




Article

Diversity and Molecular Evolution of Odorant Receptor in Hemipteran Insects

Jiahui Tian ^{1,2} , Youssef Dewer ³ , Haoyuan Hu ¹, Fengqi Li ^{2,*} , Shiyong Yang ^{1,4,*} and Chen Luo ²

¹ School of Ecology and Environment, Anhui Normal University, Wuhu 241002, China; tianjh0920@163.com (J.T.); haoyuanhu@126.com (H.H.)

² Beijing Key Laboratory of Environment Friendly Management on Fruit Diseases and Pests in North China, Institute of Plant and Environment Protection, Beijing Academy of Agriculture and Forestry Sciences, Beijing 100097, China; luochen1010@126.com

³ Central Agricultural Pesticide Laboratory, Agricultural Research Center, Phytotoxicity Research Department, Dokki, Giza 12618, Egypt; dewer72@yahoo.com

⁴ Collaborative Innovation Center of Recovery and Reconstruction of Degraded Ecosystem in Wanjiang Basin Co-Founded by Anhui Province and Ministry of Education, Anhui Normal University, Wuhu 241000, China

* Correspondence: pandit@163.com (F.L.); shiyan@ahnu.edu.cn (S.Y.)

Simple Summary: Insects' behavior and ecology are closely related to their chemosensory systems, during which odorant receptors (ORs) play an essential role in host recognition. Although OR gene evolution has been studied in many insect orders, a comprehensive evolutionary analysis and expression of OR gene gain and loss events among diverse hemipteran species are still needed. In this study, we identified and analyzed the OR genes from hemipteran species systematically. The number of OR genes discovered in each species ranged from less than ten to hundreds. Gene gain and loss events of OR have occurred in several species in the seven major clades classified through phylogenetic analysis. Then, we discovered the amino acid differences between species to understand the molecular evolution of OR in the order Hemiptera through positive selection. This study lays a foundation for subsequent investigations into the molecular mechanisms of Hemiptera olfactory receptors involved in host recognition.

Abstract: Olfaction is a critical physiologic process for insects to interact with the environment, especially plant-emitted volatiles, during which odorant receptors (ORs) play an essential role in host recognition. Although OR gene evolution has been studied in many insect orders, a comprehensive evolutionary analysis and expression of OR gene gain and loss events among diverse hemipteran species are still required. In this study, we identified and analyzed 887 OR genes from 11 hemipteran species. The number of OR genes discovered in each species ranged from less than ten to hundreds. Phylogenetic analysis revealed that all identified Hemiptera OR genes were classified into seven major clades. Gene gain and loss events of OR have occurred in several species. Then, by positive selection, we discovered the amino acid differences between species to understand the molecular evolution of OR in the order Hemiptera. Additionally, we discussed how evolutionary analysis can aid the study of insect–plant communication. This study lays a foundation for subsequent investigations into the molecular mechanisms of Hemiptera olfactory receptors involved in host recognition.

Keywords: hemiptera; odorant receptor; phylogeny; molecular evolution



Citation: Tian, J.; Dewer, Y.; Hu, H.; Li, F.; Yang, S.; Luo, C. Diversity and Molecular Evolution of Odorant Receptor in Hemipteran Insects. *Insects* **2022**, *13*, 214. <https://doi.org/10.3390/insects13020214>

Academic Editors: Liang Sun and Jin Zhang

Received: 21 December 2021

Accepted: 15 February 2022

Published: 21 February 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Insect behavior and ecology are closely related to their chemosensory system and are involved in various activities, such as host plant selection, mating, intra-specific communication, and avoidance behaviors [1,2]. A key step in insect chemosensation is the detection of chemicals by receptor proteins located on the membranes of peripheral sensory neurons, which convert chemical signals into electrical signals that travel into the insect's central

nervous system, ultimately resulting in behavioral responses. Chemoreceptors primarily include three main receptor families [3], gustatory receptors (GRs), ionotropic receptors (IRs), and odorant receptors (ORs). GRs contribute to detecting carbon dioxide, sugars, bitter components, salts, and several gustatory pheromones [3–6]. IRs mainly detect acids and amines [7], and ORs are key receptors involved in the detection of volatile compounds, including host plant volatiles [6].

The insect odorant receptors are seven-transmembrane domain proteins [8] that are classified into an odorant receptor co-receptor (Orco) and conventional ORx. The Orco and ORx form a heterotetrameric receptor complex on the membrane of olfactory sensory neurons, which is involved in olfactory recognition [9,10]. Orco is widely expressed and highly conserved in insects, highlighting its critical role in olfaction [11]. The conventional ORx varies in insects [12]. ORs have been discovered in various orders of insects, including Diptera, Hymenoptera, Lepidoptera, Coleoptera, and Hemiptera since the discovery of the first insect OR in *Drosophila melanogaster* [13,14] by using the whole-genome sequencing method. Additionally, the number of ORs in different species varies greatly [15]. For example, *Drosophila melanogaster* has 62 ORs [14] while *Tribolium castaneum* has 259 ORs [16], indicating a variable evolution of impact OR genes [17]. In Coleoptera, the diversity of OR genes in a given species is correlated with its ecology, and the number of ORs is also parallel to the host [15].

The evolution of ORs between different insect species remains unclear. This is partly due to the lack of genomic annotation information in some species, especially non-model species. Although some species have transcriptome data, OR genes have high and rapid diversity and low expression levels in different tissues, making OR annotation incomplete [18]. With the increase in genome annotation data, some gene families and genera have been annotated and compared successively in Diptera, Lepidoptera, Hymenoptera, and Coleoptera [15,19–21]. According to these findings, the OR gene family has been gained and lost numerous times in insects and has undergone rapid evolution, associated with host specificity and insect sociality [21,22]. In Hemiptera, although ORs have been identified and annotated in several species [12,23], the gain, loss, and evolution of OR genes among different species have not been well-explained.

In Hemiptera, the types and numbers of hosts in different species vary significantly. In the Delphacidae family, for example, *Nilaparvata lugens* is a monophagous pest that only feeds on rice, whereas *Laodelphax striatellus* is an oligophagous insect, which mainly feeds on rice but also feeds on other gramineous plants, including corn and tares [24]. The *Aphis gossypii* and the *Bemisia tabaci* in the Aphididae family are polyphagous insects. They host over 700 plant species globally [25,26]. The Hemiptera includes insects that feed on vertebrate blood, such as *Rhodnius prolixus*. It also includes *Cimex lectularius*, which feeds on human blood [12,27]. In the Hemiptera with such complex hosts, how the olfactory receptor adapts to different hosts has not been well-explained.

It is unclear whether the number of olfactory receptors varies in hemipteran species or whether there is any gene loss or gain as in other insect orders and adaptive evolution. Therefore, this study referred to the bioinformatics pipeline to systematically annotate the OR of 11 species from nine Hemiptera families, one species from Thysanoptera, and one from Corrodentia, closest to Hemiptera. We compared the sequence diversity and evolution of ORs in different species through phylogenetic and selection pressure analysis. Results of the present study may lay a foundation for further study of the function of Ors.

2. Materials and Methods

2.1. Genome Data and Integrity Assessment

For hemipteran insects, we selected 11 species (*Nilaparvata lugens*, *Halyomorpha halys*, *Cimex lectularius*, *Apolygus lucorum*, *Gerris buenoi*, *Rhodnius prolixus*, *Diaphorina citri*, *Acyrtosiphon pisum*, *Aphis gossypii*, *Trialeurodes vaporariorum*, and *Bemisia tabaci*), of which three cryptic species (Middle East-Asia Minor 1 (MEAM1), Mediterranean (MED), and Sub-Saharan Africa-East and Central Africa (SSA-ECA) were selected by *B. tabaci*. Furthermore,

we selected *Pediculus humanus* of Corrodentia and *Frankliniella occidentalis* of Thysanoptera as outgroups of Hemiptera. Table S1 shows the genome assemblies of these species.

We assessed the genomics of each species using BUSCO v5.1.2 [28]. In *F. occidentalis* and *P. humanus*, we used the insecta_odb10 database (https://busco-data.ezlab.org/v4/data/lineages/insecta_odb10.2020-09-10.tar.gz) (accessed on 8 February 2022), and in the hemipteran insects, we used the hemiptera_odb10 database (https://busco-data.ezlab.org/v4/data/lineages/hemiptera_odb10.2020-08-05.tar.gz) (accessed on 8 February 2022).

2.2. Species Tree

The genome-annotated proteins of each species were processed and filtered using OrthoFinder (v.2.3.8) [29] to provide information about orthologous gene families. Orthoclades in OrthoFinder are defined as homologous genes of a single gene from the last universal common ancestor of the species under examination. The 1083 orthoclades were used to construct the species tree with a minimum of 83.3% of species having single-copy genes in any orthoclade [30,31].

2.3. Identification and Annotation of Odorant Receptor Family

We referred to a bioinformatics pipeline [32,33] to identify and annotate the members of the ORs in Hemiptera. We used the protein sequences of the known gene family in insects, downloaded from the National Center for Biotechnology Information (NCBI) as a database [12,23,25,34–40], and the database was used to construct a hidden Markov Model (HMM) profile for the family. The genome sequence, proteome sequence, and general feature format (GFF) annotation file of each identified hemipteran species were used as input files.

For the specific process, we first launched BLASTP and HMMER to search for OR gene families already present in the input GFF file. We filtered BLAST and HMMER hits with a default cut-off E-value of 1×10^{-5} . The screen and retain sequences covered two-thirds of the query sequences or over 80% of the complete proteins. Results of the two searches were pooled to generate a new GFF file. Then, we used TBLASTN to search for unannotated regions of protein homologs in the genome sequence. The gene model mapper (GeMoMa) tools based on homology and intron position conservation were employed to develop new gene models. Additionally, we used the input HMM configuration file to check the existence of gene family-specific protein domains. Subsequently, we utilized all of the proteins obtained in the first two steps as input data to perform a second search. NR annotation was applied to all of these putative OR genes, and all of the non-ORs were excluded from analysis. Finally, OR genes were finally identified using their specific seven transmembrane domains.

We utilized MAFFT v.7.037b [41], IQ-TREE [42] to construct the maximum likelihood tree of OR genes of each species and NCBI search to verify the accuracy of OR genes with the obtained non-redundant data set. OR genes were nominated according to the position of OR genes in the phylogenetic tree of each species.

2.4. Phylogenetic Analysis of OR Families

We aligned the identified protein sequences of the OR family in all species utilizing MAFFT (“-auto” option) [41]. IQ-TREE (v.2.1.2) was employed to estimate the best fit substitution models (best-fit model: LG + I + G4 chosen according to BIC) and gene family trees [42]. Additionally, node support was estimated with 1000 bootstrap replicates [43]. We created the tree using FigTree (v.1.3.1) and rooted it in the Orco lineage.

2.5. Molecular Evolutionary Analysis

Genes subjected to positive selection were analyzed using the branch-site model in the codeml program in PAML v.4.9 software package [44]. Then, we aligned the nucleotide sequences using ClustalW (codons) [45] and then submitted the maximum likelihood tree and multiple sequence alignment files to the PAML software to calculate the ratio of dN

(non-synonymous substitution rate) to dS (synonymous substitution rate) at each site. The magnitude of dN/dS value (ω) represents the type of selection: $\omega < 1$ for negative selection, $\omega = 1$ for neutral selection and $\omega > 1$ for positive selection [46]. We employed the branch and branch-site models to check the positive selection of the branch, where the Aleyrodidae is located in clades two and four, and the branch where *H. halys* is located in clades three and four. In the branch model, the hypothesis was that the two-ratio model was significantly different from the one-ratio model (null, M0 one-ratio model), where only one ω value was estimated across the phylogeny. We performed the likelihood ratio test to check positive selection in the branch-site model. Finally, we mapped the amino acids under positive selection onto the odorant receptor topology to examine the distribution of the inferred positively selected sites using TMHMM.

3. Results

3.1. The Number of OR Genes in Different Species of Hemiptera Varied

We used BUSCO to evaluate the integrity of 13 species downloaded from Insect Base, NCBI, Ensembl, and whitefly genome databases, among which *B. tabaci* includes the genomes of three cryptic species. Except for the whole BUSCO genes of *D. citri*, which was 88%, the complete BUSCO genes of the other selected species were above 90% (Figure 1). This result verifies the integrity of the genome used for subsequent identification of ORs.

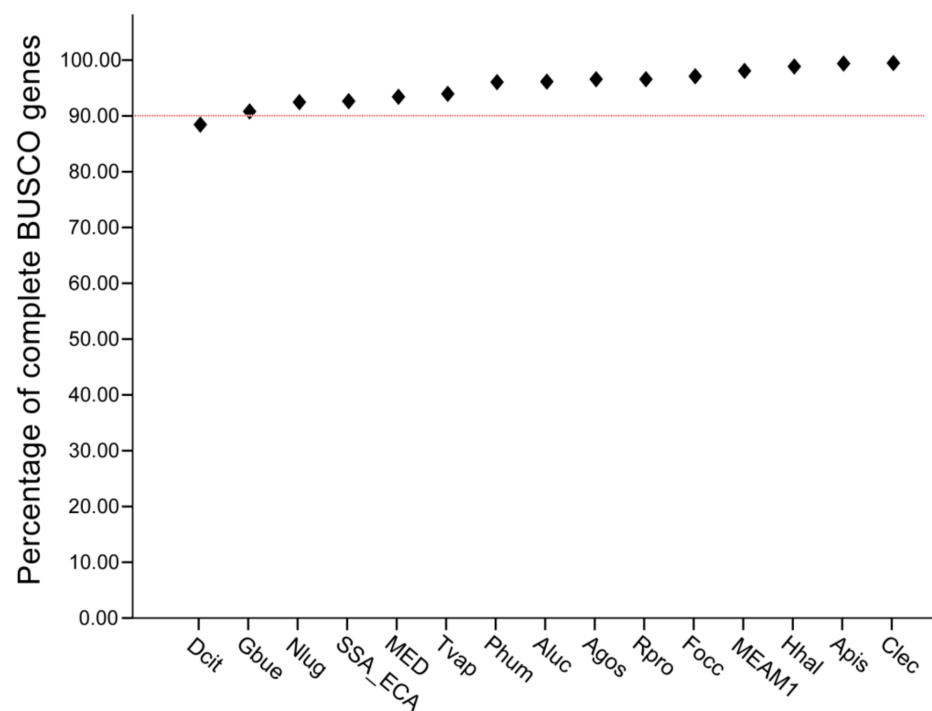


Figure 1. The completeness threshold (% of complete BUSCO genes) of the species genome for OR gene identification. *Pediculus humanus* (Phum), *Frankliniella occidentalis* (Focc), *Nilaparvata lugens* (Nlug), *Halyomorpha halys* (Hhal), *Cimex lectularius* (Clec), *Apolygus lucorum* (Aluc), *Rhodnius prolixus* (Rpro), *Gerris buenoi* (Gbue), *Diaphorina citri* (Dcit), *Acyrtosiphon pisum* (Apis), *Aphis gossypii* (Agos), *Trialeurodes vaporariorum* (Tvap), and *Bemisia tabaci* (MEAM1, MED, and SSA-ECA).

We annotated 887 OR genes from the genomes of 11 hemipteran species (including three cryptic species in *B. tabaci*) (Figure 2, Table S4). The genetic relationship between species in the species tree was consistent with a previous report [47]. The number of OR genes in different species ranges from 9 to 155 (Figure 2). We identified 11 and 39 OR genes in the genomes of *P. humanus* and *F. occidentalis*, respectively. Additionally, we identified 54, 152, and 155 OR genes in genomes of the herbivorous hemipteran insects *N. lugens*, *H. halys*, and *A. lucorum*, respectively, and 102,100 and 47 OR genes in *D. citri* in Chermidae, *A. pisum*,

and *A. gossypii* in Aphididae, respectively. In Aleyrodidae, we identified 20 OR genes in the genome of *T. vaporariorum* and 13, 12, and 9 OR genes in genomes of the whitefly SSA-ECA, MEAM1, and MED, respectively. In genomes of the blood-sucking hemipteran insects *R. prolixus* and *C. lectularius*, 129 and 63 OR genes were identified, respectively. In the genome of *G. buenoi* that feeds on body fluids of small worms and dead fish, 30 OR genes were identified. Based on the number of exons and introns in the identified sequences, we found that the number of exons in the complete OR sequence varied from 5 to 9 (Table S2). Protein sequences, nucleotide sequences, and the GFF file of all ORs used in this study are available in Table S4.

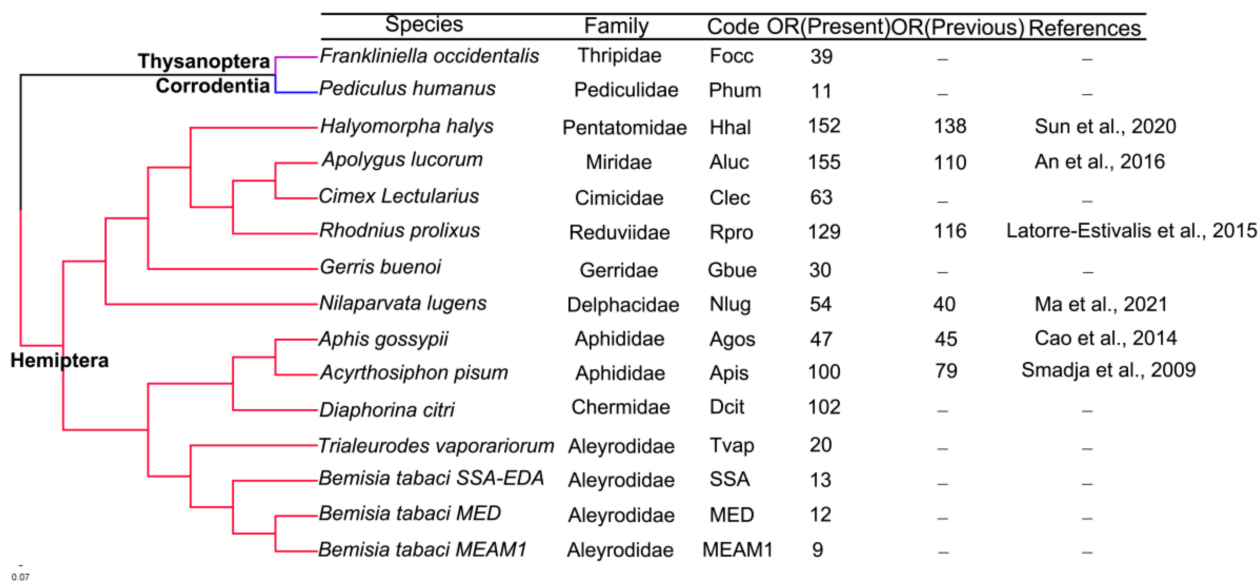


Figure 2. Numbers of odorant receptor genes annotated from genomes of one species of Corrodentia, one species of Thysanoptera and 10 species of Hemiptera. Phylogeny and taxonomy of the study species (left) are constructed using OrthoFinder.

3.2. OR Genes in Different Species Have Specific Expansion and Contraction

Seven major subfamilies of ORs that contain the ORs of Hemiptera were detected in our analysis (Figure 2). The Orco genes of all species are clustered together due to their highly conserved amino acid sequences and are located at the base of the entire phylogenetic tree, forming the root of the phylogenetic tree. Except for clades three and six, which had less than 50% support, other clades received stronger support (Figures 3 and S1).

The OR genes among the seven subfamilies showed systemic gain and loss. Clade One contained only six ORs, each of which comes from *N. lugens*, *H. halys*, *A. lucorum*, *R. prolixus*, *C. lectularius*, and *G. buenoi*. Clade Two contained 43 OR genes that were separated into two main branches, one of which had 20 ORs and 19 of which were from the Aleyrodidae. Another branch had 23 OR genes and 14 of them were from *C. lectularius*. Clade Three contained 163 ORs, which were separated into two main branches. One branch contained 46 OR genes and 41 were from *H. halys*; the other branch contained 117 OR genes, and all were from *A. pisum* and *A. gossypii*. Clade Four contained 88 ORs, which were separated into three branches: the first branch had 36 OR genes and 19 of them were from the Aleyrodidae; the second branch had 20 OR genes and 12 of them were from *A. lucorum*; the third branch had 32 OR genes and 25 of them were from *H. halys*. Clade Five had 144 ORs, including a branch with *R. prolixus*-specific expansion. Clade Six had 115 ORs, and 91 of them were *D. citri*, accounting for 89% of the ORs identified in this clade. Clade Seven had 315 ORs, including a branch with *N. lugens*-specific expansion and two branches with *A. lucorum*-specific expansion.

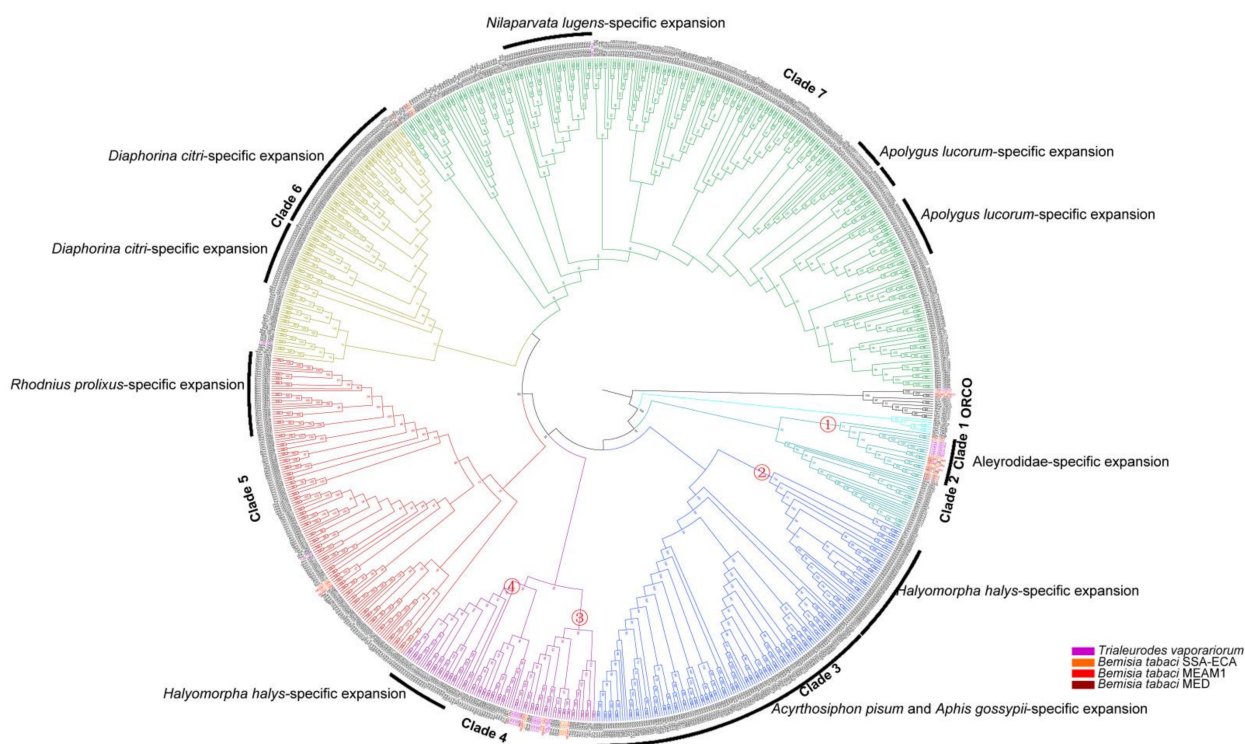


Figure 3. Phylogenetic relationships of the odorant receptor genes from ten species of Hemiptera: *Nilaparvata lugens* (Nlug), *Halyomorpha halys* (Hhal), *Cimex lectularius* (Clec), *Apolygus lucorum* (Aluc), *Rhodnius prolixus* (Rpro), *Gerris buenoi* (Gbue), *Diaphorina citri* (Dcit), *Acyrthosiphon pisum* (Apis), *Aphis gossypii* (Agos), *Trialeurodes vaporariorum* (Tvap), and *Bemisia tabaci* (MEAM1, MED, SSA-ECA). The phylogenetic tree was constructed from the amino acid sequence using the maximum likelihood. The tree was rooted with conserved olfactory coreceptor (Orco) genes. Nodes with support >50% were displayed in the tree.

3.3. OR Genes in Aleyrodidae and *Halyomorpha Halys* Undergo Positive Selection

The number of ORs in *H. halys* is 152, with obvious gene expansion. In contrast, the ORs genes in Aleyrodidae had a specific contraction. Therefore, the OR genes of *H. halys* and Aleyrodidae were selected for selection pressure analysis. The branch and branch-site models of the CODEML program [44] in the PAML v4.9 package are the most commonly used models for selection pressure analysis. We selected two branches of the Aleyrodidae OR gene in Clades two and four and two branches of *H. halys* OR gene in Clades three and four for labeling (Figure 3). For the branch model, likelihood ratio tests revealed that log-likelihood values under the free ratio model (M1) were significantly higher ($p < 0.001$) than those under the one ratio model (M0) when branches one, two, three, and four were selected as foreground branches (Table 1). These results indicated that selective pressure varies among branches in each targeted clade and that some of these proteins have evolved under positive selection.

In the branch-site model, we failed to detect any site with a posterior probability greater than 0.95 when Branches One, Three, and Four were selected as foreground branches (Table 2). In contrast, we detected nine sites (90 G, 206 I, 222 E, 271 V, 699 C, 711 A, 761 G, 775 G, and 792 E) that had undergone positive selection, and the posterior probability values were greater than 95% only when Branch Two was selected as the foreground branch (Table 2). This means that positive selection sites can only be identified in Branch Two. For ORs, the TMHMM algorithm predicted the NH2 tail intracellular direction and the COOH tail extracellular direction, which supports the “reverse” membrane topology of insect chemoreceptors [8,48,49]. We mapped all assumed positive selection sites in Branch Two onto the topology of OR. When PP > 95%, we observed that three of the nine positive selection sites located outside the transmembrane region (ER), four of the nine

positive selection sites inside the transmembrane region (IR) and two sites in the sixth transmembrane region of the OR gene (Figure 4).

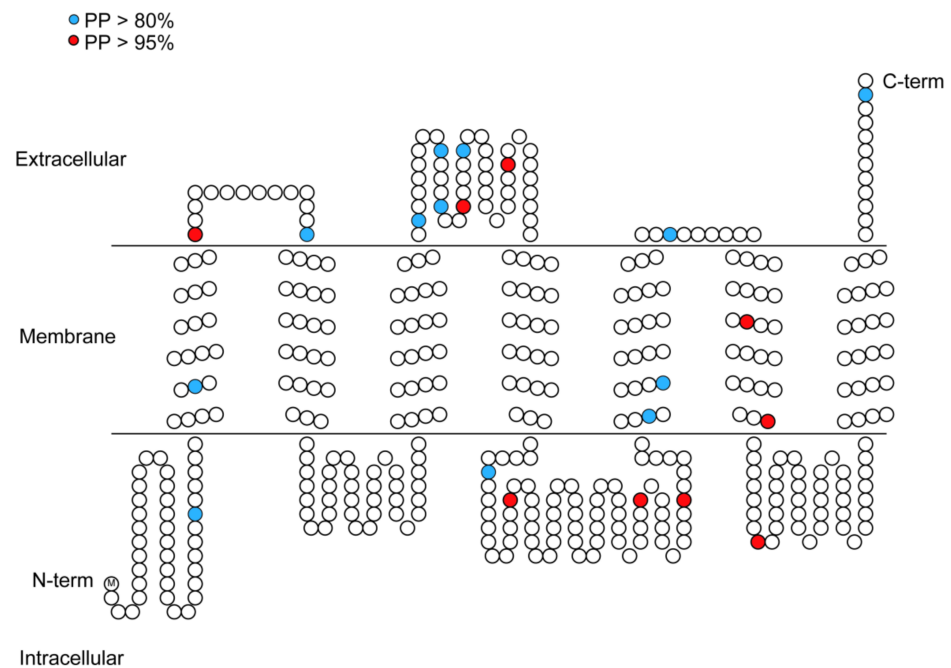


Figure 4. Positively selected sites on the structural topology of odorant receptors (ORs). Circles indicate the amino acids (white neutrally selected sites; blue positively selected sites with PPs 80%; red positively selected sites with PPs 95%).

Table 1. Likelihood ratio test of the branch model for Clade 2, Clade 3, and Clade 4.

Branch	Models	LnL	2ΔL
1	M0 (one ratio)	−36,525.665006	353.227994 **
	M1 (free ratio)	−36,349.051009	
2	M0 (one ratio)	−35,609.079852	219.037414 **
	M1 (free ratio)	−35,499.561145	
3	M0 (one ratio)	−35,793.429897	420.317102 **
	M1 (free ratio)	−35,583.271346	
4	M0 (one ratio)	−78,349.502738	700.319204 **
	M1 (free ratio)	−77,999.343136	

2ΔL is twice the log-likelihood difference between models M1 and M0. ** Chi-square test indicates the difference at the high significance level of $p < 0.01$.

Table 2. Summary of statistics for the detection of positive selection for Clade 2, Clade 3, and Clade 4 homology using the branch-site model.

Branch	LnL (nuLL)	LnL (Alternative)	2ΔL	P	Sites					
1	−36,382.9	−36,379.1	7.4833	0.006227	NONE					
					90 G 0.998 **					
					206 I 0.970 *					
					222 E 0.983 *					
					271 V 0.954 *					
					699 C 0.979 *					
2	−35,267.5	−35,258.1	18.82128	1.44×10^{-5}	711 A 0.980 *					
					761 G 0.958 *					
					775 G 0.989 *					
					792 E 0.979 *					
					3	−35,664.3	−35,673.7	18.89361	1.38×10^{-5}	NONE
					4	−78,177.5	−78,168.6	17.93291	2.29×10^{-5}	NONE

2ΔL is twice the log-likelihood difference between models M1 and M0. ** Chi-square test indicates the difference at the high significance level of $p < 0.01$. * Chi-square test indicates the difference at the significance level of $p < 0.05$.

4. Discussion

We evaluated the relative effectiveness of our method for OR gene detection in genome. We evaluated the relative effectiveness of our method for detecting OR gene pools in the genome. By comparing the number of OR genes retrieved in this study with those found in previous studies in Hemiptera, *H. halys* [39], *R. prolixus* [12], *A. lucorum* [17], *A. pisum* [23], *A. gossypii* [25], and *N. lugens* [50], we detected more OR genes than previously reported (Figure 2). The use of antenna transcriptome sequences in previous studies contributed to an increasing number of ORs in *H. halys* [39]. Among the species reported using genome identification, the most recent genome version has much higher quality than the older versions used for OR identification. Furthermore, the applicability of the bioinformatics pipeline and its advantages in annotating gene families in genomes of varying quality have been proven in nine species to identify chemosensory receptor genes, including the model organism *Drosophila melanogaster* [32,51,52]. The results obtained through the bioinformatics pipeline were more accurate than those annotated by augustus-ppx. In addition, to overcome putative gene model errors, the pipeline method implements some filtering steps to determine the correctness of predicted coding sequences. For example, hmmer searches were performed to identify protein family domains in all newly annotated sequences, resulting in more complete and sufficient identification of OR genes [32]. All of these indicate that this process is reliable and feasible in the identification of insect OR genes.

The number of OR genes in different Hemipteran species has obvious gain and loss events. Gene family expansion and contraction are considered the outcome of random genomic drift or environmental adaptation [53,54]. Studies have shown that the number of OR has an association with the breadth of its host. For example, in Coleoptera, the stenophagous *Agrilus planipennis* has fewer ORs than the oligophagous *Leptinotarsa decemlineata* and *Dendroctonus ponderosae*, which have smaller complements of ORs than the polyphagous or scavenging species *Anoplophora glabripennis*, *Onthophagus taurus*, and *Tribolium castaneum* [22]. However, the number of OR genes in Hemiptera is not parallel to that of host breadth. The biotope of different species differs considerably. Some species live on land, while others perched on static water, such as *G. buenoi*. In terms of feeding habits, some are herbivorous insects, some feed on blood (*R. prolixus*), and others feed on small insects (*G. buenoi*). Among the herbivorous species, there are monophagous (*N. lugens*), oligophagous (*L. striatellus*), and polyphagous (*B. tabaci*) species. All of these factors may be linked to the differences in ORs number. Although both *R. prolixus* and *C. lectularius* feed on vertebrate blood [12,27], *R. prolixus* ORs are twice as numerous as *C. lectularius*. In the herbivorous Hemiptera, both *H. halys* and *B. tabaci* are polyphagous [55,56]. A total of 152 OR genes were identified in *H. halys*, while only 9, 12, and 13 were identified among the three cryptic species of *B. tabaci*, respectively. It demonstrates that in hemipteran insects, the number of OR genes in the species is not directly associated with the type and number of hosts.

Some orthologous genes showing high sequence homology may have similar spatial structures and biological functions [20]. In Hemiptera, the functional studies of OR receptors are relatively less. However, previous studies have characterized the function of OR genes in some species of Hemiptera (Table S3). In *A. pisum*, Apis_OR17 and Apis_OR28 mainly identify plant volatiles [57,58], and Apis_OR6 is mainly linked to the alarm pheromone EBF and geranyl acetate [59]. Among them, Apis_OR28 was observed in Clade Three, where the OR gene specifically expanded into *A. pisum*. Apis_OR6 and Apis_OR17 were mainly observed in Clade Seven. In *A. lucorum*, Aluc_OR29 and Aluc_OR39 mainly recognize compounds that attract females [60,61], and Aluc_OR77 as a sex pheromone receptor mainly recognizes (E)-2-hexenyl butyrate and hexyl butyrate [17]. Aluc_OR39 was observed in Clade Three, Aluc_OR29 in Clade Six, and the sex pheromone receptor Aluc_OR77 in Clade Seven. In *C. lectularius* [62], Clec_OR26 mainly recognizes the aggregation pheromone, and Clec_OR16 and Clec_OR49 mainly recognize human volatiles for host location; Clec_OR28, Clec_OR34, and Clec_OR37 mainly recognize the repellent

compounds terpenes and terpenoids. We observed Clec_OR16, Clec_OR34, and Clec_OR37 in Clade Four; Clec_OR28 in Clade Five; and Clec_OR26 and Clec_OR49 in Clade Seven. A putative inference is that the types of ligands that additional OR genes with high homology may bind by combining the function of OR genes in Hemiptera with their positions in the phylogenetic tree. Additionally, through phylogenetic analysis, it is speculated that some new genes in some species have similar functions to known genes, which will provide a reference for subsequent functional studies.

Interestingly, the specific expansion of OR genes may be associated with the existence of positive selection, and the acquisition of new hosts may be driven by positive selection [20]. *H. halys* has specific expansion into Clades Three and Four. We discovered that these two branches were being driven by positive selection, and nine positive selection sites were identified in the OR gene of *H. halys* in Clade Three. Two of the sites with PP > 95% are located at the sixth transmembrane position of OR, and the amino acids of this transmembrane region are considered to be highly conserved [10]. Additionally, the specific expansion of the OR gene in *A. pisum* undergoes a positive selection [23]. This finding indicates that gene duplication is necessary for hemipteran insects to acquire new chemosensory functions and is driven by changes in ligands [63]. For example, lineage-specific duplication of OR genes contributes to adaptation to specific ecological conditions in *Drosophila* [64]. As we all know, hemipteran insects rely mainly on their chemosensory system to recognize chemicals in the ambient environment, and OR genes under positive selection pressure may drive the acquisition of new hosts. This may also explain why *H. halys* and *B. tabaci* have a wide host range. Additionally, this provides valuable information for the various odors and tastants encountered by hemipteran insects when they explore new habitats and environments.

5. Conclusions

In summary, we identified the OR gene library and found seven major subfamilies in 887 OR genes in the 11 species of Hemiptera (*B. tabaci* contains three cryptic species). We discovered that OR gene has a specific expansion in *N. lugens*, *H. halys*, *A. lucorum*, *D. citri*, *A. pisum*, and *A. gossypii*. In Aleyrodidae, the OR gene has obvious gene loss. The expansion process of the OR gene in *H. halys* has undergone positive selection, resulting in the discovery of candidate genes involved in olfactory recognition under positive selection pressure. Results of the present study may lay the foundation for future investigation into the function of ORs in hemipteran insects.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/insects13020214/s1>, Table S1: Genome assemblies in hemipteran species; Table S2: Functional characterization of OR genes, Table S3: Functional characterization of OR genes; Table S4: Protein sequences, nucleotide sequences, and the GFF file of all ORs used in this study; Figure S1: Phylogenetic relationships of the odorant receptor genes from ten species of Hemiptera.

Author Contributions: J.T., S.Y. and F.L. conceived the study; J.T. and H.H. performed the analyses; J.T., F.L. and H.H. conceived and made graphs and figures; J.T., Y.D. and S.Y. wrote the manuscript; H.H., F.L. and C.L. writing—review & editing. All authors discussed the results and commented on the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: This study was supported by the projects of The National Key R&D Program of China (2019YFD1002100), Lure Research and application technology of Agricultural sucking pests in Ningxia of China (NGSB-2021-10-02), Beijing Municipal Natural Science Foundation (6222008), National Science Foundation of China (31070338).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available in supplementary material.

Acknowledgments: We thank Julio Rozas provided valuable guidance for using the bioinformatics pipeline to identify ORs.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Suh, E.; Bohbot, J.D.; Zwiebel, L.J. Peripheral olfactory signaling in insects. *Curr. Opin. Insect Sci.* **2014**, *6*, 86–92. [[CrossRef](#)]
2. Gadenne, C.; Barrozo, R.-B.; Anton, S. Plasticity in Insect Olfaction: To Smell or Not to Smell? *Annu. Rev. Entomol.* **2016**, *61*, 17–333. [[CrossRef](#)] [[PubMed](#)]
3. Dahanukar, A.; Foster, K.; Naters, W.M.V.D.G.V.; Carlson, J.R. A Gr receptor is required for response to the sugar trehalose in taste neurons of *Drosophila*. *Nat. Neurosci.* **2001**, *4*, 1182–1186. [[CrossRef](#)] [[PubMed](#)]
4. Moon, S.J.; Köttgen, M.; Jiao, Y.; Xu, H.; Montell, C. A Taste Receptor Required for the Caffeine Response In Vivo. *Curr. Biol.* **2006**, *16*, 1812–1817. [[CrossRef](#)] [[PubMed](#)]
5. Jones, W.; Cayirlioglu, P.; Kadow, I.G.; Vosshall, L. Two chemosensory receptors together mediate carbon dioxide detection in *Drosophila*. *Nature* **2006**, *445*, 86–90. [[CrossRef](#)] [[PubMed](#)]
6. Fleischer, J.; Pregitzer, P.; Breer, H.; Krieger, J. Access to the odor world: Olfactory receptors and their role for signal transduction in insects. *Cell. Mol. Life Sci.* **2017**, *75*, 485–508. [[CrossRef](#)] [[PubMed](#)]
7. Zhang, J.; Bisch-Knaden, S.; Fandino, R.A.; Yan, S.; Obiero, G.F.; Grosse-Wilde, E.; Hansson, B.S.; Knaden, M. The olfactory coreceptor IR8a governs larval feces-mediated competition avoidance in a hawkmoth. *Proc. Natl. Acad. Sci. USA* **2019**, *116*, 21828–21833. [[CrossRef](#)]
8. Smart, R.; Kiely, A.; Beale, M.; Vargas, E.; Carraher, C.; Kralicek, A.-V.; Christie, D.L.; Chen, C.; Newcomb, R.D.; Warr, C.G. *Drosophila* odorant receptors are novel seven trans-membrane domain proteins that can signal independently of heterotrimeric G proteins. *Insect Biochem. Mol. Biol.* **2008**, *38*, 70–80. [[CrossRef](#)]
9. Leal, W.S. Odorant Reception in Insects: Roles of Receptors, Binding Proteins, and Degrading Enzymes. *Annu. Rev. Entomol.* **2013**, *58*, 373–391. [[CrossRef](#)]
10. Butterwick, J.A.; del Marmol, J.; Kim, K.H.; Kahlson, M.; Rogow, J.; Walz, T.; Ruta, V. Cryo-EM structure of the insect olfactory receptor Orco. *Nature* **2018**, *560*, 447–452. [[CrossRef](#)]
11. Missbach, C.; Dweck, H.; Vogel, H.; Vilcinskas, A.; Stensmyr, M.; Hansson, B.-S.; Grosse-Wilde, E. Evolution of insect olfactory re-ceptors. *eLife* **2014**, *3*, e02115. [[CrossRef](#)] [[PubMed](#)]
12. Latorre-Estivalis, J.-M.; de Oliveira, E.-S.; Beiral Esteves, B.; Santos Guimaraes, L.; Neves Ramos, M.; Lorenzo, M.-G. Patterns of ex-pression of odorant receptor genes in a Chagas disease vector. *Insect Biochem. Mol. Biol.* **2015**, *69*, 1–81.
13. Gao, Q.; Chess, A. Identification of candidate *Drosophila* olfactory receptors from genomic DNA sequence. *Genomics* **1999**, *60*, 1–39. [[CrossRef](#)]
14. Clyne, P.-J.; Warr, C.-G.; Freeman, M.-R.; Lessing, D.; Carlson, J.-R. A novel family of divergent seven-transmembrane proteins: Can-didate odorant receptors in *Drosophila*. *Neuron* **1999**, *22*, 27–338. [[CrossRef](#)]
15. Mitchell, R.F.; Schneider, T.M.; Schwartz, A.M.; Andersson, M.N.; McKenna, D.D. The diversity and evolution of odorant receptors in beetles (Coleoptera). *Insect Mol. Biol.* **2019**, *29*, 77–91. [[CrossRef](#)]
16. Engsontia, P.; Sanderson, A.-P.; Cobb, M.; Walden, K.-K.; Robertson, H.-M.; Brown, S. The red flour beetle’s large nose: An expanded odorant receptor gene family in *Tribolium castaneum*. *Insect Biochem. Mol. Biol.* **2008**, *38*, 87–97. [[CrossRef](#)]
17. An, X.-K.; Sun, L.; Liu, H.-W.; Liu, D.-F.; Ding, Y.-X.; Li, L.-M.; Zhang, Y.-J.; Guo, Y.-Y. Identification and expression analysis of an olfactory receptor gene family in green plant bug *Apolygus lucorum* (Meyer-Dür). *Sci. Rep.* **2016**, *6*, 37870. [[CrossRef](#)]
18. Schoville, S.D.; Chen, Y.H.; Andersson, M.N.; Benoit, J.B.; Bhandari, A.; Bowsher, J.H.; Brevik, K.; Cappelle, K.; Chen, M.-J.M.; Childers, A.K.; et al. A model species for agricultural pest genomics: The genome of the Colorado potato beetle, *Leptinotarsa decemlineata* (Coleoptera: Chrysomelidae). *Sci. Rep.* **2018**, *8*, 1–18. [[CrossRef](#)]
19. Gardiner, A.; Barker, D.; Butlin, R.-K.; Jordan, W.-C.; Ritchie, M.-G. *Drosophila* chemoreceptor gene evolution: Selection, specialization and genome size. *Mol. Ecol.* **2008**, *17*, 1648–1657. [[CrossRef](#)]
20. Engsontia, P.; Sangket, U.; Chotigeat, W.; Satasook, C. Molecular evolution of the odorant and gustatory receptor genes in lepi-dopteran insects: Implications for their adaptation and speciation. *J. Mol. Evol.* **2014**, *79*, 1–39. [[CrossRef](#)] [[PubMed](#)]
21. Zhou, X.; Rokas, A.; Berger, S.; Liebig, J.; Ray, A.; Zwiebel, L.J. Chemoreceptor Evolution in Hymenoptera and Its Implications for the Evolution of Eusociality. *Genome Biol. Evol.* **2015**, *7*, 2407–2416. [[CrossRef](#)] [[PubMed](#)]
22. Andersson, M.N.; Keeling, C.I.; Mitchell, R.F. Genomic content of chemosensory genes correlates with host range in wood-boring beetles (*Dendroctonus ponderosae*, *Agrilus planipennis*, and *Anoplophora glabripennis*). *BMC Genom.* **2019**, *20*, 690. [[CrossRef](#)]
23. Smadja, C.; Shi, P.; Butlin, R.K.; Robertson, H.M. Large Gene Family Expansions and Adaptive Evolution for Odorant and Gustatory Receptors in the Pea Aphid, *Acyrtosiphon pisum*. *Mol. Biol. Evol.* **2009**, *26*, 2073–2086. [[CrossRef](#)] [[PubMed](#)]
24. He, P.; Engsontia, P.; Chen, G.-L.; Yin, Q.; Wang, J.; Lu, X.; Zhang, Y.-N.; Li, Z.-Q.; He, M. Molecular characterization and evolution of a chemosensory receptor gene family in three notorious rice planthoppers, *Nilaparvata lugens*, *Sogatella furcifera* and *Laodelphax striatellus*, based on genome and transcriptome analyses. *Pest Manag. Sci.* **2018**, *74*, 2156–2167. [[CrossRef](#)] [[PubMed](#)]
25. Cao, D.; Liu, Y.; Walker, W.B.; Li, J.; Wang, G. Molecular Characterization of the *Aphis gossypii* Olfactory Receptor Gene Families. *PLoS ONE* **2014**, *9*, e101187. [[CrossRef](#)] [[PubMed](#)]

26. Li, S.-J.; Xue, X.; Ahmed, M.Z.; Ren, S.-X.; Du, Y.-Z.; Wu, J.-H.; Cuthbertson, A.G.S.; Qiu, B.-L. Host plants and natural enemies of *Bemisia tabaci* (Hemiptera: Aleyrodidae) in China. *Insect Sci.* **2011**, *18*, 101–120. [[CrossRef](#)]
27. Potter, M.F. The Perfect Storm: An Extension View on Bed Bugs. *Am. Entomol.* **2006**, *52*, 102–104. [[CrossRef](#)]
28. Waterhouse, R.M.; Seppely, M.; Simão, F.A.; Manni, M.; Ioannidis, P.; Klioutchnikov, G.; Kriventseva, E.V.; Zdobnov, E.M. BUSCO Applications from Quality Assessments to Gene Prediction and Phylogenomics. *Mol. Biol. Evol.* **2018**, *35*, 543–548. [[CrossRef](#)]
29. Emms, D.M.; Kelly, S. OrthoFinder: Phylogenetic orthology inference for comparative genomics. *Genome Biol.* **2019**, *20*, 1–14. [[CrossRef](#)]
30. Emms, D.-M.; Kelly, S. STRIDE: Species Tree Root Inference from Gene Duplication Events. *Mol. Biol. Evol.* **2017**, *34*, 267–3278. [[CrossRef](#)]
31. Emms, D.-M.; Kelly, S. OrthoFinder2: Fast and accurate phylogenomic orthology analysis from gene sequences. *bioRxiv.* 2018.
32. Vizueta, J.; Sánchez-Gracia, A.; Rozas, J. bitacora: A comprehensive tool for the identification and annotation of gene families in genome assemblies. *Mol. Ecol. Resour.* **2020**, *20*, 1445–1452. [[CrossRef](#)] [[PubMed](#)]
33. Vizueta, J.; Escuer, P.; Sánchez-Gracia, A.; Rozas, J. Genome mining and sequence analysis of chemosensory soluble proteins in arthropods. *Methods Enzymol.* **2020**, *642*, 1–20.
34. Bohbot, J.; Pitts, R.J.; Kwon, H.-W.; Rützler, M.; Robertson, H.M.; Zwiebel, L.J. Molecular characterization of the *Aedes aegypti* odorant receptor gene family. *Insect Mol. Biol.* **2007**, *16*, 525–537. [[CrossRef](#)]
35. Hill, C.A.; Fox, A.N.; Pitts, R.J.; Kent, L.B.; Tan, P.L.; Chrystal, M.A.; Cravchik, A.; Collins, F.H.; Robertson, H.M.; Zwiebel, L.J. G Protein-Coupled Receptors in *Anopheles gambiae*. *Science* **2002**, *298*, 176–178. [[CrossRef](#)]
36. Wanner, K.W.; Anderson, A.R.; Trowell, S.C.; Theilmann, D.A.; Robertson, H.M.; Newcomb, R.D. Female-biased expression of odourant receptor genes in the adult antennae of the silkworm, *Bombyx mori*. *Insect Mol. Biol.* **2007**, *16*, 107–119. [[CrossRef](#)]
37. Xu, P.; Choo, Y.-M.; Pelletier, J.; Sujimoto, F.R.; Hughes, D.T.; Zhu, F.; Atungulu, E.; Cornel, A.J.; Luetje, C.W.; Leal, W.S. Silent, generic and plant kairomone sensitive odorant receptors from the Southern house mosquito. *J. Insect Physiol.* **2013**, *59*, 961–966. [[CrossRef](#)]
38. Robertson, H.-M.; Warr, C.-G.; Carlson, J.-R. Molecular evolution of the insect chemoreceptor gene superfamily in *Drosophila melanogaster*. *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 4537–4542. [[CrossRef](#)]
39. Sun, D.; Huang, Y.; Qin, Z.; Zhan, H.; Zhang, J.; Liu, Y.; Yang, S. Identification of Candidate Olfactory Genes in the Antennal Transcriptome of the Stink Bug *Halyomorpha halys*. *Front. Physiol.* **2020**, *11*, 76. [[CrossRef](#)]
40. Koenig, C.; Hirsh, A.; Bucks, S.; Klinner, C.; Vogel, H.; Shukla, A.; Mansfield, J.H.; Morton, B.; Hansson, B.S.; Grosse-Wilde, E. A reference gene set for chemosensory receptor genes of *Manduca sexta*. *Insect Biochem. Mol. Biol.* **2015**, *66*, 51–63. [[CrossRef](#)]
41. Katoh, K.; Standley, D.-M. MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Molecular biology and evolution.* **2013**, *30*, 72–780. [[CrossRef](#)] [[PubMed](#)]
42. Nguyen, L.-T.; Schmidt, H.A.; Von Haeseler, A.; Minh, B.Q. IQ-TREE: A Fast and Effective Stochastic Algorithm for Estimating Maximum-Likelihood Phylogenies. *Mol. Biol. Evol.* **2015**, *32*, 268–274. [[CrossRef](#)]
43. Hoang, D.T.; Chernomor, O.; Von Haeseler, A.; Minh, B.Q.; Vinh, L.S. UFBoot2: Improving the Ultrafast Bootstrap Approximation. *Mol. Biol. Evol.* **2018**, *35*, 518–522. [[CrossRef](#)] [[PubMed](#)]
44. Yang, Z. PAML 4: Phylogenetic Analysis by Maximum Likelihood. *Mol. Biol. Evol.* **2007**, *24*, 1586–1591. [[CrossRef](#)]
45. Thompson, J.D.; Gibson, T.J.; Higgins, D.G. Multiple Sequence Alignment Using ClustalW and ClustalX. *Curr. Protoc. Bioinform.* **2003**, 2–3. [[CrossRef](#)]
46. Yang, Z.; Wong, W.S.; Nielsen, R. Bayes Empirical Bayes Inference of Amino Acid Sites Under Positive Selection. *Mol. Biol. Evol.* **2005**, *22*, 1107–1118. [[CrossRef](#)] [[PubMed](#)]
47. Johnson, K.-P.; Dietrich, C.-H.; Friedrich, F.; Beutel, R.-G.; Wipfler, B.; Peters, R.-S.; Allen, J.M.; Petersen, M.; Donath, A.; Walden, K.K.; et al. Phylogenomics and the evolution of hemipteroid insects. *Proc. Natl. Acad. Sci. USA* **2018**, *115*, 2775–12780. [[CrossRef](#)] [[PubMed](#)]
48. Sato, K.; Pellegrino, M.; Nakagawa, T.; Nakagawa, T.; Vossell, L.B.; Touhara, K. Insect olfactory receptors are heteromeric ligand-gated ion channels. *Nature* **2008**, *452*, 1002–1006. [[CrossRef](#)] [[PubMed](#)]
49. Wicher, D.; Schäfer, R.; Bauernfeind, R.; Stensmyr, M.; Heller, R.; Heinemann, S.H.; Hansson, B.S. *Drosophila* odorant receptors are both ligand-gated and cyclic-nucleotide-activated cation channels. *Nature* **2008**, *452*, 1007–1011. [[CrossRef](#)]
50. Ma, W.; Xu, L.; Hua, H.; Chen, M.; Guo, M.; He, K.; Zhao, J.; Li, F. Chromosomal-level genomes of three rice planthoppers provide new insights into sex chromosome evolution. *Mol. Ecol. Resour.* **2020**, *21*, 226–237. [[CrossRef](#)]
51. Vizueta, J.; Rozas, J.; Sánchez-Gracia, A. Comparative Genomics Reveals Thousands of Novel Chemosensory Genes and Massive Changes in Chemoreceptor Repertoires across Chelicerates. *Genome Biol. Evol.* **2018**, *10*, 1221–1236. [[CrossRef](#)] [[PubMed](#)]
52. Vizueta, J.; Frias-López, C.; Macías-Hernández, N.; Arnedo, M.; Sánchez-Gracia, A.; Rozas, J. Evolution of chemosensory gene families in arthropods: Insight from the first inclusive comparative transcriptome analysis across spider appendages. *Genome Biol. Evol.* **2016**, *9*, 178–196. [[CrossRef](#)] [[PubMed](#)]
53. Sánchez-Gracia, A.; Vieira, F.G.; Almeida, F.C.; Rozas, J. Comparative Genomics of the Major Chemosensory Gene Families in Arthropods. In *eLS*; Wiley: Hoboken, NJ, USA, 2011.
54. Dong, D.; He, G.; Zhang, S.; Zhang, Z. Evolution of Olfactory Receptor Genes in Primates Dominated by Birth-and-Death Process. *Genome Biol. Evol.* **2009**, *1*, 258–264. [[CrossRef](#)] [[PubMed](#)]

55. Bayhan, E.; Ulusoy, M.-R.; Brown, J.-K. Host range, distribution, and natural enemies of *Bemisia tabaci* 'B biotype' (Hemiptera: Aleyrodidae) in Turkey. *J. Pest Sci.* **2006**, *79*, 33–240. [[CrossRef](#)]
56. De Barro, P.-J.; Liu, S.-S.; Boykin, L.-M.; Dinsdale, A.-B. *Bemisia tabaci*: A Statement of Species Status. *Annu. Rev. Entomol.* **2011**, *56*, 1–19. [[CrossRef](#)]
57. Zhang, R.; Liu, Y.; Yan, S.-C.; Wang, G.-R. Identification and functional characterization of an odorant receptor in pea aphid, *Acyrtosiphon pisum*. *Insect Sci.* **2017**, *26*, 58–67. [[CrossRef](#)]
58. Huang, T.-Y.; Zhang, R.-B.; Yang, L.-L.; Cao, S.; Frederic, F.; Wang, G.-R. Identification and functional characterization of a novel odorant receptor in pea aphid *Acyrtosiphon pisum*. *J. Integr. Agric.* **2021**, *20*, 2–11.
59. Zhang, R.; Wang, B.; Grossi, G.; Falabella, P.; Liu, Y.; Yan, S.; Lu, J.; Xi, J.; Wang, G. Molecular Basis of Alarm Pheromone Detection in Aphids. *Curr. Biol.* **2016**, *27*, 55–61. [[CrossRef](#)]
60. Zhang, Z.; Zhang, M.; Yan, S.; Wang, G.; Liu, Y. A Female-Biased Odorant Receptor from *Apolygus lucorum* (Meyer-Dür) Tuned to Some Plant Odors. *Int. J. Mol. Sci.* **2016**, *17*, 1165. [[CrossRef](#)]
61. Yan, S.-W.; Zhang, J.; Liu, Y.; Li, G.-Q.; Wang, G.-R. An olfactory receptor from *Apolygus lucorum* (Meyer-Dur) mainly tuned to vol-atiles from flowering host plants. *J. Insect Physiol.* **2015**, *79*, 6–41. [[CrossRef](#)]
62. Liu, F.; Chen, Z.; Liu, N. Molecular Basis of Olfactory Chemoreception in the Common Bed Bug, *Cimex lectularius*. *Sci. Rep.* **2017**, *7*, 45531. [[CrossRef](#)] [[PubMed](#)]
63. Nei, M.; Niimura, Y.; Nozawa, M. The evolution of animal chemosensory receptor gene repertoires: Roles of chance and necessity. *Nat. Rev. Genet.* **2008**, *9*, 951–963. [[CrossRef](#)] [[PubMed](#)]
64. Guo, S.; Kim, J. Molecular Evolution of Drosophila Odorant Receptor Genes. *Mol. Biol. Evol.* **2007**, *24*, 1198–1207. [[CrossRef](#)] [[PubMed](#)]