Behavioral and Genomic Sensory Adaptations Underlying the Pest Activity of *Drosophila suzukii*

Sylvia M. Durkin,[1,2] Mahul Chakraborty,[3] Antoine Abrieux,[4] Kyle M. Lewald,[4] Alice Gadau,[1] Nicolas Svetec (iD),[1] Junhui Peng,[1] Miriam Kopyto,[1] Christopher B. Langer,[1] Joanna C. Chiu,[4] J.J. Emerson,[3] and Li Zhao (iD)*[1]

[1]Laboratory of Evolutionary Genetics and Genomics, The Rockefeller University, New York, NY, USA
[2]Department of Integrative Biology and Museum of Vertebrate Zoology, University of California, Berkeley, Berkeley, CA, USA
[3]Department of Ecology and Evolutionary Biology, University of California, Irvine, CA, USA
[4]Department of Entomology and Nematology, College of Agricultural and Environmental Sciences, University of California, Davis, Davis, CA, USA

*Corresponding author: E-mail: lzhao@rockefeller.edu.
Associate editor: Rebekah Rogers

Abstract

Studying how novel phenotypes originate and evolve is fundamental to the field of evolutionary biology as it allows us to understand how organismal diversity is generated and maintained. However, determining the basis of novel phenotypes is challenging as it involves orchestrated changes at multiple biological levels. Here, we aim to overcome this challenge by using a comparative species framework combining behavioral, gene expression, and genomic analyses to understand the evolutionary novel egg-laying substrate-choice behavior of the invasive pest species *Drosophila suzukii*. First, we used egg-laying behavioral assays to understand the evolution of ripe fruit oviposition preference in *D. suzukii* compared with closely related species *D. subpulchrella* and *D. biarmipes* as well as *D. melanogaster*. We show that *D. subpulchrella* and *D. biarmipes* lay eggs on both ripe and rotten fruits, suggesting that the transition to ripe fruit preference was gradual. Second, using two-choice oviposition assays, we studied how *D. suzukii*, *D. subpulchrella*, *D. biarmipes*, and *D. melanogaster* differentially process key sensory cues distinguishing ripe from rotten fruit during egg-laying. We found that *D. suzukii*'s preference for ripe fruit is in part mediated through a species-specific preference for stiff substrates. Last, we sequenced and annotated a high-quality genome for *D. subpulchrella*. Using comparative genomic approaches, we identified candidate genes involved in *D. suzukii*'s ability to seek out and target ripe fruits. Our results provide detail to the stepwise evolution of pest activity in *D. suzukii*, indicating important cues used by this species when finding a host, and the molecular mechanisms potentially underlying their adaptation to a new ecological niche.

Key words: Drosophila suzukii, Drosophila subpulchrella, egg-laying, novel behavior, sensory adaptation, population genomics.

Introduction

Novel phenotypes can give species the opportunity to occupy a new ecological niche (Mayr 1960; Muller and Wagner 1991; Moczek 2008). Understanding how and when these phenotypes arise is an exciting question in evolutionary biology. Adaptive traits can present as changes to an organism's behavior, physiology, or morphology and arise through a variety of different molecular mechanisms. In the context of pest species, adaptation to a new ecological niche can come with damaging environmental, agricultural, and economic consequences. Understanding the basis of novel pest behavior can also shed light on the ecological impact of invasive species. Unlike the majority of *Drosophila* flies, the species *Drosophila suzukii* prefers to lay its eggs in ripe as opposed to rotten fruit, causing substantial crop damage and leading to economic losses in the fruit industry (Rota-Stabelli et al. 2013). Although originally categorized in Japan and likely to be native to East Asia, within the past decade *D. suzukii* has spread throughout Europe and North America (Adrion et al. 2014; Fraimout et al. 2017; Paris et al. 2020). Preference for ripe fruit is thought to have evolved along the lineage leading to *D. suzukii*, as *D. melanogaster* and other outgroup species, including *D. takahashii*, *D. eugracilis*, and *D. ananassae*, all have an oviposition preference for rotten fruit (Karageorgi et al. 2017). *D. suzukii* has both physical and behavioral traits that allow it to selectively target ripe fruit as an oviposition substrate (Atallah et al. 2014). Morphologically, *D. suzukii* has evolved an enlarged, serrated ovipositor allowing females to puncture hard surfaces and insert their eggs into ripe fruit. Behaviorally, they have evolved the ability to seek out and selectively target ripe fruits for oviposition through changes in multiple sensory systems (Karageorgi et al. 2017). Thus, the exploitation of the ripe fruit niche by *D. suzukii* requires orchestrated changes at multiple biological levels. Investigating the behavioral and molecular underpinnings of these changes

Open Access

will advance our understanding of the forces and mechanisms that drove *D. suzukii*'s evolution into an invasive pest.

Adaptive shifts in host preference, such as that described for *D. suzukii*, are often mediated through sensory system evolution. Changes in the function and sequence of chemo-receptor genes—including odorant receptors (ORs), iono-tropic receptors (IRs), and gustatory receptors (GRs)—and mechanosensory receptor genes (MRs) underlie various species-specific host preference differences in insects (Dekker et al. 2006; Vosshall and Stocker 2007; Goldman-Huertas et al. 2015; Karageorgi et al. 2017; Mansourian et al. 2018; Auer et al. 2020). For example, modifications to sensory receptor genes are linked to the transition to herbivory in *Scaptomyza flava* (Goldman-Huertas et al. 2015), the ovipo-sition preference for morinda fruits (*Morinda citrafolia*) in *D. sechellia* (Dekker et al. 2006; Auer et al. 2020), and the specialization on marula fruit in ancient *D. melanogaster* (Mansourian et al. 2018). Additionally, through knockout experiments in *D. melanogaster*, sensory gene function has been linked to various behaviors relevant to *D. suzukii* pest activity, including preference of acetic acid containing ovipo-sition sites (Joseph et al. 2009), the avoidance of stiff sub-strates (Jeong et al. 2016; Zhang et al. 2016, 2020), and the detection of substances that are at different concentrations in ripe and rotten fruit—such as CO_2, acetic acid, and sugar (Kwon et al. 2007; Fujii et al. 2015; Rimal et al. 2019).

Although progress has been made in understanding the evolution of olfactory genes in *D. suzukii* (Keesey et al. 2015; Ramasamy et al. 2016), very little is known about the evolu-tionary history of other sensory genes in this system, despite *D. suzukii*'s oviposition site preference being mediated via multiple sensory systems (Karageorgi et al. 2017). Finally, stud-ies of adaptive sensory gene evolution in *D. suzukii* often focus only on changes to the coding sequence of these genes, de-spite differential gene expression playing a prominent role in the evolution of adaptive phenotypes (Carroll 2005; Wray 2007; Jones et al. 2012; Phifer-Rixey et al. 2018). For example, changes in expression of ORs and odorant-binding proteins in the olfactory organs of *D. sechellia* are thought to in part mediate this species' specialization on morinda fruits (Kopp et al. 2008).

Although the phenotypic and behavioral innovations leading to the pest status of *D. suzukii* are overall well defined, the specific sensory cues used by *D. suzukii* to target ripe fruits as well as the molecular changes accompanying these sensory changes remain unknown. Here we examine these knowledge gaps by investigating the behavioral, genomic, and gene regulatory mechanisms underlying the pest activity of *D. suzukii*.

First, we provide new insights into the stepwise evolution of ripe fruit preference in *D. suzukii* through the in-depth behavioral and genomic examination of two closely related species, *D. subpulchrella* and *D. biarmipes*. The divergence times of *D. subpulchrella* and *D. biarmipes* to *D. suzukii* are about 2–3 Ma (Suvorov et al. 2020) and 6–9 Ma (Ometto et al. 2013), respectively. Second, we identified key sensory cues of ripe and rotten fruit differentially processed by

D. suzukii, *D. subpulchrella*, *D. biarmipes*, and *D. melanogaster* in the context of egg-laying preference. Third, to infer the most recent changes in genome content in *D. suzukii*, we sequenced the genome of the closely related species, *D. subpulchrella* and performed comparative geno-mic analyses. Lastly, we identified candidate genes involved in *D. suzukii*'s ability to seek out and target ripe fruits through population level analyses of OR, IR, GR, and MR genes with differential expression and signatures of positive selection in *D. suzukii* compared with *D. subpulchrella*, *D. biarmipes*, and *D. melanogaster*. We provide a view, from multiple levels of analysis, of the origin of pest activity in *D. suzukii*, indicating important sensory cues used by this species during egg-laying, and the gene expression and genomic changes potentially underlying this novel pest behavior.

Results and Discussion

Ripe Fruit Preference in *D. suzukii* Evolved after the Split with *D. subpulchrella*

To clarify the evolutionary history of *D. suzukii*'s oviposition behavior, we built upon previous work that established a stepwise evolution of ripe fruit oviposition preference in *D. suzukii*, with *D. melanogaster* preferring rotten fruit and *D. biarmipes*, showing an equal preference for ripe and rotten fruit (Karageorgi et al. 2017) by measuring the oviposition preference of *D. suzukii*'s sister species, *D. subpulchrella*. It was previously hypothesized that *D. subpulchrella* prefers ripe fruit for oviposition, as *D. suzukii* does (Karageorgi et al. 2017), because they share the trait of an enlarged ovipositor and the ability to lay eggs in ripe fruits (Atallah et al. 2014). However, this assumption had not been empirically tested before this study. To test this hypothesis, we placed female flies in a cage with both a ripe and rotten whole strawberry and allowed them to lay eggs overnight. Afterward, the num-ber of eggs laid in each fruit was counted and an oviposition preference index (OPI) was calculated (fig. 1A and supple-mentary file 1, Supplementary Material online).

We recapitulated the oviposition preference results for *D. suzukii*, *D. biarmipes*, and *D, melanogaster* and found that, contrary to the above hypothesis, *D. subpulchrella* has no distinct preference for ripe or rotten fruit, displaying an intermediate oviposition behavior similar to *D. biarmipes* (OPI $= 0.196 \pm 0.138$ standard error of the mean [SEM]) (fig. 1B).

To investigate the relative roles of stiffness and chemo-sensory differentiation in driving the intermediate preference of *D. subpulchrella*, we measured the oviposition preference of a cut-ripe versus rotten strawberry for both *D. suzukii* and *D. subpulchrella* (supplementary fig. S1A, Supplementary Material online). Although both species can puncture and lay eggs in ripe fruit, *D. suzukii* can utilize a wider range of stiff egg-laying substrates as demonstrated by their ability to pierce the skin of grapes (Atallah et al. 2014). It was previously established that *D. biarmipes* and *D. melanogaster* cannot puncture ripe fruits, so we did not measure them here (Karageorgi et al. 2017). If *D. subpulchrella* oviposition site
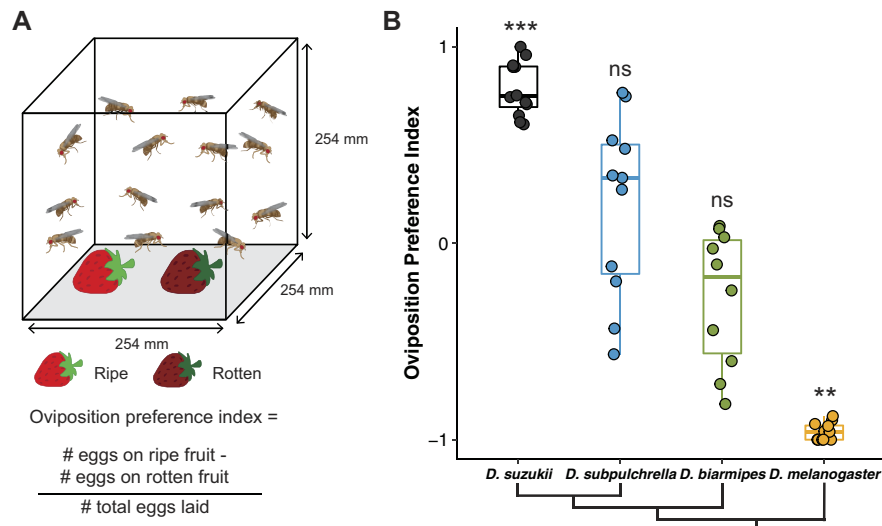
Fig. 1. The evolution of ripe fruit preference in *D. suzukii* occurred gradually, with *D. subpulchrella* and *D. biarmipes* as intermediates. (*A*) Schematic of two-choice oviposition preference assay. Ten females and five males were placed in a cage with a ripe and rotten strawberry for 19 h. Number of eggs on each fruit was then counted. (*B*) OPI of four focal species. *D. suzukii* prefers to oviposit in ripe fruit, *D. subpulchrella* and *D. biarmipes* show no preference for either fruit, and *D. melanogaster* prefers rotten fruit. Each data point represents one experimental trial, and data dispersion is represented by a boxplot. Preference *P* values were calculated using a Wilcoxon-signed rank test against a theoretical value of 0 (no preference). ***$P \leq 0.001$; **$P \leq 0.01$; *$P \leq 0.05$; ns, $P > 0.05$ for all future figures.

choice is not based on stiffness, we would expect a similar result to our whole fruit assay, in which *D. subpulchrella* equally prefers cut-ripe and rotten fruit. We found that *D. suzukii* still preferred the ripe fruit (Wilcoxon-signed rank test, $P = 0.0004$), whereas *D. subpulchrella* displayed no significant preference for the ripe or rotten fruit (Wilcoxon-signed rank test, $P = 0.06$) (supplementary fig. S1B, Supplementary Material online). However, the mean OPI for *D. subpulchrella* in this assay (OPI = 0.341 ± 0.166 SEM) was higher as compared to their whole fruit OPI (OPI = 0.196 ± 0.138 SEM), and *D. suzukii* and *D. subpulchrella* preference distributions were not significantly different from one another (pairwise Wilcoxon test, $P = 0.27$). This suggests that there are biologically meaningful differences between the cut-ripe and the whole fruit assay, pointing to both stiffness and chemosensory differentiation being involved in *D. subpulchrella* oviposition site choice. Additionally, both *D. suzukii* and *D. subpulchrella* strongly disfavor laying eggs on the exposed, white flesh of the cut fruit, further suggesting that texture differentiation is an important facet of oviposition site choice in these species (supplementary fig. S1C, Supplementary Material online).

D. subpulchrella's intermediate preference suggests that ripe fruit preference in *D. suzukii* is due to factors beyond the enlargement of the ovipositor and ability to puncture the skin of ripe fruit and evolved after the split between *D. subpulchrella* and *D. suzukii*. Further, because *D. suzukii* is an invasive pest, whereas *D. subpulchrella* is not, differences between these sister species may point to traits that have contributed to the evolution of pest behavior in *D. suzukii* and may be important for future efforts in mitigating damage from this invasive pest.

D. suzukii Displays a Preference for Stiff Oviposition Substrates, an Aversion to Acetic Acid, and Lack of Preference for Sucrose

Previous studies indicate that *D. suzukii*'s preference for ripe fruit evolved through changes in multiple sensory systems, including the olfactory, gustatory, and mechanosensory systems (Keesey et al. 2015; Ramasamy et al. 2016; Karageorgi et al. 2017). However, the evolutionary history of these sensory changes and the specific sensory cues involved in *D. suzukii*'s oviposition site choice remain unknown. To investigate these questions, we tested the oviposition preference of *D. suzukii*, *D. subpulchrella*, *D. biarmipes*, and *D. melanogaster* females for four sensory cues that change over the course of fruit maturation from ripening to rotting: ethanol, sucrose, acetic acid, and stiffness (using agarose concentration as a proxy). At early preripe and ripe maturation stages, fruit is firm and contains low levels of ethanol and total acid, both ranging from concentrations of 0% to 1% in strawberries and other fruits (Montero et al. 1996; Dudley 2004; Hidalgo et al. 2013). In rotten fruit, as fermentation and acidification occur, ethanol and acid concentrations both rise to about 7% for ethanol (Dudley 2004; Hidalgo et al. 2013) and about 3.5% for acetic acid (Hidalgo et al. 2013). Sugar content is highest in ripe fruit and differs greatly in various fruit species ranging from concentrations of about 7–20% (translating to 200 mM to 600 mM sucrose) (Montero et al. 1996; Dudley 2004; Basson et al. 2010; Hidalgo et al. 2013). We tested oviposition preference for each sensory cue at three biologically relevant levels in the context of fruit maturation (see Materials and Methods) by placing three to four females of each species in custom built egg-laying chambers and giving them the choice between two agarose egg-laying substrates: one control and one containing the experimental substrate

(fig. 2A and supplementary file 2, Supplementary Material online). Afterward, the number of eggs laid on each substrate was counted and the OPI was calculated.

When given the choice between 0.25% and 1%, 1.5%, or 2% agarose, *D. suzukii* consistently preferred the stiffer substrate (representative of early fruit maturation stages), whereas *D. subpulchrella*, *D. biarmipes*, and *D. melanogaster* preferred the softer substrate (fig. 2B and supplementary fig. S2A and supplementary file 2, Supplementary Material online, Wilcoxon-signed rank test, all *P* values <0.01). *D. suzukii*'s striking divergence in stiff substrate preference suggests that the species-specific targeting of stiff oviposition sites is associated with the early fruit stage preference only exhibited by *D. suzukii*. Further, the contrasting stiffness preference of *D. subpulchrella* and *D. suzukii* may explain why *D. suzukii* is able to target a broader range of ripe, soft-skinned fruit despite both species being able to puncture ripe fruits with their enlarged ovipositors (Atallah et al. 2014).

We next measured the oviposition preference for acetic acid and ethanol, which are both products of fermentation that increase as fruit rots and are attractive oviposition cues for *D. melanogaster* (Eisses 1997; Dudley 2004; Joseph et al. 2009; Azanchi et al. 2013; Kacsoh et al. 2013; Jouandet and Gallio 2015). Using a linear regression approach comparing the difference in preference curve *y*-intercepts, we found that overall oviposition preference for ethanol did not differ between the four focal species (fig. 2C, see Materials and Methods for details). This lack of behavioral difference despite clear differences in ripe versus rotten fruit oviposition preference may be due to the key role of yeast, which converts fruit sugars to ethanol during fermentation, in guiding natural oviposition preference. Both *D. suzukii* and *D. melanogaster* display yeast-mediated oviposition preference, with yeast being a major attractant, more so than fruit volatiles, for *D. melanogaster* (Becher et al. 2012), and *D. suzukii* exhibiting competition dependent yeast attraction, choosing to oviposit in substrates without yeast only when other species are present (Kidera and Takahashi 2020).

D. melanogaster showed a strong preference for acetic acid at each concentration tested (Wilcoxon-signed rank test, all *P* values <0.001), whereas *D. suzukii*, *D. subpulchrella*, and *D. biarmipes* avoided acetic acid at high concentrations, which are representative of late fruit maturation stages (fig. 2D and supplementary S2C and supplementary file 2, Supplementary Material online, Wilcoxon-signed rank test at 3.5% acetic acid, all *P* values <0.01). The preference for acetic acid potentially represents a unique shift in preference for *D. melanogaster*, as opposed to a loss of preference in the *D. suzukii* lineage, as other species in the *D. melanogaster* subgroup, including *D. yakuba*, *D. simulans*, and *D. mauritiana*, have a decreased tolerance of acetic acid (McKenzie and McKechnie 1979; Montooth et al. 2006). Investigating the acetic acid oviposition preference of more species within and outside of the *D. melanogaster* subgroup may help elucidate the role of acetic acid avoidance in conferring ripe versus rotten fruit preference.

Last, we tested the oviposition preference for sucrose in each of the four species. We found that while *D. melanogaster* and *D. biarmipes* have a preference for sucrose at each concentration measured (Wilcoxon-signed rank test, all *P* values <0.01), *D. suzukii* and *D. subpulchrella* show no preference or aversion to sucrose, with *D. suzukii* having no preference at each concentration measured (Wilcoxon-signed rank test, all *P* values >0.05), and *D. subpulchrella* having no preference at 400 and 600 mM sucrose (Wilcoxon-signed rank test, all *P* values >0.05) (fig. 2E and supplementary fig. S2D and supplementary file 2, Supplementary Material online). Because sugar increases during ripening, high sugar content indicates that rotting will begin imminently, perhaps explaining why *D. melanogaster* chooses high sugar substrates despite the fact that their preferred substrate of rotten fruit contains less sugar than ripe fruit. Similarly, *D. suzukii*'s indifference toward sucrose may be associated with their transition toward ovipositing on early fruit maturation stages. While in a two-choice scenario *D. suzukii* prefers ripe fruit (higher sugar) over rotten fruit (lower sugar), when given the option of earlier maturation stages, they equally target preripe and ripe fruits for egg-laying (Karageorgi et al. 2017). Sugar is still relatively low during preripe stages (Montero et al. 1996; Basson et al. 2010), suggesting that the loss of sucrose preference in *D. suzukii* could be associated with the selection of early fruit stage oviposition sites. It was previously reported that the chemical composition and contrasting stiffness of ripe and rotten fruit together explain the difference in oviposition preference in *D. suzukii*, *D. biarmipes*, and *D. melanogaster* (Karageorgi et al. 2017). Here, we have identified discrete differences in the roles of sensory cues associated with the fermentation and acetification process of fruit rotting in guiding egg-laying decisions in *D. suzukii* compared with closely related nonpest species, *D. subpulchrella* and *D. biarmipes*, as well as *D. melanogaster*.

Sequencing, Assembly, and Annotation of the *D. subpulchrella* Genome

The results from our phenotypic assays suggesting strongly anchored behavioral differences between species motivated a deeper study of the genetic factors underlying these behavioral differences. In order to apply comparative genomic and genetic approaches, we generated a near-chromosomal level assembly of the *D. subpulchrella* genome using PacBio sequencing (see Materials and Methods). The genome size is about 265 megabases (Mb) and has a contig N50 of 11.59 Mb. Specifically, the longest 6 contigs are 29.67, 29.50, 26.21, 22.17, 20.23, and 11.59 Mb, accounting to a total of 139.38 Mb (for details see supplementary table S1, Supplementary Material online). We assessed the gene content of the *D. subpulchrella* genome and found that 98.11% (2,746 out of 2,799) of the Diptera BUSCO genes are present in the genome, among which 97.43% (2,727) are complete and single-copy genes. When using the 303 BUSCO eukaryota genes, 302 complete genes were found in the genome. This suggests we have assembled a highly complete genome with relatively low levels of redundancy. We then performed three rounds of genome
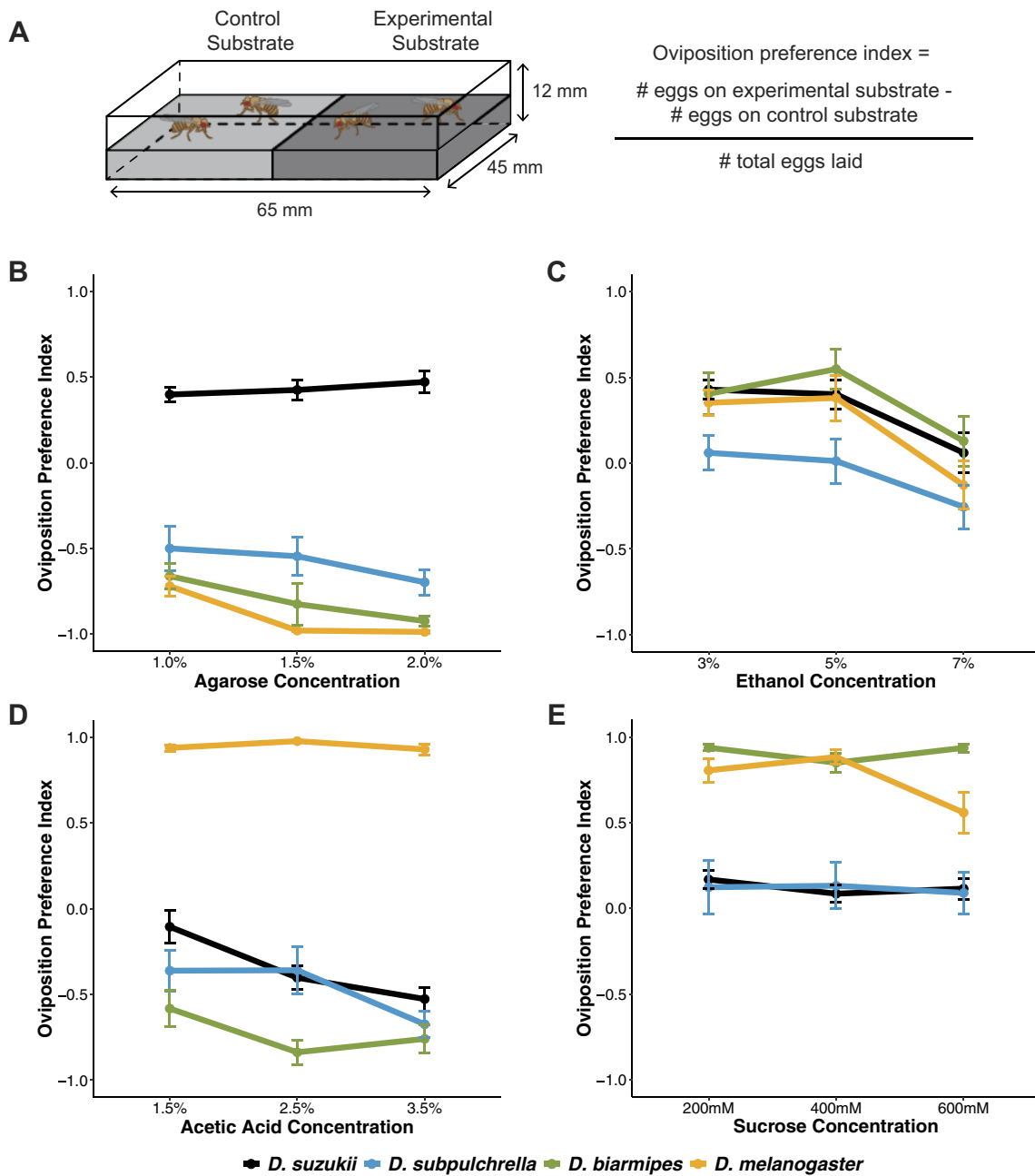
FIG. 2. Oviposition preference for substrates associated with fruit maturation differ among focal species. (*A*) Schematic of egg-laying chamber for the substrate gradient two-choice oviposition preference assay. Three to four females were placed in an arena with a choice between two substrates for 19 h. Number of eggs on each substrate was then counted. Experimental substrate was varying concentrations of either ethanol, sucrose, acetic acid, or agarose. $n \geq 15$ for each species at each concentration. *P* values calculated through linear regression analysis through pairwise comparisons of the *y*-intercepts of each species' preference curve across the three concentrations (see Materials and Methods). Data points are mean $\pm$ SEM. (*B*) Stiffness preference separated *D. suzukii* from the other three species, as it was the only species that consistently prefers stiffer oviposition substrates. There were significant differences in the preference curves of *D. suzukii* and *D. subpulchrella* (*), *D. suzukii* and *D. biarmipes* (**), and *D. suzukii* and *D. melanogaster* (***). (*C*) Ethanol oviposition preference did not differ between *D. suzukii*, *D. biarmipes*, or *D. melanogaster*, with all species showing a preference for ethanol at 3% and 5%, and a neutral response to ethanol at 7%. *D. subpulchrella* displayed a neutral response to ethanol at each concentration measured. There were no significant differences in preference curves between any two species. (*D*) *D. melanogaster* had a strong oviposition preference for acetic acid at each concentration measured, whereas *D. biarmipes*, *D. subpulchrella*, and *D. suzukii* had an aversion to acetic acid containing substrates. There were significant differences in the preference curves of *D. suzukii* and *D. biarmipes* (**), *D. suzukii* and *D. melanogaster* (***), *D. subpulchrella* and *D. melanogaster* (***), and *D. biarmipes* and *D. melanogaster* (***). (*E*) *D. suzukii* and *D. subpulchrella* did not show an aversion to or preference for sucrose at any concentration measured, whereas *D. biarmipes* and *D. melanogaster* prefer sucrose-containing substrates. There were significant differences in the preference curves of *D. suzukii* and *D. biarmipes* (***), *D. suzukii* and *D. melanogaster* (***), *D. subpulchrella* and *D. biarmipes* (**), and *D. subpulchrella* and *D. melanogaster* (**).

annotation by training the annotation program using MAKER2. In total, we annotated 15,225 genes. Among which, 13,435 genes have reciprocal best hits in *D. suzukii*. We used this gene set for downstream comparative genomic and transcriptomic analyses. This high-quality genome makes it possible to study the genomic differences and genome evolution of *D. suzukii* and *D. subpulchrella*, and how they compare with other closely related species.

The genome and assembly of *D. subpulchrella* is now available at NCBI through GCF_014743375.2 or JACOEE01. In total, 15,028 protein-coding genes were annotated, which is very similar to our MAKER2 annotation. About 29,037 transcripts were annotated with a median length of 1.9 kb, which is similar to the median transcript length of *D. melanogaster*. The median number of transcripts per gene was 1.67, and the median number of exons per transcript was 6.16. In addition, a total of 2,481 noncoding genes were annotated. We found that 23.95% of the genome was repetitive sequences using RepeatMasker, which is considerably higher than other *Drosophila* species such as *D. melanogaster* (Kaminker et al. 2002), the *simulans* species complex (Chakraborty et al. 2021), and *D. hydei* (Zhao and Begun 2017). Repetitive regions were comprised of 14.09% retroelements, 0.41% DNA transposons, and 3.01% simple repeats, among other types of repetitive sequences. Notably, 6.82% belongs to the Gypsy family, and 1.33% belongs to R1/LOA/Jockey retrotransposons; it would therefore be an interesting future direction to investigate if the telomere elongation mechanism of *D. subpulchrella* is similar to *D. biarmipes* (Saint-Leandre et al. 2019).

Adaptive Evolution of Sensory Receptors Implicated in *D. suzukii* Oviposition Site Preference

Our behavioral analyses show that changes along the *D. suzukii* lineage, in both chemosensory and mechanosensory systems, are involved in the stepwise transition to ripe fruit preference in *D. suzukii*. Therefore, to investigate the genetic changes that led to *D. suzukii*'s unique oviposition behavior, we analyzed the repertoire of *Drosophila* OR, IR, GR, and MR genes for signals of positive selection using population-level genomic data from over 200 *D. suzukii* females. Specifically, we performed a McDonald–Kreitman test (MK) to identify candidate ORs, GRs, IRs, and MRs evolving under positive selection in *D. suzukii* compared with *D. subpulchrella*, *D. biarmipes*, and *D. melanogaster*, separately (fig. 3). The MK test infers the presence of positive selection by comparing the numbers of fixed and segregating, synonymous and nonsynonymous substitutions in the genomes of a focal population (McDonald and Kreitman, 1991). Although there is previous knowledge on ORs evolving under positive selection in *D. suzukii* compared to *D. melanogaster* (Ramasamy et al. 2016), our analysis has the additional power of population level data, the additional pairwise comparisons of *D. suzukii* to both *D. biarmipes* and *D. subpulchrella*, and the addition of GRs, IRs, and MRs.

We used three independent methods to perform the MK test (see Materials and Methods) and consistently identified about 3,000 genes evolving under positive selection in *D. suzukii*, compared to the other three focal species. This number is very large compared to positively selected genes in other *Drosophila* species when inferred through whole genome and population-level genomic analyses, which range from 500 to 1,000 positively selected genes (Begun et al. 2007; Langley et al. 2012; Zhao and Begun 2017). Specifically, we found that d_N/d_S is significantly larger than p_N/p_S and that both d_N and d_S are large (supplementary fig. S3, Supplementary Material online). Since *D. suzukii* recently invaded North America, its evolutionary history suggests that the overrepresentation of positively selected genes may be caused by the bottleneck of invasion and the subsequent increase in population size and likely effective population size (Eyre-Walker, 2002; McDonald and Kreitman, 1991). Even if the number of genes under positive selection is overestimated, one would expect that there is no bias in which gene categories are enriched for signals of selection. For the majority of species pairwise comparisons between the four focal *Drosophila* species, the sensory receptors analyzed are evolving under positive selection in *D. suzukii* to a greater degree than other regions of the genome (supplementary fig. S4, Supplementary Material online). This overrepresentation for signals of positive selection in ORs, GRs, and MRs suggests that these genes are at least partly contributing to adaptive evolution in *D. suzukii*. On the other hand, there is also a likelihood of false-negative results, and sensory genes that are not significant in our MK test may still be important for adaptive evolution, as changes to a small number of amino acids can cause strong shifts in receptor affinity and specificity, a signal that would not be captured by our analysis. Thus, we focus on identifying candidate sensory genes that show expression divergence or adaptive gene evolution, rather than analyzing them as a whole.

Receptors that are under positive selection in *D. suzukii* in each of the three pairwise species comparisons may underlie *D. suzukii*'s novel oviposition behavior (fig. 3). Such genes include *Or22a*, which is involved in host choice evolution in various *Drosophilid* species (Goldman-Huertas et al. 2015; Mansourian et al. 2018; Auer et al. 2020), *Or85a*, a gene implicated in *D. sechellia*'s specialization on morinda fruit (Auer et al. 2020) that is also thought to either be a pseudogene (Hickner et al. 2016) or have changed function in *D. suzukii* (Ramasamy et al. 2016), and *Or85d*, which is linked to the detection of yeast volatiles in *D. melanogaster* (Tichy et al. 2008). Gustatory and ionotropic receptors of interest include *Gr63a*, which is involved in the detection of CO_2, a volatile emitted by ripening fruit (Jones et al. 2007; Kwon et al. 2007; Krause Pham and Ray 2015), and *Ir7a*, an ionotropic receptor linked to acetic acid consumption avoidance in *D. melanogaster* (Rimal et al. 2019). Lastly, mechanosensory receptors of note include *nan*, which is involved in substrate stiffness detection during egg-laying in *D. melanogaster* (Jeong et al. 2016; Zhang et al. 2020), *ppk*, a gene involved in acetic acid attraction during oviposition choice in *D. melanogaster* (Gou et al. 2014), and *Trpγ*, which is also linked to CO_2 detection (Badsha et al. 2012).
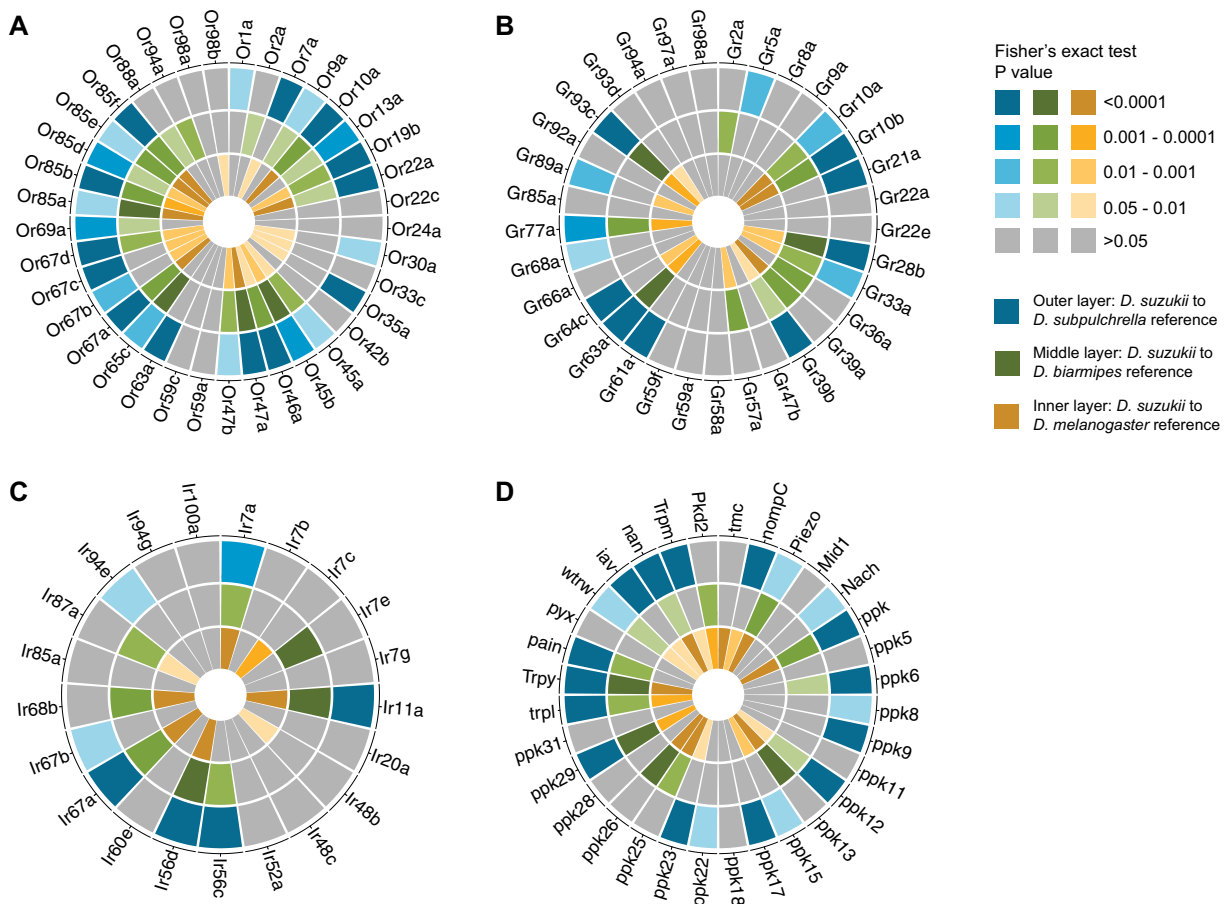
Fig. 3. MK test for sensory receptors evolving under positive selection in *D. suzukii*, as compared with *D. subpulchrella*, *D. biarmipes*, and *D. melanogaster*. Odorant receptors (A), gustatory receptors (B), divergent ionotropic receptors (C), and mechanosensory receptors (D) undergoing adaptive evolution in *D. suzukii*. MK test conducted using population data from >200 *D. suzukii* genomes. *P* values calculated using Fisher's exact test.

Differential Expression of Sensory Receptors Implicated in *D. suzukii* Oviposition Site Preference

To further understand the molecular changes associated with *D. suzukii*'s oviposition preference, we sequenced full transcriptomes and analyzed gene expression data from female adult heads of *D. suzukii*, *D. subpulchrella*, *D. biarmipes*, and *D. melanogaster*. We analyzed expression across all genes and determined differential gene expression in *D. suzukii* compared with the three other focal species. Significantly differentially expressed genes in each species comparison and the GO enrichment for these gene sets can be found in supplementary file 3, Supplementary Material online. To determine the strongest candidate genes, we focused on the same set of sensory receptors (supplementary tables S2–S5, Supplementary Material online) included in our MK analysis (supplementary table S6, Supplementary Material online), and built upon our positive selection population genomic analysis to curate a final list of genes which exhibit both significant MK tests (in all three species pairwise comparisons), and a significant difference in gene expression (in at least one species comparison). These criteria generated 15 candidate

receptor genes potentially underlying *D. suzukii*'s oviposition behavior (fig. 4). These genes include previously mentioned receptors *Or85a*, *Or85d*, *Gr63a*, and *Trpγ*, strengthening the potential role of CO_2 detection and host fruit specialization in *D. suzukii*'s ripe fruit preference. This list also includes *Ir56d*, another receptor implicated in the response to carbonation and CO_2 (Sánchez-Alcañiz et al. 2018). Other ORs implicated are *Or10a* and *Or85f*, which are both involved in response to benzaldehydes in *D. melanogaster* (Rollmann et al. 2010), a major volatile emitted by fruits during ripening (Girard and Kopp 1998). Lastly, candidate MRs include *wtrw*, which is involved in humidity sensation in *D. melanogaster* (Liu et al. 2007), *trpl*, a cation channel shown to mediate gradual dietary shifts in *D. melanogaster* (Zhang et al. 2013), and *Piezo*, which interacts with the receptor *nan*, mentioned above, to sense substrate stiffness differences in *D. melanogaster* (Zhang et al. 2020).

It is important to note that our results represent data from a single, heterogeneous tissue and cannot capture many elements of gene expression evolution along the *D. suzukii* lineage. To address this, we analyzed gene expression in the
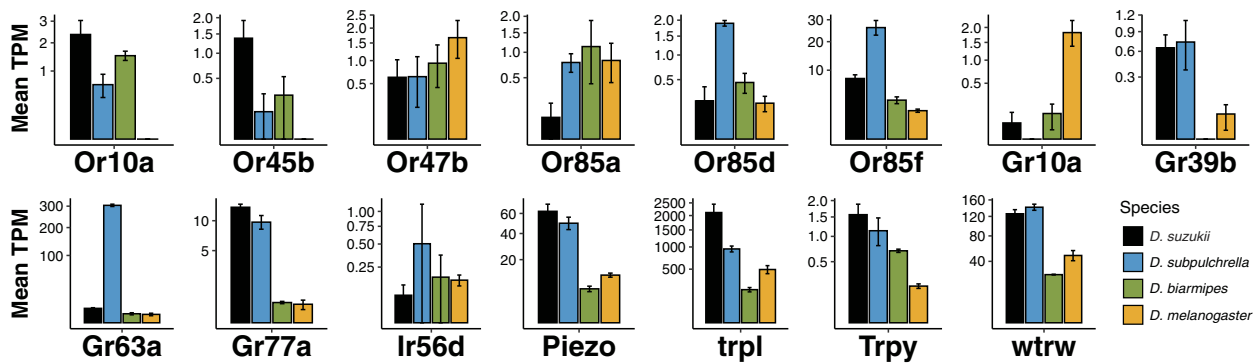
FIG. 4. Gene expression of candidate sensory receptors potentially underlying *D. suzukii*'s novel oviposition behavior. Fifteen candidate ORs, GRs, IRs, and MRs were identified as both evolving under positive selection in *D. suzukii* compared with *D. subpulchrella*, *D. biarmipes*, and *D. melanogaster* and significantly differentially expressed ($P < 0.05$) in at least one species as compared with *D. suzukii* in the gene expression analysis. The *y*-axes are mean TPM (Transcripts Per Million) normalized between all species and each *y*-axis is square root transformed. *P* values listed in supplementary tables S2–S5, Supplementary Material online, and calculated using species pairwise *t*-tests. Histogram bars are mean ± SD.

same set of sensory receptors included in our MK analysis (supplementary table S6, Supplementary Material online) using previously published transcriptome data from the abdominal tip of our four focal species (Crava et al. 2020). Overall, fewer sensory receptor genes were expressed in the abdominal tip (supplementary table S7, Supplementary Material online). However, several of our candidate genes, including *Piezo*, *nan*, *wtrw*, *trpl*, *Or85f*, and *Gr63a* were expressed at different levels among the four focal species, pointing to potentially interesting future directions. Additionally, to ascertain how well our head RNA-seq data set captures expression information from relevant sensory organs, we analyzed the correlation between our *D. suzukii* head RNA-seq data and published RNA-seq data sets from the proboscis + maxillary palp and antennae of *D. suzukii* females (Paris et al. 2020) among genes in our sensory receptor gene set (supplementary table S6, Supplementary Material online). We found that the relative expression of sensory genes in our *D. suzukii* head RNA-seq data set and in the proboscis + maxillary palp and antennae data sets were significantly correlated (Spearman's $r = 0.32$ and 0.67, respectively, $P = 0.003$ and 5.5×10^{-12}, respectively) (supplementary fig. S5, Supplementary Material online), suggesting that our sample collection method can relatively accurately measure the expression of relevant sensory genes in the context of this study.

Although we focus on behavioral and genomic changes to the peripheral sensory system here, we acknowledge that differences in oviposition choice between species could be due to changes in central-brain processing (Seeholzer et al. 2018), and this would be an interesting direction for future studies on *D. suzukii* oviposition preference. In the long-term, it is important to understand how stiffness sensation influences species-specific behavioral differences. We hope the mechanosensory genes observed in this work, such as *Piezo* and *nan*, among other genes (Zhang et al. 2020), will help shed light on the evolution of mechanosensation in this species group.

Conclusions

Here we present an investigation of the behavioral patterns, sensory modalities, genetic factors, and evolutionary forces contributing to the emergence of ripe fruit preference in *D. suzukii*, a newly invasive and rapidly spreading fruit pest. We show that *D. suzukii* differs from closely related species *D. subpulchrella*, *D. biarmipes*, and *D. melanogaster* in discrete and important ways as it pertains to oviposition preference for whole fruit, and common substances associated with fruit maturation and rotting. Compared with the other *Drosophila* species studied, *D. suzukii* prefers to oviposit in ripe fruit, displays an indifference toward sucrose, an aversion to acetic acid, and a preference for stiff oviposition substrates. The species-specific egg-laying behavior of *D. suzukii* has been shown to be associated with the enlargement and strengthening of the ovipositor allowing females to puncture stiff fruit, a trait shared by its sister species, *D. subpulchrella* (Atallah et al. 2014). Previous work established a stepwise model for the evolution of *D. suzukii* as a pest species in relation to *D. melanogaster* and *D. biarmipes*, focusing on ovipositor size and fruit stage preference (Karageorgi et al. 2017). Our work clarifies and builds upon this model with the addition of empirical fruit preference data in *D. subpulchrella*, a species representing an intermediate step toward the exploitation of ripe fruit as an egg-laying substrate, as well as an in-depth analysis of sensory cues used for the discrimination of fruit maturation stages (fig. 5).

In addition to clarifying the egg-laying preference of *D. subpulchrella*, we generated a high-quality genome for this species. This high-quality genome, together with the genomes of *D. suzukii* (Chiu et al. 2013; Paris et al. 2020), including a PacBio-based assembly (Paris et al. 2020), will benefit evolutionary studies of the *suzukii* species group. Future comparative studies focused on *D. subpulchrella* may help reveal how *D. suzukii* evolved into an invasive pest, whereas other closely related species did not. Finally,
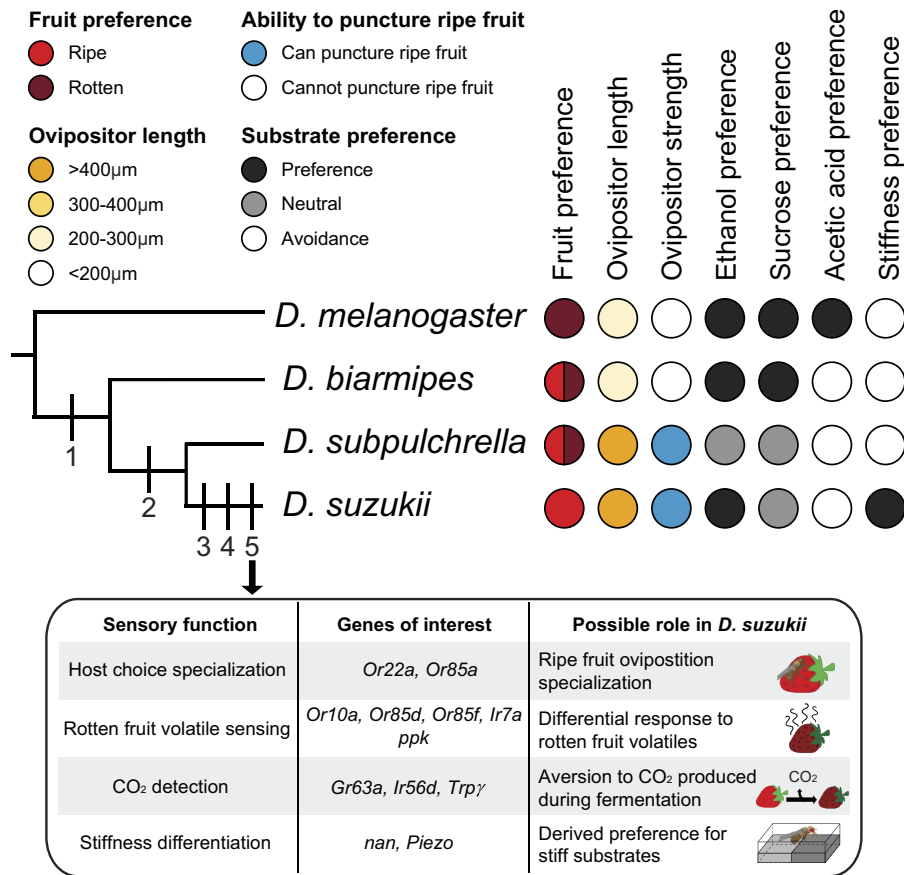
Fig. 5. Model of pest behavior evolution in *D. suzukii*. The traits acquired stepwise leading to the specialization on ripe fruit in this model are: 1. Relaxation of rotten fruit preference, 2. Enlarged ovipositor and ability to puncture ripe fruits, 3. Ripe fruit specialization, 4. Preference for stiff substrates, and 5. *D. suzukii*-specific molecular sensory changes (highlighted in the inset table). Ovipositor strength and length data from Atallah et al. (2014).

we analyzed the genomic changes associated with *D. suzukii*'s sensory evolution and generated a list of candidate ORs, IRs, GRs, and MRs with signals of both differential gene expression and positive selection (fig. 5). A substantial number of these genes seem to be evolving under positive selection and are differentially expressed in *D. suzukii* compared with the other species analyzed, despite being recently diverged.

Similar instances of rapid sensory receptor gene evolution can be seen across a wide variety of taxa during the adaptation to new ecological niches. Within the family *Drosophilidae*, sensory receptors have undergone lineage- and species-specific genomic changes driven by positive selection that are linked to ecological specialization (Guo and Kim 2007). As an example, *Or22a* has undergone several independent molecular evolution events, including instances of neuronal expansion in the morinda fruit specialist, *D. sechellia* (Dekker et al. 2006; Auer et al. 2020) and the *Pandanus* fruit specialist, *D. erecta* (Linz et al. 2013), and instances of gene deletion in the herbivorous fly, *S. flava* (Goldman-Huertas et al. 2015). Additionally, *Or22a* displays high levels of genetic differentiation among ancient populations of *D. melanogaster* specializing on marula fruit (Mansourian et al. 2018). The implication of rapid sensory receptor evolution in the adaptation to new ecological niches is pervasive beyond the family *Drosophilidae*, such as in the unique pheromone system of

orchid bees (Brand et al. 2015), the specialization of human blood feeding in *Aedes aegypti* mosquitos (McBride et al. 2014), and ecotype adaptation across phylogenetically and ecologically diverse mammals (Hayden et al. 2010). This representation across the Kingdom Animalia highlights the importance of sensory receptor gene evolution in mediating ecological shifts.

Previous work on our identified candidate genes in *D. melanogaster* and other *Drosophilid* species suggests that changes in CO_2 detection, stiffness differentiation, rotten fruit volatile sensing, and host choice specialization in part underlie *D. suzukii*'s oviposition preference for ripe fruit (fig. 5). Further work is required to truly understand the functional implications of the genetic changes seen in *D. suzukii*, and the results outlined here are a valuable resource for future studies aimed at understanding the behavioral and molecular basis of pest activity in this species.

Materials and Methods

Fly Stocks and Husbandry

All flies were reared on standard cornmeal medium at 24 °C, 55% relative humidity, on a 12-h light-dark cycle (lights on at 8:00 AM). Egg-laying experiments were conducted under the same conditions. For behavioral assays, we used a set of wild

type strains: Canton S for *D. melanogaster*, the genome strain raj3* (Chen et al. 2014) for *D. biarmipes*, #NGN6 from Japan for *D. subpulchrella* (Ehime Stock Center), and the genome strain WT3 (WT3, Chiu et al. 2013) for *D. suzukii*. We tested different lines within species and found all the behaviors are consistent. For example, in *D. melanogaster*, w ^{1118} and Canton S show the same results, suggesting that the traits tested here are likely to be fixed in each species. For genomic analyses, we used the genome reference strain of *D. melanogaster* (BDSC #2057, Adams et al. 2000), the genome strain of *D. biarmipes* (raj3*) (Chen et al. 2014), the genome strain of *D. suzukii* (WT3, Chiu et al. 2013), and our lab inbred strain of the wild caught line *D. subpulchrella* (Ehime Stock Center #NGN6).

Egg-Laying Assays

Whole Fruit Two-Choice Oviposition Assay

Flies were collected as virgins and aged for 7–8 days in food vials containing approximately ten males and ten females. For each trial, ten females and five males were placed in a mesh experimental cage (25.4 × 25.4 × 25.4 cm, BugDorm 4F2222) which contained both a whole ripe and whole rotten strawberry without anesthesia using an aspirator. Flies were allowed to lay eggs for 19 hours from late afternoon to the next morning, after which the total number of eggs laid in each fruit was counted, and an OPI was calculated as follows: (number of eggs on ripe fruit − number of eggs on rotten fruit)/(number of eggs on ripe fruit + number of eggs on rotten fruit). Ripe strawberries (always of the same variety) were purchased from a local supermarket the day of the experiment, and rotten strawberries (same variety, purchased from the same supermarket) were allowed to rot in a 24 °C, 55% relative humidity room for 4 days prior to the experiment. Only intact fruits without any damage were used in experiments. Between 10 and 12 replicate assays were performed for each species. In total, 45 assays were performed. Cut-ripe fruit versus rotten fruit assays (supplementary fig. S1, Supplementary Material online) were performed in the same way as whole-fruit choice assays, except the ripe fruit was cut in half before the trail, and the exposed flesh was placed facing up in the behavioral chamber.

Substrate Two-Choice Oviposition Assays

For all substrate choice assays, flies were collected as virgins, separated by sex, and aged separately for 3–4 days in food vials. About 2–3 days prior to the experiment, males and females were placed in a new food vial supplemented with yeast paste (1.5 ml ddH$_2$O + 1 g live active yeast) to mate and produce eggs. For each trial, three to four females were placed in a custom laser cut egg-laying chamber (see description below) containing an agarose pad of the experimental substrate and an agarose pad of the control substrate. Flies were inserted through a trap door with an aspirator without the use of anesthesia. Flies were allowed to lay eggs for 19 h, after which the total number of eggs laid on each agarose pad was counted, and the OPI was calculated as follows: (number of eggs on experimental substrate − number of eggs on control substrate)/(number of eggs on experimental substrate + number of eggs on control substrate). Only assays where flies had laid a total of ten eggs were included in the final analyses, and between 15 and 32 replicate assays were performed for each species at each concentration point. In total, 905 assays were performed.

Chamber Design. Egg-laying chambers for substrate choice assays were custom made using laser cut acrylic plastic. Each chamber contained two separate egg-laying arenas separated by 6 mm of plastic, so two trials could be conducted in one chamber. Each arena contained two wells, into which agarose substrates were poured. Wells measured 32 × 45 × 12.7 mm and were divided by 5 mm of plastic. Each arena contained a trap door through which flies could be inserted without the use of anesthesia. A 100 × 75 mm glass sheet covers the entire chamber to allow light into the chamber and to prevent flies from escaping.

Sucrose. For sucrose preference assays, experimental substrates were 1% agarose and contained 1% ethanol and increasing concentrations (200, 400, or 600 mM) of sucrose (ThermoFisher #S5-3); control substrates were 1% agarose and contained 1% ethanol.

Ethanol. For ethanol preference assays, experimental substrates were 1% agarose and contained 75 mM sucrose and increasing concentrations (3%, 5%, and 7%) of ethanol (ThermoFisher #BP2818); control substrates were 1% agarose and contained 75 mM sucrose.

Acetic Acid. For acetic acid preference assays, experimental substrates were 1% agarose and contained 75 mM sucrose, 1% ethanol and increasing concentrations (1.5%, 2.5%, and 3.5%) of acetic acid (ThermoFisher #A465-250); control substrates were 1% agarose and contained 75 mM sucrose and 1% ethanol.

Agarose. For agarose/stiffness preference assays, experimental substrates contained 75 mM sucrose, 1% ethanol, and increasing concentrations (1%, 1.5%, and 2%) of agarose (Lonza SeaKem LE Agarose #50001); control substrates were 0.25% agarose and contained 75 mM sucrose and 1% ethanol.

Statistics

All statistical analyses were performed using R (RStudio version 1.2.1335). For the whole fruit oviposition assays, a Wilcoxon-signed rank test was used with the null hypothesis set to zero, signifying no preference. For the substrate gradient preference oviposition assays, a linear regression approach was performed for each substrate using the lme4 package in R to find the overall preference difference between species across the concentrations tested. To determine if the preference curve across the three concentrations differed significantly between the four focal species, we used the lm function in the lme4 package with the response term of OPI and predictor terms of the cross of concentration and species. The reference group, the species to which the other species are compared, was manipulated using the relevel function in base R to perform pairwise comparisons between

the preference curves of each species. Effectively, this sets each of the four different species as the baseline OPI response to the concentration tested and compares this baseline to the other species analyzed. P values refer to the difference between pairwise comparisons of the y-intercepts of each species' preference curve across the three concentrations. The P values represent the level of significance for the difference between the slope of each species' regression and zero. Error bars in all figures are mean $\pm$ SEM unless otherwise noted.

Genomic Analysis

Drosophila subpulchrella Genome Library Preparation and Sequencing

D. subpulchrella inbred line generation. To generate an inbred line for PacBio sequencing, *D. subpulchrella* flies (Ehime Stock Center #NGN6) were inbred via sib-mating for ten generations to generate the strain denominated "*D. subpulchrella* 33 F10 #4."

DNA Extraction and Sequencing. We extracted DNA from adult females following the protocol of Chakraborty et al. (2016). The DNA was sheared using 20 plunges of a 21-gauge needle and size selected using the 30 kb lower cutoff size on Blue Pippin size selection system (Sage Science). Thirty kilobase SMRTbell template library was prepared from the size selected DNA and was sequenced on four SMRTcells of Pacific Biosciences Sequel I platform. We also sequenced this sheared DNA sample with 150 bp paired-end library on Illumina Hiseq 4000. All sequencing was performed at UCI GHTF.

Genome Assembly. We generated 45.3 Gb of long reads, in which 50% of the sequence is contained within sequences longer than 33.8 kb (the longest sequence is 160 kb) and 149.40 million 150 bp paired-end Illumina reads. The reads were corrected and assembled with canu v1.7 using genomeSize = 220M (Koren et al. 2017). The assembly was polished twice with arrow (smrtanalysis v5.1.0) using the long reads and twice with Pilon using the Illumina reads (Walker et al. 2014). The size of the final assembly was 265 Mb, 50% of which is contained within reads that are 11.59 Mb or longer (assembly contig N50 11.59 Mb). The genome assembly can be found on NCBI through WGS project# JACOEE01 or BioProject# PRJNA557422. We ran RepeatMasker against HMM-Dfam_3.2 lib to estimate the number of repetitive sequences in *D. subpulchrella*.

D. subpulchrella Genome Annotation

To evaluate the genome assembly quality, we used BUSCO (Simão et al. 2015) to estimate the proportion of the 2799 Diptera orthologous genes and the 303 eukaryotic genes that were completely or partially assembled in the genome. We then used the MAKER2 (Holt et al. 2011) genome annotation pipeline for genome annotation. To improve the annotation quality, we trained the Hidden Markov Model (HMM) for three times before using it for the final annotation (Zhao and Begun 2017). After that, we used OrthoMCL (Li et al. 2003) to find homologous genes between *D. subpulchrella* and the other *Drosophila* species. For multiple-copy genes, we assigned their orthologous genes by using reciprocal best hits through BlastP. The NCBI annotation can be found through GenBank assembly accession: GCA_014743375.2 and RefSeq assembly accession: GCF_014743375.2.

MK Test for Positive Selection

A high-quality population genome data set was aligned to the *D. suzukii* genome using bowtie2. After that we called bi-allelic single nucleotide polymorphisms (SNPs) using ANGSD version 0.920 (Korneliussen et al. 2014). On average, the coverage of locations with SNPs is 178. We used bi-alleles that met the following criteria: minimum mapping quality of 30, minimum allele frequency of 0.05, and a minimum coverage larger than 10. We then created alternative references using the set of SNPs. We reextracted the coding sequence of each gene from alternate references, then realigned using PRANK with the -codon function for each *D. suzukii* gene and their orthologous gene in *D. melanogaster*, *D. biarmipes*, and *D. subpulchrella*. We only carried out MK tests for genes that showed at least one variant in each of four categories, polymorphic, fixed, synonymous, and nonsynonymous (Begun et al. 2007). For genes that passed the above criteria, we carried out unpolarized MK tests using the MK.pl script (Begun et al. 2007) and a version of our own python scripts independently, followed by a Fisher's exact test for each gene within each species pairwise comparison. A third method of MK test using SNPs directly from individuals was also used for confirming the results. For each gene, we estimated the proportion of adaptive amino acid fixations (α) (Smith and Eyre-Walker 2002) and the Direction of Selection index (Stoletzki and Eyre-Walker 2011).

Lists of ORs, GRs, IRs, and MRs were compiled using FlyBase assigned gene groups for *D. melanogaster* and their orthologous genes in the other three species were inferred using OrthoMCL or BlastP. The ORs, GRs, IRs, and MRs with orthologs in each of the three species and that met the MK test criteria described above were used for genomic and transcriptomic analyses (supplementary table S6, Supplementary Material online). While we did not have an explicit protein length requirement for the MK test, short-protein sequences often do not fit the MK test criteria (must have at least one variant in each of four categories, polymorphic, fixed, synonymous, and nonsynonymous) due to lack of polymorphisms and therefore some shorter proteins are not included in our downstream analyses.

Simulations were used to compare the average degree of adaptive evolution occurring in receptor gene groups to the genome-wide average. We randomly sampled gene sets of identical size to the receptor gene set of interest for each separate species comparison (*D. suzukii*–*D. melanogaster*, *D. suzukii*–*D. biarmipes*, and *D. suzukii*–*D. subpulchrella*) and computed the mean Fisher's exact test P value for that random gene set. Simulations were run 10,000 times for each set of interest, and P values represent the proportion of the simulated distribution of means below the observed mean P value for the receptor gene set of interest.

RNA Library Preparation and Sequencing

We generated head transcriptomes from female *D. suzukii*, *D. melanogaster*, *D. biarmipes*, and *D. subpulchrella*. All individuals were mated and precisely 3 days old. Female flies were very briefly anesthetized with CO_2 and heads were collected with a clean razor blade. Dissections were performed within a 1-h window always at the same circadian zeitgeber time (ZT1-ZT2; with light turning on at 8 AM, and the dissections being performed between 9 AM and 10 AM). We collected 3 biological replicates for each species, each sample contained 15 heads. Dissected heads were immediately transferred into a low retention Eppendorf tube containing 100 µL TRIzol (Invitrogen), and RNA was extracted immediately postdissection.

All RNA extractions were performed according to TRIzol manufacturer protocol and immediately followed by a DNase treatment using the TURBO DNase from Invitrogen. RNA quality was assessed by a Bioanalyzer run of an Agilent Eukaryote Total RNA Pico chip while RNA quantity was measured with a Nanodrop One (ABI). About 1 µg total RNA was used for library preparation. Libraries were fragmented and enriched for mRNA using NEBNext Poly(A)$^+$ Magnetic Isolation Module (NEB #E7490) and prepared using NEBNext Ultra II Directional RNA Library Prep Kit (NEB #E7765) and single indexing from the NEBNext Multiplex Oligos kit (NEB #E7555L) following manufacturer protocol including beads size selection for 200 bp inserts. Library quality was first assessed on Agilent D1000 ScreenTapes for Tapestation and then by Qubit and Agilent Bioanalyzer. Finally, 150 bp paired-ends libraries were sequenced on an Illumina Nextseq500 platform.

Identification of Differentially Expressed Sensory Receptor Genes

Adaptors and low-quality bases from RNA-seq reads were trimmed using Trimmomatic (Bolger et al. 2014) using the setting LEADING : 1 TRAILING: 1 SLIDINGWINDOW : 20:25, and MINLEN : 36. Bowtie2 (Langmead and Salzberg 2012) was then used to align reads to the reference genome of the species being analyzed. Gene and transcript expression levels (transcripts per million [TPM]) were then quantified using StringTie (Pertea et al. 2015). We then obtained a list of homologous genes of *D. suzukii*, *D. biarmipes*, and *D. subpulchrella* using OrthoMCL and reciprocal best hits. To compare gene expression between replicates from different species we used trimmed mean M values (TMM) normalized TPM values. TMM normalization was implemented with the R package edgeR (Robinson et al. 2010). We then conducted a Log_2 transformation across all replicates and calculated a P value using a Student's t-test between *D. suzukii* and each of the other focal species to determine if each gene is differential expressed between two species. We specifically queried the expression patterns of sensory receptor genes inferred by homology with *D. melanogaster* sensory receptor genes extracted from FlyBase, using the same gene list as that described for the MK test above (supplementary table S6, Supplementary Material online) (Thurmond et al.

2019). GO enrichment of significantly differentially expressed genes ($P < 0.01$) was performed using PANTHER (Mi et al. 2019).

Evaluation of Sensory Receptor Gene Expression in Other Relevant Tissues

Raw RNA-seq reads from the female abdominal tip of the four focal species were downloaded from the Genbank SRA database (BioProject# PRJNA526247) (Crava et al. 2020). Raw reads of adult female proboscis + maxillary palp and female antennae for *D. suzukii* were downloaded from NCBI SRA SRR10716767 and SRR10716770, respectively (Paris et al. 2020). Reads were aligned and analyzed similar to that described above, with the only difference that no statistical analysis was performed because the data do not have biological replicates.

Correlation between the head RNA-seq data set and the maxillary palp + proboscis and antennae RNA-seq data sets was done by first ranking the expression of sensory receptor genes (supplementary table S6, Supplementary Material online) in each of the three data sets. Correlation of the rankings was calculated using the Spearman correlation coefficient in R. Genes that are not expressed were ranked as 1. We used rankings to compare these data sets because the relative raw expression levels could be influenced by technical differences that do not reflect biologically meaningful information.

Supplementary Material

Supplementary data are available at *Molecular Biology and Evolution* online.

Acknowledgments

Author Contribution

S.M.D, L.Z., and N.S. conceived the study. S.M.D performed the behavioral analyses with the help of M.K., C.B.L., and A.G. performed cut-ripe fruit behavioral experiments for the revision of the manuscript. S.M.D. generated the genome strain of *D. subpulchrella* and performed RNA sequencing library construction. M.C. and J.J.E. sequenced the *D. subpulchrella* genome. A.A., K.M.L., and J.C.C. generated the *D. suzukii* population genome data. L.Z., S.M.D., A.G., J.P., and N.S. performed the computational analysis. S.M.D. and L.Z. wrote the manuscript with the input from all authors.

Data Availability

All single-molecule sequence data and RNA-seq data have been deposited to the NCBI Sequence Read Archive (SRA) and can be found under BioProject accession PRJNA557422. The *D. subpulchrella* genome assembly has been deposited in the NCBI Assembly database under accession JACOEE01. The vcf file for the population genome data, genomes, and SNPs used in the analysis can be found in figshare: https://doi.org/10.6084/m9.figshare.c.5237618.v2 (last accessed February 24, 2021). Another copy of files for MK analysis and raw behavioral data can be found at the GitHub page https://github.com/LiZhaoLab/D.suzukii_genomes_and_analysis (last accessed February 24, 2021).

References

Adams MD, Celniker SE, Holt RA, Evans CA, Gocayne JD, Amanatides PG, Scherer SE, Li PW, Hoskins RA, Galle RF, et al. 2000. The genome sequence of *Drosophila melanogaster*. *Science* 287(5461):2185–2195.

Adrion JR, Kousathanas A, Pascual M, Burrack HJ, Haddad NM, Bergland AO, Machado H, Sackton TB, Schlenke TA, Watada M, et al. 2014. *Drosophila suzukii*: the genetic footprint of a recent, worldwide invasion. *Mol Biol Evol*. 31(12):3148–3163.

Atallah J, Teixeira L, Salazar R, Zaragoza G, Kopp A. 2014. The making of a pest: the evolution of a fruit-penetrating ovipositor in *Drosophila suzukii* and related species. *Proc R Soc B*. 281(1781):20132840.

Auer TO, Khallaf MA, Silbering AF, Zappia G, Ellis K, Álvarez-Ocaña R, Arguello JR, Hansson BS, Jefferis GSXE, Caron SJC, et al. 2020. Olfactory receptor and circuit evolution promote host specialization. *Nature* 579(7799):402–408.

Azanchi R, Kaun KR, Heberlein U. 2013. Competing dopamine neurons drive oviposition choice for ethanol in Drosophila. *Proc Natl Acad Sci U S A*. 110(52):21153–21158.

Badsha F, Kain P, Prabhakar S, Sundaram S, Padinjat R, Rodrigues V, Hasan G. 2012. Mutants in Drosophila TRPC channels reduce olfactory sensitivity to carbon dioxide. *PLoS One*. 7(11):e49848.

Basson CE, Groenewald JH, Kossmann J, Cronjé C, Bauer R. 2010. Sugar and acid-related quality attributes and enzyme activities in strawberry fruits: invertase is the main sucrose hydrolysing enzyme. *Food Chem*. 121(4):1156–1162.

Becher PG, Flick G, Rozpędowska E, Schmidt A, Hagman A, Lebreton S, Larsson MC, Hansson BS, Piškur J, Witzgall P, et al. 2012. Yeast, not fruit volatiles mediate *Drosophila melanogaster* attraction, oviposition and development. *Funct Ecol*. 26(4):822–828.

Begun DJ, Holloway AK, Stevens K, Hillier LW, Poh Y-P, Hahn MW, Nista PM, Jones CD, Kern AD, Dewey CN, et al. 2007. Population genomics: whole-genome analysis of polymorphism and divergence in *Drosophila simulans*. *PLoS Biol*. 5(11):e310.

Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30(15):2114–2120.

Brand P, Ramírez SR, Leese F, Quezada-Euan JJG, Tollrian R, Eltz T. 2015. Rapid evolution of chemosensory receptor genes in a pair of sibling species of orchid bees (Apidae: euglossini). *BMC Evol Biol*. 15:1–16.

Carroll SB. 2005. Evolution at two levels: on genes and form. *PLoS Biol*. 3(7):e245.

Chakraborty M, Baldwin-Brown JG, Long AD, Emerson JJ. 2016. Contiguous and accurate de novo assembly of metazoan genomes with modest long read coverage. *Nucleic Acids Res*. 44:147.

Chakraborty M, Chang C-H, Khost DE, Vedanayagam J, Adrion JR, Liao Y, Montooth KL, Meiklejohn CD, Larracuente AM, Emerson JJ. 2021. Evolution of genome structure in the *Drosophila simulans* species complex. *Genome Res*. Doi: 10.1101/gr.263442.120..

Chen ZX, Sturgill D, Qu J, Jiang H, Park S, Boley N, Suzuki AM, Fletcher AR, Plachetzki DC, FitzGerald PC, et al. 2014. Comparative validation of the D. melanogaster modENCODE transcriptome annotation. *Genome Res*. 24(7):1209–1223.

Chiu JC, Jiang X, Zhao L, Hamm CA, Cridland JM, Saelao P, Hamby KA, Lee EK, Kwok RS, Zhang G, et al. 2013. Genome of *Drosophila suzukii*, the spotted wing drosophila. *G3 (Bethesda)* 3:2257–2271.

Crava CM, Zanini D, Amati S, Sollai G, Crnjar R, Paoli M, Rossi-Stacconi MV, Rota-Stabelli O, Tait G, Haase A, et al. 2020. Structural and transcriptional evidence of mechanotransduction in the *Drosophila suzukii* ovipositor. *J Insect Physiol*. 125:104088.

Dekker T, Ibba I, Siju KP, Stensmyr MC, Hansson BS. 2006. Olfactory shifts parallel superspecialism for toxic fruit in *Drosophila melanogaster* sibling, *D. sechellia*. *Curr Biol*. 16(1):101–109.

Dudley R. 2004. Ethanol, fruit ripening, and the historical origins of human alcoholism in primate Frugivory 1. *Integr Comp Biol*. 44(4):315–323.

Eisses KT. 1997. The influence of 2-propanol and acetone on oviposition rate and oviposition site preference for acetic acid and ethanol of *Drosophila melanogaster*. *Behav Genet*. 27(3):171–180.

Eyre-Walker A. 2002. Changing effective population size and the McDonald-Kreitman test. *Genetics* 162(4):2017–2024.

Fraimout A, Debat V, Fellous S, Hufbauer RA, Foucaud J, Pudlo P, Marin J-M, Price DK, Cattel J, Chen X, et al. 2017. Deciphering the routes of invasion of *Drosophila suzukii* by means of ABC random forest. *Mol Biol Evol*. 34(4):980–996.

Fujii S, Yavuz A, Slone J, Jagge C, Song X, Amrein H. 2015. Drosophila sugar receptors in sweet taste perception, olfaction, and internal nutrient sensing. *Curr Biol*. 25(5):621–627.

Girard B, Kopp TG. 1998. Physicochemical characteristics of selected sweet cherry cultivars. *J Agric Food Chem*. 46(2):471–476.

Goldman-Huertas B, Mitchell RF, Lapoint RT, Faucher CP, Hildebrand JG, Whiteman NK. 2015. Evolution of herbivory in Drosophilidae linked to loss of behaviors, antennal responses, odorant receptors, and ancestral diet. *Proc Natl Acad Sci U S A*. 112(10):3026–3031.

Gou B, Liu Y, Guntur AR, Stern U, Yang CH. 2014. Mechanosensitive neurons on the internal reproductive tract contribute to egg-laying-induced acetic acid attraction in Drosophila. *Cell Rep*. 9(2):522–530.

Guo S, Kim J. 2007. Molecular evolution of Drosophila odorant receptor genes. *Mol Biol Evol*. 24(5):1198–1207.

Hayden S, Bekaert M, Crider TA, Mariani S, Murphy WJ, Teeling EC. 2010. Ecological adaptation determines functional mammalian olfactory subgenomes. *Genome Res*. 20(1):1–9.

Hickner PV, Rivaldi CL, Johnson CM, Siddappaji M, Raster GJ, Syed Z. 2016. The making of a pest: insights from the evolution of chemosensory receptor families in a pestiferous and invasive fly, *Drosophila suzukii*. *BMC Genomics* 17(1):648.

Hidalgo C, Torija MJ, Mas A, Mateo E. 2013. Effect of inoculation on strawberry fermentation and acetification processes using native strains of yeast and acetic acid bacteria. *Food Microbiol* 34:88–94.

Holt C, Yandell M, Adams M, Celniker S, Holt R, Evans C, Gocayne J, Amanatides P, Scherer S, Li P, et al. 2011. MAKER2: an annotation pipeline and genome-database management tool for second-generation genome projects. *BMC Bioinformatics* 12(1):491.

Jeong YT, Oh SM, Shim J, Seo JT, Kwon JY, Moon SJ. 2016. Mechanosensory neurons control sweet sensing in Drosophila. *Nat. Commun.* 7:1–9.

Jones FC, Grabherr MG, Chan YF, Russell P, Mauceli E, Johnson J, Swofford R, Pirun M, Zody MC, White S, et al. 2012. The genomic basis of adaptive evolution in threespine sticklebacks. *Nature.* 484(7392):55–61.

Jones WD, Cayirlioglu P, Grunwald Kadow I, Vosshall LB. 2007. Two chemosensory receptors together mediate carbon dioxide detection in Drosophila. *Nature.* 445(7123):86–90.

Joseph RM, Devineni AV, King IFG, Heberlein U. 2009. Oviposition preference for and positional avoidance of acetic acid provide a model for competing behavioral drives in Drosophila. *Proc. Natl. Acad. Sci. U. S. A.* 106(27):11352–11357.

Jouandet GC, Gallio M. 2015. Catching more flies with vinegar. *Elife.* 4:

Kacsoh BZ, Lynch ZR, Mortimer NT, Schlenke TA. 2013. Fruit flies medicate offspring after seeing parasites. *Science (80-.).* 339(6122):947–950.

Kaminker JS, Bergman CM, Kronmiller B, Carlson J, Svirskas R, Patel S, Frise E, Wheeler DA, Lewis SE, Rubin GM, et al. 2002. The transposable elements of the Drosophila melanogaster euchromatin: a genomics perspective. *Genome Biol.* 3(12):research0084.1.

Karageorgi M, Bräcker LB, Lebreton S, Minervino C, Cavey M, Siju KP, Grunwald Kadow IC, Gompel N, Prud'homme B. 2017. Evolution of Multiple Sensory Systems Drives Novel Egg-Laying Behavior in the Fruit Pest Drosophila suzukii. *Curr. Biol.* 27(6):847–853.

Keesey IW, Knaden M, Hansson BS. 2015. Olfactory specialization in Drosophila suzukii supports an ecological shift in host preference from rotten to fresh fruit. *J Chem Ecol.* 41(2):121–128.

Kidera H, Takahashi KH. 2020. Chemical cues from competitors change the oviposition preference of *Drosophila suzukii. Entomol Exp Appl.* 168(4):304–310.

Kopp A, Barmina O, Hamilton AM, Higgins L, McIntyre LM, Jones CD. 2008. Evolution of gene expression in the Drosophila olfactory system. *Mol. Biol. Evol.* 25(6):1081–1092.

Koren S, Walenz BP, Berlin K, Miller JR, Bergman NH, Phillippy AM. 2017. Canu: scalable and accurate long-read assembly via adaptive κ-mer weighting and repeat separation. *Genome Res.* 27(5):722–736.

Korneliussen TS, Albrechtsen A, Nielsen R. 2014. ANGSD: analysis of Next Generation Sequencing Data. *BMC Bioinformatics.* 15(1):356.

Krause Pham C, Ray A. 2015. Conservation of Olfactory Avoidance in Drosophila Species and Identification of Repellents for Drosophila suzukii. *Sci. Rep.* 5:

Kwon JY, Dahanukar A, Weiss LA, Carlson JR. 2007. The molecular basis of CO2 reception in Drosophila. *Proc. Natl. Acad. Sci. U. S. A.* 104(9):3574–3578.

Langley CH, Stevens K, Cardeno C, Lee YCG, Schrider DR, Pool JE, Langley SA, Suarez C, Corbett-Detig RB, Kolaczkowski B, et al. 2012. Genomic Variation in Natural Populations of Drosophila melanogaster. *Genetics.* 192(2):533–598.

Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. *Nat Methods* . 9(4):357–359.

Li L, Stoeckert CJ, Roos DS. 2003. OrthoMCL: identification of ortholog groups for eukaryotic genomes. *Genome Res.* 13(9):2178–2189.

Linz J, Baschwitz A, Strutz A, Dweck HKM, Sachse S, Hansson BS, Stensmyr MC. 2013. Host plant-driven sensory specialization in *Drosophila erecta. Proc R Soc B.* 280(1760):20130626.

Liu L, Li Y, Wang R, Yin C, Dong Q, Hing H, Kim C, Welsh MJ. 2007. Drosophila hygrosensation requires the TRP channels water witch and nanchung. *Nature.* 450(7167):294–298.

Mansourian S, Enjin A, Jirle EV, Ramesh V, Rehermann G, Becher PG, Pool JE, Stensmyr MC. 2018. Wild African Drosophila melanogaster Are Seasonal Specialists on Marula Fruit. *Curr. Biol.* 28(24):3960–3968. e3.

Mayr E. 1960. The emergence of evolutionary novelties. In: Evolution after Darwin. Chicago: University of Chicago Press. p. 349–380.

McBride CS, Baier F, Omondi AB, Spitzer SA, Lutomiah J, Sang R, Ignell R, Vosshall LB. 2014. Evolution of mosquito preference for humans linked to an odorant receptor. *Nature.* 515(7526):222–227.

McDonald JH, Kreitman M. 1991. Adaptive protein evolution at the Adh locus in Drosophila. *Nature.* 351(6328):652–654.

McKenzie JA, McKechnie SW. 1979. A comparative study of resource utilization in natural populations of Drosophila melanogaster and D. simulans. *Oecologia.* 40(3):299–309.

Mi H, Muruganujan A, Huang X, Ebert D, Mills C, Guo X, Thomas PD. 2019. Protocol Update for large-scale genome and gene function analysis with the PANTHER classification system (v.14.0). *Nat Protoc.* 14(3):703–721.

Moczek AP. 2008. On the origins of novelty in development and evolution. *BioEssays.* 30(5):432–447.

Montero TM, Mollá EM, Esteban RM, López-Andréu FJ. 1996. Quality attributes of strawberry during ripening. *Sci. Hortic. (Amsterdam).* 65(4):239–250.

Montooth KL, Siebenthall KT, Clark AG. 2006. Membrane lipid physiology and toxin catabolism underlie ethanol and acetic acid tolerance in Drosophila melanogaster. *J. Exp. Biol.* 209(19):3837–3850.

Muller GB, Wagner GP. 1991. Novelty in Evolution: restructuring the Concept. *Annu Rev Ecol Syst.* 22(1):229–256.

Ometto L, Cestaro A, Ramasamy S, Grassi A, Revadi S, Siozios S, Moretto M, Fontana P, Varotto C, Pisani D, et al. 2013. Linking genomics and ecology to investigate the complex evolution of an invasive Drosophila pest. *Genome Biol. Evol.* 5(4):745–757.

Paris M, Boyer R, Jaenichen R, Wolf J, Karageorgi M, Green J, Cagnon M, Parinello H, Estoup A, Gautier M, et al. 2020. Near-chromosome level genome assembly of the fruit pest Drosophila suzukii using long-read sequencing. *Sci Rep.* 10(1):11227.

Pertea M, Pertea GM, Antonescu CM, Chang T-C, Mendell JT, Salzberg SL. 2015. StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. *Nat Biotechnol.* 33(3):290–295.

Phifer-Rixey M, Bi K, Ferris KG, Sheehan MJ, Lin D, Mack KL, Keeble SM, Suzuki TA, Good JM, Nachman MW. 2018. The genomic basis of environmental adaptation in house mice. *PLOS Genet.* 14(9):e1007672.

Ramasamy S, Ometto L, Crava CM, Revadi S, Kaur R, Horner DS, Pisani D, Dekker T, Anfora G, Rota-Stabelli O. 2016. The Evolution of Olfactory Gene Families in Drosophila and the Genomic Basis of chemical-Ecological Adaptation in Drosophila suzukii. *Genome Biol Evol.* 8(8):2297–2311.

Rimal S, Sang J, Poudel S, Thakur D, Montell C, Lee Y. 2019. Mechanism of Acetic Acid Gustatory Repulsion in Drosophila. *Cell Rep.* 26(6):1432–1442. e4.

Robinson MD, McCarthy DJ, Smyth GK. 2010. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics.* 26(1):139–140.

Rollmann SM, Wang P, Date P, West SA, Mackay TFC, Anholt RRH. 2010. Odorant receptor polymorphisms and natural variation in olfactory behavior in Drosophila melanogaster. *Genetics.* 186(2):687–697.

Rota-Stabelli O, Blaxter M, Anfora G. 2013. Drosophila suzukii. *Curr Biol.* 23(1):R8–R9.

Saint-Leandre B, Nguyen SC, Levine MT. 2019. Diversification and collapse of a telomere elongation mechanism. *Genome Res.* 29(6):920–931.

Sánchez-Alcañiz JA, Silbering AF, Croset V, Zappia G, Sivasubramaniam AK, Abuin L, Sahai SY, Münch D, Steck K, Auer TO, et al. 2018. An expression atlas of variant ionotropic glutamate receptors identifies a molecular basis of carbonation sensing. *Nat Commun.* 9:1–14.

Seeholzer LF, Seppo M, Stern DL, Ruta V. 2018. Evolution of a central neural circuit underlies Drosophila mate preferences. *Nature* 559(7715):564–569.

Simão FA, Waterhouse RM, Ioannidis P, Kriventseva EV, Zdobnov EM. 2015. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics* 31(19):3210–3212.

Smith NGC, Eyre-Walker A. 2002. Adaptive protein evolution in Drosophila. *Nature* 415(6875):1022–1024.

Stoletzki N, Eyre-Walker A. 2011. Estimation of the neutrality index. *Mol Biol Evol.* 28(1):63–70.

Suvorov A, Kim BY, Wang J, Armstrong EE, Peede D, D'Agostino ERR, Price DK, Wadell P, Lang M, Courtier-Orgogozo V, et al. 2020. Widespread introgression across a phylogeny of 155 Drosophila genomes. *BioRxiv*. doi: 10.1101/2020.12.14.422758.

Thurmond J, Goodman JL, Strelets VB, Attrill H, Gramates LS, Marygold SJ, Matthews BB, Millburn G, Antonazzo G, Trovisco V, et al. 2019. FlyBase 2.0: the next generation. *Nucleic Acids Res.* 47(D1):D759–D765.

Tichy AL, Ray A, Carlson JR. 2008. A new Drosophila POU gene, pdm3, acts in odor receptor expression and axon targeting of olfactory neurons. *J Neurosci.* 28(28):7121–7129.

Vosshall LB, Stocker RF. 2007. Molecular Architecture of Smell and Taste in Drosophila. *Annu Rev Neurosci.* 30(1):505–533.

Walker BJ, Abeel T, Shea T, Priest M, Abouelliel A, Sakthikumar S, Cuomo CA, Zeng Q, Wortman J, Young SK, et al. 2014. Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. *PLoS One* 9(11):e112963.

Wray GA. 2007. The evolutionary significance of cis-regulatory mutations. *Nat Rev Genet.* 8(3):206–216.

Zhang L, Yu J, Guo X, Wei J, Liu T, Zhang W. 2020. Parallel mechanosensory pathways direct oviposition decision-making in Drosophila. *Curr Biol.* 30(16):3075–3088.e4.

Zhang YV, Aikin TJ, Li Z, Montell C. 2016. The basis of food texture sensation in Drosophila. *Neuron* 91(4):863–877.

Zhang YV, Raghuwanshi RP, Shen WL, Montell C. 2013. Food experience-induced taste desensitization modulated by the Drosophila TRPL channel. *Nat Neurosci.* 16(10):1468–1476.

Zhao L, Begun DJ. 2017. Genomics of parallel adaptation at two timescales in Drosophila. *PLOS Genet.* 13(10):e1007016.