Odorant-Binding Proteins OBP57d and OBP57e Affect Taste Perception and Host-Plant Preference in Drosophila sechellia

Takashi Matsuo*, Shigeru Sugaya†, Jyunichiro Yasukawa, Toshiro Aigaki, Yoshiaki Fuyama
Department of Biological Sciences, Tokyo Metropolitan University, Hachioji, Tokyo, Japan

Despite its morphological similarity to the other species in the Drosophila melanogaster species complex, D. sechellia has evolved distinct physiological and behavioral adaptations to its host plant Morinda citrifolia, commonly known as Tahitian Noni. The odor of the ripe fruit of M. citrifolia originates from hexanoic and octanoic acid. D. sechellia is attracted to these two fatty acids, whereas the other species in the complex are repelled. Here, using interspecies hybrids between D. melanogaster and D. sechellia, we showed that the Odorant-binding protein 57e (Obp57e) gene is involved in the behavioral difference between the species. D. melanogaster knock-out flies for Obp57e and Obp57d showed altered behavioral responses to hexanoic acid and octanoic acid. Furthermore, the introgression of Obp57d and Obp57e from D. simulans and D. sechellia shifted the oviposition site preference of D. melanogaster Obp57d/eKO flies to that of the original species, confirming the contribution of these genes to D. sechellia’s specialization to M. citrifolia. Our finding of the genes involved in host-plant determination may lead to further understanding of mechanisms underlying taste perception, evolution of plant–herbivore interactions, and speciation.

Introduction

Every animal must locate and identify sufficient food to meet its biological requirements. For herbivorous insects, this results in an endless battle with their host plants [1]. For example, some plants develop a chemical defense system that causes toxicity to generalist herbivores [2]. In response, generalist herbivores may then evolve a behavioral system to avoid such toxic plants. If an insect species acquires resistance to a plant toxin, however, it may reap an ecological advantage by gaining exclusive access to the toxic plant and may subsequently evolve as a specialist herbivore with a specific preference towards that plant. Such physiological and behavioral specialization plays an important role in the evolution of divergent ecological interactions between herbivores and their host plants. Nevertheless, it does not necessarily follow that ecological specialization for a particular host plant drives speciation of herbivores itself. Such specialization may not be sufficient to maintain divergence between populations at an early stage of speciation, in the face of potential gene flow via hybridization between evolving populations. The role of ecological specialization in speciation remains, therefore, to be proven [3]. Thus, it is necessary to identify the genes and molecular mechanisms responsible for ecological adaptation if we are to understand whether ecological adaptation can be a cause, or merely a consequence, of speciation [4].

Behavioral adaptation of herbivorous insects to their host plants involves the evolution of the chemosensory system [5–7]. With the recent identification of olfactory and gustatory receptors [8], knowledge of the genetic and molecular mechanisms of insect olfactory and gustatory system markedly progressed. Recent analysis of genomic information from several insect species has also revealed that the number of genes encoding these receptors varies considerably between species, indicating a close relationship between the genomic constitution of chemoreceptor gene families and the species-specific lifestyles of insects [9–11]. Thus, it is likely that the genes responsible for ecological adaptation are to be found among these receptor-encoding and receptor-related genes.

Genetic studies of Drosophila have also contributed to a substantial amount of our knowledge of “speciation genes” [4]. However, these studies have primarily focused on genes that cause reproductive isolation, and genetic analysis of ecological adaptation is relatively rare. This is, in part, due to the surprisingly limited information about Drosophila in the wild, compared with those flies reared in the laboratory as a sophisticated model system of genetics. In fact, we know little about their natural foods in the wild, except for a few species.

Drosophila sechellia is a specialist of Morinda citrifolia, which is commonly known as Tahitian Noni [12]. Although D. sechellia shows a preference for and resistance to the ripe fruit of M. citrifolia, its most closely related species, D. simulans and D. mauritiana, as well as D. melanogaster, are generalists and die upon contact with M. citrifolia, and thus avoid the fruit [13,14].

Citation: 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(5): e118. doi:10.1371/journal.pbio.0050118

Copyright: © 2007 Matsuo et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Abbreviations: bp, base pair; GFP, green fluorescent protein; HA, hexanoic acid; MH, methyl hexanoate; OA, octanoic acid; OBP, odorant-binding protein; ORF, open reading frame; RT-PCR, reverse-transcriptase polymerase chain reaction

* To whom correspondence should be addressed. E-mail: mts@comp.metro-u.ac.jp
† Current address: School of Medicine, Chiba University, Chiba-shi, Chiba, Japan

Academic Editor: Mohamed A. F. Noor, Duke University, United States of America
Received January 31, 2006; Accepted February 22, 2007; Published April 24, 2007

PLoS Biology | www.plosbiology.org May 2007 | Volume 5 | Issue 5 | e118
Because of genetic resources available for D. melanogaster and D. simulans, D. sechellia is an ideal organism with which to explore the genetics of ecological specialization. Analysis of quantitative trait loci (QTL) between D. sechellia and D. simulans has already identified the chromosomal regions responsible for the interspecies difference in resistance to the toxicity of M. citrifolia [15]. However, D. sechellia’s preference for M. citrifolia was explained only by the transformation of olfactory sensilla resulting in an increase of the ab3 subtype that responds to the host odorant methyl hexanoate (MH) [16]. These findings successfully describe the present status of D. sechellia’s specialization for M. citrifolia, but the evolutionary history, especially how an ancestral population started to use the toxic plant as its host, has been unexplained.

Here, for the first time, we have identified genes involved in D. sechellia evolution. These genes are responsible for the behavioral differences between species in their responses to hexanoic acid (HA) and octanoic acid (OA), the toxins contained in the ripe fruit of M. citrifolia, which give it its characteristic odor. Having identified the genetic factors constituting D. sechellia’s adaptation to M. citrifolia, we are now able to discuss more confidently whether host-plant specialization can drive D. sechellia speciation.

Results

Mapping of Locus Responsible for Interspecies Difference in Avoidance of HA

We previously reported that the behavioral difference (preference/avoidance) between D. sechellia and D. simulans in response to HA, one of the main components of odor from the ripe fruit of M. citrifolia, is controlled by at least one gene on the second chromosome [17]. Further analysis of the introgression lines between D. sechellia and the D. simulans second chromosome marker strain (net b sd pm) indicated that the behavioral difference is linked to the marker pm, which is on the distal end of the right arm of the second chromosome (I. Higa and Y. Fuyama, unpublished data).

Considering the fact that the overall structure of the second chromosome is conserved between D. simulans and D. melanogaster, we mapped the locus in more detail using a series of D. melanogaster deficiency strains lacking a terminal part of the right arm of the second chromosome. Because D. sechellia’s preference for HA is a recessive trait to D. melanogaster’s avoidance [17], the interspecies hybrids between D. sechellia and D. melanogaster deficiency strains that lack a region containing the responsible gene(s) were expected to show the D. sechellia–like phenotype, i.e., preference for HA.

Two deficiency strains, Df(2R)exu1 and Df(2R)AA21, showed preference for HA when they were crossed with D. sechellia, defining the responsible locus within a very small chromosomal region, in combination with Df(2R)exu2, which showed avoidance to HA when crossed with D. sechellia (Figure 1A). Because the break points of these deficiency chromosomes had been deduced from cytological observations, we determined the position of these break points precisely by PCR-direct sequencing of genomic DNA from hybrids between D. melanogaster deficiency strains and D. sechellia (Figure 1B). According to the left break point of Df(2R)exu1 and the left break point of Df(2R)exu2, the locus was narrowed down within about 200 kilobases (kb) of the genomic region that contains 24 predicted genes. There is no large deleted region in the Df(2R)AA21 chromosome around this area, which is inconsistent with the result that Df(2R)AA21 also showed preference for HA when crossed with D. sechellia. While examining the marker sequences used in break-point determination of Df(2R)AA21, however, we incidentally found that this chromosome has a small, ten–base pair (bp) deletion in the first exon (open reading frame [ORF]) of the Odorant-binding protein 57e (Obp57e) gene resulting in a frame-shift mutation (Figure 1C). Insect OBP is a protein secreted into the lymph of chemosensory hairs, and it has been shown to play a crucial role in chemosensation [18]. Thus, it seemed likely that Obp57e is a gene responsible for the interspecies difference in response to HA. However, when Obp57e ORF sequences from D. melanogaster, D. simulans, and D. sechellia are compared, there is no D. sechellia–specific alteration except for L11I, which does not affect the result of signal peptide–sequence prediction (Figure 1D). Thus, D. sechellia Obp57e ORF is supposed to be functionally intact, suggesting that the interspecies difference is not in the structure of the gene product, but rather in gene expression.

Altered Expression Control of Obp57e in D. sechellia

Quantitative reverse-transcriptase polymerase chain reaction (RT-PCR) analysis revealed that the level of Obp57e transcripts is higher in the legs of D. sechellia than in D. melanogaster and D. simulans (Figure 2). This could be due to an elevated transcription activity in particular cells and/or a widened expression pattern. According to the lacZ reporter experiment, D. melanogaster Obp57e is expressed only in four cells associated with chemosensory hairs on the fourth and fifth segments of each tarsus, the most terminal part of an insect leg [19]. We confirmed that as short as 450 bp of the upstream region of Obp57e completely reproduces the reported expression pattern (Figure 3A–3C). We then cloned the corresponding region from D. simulans and D. sechellia, and introduced it into D. melanogaster with a green fluorescent protein (GFP) reporter gene. The D. simulans sequence successfully reproduced the same expression pattern as observed in D. melanogaster (Figure 3D). However, the D. sechellia sequence failed to drive GFP expression in any parts of the fly body.
indicating that the function of the \( D. \) \textit{sechellia} sequence to promote gene expression is altered. Indeed, when the upstream sequence of \( \text{Obp57e} \) is compared between species, a 4-bp insertion was found in the \( D. \) \textit{sechellia} \( \text{Obp57e} \) upstream sequence (Figure 3H). GFP expression was restored by removing the inserted 4-bp nucleotides from the \( D. \) \textit{sechellia} sequence, showing that this 4-bp insertion abolishes the function of the \( D. \text{sechellia} \) \( \text{Obp57e} \) promoter sequence in \( D. \) \textit{melanogaster} (Figure 3F and 3G). Nevertheless, the results of GFP reporter experiments are inconsistent with that of quantitative RT-PCR analysis, thus, the exact expression pattern of \( \text{Obp57e} \) in \( D. \) \textit{sechellia} remains unclarified. Therefore, it is necessary to evaluate using more direct methods whether \( \text{Obp57e} \) is truly responsible for the interspecies difference in behavioral response to HA.

**Targeted Mutagenesis of \( \text{Obp57d/e} \) Knock-Out Flies**

We generated \( D. \) \textit{melanogaster} knock-out flies for \( \text{Obp57e} \), as well as for its neighbor \( \text{Obp57d} \), and for both \( \text{Obp57d} \) and \( \text{Obp57e} \), by gene targeting (Figure 4). The ends-out method was employed to achieve precise gene replacement in the gene-dense \( \text{Obp57d/e} \) region (Figure 4A). To avoid side effects on transcription of surrounding genes, the marker gene (3 kb) was excised by Cre recombinase, leaving only 34 bp of the \textit{loxP} sequence. Each donor construct was designed such that the \textit{ORF} was removed exactly from the \textit{ATG} translation initiation site, but a putative poly-A additional signal was left intact, ensuring the termination of residual transcription that may affect the expression of downstream genes via read-through events (Figure 4B).

The loss of transcripts from the targeted gene was confirmed by quantitative RT-PCR in each knock-out strain (Figure 2). We observed, however, an unexpected interaction between \( \text{Obp57d} \) and \( \text{Obp57e} \) in their expression control. The amount of \( \text{Obp57e} \) transcripts was higher in \( \text{Obp57dKO} \) flies than in the \( \text{w1118} \) control strain. On the other hand, the amount of \( \text{Obp57d} \) transcripts decreased in the legs of \( \text{Obp57eKO} \) flies. Because each knock-out strain lacks the intron and the \textit{ORF}, these regions may contain elements that influence the expression of downstream genes via read-through events (Figure 4B).

The loss of transcripts from the targeted gene was confirmed by quantitative RT-PCR in each knock-out strain (Figure 2). We observed, however, an unexpected interaction between \( \text{Obp57d} \) and \( \text{Obp57e} \) in their expression control. The amount of \( \text{Obp57e} \) transcripts was higher in \( \text{Obp57dKO} \) flies than in the \( \text{w1118} \) control strain. On the other hand, the amount of \( \text{Obp57d} \) transcripts decreased in the legs of \( \text{Obp57eKO} \) flies. Because each knock-out strain lacks the intron and the \textit{ORF}, these regions may contain elements that influence the expression of the other gene.

**Altered Behavioral Responses to HA and OA in the Knock-Out Flies**

Each knock-out strain responded to HA differently from the control strain in the trap assay (Figure 5). \( \text{Obp57dKO} \) and

**Figure 1.** The Locus Responsible for Interspecies Difference in HA Avoidance Is Mapped to \( \text{Obp57e} \)

\((A)\) Behavioral screening of interspecies hybrids between \( D. \) \textit{melanogaster} deficiency strains and \( D. \) \textit{sechellia}. Response to HA was measured by the trap assay [17]. Response index \((\text{RI}) = (N_h - N_w)/(N_h + N_w)\), where \( N_h \) is the number of individuals trapped in 0.5% HA solution, and \( N_w \) is that of individuals trapped in distilled water. Error bars indicate 95% confidence intervals determined by the binomial test of summed data from five replications of the assay with 100 females for each replication.

\((B)\) Determination of break points in deficiency chromosomes. A filled circle indicates that the deficiency-chromosome–derived sequence was detected, and an open circle indicates that the deficiency-chromosome–derived sequence was not detected at that position.

\((C)\) The \( \text{DF(2R)AA21} \) chromosome has a 10-bp deletion in the first exon of the \( \text{Obp57e} \) gene. A genomic sequence of \( \text{DF(2R)AA21} \) is aligned with that of the wild-type strain (CS). Predicted ORFs are boxed and capitalized. Arrows indicate the position and direction of translation start sites (ATG).

\((D)\) Comparison of \( \text{Obp57e} \) structure between \( D. \) \textit{melanogaster} (mel), \( D. \) \textit{simulans} (sim), and \( D. \) \textit{sechellia} (sec). Predicted signal peptide sequence is boxed. Altered amino acid residues are highlighted.

\( \text{doi:10.1371/journal.pbio.0050118.g001} \)
Obp57eKO avoided HA, whereas females of Obp57deKO preferred it, suggesting that not only Obp57e, but also Obp57d, is involved in the behavioral difference observed in the screening assay. In fruit flies, host plants are largely determined by the oviposition site preference of adults. Thus, we also examined the oviposition site preference of knock-out flies in response to HA. Indeed, Obp57eKO and Obp57deKO seem to prefer lower concentrations of HA than the control flies, although the difference is not statistically significant (Figure 6, Tables 1–4). The direction of behavioral alteration was, however, not the same as that found in the trap assay for Obp57deKO. We also examined oviposition site preference in response to OA, the main toxic component in Morinda fruit. Because of its toxicity at high concentrations, the oviposition assay was carried out at concentrations lower than those of HA. Obp57dKO and Obp57eKO preferred higher concentrations of OA. This preference was particularly obvious for Obp57dKO, which was comparable to that of D. sechellia. Contrary to the responses to HA and OA, knock-out strains preferred concentrations of acetic acid and butyric acid similar to those preferred by control flies, showing that the alteration of behavioral responses in these knock-out strains is specific to HA and OA.

Our observation of the behavior of Obp57dKO, Obp57eKO, and Obp57deKO revealed that these strains are qualitatively different from each other in their responses to HA and OA. This strongly suggests that Obp57d, as well as Obp57e, is involved in D. sechellia’s behavioral adaptation to M. citrifolia. Nevertheless, none of these knock-out strains was identical to D. sechellia in behavior. This is consistent with the results of quantitative RT-PCR analysis in which no knock-out strain exhibited an expression profile identical to that of D. sechellia, proving that this species is not a simple null mutant of Obp57d and/or Obp57e. Rather, D. sechellia seems to be a neomorphic mutant with an altered expression control of these genes.

Replacement of Obp57d/e Region Altered Oviposition Behavior

To examine the functions of Obp57d and Obp57e in D. simulans and D. sechellia, we cloned these genes from D. simulans and D. sechellia and introduced them into the D. melanogaster Obp57d/eKO strain. Because an interaction between the two genes was observed with respect to their expression control, a genomic fragment spanning both Obp57d and Obp57e was used for genetic transformation. The resulting transformant flies showed altered responses to HA and OA in the oviposition site–preference assay (Figure 6; Tables 3 and 4). Obp57deKO; simObp57d/e flies avoided HA as D. simulans does. Conversely, Obp57deKO; secObp57d/e flies preferred high concentrations of OA as D. sechellia does. These results clearly showed that the Obp57de genomic region contains genetic information responsible for, at least in part, the interspecies differences in behavioral responses to HA and OA.

However, these transgenic flies are not complete mimics of the original species. Although D. simulans avoided OA, as well as HA, the response of Obp57deKO; simObp57de/e flies to OA was not significantly different from that of the D. melanogaster control strain (Figure 6; Table 4). The responses of these two transgenic strains in the trap assay were also different from...
that of the original species (Figure 5). Consistent with the results of the oviposition assay, *D. simulans* avoided HA and *D. sechellia* preferred it. Obp57d/eKO; simObp57d/e females, however, did not avoid HA, and both sexes of Obp57d/eKO; secObp57d/e flies did not prefer it. Indeed, the expression profiles of *Obp57d* and *Obp57e* were not exactly the same between the transgenic strains and the corresponding original species (Figure 2). Although the genomic fragments seemed to reproduce the native expression better than the GFP reporters, there still remains significant differences in expression profile, particularly between *Obp57d/eKO*; simObp57d/e and *D. simulans*. These differences suggest a contribution of additional loci to *Obp57d/e* expression, and thus to the interspecies differences in behavioral responses to HA and OA.

Nevertheless, the *Obp57d/e* genomic region from *D. simulans* and *D. sechellia* could reproduce, at least in part, the behavioral pattern of the original species in an otherwise *D. melanogaster* genomic background, proving that a genetic difference in this region is actually involved in interspecies differences in behavioral responses to odorants contained in *M. citrifolia*. It should be particularly noted that the *Obp57d/e* region is alone sufficient for the strong avoidance of HA by *D. simulans*, which is a key trait in the evolution of *D. sechellia’s* adaptation to *M. citrifolia*, as discussed below.

**Discussion**

**Molecular Functions of OBPs57d/e**

LUSH (OBP76a), the best studied OBP in *D. melanogaster*, functions as an adaptor molecule in vaccenyl acetate (VA) recognition by an odorant receptor, OR67d [20]. Mutants lacking LUSH lose their neuronal response to VA; thus, they...
Figure 4. Generation of Obp57d/e Knock-Out Flies by Gene Targeting

(A) Targeted gene replacement by the ends-out method. A donor transgene integrated into the other chromosome by P element–based transformation was excised by the FLP recombination enzyme at FLP recognition target (FRT) sites. Resulting circular DNA was linearized by I-SceI, inducing a precise replacement of a target gene with a marker gene. Finally, a marker gene was excised by Cre recombinase that recognizes \texttt{loxP} sequences, leaving a single 34-bp \texttt{loxP} sequence.

(B) Vector structures for \texttt{Obp57d/e}–targeted mutagenesis.

doi:10.1371/journal.pbio.0050118.g004
do not respond to VA behaviorally [18]. However, our \( \text{Obp57d}^{-}\) flies retained their behavioral responses to HA and OA, suggesting that OBP57dle do not function as adaptors for HA and OA. Rather, they seem to modulate dose-dependent responses to HA and OA, which might be achieved by other proposed functions of OBP, such as the titration or degradation of ligands [21].

There are qualitative differences in the behavioral responses to HA and OA between \( \text{Obp57d}^{-}\) and \( \text{Obp57e}^{-}\) flies. These differences might reflect functional diversification between Obp57d and Obp57e. However, the elimination of either \( \text{Obp57d} \) or \( \text{Obp57e} \) affected the expression level of the other in these knock-out flies. \( \text{Obp57d}^{-} \) removal by gene targeting increased \( \text{Obp57e} \) expression level, and \( \text{Obp57e}^{-} \) removal repressed \( \text{Obp57d} \) expression. Thus, we cannot exclude the possibility that the behavioral differences between the knock-out strains are caused by an altered expression level of either gene. A more operative method such as the Gal4-UAS system could be used to separate promoters from ORFs, thus minimizing the interaction between these two genes in expression control. It would then be possible to examine the molecular functions of Obp57d and Obp57e independently.

Expression Control of Obp57d and Obp57e

The results from our GFP reporter experiments and quantitative RT-PCR analysis are inconsistent. This inconsistency is also a feature of previous studies. Galindo and Smith [19] showed that the reporter constructs with 3 kb of upstream sequence from \( \text{Obp57d} \) and \( \text{Obp57e} \) were expressed in four cells in each leg, which matches the results of our GFP reporter experiments. However, using RT-PCR analysis, Takahashi and Takano-Shimizu [22] detected the transcripts not only in tarsi, but also in labella and wings. Together with the results of our quantitative RT-PCR analysis, it is clear that the reporter constructs do not reflect the complete expression pattern of \( \text{Obp57dle} \). The differences could be, at least in part, due to the lack of coding region in the reporter constructs. In fact, the elimination of a coding region of either \( \text{Obp57d} \) or \( \text{Obp57e} \) affected the expression level of the other gene in \( \text{Obp57dle}^{-} \) and \( \text{Obp57e}^{-} \), suggesting the involvement of ORFs and/or an intron in expression control (Figure 2). Furthermore, the introduction of the \( \text{Obp57dle} \) genomic region from \( \text{D. simulans} \) and \( \text{D. sechellia} \) reproduced the expression of \( \text{Obp57dle} \) in the head as well as in the legs, which was not observed in GFP reporter experiments.

Although the \( \text{Obp57dle} \) genomic region contains a considerable part of the genetic information that controls \( \text{Obp57dle}^{-} \),...
expression, it is still not sufficient to explain all the differences in the expression profile between the species; genetic factors at loci other than Obp57d/e are also likely to be responsible. There are two possibilities for such factors: (1) Trans-acting factors such as a transcription factor that is necessary for Obp57d/e expression, could control expression by determining which type of cell expresses Obp57d/e, or by determining transcription level in particular Obp57d/e-expressing cells. (2) Developmental factors determining the cell fate to become Obp57d/e-expressing cells, could increase/decrease the number of Obp57d/e-expressing cells by transforming cell fate at the expense of other cell types. In fact, ab1 and ab2 sensilla on antennae are transformed into ab3 sensilla in D. sechellia [16]. Such cell-type transformation might have occurred also in Obp57d/e-expressing cells. Genes of these two categories could be identified by, for example, screening of mutants that alter the Obp57d/e > GFP expression pattern.

Genetic Factors Constituting D. sechellia’s Adaptation to M. citrifolia

D. sechellia’s adaptation to M. citrifolia consists of genetic changes at many loci. Although there are likely to be additional genetic differences between D. sechellia and D. simulans, the present status of D. sechellia’s adaptation to M. citrifolia can be explained by alterations in three classes of genetic factors, as follows.

Factors responsible for resistance to the host-plant toxin OA: genes of this class are mapped to at least five loci scattered over all major chromosome arms [15], suggesting

---

**Table 1. Statistical Analysis of Results of Oviposition Site–Preference Assay (Acetic Acid)**

| Strain or Species | D. melanogaster | D. simulans | D. sechellia |
|------------------|-----------------|-------------|--------------|
|                  | w¹¹⁸ | Obp57dKO | Obp57eKO | Obp57d/eKO | +simObp57d/e | +secObp57d/e |
| w¹¹⁸             | ns | ns       | ns       | ns         | ns           | ns           |
| Obp57dKO         | ns | ns       | ns       | ns         | ns           | ns           |
| Obp57eKO         | ns | ns       | ns       | ns         | ns           | ns           |
| Obp57d/eKO       | ns | ns       | ns       | ns         | ns           | ns           |
| Obp57d/eKO;simObp57d/e | ns | ns | ns | ns | ns |
| Obp57d/eKO;secObp57d/e | ns | ns | ns | ns | ns |
| D. simulans      | ns | *        | ns       | ns         | ns           | ns           |
| D. sechellia     | ns | ns       | ns       | ns         | ns           | ns           |

* *p < 0.05; **p < 0.01 by the Wilcoxon rank sum test with p-value adjustment for multiple comparisons by the Holm method. ns, not significant.

doi:10.1371/journal.pbio.0050118.t001
that the alleles at these loci were fixed independently from each other during the course of D. sechellia evolution.

Factors responsible for the olfactory preference for M. citrifolia: D. sechellia can detect Morinda fruit from as far as 150 m away [23]. Consistent with this observation, the number of olfactory sensilla specifically tuned to the host odor MH is increased in D. sechellia [16] (but also note that MH is commonly found in many other plants). In contrast, however, there are no data showing that D. simulans avoids Morinda fruit purely on the basis of olfactory cues; all behavioral assays, including our trap assay, enable flies to come in direct contact with the odor source. There is also no neural response to HA and OA from the antennae of either D. melanogaster or D. sechellia [16]. We therefore assume that the olfactory cues from Morinda fruit are fundamentally attractive to Drosophila, and not repulsive even to D. simulans. D. sechellia has an enhanced preference specifically tuned to the Morinda blend of olfactory cues, in which MH is a functionally major component. Genes responsible for this enhanced preference are thought to function in cell fate determination during neuronal development [16], but the number of genes involved is not yet known.

Factors responsible for the D. simulans’ avoidance of Morinda fruit: we found this behavior to be based on gustatory cues, and confirmed that the introduction of the Obp57de region from D. simulans made D. melanogaster avoid HA in the same manner as D. simulans (Figure 6), proving that D. simulans’ avoidance of HA-containing media as an oviposition site is determined by Obp57de. These two genes are physically close to each other and are thus treated as a single locus in the following discussions.

Historical Order of Allele Fixation during the Course of D. sechellia’s Evolution

Here, we discuss the order of allele fixation at the loci responsible for D. sechellia’s adaptation to M. citrifolia. In particular, we focus on the issue of which mutation was the first to be fixed, because it must have played a key role in D. sechellia’s evolution.

We speculate on this with respect to the ecological validity of each phenotype in light of natural selection. We assume that the first mutation arose at a single locus, and examine the resulting phenotype in an ecological context. (1) If the first mutation occurred at a resistance QTL, the resulting phenotype would be partially resistant to M. citrifolia. However, this phenotype is ecologically “silent” because these flies avoid Morinda fruit and may not lay eggs on it (a mutation on the resistance QTL cannot be advantageous unless a fly’s behavior is changed). (2) If the first mutation was for the enhanced preference for the host odorant, flies should

---

**Table 2. Statistical Analysis of Results of Oviposition Site–Preference Assay (Butyric Acid)**

| Strain or Species | D. melanogaster | D. simulans | D. sechellia |
|------------------|-----------------|-------------|-------------|
|                  | w<sup>118</sup> | Obp57d<sup>ko</sup> | Obp57e<sup>ko</sup> | Obp57d/e<sup>ko</sup> | | simObp57d/e | secObp57d/e |
| w<sup>118</sup>  |                 |             |             |                         |             |               |             |
| Obp57d<sup>ko</sup> | ns              |             |             |                         |             |               |             |
| Obp57e<sup>ko</sup> | ns              | ns          |             |                         |             |               |             |
| Obp57d/e<sup>ko</sup> | ns              | ns          | ns          |                         |             |               |             |
| Obp57d/e<sup>ko</sup>;simObp57d/e | ns | ns | ns | ns | ns | ns | ns |
| Obp57d/e<sup>ko</sup>;secObp57d/e | ns | ns | ns | ns | ns | ns | ns |
| D. simulans       | ns              | ns          | ns          | ns                      | ns          | ns            | ns          |
| D. sechellia      | ns              | ns          | ns          | ns                      | ns          | ns            | ns          |

*p < 0.05; **p < 0.01 by the Wilcoxon rank sum test with p-value adjustment for multiple comparisons by the Holm method. ns, not significant.

doi:10.1371/journal.pbio.0050118.t002

**Table 3. Statistical Analysis of Results of Oviposition Site–Preference Assay (Hexanoic Acid)**

| Strain or Species | D. melanogaster | D. simulans | D. sechellia |
|------------------|-----------------|-------------|-------------|
|                  | w<sup>118</sup> | Obp57d<sup>ko</sup> | Obp57e<sup>ko</sup> | Obp57d/e<sup>ko</sup> | | simObp57d/e | secObp57d/e |
| w<sup>118</sup>  |                 |             |             |                         |             |               |             |
| Obp57d<sup>ko</sup> | ns              |             |             |                         |             |               |             |
| Obp57e<sup>ko</sup> | ns              | ns          |             |                         |             |               |             |
| Obp57d/e<sup>ko</sup> | ns              | ns          | ns          |                         |             |               |             |
| Obp57d/e<sup>ko</sup>;simObp57d/e | ** | ** | ** | ** | ** | ** | ** |
| Obp57d/e<sup>ko</sup>;secObp57d/e | ns | ns | * | * | ** | ** | ** |
| D. simulans       | **              | **          | **          | ns                      | ns          | ns            | ns          |
| D. sechellia      | **              | **          | **          | **                      | ns          | ns            | ns          |

*p < 0.05; **p < 0.01 by the Wilcoxon rank sum test with p-value adjustment for multiple comparisons by the Holm method. ns, not significant.

doi:10.1371/journal.pbio.0050118.t003
Table 4. Statistical Analysis of Results of Oviposition Site–Preference Assay (Octanoic Acid)

| Strain or Species | D. melanogaster | D. simulans | D. sechellia |
|-------------------|----------------|-------------|-------------|
|                   | D. melanogaster | D. simulans | D. sechellia |
| w^118             |                |             |             |
| Obp57d^K0         | **             | ns          |             |
| Obp57e^K0         |                | **          |             |
| Obp57d/e^K0       | ns             | **          | ns          |
| Obp57d/e^K0;simObp57d/e | ns | ns          | ns          |
| Obp57d/e^K0;secObp57d/e | ns | ns          | ns          |
| D. simulans       | **             | **          | **          |
| D. sechellia      | ns             | ns          | ns          |

*p < 0.05; **p < 0.01 by the Wilcoxon rank sum test with p-value adjustment for multiple comparisons by the Holm method. ns, not significant.

Materials and Methods

Fly preparation. The fly strains used were w^118 for D. melanogaster, S357 for D. simulans, and SS86 for D. sechellia [17]. Adult flies were collected immediately after eclosion, and staged for 5 d at 25 °C with a cotton plug soaked with liquid medium (5% yeast extract and 5% sucrose). Staged flies were used for the trap assay, the oviposition site–preference assay, and quantitative RT-PCR analysis.

Trap assay. A 30-ml glass flask containing 20 ml of HA solution in 0.05% Triton-X and a control flask containing the same amount of 0.05% Triton-X were placed in a plastic cage covered with a lid made of wire mesh. Up to 100 staged flies were introduced into the cage and kept in a dark, ventilated chamber overnight at 25 °C. The response index was calculated as RI = (Nw - Nh(0))/(Nh(0) + Nh), where Nh is the number of flies trapped in HA solution and Nw is that of flies in the control trap.

Determination of break points in deficiency chromosomes. The PCR primers used are listed in Table 5. To amplify a fragment of about 300–600 bp from genomic DNA extracted from the inter-species hybrids between D. melanogaster deficiency strains and D. sechellia, each primer was designed within the ORF of predicted genes, with the expectation that there is enough conservation of sequences between the two species. PCR products were subjected to direct sequencing with the same primer used for amplification. The deficiency chromosome was considered to cover the position when the sequence derived from D. melanogaster or those from both D. melanogaster and D. sechellia were detected, and it was not considered to cover when only the D. sechellia sequence was detected.

Signal peptide–sequence prediction. Signal peptide sequence was predicted using SignalP 3.0 [24].

GFP reporter analysis. The genomic sequence upstream of Obp57e was PCR amplified with the primer pair 5’-(NotI) GGGCGGCCG-CCGCCGCGGCCGCGCCGCGCGGACCCAAAAATTAAATGAGC-3’ and 5’-(BamHI) AAAGATCC-ACTTGCTATCTCAGGGGAA-3’. PCR products were cloned into pGreenPelican [25], and then introduced into D. melanogaster by the established P element–based transformation method. To remove the inserted 4 bp from the sechellia > GFP construct, the vector was PCR amplified using the KOD-plus enzyme (Toyobo, http://www.toyobo.co.jp)[el] that does not append a T on the ends with the primers 5’-GATATCATATATATGAAATTTAATTGGC-3’ and 5’-ACATTTTTTAATGCGACACACATCAGC-3’, and self-ligated after phosphorylation. At least five independent transformant lines for each construct were examined for GFP expression.

Gene targeting. Disruption of Obp57d and Obp57e was carried out by the ends-out method using the vectors provided by Dr. Sekelsky [26]. A hsp70-white marker gene was excised from pBS-70w with SphI and Xhol and subcloned into the Smal site of pBSII by blunting to obtain pBS-II-70b. The Obp57d upstream region amplified with the primer pair 5’-(EcoR) AAAAAATCTCGAGTATATCC-CACGCCAAATGATG-3’ (P1) and 5’-(BamHI)AAAGATCC-ACTTGCTATCTCAGGGGAA-3’ (P2) with the primer pair 5’-(PstI-1oxP) CGTGAAGATATCTTGTAAGATG-3’ and the downstream region amplified with the primer pair 5’-(PstI-1oxP)CTGCAGATATCCCTGAGATGTTGATG-3’ -3’ and 5’-(NotI)GGGCGGCCGCTATGCGACACACATCAGC-3’ (P2) were subcloned into pLP[EndsOut2].
Table 5. Sequence Primers Used for Determination of Break Points in Deficiency Chromosomes

| Gene Name | Primer Sequence (5’ to 3’) | Position of 5′ Terminal\(^a\) |
|-----------|---------------------------|-----------------------------|
| CG1109    | CGCAGGTTGGTTGTTTGGCGCT   | 15,752,566                  |
|           | AACACTGATAGTAACAGGAGAG   | 15,752,561                  |
| CG13869   | AAAAACCTGTTGAGCAGCTTTG   | 15,753,644                  |
| CG11048   | GAGAAAGCATTGGCGGCACTT    | 15,753,958                  |
| CG11208   | AGGCAGCGCAAAATTAATGAGTA  | 15,757,617                  |
| CG11209   | CCACTATGTTGCGACATTG     | 15,757,942                  |
| ppk6      | ATGCCTGTACACCACCTCAGT    | 15,761,675                  |
| CG8914    | GCCGAAGAATTTTTCGAGAGAG  | 15,804,593                  |
| C4HJ2     | CTGGCGGAAACCTTCTGGTGG    | 15,805,092                  |
| CG13868   | TTAACCCGGCTGACACTACGAG  | 15,824,692                  |
| CG16739   | CTTTCGCTGGCTTCCGAGT      | 15,856,991                  |
| CG11192   | GTTGATCACCACGACACCTCAG  | 15,869,367                  |
| CG33041-1 | ACCATCGTTGTCACCACTC     | 15,898,994                  |
|           | GCTTGTGTTACACGACACCTC   | 15,939,503                  |
| CG33041-2 | CTCTCTCTCCATTGCGGACAG   | 15,944,207                  |
| CG3303-1  | CATCCTCCTCCGATGTTCC      | 15,944,639                  |
| CG13423   | ACCAGCTGAAAGTCGCGCGA    | 15,984,954                  |
| CG13424   | CCACCAGCTTCCCTCTCGG      | 15,985,451                  |
| CG13421   | GTAGAAGCATTATTCTAGGAGGCTG| 16,018,443                  |
| Obp57c    | TTTAGACGATTAGTTTCTTTAG   | 16,019,052                  |
| CG38042   | GGCCCTCTGATTGGTTTGTGC    | 16,019,156                  |
| Obp57b    | GACCTAATCCTAGGAGGTTG     | 16,020,032                  |
| CG38041   | ATGGTACATACGACATCTG      | 16,020,568                  |
| Obp57a    | AACAGATCCTCGGCGGTTC     | 16,027,975                  |
| CG31426   | CCCCTACGCTGATATGTTG      | 16,027,600                  |
| Obp57d    | ATGGACAGCCTACTTCTCTGCG   | 16,027,975                  |
| CG31423   | CACAAGGTGCCGAGGCAAGC      | 16,051,664                  |
| Obp57e    | CCGCTGTATAATCTACCGAGC   | 16,052,354                  |
| CG31424   | AAAATCGCTTCCGGCGCGAGA    | 16,062,700                  |
| Obp57e    | GTTGACAAACTATACCTG       | 16,063,152                  |
| CG31423   | CTCGCAAGCATATAACGAGG    | 16,072,990                  |
| Obp57c    | CGAGGACTCCGGTATTTGCG     | 16,073,464                  |
| CG31431   | TCTCGCTGACGCTTCTGGG     | 16,075,839                  |
| Mnt1      | AACATATCGTGGGATACG       | 16,076,305                  |
| CG31432   | GAATCAGAGCTAGACGAGGCA    | 16,078,377                  |
| iI20507   | AGTGGACGTCGACCCGCTGCC   | 16,078,873                  |
| CG31425   | CTACAGCTGTCAAGCCTGCC    | 16,120,193                  |
| bi        | CTCGTTGCCAGCAGACATGCG    | 16,120,714                  |
| CG31441   | TCATCTCTGGTCAGCTGCAATC  | 16,200,479                  |
| Grs7a     | GAATAGGCTGTACCATGCACCTACA | 16,201,809           |

\(^a\)Base position on 2R euchromatin, release 4.2.

doi:10.1371/journal.pbio.0050118.t005

with hsp70-white excised from the pBSII-7bw with BamHI and PstI. The Obp57e upstream region amplified with the primer pair 5′- (BamHI-loxP) GGATCC-ATAACACCTGGTATATAATGCTATACGAGTTATAGGACAGATATCCAAGTCAGTAT-3′ and P1 and the downstream region amplified with the primer pair primers 5′- (PstI-
6. Jacquin-Joly E, Merlin C (2004) Insect olfactory receptors: Contributions of molecular biology to chemical ecology. J Chem Ecol 30: 2359–2397.
7. Bruce TJA, Wadhams LJ, Woodcock CM (2005) Insect host location: A volatile situation. Trends Plant Sci 10: 269–274.
8. Hallem EA, Dahanukar A, Carlson JA (2006) Insect odor and taste receptors. Ann Rev Entomol 51: 113–135.
9. Adams MD, Celniker SE, Holt RA, Evans CA, Gocayne JD, et al. (2000) The genome sequence of Drosophila melanogaster. Science 287: 2185–2195.
10. Holt RA, Subramanian GM, Halpern A, Sutton GG, Charlab R, et al. (2002) The genome sequence of the malaria mosquito Anopheles gambiae. Science 298: 129–149.
11. Robertson HM, Gordon KHJ (2006) The chemoreceptor superfamily in the honey bee, Apis mellifera: Expansion of the odorant, but not gustatory, receptor family. Genome Res 16: 1395–1403.
12. Jones CD (2005) The genetics of adaptation in Drosophila sechellia. Genetica 123: 137–145.
13. Legal I, Chappe B, Jallon JM (1994) Molecular basis of Morinda citrifolia, (L.): toxicity on Drosophila. J Chem Ecol 20: 1951–1943.
14. Legal I, Moulin B, Jallon JM (1999) The relation between structures and toxicity of oxygenated aliphatic compounds homologous to the insecticide octanoic acid and the chemotaxis of two species of Drosophila. Pesticide Biochem Physiol 65: 90–101.
15. Jones CD (1998) The genetic basis of Drosophila sechellia’s resistance to a host plant toxin. Genetics 149: 1899–1908.
16. Dekker T, Ibbi I, Siju KP, Stenson MC, Hansson BS (2006) Olfactory shifts parallel superspecialism for toxic fruit in Drosophila melanogaster sibling, D. sechellia. Curr Biol 16: 101–109.
17. Higa I, Fuyama Y (1993) Genetics of food preference in Drosophila sechellia. I. Responses to food attractants. Genetica 88: 129–136.
18. Xu P, Atkinson R, Jones DNM, Smith DP (2005) Drosophila OBP LUSH is required for activity of pheromone-sensitive neurons. Neuron 45: 193–200.
19. Galindo K, Smith DP (2001) A large family of divergent Drosophila odorant-binding proteins expressed in gustatory and olfactory sensilla. Genetics 159: 1059–1072.
20. Ha TS, Smith DP (2006) A pheromone receptor mediates 11-cis-vaccenyl acetate-induced responses in Drosophila. J. Neurosci. 26: 8727–8733.
21. Blomquist GJ, Vogt RG, editors (2005) Insect pheromone biochemistry and molecular biology: The biosynthesis and detection of pheromones and plant volatiles. London: Elsevier Academic Press. 768 p.
22. Takahashi A, Takano-Shimizu T (2005) A high-frequency null mutant of an odorant-binding protein gene, Obp57e, in Drosophila melanogaster. Genetics 170: 709–718.
23. R’Kha S, Gapy P, David JR (1991) Host-plant specialization in the Drosophila melanogaster species complex: A physiological, behavioral, and genetical analysis. Proc Natl Acad Sci U S A 88: 1835–1839.
24. Bendtsen JD, Nielsen H, Von Heijne G, Brunak S (2004) Improved prediction of signal peptides: SignalP 3.0. J Mol Biol 340: 785–795.
25. Barolo S, Carver LA, Posakony JW (2000) GFP and β-galactosidase transformation vectors for promoter/enhancer analysis in Drosophila. BioTechniques 29: 726–732.
26. Radford SJ, Goley E, Baxter K, McMahan S, Sekelsky J (2005) Drosophila ERC1 is required for a subset of MEL-9-dependent meiotic crossovers. Genetics 170: 1737–1745.