

## Article

# Identification and Tissue Expression Profiles of Odorant Receptor Genes in the Green Peach Aphid *Myzus persicae*

Jingtao Liu <sup>1,2</sup>, Jiaoxin Xie <sup>2</sup>, Adel Khashaveh <sup>2</sup> , Jingjiang Zhou <sup>3</sup>, Yongjun Zhang <sup>2</sup> , Hui Dong <sup>1</sup> , Bin Cong <sup>1,\*</sup> and Shaohua Gu <sup>4,\*</sup> 

- <sup>1</sup> College of Plant Protection, Shenyang Agricultural University, Shenyang 110866, China; knighteva@163.com (J.L.); biocontrol@163.com (H.D.)  
<sup>2</sup> State Key Laboratory for Biology of Plant Disease and Insect Pest, Institute of Plant Protection, China Academy of Agricultural Science, Beijing 100193, China; xiejiaoxin18@126.com (J.X.); akhashaveh@caas.cn (A.K.); yjzhang@ippcaas.cn (Y.Z.)  
<sup>3</sup> State Key Laboratory of Green Pesticide and Agricultural Bioengineering, Guizhou University, Guiyang 550025, China; jjzhouchina@163.com  
<sup>4</sup> Department of Entomology, China Agricultural University, Beijing 100193, China  
\* Correspondence: bin1956@163.com (B.C.); gushaohua@cau.edu.cn (S.G.);  
Tel.: +86-24-88487148 (B.C.); +86-10-62733313 (S.G.)

**Simple Summary:** The green peach aphid is a pest that mainly endangers vegetables. It spreads a variety of plant viruses through pricking and sucking plant juice. At the same time, it also secretes honeydew, which causes mildew and pollution and leads to the loss of many crops. Olfaction plays a very important role in the aphids' search for hosts, avoidance of natural enemies, mating and oviposition. However, the molecular mechanism of olfactory recognition in the green peach aphid is not clear. Odorant receptors (ORs) are a type of olfactory neuron membrane receptor expressed on insect antennae and they play a role in identifying odorants from the environment. We identified 33 OR genes from the genome and transcriptomes of the green peach aphid and analyzed their phylogenetic relationship and genomic structure. We further examined the expression patterns in different tissues by real-time quantitative PCR (RT-qPCR). This study enriches our understanding of the olfactory system of the green peach aphid and will help to better apply it to the integrated control of a variety of pests.

**Abstract:** The green peach aphid *Myzus persicae* (Hemiptera: Aphididae) relies heavily on its olfactory system to locate plant hosts, find mates, and avoid parasitoids or predators. The insect odorant receptors (ORs) have been proven to play a critical role in the perception of odorants from the environment. In the present study, 33 odorant receptor candidate genes including the Orco gene were identified from the antennal, head, legs and body transcriptomes of *M. persicae*. Phylogenetic analysis of ORs from seven different orders of insect species suggests that ORs from different insect species are highly divergent and most ORs from the same species formed monophyletic groups. In addition, the aphid ORs were clustered into six different sub-clades in the same clade. Furthermore, the genomic structure of the OR genes also tends to be consistent, suggesting that ORs from the family Aphididae have a relatively close evolutionary relationship. Reads per kilobase per million (RPKM) and tissue expression profiles analyses revealed that 27 out of the 33 MperORs were uniquely or primarily expressed in the antennae, indicating their putative roles in chemoreception. This work provides a foundation to further investigate the molecular and ecological functions of MperORs in the aphid–aphid, aphid–plant and aphid–natural enemy interactions.

**Keywords:** *Myzus persicae*; insect olfaction; odorant receptor; expression profile; genomic structure



**Citation:** Liu, J.; Xie, J.; Khashaveh, A.; Zhou, J.; Zhang, Y.; Dong, H.; Cong, B.; Gu, S. Identification and Tissue Expression Profiles of Odorant Receptor Genes in the Green Peach Aphid *Myzus persicae*. *Insects* **2022**, *13*, 398. <https://doi.org/10.3390/insects13050398>

Academic Editor: Daniel Doucet

Received: 9 March 2022

Accepted: 19 April 2022

Published: 20 April 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

The green peach aphid, *Myzus persicae* (Hemiptera: Aphididae), is one of the most economically important crop pests worldwide. This aphid has an exceptional ability to colonize more than 400 plant species from more than 40 families for parthenogenetic reproduction, and uses peach, its primary host, for sexual reproduction [1–3]. Indeed, the status of this species as a pest is enhanced by its global distribution, remarkable efficiency as a vector of more than 100 different plant viruses, and an extremely broad host range [4]. The wide distribution of *M. persicae* in the world is due to its extremely high population adaptability to the environment, its wide genetic variability, and broad phenotypic plasticity. In addition, *M. persicae* is a typical host-alternating aphid species and usually heteroecious and holocyclic, with the winter sexual phase spent on the branches of peach trees and the parthenogenetic (asexual) summer generations spent on a wide number of secondary herbaceous hosts [3], for which chemical communication plays a critical role for their fast transference and precise positioning among different host-plants [1,5]. The genome of *M. persicae* was sequenced in 2017 [6] in order to better study the pest at the molecular level.

Olfaction is an important sensory modality of chemical communication that allows insects to perceive semiochemicals in a complex environment [7,8]. On the antennae, a variety of sensilla are distributed, within which usually contain two or more olfactory sensory neurons (OSNs). Odorant receptors (ORs) are expressed on the dendritic membrane of OSNs for the identification of a variety of volatile chemicals, including inter- and intraspecies semiochemicals and host odorants [9,10]. Insect ORs are seven-transmembrane domain G-protein-coupled receptors (GPCRs) and have an inverted membrane topology with an intracellular N-terminus and an extracellular C-terminus compared with vertebrate ORs [11–13]. Insect ORs usually form heteromeric complexes with a well-conserved odorant receptor co-receptor (Orco), which forms an odorant-gated ion channel with tuning ORs [14–16].

The first members of the OR family were discovered two decades ago [17–19]. Since then, insect ORs have been identified in many species, including *Tribolium castaneum* [20], *Locusta migratoria* [21], *Drosophila melanogaster* [22], *Bombyx mori* [23], *Apolygus lucorum* [24], *Microplitis mediator* [25], and *Manduca sexta* [26], with a high degree of divergence. The number of ORs is considerably varied between species, for example, there are 259 ORs in *T. castaneum* [20], 110 ORs in *A. lucorum* [24] and 79 ORs in *Acyrtosiphon pisum* [27].

In the present study, we identified 33 ORs genes from the antennae, head, legs and body transcriptomes of *M. persicae* combined with the data from the *M. persicae* genome ([https://bipaa.genouest.org/is/aphidbase/myzus\\_persicae/](https://bipaa.genouest.org/is/aphidbase/myzus_persicae/), accessed on 13 May 2020) [27], and phylogenetically compared them with other insect ORs. The expression profiles of the ORs among different tissues were investigated using real-time quantitative PCR (RT-qPCR). In addition, the genomic structure of *M. persicae* ORs was analyzed. This work provides basic information that could be useful for the functional clarification of these odorant genes at the molecular level.

## 2. Materials and Methods

### 2.1. Insect Rearing and Sample Collection

A *M. persicae* colony was collected from infested Chinese cabbage (*Brassica rapa* L. ssp. *pekinensis*) leaves at the Langfang Experimental Station of the Chinese Academy of Agricultural Sciences, Hebei Province, China and established in the laboratory under the following conditions: 18–24 °C, 65–75% relative humidity, and a 16 h:8 h light:dark photoperiod. The colony was reared on tobacco (*Nicotiana tabacum* L.) seedlings in environmental chambers. For the tissue expression profile studies, around 2000 apterous adult aphids were dissected on ice under anatomical microscope (Olympus Corporation, Tokyo, Japan). Four different tissues (antennae, head without antennae, legs and body without head and legs) were collected separately in 1.5 mL centrifuge tubes and immediately frozen in liquid nitrogen and stored at –80 °C until use. Samples of each tissue were collected three times in the same way as three biological replicates.

## 2.2. RNA Extraction, cDNA Library Construction and Illumina Sequencing

Each tissue sample (50 mg) was used for total RNA extraction using Trizol reagent (Life Technologies, Carlsbad, CA, USA) following the manufacturer's protocol. The concentration of the total RNA samples was measured on a NanoDrop 2000 spectrophotometer (Thermo Fisher Scientific, Waltham, MA, USA) and the RNA integrity was further confirmed using a 2100 Bioanalyzer (Agilent Technologies, CA, USA). The cDNA library of each sample was constructed using NEB Next Ultra RNA Library Prep Kit for Illumina (New England Biolabs, Ipswich, MA, USA) according to the manufacturer's protocol. Briefly, messenger RNAs were isolated from 10 µg of total RNA of each tissue using oligo (dT) magnetic beads and fragmented into short nucleotides using the fragmentation buffer supplied with the kit at 94 °C for 5 min. Each cleaved mRNA was transcribed to first strand cDNA using random hexamer primer and M-MuLV Reverse Transcriptase (RnaseH). The second strand cDNA was subsequently synthesized using DNA polymerase I, dNTPs and RNase H. After the end-repair, dA-tailing and adaptor ligation, the products were amplified by PCR and purified using the QIA quick PCR purification kit (Qiagen, Valencia, CA, USA) to create the final sequencing library. The cDNA libraries were pair-end sequenced using a PE150 strategy on an Illumina HiSeq2500 platform (Illumina, San Diego, CA, USA).

## 2.3. Bioinformatics Analysis

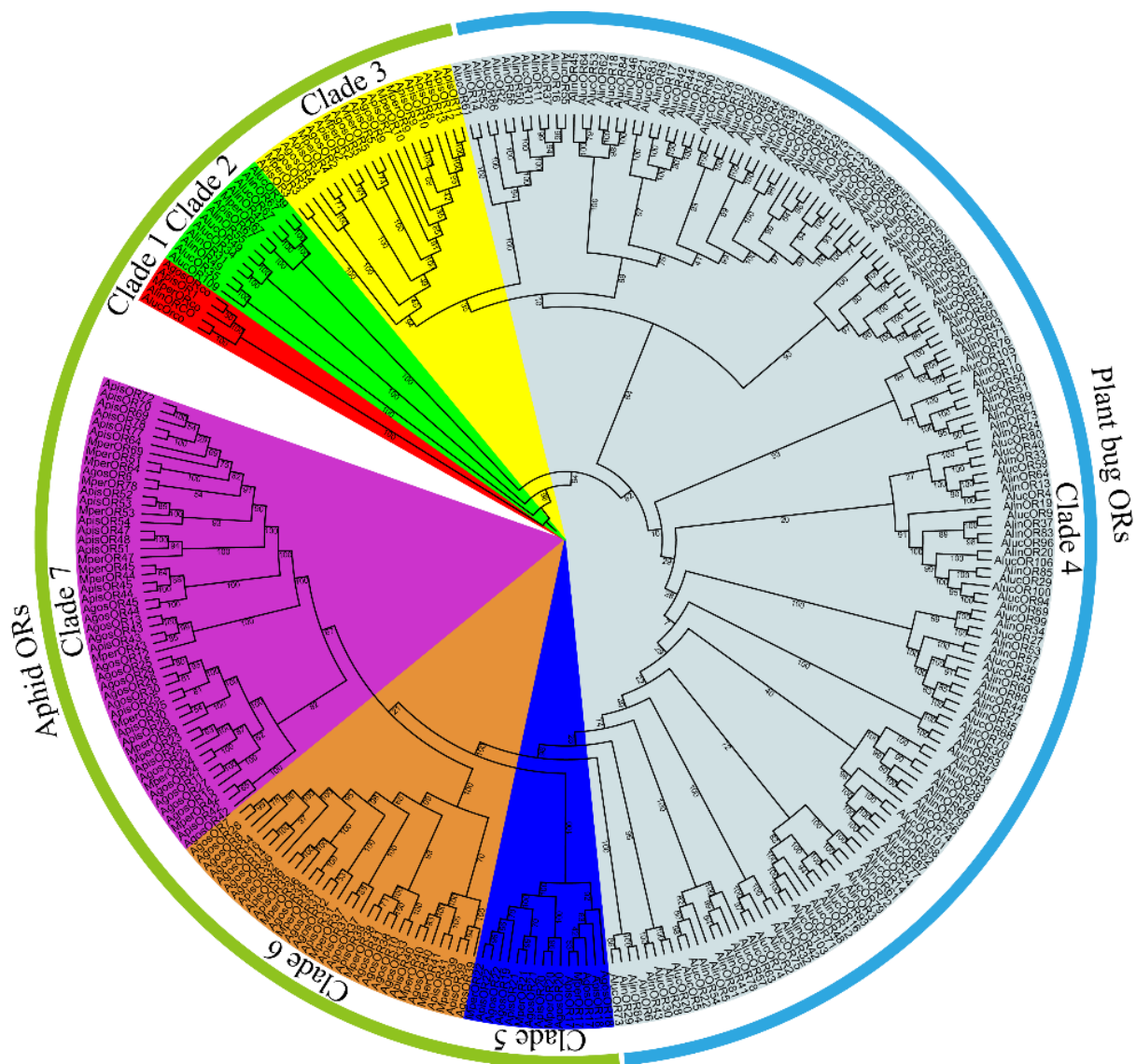
The raw reads were trimmed with Trimmomatic software (version 0.32). Clean reads were de novo assembled into unigenes with the Trinity short read assembler (version 20121005) using default parameters [28]. To annotate the unigenes, we performed a BLASTx search against Nr and Nt databases at the National Center for Biotechnology Information (NCBI) with an *E*-value cut-off of  $1.0 \times 10^{-5}$  [29]. Gene names were assigned to each unigene based on the best BLASTx hit with the highest score value. We use the BWA-MEM alignment algorithm [30] and HTSEQ program (version 0.6.1) [31] to align the RNA reads with the assembly and count the read numbers mapped to each unigene. Reads per kilobase per million (RPKM) were calculated for the assembled transcripts based on their mapping data according to the formula published by Mortazavi [32]. Thus, the RPKM of each unigene was calculated based on the length of the gene and read count mapped to this gene.

## 2.4. Identification of *M. persicae* ORs by Bioinformatics

The previously described ORs sequences in the *M. persicae* clone G006 assembly v2 were downloaded from AphidBase at the Bioinformatics Platform for Agroeco system Arthropods ([https://bipaa.genouest.org/is/aphidbase/myzus\\_persicae/](https://bipaa.genouest.org/is/aphidbase/myzus_persicae/), accessed on 13 May 2020). Genomic scaffold sequences found by tblastn (*E*-value= $e^{-100}$ ) were used to construct putative OR sequences manually using Sequencher v4.5 (Gene Codes, Inc., Ann Arbor, MI, USA) and refined using SplicePredictor (<http://deepc2.psi.iastate.edu/cgi-bin/sp.Cgi/>, accessed on 17 June 2021). All *M. persicae* ORs identified in this manner were in turn used in successive tblastn searches to identify other candidate sequences. The green peach aphid OR genes were named "*M. persicae* ORx" with a number consistent with the pea aphid ORs [27]. Fragments with high similarity with MperORs were filtered as redundant sequences. Sequences possessing only one or two recognizable exons and encoding proteins less than 200 amino acids were discarded. Because of high divergence and/or discontinuity among scaffolds, not all detected OR genes could be entirely annotated. In these cases, amino acid sequences shorter than 200 amino acids were discarded as probable gene fragments. All genes reported in this study are full length OR genes. The *M. persicae* ORs were further confirmed by their transmembrane-domain structures using TMHMM [33] server version 2.0 (<http://www.cbs.dtu.dk/services/TMHMM>, accessed on 18 June 2021). The nucleotide sequences of identified MperORs have been deposited in GenBank under the accession numbers (OM628748-OM628780).

### 2.5. Phylogenetic Analysis

Amino acid sequences of candidate *M. persicae* ORs were aligned by ClustalX 2.0 using default settings with the OR sequences from *T. castaneum* [20], *L. migratoria* [21], *D. melanogaster* [22], *Bombyx mori* [23], *A. lucorum* [24] and *M. mediator* [25] (Figure S1). Phylogenetic trees were constructed based on the maximum likelihood by RAXML version 8 [34] with LG substitution matrix selected by the PROTTEST 3 program [35]. Node support was assessed using a bootstrap procedure of 1000 replicates. The phylogenetic tree of 298 ORs from three different aphid species (*M. persicae*, *A. gossypii* and *A. pisum*), and two plant bug species (*A. lucorum* and *Adelphocoris lineolatus*) [36] (Figure 1) was constructed using the same approach. The OR names and sequences that were used in phylogenetic analyses are listed in the Supplementary Materials, Table S3 and S4.



**Figure 1.** Phylogenetic tree of 298 ORs from three different aphid species (*M. persicae*, *A. gossypii* and *A. pisum*), and two plant bug species (*A. lucorum* and *A. lineolatus*). The protein names and sequences that were used in the phylogenetic analysis are listed in Supplementary Materials, Table S4.

### 2.6. Verification of OR Sequences by Cloning and Sequencing

The open reading frames (ORFs) of each identified OR sequence was predicted by Open Reading Frame Finder (<https://www.ncbi.nlm.nih.gov/orffinder/>, accessed on 16 June 2020). Then gene-specific primers (Table S1) were designed using Primer Premier



5.0 software to clone the ORF of each MperOR gene. The template cDNA was synthesized from antennal RNA using the Fast Quant RT Kit (TianGen, Beijing, China) according to the manual. PCR reactions were carried out with 200 ng antennal cDNAs with 0.5 units of Ex Taq DNA Polymerase (TaKaRa, Dalian, China). The PCR amplification conditions were set as 95 °C for 2 min, followed by 36 cycles of 94 °C for 45 s, 56 °C for 1 min, 72 °C for 2 min, and a final extension at 72 °C for 10 min. The PCR products were gel-purified and subcloned into the pGEM-T Easy vector (Promega, Madison, WI, USA), and the insert was sequenced using an ABI3730XL automated sequencer (Applied Biosystems, Carlsbad, CA, USA.) with standard M13 primers.

### 2.7. Genomic Structure Analysis of *M. persicae* ORs

The genomic DNA sequences of *M. persicae* OR genes were extracted with BLASTn using the mRNA sequences of *M. persicae* ORs as 'query'. The mRNA-to-genomic DNA alignment of each OR gene was analyzed using the Splign online program (<http://www.aphidbase.com/aphidbase/downloads>, accessed on 21 May 2021). The predicted structure of each *M. persicae* OR gene was reconstructed and recorded as described above.

### 2.8. Expression Profiles of *M. persicae* ORs in Different Tissues

RT-qPCR was performed to examine the expression of *M. persicae* ORs genes. The cDNA of antennae, heads, legs and the body parts were synthesized using the PrimeScript RT Reagent with gDNA Eraser (TaKaRa, Dalian, China). An equal amount of cDNA (150 ng) was used as the RT-qPCR templates. RT-qPCR was carried out on an ABI 7500 Real-Time PCR system (Applied Biosystems, Carlsbad, CA, USA). The  $\beta$ -actin (GenBank Acc. XM\_022309797) and GAPDH (GenBank Acc. XM\_022315441) genes were used as reference genes for normalizing the expression of target genes and correcting the sample-to-sample variations (Figure S1). The primers used for the RT-qPCR were designed using Beacon Designer 7.90 (PREMIER Biosoft International) (Table S2).

To confirm the uniformity of the amplification efficiencies of the target and reference genes, a pilot experiment was conducted to examine the variation of  $\Delta$ Ct (Ct, Target–Ct,  $\beta$ -actin/GAPDH) with template dilution. Briefly, we used five serial dilutions of cDNA from each sample. For each dilution, amplifications were performed in triplicate using primers for the target gene and  $\beta$ -actin or GAPDH. The mean Ct was determined for both target gene and  $\beta$ -actin or GAPDH,  $\Delta$ Ct (the difference between Ct of target gene and Ct of reference gene) was calculated, and log cDNA dilution vs.  $\Delta$ Ct was plotted. The LinRegPCR program (version 11.0) [37] was used to calculate the RT-qPCR amplification efficiencies of the target and reference genes (Figure S2).

Each RT-qPCR reaction was conducted in a 25  $\mu$ L mixture containing 12.5  $\mu$ L of SuperReal PreMix Plus (TianGen, Beijing, China), 0.75  $\mu$ L of each primer (10  $\mu$ M), 1  $\mu$ L of sample cDNA (150 ng/ $\mu$ L), 0.5  $\mu$ L of Rox Reference Dye and 9.5  $\mu$ L of sterilized water. The RT-qPCR cycling parameters were 95 °C for 15 min, followed by 40 cycles of 95 °C for 10 s and 60 °C for 32 s. Then, the PCR products were heated to 95 °C for 15 s, cooled to 60 °C for 1 min, heated again to 95 °C for 30 s and cooled to 60 °C for 15 s to measure the melt curves. The negative control, which used only sterilized water, was included in each experiment. To ensure reproducibility, each RT-qPCR reaction for each sample was performed in three technical replicates and three biological replicates. The expression level of each *M. persicae* OR genes relative to  $\beta$ -actin and GAPDH were calculated by using the comparative  $2^{-\Delta\Delta$ Ct method [38]. The comparative analyses of each target gene among various tissues were determined using a one-way nested analysis of variance, followed by Tukey's honest significance difference test using the SPSS statistics 18.0 software (SPSS Inc., Chicago, IL, USA).

### 3. Results

#### 3.1. Identification of *M. persicae* ORs

We identified 33 putative OR genes (32 typical ORs and one atypical co-receptor Orco) from the transcriptome and genome of *M. persicae* (Table 1). All *M. persicae* ORs are full-length with open reading frames (ORFs) ranging from 1104 to 1626 bp and a mean length of 1253 bp. The completeness of the MperORs' length was judged by amino acid alignment with ApisORs. All MperORs showed the presence of predicted multiple transmembrane domains with an inside N-terminus, as usually observed for insect ORs. The TMHMM prediction showed that 7 MperORs (MperORco, MperOR10, MperOR21, MperOR23, MperOR37, MperOR42 and MperOR47) had seven-transmembrane domains (Table 1). The average identity of MperOR-Demlor pairs is 12%. However, comparing the MperORs with ORs from closely related species, *A. pisum* and *A. gossypii*, demonstrates a not-so divergent relationship as the average identity of MperOR-ApisOR pairs is 78% and MperOR-AgosOR pairs is 64%. Furthermore, MperORs are very closely related and have an amino acid identity of 43–96% similar to ApisORs (Table 1, Table S5–S8).

**Table 1.** List of OR genes in *M. persicae*.

Gene Name	ORF (bp)	Amino Acids	Scaffold *	Antennae RPKM	Head RPKM	Legs RPKM	Body RPKM	Homology Search with <i>A. Pisum</i>		
								Most Similar	%Identity	TM Domains
MperORco	1494	498	173 (307,518–314,584)	16.04	1.32	0.15	1.05	ApisOR1	96	7
MperOR2	1203	401	200 (329,352–332,846)	4.19	0.14	0.07	0.14	ApisOR2	93	5
MperOR3	1206	402	27 (538,815–544,733)	2.76	0.14	0.04	0.08	ApisOR3	94	8
MperOR4	1167	389	2 (1,669,537–1,680,404)	0.66	0.25	0.08	0.12	ApisOR4	94	6
MperOR5	1104	368	185 (390,783–396,720)	1.28	0.08	0.00	0.02	ApisOR5	87	8
MperOR9	1107	369	52 (795,957–799,660)	1.23	0.03	0.03	0.03	ApisOR9	82	6
MperOR10	1122	374	448 (232,245–234,867)	1.23	0.03	0.03	0.03	ApisOR10	68	7
MperOR17	1296	432	54 (655,522–659,088)	1.36	4.15	1.44	5.60	ApisOR17	84	6
MperOR20	1266	422	27 (295,168–299,537)	3.45	0.14	0.62	0.32	ApisOR20	89	6
MperOR21	1272	424	27 (273,498–275,759)	1.39	0.20	0.05	0.05	ApisOR21	84	7
MperOR22	1290	430	27 (286,505–290,732)	0.09	0.00	0.00	0.00	ApisOR22	80	6
MperOR23	1248	416	128 (606,788–609,435)	0.42	0.00	0.00	0.02	ApisOR23	90	7
MperOR24	1224	408	128 (619,838–621,712)	1.49	5.02	1.90	4.38	ApisOR23	43	4
MperOR25	1203	401	1011 (10,792–14,561)	1.30	0.01	0.01	0.05	ApisOR25	80	5
MperOR29	1242	414	495 (178,577–182,414)	0.21	0.06	0.00	0.01	ApisOR29	72	6
MperOR35	1221	407	139 (21,194–26,450)	1.32	0.01	0.01	0.08	ApisOR35	59	6
MperOR36	1221	407	0 (1,027,879–1,031,452)	0.19	0.02	0.00	0.00	ApisOR36	75	6
MperOR37	1251	417	533 (71,681–75,764)	1.95	0.01	0.05	0.04	ApisOR37	88	7
MperOR38	1221	407	16 (1,217,873–1,219,749)	0.26	0.03	0.07	0.05	ApisOR38	84	6
MperOR39	1188	396	257 (222,543–224,270)	0.12	0.00	0.02	0.03	ApisOR39	93	5
MperOR40	1266	422	268 (227,342–230,883)	0.53	0.00	0.01	0.00	ApisOR40	86	6
MperOR41	1248	416	268 (236,969–239,632)	0.61	0.04	0.18	0.28	ApisOR41	84	6
MperOR42	1290	430	129 (610,966–615,169)	0.12	2.27	0.21	0.65	ApisOR42	75	7
MperOR43	1254	418	138 (86,064–89,109)	1.32	0.15	0.01	0.08	ApisOR43	92	8
MperOR44	1140	380	115 (575,915–582,437)	0.10	0.05	0.00	0.18	ApisOR44	59	8
MperOR45	1296	432	13 (1,040,637–1,046,400)	0.50	0.22	0.05	0.64	ApisOR45	58	5
MperOR47	1302	434	109 (556,358–561,734)	0.08	0.00	0.05	0.13	ApisOR47	73	7
MperOR51	1293	431	1192 (10,158–17,868)	0.53	0.06	0.00	0.03	ApisOR70	70	6
MperOR53	1281	427	700 (104,988–111,553)	0.84	0.04	0.00	0.02	ApisOR53	70	6
MperOR64	1290	430	380 (259,518–265,853)	0.53	0.06	0.00	0.03	ApisOR64	71	5
MperOR67	1239	413	354 (145,937–151,752)	5.87	0.06	0.05	0.09	ApisOR67	81	6
MperOR69	1626	542	304 (223,678–231,005)	0.36	0.04	0.06	0.02	ApisOR69	59	5
MperOR78	1290	430	140 (453,957–457,387)	0.36	0.05	0.03	0.08	ApisOR78	62	6

\* AphidBase scaffold ID (start . . . stop nt, orientation).

#### 3.2. Phylogenetic Analyses of MperORs

The phylogenetic tree was built using 606 OR sequences from seven different insect species (*M. persicae*, *A. lucorum*, *B. mori*, *D. melanogaster*, *L. migratoria*, *M. mediator* and *T. castaneum*) in six different orders including Hemiptera, Lepidoptera, Hymenoptera, Diptera and Coleoptera (Figure S3). As illustrated in the phylogenetic tree, ORs are highly divergent between species and most ORs from the same species formed monophyletic

groups (Figure S3). In the phylogenetic tree, 25 MperORs (OR17, OR20–25, OR29, OR35–45, OR47, OR51, OR53, OR64, OR69 and OR78) were clustered in a species-specific subgroup. In addition, 4 MperORs (OR4, OR5, OR9 and OR10) were grouped with TcasOR187 and MmedOR53, and MperOR2 and MperOR3 were most closely related to the LmigORs superfamily. Analogously, MperOR67 was clustered with two AlucORs and most closely related to the MmedORs superfamily. The extremely conserved Orco subfamily in the seven species was clustered in one branch with clear orthologous relationships.

The phylogenetic tree of Hemiptera ORs was built using 298 OR sequences from three different aphid species (*M. persicae*, *A. gossypii* and *A. pisum*), and two plant bug species (*A. lucorum* and *A. lineolatus*) (Figure 1). All aphid ORs are clustered to six homologous clades (lineages), and clearly separated from ORs from plant bugs. Interestingly, the OR67s from the aphid and plant bugs are clustered into one same clade (Clade 2). The Orco sequences of the aphid and plant bugs are highly conserved and clustered into one clade (Clade 1) with a bootstrap value of 100 (Figure 1).

### 3.3. Genomic Structure of *M. persicae* ORs

The genomic structures and the splice site of the intron–exon junctions of MperOR genes were analyzed based on the *M. persicae* genome annotations and the GT–AG rules [39]. The results revealed that the size of the genomic sequences of MperOR genes ranges from 1.73 to 10.87 kb with an average length of 4.57 kb. Twenty-three OR genes (*MperOR4*, *MperOR17*, *MperOR20–23*, *MperOR25*, *MperOR29*, *MperOR35–37*, *MperOR40–45*, *MperOR47*, *MperOR51*, *MperOR53*, *MperOR64*, *MperOR69*, and *MperOR78*) have 5 introns and 6 exons, and 6 OR genes (*MperOR5*, *MperOR9*, *MperOR10*, *MperOR24*, *MperOR38* and *MperOR39*) have 4 introns and 5 exons. The remaining 4 OR genes (*MperOR2*, *MperOR3*, *MperOR67* and *MperORco*) have 1, 3, 6 and 8 introns and 2, 4, 7 and 9 exons, respectively (Table 2, Figure 2). The intron lengths of each *MperOR* are variable. The average intron length is 3.32 kb, the longest intron length is 9.7 kb for *MperOR4*, and the shortest is 0.54 kb for *MperOR39* (Table 2).

**Table 2.** The introns and exons of OR genes in *M. persicae*.

Gene Name	Genomic DNA Size (bp) and Strand (+/–)	No. of Intron	Total Length of Introns (bp)	Average Intron Size (bp)	No. of Exon	Total Length of Exons (bp)	Average Exon Size (bp)
MperORco	7066(+)	8	5572	697	9	1494	166
MperOR2	3494(–)	1	2291	2291	2	1203	602
MperOR3	5918(+)	3	4712	1571	4	1206	302
MperOR4	10,867(+)	5	9700	1940	6	1167	195
MperOR5	5937(+)	4	4833	1208	5	1104	221
MperOR9	3703(+)	4	2596	649	5	1107	221
MperOR10	2622(+)	4	1500	375	5	1122	224
MperOR17	3566(–)	5	2270	454	6	1296	216
MperOR20	4369(–)	5	3103	621	6	1266	211
MperOR21	2261(–)	5	989	198	6	1272	212
MperOR22	4227(–)	5	2937	587	6	1290	215
MperOR23	2647(–)	5	1399	280	6	1248	208
MperOR24	1874(–)	4	650	163	5	1224	245
MperOR25	3769(+)	5	2566	513	6	1203	201
MperOR29	3837(+)	5	2595	519	6	1242	207
MperOR35	5256(–)	5	4035	807	6	1221	204
MperOR36	3573(–)	5	2352	470	6	1221	204
MperOR37	4083(+)	5	2832	566	6	1251	209
MperOR38	1876(+)	4	655	164	5	1221	244
MperOR39	1727(+)	4	539	135	5	1188	238
MperOR40	3541(–)	5	2275	455	6	1266	211
MperOR41	2663(+)	5	1415	283	6	1248	208

Table 2. Cont.

Gene Name	Genomic DNA Size (bp) and Strand (+/−)	No. of Intron	Total Length of Introns (bp)	Average Intron Size (bp)	No. of Exon	Total Length of Exons (bp)	Average Exon Size (bp)
MperOR42	4203(+)	5	2913	583	6	1290	215
MperOR43	3045(+)	5	1791	358	6	1254	209
MperOR44	6522(−)	5	5382	1076	6	1140	190
MperOR45	5763(+)	5	4467	893	6	1296	216
MperOR47	5376(+)	5	4074	815	6	1302	217
MperOR51	7710(−)	5	6417	1283	6	1293	216
MperOR53	6565(+)	5	5284	1057	6	1281	214
MperOR64	6335(+)	5	5045	1009	6	1290	215
MperOR67	5815(−)	6	4576	763	7	1239	177
MperOR69	7327(+)	5	5701	1140	6	1626	271
MperOR78	3430(+)	5	2140	428	6	1290	215

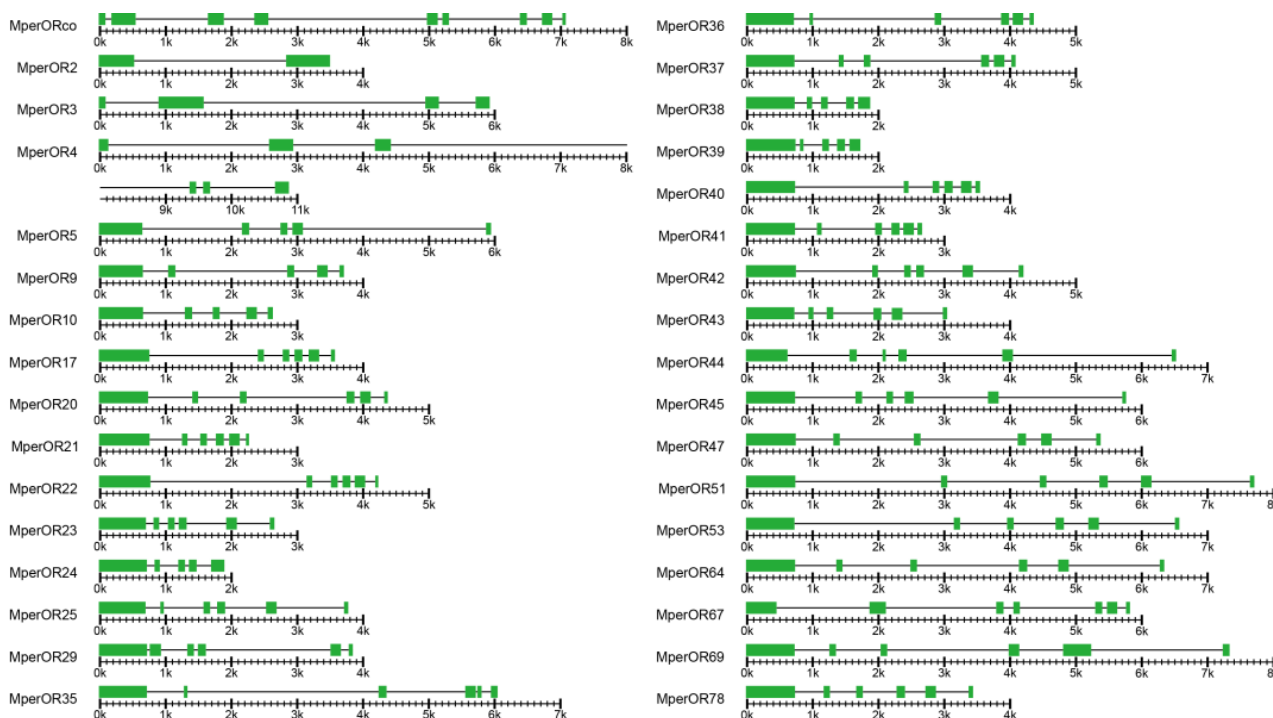


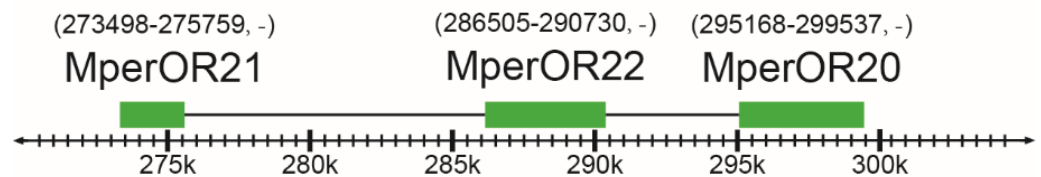
Figure 2. Genomic structure of *M. persicae* OR genes. The green rectangles and hairlines between two green rectangles represent the exons and introns, respectively. The length has been shown to scale with a scale bar under each *M. persicae* OR gene; every minor mark represents 100 bp.

We found that *MperORco* has 9 exons, 7 of which are relatively similar in size with a length of 100–241 bp. *MperOR67* has 7 exons. It is consistent with the clustering in the phylogenetic tree that *MperORco* and *MperOR67* are separately clustered. In addition, *MperOR2* and *MperOR3* have only 3 and 4 exons. *MperOR4*, *MperOR5*, *MperOR9* and *MperOR10* have 5–6 exons with relatively similar length. Again, it is consistent that *MperOR2* and *MperOR3* are clustered together in the same branch, and *MperOR4*, *MperOR5*, *MperOR9* and *MperOR10* are clustered together in another branch. The number and length of the remaining 25 OR exons are very conservative. However, for *MperOR44*, the length of the first exon is 709–781 bp, and the length of the fourth exon is 117–159 bp. The average length of the first, second, third and fourth exon is 742 bp, 84 bp, 100 bp and 123 bp, respectively. In addition, the length of the fifth and sixth exon of *MperOR69* is 426 bp and 102 bp, which is just opposite to *MperOR24*, *MperOR38* and *MperOR39*.

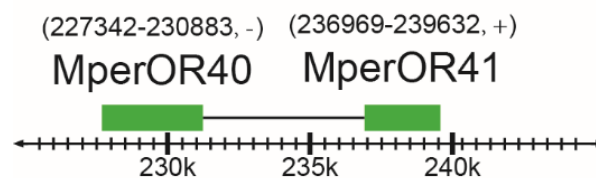


We also analyzed the genomic clustering of MperORs. The results indicated that the 28 MperORs are distributed in 33 different scaffolds (Table 1). *MperOR20*, *MperOR21* and *MperOR22* as well as *MperOR40* and *MperOR41* constitute two gene clusters on scaffold 27 and scaffold 268, respectively (Figure 3). The distances between *MperOR21* and *MperOR22*, *MperOR20* and *MperOR22*, *MperOR40* and *MperOR41* are 10.74 kb, 4.44 kb, and 6.09 kb, respectively (Figure 3).

## Scaffold 27



## Scaffold 268



**Figure 3.** The tandem arrays of OR genes in *M. persicae*. The genomic sequences of *M. persicae* were downloaded from the AphidBase at the BioInformatics Platform for Agroecosystem Arthropods (<https://bipaa.genouest.org/is/>, accessed on 13 May 2020).

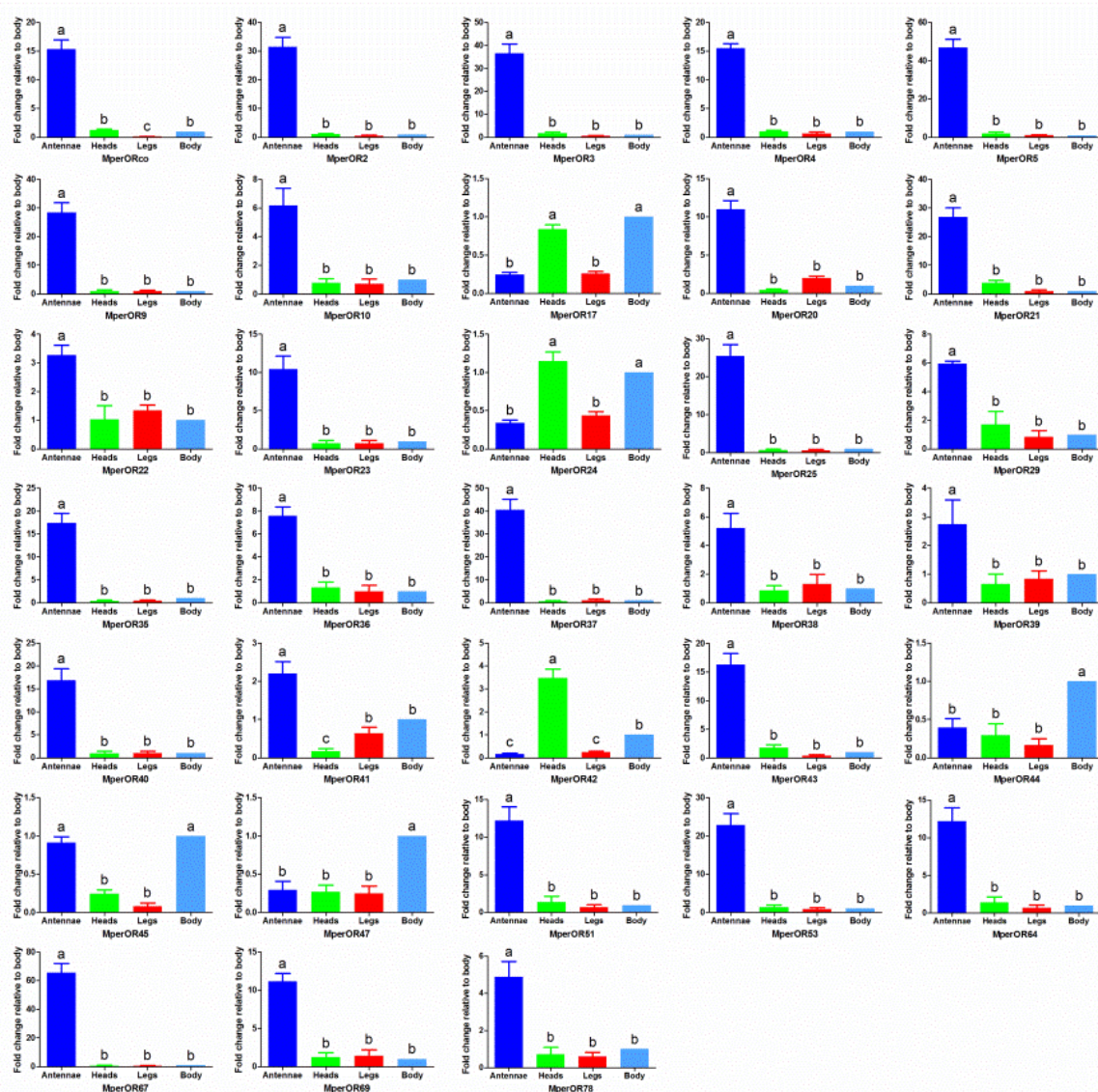
### 3.4. Expression Profiles of *M. persicae* ORs in Different Tissues

We explored the expression profiles of MperORs in different tissues using RT-qPCR. For diluted templates, the absolute values of the slope of all lines (log cDNA dilution vs.  $\Delta\text{Ct}$ ) were  $<0.1$ , and the real PCR amplification efficiencies of the target and reference genes were more than 1.90 (calculated using the LinRegPCR program and listed in the Supplementary Materials). Therefore, the efficiencies of the target and reference genes were similar in our analysis, and the  $2^{-\Delta\Delta\text{Ct}}$  calculation method can be used for the relative quantification.

The OR transcripts displayed tissue-specific expression patterns (Figure 4). The results showed that the transcripts of *MperOR2*, *MperOR3*, *MperOR4*, *MperOR5*, *MperOR9*, *MperOR10*, *MperOR20*, *MperOR21*, *MperOR22*, *MperOR23*, *MperOR25*, *MperOR29*, *MperOR35*, *MperOR36*, *MperOR37*, *MperOR38*, *MperOR39*, *MperOR40*, *MperOR41*, *MperOR43*, *MperOR51*, *MperOR53*, *MperOR64*, *MperOR67*, *MperOR69*, *MperOR78*, and *MperORco* were highly expressed in the antennae. In contrast, *MperOR42* and *MperOR44* were highly expressed in the head and the body, respectively. However, the remaining ORs showed a wide range of expression patterns. Among them, *MperOR45* was predominantly expressed in the body and the antennae, while *MperOR17* and *MperOR24* were mainly abundant in both the head and body.

Along with RT-qPCR measurements, the RPKM analysis also demonstrated the relatively high abundance of MperORs in the *M. persicae* transcriptome (Table 1). The RPKM and RT-qPCR results are in the analogical trend. The RPKM value showed that 27 ORs were highly expressed in the antennae, which was consistent with the trend of the RT-qPCR, although the specific values were slightly different. For example, the RPKM values of *MperOR17*, *MperOR24*, *MperOR42* and *MperOR45* indicated they are highly expressed in other parts of aphids, and consistent with the results of RT-qPCR. It should be noted that

although the RPKM results of *MperOR44* and *MperOR47* and RT-qPCR results show they are highly expressed in the body, the RPKM values are not significant.



**Figure 4.** The OR transcript levels of *M. persicae* in different tissues assessed by RT-qPCR. The error bars present the standard error, and the different letters (a, b, c) indicate significant differences ( $p < 0.05$ ). This figure was presented using *GAPDH* as the reference gene to normalize the target gene expression and correct sample-to-sample variation; similar results were also obtained with  *$\beta$ -actin* as the reference gene.

#### 4. Discussion

With the increase in insect genome and transcriptome sequencing projects, particularly in recent years, very large numbers of ORs have been identified in different insect species. In the present study, we profiled the transcriptomes in the antennal, head, body and legs of *M. persicae* using RNA-seq technology and annotated 33 full-length OR genes. Previous research revealed that OR genes are highly divergent in numbers and sequence among different insect species due to the requirement for each species to detect a panoply of odor signals that relate to their specific life histories. The number of *M. persicae* ORs identified here is less than the number of ORs identified in other sucking insects, for example, 79 in *A. pisum*, of which 48 OR genes encode apparently full-length, intact and potentially functional ORs [27]. Similarly, in *A. gossypii*, 45 ORs have been identified including

22 ORs full-length ORs [40]. Furthermore, MperORs are very closely related and similar to ApisORs, suggesting that aphid OR genes may be evolved from common ancestors without subsequent duplication. This suggests that they are probably homologous genes and may have similar function.

In the same species, different ORs play a role in recognizing different odorants. In the *M. persicae*, the average identity of 33 MperORs is 38%. On the other hand, compared with other insects such as *D. melanogaster*, OR genes of the *M. persicae* indeed display a high degree of divergence with DmelORs. This implies that ORs have great differences among species, although they may be homologous in functional classification. Specifically, most of the identified MperORs are highly expressed in the antennae, suggesting that they may be involved in the green peach aphid's recognition of different host odors or pheromones. Furthermore, the number of MperORs identified here is less than the number of ORs identified in *D. melanogaster* (58 ORs) [22], *B. mori* (54 ORs) [23], *M. mediator* (54 ORs) [25] and *T. castaneum* (177 ORs) [20], indicating the species uses a more narrow range of odors to detect hosts and/or conspecifics, or alternatively OR-mediated olfaction is less central to its life cycle, or that olfaction is less important than other senses. The phylogenetic analyses clearly divided most of aphid ORs into species-specific homology clades. However, OR67s of *M. persicae* and *A. pisum* are clustered with OR67s of *A. lucorum* and *A. lineolatus* with a bootstrap support of 100 (Figure 1), and with a relatively high protein identity (53%); a similar finding was also found in OBPs and CSPs between the aphid and plant bug [41], suggesting these olfactory genes may have a common function in these sucking insects.

By analyzing the genomic structure, we determined the number and length of introns and exons of 33 MperORs. The conserved genomic structure of 20 MperORs suggests that they have experienced recent gene duplication events, and the distribution of the clustering in the evolutionary tree analysis suggests that they may have similar functions.

To better understand the function of MperOR genes, tissue-specific expressions were evaluated by using RT-qPCR. MperOR genes exhibited diverse expression patterns, which could be briefly classified into five types, according to their level of expression: (1) antennae-enriched ORs, (2) head-enriched ORs, (3) body-enriched ORs, (4) head and body-enriched ORs, and (5) antennae and body-enriched ORs. It was also reported that some ORs could be expressed in a variety of tissues apart from the olfactory organs. Similarly, many ORs are also distributed in the head, especially near the mouthparts, thus OR42 could be possibly involved in the identification of odor substances at close distances [42]. As a co-receptor with ORs, *MperOrco* is highly expressed in the main olfactory organ—the antennae. The major alarm pheromone component (*E*)- $\beta$ -farnesene (*E* $\beta$ F) emitted from the cornicles of aphids [43] is recognized by large placoid sensillum neurons, which express ApisOR5 on the sixth antennal segment in *A. pisum* [44]. ApisOR5 and its orthologues in *M. persicae* (MperOR5) and *A. gossypii* (AgosOR5) show a high identity rate (77%), suggesting that their function is possibly conserved in aphids. Due to the parthenogenetic reproductive pattern of wingless aphids, the genes highly expressed in antennae, head and body may have been expressed before the birth of offspring aphids. This suggests that these genes may play a role in the embryonic development of aphids, or may have other unknown functions different from sensing smells. OR44 with a high body-specific expression may be related to the distribution of ORs at the end of the abdomen. The tissue expression profiles of each MperOR gene evaluated by the RPKM values are mostly consistent with the RT-qPCR results, but with several exceptions, this may be attributed to the different detection methods in transcriptome analysis and RT-qPCR. The RT-qPCR results can only represent the differential expression among the four tissue samples. The RPKM values, however, can compare all transcript expression levels among different tissues and each single transcript within one sample [45], for example, like most ORs, the RPKM values of all MperORs within four tissues showed most of them are enriched in the antennae. However, comparing the RPKM values in the antennae of all ORs can show that the expression depth of Orco in antennae is much higher than other OR genes, supporting the notion that Orco is widely expressed as OR-Orco complex [16].



## 5. Conclusions

In summary, we identified 33 candidate odorant receptors that may function in odorant perception in the green peach aphid, *M. persicae* from transcriptomes and genome data. By analyzing the phylogenetic relationships between other insects and the genomic structure of ORs, we speculated on the possible reasons for the evolution of OR in green peach aphid. As a crucial first step toward understanding their functions, a comprehensive examination of the expression patterns of these MperOR genes in different tissue samples was performed using RT-qPCR. We identified 27 MperOR genes that were highly expressed in antennae. The results of our study will provide a valuable foundation for further elucidating the mechanisms of olfaction in *M. persicae*, which could also help to use ORs as targets to regulate insect olfactory behavior and broaden the applications of available tools for effective control of pests.

**Supplementary Materials:** The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/insects13050398/s1>, Table S1: The gene-specific primers for MperORs cloning and sequencing. Table S2: Primers used in RT-qPCR for determination expression levels of MperORs. Table S3: The protein names and sequences of 606 ORs from 7 species of insects used in Figure S3. Table S4: The protein names and sequences of 298 ORs from 5 different species used in Figure 1. Table S5: A percent identity matrix of MperORs. Table S6: A percent identity matrix between MperORs and DmelORs. Table S7: A percent identity matrix between MperORs and AgosORs. Table S8: A percent identity matrix between MperORs and ApisORs. Figure S1: The stable expression of *M. persicae*  $\beta$ -actin and GAPDH in different tissues measured by RT-qPCR. Figure S2: The standard curves of MperORs with GAPDH as reference gene. Figure S3: Phylogenetic tree of 606 ORs from 7 species of insects.

**Author Contributions:** Conceptualization, J.L., S.G. and B.C.; methodology, J.X., A.K., S.G. and Y.Z.; software, J.L. and H.D.; formal analysis, S.G., H.D. and B.C.; data curation, J.L., J.X., J.Z. and A.K.; writing—original draft preparation, J.L. and S.G.; writing—review and editing, J.Z., B.C. and Y.Z.; visualization, J.L. and S.G.; supervision, S.G. and B.C.; funding acquisition, S.G., Y.Z. and B.C. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was funded by the National Natural Science Foundation of China (Grant No. 31772164), Open Project of State Key Laboratory for Biology of Plant Diseases and Insect Pests (Grant No. SKLOF202106).

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The data presented in this study are available in Supplementary Materials.

**Acknowledgments:** The authors are grateful to Rui Zhao and Huanhuan Wang for rearing the aphids.

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

## References

1. Weber, G. Ecological genetics of host plant exploitation in the green peach aphid, *Myzus persicae*. *Entomol. Exp. Appl.* **1986**, *40*, 161–168. [[CrossRef](#)]
2. Troncoso, A.J.; Vargas, R.R.; Tapia, D.H.; Olivares-Donoso, R.; Niemeyer, H.M. Host selection by the generalist aphid *Myzus persicae* (Hemiptera: Aphididae) and its subspecies specialized on tobacco, after being reared on the same host. *Bull. Entomol. Res.* **2005**, *95*, 23–28. [[CrossRef](#)] [[PubMed](#)]
3. Van Emden, H.F.; Eastop, V.F.; Hughes, D.P.; Way, M.J. The Ecology of *Myzus persicae*. *Annu. Rev. Entomol.* **1969**, *14*, 197–270. [[CrossRef](#)]
4. Pickett, J.A.; Glinwood, R.T. Chemical Ecology. In *Aphids as Crop Pests*; VanEmden, H.F., Harrington, R., Eds.; CAB International Press: Wallingford, UK, 2007; pp. 235–260.
5. Read, D.P.; Feeny, P.P.; Root, R.B. Habitat selection by the aphid parasite *Diaeretiella rapae* (Hymenoptera: Braconidae) and hyperparasite *Charips brassicae* (Hymenoptera: Cynipidae). *Can. Entomol.* **1970**, *102*, 1567–1578. [[CrossRef](#)]
6. Mathers, T.C.; Chen, Y.; Kaithakottil, G.; Legeai, F.; Mugford, S.T.; Baa-Puyoulet, P.; Bretaudeau, A.; Clavijo, B.; Colella, S.; Collin, O.; et al. Rapid transcriptional plasticity of duplicated gene clusters enables a clonally reproducing aphid to colonise diverse plant species. *Genome Biol.* **2017**, *18*, 1–20. [[CrossRef](#)]



7. Gadenne, C.; Barrozo, R.B.; Anton, S. Plasticity in insect olfaction: To smell or not to smell? *Annu. Rev. Entomol.* **2016**, *61*, 317–333. [[CrossRef](#)]
8. Haverkamp, A.; Hansson, B.S.; Knaden, M. Combinatorial codes and labeled lines: How insects use olfactory cues to find and judge food, mates, and oviposition sites in complex environments. *Front. Physiol.* **2018**, *9*, 49. [[CrossRef](#)]
9. Touhara, K.; Vosshall, L.B. Sensing odorants and pheromones with chemosensory receptors. *Annu. Rev. Physiol.* **2009**, *71*, 307–332. [[CrossRef](#)]
10. Dahanukar, A.; Hallem, E.A.; Carlson, J.R. Insect chemoreception. *Curr. Opin. Neurobiol.* **2005**, *15*, 423–430. [[CrossRef](#)]
11. Benton, R.; Sachse, S.; Michnick, S.W.; Vosshall, L.B. Atypical membrane topology and heteromeric function of *Drosophila* odorant receptors in vivo. *PLoS Biol.* **2006**, *4*, e20. [[CrossRef](#)]
12. Lundin, C.; Käll, L.; Kreher, S.A.; Kapp, K.; Sonnhammer, E.L.; Carlson, J.R.; von Heijne, G.; Nilsson, I. Membrane topology of the *Drosophila* OR83b odorant receptor. *FEBS Lett.* **2007**, *581*, 5601–5604. [[CrossRef](#)]
13. 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 transmembrane domain proteins that can signal independently of heterotrimeric G proteins. *Insect Biochem. Mol. Biol.* **2008**, *38*, 770–780. [[CrossRef](#)] [[PubMed](#)]
14. Sato, K.; Pellegrino, M.; Nakagawa, T.; Nakagawa, T.; Vosshall, L.B.; Touhara, K. Insect olfactory receptors are heteromeric ligand-gated ion channels. *Nature* **2008**, *452*, 1002–1006. [[CrossRef](#)] [[PubMed](#)]
15. Wicher, D.; Schäfer, R.; Bauernfeind, R.; Stensmyr, M.C.; 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](#)] [[PubMed](#)]
16. Stengl, M.; Funk, N.W. The role of the coreceptor Orco in insect olfactory transduction. *J. Comp. Physiol. A* **2013**, *199*, 897–909. [[CrossRef](#)] [[PubMed](#)]
17. Vosshall, L.B.; Amrein, H.; Morozov, P.; Rzhetsky, A.; Axel, R. A spatial map of olfactory receptor expression in the *Drosophila* antenna. *Cell* **1999**, *96*, 725–736. [[CrossRef](#)]
18. Gao, Q.; Chess, A. Identification of Candidate *Drosophila* Olfactory Receptors from Genomic DNA Sequence. *Genomics* **1999**, *60*, 31–39. [[CrossRef](#)]
19. Clyne, P.J.; Warr, C.G.; Freeman, M.R.; Lessing, D.; Kim, J.; Carlson, J.R. A novel family of divergent seven-transmembrane proteins: Candidate odorant receptors in *Drosophila*. *Neuron* **1999**, *22*, 327–338. [[CrossRef](#)]
20. Richards, S.; Gibbs, R.A.; Weinstock, G.M.; Brown, S.J.; Denell, R.; Beeman, R.W.; Gibbs, R.; Bucher, G.; Friedrich, M.; Grimmelikhuijzen, C.J. The genome of the model beetle and pest *Tribolium castaneum*. *Nature* **2008**, *452*, 949–955. [[CrossRef](#)]
21. Wang, Z.; Yang, P.; Chen, D.; Jiang, F.; Li, Y.; Wang, X.; Kang, L. Identification and functional analysis of olfactory receptor family reveal unusual characteristics of the olfactory system in the *Migratory locust*. *Cell. Mol. Life Sci.* **2015**, *72*, 4429–4443. [[CrossRef](#)]
22. Matthews, B.B.; Dos Santos, G.; Crosby, M.A.; Emmert, D.B., St.; Pierre, S.E.; Gramates, L.S.; Zhou, P.; Schroeder, A.J.; Falls, K.; Strelets, V. Gene model annotations for *Drosophila melanogaster*: Impact of high-throughput data. *G3 Genes Genomes Genet.* **2015**, *5*, 1721–1736. [[CrossRef](#)]
23. Tanaka, K.; Uda, Y.; Ono, Y.; Nakagawa, T.; Suwa, M.; Yamaoka, R.; Touhara, K. Highly selective tuning of a silkworm olfactory receptor to a key mulberry leaf volatile. *Curr. Biol.* **2009**, *19*, 881–890. [[CrossRef](#)] [[PubMed](#)]
24. 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-Dur). *Sci. Rep.* **2016**, *6*, 37870. [[CrossRef](#)] [[PubMed](#)]
25. Wang, S.-N.; Peng, Y.; Lu, Z.-Y.; Dhiloo, K.H.; Gu, S.-H.; Li, R.-J.; Zhou, J.-J.; Zhang, Y.-J.; Guo, Y.-Y. Identification and Expression Analysis of Putative Chemosensory Receptor Genes in *Microplitis mediator* by Antennal Transcriptome Screening. *Int. J. Biol. Sci.* **2015**, *11*, 737–751. [[CrossRef](#)] [[PubMed](#)]
26. Grosse-Wilde, E.; Kuebler, L.S.; Bucks, S.; Vogel, H.; Wicher, D.; Hansson, B.S. Antennal transcriptome of *Manduca sexta*. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 7449–7454. [[CrossRef](#)] [[PubMed](#)]
27. 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](#)]
28. Grabherr, M.G.; Haas, B.J.; Yassour, M.; Levin, J.Z.; Thompson, D.A.; Amit, I.; Adiconis, X.; Fan, L.; Raychowdhury, R.; Zeng, Q.D.; et al. Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat. Biotechnol.* **2011**, *29*, 644–652. [[CrossRef](#)] [[PubMed](#)]
29. Anderson, I.; Brass, A. Searching DNA databases for similarities to DNA sequences: When is a match significant? *Bioinformatics* **1998**, *14*, 349–356. [[CrossRef](#)]
30. Li, H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *arXiv* **2013**, arXiv:1303.3997v2. Available online: <https://arxiv.org/abs/1303.3997v2> (accessed on 26 May 2013).
31. Anders, S.; Pyl, P.T.; Huber, W. HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics* **2015**, *31*, 166–169. [[CrossRef](#)]
32. Mortazavi, A.; Williams, B.A.; McCue, K.; Schaeffer, L.; Wold, B. Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nat. Methods* **2008**, *5*, 621–628. [[CrossRef](#)]
33. Krogh, A.; Larsson, B.; von Heijne, G.; Sonnhammer, E.L.L. Predicting transmembrane protein topology with a hidden markov model: Application to complete genomes. *J. Mol. Biol.* **2001**, *305*, 567–580. [[CrossRef](#)]
34. Stamatakis, A. RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **2014**, *30*, 1312–1313. [[CrossRef](#)] [[PubMed](#)]

35. Darriba, D.; Taboada, G.L.; Doallo, R.; Posada, D. ProtTest 3: Fast selection of best-fit models of protein evolution. *Bioinformatics* **2011**, *27*, 1164–1165. [[CrossRef](#)] [[PubMed](#)]
36. Xiao, Y.; Sun, L.; Ma, X.Y.; Dong, K.; Liu, H.W.; Wang, Q.; Guo, Y.Y.; Liu, Z.W.; Zhang, Y.J. Identification and characterization of the distinct expression profiles of candidate chemosensory membrane proteins in the antennal transcriptome of *Adelphocoris lineolatus* (Goeze). *Insect Mol. Biol.* **2017**, *26*, 74–91. [[CrossRef](#)]
37. Livak, K.J.; Schmittgen, T.D. Analysis of Relative Gene Expression Data Using Real-Time Quantitative PCR and the  $2^{-\Delta\Delta Ct}$  Method. *Methods* **2001**, *25*, 402–408. [[CrossRef](#)] [[PubMed](#)]
38. Ramakers, C.; Ruijter, J.M.; Deprez, R.H.; Moorman, A.F. Assumption-free analysis of quantitative real-time polymerase chain reaction (PCR) data. *Neurosci. Lett.* **2003**, *339*, 62–66. [[CrossRef](#)]
39. Modrek, B.; Lee, C. A genomic view of alternative splicing. *Nat. Genet.* **2002**, *30*, 13–19. [[CrossRef](#)]
40. 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](#)]
41. Wang, Q.; Zhou, J.J.; Liu, J.T.; Huang, G.Z.; Xu, W.Y.; Zhang, Q.; Chen, J.L.; Zhang, Y.J.; Li, X.C.; Gu, S.H. Integrative transcriptomic and genomic analysis of odorant binding proteins and chemosensory proteins in aphids. *Insect Mol. Biol.* **2019**, *28*, 1–22. [[CrossRef](#)]
42. Kwon, H.W.; Lu, T.; Rutzler, M.; Zwiebel, L.J. Olfactory responses in a gustatory organ of the malaria vector mosquito *Anopheles gambiae*. *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 13526–13531. [[CrossRef](#)] [[PubMed](#)]
43. Edwards, L.J.; Siddall, J.B.; Dunham, L.L.; Uden, P.; Kislw, C.J. *Trans*- $\beta$ -farnesene, Alarm Pheromone of the Green Peach Aphid, *Myzus persicae* (Sulzer). *Nature* **1973**, *241*, 126–127. [[CrossRef](#)]
44. 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.* **2017**, *27*, 55–61. [[CrossRef](#)] [[PubMed](#)]
45. Gu, S.H.; Wu, K.M.; Guo, Y.Y.; Field, L.M.; Pickett, J.A.; Zhang, Y.J.; Zhou, J.J. Identification and Expression Profiling of Odorant Binding Proteins and Chemosensory Proteins between Two Wingless Morphs and a Winged Morph of the Cotton Aphid *Aphis gossypii* Glover. *PLoS ONE* **2013**, *8*, e73524. [[CrossRef](#)]