Intersecting experimental evolution and CRISPR screens to identify novel insecticide resistance loci

Michele Marconcini[1], Steeve Cruchet[1], Srishti Goswami[2], Raghuvir Viswanatha[2], Matthew Butnaru[2,3], Joydeep De[4], Camilla Roselli[4], Dafni Hadjieconomou[4], Norbert Perrimon[2,3], Stephanie E. Mohr[2], Richard Benton[1,*]

[1]Center for Integrative Genomics
Faculty of Biology and Medicine
University of Lausanne
CH-1015
Lausanne
Switzerland

[2]Department of Genetics
Blavatnik Institute
Harvard Medical School
Boston
MA 02115
USA

[3]Howard Hughes Medical Institute
Boston
MA 02115
USA

[4]Institut du Cerveau-Paris Brain Institute
Sorbonne Université
Inserm
CNRS
Hôpital Pitié-Salpêtrière
75013
Paris
France

*Corresponding author:
E: richard.benton@unil.ch
T: ++41(0)21 692 3932

Abstract

Understanding the evolution of insect resistance to natural and artificial poisons is important both to appreciate the adaptation of species to toxic ecological niches[1] and manage pest insects[2]. While resistance to chemicals that target neuronal ion channels is well-appreciated[3,4], dissecting toxin resistance mechanisms is hampered by its complex genetic basis. Here, we investigated how *Drosophila sechellia* has evolved resistance to octanoic acid (OA), a historically-important paradigm of adaptation to a natural toxin[5,6]. OA is an abundant component of *Morinda citrifolia* noni fruit – the exclusive niche of *D. sechellia* – and is toxic to many other insects, including its close relatives *Drosophila simulans* and *Drosophila melanogaster*[5,6]. We experimentally evolved *D. simulans* populations with increased OA resistance, identifying hundreds of loci displaying signatures of selection. Cross-referencing of these loci with the results of a genome-wide, CRISPR knockout screen for OA tolerance genes in *D. melanogaster* S2R+ cells highlighted two proteins of interest: Kraken, a putative detoxification enzyme of the serine hydrolase family expressed most strongly in the digestive and renal systems[7,8], and Alkbh7, a ubiquitously-expressed, mitochondrial-localised protein, whose mammalian orthologue functions in fatty acid metabolism[9,10]. In *D. sechellia*, *kraken* and *Alkbh7* display both signs of selection and elevated expression, mirrored by their higher transcription in artificially-evolved, OA-resistant *D. simulans*. In *D. melanogaster*, loss-of-function of *kraken*, but not *Alkbh7*, led to increased susceptibility to OA; conversely, overexpression of *Alkbh7*, but not *kraken*, increased OA resistance. Importantly, mutation of these genes in *D. sechellia* demonstrated a contribution of these proteins to the natural resistance of this species, identifying the first specific loci underlying this multigenic trait. Our results suggest an adaptive role of these genes in shaping OA resistance both under laboratory selection and during *D. sechellia*'s evolutionary history. More generally, this work emphasises the experimental power of drosophilids to dissect multilayered molecular mechanisms of toxin resistance, with applications in the development of novel insecticides.

Introduction

How insects resist toxic chemicals in the environment is of both basic scientific interest – reflecting insects' arms race with plant food sources[1,11] or venomous predators[12] – and applied interest, because humans have long used chemicals to combat insect disease vectors and agricultural pests[2]. The best-understood insecticides are those that bind to and modulate neuronal ion channels, where resistance arises through changes in these target proteins[3,4]. For example, dichlorodiphenyltrichloroethane (DDT) and pyrethroids prevent closure of voltage-gated sodium channels, resulting in neuronal excitotoxicity and organismal paralysis; diverse insect species displaying resistance for these chemicals have independently-arising mutations in genes encoding these channels[13,14]. Similarly, neonicotinoids bind to and hyperactivate nicotinic acetylcholine receptors, leading to neuronal death; mutations in various receptor subunits can confer resistance[15]. Even in such well-studied examples, however, it is thought that other neuronal targets exist[16] and/or that the chemicals impact receptors in non-neuronal tissues[17]. For many insecticides, the mode of action is incompletely, or not at all, characterised[4].

Resistance observed in nature is only partially understood, reflecting the contribution of many genes beyond those encoding target proteins, which can underlie distinct defence mechanisms[18-20]. For example, thickening of the cuticle to reduce entry of insecticides into the body is widely considered an effective resistance mechanism[21,22]. Cuticular changes have been linked to overexpression of proteins implicated in exoskeleton development (e.g., cuticle proteins, ABC transporters) but their precise roles, and how they are upregulated, remain unclear[22]. Increased expression of metabolic enzymes that degrade insecticides is another commonly observed adaptation in resistant strains, sometimes resulting from copy-number increases in the corresponding genes[23]. Duplicated metabolic genes can also undergo neofunctionalisation to produce enzymes with altered specificity, as illustrated by Cytochrome P450s in strains of the brown planthopper that recently evolved resistance to the neonicotinoid imidacloprid[24], as well as by more ancient evolution of glutathione S-transferases during transition of the *Scaptomyza* genus of Drosophilidae to herbivory to counter toxic isothiocyanates[25]. However, the lack of experimental resources in many insect species renders it difficult to characterise the causal contribution of specific genetic changes.

A compelling natural model clade to study the evolution of toxin resistance is the trio of closely-related drosophilid species, *Drosophila melanogaster*, *Drosophila simulans* and *Drosophila sechellia* (Figure 1a). *D. sechellia* is an extreme specialist on the toxic noni fruit of the *Morinda citrifolia* shrub[5,6]. Amongst *D. sechellia*'s numerous adaptations[6,26-30], this species has evolved high resistance to octanoic acid (OA), an abundant noni chemical that is toxic to other drosophilids and more divergent insects[6,31-36]. Previous studies of *D. sechellia*'s resistance to OA has been studied through genomic mapping-based approaches in *D. sechellia*/*D. simulans* hybrids, which revealed this trait to be polygenic, comprising at least five genomic regions[31,37-39]. The main effect locus encompasses a cluster of 18 genes including members of the *Osiris* (*Osi*) family, which are thought to function in cuticle patterning[40,41]. RNA interference (RNAi) knockdown of some *Osi* genes in *D. melanogaster* led to reduced tolerance of OA[42,43]; however, recent analyses indicate a general decrease in viability of *Osi* null mutant *D. melanogaster*, even in the absence of OA[40]. Moreover, genetic evidence for the contribution of these genes to OA resistance in *D. sechellia* is lacking.

Given the challenges of identifying candidate genes through genomic mapping-based strategies, in this work we took two complementary approaches to investigate the evolution of OA resistance in this drosophilid trio: experimental evolution in *D. sechellia*'s closest relative, *D. simulans*, through an Evolve-and-Resequence strategy[44], and CRISPR/Cas9-based genome-wide genetic screening in cultured *D. melanogaster* cells[45]. We show how the intersection of these approaches can identify novel genes underlying OA resistance in *D. sechellia*, establishing these methods and this drosophilid clade as a powerful paradigm to investigate the evolution and mechanisms of insecticide resistance.

Results

Experimental evolution of OA resistance in *D. simulans*

The Evolve-and-Resequence approach uses high-throughput sequencing to analyse the entire genome of populations before and after exposure to specific selective pressures, thereby facilitating the identification of adaptive loci from standing genetic variation in controlled and replicable settings[44,46-49]. For studies of insecticide resistance, it has been predominantly used to examine changes in frequency of alleles of known genes rather than as a means to identify novel loci[50,51].

The susceptibility/resistance of drosophilids to OA can be easily assessed by exposing flies to the toxin in a culture tube and counting the survivors after 24 h (Figure 1b). In line with previous observations[37,52], we found that a dose of 3 μl of pure OA (mixed with 1 ml glucose solution) leads to death of ~60% of *D. simulans* but has no effect on *D. sechellia* (Figure 1c). By collecting the *D. simulans* survivors, we reasoned that we could allow these to mate and establish the next generation (Figure 1b), thus enriching genetic variants conferring higher resistance. Repeating the process over multiple generations and multiple replicates would allow us to track phenotypic changes in resistance and investigate the associated trajectories in allele frequencies.

Standing genetic variation should be present in the evolving populations for natural selection to act upon[53,54]. As the spontaneous mutation rate in drosophilids is too low for artificial selection[55], we generated ten independent polymorphic base populations of *D. simulans* from 93 isogenic strains[56] by round-robin crossing (Figure 1d). To minimise bottleneck effects, each population was kept at a size of ~1000 flies. At each generation, ~500 flies from each population were randomly collected and selected for resistance to 3 μl OA. Over the course of 25 generations, all the populations showed increased levels of survivability, suggesting parallel adaptive evolution (Figure 1e). Given the consistent responses across all populations, we also analysed the phenotypic changes grouping all replicates, finding that the mean survivability significantly increased from ~43% at generation 0 to ~78% at generation 25 (Figure 1f).

As we kept the same amount of OA throughout the 25 generations of selection, the selective pressure lowered as the populations' survivability increased, evidenced as a larger fraction of flies surviving and reproducing. Viewing the phenotype of populations both individually (Figure 1e) and collectively (Figure 1f), it was apparent that the change in phenotype plateaued at around generation 15. For continued evolution, we therefore gradually increased the amount of OA during generations 26-31 (Figure 1g), reaching a selection pressure comparable to the initial phase of the experiment with 9 μl OA (~38% mean survivability across all populations) (Figure 1g). We then performed a second round of selection at this dose, observing an increase in

survivability across all populations, reaching ~70% by generation 50 (Figure 1h). As the populations adapted to the new condition, a second plateau started around generation 40 (Figure 1h-i). Over 50 generations we were therefore able to evolve *D. simulans* to have approximately 5-fold higher survivability to 9 µl OA compared to the unevolved populations (Figure 1j), though still falling short of the survival exhibited by *D. sechellia* (Figure 1j). OA is only one component of noni fruit, albeit probably the major toxic chemical[34]. We therefore tested three of our evolved populations for resistance to noni fruit pulp, finding that they exhibited much higher survival on this substrate compared to the same original populations maintained under neutral conditions (i.e., no OA selection) (Figure 1k).

Specialisation of *D. sechellia* on toxic noni is associated with a reduced egg-laying capacity[57,58] and a shorter lifespan[29,59,60]. Our success in evolving *D. simulans* with a partial *D. sechellia*-like resistance trait begs the question as to whether these selected strains exhibit other phenotypic changes as a trade-off for higher toxin resistance. We tested whether the evolved resistant populations of *D. simulans* exhibit similar phenotypic changes in control conditions. However, none of the three focal populations tested exhibited significant decreases in the number of eggs laid (Figure 1l), developmental viability (Figure 1m) or lifespan (Figure 1n). Thus, at least within the timeframe of our experimental evolution experiment, *D. simulans* did not display obvious phenotypic trade-offs for the gain in OA resistance.

Dynamic directional selection of allele frequencies

We hypothesised that the observed phenotypic changes in *D. simulans* depended on the directional selection of alleles of consistent loci across the ten populations, while most of the genome evolves under neutrality. Allele frequencies at the selected loci should therefore consistently change across all populations, allowing us to statistically distinguish them from randomly evolving loci (Figure 2a). We sequenced the genomes of the populations of evolved (generation 25 and generation 50) and unevolved (generation 0) flies, identifying 5,127,840 high-quality variants distributed across all chromosomes. Pairwise identity-by-state analysis indicated that as the number of generations increased, the proportions of shared alleles decreased, both between time points and between pairs of populations (left-to-right and top-to-bottom, respectively, in Figure 2b).

To identify the selected genetic variants underlying the phenotypic adaptation, we analysed the genome-wide allele frequency changes both between generations 0 and 25, and between generations 0 and 50, across all populations. The identification of significant changes in frequency is non-trivial, and many methods have been developed[61-63]. We therefore assessed the significance of frequency changes using two complementary approaches: the Cochran–Mantel–Haenszel (CMH) test[64,65], a widely-used method that tests for allele frequency changes between two time points, and Bait-ER[63], a Bayesian method that tests for selection while explicitly modelling genetic drift. Both tests identified large portions of the genome under positive selection at generation 25 and 50 (Figure 2c,d). The majority of prominent peaks were found by both methods and across time points (Figure 2e).

Based on the analysis of genome-wide linkage disequilibrium decay (Extended Data Figure 1a), we grouped the significant SNPs identified by each of the two methods into genomic blocks according to their proximity (i.e., <50 kb). We retrieved all the candidate genes that fall within the block intervals, obtaining a total of 439 and 401 genes for the CMH and Bait-ER, respectively, at generation 25, and 1238 and 174

genes at generation 50 (Data S3). These genes likely include true targets of selection but also neighbouring genes affected by hitchhiking. The increase in CMH candidate genes over time – in contrast to the decrease observed with Bait-ER – might reflect the CMH test's tendency to overestimate significance and to be insensitive to divergence among replicate populations.

The overlap of all significant genes between the two time points is <30%. This might reflect that a large fraction of the candidate genes results from noise caused by genetic drift and draft (hitchhiking) and/or reflect a shift in the phenotypic contribution of genes under the stronger selection (i.e., the higher dose of OA) during generations 27-50. Moreover, previous work on *Drosophila* using temporal sampling showed that adaptation involving complex traits can lead to non-overlapping sets of selected SNPs, in contrast with the simple expectation of continuous increases in allele frequency until fixation[44,66].

A genome-wide, pooled CRISPR screen for OA susceptibility and resistance genes in *D. melanogaster* S2R+ cells

Although the experimental evolution experiment pointed to a number of genomic regions relevant for OA resistance, they contained too many genes to enable facile choice of candidates for functional follow-up. We therefore sought an orthogonal screening approach. Motivated by previous observations of OA toxicity for insect Sf9 cultured cells[36], we found that OA also leads to reduced viability of *D. melanogaster* S2R+ cells (Figure 3a). This phenotype allowed us to perform a selection experiment in this cell line through exploitation of a genome-wide, pooled CRISPR/Cas9 mutagenesis approach (Figure 3b)[45]. In brief, a library of sgRNA-encoding plasmids (6-8 sgRNAs/gene) was transfected into Cas9-expressing cells where site-specific plasmid integration predominantly limits targeting to a single gene per cell. Two replicate cultures of this pool of cells were either left untreated, or exposed to OA in the medium, at two different concentrations (Figure 3b). After 26 days in culture, DNA was extracted from the cell pool, then sgRNA sequences were PCR amplified and sequenced to determine those that were enriched after OA selection (reflecting genes whose loss enhances cell viability, referred to hereafter as "susceptibility genes") or depleted (reflecting genes whose loss leads to lower OA resistance, hereafter "resistance genes").

To identify candidate OA susceptibility/resistance genes, we first filtered out essential genes[45] and genes that were not detectably expressed in S2R+ cells (using DGET[67]). After comparing results from each condition and replicate, we identified the top 5% susceptibility genes (Figure 3c) and top 5% resistance genes (Figure 3d). Amongst the 87 susceptibility genes, there was no single stand-out hit, suggesting that – at least in S2R+ cells – OA does not have a single major cellular target (unlike, for example, insecticidal bacterial toxins that bind a specific cell surface receptor[68]). Gene-ontology analysis of these susceptibility genes revealed a significant enrichment of genes involved membrane trafficking (Figure 3c). As OA has previously been hypothesised to disrupt cellular membranes[36,69], the recovery of these hits suggests that general reduction in membrane transport reduces OA uptake into cells and/or renders membranes less susceptible to OA treatment. 22 OA resistance genes were identified, encoding diverse types of proteins (Figure 3d), suggesting there are multiple mechanisms by which cells can resist the toxic effects of OA.

We next examined which of the genes identified in the S2R+ cell selection experiments might be relevant for evolution of OA resistance in whole animals by

determining those also present within significant blocks of the experimental evolution analyses (Figure 3e,f). Four susceptibility genes were identified: CG1620 (a predicted SANT-MYB domain transcription factor), CG4291 (a putative component of the splicing machinery), *Nph* (a chromatin remodeling factor[70]) and *Debcl* (a pro-apoptotic member of the Bcl-2 family[71]); these are all likely to have indirect roles in conferring susceptibility to OA given their broad functions. By contrast, two resistance genes have potentially direct roles in conferring tolerance to OA, *kraken* and CG14130. *kraken* encodes a putative serine hydrolase[8] and stood out both because it was the top hit in the resistance gene screen (Figure 3d), and because of its location in one of the most prominent peaks in the experimental evolution at generations 25 and 50 (Figure 3f and Figure 2c-d). CG14130 encodes the orthologue of mammalian *Alkbh7*, which is involved in regulating fatty acid metabolism[9,10], and is located within a significant peak at generation 25 (Figure 3f and Figure 2c). Therefore, intersecting the experimental evolution and CRISPR screening results allowed us to narrow down tens or hundreds of potential candidates to two focal genes for downstream investigations.

Sequence, expression and evolutionary analyses of candidate OA resistance genes

To assess the candidacy of *kraken* and CG14130 (hereafter *Alkbh7*) in contributing to OA resistance in animals, we first examined their sequence and expression in *D. melanogaster*. Kraken belongs to a large family of serine hydrolases, which encompass lipases, esterases and proteases that use a nucleophilic serine residue in the active site for hydrolysis of substrates[8]. Consistent with the conservation of this serine in Kraken (Figure 4a and Extended Data Figure 2a), an activity-based proteomic screen of the serine hydrolase superfamily in *D. melanogaster* provided evidence for *in vivo* catalytic activity of Kraken[7]. *kraken* is broadly expressed in adults, with particular enrichment in the gut and Malpighian tubules (Figure 4b), similar to previous observations of *kraken* expression in the embryo by RNA *in situ* hybridisation[8]. Using the Fly Cell Atlas[72], we examined sub-tissue adult expression of *kraken*: in the Malpighian tubules, *kraken* is expressed at highest levels in the principal cells (Figure 4c), which are involved in ion transport and detoxification. In the gut, *kraken* transcripts are most robustly detected in various types of enterocytes, which secrete digestive enzymes[73] (Figure 4c). However, Kraken lacks an N-terminal signal sequence[74], suggesting that the protein remains in the cytoplasm of these cells, unlike canonical digestive enzymes[73], and more similar to detoxification enzymes of the Cytochrome P450 and Glutathione S-transferase families[75]. To examine *kraken* expression *in situ*, we generated a *kraken* promoter-Gal4 driver line, which induced expression of a fluorescent reporter in cells of the proventriculus, the crop duct and the hindgut (Figure 4d). No GFP expression was detected in Malpighian tubules (data not shown), which we suspect reflects the lack of necessary regulatory sequences in the driver transgene.

Alkbh7 is a member of the alpha-ketoglutarate-dependent hydroxylase (Alkbh) family (Figure 4a), which are enzymes that act on diverse substrates in processes such as biosynthesis, metabolism and epigenetic regulation[76]. As in mammals[10], *D. melanogaster Alkbh7* displays broad tissue expression (Figure 4b). Murine Alkbh7 is localised to the mitochondrial matrix[10]. This property appears to be conserved in the *Drosophila* protein, which has a mitochondrial targeting sequence (Figure 4a and Extended Data Figure 2b)[77]; consistently, we identified Alkbh7 within *D. melanogaster* proteomes of the mitochondrial matrix[78,79].

Although the sequences of both Kraken and Alkbh7 are overall highly similar across *D. melanogaster*, *D. simulans* and *D. sechellia* (Extended Data Figure 2a,b), we wondered whether we could detect signs of selection at these loci in *D. sechellia* employing population sequence datasets of *D. sechellia* and *D. simulans*[80]. Using the McDonald-Kreitman test[81] on the coding sequences of both genes, we found evidence for gene-level positive selection on *kraken* but not *Alkbh7* (Extended Data Figure 2c). Because selection might act pervasively at individual sites rather than along the entire protein coding region, we also implemented a codon-based analysis using FUBAR (Fast, Unconstrained Bayesian AppRoximation)[82], which identified 1 and 4 codons under diversifying positive selection in *kraken* and *Alkbh7,* respectively (Extended Data Figure 2d). Although their functional relevance remains unclear, such mutations might reflect underlying evolutionary dynamics at the protein level.

We next compared the expression levels of these genes both in laboratory strains of *D. sechellia* and *D. simulans* (Figure 4e), as well as in RNA-seq datasets of various strains of *D. sechellia*, *D. simulans* and *D. melanogaster*[83-85] (Figure 4f, Extended Data Figure 3a,b). While there were some exceptions, overall both *kraken* and *Alkbh7* show higher expression levels in *D. sechellia* adults and larvae compared to the other species. Importantly, in our experimentally-evolved *D. simulans* populations, we also found that both *kraken* and *Alkbh7* were more highly expressed compared to unselected populations (Figure 4g). The expression differences between species, and between unevolved and evolved *D. simulans*, reflect variation in basal levels of expression, as prior exposure to OA did not lead to changes in expression in either *D. sechellia* or *D. simulans* (Extended Data Figure 3c).

Functional contributions of *kraken* and *Alkbh7* to OA resistance

Given the expression differences observed in the naturally and experimentally evolved OA-resistant drosophilids, we asked whether artificially modulating gene expression would impact OA tolerance. We first performed ubiquitous RNAi of these genes in *D. melanogaster* (Extended Data Figure 4a) and assessed the effects on resistance to OA using a short-term, plate-based assay[31] (see Methods). *kraken*^{RNAi} led to significantly diminished OA resistance compared to genetic controls, while *Alkbh7*^{RNAi} did not have a discernable effect (Figure 5a). We next generated deletion mutants for these genes in *D. melanogaster* (Extended Data Figure 4b). Similar to the RNAi experiments, *kraken* mutants, but not *Alkbh7* mutants, displayed reduced resistance to OA (Figure 5b). All genotypes remained fully viable when no OA was added (Extended Data Figure 5).

Loss-of-function analyses might fail to uncover a role for *Alkbh7* in whole animals if there are redundant resistance mechanisms absent from S2R+ cells. We therefore took a complementary approach by overexpressing these genes in *D. melanogaster*. Strikingly, overexpression of *Alkbh7* resulted in survivability of *D. melanogaster* to levels close to those observed for *D. sechellia* under the same experimental conditions (Figure 5c, Extended Data Figure 4c), indicating that this gene is sufficient to confer high-level OA resistance. By contrast, *kraken* overexpression did not have any detectable impact on OA resistance (Figure 5c); a similar lack of effect was seen when overexpressing *D. sechellia kraken* (Extended Data Figure 2e). The absence of a phenotypic effect upon *kraken* overexpression is surprising, given its elevated expression in *D. sechellia* and in evolved *D. simulans* but this could reflect either insufficient overexpression levels (Extended Data Fig. 4c) and/or genotype-by-

genotype interactions that are not recapitulated when manipulating a single locus in a *D. melanogaster* background.

A key question is whether these genes contribute to the natural OA resistance of *D. sechellia*. Implementing CRISPR/Cas9-based genome-engineering in this species[6,30], we generated and validated null mutants in these genes in *D. sechellia* (Extended Data Figure 4b) and tested their resistance to OA in the plate assay. Although we did not detect any significant difference to genetically-matched wild-type flies (Figure 5d), this was anticipated given the polygenic nature of OA resistance in *D. sechellia*[31,37-39], and the modest or absent phenotypes of equivalent mutants in *D. melanogaster* (Figure 5b). We reasoned that more subtle differences might be detected by monitoring *D. sechellia*'s survivability over longer periods of time, when repeated contact with OA was enforced by blending it with the food. Indeed, we observed higher mortality rates of both *kraken* and *Alkbh7* mutants compared to genetically-matched wild-type flies over a period of 6 days (Figure 5e). By contrast, both mutants survived equally well as controls in the absence of OA over this time period (Figure 5e). These data provide the first evidence for the contribution of specific genes to the natural OA resistance of *D. sechellia*.

Discussion

Identifying specific loci underlying phenotypic evolution is challenging due to the polygenicity of the vast majority of adaptive traits. Investigating the evolution of chemical toxicity is an attractive model trait because of the relatively easy phenotypic read-out. To gain insight into the genetic basis of resistance to a natural insecticide, OA, we have intersected the results of selection over vastly different timescales: weeks in cell lines, years in experimentally-evolved laboratory strains, and millennia during speciation.

Unlike insecticides that activate or antagonise broadly-expressed neuronal ion channels to cause organismal death, the mode of action of OA appears to be mechanistically very different, capable of inducing death at the level both of whole organisms and of cultured cells. Beyond insect cell lines (our work and[36]), OA is also toxic to budding yeast[86] and bacteria[87,88], suggesting that this toxin acts via a more fundamental cellular mechanism. Indeed, a genome-wide knock-out screen in yeast for genes required for tolerance of OA highlighted contributions of several core cellular processes, such as vesicle trafficking, mitochondrial function or chromatin regulation[86], reminiscent of our CRISPR screen results. Consistent with this idea, the results of our experimental evolution and CRISPR screening, as well as earlier mapping efforts in *D. sechellia*/*D. simulans* hybrids[31,37-39] revealed multiple genomic regions/genes contributing to the resistance to OA.

Considered separately, the experimental evolution and cell-based CRISPR screens identified many candidates. That there was very limited overlap between the screen results is unsurprising given their very different designs. While both lists are likely to contain multiple genes relevant to understanding OA resistance (and hence mode(s) of action of OA), here we prioritised two common hits, which point to new defence mechanisms underlying OA resistance.

Although both Kraken and Alkbh7 are enzymes, the (partial) necessity, but insufficiency, of *kraken*, and lack of necessity but sufficiency of *Alkbh7* to confer OA resistance implies fundamentally different roles of these proteins. The expression of *kraken* in the gut and renal system supports a hypothesis of this protein as a detoxification enzyme, although future work will be necessary to determine whether

Kraken has a direct role in OA degradation. The mild loss-of-function phenotype of *kraken* in both *D. melanogaster* and *D. sechellia* likely reflects substantial redundancy of defence mechanisms, including other potential detoxification mechanisms, especially given the plethora of enzymes expressed in the fly gut and Malpighian tubules[73,89]. Indeed, extremely few associations between such enzymes and substrates and/or physiological roles have previously been defined. One exception is the Cytochrome P450 gene *Cyp6g1*, which has a key role in Malpighian tubules in conferring resistance to DDT[75]. Our data on Kraken therefore provide one of the first links between a natural toxic chemical and a putative detoxification enzyme.

The mitochondrial Alkbh7 enzyme likely has a distinct role in OA resistance. Insights from mammals are likely informative: mice lacking *Alkbh7* are obese, with higher body fat levels, proposed to result from a defect in metabolism of fatty acids[10]. Molecularly, one study suggests that Alkbh7's metabolic role stems from its regulation of RNA editing to control mitochondrial protein translation[9], although it is unclear whether this is its only function. While the *Alkbh7* loss-of-function phenotype is mild in drosophilids, it is striking that artificially elevated expression in *D. melanogaster* is sufficient to produce *D. sechellia*-like levels of OA resistance. This result implies a potent effect of Alkbh7 on the mitochondrial metabolic network to modulate toxin resistance, which can be explored in future studies. Intriguingly, our implication of Alkbh7 in the niche adaptation of *D. sechellia* has a potential parallel in bats, where this gene displays evidence of positive selection during dietary diversification of different species[90].

Taken together, our work has demonstrated the power of orthogonal selection approaches in the laboratory to understand the evolution of toxin resistance in a historic natural paradigm of this phenomenon[26]. Moreover, the methodology and concepts are now applicable to explore the molecular mechanisms of susceptibility and resistance of insects to the wide-range of artificial toxic chemicals used in the environment, encompassing those used to specifically target pest insects, as well as those that indirectly impair the health of insect species[91]. Such knowledge will have application in the future design and deployment of novel types of insecticides.

Author contributions

M.M. and R.B. conceived the project. M.M. designed and performed the experimental evolution experiment, generated and analysed all transgenic and mutant flies, and performed the comparative gene expression and molecular evolutionary analyses. S.C. generated the base populations and provided technical support for the experimental evolution experiment. M.M., R.B., R.V. and S.E.M. designed the cell screen. S.G. and R.V. tested and defined cell screen assay conditions. S.G. and M.B. performed the CRISPR/Cas9 screen and prepared DNA for sequencing. R.V. analysed the screen data, and M.M., R.B., R.V., S.E.M and N.P. contributed to interpretation of screen results. J.D. and C.R. performed histological experiments, supervised by D.H. M.M. and R.B. wrote the paper with input from other authors. All authors approved the final version of the manuscript.

Acknowledgements

We thank Tadeusz Kawecki and Ana Marija Jakšić for advice on experimental evolution experiments, Tina Lenče and Jean-Yves Roignant for the gift of the *D. melanogaster Alkbh7* sgRNA transgenic line, the Bloomington *Drosophila* Stock

Center (NIH P40OD018537) and the Vienna *Drosophila* Resource Center for *D. melanogaster* stocks, Christian Schlötterer and Neda Barghi for the *D. simulans* lines, Blaise Tissot-Dit-Sanfin for maintenance of *M. citrifolia* plants and Ambra Masuzzo for providing noni fruits. Figure icons were created with Biorender.com. We are grateful to Tadeusz Kawecki, Philippe Reymond, Michael Shahandeh and members of the Benton laboratory for discussions throughout the project and comments on the manuscript. D.H. is funded by an ERC Starting Grant (101117267). N.P. is an Investigator of the Howard Hughes Medical Institute. Research in R.B.'s laboratory was supported by the University of Lausanne, an ERC Advanced Grant (833548) and the Swiss National Science Foundation (310030_219185).

Declaration of interests

The authors declare no competing interests.

Data availability

Source data is provided in the supplementary data files. All other data, new materials and code are available from the corresponding author upon reasonable request.

Methods

Drosophila strains and husbandry

Drosophila stocks were maintained on standard wheat flour–yeast–fruit–juice medium under a 12 h light:12 h dark cycle at 25°C. Wild-type strains, as well as mutant and transgenic lines used in this study are listed in Table S1. Experimental evolution was performed using a subset of published *D. simulans* isogenic strains[56] (Table S2).

Generation of mutant and transgenic flies

Loss-of-function mutants: we generated *D. melanogaster* mutants for *kraken* and *Alkbh7* using in vivo CRISPR/Cas9 mutagenesis[92] by crossing flies expressing Cas9 in the germline (*nanos-Cas9*) to those expressing sgRNAs targeting the desired gene. For *kraken*, we used *P{hsFLP}1,y^{1},w^{1118};P{HD_CFD02326}attP40* as a source of sgRNAs driven by *Act5C-GAL4*; for *Alkbh7*, we used the CG14130 sgRNA line described in[93]. Mutants were validated by genomic PCR and sequencing (Extended Data Figure 4b).

D. sechellia mutant flies were generated by WellGenetics Inc. using modified methods of[94]. For *kraken* (GM16772/LOC6617307) the upstream sequence CTGTAATGCTGGACGCGGAT[TGG] (PAM in square brackets) and the downstream sequence CCTGGTCACCCCGGATCGAG[TGG] were cloned separately into a U6 promoter plasmid. The 3xP3-RFP cassette, which contains a floxed 3xP3-RFP flanked by a 5' homology arm (999 bp: from -1043 nt to -45 nt relative to the *kraken* ATG) and a 3' homology arm (1063 bp: +1388 nt to +2450 nt) were cloned into *pUC57-Kan* as a donor template for repair. For *Alkbh7* (GM25300/LOC6605344) the upstream sequences ACTAATAAGAGTCGACCGGG[TGG], AACAGTCAAACAGCTGTTCC[TGG] and downstream sequences GTCCCAGTTTCAGGGTACGC[TGG], ATGGAAATGCGACGTGTCCT[GGG] were cloned separately into a U6 promoter plasmid. The 3xP3-RFP cassette, which contains a floxed 3xP3-RFP, 5' homology arm (1037 bp: -1121 nt to -85 nt relative to the *Alkbh7* ATG) and 3' homology arm (964 bp: +940 nt to +1903 nt) were cloned into *pUC57-Kan*. DNA plasmids encoding the sgRNAs and the corresponding homology-directed repair template were microinjected into embryos of *Dsec.07 pBAC(nos:Cas9,3P3-YFP)*[30]. In both cases, a single F1 line expressing the 3xP3-RFP selection marker was obtained, which was validated by genomic PCR and sequencing (Extended Data Figure 4b).

D. melanogaster and *D. sechellia* mutant flies were backcrossed for six generations to Canton-S and *Dsec*07, respectively.

Transgenic flies: to make *kraken-Gal4*, we amplified putative 5' regulatory regions for *D. melanogaster kraken* (encompassing the entire intergenic region between the last exon of the upstream gene and the first exon of *kraken*) from *w^{1118}* genomic DNA (all oligonucleotides are listed in Table S6), and subcloned the PCR product into *pGal4-attB*[95]. To make the *UAS-Dmel\kraken* and *UAS-Dsec\kraken* transgenes, we amplified the regions from the start to the stop codons from *w^{1118}* and *Dsec07* genomic DNA and subcloned them into a *pUASTattB* vector[96]. Transgenic flies were obtained by phiC31 integrase-mediated transgenesis into attP2 performed by BestGene Inc. and verified by sequencing.

545 **Toxin resistance assays**
546
547 *Tube assay*: resistance to OA for the experimental evolution setup was scored
548 adapting a previous protocol[37,52]. In brief, flies were lightly anesthetised with CO_2 and
549 ~50 animals were placed in standard cultured tubes (25 mm diameter × 95 mm height,
550 Milian SA) and provided with 1 ml of a 3% glucose solution on a tissue paper (Kimtech,
551 7552). 3 μl of pure OA (Sigma-Aldrich, CAS 124-07-2, C2875) were directly pipetted
552 onto the tissue while the flies remained anesthetised. Direct contact with the acid was
553 scrupulously avoided until all flies recovered. The fraction of alive flies was scored
554 after 24 h. For the experimental evolution experiment, the surviving flies were collected
555 (see below).
556
557 *Plate assay*: to measure OA resistance in the genetically-manipulated *D.*
558 *melanogaster* and *D. sechellia* strains, we used a plate assay[31]. Compared to the
559 experiments in tubes, this assay offers greater temporal resolution and higher dynamic
560 range (in contrast to the experimental evolution experiment, we did not strive to
561 maintain a certain fraction of surviving flies), as well as greater statistical power from
562 tests of fewer animals. Ten flies were lightly anesthetised with CO_2 and placed in a
563 Petri dish (60 mm diameter × 15 mm height) in which 2 μl of pure OA (or 30 μl for *D.*
564 *sechellia*) were spotted in the centre of the underside of the lid. The flies were allowed
565 to recover (typically <5 min), and immobile/dead were scored every 5 min for 1 h. The
566 survival curves were calculated and plotted using the R package survminer
567 (10.32614/CRAN.package.survminer). The statistical analysis was conducted fitting a
568 mixed effect Cox regression model as implemented in the R package coxme[97].
569 Genotype was included as a fixed effect, and replicate dishes were modelled as a
570 random effect.
571
572 *Noni resistance assay*: noni fruit was collected from *M. citrifolia* plants grown in the
573 University of Lausanne greenhouses. Five ripe (pale yellow) fruits were homogenised
574 together, and the pulp puree frozen at -20°C. The noni assay was performed with the
575 same setup as for tube assay described above, using 2 g of thawed noni puree placed
576 on the bottom of the tube.
577
578 *Long-term assay*: To assess the long-term effects of OA (i.e., beyond 24 h), two-day-
579 old mated female flies were individually housed in vials containing food supplemented
580 with OA. OA was thoroughly mixed into Formula 4-24® Instant Drosophila Medium,
581 Blue (Carolina Biological Supply) to a final concentration of 0.9%. Fresh OA-containing
582 food was provided every two days, and mortality was monitored daily until all flies in
583 the experimental group died. Statistical analysis was run as for the plate assay.
584
585 *Fecundity* and *developmental viability*: 10 mated adult female flies (4-7 days old) were
586 allowed to lay eggs for 24 h on agar plates (2% agar, 2.5% sucrose in a 1:3 grape
587 juice:water mix). The total number of eggs on each plate was counted. From each
588 plate, 100 eggs were collected and transferred to standard fly food. Emerging adults
589 were counted and removed daily until no further flies emerged.
590
591 *Lifespan*: Two-day-old mated female flies were individually housed in vials containing
592 standard fly food and monitored daily until death. Fresh food was provided every 2–3
593 days.
594

Experimental evolution

Selection procedure: to generate the base populations for experimental evolution, we used 93 isogenic lines (Table S2) of the *D. simulans* panel established by[56]. We first crossed each of the lines together following a round-robin design to ensure equal contribution of each line to the starting genetic pool; 5 males and 5 females from each cross were then combined in ten independent replicates (Figure 1d). These ten populations were maintained in bottles (57 mm diameter ×103 mm height, Milian SA) of standard fly food at high population size (~1000 individuals per population) to minimise the effect of genetic drift. Each population was subdivided into smaller groups of ~250 individuals per bottle (i.e., four bottles per population) to reduce competition. At each generation, individuals from the four bottles were mixed to prevent bottle-specific genetic drift. The experimental evolution experiment was initiated 10 generations after the creation of these populations.

From each population, 500 individuals (2-7 day old) were randomly collected and further split into 10 groups of 50 mixed-sex flies (i.e., ~5000 flies total). Each group of 50 flies was then placed under selection with 3 μl of OA for 24 h (as described above); the survivors were collected and allowed to breed in food vials for 6-8 days, transferring them to fresh vials every 2-3 days. In the next generation, 500 flies per population were collected only from the last seeded vial, to maximise the probability of these flies being parented by the OA-resistant males and females and not offspring generated by females that mated before the selection procedure. These flies were then subjected to the same selection procedure. This process was repeated for 25 generations, and survivability was scored at each generation.

A second round of experimental evolution was performed by increasing the amount of OA to 9 μl between generations 31 and 50; this amount was determined empirically during generation 26-31 to re-establish the survivability close to the initial level (Figure 1g). All successive generations were maintained following the same procedure.

As a control, the ten original populations were maintained in parallel in the same conditions as previously described (i.e. ~1000 individuals per population subdivided into smaller groups of ~250 individuals per bottle).

DNA sequencing and analysis: we pooled ~200 individual flies per population per time point (generations 0, 25 and 50). Genomic DNA (and total RNA for qPCR, as described below) extraction was performed using the Quick DNA/RNA Miniprep plus kit (Zymo Research, D7003) following the manufacturer's instructions. Samples were stored at -80°C.

We performed paired-end whole genome sequencing on an Illumina NovaSeq 6000 following Nextera DNA flex preparation, at an average coverage of >200×. Read quality was assessed with Fastqc[98] and adapters were trimmed with Trimmomatic[99] providing the Nextera adapters sequences. We aligned the trimmed reads to the *D. simulans* reference genome (NCBI GCF_016746395.2) using bwa-mem[100] and the results were sorted and compressed using samtools[101]. Duplicate reads were then removed with Picard MarkDuplicates (https://broadinstitute.github.io/picard/). Unplaced scaffolds and alignments with mapping quality <20 were removed with samtools. We then combined all the aligned sequences into a samtools pileup file and split this into individual chromosomes. Regions with zero coverage were removed with custom scripts. Sync files for each chromosome were generated with the

644 mpileup2sync.pl script implemented in Popoolation2[102] and used as input files for all
645 the Evolve-and-Resequence analyses.
646 To explore the genetic variation of our samples we generated vcf files with
647 vcftools 0.1.17[103] and binary files with plink1.9[104,105]. We calculated the identity-by-
648 state (IBS) matrix as implemented in plink1.9 (function -distance 'ibs').
649
650 *Evolve-and-Resequence analysis*: allele frequency changes between generation 0
651 and generations 25 or 50 were investigated via Popoolation2 and Bait-ER[63].
652 We performed a Cochran–Mantel–Haenszel (CMH) test via the cmh-test.pl
653 script implemented in Popoolation2 comparing base and evolved populations with
654 parameters --min-count 12 and --min-coverage 50. We estimated a plausible
655 significance threshold approximating the top 0.000005% of the simulated p-values
656 based on the effect of genetic drift alone by running genome-wide forward simulations
657 under neutrality in mimicree2[106]. We used the same genomic information from the
658 original 93 lines as input and the average real survivability ratios observed throughout
659 the selection experiment as –selection-regime parameter and 500 as –population-size
660 parameter (Extended Data Figure 1b).
661 For Bait-ER, the input sync files were filtered to match --min-count 12 and --
662 min-coverage 50 and scaled to a uniform coverage of 100 as in[61] to avoid numerical
663 problems due to high coverage. We ran the algorithm on each chromosome
664 separately. Bait-ER models the effect of genetic drift directly using estimates of the
665 effective population size (N_e). We calculated N_e estimates for each chromosome in
666 each replicate population with the R package poolSeq[107] and used the median value
667 of each chromosome as "Population_size" parameter (Table S3). Other parameters
668 included "Sampling correction" = 0 (i.e., binomial) and "Prior parameters" =0.001
669 (shape), 0.001 (rate). Data points were considered significant when the absolute
670 natural logarithm of the Bayes Factor exceeded the conservative threshold of
671 log(0.99/0.01) as in[63].
672 The linkage disequilibrium decay in the 10 evolved populations was analysed
673 across the whole genome using PopLDdecay[108]. Linkage disequilibrium maps of focal
674 regions were generated with LDBlockShow[109] using the Solid Spine method based on
675 D' statistics originally introduced in Haploview[110].
676 For the genome-wide screen, all the genes within selection blocks were
677 considered potential candidates. Based on the results on linkage disequilibrium decay,
678 a block was arbitrarily defined as a genomic interval where at least two significant
679 SNPs were found at a distance less than 50 kb. Consecutive SNPs respecting this
680 definition were grouped into larger blocks and visualised as circus plots using the R
681 package circlize[111]. The genes overlapping such blocks were extracted using
682 BEDtools intersect[112]. Orthologous genes in *D. melanogaster* were retrieved using a
683 combination of the OMA browser[113], DAVID[114,115], BLAST[116], NCBI[117] and FlyBase[118],
684 and subsequent manual curation.
685
686 **Genome-wide CRISPR/Cas9 screen in *D. melanogaster* S2R+ cells**
687
688 *Library design, construction, and transfection*: design, synthesis and cloning of the
689 genome-wide sgRNA library are previously described[119]. To determine sub-lethal
690 doses of OA appropriate for screening, the effect of a gradient of OA supplementation
691 into media on attP+ Cas9+ cells was determined using Cell Titer Glo (Promega)
692 following the manufacturer's protocol; luminescence measurements were made using
693 a Spectramax Paradigm (Molecular Devices).

694 For the screen, attP+ Cas9+ *D. melanogaster* S2R+ cells (RRID:CVCL_UD30)
695 were transfected using Effectene and a 1:1 molar ratio of sgRNA library and pIntAC,
696 as described[119]. Briefly, transfected cells were distributed in 5 ml aliquots across ten
697 100 mm culture dishes and incubated overnight, then 5 ml of fresh media was added
698 to each plate. After 4 days, cells were subjected to puromycin selection for 12 days,
699 subculturing every 4 days, to establish a stable sgRNA-expressing cell library. For the
700 selection assay, 2.5×10^7 cells were plated onto four 15 cm diameter cell culture
701 dishes to ensure sufficient sgRNA coverage, with each sgRNA represented
702 approximately 1000 times[119].
703
704 *OA selection assay*: library-containing cells were left untreated or treated with either
705 0.075 µg/µl or 0.05 µg/µl OA (Sigma-Aldrich, CAS 124-07-2, C2875). Each replicate
706 consisted of four 15 cm diameter cell culture dishes seeded with cells, achieving
707 ~1000× coverage. Two replicate sets were included for each of the two OA treatment
708 conditions and one replicate for the untreated control. During the selection period, cells
709 were split 1-2 times per week to avoid overgrowth. An aliquot of cells from each
710 replicate was collected and stored for genomic prep following each split and at the end
711 of the assay. The final samples were collected after 26 days of OA or control treatment.
712
713 *Genomic DNA extraction and sequencing*: genomic DNA was extracted from samples
714 collected during the screen and at the final endpoint using a Zymo Quick-gDNA
715 Miniprep kit (Zymo Research, D3025). DNA fragments containing the sgRNA
716 sequences were amplified by PCR using what we calculated to be at least 1000
717 genomes per sgRNA as templates for each sample. PCR fragments were in-line
718 barcoded using a previously described approach[45,119,120]. Next-generation sequencing
719 (Illumina NextSeq) was performed at the Biopolymers Facility at Harvard Medical
720 School (RRID:SCR_007175). FASTQ sequence files were demultiplexed using in-line
721 barcodes as sequence tags as described previously[45,119]. All further processing was
722 done using MAGeCK (v.0.5.4 and 0.5.9.4)[121] to generate readcount files and perform
723 robust rank aggregation (RRA) analysis and Maximum likelihood estimation directly
724 on the generated readcount files using default parameters. Plasmid readcounts were
725 used to normalise RRA values.
726 After removing essential genes, we considered as candidates the top 5% genes
727 overlapping across the two conditions that were over-represented or under-
728 represented in the OA-exposed cell populations compared to the control cells (i.e.,
729 representing "susceptibility" and "resistance" genes, respectively). The gene-set
730 enrichment analysis was performed with PANGEA 1.1 Beta using the *Drosophila* GO
731 subsets (GO slim) and a p-value cut-off of 0.05[122].
732
733 **qPCR analysis**
734
735 Reverse transcription of total RNA (extracted as described above) was performed with
736 the PrimeScript first-strand cDNA synthesis kit (Takara, 6110A) following the
737 manufacturer's instructions. Samples were stored at -80°C.
738 Quantitative PCR experiments were conducted on a QuantStudio 6 Flex Real-
739 Time PCR System using the PowerUp™ SYBR™ Green Master Mix following the
740 manufacturer's instructions. Each sample consisted of at least three biological
741 replicates. We performed three technical replicates per reaction and removed
742 replicates where the standard deviation was equal or greater than 0.5. We determined
743 target specific amplification efficiencies for each gene (Table S4, Table S6). Relative

quantification and statistical analysis were performed with the software qbase+ v3.4[123] using Calibrated Normalised Relative Quantities (CNRQs). Normalisation was performed on the geometric mean of the relative quantities for the three most commonly-used reference targets: *Act5C*, *αTub84B* and *RpL32*[124]. We assessed reference target stability by the geNorm expression stability value (M) and the coefficient of variation of the normalised reference gene relative quantities (CV) (Table S5)[123,125].

Population genetics analysis

Publicly available WGS data for 41 *D. sechellia* and 20 *D. simulans* genomes were downloaded from the Sequence Read Archive (PRJNA215932 and PRJNA395473)[80,126] with fasterq-dump[127]. FASTQ files were processed and aligned as described in the "Experimental evolution" section. Aligned files were phased with samtools phase[101] and the genes of interest were extracted with pilon[128] to find variation among populations. CDS sequences were aligned based on amino acid information using translatorX[129]. Gene sequences were aligned using Clustal Omega[130]. McDonald-Kreitman test, Tajima's D and Fu and Li's statistics were calculated in DNAsp 6[131]. FUBAR analyses (Fast, Unconstrained Bayesian AppRoximation) were performed in HyPhy 2.5[132]. The consensus sequences of *D. simulans* and *D. sechellia* were computed by SnapGene (www.snapgene.com) "export consensus" with a >50% threshold. The sequence alignment and secondary structures were depicted using ESPript 3[133].

Histology

Intestinal tissues and Malpighian tubules were dissected in phosphate-buffered saline (PBS) and fixed at room temperature for 45 min in PBS, 4% paraformaldehyde. Staining of muscle was performed using a mushroom phalloidin AlexFluor647 probe (Life Technologies, A22287, diluted 1:200 in PBS + 0.2% Triton X-100) for 45 min. Samples were mounted in Vectashield with DAPI (Vector Labs, H-1200-10).

A Leica STELLARIS-DIVE confocal microscope was used to generate all confocal images, using a water-immersive 40× objective or an oil-immersive 63× objective. The images were acquired using both HyDs as well as PMTs tailored for the fluorophores of each sample.

Statistics and reproducibility

Data were analysed and plotted in R 4.0.3 (R Core Team 2021).

Figure legends

Figure 1. Experimental evolution of *D. simulans* with increased OA resistance.
a Phylogeny of the drosophilid trio studied in this work. Mya: million years ago.
b Schematic of octanoic acid (OA) resistance tube assay (see Methods) and selection procedure.
c Species differences in resistance of *D. simulans* 04 and *D. sechellia* 07 strains to 3 µl OA.
d Schematic illustrating the generation of the base populations for Evolve-and-Resequence (see Methods).
e Changes in survivability to 3 µl OA of the ten *D. simulans* populations over generations 0-25. Box plots represent the interquartile ranges of the survivability for all the replicates; blue lines shows the generalised additive model (GAM) regression; grey shading represents the 95% confidence interval for the mean responses at each time point.
f As in **e** for the 10 populations combined.
g OA resistance for the combined 10 populations with increasing OA volumes during generation 25-31.
h As in **e**, with 9 µl OA during generations 31-50.
i As in **h** for the 10 populations combined.
j Comparison of OA resistance of unevolved *D. simulans* (i.e., the original 10 outcrossed populations maintained on standard fly food; see Methods), evolved (generation 50 (G50)) *D. simulans*, and *D. sechellia*.
k-n Phenotypic comparison of unevolved and evolved (G50+) *D. simulans* populations 1-3 for **k** noni resistance (G54), **l** fecundity (G54), **m** developmental viability (G57), and **n** lifespan (G57). For **k-m**, significance was assessed using the unpaired t-test correcting for multiple testing. NS $p>0.05$, ***$p<0.001$. For **l-n**, tests were run on standard media. No significant differences between unevolved and evolved flies were detected using the Cox regression model. See Methods for details and Data S1 for raw data.

Figure 2. Dynamic directional selection of allele frequencies during experimental evolution.
a Schematic of the analysis.
b Identity-by-state analysis. The heatmap depicts the proportion of shared alleles between populations, with values ranging from 0.8 (blue, lowest shared proportion) to 1 (red, identical alleles).
c,d Manhattan plot showing the probability of allele frequency changes between **c** generations 0 and 25, and **d** generations 0 and 50. Top panel shows the results of the CMH test; orange dots indicate significant variants at a threshold of $-\log_{10}(p) \geq 40$, a conservative approximation of the top 0.000005% simulated SNPs under neutrality (Extended Data Figure 1b). Bottom panel shows the absolute log-Bayes factors estimated by Bait-ER; orange dots indicate significant variants at the conservative threshold log(0.99/0.01). Raw data in Data S2.
e Circular plots illustrating overlap of predicted genomic regions under directional selection between the two statistical approaches at generation 25 (left) and generation 50 (middle), or between generation 25 and 50 (right). Raw data in Data S3.

834
835
836 **Figure 3. A genome-wide pooled CRISPR screen for OA susceptibility and**
837 **resistance genes in *D. melanogaster* cultured cells.**
838 **a** Left: images of S2R+ cells cultured in the absence or presence of OA. Scale bar, 50
839 μm. Right: dose-dependent effects of OA on S2R+ cell viability, through measurement
840 of ATP levels with a CellTiter-Glo assay (see Methods).
841 **b** Schematic of CRISPR screen design.
842 **c** Top 5% "susceptibility" genes (see Methods and Data S4). Box plots represent the
843 interquartile ranges for all the replicates of the Z-scores obtained through maximum
844 likelihood estimation to test for selection. The inset (bottom-right) displays the gene-
845 ontology analysis of these genes.
846 **d** As in **c**, for the top 5% "resistance" genes.
847 **e** Venn diagram showing the overlap between all the genes identified in the
848 experimental evolution at G25 and G50, and genes associated with susceptibility and
849 resistance in the CRISPR screen.
850 **f** Subportion of the top panel (CMH test) of the Manhattan plot from Figure 2c for
851 chromosome arms 2L and 3L, with the red arrows highlighting the regions magnified
852 on the right showing the significant variants and linkage disequilibrium blocks
853 (triangles) for *kraken* and *Alkbh7* (CG14130).
854
855
856 **Figure 4. Sequence and expression analyses of *kraken* and *Alkbh7*.**
857 **a** AlphaFold protein models for *D. melanogaster* Kraken (AF-O18391-F1-model_v4)
858 and Alkbh7 (AF-Q9VTP1-F1-model_v4)[134,135], highlighting a conserved putative
859 catalytic serine in the active site in Kraken, and the mitochondrial targeting sequence
860 in Alkbh7 (predicted by MitoFates[77]). See also Extended Data Figure 2a,b.
861 **b** Tissue-specific expression of *D. melanogaster kraken* and *Alkbh7* from bulk RNA-
862 sequencing data from the Fly Atlas 2.0[136].
863 **c** tSNE plots illustrating *D. melanogaster kraken* expression in single-cell
864 transcriptomes of the Malpighian tubules and gut from the Fly Cell Atlas (Stringent 10x
865 datasets)[72].
866 **d** Left: whole gut of an adult female expressing GFP (green, endogenous fluorescence
867 signal) under the control of a *kraken* promoter Gal4 driver (genotype: *kraken-*
868 *Gal4*/*UAS-FB1.1*, aged 4-5 days); muscles are labelled by phalloidin (red) to reveal
869 tissue morphology. Imaging was performed with a 40× objective; scale bars, 200 μm.
870 Right: higher magnification images of the regions indicated with dashed boxes.
871 Imaging was performed with a 63× objective; scale bars, 100 μm.
872 **e** Expression levels of the *kraken* and *Alkbh7* in the indicate strains of adult female *D.*
873 *sechellia* and *D. simulans* measured by qPCR (Data S6). Expression is represented
874 as calibrated and normalised relative quantities (CNRQ). Significance was assessed
875 using the unpaired t-test correcting for multiple testing and represented using letter
876 codes.
877 **f** Expression levels of *kraken* and *Alkbh7* in the indicated strains and species at adult
878 or larval stages as indicated from published RNA-seq data[83-85]. Significance was
879 assessed using the unpaired t-test correcting for multiple testing and represented
880 using letter codes.
881 **g** Expression levels of the candidate genes at generation 0, 25 and 50 of the
882 experimentally-evolved *D. simulans* populations measured by qPCR. Significance was

assessed using the paired t-test correcting for multiple testing and represented using letter codes.

Figure 5. Functional validation of the contribution of *kraken* and *Alkbh7* to OA resistance.

a-c Survival curves of adult flies of the indicated genotypes tested in the plate assay with 2 µl OA for *kraken* and *Alkbh7* RNAi (**a**), mutants (**b**) and overexpression (**c**). Each replicate consisted of 10 2-7 day-old female flies. N replicates ≥6, n flies ≥60. Shading indicates 95% confidence intervals. Genotypes: **a** *act-Gal4/+;UAS-kraken^{RNAi}/+, UAS-kraken^{RNAi}/+, act-Gal4/+, act-Gal4/+;UAS-Alkbh7^{RNAi}/+, UAS-Alkbh7^{RNAi}/+, act-Gal4/+* **b** *kraken^{1}*, *Alkbh7^{1}, Canton-S* (wild-type genetic background control for both mutants), **c** *act-Gal4/+;UAS-kraken$^{o/e}$/+*, *UAS-kraken$^{o/e}$/+*, *act-Gal4/+*, *act-Gal4/+;UAS-Alkbh7$^{o/e}$/+*, *UAS-Alkbh7$^{o/e}$/+*, *act-Gal4/+*.
Significance is based on the resulting mixed effect Cox regression model. NS $p>0.05$, $^{*}p<0.05$, $^{**}p<0.01$, $^{***}p<0.001$. The same data for *D. sechellia* are shown in both plots in **c**.

d Survival curves of wild-type (*Dsec07*), and *kraken* (*Dsec\kraken^{RFP}*) and *Alkbh7* (*Dsec\Alkbh7^{RFP}*) mutant *D. sechellia* in the plate assay with 30 µl OA. Statistics as in **a**.

e Long-term survivability of the same genotypes as in **d** (n = 60 per condition) in control conditions (no OA) (left) and food supplemented with 0.9% OA (right). Statistics as in **a**.

Raw data in Data S7.

909 **Supplementary Information**
910
911 **Supplementary Tables**
912
913 **Table S1. *Drosophila* strains.**
914

Genotype	Reference/Source
D. sechellia 07 (*Dsec07*)	*Drosophila* Species Stock Center (DSSC):14021-0248.07
D. simulans 04 (*Dsim04*)	DSSC:14021-0251.004
D. melanogaster Canton-S *(DmelCS)*	
y[1] w[*]; P{Act5C-GAL4-w}E1/CyO (act-Gal4)	RRID:BDSC_25374
y[1] w[1118] (used as background for act-Gal4)	RRID:BDSC_6598
y[1] w[67c23] (background of Alkbh7$^{o/e}$)	RRID:BDSC_6599
y[1] w[67c23]; P{y[+mDint2] w[+mC]=EPgy2}EY05855 (*Alkbh7$^{o/e}$*)	RRID:BDSC_16679
y,w[1118];P{attP,y[+],w[3']} (background of UAS-kraken^{RNAi} and UAS-Alkbh7^{RNAi})	VDRC:v60100
P{KK100356}VIE-260B (UAS-kraken^{RNAi})	VDRC:v105604
P{KK102727}VIE-260B (UAS-Alkbh7^{RNAi})	VDRC:v109669
y[1] M{w[+mC]=nanos-Cas9.P}ZH-2A w[*] (nanos-cas9)	RRID:BDSC_54591
P{hsFLP}1, y1 w1118; P{HD_CFD02326}attP40/CyO-GFP (sgRNA line for Dmel\kraken CRISPR mutant)	VDRC:v342579
CG14130 cris (gRNA 1+2)/CyO *(sgRNA line for Dmel\Alkbh7 CRISPR mutant)*	[93]
y[1] w[67c23];;UAS-kraken (kraken$^{o/e}$)	This study
y[1] w[67c23];;UAS-Dsec\kraken (Dsec\kraken$^{o/e}$)	This study
y[1] w[67c23]; P{y[+t7.7]=CaryP}attP2 (background of kraken$^{o/e}$)	This study
Dmel\kraken-Gal4	This study
UAS-FB1.1 (GFP reporter)	[137]
Dmel\kraken^{1}	This study
Dmel\Alkbh7^{1}	This study
Dsec\kraken^{RFP}	This study
Dsec\Alkbh7^{RFP}	This study

915
916
917 **Table S2. *D. simulans* lines used to generate the base populations.**
918 *- see separate Excel file.*
919
920

921 **Table S3. Effective population size estimates from the experimentally-evolved**
922 **populations.**
923

	2L	2R	3L	3R	4	X
pop1	80.65681	101.5796	107.7145	100.5748	369.3642	107.3616
pop2	135.7243	56.29977	76.40593	91.41878	333.638	53.66758
pop3	74.08474	119.4491	95.20849	128.4216	193.5982	99.40552
pop4	83.05221	89.62312	112.3929	116.7979	171.2875	112.1581
pop5	123.0386	106.8179	107.448	81.31624	364.4287	83.55315
pop6	98.8018	77.03804	60.71772	147.1788	124.9405	75.48775
pop7	119.4038	101.4515	96.22689	114.8849	431.7489	114.6792
pop8	136.6513	176.4308	70.8519	130.9822	639.7962	78.70108
pop9	95.58733	101.7351	122.8352	78.80166	109.3757	69.50865
pop10	141.862	101.6789	107.512	83.0109	376.8733	71.34583
mean	108.8863	103.2104	95.73134	107.3388	311.5051	86.58684
median	109.1028	101.6293	101.8374	107.7299	349.0333	81.12711

924
925 **Table S4. qPCR primer efficiency.**
926

Gene	**Efficiency**	**SE**
kraken	1.998	0.041
Alkbh7	1.977	0.016
Act5C	1.976	0.022
αTub84B	2.027	0.023
RpL32	2.053	0.031

927
928 **Table S5. geNorm analysis of reference target stability.**
929 *- see separate Excel file.*
930
931 **Table S6. Oligonucleotides.**
932 *- see separate Excel file.*
933
934

Extended Data Figure legends

Extended Data Figure 1. Linkage disequilibrium decay, and comparison of simulated and observed p-value distributions in the experimentally evolved populations.

a Linkage disequilibrium decay at generations 25 (left) and 50 (right). Black curves represent the trajectory of r2 values (y-axis left) as genomic distances increase (x-axis). Red curves represent the same but for D' values (y-axis right).

b Violin plots showing the distribution of p-values from the CMH test, comparing allele frequency changes between generation 25 and 0 (left), and generation 50 and 0 (right). Blue violins represent p-values derived from forward evolution simulations, while orange violins correspond to p-values from experimentally evolved populations. Red dashed lines mark the top 0.000005% of simulated p-values, serving as the significance threshold for the experimental data.

Extended Data Figure 2. Sequence and molecular evolutionary analyses of *kraken* and *Alkbh7*.

a,b Protein sequence alignment and secondary structure of Kraken (**a**) and Alkbh7 (**b**) from the reference sequence of *D. melanogaster* (NP_477265.1 and NP_648511.2) and the consensus sequences of *D. simulans* and *D. sechellia*. Sequence data from PRJNA215932 and PRJNA395473 (see Methods). Black triangles indicate the residues where signs of selection were detected by the FUBAR analysis (see **d**).

c Results from the McDonald-Kreitman test. NI = neutrality index; alpha = proportion of substitutions driven by positive selection; p-value = p-value resulting from a two-tailed Fisher's exact test.

d FUBAR analysis to detect pervasive selection as implemented in HyPhy. position = codon position; alpha = mean posterior synonymous substitution rate; beta = mean posterior nonsynonymous substitution rate; posterior prob = posterior probability of positive selection; codons = alternative codons, Freq Dsec = frequency of the alternative codons in *D. sechellia* in the same order as codons; Freq Dsim = frequency of the alternative codons in *D. simulans* in the same order as codons.

e Survival curves of *D. melanogaster* overexpressing *Dmel\kraken* and *Dsec\kraken* exposed to 2 µl OA. Each replicate consisted of 10 2-7 day-old female flies. Genotypes: *act-Gal4/+;UAS-Dmel\kraken/+*, *act-Gal4/+;UAS-Dsec\kraken/+*. N replicates ≥6, n flies ≥60. Shading indicates 95% confidence intervals. Significance is based on the resulting mixed effect Cox regression model. NS p>0.05. Raw data in Data S5.

Extended Data Figure 3. Additional expression analyses of *kraken* and *Alkbh7*.

a Bar plots showing the expression levels of *kraken* and *Alkbh7* across various adult *Drosophila* tissues and species. Data from[138].

b Bar plots showing the expression levels of *kraken* and *Alkbh7* across various larval *Drosophila* tissues and species. Data from[83] ("medium diet" condition).

c Bar plots showing the effect of exposure to 9 µl OA (in the tube assay) on *kraken* and *Alkbh7* expression in *D. sechellia* and three evolved (G50) *D. simulans* populations.

985 **Extended Data Figure 4. Validation of RNAi and overexpression lines and**
986 **generation of mutants for *kraken* and *Alkbh7*.**
987 **a** Quantitative RT-PCR analysis of RNAi-mediated knockdown of *kraken* and *Alkbh7*
988 in *D. melanogaster*. Genotypes as in Figure 5a.
989 **b** Schematics of the CRISPR/Cas9 design to generate *D. melanogaster* and *D.*
990 *sechellia* mutants for *kraken* and *Alkbh7*. Validation was performed by PCR and
991 Sanger sequencing. The *Dsec\kraken*^{RFP} mutant comprised a 1431 bp deletion (-44 to
992 +1,387 nt relative to the ATG), replaced with the 3xP3-RFP cassette. The
993 *Dsec\Alkbh7*^{RFP} mutant was a 1023 bp deletion (-84 nt to +939 nt relative to the ATG),
994 replaced with the 3xP3-RFP cassette.
995 **c** Quantitative RT-PCR analysis of Gal4/UAS-mediated overexpression of *D.*
996 *melanogaster kraken* and *Alkbh7*. Genotypes as in Figure 5c.
997
998
999 **Extended Data Figure 5. Control survival curves in the absence of OA.**
1000 Survival curves of adult flies of the indicated genotypes tested in the plate assay
1001 without OA. Each replicate consisted of 10 2-7 day-old female flies. N replicates ≥5, n
1002 flies ≥50. All genotypes display full viability, rendering datapoints invisible beneath the
1003 topmost displayed genotype (*UAS-kraken*^{RNAi}). Genotypes as in Figure 5a-b.
1004
1005
1006 **Data S1.** Source data for Figure 1
1007 **Data S2.** Experimental evolution data
1008 **Data S3.** Source data for Figure 2
1009 **Data S4.** Source data for Figure 3
1010 **Data S5.** Source data for Extended Data Figure 2
1011 **Data S6.** Source data for Figure 4
1012 **Data S7.** Source data for Figure 5
1013

1014 **References**
1015

1016 1 Erb, M. & Reymond, P. Molecular Interactions Between Plants and Insect
1017 Herbivores. *Annu Rev Plant Biol* **70**, 527-557, doi:10.1146/annurev-arplant-
1018 050718-095910 (2019).
1019 2 Umetsu, N. & Shirai, Y. Development of novel pesticides in the 21st century. *J*
1020 *Pestic Sci* **45**, 54-74, doi:10.1584/jpestics.D20-201 (2020).
1021 3 Raisch, T. & Raunser, S. The modes of action of ion-channel-targeting
1022 neurotoxic insecticides: lessons from structural biology. *Nat. Struct. Mol. Biol.*
1023 **30**, 1411-1427, doi:10.1038/s41594-023-01113-5 (2023).
1024 4 Sparks, T. C. *et al.* Insecticides, biologics and nematicides: Updates to IRAC's
1025 mode of action classification - a tool for resistance management. *Pesticide*
1026 *biochemistry and physiology* **167**, 104587, doi:10.1016/j.pestbp.2020.104587
1027 (2020).
1028 5 Jones, C. D. The genetics of adaptation in *Drosophila sechellia*. *Genetica*
1029 **123**, 137-145 (2005).
1030 6 Auer, T. O., Shahandeh, M. P. & Benton, R. *Drosophila sechellia*: A Genetic
1031 Model for Behavioral Evolution and Neuroecology. *Annu Rev Genet* **55**, 527-
1032 554, doi:10.1146/annurev-genet-071719-020719 (2021).
1033 7 Kumar, K., Mhetre, A., Ratnaparkhi, G. S. & Kamat, S. S. A Superfamily-wide
1034 Activity Atlas of Serine Hydrolases in *Drosophila melanogaster*. *Biochemistry*
1035 **60**, 1312-1324, doi:10.1021/acs.biochem.1c00171 (2021).
1036 8 Edwin Chan, H. Y., Harris, S. J. & O'Kane, C. J. Identification and
1037 characterization of kraken, a gene encoding a putative hydrolytic enzyme in
1038 *Drosophila melanogaster*. *Gene* **222**, 195-201, doi:10.1016/s0378-
1039 1119(98)00497-1 (1998).
1040 9 Zhang, L. S. *et al.* ALKBH7-mediated demethylation regulates mitochondrial
1041 polycistronic RNA processing. *Nat Cell Biol* **23**, 684-691, doi:10.1038/s41556-
1042 021-00709-7 (2021).
1043 10 Solberg, A. *et al.* Deletion of mouse Alkbh7 leads to obesity. *J Mol Cell Biol* **5**,
1044 194-203, doi:10.1093/jmcb/mjt012 (2013).
1045 11 Jones, A. C., Felton, G. W. & Tumlinson, J. H. The dual function of elicitors
1046 and effectors from insects: reviewing the 'arms race' against plant defenses.
1047 *Plant Mol Biol* **109**, 427-445, doi:10.1007/s11103-021-01203-2 (2022).
1048 12 Smith, J. J., Herzig, V., King, G. F. & Alewood, P. F. The insecticidal potential
1049 of venom peptides. *Cell Mol Life Sci* **70**, 3665-3693, doi:10.1007/s00018-013-
1050 1315-3 (2013).
1051 13 Busvine, J. R. Mechanism of resistance to insecticide in houseflies. *Nature*
1052 **168**, 193-195, doi:10.1038/168193a0 (1951).
1053 14 Dong, K. Insect sodium channels and insecticide resistance. *Invert Neurosci*
1054 **7**, 17-30, doi:10.1007/s10158-006-0036-9 (2007).
1055 15 Matsuda, K., Ihara, M. & Sattelle, D. B. Neonicotinoid Insecticides: Molecular
1056 Targets, Resistance, and Toxicity. *Annu Rev Pharmacol Toxicol* **60**, 241-255,
1057 doi:10.1146/annurev-pharmtox-010818-021747 (2020).
1058 16 Soderlund, D. M. & Bloomquist, J. R. Neurotoxic actions of pyrethroid
1059 insecticides. *Annu. Rev. Entomol.* **34**, 77-96,
1060 doi:10.1146/annurev.en.34.010189.000453 (1989).
1061 17 Di Prisco, G. *et al.* Neonicotinoid clothianidin adversely affects insect
1062 immunity and promotes replication of a viral pathogen in honey bees.

Proceedings of the National Academy of Sciences of the United States of America **110**, 18466-18471, doi:10.1073/pnas.1314923110 (2013).

18 Clarkson, C. S., Temple, H. J. & Miles, A. The genomics of insecticide resistance: insights from recent studies in African malaria vectors. *Curr Opin Insect Sci* **27**, 111-115, doi:10.1016/j.cois.2018.05.017 (2018).

19 Lucas, E. R. *et al.* Genome-wide association studies reveal novel loci associated with pyrethroid and organophosphate resistance in *Anopheles gambiae* and *Anopheles coluzzii*. *Nat Commun* **14**, 4946, doi:10.1038/s41467-023-40693-0 (2023).

20 Catteruccia, F. Malaria-carrying mosquitoes get a leg up on insecticides. *Nature* **577**, 319-320, doi:10.1038/d41586-019-03728-5 (2020).

21 Balabanidou, V. *et al.* Mosquitoes cloak their legs to resist insecticides. *Proc Biol Sci* **286**, 20191091, doi:10.1098/rspb.2019.1091 (2019).

22 Balabanidou, V., Grigoraki, L. & Vontas, J. Insect cuticle: a critical determinant of insecticide resistance. *Curr Opin Insect Sci* **27**, 68-74, doi:10.1016/j.cois.2018.03.001 (2018).

23 Weetman, D., Djogbenou, L. S. & Lucas, E. Copy number variation (CNV) and insecticide resistance in mosquitoes: evolving knowledge or an evolving problem? *Curr Opin Insect Sci* **27**, 82-88, doi:10.1016/j.cois.2018.04.005 (2018).

24 Zimmer, C. T. *et al.* Neofunctionalization of Duplicated P450 Genes Drives the Evolution of Insecticide Resistance in the Brown Planthopper. *Curr Biol* **28**, 268-274 e265, doi:10.1016/j.cub.2017.11.060 (2018).

25 Gloss, A. D. *et al.* Evolution in an ancient detoxification pathway is coupled with a transition to herbivory in the drosophilidae. *Mol. Biol. Evol.* **31**, 2441-2456, doi:10.1093/molbev/msu201 (2014).

26 R'Kha, S., Capy, P. & David, J. R. Host-plant specialization in the *Drosophila melanogaster* species complex: a physiological, behavioral, and genetical analysis. *Proceedings of the National Academy of Sciences of the United States of America* **88**, 1835-1839, doi:10.1073/pnas.88.5.1835 (1991).

27 Dekker, T., Ibba, I., Siju, K. P., Stensmyr, M. C. & Hansson, B. S. Olfactory shifts parallel superspecialism for toxic fruit in *Drosophila melanogaster* sibling, *D. sechellia*. *Curr Biol* **16**, 101-109 (2006).

28 Alvarez-Ocana, R. *et al.* Odor-regulated oviposition behavior in an ecological specialist. *Nat Commun* **14**, 3041, doi:10.1038/s41467-023-38722-z (2023).

29 Shahandeh, M. P. *et al.* Circadian plasticity evolves via regulatory changes in a neuropeptide gene. *Nature* **635**, 951-959, doi:10.1101/2023.07.05.547553 (2024).

30 Auer, T. O. *et al.* Olfactory receptor and circuit evolution promote host specialization. *Nature* **579**, 402-408, doi:10.1038/s41586-020-2073-7 (2020).

31 Jones, C. D. The genetic basis of *Drosophila sechellia*'s resistance to a host plant toxin. *Genetics* **149**, 1899-1908 (1998).

32 Legal, L., Moulin, B. & Jallon, J. M. The Relation between Structures and Toxicity of Oxygenated Aliphatic Compounds Homologous to the Insecticide Octanoic Acid and the Chemotaxis of Two Species of *Drosophila*. *Pesticide biochemistry and physiology* **65**, 90-101 (1999).

33 Markow, T. A. Host use and host shifts in *Drosophila*. *Curr Opin Insect Sci* **31**, 139-145, doi:10.1016/j.cois.2019.01.006 (2019).

34 Legal, L., Chappe, B. & Jallon, J. M. Molecular basis of *Morinda citrifolia* (L.): Toxicity on *Drosophila*. *J Chem Ecol* **20**, 1931-1943, doi:10.1007/BF02066234 (1994).
35 Kaczmarek, A., Katarzyna Wronska, A. & Irena Bogus, M. Octanoic acid kills Lucilia sericata (Diptera: Calliphoridae) by affecting two major defence systems: cuticular free fatty acids and immunocompetent cells. *J Invertebr Pathol* **206**, 108165, doi:10.1016/j.jip.2024.108165 (2024).
36 Kaczmarek, A., Wronska, A. K., Kazek, M. & Bogus, M. I. Octanoic Acid-An Insecticidal Metabolite of *Conidiobolus coronatus* (Entomopthorales) That Affects Two Majors Antifungal Protection Systems in *Galleria mellonella* (Lepidoptera): Cuticular Lipids and Hemocytes. *International journal of molecular sciences* **23**, doi:10.3390/ijms23095204 (2022).
37 Amlou, M., Pla, E., Moreteau, B. & David, J. R. Genetic analysis by interspecific crosses of the tolerance of *Drosophila sechellia* to major aliphatic acids of its host plant. *Genet Sel Evol* **29**, 511-522 (1997).
38 Amlou, M., Moreteau, B. & David, J. R. Larval tolerance in the *Drosophila melanogaster* species complex toward the two toxic acids of the *D. sechellia* host plant. *Hereditas* **129**, 7-14, doi:10.1111/j.1601-5223.1998.00007.x (1998).
39 Huang, Y. & Erezyilmaz, D. The Genetics of Resistance to *Morinda* Fruit Toxin During the Postembryonic Stages in *Drosophila sechellia*. *G3 (Bethesda)* **5**, 1973-1981, doi:10.1534/g3.114.015073 (2015).
40 Sun, Z., Inagaki, S., Miyoshi, K., Saito, K. & Hayashi, S. Osiris gene family defines the cuticle nanopatterns of *Drosophila*. *Genetics* **227**, doi:10.1093/genetics/iyae065 (2024).
41 Ando, T. *et al.* Nanopore Formation in the Cuticle of an Insect Olfactory Sensillum. *Curr Biol* **29**, 1512-1520 e1516, doi:10.1016/j.cub.2019.03.043 (2019).
42 Andrade Lopez, J. M. *et al.* Genetic basis of octanoic acid resistance in *Drosophila sechellia*: functional analysis of a fine-mapped region. *Mol. Ecol.* **26**, 1148-1160, doi:10.1111/mec.14001 (2017).
43 Lanno, S. M., Shimshak, S. J., Peyser, R. D., Linde, S. C. & Coolon, J. D. Investigating the role of Osiris genes in *Drosophila sechellia* larval resistance to a host plant toxin. *Ecology and evolution* **9**, 1922-1933, doi:10.1002/ece3.4885 (2019).
44 Schlotterer, C., Kofler, R., Versace, E., Tobler, R. & Franssen, S. U. Combining experimental evolution with next-generation sequencing: a powerful tool to study adaptation from standing genetic variation. *Heredity (Edinb)* **116**, 248, doi:10.1038/hdy.2015.85 (2016).
45 Viswanatha, R., Li, Z., Hu, Y. & Perrimon, N. Pooled genome-wide CRISPR screening for basal and context-specific fitness gene essentiality in *Drosophila* cells. *Elife* **7**, doi:10.7554/eLife.36333 (2018).
46 Barghi, N. *et al.* Genetic redundancy fuels polygenic adaptation in *Drosophila*. *PLOS Biol.* **17**, e3000128, doi:10.1371/journal.pbio.3000128 (2019).
47 Turner, T. L. & Miller, P. M. Investigating natural variation in *Drosophila* courtship song by the evolve and resequence approach. *Genetics* **191**, 633-642, doi:10.1534/genetics.112.139337 (2012).
48 Kelly, J. K. & Hughes, K. A. Pervasive Linked Selection and Intermediate-Frequency Alleles Are Implicated in an Evolve-and-Resequencing Experiment

of *Drosophila simulans*. *Genetics* **211**, 943-961, doi:10.1534/genetics.118.301824 (2019).

49 Shahrestani, P. *et al.* The molecular architecture of *Drosophila melanogaster* defense against *Beauveria bassiana* explored through evolve and resequence and quantitative trait locus mapping. *G3 (Bethesda)* **11**, doi:10.1093/g3journal/jkab324 (2021).

50 Zoh, M. G. *et al.* Experimental evolution supports the potential of neonicotinoid-pyrethroid combination for managing insecticide resistance in malaria vectors. *Sci Rep* **11**, 19501, doi:10.1038/s41598-021-99061-x (2021).

51 Sadia, C. G. *et al.* The impact of agrochemical pollutant mixtures on the selection of insecticide resistance in the malaria vector *Anopheles gambiae*: insights from experimental evolution and transcriptomics. *Malar J* **23**, 69, doi:10.1186/s12936-023-04791-0 (2024).

52 Colson, I. *Drosophila simulans*' response to laboratory selection for tolerance to a toxic food source used by its sister species *D. sechellia*. *Evolutionary Ecology* **18**, 15-28, doi:10.1023/B:EVEC.0000017669.56353.cb (2004).

53 Long, A., Liti, G., Luptak, A. & Tenaillon, O. Elucidating the molecular architecture of adaptation via evolve and resequence experiments. *Nat. Rev. Genet.* **16**, 567-582, doi:10.1038/nrg3937 (2015).

54 Teotonio, H., Chelo, I. M., Bradic, M., Rose, M. R. & Long, A. D. Experimental evolution reveals natural selection on standing genetic variation. *Nat Genet* **41**, 251-257, doi:10.1038/ng.289 (2009).

55 Keightley, P. D., Ness, R. W., Halligan, D. L. & Haddrill, P. R. Estimation of the spontaneous mutation rate per nucleotide site in a *Drosophila melanogaster* full-sib family. *Genetics* **196**, 313-320, doi:10.1534/genetics.113.158758 (2014).

56 Signor, S. A., New, F. N. & Nuzhdin, S. A Large Panel of *Drosophila simulans* Reveals an Abundance of Common Variants. *Genome Biol. Evol.* **10**, 189-206, doi:10.1093/gbe/evx262 (2018).

57 R'Kha, S., Moreteau, B., Coyne, J. A. & David, J. R. Evolution of a lesser fitness trait: egg production in the specialist *Drosophila sechellia*. *Genet Res* **69**, 17-23, doi:10.1017/s0016672396002546 (1997).

58 Markow, T. A., Beall, S. & Matzkin, L. M. Egg size, embryonic development time and ovoviviparity in *Drosophila* species. *J Evol Biol* **22**, 430-434, doi:10.1111/j.1420-9101.2008.01649.x (2009).

59 Abe, M. *et al.* Shortened lifespan induced by a high-glucose diet is associated with intestinal immune dysfunction in *Drosophila sechellia*. *J Exp Biol* **225**, doi:10.1242/jeb.244423 (2022).

60 Watada, M. *et al.* Divergence of *Drosophila* species: Longevity and reproduction under different nutrient balances. *Genes Cells* **25**, 626-636, doi:10.1111/gtc.12798 (2020).

61 Vlachos, C. *et al.* Benchmarking software tools for detecting and quantifying selection in evolve and resequencing studies. *Genome Biol* **20**, 169, doi:10.1186/s13059-019-1770-8 (2019).

62 Kojima, Y., Matsumoto, H. & Kiryu, H. Estimation of population genetic parameters using an EM algorithm and sequence data from experimental evolution populations. *Bioinformatics* **36**, 221-231, doi:10.1093/bioinformatics/btz498 (2020).

63 Barata, C., Borges, R. & Kosiol, C. Bait-ER: A Bayesian method to detect targets of selection in Evolve-and-Resequence experiments. *J Evol Biol* **36**, 29-44, doi:10.1111/jeb.14134 (2023).
64 Cochran, W. G. Some Methods for Strengthening the Common Chi-Squared Tests. *Biometrics* **10**, 417-451, doi:10.2307/3001616 (1954).
65 Mantel, N. & Haenszel, W. Statistical aspects of the analysis of data from retrospective studies of disease. *J Natl Cancer Inst* **22**, 719-748 (1959).
66 Orozco-terWengel, P. *et al.* Adaptation of *Drosophila* to a novel laboratory environment reveals temporally heterogeneous trajectories of selected alleles. *Mol. Ecol.* **21**, 4931-4941, doi:10.1111/j.1365-294X.2012.05673.x (2012).
67 Hu, Y., Comjean, A., Perrimon, N. & Mohr, S. E. The *Drosophila* Gene Expression Tool (DGET) for expression analyses. *BMC Bioinf.* **18**, 98, doi:10.1186/s12859-017-1509-z (2017).
68 Xu, Y. *et al.* CRISPR screens in *Drosophila* cells identify Vsg as a Tc toxin receptor. *Nature* **610**, 349-355, doi:10.1038/s41586-022-05250-7 (2022).
69 Borrull, A., Lopez-Martinez, G., Poblet, M., Cordero-Otero, R. & Rozes, N. New insights into the toxicity mechanism of octanoic and decanoic acids on *Saccharomyces cerevisiae*. *Yeast* **32**, 451-460, doi:10.1002/yea.3071 (2015).
70 Emelyanov, A. V. *et al. Drosophila* TAP/p32 is a core histone chaperone that cooperates with NAP-1, NLP, and nucleophosmin in sperm chromatin remodeling during fertilization. *Genes Dev* **28**, 2027-2040, doi:10.1101/gad.248583.114 (2014).
71 Galindo, K. A., Lu, W. J., Park, J. H. & Abrams, J. M. The Bax/Bak ortholog in *Drosophila*, Debcl, exerts limited control over programmed cell death. *Development* **136**, 275-283, doi:10.1242/dev.019042 (2009).
72 Li, H. *et al.* Fly Cell Atlas: A single-nucleus transcriptomic atlas of the adult fruit fly. *Science* **375**, eabk2432, doi:10.1126/science.abk2432 (2022).
73 Lemaitre, B. & Miguel-Aliaga, I. The digestive tract of *Drosophila melanogaster*. *Annu Rev Genet* **47**, 377-404, doi:10.1146/annurev-genet-111212-133343 (2013).
74 Teufel, F. *et al.* SignalP 6.0 predicts all five types of signal peptides using protein language models. *Nat. Biotechnol.* **40**, 1023-1025, doi:10.1038/s41587-021-01156-3 (2022).
75 Yang, J. *et al.* A *Drosophila* systems approach to xenobiotic metabolism. *Physiol Genomics* **30**, 223-231, doi:10.1152/physiolgenomics.00018.2007 (2007).
76 Xu, B., Liu, D., Wang, Z., Tian, R. & Zuo, Y. Multi-substrate selectivity based on key loops and non-homologous domains: new insight into ALKBH family. *Cell Mol Life Sci* **78**, 129-141, doi:10.1007/s00018-020-03594-9 (2021).
77 Fukasawa, Y. *et al.* MitoFates: improved prediction of mitochondrial targeting sequences and their cleavage sites. *Molecular & cellular proteomics : MCP* **14**, 1113-1126, doi:10.1074/mcp.M114.043083 (2015).
78 Chen, C. L. *et al.* Proteomic mapping in live *Drosophila* tissues using an engineered ascorbate peroxidase. *Proceedings of the National Academy of Sciences of the United States of America* **112**, 12093-12098, doi:10.1073/pnas.1515623112 (2015).
79 Sen, A. & Cox, R. T. Loss of *Drosophila* Clueless differentially affects the mitochondrial proteome compared to loss of Sod2 and Pink1. *Frontiers in physiology* **13**, 1004099, doi:10.3389/fphys.2022.1004099 (2022).

80 Schrider, D. R., Ayroles, J., Matute, D. R. & Kern, A. D. Supervised machine learning reveals introgressed loci in the genomes of *Drosophila simulans* and *D. sechellia*. *PLOS Genet.* **14**, e1007341, doi:10.1371/journal.pgen.1007341 (2018).
81 McDonald, J. H. & Kreitman, M. Adaptive protein evolution at the Adh locus in *Drosophila*. *Nature* **351**, 652-654, doi:10.1038/351652a0 (1991).
82 Murrell, B. *et al.* FUBAR: a fast, unconstrained bayesian approximation for inferring selection. *Mol. Biol. Evol.* **30**, 1196-1205, doi:10.1093/molbev/mst030 (2013).
83 Watanabe, K. *et al.* Interspecies Comparative Analyses Reveal Distinct Carbohydrate-Responsive Systems among *Drosophila* Species. *Cell Rep* **28**, 2594-2607 e2597, doi:10.1016/j.celrep.2019.08.030 (2019).
84 Ma, S. *et al.* Comparative transcriptomics across 14 *Drosophila* species reveals signatures of longevity. *Aging Cell* **17**, e12740, doi:10.1111/acel.12740 (2018).
85 Kalra, S., Lanno, S., Sanchez, G. & Coolon, J. D. cis- and trans-regulatory contributions to a hierarchy of factors influencing gene expression variation. *Communications biology* **7**, 1563, doi:10.1038/s42003-024-07255-6 (2024).
86 Mota, M. N., Matos, M., Bahri, N. & Sa-Correia, I. Shared and more specific genetic determinants and pathways underlying yeast tolerance to acetic, butyric, and octanoic acids. *Microb Cell Fact* **23**, 71, doi:10.1186/s12934-024-02309-0 (2024).
87 Liu, P. *et al.* Membrane stress caused by octanoic acid in *Saccharomyces cerevisiae*. *Appl Microbiol Biotechnol* **97**, 3239-3251, doi:10.1007/s00253-013-4773-5 (2013).
88 *National Center for Biotechnology Information PubChem Compound Summary for CID 379, Octanoic Acid. Retrieved January 19, 2024*, <https://pubchem.ncbi.nlm.nih.gov/compound/Octanoic-Acid>
89 Dow, J. A. T., Simons, M. & Romero, M. F. *Drosophila melanogaster*: a simple genetic model of kidney structure, function and disease. *Nat Rev Nephrol* **18**, 417-434, doi:10.1038/s41581-022-00561-4 (2022).
90 Gutierrez-Guerrero, Y. T. *et al.* Genomic consequences of dietary diversification and parallel evolution due to nectarivory in leaf-nosed bats. *Gigascience* **9**, doi:10.1093/gigascience/giaa059 (2020).
91 Gandara, L. *et al.* Pervasive sublethal effects of agrochemicals on insects at environmentally relevant concentrations. *Science* **386**, 446-453, doi:10.1126/science.ado0251 (2024).
92 Port, F. *et al.* A large-scale resource for tissue-specific CRISPR mutagenesis in *Drosophila*. *Elife* **9**, doi:10.7554/eLife.53865 (2020).
93 Lenče, T. *PhD Thesis: The role of m6A modification on mRNA processing in Drosophila melanogaster* PhD thesis, Johannes Gutenberg University Mainz, (2021).
94 Kondo, S. & Ueda, R. Highly Improved Gene Targeting by Germline-Specific Cas9 Expression in *Drosophila*. *Genetics*, doi:10.1534/genetics.113.156737 (2013).
95 Croset, V. *et al.* Ancient protostome origin of chemosensory ionotropic glutamate receptors and the evolution of insect taste and olfaction. *PLOS Genet.* **6**, e1001064, doi:10.1371/journal.pgen.1001064 (2010).
96 Bischof, J., Maeda, R. K., Hediger, M., Karch, F. & Basler, K. An optimized transgenesis system for *Drosophila* using germ-line-specific phiC31

integrases. *Proceedings of the National Academy of Sciences of the United States of America* **104**, 3312-3317, doi:10.1073/pnas.0611511104 (2007).

97 Therneau, T. M. *coxme: Mixed Effects Cox Models. R package version 2.2-22*, <https://CRAN.R-project.org/package=coxme> (2024).

98 Wood, S. N. Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **73**, 3-36, doi:https://doi.org/10.1111/j.1467-9868.2010.00749.x (2011).

99 Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114-2120, doi:10.1093/bioinformatics/btu170 (2014).

100 Li, H. Aligning Sequence Reads, Clone Sequences and Assembly Contigs with BWA-MEM. *arXiv*, doi:1303.3997 (2013).

101 Danecek, P. *et al.* Twelve years of SAMtools and BCFtools. *Gigascience* **10**, doi:10.1093/gigascience/giab008 (2021).

102 Kofler, R., Pandey, R. V. & Schlotterer, C. PoPoolation2: identifying differentiation between populations using sequencing of pooled DNA samples (Pool-Seq). *Bioinformatics* **27**, 3435-3436, doi:10.1093/bioinformatics/btr589 (2011).

103 Danecek, P. *et al.* The variant call format and VCFtools. *Bioinformatics* **27**, 2156-2158, doi:10.1093/bioinformatics/btr330 (2011).

104 Purcell, S. *et al.* PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* **81**, 559-575, doi:10.1086/519795 (2007).

105 Purcell, S. *PLINK v1.90b6.21*, <http://pngu.mgh.harvard.edu/purcell/plink/> (2020).

106 Vlachos, C. & Kofler, R. MimicrEE2: Genome-wide forward simulations of Evolve and Resequencing studies. *PLoS Comput Biol* **14**, e1006413, doi:10.1371/journal.pcbi.1006413 (2018).

107 Taus, T., Futschik, A. & Schlotterer, C. Quantifying Selection with Pool-Seq Time Series Data. *Mol. Biol. Evol.* **34**, 3023-3034, doi:10.1093/molbev/msx225 (2017).

108 Zhang, C., Dong, S. S., Xu, J. Y., He, W. M. & Yang, T. L. PopLDdecay: a fast and effective tool for linkage disequilibrium decay analysis based on variant call format files. *Bioinformatics* **35**, 1786-1788, doi:10.1093/bioinformatics/bty875 (2019).

109 Dong, S. S. *et al.* LDBlockShow: a fast and convenient tool for visualizing linkage disequilibrium and haplotype blocks based on variant call format files. *Briefings in bioinformatics* **22**, doi:10.1093/bib/bbaa227 (2021).

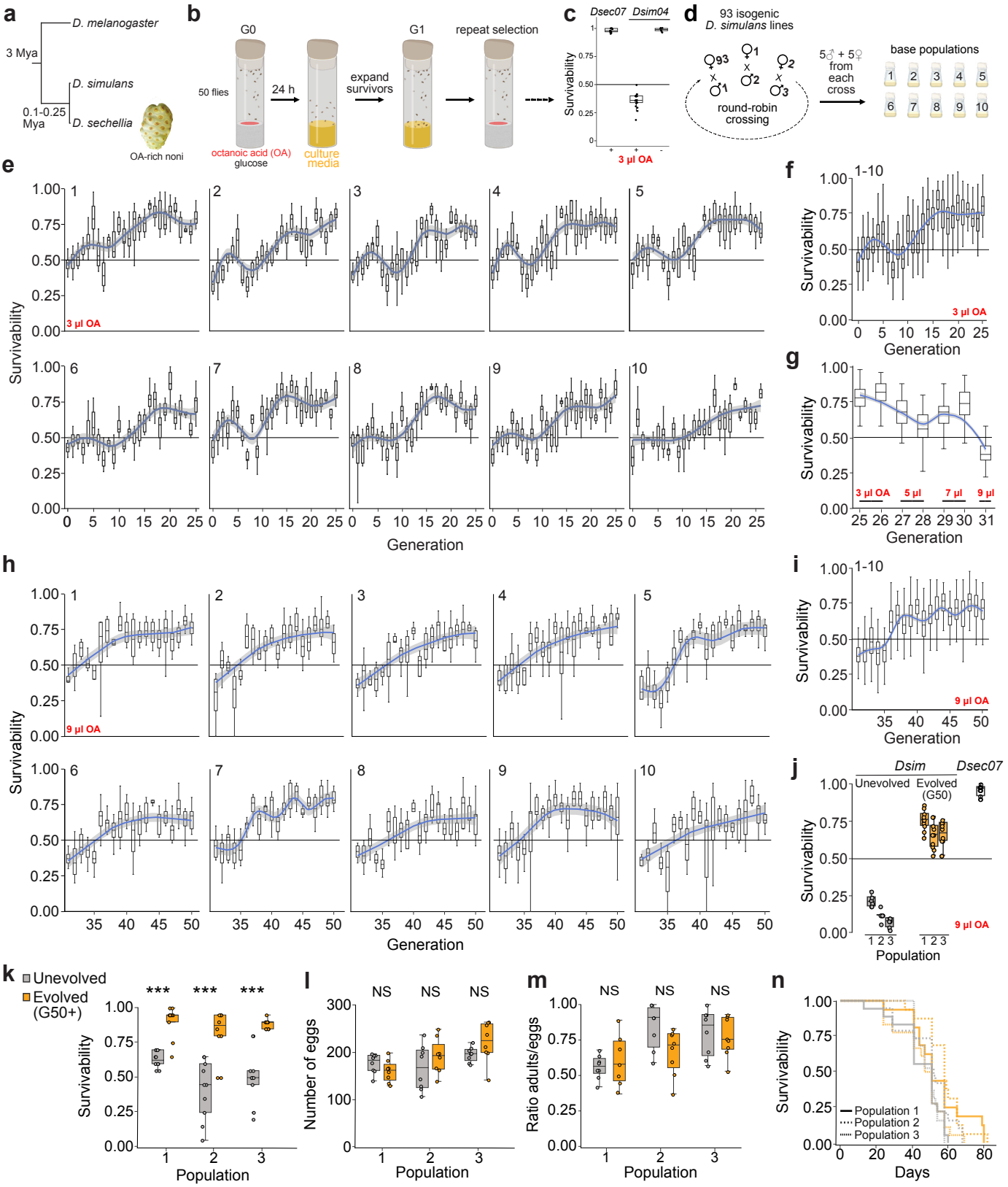
110 Barrett, J. C., Fry, B., Maller, J. & Daly, M. J. Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics* **21**, 263-265, doi:10.1093/bioinformatics/bth457 (2005).

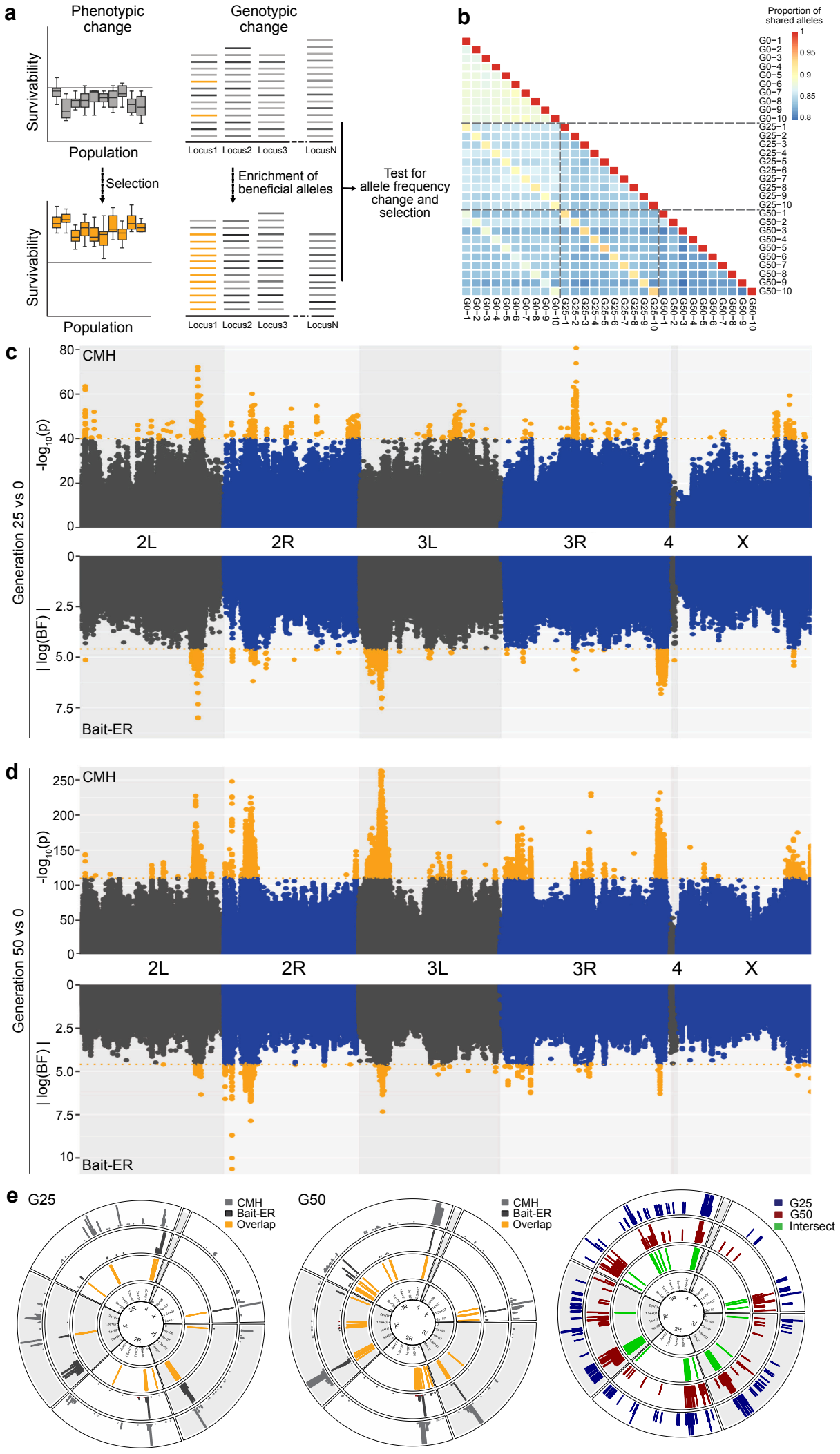
111 Gu, Z., Gu, L., Eils, R., Schlesner, M. & Brors, B. circlize Implements and enhances circular visualization in R. *Bioinformatics* **30**, 2811-2812, doi:10.1093/bioinformatics/btu393 (2014).

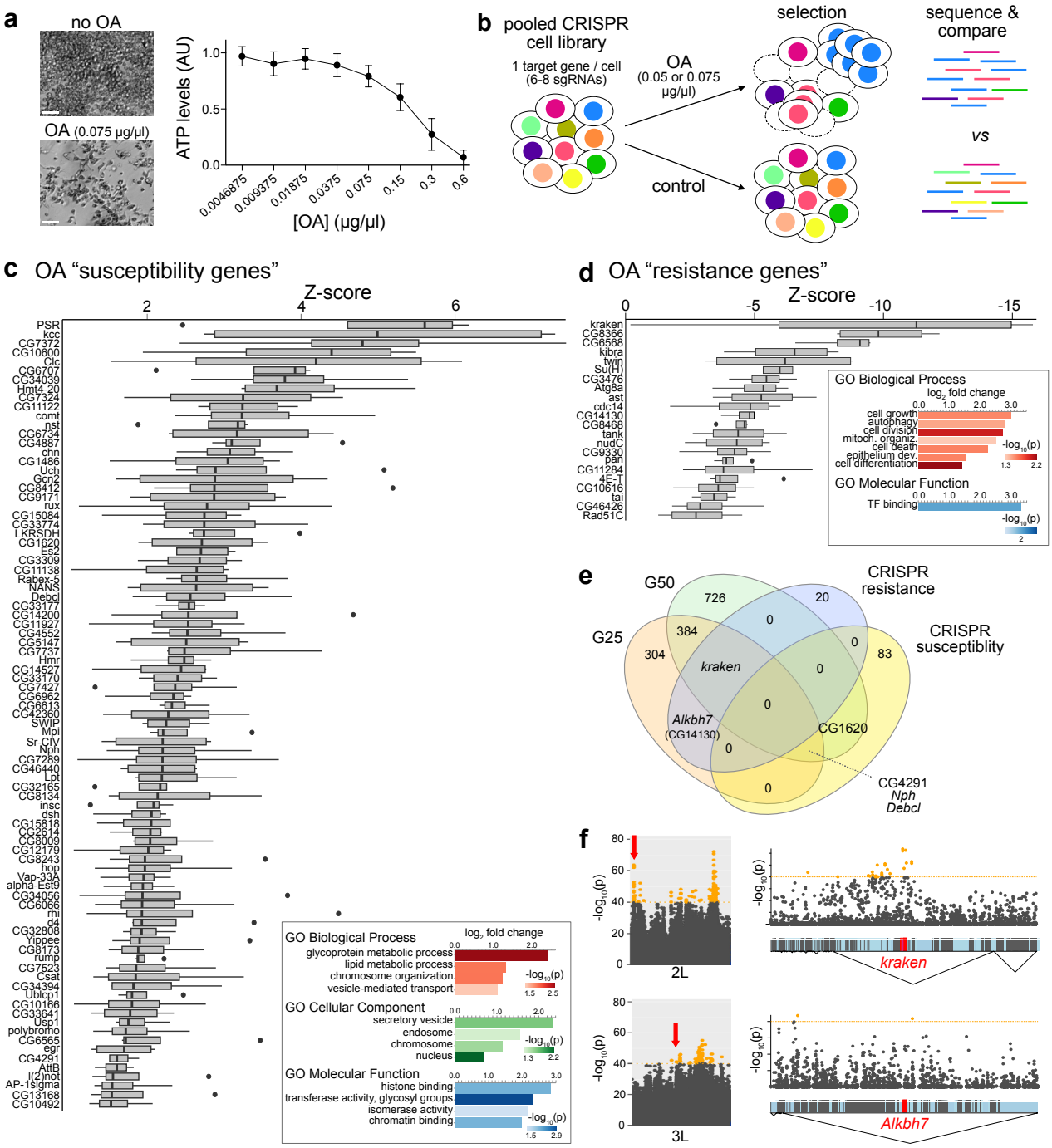
112 Quinlan, A. R. & Hall, I. M. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841-842, doi:10.1093/bioinformatics/btq033 (2010).

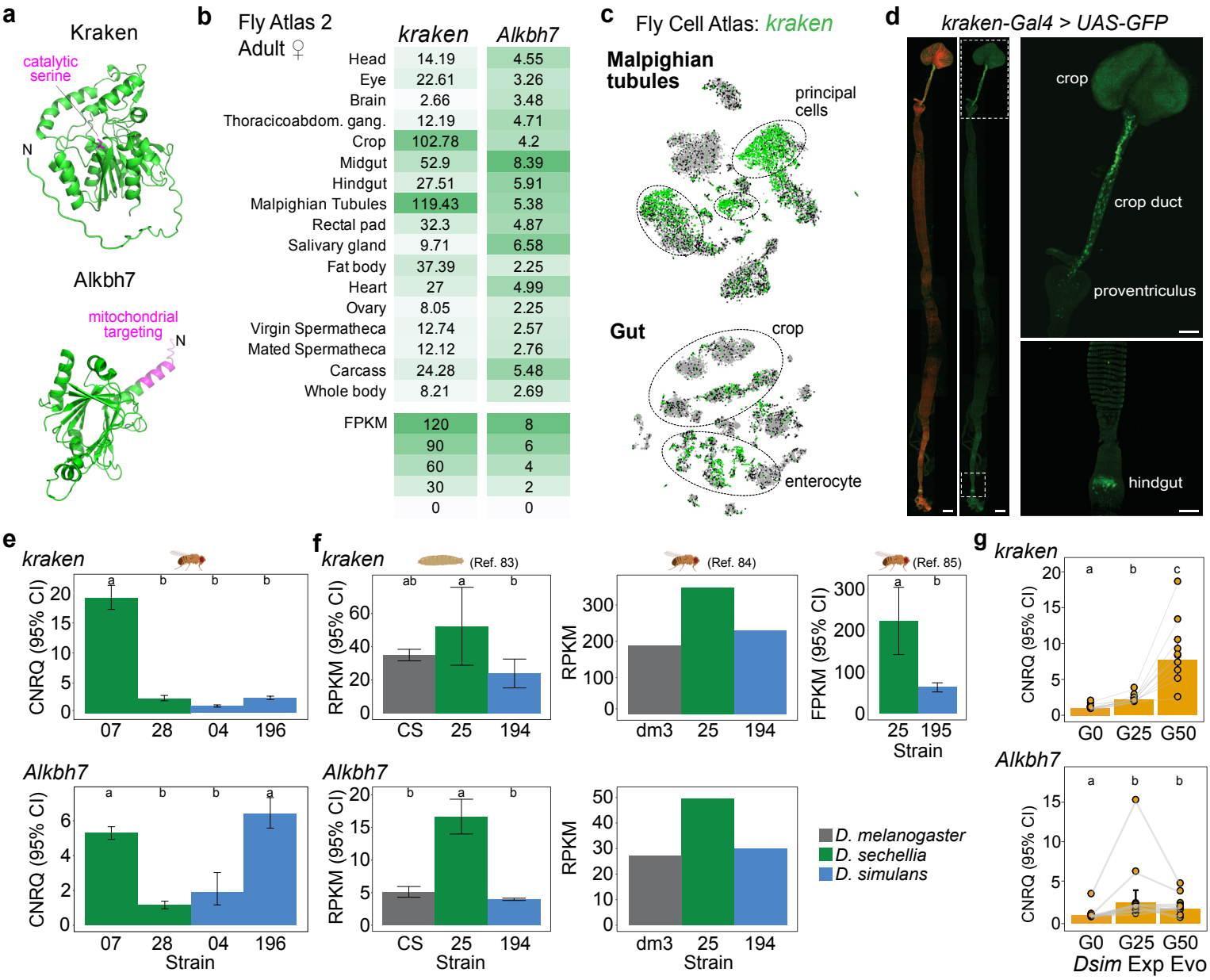
113 Altenhoff, A. M. *et al.* OMA orthology in 2021: website overhaul, conserved isoforms, ancestral gene order and more. *Nucleic Acids Res.* **49**, D373-D379, doi:10.1093/nar/gkaa1007 (2021).
114 Sherman, B. T. *et al.* DAVID: a web server for functional enrichment analysis and functional annotation of gene lists (2021 update). *Nucleic Acids Res.* **50**, W216-W221, doi:10.1093/nar/gkac194 (2022).
115 Huang da, W., Sherman, B. T. & Lempicki, R. A. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc* **4**, 44-57, doi:10.1038/nprot.2008.211 (2009).
116 Camacho, C. *et al.* BLAST+: architecture and applications. *BMC Bioinf.* **10**, 421, doi:10.1186/1471-2105-10-421 (2009).
117 Sayers, E. W. *et al.* Database resources of the National Center for Biotechnology Information in 2025. *Nucleic Acids Res.* **53**, D20-D29, doi:10.1093/nar/gkae979 (2025).
118 Ozturk-Colak, A. *et al.* FlyBase: updates to the *Drosophila* genes and genomes database. *Genetics* **227**, doi:10.1093/genetics/iyad211 (2024).
119 Viswanatha, R. *et al.* Higher resolution pooled genome-wide CRISPR knockout screening in *Drosophila* cells using integration and anti-CRISPR (IntAC). *Nat Commun* **16**, 6498, doi:10.1038/s41467-025-61692-3 (2025).
120 Viswanatha, R. *et al.* Pooled CRISPR Screens in *Drosophila* Cells. *Current protocols in molecular biology* **129**, e111, doi:10.1002/cpmb.111 (2019).
121 Li, W. *et al.* MAGeCK enables robust identification of essential genes from genome-scale CRISPR/Cas9 knockout screens. *Genome Biol* **15**, 554, doi:10.1186/s13059-014-0554-4 (2014).
122 Hu, Y. *et al.* PANGEA: a new gene set enrichment tool for *Drosophila* and common research organisms. *Nucleic Acids Res.* **51**, W419-W426, doi:10.1093/nar/gkad331 (2023).
123 Hellemans, J., Mortier, G., De Paepe, A., Speleman, F. & Vandesompele, J. qBase relative quantification framework and software for management and automated analysis of real-time quantitative PCR data. *Genome Biol* **8**, R19, doi:10.1186/gb-2007-8-2-r19 (2007).
124 Lu, J., Yang, C., Zhang, Y. & Pan, H. Selection of Reference Genes for the Normalization of RT-qPCR Data in Gene Expression Studies in Insects: A Systematic Review. *Frontiers in physiology* **9**, 1560, doi:10.3389/fphys.2018.01560 (2018).
125 Vandesompele, J. *et al.* Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biol* **3**, RESEARCH0034, doi:10.1186/gb-2002-3-7-research0034 (2002).
126 Rogers, R. L. *et al.* Landscape of standing variation for tandem duplications in *Drosophila yakuba* and *Drosophila simulans*. *Mol. Biol. Evol.* **31**, 1750-1766, doi:10.1093/molbev/msu124 (2014).
127 Leinonen, R. *et al.* Improvements to services at the European Nucleotide Archive. *Nucleic Acids Res.* **38**, D39-45, doi:10.1093/nar/gkp998 (2010).
128 Walker, B. J. *et al.* Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. *PLOS ONE* **9**, e112963, doi:10.1371/journal.pone.0112963 (2014).
129 Abascal, F., Zardoya, R. & Telford, M. J. TranslatorX: multiple alignment of nucleotide sequences guided by amino acid translations. *Nucleic Acids Res.* **38**, W7-13, doi:10.1093/nar/gkq291 (2010).

130 Madeira, F. *et al.* The EMBL-EBI Job Dispatcher sequence analysis tools framework in 2024. *Nucleic Acids Res.* **52**, W521-W525, doi:10.1093/nar/gkae241 (2024).
131 Rozas, J. *et al.* DnaSP 6: DNA Sequence Polymorphism Analysis of Large Data Sets. *Mol. Biol. Evol.* **34**, 3299-3302, doi:10.1093/molbev/msx248 (2017).
132 Kosakovsky Pond, S. L. *et al.* HyPhy 2.5-A Customizable Platform for Evolutionary Hypothesis Testing Using Phylogenies. *Mol. Biol. Evol.* **37**, 295-299, doi:10.1093/molbev/msz197 (2020).
133 Robert, X. & Gouet, P. Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res.* **42**, W320-324, doi:10.1093/nar/gku316 (2014).
134 Varadi, M. *et al.* AlphaFold Protein Structure Database: massively expanding the structural coverage of protein-sequence space with high-accuracy models. *Nucleic Acids Res.* **50**, D439-D444, doi:10.1093/nar/gkab1061 (2022).
135 Jumper, J. *et al.* Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583-589, doi:10.1038/s41586-021-03819-2 (2021).
136 Krause, S. A., Overend, G., Dow, J. A. T. & Leader, D. P. FlyAtlas 2 in 2022: enhancements to the *Drosophila melanogaster* expression atlas. *Nucleic Acids Res.* **50**, D1010-D1015, doi:10.1093/nar/gkab971 (2022).
137 Hadjieconomou, D. *et al.* Flybow: genetic multicolor cell labeling for neural circuit analysis in *Drosophila melanogaster*. *Nat Methods* **8**, 260-266, doi:nmeth.1567 [pii] 10.1038/nmeth.1567 (2011).
138 Bontonou, G. *et al.* Evolution of chemosensory tissues and cells across ecologically diverse Drosophilids. *Nat Commun* **15**, 1047, doi:10.1038/s41467-023-44558-4 (2024).


a
D. melanogaster
3 Mya
D. simulans
0.1-0.25 Mya
D. sechellia
OA-rich noni
b
G0
50 flies
24 h
octanoic acid (OA)
glucose
culture media
expand survivors
G1
repeat selection
c
Dsec07
Dsim04
Survivability
3 µl OA
d
93 isogenic D. simulans lines
round-robin crossing
5♂ + 5♀ from each cross
base populations
e
Survivability
3 µl OA
Generation
f
1-10
Survivability
3 µl OA
Generation
g
Survivability
3 µl OA
5 µl
7 µl
9 µl
Generation
h
Survivability
9 µl OA
Generation
i
1-10
Survivability
9 µl OA
Generation
j
Dsim
Dsec07
Unevolved
Evolved (G50)
Survivability
9 µl OA
Population
k
Unevolved
Evolved (G50+)
Survivability
Population
l
Number of eggs
Population
m
Ratio adults/eggs
Population
n
Survivability
Population 1
Population 2
Population 3
Days

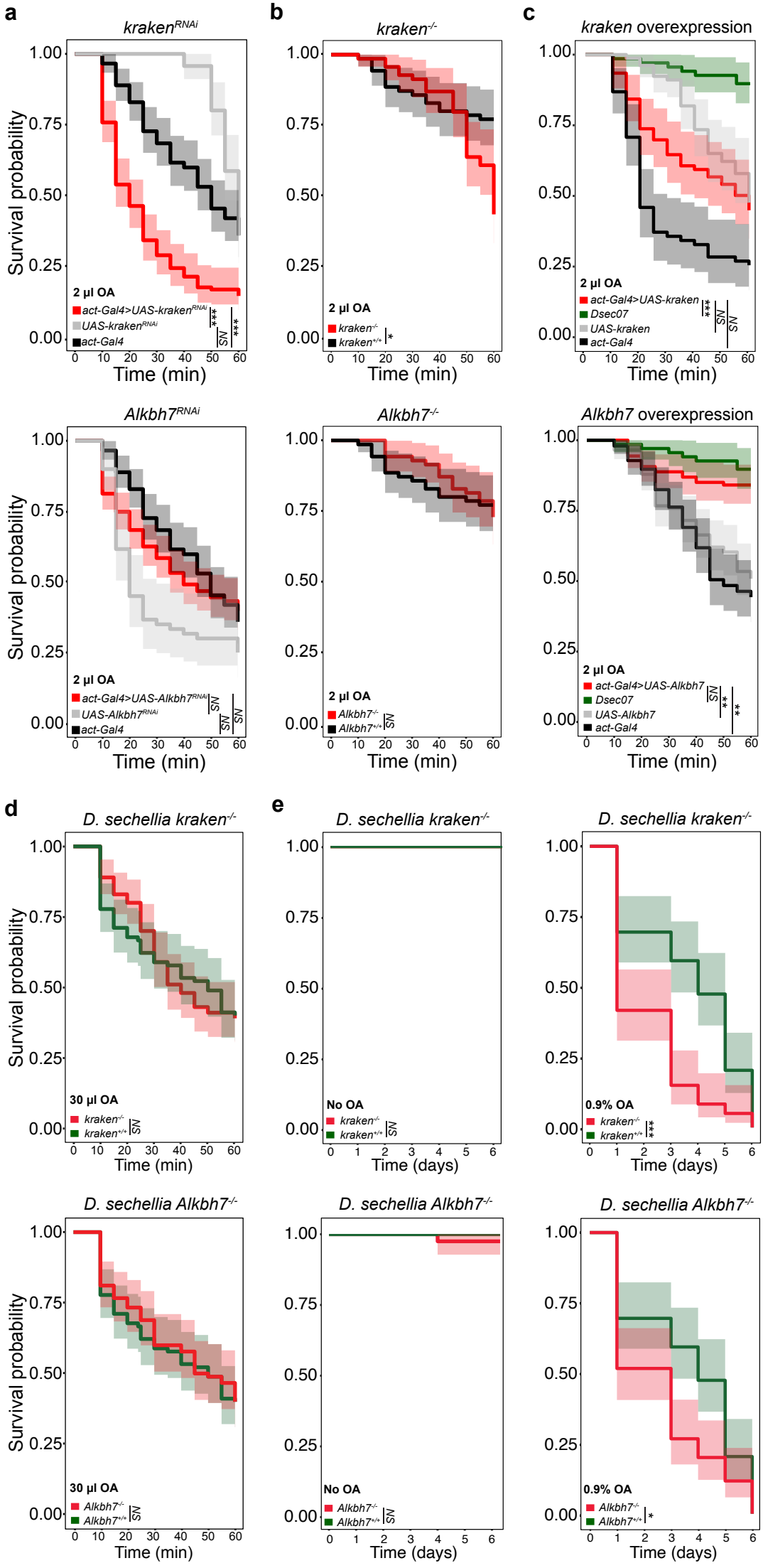

a
Phenotypic change
Genotypic change
Survivability
Population
Selection
Enrichment of beneficial alleles
Locus1 Locus2 Locus3 LocusN
Test for allele frequency change and selection
b
Proportion of shared alleles
c
CMH
-log10(p)
Generation 25 vs 0
2L 2R 3L 3R 4 X
| log(BF) |
Bait-ER
d
CMH
Generation 50 vs 0
Bait-ER
e
G25
G50
CMH
Bait-ER
Overlap
Intersect

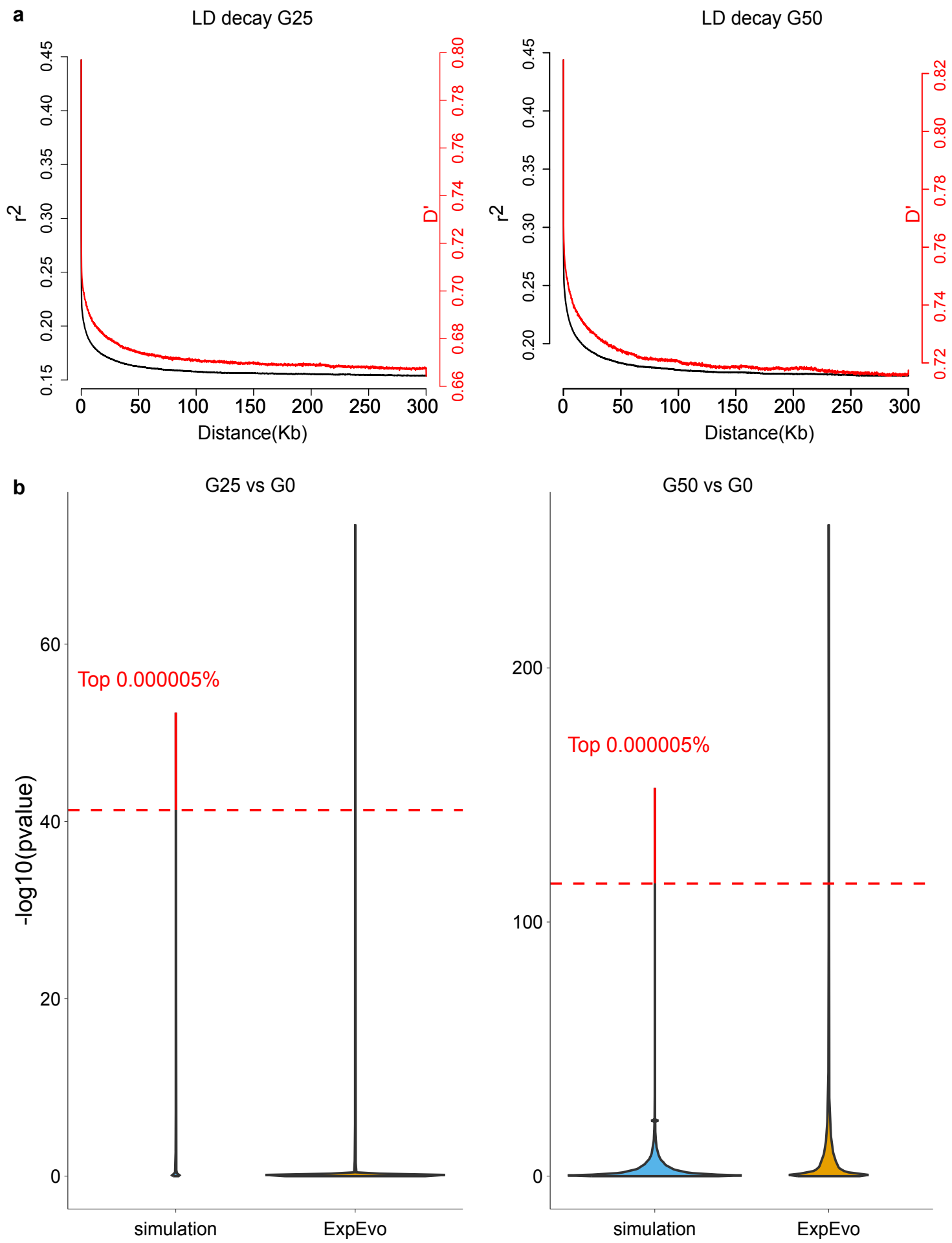

a
no OA
OA (0.075 µg/µl)
ATP levels (AU)
1.0
0.5
0.0
0.0046875
0.009375
0.01875
0.0375
0.075
0.15
0.3
0.6
[OA] (µg/µl)
b
pooled CRISPR cell library
1 target gene / cell (6-8 sgRNAs)
OA (0.05 or 0.075 µg/µl)
control
selection
sequence & compare
vs
c
OA "susceptibility genes"
Z-score
2
4
6
PSR
kcc
CG7372
CG10600
Clc
CG6707
CG34039
Hmt4-20
CG7324
CG11122
comt
nst
CG6734
CG4887
chn
CG1486
Uch
Gcn2
CG8412
CG9171
nux
CG15084
CG33774
LKRSDH
CG1620
Es2
CG3309
CG11138
Rabex-5
NANS
Debcl
CG33177
CG14200
CG11927
CG4552
CG5147
CG7737
Hmr
CG14527
CG33170
CG7427
CG6962
CG6613
CG42360
SWIP
Mpi
Sr-CIV
Nph
CG7289
CG46440
Lpt
CG32165
CG8134
insc
dsh
CG15818
CG2614
CG8009
CG12179
CG8243
hop
Vap-33A
alpha-Est9
CG34056
CG6066
rhi
d4
CG32808
Yippee
CG8173
rump
CG7523
Csat
CG34394
Ubiquitin-conjugating enzyme 9
CG10166
CG33641
Usp1
polybromo
CG6565
egr
CG4291
AttB
l(2)not
AP-1sigma
CG13168
CG10492
GO Biological Process
log2 fold change
glycoprotein metabolic process
lipid metabolic process
chromosome organization
vesicle-mediated transport
-log10(p)
GO Cellular Component
secretory vesicle
endosome
chromosome
nucleus
GO Molecular Function
histone binding
transferase activity, glycosyl groups
isomerase activity
chromatin binding
d
OA "resistance genes"
Z-score
0
-5
-10
-15
kraken
CG8366
CG6568
kibra
twin
Su(tr)
CG3476
Atg8a
ast
cdc14
CG14130
CG8468
tank
nudC
CG9330
pan
CG11284
4E-T
CG10616
tai
CG46426
Rad51C
GO Biological Process
log2 fold change
cell growth
autophagy
cell division
mitoch. organiz.
cell death
epithelium dev.
cell differentiation
-log10(p)
GO Molecular Function
TF binding
e
G50
G25
CRISPR resistance
CRISPR susceptiblity
726
20
384
304
83
kraken
Alkbh7 (CG14130)
CG1620
CG4291
Nph
Debcl
f
-log10(p)
2L
3L
kraken
Alkbh7

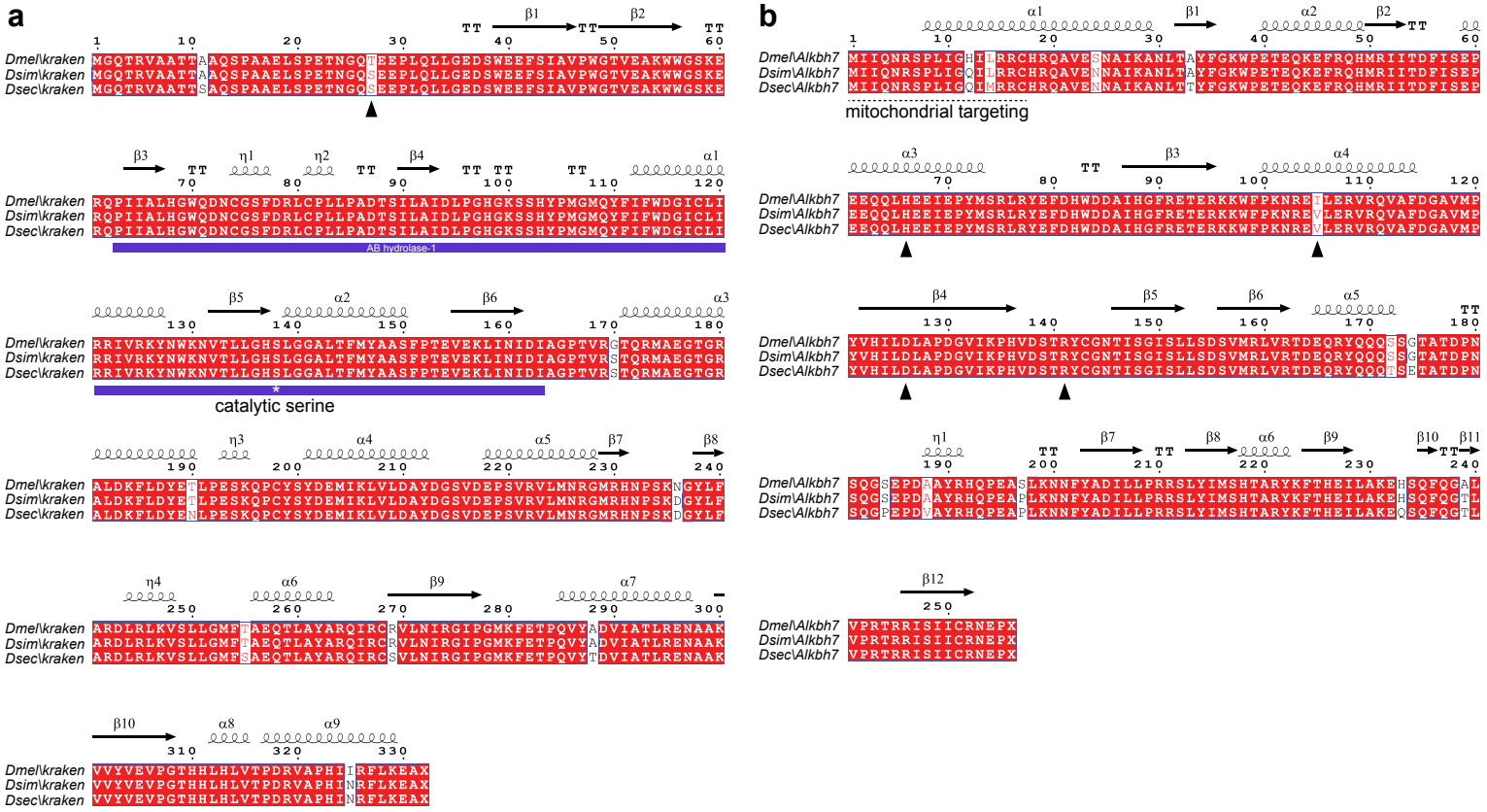


	kraken	Alkbh7
Head	14.19	4.55
Eye	22.61	3.26
Brain	2.66	3.48
Thoracicoabdom. gang.	12.19	4.71
Crop	102.78	4.2
Midgut	52.9	8.39
Hindgut	27.51	5.91
Malpighian Tubules	119.43	5.38
Rectal pad	32.3	4.87
Salivary gland	9.71	6.58
Fat body	37.39	2.25
Heart	27	4.99
Ovary	8.05	2.25
Virgin Spermatheca	12.74	2.57
Mated Spermatheca	12.12	2.76
Carcass	24.28	5.48
Whole body	8.21	2.69
FPKM	120	8
	90	6
	60	4
	30	2
	0	0

Marconcini, Figure 5




a
LD decay G25
r2
D'
Distance(Kb)
LD decay G50
b
G25 vs G0
G50 vs G0
Top 0.000005%
-log10(pvalue)
simulation
ExpEvo



c

McDonald & Kreitman Test

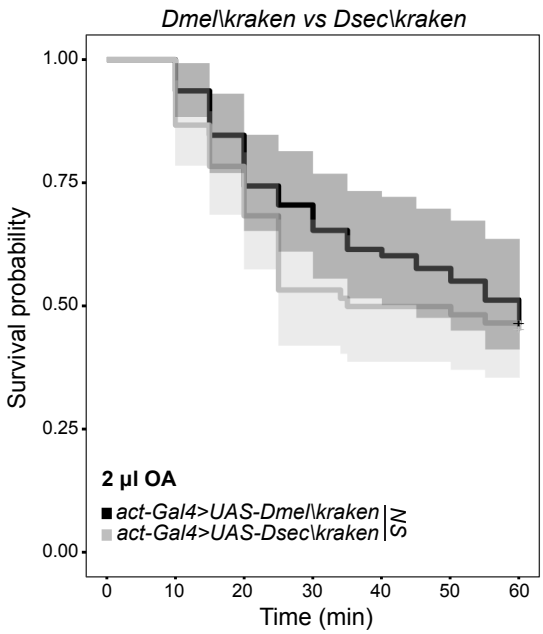
	NI	alpha	p-value
kraken	0.083	0.917	0.018*
Alkbh7	0.544	0.456	0.477

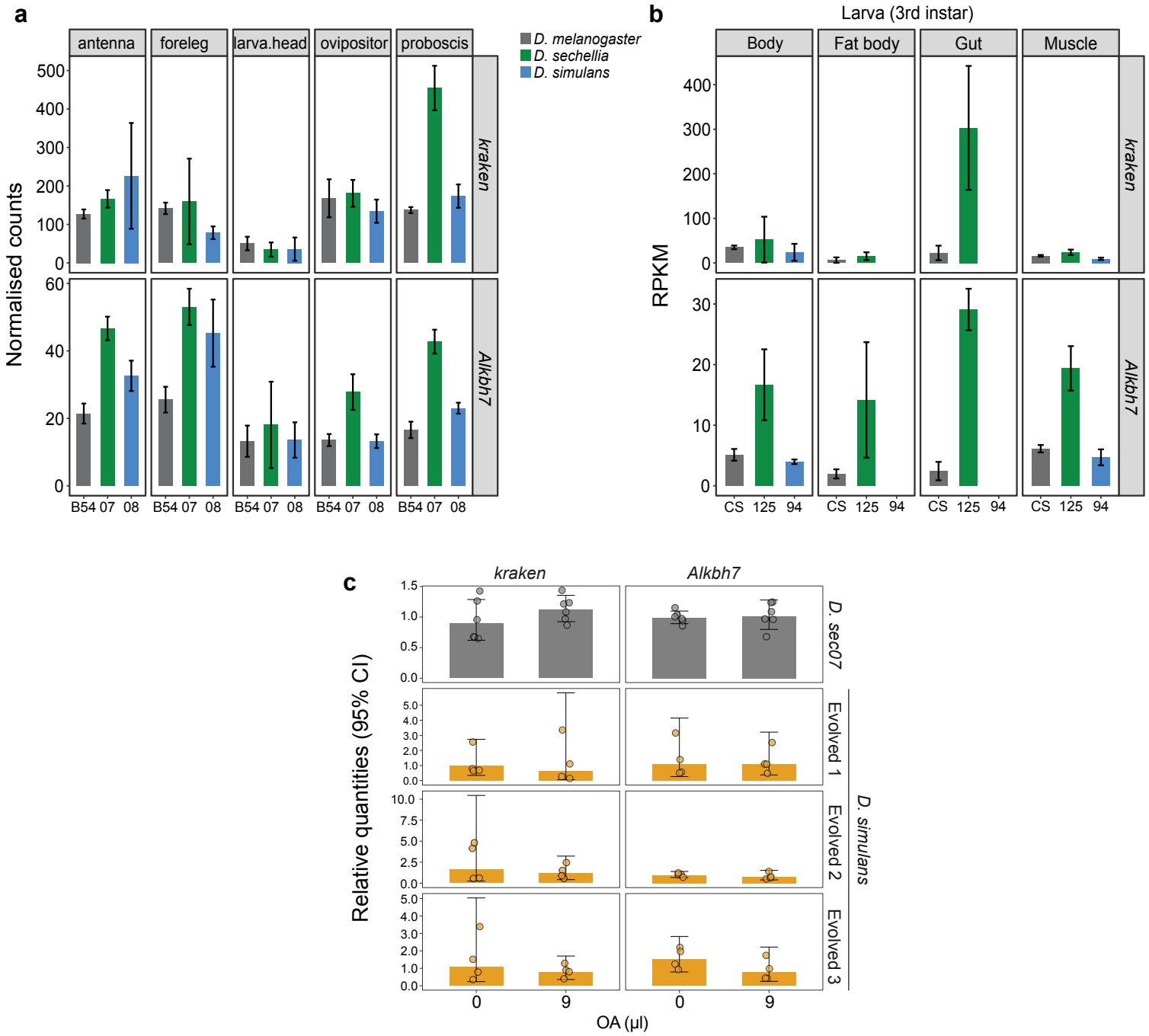
d

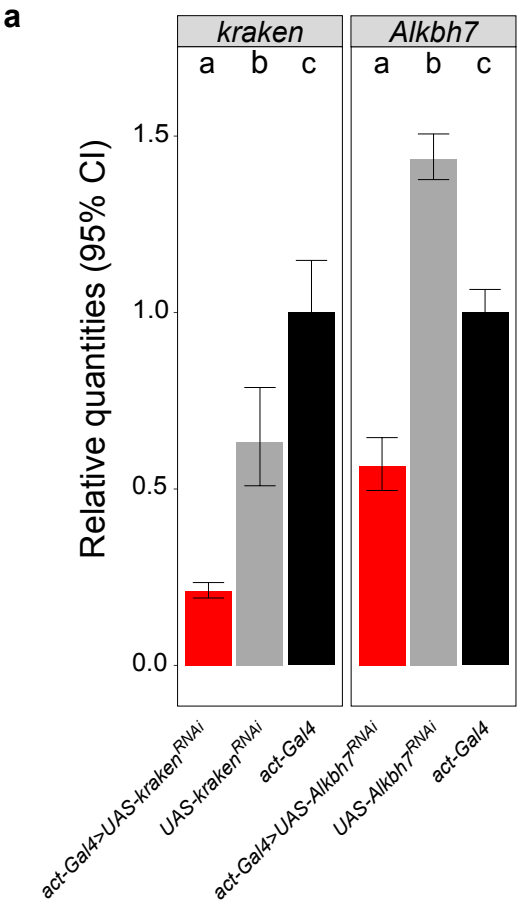
HyPhy FUBAR

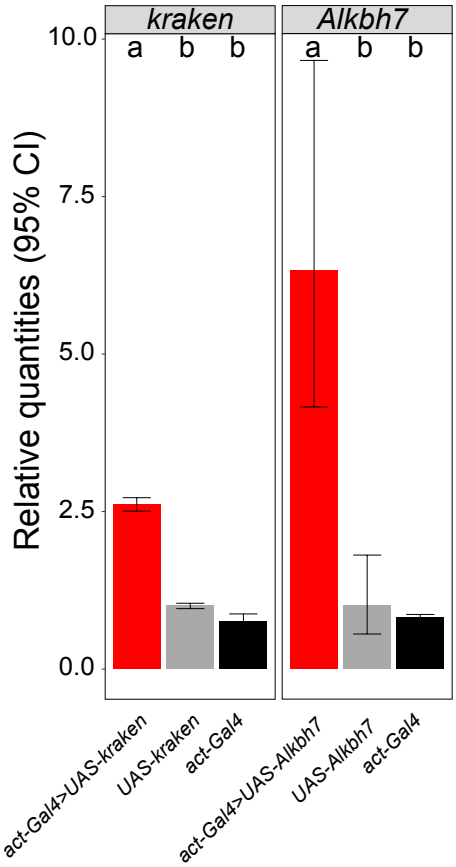
	position	alpha	beta	post. prob.	codons	Freq *Dsec*	Freq *Dsim*
kraken	27	2.503	40.230	0.979	S / A / T	1 / 0 / 0	0.658 / 0.237 / 0.105
Alkbh7	66	0.768	9.440	0.977	H / Y	1 / 0	0.806 / 0.194
	105	0.665	7.158	0.957	V / I / L	1 / 0 / 0	0.839 / 0.097 / 0.064
	126	0.783	13.505	0.989	D / E	1 / 0	0.774 / 0.226
	141	0.723	9.841	0.982	R / K	1 / 0	0.871 / 0.129

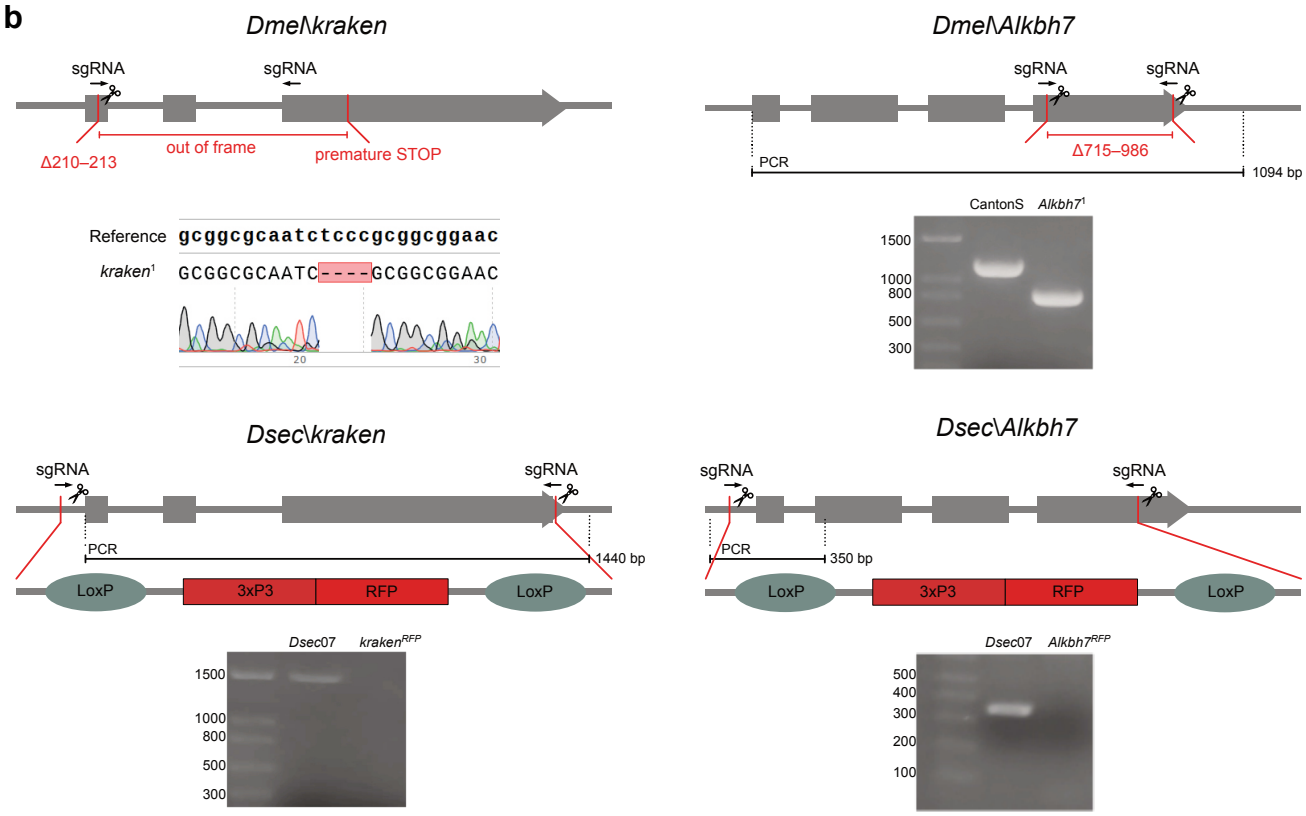
e

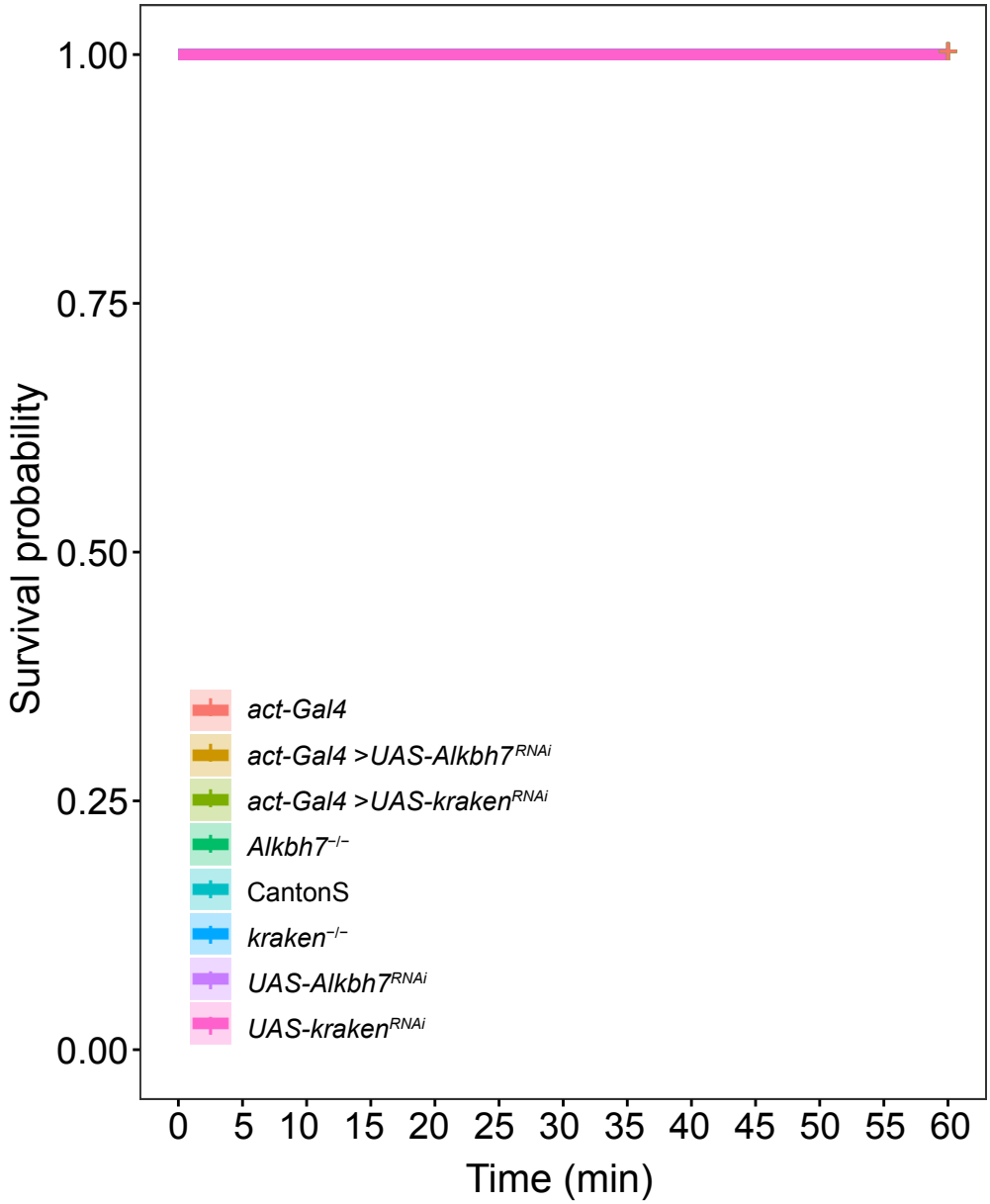



a
antenna
foreleg
larva.head
ovipositor
proboscis
D. melanogaster
D. sechellia
D. simulans
kraken
Alkbh7
Normalised counts
B54 07 08
b
Larva (3rd instar)
Body
Fat body
Gut
Muscle
RPKM
CS 125 94
c
kraken
Alkbh7
D. sec07
Evolved 1
Evolved 2
Evolved 3
D. simulans
Relative quantities (95% CI)
OA (µl)


a
kraken
Alkbh7
a b c
a b c
Relative quantities (95% CI)
1.5
1.0
0.5
0.0
act-Gal4>UAS-kraken^RNAi
UAS-kraken^RNAi
act-Gal4
act-Gal4>UAS-Alkbh7^RNAi
UAS-Alkbh7^RNAi
act-Gal4


c
kraken
Alkbh7
a b b
a b b
Relative quantities (95% CI)
10.0
7.5
5.0
2.5
0.0
act-Gal4>UAS-kraken
UAS-kraken
act-Gal4
act-Gal4>UAS-Alkbh7
UAS-Alkbh7
act-Gal4


b
Dme\kraken
sgRNA
sgRNA
Δ210–213
out of frame
premature STOP
Reference gcggcgcaatctcccgcggcggaac
kraken¹ GCGGCGCAATC----GCGGCGGAAC
20
30
Dme\Alkbh7
sgRNA
sgRNA
PCR
Δ715–986
1094 bp
CantonS
Alkbh7¹
1500
1000
800
500
300
Dsec\kraken
sgRNA
sgRNA
PCR
1440 bp
LoxP
3xP3
RFP
LoxP
Dsec07
kraken^RFP
1500
1000
800
500
300
Dsec\Alkbh7
sgRNA
sgRNA
PCR
350 bp
LoxP
3xP3
RFP
LoxP
Dsec07
Alkbh7^RFP
500
400
300
200
100


1.00
0.75
0.50
0.25
0.00
Survival probability
0
5
10
15
20
25
30
35
40
45
50
55
60
Time (min)
act-Gal4
act-Gal4 >UAS-Alkbh7RNAi
act-Gal4 >UAS-krakenRNAi
Alkbh7−/−
CantonS
kraken−/−
UAS-Alkbh7RNAi
UAS-krakenRNAi