

## RESEARCH ARTICLE

# Ethanol resistance in *Drosophila melanogaster* has increased in parallel cold-adapted populations and shows a variable genetic architecture within and between populations

Quentin D. Sprengelmeyer | John E. Pool 

Laboratory of Genetics, University of Wisconsin-Madison, Madison, Wisconsin, USA

**Correspondence**

John E. Pool, Laboratory of Genetics, University of Wisconsin-Madison, 425-G Henry Mall, Madison, WI 53706, USA.  
Email: jpool@wisc.edu

**Funding information**

National Institute of General Medical Sciences, Grant/Award Number: R01 GM111797, R35 GM136306 and T32 GM007133; Division of Environmental Biology, Grant/Award Number: DEB 1754745

**Abstract**

Understanding the genetic properties of adaptive trait evolution is a fundamental crux of biological inquiry that links molecular processes to biological diversity. Important uncertainties persist regarding the genetic predictability of adaptive trait change, the role of standing variation, and whether adaptation tends to result in the fixation of favored variants. Here, we use the recurrent evolution of enhanced ethanol resistance in *Drosophila melanogaster* during this species' worldwide expansion as a promising system to add to our understanding of the genetics of adaptation. We find that elevated ethanol resistance has evolved at least three times in different cooler regions of the species' modern range—not only at high latitude but also in two African high-altitude regions. Applying a bulk segregant mapping framework, we find that the genetic architecture of ethanol resistance evolution differs substantially not only between our three resistant populations, but also between two crosses involving the same European population. We then apply population genetic scans for local adaptation within our quantitative trait locus regions, and we find potential contributions of genes with annotated roles in spindle localization, membrane composition, sterol and alcohol metabolism, and other processes. We also apply simulation-based analyses that confirm the variable genetic basis of ethanol resistance and hint at a moderately polygenic architecture. However, these simulations indicate that larger-scale studies will be needed to more clearly quantify the genetic architecture of adaptive evolution and to firmly connect trait evolution to specific causative loci.

**KEYWORDS**

adaptive evolution, alcohol, *Drosophila melanogaster*, genetic architecture, genetic differentiation, local adaptation, quantitative trait locus mapping

## 1 | INTRODUCTION

The genetic basis of adaptive trait evolution is an area of great interest to biologists and has raised several key questions. There are two

questions that are of particular interest to this study. For example, how polygenic is trait evolution (Wellenreuther & Hansson, 2016)? And do favored variants tend to reach fixation, or stop rising because selective pressures change or traits reach their new optima (Thornton, 2019)?

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2021 The Authors. *Ecology and Evolution* published by John Wiley & Sons Ltd.

Early theory suggested that adaptive trait evolution is the result of many genes with small effect (Fisher, 1930), or mutations with intermediate effect size (Kimura, 1983). A more recent hypothesis proposes that depending on where a population is relative to the phenotypic optimum will dictate whether few mutations with large effect or many small effect mutations will be favored (Orr, 1998). This model argues that when an organism first encounters a novel environment, genes of large effect size would be most abundant and as the population moves closer to an optimal phenotype, the effect size would decrease, with an overall geometric distribution of effect sizes predicted. Alternatively, migration–selection balance may favor larger effect sizes underlying local adaptation (Yeaman & Whitlock, 2011), whereas an important role for previously deleterious standing variation may lead to a greater role for smaller effects instead (Dittmar et al., 2016).

There have been many studies that investigate the genetic architecture of novel traits. A classic example is the body color of peppered moths found in England. It was discovered that a single gene is responsible for the light and dark morph (van't Hof et al., 2011). While a number of other cases of a simple genetic basis to trait evolution are known, it is unclear how common they are in light of ascertainment bias (Rockman, 2012). As a contrasting example, studies that have looked at human height differences between populations found that this trait may be due to many genes (Gudbjartsson et al., 2008; Turchin et al., 2012). However, recent studies have called into question these results and suggest that population stratification has led to overestimating the number of genes involved (Berg et al., 2019; Sohail et al., 2019). There is growing evidence for polygenic adaptation underlying trait evolution (Barghi et al., 2020). However, the term “polygenic adaptation” actually encompasses a broad array of possible evolutionary scenarios in terms of the number of loci involved, the magnitudes of their effects on evolving traits, and their frequencies before and after being selectively favored.

Theoretic and empirical studies have led to testable predictions about polygenic adaptation. First, depending in part on the number and effect sizes of beneficial alleles, they may not always reach fixation (Barghi et al., 2020; Barghi & Schlötterer, 2020; Höllinger et al., 2019; John and Stephan, 2020; Stephan, 2016; Thornton, 2019). Experimental evolution studies, including in *Drosophila*, have begun to provide evidence for genetically heterogeneous responses to selection. Considerable variability among replicates in the loci that respond to selection on environmental tolerances (Barghi et al., 2019; Griffin et al., 2017) suggests an abundance of segregating variation in natural populations for selection in natural populations to act upon, not all of which is needed for individuals to display an evolved trait. And when another study selected for viral resistance, plateaus in the frequencies of initially favored alleles were observed (Faria et al., 2018). If similar dynamics underlie natural instances of trait evolution, then different individuals from the same population may often have distinct genetic architectures underlying the same evolved trait. As one example, a previous study found that melanism had a variable genetic basis within three different highland African populations of *D. melanogaster* (Bastide et al., 2016).

*Drosophila melanogaster* originated in woodland environments of southern-central Africa and then expanded throughout Africa beginning ~13,000 years ago (Sprengelmeyer et al., 2020). The species appears to have crossed the Sahara relatively soon after their expansion started and may have only reached Europe ~1800 years ago. During the migration out of their ancestral habitat, populations of *D. melanogaster* encountered many novel environments, which included equatorial tropical rainforest, northern temperate grassland, and high-altitude alpine regions. Each of these different ecosystems provides unique selection pressures that may have forced local populations to acquire novel traits in order to survive. For example, increased ultraviolet radiation found at higher altitudes might have caused populations to evolve darker cuticle pigmentation (Bastide et al., 2014). And populations at high latitude and altitude have independently evolved elevated cold tolerance (Pool et al., 2017).

Ethanol resistance is another trait that has evolved in *D. melanogaster*. When compared to its sister species *D. simulans*, *D. melanogaster* are more ethanol resistant (McKenzie & Parsons, 1972). Within *D. melanogaster*, ethanol resistance has shown a positive correlation with latitude (Cohan & Graf, 1985; David & Bocquet, 1975) with populations living in breweries and wine cellars of France and Spain being the most resistant (McKenzie & Parsons, 1972; Merçot et al., 1994). Female flies lay their eggs on ethanol-producing fermenting fruit and having a higher ethanol resistance may provide more available resources (Merçot et al., 1994). There is evidence that *D. melanogaster* prefers to lay their eggs on medium that contains alcohol (McKenzie & Parsons, 1972), which can be a defense against parasitoids (Kacsoh et al., 2013; Milan et al., 2012).

Alcohol metabolism in *D. melanogaster* involves ethanol being converted to acetaldehyde by *ADH* (Greer et al., 1993). *ADH* and *ALDH* convert acetaldehyde to acetate. Acetate can be turned into acetyl-CoA, which can be used in the production of fatty acids, the citric acid cycle, and other pathways. Differences at the *Adh* gene are correlated with improved alcohol resistance (David & Bocquet, 1976), with the “fast” allele having a higher resistance compared with the “slow” allele. David and Bocquet (1975) found a latitudinal gradient and populations at higher latitudes tend to be more resistant and also have a higher *Adh*<sub>fast</sub> frequency. However, *D. funebris*, *D. littoralis*, and *D. mercatorum* all display ethanol resistance but low *ADH* activity, whereas in spite of high *ADH* activity, *D. ercepeae* are classified as being sensitive to alcohol (Merçot et al., 1994). It has been hypothesized that the *Adh*<sub>fast</sub> and *Adh*<sub>slow</sub> polymorphism has been maintained by a temperature dependent balancing selection (Van Delden et al., 1978). However, Siddiq and Thornton (2019) found *Adh*<sub>fast</sub> protein is neither less stable nor active at high temperatures and will increase ethanol resistance along with survivorship at all temperatures. Further, when they analyzed a population genomic data set, there was not a signature of balancing selection in the *Adh* gene.

Changes at *ALDH* can also increase ethanol resistance (Fry & Saweikis, 2006). Fry et al. (2008) also showed that there is an amino acid difference between more resistant populations found in higher latitudes and less resistant flies found in lower latitudes.

It has also been found that European flies can have higher *ALDH* enzyme activity compared with less resistant African flies even without the amino acid polymorphism (Fry, 2014). Chakraborty and Fry (2016) found that polymorphisms in *ALDH* are maintained by environmental conditions, and their transgenic experiments confirmed effects on lifetime fitness on ethanol-supplemented medium specifically.

Although *ADH* and *ALDH* play an important role, they are not the only genes involved in ethanol resistance. Other genes linked to ethanol resistance encompass a wide range of functions such as lipid membrane physiology (Montooth et al., 2006), ion channels (Cowmeadow et al., 2005), central nervous system (Chandler et al., 1998), zinc retention (Zhao et al., 2009), and feeding behavior and behavioral responses to ethanol (Fochler et al., 2017). Signor and Nuzhdin (2018) found that many genes display plasticity in expression and splicing in response to ethanol exposure. Other studies that focused on changes in gene expression (Morozova et al., 2006) or histone modification (Ghezzi et al., 2013) have also found that numerous genes respond to ethanol exposure.

In the present study, we use ethanol resistance in *D. melanogaster* to further our understanding of the genetic architecture of adaptive trait evolution. Although a number of prior studies have studied variation in ethanol resistance both within and between *D. melanogaster* populations (as indicated above), the present study integrates resources and approaches not previously deployed in this pursuit. We use wild populations from their ancestral range (Zambia), along with multiple populations that display elevated ethanol resistance: from high altitude sub-Saharan Africa (Ethiopia and South Africa) and from high latitude (France). We note that each of the ethanol-resistant populations has also evolved elevated cold tolerance, and in light of the species' expansion history, this trait change is thought to have arisen independently in each of these three populations' histories (Pool et al., 2017). To detect QTLs that are involved in this adaptive trait evolution, we perform bulk segregant analysis (Pool, 2016). We then leverage population genomic data to identify signatures of local adaptation within QTLs, while also applying Gene Ontology (GO) enrichment and genotype-phenotype association testing to search for potential causative genes. We also perform simulations to explore the parameters (number of detected QTLs, environmental variance, and QTL strength) involved in the genetic architecture of this adaptive trait change.

## 2 | MATERIALS AND METHODS

### 2.1 | Experimental populations

All flies used in the experiment had been inbred for 8 generations from wild-caught isofemale lines (Lack et al., 2015). The sub-Saharan African populations came from Fiche, Ethiopia (EF, 9.81°N, 38.63°E, alt. 3070 m), Dullstroom, South Africa (SD, 25.42°S, 30.10°E, alt. 2000 m), and Siavonga, Zambia (ZI, 16.54°S, 28.72°E, alt. 530 m). The French samples are from Lyon, France (FR, 45.77°N, 4.86°E, alt.

175 m). Flies were all raised at 20°C on medium prepared in batches of 4.5 L water, 500 ml cornmeal, 500 ml molasses, 200 ml yeast, 54 g agar, 20 ml propionic acid, and 45 ml tegosept 10% (in 95% ethanol).

### 2.2 | Ethanol resistance

To test for population differences in ethanol resistance, we measured mobility over a 6-h period. We collected data from the offspring of inbred lines from the same population. The number of flies and pairs of lines used was as follows: FR: 10 lines,  $N = 300$ , EF: 10 lines,  $N = 300$ , SD: 10 lines,  $N = 300$  and ZI: 10 lines,  $N = 300$ . We then placed 3- to 5-day-old female flies into 50-ml falcon tubes with a single tissue placed in the bottom that was saturated with 1.5 ml of 3% sucrose (molasses) solution that contained 8% ethanol (Fry, 2014). Every 15 min, we visually checked tubes and scored flies that did not move after the vial was flicked as "immobile."

To test for evidence that local adaptation has influenced population differences in ethanol resistance, we calculated  $Q_{ST}$  values (Lande, 1992; Miller et al., 2008; Spitze, 1993) based on two-population comparisons of the above phenotypic data. We compared these  $Q_{ST}$  values against  $F_{ST}$  values from short windows based on population genomic data from the same population samples available from the *Drosophila* Genome Nexus (Lack et al., 2016). For  $F_{ST}$  calculations, we used the same number of inbred genomes that were phenotyped ( $n = 10$ ) and used diversity-scaled windows that averaged ~700 bp in length (specifically, based on 20 non-singleton SNPs from the full Zambia ZI sample).  $F_{ST}$  for individual short windows should have a conservatively higher neutral variance than  $Q_{ST}$  for a polygenic trait.

### 2.3 | Bulk segregant analysis

To ascertain areas of local adaptation responsible for higher ethanol resistance, bulk segregant analysis was performed to detect quantitative trait loci (QTL) (Pool, 2016). Population cages were started from reciprocal crosses between eight inbred parental individuals of low resistant (Zambia) and one each of the more resistant African populations (Ethiopia and South Africa) lines and strains from two French populations (Table S1). From each reciprocal cross, 125 F1 offspring of each sex were used to establish the second generation. These mapping populations for the rest of the (non-overlapping) generations were maintained at ~1200 individuals. The flies were housed in 28 × 14 × 15 cm plastic cages that contain 14 vials with a medium that contains molasses, cornmeal, yeast, agar, and antimicrobial agents at ~20°C. Adult flies were allowed to lay eggs on the food for one week before being removed. The food vials were replaced when adult flies in the cage were 7–10 days old. At the 15th generation, 600 3- to 5-day-old female flies from each population cage were exposed to the 8% ethanol mobility assay described above. Twelve batches of 50 flies were simultaneously placed into twelve inverted 50-ml falcon tubes containing a Kimwipe (resting

on the cap) saturated with 1.5 ml of an aqueous solution containing 8% ethanol and 3% molasses. When flies became immobile, they were collected by removing the bottom cap. The cap was then replaced with a new cap, with a Kimwipe containing the same solution. Female flies were placed into two pools to be sequenced, the 10% least resistant ( $N = 60$ ) representing the first flies to lose mobility, and the 10% most resistant ( $N = 60$ ) representing the last flies still mobile. The remaining flies were discarded.

## 2.4 | Genome preparation

We sequenced the genomes of pooled samples ( $N = 30$  individuals) for the parental lines and two such pools for each of the low and high resistant groups (0%–5% and 5%–10% extremes for each direction, summing to  $N = 60$  total for each extreme). Genomic DNA was obtained using a chloroform extraction and ethanol precipitation protocol. The DNA was fragmented with a Bioruptor sonicator (Diagenode) and paired-end libraries with ~300 bp inserts prepared using NEBNext DNA Library Prep Reagent Set for Illumina (New England Biolabs no. E6000L). Each library's concentration and quality was analyzed with an Agilent 2100 Bioanalyzer (Agilent Technologies, Inc.). The prepared libraries were sequenced at UW-Madison Biotechnology Center on the Illumina HiSeq 2000 platform. Having concluded that the full 10% extremes would best be analyzed together (Pool, 2016), we merged reads from the 0%–5% and 5%–10% pools (similar numbers of reads were obtained from these pools in each case) before proceeding with the analysis.

## 2.5 | Genome alignment

All the raw data that passed the Illumina filters were processed using a Perl-scripted pipeline. Reads from each sequenced genome were mapped to the *D. melanogaster* reference genome (release 5.57) obtained from Flybase ([www.flybase.org](http://www.flybase.org)), with the default parameters in BWA ver. 0.6.2-r126 (Li & Durbin, 2009). Using Stampy ver. 1.0.21 (Lunter & Goodson, 2011), the BAM files were then remapped in order to mirror the *Drosophila* Genome Nexus pipeline used for the parental strain genomes (Lack et al., 2015). With samtools ver. 0.1.18 (Li et al., 2009), reads were filtered for a mapping quality of 20 and for proper pairs. The BAM files were further processed by removing unmapped reads and sorted by coordinate, and PCR duplicates were marked using Picard ver. 1.109 (<http://picard.sourceforge.net>). To improve the alignment around indels, we used GATK ver. 3.2 (McKenna et al., 2010). The average depth of coverage per genome was calculated for the parental lines and the low and high resistance lines (Table S1).

## 2.6 | Quantitative trait locus mapping

The PoPoolation2 ver. 1.201 software package (Kofler et al., 2011) was used to create synchronized mpileup files for the aligned

genomes. For each biallelic SNP, an ancestry difference ( $a_d$ ) was calculated (Bastide et al., 2016), which was calculated as:

$$a_d = (f_H - f_L) / (p_H - p_L) \quad (1)$$

where  $p_H$  is the frequency of parental high resistant allele,  $p_L$  is the low resistant parental allele,  $f_H$  is the frequency of parental high resistant allele in the F12 offspring, and  $f_L$  is the frequency of parental low resistant allele F12 offspring. The five analyzed chromosomal arms (X, 2L, 2R, 3L, and 3R) were divided into windows based on 200 non-singleton SNPs in the Zambia population (Lack et al., 2015), which created 2728, 3131, 2357, 2956, and 2935 windows, respectively, each roughly 8.4-kb in size on average. Only sites that had a parental allele frequency difference of  $\geq 0.25$  were used in the analysis. A simulation-based inference for BSA mapping (SIBSAM) was performed (Pool, 2016) to identify significant QTLs and calculate their confidence intervals and effect sizes. The custom scripts used for SIBSAM can be found at: <http://github.com/JohnEPool/SIBSAM1>. SIBSAM is able to evaluate both primary QTL peaks and flanking secondary QTL peaks, evaluating whether ragged peaks contain significant evidence for more than one QTL. Forward simulations incorporate recombination in multiple individuals for multiple generations, selection on phenotype in the final generation with additivity, plus environmental variance, and then the sampling of sequence reads to obtain  $a_d$ .

## 2.7 | Genetic differentiation and Gene Ontology enrichment analysis

To find evidence of local adaptation and produce a list of candidate genes found within the significant QTLs, window  $F_{ST}$  and maximum SNP  $F_{ST}$  per window (hereafter “SNP  $F_{ST}$ ”), and the haplotype statistic  $\chi_{MD}$  (Lange & Pool, 2016) were analyzed. Pairwise comparisons were made between the low ethanol-resistant population Zambia and each of the high resistance populations. Genomes from Zambia ( $n = 197$ ), South Africa ( $n = 61$ ), Ethiopia ( $n = 68$ ), and France ( $n = 96$ ) were used from the *Drosophila* Genome Nexus (Lack et al., 2015). The  $\chi_{MD}$  compares length of identical haplotype blocks among individuals in one population versus another. The comparisons were made within each of the five chromosomal arms (X, 2L, 2R, 3L, and 3R), which were divided into windows based on SNP density (Lack et al., 2015). The idea behind  $\chi_{MD}$  is that in a recently selected population, longer stretches of identical haplotypes will not have had time for recombination or mutation to break up longer identical tracts. A chromosomal arm quantile outlier approach was used to focus on genes with an extreme population genetic signal. Only windows that were in the top 2.5% quantile in any of the three statistics were classified as outliers. To form an outlier region, a maximum of two non-outlier windows were allowed between two outlier windows. Genes associated with outlier windows (overlapping them or the nearest gene in either direction) were retained for subsequent analysis. The outlier genes identified in significant QTL regions were used for window-based gene ontology (GO) enrichment analysis (as

implemented in Pool et al., 2012) to identify functional categories that differ between low and high resistance populations. A  $p$  value was calculated based on the probability of observing a given number of outlier genes from a GO category.  $p$  values were obtained from permutation in which outlier region was randomly reassigned 10,000 times.

## 2.8 | Genotype–phenotype association testing

Phenotypic data were collected on 51 France inbred strains with sequenced genomes from the Drosophila Genome Nexus genomic resource (Lack et al., 2016). Here, we followed the ethanol assay described above, except that 50 flies from each strain were placed in a single falcon tube and the immobilization times for each individual were recorded (and subsequently averaged for the strain). To capture the variation in ethanol resistance found in the France population, the ethanol concentration used was 18%. The assay was otherwise identical to the inter-population assay described above. Genotype–phenotype associations were analyzed with the R package rrBLUP version 3.1 (Endelman, 2011). Only regions within the QTL peaks of <2 Mb in length in the two France crosses were examined. Within these peaks, only SNPs that had a called allele of >25% and a minor allele frequency >5% among the 51 France strains were analyzed. One thousand permutations of the phenotypic data were used to calculate the significant threshold.

## 2.9 | Simulations of genetic architecture and association testing power

We performed simulations to better understand the genetic architecture of this adaptive trait, using modified versions of SIBSAM scripts. These simulations involved three steps. First, we calibrated the number and strengths of QTLs to match the empirical data from the two France/Zambia crosses. To do this, we analyzed a range of values for three different parameters: (1) the number of detected QTLs (10, 20, 30, 40, and 50); (2) environmental variance, how much of the phenotypic trait is caused by factors other than genetic factors (0.5, 0.6, 0.7, 0.8, and 0.9); and (3) QTL strength. Here, we used a gamma distribution (shape parameter 0.5, 1, 2, 4, and 8, and scale parameter fixed at 1 because it is not relevant in this relative context). We performed 10,000 simulation replicates for all parameter combinations for both France crosses.

For these comparisons between empirical and simulated QTL mapping data, we used a simplified set of QTL criteria in order to avoid the computational infeasible requirement of running full SIBSAM inference to identify significant QTLs from each simulated replicate. Specifically, we defined QTLs as having ancestry difference greater than 0.16. The flanking secondary QTL peaks were defined as having a secondary deviation (the magnitude of ancestry difference recovery from a local valley; Pool, 2016) greater than

0.16. These criteria were chosen to largely recapitulate the same QTLs found to be significant from the empirical data.

We then looked at four summary statistics: the mean ancestry difference and its standard deviation across all windows, the mean QTL peak height, and the number of QTLs. We calculated the relative error sum of all the replicates for each combination using the empirical values: mean peak height 0.256, number of QTLs 18, mean ancestry difference 0.041, and standard deviation 0.083. The parameter combination with the lowest mean relative error sum was then used to perform the next step to calibrate the frequency of all QTLs. To analyze how well the top model performed, we performed bootstrapping among the 10,000 replicate simulations from both the top model and one of the other 125 parameter combinations, monitoring the proportion of 10,000 bootstrap replicates in which the top model still had a lower error.

Next, we wanted to see which QTL frequency along with fixed parameters from the previous step would match the proportion of empirical QTL peaks overlapping between the two crosses. We ran 10,000 replicates of each of the different frequency values: 0.05, 0.10, 0.15, 0.20, 0.25, 0.30, 0.35, 0.40, 0.45, and 0.50—where these values indicated the probability that a given QTL in one cross would also be present in the second cross. Each QTL was considered to overlap if its peak fell within a simplified “QTL region” from another cross (defined as the area in which a peak exceeds an ancestry difference or secondary deviation of 0.16).

Finally, we estimated association testing power for different scenarios involved with sample sizes and allele frequencies. We used sample sizes of 50, 100, 200, 500, and 1000 and allele frequencies of 5%, 10%, 20%, 30%, 40%, and 50%. For each parameter combination, we created genotypes and phenotypes. Genotypes were assigned by first determining if the individual was either homozygous or heterozygous based on empirical residual heterozygosity levels of 35% in the France population (Lack et al., 2016). If an individual was homozygous, then they had one draw of getting either the ancestral or derived allele and it was added twice. If an individual was heterozygous, then they had two independent draws of getting either the ancestral or derived allele. We translated the QTL frequency identified in the previous into allele frequency using the equation:

$$0.65q + 0.35(q_2 + 2q(1 - q)) \quad (2)$$

where 0.65 is the frequency of being homozygous and 0.35 is the frequency of being heterozygous and  $q$  is the frequency of the derived allele. Once the genotypes were established, the phenotypic trait values could be assigned. If the individual genotype at the locus was homozygous for the derived allele, then the full QTL strength was added. If the genotype was heterozygous, then half the QTL strength was added, and if the genotype was homozygous for the ancestral allele, then nothing was added to the trait value (i.e., assuming additivity). In light of the replicated phenotyping of individuals from each inbred line, no environmental variance was simulated. We performed 1000 simulated genotype–phenotype association replicates for each parameter combination and recorded the

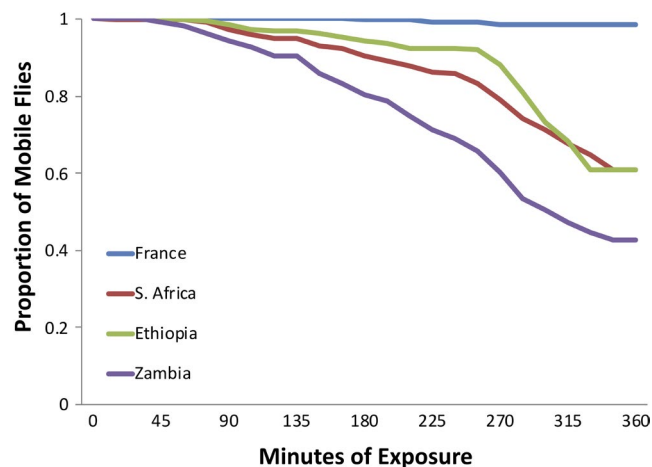
proportion of total alleles that exceeded the empirical permutation  $-\log(P)$  threshold of 6.17.

### 3 | RESULTS

#### 3.1 | Population differences in ethanol resistance

We performed a phenotypic assay of adult exposure to 8% ethanol vapor, using flies from ten independent inbred strains from each of four population samples. This assay revealed variation in ethanol resistance among the populations studied (Figure 1). As expected, the France population (David et al., 1986) had the highest resistance with only 1% of the individuals immobile after six hours of exposure (mean 379.5 min of mobility, SD 16.7). The Zambia population from the ancestral range had the lowest resistance; after six hours, there was 57% immobility (mean 267.4 min, SD 93.5). The two high altitude African populations, Ethiopia and South Africa, were not as resistant as the France population, but more resistant than Zambia. Both South Africa (mean 307.5 min, SD 78.5) and Ethiopia (mean 318.5 min, SD 63.7) had ~40% of the individuals immobile at 6 h. We performed an individual-based ANOVA between each of the more resistant population (Ethiopia, S. Africa, and France) and the least resistant Zambian population, using the 6-h data. We found that in spite of substantial within-population variation among strains, the three former populations were each significantly more resistant than the Zambian population ( $p \leq 3.21 \times 10^{-11}$  in each case; Table S2).

We used  $Q_{ST}$  values (Lande, 1992; Miller et al., 2008; Spitze, 1993) to quantify phenotypic differentiation between the three ethanol-resistant populations and the least tolerant Zambian population. For all three comparisons, the  $Q_{ST}$  values were much greater than the  $F_{ST}$  values. The France–Zambia comparison had a



**FIGURE 1** *Drosophila melanogaster* populations differ strongly in their resistance to concentrated ethanol vapor. The percentage of flies that remained mobile after being exposed to 8% ethanol is shown across a 6-h interval. Here, Zambia represents a population from the ancestral range of the species, South Africa and Ethiopia derive from relatively cooler high elevation locations, and France represents a cooler temperate latitude

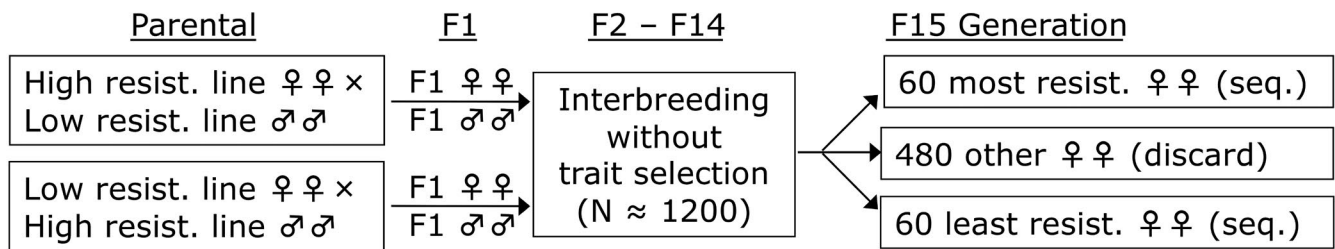
$Q_{ST}$  value of 0.548 compared with a mean  $F_{ST}$  value of 0.226, with 0.015% of genomic windows having an  $F_{ST} > Q_{ST}$ . The Ethiopia–Zambia comparison had a  $Q_{ST}$  value of 0.216 compared with  $F_{ST}$  value of 0.137, with 0.16% of genomic windows having an  $F_{ST}$  greater than  $Q_{ST}$ . The South Africa–Zambia comparison had a  $Q_{ST}$  value of 0.115 compared with  $F_{ST}$  value of 0.024, with 0.027% of genomic windows have an  $F_{ST}$  greater than  $Q_{ST}$ . If genetic differentiation across most of the genome is primarily the result of neutral processes, then high values of  $F_{ST}$  or  $Q_{ST}$  that fall outside the genomic distribution of  $F_{ST}$  may reflect local adaptation rather than population history alone. Given that ethanol resistance  $Q_{ST}$  values are indeed greater than the vast majority of  $F_{ST}$  values, it appears that this trait has evolved under the influence of local adaptation, acting either upon this trait directly or upon pleiotropically correlated traits.

#### 3.2 | Quantitative trait locus mapping

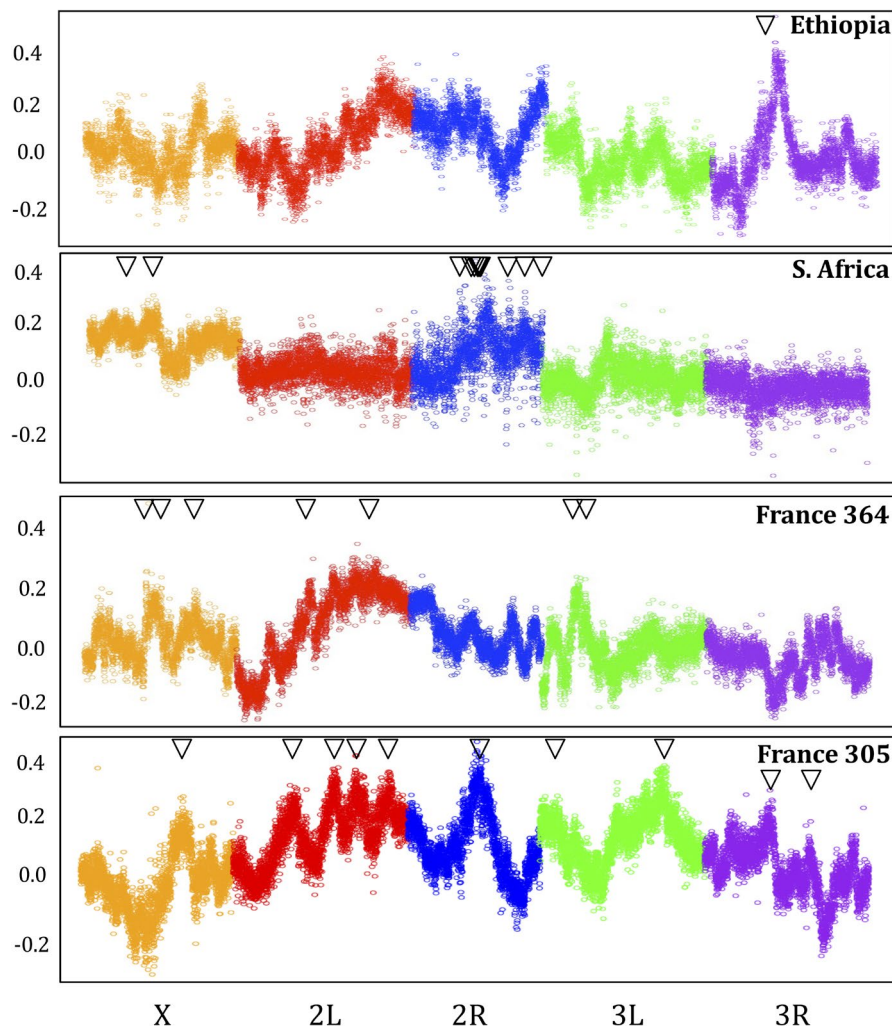
We performed QTL mapping using four different between-population crosses using individual inbred strains, each of which involved the low-resistance Zambia population. Of the higher resistance parental strains, two were independent France strains, and one each was from the African high altitude populations, Ethiopia and South Africa. We allowed offspring of reciprocal crosses to interbreed without selection at a fairly large population size ( $N \approx 1200$ ) until the 15th generation, at which time 600 adult females were exposed to ethanol vapor and the top and bottom 10% of individuals were isolated and subject to pooled genomic sequencing (Figure 2; *Materials and Methods*). Primary and secondary QTL peaks, along with their estimated effect sizes and genomic confidence intervals, were then identified using SIBSAM (Pool, 2016).

The four mapping crosses revealed a total of 32 significant peaks (Figure 3; Table S3). Whereas the Ethiopia cross had just three significant QTLs with estimated effect sizes between 15% and 20%, the South Africa cross had a total of 12 significant peaks, ten of which were on chromosome arm 2R and two on the X chromosome, and these 12 QTLs had estimated effects sizes between 7% and 13%. Between the two France crosses there were 17 peaks, ten from the cross involving strain FR305N and seven for FR364N, which collectively ranged in estimated effect size from 6% to 27%. Encouragingly, the highest peaks in each cross were estimated to have narrow genomic confidence intervals (Table S3). We note that under some scenarios, effect size may be overestimated (Pool, 2016), and consistent with this possibility, the point estimates of QTL effect sizes sum to more than 100% for three out of four crosses (Table S3).

Overlap between QTL peaks may occur by chance or due to a shared genetic basis of ethanol resistance differences between crosses. Between the two France crosses, there were six regions where QTL peaks overlapped with genomic confidence intervals involving a total of 6 out of the 17 QTLs (Figure 4). In a few cases, overlapping QTLs were found between crosses from different populations. Ethiopia shared two distinct QTLs with each of the

**BULK SEGREGANT ANALYSIS (BSA):**

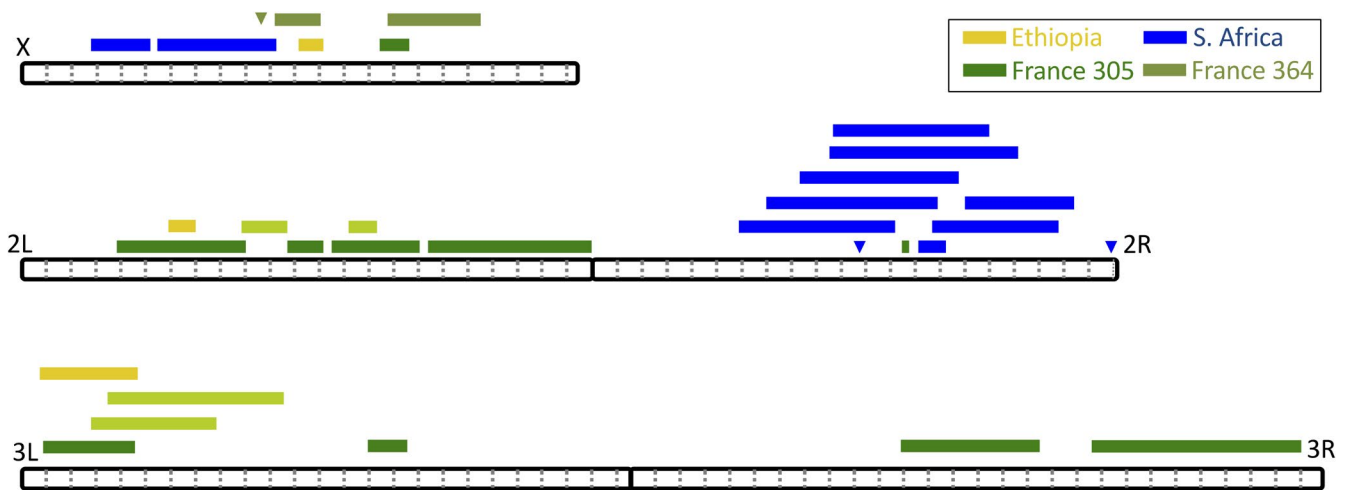
**FIGURE 2** The bulk QTL mapping experimental design is illustrated. Population cages were started from reciprocal crosses between eight inbred parental individuals from a single low-resistance Zambia strain and an equal number of individuals from a single strain from one of the more resistant populations (Ethiopia, France, or South Africa). 125 F1 offspring from each reciprocal cross were combined, and the mapping population was allowed to interbreed with a population size of roughly 1200 until the 15th generation. At that point, females were tested for ethanol resistance, and the first 10% and last 10% to become immobile were isolated for pooled sequencing



**FIGURE 3** Significant QTL peaks for Ethiopia, South Africa, and France crosses are shown, based on bulk segregant mapping from 15th generation cross offspring. A point for each ~8 kb window corresponds to the average difference in the frequency of the resistant parental strain's allele between the high- and low-resistance F15 pools (i.e., "ancestry difference," y-axis). Significant QTLs are denoted with an arrow. The South Africa cross includes a total of 10 significant QTLs on chromosome arm 2R. The significance threshold for primary peaks is approximately 0.16

France crosses, while South Africa shared one QTL with FR364N. The two high-altitude populations, Ethiopia and South Africa, did not share any peaks. Hence, while there is some unconfirmed

potential for genetic parallelism between ethanol resistance in different *D. melanogaster* populations, most QTLs tend to be unique between a given pair of crosses—even when two crosses involve the



**FIGURE 4** QTL locations vary within and between ethanol-resistant populations. The locations of significant QTLs on the five euchromatic chromosome arms of *D. melanogaster* are shown. The colors indicate ethanol resistance mapping crosses involving Ethiopia, South Africa, and France lines 305 and 364. The width of each box indicates the 90% C.I. of each QTL. Intervals that are <10 kb in width are marked with triangles. Dotted gray lines indicate Mb increments

same France and Zambia populations. While chance false-positive and false-negative results may contribute to differences in QTL detection, distinct genetic paths to ethanol resistance in different populations, as well as genetic heterogeneity in the architecture of ethanol resistance within populations, may contribute to these results as well, as further explored below.

### 3.3 | Potential targets of local adaptation within QTL regions

Strong differences in genetic variation between the least resistant Zambia population and one of the more resistant Ethiopia, South Africa, and France populations may signify genes subject to local adaptation, and some of these signals could relate to the trait in question. Therefore, to identify possible candidate genes for ethanol resistance evolution within the significant QTLs, we used three population genetic statistics, window  $F_{ST}$ , maximum SNP  $F_{ST}$  within a window, and the haplotype statistic  $\chi_{MD}$  (Table S4). These statistics may have differing power to detect local adaptation depending on whether selective sweeps are complete or incomplete, or hard versus soft (Lange & Pool, 2016). We used a quantile approach focusing on regions that had one of the three statistics with a quantile below 0.025 (Table S5). This analysis yielded both genes with known functions that may relate to our trait, and genes with no such known functions. While any of these genes might contribute to ethanol resistance evolution, we mention below a few plausible candidates.

Within the South Africa QTLs, peaks on chromosome arms X and 2R each have one outlier redox gene, *Pp2C1* and *Nox*, respectively. Genes involved in regulating oxidative stress have previously been implicated in *Drosophila* ethanol resistance (Awofala et al., 2012). Of potential relevance in light of our aerosol ethanol assay, several genes involved in the development of chitin also have population

genetic signals: *ovo*, *mgl*, and *CG1367*. Potential candidate genes found in Ethiopia QTLs include the following: *Shab* and *Teh2* (ion channels), and *m* (cuticle development).

Notably in one of the France crosses, *Adh* was identified as a population genetic outlier in a QTL region. We note that the Fast allele associated with greater ethanol resistance is at a 88% frequency in our France population sample, compared to 0% in Zambia. Other genes found in one of the two France crosses included several potentially involved in alcohol metabolism (*CG5065*, *CG6650*, *CG8303*, *CG9521*, *CG13091*, *CG15601*, *CG43658*, *Pis*), as well as ion channels (*para*, *ppk*, *sh*) and other genes involved in neurotransmission (*be*, *CG33639*). Diverse aspects of nervous system function have previously been linked to alcohol resistance (Morozova et al., 2015; Park et al., 2017).

Between the two France crosses, shared candidate genes included the following: *CG45065* (alcohol metabolism), *CG9503* (choline/aldehyde metabolism), *bgm* and *pgdy* (fatty acid metabolism), *hiw* (synapse organization), and *eag* (ion channel, response to ether). South Africa and FR364N had two candidate genes of interest, *CG32698* (carbonate dehydratase) and *CG1986* (lipase). Lipid levels are known to influence ethanol resistance (Geer et al., 1991; Lieber & Savolainen, 1984).

### 3.4 | Gene Ontology enrichment

While local adaptation outliers within QTLs represent promising candidates for causative loci, our suspicion would be strengthened if particular types of genes recurrently appeared in the intersection of our mapping and population genetic analyses. Therefore, as a hypothesis-generating exercise, we conducted a GO enrichment analysis on the set of genes both located within a QTL region from any of our crosses and also associated with a population genetic outlier region for that same resistant population. Alcohol metabolism



genes were enriched in this analysis ( $p = .00356$ ; Table S6). The categories showing the strongest enrichment ( $p$  values below 0.001) corresponded to functions previously linked to ethanol response: spindle localization (Hass et al., 2019), sterol biosynthesis (Mo et al., 2019; Stanley et al., 2010), and microvillus membrane (Bjorkman & Jessop, 1994). Other enriched categories related to the perception of sound and light, cuticle development, response to hypoxia, histone H4 acetylation (Ghezzi et al., 2013), and zinc transport (Zhao et al., 2009).

### 3.5 | Genotype–phenotype association testing

Given that the genetic architecture of ethanol resistance in the France population appears to be genetic variable, a complementary approach to identify genes within QTLs responsible for trait evolution could be to look for genotype–phenotype associations among France strains. We collected phenotype data from 51 France inbred lines with previously sequenced genomes (Lack et al., 2016) in order to perform genotype–phenotype association testing. This sample size would not be adequate for genome-wide association testing, and so we restricted our focus to France QTL regions of less than 2 Mb in length. We performed this analysis either on all SNPs within these QTLs (120,243 SNPs), or focusing more specifically on SNPs within-population genetic outliers windows (9480 SNPs). Genome-wide significance, assessed via permutations, was not reached by any SNP in either analysis (Table S7). From the more inclusive analysis, the highest marker had a  $-\log p$  value of 4.43, whereas the permutation significance threshold was 6.17. From the population genetic outlier analysis, the highest marker had a  $-\log p$  value of 3.56 with permutation significance threshold of 4.44.

### 3.6 | Simulation-driven investigation of genetic architecture and association testing power

We then considered which genetic architectures our QTL mapping data might provide evidence for, and whether they might account for

our negative association testing results. Although full model inference of adaptive evolution at the genetic level is beyond the scope of the present study, we conducted an exploratory simulation analysis in three stages, focusing on the two France crosses.

First, we wanted to assess the number and strength of QTLs that our mapping data were most consistent with. Our simulations used a modified version of SIBSAM, which simulates the full mapping experiment (including recombination, phenotypic selection, and sequencing read sampling). We varied the number of QTLs present in each cross, their distribution of effect sizes as a function of the gamma distribution shape parameter, and the proportion of trait variation contributed by environmental/random effects rather than these QTLs. And we quantified properties of QTL peaks and genome-wide ancestry in the simulated data and compared it with our empirical observations using mean relative error. The parameter combination with the lowest average mean relative error was 10 QTLs per cross, a gamma shape parameter of 4, and 70% environmental variation (Figure 5; Table S7). An otherwise identical parameter combination with a gamma shape parameter of 8 matched the empirical data almost equally as well, and so we chose an intermediate shape parameter of six in further analyses. Parameter combinations involving a wide range of QTL numbers and gamma shape values were non-significantly worse than the above combination. Thus, larger data sets (especially with greater numbers of crosses) will be needed to make formal inferences of this type about the genetic architecture of a trait's adaptive evolution (Table S8).

Second, we assessed whether the degree of QTL overlap between the two France crosses provides information about the frequency of ethanol resistance alleles in this population. The QTL frequency that resulted in the average overlap of peaks that best matched the empirical data was 5% (Table S9), which resulted in ~54% overlap compared to the empirical ~44%. However, only the highest frequency values (90% and above) had confidence intervals that marginally excluded the empirical frequency. Here again, larger data sets with more independent crosses might help us to gain further resolution about the frequencies of adaptive variants.

We therefore investigated a wide range of frequencies (5% to 50%) in assessing the power of our association testing analysis. The

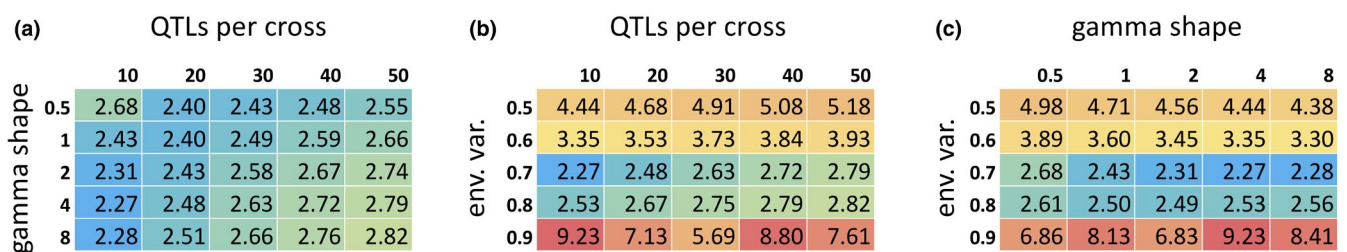
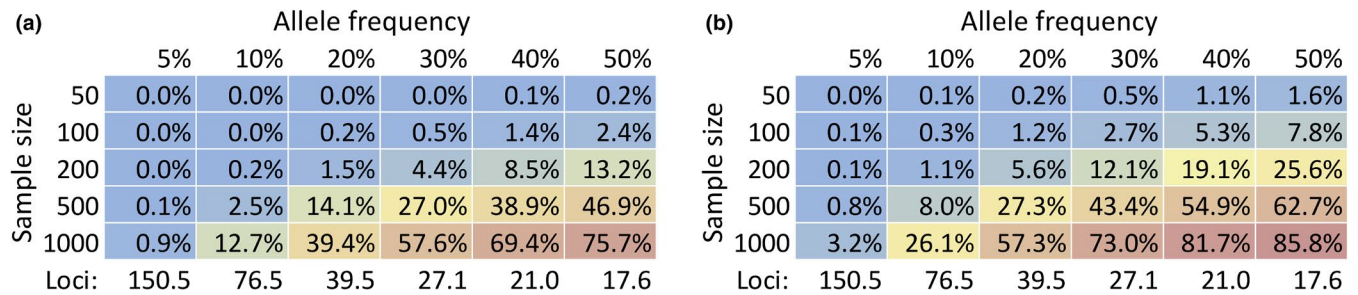


FIGURE 5 QTL simulations provide limited resolution on parameters underlying the evolution of ethanol resistance in the France population. Heat maps depict the mean relative error between empirical data (from France crosses) and selected simulated data sets, based on the QTL mapping summary statistics compared (see *Materials and Methods*). These plots each fix one of the three parameters with its value from the best-matching parameter combination: (a) shows the case of an environmental variance of 0.7 (70% of individual trait variation explained by non-genetic factors), (b) gamma shape parameter of 4 (i.e., shifted toward larger effect loci relative to an exponential distribution), and (c) 10 distinct QTLs present in each of the two crosses



**FIGURE 6** Larger sample sizes would be needed to perform powerful genotype–phenotype association within QTL regions, even if restricted to local adaptation outlier windows. Heat maps depict power to detect genotype–phenotype associations after multiple testing correction (the proportion of SNPs meeting an analysis-wide significance threshold), as a function of the frequency of causative variants and the sample size of individuals/strains. These results are based on simulations (using a gamma shape parameter of six for the distribution of effect sizes) along with the  $p$  value thresholds identified from the empirical analysis. (a) corresponds to the scenario in which full QTL regions were tested, while (b) corresponds to the scenario in which only population genetic outlier windows within QTL regions were tested. The mean population-wide numbers of loci that each simulation scenario required are also indicated below each panel

power analysis revealed that there is little power to detect causative SNPs that segregate at lower frequencies. Small-to-moderate population sizes ( $n = 50, 100,$  and  $200$ ) had low power to detect SNPs at any frequency (Figure 6). It was not until QTL frequency reached 50% that there was a greater than 10% detection power. Even when there was a large number of individuals used, for example,  $n = 1000$ , only ~0.9% of the causative SNPs at a 5% QTL frequency met the empirical threshold. Only once QTL frequency reached 30% was there greater than 50% detection rate. However, the detection power did improve when the lower outlier region threshold was used. The small population sizes had a greater than 10% detection power when QTL frequency reached 30%. The detection power at 5% QTL frequency for large population size of 1000 improved to ~3% and had a greater than 50% detection rate when QTL frequency reached 20%. Hence, significantly larger sample sizes would be needed to identify variants underlying polymorphic architectures of adaptive evolution, unless the number of suspected variants (and hence the multiple testing burden) could be further reduced.

## 4 | DISCUSSION

We have shown that there is a range of ethanol resistance found in wild populations of *D. melanogaster*. The Zambia population, which inhabits the species' ancestral range (Sprengelmeyer et al., 2020), is the least resistant with the recently diverged populations becoming more resistant. We show that elevated ethanol resistance is present not only at high latitudes, but also at high altitudes within Africa. When ethanol resistance for either a France, Ethiopia, or South Africa population is compared against Zambia, phenotypic differentiation ( $Q_{ST}$ ) strongly exceeds genetic differentiation ( $F_{ST}$ )—suggesting that local adaptation has operated on this or a pleiotropically correlated trait. In agreement with other studies (David & Bocquet, 1975), we found the higher latitude France population to be highly resistant. Interestingly, *Aldh* (Fry et al., 2008) is contained in one of the France mapping crosses and *Adh* (Cohen & Graf, 1985) is contained in the other. However, only *Adh* has a population genetic signal of local

adaptation based on our criteria. Both genes have been implicated in the latitudinal cline of increased ethanol resistance.

The BSA performed on the four different crosses revealed 32 significant QTLs with the largest estimated effect size for each cross between 12% and 27%. These data taken together suggest that ethanol resistance is moderately polygenic with moderate to large effect QTLs present (whereas smaller QTLs may elude our detection power; Pool, 2016). We found that there are no QTLs overlapping between all three high resistant populations. However, each of the high altitude populations shares QTLs with the France strains, whereas the two high altitude populations, South Africa and Ethiopia, do not have any QTLs in common with each other. Given the unequal levels of ethanol resistance between populations (Figure 1), some differences in genetic architecture would of course be expected.

In interpreting the observed levels of QTL overlap between populations, it is important to keep in mind that even between two crosses from the same resistant population (France), QTL overlap was modest. Of the seventeen significant QTL peaks between the two France crosses, they shared only six QTLs (and even some of these could reflect random overlap in light of the QTL sizes). For example, the strongest QTL in either France cross with an estimated effect size of 27% from FR305N on chromosome arm 3L is completely missing in FR364N. Our experiment should have very high power to detect a QTL with an effect size even roughly this large if it existed in a second cross (Pool, 2016). Those results suggest a genetically heterogeneous architecture of ethanol resistance evolution not only between populations but also within a resistant population. Notably, very similar patterns, both within and between populations, were also observed in similar experiments focused on the evolution of melanism within this species (Bastide et al., 2016). The implication that causative variants have not been fixed has multiple potential explanations, including ongoing adaptation, balancing selection, or that a trait has reached a new optimum value or exceeded a new threshold value. Persistent variability in the genetic basis of an adaptive trait might be expected when populations start with abundant standing genetic variation, as might be expected for *D. melanogaster*.

Still, our simulation results clearly show that larger experiments will be needed to gain quantitative resolution on key parameters that describe the genetic architecture of adaptive evolution. Studies with larger numbers of QTL mapping crosses may allow clearer estimation of the number of QTLs per cross, the distribution of QTL effect sizes, and the frequencies of causative variants in an evolved population. The utility of genotype–phenotype association testing will depend on either much larger population samples of sequenced inbred line genomes becoming available, or else further progress in restricting the number of SNPs to be tested. Candidate SNPs might be further limited by more precise QTL mapping (more generations, more individuals), functional genomic data, or complementary population genomic analysis such as genotype–environment association. Regardless of the strategy chosen, the results of this study and others suggest that future efforts to understand the genetic basis of adaptive trait evolution should allow for the likelihood of a genetically variable trait architecture among individuals such as that detected here.

## ACKNOWLEDGMENTS

This project benefitted from funding from NIH grants R01 GM111797, R35 GM136306, and T32 GM007133, and NSF grant DEB 1754745. We would like to thank Justin Lack, Jeremy Lange, and Tiago Ribeiro for their help with bioinformatics. We also thank the UW-Madison Center for High Throughput Computing (CHTC) for computational resources and assistance.

## CONFLICT OF INTEREST

The authors declare that no conflict of interest exists.

## AUTHOR CONTRIBUTION

**Quentin D. Sprengelmeyer:** Conceptualization (supporting); Data curation (lead); Formal analysis (lead); Funding acquisition (supporting); Investigation (lead); Methodology (supporting); Project administration (supporting); Visualization (lead); Writing-original draft (lead); Writing-review & editing (supporting). **John E. Pool:** Conceptualization (lead); Formal analysis (supporting); Funding acquisition (lead); Investigation (supporting); Methodology (lead); Project administration (lead); Supervision (lead); Visualization (supporting); Writing-original draft (supporting); Writing-review & editing (lead).

## DATA AVAILABILITY STATEMENT

Raw sequence read data are available in the NIH SRA under accession number PRJNA686135.

## ORCID

John E. Pool  <https://orcid.org/0000-0003-2968-9545>

## REFERENCES

- Awofala, A. A., Davies, J. A., & Jones, S. (2012). Functional roles for redox genes in ethanol sensitivity in *Drosophila*. *Functional & Integrative Genomics*, 12(2), 305–315. <https://doi.org/10.1007/s10142-012-0272-5>
- Barghi, N., Hermisson, J., & Schlötterer, C. (2020). Polygenic adaptation: A unifying framework to understand positive selection. *Nature Reviews Genetics*, 21(12), 769–781. <https://doi.org/10.1038/s41576-020-0250-z>
- Barghi, N., & Schlötterer, C. (2020). Distinct patterns of selective sweep and polygenic adaptation in evolve and resequence studies. *Genome Biology and Evolution*, 12(6), 890–904. <https://doi.org/10.1093/gbe/evaa073>
- Barghi, N., Tobler, R., Nolte, V., Jakšić, A. M., Mallard, F., Otte, K. A., Dolezal, M., Taus, T., Kofler, R., & Schlötterer, C. (2019). Genetic redundancy fuels polygenic adaptation in *Drosophila*. *PLoS Biology*, 17(2), e3000128. <https://doi.org/10.1371/journal.pbio.3000128>
- Bastide, H., Lange, J. D., Lack, J. B., Yassin, A., & Pool, J. E. (2016). A variable genetic architecture of melanic evolution in *Drosophila melanogaster*. *Genetics*, 204(3), 1307–1319.
- Bastide, H., Yassin, A., Johanning, E. J., & Pool, J. E. (2014). Pigmentation in *Drosophila melanogaster* reaches its maximum in Ethiopia and correlates most strongly with ultra-violet radiation in sub-Saharan Africa. *BMC Evolutionary Biology*, 14, 179. <https://doi.org/10.1186/s12862-014-0179-y>
- Berg, J. J., Harpak, A., Sinnott-Armstrong, N., Joergensen, A. M., Mostafavi, H., Field, Y., Boyle, E. A., Zhang, X., Racimo, F., Pritchard, J. K., & Coop, G. (2019). Reduced signal for polygenic adaptation of height in UK Biobank. *Elife*, 8, e39725. <https://doi.org/10.7554/eLife.39725>
- Bjorkman, D. J., & Jessop, L. D. (1994). Effects of acute and chronic ethanol exposure on intestinal microvillus membrane lipid composition and fluidity. *Alcoholism: Clinical and Experimental Research*, 18(3), 560–565. <https://doi.org/10.1111/j.1530-0277.1994.tb00910.x>
- Chakraborty, M., & Fry, J. D. (2016). Evidence that environmental heterogeneity maintains a detoxifying enzyme polymorphism in *Drosophila melanogaster*. *Current Biology*, 26(2), 219–223. <https://doi.org/10.1016/j.cub.2015.11.049>
- Chandler, L. J., Harris, R. A., & Crews, F. T. (1998). Ethanol tolerance and synaptic plasticity. *Trends in Pharmacological Sciences*, 19(12), 491–495. [https://doi.org/10.1016/S0165-6147\(98\)01268-1](https://doi.org/10.1016/S0165-6147(98)01268-1)
- Cohan, F. M., & Graf, J.-D. (1985). Latitudinal Cline in *Drosophila melanogaster* for knockdown resistance to ethanol fumes and for rates of response to selection for further resistance. *Evolution*, 39(2), 278–293.
- Cowmeadow, R. B., Krishnan, H. R., & Atkinson, N. S. (2005). The slowpoke gene is necessary for rapid ethanol tolerance in *Drosophila*. *Alcoholism: Clinical and Experimental Research*, 29(10), 1777–1786. <https://doi.org/10.1097/01.alc.0000183232.56788.62>
- David, J. R., & Bocquet, C. (1975). Similarities and differences in latitudinal adaptation of two *Drosophila* sibling species. *Nature*, 257(5527), 588–590. <https://doi.org/10.1038/257588a0>
- David, J. R., Bocquet, C., Arens, M. F., & Fouillet, P. (1976). Biological role of alcohol dehydrogenase in the tolerance of *Drosophila melanogaster* to aliphatic alcohols: Utilization of an ADH-null mutant. *Biochemical Genetics*, 14(11–12), 989–997. <https://doi.org/10.1007/BF00485131>
- David, J. R., Mercot, H., Capy, P., McEvey, S. F., & Van Herrewege, J. (1986). Alcohol tolerance and Adh gene frequencies in European and African populations of *Drosophila melanogaster*. *Génétique, Sélection, Évolution*, 18(4), 405. <https://doi.org/10.1186/1297-9686-18-4-405>
- Dittmar, E. L., Oakley, C. G., Conner, J. K., Gould, B. A., & Schemske, D. W. (2016). Factors influencing the effect size distribution of adaptive substitutions. *Proceedings of the Royal Society B: Biological Sciences*, 283(1828), 20153065. <https://doi.org/10.1098/rspb.2015.3065>
- Endelman, J. B. (2011). Ridge regression and other kernels for genomic selection with R package rrBLUP. *The Plant Genome*, 4(3), 250–255. <https://doi.org/10.3835/plantgenome2011.08.0024>

- Faria, V. G., Martins, N. E., Schlötterer, C., & Sucena, É. (2018). Readapting to DCV infection without wolbachia: frequency changes of drosophila antiviral alleles can replace endosymbiont protection. *Genome Biology and Evolution*, 10(7), 1783–1791. <https://doi.org/10.1093/gbe/evy137>
- Fisher, R. A. (1930). *The genetical theory of natural selection: A complete. Variorum* edn. Oxford University Press.
- Fochler, S., Morozova, T. V., Davis, M. R., Gearhart, A. W., Huang, W., Mackay, T. F. C., & Anholt, R. R. H. (2017). Genetics of alcohol consumption in *Drosophila melanogaster*. *Genes, Brain and Behavior*, 16(7), 675–685.
- Fry, J. D. (2014). Mechanisms of naturally evolved ethanol resistance in *Drosophila melanogaster*. *Journal of Experimental Biology*, 217(22), 3996–4003.
- Fry, J. D., Donlon, K., & Saweikis, M. (2008). A worldwide polymorphism in aldehyde dehydrogenase in *Drosophila melanogaster*: evidence for selection mediated by dietary ethanol. *Evolution*, 62(1), 66–75. <https://doi.org/10.1111/j.1558-5646.2007.00288.x>
- Fry, J. D., & Saweikis, M. (2006). Aldehyde dehydrogenase is essential for both adult and larval ethanol resistance in *Drosophila melanogaster*. *Genetical Research*, 87(2), 87–92.
- Geer, B. W., Heinstra, P. W., & McKechnie, S. W. (1993). The biological basis of ethanol tolerance in *Drosophila*. *Comparative Biochemistry and Physiology Part B: Comparative Biochemistry*, 105(2), 203–229. [https://doi.org/10.1016/0305-0491\(93\)90221-P](https://doi.org/10.1016/0305-0491(93)90221-P)
- Geer, B. W., McKechnie, S. W., Heinstra, P. W. H., & Pyka, M. J. (1991). Heritable variation in ethanol tolerance and its association with biochemical traits in *Drosophila melanogaster*. *Evolution*, 45(5), 1107–1119.
- Ghezzi, A., Krishnan, H. R., Lew, L., Prado, F. J. III, Ong, D. S., & Atkinson, N. S. (2013). Alcohol-induced histone acetylation reveals a gene network involved in alcohol tolerance. *PLoS Genetics*, 9(12), e1003986. <https://doi.org/10.1371/journal.pgen.1003986>
- Griffin, P. C., Hangartner, S. B., Fournier-Level, A., & Hoffmann, A. A. (2017). Genomic trajectories to desiccation resistance: convergence and divergence among replicate selected *Drosophila* lines. *Genetics*, 205(2), 871–890.
- Gudbjartsson, D. F., Walters, G. B., Thorleifsson, G., Stefansson, H., Halldorsson, B. V., Zusmanovich, P., Sulem, P., Thorlacius, S., Gylfason, A., Steinberg, S., Helgadóttir, A., Ingason, A., Steinthorsdóttir, V., Olafsdóttir, E. J., Olafsdóttir, G. H., Jonsson, T., Borch-Johnsen, K., Hansen, T., Andersen, G., ... Stefansson, K. (2008). Many sequence variants affecting diversity of adult human height. *Nature Genetics*, 40(5), 609–615. <https://doi.org/10.1038/ng.122>
- Haas, R., Horev, G., Lipkin, E., Kesten, I., Portnoy, M., Buhnik-Rosenblau, K., Soller, M., & Kashi, Y. (2019). Mapping ethanol tolerance in budding yeast reveals high genetic variation in a wild isolate. *Frontiers in Genetics*, 10, 998. <https://doi.org/10.3389/fgene.2019.00998>
- Höllinger, I., Pennings, P. S., & Hermisson, J. (2019). Polygenic adaptation: From sweeps to subtle frequency shifts. *PLoS Genetics*, 15(3), e1008035. <https://doi.org/10.1371/journal.pgen.1008035>
- John, S., & Stephan, W. (2020). Important role of genetic drift in rapid polygenic adaptation. *Ecology and Evolution*, 10(3), 1278–1287. <https://doi.org/10.1002/ece3.5981>
- Kacsoh, B. Z., Lynch, Z. R., Mortimer, N. T., & Schlenke, T. A. (2013). Fruit flies medicate offspring after seeing parasites. *Science* 339(6122), 947–950.
- Kimura, M. (1983). *The neutral theory of molecular evolution*. Cambridge University Press.
- Kofler, R., Pandey, R. V., & Schlötterer, C. (2011). PoPoolation2: identifying differentiation between populations using sequencing of pooled DNA samples (Pool-Seq). *Bioinformatics*, 27(24), 3435–3436. <https://doi.org/10.1093/bioinformatics/btr589>
- Lack, J. B., Cardeno, C. M., Crepeau, M. W., Taylor, W., Corbett-Detig, R. B., Stevens, K. A., Langley, C. H., & Pool, J. E. (2015). The *Drosophila* genome nexus: a population genomic resource of 623 *Drosophila melanogaster* genomes, including 197 from a single ancestral range population. *Genetics*, 199(4), 1229–1241.
- Lack, J. B., Lange, J. D., Tang, A. D., Corbett-Detig, R. B., & Pool, J. E. (2016). A thousand fly genomes: an expanded *Drosophila* genome nexus. *Molecular Biology and Evolution*, 33(12), 3308–3313.
- Lande, R. (1992). Neutral theory of quantitative genetic variance in an island model with local extinction and colonization. *Evolution*, 46(2), 381–389.
- Lange, J. D., & Pool, J. E. (2016). A haplotype method detects diverse scenarios of local adaptation from genomic sequence variation. *Molecular Ecology*, 25(13), 3081–3100. <https://doi.org/10.1111/mec.13671>
- Li, H., & Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*, 25(14), 1754–1760. <https://doi.org/10.1093/bioinformatics/btp324>
- Lieber, C. S., & Savolainen, M. (1984). Ethanol and lipids. *Alcoholism: Clinical and Experimental Research*, 8(4), 409–423. <https://doi.org/10.1111/j.1530-0277.1984.tb05689.x>
- Lunter, G., & Goodson, M. (2011). Stampy: A statistical algorithm for sensitive and fast mapping of Illumina sequence reads. *Genome Research*, 21(6), 936–939. <https://doi.org/10.1101/gr.111120.110>
- McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., Garimella, K., Altshuler, D., Gabriel, S., Daly, M., & DePristo, M. A. (2010). The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Research*, 20(9), 1297–1303. <https://doi.org/10.1101/gr.107524.110>
- McKenzie, J. A., & Parsons, P. A. (1972). Alcohol tolerance: An ecological parameter in the relative success of *Drosophila melanogaster* and *Drosophila simulans*. *Oecologia*, 10(4), 373–388. <https://doi.org/10.1007/BF00345738>
- Merçot, H., Defaye, D., Capy, P., Pla, E., & David, J. R. (1994). Alcohol tolerance, Adh activity, and ecological niche of *Drosophila* species. *Evolution*, 48(3), 746–757.
- Milan, N. F., Kacsoh, B. Z., & Schlenke, T. A. (2012). Alcohol consumption as self-medication against blood-borne parasites in the fruit fly. *Current Biology*, 22(6), 488–493. <https://doi.org/10.1016/j.cub.2012.01.045>
- Miller, J. R., Wood, B. P., & Hamilton, M. B. (2008).  $F_{st}$  and  $Q_{st}$  under neutrality. *Genetics*, 180(2), 1023–1037.
- Mo, W., Wang, M., Zhan, R., Yu, Y., He, Y., & Lu, H. (2019). *Kluyveromyces marxianus* developing ethanol tolerance during adaptive evolution with significant improvements of multiple pathways. *Biotechnology for Biofuels*, 12(1), 63. <https://doi.org/10.1186/s13068-019-1393-z>
- Montooth, K. L., Siebenthal, K. T., & Clark, A. G. (2006). Membrane lipid physiology and toxin catabolism underlie ethanol and acetic acid tolerance in *Drosophila melanogaster*. *Journal of Experimental Biology*, 209(19), 3837–3850
- Morozova, T. V., Anholt, R. R., & Mackay, T. F. (2006). Transcriptional response to alcohol exposure in *Drosophila melanogaster*. *Genome Biology*, 7(10), R95.
- Morozova, T. V., Huang, W., Pray, V. A., Whitham, T., Anholt, R. R. H., & Mackay, T. F. C. (2015). Polymorphisms in early neurodevelopmental genes affect natural variation in alcohol sensitivity in adult *Drosophila*. *BMC Genomics*, 16(1), 865. <https://doi.org/10.1186/s12864-015-2064-5>
- Orr, H. A. (1998). The population genetics of adaptation: the distribution of factors fixed during adaptive evolution. *Evolution*, 52, 935–949.
- Park, A., Ghezzi, A., Wijesekera, T. P., & Atkinson, N. S. (2017). Genetics and genomics of alcohol responses in *Drosophila*. *Neuropharmacology*, 122, 22–35. <https://doi.org/10.1016/j.neuropharm.2017.01.032>

- Pool, J. E. (2016). Genetic mapping by bulk segregant analysis in *Drosophila*: experimental design and simulation-based inference. *Genetics*, 204(3), 1295–1306.
- Pool, J. E., Braun, D. T., & Lack, J. B. (2017). Parallel evolution of cold tolerance within *Drosophila melanogaster*. *Molecular Biology and Evolution*, 34(2), 349–360.
- Pool, J. E., Corbett-Detig, R. B., Sugino, R. P., Stevens, K. A., Cardeno, C. M., Crepeau, M. W., Duchon, P., Emerson, J. J., Saelao, P., Begun, D. J., & Langley, C. H. (2012). Population genomics of sub-Saharan *Drosophila melanogaster*: African diversity and non-African admixture. *PLoS Genetics*, 8(12), e1003080. <https://doi.org/10.1371/journal.pgen.1003080>
- Rockman, M. V. (2012). The Qtn program and the alleles that matter for evolution: all that's gold does not glitter. *Evolution*, 66(1), 1–17. <https://doi.org/10.1111/j.1558-5646.2011.01486.x>
- Siddiq, M. A., & Thornton, J. W. (2019). Fitness effects but no temperature-mediated balancing selection at the polymorphic *Adh* gene of *Drosophila melanogaster*. *Proceedings of the National Academy of Sciences*, 116(43), 21634–21640.
- Signor, S., & Nuzhdin, S. (2018). Dynamic changes in gene expression and alternative splicing mediate the response to acute alcohol exposure in *Drosophila melanogaster*. *Heredity*, 121(4), 342–360. <https://doi.org/10.1038/s41437-018-0136-4>
- Sohail, M., Maier, R. M., Ganna, A., Bloemendal, A., Martin, A. R., Turchin, M. C., Chiang, C. W., Hirschhorn, J., Daly, M. J., Patterson, N., Neale, B., Mathieson, I., Reich, D., & Sunyaev, S. R. (2019). Polygenic adaptation on height is overestimated due to uncorrected stratification in genome-wide association studies. *Elife*, 8, e39702. <https://doi.org/10.7554/eLife.39702>
- Spitze, K. (1993). Population structure in *Daphnia obtusa*: Quantitative genetic and allozymic variation. *Genetics*, 135(2), 367–374. <https://doi.org/10.1093/genetics/135.2.367>
- Sprengelmeyer, Q. D., Mansourian, S., Lange, J. D., Matute, D. R., Cooper, B. S., Jirle, E. V., Stensmyr, M. C., & Pool, J. E. (2020). Recurrent collection of *Drosophila melanogaster* from wild African environments and genomic insights into species history. *Molecular Biology and Evolution*, 37(3), 627–638. <https://doi.org/10.1093/molbev/msz271>
- Stanley, D., Bandara, A., Fraser, S., Chambers, P. J., & Stanley, G. A. (2010). The ethanol stress response and ethanol tolerance of *Saccharomyces cerevisiae*. *Journal of Applied Microbiology*, 109(1), 13–24.
- Stephan, W. (2016). Signatures of positive selection: From selective sweeps at individual loci to subtle allele frequency changes in polygenic adaptation. *Molecular Ecology*, 25(1), 79–88. <https://doi.org/10.1111/mec.13288>
- Thornton, K. R. (2019). Polygenic adaptation to an environmental shift: Temporal dynamics of variation under Gaussian stabilizing selection and additive effects on a single trait. *Genetics*, 213(4), 1513–1530. <https://doi.org/10.1534/genetics.119.302662>
- Turchin, M. C., Chiang, C. W., Palmer, C. D., Sankararaman, S., Reich, D., & Hirschhorn, J. N. (2012). Evidence of widespread selection on standing variation in Europe at height-associated SNPs. *Nature Genetics*, 44(9), 1015–1019. <https://doi.org/10.1038/ng.2368>
- Van Delden, W., Boerema, A. C., & Kamping, A. (1978). The alcohol dehydrogenase polymorphism in populations of *Drosophila melanogaster*. I. Selection in different environments. *Genetics*, 90(1), 161–191.
- van't Hof, A. E., Edmonds, N., Dalikova, M., Marec, F., & Saccheri, I. J. (2011). Industrial melanism in British peppered moths has a singular and recent mutational origin. *Science*, 332(6032), 958–960. <https://doi.org/10.1126/science.1203043>
- Wellenreuther, M., & Hansson, B. (2016). Detecting polygenic evolution: Problems, pitfalls, and promises. *Trends in Genetics*, 32(3), 155–164. <https://doi.org/10.1016/j.tig.2015.12.004>
- Yeaman, S., & Whitlock, M. C. (2011). The genetic architecture of adaptation under migration–selection balance. *Evolution*, 65(7), 1897–1911. <https://doi.org/10.1111/j.1558-5646.2011.01269.x>
- Zhao, X. Q., Xue, C., Ge, X. M., Yuan, W. J., Wang, J. Y., & Bai, F. W. (2009). Impact of zinc supplementation on the improvement of ethanol tolerance and yield of self-flocculating yeast in continuous ethanol fermentation. *Journal of Biotechnology*, 139(1), 55–60. <https://doi.org/10.1016/j.jbiotec.2008.08.013>

## SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

**How to cite this article:** Sprengelmeyer, Q. D., & Pool, J. E. (2021). Ethanol resistance in *Drosophila melanogaster* has increased in parallel cold-adapted populations and shows a variable genetic architecture within and between populations. *Ecology and Evolution*, 11, 15364–15376. <https://doi.org/10.1002/ece3.8228>