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 [1], and John E. Pool[1]

[1]*Laboratory of Genetics, University of Wisconsin-Madison, Madison, WI, 53706*

Corresponding author:

John E. Pool

425-G Henry Mall

Madison, WI 53706

+1-608-265-1036

jpool@wisc.edu

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 - and that ethanol and cold resistance may have a partially shared genetic basis. 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.

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)? Is there genetic predictability between populations (Stern & Orgogozo 2008)? And do favored variants tend to reach fixation, or stop rising because selective pressures change or traits reach their new optima (Thornton 2019)?

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).

Studies have shown that there can be genetic predictability underlying parallel trait evolution. The same genes with large-effect have been found to cause armor loss in different Alaskan populations of three-spine stickleback (Cresko et al. 2004). However, the same genes may not always be responsible for the adaptive change. Dark pigmentation found in African populations of *D. melanogaster* may be due in part to the result of

67 population specific genes (Bastide et al. 2016). Differences in coat color found in
68 populations of pocket mice also are the result of different genetic changes (Nachman et al.
69 2003).

70 *D. melanogaster* originated in woodland environments of southern-central Africa
71 and then expanded throughout Africa beginning ~13,000 years ago (Sprengelmeyer et al.
72 2020). The species appears to have crossed the Sahara relatively soon after their
73 expansion started and may have only reached Europe ~1,800 years ago. During the
74 migration out of their ancestral habitat, populations of *D. melanogaster* encountered many
75 novel environmental habitats, which included equatorial tropical rainforest, northern
76 temperate grassland, and high altitude alpine regions. Each of these different ecosystems
77 provides unique selection pressures that may have forced local populations to acquire
78 novel traits in order to survive. For example, increased ultra-violet radiation found at
79 higher altitudes might have caused populations to evolve darker cuticle pigmentation
80 (Bastide et al. 2014). And populations at high latitude and altitude have independently
81 evolved elevated cold tolerance (Pool et al. 2017).

82 Ethanol resistance is another trait that has evolved in *D. melanogaster.* When
83 compared to its sister species *D. simulans*, *D. melanogaster* are more ethanol resistant
84 (McKenzie and Parsons 1972). Within *D. melanogaster*, ethanol resistance has shown a
85 positive correlation with latitude (David & Bocquet 1975 and Cohan & Graf 1985) with
86 populations living in breweries and wine cellars of France and Spain being the most
87 resistant (McKenzie & Parsons 1974 and Mercot et al. 1994). Female flies lay their eggs
88 on ethanol-producing fermenting fruit and having a higher ethanol resistance may
89 provide more available resources. There is evidence that *D. melanogaster* prefers to lay their

eggs on medium that contains alcohol (McKenzie and Parsons 1972), which can be a defense against parasitoids (Milan et al. 2012; Kacsoh et al. 2013).

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 and Bocquet 1976), with the "fast" allele having a higher resistance compared to the "slow" allele. David et al. (1975) found a latitudinal gradient and populations at higher latitudes tend to be more resistant and also have a higher Adh_{fast} 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 (Mercot et al. 1994). It has been hypothesized that the Adh_{fast} and Adh_{slow} polymorphism has been maintained by a temperature dependent balancing selection (Van Delden et al. 1978). However, Siddiq and Thorton (2019) found Adh_{fast} 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 and 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 to less resistant African flies even without the amino acid polymorphism (Fry 2014). Chakraborty and Fry (2016) found that polymorphisms in *ALDH* are maintained by

114 environmental conditions. Transgenic experiments confirmed there is an increase in
115 lifetime fitness on ethanol-supplemented medium specifically.

116 Although *ADH* and *ALDH* play an important role, they are not the only genes
117 involved in ethanol resistance. Other genes linked to ethanol resistance encompass a
118 wide range of functions such as lipid membrane physiology (Montooth et al. 2006), ion
119 channels (Cowmeadow et al. 2005), central nervous system (Chandler et al. 1998), zinc
120 retention (Zhao et al. 2009), and feeding behavior and behavioral responses to ethanol
121 (Fochler et al. 2017). Signor and Nuzhdin (2018) found that many genes display plasticity
122 in expression and splicing in response to ethanol exposure. Other studies that focused on
123 changes in gene expression (Morozova et al. 2006) or histone modification (Ghezzi et al.
124 2013) have also found that numerous genes respond to ethanol exposure.

125 Many of these studies investigate lab strains or compare single populations
126 between regions of contrasting ethanol resistance. The goal of this study is to understand
127 the genetic architecture of ethanol resistance, an adaptive trait in *D. melanogaster*. To
128 investigate we use wild populations from their ancestral range (Zambia), along with
129 multiple populations that display elevated ethanol resistance: from high altitude sub-
130 Saharan Africa (Ethiopia and South Africa) and from high latitude (France). Each of these
131 populations has also evolved elevated cold tolerance, and in light of the species' expansion
132 history, these trait changes are thought to arisen independently in the Ethiopia, France,
133 and South Africa populations (Pool et al. 2017). To detect QTLs that are involved in this
134 adaptive trait evolution, we performed bulk segregant analysis (Pool 2016). We used
135 population genetic statistics, Gene Ontology enrichment and genotype-phenotype
136 association testing to find evidence of local adaptation and candidate genes. We also

performed simulations to explore the parameters involved in the genetic architecture of this adaptive trait change.

Material and Methods

Experimental Populations

All flies used in the experimental 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).

Ethanol Resistance

To test for population differences in ethanol resistance, we measured mobility over a 6-hour period. To help reduce any adverse effects due to inbreeding each line was outcrossed with to a unique inbred line from the same population. The number of flies and pairs of lines used was: FR: 5 lines, N=50, EF: 5 lines, N=46, SD: 3 lines, N=30 and ZI: 4 lines, N=40. We then placed 3-5-day-old outcrossed female flies into 50ml falcon tubes with a single tissue placed in the bottom that was saturated with 1.5ml of 3% sucrose (molasses) solution that contained 8% ethanol (Fry 2014). We visually scored flies that did not move after the vial was flicked as “immobile”.

161 *Bulk Segregant Analysis*

162 To ascertain areas of local adaptation responsible for higher ethanol resistance,
163 bulk segregant analysis was performed to detect quantitative trait loci (QTL) (Pool 2016).
164 Population cages were started from reciprocal crosses between eight inbred parental
165 individuals of low resistant (Zambia) and one each of the more resistant African
166 populations (Ethiopia and South Africa) lines and strains from two French populations
167 (Supplemental Material, S1). From each reciprocal cross, 125 F1 offspring of each sex
168 were used to establish the second generation. These mapping populations for the rest of
169 the (non-overlapping) generations were maintained at ~1200 individuals. The flies were
170 housed in 28 x 14 x 15 cm plastic cages that contain 14 vials with a medium that contains
171 molasses, corn meal, yeast, agar, and antimicrobial agents at ~20° C. Adult flies were
172 allowed to lay eggs on the food for one week before being removed. The food vials were
173 replaced when adult flies in the cage were 7-10 day old. At the 15th generation, 600 3-5-
174 day-old female flies from each population cage were exposed to the 8% ethanol mobility
175 assay described above. The flies were placed into two pools, 10% least resistant (N=60)
176 and 10% most resistant (N=60).

177

178 *Genome Preparation*

179 We sequenced the genomes of pooled samples (N=30 individuals) for the parental
180 lines and two such pools for each of the low and high resistant groups (N=60 total for
181 each extreme). Genomic DNA was obtained using a chloroform extraction and ethanol
182 precipitation protocol. The DNA was fragmented with a Bioruptor sonicator
183 (Diagenode), and paired-end libraries with ~300 bp inserts prepared using NEBNext
184 DNA Library Prep Reagent Set for Illumina (New England Biolabs no. E6000L). Each

185 library's concentration and quality was analysed with an Agilent 2100 Bioanalyzer
186 (Agilent Technologies, Inc.). The prepared libraries were sequenced at UW-Madison
187 Biotechnology Center on the Illumina HiSeq 2000 platform.
188

189 *Genome Alignment*

190 All the raw data that passed the Illumina filters were processed using a Perl-
191 scripted pipeline. Reads from each sequenced genome were mapped to the *D.*
192 *melanogaster* reference genome (release 5.57) obtained from Flybase (www.flybase.org), with
193 the default parameters in BWA ver. 0.6.2-r126 (Li and Durbin 2009). Using Stampy ver.
194 1.0.21(Lunter and Goodson 2011) the BAM files were then remapped. With samtools ver.
195 0.1.18 (Li et al. 2009) reads were filtered for a mapping quality of 20 and for proper pairs.
196 The BAM files were further processed by removing unmapped reads and sorted by
197 coordinate, and PCR duplicates were marked using Picard ver. 1.109
198 (http://picard.sourceforge.net). To improve the alignment around indels we used GATK
199 ver. 3.2 (McKenna et al. 2010). The average depth of coverage per genome was
200 calculated for the parental lines and the low and high resistance lines (Table S1).
201

202 *Quantitative Trait Locus (QTL) Mapping*

203 The PoPoolation2 ver. 1.201 software package (Kofler et al. 2011) was used to
204 create synchronised mpileup files for the aligned genomes. For each biallelic SNP an
205 ancestry difference (a_d) was calculated (Bastide et al. 2016). This quantity was calculated
206 as the proportion of ethanol resistant pooled sequence reads carrying the fixed allele in
207 the high resistant parental line and absent in the low resistant parental line.

208 *Equation 1:* $a_d = (f_H - f_L) / (p_H - p_L)$

209 Where p_H is the frequency of parental high resistant allele, p_L is the low resistant parental
210 allele, f_H is high resistant F12 offspring and f_L is low resistant F12 offspring. The five
211 chromosomal arms (X, 2L, 2R, 3L, and 3R) were divided into windows based on SNP
212 density (Lack et al. 2015) which created 2728, 3131, 2357, 2956, and 2935 windows
213 respectively each roughly 8.4-kb in size on average. Across the five chromosomal arms
214 only sites that had a frequency difference of ≥ 0.25 were used in the analysis. A
215 simulation-based inference for BSA mapping (SIBSAM) was performed (Pool 2016) to
216 identify significant QTL and calculate their confidence intervals and effect sizes. The
217 custom scripts used for SIBSAM can be found at:
218 http://github.com/JohnEPool/SIBSAM1. SIBSAM is able to evaluate both primary
219 QTL peaks and flanking secondary QTL peaks, evaluating whether ragged peaks contain
220 significant evidence for more than one QTL. Forward simulations incorporate
221 recombination in multiple individuals for multiple generations, selection on phenotype in
222 the final generation with additivity, plus environmental variance, and then the sampling
223 of sequence reads to obtain a_d.
224
225 *Genetic differentiation and Gene Ontology (GO) enrichment analysis*
226 To find evidence of local adaptation and produce a list of candidate genes found
227 within the significant QTLs, window F_{ST} and maximum SNP F_{ST} per window (hereafter
228 "SNP F_{ST}"),and the haplotype statistic χ_{MD} (Lange & Pool 2016) were analyzed. Genomes
229 from Zambia (n=197), South Africa (n=61), Ethiopia (n=68) and France (n=96) were
230 used from the *Drosophila* Genome Nexus (Lack et al. 2015). The χ_{MD} compares length of

231 identical haplotype blocks among individuals in one population versus another. The
232 comparisons were made within each of the five chromosomal arms (X, 2L, 2R, 3L, and
233 3R), which were divided into windows based on SNP density (Lack et al. 2015). The idea
234 behind χ_{MD} is that in a recently selected population, longer stretches of identical
235 haplotypes will not have had time for recombination or mutation to break up longer
236 identical tracts. A chromosomal arm quantile outlier approach was used to focus on genes
237 with an extreme population genetic signal. Only windows that were in the top 2.5%
238 quantile in any of the three statistics were classified as outliers. To form an outlier region,
239 a maximum of two non-outlier windows were allowed between two outlier windows.
240 Genes associated with outlier windows (overlapping them or the nearest gene in either
241 direction) were retained for subsequent analysis. The outlier genes identified in significant
242 QTL regions were used for window-based gene ontology (GO) enrichment analysis (as
243 implemented in Pool et al. 2012) to identify functional categories that differ between low
244 and high resistance populations. A *P* value was calculated based on the probability of
245 observing a given number of outlier genes from a GO category. *P* values were obtained
246 from permutation in which outlier region were randomly reassigned 10,000 times.
247
248 *Genotype-Phenotype Association Testing*
249 Phenotypic data was collected on 51 France inbred strains with sequenced
250 genomes from the Drosophila Genome Nexus genomic resource (Lack et al. 2016)
251 following the ethanol assay described above. To capture the variation in ethanol
252 resistance found in France population, the ethanol concentration used was 18%.
253 Genotype-phenotype associations were analyzed with the R package rrBLUP version 3.1
254 (Endelman 2011). Only regions within the QTL peaks of less than 2 Mb in length in the

two France crosses were examined. Within in these peaks, only SNPs that had a called allele of greater than 25% and a minor allele frequency greater than 5% were analyzed. One thousand permutations of the phenotypic data were used to calculate the significant threshold.

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 population and allele frequencies. We used sample sizes of 50, 100, 200, 500, and 1,000, 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:

Equation 2 $0.65q + 0.35(q^2 + 2q(1-q))$,

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. In light of the replicated phenotyping of individuals from each inbred line, no environmental variance. We performed 1,000 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.

Results

Population Differences in Ethanol Resistance

We performed a phenotypic assay of adult exposure to 8% ethanol vapor, using outbred flies from multiple independent within-population crosses 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 4% of the individuals immobile after six hours of exposure. The Zambia population from the ancestral range had the lowest resistance;

after four hours nearly all the individuals were immobile and after six hours there was 100% immobility. 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 and Ethiopia had ~40% of the individuals immobile at 4 hours. However, South Africa had 100% immobility at 6 hours while Ethiopia had 72% immobility.

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 were 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 ($N \approx 1{,}200$) until the 15^{th} 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 S2). 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 X% to 27%. Encouragingly, the

highest peaks in each cross were estimated to have narrow genomic confidence intervals (Table S2).

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 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 same France and Zambia populations. While chance false positive and 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.

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 χ_{MD}. These statistics

375 may have differing power to detect local adaptation depending on whether selective
376 sweeps are complete or incomplete, or hard versus soft (Lange & Pool 2016). We used a
377 quantile approach focusing on regions that had one of the three statistics with a quantile
378 below 0.025 (Table S3). This analysis yielded both genes with known functions that may
379 relate to our trait, and genes with no such known functions. While any of these genes
380 might contribute to ethanol resistance evolution, we mention below a few plausible
381 candidates.

382 Within the South Africa QTLs, peaks on chromosome arms X and 2R each have
383 one outlier redox gene, *Pp2C1* and *Nox*, respectively. Genes involved in regulating
384 oxidative stress have previously been implicated in *Drosophila* ethanol resistance (Awofala
385 et al. 2012). Of potential relevance in light of our aerosol ethanol assay, several genes
386 involved in the development of chitin also have population genetic signals: *ovo*, *mgl*, and
387 *CG1367*. Potential candidate genes found in Ethiopia QTLs include: *Shab* and *Teh2* (ion
388 channels), and *m* (cuticle development). Genes found in one of the two France crosses
389 included some potentially involved in alcohol metabolism (*CG5065*, *CG6650*, *CG8303*,
390 *CG9521*, *CG13091*, *CG15601*, *CG43658*, *Pis*), as well as ion channels (*para*, *ppk*, *sh*) and
391 other genes involved in neurotransmission (*be*, *CG33639*). Diverse aspects of nervous
392 system function have previously been linked to alcohol resistance (*e.g.* Morozova et al.
393 2015; Park et al. 2017).

394 Between the two France crosses, shared candidate genes included: *CG45065*
395 (alcohol metabolism), *CG9503* (choline/aldehyde metabolism), *bgm* and *pgdy* (fatty acid
396 metabolism), *hiw* (synapse organization), and *eag* (ion channel, response to ether). South
397 Africa and FR364N had two candidate genes of interest; *CG32698* (carbonate

dehydratase) and *CG1986* (lipase). Lipid levels are known to influence ethanol resistance (Lieber and Savolainen 1984, Geer et al. 1991).

Gene Ontology (GO) Enrichment

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 ($P = 0.00356$; Table S4). 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 (Stanley et al. 2010; Mo et al. 2019), and microvillus membrane (Bjorkman et al. 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).

Genotype-Phenotype Association Testing

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 (9,480 SNPs). Genome-wide significance, assessed via permutations, was not reached by any SNP in either analysis (Table S5). From the more inclusive analysis, the highest marker had a –

422 log *P* value of 4.43, whereas the permutation significance threshold was 6.17. From the
423 population genetic outlier analysis, the highest marker had a −log *P* value of 3.56 with
424 permutation significance threshold of 4.44.
425
426 *Simulation-Driven Investigation of Genetic Architecture and Association Testing Power*
427 We then considered which genetic architectures our QTL mapping data might
428 provide evidence for, and whether they might account for our negative association testing
429 results. Although full model inference of adaptive evolution at the genetic level is beyond
430 the scope of the present study, we conducted an exploratory simulation analysis in three
431 stages, focusing on the two France crosses.
432 First, we wanted to assess the number and strength of QTLs that our mapping
433 data were most consistent with. Our simulations used a modified version of SIBSAM,
434 which simulates the full mapping experiment (including recombination, phenotypic
435 selection, and sequencing read sampling). We varied the number of QTLs present in each
436 cross, their distribution of effect sizes as a function of the gamma distribution shape
437 parameter, and the proportion of trait variation contributed by environmental/random
438 effects rather than these QTLs. And we quantified properties of QTL peaks and genome-
439 wide ancestry in the simulated data and compared it with our empirical observations
440 using mean relative error. The parameter combination with the lowest average mean
441 relative error was 10 QTLs per cross, a gamma shape parameter of 4, and 70%
442 environmental variation (Figure 5; Table S6). An otherwise identical parameter
443 combination with a gamma shape parameter of 8 matched the empirical data almost
444 equally as well, and so we chose an intermediate shape parameter of 6 in further analyses.
445 Parameter combinations involving a wide range of QTL numbers and gamma shape

446 values were non-significantly worse than the above combination, indicating that larger
447 data sets will be needed to make formal inferences about the genetic architecture of
448 adaptive evolution (Table S6).
449 Second, we assessed whether the degree of QTL overlap between the two France
450 crosses provides information about the frequency of ethanol resistance alleles in this
451 population. The QTL frequency that resulted in the average overlap of peaks that best
452 matched the empirical data was 5% (Table S7), which resulted in ~54% overlap
453 compared to the empirical ~44%. However, only the highest frequency values (90% and
454 above) had confidence intervals that marginally excluded the empirical frequency,
455 suggesting that larger data sets will be needed to gain further resolution about the
456 frequencies of adaptive variants, and underscoring the exploratory nature of our
457 simulations.
458 We therefore investigated a wide range of frequencies (5% to 50%) in assessing the
459 power of our association testing analysis. The power analysis revealed that there is little
460 power to detect causative SNPs that segregate at lower frequencies. Small to moderate
461 population sizes (n=50, 100, and 200) had low power to detect SNPs at any frequency
462 (Figure 6). It was not until QTL frequency reached 50% that there was a greater than
463 10% detection power. Even when there was a large number of individuals used, e.g.
464 n=1000, only ~0.9% of the causative SNPs at a 5% QTL frequency met the empirical
465 threshold. Only once QTL frequency reached 30% was there greater than 50% detection
466 rate. However, the detection power did improve when the lower outlier region threshold
467 was used. The small population sizes had a greater than 10% detection power when QTL
468 frequency reached 30%. The detection power at 5% QTL frequency for large population
469 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 tested variants could be further reduced.

Testing for pleiotropy between ethanol and cold resistance

Cold resistance has evolved in these same three populations for which we detected elevated ethanol resistance (France and high altitude Ethiopia and South Africa; Pool et al. 2017). These observations raise the question of whether cold and ethanol resistance are genetically correlated (pleiotropically connected), or if instead their geographic co-occurrence is due to ecological correlation (greater exposure to ethanol in cold environments). Cold resistance, in the sense of mobility after long-term cold exposure, was previously estimated for a panel of France inbred lines. For 37 of those same lines, we had also collected ethanol resistance data for the genotype-phenotype analysis described above. In testing for a correlation between these two sets of trait measurements, we found a mild positive correlation ($r = 0.225$; one-tailed $P = 0.0897$). While some strains did show high cold resistance but low ethanol resistance, there appeared to be a lack of strains showing high ethanol resistance but low cold resistance (Figure 7). This analysis provides a tentative hint that pleiotropy may indeed exist between the cold resistance and ethanol resistance that have evolved recurrently within *D. melanogaster*. However, larger-scale experiments will be needed to confirm this result, and linkage (*e.g.* mediated by inversions) might also contribute to such a trait correlation.

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. In agreement with other studies (David and Bocquet 1975) we found the higher latitude France population to be highly resistant. Interestingly, none of the France QTL peaks contain *Adh* (Cohen and Graf 1985) or *Aldh* (Fry et al. 2008), both implicated in the latitudinal cline of increased ethanol resistance. We have also reported for the first time that populations at higher altitudes have increased ethanol resistance.

The France population and the high altitude Ethiopia and South Africa populations have all evolved increased cold tolerance (Pool *et al.* 2017). The geographic cooccurrence between resistance to cold and ethanol might be due to either genetic correlation (*i.e.* pleiotropy) or else ecological correlation. Above, we found initial evidence that pleiotropy may exist between these evolved traits in the France population. There is also prior evidence suggesting that genes that are involved with ethanol resistance may have a pleiotropic effect with cold resistance: increased lipid concentration in membranes can increase cold resistance and make the cell more stable, which in turn also increases ethanol resistance (Montooth et al. 2006). However, ecological correlation may still play a role in the geographic co-occurrence of these traits. Flies in colder environments might encounter greater concentrations of ethanol, leading to selection for ethanol resistance variants that may be independent of those conferring cold adaptation. In warmer environments, both ethanol and water vapor evaporate more rapidly. There, fruit may desiccate before it would accumulate enough ethanol, and ethanol that is produced may

516 dissipate more quickly. In colder environments, fruit may retain moisture for a longer
517 period of time, allowing ethanol to build up through microbial fermentation, and this
518 ethanol. However, it is worth noting that altitude also serves to increase evaporation rates.
519 Future genetic and physiological studies are needed to more clearly discern the potential
520 genetic and ecological correlations that may underlie the geographic co-occurrence of
521 these traits.

522 The BSA performed on the four different crosses revealed 32 significant QTLs
523 with the largest estimated effect size for each cross between 12% and 27%. These data
524 taken together suggests that ethanol resistance is moderately polygenic with moderate to
525 large effect QTLs present (whereas smaller QTLs may elude our detection power; Pool
526 2016). We found that there are no QTLs overlapping between all three high resistant
527 populations. However, each of the high altitude populations shares QTLs with the France
528 strains, whereas the two high altitude populations, South Africa and Ethiopia, do not
529 have any QTLs in common with each other.

530 In interpreting the observed levels of QTL overlap between populations, it is
531 important to keep in mind that even between two crosses from the same resistant
532 population (France), QTL overlap was modest. Of the seventeen significant QTL peaks
533 between the two France crosses, they shared only six QTLs (and even some of these could
534 reflect random overlap in light of the QTL sizes). For example, the strongest QTL in
535 either France cross with an estimated effect size of 27% from FR305N on chromosome
536 arm 3L, is completely missing in FR364N. Our experiment should have very high power
537 to detect a QTL with an effect size this large if it existed in a second cross (Pool 2016).
538 Those results suggest a genetically heterogenous architecture of ethanol resistance
539 evolution not only between populations but 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.

Data Accessibility

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

Acknowledgments

This project benefitted from funding from NIH grants R01 GM111797 and R35 GM136306, 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.

Literature Cited

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.

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.

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.

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.

587 Chakraborty, M., & Fry, J. D. (2016). Evidence that Environmental Heterogeneity
588 Maintains a Detoxifying Enzyme Polymorphism in Drosophila melanogaster. *Current*
589 *Biology*, *26*(2), 219–223.
590
591 Chandler, L. J., Harris, R. A., & Crews, F. T. (1998). Ethanol tolerance and synaptic
592 plasticity. *Trends in Pharmacological Sciences*, *19*(12), 491-495.
593
594 Cohan, F. M., & Graf, J.-D. (1985). Latitudinal Cline in Drosophila melanogaster for
595 Knockdown Resistance to Ethanol Fumes and for Rates of Response to Selection for
596 Further Resistance. *Evolution*, *39*(2), 278–293.
597
598 Cowmeadow, R. B., Krishnan, H. R., & Atkinson, N. S. (2005). The slowpoke gene is
599 necessary for rapid ethanol tolerance in Drosophila. *Alcoholism: Clinical and Experimental*
600 *Research*, *29*(10), 1777-1786.
601
602 Cresko, W. A., Amores, A., Wilson, C., Murphy, J., Currey, M., Phillips, P., Bell, M. A.,
603 Kimmel, C. B., & Postlethwait, J. H. (2004). Parallel genetic basis for repeated evolution
604 of armor loss in Alaskan threespine stickleback populations. *Proceedings of the National*
605 *Academy of Sciences*, *101*(16), 6050–6055.
606
607 David, J. R., & Bocquet, C. (1975). Similarities and differences in latitudinal adaptation of
608 two Drosophila sibling species. *Nature*, *257*(5527), 588-590.
609

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.

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.

Dittmar, E. L., Oakley, C. G., Conner, J. K., Gould, B. A., & Schemske, D. W. (2016, April). Factors influencing the effect size distribution of adaptive substitutions. In *Proc. R. Soc. B* (Vol. 283, No. 1828, p. 20153065). The Royal Society.

Endelman, J. B. (2011). Ridge regression and other kernels for genomic selection with R package rrBLUP. *The Plant Genome*, *4*(3), 250-255.

Fisher, R. A. (1930). *The genetical theory of natural selection: a complete variorum edition*. 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., & Saweikis, M. (2006). Aldehyde dehydrogenase is essential for both adult and larval ethanol resistance in Drosophila melanogaster. *Genetical research*, *87*(02), 87-92.

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.

Fry, J. D. (2014). Mechanisms of naturally evolved ethanol resistance in Drosophila melanogaster. *Journal of Experimental Biology*, *217*(22), 3996-4003.

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.

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.

Ghezzi, A., Krishnan, H. R., Lew, L., Prado III, F. J., Ong, D. S., & Atkinson, N. S. (2013). Alcohol-induced histone acetylation reveals a gene network involved in alcohol tolerance. *PLoS Genet*, *9*(12), e1003986

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*.

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.

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.

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.

Li, H., & Durbin, R. (2009). Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics*, *25*(14), 1754–1760.

Lieber, C. S., & Savolainen, M. (1984). Ethanol and Lipids. *Alcoholism: Clinical and Experimental Research*, *8*(4), 409–423.

Lunter, G., & Goodson, M. (2011). Stampy: A statistical algorithm for sensitive and fast mapping of Illumina sequence reads. *Genome Research*, *21*(6), 936–939.

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.

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.

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.

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.

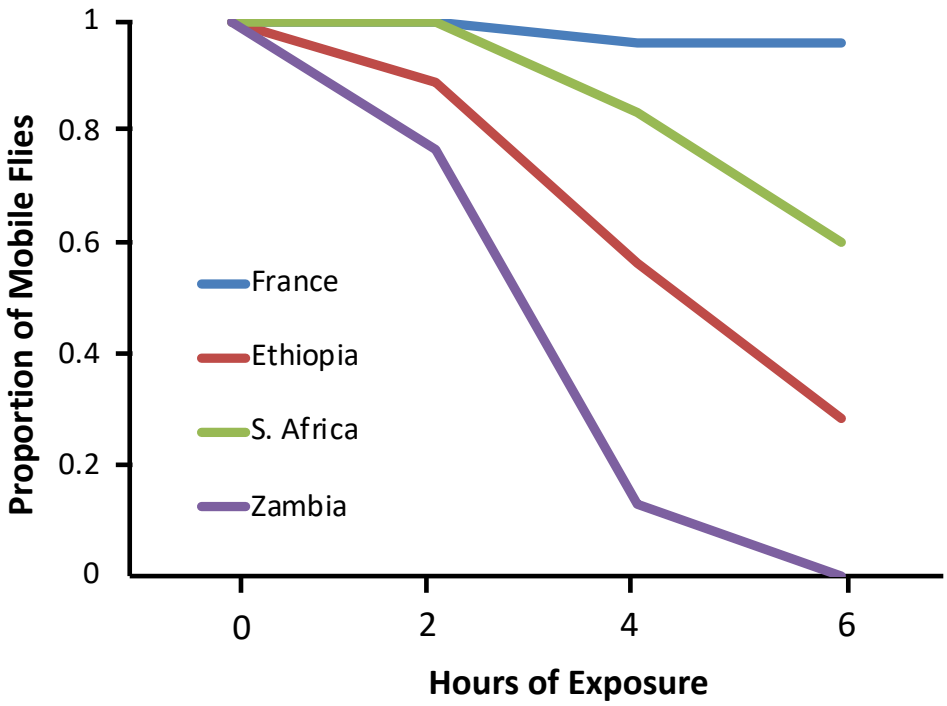
706

707 Montooth, K. L., Siebenthall, K. T., & Clark, A. G. (2006). Membrane lipid physiology
708 and toxin catabolism underlie ethanol and acetic acid tolerance in Drosophila
709 melanogaster. *Journal of experimental biology*, *209*(19), 3837-3850.
710
711 Morozova, T. V., Anholt, R. R., & Mackay, T. F. (2006). Transcriptional response to
712 alcohol exposure in Drosophila melanogaster. *Genome biology*, *7*(10), R95.
713
714 Morozova, T. V., Huang, W., Pray, V. A., Whitham, T., Anholt, R. R. H., & Mackay, T.
715 F. C. (2015). Polymorphisms in early neurodevelopmental genes affect natural variation
716 in alcohol sensitivity in adult drosophila. *BMC Genomics*, *16*(1), 865.
717
718 Nachman, M. W., Hoekstra, H. E., & D'Agostino, S. L. (2003). The genetic basis of
719 adaptive melanism in pocket mice. *Proceedings of the National Academy of Sciences*, *100*(9),
720 5268–5273.
721
722 Orr, H. A. (1998). The population genetics of adaptation: the distribution of factors fixed
723 during adaptive evolution. *Evolution*, 935-949.
724
725 Park, A., Ghezzi, A., Wijesekera, T. P., & Atkinson, N. S. (2017). Genetics and genomics
726 of alcohol responses in Drosophila. *Neuropharmacology*, *122*, 22–35.
727
728 Pool, J. E., Corbett-Detig, R. B., Sugino, R. P., Stevens, K. A., Cardeno, C. M.,
729 Crepeau, M. W., Duchen, P., Emerson, J. J., Saelao, P., Begun, D. J., & Langley, C. H.

730 (2012). Population Genomics of Sub-Saharan Drosophila melanogaster: African Diversity
731 and Non-African Admixture. *PLOS Genetics*, *8*(12), e1003080.
732
733 Pool, J. E. (2016). Genetic mapping by bulk segregant analysis in Drosophila:
734 experimental design and simulation-based inference. *Genetics*, *204*(3), 1295-1306.
735
736 Pool, J. E., Braun, D. T., & Lack, J. B. (2017). Parallel evolution of cold tolerance within
737 Drosophila melanogaster. *Molecular biology and evolution*, *34*(2), 349-360.
738
739 Siddiq, M. A., & Thornton, J. W. (2019). Fitness effects but no temperature-mediated
740 balancing selection at the polymorphic Adh gene of Drosophila melanogaster. *Proceedings*
741 *of the National Academy of Sciences*, *116*(43), 21634–21640.
742
743 Signor, S., & Nuzhdin, S. (2018). Dynamic changes in gene expression and alternative
744 splicing mediate the response to acute alcohol exposure in Drosophila melanogaster.
745 *Heredity*, *121*(4), 342–360.
746
747 Sprengelmeyer, Quentin D., Suzan Mansourian, Jeremy D. Lange, Daniel R. Matute,
748 Brandon S. Cooper, Erling V. Jirle, Marcus C. Stensmyr, & John E. Pool. (2020).
749 Recurrent Collection of Drosophila melanogaster from Wild African Environments and
750 Genomic Insights into Species History. *Molecular biology and evolution* 37(3). 627-638.
751

752 Stanley, D., Bandara, A., Fraser, S., Chambers, P. J., & Stanley, G. A. (2010). The
753 ethanol stress response and ethanol tolerance of Saccharomyces cerevisiae. *Journal of*
754 *applied microbiology*, *109*(1), 13-24.
755
756 Stern, D. L., & Orgogozo, V. (2008). The Loci of Evolution: How Predictable Is Genetic
757 Evolution? *Evolution*, *62*(9), 2155–2177.
758
759 Thornton, K. R. (2019). Polygenic Adaptation to an Environmental Shift: Temporal
760 Dynamics of Variation Under Gaussian Stabilizing Selection and Additive Effects on a
761 Single Trait. *Genetics*, *213*(4), 1513–1530.
762
763 Van Delden, W., Boerema, A. C., & Kamping, A. (1978). The alcohol dehydrogenase
764 polymorphism in populations of Drosophila melanogaster. I. Selection in different
765 environments. *Genetics*, *90*(1), 161-191.
766
767 Wellenreuther, M., & Hansson, B. (2016). Detecting Polygenic Evolution: Problems,
768 Pitfalls, and Promises. *Trends in Genetics*, *32*(3) 155–164.
769
770 Yeaman, S., & Whitlock, M. C. (2011). The genetic architecture of adaptation under
771 migration–selection balance. *Evolution*, *65*(7), 1897-1911.
772
773 Zhao, X. Q., Xue, C., Ge, X. M., Yuan, W. J., Wang, J. Y., & Bai, F. W. (2009). Impact
774 of zinc supplementation on the improvement of ethanol tolerance and yield of self-

775 flocculating yeast in continuous ethanol fermentation. *Journal of Biotechnology*, *139*(1), 55–
776 60.
777



778

779

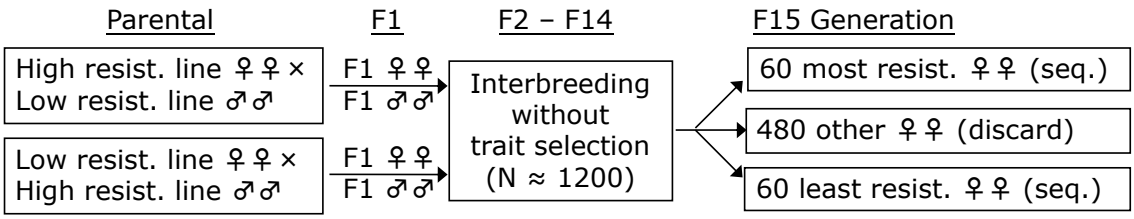
780 **Figure 1**. Population differences in resistance to concentrated ethanol vapor. The

781 percentage of flies that was mobile after being exposed to 8% ethanol is shown across a 6

782 hour interval. Ethiopia (EF) n=46, Zambia (ZI) n=39, France (FR) n=50, South Africa

783 (SD) n=30.

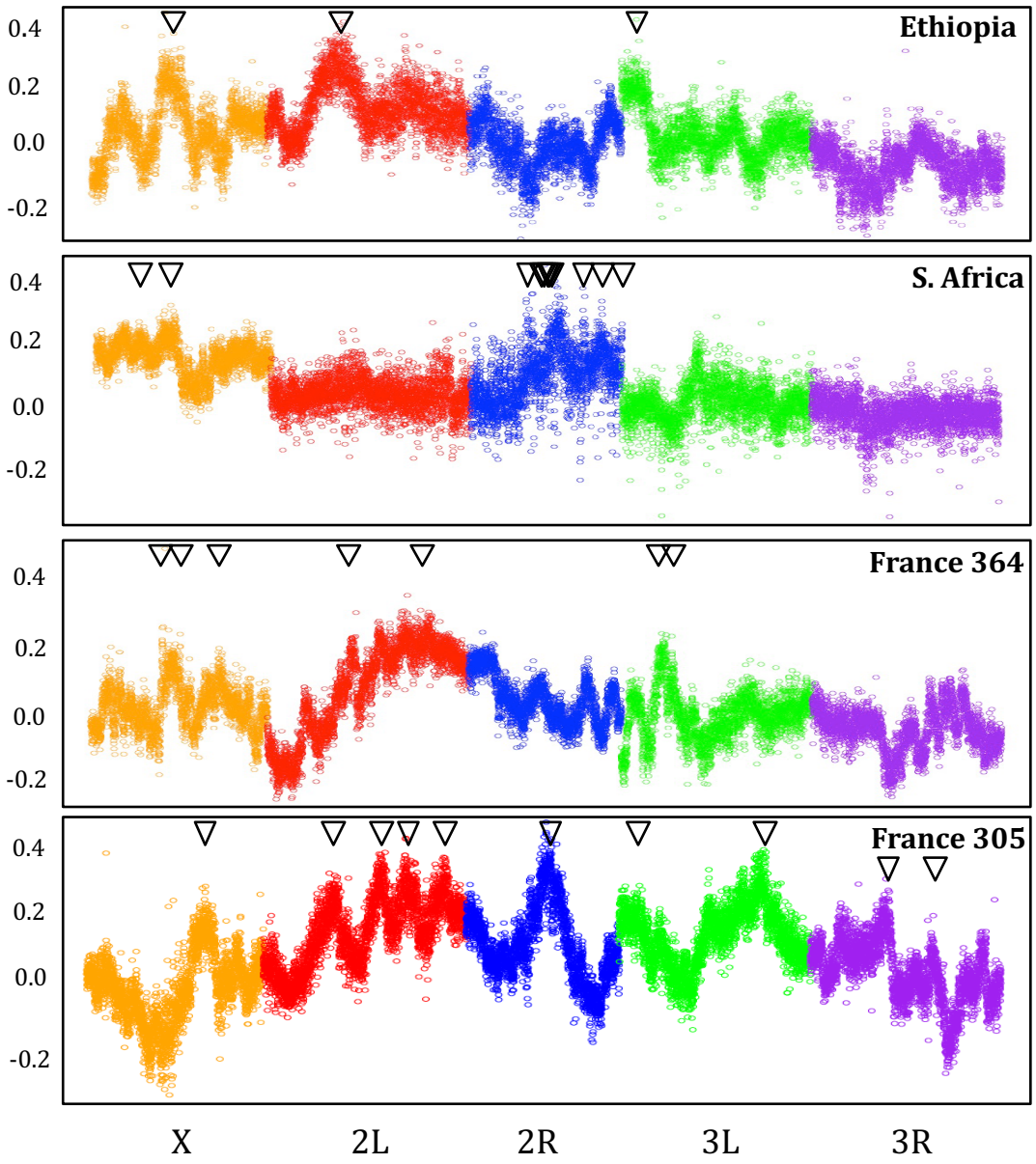
784



785
786

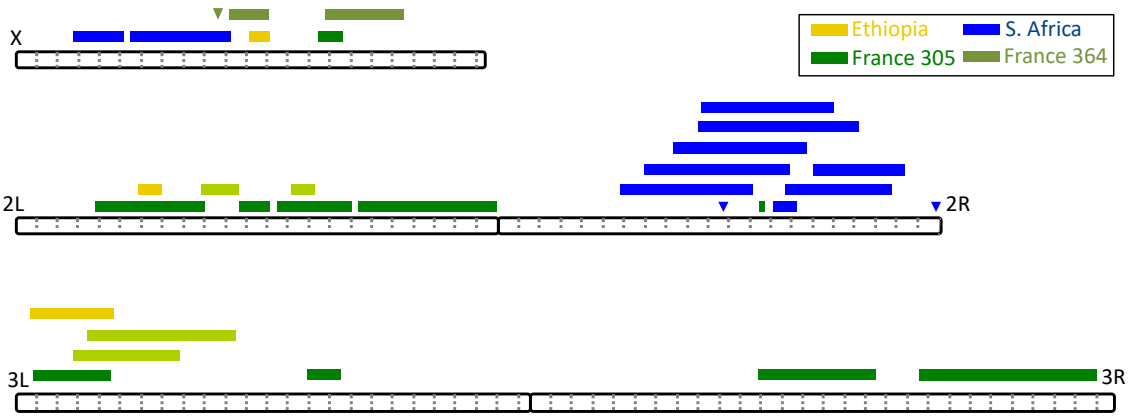
787 **Figure 2**. The bulk QTL mapping experimental design is illustrated, as further

788 described in the Materials and Methods.



789

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



796

797

798 **Figure 4**. The locations of significant QTLs on the five euchromatic chromosome arms

799 of *D. melanogaster*. The colors indicate ethanol resistance mapping crosses involving

800 Ethiopia, South Africa, and France 305 and 364. The width of each box indicates the

801 90% C.I. of each QTL. Intervals that are less than 10 kb in width are marked with

802 triangles. Dotted gray lines indicate Mb increments.

803

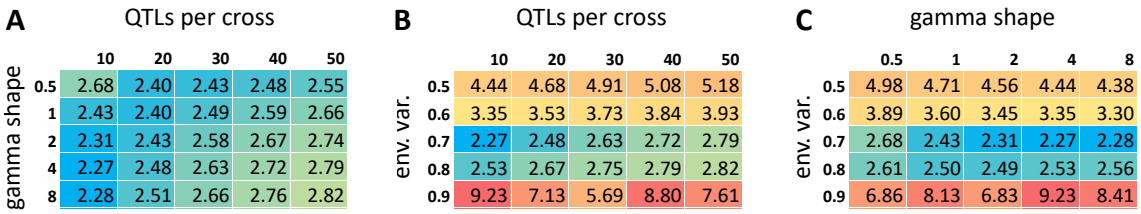


Figure 5. Heat map depicting mean relative error between empirical data (from France crosses) and selected simulated data sets, based on the QTL mapping summary statistics compared. These plots each fix one of the three parameters with its value from the best-matching parameter combination: (A) environmental variance of 0.7, (B) gamma shape parameter of 6, and (C) 10 QTLs per cross.

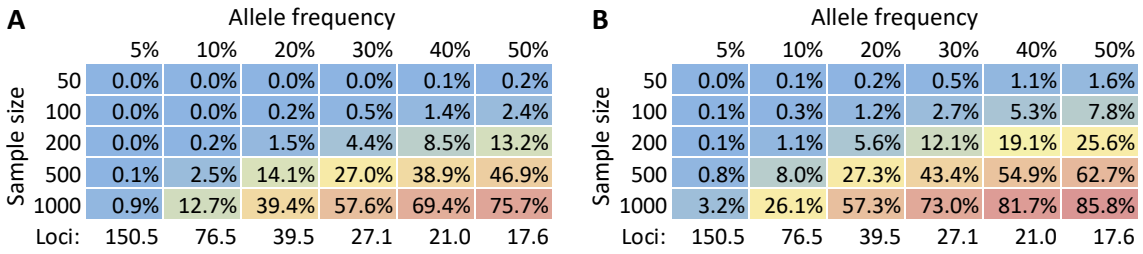
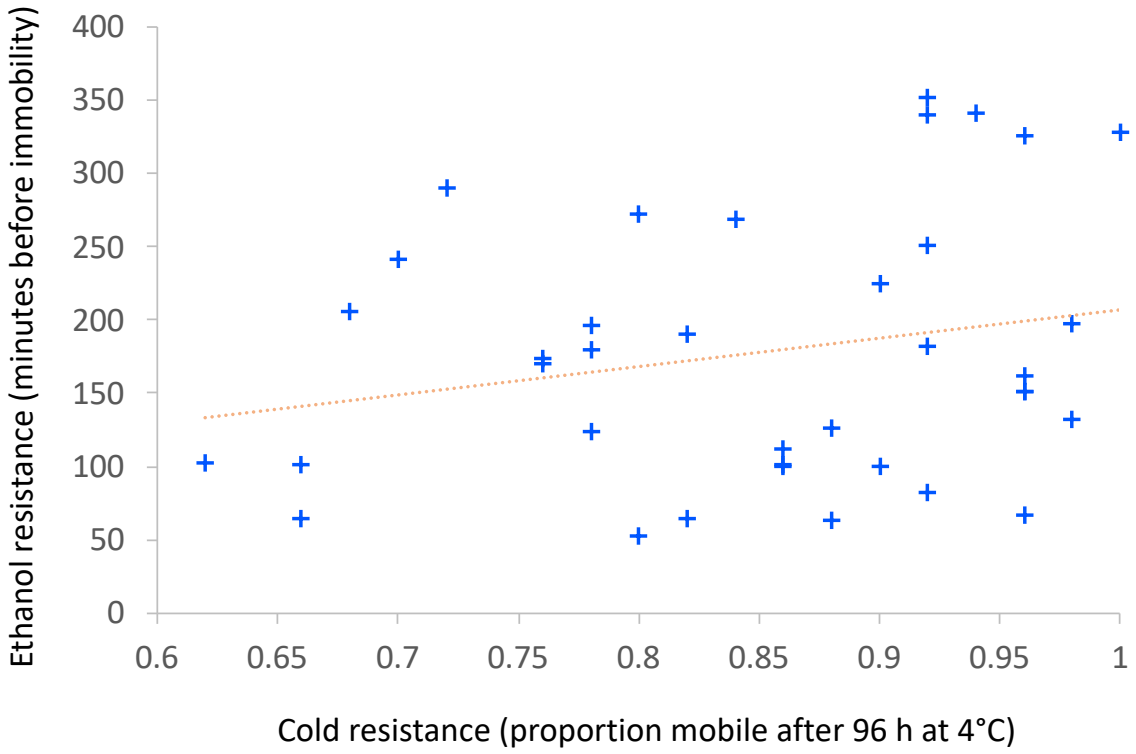


Figure 6. Heat map depicting power to detect genotype-phenotype associations after multiple testing correction, as a function of the frequency of causative variants and the sample size of individuals/strains, based on simulations (within a gamma shape parameter of 6 for the distribution of effect sizes) and 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.



823

824

825 **Figure 7**. A mild positive correlation between ethanol resistance and cold resistance

826 among 37 independent France inbred lines (r = 0.225; P = 0.0897), illustrating a potential

827 signal of pleiotropy between these traits.