Evolution of Gene Expression in the Drosophila Olfactory System

Artyom Kopp,* Olga Barmina,* Andrew M. Hamilton,* Laura Higgins,†‡ Lauren M. McIntyre,§∥ and Corbin D. Jones¶

*Section of Evolution and Ecology, University of California, Davis; †Department of Biology, Indiana University; ‡NASA, Ames Research Center, Moffett Field, CA; §University of Florida Genetics Institute, University of Florida; ||Department of Molecular Genetics and Microbiology, University of Florida; and ¶Department of Biology and Carolina Center for Genome Science, University of North Carolina, Chapel Hill

Host plant shifts by phytophagous insects play a key role in insect evolution and plant ecology. Such shifts often involve major behavioral changes as the insects must acquire an attraction and/or lose the repulsion to the new host plant’s odor and taste. The evolution of chemotactic behavior may be due, in part, to gene expression changes in the peripheral sensory system. To test this hypothesis, we compared gene expression in the olfactory organs of Drosophila sechellia, a narrow ecological specialist that feeds on the fruit of Morinda citrifolia, with its close relatives Drosophila simulans and Drosophila melanogaster, which feed on a wide variety of decaying plant matter. Using whole-genome microarrays and quantitative polymerase chain reaction, we surveyed the entire repertoire of Drosophila genes in D. sechellia compared both with the genome average in that species and with the rate of OR and OBP evolution in the other species. However, some of the gene expression changes that correlate with D. sechellia’s increased sensitivity to Morinda odorants may predate its divergence from D. simulans. Interspecific divergence of olfactory gene expression cannot be fully explained by changes in the relative abundance of different sensilla as some ORs and OBPs have evolved independently of other genes expressed in the same sensilla. A number of OR and OBP genes are upregulated in D. sechellia compared with its generalist relatives. These genes include Or22a, which likely responds to a key odorant of M. citrifolia, and several genes that are yet to be characterized in detail. Increased expression of these genes in D. sechellia may have contributed to the evolution of its unique chemotactic behavior.

Introduction

The adaptation of phytophagous insects to host plant chemistry influences both insect evolution and plant community structure (Becerra 1997, 2007; Renwick 2001). Host plant shifts are particularly important in insect speciation and ecological diversification, leading, for example, to the origin of new agricultural pests (Hawthorne and Via 2001; Dres and Mallet 2002; Nosil et al. 2002; Linn et al. 2003). Changes in chemotactic behavior are an integral part of these shifts as the insect must acquire an attraction and/or overcome the repulsion to the new food source.

Drosophila use both taste and smell to find food, mates, and appropriate environments for their offspring. Olfactory receptors (ORs) are important for locating suitable food sources at a distance; once at a food source or oviposition site, both gustatory and olfactory systems contribute to the fly’s behavior (Takamura and Fuyama 1980; Hallem et al. 2006; Ebbs and Amrein 2007). Here, we examine the evolution of olfactory system by contrasting gene expression in Drosophila sechellia, a narrow host plant specialist, with that of its generalist relatives Drosophila simulans and Drosophila melanogaster.

Physiology and Biochemistry of Taste and Smell in Drosophila

The role of the Drosophila “nose” is played by the third antennal segment, which carries several hundred olfactory sensilla, and the maxillary palp, which carries ∼60 sensilla (Shanbhag et al. 1999). These sensilla have 3 major morphological types—basiconic, trichoid, and coeloconic—each of which can be further subdivided into several subtypes based on innervation patterns and neurophysiological responses. For example, the large basiconic sensilla on the antennae are classified into ab1, ab2, and ab3 subtypes (Shanbhag et al. 1999; Stocker 2001; Couto et al. 2005). Each sensillum is innervated by 1–4 olfactory neurons, which project into the antennal lobe of the brain (Couto et al. 2005; Fishilevich and Voshall 2005). The sensitivity of each chemoreceptive neuron to a range of volatile compounds is determined by transmembrane OR proteins. Most neurons in the antennal sensory organs express a single OR, although some types of neurons express 2 or 3 (Voshall et al. 1999; Dobritsa et al. 2003; Yao et al. 2005). Each OR responds to a subset of volatile compounds with a concentration-dependent manner, with different ORs having distinct but overlapping response spectra (de Bruyne and Warr 2005; Hallem et al. 2006; Smith 2007). Most ORs are sensitive to alcohols, ketones, fatty acids, esters, and other volatile compounds present in the fruit (Yao et al. 2005; Hallem and Carlson 2006), whereas a few are responsible for the perception of fly pheromones (Kurtovic et al. 2007; van der Goes van Naters and Carlson 2007).

Contact chemoreception (i.e., taste) is mediated by gustatory organs located on the legs, proboscis, wings, and genitalia (Dunipace et al. 2001; de Bruyne and Warr 2005; Hallem et al. 2006; Smith 2007). These organs express a different set of transmembrane proteins called gustatory receptors (GRs) (Scott et al. 2001; Thorne et al. 2005; Ebbs and Amrein 2007). Both olfactory and gustatory organs also express odorant-binding proteins (OBPs), which are secreted proteins that function in the lymph fluid that fills the sensilla lumen (Galindo and Smith 2001; Graham and Davies 2002; Hekmat-Scafe et al. 2002). One of their functions may be to chaperone the ligands through the
lymph to the neuron dendrites where they can be bound by an OR or GR. Other hypothesized functions of OBPs include concentration of odorant molecules in the sensillum lymph, removal or detoxification of odorants, and coinervation of signal transduction through OR and GR receptors (Rutzel and Zwiebel 2005; Hallem et al. 2006). In general, little is known about the binding specificity of OBPs. The only Drosophila OBP characterized in detail, Obp76a (also called *lush*), is essential for the sensitivity of olfactory neurons to cis-vaccenyl acetate, which acts as a sex and aggregation pheromone in *D. melanogaster* (Xu et al. 2005; Ha and Smith 2006).

**Host Plant Choice in Drosophila**

*Drosophila* exploit a wide variety of food sources, including fungi, flowers, cambium and sap fluxes of woody plants, cacti, and both fleshy and dry fruits (Markow and O’Grady 2005). In the *D. melanogaster* species group, most species are ecological generalists that feed on decaying fruit and other decaying vegetation. However, a few species in this group have evolved into narrow specialists restricted to a small number of host plants, such as *Eucalyptus* and *Eugenia* for *Drosophila flavohirta*, *Pandanus* for *Drosophila erecta*, and *Morinda* citrifolia for *D. sechellia* (Lachaise et al. 1988; McEvey et al. 1989).

*Drosophila sechellia*, in particular, is a promising model for understanding the genetic and molecular mechanisms underlying the evolution of host plant preferences. Whereas its closest relatives—*D. melanogaster*, *D. simulans*, and *Drosophila mauritiana*—are generalists, *D. sechellia* almost exclusively uses relatively unspoiled ripe fruit of *Morinda* (Lachaise et al. 1988). *Drosophila sechellia* preference for this plant typifies the complex interplay of smell and taste in fly behavior. Field and laboratory studies have shown that volatile compounds present in the ripe *Morinda* fruit are sufficient to attract *D. sechellia* and repel *D. simulans* and *D. melanogaster* over large distances, suggesting that long-range host plant choice is mediated by ORs (Louis and David 1986; R’Kha et al. 1991; Higa and Fuyama 1993; Amlou et al. 1998; Legal et al. 1999). Part of this attraction is likely due to the octanoic and hexanoic acids that are abundant in *Morinda*, although other medium-chain fatty acids and derived compounds may also play a role (Amlou et al. 1998; Legal et al. 1999). Once on its host, contact-mediated gustatory responses become important, especially during oviposition. Recent work has implicated Obp57d/e, expressed in leg gustatory organs, in behavioral differences between *D. sechellia* and its relatives. Replacement of *D. melanogaster* Obp57d/e with the *D. sechellia* ortholog causes transgenic flies to stop avoiding oviposition on media containing compounds from *Morinda* fruit (Matsuo et al. 2007).

Although not as extreme as *D. sechellia*, both *D. melanogaster* and *D. simulans* also exhibit species-specific food and oviposition site preferences (David et al. 2004). Even within *D. melanogaster*, there is evidence for considerable variation in odor-driven food preference (van Delden and Kamping 1990; Possidente et al. 1999).

Evolution of the Olfactory System in *D. sechellia* and its relatives

Morphology of the third antennal segment and the distribution of morphologically distinguishable sensilla are well conserved among the close relatives of *D. melanogaster* (the *melanogaster* species subgroup), although *D. sechellia* exhibits subtle changes in the relative numbers and fine structure of some sensilla (Stensmyr et al. 2003; Dekker et al. 2006). However, neurophysiological recordings have shown that *D. sechellia* is unique in this subgroup in not having any basiconic sensilla with the ab2 response spectrum, with a concomitant increase in the number of ab3 and a decrease in the number of ab1 sensilla (Stensmyr et al. 2003; Dekker et al. 2006). In addition, the dominant ligand for the ab3 receptors has shifted from ethyl hexanoate in *D. melanogaster* and most other species to methyl hexanoate in *D. sechellia* and *D. simulans*. *Drosophila sechellia* ab3 receptors show increased sensitivity to hexanoate esters compared with *D. melanogaster*. These changes may have been a key factor in the ecological specialization of *D. sechellia* as hexanoates are major components of the *Morinda* fruit headspace—the volatile odors around the fruit (Dekker et al. 2006).

Olfactory and gustatory systems of *D. sechellia* have also undergone considerable changes at the molecular level. Six OR and 13 GR genes have been inactivated by loss-of-function mutations in this species—a rate of gene loss much higher than in its generalist relatives. Coding sequence evolution in the remaining OR and GR genes has also been accelerated in *D. sechellia* compared with *D. simulans* (McBride 2007; Nozawa and Nei 2007; Tunstall et al. 2007). Interestingly, genes inactivated in *D. sechellia* include Or22b, one of the receptor proteins expressed in the hexanoate-sensitive ab3 sensilla. These gene losses may have contributed to the host plant shift in *D. sechellia*, although for any given gene it is unclear whether its loss was a cause or a consequence of ecological specialization. As *D. sechellia* is an island endemic with a small effective population size, the relative roles of selective and demographic factors in gene loss are not always clear.

In the evolution of an ecological specialist from a generalist ancestor, some genes and receptors may be rendered superfluous, whereas others may take on increasingly prominent roles. Although most work to date has focused on the former (McBride 2007; Nozawa and Nei 2007; Tunstall et al. 2007), it is necessary to elucidate both sides of this process to understand the molecular mechanisms of an ecological shift. Accelerated evolution of many ORs in *D. sechellia* may in part reflect functional changes increasing or decreasing their sensitivity to ecologically relevant odorants. However, changes in the relative abundance of different receptor proteins in olfactory neurons may also affect their response spectra and contribute to the unique chemotactic behavior of this species. In this report, we examine the divergence of gene expression in the olfactory system between *D. sechellia* and its generalist relatives *D. simulans* and *D. melanogaster*, focusing particularly on OR and OBP genes. We hypothesized that *D. sechellia* may have experienced “gain-of-function” expression changes, in addition to the loss-of-function changes already
Materials and Methods

Gene expression was compared among 4 genotypes: D. sechellia isofemale line 81, multilemale D. simulans strain from Winters, CA, and 2 isogenic strains of D. melanogaster, W158 and W189 (Yang and Nuzhdin 2003). Males and females were analyzed separately. For each genotype and sex, we used 3 independent biological replicates of 32–35 individuals each, for a total of 24 samples. All dissections were performed between 1 and 2 h after eclosion. From each fly, the third antennal segment and arista were removed with microsurgical scissors (Storz Ophthalmic Instruments, San Dimas, CA) and placed immediately in Trizol (Invitrogen, Carlsbad, CA) on ice. Dissected samples were frozen in Trizol at –80 °C until RNA extraction.

RNA extraction and quality control, cDNA synthesis, and linear amplification and labeling of complementary RNA were performed as described (Barmina et al. 2005). Each replicate RNA sample isolated from the antennae contained 500–1,250 ng of total RNA. A total of 75 ng from each sample was amplified to obtain 122–172 μg of biotin-labeled complementary RNA. A total of 15 μg from each labeled sample were hybridized to Affymetrix Drosophila genome v2.0 microarrays.

Signal intensities were quantified using MAS 5.0 (http://www.affymetrix.com). Microarray probes that did not match perfectly the genome sequences of all 3 species were removed from the analysis (see Supplementary Information online). Normalization was performed by centering natural log-transformed signal intensities to the chip median. Probe sets that produced “absent” calls for all replicates were excluded from further analysis. For each remaining probe set, we fitted the following analysis of variance (ANOVA) model:

$$Y_{jk} = \mu + l_j + s_k + ls_{jk} + \epsilon_{jk},$$

where $Y_{jk}$ is the log-normalized transcript abundance for the $j$th species ($j$ = melanogaster, simulans, or sechellia), $k$th sex ($k$ = male or female), and $r$th replicate ($r = 1, 2, 3$). The $l_j$ in this model is the effect of species, $s_k$ is the effect of sex, and $ls_{jk}$ is the species $\times$ sex interaction. Normality of the residuals ($\epsilon_{jk}$) for each probe set was checked using Shapiro–Wilk’s test. All effects with $P$ values below the false discovery rate (FDR) of 0.05 (Benjamini and Hochberg 1995) were considered significant (for a review, see Verhoeven et al. 2005).

We first examined each probe set for the effects of species, sex, and species $\times$ sex interaction. Genes that showed significant sex or species $\times$ sex interaction effects were considered sexually dimorphic. Further analyses were performed separately for sexually dimorphic and monomorphic genes. For the sexually monomorphic genes, we used a reduced ANOVA model $Y_{im} = \mu + l_i + \epsilon_{im}$ to test for the effect of species in each of 4 separate pairwise comparisons: D. sechellia versus D. simulans, D. sechellia versus D. melanogaster, D. simulans versus D. melanogaster, and D. melanogaster W158 versus D. melanogaster W189. Within each genotype, gene expression measurements from both sexes were combined for this analysis. The same 4 pairwise comparisons were performed for the sexually dimorphic genes using the same reduced ANOVA model, but the ANOVA was performed separately for males and females.

Confirmation of Microarray Data by Quantitative Polymerase Chain reaction

To confirm the microarray data, we measured the expression of 34 antennal OR and OBP genes using quantitative rt–polymerase chain reaction (quantitative rt-PCR) as described in the Supplementary Information online. We then compared the microarray and qPCR results for each gene and for each pair of species, for a total of 99 tests (supplementary table 1, Supplementary Material online). The qPCR data were corrected for multiple testing using an FDR level of 0.05. In 75 out of 99 comparisons, qPCR and microarray hybridization showed expression differences in the same direction; most of these comparisons were significant at FDR = 0.05 according to both microarray and qPCR results (supplementary table 1, Supplementary Material online). This level of concordance compares favorably with the results of cross-platform validation experiments in which identical RNA samples were examined on different microarrays (Bammler et al. 2005; Dallas et al. 2005; Morey et al. 2006; Wang et al. 2006; Stafford and Brun 2007) or quantified by microarrays and massively parallel sequencing (Liu et al. 2007). Given the different technical biases of microarray hybridization and qPCR, we have strong evidence that the genes showing concordant expression differences in both analyses have diverged among species and strains. In cases where qPCR and microarray data disagree (~25% of comparisons), no conclusion can be drawn about the direction and magnitude of change because we do not know which method provides a more accurate estimate of transcript abundance under our experimental conditions.

Reconstruction of Ancestral Gene Expression Levels

Ancestral gene expression states were reconstructed from the microarray data using AncML v 1.0 (Schluter et al. 1998). For each species and strain, gene expression levels were estimated as the least-squares mean of normalized, log-transformed signal intensities across all samples (6 samples each for D. sechellia, D. simulans, and each strain of D. melanogaster). Male and female samples were combined for this analysis. A neutral phylogenetic tree for reconstructing ancestral character states was obtained by selecting from published genome sequences 10 noncoding regions evenly spaced along each of the major chromosome arms (50 loci total, mean length ~500 bases). Sequences from Drosophila yakuba were used as outgroups to root
the trees. Phylogenetic reconstruction was performed for each locus using MrBayes version 3.1.2 (Huelsenbeck and Ronquist 2001) under a GTR + Γ model. We chose this general model rather than estimating a specific model for each locus to ensure that branch lengths were comparable across genes. We did not observe any particular biases among the loci in either tree topology or branch lengths. For ancestral reconstruction, we used the median branch length across all loci for each species so that the tree was as credible across genes. We did not observe any particular biases among the loci in either tree topology or branch lengths.

Ancestral character reconstruction relies on a number of assumptions, including that all characters are independent and that each character follows a Brownian motion model—that is, each subsequent change in character value is independent of all previous changes (Felsenstein 1985). The first assumption is almost certainly violated for gene expression, which is controlled by a limited number of transcription factors. The second assumption may be violated, as well, for genes evolving under directional selection. Thus, the results of comparative analysis for individual genes should be interpreted with caution. Finally, because D. sechellia and D. simulans were each represented by only 1 strain in this study, inferences of interspecific divergence may be confounded with intraspecific variation. Despite these qualifications, genome-wide patterns of expression divergence revealed by this analysis may provide valuable insights into the evolutionary history of each species (e.g., Holloway et al. 2007).

### Results and Discussion

Extensive Interspecific Differences in Olfactory Gene Expression

A total of 51 out of 60 known OR genes and 38 out of 53 annotated OBPs were detected in RNA samples isolated from the third antennal segment (table 1 and supplementary tables 2 and 3 [Supplementary Material online]). This includes 38 out of 38 ORs and 15 out of 16 OBPs with previously described expression patterns in the antennae (Galindo and Smith 2001; Couto et al. 2005). OR transcripts are present in the third antennal segment at levels similar to the genome-wide distribution, whereas OBPs transcripts are more abundant than the genome average (for details, see Supplementary Information online). Thus, our comparative analysis captures the entire or nearly entire set of antennal OR and OBP genes.

Gene expression in the antennae has diverged extensively among the 3 species. For the genome as a whole, 80.2% of genes show a significant effect of genotype in the ANOVA at $FDR = 0.05$ (table 1), indicating that the expression levels of these genes differ significantly across species. In pairwise comparisons, 24.7% of genes differ between D. sechellia and D. simulans, 59.1% differ between D. sechellia and D. melanogaster, 55.8% differ between D. simulans and D. melanogaster, and 14.6% differ between the 2 isogenic strains of D. melanogaster at $P < 0.001$ (table 1).

Divergence of olfactory gene expression is accelerated relative to genome-wide expression divergence (Fisher’s exact test, $P = 0.0004$). A total of 92% of detected ORs (46 out of 51) and 97.4% of OBPs (37 out of 38) showed a significant effect of genotype at $FDR = 0.05$. Among these genes, 14 ORs and 19 OBPs differ between D. sechellia and D. simulans, 19 ORs and 21 OBPs differ between D. sechellia and D. melanogaster, 27 ORs and 20 OBPs differ between D. simulans and D. melanogaster, and 10 ORs and 3 OBPs differ between the 2 isogenic strains of D. melanogaster at $P < 0.001$. Of the 34 OR and OBP genes examined by qPCR (see Materials and Methods), 20 showed significant differences in expression between D. sechellia

### Table 1

<table>
<thead>
<tr>
<th>Gene Class</th>
<th>Total Number of Detected Genes</th>
<th>Number and Percentage of Genes Significant for the Main Effect of Genotype at FDR &lt; 0.05 across All Samples</th>
<th>Number and Percentage of Genes Significant for the Main Effect of Genotype at $P &lt; 0.001$ in Pairwise Comparisons</th>
</tr>
</thead>
<tbody>
<tr>
<td>ORs</td>
<td>51</td>
<td>46 (92.0%)</td>
<td>14 (30.4%) Drosophila sechellia versus Drosophila simulans, 19 (41.3%) Drosophila sechellia versus Drosophila melanogaster, 27 (58.7%) Drosophila simulans versus D. melanogaster, 10 (21.7%) Between Strains of D. melanogaster</td>
</tr>
<tr>
<td>OBPs</td>
<td>38</td>
<td>37 (97.4%)</td>
<td>19 (50.0%) Drosophila sechellia versus Drosophila simulans, 21 (55.3%) Drosophila sechellia versus Drosophila melanogaster, 20 (52.6%) Drosophila simulans versus D. melanogaster, 3 (7.9%) Between Strains of D. melanogaster</td>
</tr>
<tr>
<td>ORs and OBPs</td>
<td>89</td>
<td>83 (93.5%)</td>
<td>33 (37.1%) Drosophila sechellia and D. simulans, 40 (44.9%) Drosophila sechellia and D. melanogaster, 47 (52.8%) Drosophila simulans and D. melanogaster, 13 (14.6%) Between Strains of D. melanogaster</td>
</tr>
<tr>
<td>Whole genome</td>
<td>9,830</td>
<td>7,882 (80.2%)</td>
<td>1,946 (24.7%) Drosophila sechellia versus Drosophila simulans, 4,657 (59.1%) Drosophila sechellia versus Drosophila melanogaster, 4,399 (55.8%) Drosophila simulans versus D. melanogaster, 1,153 (14.6%) Between Strains of D. melanogaster</td>
</tr>
</tbody>
</table>
and D. simulans, 23 differed between D. sechellia and D. melanogaster, 17 differed between D. simulans and D. melanogaster, and 15 differed between the 2 strains of D. melanogaster. Thus, in contrast to electrophysiological analysis (Stensmyr et al. 2003; Dekker et al. 2006), we find that significant changes in the olfactory system are not limited to the specialist D. sechellia and may affect ecologically similar generalist species such as D. simulans and D. melanogaster.

Few genes showed evidence of sex-biased expression in the antennae; however, the proportion of sexually dimorphic genes was higher for olfactory genes than for the genome as a whole. At a lenient significance threshold (FDR\textunderscore 0.2), 32% of ORs and 11.8% of OBPs were expressed in a sex-biased manner, compared with the genome-wide level of 4.1% (Fisher’s exact test, P < 0.0001). At least some of these differences may be attributed to the quantitative differences between males and females in the relative number of different sensilla subtypes (Shanbhag et al. 1999). We found that sexually dimorphic genes showed greater interspecific divergence and that this was true both for olfactory and for other genes. For the genome as a whole, 79.7% of sexually monomorphic genes but 91.9% of sexually dimorphic (at FDR = 0.2) genes showed a significant effect of genotype (Fisher’s exact test, P < 0.0001). For the olfactory genes, the corresponding proportions were 92.2% and 100% (table 2). Increased divergence of sex-biased genes was observed in all 3 interspecific comparisons (table 2).

Evolution of Olfactory Gene Expression in D. sechellia

Differences in gene expression between 2 extant species can, in principle, be due to evolutionary changes in either, or both, phylogenetic lineages following their divergence from their latest common ancestor. Drosophila sechellia is an island endemic with narrow host plant specificity, suggesting that olfactory gene expression may have undergone greater changes on the D. sechellia branch than on the D. simulans branch. Our analysis confirms this prediction. A total of 37.1% of ORs and OBPs differ between D. sechellia and D. simulans at P < 0.001, compared with 24.7% for the genome as a whole (table 1) (Fisher’s exact test, P < 0.0001). In contrast, the corresponding numbers are 44.9% versus 59.1% for the D. sechellia–D. melanogaster comparison, 52.8% versus 55.8% for the D. simulans–D. melanogaster comparison, and 14.6% versus 14.6% for the comparison between 2 strains of D. melanogaster (table 1).

We used a phylogenetic approach to explore the origins of gene expression differences between D. sechellia and D. simulans. For each gene, we reconstructed the mean and standard deviation (SD) of gene expression level in the latest common ancestor of these species in a maximum likelihood framework (see Materials and Methods for a discussion of this approach). We then asked whether the expression of each gene has undergone a greater absolute change in the D. sechellia or in the D. simulans lineage. On average, we expect greater changes in D. sechellia than in D. simulans under any mode of evolution as D. sechellia has the longer branch in the phylogenetic tree used for character reconstruction. Indeed, for the genome as a whole, 74% of genes show greater absolute change in D. sechellia (table 3). Among ORs and OBPs, however, 94% show greater change on the D. sechellia than on the D. simulans branch. This proportion is significantly greater than the genome average (Fisher’s exact test, P = 0.0029). For 25 OR and OBP genes, the amount of change on the D. sechellia lineage exceeded 1 SD of the estimated ancestral value and for 15 genes it exceeded 2 SDs. In contrast, the amount of change on the D. simulans branch exceeded 1 SD of the ancestral estimate for only 14 genes and none exceeded 2 SDs (table 3). This analysis suggests that olfactory gene expression has evolved considerably faster on the branch leading to D. sechellia than on the D. simulans branch.

We also compared the absolute amount of change that occurred on the D. sechellia branch between olfactory and nonolfactory genes. This distribution is considerably steeper for olfactory genes, indicating that the evolution of OR and OBP expression has been accelerated in D. sechellia compared with the genome average (fig. 1). A similar though weaker trend is also observed on the D. simulans branch (data not shown), suggesting that olfactory genes may have a general tendency for rapid expression divergence. On average, OBPs have undergone greater changes than ORs (data not shown). These differences remain even after divergence is normalized by mean expression levels, showing that accelerated divergence among some genes is not an artifact of their higher overall expression.

Table 2

<table>
<thead>
<tr>
<th>Effect of Sexual Dimorphism on Interspecific Divergence of Gene Expression</th>
<th>Total Number of Detected Genes</th>
<th>Number and Percentage of Genes Significant for the Main Effect of Genotype at FDR &lt; 0.05 across All Samples</th>
<th>Number and Percentage of Genes Significant for the Main Effect of Genotype at P &lt; 0.001 in Pairwise Comparisons</th>
</tr>
</thead>
<tbody>
<tr>
<td>Drosophila sechellia versus Drosophila simulans</td>
<td>406</td>
<td>373 (91.9%)</td>
<td>111 (27.3%)</td>
</tr>
<tr>
<td>Drosophila sechellia versus Drosophila melanogaster</td>
<td>9,424</td>
<td>7509 (79.7%)</td>
<td>1835 (19.5%)</td>
</tr>
<tr>
<td>Drosophila simulans versus D. melanogaster</td>
<td>20</td>
<td>20 (100%)</td>
<td>10 (50%)</td>
</tr>
<tr>
<td>D. sechellia versus D. simulans</td>
<td>64</td>
<td>59 (92.2%)</td>
<td>20 (33.9%)</td>
</tr>
</tbody>
</table>
The amount of expression change on the *D. sechellia* and *D. simulans* branches is positively correlated both for olfactory genes ($r = 0.72, P < 0.0001$) and for the genome as a whole ($r = 0.54, P < 0.0001$) (fig. 2). Because the ancestral values are reconstructed rather than measured independently, positive correlation between branches is expected under most evolutionary scenarios. However, several ORs and OBPs have undergone a large amount of change on the *D. sechellia* branch but virtually none on the *D. simulans* branch. Such disparity is unlikely under neutral evolution, suggesting that these genes may be involved in the unique ecological adaptations of this species. The most prominent example of such genes is Or22a (arrow in fig. 2). The receptor encoded by this gene is sensitive to methyl hexanoate, consistent with an increased sensitivity to this chemical in *D. sechellia* (Dekker et al. 2006).

OR expression divergence also correlates with protein sequence evolution. McBride (2007) has estimated lineage-specific nucleotide and amino acid substitution rates for the OR genes in *D. sechellia* and *D. simulans*. We compared the estimated amount of change in gene expression on the *D. sechellia* branch with the rate of amino acid replacement ($K_a$) on the same branch. The 2 parameters are positively correlated ($r = 0.38, P = 0.0071$) (fig. 3A). In contrast, we see no correlation between the rate of silent nucleotide substitutions ($K_s$) and the amount of gene expression change on the *D. sechellia* branch ($r = -0.04, P = 0.598$) (fig. 3B). These observations suggest that rapid evolution of OR expression in *D. sechellia* may have been driven by positive selection, although relaxed purifying selection (e.g., reduced constraint on amino acid composition) is also a plausible explanation. In this analysis as well, Or22a is highly unusual. Expression of this gene has changed dramatically in *D. sechellia*, even though it did not undergo any changes in protein sequence (arrow in fig. 3A).

***Table 3***  
Gene Expression Changes on the *Drosophila sechellia* and *Drosophila simulans* Branches

<table>
<thead>
<tr>
<th>Phylogeny Branch</th>
<th>Between <em>D. sechellia</em> and MRCA of <em>D. simulans</em></th>
<th>Between <em>D. simulans</em> and MRCA of <em>D. sechellia</em> and <em>D. simulans</em></th>
</tr>
</thead>
<tbody>
<tr>
<td>All genes</td>
<td>Greater absolute change on branch 2,944 (73.7%)</td>
<td>1,051 (26.3%)</td>
</tr>
<tr>
<td></td>
<td>Expression change on branch exceeds 1 SD of MRCA estimate 2,881 (72.1%)</td>
<td>1,848 (46.3%)</td>
</tr>
<tr>
<td></td>
<td>Expression change on branch exceeds 2 SDs of MRCA estimate 1,115 (27.9%)</td>
<td>0</td>
</tr>
<tr>
<td>OR and OBP genes</td>
<td>Greater absolute change on branch 31 (93.9%)</td>
<td>2 (6.1%)</td>
</tr>
<tr>
<td></td>
<td>Expression change on branch exceeds 1 SD of MRCA estimate 25 (75.8%)</td>
<td>14 (42.4%)</td>
</tr>
<tr>
<td></td>
<td>Expression change on branch exceeds 2 SDs of MRCA estimate 15 (45.5%)</td>
<td>0</td>
</tr>
</tbody>
</table>

**Fig. 1.** Amount of gene expression change on the *Drosophila sechellia* branch for all transcripts in the genome (gray line) and for OR and OBP genes (black line). Expression change ($y$ axis) is expressed in standardized units that refer to log-transformed expression levels from the microarray analysis.

**Fig. 2.** Amount of gene expression change on the *Drosophila sechellia* and *Drosophila simulans* branches for all transcripts in the genome (A) and for OR and OBP genes (B). Black lines indicate 1:1 diagonals. Expression divergence is expressed as in figure 1. Arrow indicates Or22a.
We find that OBPs, like the OR genes (McBride 2007; Tunstall et al. 2007), evolve more rapidly in *D. sechellia* than in *D. simulans* (supplementary tables 4 and 5, Supplementary Material online). In contrast to OR genes (fig. 3), we find no significant correlation between the rate of expression divergence and amino acid replacement rate among OBP genes (*r* = 0.10, *P* = 0.296). Thus, protein sequence and expression divergence of OBPs do not appear to be under the same selective regime as they are for OR genes. Obp50a and Obp56c may be an exception as they show elevated *dS/dN* ratios in *D. sechellia* but not in *D. simulans* (supplementary table 5, Supplementary Material online). Interestingly, Obp50a expression is greatly increased in *D. sechellia* relative to the other species (see below), suggesting that this gene could be involved in sensory specialization.

Gene Expression Changes in ab2 and ab3 Sensilla

Electrophysiological analysis shows that large basiconic sensilla of the ab2 subtype are present in all species of the *melanogaster* subgroup except *D. sechellia*. The loss of ab2 receptors in this species is accompanied by an increase in the number of ab3 receptors (Stensmyr et al. 2003; Dekker et al. 2006). One might expect that the ORs expressed in ab2 neurons (Or59b and Or85a) would not be expressed in *D. sechellia*, whereas the ORs expressed in ab3 neurons (Or22a, Or22b, and Or85b) would be overexpressed in *D. sechellia* but expressed at similar levels in *D. simulans* and *D. melanogaster*. We find, however, that Or59b and Or85a do not differ significantly among the 3 species (fig. 4A). Thus, there is no strict correspondence between the levels of OR expression, as measured in the entire third antennal segment, and the prevalence of physiologically defined receptor subtypes. In the future, it will be interesting to determine where Or59b and Or85a are expressed in *D. sechellia* and what functions they perform in this species. One possibility is that they have become expressed in the ab1 sensilla because ab1A neurons in *D. sechellia* show a strong response to ethyl-3-hydroxybutyrate, a key ligand for Or85a and the ab2B neuron in which it is expressed in *D. melanogaster* (Hallem et al. 2004; Dekker et al. 2006). Consistent with this explanation, Dekker et al. (2006) found the response spectra of large basiconic sensilla to be considerably more variable in *D. sechellia* than in *D. melanogaster*.

The expression of Or22a, Or22b, and Or85b is markedly increased in *D. sechellia*, but Or22a and Or22b also differ strongly between *D. simulans* and *D. melanogaster*, and Or22a shows significant differences between the 2 strains of *D. melanogaster*, as well (fig. 4A and supplementary table 1 [Supplementary Material online]). This suggests that some evolutionary changes in gene expression in the ab3 sensilla may have occurred prior to the divergence between *D. sechellia* and *D. simulans*. Importantly, interspecific divergence in Or22a, Or22b, and Or85b expression cannot be fully explained by the increase in the number of ab3 sensilla in *D. sechellia*. Or22a and Or22b are upregulated in both *D. sechellia* and *D. simulans* compared with *D. melanogaster*, whereas Or85b is only upregulated in *D. sechellia* and does not differ significantly between *D. simulans* and *D. melanogaster* (fig. 4A). Thus, the ratio of Or22a to Or85b expression is higher in *D. sechellia* and *D. simulans* than in *D. melanogaster* (*t*-test, *P* < 10−6) (fig. 4B). This change may have important implications for chemotactic behavior, although it is not immediately clear how to reconcile it with neurophysiological observations. Or22a is highly sensitive to both methyl hexanoate and ethyl hexanoate, whereas Or85b is much more sensitive to methyl hexanoate (Hallem and Carlson 2006) so that increased expression of Or22a relative to Or85b would be expected to increase the ab3 chemoreceptor’s response to ethyl hexanoate compared with methyl hexanoate. In fact, the opposite pattern is observed as the ab3 sensilla show increased sensitivity to methyl hexanoate relative to ethyl hexanoate in *D. sechellia* and *D. simulans* compared with *D. melanogaster* and other tested *Drosophila* species (Stensmyr et al. 2003).

One potential explanation may involve Or22b. This gene has undergone a large deletion resulting in a loss of function in the *D. sechellia* lineage (McBride 2007) and has experienced positive selection in the *D. simulans* lineage (Tunstall et al. 2007). Dobritsa et al. (2003) have found that Or22a is sufficient for all olfactory functions of the
ab3A neuron in the Canton-S strain of *D. melanogaster*, whereas Or22b does not rescue or change sensitivity of that neuron and appears nonfunctional. However, both Or22a and Or22b encode functional receptors in a different wild-type strain of *D. melanogaster*, Oregon-R (Ray et al. 2007). Moreover, the evolutionary conservation of Or22b in *D. simulans* as well as more distantly related species such as *Drosophila pseudoobscura* suggests that Or22b is likely to perform some functional role. It is possible that changes in Or22b have contributed to the increased sensitivity to methyl hexanoate in *D. sechellia* and *D. simulans*. In any event, it appears that at least some genetic changes that contribute to the distinct sensory profile of *D. sechellia* may actually have occurred in the common ancestor of *D. simulans* and *D. sechellia*. The role of these changes in the adaptation of *D. sechellia* to Morinda fruit remains to be tested by behavioral experiments.

Evolution of other Olfactory Genes

In general, more ORs and OBPs are overexpressed than underexpressed in *D. sechellia*. Of the genes analyzed by both microarrays and qPCR, only Or9a and Obp99c showed decreased expression in *D. sechellia* (fig. 5). In contrast, 6 genes (Or35a, Or49a, Or82a, Or85c, Obp50a, and Obp83ef) were upregulated in *D. sechellia*. All these genes also showed significant differences in expression between *D. simulans* and *D. melanogaster*. In most cases (Or85c, Or49a, Obp50a, and Obp83ef), expression level in *D. simulans* was intermediate between *D. sechellia* and *D. melanogaster* (fig. 5). However, Or82a did not differ between *D. sechellia* and *D. simulans*, whereas Or35a expression was strongly decreased in *D. simulans* compared with both *D. sechellia* and *D. melanogaster* (fig. 5). Finally, several genes (Or49a, Or85c, Obp83ef, and Obp99d) showed significant differences between the 2 strains of *D. melanogaster* (fig. 5). Thus, we find evidence for evolutionary changes in olfactory gene expression on all branches of the phylogeny.

The OR genes upregulated in *D. sechellia* are expressed in coeloconic (Or35a) and small basiconic (Or82a and Or49a) sensilla (Couto et al. 2005; Yao et al. 2005). Or82a is a narrowly tuned receptor that shows strong response only to geranyl acetate among the tested odorants.
Hallem and Carlson 2006). Or35a is a broadly tuned receptor with a spectrum that is generally similar to Or22a and Or85b and includes medium-chain alcohols and aldehydes (e.g., hexanol, hexanal, and their unsaturated equivalents) as well as some short-chain esters (including hexyl acetate, hexenyl acetate, and hexyl butyrate). However, Or35a does not respond to medium-chain esters such as methyl hexanoate or other hexanoate or octanoate esters (Hallem and Carlson 2006). With the exception of Or22a, none of the ORs upregulated in D. sechellia show sensitivity to hexanoic or octanoic acids. This is consistent with the neurophysiological analysis that found no differences between D. melanogaster and D. sechellia in olfactory sensitivity to the hexanoic acid (Dekker et al. 2006), despite its importance in D. sechellia’s chemotaxis and oviposition behavior (Higa and Fuyama 1993; Matsuo et al. 2007). These results suggest that the adaptation of D. sechellia to its host involved changes in its behavioral response to hexanoic and octanoic acids, rather than increased ability to detect these compounds.

Four ORs expressed in trichoid sensilla have been implicated in the reception of pheromones. Or47b and Or88a are sensitive to both male and female cuticular extracts, whereas Or67d and Or65a respond specifically to the male pheromone cis-vaccenyl acetate (Kurtovic et al. 2007; van der Goes van Naters and Carlson 2007). We find that Or47b
and Or67d are expressed at a slightly higher level in males than in females, whereas Or88a and Or65a show no evidence of sex-biased expression (supplementary table 1, Supplementary Material online). This observation cannot be fully explained by differences in the abundance of different trichoid sensilla. Or67d is the only OR expressed in the at1 subtype. However, Or47b, Or65a, and Or88a are all expressed in the at4 subtype (Couto et al. 2005). This suggests that Or47b may be regulated by the sex determination pathway independently of other genes expressed in the at4 sensilla and that males may have more copies of Or47b per neuron than females. Interestingly, Or47b shows by far the highest $K_s/K_a$ ratio on the D. sechellia branch of all OR genes (1.28), whereas Or67d, Or88a, and Or65a have medium to low $K_s/K_a$ ratios (0.07–0.19). Thus, it is possible that Or47b has evolved under sexual selection.

No Evidence for Coordinated Evolution of Functionally Similar or Physically Clustered Genes

To test whether OR genes with similar functional specificities show coordinated evolutionary changes in expression, we first estimated pairwise covariances and correlations of gene expression levels using an independent contrasts approach (Felsenstein 1985, 2005). This allowed us to examine the correlations of gene expression across species after accounting for their phylogenetic relationships. Many ORs and OBPs occur in tandem clusters in the genome (Graham and Davies 2002; Hekmat-Scafe et al. 2002). We find that physically clustered genes do not show any consistent tendency to coevolve in expression (supplementary fig. 1, Supplementary Material online). The qPCR analysis confirms this observation. For example, Obp50a differs very strongly in expression between D. sechellia and the other 2 species (supplementary table 1 and fig. 4, Supplementary Material online). In contrast, the adjacent genes Obp50b and Obp50c do not show any interspecific differences (data not shown).

There is also no evidence for coordinated changes in the expression of receptors expressed in the same olfactory organ subtype. For example, Or43b, expressed in the basiconic a/b sensilla (Couto et al. 2005), shows a negative correlation (−0.7) with the other a/b gene, Or9a, but a high positive correlation (0.99) with Or47b, which is expressed in the trichoid at4 sensilla (supplementary fig. 1, Supplementary Material online). This indicates that not all interspecific divergences in OR expression can be attributed to changes in the relative numbers of different sensilla subtypes. Rather, increased or decreased expression of at least some OR genes must be due to evolutionary changes either in their cis-regulatory elements or in the transcription factors that control their expression in olfactory neurons. Consistent with this model, the development of different receptors housed in the same sensilla can be controlled by different transcription factors (Ray et al. 2007).

Hallem and Carlson (2006) examined the response of 24 OR genes to each of 110 volatile compounds. Functional similarities among these ORs can be represented as pairwise Euclidean distances in a 110-dimensional space (Hallem and Carlson 2006). We tested whether these distances correlated with the phylogenetic covariances of gene expression levels estimated by the independent contrasts method (supplementary fig. 2, Supplementary Material online). No significant correlation was observed ($r = −0.012$, rank correlation $= −0.002$, $P = 0.522$). Thus, we find no evidence for coordinated evolution of functionally similar ORs.

Some ORs respond to a wide range of chemicals, whereas others respond to only a few (Hallem and Carlson 2006). We asked whether the more narrowly or the more broadly tuned receptors show greater evolutionary changes in expression. Breadth of tuning was quantified as the number of chemicals that produced an excitation response above 100 spikes per second (other cutoffs, ranging from 50 to 150 spikes per second, produced similar results). Gene expression divergence was quantified as the variance of mean expression levels across the 4 genotypes (D. sechellia, D. simulans, and the 2 strains of D. melanogaster). We found a positive correlation between expression divergence and the number of odorants eliciting receptor response ($r = 0.46$, rank correlation $= 0.44$, $P = 0.011$) (supplementary fig. 3A, Supplementary Material online). Separately, we quantified gene expression divergence as the estimated amount of change on the D. sechellia branch. Again, we found a positive correlation between the breadth of tuning and the rate of divergence ($r = 0.59$, rank correlation $= 0.61$, $P = 0.0012$) (supplementary fig. 3B, Supplementary Material online). However, both correlations are due to only 5 rapidly evolving ORs with the broadest functional specificity: Or67a, Or22a, Or85b, Or9a, and Or35a. Aside from these genes, there is not consistent relationship between the breadth of tuning and the rate of expression divergence.

In sum, these 4 analyses show that the patterns of gene expression divergence are not by-products of genomic location, anatomical organization, or broad functional specificity. It is likely that at least some evolutionary changes, particularly the increased expression of several olfactory genes, reflect a specific adaptation of D. sechellia to a new host plant. Behavioral experiments will be needed to test this hypothesis.

Conclusions

We have carried out a genome-wide survey of interspecific divergence in olfactory gene expression between D. sechellia, an island endemic with narrow host plant specificity, and its generalist relatives D. simulans and D. melanogaster. Our results reveal a number of general evolutionary patterns and suggest several candidate genes that may have contributed to the evolution of host plant preferences.

1. Evolution of OR and OBP expression was accelerated in D. sechellia compared both with the genome average in that species and with the rate of OR and OBP evolution in the other species. Positive selection may have played a role in the accelerated evolution of OR expression, although relaxed purifying selection is also a plausible explanation in some cases.
2. Interspecific divergence in olfactory gene expression cannot be fully explained by changes in the relative abundance of different sensilla. Expression of some ORs and OBPs has evolved independently of other genes expressed in the same sensilla.

3. There is no general tendency for functionally similar or physically clustered genes to coevolve in expression.

4. A number of OR and OBP genes are upregulated in D. sechellia compared with its generalist relatives. This striking result suggests that the evolution of host plant specialization in D. sechellia likely included gain-of-function changes in addition to the fixation of loss-of-function alleles. Genes with increased expression in D. sechellia include Or22a, which is highly sensitive to methyl hexanoate, as well as several genes that are yet to be characterized in detail.

5. Some of the gene expression changes that contribute to the distinct neurophysiological phenotype of D. sechellia may predate its divergence from its generalist relative D. simulans.

Supplementary Material

Supplementary tables 1–5 and figures 1–4 are available online (http://www.mbe.oxfordjournals.org/).

Acknowledgments

This work was supported by National Institute of Health (NIH) grant 1R24GM065513-01 to A.K. and L.M.M., National Science Foundation (NSF) grant IOB-0518654 to A.K., NIH grant 1R01GM077618-01A1 to L.M.M., and NSF Grant MCB-0618691 and Funds from the Carolina Center for Genome Sciences to C.D.J. The qPCR equipment was provided by the University of California—Davis Clinical Nutrition Research Unit (DK35747). We would like to thank Ryan Bickel, Deniz Erzyilmaz, Lindy McBride, Carson Miller, and 3 anonymous reviewers for comments that improved the manuscript.

Literature Cited


Morey JS, Ryan JC, Van Dolah FM. 2006. Microarray validation: factors influencing correlation between oligonucleotide microarrays and real-time PCR. Biol Proced Online. 8:175–193.


