PNAS RESEARCH ARTICLE | GENETICS OPEN ACCESS

Longitudinal sequencing reveals polygenic and epistatic nature of genomic response to selection

Simon K. G. Forsberg[a,b,1], Diogo Melo[a,b,1], Scott W. Wolf[a,b], Jennifer K. Grenier[c], Minjia Tang[a,b], Lucas P. Henry[a,b], Luisa F. Pallares[a,b,2], Andrew G. Clark[c,3], and Julien F. Ayroles[a,b,3]

Affiliations are included on p. 9.

Contributed by Andrew G. Clark; received June 15, 2024; accepted March 25, 2025; reviewed by Corbin D. Jones and Jason B. Wolf

Evolutionary adaptation to new environments likely results from a combination of selective sweeps and polygenic shifts, depending on the genetic architecture of traits under selection. While selective sweeps have been widely studied, polygenic responses are thought to be more prevalent but remain challenging to quantify. The infinitesimal model makes explicit the hypothesis about the dynamics of changes in allele frequencies under selection, where only allelic effect sizes, frequencies, linkage, and gametic disequilibrium matter. Departures from this, like long-range correlations of allele frequency changes, could be a signal of epistasis in polygenic response. We performed an Evolve & Resequence experiment in *Drosophila melanogaster* exposing flies to a high-sugar diet for over 100 generations. We tracked allele frequency changes in >3000 individually sequenced flies and population pools and searched for loci under selection by identifying sites with allele frequency trajectories that differentiated selection regimes consistently across replicates. We estimate that at least 4% of the genome was under positive selection, indicating a highly polygenic response. The response was dominated by small, consistent allele frequency changes, with few loci exhibiting large shifts. We then searched for signatures of selection on pairwise combinations of alleles in the new environment and found several strong signals of putative epistatic interactions across unlinked loci that were consistent across selected populations. Finally, we measured differentially expressed genes (DEGs) across treatments and show that DEGs are enriched for selected SNPs. Our results suggest that epistatic contributions to polygenic selective response are common and lead to detectable signatures.

polygenic selection | epistasis | *Drosophila* | evolve and resequence | artificial selection

Significance

How adaptations are built genetically depends on if and how genes interact with one another. To investigate this question, we tracked genetic changes over 100 generations in fruit flies adapting to high-sugar diets, showing that adaptation is not only highly polygenic but also shaped by interactions between genes (epistasis). Our longitudinal approach reveals that natural selection favors specific gene combinations, as seen in coordinated genetic shifts and unexpected patterns of genetic variation. These findings provide a framework for identifying the genomic footprint of epistasis in evolving populations and offer insights into the complex genetic architecture that underlies adaptation to environmental challenges.

The genetic basis of evolutionary adaptation depends fundamentally on the architecture of traits under selection. While early research focused heavily on identifying individual genes with large effects that undergo selective sweeps (1, 2), we now understand that most adaptive responses involve more subtle changes across many loci (3, 4). This emphasis on single-gene effects has created a biased view of adaptation in the literature (4), even though cataloged examples exist in both humans and other species (5, 6).

Beyond identifying and just counting the number of genes involved, understanding adaptation requires considering how genes interact with each other (epistasis) and with the environment (7). The extent to which epistasis is important for polygenic response is still an open question (7–11), but theory suggests it should leave two detectable signatures: correlated changes in allele frequencies between interacting loci, even when unlinked (12), and gametic disequilibrium in adapted populations as allelic combinations are selected for or against—resulting in deviations from two-locus Hardy–Weinberg proportions between pairs of unlinked loci (13, 14).

Evolve and Resequence (E&R) experiments provide an ideal framework for studying these complex adaptive processes (15). By tracking allele frequencies over time in both selected and control populations, we can distinguish between adaptation to laboratory conditions and responses to specific selective pressures. Environmental stress can also reveal previously hidden genetic interactions (16, 17), allowing us to identify epistatic combinations that become advantageous in new conditions.

Dietary change represents a particularly informative source of environmental stress. In *Drosophila melanogaster*, exposure to high dietary sugar triggers complex metabolic and behavioral responses (18) involving coordinated changes across multiple systems: sugar absorption and metabolism (19), feeding behaviors (20), and gene expression patterns, particularly in digestive and metabolic pathways (21, 22). Chronic high-sugar exposure can lead to obesity, diabetes-like symptoms, cardiovascular problems, and reduced lifespan (23, 24). This broad physiological impact suggests that adaptation to

Author contributions: S.K.G.F., J.K.G., A.G.C., and J.F.A. designed research; S.K.G.F., D.M., J.K.G., M.T., L.P.H., and J.F.A. performed research; S.K.G.F., J.K.G., A.G.C., and J.F.A. contributed new reagents/analytic tools; S.K.G.F., D.M., S.W.W., L.P.H., L.F.P., and J.A. analyzed data; and S.K.G.F., D.M., A.G.C., and J.F.A. wrote the paper.

Reviewers: C.D.J., The University of North Carolina at Chapel Hill; and J.B.W., University of Bath.

The authors declare no competing interest.

[1]S.K.G.F. and D.M. contributed equally to this work.

[2]Present address: Friedrich Miescher Laboratory of the Max Planck Society, Tübingen 72076, Germany.

[3]To whom correspondence may be addressed. Email: ac347@cornell.edu or jayroles@princeton.edu

This article contains supporting information online at https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2410452122/-/DCSupplemental.

Published June 18, 2025.

high-sugar should involve multiple interacting gene networks under strong selection.

Here, we conducted an E&R experiment exposing replicate *D. melanogaster* populations to either high-sugar or control diets for 100 generations, starting from the same base population. We performed whole-genome sequencing on flies from six populations at generations 1, 11, 25, and 100, giving a total of almost 3,000 sequenced individuals (Fig. 1*A*). We identified two major directions of genetic change: a primary response shared across all populations, reflecting laboratory adaptation, and a secondary response specific to high-sugar selection. While at least 4% of the genome showed signatures of positive selection under high-sugar conditions, the selected loci do not show archetypal signals of selective sweeps. This highly polygenic response appears to operate largely through regulatory changes, as selected regions were enriched among differentially expressed genes. Additionally, we found evidence for epistatic selection through correlated allele frequency changes and gametic disequilibrium between unlinked loci, patterns that Wright–Fisher simulations suggest are unlikely without epistatic interactions. Though we cannot directly measure epistatic effects on phenotypes, our results indicate that epistatic interactions commonly contribute to polygenic adaptation and leave detectable genomic signatures.

Results

Polygenic Selection Response. To study the effect of long-term selection in a stressful environment, we maintained three replicate populations of flies under high-sugar stress and three under control conditions for 100 generations. To assess adaptation to the high-sugar environment, we performed a factorial egg-lay experiment, measuring the fecundity of both control and high-sugar (HS) populations on both diets. Both populations showed higher fecundity on their corresponding diet (Fig. 1*D*), indicating successful adaptation. We collected data from 100 individuals at four time points, obtaining both allele frequencies from Pool-seq and genotypes from individual sequencing (Fig. 1 *A* and *C*). This time series genomic data allowed us to analyze allele frequency changes and identify the primary drivers of genetic change in response to the stressful environment. After quality control (*Methods*), we obtained allele frequency estimates for ~1.76 M SNPs, representing 4 time points × 3 replicate populations × 2 treatments × 1.76 M SNP genotype calls. To identify the main drivers of genetic change without prior assumptions, we performed Principal Component Analysis (PCA) of the allele frequencies across the entire selection experiment. The first two principal components (PCs), explaining 17% and 13% of the variance, corresponded to time and selection regime, respectively (Fig. 2, *SI Appendix*, Fig. S3). This unsupervised

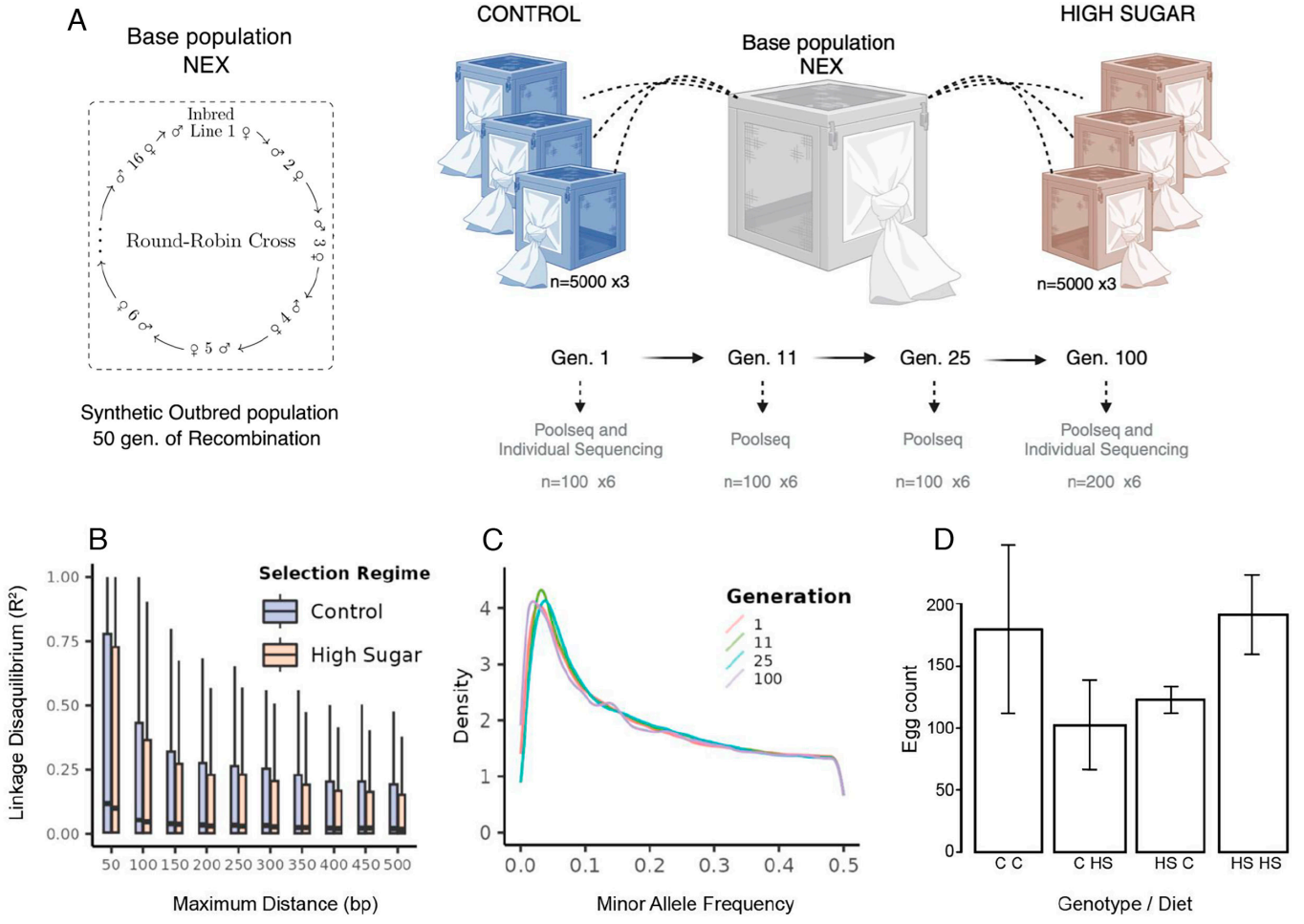

Fig. 1. Selection experimental design. (*A*) Scheme of the experimental design. A synthetic outbred population was created by a round-robin cross of 16 inbred lines from the Netherlands. This population (NEX) was kept as an outbred population for over 50 generations before the start of the selection experiment. Starting from NEX, 3 control (CONTROL) and 3 treatment (HS) populations of around 5,000 individuals were kept for 100 generations. Samples of one hundred individuals were taken at generations 1, 11, 25, and 100 for allele frequency tracking. (*B*) Linkage disequilibrium (LD) decay across HS and CONTROL populations at generation 100. NEX derived populations have very low levels of LD. (*C*) Minor allele frequency across generations. (*D*) Egg-lay measurements after selection showing the adaptive response to the high-sugar environment in HS populations.

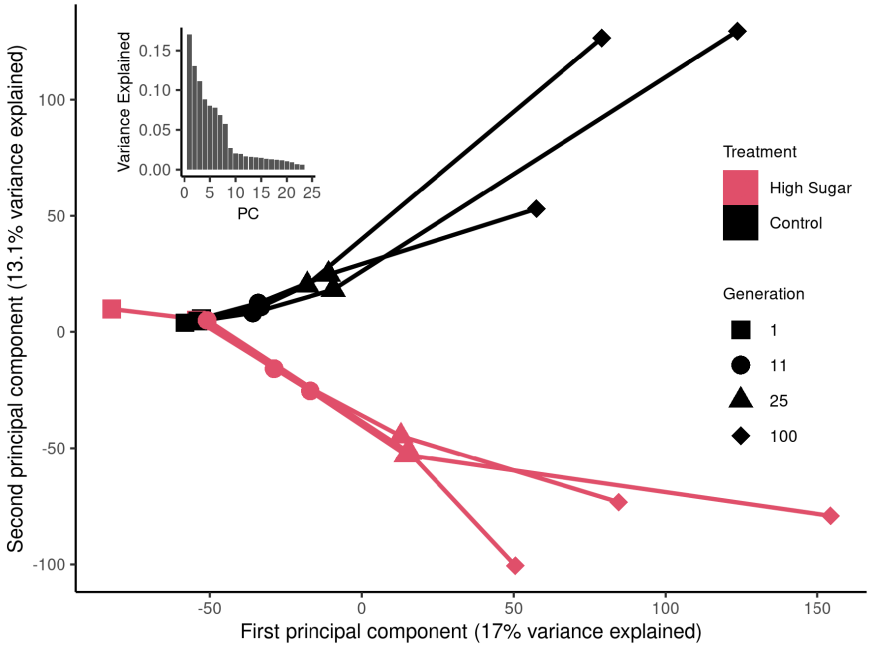

Fig. 2. PCs one (*x*-axis) and two (*y*-axis) from the PCA on the genome wide allele frequencies across the entire selection experiment. Each line corresponds to one of the six experimental populations, red indicating high-sugar treatment and black control, with symbols marking the mean scores for each population and time point. The variance explained by each of the first 23 PCs is shown in the *Inset*.

approach identified time and high-sugar selection as the two main drivers of genome-wide genetic change. Notably, the time dimension, captured by the first PC, explained more variance than the selection regime, captured by the second PC. This indicates that all six populations experienced common selection pressures, presumably related to a shift to a laboratory cage environment and appearance of many other competing genotypes with high-sugar exposure representing the second largest driver of genetic change (Fig. 2, *SI Appendix*, Fig. S3).

High-Sugar Selection on Individual Loci. To identify individual loci under selection, we fitted a univariate regression model for each SNP, incorporating allele frequencies across all time points, replicate populations, and selection regimes (*Methods*). This model identifies SNPs whose allele frequency changes in the same direction over time in all replicate populations. The time coefficient in the model captures changes that are similar across all six populations (Fig. 3*A*), while the time-by-selection-regime coefficient captures changes that are unique to one selection treatment (Fig. 3 *B* and *C*). The *P*-values of the time coefficient were highly correlated with SNP loadings onto PC1 (cor = 0.59, $P < 10^{-16}$, *SI Appendix*, Fig. S2), while those of the time-by-regime coefficient were highly correlated with SNP loadings onto PC2 (cor = 0.68, $P < 10^{-16}$, *SI Appendix*, Fig. S2), consistent with the first two PCs capturing time and selection regime effects.

Different SNPs exhibited distinct allele frequency trajectories over time. Some responded similarly to selection in all replicate populations regardless of selection regime (Fig. 3*A*), while others responded in opposite directions (Fig. 3*C*) or in only one regime (Fig. 3*B*). Our regression model allowed us to distinguish among these scenarios, and in subsequent analyses, we focus on selection signatures unique to the high-sugar selection regime. The Manhattan profile in Fig. 2*D*, showing the time-by-regime *P*-values, suggests a polygenic selection response, consistent with our observation that time and selection regime are the two main drivers of genetic change genome-wide (Fig. 2).

To relate the locus-specific results (Fig. 3) to the genome-wide signal quantified by the PCA (Fig. 2, *SI Appendix*, Fig. S1), we repeated the PCA after excluding SNPs with regression *P*-values below various thresholds, effectively removing SNPs associated with the selection regime. Varying the significance threshold allowed us to evaluate the effect of the filtered SNPs on the PCA. Using a very conservative threshold that excluded only the most strongly selected SNPs, the PCA results remained largely unchanged, indicating that the PCA signal was not driven by a few loci under strong selection (*SI Appendix*, Fig. S2). We used these changes in the PCA as a heuristic to pick a *P*-value threshold of 8×10^{-12}, since PC2 no longer distinguish the different selection regimes when excluding SNPs with a *P*-value below this threshold (Fig. 8*C*). SNPs passing this significance threshold are thus driving the majority of the selection response to high-sugar stress that we observe in the PCA.

What Proportion of the Genome Is Responding to Selection? Using this conservative threshold, ~45 k SNPs showed a signature of positive selection unique to the high-sugar selection regime. Considering 200 bp around every selected SNP, corresponding to an average r^2 of 0.2 (Fig. 1*B*), these SNPs span ~5.6 Mb, or ~4% of the mappable genome of *D. melanogaster*. Since the linkage disequilibrium (LD) around the selected loci is expected to be larger than the genome-wide average, this represents a conservative estimate. The magnitudes of allele frequency changes were relatively small. Between generations 1 and 100, the mean change across all SNPs on the high-sugar selection regime was 0.11, while the mean change among selected SNPs was 0.25 (Fig. 4). Among all 1.76 M SNPs, only 4,753 showed a pattern where the minor allele at generation 1 reached fixation at generation 100 in at least one high-sugar population. Additionally, many SNPs displayed a delayed selection response, with the largest allele frequency changes occurring after generation 25 (Fig. 4). This pattern is consistent with theoretical predictions for polygenic adaptation involving independent loci (25, 26), but could also result from epistatic effects (27).

Do the Selected Alleles Show a Detectable Sweep Signature? Next, we asked whether the identified selection signatures coincided with the genomic footprint of selective sweeps.

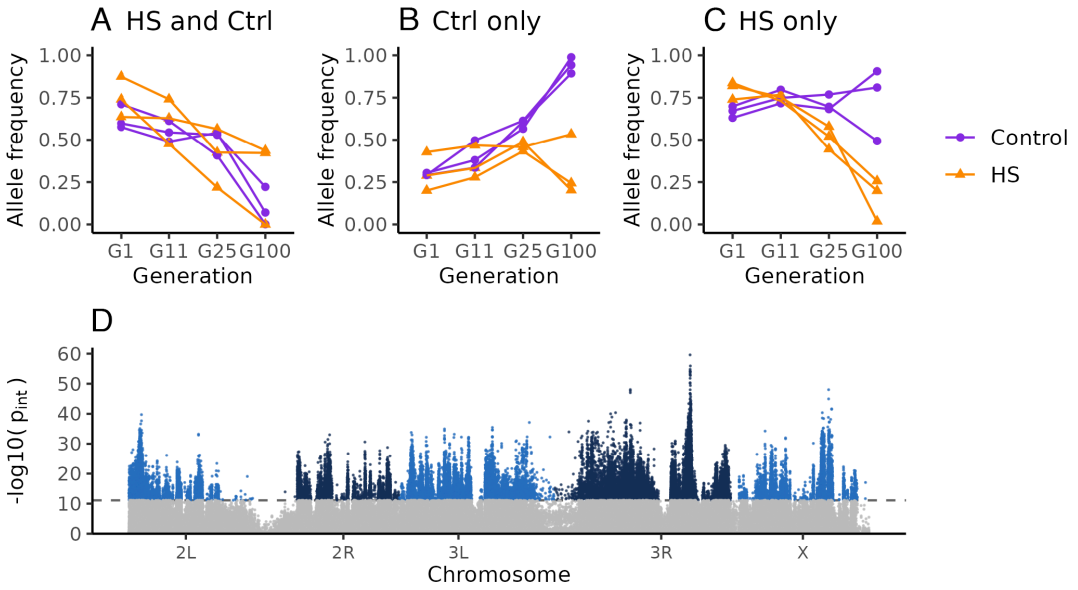

Fig. 3. Results from the per SNP regression model. Panels *A–C* show possible patterns of relevant and consistent allele frequency change across the six populations. We chose SNPs with strong signals for illustration, but many significant SNPs show more subtle allele frequency changes. Plotted are allele frequency trajectories of SNPs with significant linear trends under the specified model. (*A*) Consistent change in CONTROL and HS (*B*) consistent change in CONTROL only (*C*) HS and CONTROL differ. Both the examples shown in (*B*) and (*C*) would lead to a significant interaction term between time and treatment, but we filter SNPs that change only in Control (like in panel *B*); (*D*) Manhattan plot showing negative log10 transformed *P*-values from the regression analysis of allele frequency over time. The *P*-values correspond to the time-by-selection regime interaction coefficient in the model. A significant *P*-value indicates different trajectories in the two treatments. SNPs showing a selection response primarily in the control regime were excluded and are not shown.

Using a core set of 20 k high-confidence SNPs, we estimated individual haplotypes at generation 100. These haplotypes were used to calculate the integrated Haplotype Score (iHS) (28) in the HS populations. A large iHS indicates an extended haplotype associated with one allele at a given SNP, a pattern characteristic of a selective sweep. The estimated iHS and the *P*-values from our regression model showed a small but significant correlation (cor = 0.07, $P = 2.1 \times 10^{-17}$, *SI Appendix,* Fig. S4), indicating a tendency for longer haplotypes at selected loci. The correlation was, however, very modest, suggesting that loci under selection based on our time series data did not display strong sweep-like patterns after 100 generations of selection. At a nominal significance threshold of $P < 0.05$, only 4.7% of the selected loci, as inferred from the regression analysis, displayed a significant iHS. These observations support a polygenic model of adaptation through subtle shifts in allele frequency at many loci, with selection acting primarily on standing genetic variation rather than novel mutations.

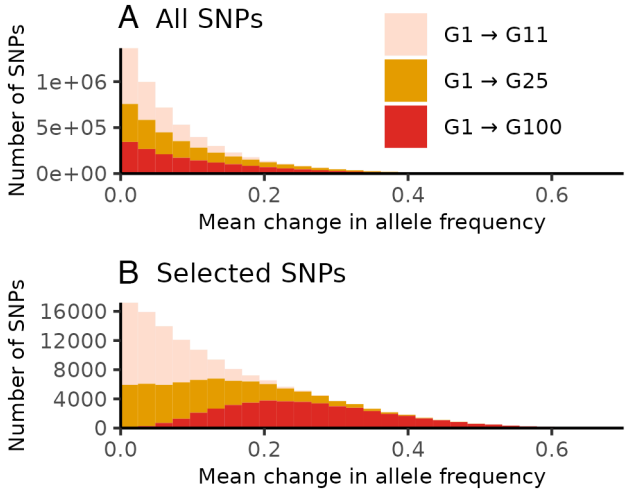

Fig. 4. Histograms showing mean changes in allele frequency in the populations exposed to the high-sugar selection regime, between generation 1 and 11, 1 and 25, and 1 and 100. Panel (*A*) includes all 1.76 M SNPs and panel (*B*) includes the 45 k SNPs that show a signature of positive selection unique to the high-sugar selection regime.

Effects of Polygenic Adaptation on Gene Expression. Much of the genetic variation for complex traits resides in gene-regulatory regions (29). Selection on complex traits would therefore likely act on this regulatory variation, resulting in gene expression changes. To characterize the effect of selection on gene expression, we performed a full reciprocal experiment where flies adapted to either high-sugar or control selection regimes were reared on either high-sugar or control conditions (Fig. 5). This design allowed us to account for both short-term plastic changes due to different diets and long-term selection effects. For each of the four experimental groups, we performed RNA-seq separately on bodies and heads (n ~ 40 per group; *Methods*). After quality control, we obtained expression data for 8,397 genes from body samples and 8,298 genes from head samples.

Differentially expressed (DE) genes between flies adapted to the respective selection regimes were measured separately for head and body samples. At an FDR < 0.01, 1,155 and 578 genes showed differential expression in body and head samples, respectively. We then examined how many of these DE genes occurred in regions with selection signatures. In both body and head samples, we found an enrichment of selected SNPs among DE genes (Fig. 5*B*). Starting at a *P*-value of 10^{-5} for the selection term, this enrichment became more pronounced with increasingly stringent selection *P*-values. The enrichment substantially exceeded the expected random overlap between selection signals and DE genes, as estimated from permutation tests, indicating that adaptation acted on regulatory genetic variants. While clearly regulatory variation does not account for all the variation, this proportional increase in enrichment with increasing signal strength (as measured by *P*-values) further supports our argument that regulatory variation played a key role in adaptation to high-sugar.

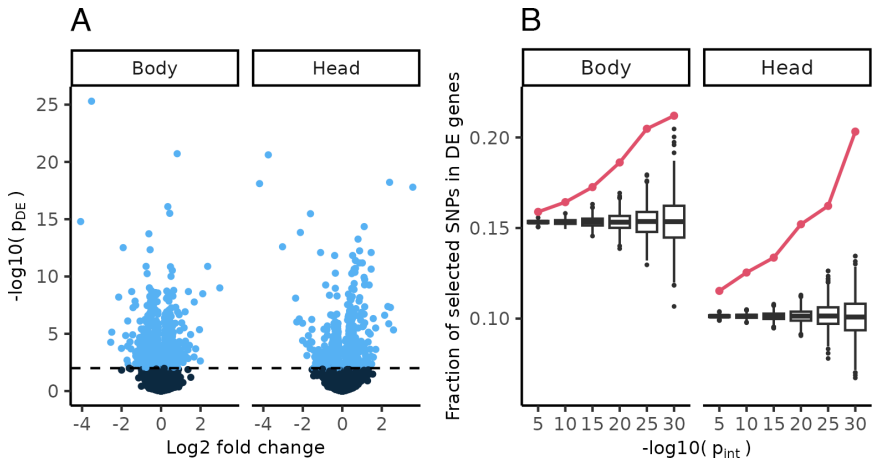

Fig. 5. Differential expression after selection. (*A*) Volcano plots showing differential gene expression between flies adapted to the high-sugar versus control selection regime, after controlling for the plastic effects related to each diet. Each point corresponds to one gene. The Y-axis shows the negative log10 transformed *P*-value of the differential expression, and *x*-axis shows the log2 transformed fold change. The two panels correspond to expression in body and head tissue. (*B*) Fraction of SNPs under positive selection on the high-sugar selection regime that coincide with a differentially expressed gene (*y*-axis), at different *P*-values for the selection signature (*x*-axis). Solid red lines show the observed fraction of SNPs overlapping DE genes and box-plots show the empirical null distributions obtained by measuring the overlap of the selected genes with random genomic regions of the same size as the DE genes. The two panels correspond to expression in body and head tissue.

Epistasis Across Selected Loci. The contribution of epistatic interactions to long-term selection response remains contentious, but several mechanisms could enable epistasis to influence polygenic response. For example, diminishing returns epistasis—where increasing frequency of one allele decreases the fitness effects of alleles at other loci—has been empirically observed (30) and is implicit in stabilizing selection models describing population movement toward an optimum (25, 31, 32). Our access to allele frequency time-series and haplotype information allowed us to examine two distinct signatures of epistatic contributions to the observed polygenic selection response. In the presence of fitness epistasis, the trajectory of a given allele during selection depends on the trajectories of alleles at interacting loci (27). This dependence should result in two observable patterns: 1) correlations between allele frequencies at interacting loci, as changes at one locus are accompanied by corresponding changes at interacting loci; and 2) gametic disequilibrium in adapted populations, where selective removal of unfavorable allelic combinations results in deviations from the two-locus Hardy–Weinberg proportions expected for unlinked loci (13).

To explore the expected genomic footprint of selection under fitness epistasis versus strict additivity, we performed Wright–Fisher simulations using the SLiM modeling framework (33). The simulations mimicked key aspects of our experiment, starting with neutral populations containing ~3 k segregating SNPs in mutation-drift equilibrium distributed along two unlinked chromosomes. From these starting populations, we sampled 1,000 segregating mutations to serve as quantitative trait loci (QTLs) contributing to the trait under truncation selection. We then simulated two scenarios: 1) an additive scenario, where the trait value depended only on additive QTL effects, and 2) an epistatic scenario, where we included both additive effects and additive-by-additive epistatic effects between 200 pairs of QTLs (one from each chromosome). After 100 generations of selection, we quantified gametic disequilibrium and correlated changes in allele frequency. We analyzed the same QTL pairs in both scenarios, with epistatic interactions present only in the second scenario. Gametic disequilibrium between unlinked loci was substantially higher under epistasis, 3 SD higher than random pairs in the epistatic scenario, equal in the additive scenario. Similarly, correlations between allele frequency trajectories of QTL pairs on different chromosomes differed from zero only in the epistatic scenario (Fig. 6*C*).

Next, we searched for similar genomic footprints in our empirical data. We focused on SNPs under strong selection unique to the high-sugar regime (time-by-regime $P < 8 \times 10^{-12}$) with sufficient coverage to confidently genotype many individual flies, yielding a set of 1.3 k SNPs. We estimated allele-frequency correlations between these SNPs on the high-sugar regime, producing a SNP×SNP correlation matrix that revealed which SNPs moved in unison through time across all replicate populations (Fig. 6 *A*, *Lower* triangle). The corresponding SNP×SNP gametic disequilibrium matrix showed which SNP pairs deviated from the expected proportions for unlinked loci in high-sugar selected populations (Fig. 6 *A*, *Upper* triangle)

Both allele frequency correlations and gametic disequilibrium were substantial between physically linked SNPs, as expected (Fig. 6*A*, elements near the diagonal). We also observed numerous correlations between physically distant SNPs, indicating their allele frequencies changed similarly during selection. While this pattern could result from either epistasis or independent selection pressures acting on SNP pairs, we identified multiple cases of gametic disequilibrium between physically distant SNPs, specifically indicating epistatic selection. Comparing gametic disequilibrium between high-sugar and control populations for SNPs showing positive selection unique to the high-sugar regime revealed a small but highly significant negative correlation ($r = -0.06$, $P < 10^{-16}$). In contrast, SNPs showing similar selection signatures in both treatments, likely due to lab adaptation, showed a positive correlation ($r = 0.14$, $P < 10^{-16}$). These patterns suggest two distinct mechanisms: First, epistatic selection causes gametic disequilibrium between physically distant SNPs when allelic combinations are selected on the high-sugar regime but neutral in controls. Second, allelic combinations beneficial in both regimes are similarly selected across all populations, producing comparable gametic disequilibrium patterns in both high-sugar and control populations.

We identified 1,413 SNP pairs in the HS populations located on different chromosomes that displayed both gametic disequilibrium (chi-square test, $P < 5.7 \times 10^{-8}$) and correlated allele frequencies (correlation test, $P < 0.001$). To rule out population structure as a cause of gametic disequilibrium, we examined each

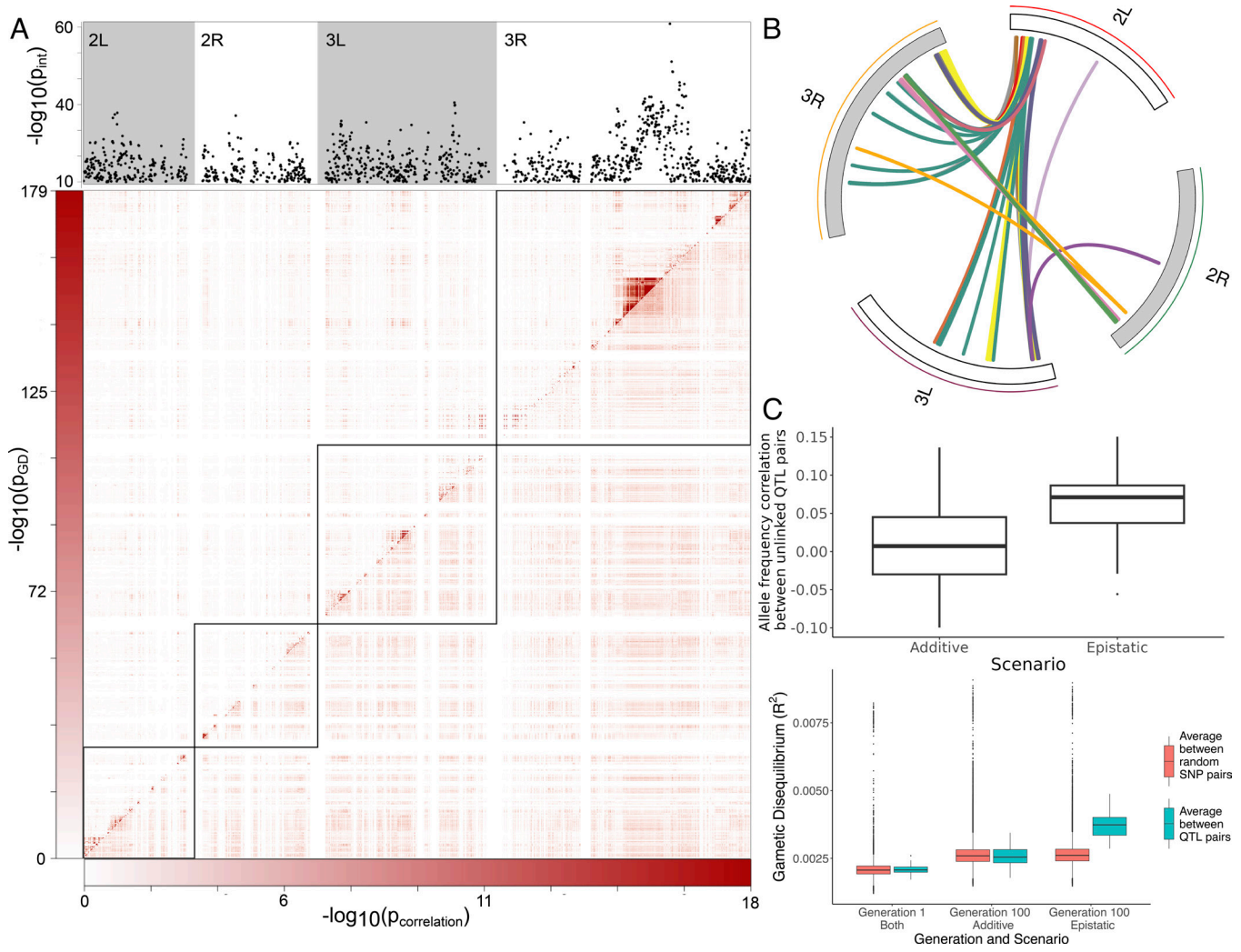

Fig. 6. Signatures of epistasis in experimental data and simulations. (*A*) *Top* panel shows negative log10 *P*-values (*y*-axis) from the regression analysis of allele frequency over time. The *P*-values correspond to the time-by-selection regime interaction coefficient in the model. SNP positions are scaled for visualization. Bottom panel shows a heatmap of pairwise SNP analyses performed in the populations exposed to the high-sugar selection regime. Negative log10 *P*-values of the gametic disequilibrium, given by a chi-square test, are shown above the diagonal. Negative log10 transformed *P*-values of the correlation in allele frequencies over time are shown below the diagonal. (*B*) Locus pairs showing both genotype ratio distortions (chi-square test, $P < 5.7 \times 10^{-8}$) and correlated allele frequencies (r, $P < 0.001$). The outer circle represents the chromosome arms, and each link represents a locus pair. Colors correspond to our LD clumping procedure, where links with the same color involve the same locus at one end, supported by multiple locally linked SNPs. (*C*) Comparison of allele frequency correlations and gametic disequilibrium in the simulations, across additive, and epistatic scenarios.

replicate population individually. While this approach had lower statistical power due to smaller sample sizes compared to analyzing all replicates jointly, it allowed us to control for population structure, which by design cannot exist within individual replicate populations. By retaining only signals supported by multiple linked SNPs at both chromosomal locations, and where gametic disequilibrium replicated in at least two replicate populations, we identified 11 pairwise epistatic selection signatures, each supported by between 5 and 63 SNP pairs (Fig. 6*B*).

Discussion

Here, we tracked allele frequency trajectories of 1.7 million SNPs in replicate *D. melanogaster* populations exposed to high-sugar stress, uncovering evidence for epistatic contributions to polygenic adaptation. While selection drove changes in allele frequencies across thousands of loci, the patterns of these changes revealed complex interactions among selected variants. Our key finding is that adaptation involved coordinated changes between physically unlinked loci, manifesting as both correlated allele frequency trajectories and gametic disequilibrium. Notably, these signatures were specific to high-sugar populations, consistently observed across replicates, and absent in controls, strongly suggesting they result from epistatic selection rather than demographic effects.

PCA of the SNP allele frequencies in these population samples revealed a striking pattern: PC1 corresponded with generation time for both treatments, whereas PC2 cleanly separated the control from the high-sugar diet populations. First, we found that adaptation to laboratory conditions drove stronger genomic changes than the high-sugar treatment itself, highlighting the importance of controlling for laboratory adaptation in experimental evolution studies. Second, our results demonstrate that adaptation to high-sugar environments occurs through subtle allele frequency changes at thousands of loci rather than strong selective sweeps, consistent with a highly polygenic response. Third, we show this polygenic response operates in many cases through regulatory changes, as evidenced by the enrichment of selected SNPs near differentially expressed genes.

Recent advances in genomic technologies have revealed the extensive genetic complexity underlying quantitative traits. Large-scale genome-wide association studies (34, 35), experimental evolution

studies (3), and gene-regulatory network analyses (36) consistently show that adaptation involves changes at many loci. This empirical evidence aligns with theoretical predictions about polygenic adaptation—particularly that mutational target size determines the number of loci involved in adaptive response (37, 31). Our results support Höllinger's intermediate scenario, where adaptation proceeds through partial sweeps at many loci rather than classic selective sweeps. The broad genomic response mirrors the known physiological complexity of dietary adaptation in *Drosophila*, where high-sugar environments affect multiple traits including metabolism, feeding behavior, and lifespan. While we do not have access to the phenotypes under selection, we are able to probe these putative differences by analyzing differential gene expression. Selected lines show hundreds of DE genes, and these are indeed enriched for selected SNPs, suggestive of a link between regulatory *cis*-eQTL variation and the response to selection. These widespread regulatory changes are expected to percolate through metabolic networks (38), leading to interactions between connected genes, which potentially complicates the adaptive architecture (38).

We attempt to further characterize the adaptive architecture by searching for possible signatures of epistatic interactions between the SNPs under selection in the HS populations. While epistasis can store additive genetic variation (39, 40), its effect on long-term selection response is generally considered modest (27). However, epistasis becomes important when it is directional—where epistatic effects consistently enhance or buffer additive variation (41–43). Although measuring epistasis is challenging, making general patterns of directional epistasis difficult to document (9), it has been observed in several model organisms (44, 45). When multiple traits are selected simultaneously, epistatic effects on trait associations can significantly influence selection response (46–48).

Recent E&R experiments highlight two ways epistasis shapes adaptation. First, through allelic redundancy, where different combinations of alleles at multiple loci can produce similar fitness effects, leading to variable responses across replicates (3). Under stabilizing selection, such interactions necessarily create fitness epistasis (31, 49). Second, epistatic interactions can increase parallel evolution by constraining adaptive paths through genotype space (16, 30, 49). Given these mechanisms, we examined our high-sugar populations for epistatic signatures, specifically looking for correlated allele frequencies and gametic disequilibrium between unlinked loci. While we cannot determine the precise mechanisms underlying these epistatic interactions, their consistent signature across replicates provides strong evidence for epistasis contributing to adaptation to high-sugar stress.

Our findings highlight the power of E&R experiments to dissect complex adaptive responses that would be difficult to detect through other approaches. The combination of replicated evolution, temporal sampling, and multiple genomic analyses revealed not only the extensive nature of polygenic adaptation but also underscored the significance of epistasis in shaping evolutionary trajectories. This experimental framework enables fine-scale resolution into how organisms adapt to environmental challenges by allowing us to track the concurrent evolution of thousands of loci and their interactions.

Methods

Mapping Population. To allow the detection of allelic effects that would be hidden in natural populations due to low frequency, we created a synthetic outbred mapping population. To create this population, we selected 16 inbred lines from the Netherlands population (NEX) from the GLOBAL DIVERSITY Lines (50). The lines were selected based upon their low frequency of inversions to reduce the suppression of recombination associated with inversions (3). To establish the population from these lines, we performed a round-robin cross on the initial lines (1 × 2, 2 × 3, …, 16 × 1) and subsequently performed a round-robin cross on the F1s to ensure parental representation and that no chromosome was lost. The resulting F2 individuals were placed in 42 h × 42w × 42d cm cages and allowed to recombine freely for more than 50 generations. This design increases the allele frequency of rare variants by replicating and randomizing throughout the population (Fig. 1*C*).

Selection Regime. We performed a laboratory natural selection experiment (51) on high-sugar diets without selecting for any phenotype. High-sugar diets are known to have high fitness costs (23)and by allowing our populations to directly evolve under this physiological stress, we explored the adaptation to this deleterious effect. To do this, we subdivided our mapping population into 6 replicate populations, 3 of which were placed on a standard medium and 3 of which were placed on high-sugar medium. The standard medium consists of 8% glucose, 8% yeast, 1.2% agar, 0.04% phosphoric acid, and 0.4% propionic acid. High-sugar medium follows the same recipe as standard medium with the addition of 12% glucose resulting in a total of 20% glucose. Each population was placed in a population cage (BugDorm #4F3030) and maintained at ~5,000 individuals for ~120 generations. Each generation was seeded from an egg lay, on fresh bottles of the respective diet, at 5 to 6 d posteclosion. After pupation but before eclosion, bottles were cleared of adults, moved to new cages and opened. Following each egg lay, individuals were collected and stored at −80 C for subsequent sequencing.

Factorial Egg Lay After Selection. To assess whether selected populations had adapted to the stressful high-sugar environment, we performed a factorial egg lay experiment, measuring the fecundity of both CONTROL and HS populations in the control and high-sugar diet. Both populations show higher fitness in the diet to which they had adapted (Fig. 1*D*).

Library Preparation and Sequencing. Flies from generation 1, 11, 25, and 100 were selected from each population for sequencing and distributed into 96-well plates. One 2.8 mm stainless steel grinding bead (OPS diagnostics, #089-5000-11) and 100 μL of lysis buffer were added to each well. Flies were homogenized for 10 min at maximum speed in a Talboys High Throughput Homogenizer (#930145). The resulting lysate was moved to a new 96-well plate for DNA extraction, using a Multi-Well Plate Vacuum Manifold (Pall Life Sciences #5017) and Acroprep advance 1 mL DNA binding plates (Pall Life Sciences #8132).

Library prep was performed using a liquid handling robot (CyBio® FeliX, Analitik Jena) to ease the processing of many samples and reduce variability from manual handling of samples. The protocol broadly followed the strategy described by Picelli et al (52). Specifically, we added 10 μL (100 μM) of forward oligo adapter A and 10 μl (100 μM) of reverse oligo adapter (Tn5MERev) to 80 μL of reassociation buffer (10 mM Tris pH 8.0, 50 mM NaCl, 1 mM EDTA). Following this, we annealed in a thermocycler with the following program: 95 °C for 10 min, 90 °C for 1 min, reduce the temperature by 1 °C per cycle for 60 cycles, and then hold at 4 °C. The process was repeated for oligo adapter B. To load the adapters onto Tn5, we mixed 5 μL of Tn5, 9 μL of preannealed adapter A, and 9 μL of preannealed adapter B then incubated this mixture in a thermocycler at 37 °C for 30 min. The resulting precharged Tn5 was then diluted with a 1:1 solution of reassociation buffer and glycerol to 1:1 reassociation buffer:glycerol to precharged Tn5.

Mapping of Reads, SNP Calling, and Estimation of Allele Frequencies. Following sequencing, we mapped reads to the *D. melanogaster* reference genome (v6.14) using BWA (3) 53, retained only uniquely mapped reads, and removed PCR generated duplicates using Picard ("Picard Toolkit," 2019). SNPs were called jointly in batch 1 (generations 1,11,25; 1728 samples) and batch 2 (generation 100; 1,116 samples) using the haplotype-based variant detector Freebayes (54), ignoring indels and multiallelic SNPs. Any SNPs with a quality score less than 30, or with a coverage smaller than 28× in any population in batch 1, or smaller than 80× in any population in batch 2, were excluded. We also excluded SNPs with coverage above the genome wide baseline of 167× in any population in batch 1, or above 333× in any population in batch 2, since such highly covered SNPs might be indicative of collapsed repeats. After these filtering steps, we retained 1,741,428 SNPs on the major chromosomes (2L, 2R, 3L, 3R, and X) for subsequent analyses.

Allele frequencies inferred from pooled sequencing can be biased if the coverage per individual in the pool is uneven. Our individually barcoded DNA-libraries allowed us to identify from which individual any given read originates, thereby avoiding this problem. The variance in per sample read depth was substantial, suggesting that allele frequency estimated from Pool-seq, agnostic to read origin, might be error prone. We corrected for uneven coverage when estimating allele frequencies using the formula: $f_{ik} = \frac{1}{n_k} \sum_j \left[AO_{ij} / \left(AO_{ij} + RO_{ij} \right) \right]$, where f_{ik} is the estimated allele frequency of SNP i in pool k, n is the number of individuals in the pool, and AO_{ij} and RO_{ij} are the number of observations of the alternative and reference alleles respectively at SNP i in individual j. The sum is taken over all individuals in pool k, where k is one of the 24 [4 (timepoints) × 3 (replicate populations) × 2 (treatments)] population samples. This strategy corrects for unequal coverage and should be referred over naive Pool-seq estimates. We thus obtained estimates of allele frequency per SNP, in each of the 24 population samples.

We minimized variability in sequence coverage introduced by manual handling of samples by preparing libraries with a CyBio® FeliX liquid handling robot. We pooled libraries according to DNA concentration, and repooled the DNA libraries after preliminary sequencing on the Illumina Miseq platform to normalize coverage.

Inference of Patterns of Polygenic Adaptation Using PCA. To explore the main drivers of genetic changes during the course of the experiment, we performed a PCA of the genome wide allele frequencies. For this analysis, we first assembled a sample-by-SNP matrix P, containing the genome-wide allele frequencies in all 24 population samples (4 time points × 2 treatments × 3 replicate populations). We then computed the PCs of this matrix allowing us to identify the main drivers of allele frequency changes in an unsupervised fashion.

Inference of Individual Loci Selection Signatures. To detect signals of positive selection, we fitted the following logistic regression model:

$$log\left(\frac{p_i}{1-p_i}\right) = \beta_t t_i + \beta_{HS} HS_i + \beta_{HS*t} t_i HS_i + e_i,$$

where p_i denotes the allele frequency at a given locus in a given population and time point; t_i is a numerical variable corresponding to the 4 sampled timepoints (generations 1, 11, 25, and 100 are numerically coded as {1, 2, 3, 4}, and so the corresponding coefficient (β_t) measures the average allele frequency change across all time-points); HS_i is an indicator variable corresponding to the two treatments $HS = 1$ in high-sugar (HS) and $HS = 0$ in control (CONTROL); e_i is an error term. The β parameters are the corresponding regression coefficients. This allowed us to model the allele frequency for every locus across the entire selection experiment in one joint statistical framework. We focus primarily on the interaction effect β_{HS*t}, which quantifies the degree to which the allele frequency trajectory in the control regime differs from the one on the high-sugar regime.

After fitting this model for all SNPs (as described above), we obtained estimates of the effect of time separately for the control and high-sugar selection regimes. This was done using the *emtrends* function in the R package *emmeans* (55). In order to exclude selection signatures that did not correspond to high-sugar adaptation, we disregarded SNPs where the effect of time on the high-sugar selection regime showed a P-value above 10^{-4}.

Individual Level Genotypes. To obtain individual-level genotypes rather than allele frequencies from our low coverage data at generation 100, we first filtered our SNP data more stringently. Having already applied the filter described above, we retained SNPs with called genotypes in more than 90% of the individuals, each genotype being called with a minimum depth of 3. We also excluded individuals with more than 50% missing genotypes. This filtering was done separately for each chromosome, giving a set of 51 k SNPs called in 412 individuals in the control populations, and 52 k SNPs called in 439 individuals on the high-sugar populations. These genotypes were used to estimate linkage disequilibrium (Fig. 1*B*) and to search for signatures of selective sweeps and gametic disequilibrium.

Detecting Selective Sweeps. To detect signatures of selective sweeps, we first used the software *shapeit* (55) to phase the individual genotypes into haplotypes. The estimated haplotypes were then used to calculate the iHS (28). Briefly, iHS measures the length of haplotype homozygosity around a given allele, compared to its alternative allele. A recent selective sweep is expected to leave a genomic footprint of extended homozygosity around the selected allele, whereas selection on standing genetic variation and/or polygenic selection might not leave such a footprint (56). iHS was calculated using the R-package *rehh* (57), and scores were standardized per allele frequency bin as described by Voight et al. (28). We calculated iHS at generation 100 on the high-sugar selected populations, using the 3 replicate populations. To compare selection signatures inferred from our regression model to selective sweeps inferred by iHS, we contrasted the regression P-values to iHS on a SNP-by-SNP basis (*SI Appendix*, Fig. S4).

Transcriptional Changes Associated with Adaptation to High-Sugar Diet.
Experimental design. To identify transcriptional changes associated with genetic adaptation to high-sugar, we performed an experiment that allowed us to robustly differentiate gene expression differences due to the adaptation regime from the plastic response due to short-term changes in dietary condition. For this, we used a full reciprocal design where flies from each replicate cage from generation 170 were allowed to lay eggs in either the dietary condition they evolved in (i.e., HS evolved flies on high-sugar food, CONTROL evolved flies on control food), or in the alternative diet (i.e., HS evolved flies on control diet, CONTROL evolved flies on high-sugar diet). Female flies were collected 7 to 11 d after eclosion, and head and body were separated and plated each in two 96-well plates with each plate containing samples for only one tissue and all four experimental combinations. Plates were stored at −80°C until further processing.
RNA extraction and sequencing. Plates containing heads and bodies were processed in the same way: Sample homogenization was done as described above for DNA samples, and mRNA extraction as described in Suppl. File 2 of (58) using Dynabeads™ mRNA DIRECT™ Purification kit (ThermoFisher), and a final elution of 10 µL and 30 µL Tris-HCl for heads and body, respectively. 3′-enriched RNAseq libraries were prepared following the TM3′seq pipeline (58). In brief, 10 µL of input mRNA was used in the first strand cDNA synthesis reaction which was primed with Tn5Me-B-30 T oligo that binds to the polyA tail of mRNA molecules resulting in 3′ enriched libraries. cDNA was amplified in three rounds of PCR and tagmented using homemade Tn5 transposase. 12 PCR cycles were used for final library amplification using Illumina's i5 and i7 primers. The step by step TM3′ seq protocol can be found in Suppl. File 1 of (58). All libraries within a plate were pooled using 5 µL or 2 µL per head and body library, respectively, and cleaned and size-selected using the double-sided Agencourt AMPure XP bead (Beckman Coulter) cleanup approach described for DNA-seq libraries. The resulting four plate-level libraries were pooled in equal proportions and sequenced on the Illumina NovaSeq S2 platform at the Genomics Core Facility of the Lewis-Sigler Institute for Integrative Genomics at Princeton University. RNA extraction, cDNA synthesis, and library preparation were done in the CyBio® FeliX liquid handling robot.
Processing of RNAseq data. Raw RNA-seq reads were trimmed to remove low quality bases, adapter sequences, and to exclude posttrimmed reads shorter than 20 nt using Trimmomatic 0.32 (59) and the following parameters: SE ILLUMINACLIP:1:30:7 LEADING:3 TRAILING:3 SLIDINGWINDOW:4:15 MINLEN:20. The trimmed reads were mapped to the *D. melanogaster* genome r6.14 using STAR (60), and uniquely mapped reads were assigned to genes using featureCounts from the package Subread (61) and the following parameters: -t exon –g gene_id]. Samples with fewer than 500 k or more than 20 M gene counts, and genes with mean CPM < 1 were removed. After this filtering, the final dataset used in further analysis consisted of 161 head samples with a median of 3.45 M gene counts covering 8,460 genes, and 171 body samples with a median of 2.3 M gene counts covering 8,360 genes.
Differential expression analysis. To identify the transcriptional differences due to adaptation to high-sugar, we performed a differential expression analysis between flies evolved in HS diet and flies evolved in CONTROL diet while accounting for the dietary condition the flies were exposed to for one generation. For each tissue separately, we used a Wald test in DESeq2 (62) and the following design: Expression ~ Plate + Diet + Genotype, where Plate indicates the 96-well plate in which samples were processed from sample collection through library preparation; Diet represents the dietary condition the flies were exposed to for one generation (HS or CONTROL); Genotype represents the diet flies evolved in (HS or CONTROL). Sample size for each of the four groups in body and head, respectively: *n*(genotype HS, diet HS) = 41, 38; n(genotype HS, diet CONTROL) = 46, 42; *n*(genotype CONTROL, diet CONTROL) = 45, 41;

n(genotype CONTROL, diet HS) = 39, 40. *P*-values were estimated for the null hypothesis lfcThreshold = 0 and alpha = 0.05, and adjusted using Benjamini & Hochberg FDR method.

Differentially expressed genes and selection. To investigate the relationship between transcriptional and genomic responses to selection, we analyzed the spatial association between differentially expressed (DE) genes and genomic regions showing signatures of selection. Using the R-package *GenomicRanges*, we identified SNPs within ±5 kb of DE genes that showed significant time-by-selection interactions in our regression analysis. We assessed this overlap across multiple stringency thresholds, using *P*-values ranging from 10^{-5} to 10^{-3} for the time-by-regime effect (Fig. 5*B*, red lines and points).

Given the large number of both DE genes and selected SNPs in our dataset, we expected some overlap to occur by chance. To quantify this background expectation, we implemented a permutation testing approach. For each *P*-value threshold, we randomly sampled genome windows of the same size as our DE genes. We then calculated the fraction selected SNPs that fell within the ±5 kb around these randomly selected windows. We repeated this process 1,000 times for each *P*-value threshold to generate empirical null distributions for the expected overlap (Fig. 5*B*, boxplots). This approach allowed us to assess whether the observed spatial association between DE genes and selected SNPs exceeded what would be expected by chance alone.

Epistatic Selection Signatures.

Wright–fisher model with selection for epistatic QTLs. We used an individual based Wright–Fisher model to investigate the effect of epistatic interactions in the interchromosomal LD in our selection experiment. The simulation was based on the code from (63), using the SLiM modeling framework. We start by creating a neutral burn-in population with 5000 individuals, two equal chromosomes with 300 k sites, a base mutation rate of 1.5×10^{-9}, and a between-site recombination rate of 10^{-8}. This burn-in population is allowed to evolve under a Wright–Fisher neutral model for 50 k generations. With these parameters we expect about 0.5 recombinations per generation, and after 50 k generations we have about 3 k segregating neutral SNPs in mutation-drift equilibrium with a minor allele frequency above 5%. The burn-in process was repeated for each simulation replicate, so each replicate simulation started with a different initial population. For each initial population, we also did experimental replicates, which started from the same starting population.

Using these starting populations, we sampled 1,000 of the segregating mutations to be QTLs contributing to the phenotypic effect of a polygenic trait, with all QTLs having the same phenotypic effect and 2 alleles. We then considered two scenarios: 1) an additive scenario, where the value of the trait is only given by these additive QTLs, and 2) an epistatic scenario, where, in addition to the additive effects, we sampled 200 pairs of QTLs (one member of the pair in each chromosome) to have an additive-by-additive epistatic effect. In this epistatic scenario, the value of the trait depended on these epistatic interactions. Both scenarios proceed with truncation selection on the polygenic trait for 100 generations, with the 500 individuals with the smallest trait value being removed before reproduction in each generation. This strength of selection was chosen so that at the end of the simulation we had only a few fixations (around 10), mimicking the observation in our fly selection experiment.

During the selection phase of the simulation we sampled allele frequencies at regular intervals (every 10 generations) and used these to calculate the correlation between allele frequencies at the QTL pairs. At generation 100, we also measured the gametic disequilibrium between the same pairs of QTLs in both simulations, with the only difference being the presence of the epistatic interaction in one of the scenarios. We compared the mean gametic disequilibrium between these QTL pairs to a distribution of the mean gametic disequilibrium between random pairs of SNPs across chromosomes. To create this distribution, we sampled 200 SNPs in each chromosome and calculated the gametic disequilibrium between the pairs, this process was repeated 10 k times. We created a separate distribution for each scenario.

Identifying well-supported epistatic SNP pairs. To detect potential instances of epistatic selection, we looked for two types of signals: 1) correlations between the allele frequencies at different loci; 2) gametic disequilibrium after 100 generations of selection. The former could be due to epistasis or due to similar but independent selection coefficients at the respective loci, while the latter is only expected under epistasis. The correlations were estimated by the Pearson correlation coefficient, using allele frequencies in all generations on the high-sugar populations ($n = 12$ per SNP pair). The gametic disequilibrium was quantified separately in high-sugar and control populations, using the individual genotypes at generation 100 described above. For each SNP pair, we tested the deviation from independent segregation using a chi square test ($n \sim 420$ per SNP pair). Having identified candidate pairs where the two SNPs displayed gametic disequilibrium and were located on different chromosomes, we attempted to find well supported signals by clustering physically close SNPs with a similar signal of gametic disequilibrium. To do this, we applied a clustering procedure akin to the LD clumping algorithm implemented in PLINK (64). The algorithm works as follows:

i) Starting with the SNP pair with the smallest *P*-value for gametic disequilibrium, assign to the same cluster all other SNP pairs with one SNP on chromosome 2 that are within 250 kb, in linkage disequilibrium, and have not already been clustered. Repeat until there are no more SNP-pairs to assign to clusters. Thus, each cluster contains SNP pairs sharing proximal and linked SNPs on chromosome 2.
ii) For each cluster identified in i), perform a second round of clustering for the SNP pairs within that cluster. This is done by assigning to the same cluster all SNP-pairs with one SNP on chromosome 3 that are within 250 kb, in linkage disequilibrium, and have not already been clustered.

We thus obtain hierarchical clusters, where each "chromosome 2 end" cluster contains one or several "chromosome 3 end" clusters. The algorithm is greedy, so each SNP pair will only end up in one cluster, if at all. Finally, we keep the clusters with at least three linked SNPs at "each end". We thus identify epistatic selection signatures supported by multiple SNP pairs, with multiple linked SNPs at each end of the putative interaction. Through the nature of our clustering procedure, each such signature will involve one locus on chromosome 2, and one or several loci on chromosome 3 (Fig. 6*B*).

Data, Materials, and Software Availability. Genomic Sequences data is deposited in SRA under PRJNA1271221 (65). Code to reproduce figures in this paper is available at https://doi.org/10.5281/zenodo.15390444 (66).

ACKNOWLEDGMENTS. We would like to thank the reviewer for the time and substance comments. A genuine thank you to Grace Chi for her help maintaining the population cages for many years. We acknowledge our funders: S.K.G.F. supported by the Swedish Research Council (2017-06229) and European Molecular Biology Organization (632-2017), J.F.A. funded by NIH-National Institute of General Medical Sciences R35GM124881-04 and NIH-National Institute of Environmental Health Sciences R01ES029929, and L.F.P. supported by a Long-Term Postdoctoral Fellowship from the Human Frontiers Science Program. We also acknowledge that the work reported in this paper was substantially performed using the Princeton Research Computing resources at Princeton University, which is a consortium of groups led by the Princeton Institute for Computational Science and Engineering and Office of Information Technology's Research Computing. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Author affiliations: [a]Lewis-Sigler Institute for Integrative Genomics, Princeton University, Princeton, NJ 08544; [b]Department of Ecology and Evolutionary Biology, Princeton University, Princeton, NJ 08544; and [c]Department of Molecular Biology and Genetics, Cornell University, Ithaca, NY 14850

1. P. Pavlidis, N. Alachiotis, A survey of methods and tools to detect recent and strong positive selection. *J. Biol. Res-Thessaloniki* **24**, 7 (2017).
2. J. M. Smith, J. Haigh, The hitch-hiking effect of a favourable gene. *Genet. Res.* **23**, 23–35 (1974).
3. N. Barghi *et al.*, Genetic redundancy fuels polygenic adaptation in *Drosophila*. *PLoS Biol.* **17**, e3000128 (2019).
4. J. K. Pritchard, A. Di Rienzo, Adaptation – not by sweeps alone. *Nat. Rev. Genet.* **11**, 665–667 (2010).
5. A. F. Feder *et al.*, More effective drugs lead to harder selective sweeps in the evolution of drug resistance in HIV-1. *eLife* **5**, e10670 (2016).
6. N. R. Garud, P. W. Messer, E. O. Buzbas, D. A. Petrov, Recent selective sweeps in North American *Drosophila melanogaster* show signatures of soft sweeps. *PLoS Genet.* **11**, e1005004 (2015).

7. T. F. Hansen, The evolution of genetic architecture. *Annu. Rev. Ecol. Evol. Syst.* **37**, 123–157 (2006).
8. J. F. Crow, On epistasis: Why it is unimportant in polygenic directional selection. *Phil. Trans. R. Soc. B* **365**, 1241–1244 (2010).
9. A. Le Rouzic, Estimating directional epistasis. *Front. Genet.* **5**, 198 (2014).
10. P. C. Phillips, Epistasis – the essential role of gene interactions in the structure and evolution of genetic systems. *Nat. Rev. Genet.* **9**, 855–867 (2008).
11. D. M. Weinreich, Y. Lan, C. S. Wylie, R. B. Heckendorn, Should evolutionary geneticists worry about higher-order epistasis? *Curr. Opin. Gene. Dev.* **23**, 700–707 (2013).
12. K. Csilléry, A. Rodríguez-Verdugo, C. Rellstab, F. Guillaume, Detecting the genomic signal of polygenic adaptation and the role of epistasis in evolution. *Mol. Ecol.* **27**, 606–612 (2018).
13. R. B. Corbett-Detig, J. Zhou, A. G. Clark, D. L. Hartl, J. F. Ayroles, Genetic incompatibilities are widespread within species. *Nature* **504**, 135–137 (2013).
14. L. Boyrie, C. Moreau, F. Frugier, C. Jacquet, M. Bonhomme, A linkage disequilibrium-based statistical test for genome-wide epistatic selection scans in structured populations. *Heredity* **126**, 77–91 (2021).
15. N. Barghi, J. Hermisson, C. Schlötterer, Polygenic adaptation: A unifying framework to understand positive selection. *Nat. Rev. Genet.* **21**, 769–781 (2020).
16. S. G. Das, S. O. Direito, B. Waclaw, R. J. Allen, J. Krug, Predictable properties of fitness landscapes induced by adaptational tradeoffs. *eLife* **9**, e55155 (2020).
17. C. B. Ogbunugafor, The mutation effect reaction norm (mu-rn) highlights environmentally dependent mutation effects and epistatic interactions. *Evolution* **76**, 37–48 (2022).
18. W. A. Chng, V. Hietakangas, B. Lemaitre, Physiological adaptations to sugar intake: New paradigms from *Drosophila melanogaster*. *Trends Endocrinol. Metab.* **28**, 131–142 (2017).
19. D. A. Hickey, B. Benkel, Regulation of amylase activity in *Drosophila melanogaster:* Effects of dietary carbohydrate. *Biochem. Genet.* **20**, 1117–1129 (1982).
20. M. Dus *et al.*, Nutrient sensor in the brain directs the action of the brain-gut axis in *Drosophila*. *Neuron* **87**, 139–151 (2015).
21. W. A. Chng, M. S. B. Sleiman, F. Schüpfer, B. Lemaitre, Transforming growth factor β/Activin signaling functions as a sugar-sensing feedback loop to regulate digestive enzyme expression. *Cell Rep.* **9**, 336–348 (2014).
22. J. Mattila *et al.*, Mondo-Mlx mediates organismal sugar sensing through the gli-similar transcription factor sugarbabe. *Cell Rep.* **13**, 350–364 (2015).
23. L. Palanker Musselman *et al.*, A high-sugar diet produces obesity and insulin resistance in wild-type *Drosophila*. *Disease Models Mech.* **4**, 842–849 (2011).
24. J. Na *et al.*, A *Drosophila* model of high sugar diet-induced cardiomyopathy. *PLoS Genet.* **9**, e1003175 (2013).
25. L.-M. Chevin, F. Hospital, Selective sweep at a quantitative trait locus in the presence of background genetic variation. *Genetics* **180**, 1645–1660 (2008).
26. P. Pavlidis, D. Metzler, W. Stephan, Selective sweeps in multilocus models of quantitative traits. *Genetics* **192**, 225–239 (2012).
27. T. Paixão, N. H. Barton, The effect of gene interactions on the long-term response to selection. *Proc. Natl. Acad. Sci. U.S.A.* **113**, 4422–4427 (2016).
28. B. F. Voight, S. Kudaravalli, X. Wen, J. K. Pritchard, A map of recent positive selection in the human genome. *PLoS Biol.* **4**, e72 (2006).
29. F. W. Albert, L. Kruglyak, The role of regulatory variation in complex traits and disease. *Nat. Rev. Genet.* **16**, 197–212 (2015).
30. S. Kryazhimskiy, D. P. Rice, E. R. Jerison, M. M. Desai, Global epistasis makes adaptation predictable despite sequence-level stochasticity. *Science* **344**, 1519–1522 (2014).
31. I. Höllinger, P. S. Pennings, J. Hermisson, Polygenic adaptation: From sweeps to subtle frequency shifts. *PLoS Genet.* **15**, e1008035 (2019).
32. K. Jain, W. Stephan, Response of polygenic traits under stabilizing selection and mutation when loci have unequal effects. *G3 Genes|Genomes|Genetics* **5**, 1065–1074 (2015).
33. B. C. Haller, P. W. Messer, SLiM 3: Forward genetic simulations beyond the wright-fisher model. *Mol. Biol. Evol.* **36**, 632–637 (2019).
34. L. F. Pallares *et al.*, Dietary stress remodels the genetic architecture of lifespan variation in outbred *Drosophila*. *Nat. Genet.* **55**, 123–129 (2023).
35. L. Yengo *et al.*, A saturated map of common genetic variants associated with human height. *Nature* **610**, 704–712 (2022).
36. U. Võsa *et al.*, Large-scale cis- and trans-eQTL analyses identify thousands of genetic loci and polygenic scores that regulate blood gene expression. *Nat. Genet.* **53**, 1300–1310 (2021).
37. H. Götsch, R. Bürger, Polygenic dynamics underlying the response of quantitative traits to directional selection. *Theor. Popul. Biol.* **158**, 21–59 (2023), http://biorxiv.org/lookup/doi/10.1101/2023.02.23.529647.
38. S. K. G. Forsberg, J. S. Bloom, M. J. Sadhu, L. Kruglyak, Ö. Carlborg, Accounting for genetic interactions improves modeling of individual quantitative trait phenotypes in yeast. *Nat. Genet.* **49**, 497–503 (2017).
39. J. M. Cheverud, E. J. Routman, Epistasis and its contribution to genetic variance components. *Genetics* **139**, 1455–1461 (1995).
40. J. M. Cheverud, E. J. Routman, Epistasis as a source of increased additive genetic variance at population bottlenecks. *Evolution* **50**, 1042 (1996).
41. N. H. Barton, How does epistasis influence the response to selection? *Heredity* **118**, 96–109 (2017).
42. T. F. Hansen, Why epistasis is important for selection and adaptation: Perpective. *Evolution* **67**, 3501–3511 (2013).
43. T. F. Hansen, G. P. Wagner, Modeling genetic architecture: A multilinear theory of gene interaction. *Theor. Popul. Biol.* **59**, 61–86 (2001).
44. A. Le Rouzic, M. Roumet, A. Widmer, J. Clo, Detecting directional epistasis and dominance from cross-line analyses in alpine populations of *Arabidopsis thaliana*. *J. Evol. Biol.* **37**, 838–847 (2023), http://biorxiv.org/lookup/doi/10.1101/2023.04.19.537438.
45. M. Pavlicev, A. Le Rouzic, J. M. Cheverud, G. P. Wagner, T. F. Hansen, Directionality of epistasis in a murine intercross population. *Genetics* **185**, 1489–1505 (2010).
46. A. G. Jones, R. Bürger, S. J. Arnold, Epistasis and natural selection shape the mutational architecture of complex traits. *Nat. Commun.* **5**, 3709 (2014).
47. D. Melo, G. Marroig, Directional selection can drive the evolution of modularity in complex traits. *Proc. Natl. Acad. Sci. U.S.A.* **112**, 470–475 (2015).
48. M. Pavlicev, J. M. Cheverud, G. P. Wagner, Evolution of adaptive phenotypic variation patterns by direct selection for evolvability. *Proc. R. Soc. B.* **278**, 1903–1912 (2011).
49. C. Bank, Epistasis and adaptation on fitness landscapes. *Annu. Rev. Ecol. Evol. Syst.* **53**, 457–479 (2022).
50. J. K. Grenier *et al.*, Global diversity lines–A five-continent reference panel of sequenced *Drosophila melanogaster* Strains. *G3 Genes|Genomes|Genetics* **5**, 593–603 (2015).
51. R. C. Fuller, How and when selection experiments might actually be useful. *Integr. Comp. Biol.* **45**, 391–404 (2005).
52. S. Picelli *et al.*, Tn5 transposase and tagmentation procedures for massively scaled sequencing projects. *Genome Res.* **24**, 2033–2040 (2014).
53. H. Li, R. Durbin, Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).
54. E. Garrison, G. Marth, Haplotype-based variant detection from short-read sequencing. arXiv [Preprint] (2012), http://arxiv.org/abs/1207.3907 (Accessed 17 July 2012).
55. O. Delaneau, C. Coulonges, J.-F. Zagury, Shape-IT: New rapid and accurate algorithm for haplotype inference. *BMC Bioinform.* **9**, 540 (2008).
56. M. Lynch, W. Bruce, *Genetics and Analysis of Quantitative Traits* (Sinauer Sunderland, MA, 1998).
57. M. Gautier, R. Vitalis, *rehh*: An R package to detect footprints of selection in genome-wide SNP data from haplotype structure. *Bioinformatics* **28**, 1176–1177 (2012).
58. L. F. Pallares, S. Picard, J. F. Ayroles, TM3'seq: A tagmentation-mediated 3′ sequencing approach for improving scalability of RNAseq experiments. *G3 Genes|Genomes|Genetics* **10**, 143–150 (2020).
59. A. M. Bolger, M. Lohse, B. Usadel, Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120 (2014).
60. A. Dobin *et al.*, STAR: Ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).
61. Y. Liao, G. K. Smyth, W. Shi, The Subread aligner: Fast, accurate and scalable read mapping by seed-and-vote. *Nucleic Acids Res.* **41**, e108 (2013).
62. M. I. Love, W. Huber, S. Anders, Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).
63. R. N. Lou, N. O. Therkildsen, P. W. Messer, The effects of quantitative trait architecture on detection power in short-term artificial selection experiments. *G3 Genes|Genomes|Genetics* **10**, 3213–3227 (2020).
64. S. Purcell *et al.*, PLINK: A tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* **81**, 559–575 (2007).
65. S. K. G. Forsberg *et al.*, Longitudinal sequencing reveals polygenic and epistatic nature of genomic response to selection. NCBI Sequence Read Archive. https://www.ncbi.nlm.nih.gov/bioproject/PRJNA1271221. Deposited 2 June 2025.
66. D. Melo *et al.*, Drosophila high sugar selection. Zenodo. https://doi.org/10.5281/zenodo.15390444. Deposited 12 May 2025.