



Supplementary Information for

Fitness effects but no balancing selection at the polymorphic *Adh* gene of *D. melanogaster*

Mohammad A. Siddiq, Joseph W. Thornton

Joseph W. Thornton

Email: joet1@uchicago.edu

This PDF file includes:

Supplementary text

Figs. S1 to S4

Tables S1 to S4

Other supplementary materials for this manuscript include the following:

Datasets S1

SI Methods

Enzyme Purification:

E. coli BL21 (DE3) Rosetta cells harboring expression plasmids were first grown overnight in 5 mL culture tubes at 37 degrees. These 5mL cultures were diluted into large 500 mL cultures and subsequently grown at 37 degrees with shaking for approximately 3 hours until reaching an $OD_{600} = \sim 0.6$, at which point protein expression was induced with 0.5 mM IPTG. The cultures were maintained at 37 degrees with shaking for approximately 3-4 hours until they reached an $OD_{600} = 1.7-2.0$. The cells were then harvested by centrifugation and the pellets were frozen at -20°C . The frozen pellets were lysed using Bper, DNaseI, and lysozyme and centrifuged at 10°C . The lysate was then passed over a nickel-charged HisTrapHP (GE Healthcare) column to isolate the (His₆)MBP-ADH fusion. The (His₆)MBP was removed by treating the fusion protein with tobacco etch virus (TEV) protease overnight and repassing the mixture over nickel-charged and cation exchange (GE HiTrap SP) columns. Protein samples were flash-frozen in 10% glycerol and stored at -80°C until they were ready to be used.

Transgenic Organisms:

Transgenic *D. melanogaster* that differ only at their *Adh* locus were made using the PhiC31-attP transgenesis system. Injection constructs were prepared with three different alleles: a naturally occurring Slow (K192) allele derived from a Canton-S fly strain, the Slow allele modified with the K192T mutation, and a naturally occurring Fast allele derived from a Florida strain (FL9, Bloomington Drosophila stock center #2374, provided by D. Loehlin), which differs from the Slow allele by the nonsynonymous change in codon 192 and several differences at synonymous, intronic, and 5' and 3' flanking sites. These constructs have identical boundaries and are the same ones used in refs. 1 and 2. They include the entire ADH coding sequence, the co-transcribed ADHR coding sequence, and 2.9 kb upstream and 1.6 kb downstream sequence, as described in ref. 3. The sequences for the Fast and Slow haplotype are provided in Supplemental Dataset 1 and were uploaded into GenBank as a part of sequences from ref. 1 (MH614199–MH614205). Constructs were transformed into *Adh*-null flies at the same attP landing site and bred to homozygosity (detailed in ref. 3). Insertion sequence and location were validated by PCR and Sanger sequencing. Specific injection strains and sequences of primers used to generate and validate these transformant lines have been described in detail in ref. 3.

Ethanol survivorship:

Adult ethanol tolerance was measured by placing adults aged 2 to 4 days into vials with Whatman paper soaked with 600ul of 3% sucrose solution and a variable ethanol concentration (0 – 10%). For each genotype, 4 to 8 replicates of 25 to 30 adults each were assayed at each ethanol concentration for each sex. The proportion of surviving flies after 48 hours was measured at each concentration. Larval ethanol tolerance was measured by picking larvae approximately at the second to third instar transition and placing them in vials containing *Drosophila* media (Genesee Scientific Molasses Formulation) with varying concentrations of ethanol (0-15%). For each genotype, 8-10 replicates of 30 larvae were assayed at each ethanol concentration, and the proportion of larvae surviving to eclosion recorded. LD₅₀ was estimated and the relationships between mortality and ethanol concentration, genotype, and sex were assessed using logistic regression and the quasibinomial link function, as implemented in glm() in R.

Heat stress:

The effect of heat stress on adult ethanol tolerance was measured by placing flies 2-4 days PE into vials with different concentrations of ethanol (0% or 5%) and various heat treatments. For each genotype, 4-8 replicates of 25-30 adults were assayed under each condition, and the proportion of adults surviving after 48 hours was recorded. For larvae, developmentally staged larvae were raised at 27°C , placed in vials with different concentrations of ethanol (0% or 8.5%)

and heat treatments (no treatment, chronic: 32°C for 48 hours, acute: 40°C for 30 minutes), and then returned to 22°C following the heat-treatments. 8-10 replicates of 30 larvae were assayed under each condition, and the proportion of larvae surviving to eclosion recorded. Effect and statistical significance of ethanol, genotype, heat-stress, and interaction terms were estimated using logistic regression in R; only interaction terms that significantly improved fit (*F*-test) were included in the final fitted model.

Supplementary References:

1. Loehlin DW, Ames AR, Vaccaro K, Carroll SB (2019) A major role for noncoding regulatory mutations in the evolution of enzyme activity. *Proceedings of the National Academy of Sciences* 116(25):12383–12389.
2. Siddiq MA, Loehlin DW, Montooth KL, Thornton JW (2017) Experimental test and refutation of a classic case of molecular adaptation in *Drosophila melanogaster*. *Nature Ecology and Evolution* 1(2):1–6.
3. Loehlin DW, Carroll SB (2016) Expression of tandem gene duplicates is often greater than twofold. *Proceedings of the National Academy of Sciences* 113(21):5988–5992.

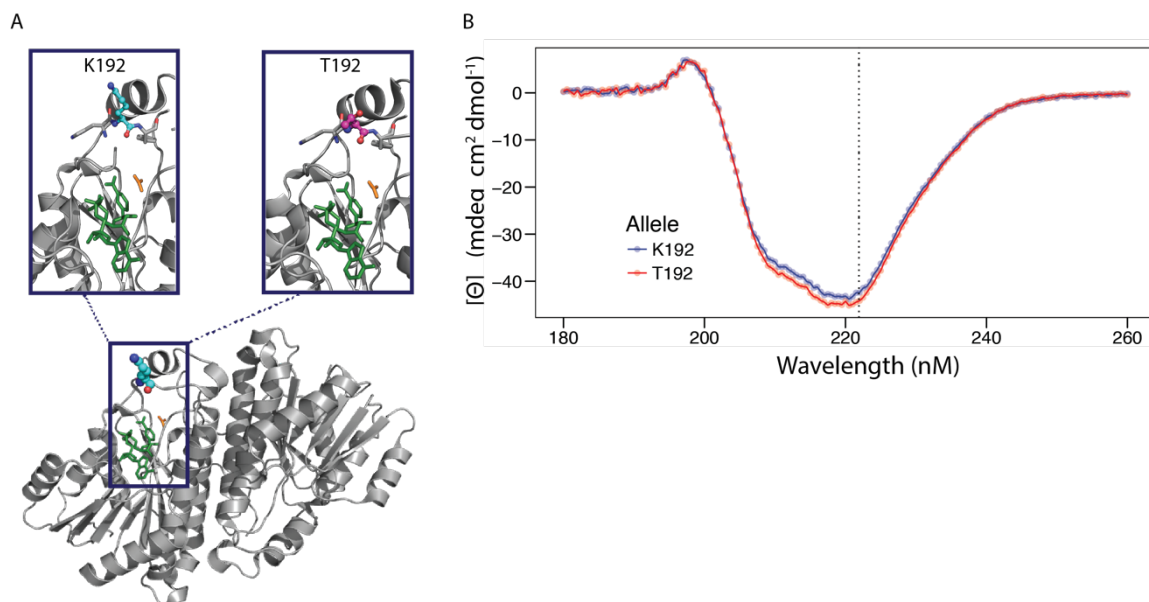


Fig. S1. Structure and stability of *Drosophila melanogaster* ADH. (A) Residue 192 does not contact the substrate or co-factor in the enzyme activity site nor change the secondary structure of the enzyme. X-ray crystal structure of ADH with K192 is shown with substrate (orange), co-factor (green), and residue 192 (spheres) for one of the monomers. Insets show magnification of the active site with residues 191-193 shown as sticks and residue 192 as ball and sticks with either a Lys (left, light blue backbone) or modeled Thr (right, magenta backbone). Very brief Method for modeling: The PyMol program was used to introduce the most probable rotamer of Thr192 onto the available *D. melanogaster* ADH crystal structure (PDB: 1MG5). Atoms for positions are colored by element: Blue—nitrogen, Red—oxygen. PDB: 1MG5. (B) Circular dichroism spectral profile of K192 (blue) and T192 (red) at 22 degrees. Dotted lines corresponds to 222nm, the wavelength monitored to fit melting curves and temperatures.

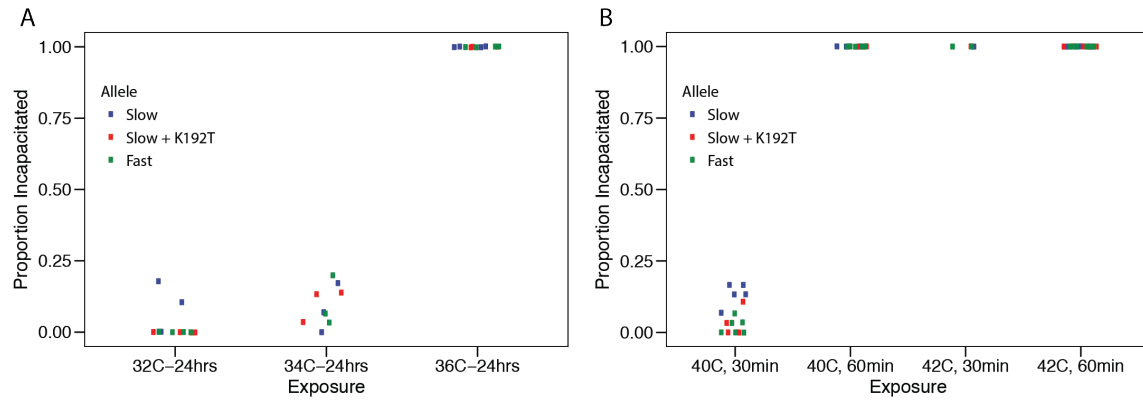


Fig. S2. Effect of heat-stress on transgenic adult flies. Populations of adult flies were treated with acute (*A*) or chronic (*B*) heat shocks and measured for proportion of flies incapacitated. Each point represents a measurement from a population of 25-30 adult flies.

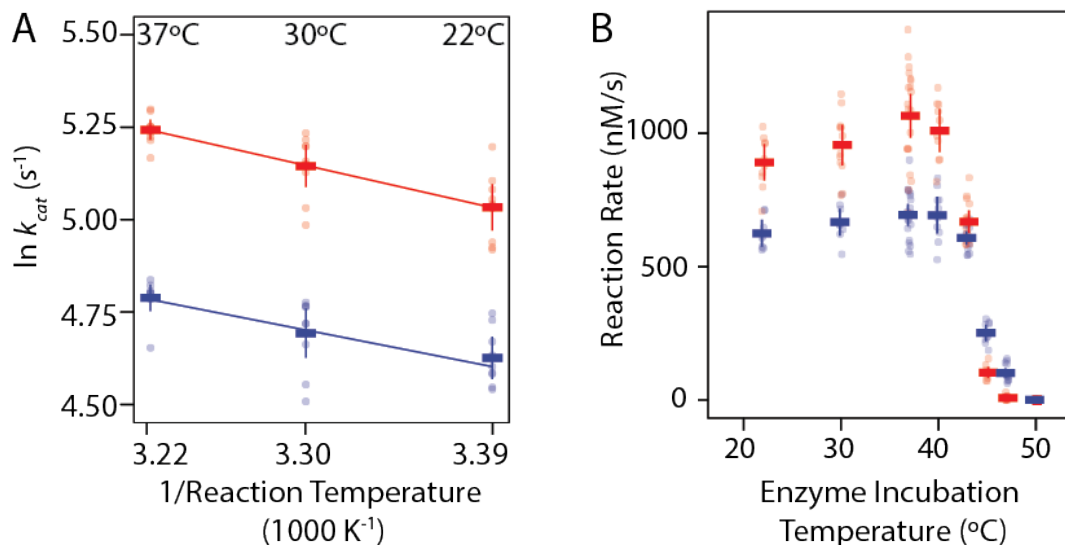


Fig. S3. Effect of K192T on enzyme activity against isopropanol. (A) The reaction mixture was incubated and reaction rates were measured out at the temperature listed. Solid line, best-fit linear regression using expected Arrhenius relationship between reaction rate and temperature. There is a significant effect of temperature (*F-test*, and $p < 0.01$) and genotype ($p < 0.01$), but no genotype by temperature interaction effect ($p = 0.56$). (B) Enzyme was incubated for 1 hour at the temperature plotted, and the reaction rate was then measured at 22 $^{\circ}C$.

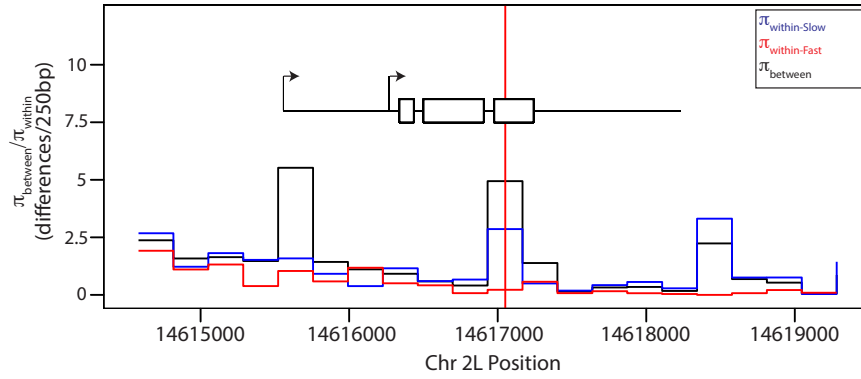


Fig. S4. Genetic variation within and between haplotypes at the *Adh* locus. Within- and between-class nucleotide diversity when 163 *D. melanogaster* haplotypes from Raleigh, NC, are classified by the amino acid at site 192. For each 250-bp window across the locus, the mean number of pairwise nucleotide differences is plotted for all pairs of observed sequences within each class (K192 class, blue, $\pi_{\text{within-Slow}}$; T192, $\pi_{\text{within-Fast}}$; red). Mean pairwise diversity between classes is also shown ($\pi_{\text{between/within}}$; black). The structure of the *Adh* locus is shown, with exons (rectangles), adult and larval transcriptional start sites (arrows). Red vertical line, amino acid 192.

Table S1: Logistic regression analysis of the effects of ethanol and genotype on larval mortality at 22°C and 27°C. Coefficient (with standard error) shows the estimated effect of each variable on the log-odds of mortality, relative to the baseline treatment of larvae with the Slow genotype at 0% ethanol. Fast, entire Fast *Adh* allele; Slow+K192T, Slow genotype modified to contain Thr at amino acid site 192. *t*, standardized coefficient. *P*, probability of observing *t* of equal or greater absolute value under the null hypothesis of no effect.

22°C

Variable	Coefficient	Std. error	<i>t</i>	<i>P</i>
(Intercept)	-4.45423	0.27661	-16.103	<2e-16
Ethanol	0.56524	0.03206	17.629	<2e-16
Slow + K192T	-0.09063	0.17130	-0.529	0.5973
Fast	-0.33516	0.17287	-1.939	0.0539

27°C

Variable	Coefficient	Std. error	<i>t</i>	<i>P</i>
(Intercept)	-2.67217	0.13972	-19.125	< 2e-16
Ethanol	0.38830	0.01704	22.784	< 2e-16
Slow + K192T	-0.33320	0.09885	-3.371	0.000844
Fast	-0.32398	0.09960	-3.253	0.001267

Table S2: Logistic regression analysis of the effects of ethanol, genotype, and sex on adult mortality at 22°C and 27°C. Coefficient (with standard error) shows the estimated effect of each variable on the log-odds of mortality, relative to the baseline treatment of females with the Slow genotype at 0% ethanol. Fast, entire Fast *Adh* allele; Slow+K192T, Slow genotype modified to contain Thr at amino acid site 192. *t*, standardized coefficient. *P*, probability of observing *t* of equal or greater absolute value under the null hypothesis of no effect.

22°C

Variable	Coefficient	Std. error	<i>t</i>	<i>P</i>
(Intercept)	-7.4078	0.8427	-8.791	4.52e-16
Ethanol	1.2829	0.1346	9.534	< 2e-16
Slow + K192T	-0.2288	0.3176	-0.720	0.4721
Fast	-2.2636	0.3729	-6.070	5.62e-09
Male	-0.6964	0.2747	-2.535	0.0119

27°C

Variable	Coefficient	Std. error	<i>t</i>	<i>P</i>
(Intercept)	-5.5011	0.7138	-7.707	3.28e-13
Ethanol	1.2636	0.1294	9.764	< 2e-16
Slow + K192T	-0.1303	0.3366	-0.387	0.69892
Fast	-1.7634	0.3426	-5.147	5.47e-07
Male	-1.0030	0.2771	-3.619	0.00036

Table S3. Logistic regression analysis of the effects of ethanol treatment, heat-shock, and genotype on mortality of larvae. Coefficient (with standard error) shows the estimated effect of each variable on the log-odds of mortality relative to the baseline of larvae with the Slow genotype, 0% ethanol, and no heat shock. Fast, entire Fast *Adh* allele; Slow+K192T, Slow genotype modified to contain Thr at amino acid site 192. *t*, standardized coefficient. *P*, probability of observing *t* of equal or greater absolute value under the null hypothesis of no effect. Interaction terms included in the model were those that significantly improved the fit of the model to the data (*F*-test).

Variable	Coefficient	Std. Error	<i>t</i>	<i>P</i>
(Intercept)	-1.98092	0.18639	-10.628	< 2e-16
Ethanol treatment (8.5%)	2.22182	0.19218	11.561	< 2e-16
Acute heat shock	1.04934	0.20971	5.004	1.32e-06
Chronic heat shock	2.88008	0.20328	14.168	< 2e-16
Slow + K192T	-0.50852	0.09975	-5.098	8.60e-07
Fast	-0.48078	0.10107	-4.757	4.00e-06
Ethanol treatment:Acute heat shock	-1.17226	0.23997	-4.885	2.27e-06
Ethanol treatment:Chronic heat shock	-1.27700	0.25023	-5.103	8.39e-07

Table S4. Logistic regression analysis of the effects of ethanol, heat-shock, genotype, and sex on mortality of adult flies. Coefficient (with standard error) shows the estimated effect of each variable on the log-odds of mortality relative to the baseline of females with the Slow genotype, 0% ethanol, and no heat shock. Fast, entire Fast *Adh* allele; Slow+K192T, Slow genotype modified to contain Thr at amino acid site 192. *t*, standardized coefficient. *P*, probability of observing *t* of equal or greater absolute value under the null hypothesis of no effect. Interaction terms included in the model were those that significantly improved the fit of the model to the data (*F*-test).

Variable	Coefficient	Std. Error	<i>t</i>	<i>P</i>
(Intercept)	-4.2476	0.5566	-7.632	2.56e-12
Ethanol treatment (5%)	2.1826	0.5594	3.902	0.000144
Acute heat shock	-0.6162	1.0459	-0.589	0.556631
Chronic heat shock	-1.4064	1.6894	-0.833	0.406442
Slow + K192T	-0.0349	0.1838	-0.190	0.849648
Fast	-1.0615	0.2059	-5.155	7.93e-07
Male	0.2337	0.1574	1.485	0.139718
Ethanol treatment:Acute heat shock	1.0607	1.0665	0.995	0.321548
Ethanol treatment:Chronic heat shock	4.7369	1.7004	2.786	0.006036