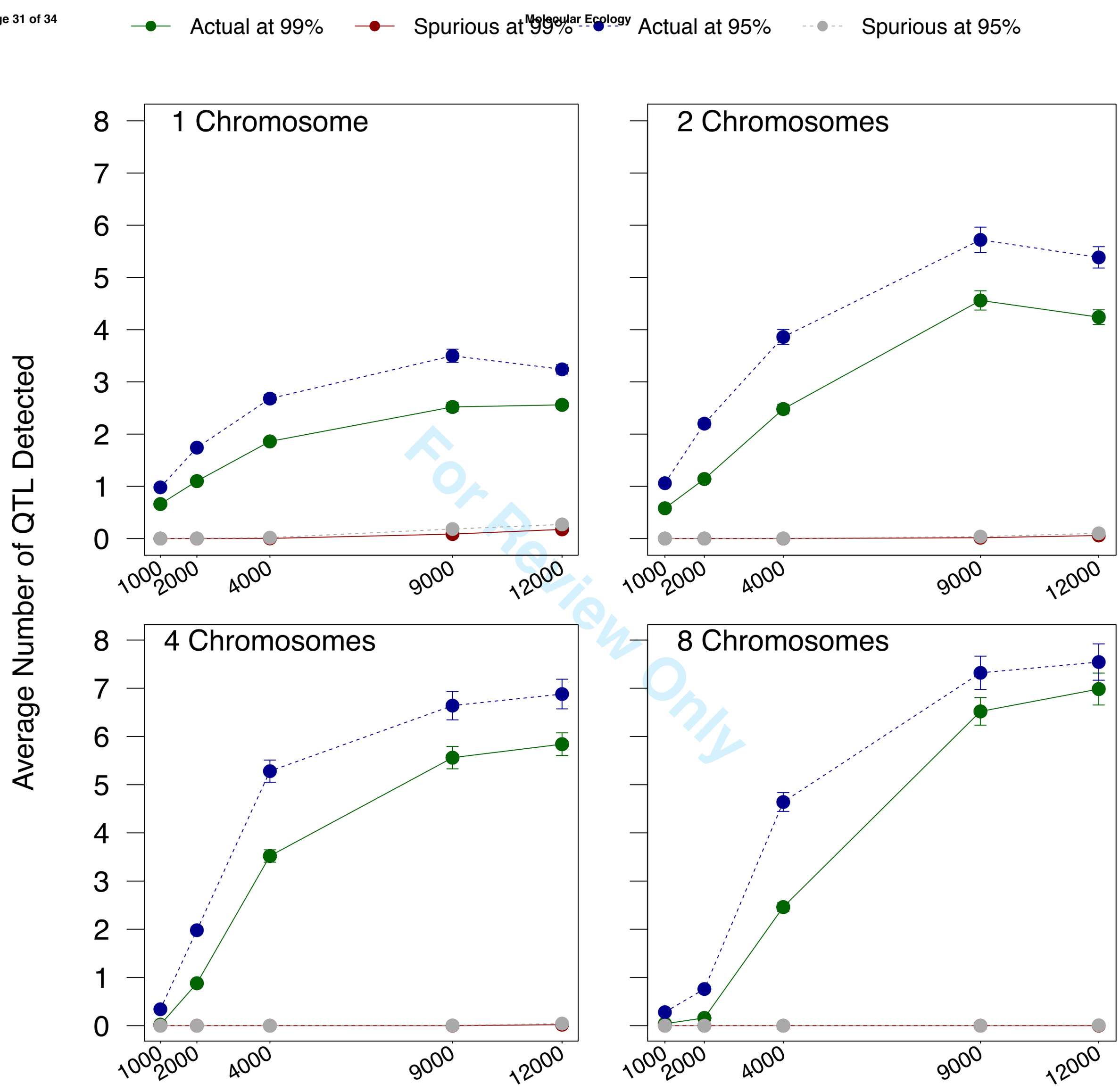
selection strengths were implemented starting in the first of the five experimental generations. Weak viability selection occurred ($\omega_{VI}^2 = \omega_{VE}^2 = 500$). The carrying capacity was 5000 and adults and offspring each had a sample size of 4000. Each individual had 4 chromosomes each with 1000 marker loci, and the total number of quantitative trait loci was distributed equally among the chromosomes. These data are averages and standard errors from the five experimental generations of 10 replicates.

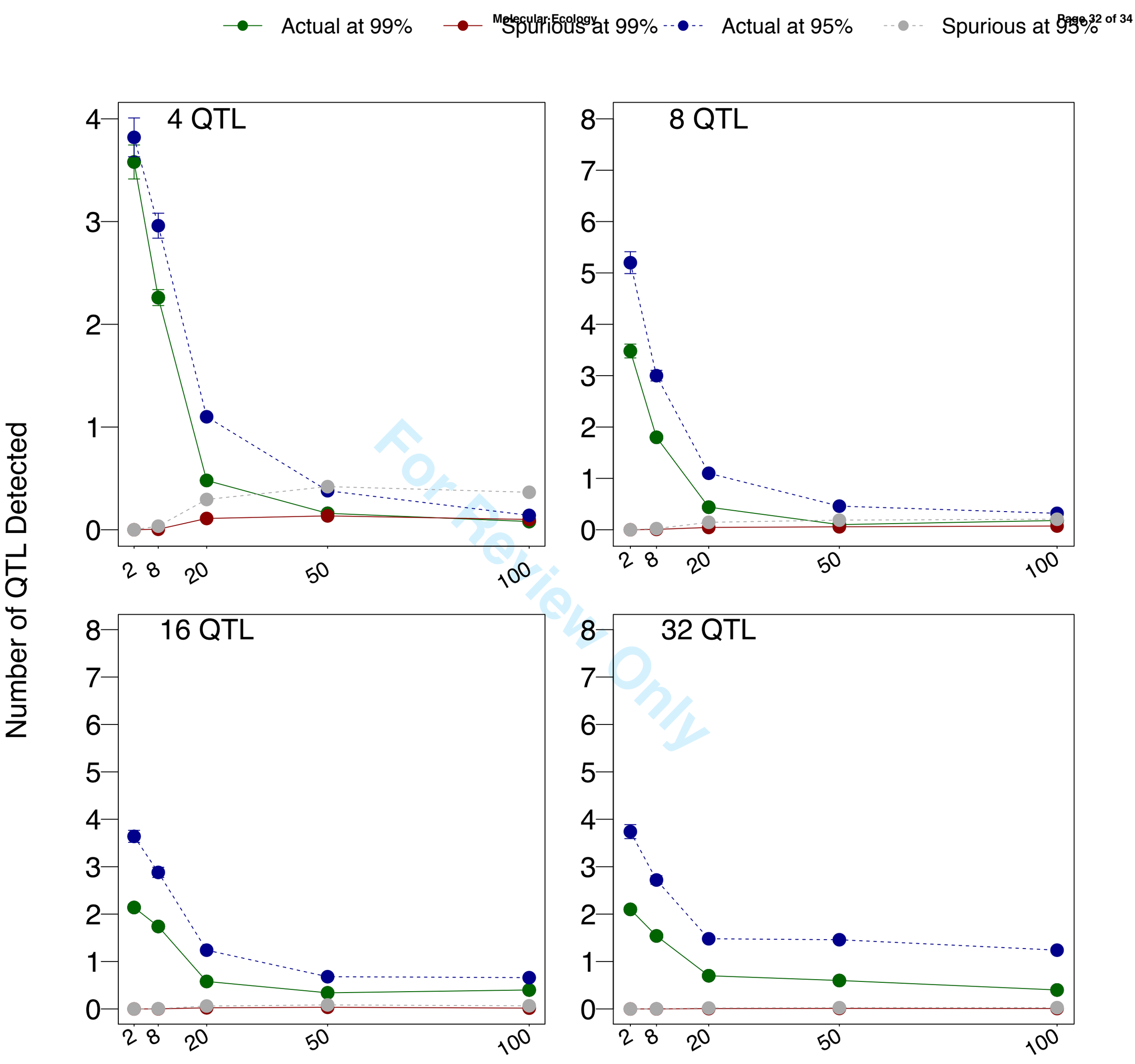
Figure 5. The effects of environmental variation on the ability to detect quantitative trait loci using an outlier F_{ST} approach. Environmental variation was implemented by drawing a number from a normal distribution with variances of 0, 0.1, 0.5, 1, 2, 8, 12, and 20 and adding that value to the phenotype of the individual. There was a carrying capacity of 5000 and sample sizes of 4000 adults and 4000 offspring. Each individual had 1000 marker loci on each of 4 chromosomes, and the quantitative trait loci were distributed equally among the chromosomes. Ten replicates of the model were run with 200 initial generations and 5 experimental generations. The number of real and spurious peaks were averaged over all five experimental generations for all ten replicates. Viability selection was weak during initial and experimental generations ($\omega_{VI}^2 = \omega_{VE}^2 = 500$), and sexual selection was introduced during the experimental generations (ω_{SI}^2 = random mating, $\omega_{SE}^2 = 2$). The heritability of the trait is included in the panel with 32 QTL, and a similar pattern was observed for all QTL numbers. Error bars are the standard error of the mean.

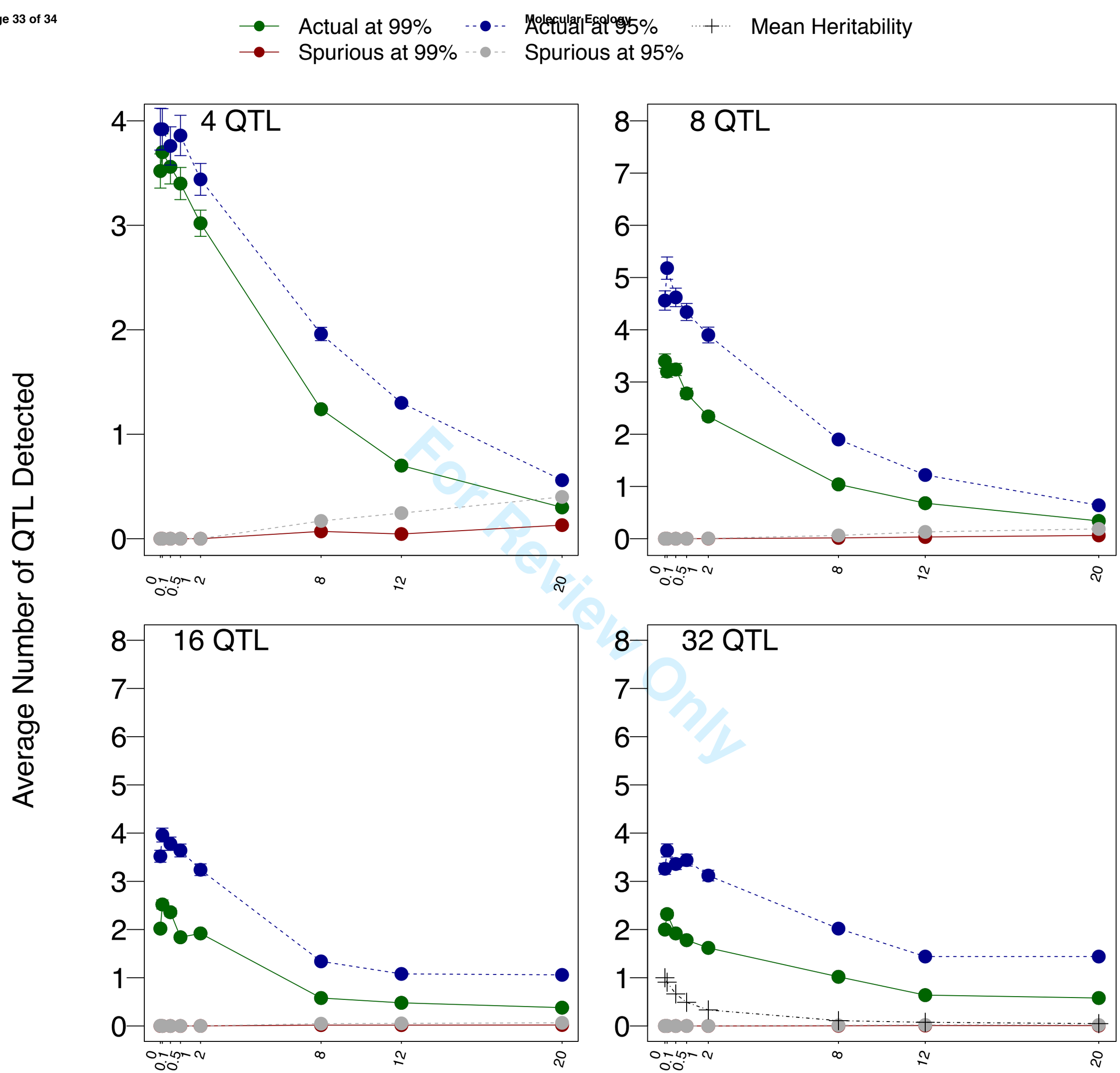
Figure 6. Sampling multiple populations can improve detection of real quantitative trait loci, if the same loci underlie the trait affected by selection. The model was run with 10 replicates, but in each replicate the 8 same quantitative trait loci were designated, rather than being randomly assigned. Thus, each replicate was essentially another population without gene flow. The model was run with the default parameters (Table 1), so that there were 200 initial generations with weak viability selection and no sexual selection ($\omega_{VI}^2 = 500$, ω_{SI}^2 = random mating) followed by 5 experimental generations with weak viability selection and strong sexual selection ($\omega_{VE}^2 = 500$, $\omega_{SE}^2 = 2$). The carrying capacity of the population was 5000 and the sample size was 4000. Each individual had 4 chromosomes, each with 1000 marker loci and 2 quantitative trait loci. Comparing the significant weighted F_{ST} values uncovered in each population, it is obvious that the peaks that reappear in each population are the quantitative trait loci (whose locations are designated by red asterisks). The four chromosomal regions are delineated by different colored backgrounds, and five different genome-wide confidence intervals (99%, 95%, 90%, 85%, and 80%) are depicted on the graph.

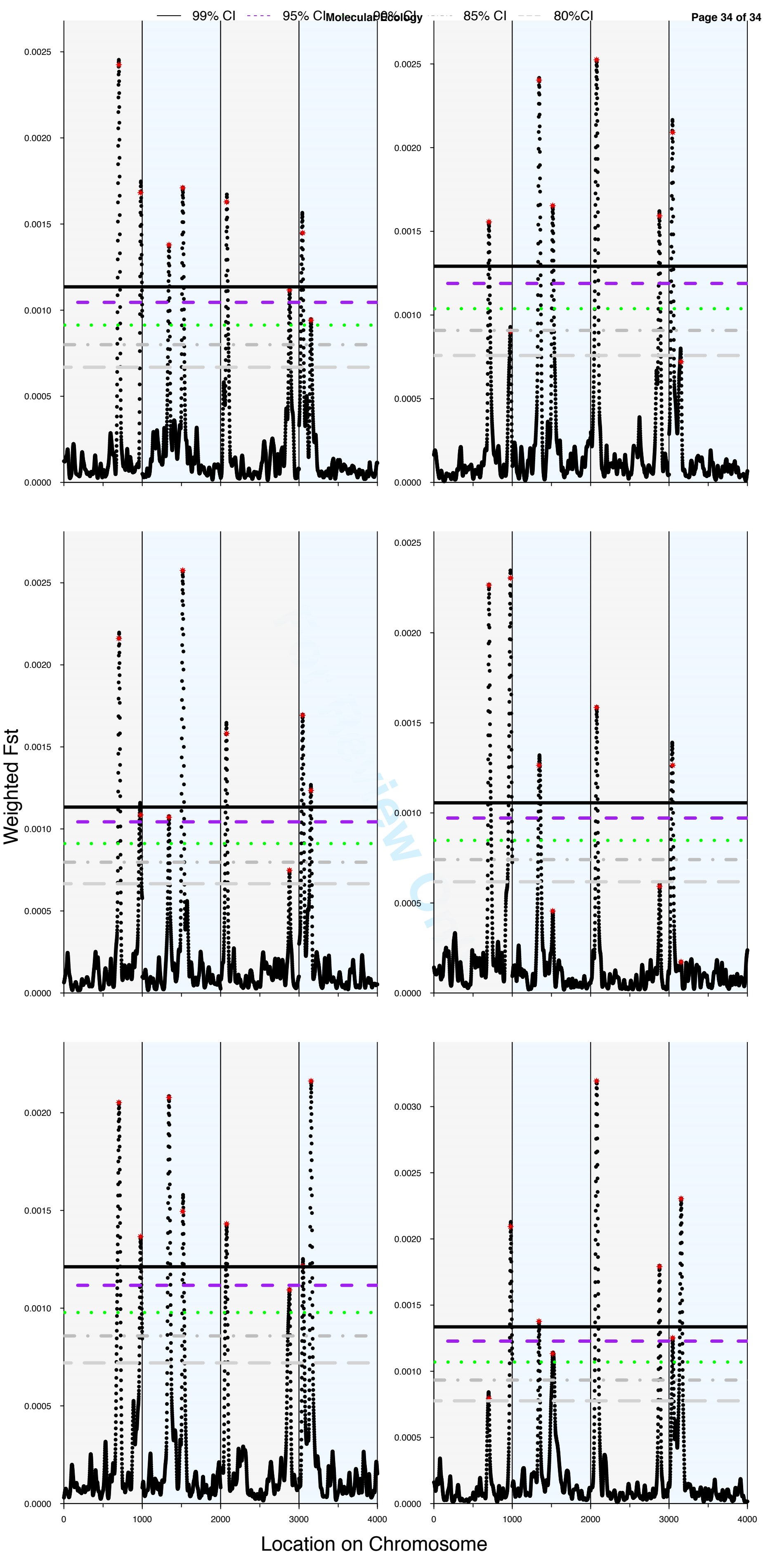











99% CI
95% CI
85% CI
80%CI
Weighted Fst
Location on Chromosome
0.0000
0.0005
0.0010
0.0015
0.0020
0.0025
0.0030
0
1000
2000
3000
4000