nature ecology & evolution

ARTICLES

PUBLISHED: 13 JANUARY 2017 | VOLUME: 1 | ARTICLE NUMBER: 0025

Experimental test and refutation of a classic case of molecular adaptation in *Drosophila melanogaster*

Mohammad A. Siddiq[1], David W. Loehlin[2,3], Kristi L. Montooth[4] and Joseph W. Thornton[1,5*]

Identifying the genetic basis for adaptive differences between species requires explicit tests of historical hypotheses concerning the effects of past changes in gene sequence on molecular function, organismal phenotype and fitness. We address this challenge by combining ancestral protein reconstruction with biochemical experiments and physiological analysis of transgenic animals that carry ancestral genes. We tested a widely held hypothesis of molecular adaptation—that changes in the alcohol dehydrogenase protein (ADH) along the lineage leading to *Drosophila melanogaster* increased the catalytic activity of the enzyme and thereby contributed to the ethanol tolerance and adaptation of the species to its ethanol-rich ecological niche. Our experiments strongly refute the predictions of the adaptive ADH hypothesis and caution against accepting intuitively appealing accounts of historical molecular adaptation that are based on correlative evidence. The experimental strategy we employed can be used to decisively test other adaptive hypotheses and the claims they entail about past biological causality.

A central goal of molecular evolutionary biology is to identify the genes and biological mechanisms that mediated historical adaptation. Rigorously testing hypotheses in this area has been a major challenge. Many studies infer past selection from statistical signatures in genes that are involved in biological processes that might have suited species to their environments[1–4]. But sequence signatures of selection can be forged by chance or demographic processes and it is difficult to predict from sequence alone how genetic changes affect phenotypes and fitness[5–8]. Compelling evidence for molecular adaptation therefore requires formulating and testing explicit hypotheses about the causal links between specific evolutionary changes in gene sequence and the resulting changes in molecular function, organismal phenotype and fitness[6–10]. Advances in genetic mapping, experimental studies of molecular function and transgenic engineering have allowed hypotheses of molecular adaptation between recently diverged populations to be tested with increasing rigour[11–16]. But hypotheses about adaptive divergence between species or at higher taxonomic levels are explicitly historical, so testing them requires the effect of genetic changes that occurred on phenotype and fitness in specific evolutionary lineages from the distant past to be measured. Here we address this challenge by combining ancestral protein reconstruction[17] with biochemical experiments and physiological analysis of transgenic animals that carry ancestral genes.

We applied this approach to a longstanding hypothesis of molecular adaptation—that changes in the alcohol dehydrogenase (ADH) protein of the fruit fly *Drosophila melanogaster* increased the catalytic activity of the enzyme and thereby contributed to the adaptation of the species to its ethanol-rich ecological niche[18–20]. This hypothesis was articulated decades ago[18,21] and became widely accepted[19,20,22,23] on the basis of several observations that were consistent with it, but did not directly address the putative causal links among historical changes in protein sequence, function and fitness. First, *D. melanogaster* evolved to colonize ethanol-rich habitats in rotting fruit after it split from its sister species, *D. simulans*, some two to four million years ago[24,25]. Second, fractionated cell extracts from *D. melanogaster* catalyse alcohol turnover more rapidly than those from *D. simulans*[18,26,27]. Third, the first-ever application of the McDonald–Kreitman (MK) test detected an excess of non-synonymous substitutions in an alignment of the ADH coding sequences of *D. melanogaster* and closely related species[28], which was interpreted as evidence for adaptive evolution driving the divergence of the ADH protein between *D. melanogaster* and *D. simulans*[21,22,29,30]. These observations were integrated into a narrative in which adaptation to ethanol-rich habitats was driven by selection on the ADH protein sequence for increased catalytic activity. Other factors—particularly increases in the expression level[26,31–33] of ADH, changes at other genetic loci[34–36] and within-species polymorphisms[37–39]—also probably contributed to ethanol adaptation in *D. melanogaster*, but they are independent of and cannot explain the selection signature on the protein-coding sequence of the ADH enzyme found in the MK test.

We focused on the hypothesis of adaptive ADH protein evolution because it is widely accepted on the basis of correlated forms of variation in extant species and because it is particularly amenable to testing using the experimental approaches of ancestral reconstruction, biochemical characterization and engineering of transgenic organisms. The ADH adaptive hypothesis entails specific, testable predictions about how genetic changes that occurred in the ADH protein sequence during the historical divergence of *D. melanogaster* affect the phenotype at several levels, including molecular function (catalytic turnover of ethanol by pure ADH protein), physiology (ethanol catabolism in the tissues of *D. melanogaster*) and fitness components (survival in the presence of ethanol) (Fig. 1a). We tested these predictions by reconstructing the ADH protein from the last common ancestor of *D. melanogaster* and *D. simulans* (AncMS) and experimentally characterizing how changes in ADH

[1]Department of Ecology and Evolution, University of Chicago, Chicago, Illinois, USA. [2]Laboratory of Cell & Molecular Biology, University of Wisconsin-Madison, Madison, Wisconsin, USA. [3]Howard Hughes Medical Institute, University of Wisconsin-Madison, Madison, Wisconsin, USA. [4]School of Biological Sciences, University of Nebraska, Lincoln, Nebraska, USA. [5]Department of Human Genetics, University of Chicago, Chicago, Illinois, USA. *e-mail: joet1@uchicago.edu

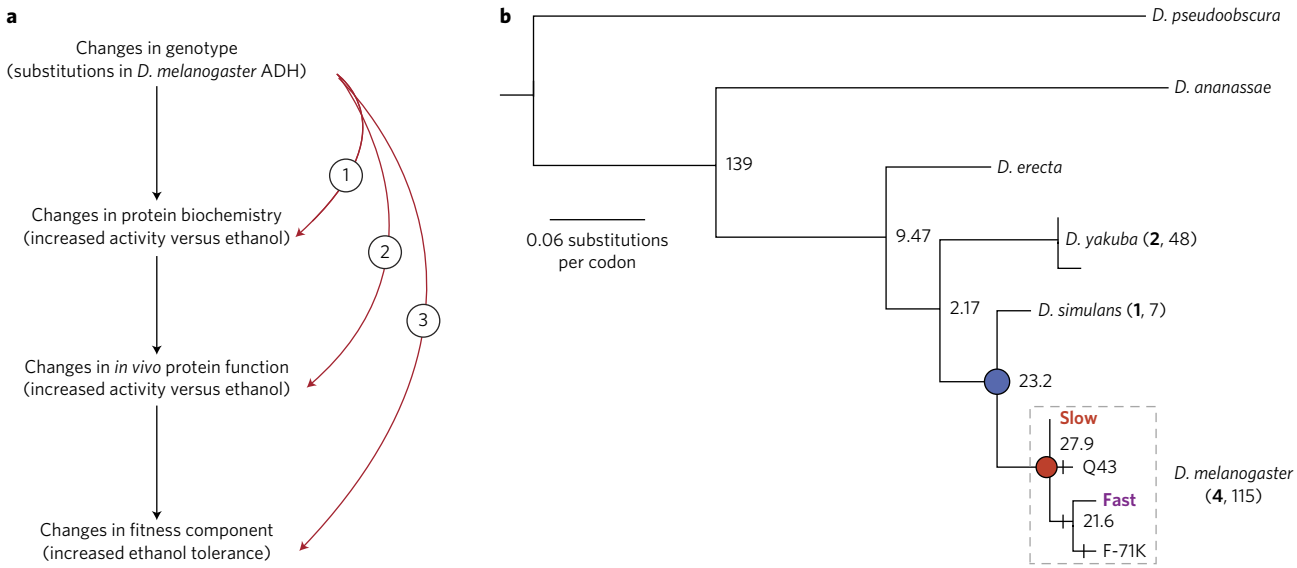

Figure 1 | Predictions of the classic hypothesis of ADH adaptive evolution. a, Hypotheses of molecular adaptation entail putative causal links (black arrows) between evolutionary change in genes and effects on molecular function, organismal phenotype and fitness. The specific links that comprise the classic hypothesis of ADH adaptive evolution in *D. melanogaster* are in parentheses. Red arrows represent testable predictions entailed by this hypothesis, including effects on enzyme activity (1), physiology (2) and fitness components (3). **b**, Maximum-likelihood phylogeny of ADH protein sequences. The classic ADH adaptive hypothesis predicts functional divergence on the branch connecting AncMS (blue) to the ancestral *D. melanogaster* ADH allele (red). The box shows polymorphic ADH protein variants in *D. melanogaster*, including the slow allele (which is identical in sequence to the ancestral *D. melanogaster* sequence) and the derived fast allele. The number of distinct segregating protein alleles (bold numbers) followed by the number of sampled alleles is displayed in parentheses for species with available polymorphism data. Node labels show statistical support as approximate likelihood ratios.

sequence along the *D. melanogaster* lineage affected ADH function, physiology and fitness.

Results

We generated a large alignment of ADH sequences, determined the best-fit evolutionary model, inferred the maximum likelihood phylogeny and calculated the posterior probability distribution of amino acid states at key ancestral nodes. We synthesized coding sequences for the maximum a posteriori sequence of AncMS, which was inferred with high confidence and only one ambiguously reconstructed amino acid (Fig. 1b, Supplementary Fig. 1), and for an alternative version of AncMS (Alt-AncMS), which contained the other plausible state at the ambiguous site and was identical to *D. simulans* ADH. We also characterized the inferred ancestral *D. melanogaster* ADH, the amino acid sequence of which is identical to that of the 'slow' allele present in extant populations, which is known to be older than other ADH variants[40]. The adaptive ADH hypothesis predicts differences in ethanol catalysis between the AncMS ADH and the ancestral *D. melanogaster* ADH. In addition, we characterized the 'fast' allele, a more recently derived ADH variant, to determine whether the assays we used were sensitive enough to detect previously identified phenotypic differences thought to be of selective importance in some natural populations of *D. melanogaster*[8,38,41].

We first tested the prediction that genetic change in the ADH protein along the *D. melanogaster* lineage should enhance ethanol catabolism *in vitro*. Unlike the studies performed decades ago on fractionated homogenates from present-day flies, we were able to directly measure the functional effects of specific historical changes in protein sequence by using heterologously expressed ancestral proteins and improved methods for purification and quantification. We found, contrary to the prediction of the adaptive ADH hypothesis, that both the maximal catalytic turnover rate of ethanol per enzyme molecule (k_{cat}) and the Michaelis–Menten constant (K_m; a measure of the performance of the enzyme when substrate concentration is limiting) were indistinguishable among AncMS, *D. melanogaster* and *D. simulans* ADH proteins (Fig. 2). The assay was sensitive enough, however, to detect the expected increase in ADH catalytic function of the fast allele. When enzyme activity was measured using isopropanol (a higher-activity ADH substrate not thought to be ecologically important) we again observed no difference between the ancestral and *D. melanogaster* alleles, whereas the fast allele displayed enhanced activity (Supplementary Fig. 2; Supplementary Table 1).

Second, the adaptive ADH hypothesis predicts that sequence evolution in *D. melanogaster* should enhance ethanol catabolism *in vivo*. Differences in solubility, translational efficiency or accuracy, post-translational modifications, stability or the presence of other cellular co-factors could cause ADH proteins to behave differently when produced *in vivo*. To test whether divergence of the ADH protein sequence caused biochemical differences in ethanol catabolism *in vivo*, we genetically transformed Adh-null *D. melanogaster* with ancestral or extant ADH alleles that differed only in their amino acid sequences. We raised these transgenic flies to adulthood and measured the catabolism of ethanol by homogenates from each genotype under maximum velocity conditions. Contrary to the prediction of the adaptive hypothesis, homogenates from flies expressing the *D. melanogaster* ADH allele did not have higher rates of ethanol turnover than AncMS or Alt-AncMS. Again, the derived fast allele was associated with significantly faster ethanol turnover (Fig. 3a).

Finally, the adaptive ADH hypothesis predicts that divergence of the ADH protein along the *D. melanogaster* lineage should improve fitness by increasing survival in ethanol-rich environments. We found that at both larval and adult stages, transgenic flies carrying AncMS, Alt-AncMS or *D. melanogaster* ADH alleles had statistically indistinguishable ethanol tolerance, measured as the dose of ethanol that caused a 50% probability of death (LD_{50}) (Fig. 3b,c). In contrast, the fast allele conferred higher ethanol tolerance in larvae (Fig. 3b). Thus, divergence of the ADH protein sequence along the *D. melanogaster* lineage had no detectable effect on survival in the presence of ethanol.

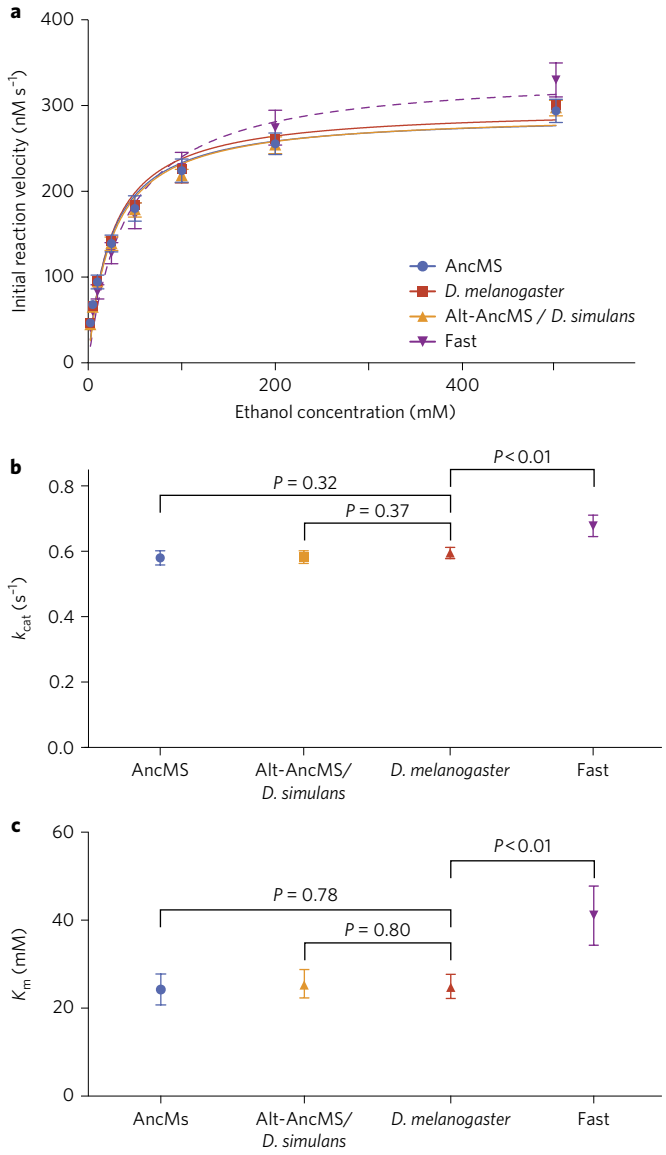

Figure 2 | Effects of ADH sequence divergence on the activity of purified enzymes. *In vitro* assays of bacterially expressed protein show no divergence in catalytic properties of ADH between the ancestral form (AncMS) and the *D. melanogaster* protein. Alternative reconstructed sequence (Alt-AncMS, identical in sequence to *D. simulans* ADH) and the derived fast allele are also shown. **a**, Initial reaction velocity across ethanol concentrations. Points and error bars show the mean and 95% confidence interval of nine measurements at each concentration. **b**, Estimated k_{cat} for each allele. **c**, Estimated K_m for each allele. In **b** and **c**, points and error bars show the estimated parameter and 95% confidence interval calculated by nonlinear regression from the data in **a**. *P* values for differences are from likelihood ratio tests that compare a global model in which a single value of the parameter of interest is estimated from the data for both genotypes versus a free model with separately estimated parameters.

These experiments indicate that historical substitutions in the ADH coding sequence along the *D. melanogaster* lineage caused none of the predicted effects on biochemical function, physiology or fitness components, refuting the widely held hypothesis of adaptive ADH divergence. Why then did the original statistical analysis[28] of the ADH coding sequence suggest positive selection? We considered two possibilities. First, the inference of positive selection might have been a stochastic error due to sparse sampling of polymorphisms; we therefore repeated the MK test using a much expanded contemporary data set[42], with greater sampling of polymorphism. We found that the signature remained (Supplementary Table 2). Second, the signature of selection might come from lineages other than *D. melanogaster*, because the MK test in its standard form does not apportion sequence changes onto phylogenetic lineages. We therefore conducted a polarized MK test on the expanded data set by assigning substitutions to specific branches on the phylogeny; we also conducted a standard MK test, but with individual species removed. We found no signature of positive selection on the *D. melanogaster* lineage and removing *D. melanogaster* from the analysis did not affect the MK result (Fig. 4a; Supplementary Fig. 3; Supplementary Table 2). In fact, there was only one non-synonymous substitution along the putatively adaptive *D. melanogaster* branch, at N-terminal residue 1 of the mature protein, in a solvent-exposed loop far from the active site (Fig. 4b). The detected signature of selection came primarily from the lineage leading to *D. yakuba*, where we observed a marginally significant excess of non-synonymous divergence ($P = 0.047$). Whether this result reflects adaptive evolution, relaxed constraint, sampling error or drift is unknown. The ethanol tolerance of *D. yakuba* is no different from that of closely related species and is lower than that of *D. melanogaster*[34,43].

Discussion

A strength of the ADH adaptive hypothesis was that it entailed specific predictions about the effects of genetic divergence along the lineage leading to *D. melanogaster* on protein function, organismal phenotype and components of fitness. Ancestral sequence reconstruction, engineering of transgenic organisms, and biochemical/physiological assays allowed us to test these predictions directly. Our experiments show that none of these predictions hold.

We did not test any of the innumerable other hypotheses that have been or could be proposed concerning fruit fly adaptation to rotting fruit or ADH evolution. For example, evidence suggests that the increased ethanol tolerance of *D. melanogaster* may have evolved because of substitutions at other loci[34,36] or in regulatory regions[32,33] of *Adh* and it is possible that these changes were positively selected. The single amino acid replacement that occurred along the *D. melanogaster* lineage could have affected functions other than ethanol catabolism, such as the breakdown of other substrates, and, if it did, these changes may or may not have increased fitness. For these or any other claims of molecular adaptation, further work would be required to formulate specific adaptive hypotheses and test their causal predictions.

Our experiments also provide information relevant to a different question: how ADH alleles segregating in present-day *D. melanogaster* populations affect fitness. Our data show that the amino acid polymorphism distinguishing the fast and slow alleles does confer measurable differences in ethanol catabolism and ethanol tolerance, even in the absence of other linked and functionally important genetic variants[44]. These results provide an initial corroboration of the hypothesis that the fast/slow polymorphism in the protein sequence is biologically and ecologically important in present-day populations[38,41,45,46]. The differences in ethanol tolerance that we observed between transgenic flies carrying fast and slow coding alleles, however, were small relative to the large range of ethanol tolerances observed among *D. melanogaster*[34]; further, the amino acid changes in ADH were not sufficient to explain the extent of variation in ethanol tolerance within this species. Additional work is required to propose and test specific causal hypotheses about why these alleles are distributed in clines that correlate with latitude[38,46,47] and why the polymorphism is balanced in *D. melanogaster*[45].

The strategy we employed may be useful in efforts to increase the rigour of scientific inferences about adaptation[6,7]. A hypothesis of molecular adaptation is a conjecture that particular changes in genotype during history caused particular evolutionary changes in phenotype that enhanced fitness: a signature of selection in a gene sequence

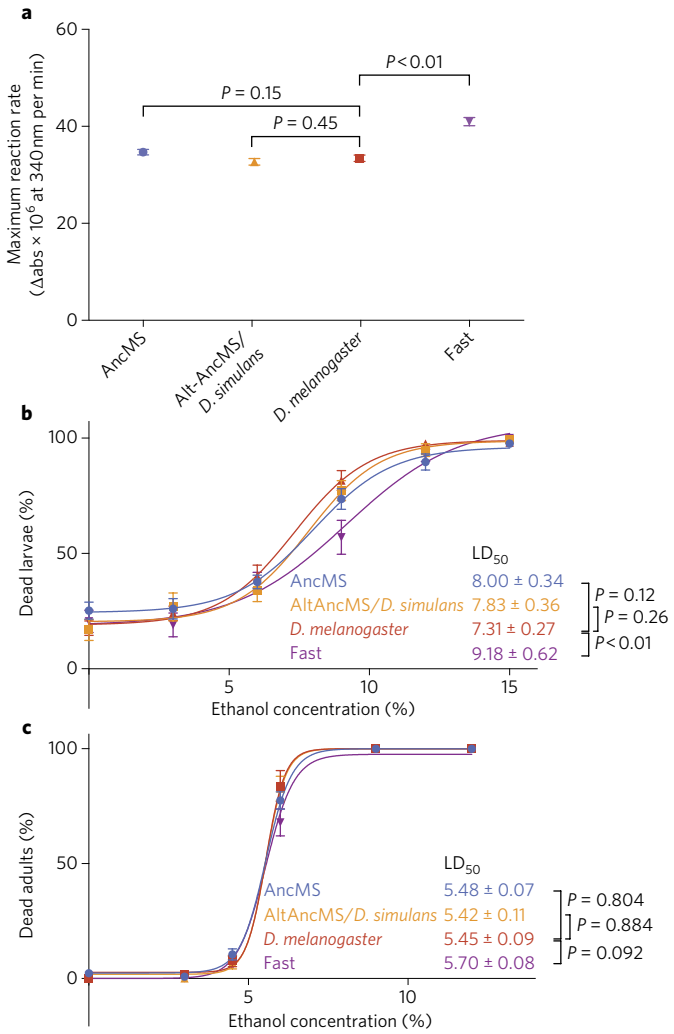

Figure 3 | Effects of ADH sequence divergence on ethanol catabolism and fitness in transgenic flies. Adh-null *D. melanogaster* were genetically transformed to express coding sequences of ancestral or extant ADH proteins; genotypes were otherwise identical. **a**, Ancestral and *D. melanogaster* ADH alleles do not confer differences in ADH catabolism. Animals of each transgenic genotype were homogenized and the soluble fraction assayed for ethanol turnover rate under saturating substrate conditions. The graph shows maximum reaction rate normalized per milligram of total protein content of the homogenate. Points and error bars show the mean and standard error of the mean of 30 replicate homogenates. *P* values are from *t*-tests for differences in means for genotype pairs of interest (see Methods). **b**,**c**, Effect of ADH genotype on ethanol tolerance. Transgenic larvae (**b**) and adults (**c**) were assayed for survival in the presence of increasing ethanol concentration. Points and error bars show mean and standard error of the mean. For larvae, ten replicate groups were measured at each dose, except for in the fast genotype, which had eight replicate groups. For adults, there were four replicate groups at 0%, 9% and 12% ethanol, six replicate groups at 3% and 6% ethanol, and eight replicate groups at 4.5% ethanol. The estimated LD_{50} (adjusted for baseline mortality) and 95% confidence interval is shown for each genotype. *P* values for comparisons are from likelihood ratio tests comparing a global model in which a single LD_{50} is fit to the pooled data from both genotypes to a free model with an independent LD_{50} for each genotype.

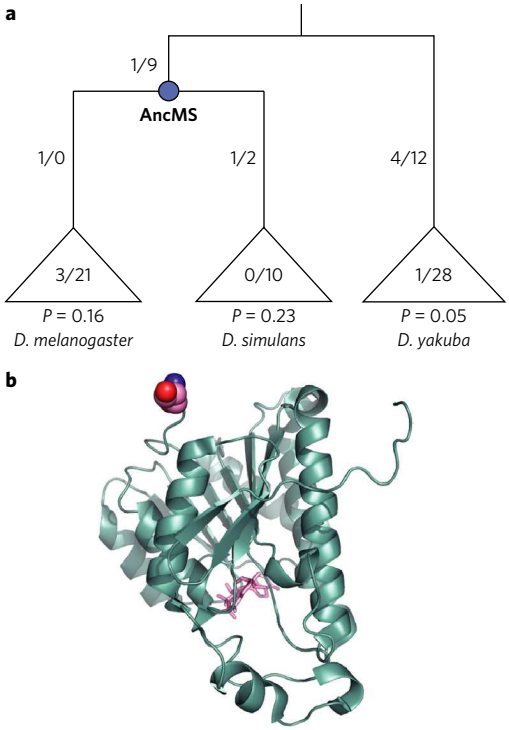

Figure 4 | Sequence evolution on the phylogeny of *D. melanogaster* and closely related species. a, The signature of adaptive evolution in the ADH coding sequence is not caused by substitutions on the *D. melanogaster* lineage. Substitutions were assigned to specific lineages on the basis of maximum-likelihood ancestral reconstructions at each node. Labels show non-synonymous/synonymous substitutions (on branches) and polymorphisms (in triangles). MK tests were conducted separately for each species, using substitutions that occurred on the branch leading to each species and polymorphisms in extant populations of that species. *P* values show the significance of the test for each species. **b**, The non-synonymous substitution that occurred during the divergence of *D. melanogaster* from AncMS. The structure of *D. melanogaster* ADH (Protein Data Bank 1MG5) is shown in cyan. The spheres show the Ala1Ser substitution near the N-terminus of the protein, far from the active site and substrate (pink sticks).

may suggest such a hypothesis but cannot test it. The case of ADH shows that the existence of variation between present-day species in genotype, phenotype and fitness is also insufficient to test a hypothesis of molecular adaptation, even if it is consistent with it, because covariation alone does not demonstrate the hypothesized causal links among these forms of variation or establish the historical direction of the evolutionary trajectory that produced them. We were able to directly test the predictions of the adaptive ADH hypothesis by combining ancestral protein reconstruction with biochemical studies of recombinant proteins and transgenic engineering of organisms carrying ancestral alleles. A similar approach could be applied to test many other adaptive hypotheses. This strategy has some limitations: not all ancestral sequences can be reconstructed with confidence, only some phenotypes can be characterized experimentally and laboratory experiments cannot detect all fitness differences. Furthermore, manipulative experiments can never account for the full range of genomic and environmental variables that affect the biology and evolution of an organism. Some notions concerning adaptation will therefore remain difficult to study rigorously. Nevertheless, because of technical and conceptual advances, it should now be possible to experimentally assess the causal predictions of many previously untested or weakly tested hypotheses of historical molecular adaptation, allowing them to be corroborated or, like the classic hypothesis of ADH divergence in *D. melanogaster*, decisively refuted.

Methods

Phylogenetics and inference of ancestral sequences. Coding DNA sequences for species in the *D. melanogaster* group, as well as from outgroup species *D. pseudoobscura*, were obtained from Genbank, the DPGP2 consortium[42] and D. Matute (University of North Carolina). DNA coding sequences of alleles that differed in their protein sequence were aligned using MUSCLE and a maximum likelihood phylogeny was inferred in PhyML (v. 3.0) using the best-fit parameters and model TrN + G (Tamura-Nei with gamma-distributed among-site rate variation), as determined by the Akaike information criterion (jModelTest software, v. 2.1.7). Ancestral sequence reconstruction was performed using the maximum likelihood method[48] in PAML software (v. 4.8); sequences were analysed using the GY94 general codon model, with model = 0, nssites = 1, 3 × 4 codon frequencies and a transition/transversion ratio inferred from the data. The posterior probability distribution of ancestral states at each site was analysed at nodes that correspond AncMS and to the last common ancestor of all *D. melanogaster* alleles. Sites were considered ambiguously reconstructed if two or more states had posterior probability >0.2.

Synthesis, expression and purification of ADH alleles. For bacterial expression of ADH proteins, the coding sequence of the *D. melanogaster* slow ADH allele was generated by *de novo* synthesis (GenScript). Coding sequences for other alleles were generated by site-directed mutagenesis of the slow sequence using the QuickChange method (Stratagene) and verified by Sanger sequencing. Coding sequences were cloned into pLIC-maltose binding protein (MBP) plasmids to yield fusion proteins with the maltose binding protein and an N-terminal hexahistidine tag. Plasmids were verified by sequencing and transformed into *E. coli* BL21(DE3) Rosetta cells. Cells were grown at 37 °C and expression was induced using 1 mM isopropyl B-D-thiogalactoside (IPTG) at $OD_{600} = 0.6$. Cells were harvested by centrifugation after reaching OD_{600} of 1.5–1.8 and then frozen. To purify proteins, cells were lysed using B-PER, lysozyme and DNAse I. Lysate was passed over a nickel-affinity HIS-trap chromatography column to isolate the MBP/ADH protein. The MBP tag was removed by treating with sample tobacco etch virus (TEV) protease overnight and then the ADH protein was purified using HisTrap and cation columns. Purified ADH proteins were flash frozen in 10% glycerol solution and stored at −80 °C until they were ready to be characterized.

Transgenic organisms. To make *D. melanogaster* flies carrying *Adh* alleles that differ only in the amino acid sequences they code for, we first generated the Adh gene variants *in vitro* as described below and then transformed these constructs into flies using the ϕC31-attP transgenesis system. Primer sequences are given in Supplementary Table 4. First, a 7.8 kb segment containing the *Adh*-slow allele and all known *cis*-regulatory elements was amplified from genomic DNA of *D. melanogaster* strain Canton-S; this segment contained the entire transcriptional unit including ADH and ADHR coding sequences with their introns, plus untranscribed sequences extending 2.9 and 1.6 kb in the 5′ and 3′ directions, respectively. This PCR product was gel extracted and ligated into the AscI and NotI sites of the attB vector, pS3aG. This vector was then modified to facilitate further cloning of alternate ADH coding sequences by amplifying the vector by PCR with primers that incorporated BspQ1 restriction sites at the boundaries of the coding region; digesting with BspQ1 removed the coding region and allowed replacement with a new coding region. Variant coding sequences were produced by PCR amplification of the Canton-S slow allele ADH coding sequence (including its introns) in overlapping pieces using primers containing the desired non-synonymous mutations, then assembling the fragments in pGem-T-Easy (Promega) using Gibson Assembly Master Mix (New England Biolabs), producing a full-length amplicon of this variant coding sequence (with introns) by PCR and then inserting the amplicon into the vector containing the flanking sequences by Gibson asesmbly. Sequences of amplicons at each stage and of all final vectors were verified by Sanger sequencing. This process produced transformation vectors that coded for the ancestral *D. melanogaster*, AncMS, Alt-AncMS and fast protein alleles but were otherwise identical. Plasmid DNA for injection was prepared using the Nucleobond Xtra Midi Plus EF kit (Macherey-Nagel) and adjusted to 1 μg μl^{-1}.

Constructs were injected into the inbred recipient strain[49], 'pf86', which is null for *Adh* and contains the attP landing site *ZH-86Fb* and the phiC31 integrase (genotype: y[1] M{vas-int.Dm}ZH-2A w[*]; Adh[fn6] cn[1]; M{3xP3-RFP.attP} ZH-86Fb). Injected G0 flies were backcrossed to the pf86 strain. F1 transformants carried the w^+ allele and were identifiable by eye colour. These transformants were crossed to w^- sibs and transformant lines were made homozygous. Lines were tested for correct insertions via PCR (for primers, see Supplementary Table 4). At least two independent transformation strains were generated for each *Adh* genotype.

Enzyme assays. For enzymes purified from bacteria, the activity of 500 nM ADH enzyme was characterized in a solution containing 1 mM nicotinamide adenine dinucleotide (NAD), 50 mM sodium phosphate (pH 7.6) and ethanol or isopropanol concentrations of 2.5, 5, 10, 25, 50, 100, 200 and 500 mM, with three replicate reactions at each concentration. The rate of reaction was measured every 30 s by monitoring absorbance at 340 nm, which corresponded to the concentration of NADH, a byproduct of ethanol oxidation. The first five observations for each reaction were used to estimate the initial velocity. This procedure was repeated three times on separate days. Data were pooled and the best-fit values of K_m and k_{cat} were estimated using the MM nonlinear regression function in GraphPad Prism 7.0. The differences between parameters associated with different ADH genotypes were assessed using the extra sum of squares *F*-test as implemented in GraphPad Prism 7.0, which uses a likelihood ratio test to compare the likelihood of a model with a globally fitted parameter (k_{cat} or K_m) to one in which the parameter is fit individually to each genotype.

ADH enzyme activity from transgenic flies was measured from crude fly homogenates using the 'manual grinders' protocol[48]. For each transformation strain, we propagated three replicate cultures (broods), each of which was initiated by placing five females and two males in a vial of yeast-free food to lay eggs for two to three days. After pupation and eclosion, all 0- to 24-hours-old adult males were transferred to a fresh vial; this procedure was repeated on five separate days, yielding 15 replicate vials of flies for each transformation strain. When males in a vial reached four days old, two flies were collected and homogenized using a Potter–Elvehjem homogenizer in 400 μl of 0.1 M sodium phosphate buffer, pH 8.6, then centrifuged at 21000 × *g* for 5 min at 4 °C; supernatant from the homogenate of each vial was split among three replicate enzyme assays and three replicate protein concentration assays. Assays for ADH enzyme activity and protein concentration (Quant-IT protein assay, Thermo-Fisher) were performed[48]. Maximal reaction velocity (V_{max}) was measured using ethanol and NAD^+ concentrations more than twice those needed to generate maximum rates. The V_{max} of the homogenate of each vial was estimated as the mean of the three velocity measurements divided by the mean of the three protein quantity measurements.

Data from the 15 replicate homogenates of each transformation strain were pooled for analysis. V_{max} did not differ significantly between the transformation strains within any genotype (Supplementary Fig. 4b; Supplementary Table 3), so data from strains of each genotype were pooled for further analysis, yielding 30 replicate V_{max} estimates per genotype. Unpaired *t*-tests were used to test the hypotheses that V_{max} of homogenates from flies carrying the ancestral *D. melanogaster* ADH differed from (i) AncMS, (ii) Alt-AncMS and (iii) the ecologically relevant fast allele. Each *t*-test represented an independent hypothesis, so we did not correct for multiple testing; however, using a Bonferroni correction did not change the significance ($P < 0.05$) of any comparison. Analysis by ANOVA and Dunnett's test for multiple comparisons also found no difference between *D. melanogaster* and AncMS or Alt-AncMS, but a significant difference between fast and *D. melanogaster* (Supplementary Fig. 4c).

Ethanol survival assays. For each ADH genotype, four to five replicate pools of 25 larvae from each of two independent transformation strains were characterized for survival at each of six ethanol concentrations. Beginning at the transition from the 2nd to 3rd larval instar, 150 individuals in each population were divided into equally sized groups and reared on food containing 0, 3, 6, 9, 12, or 15% ethanol. Ethanol supplemented food was prepared by adding ethanol to a standard molasses–cornmeal *Drosophila* food to obtain the appropriate percentage ethanol in the total volume of food. To minimize the loss of ethanol due to evaporation, ethanol was added when the food had cooled as long as possible and to less than 60 °C before pouring the food into vials. Vials were plugged immediately after pouring and stored in an 11 °C refrigerator for no more than three days.

The fraction of individuals surviving to eclosion in each dose group was measured and the relationship between ethanol concentration and proportion not surviving for each genotype was assessed by fitting a Boltzmann sigmoidal model using nonlinear regression in GraphPad Prism 7.0. The LD_{50} was estimated for each genotype from eight to ten replicate pools of larvae at each concentration of ethanol. Significant differences in LD_{50} estimates among genotype pairs were assessed using an extra sum of squares *F*-test to compare the likelihoods of a constrained model with a single LD_{50} parameter fit to the data from both genotypes to that of a free model with independent LD_{50}s for each genotype.

Adult ethanol tolerance was assayed by placing 25, 2–4-day-old adult flies in vials with Whatman paper containing 1 ml of 3% sucrose solution with either 0, 3, 4.5, 6, 9, or 12% ethanol and measuring the fraction surviving after 48 h. Replication and data analyses were as described for larval tolerance assays, except only one transformation strain per genotype was used.

MK tests. The MK test[24] was applied to an alignment of sequences from *D. melanogaster*, *D. simulans* and *D. yakuba*. Analyses were restricted to these three species, as in the original study, because they are the only ones in the *D. melanogaster* group for which recently collected polymorphism data are available. The alignment included all sequences used in the original study, along with recent polymorphism data provided by the DPGP2 consortium[29] and D. Matute. We excluded variants sampled only once because rare segregating variants are known to compromise the efficacy of the MK test[49]. For lineage-specific MK tests, we counted the number of non-synonymous and synonymous divergences that occurred along a branch between reconstructed ancestral alleles, as well as the number of extant polymorphisms of each type within the species descending from that branch.

Data availability. The new sequence data used in this analysis have been deposited in GenBank with the accession codes KX976486 to KX976521. All of the other sequence

data are available as a part of the *Drosophila* population genomics project (DPGP2) from the Drosophila Genome Nexus (www.johnpool.net/genomes.html). Plasmids, primers and cell lines used in this study are available from the authors upon request.

Received 13 September 2016; accepted 1 November 2016; published 13 January 2017

References

1. Sabeti, P. C. *et al.* Genome-wide detection and characterization of positive selection in human populations. *Nature* **449,** 913–918 (2007).
2. Mathieson, I. *et al.* Genome-wide patterns of selection in 230 ancient Eurasians. *Nature* **528,** 499–503 (2015).
3. Liu, S. *et al.* Population genomics reveal recent speciation and rapid evolutionary adaptation in polar bears. *Cell* **157,** 785–794 (2014).
4. Shen, Y.-Y. *et al.* Adaptive evolution of energy metabolism genes and the origin of flight in bats. *Proc. Natl Acad. Sci. USA* **107,** 8666–8671 (2010).
5. Li, J. *et al.* Joint analysis of demography and selection in population genetics: where do we stand and where could we go? *Mol. Ecol.* **21,** 28–44 (2012).
6. Storz, J. F. & Wheat, C. W. Integrating evolutionary and functional approaches to infer adaptation at specific loci. *Evolution* **64,** 2489–2509 (2010).
7. Barrett, R. D. H. & Hoekstra, H. E. Molecular spandrels: tests of adaptation at the genetic level. *Nat. Rev. Genet.* **12,** 767–780 (2011).
8. Lewontin, R. C. Twenty-five years ago in genetics: electrophoresis in the development of evolutionary genetics: milestone or millstone? *Genetics* **128,** 657 (1991).
9. Lewontin, R. C. *The Genetic Basis of Evolutionary Change* (Columbia Univ. Press, 1974).
10. Feder, M. E. & Watt, W. B. in *Genes in Ecology* (eds Berry, R. J. *et al.*) 365–392 (Blackwell Scientific, 1992).
11. Storz, J. F. *et al.* Evolutionary and functional insights into the mechanism underlying high-altitude adaptation of deer mouse hemoglobin. *Proc. Natl Acad. Sci. USA* **106,** 14450–14455 (2009).
12. Prasad, K. V. S. K. *et al.* A gain-of-function polymorphism controlling complex traits and fitness in nature. *Science* **337,** 1081–1084 (2012).
13. Tishkoff, S. A. *et al.* Convergent adaptation of human lactase persistence in Africa and Europe. *Nat. Genet.* **39,** 31–40 (2007).
14. Linnen, C. R. *et al.* Adaptive evolution of multiple traits through multiple mutations at a single gene. *Science* **339,** 1312–1316 (2013).
15. Chan, Y. F. *et al.* Adaptive evolution of pelvic reduction in sticklebacks by recurrent deletion of a *Pitx1* enhancer. *Science* **327,** 302–305 (2010).
16. Cheviron, Z. A. *et al.* Integrating evolutionary and functional tests of adaptive hypotheses: a case study of altitudinal differentiation in hemoglobin function in an Andean sparrow, *Zonotrichia capensis*. *Mol. Biol. Evol.* **31,** 2948–2962 (2014).
17. Harms, M. J. & Thornton, J. W. Analyzing protein structure and function using ancestral gene reconstruction. *Curr. Opin. Struct. Biol.* **20,** 360–366 (2010).
18. Heinstra, P. W., Thörig, G. E., Scharloo, W., Drenth, W. & Nolte, R. J. Kinetics and thermodynamics of ethanol oxidation catalyzed by genetic variants of the alcohol dehydrogenase from *Drosophila melanogaster* and *D. simulans*. *Biochim. Biophys. Acta* **967,** 224–233 (1988).
19. Freeman, S. & Herron, J. C. *Evolutionary Analysis* (Pearson Prentice Hall, 2007).
20. Johnson, N. A. *Darwinian Detectives: Revealing the Natural History of Genes and Genomes* (Oxford Univ. Press, 2007).
21. Kreitman, M., Shorrocks, B. & Dytham, C. in *Genes in Ecology* (eds Berry, R. J. *et al.*) 281–312 (Blackwell Scientific, 1992).
22. Fox, C. W. & Wolf, J. B. *Evolutionary Genetics: Concepts and Case Studies* (Oxford Univ. Press, 2006).
23. Ruse, M. *Darwin and Design: Does Evolution Have a Purpose?* (Harvard Univ. Press, 2003).
24. Heinstra, P. W. H. Evolutionary genetics of the *Drosophila* alcohol dehydrogenase gene-enzyme system. *Genetica* **92,** 1–22 (1993).
25. McKenzie, J. A. & Parsons, P. A. Alcohol tolerance: an ecological parameter in the relative success of *Drosophila melanogaster* and *Drosophila simulans*. *Oecologia* **10,** 373–388 (1972).
26. McDonald, J. F. & Avise, J. C. Evidence for the adaptive significance of enzyme activity levels: interspecific variation in α-GPDH and ADH in *Drosophila*. *Biochem. Genet.* **14,** 347–355 (1976).
27. Heinstra, P. W. H., Scharloo, W. & Thorig, G. E. W. Physiological significance of the alcohol dehydrogenase polymorphism in larvae of *Drosophila*. *Genetics* **117,** 75–84 (1987).
28. McDonald, J. H. & Kreitman, M. Adaptive protein evolution at the *Adh* locus in *Drosophila*. *Nature* **351,** 652–654 (1991).
29. Hartl, D. L. & Clark, A. G. *Principles of Population Genetics* 4th edn (Sinauer Associates, 2007).
30. Charlesworth, B. & Charlesworth, D. Population genetics from 1966 to 2016. *Heredity* (2016).
31. Dickinson, W. J., Rowan, R. G. & Brennan, M. D. Regulatory gene evolution: adaptive differences in expression of alcohol dehydrogenase in *Drosophila melanogaster* and *Drosophila simulans*. *Heredity* **52,** 215–225 (1984).
32. Laurie, C. C., Heath, E. M., Jacobson, J. W. & Thomson, M. S. Genetic basis of the difference in alcohol dehydrogenase expression between *Drosophila melanogaster* and *Drosophila simulans*. *Proc. Natl Acad. Sci. USA* **87,** 9674–9678 (1990).
33. Thomson, M. S., Jacobson, J. W. & Laurie, C. C. Comparison of alcohol dehydrogenase expression in *Drosophila melanogaster* and *Drosophila simulans*. *Mol. Biol. Evol.* **8,** 31–48 (1991).
34. Montooth, K. L., Siebenthall, K. T. & Clark, A. G. Membrane lipid physiology and toxin catabolism underlie ethanol and acetic acid tolerance in *Drosophila melanogaster*. *J. Exp. Biol.* **209,** 3837–3850 (2006).
35. Kaun, K. R., Devineni, A. V. & Heberlein, U. *Drosophila melanogaster* as a model to study drug addiction. *Hum. Genet.* **131,** 959–975 (2012).
36. Fry, J. D. Mechanisms of naturally evolved ethanol resistance in *Drosophila melanogaster*. *J. Exp. Biol.* **217,** 3996–4003 (2014).
37. Sezgin, E. *et al.* Single-locus latitudinal clines and their relationship to temperate adaptation in metabolic genes and derived alleles in *Drosophila melanogaster*. *Genetics* **168,** 923–931 (2004).
38. Umina, P. A., Weeks, A. R., Kearney, M. R., McKechnie, S. W. & Hoffmann, A. A. A rapid shift in a classic clinal pattern in *Drosophila* reflecting climate change. *Science* **308,** 691–693 (2005).
39. Laurie, C. C. & Stam, L. F. The effect of an intronic polymorphism on alcohol dehydrogenase expression in *Drosophila melanogaster*. *Genetics* **138,** 379–385 (1994).
40. Begun, D. J., Betancourt, A. J., Langley, C. H. & Stephan, W. Is the fast/slow allozyme variation at the *Adh* locus of *Drosophila melanogaster* an ancient balanced polymorphism? *Mol. Biol. Evol.* **16,** 1816–1819 (1999).
41. Berry, A. & Kreitman, M. Molecular analysis of an allozyme cline: alcohol dehydrogenase in *Drosophila melanogaster* on the east coast of North America. *Genetics* **134,** 869–893 (1993).
42. Pool, J. E. *et al.* Population genomics of sub-Saharan *Drosophila melanogaster*: African diversity and non-African admixture. *PLoS Genet.* **8,** e1003080 (2012).
43. Parsons, P. A. Ethanol utilization: threshold differences among six closely related species of *Drosophila*. *Aust. J. Zool.* **28,** 535–541 (1980).
44. Stam, L. F. & Laurie, C. C. Molecular dissection of a major gene effect on a quantitative trait: the level of alcohol dehydrogenase expression in *Drosophila melanogaster*. *Genetics* **144,** 1559–1564 (1996).
45. Hudson, R. R., Kreitman, M. & Aguadé, M. A test of neutral molecular evolution based on nucleotide data. *Genetics* **116,** 153–159 (1987).
46. Oakeshott, J. G. *et al.* Alcohol dehydrogenase and glycerol-3-phosphate dehydrogenase clines in *Drosophila melanogaster* on different continents. *Evolution* **36,** 86–96 (1982).
47. Yang, Z., Kumar, S. & Nei, M. A new method of inference of ancestral nucleotide and amino acid sequences. *Genetics* **141,** 1641–1650 (1995).
48. Loehlin, D. W. & Carroll, S. B. Expression of tandem gene duplicates is often greater than twofold. *Proc. Natl Acad. Sci. USA* **113,** 5988–5992 (2016).
49. Eyre-Walker, A. Changing effective population size and the McDonald–Kreitman test. *Genetics* **162,** 2017–2024 (2002).

Acknowledgements

We thank L. Picton, K. O'Brien, K. Gordon and members of the C. Meiklejohn and K. Montooth laboratories for technical assistance. We thank D. Matute for providing polymorphism data for *D. yakuba*. We thank M. Kreitman, members of the J. Thornton laboratory and D. Anderson for comments and suggestions that enriched the project. The project was supported by a National Science Foundation (NSF) grant (DEB-1501877; J.W.T./M.A.S.), an NSF graduate research fellowship (M.A.S.), National Institutes of Health (NIH) grant (R01-GM104397; J.W.T.), NSF CAREER Award (1505247; K.L.M.) and an NIH training grant (T32-GM007197; M.A.S.). D.W.L. was supported by a Howard Hughes Medical Institute postdoctoral fellowship from the Life Sciences Research Foundation and an investigatorship to S. B. Carroll from the Howard Hughes Medical Institute.

Author contributions

M.A.S. and J.W.T. conceived the project. All authors participated in the experimental design. M.A.S. performed the phylogenetic and population genetic analyses. D.W.L. constructed the transgenic animals. M.A.S., D.W.L. and K.L.M. performed the functional experiments. All authors participated in data analysis and interpretation. M.A.S. and J.W.T. wrote the paper with contributions from D.W.L. and K.L.M.

Additional information

Supplementary information is available for this paper.

Reprints and permissions information is available at www.nature.com/reprints.

Correspondence and requests for materials should be addressed to J.W.T.

How to cite this article: Siddiq, M. A., Loehlin, D. W., Montooth, K. L. & Thornton, J. W. Experimental test and refutation of a classic case of molecular adaptation in *Drosophila melanogaster*. *Nat. Ecol. Evol.* **1,** 0025 (2017).

Competing interests

The authors declare no competing financial interests.