

**OPEN ACCESS**

Full open access to this and thousands of other papers at <http://www.la-press.com>.

## Evolutionary Genomics of Genes Involved in Olfactory Behavior in the *Drosophila melanogaster* Species Group

Nicolás Lavagnino<sup>1,2</sup>, François Serra<sup>3</sup>, Leonardo Arbiza<sup>3</sup>, Hernán Dopazo<sup>3</sup> and Esteban Hasson<sup>1,2</sup>

<sup>1</sup>Departamento de Ecología, Genética y Evolución; Facultad de Ciencias Exactas y Naturales; Universidad de Buenos Aires; Buenos Aires; Argentina. <sup>2</sup>Instituto de Ecología, Genética y Evolución (IEGEB), CONICET, Buenos Aires, Argentina. <sup>3</sup>Evolutionary Genomics Unit; Bioinformatics and Genomic Department; Centro de Investigación Príncipe Felipe; Valencia; España. Corresponding author email: [nlavagnino@ege.fcen.uba.ar](mailto:nlavagnino@ege.fcen.uba.ar)

**Abstract:** Previous comparative genomic studies of genes involved in olfactory behavior in *Drosophila* focused only on particular gene families such as odorant receptor and/or odorant binding proteins. However, olfactory behavior has a complex genetic architecture that is orchestrated by many interacting genes. In this paper, we present a comparative genomic study of olfactory behavior in *Drosophila* including an extended set of genes known to affect olfactory behavior. We took advantage of the recent burst of whole genome sequences and the development of powerful statistical tools to analyze genomic data and test evolutionary and functional hypotheses of olfactory genes in the six species of the *Drosophila melanogaster* species group for which whole genome sequences are available. Our study reveals widespread purifying selection and limited incidence of positive selection on olfactory genes. We show that the pace of evolution of olfactory genes is mostly independent of the life cycle stage, and of the number of life cycle stages, in which they participate in olfaction. However, we detected a relationship between evolutionary rates and the position that the gene products occupy in the olfactory system, genes occupying central positions tend to be more constrained than peripheral genes. Finally, we demonstrate that specialization to one host does not seem to be associated with bursts of adaptive evolution in olfactory genes in *D. sechellia* and *D. erecta*, the two specialists species analyzed, but rather different lineages have idiosyncratic evolutionary histories in which both historical and ecological factors have been involved.

**Keywords:** olfactory behavior, insects, evolutionary genomics, adaptation, *Drosophila melanogaster* species group

*Evolutionary Bioinformatics* 2012:8 89–104

doi: [10.4137/EBO.S8484](https://doi.org/10.4137/EBO.S8484)

This article is available from <http://www.la-press.com>.

© the author(s), publisher and licensee Libertas Academica Ltd.

This is an open access article. Unrestricted non-commercial use is permitted provided the original work is properly cited.



## Introduction

Animals rely on their senses to perceive the world; in this sense olfactory behavior is a crucial character that allows insects to perform tasks of ecological and evolutionary importance such as: finding food, mate recognition, oviposition site selection, avoidance of predators, etc. In the last 15 years a large amount of information about genetic, physiological and molecular aspects of olfaction has been accumulated in invertebrates<sup>1,2</sup> and vertebrates.<sup>3–5</sup>

In the fruit-fly *Drosophila melanogaster* (Diptera: Drosophilidae), odorants bind to olfactory receptors (*Or*) that are expressed in olfactory receptor neurons (ORNs) membranes.<sup>2</sup> At this point, odorant binding proteins (*Obp*) expressed in the aqueous lymph that fills the olfactory sensilla on the antenna<sup>6</sup> also play important roles in this interaction, for example: helping in solubilisation, transport and presentation of odorants to *Or*'s and removal and rapid inactivation of odors after reaction with *Or*'s in ORNs membranes.<sup>7</sup> This first contact of flies' olfactory molecules with odorants (that take place in the olfactory organs, the third antennal segment and the maxilar palp) constitutes the peripheral level of the *Drosophila* olfactory system.<sup>2</sup> Nevertheless, other gene products, not expressed in the peripheral level, are also necessary for the proper expression of olfactory behavior.<sup>8–14</sup> These genes, which encode proteins with quite different functions, occupy central positions in the *Drosophila* olfactory system. Therefore, to understand the genetic architecture of a phenotypic character identification of all genes, gene-to-gene interactions, gene-to-environment interactions that contribute to its expression is needed, as well as a comprehension of its variational properties.<sup>15,16</sup> Several studies have shown that olfactory behavior has a complex genetic architecture since it is orchestrated by many interacting genes that also interact with the environment.<sup>8,9,14,17–21</sup>

Comparative genomic studies on the evolution of olfactory behavior in *Drosophila* have focused on particular gene families, such as *Or*<sup>22–26</sup> and *Obp* gene families.<sup>27</sup> Here, we present a comparative genomic study of olfactory behavior in *Drosophila* including an extended set of genes, besides *Or*'s and *Obp*'s gene families, that are known to be involved in the genetic architecture of olfactory behavior. We took advantage of the recent burst of whole genomes sequences and the development of powerful statistical and

bioinformatic tools to analyze large sets of genomic data<sup>28,29</sup> and tested evolutionary hypotheses regarding the genetic architecture of olfactory behavior in the six species of the *D. melanogaster* species group for which whole genome sequences are available, namely *Drosophila melanogaster*, *Drosophila simulans*, *Drosophila sechellia*, *Drosophila erecta*, *Drosophila yakuba*, *Drosophila ananassae*.<sup>30</sup>

Even though in our analysis we did not group genes in gene families, the hypotheses we aim to test are not about individual genes but about groups of genes that share a given characteristic that could help to explain the evolution of olfactory behavior. As pointed out earlier, olfactory genes participate in olfaction at different stages of the life cycle and/or at different positions in the chain of events that builds the olfactory response.<sup>2,31,32</sup> In this context, we aim to test the following hypotheses: (i) olfactory genes that participate in the expression of both larval and adult olfactory behavior evolve at a slower rate than genes that contribute either to larval or adult olfactory behavior, since genes with ubiquitous expression are expected to be under stronger functional constraints; (ii) olfactory genes that participate in a peripheral position of the olfaction transduction pathway evolve faster than genes expressed in central positions, since central position genes tend to participate in other signal cascades, besides olfactory. This hypothesis follows from the general notion that more pleiotropic genes may be under stronger purifying selection than less pleiotropic genes.<sup>33</sup>

Comparative genomic studies between closely related species also offer the unique opportunity to compare the evolutionary rates in lineages with particular ecological and evolutionary histories. Sequenced species in the *D. melanogaster* species group can clearly be grouped in specialist and generalist species. *D. sechellia* breeds and feeds only on the decaying fruits of the shrub *Morinda citrifolia*<sup>34–36</sup> and *D. erecta* is a seasonal specialist of *Pandanus candelabrum*.<sup>37</sup> Among the remaining species, we found two known quintessential generalists: *D. melanogaster* and *D. simulans*; and a less known generalist *D. yakuba*.<sup>38,39</sup> Previous work, based on studies in *Or* and *Obp* gene families, proposed that olfactory genes in specialist lineages of the *D. melanogaster* species group should exhibit faster evolutionary rates than generalists due to directional selective pressures

from a one-host environment.<sup>23,24,27</sup> To verify these hypotheses we used an extended dataset of olfactory genes, a maximum likelihood model to estimate evolutionary rates at coding sequences and suitable tests aimed to differentiate true cases of positive selection (PS) from relaxation of selective constraints (RSC) or weak positive selection (WS). Our results show that evolutionary rates of olfactory genes are independent of the life cycle stage, and the number of stages in which they participate in olfaction. However, we discovered a relationship between the position that the gene product occupies in the olfactory signal transduction pathway and its evolutionary rate. Also, we found that in the *D. melanogaster* species group specialization to one host does not seem to trigger bursts of rapid evolution in olfactory genes, but reveals instead idiosyncratic evolutionary patterns in the two specialists species analyzed.

## Materials and Methods

We performed a genomic comparative analysis of genes known to be involved in olfactory behavior recovered from the whole genome sequences of the six species of the *D. melanogaster* species group available so far. The analysis was done with coding sequences (CDS) of 115 *D. melanogaster* genes, known to be involved in olfaction, and the CDS of the orthologs in the five remaining species; adding to a total of 664 CDS (26 genes under a process of pseudogenization or/and misaligned were excluded from the analysis, so the total number does not add to 690). CDS were downloaded from *FlyBase* (sequences were uploaded to *FlyBase* by<sup>40</sup>). CDSs were aligned with MUSCLE<sup>41,42</sup> using predicted amino acid sequences as templates. Alignment columns containing gaps were removed.

Olfactory genes were categorized in two different functional classes, life cycle stage (larval and/or adult) in which a gene is involved in olfaction and the position (peripheral or central) in which participates in the olfactory system. Regarding the categorization of genes in terms of involvement in larval and/or adult olfactory behavior, we only included those that were tested for olfactory behavior in both life cycle stages. If a gene was reported to participate in the expression of adult olfactory behavior but was never tested in larvae it was not included in the study. The categorization in terms of position in the olfactory system

is based on a conception of olfactory behavior as a complex trait orchestrated by many genes with a wide range of molecular functions. *ORs* and *OBPs*, which encode for proteins involved in the first contact of the fly with odorant stimuli, were considered as peripheral in the olfactory system. We consider as central a subset of genes with different molecular functions that are known to participate in olfaction but not in the first contact with odorants. This classification of olfactory genes is commonly used in studies of olfaction in *Drosophila*, as can be seen, for example, in a recent review of molecular aspects.<sup>2</sup> The involvement of these genes in olfaction was identified by means of different genetic approaches as loss of function by mutagenesis, rescue experiments and expression techniques like RT-PCR, in situ hybridization and immunohistochemistry. The number of genes included in the analyses along with information on functional categories and the corresponding sources are given in Table 1 (a more detailed description of each gene categorization can be found in Supplementary Table S1).

Although we acknowledge that function, timing and tissue of expression of a given gene may change across species. We assume, given the close phylogenetic relationships among species of the

**Table 1.** Olfactory genes included in the study.

| Categorization  | Nº of genes | Reference  |
|-----------------|-------------|--|
| <b>Position</b> |             |  |
| Central         | 53          | 9, 10, 11, 12, 13, 14, 64, 65, 66, 67                  |
| Peripheral      | 62          | 2 (and references therein), 68, 69, 70, 71             |
| <b>Stage</b>    |             |  |
| Larva           | 27          | 2 (and references therein), 14, 64, 68, 69             |
| Adult           | 34          | 2 (and references therein), 13, 14, 64, 68, 69, 70, 71 |
| Larva and adult | 13          | 2 (and references therein), 14, 64, 68, 69, 70, 71     |

**Notes:** Place of participation of genes in the olfactory system (peripheral: genes that participate in the first contact with odor molecules in the periphery of olfactory response; central: genes known to be involved in the proper expression of olfactory behavior in down stream positions of the olfactory system). Stage: life cycle stage of participation in olfaction (L: larvae; A: adult; LA: larvae and adult). Reference: bibliographical reference from where the information to make the categorization was obtained from: 2 (and references therein), 9, 10, 11, 12, 13, 14, 64, 65, 66, 67, 68, 69, 70, 71. Categorization was based on results obtained in *D. melanogaster*. A more detailed list of genes used in this paper can be found in Supplementary Table S1.



*D. melanogaster* species group, that the functionality of olfactory genes (as described in *D. melanogaster*) did not change in any of the branches of the phylogeny of the *D. melanogaster* species group.

### Maximum likelihood tests

Maximum Likelihood (ML) estimates of the ratio  $\omega = d_N/d_S$  ( $d_N$ : nonsynonymous substitution rate;  $d_S$ : synonymous substitution rate) for each olfactory gene were employed to estimate evolutionary rates and to infer the evolutionary forces acting upon them. Synonymous mutations do not change amino acid sequence, hence  $d_S$  is assumed neutral with respect to selective pressures on the protein product of the gene; and nonsynonymous mutations change the amino acid sequence, so  $d_N$  depends on the selective pressures acting upon the protein sequences. The ratio  $\omega = d_N/d_S$ , then becomes a measure of selective pressures. On one hand, if nonsynonymous mutations are deleterious, purifying selection will slow their fixation rate and  $\omega$  values are expected to be lesser than 0.3; when  $0.3 < \omega < 1$  it is interpreted as weak purifying selection. On the other hand, if nonsynonymous mutations are advantageous, they will fix at a higher rate than synonymous mutations, and thus,  $\omega$  will be greater than 1 and indicative of PS. A  $\omega = 1$  stands for neutral evolution. Estimates of  $\omega$  were obtained applying a free-ratio ML model using the CodeML program of the PAML 4 package.<sup>28</sup> Genes with  $d_S = 0$  were excluded from the analysis in which we compare median values of  $\omega$  between groups of categorized genes, since in a comparative framework, the resulting  $\omega$  values are not biologically informative.

Positive selection was evaluated using two different branch-site models (A and A1) and two different likelihood ratio tests (I and II),<sup>43</sup> implemented in the CodeML program of the PAML 4 package<sup>28</sup> (from here on called the Zhang-Nielsen-Yang 2005 tests). Branches in the phylogeny were defined *a priori* as foreground and background lineages. Under these models only the foreground lineage may contain events of PS. Each species, except *D. ananassae* that was used as out-group, in the *D. melanogaster* species group and the respective ancestral lineage were tested independently as the foreground lineage. Ancestral sequences were reconstructed by ML from actual species sequences in the phylogenetic tree.

In contrast to the statistical behavior of previous branch-site tests,<sup>44</sup> the Zhang-Nielsen-Yang 2005 tests are improved methods of branch-site tests using an ML approach which has proved to be more successful in differentiating PS from RSC.<sup>43</sup> Test I compares model M1a against model A. M1a assumes two classes of sites, one with  $0 < \omega_0 < 1$  and another with  $\omega_1 = 1$ , fixed in all the lineages of the phylogenetic tree. Model A considers four classes of sites: site class 0 which includes conserved codons throughout the tree ( $0 < \omega_0 < 1$  in all lineages), site class 1 that includes codons evolving neutrally throughout the tree ( $\omega_1 = 1$  in all lineages), site classes 2a and 2b which include codons conserved or evolving neutrally on the background branches but that become positively selected in the foreground branches (2a:  $0 < \omega_0 < 1$  in background lineages and  $\omega_2 > 1$  in foreground lineages and 2b:  $\omega_1 = 1$  in background lineages and  $\omega_2 > 1$  in foreground lineages). The proportion  $p_i$  of each site class ( $p_0, p_1, p_{2a}, p_{2b}$ ) and the mean value of  $\omega_2$  ( $\omega_{2a}$  and  $\omega_{2b}$ ) can be estimated from the data using ML methods. Test II compares the null model A1 against model A. Parameters in A1 are equal to those of A with the exception that site classes 2a and 2b are fixed in the foreground with  $\omega_2 = 1$ . As was demonstrated using a simulation, Test I cannot properly distinguish cases of RSC from true events of PS. On the other hand, Test II, by allowing selectively constrained sites in the background to become relaxed under the proportion of site classes with  $\omega_2 = 1$  set in the foreground of A1, can properly distinguish between these two alternatives, with an acceptable false discovery rate.<sup>43</sup> Therefore, we evaluate the results obtained with both tests to distinguish between events of PS from RSC. Since the compared models are nested, likelihood ratio tests were performed and likelihood ratio tests statistics ( $2\Delta l = -2 [\ln[\text{likelihood for null model}] - \ln[\text{likelihood for alternative model}]]$ ) were posteriorly transformed into exact *P*-values using the `pchisq` function of the R statistical package.<sup>45</sup> Likelihood ratio tests were performed using a  $\chi^2$  distribution with d.f. = 2 for Tests I and d.f. = 1 for Test II, which have been shown to be conservative under conditions of PS.<sup>43</sup> *P*-values derived from PS analyses were false discovery rate-adjusted using the method of Benjamini and Hochberg.<sup>46</sup>

We excluded from the analysis the other 6 species that not belong to *D. melanogaster* species group for which whole genome information is available, since the inclusion of highly diverged taxa may lead

to saturation in synonymous positions, which may produce unreliable estimates of  $\omega$ .

Phylogenetic relationships between species in the *D. melanogaster* species group are well resolved, with the exception of controversial results for some members of the *melanogaster* sub-group, specifically the placement of *D. erecta* and *D. yakuba* relative to the *D. melanogaster* lineage. We only show the results obtained using the best supported topology (*D. melanogaster*, (*D. erecta*, *D. yakuba*))<sup>40</sup> (Fig. 1).

For details of parameters and input data used in maximum likelihood models in CodeML see Appendix in Supplementary data.

### Multiple regression analysis

In order to rule out the possibility that differences in  $\omega$  values in our set of genes are related to other variables besides gene categorizations made by us we tested whether physical location of genes in the genome and/or the amount of expression are related to  $\omega$  by performing a multiple linear regression analysis of  $\omega$  (estimated by a free ratio model) with physical location (cytological band) and expression level (mRNA abundance in whole adult flies). mRNA abundance data in whole adult flies were obtained from *FlyAtlas*.<sup>47</sup> The multiple linear regression model we used is:

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \epsilon_i$$

where  $Y_i$  is the  $i$ -th observation of the dependent variable  $\omega$ .  $X_{1i}$  and  $X_{2i}$  are  $i$ -th value of the independent

variables physical location and expression level, respectively.  $\beta_1$  and  $\beta_2$  are unknown parameters representing the rates of change in  $Y$  per unit change in  $X_1$  and  $X_2$ , respectively.  $\beta_0$  is another unknown parameter that represents the intercept of the line.

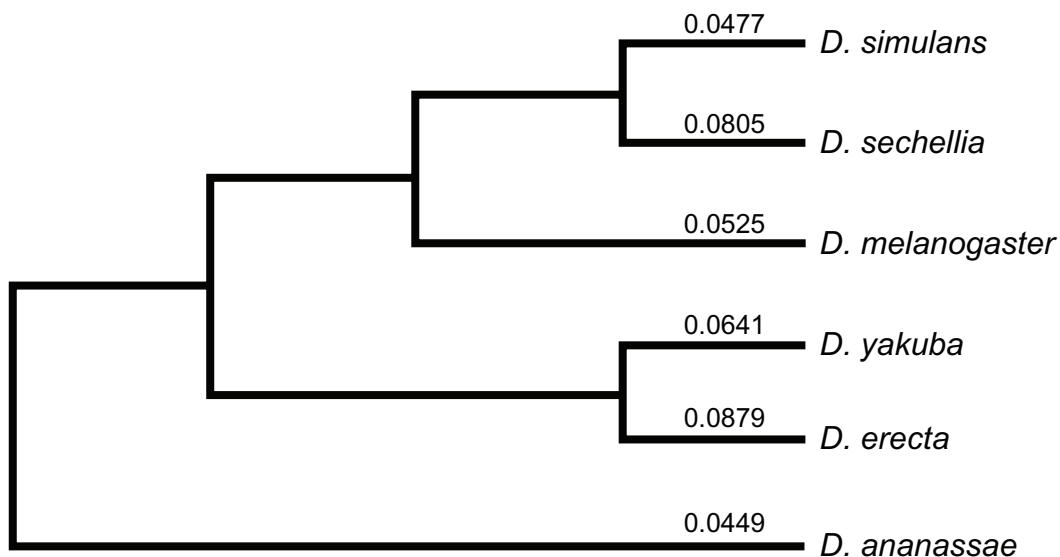
All statistical analyses and tests were done using R statistical package.<sup>45</sup>

### Results

Maximum likelihood estimates of  $\omega$  using a free ratio-model revealed that most genes included in the analysis are under purifying selection (ie,  $\omega < 0.3$ ) in all branches of the phylogeny. Median  $\omega$  values for the set of genes analyzed in each species are:  $\omega_{D. melanogaster} = 0.0525$ ,  $\omega_{D. simulans} = 0.0477$ ,  $\omega_{D. sechellia} = 0.0805$ ,  $\omega_{D. erecta} = 0.0879$ ,  $\omega_{D. yakuba} = 0.0641$ ,  $\omega_{D. ananassae} = 0.0449$  (Fig. 1; Supplementary Table S1). However, four genes exhibited values of  $\omega$  greater than 1, suggesting the occurrence of PS in their coding regions.

### Inferring the forces affecting the evolution of olfactory genes in the *D. melanogaster* species group

Since estimates of  $\omega$  from a free-ratio model are not appropriate test for detecting PS, we performed the Zhang-Nielsen-Yang 2005 tests<sup>43</sup> for sites under PS in specific branches of the tree. In this case, none of the four olfactory genes that presented a  $\omega > 1$  in the free-ratio model reached significance for PS or RSC



**Figure 1.** Maximum likelihood estimates of median  $\omega$  of all olfactory genes in each lineage of the *D. melanogaster* species group analyzed.



when tested using Zhang-Nielsen-Yang 2005 tests. The  $\omega$  values, estimated by a free ratio model, of these genes are only slightly greater than 1, hence Tests I and II both failed to accept either PS or RSC. Instead, two olfactory genes, *CG10777* in *D. sechellia* and *Xrp1* in *D. erecta*, presented signals of PS; and seven showed signs of RSC or weak selection: *CG11883*, *CG32556* and *Moesin* in *D. simulans*; *discs lost* and *Or42a* in *D. erecta*; *CG16708* and *Or82a* in *D. yakuba* (Table 2). In general, the analysis performed on the subset of five closely related species showed that only 9 out of the 565 (1.6%) olfactory genes analyzed exhibited signs of either PS or RSC. The remaining set of genes (98.4%) showed values of  $\omega$  that were not significantly different from neutral expectations or were compatible with purifying selection.

### Evolution of specific groups of olfactory genes in the *D. melanogaster* species group

The vast amount of genetic information that has accumulated for *D. melanogaster* in past decades allowed us to assign genes to *a priori* categories on the basis of: (i) the stage of the life cycle in which a particular gene participates in olfaction -genes can take part in either larval or adult olfactory behavior or in both-, and (ii) the position at which a gene product participates in the transduction pathway that characterizes the olfactory response -genes products can play a role either in the first contact with odor molecules in the periphery of the olfactory response (*Or* and *Obp* gene families) or in more central positions of the olfactory

transduction pathway. Supplementary Table S1 lists all olfactory genes analyzed and their classification in the different categories.

In the following sections we describe the results of the comparative analyses of evolutionary rates among *a priori* categorized groups of genes to test the following hypotheses: (i) olfactory genes that participate in the expression of both larval and adult olfactory behavior evolve at a slower rate than genes that contribute to olfactory behavior in only one life cycle stage and (ii) genes that participate in peripheral positions of the olfactory system evolve at a faster rate than genes involved in central positions.

### Rates of evolution of olfactory genes and their role in the life cycle

In *D. melanogaster* the median value of  $\omega$  for genes involved in larval (L) olfaction ( $\omega = 0.0742$ ) was twice larger than the median  $\omega$  estimated for genes involved in larval and adult (LA) ( $\omega = 0.0389$ ) and only adult (A) olfaction ( $\omega = 0.0359$ ) (Fig. 2). However, differences among medians were not significant (Fig. 2; Wilcoxon rank sum test, A vs. L:  $w = 614$ ,  $P = 0.0243$ ; A vs. LA:  $w = 184.5$ ,  $P = 0.3889$ ; L vs. LA:  $w = 210$ ,  $P = 0.3249$ ), suggesting that the rate of evolution is independent of the life cycle stage of expression and also of whether the gene participates in olfaction in one (either larval or adult) or in both stages of the life cycle.

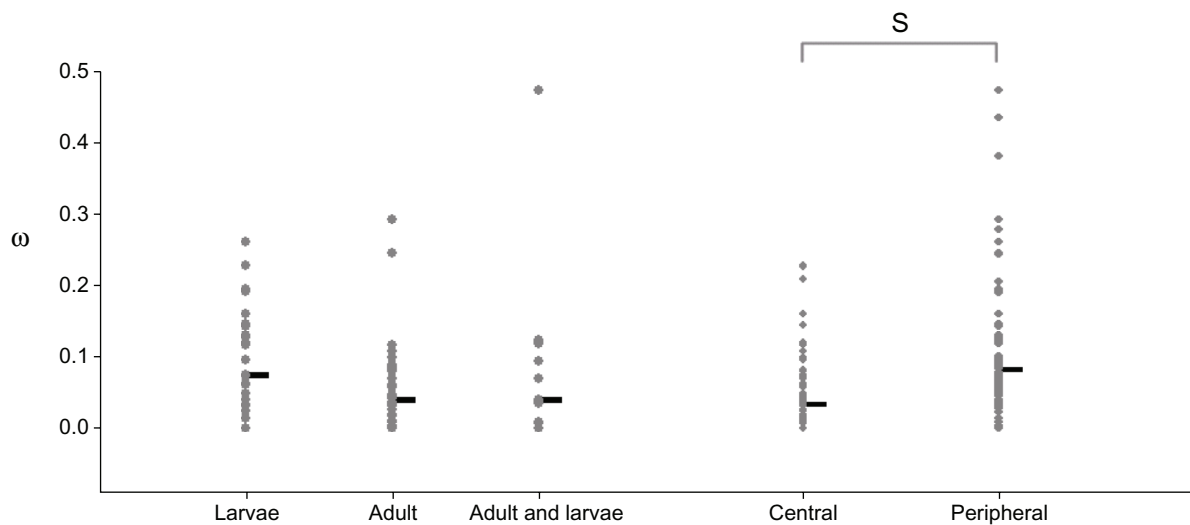
We performed the same analysis in the other species of the *D. melanogaster* species group and observed significant differences between categories in *D. simulans*, *D. erecta* and *D. ananassae*, but not in the other

**Table 2.** Positive selection and relaxation of selective constraints on olfactory genes.

| Gen  | Test I |                    | Test II |         | Biological interpretation:<br>PS or RSC-WS |
|--|--------|--------------------|---------|---------|--|
|  | d.f.   | P-value            | d.f.    | P-value |  |
| <i>CG10777</i> <sub><i>D. sec</i></sub>    | 2      | 1                  | 1       | 0.0037  | Gen under PS                               |
| <i>Xrp1</i> <sub><i>D. ere</i></sub>       | 2      | $2 \times 10^{-5}$ | 1       | 0.0007  | Gen under PS                               |
| <i>CG11883</i> <sub><i>D. sim</i></sub>    | 2      | 0.044              | 1       | 1       | Gen with signs of RSC-WS                   |
| <i>CG32556</i> <sub><i>D. sim</i></sub>    | 2      | 0.0054             | 1       | 0.1981  | Gen with signs of RSC-WS                   |
| <i>Moesin</i> <sub><i>D. sim</i></sub>     | 2      | 0.0026             | 1       | 1       | Gen with signs of RSC-WS                   |
| <i>discs lost</i> <sub><i>D. ere</i></sub> | 2      | 0.0181             | 1       | 1       | Gen with signs of RSC-WS                   |
| <i>Or42a</i> <sub><i>D. ere</i></sub>      | 2      | 0.0136             | 1       | 1       | Gen with signs of RSC-WS                   |
| <i>CG16708</i> <sub><i>D. yak</i></sub>    | 2      | 0.0329             | 1       | 1       | Gen with signs of RSC-WS                   |
| <i>Or82a</i> <sub><i>D. yak</i></sub>      | 2      | 0.0045             | 1       | 0.5590  | Gen with signs of RSC-WS                   |

**Notes:** Significant results for Zhang-Nielsen-Yang 2005 tests are shown. Gene names and species for which the ortholog gene belongs (in sub-index) are shown.

**Abbreviations:** d.f., Degrees of freedom; PS, positive selection; RSC-WS, relaxation of selective constraints or weak positive selection.



**Figure 2.** Comparisons of gene group categories median  $\omega$  in *D. melanogaster*. ML estimates of median  $\omega$  for gene group categories (life cycle stage of participation in olfaction and position of participation in the olfactory system) for *D. melanogaster* (horizontal marks).

**Notes:** Gray circles are free ratio-model  $\omega$  estimates for individual genes within a given category. To easier the lecture, the plot was truncated for  $\omega$  values higher than 0.5. Significant differences (highlighted with an S) between median  $\omega$  of peripheral position genes and central position genes were found using a Wilcoxon rank sum test. A Bonferroni multiple-comparison correction was applied.

lineages analyzed. In the first two species median  $\omega$  was significantly closer to 1 for larval genes than for adult genes (Fig. 3; Wilcoxon rank sum test: A vs. L in *D. simulans*,  $w = 539$ ,  $P = 0.0057$ ; in *D. erecta*,  $w = 609$ ,  $P = 0.0154$ ), whereas the median value of  $\omega$  for L genes was significantly closer to 1 than for LA genes in *D. ananassae* (Fig. 3; Wilcoxon rank sum test: L vs. LA,  $w = 258$ ,  $P = 0.0165$ ). These tests revealed that olfactory genes known to be expressed only in the larval stage evolve faster relative to genes expressed in adults or in both stages, in three of the six species analyzed (Fig. 3). Such difference may be a reflection of different selective pressures (or functional constraints) among species on olfactory genes which help flies to perceive the environment.

### Olfactory genes and position in the olfactory transduction pathway

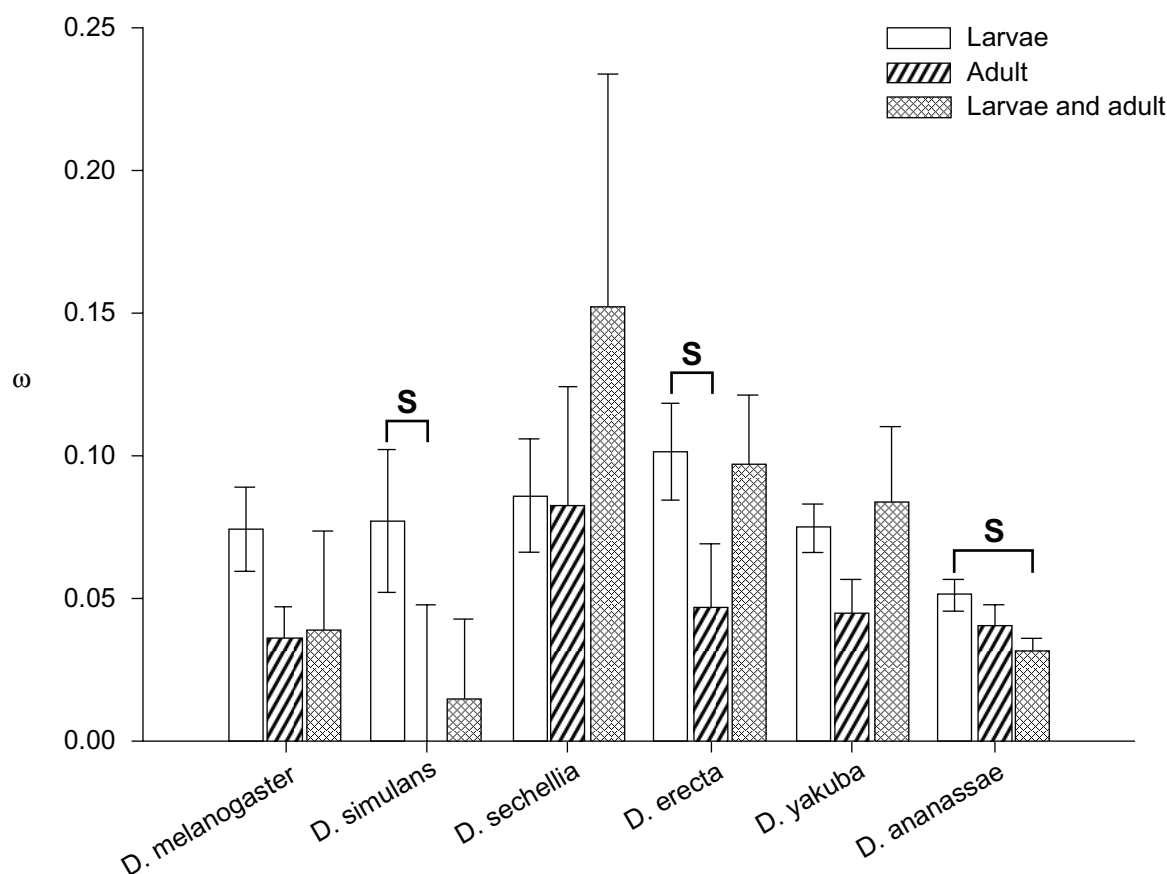
Genes expressed in a central position of the olfactory transduction pathway exhibited a median  $\omega$  ( $\omega = 0.0326$ ) that was significantly lower than for genes expressed in a peripheral position ( $\omega = 0.0785$ ) in *D. melanogaster* (Fig. 2; Wilcoxon rank sum test,  $w = 987$ ,  $P = 2.276 \cdot 10^{-4}$ ). The same is true in the other species of the *D. melanogaster* species group (Fig. 4; Wilcoxon rank sum test: in *D. sechellia*,  $w = 1696.5$ ,  $P = 0.0111$ ; in *D. yakuba*,  $w = 2407$ ,  $P = 1.82 \times 10^{-5}$ ; in *D. erecta*,  $w = 430$ ,  $P = 1.54 \times 10^{-8}$ ; in *D. ananassae*,  $w = 947$ ,  $P = 9.52 \times 10^{-5}$ ) with the only exception of

*D. simulans*. These results suggest stronger selective constraints acting on central than on peripheral genes, a pattern that seems to be independent of the particular evolutionary history and ecological context that characterize the species that belong to the *D. melanogaster* group.

However, these results may be a consequence of the physical location of the genes in the genome and/or the amount of expression and not to the position they occupy in the olfactory system. To rule out this possibility we performed a multiple regression test to evaluate in *D. melanogaster* whether there is a relationship between  $\omega$  values and cytological position and/or amount of expression and found non-significant results ( $F_{2,112} = 0.58$ ,  $P = 0.564$ ).

### Evolutionary rates in generalist vs. specialist lineages

*D. sechellia* breeds and feeds on decaying fruits of the shrub *Morinda citrifolia*<sup>34–36</sup> and *D. erecta* is a seasonal specialist on *Pandanus candelabrum*.<sup>37</sup> Previous works reported that olfactory genes evolve at a faster rate in specialists than in generalists as a consequence of directional selective pressures from a one-host environment<sup>23,24,27</sup>. In order to investigate this hypothesis we conducted two different tests. First, we compared  $\omega$  values between generalist and specialist species for *a priori* defined categories of olfactory



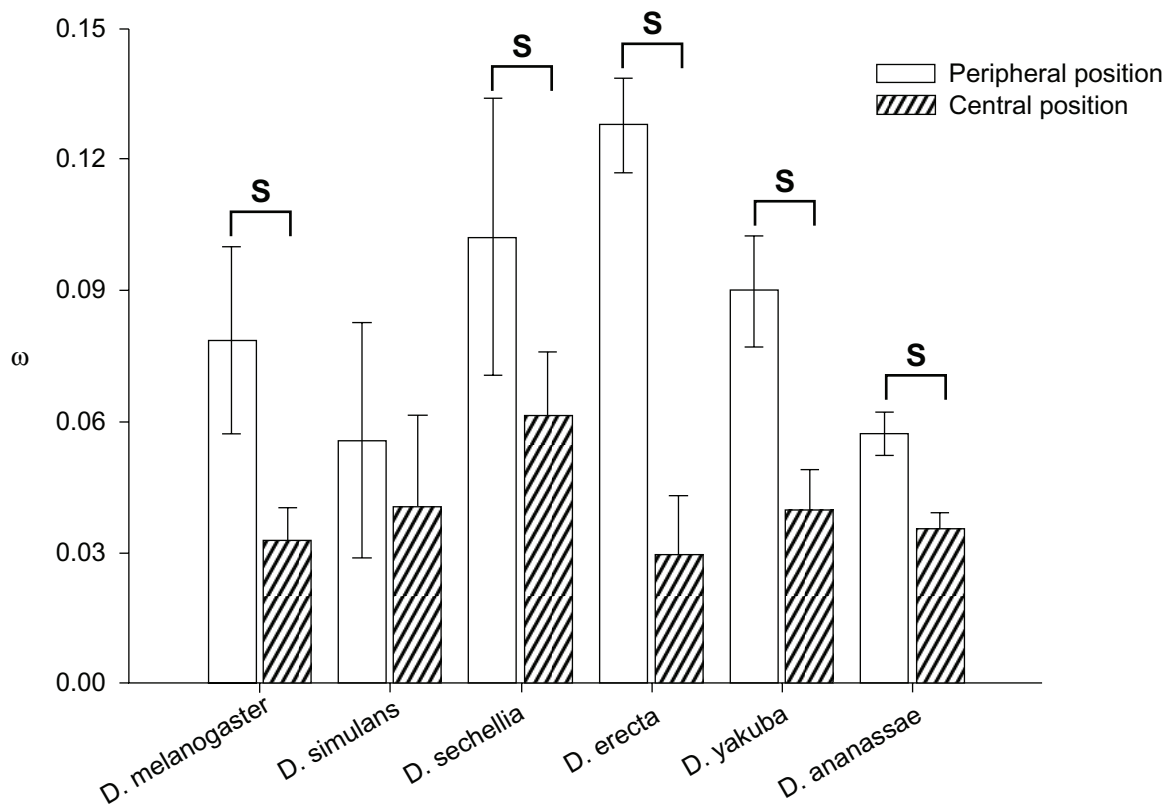
**Figure 3.** Comparisons of the  $\omega$  for olfactory genes expressed in different life cycle stages for the six species of the *D. melanogaster* species group. ML estimates of median  $\omega$  for defined gene groups according to life cycle stage of participation in olfaction for the six species of the *D. melanogaster* species group. **Notes:** Error bars are SEM values (Standard Deviation/ $N^{1/2}$ , where  $N$  is the number of genes). A Bonferroni multiple-comparison correction was applied. Significant  $P$ -values are marked with an S.

genes as described above (peripheral and central position) and used the rest of protein coding genome as control. This organization of the dataset allows us to distinguish between genome wide effects due to demographic events from local effects of PS. Specialists exhibited a significantly faster evolutionary rate than generalists when all olfactory genes were considered (Table 3; Wilcoxon Rank Sum test, Specialists vs. Generalists:  $w = 3.182$ ,  $P = 0.00146$ ). However, when we refined the analysis by dividing olfactory genes into peripheral and central, we found that differences between specialists and generalists are only accounted for by peripheral position genes (Table 3; Wilcoxon Rank Sum test, Specialists vs. Generalists:  $w = 3.793$ ,  $P < 0.001$ ) and that evolutionary rates of central position genes did not differ between specialists and generalists (Table 3; Wilcoxon Rank Sum test, Specialists vs. Generalists:  $w = 0.841$ ,  $P = 0.4$ ). Interestingly, specialists also presented a genome wide acceleration of evolutionary rates (Table 3;

Wilcoxon Rank Sum test, Specialists vs. Generalists:  $w = 25.277$ ,  $P < 0.0001$ ) relative to generalists. Moreover, comparisons between each specialist against the group of generalists revealed that *D. erecta* exhibited such acceleration (Table 3; Wilcoxon Rank Sum test, all olfactory genes:  $w = 3.122$ ,  $P = 0.0018$ ; peripheral position genes:  $w = 4.307$ ,  $P < 0.0001$ ; central position genes:  $w = 0.141$ ,  $P = 0.888$ ; rest of protein coding genome:  $w = 24.797$ ,  $P < 0.0001$ ) but not *D. sechellia* (Table 3; Wilcoxon Rank Sum test, all olfactory genes:  $w = 1.864$ ,  $P = 0.0623$ ; peripheral position genes:  $w = 1.609$ ,  $P = 0.108$ ; central position genes:  $w = 1.231$ ,  $P = 0.2183$ ; rest of protein coding genome:  $w = 14.922$ ,  $P < 0.0001$ ).

Secondly, we specifically evaluated if there has been an acceleration of non-synonymous substitution rates in the *D. erecta* and *D. sechellia* lineages. To this end, we took advantage of the genome sequences of *D. yakuba*, a generalist close relative of *D. erecta*; and *D. simulans*, a generalist close relative of *D. sechellia*,





**Figure 4.** Comparisons of the  $\omega$  for olfactory genes that participates in different positions in the olfactory system for the six species of the *D. melanogaster* species group. ML estimates of median  $\omega$  for defined gene groups according to position of participation in the olfactory transduction pathway for the six species of the *D. melanogaster* species group.

**Notes:** Error bars are SEM values (Standard Deviation/ $N^{1/2}$ , where  $N$  is the number of genes). A Bonferroni multiple-comparison correction was applied. Significant  $P$ -values are marked with an S.

to infer the sequences of the corresponding common ancestors. These analyses revealed a significant acceleration of the non-synonymous substitution rate in the branch leading to *D. erecta* for olfactory genes occupying a peripheral position in the olfactory transduction pathway. However, it should be noted that a similar trend was detected for the rest of the protein coding genome (Fig. 5; Wilcoxon Matched Pairs test, peripheral position genes:  $N = 53$ ,  $Z = 5.83$ ,  $P < 0.0001$ ; rest of protein coding genome:  $N = 8369$ ,  $Z = 28.526$ ,  $P < 0.0001$ ). Another interesting result was that a similar acceleration of the non-synonymous substitution rate was found in the branch leading to the generalist *D. yakuba* (Fig. 5; Wilcoxon Matched Pairs test, peripheral genes:  $N = 62$ ,  $Z = 3.033$ ,  $P = 0.0024$ ; central genes:  $N = 53$ ,  $Z = 2.518$ ,  $P = 0.0118$ ; rest of protein coding genome:  $N = 8369$ ,  $Z = 28.632$ ,  $P < 0.0001$ ). Therefore, there are footprints of PS for olfactory genes in *D. erecta*, however, our study also points out that the acceleration of the non-synonymous substitution rate acceleration may not be related to

specialization on one host. In contrast, the replacement substitution rate did not appear to be accelerated neither in the branch leading to the specialist *D. sechellia* (Fig. 5; Wilcoxon Matched Pairs test, peripheral genes:  $N = 62$ ,  $Z = 1.606$ ,  $P = 0.108$ ; central genes:  $N = 53$ ,  $Z = 1.171$ ,  $P = 0.241$ ), nor in the branch of the generalist *D. simulans* (Fig. 5; Wilcoxon Matched Pairs test, peripheral genes:  $N = 62$ ,  $Z = 0.382$ ,  $P = 0.702$ ; central genes:  $N = 53$ ,  $Z = 0.08$ ,  $P = 0.936$ ). Similarly, we did not detect an acceleration of the evolutionary rates of olfactory genes in the branch leading to *D. sechellia*/*D. simulans* common ancestor (Fig. 5; Wilcoxon Matched Pairs test, peripheral genes:  $N = 62$ ,  $Z = 0.022$ ,  $P = 0.982$ ; central genes:  $N = 53$ ,  $Z = 0.781$ ,  $P = 0.435$ ).

## Discussion

With regard to general evolutionary patterns we found that purifying selection is the main evolutionary process shaping olfactory genes evolution in the six species of the *D. melanogaster* species group investigated



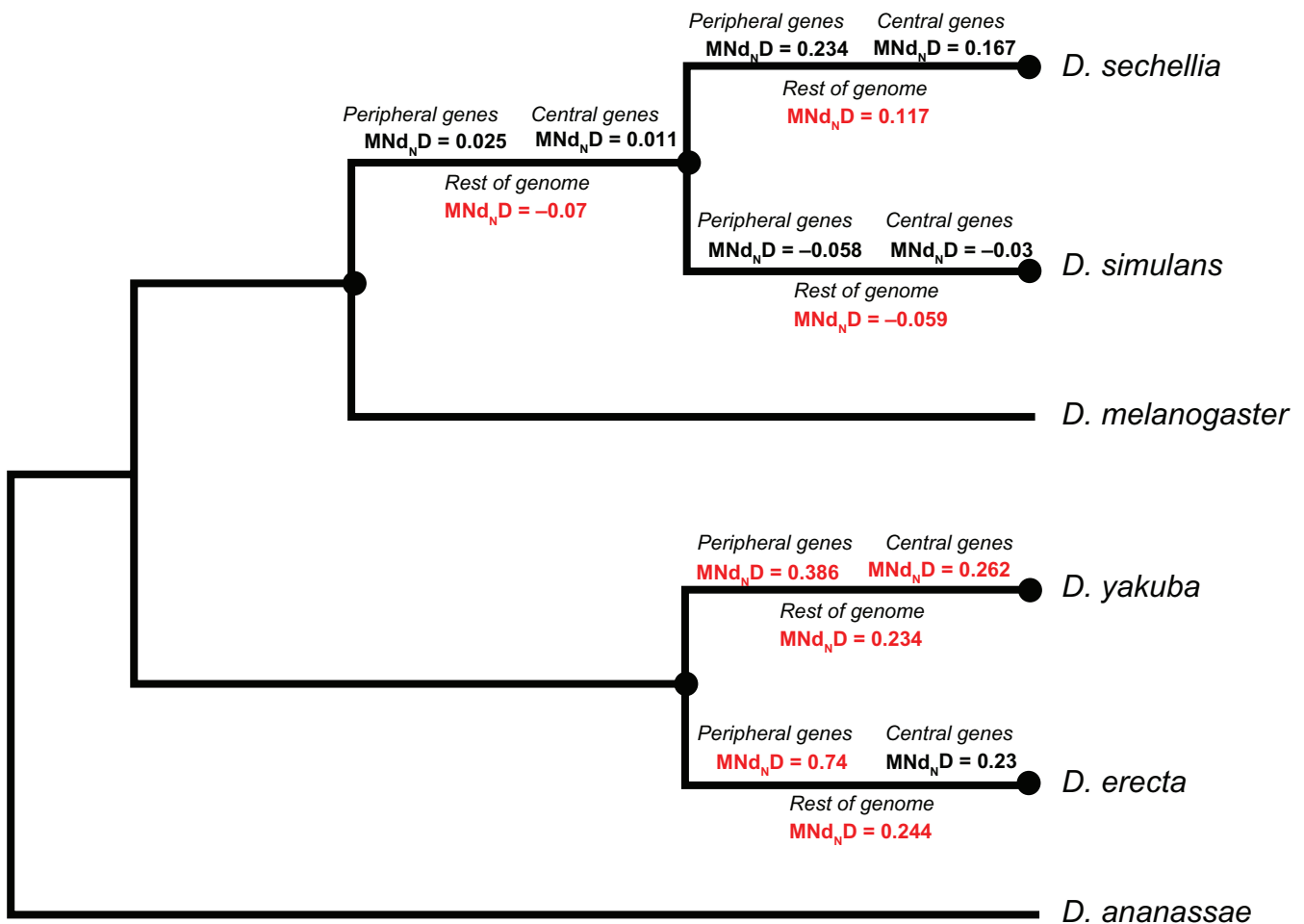
**Table 3.** Evolutionary analysis of olfactory genes in generalists vs. specialist's lineages in the *D. melanogaster* species group.

| Species                             | Peripheral position genes |             | Central position genes |             | All olfactory genes |             | Rest of protein coding genome |             |                |             |        |         |
|-------------------------------------|---------------------------|-------------|------------------------|-------------|---------------------|-------------|-------------------------------|-------------|----------------|-------------|--------|---------|
|                                     | Genes per group           | Deviation P | Genes per group        | Deviation P | Genes per group     | Deviation P | Genes per group               | Deviation P |                |             |        |         |
| Specialists vs. Generalists         | 117–779                   | 0.0426      | <0.001                 | 100–052     | 0.0047              | 0.4         | 217–731                       | 0.0315      | <b>0.00146</b> | 15968–23959 | 0.0174 | <0.0001 |
| <i>D. erecta</i> vs. Generalists    | 61–179                    | 0.0496      | <0.0001                | 53–152      | -0.0052             | 0.888       | 114–431                       | 0.0358      | <b>0.0018</b>  | 8288–23959  | 0.0184 | <0.0001 |
| <i>D. sechellia</i> vs. Generalists | 56–679                    | 0.024       | 0.108                  | 47–752      | 0.0268              | 0.2183      | 103–331                       | 0.0266      | 0.0623         | 7680–23959  | 0.016  | <0.0001 |

**Notes:** Specialist species are *D. erecta* and *D. sechellia*; generalist species are *D. melanogaster*, *D. simulans* and *D. yakuba* (*D. ananassae* was not include in this between-species comparison because, as a results of being used as out-group, the  $\omega$  estimation for this specie tend to be overestimated compared to the rest), *P*-values are from Wilcoxon Rank Sum test. A Bonferroni multiple-comparison correction was applied. Significant *P*-values are in bold.

(Fig. 1; Supplementary Table S1), in agreement with previous studies that only surveyed *Or* and *Obp* gene families.<sup>22,24–27</sup> However, applying a ML conservative test that effectively distinguishes true cases of PS from RSC, we detected signs of PS in only two olfactory genes, one in *D. sechellia* (*CG10777*) and the other in *D. erecta* (*Xrp1*) (Table 2). The remaining genes involved in the olfactory system did not show departures from neutral expectations or showed patterns compatible with purifying selection. Regarding studies reporting the occurrence of PS in olfactory genes in *D. melanogaster* there are conflicting results. Two studies did not find evidence of PS acting upon *Obp*'s and *Or*'s<sup>26,27</sup> and another two claimed that ten *Or* genes (*Or9a*, *Or10a*, *Or19a*, *Or43a*, *Or56a*, *OrN1*, *OrN2*, *Or33c*, *Or42a*, *Or85e*) exhibit the hallmark of PS.<sup>22,25</sup> In the present study we included six of the *Or* genes claimed to have evolved under PS (*Or9a*, *Or10a*, *Or33c*, *Or42a*, *Or43a*, *Or56a*) and found no trace of PS. Although ML codon-models were applied in all studies, subtle differences among studies (genes showing different results among studies are only a minor subset of the *Or* gene family) may exist because different approaches aimed at detection of positive selection may lead to slight different results even when the same, or very similar, genomic data and phylogeny are used. The aforementioned studies showed that the number of genes found to be under PS is a quite small fraction of the total number of genes analyzed. In our analysis we surveyed a large number of genes involved in olfactory behavior (not just *Or* and/or *Obp* genes) and found that only two showed traces of PS and seven exhibited signs of RSC out of a total of 565 genes in five closely related species. More than 98% of our set of genes is evolving neutrally under purifying (negative) rather than positive selection. Therefore, we concur with a recent review of genomic studies of chemosensory gene families stating that “although appealing, adaptationist interpretation might not be justified”.<sup>48</sup> Thus, the proposition of a scenario where olfactory behavior is evolving mainly by PS despite weak empirical support should be avoided.

We also tested hypotheses concerning the evolution of genes involved in olfactory behavior. A recent genome-wide study in *D. melanogaster* revealed a gradient of divergence rates over ontogeny, the earlier the expression of genes the lower



**Figure 5.** Non synonymous substitution rate acceleration analyzes in clades including specialist and generalist species. The mean normalized non-synonymous substitution rate difference (MNd<sub>N</sub>D) was calculated for specialist lineages *D. erecta* and *D. sechellia*, neighbor generalist lineages and for the reconstructed sequence of the common ancestor of *D. simulans*/*D. sechellia* and *D. erecta*/*D. yakuba*. MNd<sub>N</sub>D = (Mean d<sub>N</sub> actual lineage/Standard Deviation d<sub>N</sub> actual lineage) – (Mean d<sub>N</sub> ancestor/Standard Deviation d<sub>N</sub> ancestor). MNd<sub>N</sub>D > 0 indicate non synonymous substitution rate acceleration in derived lineages, MNd<sub>N</sub>D ≤ 0 indicate the absence of non synonymous substitution rate acceleration in derived specialist lineages. **Notes:** In red are highlighted the comparisons showing a MNd<sub>N</sub>D significantly different from zero using a Wilcoxon Matched Pairs test. A Bonferroni multiple-comparison correction was applied. Branch lengths are artificial.

the evolutionary rate (ie, embryonic < larval/pupal < adult).<sup>49</sup> Moreover, the lavish literature on olfactory behavior is consistent with the idea that larval and adult olfactory behavior are partially decoupled in terms of genetic architecture,<sup>20,31,32,50</sup> physiology,<sup>51</sup> external anatomy,<sup>52,53</sup> and molecular neuroanatomy.<sup>2</sup> These observations suggest that evolutionary rates may differ among genes that participate in olfaction in different stages of the life cycle as a result of differential functional constraints and different environmental pressures. Nevertheless, our results show that evolutionary rates of olfactory genes appear to be largely independent of both the life cycle stage and the number of stages in which they are expressed in *D. melanogaster* (Fig. 2). A similar finding was reported for

*Or* genes<sup>24</sup> and confirmed for a wide set of olfactory genes in the present study. Thus, other outcomes of evolutionary processes, rather than changes in coding sequences, should be invoked to account for the partial decoupling of larval and adult olfactory behavior. Then, gene gains and losses<sup>22,27</sup> and/or changes in expression patterns may be the key mechanisms underlying these differences.<sup>54</sup> We also evaluated the same hypothesis mentioned above in a larger dataset including all species of the *D. melanogaster* species group for which complete genomes are available. These tests revealed a significant increase in the pace of the evolution of olfactory genes expressed only in larvae (relative to genes expressed in adult and in both stages) in *D. simulans*, *D. erecta* and *D. ananassae* but



not in *D. melanogaster*, *D. sechellia* and *D. yakuba* (Fig. 3). Such among-species difference may result from different evolutionary histories reflecting differential selective pressures (or functional constraints) on olfactory genes which help flies to perceive the environment. Particularly important in fruit flies may have been the evolutionary history of host plant utilization.<sup>23,38,39</sup> Alternatively, but not mutually exclusive, differences in population structure across species, either related or unrelated to host usage evolution, may affect the relative roles of natural selection, gene flow and drift<sup>55,56</sup> and account for differences in substitution rates heterogeneity across species between genes that participate in different stages of the life cycle.

Concerning comparisons between genes that participate in different positions of the olfaction transduction pathway, stronger purifying selection was detected on central than on peripheral position genes in all species except *D. simulans* (Fig. 4), a pattern that emerges as reliable and independent of the particular historical and ecological context characterizing the species studied. The most plausible explanation for this observation is that peripheral genes (*Or* and *Obp*) only participate in the expression of olfactory behavior, while central genes tend to be more pleiotropic, since they participate not just in olfaction, but also in signal cascades or gene networks that determine other phenotypic traits. Then, more pleiotropic genes tend to be more constrained and evolve at a slower pace. It is well known that pathway-central genes are under strong functional constraints due to extensive pleiotropy.<sup>33</sup> In this sense, our present results along with empirical evidence in *Drosophila*<sup>57</sup> and fungi,<sup>58,59</sup> and also theoretical work<sup>60,61</sup> give strong support to the general hypothesis that more pleiotropic genes evolve at a slower pace.

Finally, we tested the hypothesis that evolutionary rates of olfactory genes may be accelerated in specialists (*D. sechellia* and *D. erecta*) as a response to a one host environment as compared to generalist species (*D. simulans*, *D. melanogaster*, *D. yakuba* and *D. ananassae*). Our comparisons between specialists and generalists species showed that peripheral olfactory genes (but not central) evolved faster in specialists. However, further analysis revealed that this is a genome wide trend and not a particular feature of olfactory genes. We further refine our analysis

by comparing *D. sechellia* and *D. erecta* individually to the group of generalist species and found that the acceleration in peripheral (but not central) genes can only be detected in *D. erecta*. In contrast, the genome wide acceleration as well as the lack of acceleration in central genes emerged as features common to both specialists (Table 3). These shared patterns may be a consequence of stronger functional constraints of central genes since they are highly pleiotropic. Moreover, our results concur with current views of network analysis in biology that propose that certain nodes in a network (in this case olfactory central position genes) are 'essential' and thus less prone to change.<sup>62</sup> *D. erecta* and *D. sechellia* represent two independent events of specialization<sup>38</sup> in which the evolutionary histories of olfaction genes exhibit certain subtleties that should be mentioned. On one hand, the analysis of non-synonymous substitutions rates in the clade *D. simulans/D. sechellia* revealed a pattern consistent with the idea of a genome wide acceleration restricted to the *D. sechellia* lineage, but not particularly for olfactory genes (Fig. 5). Thus, our results suggest that a scenario of directional selective pressures from a one-host-environment driving of the evolution of olfactory genes is not entirely sustainable in *D. sechellia*. Therefore, in agreement with Gardiner *et al.* (2008)<sup>26</sup> we propose that historical demographic events like small effective population size after speciation from a mainland population<sup>55</sup> and/or historical low levels of genetic variation (ie, resulting from historical events not from selective pressures)<sup>56,63</sup> cannot be ruled out as determinants of the evolution of the olfactory behavior genetic basis in the specialist *D. sechellia*. In this vein, Lachaise and Silvain argued, on the basis of historical evidence, that *M. citrifolia* may not be the ancestral host for this species, which probably evolved on *Pandanus candelabrum*, and invaded *M. citrifolia*, which is an introduced species in the Seychelles in recent times.<sup>38</sup> On the other hand, the *D. yakuba/D. erecta* clade depicted a different scenario. Comparisons of each species against the common ancestor revealed a genome wide acceleration of non-synonymous substitutions rates, as well as peripheral genes in both lineages. As in *D. sechellia*, our results argue against the one-host directional selection hypothesis. Moreover, our phylogenetic comparative analysis reveals that olfactory central genes are an exception to the genome



wide acceleration in *D. erecta*, suggesting that these genes are under stronger constraints in the lineage leading to this specialist. Although there appears to be a relationship between ecological specialization and the evolution of olfactory behavior genetic basis, this relation does not seem to be specifically related to an acceleration of evolutionary rates of olfactory genes. However, the generalized acceleration of evolutionary rates detected in *D. erecta* may provide new insights to investigate ecological specialization in a wide genomic context. Having complete genomes of other drosophilids with a habitat-specialist or narrow niche-width could help to elucidate the relationship between specialization and genome evolution.

## Conclusion

In conclusion, our study shows that evolutionary rates of olfactory genes are mostly independent of the life cycle stage and also of the number of life cycle stages in which they participate in olfaction and, more importantly, that evolutionary rates depend on the position (peripheral or central) that genes occupy in the olfactory system. Our approach is based on a conception that considers olfaction as a complex trait whose genetic bases include *Or*, *Obp* and a heterogeneous set of genes (identified as olfaction genes by means of an ample suite of methodological approaches) that do not participate in the periphery of the olfactory system. Actually, our study includes a larger set of genes identified as part of the genetic architecture of olfactory behavior than previous reports<sup>22,24–27</sup> allowing us to demonstrate that genes categorized as central in the olfactory system evolve at a slower rate than peripheral genes. In addition, we show that, at variance with expectations of the one host adaptation hypothesis, the analysis of evolutionary rates of olfactory genes does not reveal common patterns in *D. erecta* and *D. sechellia*, the two specialists investigated. Our results suggest that specialization to one host does not seem to trigger bursts of rapid evolution in olfactory genes. The relevance of our present report is twofold. Firstly, we confirm the general biological tenet that more pleiotropic genes tend to evolve at a slower pace. Secondly, on a more specific feature concerning the evolution of olfaction in *Drosophila* we found that positive selection has had a limited incidence and that acceleration of substitution rates in lineages that underwent a process of ecological specialization is a feature extended

to the entire protein coding genome not exclusive of olfactory genes.

## Acknowledgements

NL is a recipient of a scholarship from CONICET, EH is a member of Carrera del Investigador Científico of CONICET (Argentina). This work was supported by grants from Universidad de Buenos Aires, CONICET and Agencia Nacional de Promoción Científica y Técnica to EH, and grants from the Spanish Ministry of Science and Innovation's (MICINN: BFU2009-13409-C02-01, PCI2006-A7-0537) to HD.

## Disclosures

Author(s) have provided signed confirmations to the publisher of their compliance with all applicable legal and ethical obligations in respect to declaration of conflicts of interest, funding, authorship and contributorship, and compliance with ethical requirements in respect to treatment of human and animal test subjects. If this article contains identifiable human subject(s) author(s) were required to supply signed patient consent prior to publication. Author(s) have confirmed that the published article is unique and not under consideration nor published by any other publication and that they have consent to reproduce any copyrighted material. The peer reviewers declared no conflicts of interest.

## References

1. Hallem EA, Dahanukar A, Carlson JR. Insect odor and taste receptors. *Annu Rev Entomol.* 2006;51:113–35.
2. Vosshall LB, Stocker RF. Molecular architecture of smell and taste in *Drosophila*. *Annu Rev Neurosci.* 2007;30:505–33.
3. Young JM, Trask BJ. The sense of smell: genomics of vertebrate odorant receptors. *Hum Mol Genet.* 2002;11:1153–60.
4. Buck LB. The search for odorant receptors. *Cell.* 2004;116:117–9.
5. Halász N. *The vertebrate olfactory system: chemical neuroanatomy, function, and development.* Budapest: Akadémiai Kiadó; 1990.
6. Pelosi P, Maida R. Odorant-binding proteins in insects. *Comp Biochem Physiol B Biochem Mol Biol.* 1995;111:503–14.
7. Steinbrecht RA. Odorant-binding proteins: expression and function. *Ann N Y Acad Sci.* 1998;855:323–32.
8. Anholt RR, Fanara JJ, Fedorowicz GM, et al. Functional genomics of odor-guided behavior in *Drosophila melanogaster*. *Chem Senses.* 2001;26:215–21.
9. Anholt RR, Lyman RF, Mackay TF. Effects of single P-element insertions on olfactory behavior in *Drosophila melanogaster*. *Genetics.* 1996;143:293–301.
10. Goulding SE, zur Lage P, Jarman AP. *amos*, a proneural gene for *Drosophila* olfactory sense organs that is regulated by *lozenge*. *Neuron.* 2000;25:69–78.
11. zur Lage PI, Prentice DR, Holohan EE, Jarman AP. The *Drosophila* proneural gene *amos* promotes olfactory sensillum formation and suppresses bristle formation. *Development.* 2003;130:4683–93.



12. Shiraiwa T, Nitasaka E, Yamazaki T. *Geko*, a novel gene involved in olfaction in *Drosophila melanogaster*. *J Neurogenet*. 2000;14:145–64.
13. Stoltzfus JR, Horton WJ, Grotewiel MS. Odor-guided behavior in *Drosophila* requires *calreticulin*. *J Comp Physiol A Neuroethol Sens Neural Behav Physiol*. 2003;189:471–83.
14. Sambandan D, Yamamoto A, Fanara JJ, Mackay TF, Anholt RR. Dynamic genetic interactions determine odor-guided behavior in *Drosophila melanogaster*. *Genetics*. 2006;174:1349–63.
15. Anholt RR, Mackay TF. Quantitative genetic analyses of complex behaviours in *Drosophila*. *Nat Rev Genet*. 2004;5:838–49.
16. Hansen TF. The evolution of genetic architecture. *Annu Rev Ecol Evol Syst*. 2006;37:123–57.
17. Anholt RR, Dilda CL, Chang S, et al. The genetic architecture of odor-guided behavior in *Drosophila*: epistasis and the transcriptome. *Nat Genet*. 2003;35:180–4.
18. Anholt RR, Mackay TF. The genetic architecture of odor-guided behavior in *Drosophila melanogaster*. *Behav Genet*. 2001;31:17–27.
19. Fedorowicz GM, Fry JD, Anholt RR, Mackay TF. Epistatic interactions between smell-impaired loci in *Drosophila melanogaster*. *Genetics*. 1998;148:1885–91.
20. Lavagnino NJ, Anholt RR, Fanara JJ. Variation in genetic architecture of olfactory behaviour among wild-derived populations of *Drosophila melanogaster*. *J Evol Biol*. 2008;21:988–96.
21. Sambandan D, Carbone MA, Anholt RR, Mackay TF. Phenotypic plasticity and genotype by environment interaction for olfactory behavior in *Drosophila melanogaster*. *Genetics*. 2008;179:1079–88.
22. Guo S, Kim J. Molecular evolution of *Drosophila* odorant receptor genes. *Mol Biol Evol*. 2007;24:1198–207.
23. McBride CS. Rapid evolution of smell and taste receptor genes during host specialization in *Drosophila sechellia*. *Proc Natl Acad Sci U S A*. 2007;104:4996–5001.
24. McBride CS, Arguello JR. Five *Drosophila* genomes reveal nonneutral evolution and the signature of host specialization in the chemoreceptor superfamily. *Genetics*. 2007;177:1395–416.
25. Tunstall NE, Sirey T, Newcomb RD, Warr CG. Selective pressures on *Drosophila* chemosensory receptor genes. *J Mol Evol*. 2007;64:628–36.
26. Gardiner A, Barker D, Butlin RK, Jordan WC, Ritchie MG. *Drosophila* chemoreceptor gene evolution: selection, specialization and genome size. *Mol Ecol*. 2008;17:1648–57.
27. Vieira FG, Sanchez-Gracia A, Rozas J. Comparative genomic analysis of the odorant-binding protein family in 12 *Drosophila* genomes: purifying selection and birth-and-death evolution. *Genome Biol*. 2007;8:R235.
28. Yang Z. PAML 4: phylogenetic analysis by maximum likelihood. *Mol Biol Evol*. 2007;24:1586–91.
29. Yang Z. Adaptive Molecular Evolution. In: Balding D, Bishop M, Cannings C, eds. *Handbook of statistical genetics*. 2nd ed. New York: John Wiley & Sons, Ltd.; 2003:229–54.
30. Clark AG, Eisen MB, Smith DR, et al. Evolution of genes and genomes on the *Drosophila* phylogeny. *Nature*. 2007;450:203–18.
31. Andruss BF, Lu AQ, Beckingham K. Expression of calmodulin in *Drosophila* is highly regulated in a stage- and tissue-specific manner. *Development Genes and Evolution*. 1997;206:541–5.
32. Shaver SA, Varnam CJ, Hilliker AJ, Sokolowski MB. The foraging gene affects adult but not larval olfactory-related behavior in *Drosophila melanogaster*. *Behav Brain Res*. 1998;95:23–9.
33. Hodgkin J. Seven types of pleiotropy. *Int J Dev Biol*. 1998;42:501–5.
34. Louis J, David JR. Ecological specialization in the *Drosophila melanogaster* species subgroup: a case study of *Drosophila sechellia*. *Acta Oecol*. 1986;7:215–30.
35. Tsacas L, Bächli G. *Drosophila sechellia*, n. sp., huitième espèce du sous-groupe *melanogaster* des Iles Seychelles (Diptera, Drosophilidae). *Rev Fr Entomol*. 1981;3:146–50.
36. R’Kha S, Capy P, David JR. Host-plant specialization in the *Drosophila melanogaster* species complex: a physiological, behavioral, and genetical analysis. *Proc Natl Acad Sci U S A*. 1991;88:1835–9.
37. Rio B, Couturier G, Lemeunier F, Lachaise D. Evolution d’une spécialisation saisonnière chez *Drosophila erecta* (Dipt., Drosophilidae). *Annls Soc ent Fr (NS)*. 1983;19:235–48.
38. Lachaise D, Silvain JF. How two Afrotropical endemics made two cosmopolitan human commensals: the *Drosophila melanogaster*-*D. simulans* palaeogeographic riddle. *Genetica*. 2004;120:17–39.
39. Markow TA, O’Grady PM. *Drosophila* biology in the genomic age. *Genetics*. 2007;177:1269–76.
40. Pollard DA, Iyer VN, Moses AM, Eisen MB. Widespread discordance of gene trees with species tree in *Drosophila*: evidence for incomplete lineage sorting. *PLoS Genet*. 2006;2:e173.
41. Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res*. 2004;32:1792–7.
42. Edgar RC. MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics*. 2004;5:113.
43. Zhang J, Nielsen R, Yang Z. Evaluation of an improved branch-site likelihood method for detecting positive selection at the molecular level. *Mol Biol Evol*. 2005;22:2472–9.
44. Zhang J. Frequent false detection of positive selection by the likelihood method with branch-site models. *Mol Biol Evol*. 2004;21:1332–9.
45. R Development Core Team. *A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing; 2010.
46. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J Roy Stat Soc B Stat Meth*. 1995;57:289–300.
47. Chintapalli VR, Wang J, Dow JA. Using *FlyAtlas* to identify better *Drosophila melanogaster* models of human disease. *Nat Genet*. 2007;39:715–20.
48. Whiteman NK, Pierce NE. Delicious poison: genetics of *Drosophila* host plant preference. *Trends Ecol Evol*. 2008;23:473–8.
49. Artieri CG, Haerty W, Singh RS. Ontogeny and phylogeny: molecular signatures of selection, constraint, and temporal pleiotropy in the development of *Drosophila*. *BMC Biol*. 2009;7:42.
50. Ayer RK Jr, Carlson J. *acj6*: a gene affecting olfactory physiology and behavior in *Drosophila*. *Proc Natl Acad Sci U S A*. 1991;88:5467–71.
51. Tissot M, Stocker RF. Metamorphosis in *Drosophila* and other insects: the fate of neurons throughout the stages. *Prog Neurobiol*. 2000;62:89–111.
52. Shanbhag SR, Muller B, Steinbrecht RA. Atlas of olfactory organs of *Drosophila melanogaster* 1. Types, external organization, innervation and distribution of olfactory sensilla. *Int J Insect Morphol Embryol*. 1999;28:377–97.
53. Python F, Stocker RF. Adult-like complexity of the larval antennal lobe of *D. melanogaster* despite markedly low numbers of odorant receptor neurons. *J Comp Neurol*. 2002;445(4):374–87.
54. Arbeitman MN, Furlong EE, Imam F, et al. Gene expression during the life cycle of *Drosophila melanogaster*. *Science*. 2002;297:2270–5.
55. Kliman RM, Andolfatto P, Coyne JA, et al. The population genetics of the origin and divergence of the *Drosophila simulans* complex species. *Genetics*. 2000;156:1913–31.
56. Morton RA, Choudhary M, Cariou ML, Singh RS. A reanalysis of protein polymorphism in *Drosophila melanogaster*, *D. simulans*, *D. sechellia* and *D. mauritiana*: effects of population size and selection. *Genetica*. 2004;120:101–4.
57. Alvarez-Ponce D, Aguade M, Rozas J. Network-level molecular evolutionary analysis of the *insulin/TOR* signal transduction pathway across 12 *Drosophila* genomes. *Genome Res*. 2009;19:234–42.
58. Salathe M, Ackermann M, Bonhoeffer S. The effect of multifunctionality on the rate of evolution in yeast. *Mol Biol Evol*. 2006;23:721–2.
59. He X, Zhang J. Toward a molecular understanding of pleiotropy. *Genetics*. 2006;173:1885–91.
60. Otto SP. Two steps forward, one step back: the pleiotropic effects of favoured alleles. *Proc R Soc Lond B Biol Sci*. 2004;271:705–514.
61. Gu X. Evolutionary framework for protein sequence evolution and gene pleiotropy. *Genetics*. 2007;175:1813–22.
62. Proulx SR, Promislow DE, Phillips PC. Network thinking in ecology and evolution. *Trends Ecol Evol*. 2005;20:345–53.
63. Cariou ML, Solignac M, Monnerot M, David JR. Low allozyme and mtDNA variability in the island endemic species *Drosophila sechellia* (*D. melanogaster* complex). *Experientia*. 1990;46:101–4.
64. Lavagnino NJ. *Evolución de la arquitectura genética del comportamiento olfativo en Drosophila* [PhD. Thesis]. Facultad de Ciencias Exactas y Naturales. Universidad de Buenos Aires; 2011. Available at: [http://digital.bl.fcen.uba.ar/Download/Tesis/Tesis\\_4849\\_Lavagnino.pdf](http://digital.bl.fcen.uba.ar/Download/Tesis/Tesis_4849_Lavagnino.pdf).



65. Fanara JJ, Robinson KO, Rollmann SM, Anholt RR, Mackay TF. *Vanaso* is a candidate quantitative trait gene for *Drosophila* olfactory behavior. *Genetics*. 2002;162:1321–8.
66. Pielage J, Stork T, Bunse I, Klambt C. The *Drosophila* cell survival gene *discs lost* encodes a cytoplasmic Codanin-1-like protein, not a homolog of tight junction PDZ protein Patj. *Dev Cell*. 2003;5:841–51.
67. Rollmann SM, Mackay TF, Anholt RR. *Pinocchio*, a novel protein expressed in the antenna, contributes to olfactory behavior in *Drosophila melanogaster*. *J Neurobiol*. 2005;63:146–58.
68. Galindo K, Smith DP. A large family of divergent *Drosophila* odorant-binding proteins expressed in gustatory and olfactory sensilla. *Genetics*. 2001;159:1059–72.
69. Hekmat-Scafe DS, Scafe CR, McKinney AJ, Tanouye MA. Genome-wide analysis of the odorant-binding protein gene family in *Drosophila melanogaster*. *Genome Res*. 2002;12:1357–69.
70. Pikielny CW, Hasan G, Rouyer F, Rosbash M. Members of a family of *Drosophila* putative odorant-binding proteins are expressed in different subsets of olfactory hairs. *Neuron*. 1994;12:35–49.
71. Park SK, Shanbhag SR, Wang Q, Hasan G, Steinbrecht RA, Pikielny CW. Expression patterns of two putative odorant-binding proteins in the olfactory organs of *Drosophila melanogaster* have different implications for their functions. *Cell Tissue Res*. 2000;300:181–92.



## Supplementary Data

Supplementary Table S1 is available from 8484Supplementaryfile.zip

## Appendix

Parameters and input data details of maximum likelihood models in CodeML.

### Free-ratio maximum likelihood model in CodeML

This model allows  $\omega$  values to be different in each branch of a given phylogeny. Parameters: model = 1; CodonFreq = 2; ncatG = 8; NSSites = 0; fix omega = 0; clock = 0; runmode = 0; fix kappa = 0; fix alpha = 1; method = 0; Malpha = 0; aaDist = 0; RateAncestor = 0; icode = 0; alpha = 0.0; seqtype = 1; getSE = 0; Mgene = 0; omega = 0.4; kappa = 2. Last 2 parameters omega and kappa are let free however starting value should be supplied.

### Models used in the Zhang-Nielsen-Yang 2005 tests<sup>43</sup>

*Model M1a* parameters: model = 0; CodonFreq = 2; ncatG = 10; cleandata = 0; fix blength = 0; NSSites = 1; fix omega = 0; clock = 0; runmode = 0; fix kappa = 0; aaDist = 0; RateAncestor = 0; icode = 0; seqtype = 1; getSE = 0; Mgene = 0; omega = 1.3; kappa = 0.3.

*Model A* parameters: model = 2; CodonFreq = 2; ncatG = 8; cleandata = 0, NSSites = 2; fix omega = 0; clock = 0; runmode = 0; fix kappa = 0; fix alpha = 1;

method = 0; Malpha = 0; aaDist = 0; RateAncestor = 0; icode = 0; alpha = 0.0; seqtype = 1; getSE = 0; Mgene = 0; omega = 0.7; kappa = 2.

*Model A1* parameters: model = 2; CodonFreq = 2; ncatG = 8; cleandata = 0; NSSites = 2; fix omega = 1; clock = 0; runmode = 0; fix kappa = 0; fix alpha = 1; method = 0; Malpha = 0; aaDist = 0; RateAncestor = 0; icode = 0; alpha = 0.0; seqtype = 1; getSE = 0; Mgene = 0; omega = 1; kappa = 2.

How these models are used to perform the Zhang-Nielsen-Yang 2005 tests<sup>43</sup> is explained in detail in Materials and Methods.

In all cases input data were CDS's in PHYLIP format aligned with the software MUSCLE<sup>41,42</sup> using predicted amino acid sequences as templates and the input tree in newick format used was: ((*D. melanogaster*, (*D. simulans*, *D. sechellia*)), (*D. erecta*, *D. yakuba*), *D. ananassae*).

**Table S1.**  $\omega$  estimation for each olfactory gene in the six species of the *D. melanogaster* species group. Also showing gene categorizations concerning the life cycle stage (larval and/or adult) in which is involved in olfaction and position of expression in the olfactory system (peripheral or central).

### Publish with Libertas Academica and every scientist working in your field can read your article

*"I would like to say that this is the most author-friendly editing process I have experienced in over 150 publications. Thank you most sincerely."*

*"The communication between your staff and me has been terrific. Whenever progress is made with the manuscript, I receive notice. Quite honestly, I've never had such complete communication with a journal."*

*"LA is different, and hopefully represents a kind of scientific publication machinery that removes the hurdles from free flow of scientific thought."*

#### Your paper will be:

- Available to your entire community free of charge
- Fairly and quickly peer reviewed
- Yours! You retain copyright

<http://www.la-press.com>