GBE

Genome Biology and Evolution

SMBE

The Evolution of Olfactory Gene Families in *Drosophila* and the Genomic Basis of chemical-Ecological Adaptation in *Drosophila suzukii*

Sukanya Ramasamy[1,2], Lino Ometto[1], Cristina M. Crava[1], Santosh Revadi[1,3], Rupinder Kaur[1], David S. Horner[2], Davide Pisani[4], Teun Dekker[3], Gianfranco Anfora[1] and Omar Rota-Stabelli[1,*]

[1]Agricultural Entomology Unit, Research and Innovation Centre, Fondazione Edmund Mach, San Michele all'Adige, Italy

[2]Diparimento di Bioscienze, Università degli Studi di Milano, Milano, Italy

[3]Chemical Ecology Unit, Swedish University of Agricultural Sciences, Alnarp, Sweden

[4]School of Biological Sciences and School of Earth Sciences, University of Bristol, Bristol, UK

*Corresponding author: E-mail: omar.rota@fmach.it.

Accepted: June 22, 2016

Abstract

How the evolution of olfactory genes correlates with adaption to new ecological niches is still a debated topic. We explored this issue in *Drosophila suzukii*, an emerging model that reproduces on fresh fruit rather than in fermenting substrates like most other *Drosophila*. We first annotated the repertoire of odorant receptors (ORs), odorant binding proteins (OBPs), and antennal ionotropic receptors (aIRs) in the genomes of two strains of *D. suzukii* and of its close relative *Drosophila biarmipes*. We then analyzed these genes on the phylogeny of 14 *Drosophila* species: whereas ORs and OBPs are characterized by higher turnover rates in some lineages including *D. suzukii*, aIRs are conserved throughout the genus. *Drosophila suzukii* is further characterized by a non-random distribution of OR turnover on the gene phylogeny, consistent with a change in selective pressures. In *D. suzukii*, we found duplications and signs of positive selection in ORs with affinity for short-chain esters, and loss of function of ORs with affinity for volatiles produced during fermentation. These receptors—*Or85a* and *Or22a*—are characterized by divergent alleles in the European and American genomes, and we hypothesize that they may have been replaced by some of the duplicated ORs in corresponding neurons, a hypothesis reciprocally confirmed by electrophysiological recordings. Our study quantifies the evolution of olfactory genes in *Drosophila* and reveals an array of genomic events that can be associated with the ecological adaptations of *D. suzukii*.

Key words: odorant receptors, adaptation, *Drosophila suzukii*, comparative genomics.

Introduction

The Importance of the Olfactory System in Insect Evolution and Pest Management

Olfaction has a fundamental role in animal behavior and is one of the key players in niche specialization (Sanchez-Gracia et al. 2009). For this reason, targeting chemosensation, for example, by means of repellents or attractants, is important for understanding and controlling insect populations (Heuskin et al. 2011). Insect olfaction is mediated at the periphery level by an array of olfactory proteins, including odorant receptors (ORs), a sub-family of antennal expressed ionotropic receptors (aIRs), and their associated odorant binding proteins (OBPs) (Benton et al. 2009; Sanchez-Gracia et al. 2009). Insect ORs, which are not homologs to vertebrate ones, are expressed in specialized neurons that extend into sensilla on the antennae and on the maxillary palp; air-borne volatiles enter through pores present in the sensilla (Steinbrecht 1997) and most likely bind a specific OBP in the extra-cellular aqueous lumen. Then volatiles enter in contact with the surface of dendrites (Vogt and Riddiford 1981), where they bind to an OR (or aIR), resulting in a conformational change in the OR-coreceptor (ORCO) heterodimer, and cause opening of its ion channels, membrane depolarization, and a neuronal response (Sato et al. 2008; Wicher et al. 2008).

The genomic basis of olfaction has been widely studied in insects, particularly in the 12 annotated *Drosophila* genomes (Robertson et al. 2003; Vieira et al. 2007; Gardiner et al. 2008;

Sanchez-Gracia et al. 2009; Robertson 2009). Previous studies, mostly quantitative, indicated that OR and OBP gene families evolve according to a typical birth-and-death process with random lineage-specific duplications and losses, whose fate will be determined by genetic drift and selection (Guo and Kim 2007; McBride and Arguello 2007; Vieira and Rozas 2011). However, how the evolution of these genes correlates with actual adaption to new ecological niches is a mostly unexplored topic. Current examples of the role of ORs in the adaptation of Drosophilids include *Drosophila sechellia*, which oviposits only on morinda fruits (*Morinda citrifolia* L.), and *Scaptomyza flava*, which has leaf-miner larvae: in both species, the ecological switch has been correlated to shifts in the role of *Or22a* (Dekker et al. 2006; Goldman-Huertas et al. 2015).

Drosophila suzukii, an Emerging Evolutionary Model and Pest

Drosophila suzukii Matsumura (Diptera: Drosophilidae) is a fruit fly native to Southeast Asia that has recently invaded America and Europe (Rota-Stabelli et al. 2013; Asplen et al. 2015). Whereas most *Drosophila* species are attracted to, and oviposit in, fermenting fruits, *D. suzukii* uses a typical serrated ovipositor to pierce the skin of ripening soft fruits and oviposit in them. Larvae feed on fruit pulp, promoting yeast/bacterial secondary infections and causing serious economic losses to the American and European soft-fruit production (Goodhue et al. 2011; Walsh et al. 2011; Calabria et al. 2012; Cini et al. 2012). To limit such damage, *D. suzukii* populations are mainly suppressed using pesticides; this causes environmental and health concerns because treatments are performed close to harvest, with a consequently high risk of chemical residual on fruits. Therefore, the current research management agenda (Asplen et al. 2015) includes scrutinizing the neurophysiology, genetics, genomics, metagenomics, and behavior of *D. suzukii* in search of potential targets for use in integrated pest management strategies.

The shift in preference for ripe fruits in *D. suzukii* also offers a unique possibility for comparative evolutionary studies on the adaptive origin of new ecological and behavioral traits. Throughout the past decades, *Drosophila* proved to be an excellent model organism for olfactory studies (Dekker et al. 2006; Ibba et al. 2010; McBride 2007). Although it is unknown whether ancestral *Drosophila* species had a preference for fermenting and rotting resources (Begon 1982), today such a preference predominates in the Sophophora subgenus, to which *D. suzukii* belongs. The group includes several species (fig 1) for which a wealth of genetic, genomic, neurobiological, and physiological resources are available, facilitating comparative genomic studies and the interpretation of evolutionary analyses (Ometto et al. 2013; Dekker et al. 2015; Rossi-Stacconi et al. 2016). Work is ongoing to understand how *D. suzukii* is attracted (Landolt et al. 2012; Keesey et al. 2015; Revadi et al. 2015; Scheidler et al. 2015) or repelled (Krause Pham and Ray 2015) by specific odors compared with its sister species: the genetic basis of these (and likely other still undetected) chemo-ecological differences are however almost totally unexplored and only a handful of ORs have been functionally annotated (Revadi et al. 2015).

Aim of the Study

The aim of our study is to identify key chemosensory genes that accompanied the move of *D. suzukii* into a new ecological niche. We mined the genome and annotated the entire olfactory repertoire (ORs, aIRs, and OBPs) of two *D. suzukii* strains (Italian and American) and of the closely related species, *Drosophila biarmipes*. We studied these genes within a phylogenetic framework of 14 *Drosophila*, discriminating and characterizing the genomic events, the genetic changes, and the selective forces that occurred during the evolutionary history of the genus, particularly *D. suzukii*. In addition, we coupled these results with *ad hoc* physiological experiments to confirm the functional and likely adaptive role of some of the genomic events. Our results not only cast new light on the molecular basis of adaptation in *D. suzukii*, but also provide an updated look at the evolution of odorant genes in *Drosophila*.

Materials and Methods

Identification, Annotation, and Nomenclature of Chemosensory Repertoire in *D. suzukii* and *D. biarmipes*

We extracted the complete set of OR, OBP, and aIR protein sequences using two different strategies. In a first approach, we used an automatic *de novo* gene prediction in the *D. suzukii* and *D. biarmipes* genomes (accession number CAKG00000000.1 (Ometto et al. 2013) and AFFD00000000.2, respectively) using AUGUSTUS (Stanke and Waack 2003). We then queried the predicted proteomes using orthologs of the three gene families from all 12 *Drosophila* genomes from FlyBase (Drysdale et al. 2005) using iterated PSI-BLAST (Altschul et al. 1997) searches with an e-value cut-off of 10^{-5} for homology assignment. The first hit was labeled as the putative ortholog, whereas other hits, when present, were labeled as putative paralogs. In a second, more manual-based, approach, we directly searched the complete set of *D. melanogaster* OR, OBP, and aIR protein sequences against *D. suzukii* and *D. biarmipes* genomes, using TBLASTN (Altschul et al. 1997) with an e-value cut-off of 10^{-5}. Scaffolds that passed this threshold were extracted and exons were mapped onto the protein query to manually reconstruct the orthologous coding sequence (CDS). As in the first approach, the first hit was labeled as the putative ortholog, whereas other possible hits were considered putative paralogs. To assess orthology, we studied the distribution of genes on a six-species gene phylogeny (described in the

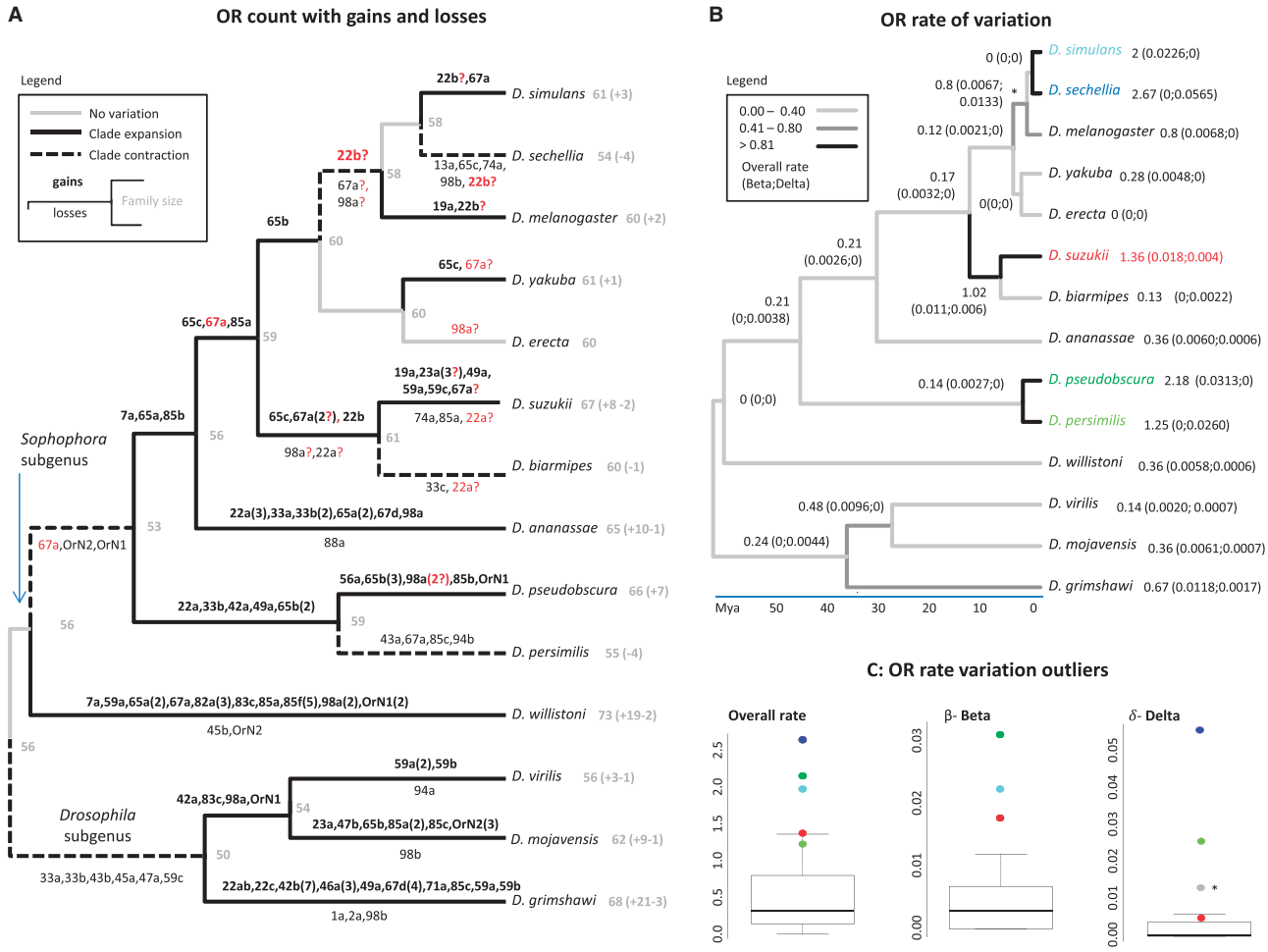

Fig. 1.—Evolution of ORs on the *Drosophila* phylogeny. A: Distribution of gene gains (above branches, in bold) and losses (below branches) on a cladogram depicting phylogeny of 14 *Drosophila* species; values at the right of each terminal or internal nodes are the number of genes calculated by BadiRate using BDI-FR-CML model. Alternative positions of genes for which there is no reconciliation of gene trees (supplementary fig S2, Supplementary Material online) with BadiRate distribution are highlighted in red. B: Distribution of the gene family size rate variation mapped on a time-tree. Each branch in the tree has overall rate of variation (rate of gain + rate of loss/(divergence times)) followed by the beta (β) and delta (δ) parameters describing, respectively, birth and death rates from the BadiRate analysis. β and δ values are rounded at the fifth integer. C: boxplots of overall rate of variation, beta, and delta.

Gene trees section, see below): using a threshold rule of bootstrap > 70, we could assess the orthology of genes previously labeled as paralogs, and define lineage-specific duplications.

We performed an exhaustive search by using a recursive approach that required various rounds of BLAST searches, annotation, alignment, and gene-trees inspection. To further verify whether missing hits in either *D. suzukii* or *D. biarmipes* genomes were false-negatives, we further used the HMMER package (Mistry et al. 2013) (v3.1b1; threshold of 10^{-5}) to perform an exhaustive search against the PFAM protein libraries (Punta et al. 2012) of chemosensory receptor (7TM_7) family (PF02949) for ORs, and of PBP/GOBP family (PF01395) for OBPs. A further validation of putative missing/incomplete genes was performed by searching the chemosensory proteins against the trace archives of the raw unassembled *D. biarmipes* and *D. suzukii* NGS data using MegaBlast (with an e-value cut-off of 10^{-10}). We also checked for possible false-positives in *D. suzukii* as a consequence of intraspecific allelic variations by doing re-blast and cross-checking the results against the *D. suzukii* American genome (Accession AWUT00000000.1, see (Chiu et al. 2013)). In few cases (supplementary table S1, Supplementary Material online), we could recover only incomplete *D. suzukii* genes, likely because of the shorter length of the scaffolds in the genome of the Italian strain compared with the American one. In these cases, we used the American genome assembly as a database to retrieve full-length gene sequences. In few other cases, genes had more than a copy in the Italian *D. suzukii* genome, but such copies were extremely similar, with only few SNPs at synonymous sites; a cross-check with

the American genome retrieved only one copy, suggesting that they were allelic variants still segregating in the Italian strain (which had been sequenced at the only partially homozygous inbred generation F5; Ometto et al. 2013).

For orthologous genes, we followed the nomenclature of the *D. melanogaster* receptor, adding a two-letter prefix corresponding to the species' name. For example, *DmOr10a* corresponds to the olfactory receptor 10a of *D. melanogaster*, and *DbObp1* to the *D. biarmipes* odorant binding protein 1. Paralogs (i.e. duplicates) in *D. suzukii* and *D. biarmipes* were named with consecutive numbers: for example, *Or67a* has 5 copies in *D. suzukii*, which are named as *DsOr67a1*, *DsOr67a2*, *DsOr67a3*, *DsOr67a4*, and *DsOr67a5*. Annotated CDS are presented in dataset S1 (Supplementary Material online).

Gene Trees

To help in the gene annotation process and to understand the evolutionary history of the genes, we constructed six-species gene phylogenies for each of the 3 gene families (OR, OBP, and aIR). The gene and protein sequences for *Drosophila erecta, Drosophila ananassae, and Drosophila pseudoobscura* were downloaded from FlyBase (Drysdale et al. 2005) and their orthologous relationships predicted by OrthoDB (Waterhouse et al. 2013). We then added the orthologs of *D. melanogaster*, *D. suzukii*, and *D. biarmipes*, and built multiple sequence alignments at both nucleotide and protein level for each of the 3 families with MUSCLE (Edgar 2004) using TranslatorX (Abascal et al. 2010); we did not use PRANK (Löytynoja and Goldman 2005) because of its unpermissive computational cost for this data set. The resulting alignments were manually checked and edited to avoid possible misleading signals in the phylogeny. In case of frame shifts (e.g., *Or22a, Or85a, Or74a*), we restored the coding frame by adding an appropriate number of single-base insertions. Phylogenies were inferred in a maximum likelihood framework using RAxML version 7.2.8 (Stamatakis 2014), bootstrapping the data set with 100 pseudo-replicates, and using protein sequences with a PROTGAMMA + LG + F model (which has been shown to significantly fit a variety of protein families better than other empirical replacement models (Le and Gascuel 2008)).

Species Trees

We mapped the evolution of each of the 3 gene families on a 14 *Drosophila*-species cladogram using the tree topology proposed by Ometto et al. (2013). We estimated the gene family size at each internal node and the family turnover rates for each branch using stochastic models implemented in BadiRate version 1.35 (Librado et al. 2011). The program uses the information of the divergence time and the number of genes in the extant species to model changes in gene family size along the phylogenetic tree: divergence times were taken from Ometto et al. (2013), whereas the data matrix of extant genes for 12 species was inferred from Gardiner et al. (2008) and Vieira et al. (2007), and, for *D. suzukii* and *D. biarmipes*, from the present study (supplementary table S2, Supplementary Material online). For BadiRate calculations, we used the BDI-FR-CML model, where a maximum likelihood model that assumes independent evolution along each branch is used to calculate the probability of a gene family to have a given size at each internal node. To define which genes had been gained/lost at each node of the *Drosophila* phylogeny, we manually mapped the gene information from the data matrix onto our phylogenetic framework. We also evaluated the overall rate of evolution on the time tree as Rate of Expansion + Contraction = (No. of gene gains + No. of gene losses)/Divergence time (to the Present in mya).

Molecular Evolution Analyses

We aligned orthologous gene sequences of *D. suzukii*, *D. biarmipes*, *D. melanogaster*, *D. erecta*, and *D. ananassae* with PRANK (Löytynoja and Goldman 2005), without providing a guide tree, using the tool TranslatorX (Abascal et al. 2010). When performing test for positive selection, we used both the raw alignment and one in which we removed regions of high complexity to minimize false signals of rapid evolution. For the latter approach, we removed gaps using Gblocks (Castresana 2000) and then used a custom perl script to remove problematic alignment regions using an approach similar to that proposed by Han et al. (2009). We translated the alignment and flagged, in each sequence *i*, those portions of length ≥ 5 amino acids with more than $f_i \times 60\%$ differences at the amino acid level and $f_i \times 50\%$ at the nucleotide level compared with the consensus. The parameter *f* was used to adjust for the species-specific divergence and was set to $f = 0.6$ for the orthologs of *D. suzukii*, *D. biarmipes*, and *D. melanogaster*; $f = 0.8$ for *D. erecta*; and $f = 1$ for *D. ananassae*. When needed, amino acids conserved across orthologs were de-flagged if at the edge of such portion. Finally, we removed the portions of the alignment flagged in at least one sequence. On average (standard deviation), this approach removed 67.2 (75.6) amino acids from the alignments, corresponding to 7.2 (7.3) % of the original alignment length. We used PAML 4.7 (Yang 2007) to estimate the rate of non-synonymous, *d*N, and synonymous substitution, *d*S, using the "free-ratio" model, which allows branch-specific values for $\omega = d\mathrm{N}/d\mathrm{S}$ over all branches of the unrooted phylogenetic tree. In case of duplications in *D. suzukii*, the analysis was done for each paralog separately. In case of paralogs in the other species, we first estimated the maximum likelihood best tree and then retained the paralog(s) with the shortest branch length and/or that was closer to the other orthologs. We tested for different selective regimes and positive selection in each gene (and, for *D. suzukii*, in each paralog) using two-codon-substitution model-based tests. In the branch test, we

compared the likelihood of a model that assumed a single ω across branches (model = 0 and NSsites = 0) with a second that assumed two ω values, one for the *D. suzukii* branch and one for the rest of the tree (model = 2 and NSsites = 0). In the branch-site test, we explicitly tested the occurrence of positive selection affecting sites along the *D. suzukii* branch (branch-site model A, test 2; model = 2 and NSsites = 2; null model has parameters fix_omega = 1, omega = 1; the positive selection model fix_omega = 0, omega = 1). In both tests, the value of twice the difference between the two alternative likelihoods ($2\Delta\lambda$) was tested using a χ^2 test with 1 degree of freedom. To account for multiple testing, we estimated the false discovery rate (FDR) of each test using the qvalue (Storey 2002) package implemented in R (R Development Core Team 2009).

Ligands Response Screening

We mined the literature and the DoOR (v.2) database (Muench and Galizia 2015) for the chemicals that elicit electrophysiological response in (or are associated with) ORs, OBPs, and aIRs in *D. melanogaster* or closely related species. We were interested in evaluating whether certain chemicals or chemical classes were over- or under-represented among those eliciting a response in the ORs duplicated (or lost) in *D. suzukii*. We therefore developed a quantitative screening: from each of the ORs listed in DoOR (v.2), we selected up to a maximum of 15 ligands eliciting them over a modeled response threshold of 0.3 (on a normalized scale from -1 to 1) and counted their occurrence (L) in the 10 ORs that experienced duplication or loss in *D. suzukii* (ORS) and in the remaining 27 ORs for which DoOR (v.2) provides accurate response data (ORR). These values were then compared by measuring the skew index $S = (L_{ORS}/10\text{-}L_{ORR}/27)/(L_{ORS}/10 + L_{ORR}/27)$, which takes values between -1 (ligands bind only to ORR) and 1 (ligands bind only to ORS).

Or85a Population Screening

To confirm the presence of two different *Or85a* alleles in *D. suzukii*, we performed PCR analysis on DNA extracted from an Italian (from Trentino, reared in our lab) and a North American population (from Oregon, provided by Dr. Vaughn Walton, Oregon State University, USA). Both populations were reared under controlled standard laboratory conditions. We designed two allele-specific forward primers (F85a.1 and F85a.2) with a single reverse primer (R85a) shared by both alleles (supplementary table S3, Supplementary Material online). F85a.1 was designed on a region present on the Italian strain, and missing in the American genome. F85a.2 was designed to confirm the missing first transmembrane domain in the American genome and to confirm the size-specific variant presence in both of the strains. This primer is covered to confirm the presence of upstream of the missing 5'UTR in American genome. PCR analysis was done in 20 µl reaction mixture using 1 µl of DNA template, 0.5 µl of 10uM primers, and GoTaq® Green Master Mix (Promega) under following conditions: a denaturing step at 95 °C for 5 min, followed by 35 cycles (95 °C for 30 s, 50 °C for 30 s, 72 °C for 30 s), and a final step of 7 min at 72 °C. PCR products were electrophoresed on a 2% agarose gel, stained with ethidium bromide, and visualized under UV light.

Single-Sensillum Recordings

We conducted all experiments on wild strains of *D. melanogaster* and *D. suzukii* collected in Trento Province (Italy), and reared on a semi-artificial diet (https://stockcenter.ucsd.edu/info/food_cornmeal.php, last accessed July 15, 2016) at 23-25°C, 65 ± 5% relative humidity (R.H.), and 16L:8D photoperiod. Flies were gently blown head first into a cut pipette tip so that the head protruded from the narrow end. The pipette tip was placed on a wax surface on a microscope slide and we used a glass micropipette to bend backwards and stably position the right antenna on a cover slip. The preparation was placed under a microscope (Olympus BX51W1), with a magnification ≤ 1500×, where a 1 l/min charcoal purified and humidified airflow was constantly blown over the fly head. To record the action potentials of antennal sensory neurons, we used tungsten microelectrodes sharpened in a KNO_2-solution, which we positioned using a motor-controlled micromanipulator (Märzhauser DC-3K, Wetzlar, Germany) equipped with a piezo unit (Märzhauser PM-10). A reference electrode was inserted into the eye with a manually controlled micromanipulator (Narishige MM33, Tokyo, Japan). A Syntech SFC-1/b stimulus controller delivered 0.5-s-long odor stimulations into the airstream at 0.5 l/min. Stimulus pipettes contained a 12.7-mm disc (Sigma-Aldrich, St. Louis, Mo, USA) onto which 5 µl of synthetic odors in paraffin oil was pipetted. After A/D conversion using an IDAC-USB (Syntech), the electrophysiological responses were fed into a PC for further analysis using AutoSpike 3.2 software (Syntech, Kirchzarten, Germany).

Y-Tube Olfactometer Bioassays

Flies were separated based on sex upon hatching, then 3-day-old females and males were put together in a vial and allowed to mate; only mated females (starved overnight) were used in the subsequent behavioral assays. Isopentyl acetate (IPA hereafter, purity > 97%; Sigma-Aldrich, Milan, Italy) was loaded on red rubber septa (Wheaton, 20-mm straight plug stopper, Millville, NJ, USA) in doses of 1, 10, and 100 µg per septum using hexane (> 99% purity, Sigma-Aldrich) as solvent. The rubber dispensers with the solution were kept for 1 hr in a climatic chamber (25 ± 2 °C and 60 ± 5% R.H.) before starting the experiment to allow solvent evaporation and to equilibrate. Behavioral bioassays were conducted using a Y-tube olfactometer to evaluate the response of mated females flies toward IPA (olfactometer size: stem = 30 cm; arm length = 20 cm; arm angle = 60°; internal diam. = 4 cm) (Revadi et al. 2015). Each dose of IPA was tested against a

control consisting of a rubber septum with hexane only. The air was filtered with activated charcoal, humidified using distilled water, and uniformly pumped through the olfactometer arms at 250 ml/min. We introduced single flies into the olfactometer at the entrance of the main stem and observed them until they made a "choice" or until 5 minutes elapsed (in this case they were recorded as "no choice"). To minimize any spatial effects, the arms were switched after having tested five females. We performed five replicates for each dose of IPA, with each replication comprising 20 flies. After every replication, the olfactometer was rinsed with water and absolute ethanol, and baked overnight at 200 °C.

Results

The Repertoire of Olfactory Genes in *D. suzukii*, *D. biarmipes*, and Other 12 *Drosophila*

We identified and manually annotated the complete repertoire of OR, OBP, and aIR genes present in the genomes of *D. biarmipes* and *D. suzukii* (supplementary table S1, dataset S1, Supplementary Material online). The *automatic* approach (based on querying predicted proteomes using the 12 *Drosophila* orthologs) and the *manual* approach (based on querying non-annotated genomes using *D. melanogaster* orthologs) retrieved similar results, although the second identified more putative orthologs in both *D. suzukii* and *D. biarmipes*. For example, in *D. suzukii*, only the *manual* approach identified four genes—*DsOr47b, DsOr35a, DsOr82a, DsOr23a4*, and two genes that lost their original function—*DsOr22a* and *DsOr74a*. By combining the two approaches, we could recover more orthologs in *D. biarmipes* (*DbOr98b, DbOr49a2, DbOr92a, DbOr67a2, DbOr67a5, DbOr42a*) compared with a recent annotation by Hopf et al. (2015). The recursive annotation strategy was successful in highlighting false-positive duplications, which were ultimately identified as allele variants in the partially heterozygous *D. suzukii* genome, as well as in revealing a duplication (*Or19a)* and two putative isoforms (*Or42a*) that were missed in a previous screening (Revadi et al. 2015). From a methodological point of view, our results indicate that for fragmented genomes like that of *D. suzukii* and *D. biarmipes*, the best annotation approach is to directly perform a BLAST search on the genome and manually assemble hits on reference orthologs, even if it is more time-consuming than the more conventional *de novo* approach.

Opposite to what is observed for ORs and OBPs, our annotations indicate conservation of the aIR gene family size among the *Drosophila* species. The OR gene family proved to be extremely dynamic in the branch leading to *D. suzukii* (fig 1), with eight gene gains (duplications of *Or19a, Or49a, Or59a, Or59c, Or67a* and quadruplication of *Or23a*), two genes that likely lost their original function (*Or85a, Or74a;* see below on how we defined a change of function), and two new isoforms (the locus of *Or42a* has three likely transcription start sites). In the branch leading to *D. suzukii* and *D. biarmipes*, we further identified a loss of function for *Or22a*, a loss of *Or98a*, duplications of *Or65c* and *Or22b*, and a quadruplication of *Or67a*. In *D. suzukii*, all OR duplications arose by tandem replication. Concerning OBPs, we identified three changes in the *D. suzukii* repertoire, namely, duplications in *Obp46a* and *Obp47a* and loss of *Obp18a*.

Species Tree: Accelerated Evolution of Olfactory Receptors in *D. suzukii*, the Obscura Subgroup, and the Simulans Complex

In *D. suzukii*, we observed a noticeable departure in the evolutionary patterns of OR and OBP gene families compared with other *Drosophila* species (figs 1 and 2). We found the overall rates to be concordant with the birth and death rates (β-Beta and δ-Delta parameters) calculated by BadiRate (figs 1B and 2B, supplementary fig S1*B*, Supplementary Material online). In the case of OR, both the overall rate of expansion and the normalized β (1.36 and 0.018, respectively) are among the highest in the phylogeny (black branch in fig 1B, boxplots in fig 1C). An overall turnover rate higher than 1 is present only in four other species, two from the *simulans* complex (*D. simulans* and *D. sechellia*, colored in blue in fig 1B) and two from the *obscura* group (*Drosophila pseudobscura* and *Drosophila persimilis*; see outliers in the boxplots of fig 1C). The overall number of events is in fact higher in *D. suzukii* ($n = 10$) than in these four species ($n = 3$ to 7), but occurred during a longer evolutionary time scale. Similarly, some internal branches are characterized by an extremely high number of events and relatively low rates (for example, the branch leading to *Drosophila grimshawi* and *Drosophila willistoni*), although the extremely incomplete taxon cautions against its information content. Turnover rates of OBPs are in general much lower than those of ORs for all the species (fig 2). *Drosophila suzukii*, and the branch leading to *D. suzukii* plus *D. biarmipes* (respectively, in red and gray/asterisk in fig 2B and 2C) are, however, outliers, as are *Drosophila yakuba, D. erecta, D. sechellia* (respectively, orange, purple, and blue in fig 2B and 2C), and the node leading to the *Sophophora* subgenus: these branches fall outside the internal quartile that groups the majority of the remaining branches. Conversely, aIRs evolve similarly in most *Drosophila* species, both in terms of rate of evolution and gene family size; we could only find one gain in *Drosophila mojavensis* and one loss in *D. sechellia* and in *D. biarmipes* (supplementary fig S1, Supplementary Material online). Notably, *D. sechellia* and *D. suzukii* are the only two species to show both a high OR and OBP turnover rate.

Because BadiRate does not take into account individual gene phylogenies, we examined the reconciliation between the single gene trees and the inferred gene birth–death distribution. Results indicate that for most subfamilies, there is perfect concordance (details are in supplementary fig S2,

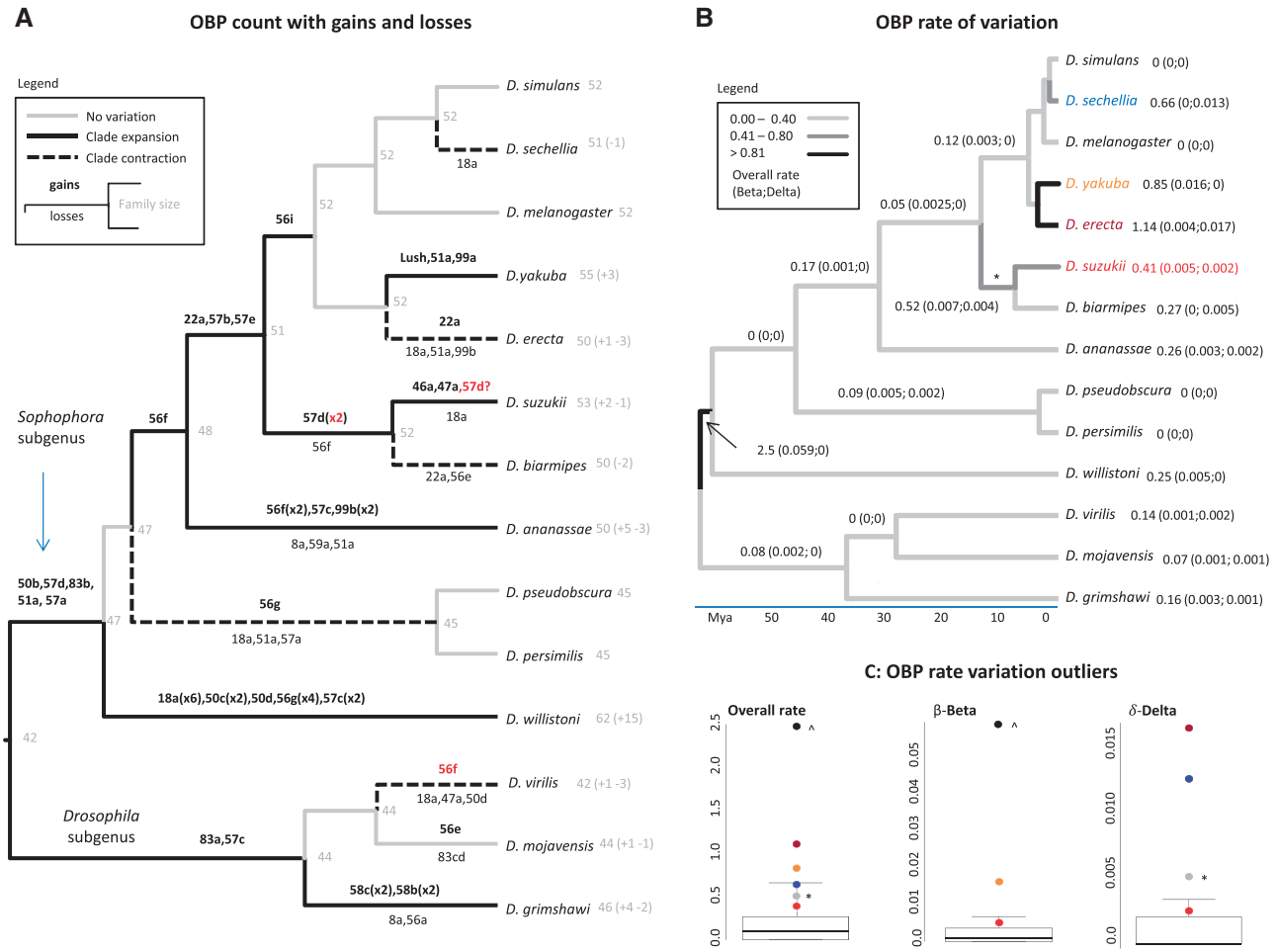

Fig. 2.—Evolution of OBPs on the *Drosophila* phylogeny. Same caption as in fig 1.

Supplementary Material online), whereas in few cases (Or22a/b, Or23a, Or67a, Or98a), the gene tree indicates an alternative distribution of gains and loss (depicted in red in fig 1A and 1B). The different interpretation of such gene distribution in *D. suzukii* would imply an extra loss in Or22a, and one less duplication of Or23a and/or Or67a and one extra duplication of Obp57d, therefore leaving the overall rate of gene family variation in *D. suzukii* similar if not higher than the one obtained with BadiRate. Even dismissing the BadiRate distribution completely, *D. suzukii* is undisputedly characterized by the highest number of ORs in the melanogaster group (67, compared with an average of 60).

Gene Trees: Evolutionary Events for Odorant Receptors Are Not Randomly Distributed in *D. suzukii*

Seven out of the 10 gains/losses that characterize ORs in *D. suzukii* are clustered in a well-supported sub-family of ORs (bootstrap support BS = 71, gray box in fig 3). This sub-family accounts for less than one-third of the whole OR family (16 out of 60), but contains the majority of gains/losses that characterize *D. suzukii*, indicating a significant departure from a random distribution of genomic events on the phylogeny (Fisher exact test, two-tailed $P = 0.002$). The only other species for which the test scored significantly are *D. ananassae* in the melanogaster group, and *D. grimshawi* and *Drosophila virilis* from the *Drosophila* subgenus (supplementary table S4, Supplementary Material online). The distribution of events for the OBPs was not assessed because their phylogenetic relationship could not be resolved with significant support (supplementary fig S3*A*, Supplementary Material online).

Signs of Positive Selection in *D. suzukii*'s Duplicated Odorant Receptors

In all species, OR, OBP, and aIR genes are under similar selective pressures, as measured by *d*N/*d*S ($P > 0.05$ for all comparisons; see supplementary table S5, Supplementary Material online). Mean *d*N and *d*N/*d*S were similar in *D. suzukii* and *D. biarmipes* in all functional classes ($P > 0.1$),

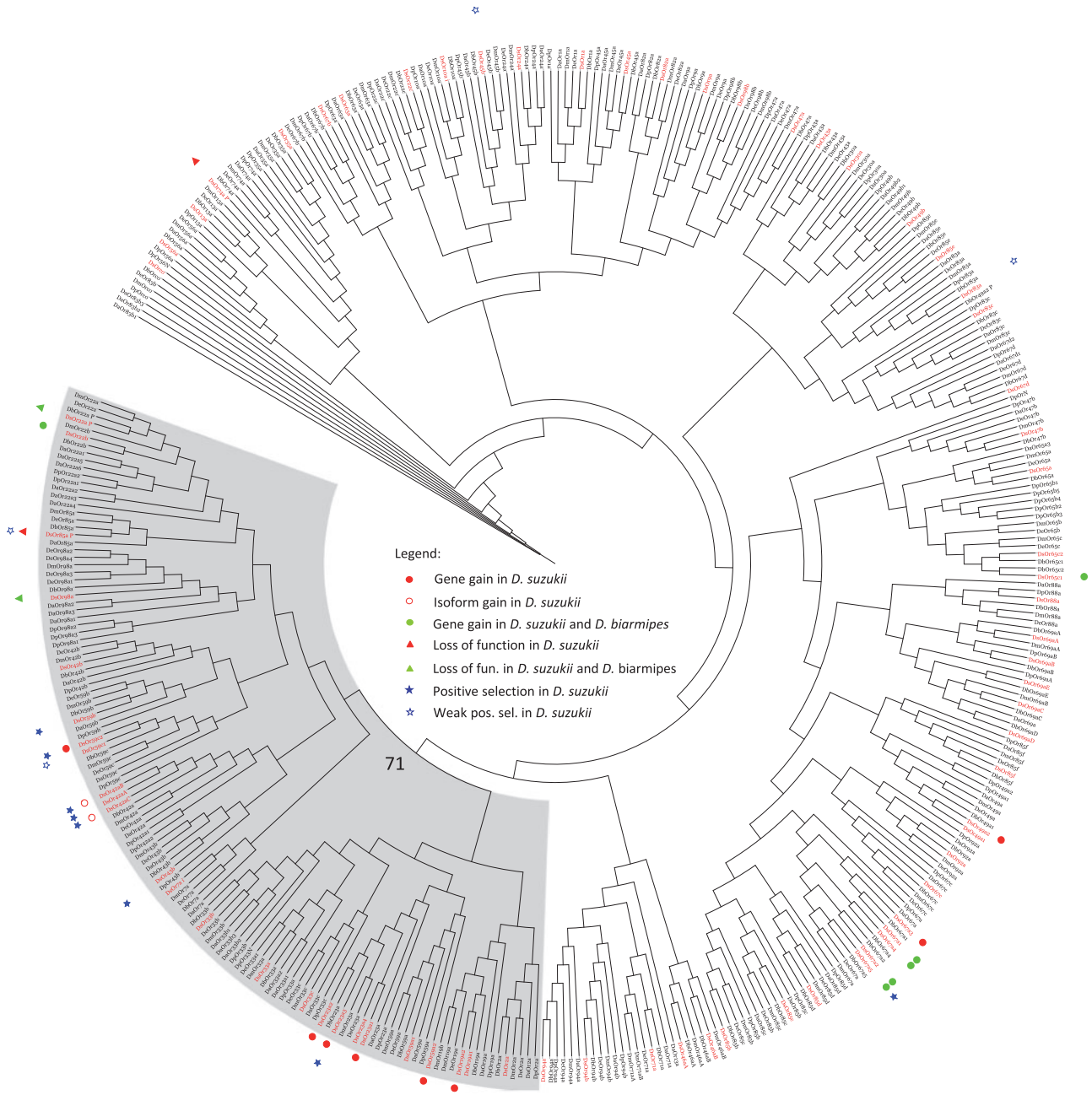

Fig. 3.—Phylogenetic tree of ORs. Most of the genomic events detected in *D. suzukii* (duplications, losses, loss of function, positive selection, see legend) cluster significantly in one subfamily highlighted with gray shade. The tree is inferred using the protein sequences from the entire gene families of 6 species (*D. melanogaster, D. erecta, D. suzukii, D. biarmipes, D. ananassae, and D. pseudoobscura*). Support at selected node is the bootstrap support from the analysis of 100 pseudo-replicates. *D. suzukii* sequences are highlighted in red.

whereas *d*S was significantly lower in *D. suzukii* than in *D. biarmipes* for ORs ($P < 0.001$; as observed at a genome-wide scale (Ometto et al. 2013)), but not for OBPs ($P = 0.351$). We found a total of 22 genes showing signs of different selective regimes and positive selection in *D. suzukii* (supplementary table S6, Supplementary Material online); after correcting for multiple testing, the number dropped to 15 genes (9 ORs, 4 OBPs, 2 aIRs; see also stars in fig 3 and boxes in fig 4). Similar results were obtained when analyzing alignments from which we removed problematic regions that could produce false-positives for fast evolution. Overall, the results of our analyses indicate that *D. suzukii* chemosensory genes are under similar evolutionary forces when compared with the closely related *D. biarmipes* and with *D. melanogaster*. In the

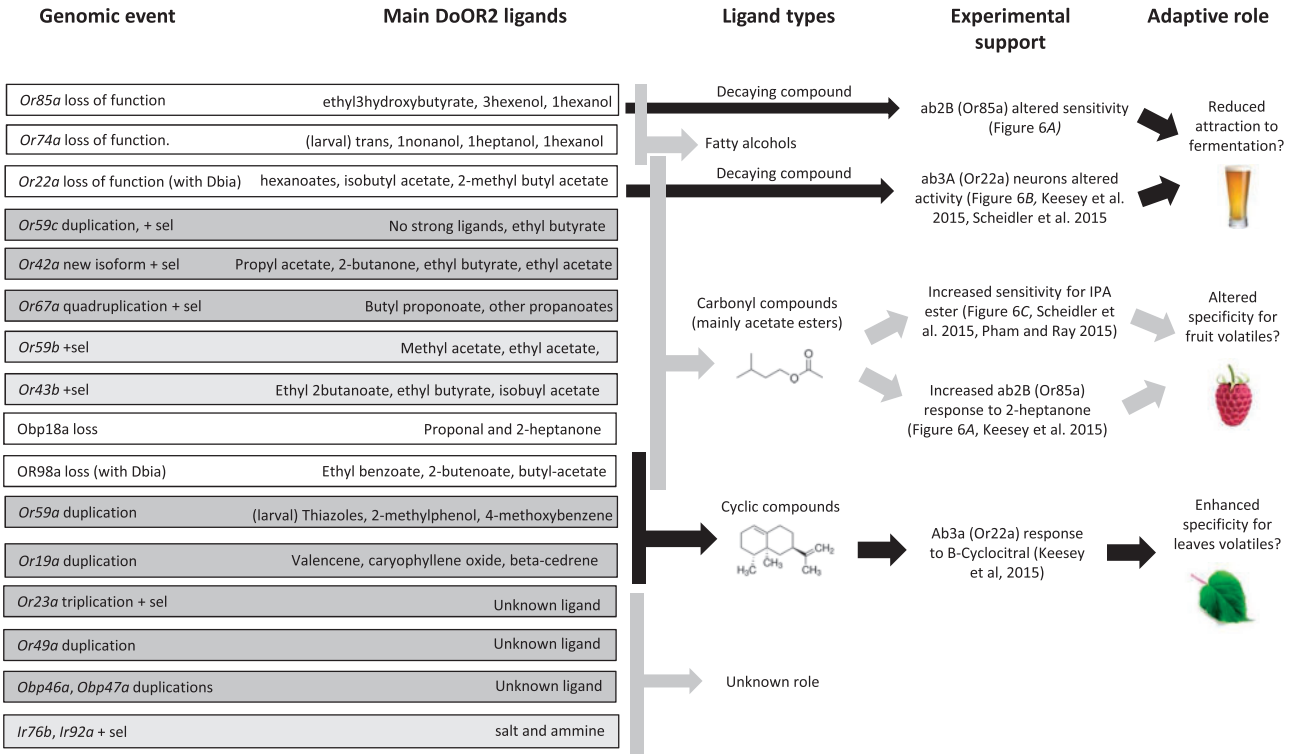

Fig. 4.—Biological and ecological interpretation of the most relevant genomic events in *D. suzukii*. Each of the chemosensory genes experiencing duplication, loss, non-functionalization, or positive selection in *D. suzukii* are listed along with the ligands they respond to in *D. melanogaster* according with the DoOR (v.2) database (Muench and Galizia 2015); a proposed behavioral ecological explanation is given. As many chemical ligands are associated with each of the receptors, we have reported only the three compounds eliciting the highest responses in the database.

case of *D. suzukii*, we found that ORs that underwent duplication are more likely to also experience positive selection (Fisher exact test, two-tailed $P = 0.012$), as expected for this type of duplicated genes (Almeida et al. 2014). This bias was, however, not observed in OBP genes ($P = 0.33$).

The Putative Chemical Responses of Duplicated/Lost Odorant Genes in *D. suzukii*

Most of the duplicated/lost/under positive selection ORs in *D. suzukii* (fig 4) respond in *D. melanogaster* to medium-sized esters (*Or22a, Or42a, Or59c, Or67a, Or98a*, but also *Obp18a)*, to similarly sized fatty alcohols (*Or74a, Or85a*), and to large although chemically unrelated cyclic compounds (*Or19a*, *Or59a*, *Or98a*). These patterns are confirmed by a quantitative screening of the DoOR (v.2) database, which reveals a variety of esters such as ethyl-butyrate, methyl-hexanoate, pentyl-acetate, and isopentyl-acetate occurring more frequently in duplicated/lost ORs of *D. suzukii* and more represented in these ORs than in all other tested ORs (see skew indexes in supplementary table S7, Supplementary Material online). Two of the ORs that lost their original function in *D. suzukii* (*Or85a* and *Or22a*) bind with high affinity to ethyl 3-hydroxybutyrate and ethyl (and methyl) hexanoate, compounds associated with yeast and bacterial fermentation (Antonelli et al. 1999). Another OBP gene, *Obp57d*, is triplicated in both *D. suzukii* and *D. biarmipes* and is involved in detecting hexanoic and octanoic acids, which are toxic for *Drosophila* in general, but not to *D. sechellia* (Matsuo et al. 2007; Harada et al. 2012).

Gene Structure Reveals Loss of Function of Key Receptors and Different Alleles in Italian and American Strains of *D. suzukii*

In *D. suzukii*, the amino acid sequences of two ORs (*Or22a* and *Or85a*) present deletions that compromise their reading frame, but otherwise retained high sequence similarity with their *D. melanogaster* orthologs. These genes are characterized by the presence of stop codons and frameshifts in the *D. suzukii* portion of the sequences matching the *D. melanogaster* exons (fig 5A and 5D). Because *Or22a* and *Or85a* are transcribed, there is an intriguing possibility that these changes did not cause a pseudogenization of the gene, but rather are associated to a change in function. The aforementioned "deleterious" changes are indeed found in portions of the exons that are missing in the transcripts (available only for the American strain, Bioproject Accession: PRJNA221549),

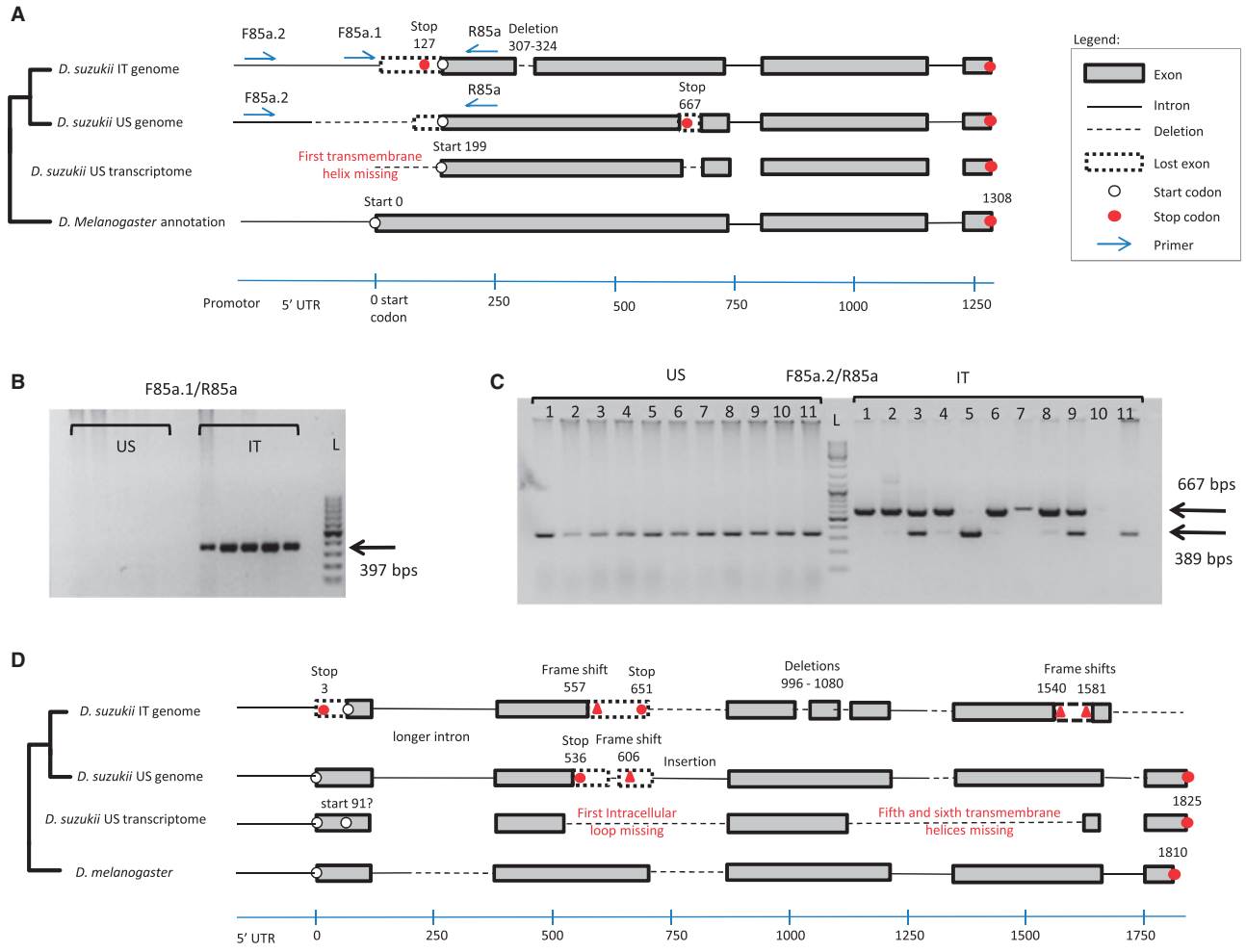

FIG. 5.—Different non-functional ORs in American and European populations. The structure of the predicted coding sequences (CDS) of *Or85a* (panel A) and *Or22a* (panel D) from the genome analysis of the Italian (IT) and American (US) strains of *D. suzukii*. For the American strain, we also provide the CDS from transcriptome (Chiu et al. 2013). Dotted lines in *D. suzukii* indicate that the CDS is missing either from the genomes or the transcriptome. B and C: agarose gel (2%) electrophoresis of different splice variants present in different individuals of American and Italian *D. suzukii* populations: US – American strain, IT – Italian strain, L – Ladder.

suggesting new exon structures and novel splicing patterns that resulted in at least one transmembrane region being lost in each gene when compared with the *D. melanogaster* orthologs. On the other hand, a third receptor, *Or74a*, is more likely a pseudogene because it retains poor similarity with the ortholog in other species (supplementary fig S4, Supplementary Material online).

Unexpectedly, the sequences of these three genes in the Italian and the American strains are characterized by a different set of putative stop codons and frame shifts. We further investigated and validated these differences in the gene *Or85a*. The F85a.1/R85a primer set, designed to confirm the presence of first transmembrane helix in both strains, amplified only in the Italian strain (fig 5B). This observation confirms that the genomic region covering the first transmembrane helix is completely absent from the American strain (dotted line in fig 5A). The primer set F85a.2/R85a further confirmed the presence of a gene size polymorphism in the American and the Italian populations: whereas all the 11 tested American flies have short alleles, we could also amplify a longer allele in the Italian population (fig 5C). Furthermore, in some of the Italian samples, the amplification of the long allele was accompanied by a very faint signal of amplification of the short one, suggesting a possible third allele in which the region close to the deletion contains a mutation preventing an efficient binding of the primer.

Sensory Physiology Shows Altered Responses of Key Neurons

Single-sensillum recordings from the large basiconic sensilla that house neurons expressing *Or85a* (ab2B) and *Or22a* (ab3A) in *D. melanogaster* demonstrated that the *D. suzukii*

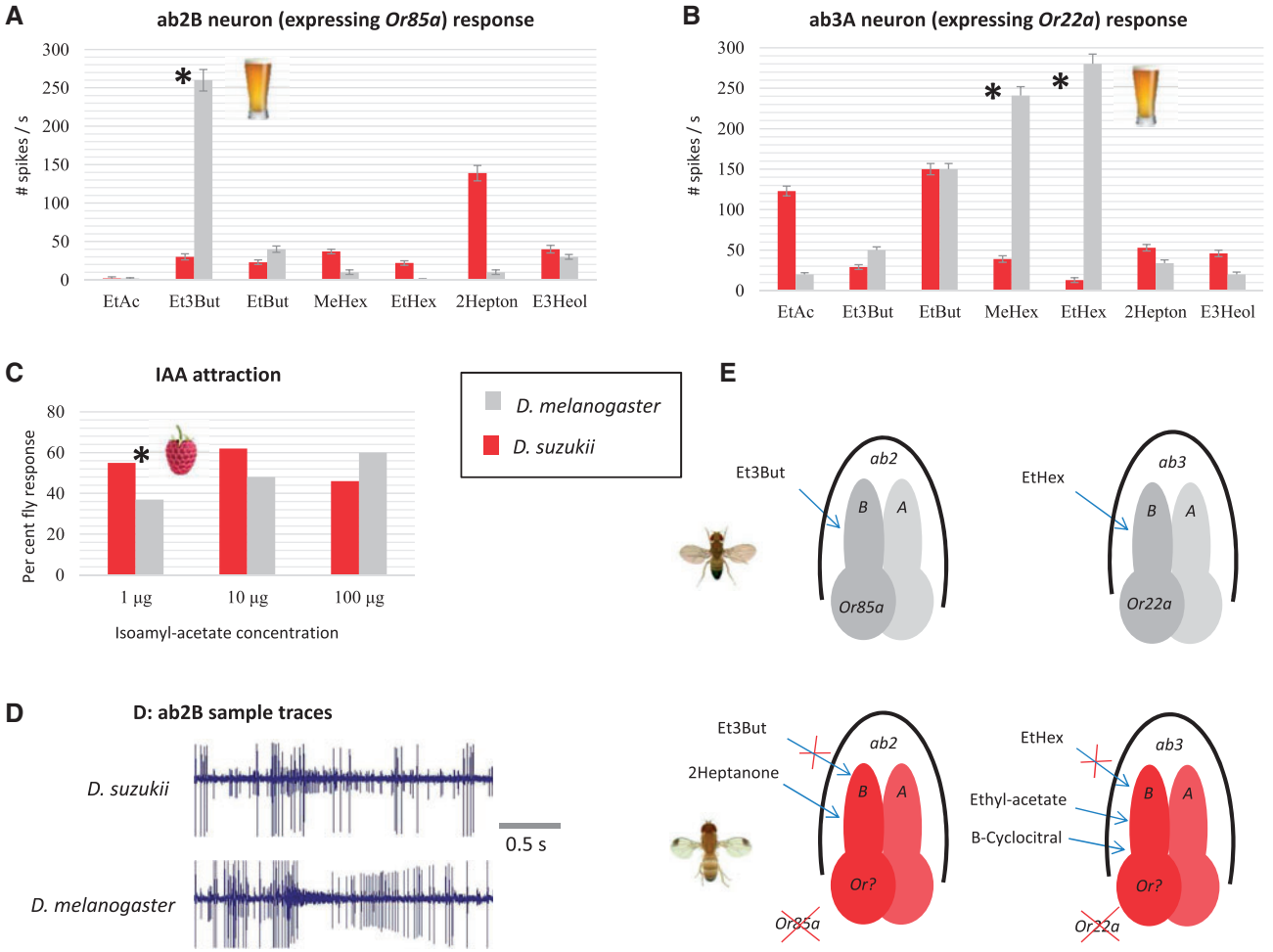

Fig. 6.—Behavior and sensory physiology. A and B: Response profiles of ab2B and ab3A neurons in *D. suzukii* compared with *D. melanogaster* support a shift of function for their *Or85a* and *Or22a* receptors. On the *x*-axis, we have EtAc – ethyl acetate, Et3But – ethyl 3-hydroxybutyrate, EtBut – ethyl butyrate, MeHex – methyl hexanoate, EtHex – ethyl hexanoate, 2Hepton – 2 heptanone, E3Heol – (*E*)-3-hexenol. C: Y-tube olfactometer bioassay shows that *D. suzukii* is significantly attracted by a reduced amount of IPA, in support of a high turnover rate for IPA receptors. D: Sample traces of aB2b neurons in *D. suzukii* and *D. melanogaster* in response to ethyl-3-hydroxy butyrate (at 10-3 dilution) 0.5-s stimulation. E: Proposed receptor replacement in *D. suzukii*.

cognate neurons have a strongly shifted response profile compared with *D. melanogaster*. In *D. melanogaster*, the ab2B neuron is tuned to oxidized esters typical of rotten fruit like ethyl 3-hydroxybutyrate, whereas our recordings demonstrate that in *D. suzukii*, this neuron does not respond to this odor (fig 6A). Conversely, in *D. suzukii*, this neuron has acquired an increased affinity for 2-heptanone, supporting a loss of function of its cognate receptor *Or85a* (see also (Keesey et al. 2015)), and a likely replacement by another OR. Similarly, whereas in *D. melanogaster*, the ab3A neuron responds strongly to ethyl and methyl hexanoate (see also (Andersson et al. 2012)), *D. suzukii* has lost its high sensitivity to these compounds and acquired an increased sensitivity for ethyl acetate (fig 6B). Response of neuron expressing *Or74a* was not tested, as in *D. melanogaster*, this receptor is expressed during the larval stage (Kreher *et al.* 2005). The results of the Y-tube olfactometer bioassay showed that *D. suzukii* is significantly more attracted by low doses of the ester IPA than *D. melanogaster* (fig 6C; Pearson χ^2 tests: 1 µg: $\chi^2 = 4.44$, df = 1, $P = 0.03$, 10 µg: $\chi^2 = 1.78$, df = 1, $P = 0.18$, 100 µg: $\chi^2 = 1.8$, df = 1, $P = 0.17$).

Discussion

Is Natural Selection Shaping the Evolution of Chemosensory Genes in *D. suzukii*?

Chemosensory gene families such as ORs are widely recognized to evolve according to a birth-and-death process (Vieira and Rozas 2011). This process assumes that genes are randomly gained or lost by local genomic events, and that duplicates can stay in the genome for long time; then their final

fate (loss, fixation of duplications, gain of function) is mostly defined by a combination of drift and selection, that is, by whether and to what extent these events affect the fitness (Vieira et al. 2007). In the case of *D. suzukii*, but also a few other species, we observe an increase in the birth-and-death rate of the OR (and to a lesser extent OBP) gene families relative to other lineages (figs 1C and 2C). Although there is not perfect reconciliation between BadiRate inferences and individual protein trees, the turnover rates remain high for *D. suzukii*, even assuming a different distribution of events. Such high rates suggest a novel selective regime permitting, or favoring, the high turnover of OR genes in *D. suzukii*. A further indication that selection may have played a major role in shaping the duplication pattern in *D. suzukii* receptor genes comes from the observation that duplications and deletions are not randomly distributed along the gene phylogenies. Instead, a subfamily of *D. suzukii* OR genes are clearly subject to higher rates of duplication/loss (fig 3): this formally violates a neutral birth-and-death process, which assumes random distribution of the mutational events. This pattern is observed, although with less significance, only in three other *Drosophila* species (*D. ananassae, D. virilis*, and *D. grimshawi*) among the 14 included in our study. Genes under diversifying selection are also non-randomly distributed and tend to cluster within the same subfamilies that experience high duplication/loss rates, consistent with a non-neutral pattern of evolution in these genes (Almeida et al. 2014).

Apart from *D. suzukii*, the only other *Drosophila* characterized by high turnover rates of both OR and OBPs is *D. sechellia*, a species for which there is ample evidence of a link between the evolution of chemosensory genes and adaptation to a new ecological niche (Dekker et al. 2006; Matsuo et al. 2007; McBride 2007; Ibba et al. 2010; Harada et al. 2012): our results suggest that such a link may also be valid for *D. suzukii*. While mutational events (deletions, duplications, point mutations) occur randomly, their fixation is not necessarily stochastic; in the case of *D. suzukii*, selective fixation of certain mutational events may have instead been favored by natural selection. We can hypothesize the effect (likely combined) of two different processes. In the first, relaxed selective pressures have allowed the fixation of gene deletions. In the second, natural selection may have favored the retention of gene duplicates. The observation that such high dynamism occurs within a single clade of ORs suggests that in *D. suzukii*, there has been a shift in the perception of the ligands that characterized such an OR clade (gray shade in fig 3). In any case, we can hypothesize that a modification of the chemosensory system, and the associated assortment of receptor genes, accompanied the change in the reproductive lifestyle of *D. suzukii*. This hypothesis is compatible with patterns of molecular evolution observed across the olfactory genes, whereby the mean level of selective pressure is similar between *D. suzukii* and other species, while some of the single duplications have undergone positive selection (see supplementary table S6, Supplementary Material online). These results, coupled with the high birth rates, suggest that ecological adaptation in *D. suzukii* occurred through an increased acceptance of gene duplications and losses and natural selection favoring the fixation of novel mutations in (some of) the duplicates.

The Ecological Significance of Duplication Events in *D. suzukii*

It is not straightforward to generalize the biological significance of the many duplications and losses that characterize ORs in *D. suzukii*, as these receptors are elicited by a large assortment of ligands (fig 4). Moreover, although being the most comprehensive source for receptor-ligand data, the DoOR (v.2) database has important limitations: first, it is not based on all possible ligands; second, it is biased toward experiments conducted on *D. melanogaster*; and finally, in some cases, it reports results of ligand concentrations that are not found in nature. Consequently, this database may be prone to both false-negatives and false-positives, so that our discussion of the ecological significance of the ORs (and their ligands) is speculative. Nonetheless, our analyses point toward a role of fatty alcohols, aromatic compounds, and especially esters, which are clearly over-represented as ligands of duplicated/lost genes (compared with all other ORs) in *D. suzukii* (supplementary table S7, Supplementary Material online). Among esters, the most represented are ethyl butyrate and IPA; the latter is present in many ripening soft fruits that host *D. suzukii* (Revadi et al. 2015), and is also released at a much higher concentration by fermenting materials such as wine and vinegar (Cha et al. 2013). Our behavioral assays demonstrate that egg-laying females of *D. suzukii* are indeed attracted by lower amount of IPA than *D. melanogaster* are (fig 6C). We speculate an adaptive scenario in which *D. suzukii* has tuned its chemosensory system to better discriminate the odor blend from ripening fresh fruit (e.g., releasing low amount of IPA), from rotting ones (releasing higher amount of IPA). Our results further point toward *Or19a* and *Or59a*, which are duplicated in *D. suzukii* and respond to different types of aromatic volatiles in *D. melanogaster*, suggesting a change in the response to cyclic/aromatic compounds in *D. suzukii*, an hypothesis that finds confirmation in the analysis by Keesey et al. (2015).

Possible Loss of Receptor Function (and Replacement) of Key Odorant Receptors

Or85a and *Or22a* are interesting cases of loss of original odorant function in, respectively, *D. suzukii* and the branch leading to *D. suzukii* plus *D. biarmipes*. In *Drosophila*, OR genes are characterized by 7 conserved transmembrane helices (7TM, Clyne et al. 1999), implying high structural constraints and pervasive purifying selection. Despite this, in *D. suzukii*, all three genes accumulated stop codons, frame shifts, and indels in regions that otherwise code for transmembrane

domains in *D. melanogaster* (and all *Drosophila* species in general). These regions are not present in the *D. suzukii* American *Or85a* and *Or22a* transcripts, suggesting that they are removed post-transcription as part of a newly formed intron or an untranslated region (5′-UTRs). Our gene annotation indicates a putative *Or85a* protein that lacks the first transmembrane domain (fig 5A), and *Or22a* proteins that lack at least two transmembrane domains (fig 5D). Because all the 7 transmembrane domains are needed for a correct and functional 7TM folding, the loss of at least one transmembrane domain is a strong indication that these genes experienced a loss of their original receptor function. Two lines of evidence support the hypothesis that *Or85a* and *Or22a* may be still functional. The first is that the coding sequences of these genes are still under selective constraints: whereas introns (including the newly formed ones) are fairly divergent between the American and the Italian genome, exons are highly conserved and did not accumulate deleterious mutations. The second is that they are transcribed; this however does not exclude that these genes may act as non-translated transcripts as in the case of competing endogenous RNAs (Welch et al. 2015). From an evolutionary perspective, our data are consistent with a model where the new splice pattern (and the consequent loss of the original OR function) evolved in the *D. suzukii* common ancestor, followed by a relaxation in selective pressure on the new non-coding region and the consequent independent accumulation of stop codons and other polymorphisms in the introns.

A loss of original function of *Or85a* and *Or22a* is confirmed by our observation that the corresponding ab2B and ab3A neurons shifted their affinity from volatiles typically produced by yeast during fruit fermentation to volatiles more typical of ripening fruit; this is in accordance with Keesey et al. (2015), who further found an ab3A affinity for the leaf volatile cyclic compound beta-cyclocitral. Because they lack at least one transmembrane domain, it is very unlikely that the new isoforms of *Or85a* and *Or22a* are responsible for the new ligand affinity. Our results rather point toward a scenario in which *Or85a* and *Or22a* have been replaced in *D. suzukii* by other ORs in the corresponding ab2B and ab3A neurons (fig 6E). Our data do not allow proposing any specific ORs, but we cautiously suggest as candidates the various genes that experience duplication or positive selection in *D. suzukii* (fig 4). Overall, our results suggest that *D. suzukii* changed its response to some of the compounds typical of decaying materials in general, which are the primary oviposition sites of most other *Drosophila* species. This was achieved, at least partially, by losing the original function of *Or85a*, a receptor that is otherwise widely conserved among *Drosophila* (de Bruyne et al. 2010), because it is linked with fermented foods, the feeding source of most *Drosophila* species. As *D. suzukii* oviposits on ripe fruits, but feeds on rotten substrate, the loss of *Or85a* may primarily be involved in avoiding oviposition in rotten fruits.

Utility of Identified Genes and Chemicals for Downstream Applications

Our analyses revealed a list of ORs and binding proteins likely involved in the unique biology of *D. suzukii.* Further research should focus on functional analyses of new (duplicated) and putatively lost/replaced receptors, and on mapping their expression pattern in the antenna, palp, and dorsal organ of larvae. In this work, we have assayed the behavioral role of one ester, IPA, and demonstrated that *D. suzukii* is attracted to lower concentrates of IPA compared with *D. melanogaster*, suggesting a possible use for species-specific, dosage-controlled trapping systems. The results presented here will help direct research efforts in the development of more targeted odor-based trapping and control methods. Future works should test those ligands for which there has been a shift in chemosensation, particularly 1-hexanol, 2-heptanone, and beta-cyclocitral, the two latter being putatively new ligands of, respectively, ab2B and ab3A neurons in *D. suzukii.*

Evolution of Chemosensory Genes in *Drosophila*

While previous works on chemosensory genes have often concentrated on either genomics or physiology (Robertson et al. 2003; Guo and Kim 2007; Vieira and Rozas 2011; Becher et al. 2012; Swarup et al. 2014; Keesey et al. 2015), in this study, we combined the two to gain a broader and more in-depth knowledge of their evolution (Goldman-Huertas et al. 2015). Furthermore, by comparing the birth–death trees with their corresponding gene trees, we could assess which genes have been lost or gained in each of the *Drosophila* species (detailed in figs 1 and 2), thus obtaining much more information than if concentrating only on the quantitative aspect of the evolution of duplicated/lost genes on the gene phylogeny (Guo and Kim 2007; McBride and Arguello 2007; Vieira and Rozas 2011, but see Robertson 2009). These results may serve for future chemical ecological studies involving the various *Drosophila* species we have studied. From a quantitative point of view, our results confirm that the evolution of chemosensory genes in *Drosophila* is quite dynamic, with significant variation in the birth and death processes affecting some lineages and gene subfamilies: OBPs and especially ORs are extremely variable families, while aIR are fairly conserved throughout the species tree. Our distribution of gene gain/loss along *Drosophila* phylogeny slightly differs from inferences made on a more restricted sample of (Guo and Kim 2007, McBride and Arguello 2007). For example, some of the gains (*Or67a, Or65c*, and *Or85a*) that were previously located on the branch subtending the *melanogaster* subgroup, in our analysis, are located on the branch subtending the whole *melanogaster* group.

It has to be pointed out that, like in other studies, our comparative analysis is biased by the available sampling of species, which is extremely poor outside the Sophophora lineage. Therefore, although the changes we observe, for

example, in the branch subtending *D. sechellia* are good descriptors of the evolutionary events that characterize this species, the events on the branch subtending *D. ananassae* are not species-specific, but rather characterize the whole *ananassae* subgroup. The case of *D. suzukii* is somehow in-between, because *D. biarmipes* is a fairly closely related species. Another of its closely related species, *Drosophila subpulchrella*, has not been included in our analyses: if *D. subpulchrella* is the actual sister species to *D. suzukii*, then some of the evolutionary events we have ascribed to *D. suzukii* may instead be shared by both species. The genome of the *D. subpulchrella* is being analyzed in our lab and annotations of its gene repertoire with that of *D. suzukii* will clarify this issue.

Conclusions

Our results describe the genome evolution behind some of the peculiar biology of an emerging pest and further instruct us over the general evolution of chemosensation in animals. Results indicate that the evolution of the *D. suzukii*'s olfactory genes repertoire, particularly ORs, is different from that of most other *Drosophila* species: we have shown that *D. suzukii* is the only species to show both a high OR turnover rate and a non-random distribution of OR events, suggesting distinct selective forces possibly imposed by a shift in their chemo-ecological environment. The most convincing cases we found for *D. suzukii* are (i) a burst of duplications for genes with affinity for some type of ligands, particularly esters, which may have resulted in enhanced sensitivity for small dosages of IPA; and (ii) a loss of function for receptors with high affinity for volatiles associated with fermentations. These genes, as well as all the other genes listed in fig 4 and their putative ligands, are good candidates for downstream applied physiological and behavioral experiments.

Supplementary Material

Supplementary dataset S1, figures S1–S4, and tables S1–S7 are available at *Genome Biology and Evolution* online (http://www.gbe.oxfordjournals.org/).

Acknowledgment

The authors would like to thank 4 anonymous reviewers and GBE associate editor for their useful comments and suggestions. Sukanya Ramasamy and Rupinder Kaur were funded by PhD fellowships from the FIRS>T (FEM International Research School-Trentino) programme at Fondazione Edmund Mach, Italy. Cristina M. Crava is recipient of a Maire-Curie FP7-PEOPLE-2013-IEF grant. This work was partly funded by Project LExEM (Laboratory of excellence for epidemiology and modeling, http://www.lexem.eu).

Literature Cited

Abascal F, Zardoya R, Telford MJ. 2010. TranslatorX: multiple alignment of nucleotide sequences guided by amino acid translations. Nucleic Acids Res. 38:W7–W13.

Almeida FC, Sánchez-Gracia A, Campos JL, Rozas J. 2014. Family size evolution in Drosophila chemosensory gene families: a comparative analysis with a critical appraisal of methods. Genome Biol Evol. 6(7):1669–1682.

Altschul SF, et al. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. 25:3389–3402.

Andersson MN, Schlyter F, Hill SR, Dekker T. 2012. What reaches the Antenna? How to calibrate Odor Flux and ligand–receptor affinities. Chem Senses 37:403–420.

Antonelli A, Castellari L, Zambonelli C, Carnacini A. 1999. Yeast influence on volatile composition of wines. J Agric Food Chem. 47:1139–1144.

Asplen MK, et al. 2015. Invasion biology of Spotted Wing Drosophila (*Drosophila suzukii*): a global perspective and future priorities. J Pest Sci. 88:469–494.

Becher PG, et al. 2012. Yeast, not fruit volatiles mediate Drosophila melanogaster attraction, oviposition and development. Funct Ecol. 26:822–828.

Begon M. 1982. Yeasts and *Drosophila*. In: Ashburner M, Carson HL, Thompson JJN, editors. The genetics and biology of Drosophila. New York: Academic Press, p. 345–384.

Benton R, Vannice KS, Gomez-Diaz C, Vosshall LB. 2009. Variant Ionotropic Glutamate Receptors as Chemosensory Receptors in Drosophila. Cell 136:149–162.

Calabria G, Maca J, Bachli G, Serra L, Pascual M. 2012. First records of the potential pest species Drosophila suzukii (Diptera: Drosophilidae) in Europe. J Appl Entomol. 136:139–147.

Castresana J. 2000. Selection of conserved blocks from multiple alignments for their use in Phylogenetic analysis. Mol Biol Evol. 17:540–552.

Cha DH, et al. 2013. Comparison of a synthetic chemical lure and standard fermented baits for trapping *Drosophila suzukii* (Diptera: Drosophilidae). Environ Entomol. 42:1052–1060.

Chiu JC, et al. 2013. Genome of *Drosophila suzukii*, the spotted wing Drosophila. G3 12:2257–2271.

Cini A, Ioriatti C, Anfora G. 2012. A review of the invasion of *Drosophila suzukii* in Europe and a draft research agenda for integrated pest management. Bull Insectol. 65(1):149–160.

Clyne PJ, et al. 1999. A novel family of divergent seven-transmembrane proteins: candidate odorant receptors in Drosophila. Neuron 22:327–338.

de Bruyne M, Smart R, Zammit E, Warr C. 2010. Functional and molecular evolution of olfactory neurons and receptors for aliphatic esters across the Drosophila genus. J Comp Physiol. 196:97–109.

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. 16:101–109.

Dekker T, et al. 2015. Loss of Drosophila pheromone reverses its role in sexual communication in *Drosophila suzukii*. Proc R Soc B. 282:

Drysdale RA, Crosby MA, Consortium TF. 2005. FlyBase: genes and gene models. Nucleic Acids Res. 33:D390–D395.

Edgar RC. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 32:1792–1797.

Gardiner A, Barker D, Butlin RK, Jordan WC, Ritchie MG. 2008. Drosophila chemoreceptor gene evolution: selection, specialization and genome size. Mol Ecol. 17:1648–1657.

Goodhue RE, Bolda M, Farnsworth D, Williams JC, Zalom FG. 2011. Spotted wing drosophila infestation of California strawberries and

raspberries: economic analysis of potential revenue losses and control costs. Pest Manag Sci. 67:1396–1402.
Goldman-Huertas B, et al. 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.
Guo S, Kim J. 2007. Molecular evolution of Drosophila odorant receptor genes. Mol Biol Evol. 24:1198–1207.
Han MV, Demuth JP, McGrath CL, Casola C, Hahn MW. 2009. Adaptive evolution of young gene duplicates in mammals. Genome Res. 19:859–867.
Harada E, et al. 2012. Functional evolution of duplicated odorant-binding protein genes, *Obp57d* and *Obp57e* in Drosophila. PLoS One 7:e29710.
Heuskin S, Verheggen FJ, Haubruge É, Wathelet JP, Lognay G. 2011. The use of semiochemical slow-release devices in integrated pest management strategies. Agron Soc Env. 15:459–470.
Hopf TA, et al. 2015. Amino acid coevolution reveals three-dimensional structure and functional domains of insect odorant receptors. Nat Commun. 6:6077.
Ibba I, Angioy AM, Hansson BS, Dekker T. 2010. Macroglomeruli for fruit odors change blend preference in Drosophila. Naturwissenschaften 97:1059–1066.
Keesey I, Knaden M, Hansson B. 2015. Olfactory specialization in *Drosophila suzukii* Supports an ecological shift in host preference from Rotten to fresh fruit. J Chem Ecol. 41:121–128.
Kreher SA, Kwon JY, Carlson JR. 2005. The molecular basis of odor coding in the Drosophila Larva. Neuron 46:445–456.
Krause Pham C, Ray A. 2015. Conservation of Olfactory avoidance in Drosophila species and identification of repellents for *Drosophila suzukii.* Sci Rep. 5:11527.
Landolt PJ, Adams T, Rogg H. 2012. Trapping spotted wing drosophila, Drosophila suzukii (Matsumura) (Diptera: Drosophilidae), with combinations of vinegar and wine, and acetic acid and ethanol. J Appl Entomol. 136:148–154.
Le SQ, Gascuel O. 2008. An improved general amino acid replacement matrix. Mol Biol Evol. 25:1307–1320.
Librado P, Vieira FG, Rozas J. 2011. BadiRate: estimating family turnover rates by likelihood-based methods. Bioinformatics 28(2): 279–281.
Löytynoja A, Goldman N. 2005. An algorithm for progressive multiple alignment of sequences with insertions. Proc Natl Acad Sci U S A. 102:10557–10562.
Matsuo T, Sugaya S, Yasukawa J, Aigaki T, Fuyama Y. 2007. Odorant-Binding Proteins OBP57d and OBP57e Affect Taste Perception and Host-Plant Preference in *Drosophila sechellia.* PLoS Biol. 5:e118.
McBride CS, Arguello JR. 2007. Five Drosophila genomes reveal nonneutral evolution and the signature of host specialization in the Chemoreceptor superfamily. Genetics 177(3):1395–1416.
McBride CS. 2007. Rapid evolution of smell and taste receptor genes during host specialization in *Drosophila sechellia.* Proc Natl Acad Sci U S A. 104:4996–5001.
Mistry J, Finn RD, Eddy SR, Bateman A, Punta M. 2013. Challenges in homology search: HMMER3 and convergent evolution of coiled-coil regions. Nucleic Acids Res. 41:e121.
Muench D, Galizia CG. 2015. DoOR 2.0 - comprehensive mapping of Drosophila melanogaster odorant responses. Sci Rep. 6:21841.
Ometto L, et al. 2013. Linking genomics and ecology to investigate the complex evolution of an Invasive Drosophila Pest. Genome Biol Evol. 5:745–757.
Punta M, et al. 2012. The Pfam protein families database. Nucleic Acids Res. 40:D290–D301.
Revadi S, et al. 2015. Olfactory responses of *Drosophila suzukii* females to host plant volatiles. Physiol Entomol. 40:54–64.
Robertson HM. 2009. The insect Chemoreceptor Superfamily in Drosophila pseudoobscura: molecular evolution of ecologically-relevant genes over 25 million years. J Insect Sci. 9:1–14.
Robertson HM, Warr CG, Carlson JR. 2003. Molecular evolution of the insect chemoreceptor gene superfamily in *Drosophila melanogaster.* Proc Natl Acad Sci U S A. 100:14537–14542.
Rossi-Stacconi MV, et al. 2016. Multiple lines of evidence for reproductive winter diapause in the invasive pest Drosophila suzukii: useful clues for control strategies. J Pest Sci. 89:689.
Rota-Stabelli O, Blaxter M, Anfora G. 2013. *Drosophila suzukii.* Curr Biol. 23:R8–R9.
Sanchez-Gracia A, Vieira FG, Rozas J. 2009. Molecular evolution of the major chemosensory gene families in insects. Heredity 103:208–216.
Sato K, et al. 2008. Insect olfactory receptors are heteromeric ligand-gated ion channels. Nature 452:1002–1006.
Scheidler NH, Liu C, Hamby KA, Zalom FG, Syed Z. 2015. Volatile codes: correlation of olfactory signals and reception in *Drosophila*-yeast chemical communication. Sci Rep. 5:14059.
Stamatakis A. 2014. RAxML Version 8: a tool for Phylogenetic analysis and post-analysis of Large Phylogenies. Bioinformatics 30(9):1312–1313.
Stanke M, Waack S. 2003. Gene prediction with a hidden Markov model and a new intron submodel. Bioinformatics 19:ii215–ii225.
Steinbrecht RA. 1997. Pore structures in insect olfactory sensilla: a review of data and concepts. Int J Insect Morphol Embryol. 26:229–245.
Storey JD. 2002. A direct approach to false discovery rates. J R Stat Soc. 64:479–498.
Swarup S, Morozova TV, Sridhar S, Nokes M, Anholt RRH. 2014. Modulation of feeding behavior by Odorant-Binding Proteins in Drosophila melanogaster. Chem Senses 39:125–132.
Vieira F, Sanchez-Gracia A, Rozas J. 2007. Comparative genomic analysis of the odorant-binding protein family in 12 Drosophila genomes: purifying selection and birth-and-death evolution. Genome Biol. 8:R235.
Vieira FG, Rozas J. 2011. Comparative genomics of the odorant-binding and chemosensory protein gene families across the Arthropoda: origin and evolutionary history of the chemosensory system. Genome Biol Evol. 3:476–490.
Vogt RG, Riddiford LM. 1981. Pheromone binding and inactivation by moth antennae. Nature 293:161–163.
Walsh DB, et al. 2011. *Drosophila suzukii* (Diptera: Drosophilidae): invasive pest of ripening soft fruit expanding its geographic range and damage potential. J Integr Pest Manag. 2:1–7.
Waterhouse RM, Tegenfeldt F, Li J, Zdobnov EM, Kriventseva EV. 2013. OrthoDB: a hierarchical catalog of animal, fungal and bacterial orthologs. Nucleic Acids Res. 41:D358–D365.
Welch JD, Baran-Gale J, Perou CM, Sethupathy P, Prins JF. 2015. Pseudogenes transcribed in breast invasive carcinoma show subtype-specific expression and ceRNA potential. BMC Genomics 16(1):113.
Wicher D, et al. 2008. Drosophila odorant receptors are both ligand-gated and cyclic-nucleotide-activated cation channels. Nature 452:1007–1011.
Yang Z. 2007. PAML 4: phylogenetic analysis by maximum likelihood. Mol Biol Evol. 24:1586–1591.

Associate editor: Josefa Gonzalez