博士論文

Identification and targeted mutagenesis of odorant receptor genes in *Ostrinia furnacalis*
(アワノメイガにおける嗅覚受容体遺伝子の同定と突然変異体の作製)

楊　斌
(Yang Bin)

Table of Contents

General introduction……1

Chapter 1. Identification of candidate odorant receptor genes using Roche 454 GS-Jr……8

Introduction……9

Materials and Methods……10

Results……17

Discussion……28

Chapter 2. Identification of candidate chemosensory genes using Illumina MiSeq……30

Introduction……31

Materials and Methods……32

Results……42

Discussion……57

Chapter 3. Targeted mutagenesis of *OfurOR4* and *OfurOrco* (*OR2*)……61

Introduction……62

Materials and Methods……65

Results……80

Discussion……89

Chapter 4. Physiological and behavioral analysis of the *OfurOrco* (*OR2*) mutants……90

Introduction……91

Materials and Methods……92

Results……96

Discussion……99

General discussion……100

Acknowledgments……104

References……105

General introduction

1. Chemical signals affect insect behavior

Many insects utilize chemical signals in feeding, courtship, mating, and oviposition. For example, insects respond to volatile organic compounds (VOCs) released by plants. VOCs such as floral scents attract pollinators (Knudsen et al, 2006). Odorants from the flower of *Datura wrightii* attract females of *Manduca sexta* for oviposition (Reisenman et al, 2010). On the other hand, herbivore-induced VOCs that are released after the plants damaged by larval feeding, repel moths for oviposition (Reisenman et al, 2012). Furthermore, plants release complex blends of VOCs that attract natural enemies of the herbivores (Moraes et al, 1998; Turlings et al, 1998; Baldwin and Preston, 1999; Paré and Tumlinson, 1999; Dicke and van, 2000; Halitschke et al, 2000; Schnee et al, 2006). Insects also use chemical signals to discriminate safe food from potentially toxic food which contaminated by pathogens by sensing CO2, geosmin, and acid (Greg et al, 2004; Turner and Ray, 2009; Stensmyr et al, 2012; Joseph et al, 2009; Ai et al, 2010).

Insects release chemical signals by themselves as pheromones. Sex pheromones allow individuals to identify mating partners of the opposite gender. Usually, females release species-specific sex pheromones at extremely low concentration, and conspecific males have corresponding pheromone receptors to perceive the signal. Sex pheromones attract males at a long distance (Leal, 2005; Blomquist and Vogt, 2003). The first sex pheromone was identified in *Bombyx mori* in 1959 (Karlson and Butenandt, 1959). Sex pheromones of *B. mori* were secreted by the pheromone grand in the terminal of abdomen of mature females (Jurenka, 2003). Males

of some diurnal species such as butterflies release male pheromones (Andersson et al, 2007). In a fruit fly *Drosophila melanogaster*, the male-specific cis-vaccenyl acetate (cVA) inhibits male-male courtship and increases female receptivity (Kurtovic et al, 2007; Benton, 2007).

Other types of pheromones, such as aggregation pheromones (Sudhakar et al, 2007), trail pheromones (Ji et al, 2013), and alarm pheromones (Bushra and Tariq, 2014) are also involved in insect behavior. Interestingly, one alarm pheromone of ants was used as a host-finding cue by a parasitoid fly (Wittea et al, 2010).

In addition to pheromones, insects also release chemical signals as allelochamicals. In some Lepidopteran species, larvae release stinks to defense themselves from predators (Takagi et al, 1995).

2. Genes involved in chemosensation

Chemical signals are perceived by chemosensory organs. For example, the antenna is a major sensory organ for volatile compounds. The surface of the antenna is covered with morphologically distinct sensilla. Odorant molecules get into sensilla though the pores on the cuticle, then bind to the odorant receptors (ORs) on the dendrite membrane of olfactory neurons, which send signals to the brain (Pelosi, 1995; Hildebrand and Shepherd, 1997; Shanbhag et al, 1999; Stengl et al, 2010). The brain integrates these signals and makes instructions for the behaviors.

Odorant receptors are classified into two major groups, pheromone receptors and general odorant receptors, primarily based on their functions (Hansson, 2014; Hansson and Stensmyr, 2011). Pheromone receptors are specialized for the perception of sex pheromones that mediate sexual communication between males and females

(Jacquin-Joly and Merlin, 2004; Johansson and Jones, 2007). Most pheromone receptors are narrowly tuned to the respective components of sex pheromones, and their sensitivity is usually high (Nakagawa et al, 2005). On the other hand, general odorant receptors are considered to function in the perception of environmental odorants such as host-plant volatiles, the detection of which is crucial for the selection of oviposition sites (Hansson and Stensmyr, 2011; King et al, 2000; Zhang et al, 2013). General odorant receptors are as important as pheromone receptors for understanding of the molecular basis of ecological characteristics of each species, but their identification and functional analysis has not been conducted to the same extent compared with those for pheromone receptors (Bruyne and Baker, 2008).

Gustatory receptors (GRs) are structurally similar to ORs and expressed primarily in gustatory neurones (Rutzler and Zwiebel, 2005; Hallem et al, 2006; Vosshall and Stocker, 2007). GRs respond to sugars and bitter substances, as well as to CO2 and some contact pheromones (Hallem et al, 2006; Montell 2009; Obiero et al, 2014).

Ionotropic receptors (IRs) were characterized by a genome-based bio-informatics screen for insect-specific genes with enriched expression in OSNs in *Drosophila melanogaster* (Benton et al, 2009). IRs are similar but not closely related to ionotropic glutamate receptors (iGluRs) (Croset et al, 2010). Comparisons between genomes of insects and mammals showed that IRs are only found in insects (Silbering and Benton, 2010). Like ORs, IRs are expressed in sensory dendrites. Expression pattern of several *Drosophila* IRs suggested that they might be tuned to an odor panel such as small amine-like volatile compounds (Benton et al, 2009; Abuin et al, 2011; Croset et al, 2010). Whereas OR repertoires are primarily species specific, many IRs are

conserved in insects (Croset et al, 2010).

Sensory neuron membrane proteins (SNMPs) are uniquely expressed in olfactory neurons and homologous to a family of two-transmembrane domain receptor proteins, which includes a human fatty-acid transporter CD36 (Calvo et al, 1998; Rogers et al, 1997 and 2001). SNMPs function in the pheromone-detecting ORNs (Benton et al, 2007; Jin et al, 2008).

Odorant degrading enzymes (ODEs) are thought to inactivate odorant molecules by enzymatic degradation in the sensillar lymph and it could restore the sensitivity of ORs for receiving new chemical signals (Leal, 2013; Ferkovich et al, 1980; Vogt and Riddiford, 1981). The first ODE was identified in the sensillar fluid of male silkmoth, Antheraea Polyphemus (Vogt and Riddiford, 1981).

3. Ostrinia species

The genus *Ostrinia* (Lepidoptera: Crambidae) comprises 21 species including important agricultural pests (Ishikawa, 1999; Huang, 1998). Asian corn borer *Ostrinia furnacalis* (Guenée), and sympatric congeners *Ostrinia scapulalis*, *Ostrinia zaguliaevi* and *Ostrinia. zealis* occurs in Asia, whereas the European corn borer, *Ostrinia nubilalis*, occurs in Europe and North America. *O. furnacalis*, is the most economically important corn stalk boring pest in Asia and is distributed throughout India, Southeast Asia, China, Korea, Japan, Australia, New Guinea, Solomon Islands, and western Micronesia (Lewvanich, 1973; Mutuura and Monroe, 1970; Huang et al, 1998). *O. furnacalis* damage many plants such as cotton, tomato, sorghum, peppers and some beans (Caasi-Lit and Fernandez, 2006; Caasi-Lit et al, 2009). Laboratory and field studies indicated that male and female *O. furnacalis* adults have high flight capacities

for dispersal (Zhai and Chen, 1989; Shirai, 1998; Wang et al, 1994).

Because of their status as a serious pest, the sex pheromone communication system of *Ostrinia* species has been studied intensely. Sex pheromones of nine *Ostrinia* species have been characterized to date (Mutuura and Munroe, 1970; Ohno et al, 2003). Six pheromone components (Z9-14:OAc, E11-14:OAc, Z11-14:OAc, E12-14:OAc, Z12-14:OAc and E11-14:OH) were identified from these species, and the respective species use different combinations of these components in different proportions for species-specific signaling (Miura et al, 2010; Takanashi et al, 2000 and 2006; Roelofs et al, 1985; Glover et al, 1989; Huang et al, 1998b; Ishikawa et al, 1999). Nine pheromone receptors (including an odorant receptor coreceptor, *Orco*) have been identified in *O. nubilalis*, *O scapulalis*, and *O. furnacalis* (Miura et al, 2009 and 2010; Wanner et al, 2010; Leary et al, 2012). Electrophysiological analyses by ectopic expression in *Xenopus* oocytes have proven that these receptors in fact respond to the pheromone components (Miura et al, 2009 and 2010; Wanner et al, 2010). Although the male sex pheromones were shown to be involved in mating acceptance by females in *O. nubilalis* (Lassance and Lofstedt, 2009), most of the previously identified pheromone receptors were reported to be expressed exclusively in the male antennae (Miura et al, 2009 and 2010; Wanner et al, 2010). Odorant receptors responsible for perception of the male sex pheromones remain to be identified. Besides pheromone receptors, general odorant receptors in *Ostrinia* species are important for understanding of the molecular mechanisms underlying their ecological adaptation, such as host-plant specialization. Considering the fact that many *Ostrinia* species are important agricultural pests, general odorant receptors have the potential to be a target for novel pest control methods. However, no general odorant receptor has been identified in these species.

4. Current problems to solve in studies of genes involved in chemosensation in pest insects

Genes involved in chemosensation have been studied in the model insects such as fruit flies and the silkworm but less in pest insects. Identification of pheromone receptor genes in *O. furnacalis* were achieved by degenerate PCR based on the conserved sequences at the 5' and 3' terminals of ORF sequences (Miura et al, 2009 and 2010). It should be noted that receptors with divergent structures may have been overlooked by this method. Therefore, it is impossible to identify the entire repertoire of genes involved in the chemosensation in this way.

Several methods have been used for functional analyses of chemosensory genes. The most widely used was an ectopic expression system, in which receptors were expressed in *Xenopus* oocytes or in insect cells derived from ovaries such as sf9 (Miura et al, 2010; Kiely et al, 2007). Although this method is suitable to examine the molecular function of the chemoreceptors, it may not reflect the *in vivo* function correctly. For example, ligand specificity, which is influenced by other gene products interacting with chemoreceptors, may not be reproduced in ectopic environments. RNA interference (RNAi) might be used to reduce expression levels of chemosensory receptors. However, because it does not turn off the expression completely, understanding the results of physiological and behavioral analysis of the RNAi treated individuals may not be easy (Howlett et al, 2012). Therefore, better solutions are required for identification and functional analysis of chemosensory genes in pest insects.

5. Aims of this study

Two new technologies were integrated in this study to prove their potential in the study of chemosensory genes in pest insects: next generation sequencing (NGS) and genome editing. These two techniques are applicable to any insect species at least in theory. I selected *O. furnacalis* as a target species because of its importance as an agricultural pest and its well-studied background in the pheromone system. In chapter 1, antennal transcriptome analysis by RNA sequencing using the Roche 454 GS-Jr sequencing system was carried out as a preliminary trial for identification of odorant receptors. In chapter 2, RNA sequencing was performed in a larger scale using the Illumina MiSeq system to identify the repertoire of genes involved in the chemoreception including ORs, IRs, GRs, SNMPs and ODEs. In chapter 3, targeted mutagenesis of two selected odorant receptors, the odorant co-receptor *OfurOrco* and a pheromone receptor *OfurOR4*, were carried out using transcription activator-like effector nuclease (TALEN). In chapter 4, *in vivo* function of *OfurOrco* was confirmed by physiological and behavioral analysis of the mutant individuals. These methods are expected to be applicable to any species for identification and functional analysis of chemosensory genes.

Chapter 1

Identification of candidate odorant receptor genes using Roche 454 GS-Jr

1.1 Introduction

ORs play a central role in chemosensory signal transduction. In *O. furnacalis*, nine pheromone receptors (including an odorant receptor coreceptor, *Orco*) have been identified but additional pheromone receptors and general odorant receptors were totally unknown. NGS were good tools to identify novel genes because of its high throughput. Among them, the Roche 454 pyrosequencing system was the first to be introduced into consumer use, and has been widely applied for whole genome sequencing, transcriptome sequencing (RNA-seq), amplicon sequencing, sequence capture and metagenomics. Roche 454 sequencers produce long reads, which was initially about 400bp in length (2012) and currently improved to 1000 bp (2015). Long reads were expected to be advantageous for assembly of gene sequences in *de novo* transcriptome analysis of the species in which genome sequences were not available. The FLX system produces 450 Mb in a single run, and the GS-Jr system produces 35 Mb.

In this chapter, a GS-Jr system was used for a preliminary analysis to identify candidate odorant receptor genes in *O. furnacalis*. Expression levels of the identified genes in various tissues were examined by the quantitative RT-PCR analysis.

1.2 Materials and Methods

1.2.1 Insect rearing

O. furnacalis were collected on the Eai river bank (38°35′40″N, 140°57′20″E), Furukawa, Japan, in June 2010. This species is not endangered or protected. Collection of unprotected insects in this area does not require any permission. The collected insects were maintained in the laboratory on the artificial diet for silkworm (Silkmate 2M, Nosan Corporation Life-Tech Department, Yokohama, Japan) at 23°C, under a 16:8 light/dark cycle. The larvae were reared in the CE screw cap bottle (88.2mm wide, 116mm high) which cap was made a hold and fixed with a metal mesh. They were cultured at a high density of more than 300 individuals per bottle to make competition and bigger ones of the fifth instar were picked up and divided into a density of 60−80 individuals per bottle until they became pupae. The pupae were collected and divided by sex based on the morphology of the abdominal terminal segments. Eclosed adults were fed with water for 2 days, then allowed to mate in a net cage containing a plastic cup as the substrate for egg laying.

1.2.2 RNA sequencing and assembly

Male and female antennae were dissected from 2-day-old adults, and frozen in liquid nitrogen. RNA was immediately isolated from the frozen antennae using the Dynabeads® mRNA DIRECT™ Kit (Life Technologies Japan Ltd., Tokyo, Japan). The antennae from more than 100 individuals were pooled for a single RNA isolation experiment for each sex. Sequencing libraries were prepared followed the manufacturer’s manual for a GS Junior Titanium Rapid Library (cDNA) Preparation Kit

(Roche, Mannheim, Germany). Briefly, RNA was fragmented to 500bp in average length using $ZnCl_2$ and transferred into cDNA using cDNA synthesis kit (Roche, Mannheim, Germany) with random primers, then ligated with adaptors. The qualities of libraries were checked using Bioanalyzer with an RNA 6000 Pico kit (Agilent Technologies, California, USA) after remove short fragments. The libraries were carried on an emulsion PCR using GS Junior Titanium emPCR Kit Lib-L (Roche, Mannheim, Germany) followed the description of manufacturer's manual for GS Junior Titanium emPCR (Lib-L) besides that the molecular numbers of libraries were using three times of capture beads, 3×10^7 in total. Sequencing was carried on using Roche 454 GS-Jr system with GS Junior Titanium Sequencing Kit (Roche, Mannheim, Germany) according to the manufacturer's manual of GS Junior Titanium Sequencing Run. Libraries from males and females were sequenced separately. The reads from male and female were pooled together and assembled using Trinity r2012-10-05 (http://trinityrnaseq.sourceforge.net/; Grabherr et al, 2011) and Newbler v2.5 (Roche, Mannheim, Germany) independently.

1.2.3 Screening of candidate odorant receptor genes and read mapping

Obtained contigs were translated into peptide sequences in six reading frames using an in-house script. Contigs were first screened by similarity to *Bombyx mori* odorant receptors (*BmorORs*) using three different methods to maximize the possibility of identifying candidate contigs of odorant receptors (**Fig. 1-1**). A total of 68 sequences of *B. mori* were obtained from the database (Tanaka et al, 2009; Jordan et al, 2009). *B. mori* sequences were used as a query in BLASTn or BLASTx searches against the contigs or translated peptide sequences. In parallel, PSI-BLAST searches were

performed using alignments of *BmorORs* in various groupings as a query (**Table 1-1**). In all the searches, the E-value cutoff was set to 0.0001. Overlapping variants were removed at this step by selecting the longest one as a representative transcript of a variant group. The results of three screenings were merged and duplications were removed. The remaining contigs were screened for the presence of transmembrane domains using SOSUI (http://harrier.nagahama-i-bio.ac.jp/sosui/; Hirokawa et al, 1998). The contigs that contained transmembrane domains were finally screened using BLASTn against the NCBI non-redundant protein database (12.07.2012), and those that had an insect odorant receptor as a top-hit homolog were considered as candidate odorant receptors. Screening of contigs from two assemblers was done independently, and the results were pooled and duplicated sequences were removed. The expression level of each candidate was estimated by mapping the raw reads to the nucleotide sequences of the candidate contigs of odorant receptors using CLC Genomics Workbench (http://www.clcbio.com/products/clc-genomics-workbench/; Orsi et al, 2013) with setting of following parameters; Mismatch cost: 2, Insertion cost: 3, deletion cost: 3, Length fraction: 0.5, Similarity fraction 0.9. Expression levels were calculated by reads numbers.

Table 1-1. Grouping of *BmorORs* used in the query of the PSI-BLAST search.

Groups	***B. mori* odorant receptors**
1	*OR26, OR45, OR46, OR47, OR48, OR50, OR51, OR57, OR58, OR59, OR63*
2	*Group1, OR10, OR11, OR12, OR13, OR15, OR16, OR24, OR25, OR33, OR34, OR41, OR61, OR62, OR65, OR66, OR67*
3	*OR1, OR5, OR7, OR18, OR35, OR37, OR38, OR39, OR43, OR68*
4	*OR5, OR7*
5	*OR27, OR29, OR49, OR53, OR54, OR55, OR56*
6	*Group2, Group3, Group5*
7	*Group6, OR3, OR23, OR28, OR42, OR64*
8	*Group7, OR40, OR44*
9	*Group8, OR2, OR8, OR19, OR20, OR21, OR22, OR30, OR36*
10	*Group9, OR4, OR6, OR9, OR14, OR17, OR32, OR52, OR60*

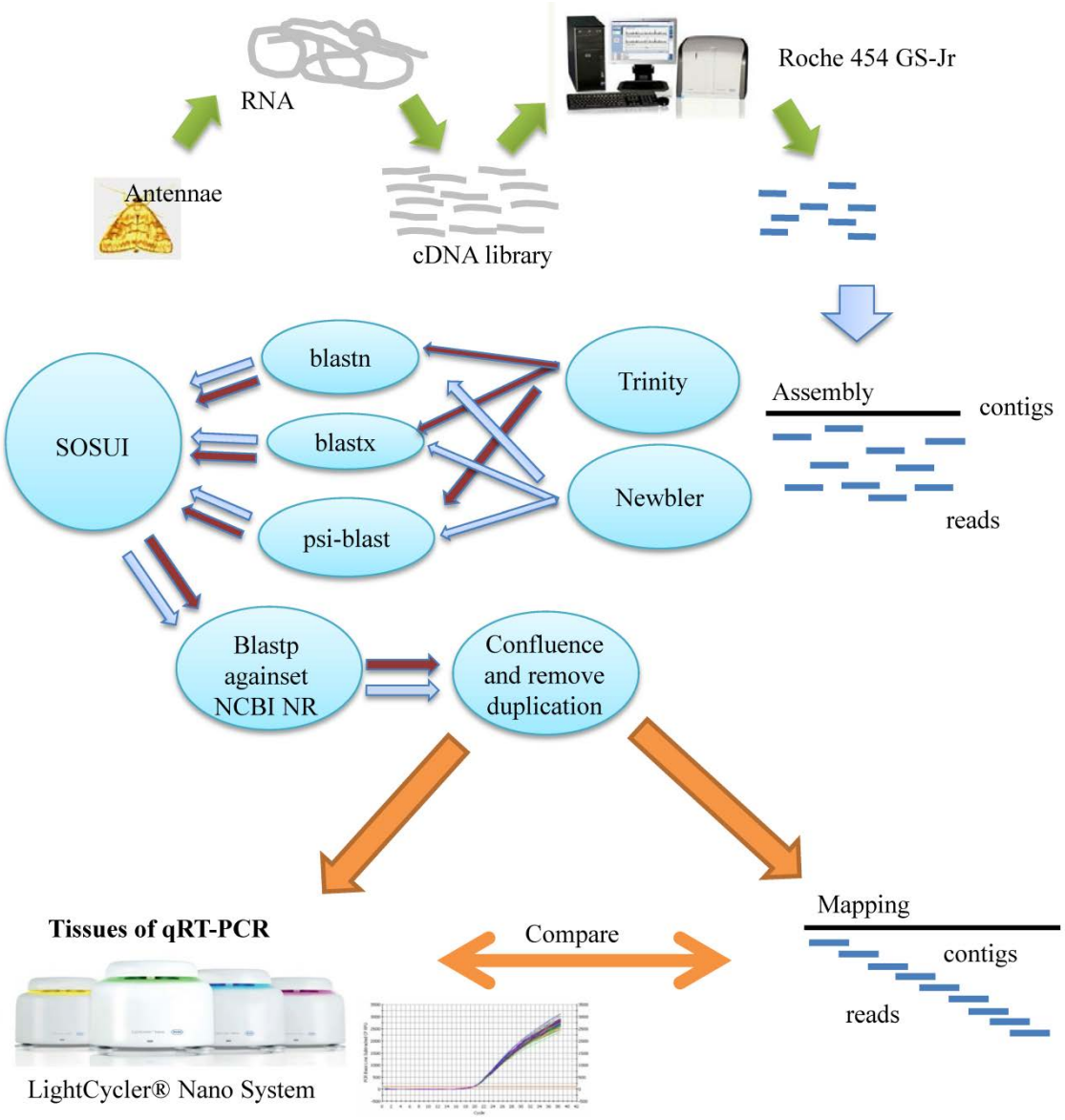


Figure 1-1. Schematic diagram of sequence data analysis.

1.2.4 Expression levels of candidate odorant receptors in various tissues

The relative expression levels of the candidate odorant receptors in various tissues (male antennae, female antennae, male proboscis, female proboscis, male foreleg, female foreleg, male mid&hindleg, female mid&hindleg, male pupae, female pupae, larvae head, larvae foreleg, larvae fatbody and larvae midgut) were examined by quantitative reverse transcription PCR (qRT-PCR). Primers were designed to amplify an approximately 500 bp-long fragment at the 3' end of each candidate (**Table 1-2**). Tissues were dissected from 2-days-old adults independently from those used in the RNA-seq analysis. mRNA was isolated using the Dynabeads® mRNA DIRECT™ Kit(Life Technologies^{TM}, Carlsbad, CA, USA) and cDNAs were transcribed using the SuperScript III First-Strand Synthesis System (Invitrogen, Carlsbad, CA, USA). Five biological repeats of antennae and three biological repeats of other tissues (independent RNA extraction and cDNA synthesis) were made. Quantitative RT-PCR was done using the LightCycler Nano system (Roche, Mannheim, Germany) with the FastStart Essential DNA Green Master Kit (Roche, Mannheim, Germany). RpS3 were included in the analysis, whose quantification cycle (Cq) value was used as an internal control. Relative expression levels of candidate receptors to the internal controls were estimated as $2^{-\Delta Cq}$, where ΔCq represents the difference of Cq between each candidate receptor and the internal control.

Table 1-2. Primers for qRT-PCR.

Name	Primers (F/R)
OR1	GTGCTGTTCCTGCTCTACAAC/GCTGAACGTTCGCAAGAACATG
Orco	GCTCATCAGTGATGGAAGCAG/GCACCAAGTACAGAAGCGAAC
OR3	AAGGACGCGGTGTACAGCATG/ATAATGGTGGCCATCGTCTGC
OR4	CTGACGACTGTAGTGTTTGGTC/TCCTGCATCTTGTGCAGCATG
OR5a	AGTTGTTGGGTTCAGAGACTG/GAGGCTGATCTTGTACTGCATC
OR5b	AGATGCTGGGTTCAGAGACTG/GAGGCTGATCTTGTACTGCATC
OR6	GGTCAGCTCATACAGATGTCAG/CCAGTCCTAATGCCTTGAGAC
OR7	GCCTGATGAACGCTGTGTACG/GCCGATGTCCACCATGTTCAG
OR8	CCAGATTCAGCTTACTTCCTG/AATATCCCGGTCATGGTGCTG
T1332	ACCAACAGTCCTACGTAAGTG/CGCACACTGAGAACATAGATG
T1692	TCCCTACGAAAGTCGATGGTG/AGTGAAGTAACTGTAGGCTCC
T2420	CCATTGCGAAAACAATGCTGC/ CCACCGGCATAAATTGAAGCC
T2863	CTGGGAACGAGTGTTAGACAC/AGCGTATACGCCATTTTGACG
T8374	GGGATGTTACATAGGAGCAATG/ CATCTTGAGGCCGGTGATTAC
T12091	GAGTAACAAAGTGGACGACGG/GTAGTAGCTGTACGACCCTTTC
T15259	ATCCATCCAGCCATCCTATGC/AGTGGGCGTTTGCAGCGTTCC
T15490	GGAATACATCAGTCGCAGTCG/CACTTCGCTGAATGGAGTGTAC
T15830	ATTAGAAGCAGCTCGCATCGC/GCATAATACGAATACGCTGCC
T15990	CATCGTCAATTAAGCCTCCTCG/AAACTGCGAGCTCATGTGCAG
T17244	TGTTACTGCGGAGACTTGGTG/TCGCTGCAAATGTGGTGAGTG
T18209	TGCCAGTTTGATGTACTCTGC/GAGTTCCATGGCGTGGTCAAG
T20049	CATGCTCATGGATGCCAGTAC/CGAAGTACGACCACGACGTGC
T21100	CCAACATCAATCTCATCCAGC/ATCGCGACTTGCATCAGTGC
T22848	AACGTGAAGAAGCGAGCCATG/GAGGTAGCAGGTGAAGAGTAC
T23271	ACTGCAGTTTATGTGCGCGAC/ACCCATGCCTAGTATCACCAG
T35457	CGAATATGCGATGCGTGTTGG/GCCTGTCTCAGCATGTTGAAC
T35545	TTTAGCTACACACGAGCCTGG/GCCTTTGATGCGATTTGTCGG
T36638	CCTGACCCTCATACTGGTTGC/CTCGTTATCGACTTGTTGTCC
T36777	ATCGCTAGCAGCATACGAGTC/TCCGTTGTATAGAGCCTGTGC
N09067	ATGTGATGCTCGGTCCCTACG/ACTTTGCTGATTCGCAGCTGG

1.3 Results

1.3.1 RNA sequencing and screening of candidate odorant receptors

From single sequencing runs with Roche 454 GS-Jr for each of two cDNA libraries, 138,964 and 127,252 reads with a total number of bases more than 61 million and 56 million from male and female antennae, respectively (**Table 1-3**). The read lengths were at a range from 40 bases to 752 bases and the average length was approximately 440 bp (**Fig 1-2**). All reads were pooled together to be assembled into 29,864 contigs using Trinity and 13,242 contigs using Newbler (**Table 4**). In the first screening, homology searches against *B. mori* odorant receptors were performed using BLASTn, BLASTx, and PSI-BLAST in parallel. Results from three screenings were merged and duplications were removed. Because many putative variants with differences in the length of terminal region were included in the contigs, criteria for selection of representative sequence were examined by comparing with the sequences of the previously identified pheromone receptors. In the most cases, longest variants were best matched with the known sequences. Therefore, the longest variant was selected from each variant group. After the removal of duplicates, 169 and 131 contigs remained from those assembled with Trinity and Newbler, respectively. In the Second screening using SOSUI, 101 and 86 sequences were found to contain transmembrane domains from contigs made assembled by Trinity and Newbler, respectively. These sequences were finally screened against the NCBI non-redundant protein database, of which 28 from Trinity and 21 from Newbler had an insect odorant receptor as the top-hit homolog. Theese contigs were pooled together, and after the removal of duplicates, 29 candidate odorant receptors were identified (**Table 1-4**). Eight of the nine previously identified

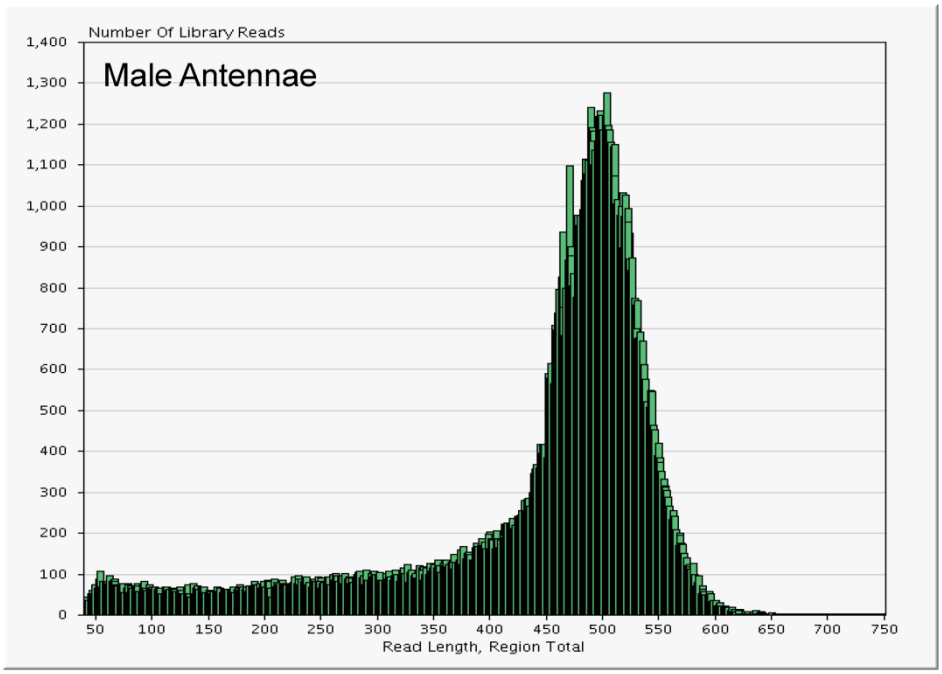
pheromone receptors were found in the candidates, in which only five were full length. Sequences of two putative duplicated receptors, *OfurOR5a* and *OfurOR5b*, are quite similar to each other, Newbler made two independent contigs for them, while Trinity combined them resulting in the one conting. *OfurOR1* was not found from the results of both assemblers. The ORF sequences obtained for the 21 newly identified receptors seemed not to be full length, and 7 of them were fragmented by stop codons (**Table 1**-**5**).

Table 1-3. Summary of sequencing results.

GACT (Library)	Antennae	
	Male	**Female**
Raw Wells	236,356	201,928
Key Pass Wells	229,960	195,465
Passed Filter Wells	138,964	127,252
Total Bases	61,218,105	56,255,509
Length Average	440.53	442.08
Length Std Deviation	119.57	117.23
Longest Reads Length	751	752
Shortest Reads Length	40	40
Median Reads Length	483	483

Table 1-4. Screening steps.

	Female		**Male**
Raw reads	127252		138964
Total reads		266216	
	Trinity		**Newbler**
Assembled contigs	29864		13242
1st screening	169		131
2nd screening (SOSUI)	101		86
3rd screening (NR database)	28		21
Total		29	



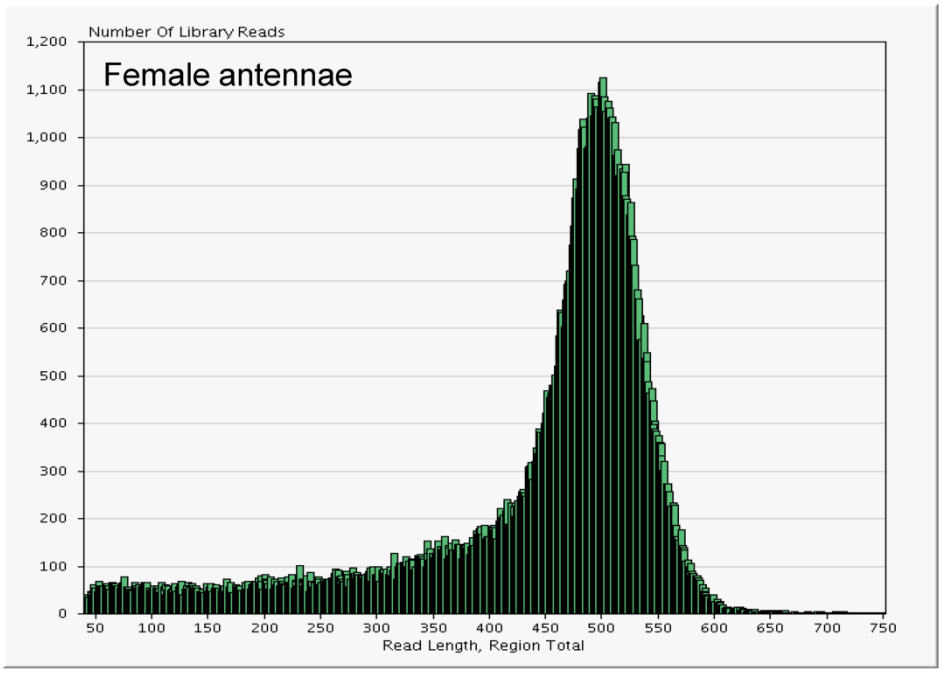


Figure 1-2. Distribution of reads length of 454 GS-Jr system.

1.3.2 Expression levels of the candidate odorant receptors estimated by read mapping

Expression level of the candidate odorant receptors were estimated by number of reads mapped to the candidate receptors. The sequence of *OfurOR1* was obtained from the database. Although the number of mapped reads for the previously identified pheromone receptors was not large, the results were basically consistent with expectation. Highest expression was observed for *OfurOrco*, and it was expressed equally in males and females, while other pheromone receptors showed male-specific expression except for *OfurOR7* (**Table 1-5**). For the newly identified receptors, it was difficult to tell anything about their sex specific expression because most of them had less than ten reads.

1.3.3 Tissues specific expression levels of the candidates confirmed by qRT-PCR

Because the numbers of mapped reads were so small for most of the receptors, qRT-PCR was used to estimate the expression level. In addition, tissue specificity was examined using various tissues other than the antennae. To evaluate the reliability of our qRT-PCR analysis, multiple replications were prepared. For both male and female antennae, nine replicates comprised of 3 biological replicates (independent sampling and cDNA library construction), each with 3 technical replicates, were made. Results of replicates were consistent with each other (**Fig. 1-3A**). For the other tissues, 3 technical replicates were made (**Fig. 1-3B**).

The results of expression analyses were summarized in **Table 1-6**. For the previously identified pheromone receptors, the results of qRT-PCR in the antennae were roughly consistent with those of read mapping. In terms of tissue specificity, most of the

candidate receptors showed antennae biased expression with some exception. Five receptors, *OfurOR7*, *T12091*, *T17244*, *T21100*, and *N09067*, were expressed in most of the tissues examined. Many of the candidate receptors, including a part of the previously reported pheromone receptors, were expressed in proboscis in addition to antennae. Three of the newly identified receptors, *T8374*, *T15490*, and *T22848* were specifically expressed in the antennae both in males and females. On the other hand, *T23271* and *T35545* were not detected by qRT-PCR in the antennae. Many candidate receptors were expressed in the pupa. The reason of it is discussed in the next section.

Table 1-5. List of candidate odorant receptors.

Name	Contigs Length (* : full length)	ORF length (aa)	Mapped Reads	
			Male	Female
OR1	1275#	425#	0	0
***T4408+N560* (*Orco*)**	1574 *	475	179	169
***T7761* (*OR3*)**	1609	426	16	0
***T3947* (*OR4*)**	1715 *	413	75	0
***N00956-2* (*OR5a*)**	1371 (gap:869-927)	289+148	6	0
***N00956-1* (*OR5b*)**	1134 (gap: 726-811)	241+107	19	0
***T9084+370* (*OR6*)**	1266 *	421	12	0
***T12315+137* (*OR7*)**	1354 *	431	7	1
***T2808* (*OR8*)**	1736 *	447	10	0
T1332	674	36+89+65	0	6
T1692	821	216	1	4
T2420	1080	274+43	4	2
T2863	1267	361	5	11
T8374	1024	294+34	2	7
T12091	1081	301+25	3	3
T15259	639	213	0	4
T15490	729	220	3	1
T15830	600	103+27	1	3
T15990	741	247	1	3
T17244	624	177	0	5
T18209	943	314	3	1
T20049	1090	242	1	4
T21100	717	239	2	2
T22848	609	203	1	2
T23271	553	120	1	2
T35457	483	33+131	0	2
T35545	628	197	2	0
T36638	489	61+15+75	2	0
T36777	405	122	0	2
N09067	687	179	1	2

#: Reported in the previous paper (Miura et al, 2010).

Table 1-6. Summary of expression levels in various tissues. Symbols were assigned by median values of replicates (X); **++**: X>0.1, **+**: 0.01<X<0.1, **-**: X<0.01.

Name	Mapped Reads		Antennae		Proboscis		Foreleg		Mid&Hindleg		Pupa		Head	Foreleg	Midgut	Fatbody
	M	F	M	F	M	F	M	F	M	F	M	F	Larvae			
OR1	0	0	+	+	++	+	-	-	-	-	++	-	-	+	++	-
Orco	179	169	++	++	-	-	-	-	-	-	+	-	-	-	-	-
OR3	16	0	++	+	+	+	-	-	-	-	-	-	-	-	-	-
OR4	75	0	++	-	-	-	-	-	-	-	-	-	-	-	-	-
OR5a	6	0	++	+	+	-	-	-	-	-	+	-	-	-	-	-
OR5b	19	0	++	+	+	-	-	-	-	-	+	-	-	-	-	-
OR6	12	0	++	+	++	+	-	-	-	-	++	-	-	+	+	-
OR7	7	1	++	++	++	++	++	++	++	++	++	++	+	++	++	++
OR8	10	0	++	+	++	++	-	-	+	-	++	-	-	+	+	-
T1332	0	6	+	+	++	+	-	-	-	-	+	-	-	+	+	-
T1692	1	4	+	+	++	+	-	-	-	-	+	-	-	-	-	-
T2420	4	2	+	+	++	+	-	-	-	-	+	-	-	-	-	-
T2863	5	11	+	++	++	+	+	+	+	-	+	+	-	+	-	-
T8374	2	7	++	++	-	-	-	-	-	-	-	-	-	-	-	-
T12091	3	3	+	+	++	++	+	+	+	++	++	+	-	++	+	+
T15259	0	4	+	+	+	-	-	-	-	-	-	-	-	-	-	-
T15490	3	1	+	+	-	-	-	-	-	-	-	-	-	-	-	-
T15830	1	3	+	+	+	+	-	-	-	-	-	-	-	-	-	-

T15990	1	3	+	+	+	+	-	-	-	-	+	-	-	-	-	-
T17244	0	5	++	++	++	++	+	-	+	+	++	+	-	+	++	-
T18209	3	1	+	-	+	-	-	-	-	-	-	-	-	-	-	-
T20049	1	4	+	++	+	+	-	-	-	-	-	-	-	-	+	-
T21100	2	2	+	+	++	++	+	+	+	+	+	-	-	+	+	-
T22848	1	2	+	+	-	-	-	-	-	-	-	-	-	-	-	-
T23271	1	2	-	-	+	-	-	-	-	-	-	-	-	-	-	-
T35457	0	2	+	+	+	+	-	-	-	-	+	-	-	+	-	-
T35545	2	0	-	-	-	-	-	-	-	-	-	-	-	-	-	-
T36638	2	0	+	+	+	+	-	-	-	-	+	-	-	-	-	-
T36777	0	2	+	+	+	+	-	-	-	-	-	-	-	-	-	-
N09067	1	2	+	++	++	++	+	-	+	+	++	+	-	+	++	+

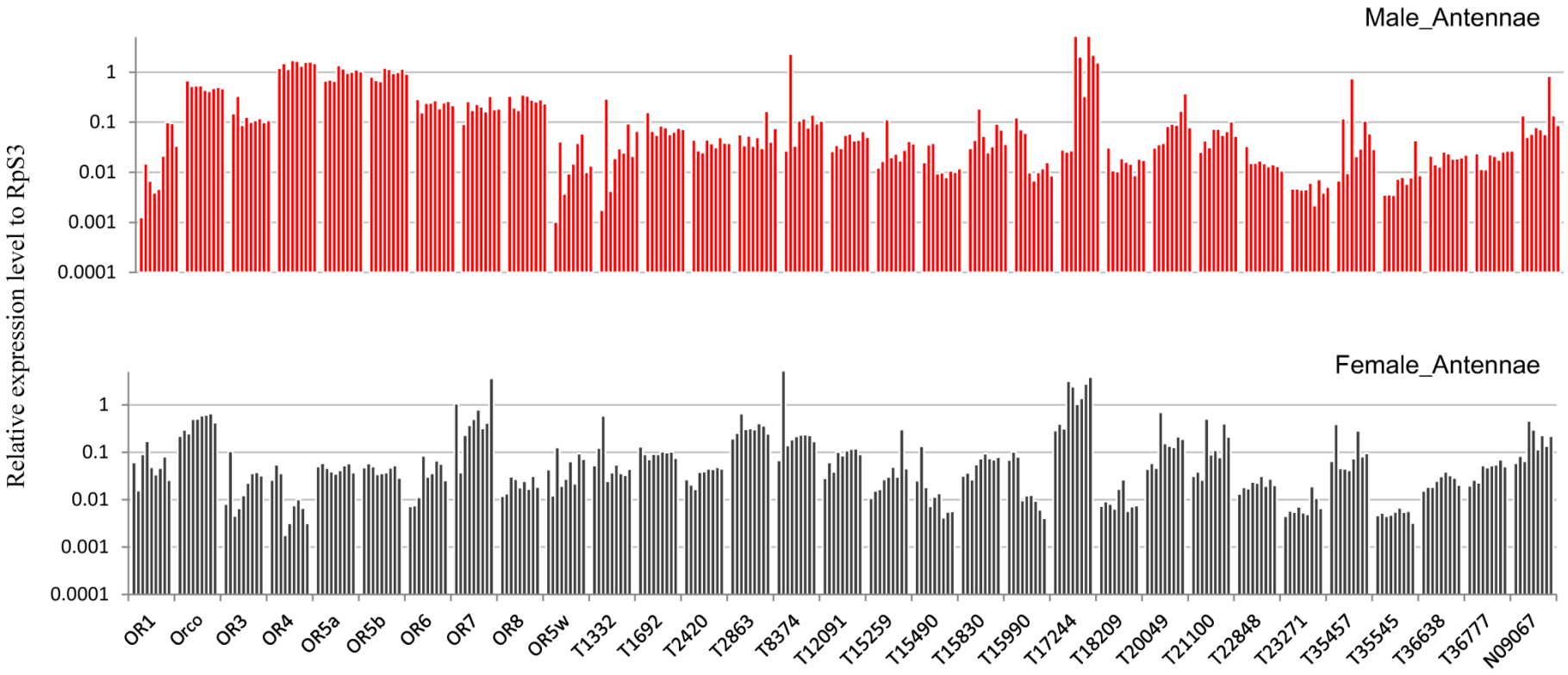


Figure 1-3A. Relative expression levels of candidate odorant receptors in the antennae. Ribosomal protein S3 (*RpS3*) was used as an internal control. Three biological replicates (three independent cDNA libraries) were made, and three technical replicates were made for each library (in total 9 replicates).

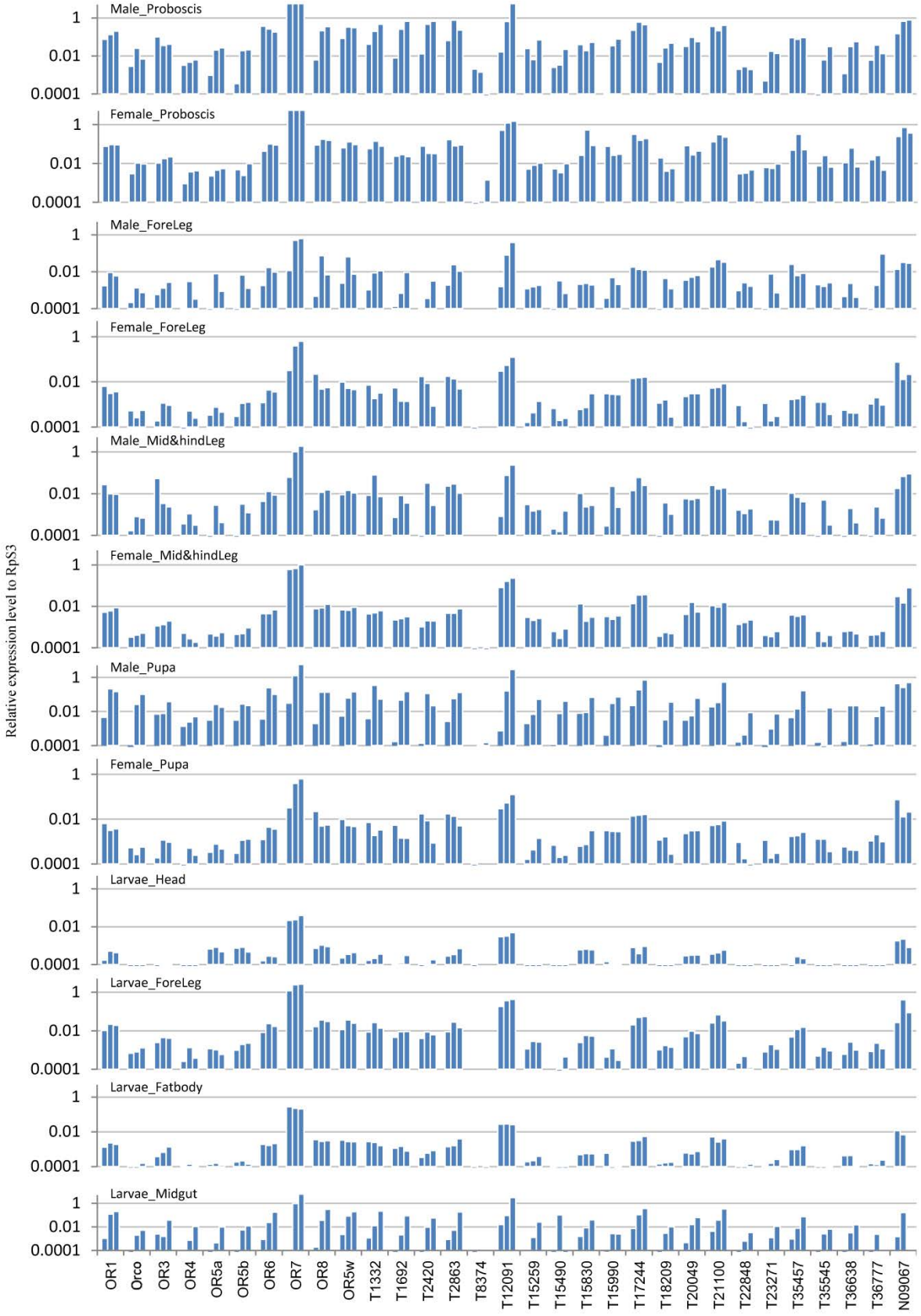


Figure 1-3B. Relative expression levels of candidate odorant receptors in various tissues. Ribosomal protein S3 (*RpS3*) was used as an internal control. Three technical replicates were made for one cDNA library.

1.4 Discussion

1.4.1 Suitability of the GS-Jr system

Although eight of the nine previously identified pheromone receptors were found in our result, full ORF sequences were not obtained for *OfurOR3*, *OfurOR5a*, and *OfurOR5b*. For all of the newly identified receptors, which showed much lower expression levels than the previously identified pheromone receptors, full length of ORF sequences were not obtained. It should be resulted because the number of reads was not enough to cover the whole ORF sequences. Expression levels of odorant receptors were low even in the antennae. Throughput of the GS-Jr system seems to be too low for our purpose.

Another weakness of the Roche 454 sequencers is the fidelity of obtained reads. Stop codons truncated the ORF sequences of *OfurOR5a* and *OfurOR5b*, and 7 newly identified receptors. They were supposed to be derived from sequencing errors. Pyrosequencing produces errors at high rate in reading homopolymers, which result in frame shifts in the obtained contigs.

The Roche 454 sequencers produce long reads. This characteristic was thought to be advantageous for *de novo* assembly of transcripts, particularly when two transcripts with similar sequences are to be discriminated in assembly. However, this advantage was obvious only with Newbler, an assembler specifically designed for the Roche 454 systems. Newbler produced contigs corresponding to each of *OfurOR5a* and *OfurOR5b*, but Trinity combined them into a single contig.

Considering above issues, GS-Jr is inadequate for identification of odorant receptors in *O. furnacalis*. Other sequencing platforms with larger throughput and

higher fidelity were required for this purpose.

1.4.2 Tissue specificity of expression

Expression of several receptors was detected in the pupae. Because the timing of sampling during the pupal stage was not controlled, pupae at the very late stage of development might be included. It is possible that these receptors were expressed in the developing antennae before eclosion.

High level expression of *OfurOR7*, *T12091*, *T17244*, *T21100*, and *N09067* was detected in many tissues. Although the PCR products were always examined by the melting curve assay to confirm homogeneity, they were not always examined by gel electrophoresis for their size. Therefore, it was not sure whether the primers used in the RT-PCR analysis could specifically amplified the target fragments in all the analyses. Because the whole genome sequence was not available for *O. furnacalis*, specificity of the primers could not be evaluated beforehand. It would be necessary to re-examine the results using alternative primer pairs if such pattern of amplification was observed in the future analysis.

Expression of other receptors was restricted to the antennae and proboscis, on which chemosensilla exist. Therefore, it would be safe to conclude that these candidate genes encode odorant receptors.

Chapter 2

Identification of candidate chemosensory genes using Illumina MiSeq

2.1 Introduction

In chapter 1, 29 candidate odorant receptors were identified using Roche 454 GS-Jr. However, full length of ORF sequences was not obtained for most of them because the number of reads corresponding to each receptor was small. To solve this problem, deeper sequencing was necessary. Illumina HiSeq systems produce up to 1.5 Tb in a single run, while up to 15 Gb can be produced by the MiSeq system. Unlike Roche 454 system, Illumina sequencing chemistry does not use beads; DNA fragments are attached to the oligo nucleotides on the flow cell by specific adaptors. The fragments are amplified by a special double strand bridge and sequenced from oligo nucleotides. Each oligo nucleotide attached with a fragment produces a read. Because the adaptor can be attached with specific tag sequence, different libraries attached with different tags can be sequenced together in a single run (multiplex). Although the length of reads was initially short (50bp), recent improvement of chemistry achieved more than 200bp of the read length in the MiSeq system.

In this chapter, considering the low expression level of chemosensory genes, a MiSeq system was used for RNA sequencing analysis. In addition to ORs, other chemosensory genes such as IRs, GRs, SNMPs, and ODEs were also included in the target of the analysis

2.2 Materials and Methods

2.2.1 Insect rearing

O. furnacalis were collected as described in chapter 1. The collected insects were maintained in the laboratory on the artificial diet for silkworm (Silkmate 2M, Nosan Corporation Life-Tech Department, Yokohama, Japan) at 23°C, under a 16:8 light/dark cycle. The larvae were reared in the insect breeding jar (100mm diameter × 80mm height, 310122; SPL Lifesciences Co. Ltd., Seoul, Korea) at a density of 60−80 individuals per bottle until they became pupae. The pupae were collected and divided by sex. Eclosed adults were fed with water for 2 days. and allowed to mate in a net cage containing a plastic cup as the substrate for egg laying.

2.2.2 RNA sequencing and assembly

Male and female antennae were dissected from 2-day-old adults, and frozen in liquid nitrogen. RNA was immediately isolated from the frozen antennae using the QuickGene RNA tissue Kit SII (RT-s2; KURABO, Neyagawa, Japan). The antennae from more than 20 individuals were pooled for a single RNA isolation experiment. Three biological repeats for each sex were made for the analysis of expression levels. Sequencing libraries were prepared using the TruSeq RNA Sample Preparation Kit v2 according to the LS protocol of the manufacturer's instructions (Illumina, Inc., San Diego, CA, USA) using 1 μg of total RNA from each sample, except for the following modifications to select the library with long inserts. Incubation time of purified mRNA fragmentation was changed from 8 min to 30 sec at 94°C, and 0.7×volume of the AMPure XP beads was used in the all purification steps. Prepared libraries were mixed

at a concentration identical to each other in a 1.5 ml tube and applied for cluster generation on the MiSeq system using the MiSeq Reagent Kit v3 600-cycle (Illumina, Inc., San Diego, CA, USA). A total of 6 libraries were indexed and applied for a single multiplex run in the 300 bp single-end mode. The raw data were deposited in the DDBJ Sequence Read Archive under accession number DRA002255. The reads were preprocessed with cutadapt v1.2.1 (Marcel, 2008) for quality trimming at QV30 with a minimum length of 50bp. The pass-through reads were pooled and assembled using Trinity r2013_08_14 (http://trinityrnaseq.sourceforge.net/; Grabherr et al, 2011). Open reading frames were extracted from the Trinity contigs with TransDecoder (http://transdecoder.sourceforge.net/; Haas et al, 2013; Liu et al, 2014; Hoeppner et al, 2014; Yang et al, 2014) using the script that came with the Trinity distribution without modification.

2.2.3 Screening of odorant receptors and read mapping

The extracted ORF sequences (referred to as the Trinity transcripts hereafter) were first screened by similarity to *Bombyx mori* odorant receptors (*BmorORs*) using two different methods to maximize the possibility of identifying candidate odorant receptors (**Fig. 2-1**). A total of 68 protein sequences of *B. mori* were obtained from the database (Tanaka et al, 2009; Jordan et al, 2009), and each was used as a query in BLASTp searches against the Trinity transcripts. In parallel, PSI-BLAST searches were performed using alignments of *BmorORs* in various groupings as a query (**Table 2-1**). In both searches, the E-value cutoff was set to 0.0001. Overlapping variants were removed at this step by selecting the longest one as a representative transcript of a variant group. The results of two screenings were merged and duplications were

removed. The remaining Trinity transcripts were screened for the presence of transmembrane domains using SOSUI (http://harrier.nagahama-i-bio.ac.jp/sosui/; Hirokawa et al, 1998) and TMHMM (http://www.cbs.dtu.dk/services/TMHMM/; Krogh et al, 2001). The transcripts that contained transmembrane domains were finally screened using BLASTp against the NCBI non-redundant protein database (16.12.2013) (**Fig. 2-1**), and those that had an insect odorant receptor as a top-hit homolog were considered as candidate odorant receptors. The expression level of each receptor was estimated by mapping the raw reads to the ORF sequences of the candidate odorant receptors using bowtie2 v2.0.6 in local mode with -a option, followed by processing with eXpress v1.5.1 (Roberts and Pachter, 2012). Expression levels were calculated as reads per kilobase of the ORF length per million total reads for each library (RPKM) (Mortazavi et al, 2008).

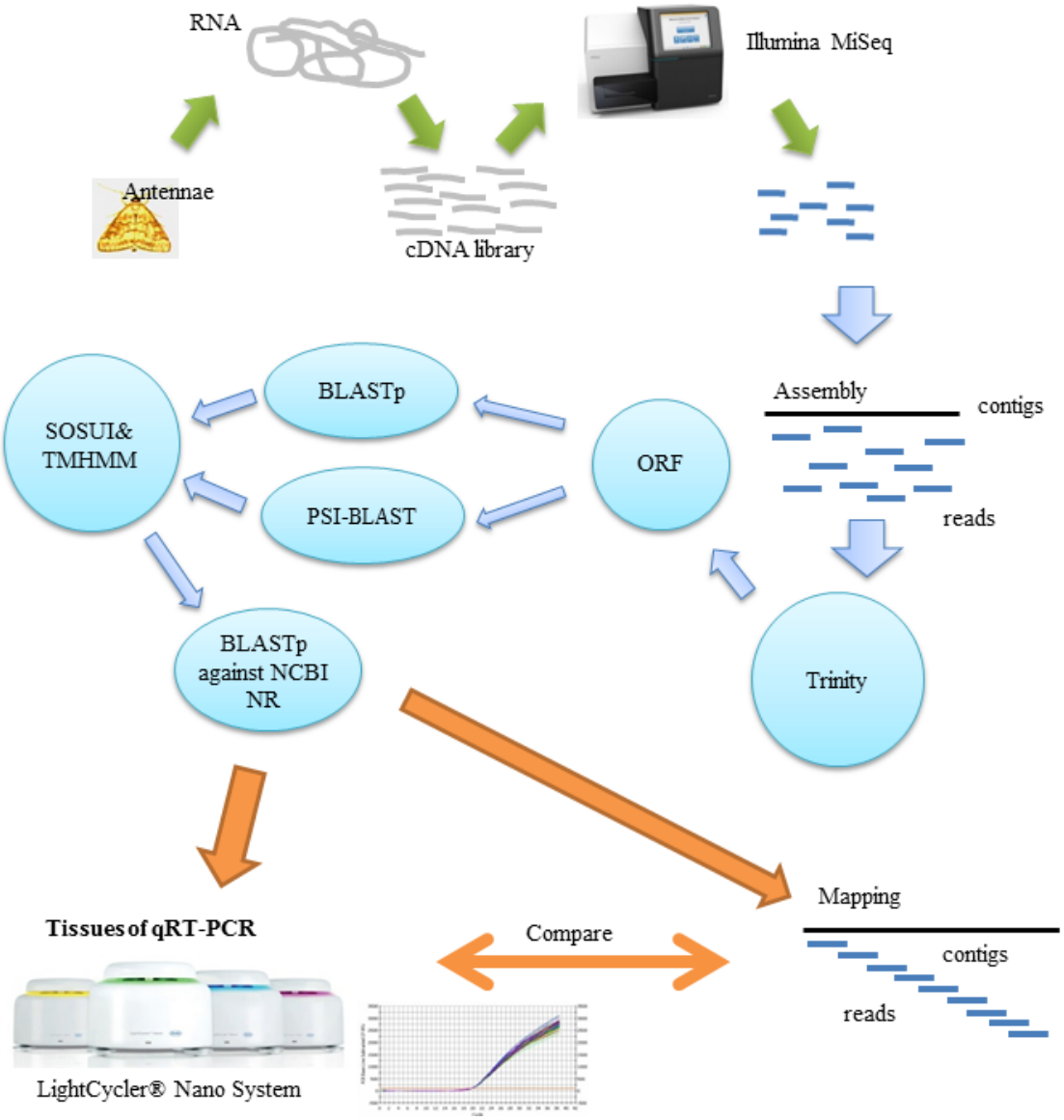


Figure 2-1. Schematic diagram of sequence data analysis.

2.2.4 Quantitative reverse transcription PCR

The relative expression levels of the candidate odorant receptors in the antennae and the thorax were validated by quantitative reverse transcription PCR (qRT-PCR). Thorax (mixture from males and females) was used as a control to examine the tissue specificity of the expression pattern. Primers were designed to amplify an approximately 200 bp-long fragment at the 3' end of the ORF of each candidate (**Table 2-2**). Tissues were dissected from 2-days-old adults independently from those used in the RNA-seq analysis. Total RNA was isolated using the QuickGene RNA tissue Kit SII (RT-s2, KURABO, Neyagawa, Japan) and cDNAs were transcribed using the SuperScript III First-Strand Synthesis System (Invitrogen, Carlsbad, CA, USA). Three biological repeats (independent RNA extraction and cDNA synthesis) were made. Quantitative RT-PCR was done using the LightCycler Nano system (Roche, Mannheim, Germany) with the FastStart Essential DNA Green Master Kit (Roche). Three genes, *RpS3*, *actin*, and *NADH dehydrogenase*, were included in the analysis, and the average quantification cycle (Cq) value of these three genes was used as an internal control. Relative expression levels of candidates to the internal controls were estimated as $2^{-\Delta Cq}$, where ΔCq represents the difference of Cq between each candidate and the internal control.

2.2.5 Phylogenetic analysis

Phylogenetic relationships of *O. furnacalis* odorant receptors (*OfurORs*) were analyzed against *BmorORs* and *Cydia pomonella* odorant receptors (*CpomORs*) (Wanner et al, 2007; Bengtsson et al, 2012). A total of 164 amino acid sequences were

aligned using MAFFT v7.130 with the option E-INS-I (Katoh and Toh, 2010). Phylogenetic relationship was deduced by the maximum likelihood method using RAxML v8.0.17 (Stamatakis et al, 2012; Vieira and Rozas, 2011) with the GAMMA model for rate heterogeneity and the WAG model for substitution matrix. In addition, the rapid hill-climbing search algorithm (–f d) was used. Model optimization precision in log likelihood units for final optimization of tree topology (–e) was set at 0.0001. The tree image was created using FigTree v1.4.1 (http://tree.bio.ed.ac.uk/software/figtree/; Lemey et al, 2009).

2.2.6 Identification of other genes involved in the odorant perception

IRs, GRs, SNMPs, and ODEs were identified by the same method used in the identification of odorant receptors, except for the screening by presence of transmembrane domains which was not applied for SNMPs and ODEs. Protein sequences of IRs, GRs, SNMPs in *B. mori* were obtained from the database (Croset et al, 2010; Sato et al, 2011; Wanner et al, 2008; Rogers et al, 2001), grouping for PSI-BLAST was shown in **Table 2-1**. SNMPs in *Ostrinia furnacalis* were also used as queries (Allen and Wanner, 2011). Because ODEs were not systematically studied in the *B. mori*, ODEs in *Sesamia inferens* were used (Zhang et al, 2014).

Table 2-1. Grouping of genes used in the query of the PSI-BLAST search.

Groups	genes from other species
ORs	***B. mori***
1	*OR26, OR45, OR46, OR47, OR48, OR50, OR51, OR57, OR58, OR59, OR63*
2	*Group1, OR10, OR11, OR12, OR13, OR15, OR16, OR24, OR25, OR33, OR34, OR41, OR61, OR62, OR65, OR66, OR67*
3	*OR1, OR5, OR7, OR18, OR35, OR37, OR38, OR39, OR43, OR68*
4	*OR5, OR7*
5	*OR27, OR29, OR49, OR53, OR54, OR55, OR56*
6	*Group2, Group3, Group5*
7	*Group6, OR3, OR23, OR28, OR42, OR64*
8	*Group7, OR40, OR44*
9	*Group8, OR2, OR8, OR19, OR20, OR21, OR22, OR30, OR36*
10	*Group9, OR4, OR6, OR9, OR14, OR17, OR32, OR52, OR60*
IRs	***B. mori***
1	*IR8a, IR25a, IR40a, IR76b, IR93a*
2	*IR7d2, IR7d3, IR87a, IR143*
3	*IR64a, IR75d, IR75p, IR75q2*
4	*Group2, IR21a, IR41a, IR68a*
5	*Group1, Group4*
6	*Group3, Group5*
GRs	***B. mori***
1	*GR39, GR41, GR42, GR43, GR44, GR45, GR46, GR, GR48, GR58, GR59, GR60, GR61, GR62*
2	*GR12, GR13, GR24, GR25, GR26, GR27, GR28, GR29, GR30, GR31, GR32, GR33, GR34, GR35, GR36, GR37, GR38, GR40, GR47, GR64, GR65, GR68*
3	*GR14, GR15, GR16, GR17, GR18, GR19, GR20, GR21, GR22, GR23, GR49, GR50, GR51, GR52, GR54, GR69*
4	*GR1, GR2, GR3, GR4, GR5, GR6, GR7, GR8, GR9, GR10, GR11, GR53, GR55, GR56, GR57, GR63, GR66, GR67*
5	*Group2, Group3*
6	*Group1, Group5*
7	*Group4, Group6*

SNMPs	***B. mori , O. furnacalis***
1	*BmorSNMP1,BmorSNMP2*
2	*OfurSNMP1, OfurSNMP2*
3	*BmorSNMP1, OfurSNMP1*
4	*BmorSNMP2, OfurSNMP2*
5	*BmorSNMP1,BmorSNMP2, OfurSNMP1, OfurSNMP2*
ODEs	***S. inferens***
1	*CXE2, CXE6, CXE12, CXE14, CXE28*
2	*CXE1, CXE18, CXE20*
3	*Group1, Group2*
4	*CXE5, CXE9, CXE11, CXE13, CXE16, CXE30*
5	*CXE3, CXE10, CXE26*
6	*Group3, Group4, Group5,CXE19*
7	*AD1, AD6*
8	*AD2, AD3, AD4, AD5*
9	*AOX1, AOX2, AOX3*

Table 2-2. Primers used in qRT-PCR.

Name	Primers
RpS3	CAGCTCCCATAGCAATCATGG/CCACGGAAGCATGATCTTTACC
Actin	CCGTCCTCCTGACCGAGGCTC/GGTGTGGGAGACACCATCTCCG
NADH	GCTGAAGGTGAGAGAGAATTAG/CGAGGTAATGTTCCTCGAACTC
OfurOR1	GTGCTGTTCCTGCTCTACAAC/GCTGAACGTTCGCAAGAACATG
OfurOrco	GCTCATCAGTGATGGAAGCAG/GCACCAAGTACAGAAGCGAAC
OfurOR3	TTGGTACTCAGAGCGAGACCC/GGTGAATGTTCGCAGTAGCATG
OfurOR4	GATGTTAGGTGCTGAGACGGAG/TTAATCATTCATTGTTTGTAGG
OfurOR5a	GGATTTACAGATGAAGTTTCGGT/GACCGTATATGAGTACAGTCATA
OfurOR5b	GGATTTACGGATGAACTTTCGGC/GACCGTATATGAGTAAAGTCAGT
OfurOR6	TGCAGTACTACGTTACGGACC/CAGTCCTAATGCCTTGAGACTG
OfurOR7	CCTTAGTCTTCGAACTGCTAGG/TAGCAATCATGGTCCTCGAGC
OfurOR8	GAGATGTTGGGTTCAGAGACTG/TCTTCAATATCCCGGTCATGG
OfurOR9	CAGAGGATGATGGATGCGTGC/TTACGCCATCATTGACCGCAG
OfurOR10	CGTACAGTGCCGATTGGATAC/CAGAAGCGTGAAGAACGAGTAC
OfurOR11	GGCTTCAATTTATGCCGGTGG/CACTGGTATGATATCAGCAGCC
OfurOR12	TTCTGTTGGCACAGCAACGAC/CACTTTGCTGATTCGCAGCTG
OfurOR13	ATTGCTGGCACAGCAACGACG/GCCACAGTGAGCTTGGTGAAC
OfurOR14	GAGTAGGTGAAGCAGTGTACTG/GAGACGTAGCAAGAGCGTCAATG
OfurOR15	GGACTTGTTGAAGAGGAGTCAG/CTCGTGGTTGACATGAACGTG
OfurOR16	GAGTGATGGATGCAAGCAAGGC/GCTGAAGCTCAACGTGGTGAC
OfurOR17	TAGCTATGGACTGCTGGACTG/CATGAGGCATTCGAAACTCAGC
OfurOR18	TTATTATACAGGCGGACCGCG/CACGACAAAAGTGTGGAGATCC
OfurOR19	CTCATCGTTTGCTACTGCAGTG/ACCATCGTAAATGTGGCTTGC
OfurOR20	GAAAGTACCCTAGTGAGCTACGG/CTGCAGCTTAATCGCAGGATC
OfurOR21	ACAGTAGAGAGCGACCGCATG/CAAAGGTGTCAAGTGAGAGCG
OfurOR22	GGCACAGTAACGAAGCTTTAG/GAGTGTAGTAGCTGTACGACC
OfurOR23	GTGGCCATGCTGCAGATTTAC/GACCACGACGTGCTAATAATC
OfurOR24	CATTAGAAGCAGCTCGCATCG/TACGCTGCCTTCATTATCGCG
OfurOR25	CTCTTCATGAGCTTGCTGCAAG/GCTCTCAAGTTGACATCTGCG
OfurOR26	ATCGCTGCTATGCTACTTCGG/GTAGTGAAGGCCGTCAAGTTC
OfurOR27	CGCTAGCAACTATGGAACAGAC/GGTTCCAGCAAGACAATGGTG
OfurOR28	CTGAAGTGCGTTTGTGAGAAC/ATCGTGTACATGGAATAAGCC
OfurOR29	ACTGCAGTTTATGTGCGCGAC/CACCGAAATGAATGGGCCTGC
OfurOR30	GAGTTAACTGCTACTAGCGAAG/ACGAACGTCTGCCTAGACATG

OfurOR31	TCGACTGTGAGCAGTCAAGTG/TCAATCTTCTCTTTGGAGCAC
OfurOR32	TGGAGCTTAGCTCTATTGAAC/TTACTCTCTCTTGTGCGTTGC
OfurOR33	ACAAGTCGATAATGAGTGCGC/ATCCTCCAGAACGGACATGAC
OfurOR34	GAGTGGCAGATGCTTTGTATA/CTATGGGTTATAAGTATTGAG
OfurOR35	GAACTGATTTGGAAGAGCACTGC/GACTGCGAATGCTTTGTAAGACC
OfurOR36	TTGACGTTCGTCGCGAGTATG/GAAGGCCTTCATGACAATCGG
OfurOR37	TATGATAGCCGGTTCAGCGTAC/CCTTCCACTTGCTGCAGCAATG
OfurOR38	CATCACTATCGAGGCAGCAAG/AACGGAGTATGCTGATTTCACG
OfurOR39	AGCGAGAGCGATCAGGTGTGC/GCGAATGTAGTAAGCGAGATGG
OfurOR40	CGTATCAGGCTTCACTGTTAC/GAGAATAAGTTCCCTTCAAGAC
OfurOR41	AGTTTCACATCTGTCGTCCAC/GAACAGGTTGAAAGCGGTGATG
OfurOR42	CATATCGCTAGCAGCATACGAG/TTAATTGACGACGTCTCGGAG
OfurOR43	AAATTGCTGACACGATGGCGC/AATGAAGAGCGTAGTTGACGC
OfurOR44	ACGATGGTGGCACAGCTGTAC/ACAAAGGTAGGCCTCGACAGG
OfurOR45	GATGCGGCATACAATAGTAAATG/CTATTCTGGAGGCTTATAAACCG
OfurOR46	GCGTGTTGGGAGATTAGGTTC/GCCTGTCTCAGCATGTTGAAC
OfurOR47	CGCTAATACAGCGATGGATGTC/GTCAGCCGAGTTCACGATTTC
OfurOR48	GACGGCTTCACCAATATTAAC/TCATTCATCATCATCGTCATC
OfurOR49	CACCTTCGACATCCTGTTCATG/TGATGGAGATTGGATGCTGCG
OfurOR50	CGATAATCTGCGAATTGCAGCG/GAGTGAAGAATGAGTAGGCCG
OfurOR51	TCGGGCTTAGTGTATCATCAGC/GCACTGCATCAGAATCACCATC
OfurOR52	GCCTAACCGATGCCATATACTC/GCAGTGCGAAGCAGATTGAAG
OfurOR53	GGAGCTATTACCTACGTGAAGC/TTAAGCGCAGGCTGCGTTCATG

2.3 Results

2.3.1 RNA sequencing and screening of candidate odorant receptors

From multiplexed sequencing with Illumina MiSeq for six cDNA libraries, more than 12 million reads were obtained, which consisted of 5,852,653 reads from female antennae and 6,167,215 reads from male antennae (**Fig. 2-2**). The average length of reads was approximately 160 bp (**Table 2-3**). All reads in a total number of 2 billion bases were pooled together to be assembled into 60,399 contigs, from which 24,629 open reading frame sequences (ORFs) were extracted (**Table 2-4**). In the first screening, 244 and 243 sequences were obtained from the homology searches against *B. mori* odorant receptors using BLASTp and PSI-BLAST, respectively. These sequences included groups of variants that were identical in their middle section but different from each other in the length of the two termini. Such variants were probably generated by sequencing errors that truncated the deduced ORF. For further analysis, the longest one was selected as a representative sequence of each group. The robustness of this method was confirmed by comparisons with the sequences of previously identified pheromone receptors (see below). After removing duplications between the two screening results (BLASTp and PSI-BLAST), 134 candidates remained. From the second screening using SOSUI and TMHMM, 117 sequences were found to contain transmembrane domains. These sequences were finally screened against the NCBI non-redundant protein database, of which 52 had an insect odorant receptor as the top-hit homolog (**Table 2-5**). Seven of the nine previously identified pheromone receptors were found in the candidates, with the exception of *OfurOR1*. Although two sequences were reported for *OfurOR5* (*OfurOR5a* and *OfurOR5b*), only one sequence was found in our candidates,

which was slightly different from either. In the other cases, the previously identified receptors and the corresponding candidates were completely identical at the amino acid level but with some differences at the nucleotide level. The 45 newly identified receptors were named from *OfurOR9* to *OfurOR53*.

Table 2-3. Summary of sequencing results.

Reads	Male			Female		
	Repeat1	**Repeat2**	**Repeat3**	**Repeat1**	**Repeat2**	**Repeat3**
Total number	1837797	1896765	2118091	2409242	1503179	2254794
Total bases (bp)	300350405	309419147	347769463	390381207	250725980	372529795
Median length (bp)	156	162	160	157	157	158
Q30 percentage	96.80%	96.61%	96.57%	96.67%	96.41%	96.72%

Table 2-4. Summary of assembly results.

	Contigs	**ORFs**
Total number	60399	24629
Total length (bp)	65431963	29858679
Mean length (bp)	1083	1212
Median length (bp)	617	921
N50 length (bp)	1865	1578

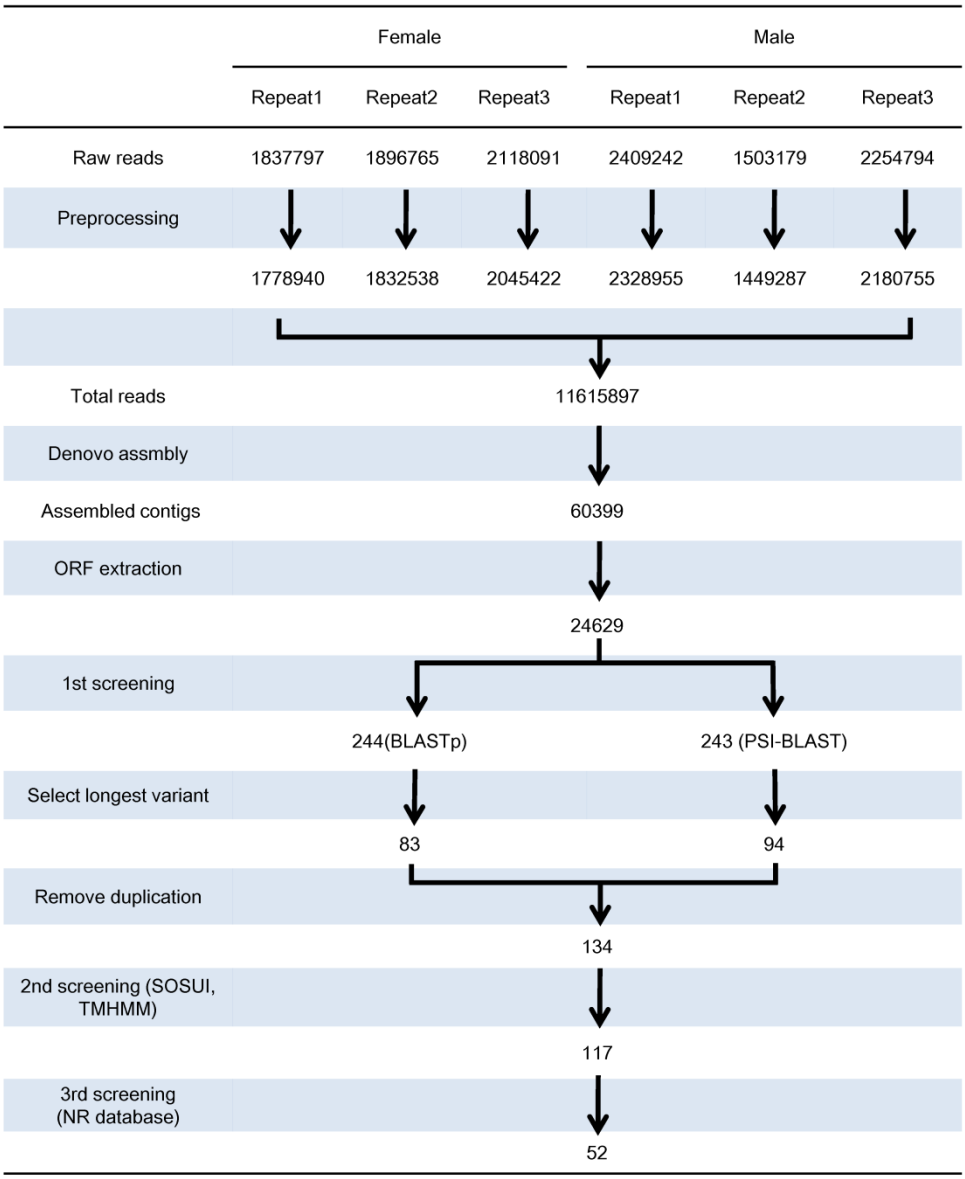


Figure 2-2. Results of sequence data analysis for odorant receptors. The numbers of reads or contigs at each step are indicated. See text for detailed explanation.

2.3.2 Expression levels of the candidate odorant receptors estimated by read mapping

To estimate the expression level of the candidate odorant receptors in males and females, the reads were mapped onto the ORF sequences of the candidate receptors. Because *OfurOR1* was not found in our candidates, its sequence was obtained from the database. Sequences for *OfurOR5a* and *OfurOR5b* were also obtained from the database and treated as independent receptors in the mapping. As expected, *OfurOrco* was expressed at the highest level in both male and female antennae (**Table 2-5, Fig. 2-3A**). Most of the previously identified pheromone receptors (*OfurOR3*, *4*, *5a*, *5b*, *6*, *7*, and *8*) showed male-specific expression, which is consistent with previous studies (Miura et al, 2009 and 2010; Wanner et al, 2010). Among these, *OfurOR4*, the receptor for the major pheromone component in *O. furnacalis*, was expressed at the highest level. *OfurOR7* was expressed not only in males but also in females at an intermediate level. Surprisingly, but consistently with the results of the candidate screening, the number of reads mapped onto *OfurOR1* was very low, suggesting that it was not expressed in our samples. None of the 45 novel candidate receptors showed strongly male-biased expression as observed with the previously identified pheromone receptors. Because the read counts were normalized by the total read number, and a large part of the reads were mapped onto the pheromone receptors in males, the RPKM values for the other receptors tended to be higher in females. Nevertheless, *OfurOR15*, *39*, *52*, and *53* should be recognized as female-biased receptors. In particular, *OfurOR15* and *OfurOR39* were expressed at the next highest levels after *OfurOrco* in female antennae.

Table 2-5. List of candidate odorant receptors in *O. furnacalis.*

Name	Accession Number	aa length	RPKM		RT-PCR**		
			Male	Female	Male	Female	Thorax
OfurOR1	AB467327*	425	0.29	0.00	0.000	0.000	0.0000
OfurOrco	LC002697	474	678.43	449.04	0.366	0.301	0.0000
OfurOR3	LC002698	426	61.07	0.15	0.150	0.000	0.0000
OfurOR4	LC002699	423	376.26	0.57	0.383	0.000	0.0000
OfurOR5a	AB508302*	408	61.85	1.21	0.199	0.000	0.0000
OfurOR5b	AB508303*	408	45.79	1.10	0.234	0.001	0.0000
OfurOR6	LC002700	422	37.69	1.01	0.168	0.002	0.0000
OfurOR7	LC002701	448	21.43	4.27	0.111	0.014	0.0000
OfurOR8	LC002702	438	44.57	0.01	0.162	0.001	0.0000
OfurOR9	LC002703	324	27.69	33.48	0.104	0.126	0.0001
OfurOR10	LC002704	404	17.35	31.60	0.046	0.066	0.0000
OfurOR11	LC002705	398	15.75	16.22	0.027	0.021	0.0000
OfurOR12	LC002706	407	13.95	16.74	0.036	0.027	0.0000
OfurOR13	LC002707	425	12.75	19.00	0.022	0.019	0.0000
OfurOR14	LC002708	424	12.32	23.11	0.027	0.038	0.0000
OfurOR15	LC002709	423	12.11	88.09	0.034	0.210	0.0000
OfurOR16	LC002710	422	10.16	22.84	0.019	0.023	0.0000
OfurOR17	LC002711	239	10.08	6.46	0.034	0.043	0.0000
OfurOR18	LC002712	471	9.72	14.54	0.012	0.012	0.0001
OfurOR19	LC002713	422	8.93	15.96	0.039	0.047	0.0000
OfurOR20	LC002714	411	8.79	12.19	0.018	0.018	0.0000
OfurOR21	LC002715	363	8.37	8.91	0.003	0.003	0.0000
OfurOR22	LC002716	409	8.27	14.28	0.022	0.028	0.0000
OfurOR23	LC002717	404	7.71	14.27	0.023	0.039	0.0000
OfurOR24	LC002718	448	6.44	8.52	0.014	0.017	0.0005
OfurOR25	LC002719	418	6.14	8.19	0.004	0.002	0.0000
OfurOR26	LC002720	433	6.01	7.07	0.011	0.015	0.0000
OfurOR27	LC002721	402	5.94	6.89	0.012	0.020	0.0001
OfurOR28	LC002722	420	5.93	11.78	0.022	0.022	0.0001
OfurOR29	LC002723	441	5.71	12.74	0.008	0.014	0.0000
OfurOR30	LC002724	402	5.58	10.89	0.010	0.014	0.0000
OfurOR31	LC002725	396	5.43	6.52	0.008	0.009	0.0001
OfurOR32	LC002726	412	5.11	14.27	0.004	0.006	0.0000

OfurOR33	LC002727	410	5.06	7.58	0.023	0.028	0.0001
OfurOR34	LC002728	439	5.01	8.82	0.008	0.008	0.0000
OfurOR35	LC002729	430	4.76	7.34	0.009	0.011	0.0003
OfurOR36	LC002730	398	4.72	5.60	0.001	0.001	0.0000
OfurOR37	LC002731	390	4.69	5.36	0.008	0.008	0.0000
OfurOR38	LC002732	415	4.59	7.55	0.010	0.009	0.0001
OfurOR39	LC002733	386	4.46	44.89	0.010	0.064	0.0000
OfurOR40	LC002734	432	4.44	10.99	0.005	0.006	0.0000
OfurOR41	LC002735	421	4.24	13.06	0.008	0.025	0.0004
OfurOR42	LC002736	346	4.16	7.13	0.010	0.015	0.0000
OfurOR43	LC002737	198	4.00	8.34	0.008	0.010	0.0003
OfurOR44	LC002738	436	3.54	4.33	0.003	0.006	0.0000
OfurOR45	LC002739	407	3.51	6.69	0.013	0.013	0.0000
OfurOR46	LC002740	431	3.05	8.02	0.015	0.016	0.0006
OfurOR47	LC002741	103	2.86	1.63	0.002	0.003	0.0000
OfurOR48	LC002742	265	2.65	7.15	0.037	0.015	0.0001
OfurOR49	LC002743	361	2.49	5.72	0.006	0.010	0.0002
OfurOR50	LC002744	355	1.88	4.52	0.003	0.006	0.0000
OfurOR51	LC002745	380	1.56	2.86	0.006	0.008	0.0000
OfurOR52	LC002746	409	0.79	6.31	0.002	0.003	0.0000
OfurOR53	LC002747	407	0.00	7.55	0.000	0.018	0.0000

*: Reported in the previous paper (Miura et al, 2010).

**: Relative expression level to the internal control.

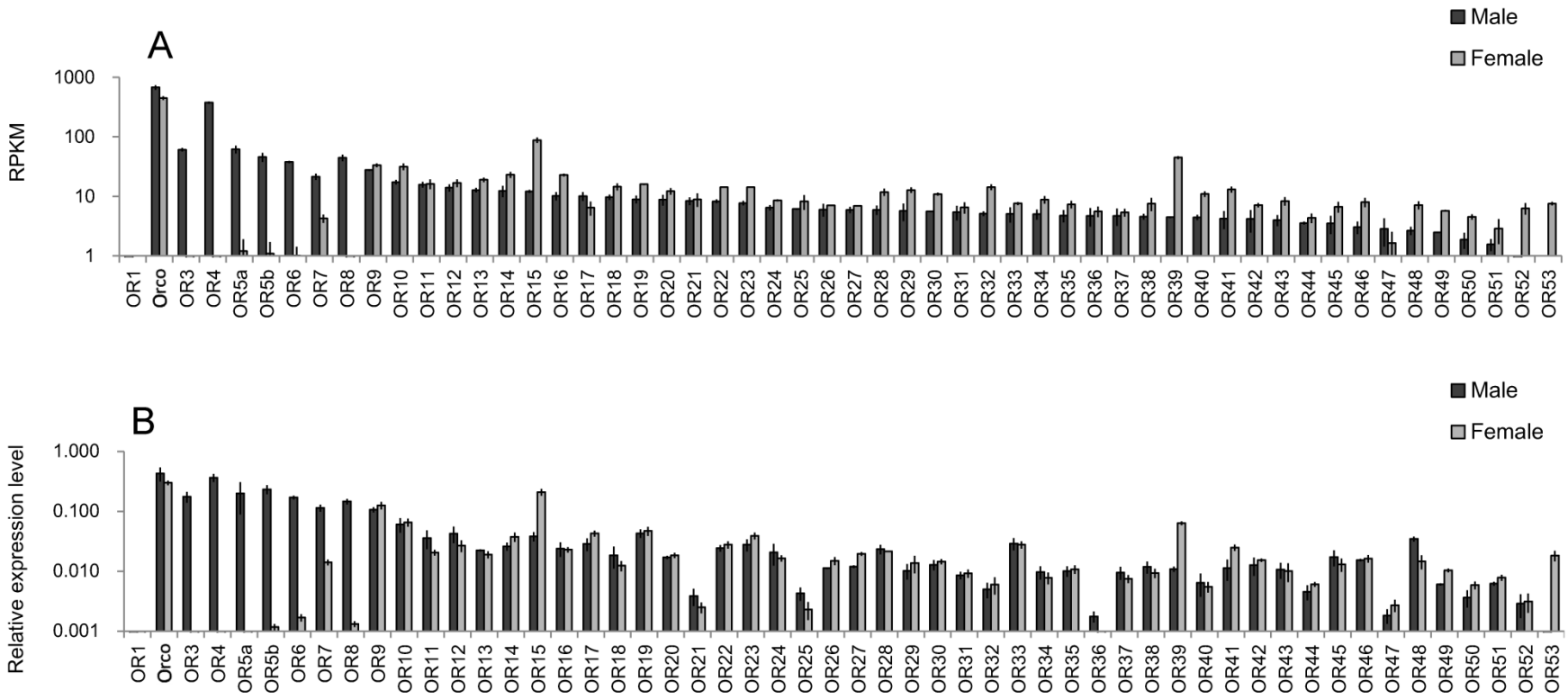


Figure 2-3. Expression level of candidate odorant receptors in adult antennae. (**A**) Estimated by read mapping. (**B**) Estimated by qRT-PCR. Error bars represent standard error calculated from the results of three biological replicates.

2.3.3 Expression levels of the candidate odorant receptors confirmed by qRT-PCR

To confirm the expression levels of the candidate odorant receptors, qRT-PCR was carried out using the independently prepared cDNA libraries. The primers for qRT-PCR were designed to specifically recognize the sequence at the 3' end of each candidate (**Table 2-2**). Thorax cDNA libraries were used as the negative control, and no expression was detected for any of the odorant receptors (**Table 2-5**). In the antennae, the results were generally consistent with those of the read mapping but with some exceptions (**Fig. 2-3B**). The inconsistency with the results of read mapping was probably caused by high sequence similarity between two receptors. Because we used -a option in the bowtie2 mapping, single reads derived from the high-homology regions were mapped to both of the receptors with a 0.5 count each, resulting in a similar RPKM value in both receptors. Such cases were likely in *OfurOR21*, *25*, and *36* that were expressed at lower levels than estimated by the read mapping. Female-biased expression was confirmed for *OfurOR15*, *39*, and *53*. In particular, *OfurOR53* was highly female-specific, suggesting its dedicated role in females. Expression of *OfurOR7* in females was also confirmed. Because the male-female expression ratio was more accurately estimated by qRT-PCR than read mapping, the difference between males and females was smaller for most receptors than that estimated by read mapping. On the other hand, qRT-PCR is less accurate compared with read mapping for comparisons between different receptors, because independent primer pairs may not always give identical amplification efficiency by PCR. In this sense, the order between the receptors determined by their expression level should be more accurately represented by the results of read mapping than that of qRT-PCR, although the above-mentioned mis-assignment problem between similar sequences should be taken into account.

2.3.4 Phylogenetic analysis for odorant receptors

Phylogenetic relationships between *OfurORs* and *BmorORs*, as well as with odorant receptors in *C. pomonella* are shown in **Figure 2-4**. As expected, *Orco* was highly conserved among the three species. All the previously identified pheromone receptors of *O. furnacalis* formed a single clade with other pheromone receptors from *B. mori* and *C. pomonella*. Within this clade, however, receptors from the same species tended to form subclusters, suggesting that pheromone receptors have undergone species-specific duplication events. The female-specific receptor *OfurOR53* formed a clade with *BmorOR30* and *CpomOR30*, among which *BmorOR30* was reported to exhibit female-specific expression (Tanaka et al, 2009; Wanner et al, 2007), whereas *CpomOR30* was not (Bengtsson et al, 2012). The other two female-biased receptors, *OfurOR15* and *OfurOR39*, belonged to independent clades. *OfurOR15* formed a clade with *OfurOR28*, *OfurOR41*, *BmorOR14*, *CpomOR14*, and *CpomOR20*. Among these, *OfurOR41* showed slightly female-biased expression (**Fig. 2-3B**), but the others were expressed both in males and females (Tanaka et al, 2009; Wanner et al, 2007; Bengtsson et al, 2012). *OfurOR39* formed a clade with *OfurOR51*, *BmorOR50*, *BmorOR51*, and *CpomOR43*. None of these were reported to be female biased (Tanaka et al, 2009; Wanner et al, 2007; Bengtsson et al, 2012).

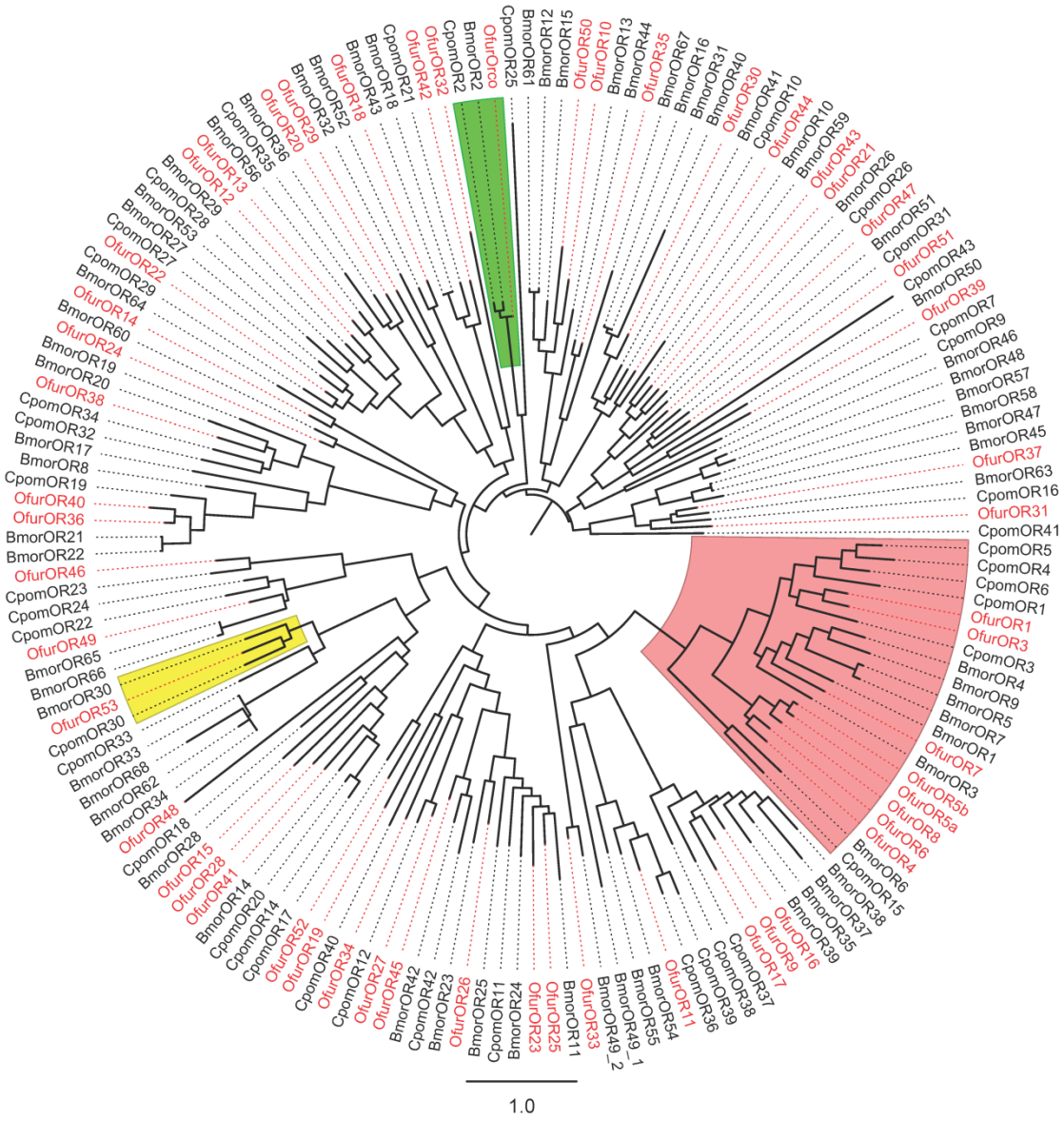


Figure 2-4. Phylogenetic relationship of *O. furnacalis* odorant receptors (*OfurORs*) with those of *B. mori* (*BmorORs*) and *C. pomonella* (*CpomORs*). The tree was constructed by the maximum likelihood method using RAxML and visualized using FigTree. *OfurORs* are indicated in red. Green, pink, and yellow shading indicates the clades of *Orco*, pheromone receptors, and *OfurOR53*, respectively.

2.3.5 Identification of other genes involved in the odorant perception

We also identified candidate genes of 21 IRs, 5 GRs, 2 SNMPs and 26 ODEs (**Table 2-6**). All the genes were novel in *O. furnacalis* except for 2 SNMPs (Allen and Wanner, 2011). The phylogenetic relationships between *OfurIRs*, *BmorIRs*, and *CpomIRs* are shown in **Figure 2-5**. ODEs were divided into three families, including eight aldehyde oxidases (*OfurAOX1* to *OfurAOX8*), fifteen carboxylesterase (*OfurCXE1* to *OfurCXE15*) and three alcohol dehydrogenase (*OfurAD1* to *OfurAD3*). Most of the identified genes were full length. However, all of the GR genes were partial, probably due to their low expression levels in the antennae.

Table 2-6. List of other candidate genes involved in olfactory perception in *O. furnacalis*.

Name	Accession Number	aa length	RPKM	
			Male	Female
OfuriGluR1	LC017780	923	5.67	7.99
OfuriGluR2	LC017781	900	5.93	6.37
OfurIR8a	LC017782	902	37.14	50.82
OfurIR21a	LC017783	849	26.88	36.66
OfurIR25a	LC017784	942	71.03	68.03
OfurIR40a	LC017785	709	3.26	5.41
OfurIR41a	LC017786	596	12.15	17.01
OfurIR64a	LC017787	606	5.79	8.17
OfurIR68a	LC017788	340	4.34	6.67
OfurIR75	LC017789	626	13.28	16
OfurIR75d	LC017790	274	2.29	3.13
OfurIR75p1	LC017791	630	4.4	10.33
OfurIR75p2	LC017792	609	5.42	10.47
OfurIR75p3	LC017793	639	3.3	0
OfurIR75q2	LC017794	637	6.99	11.58
OfurIR76b	LC017795	547	27.01	46.73
OfurIR87a	LC017796	654	5.25	7.12
OfurIR93a	LC017797	890	6.64	7.06
OfurIR1	LC017798	358	2.02	4.12
OfurIR2	LC017799	357	9.06	18.22
OfurIR3	LC017800	178	1.11	1.65
OfurGR1	LC017775	140	1.62	3.25
OfurGR2	LC017776	130	2.36	0
OfurGR3	LC017777	121	1.05	1.47
OfurGR4	LC017778	100	0	2.5
OfurGR5	LC017779	194	2.36	3.38
OfurSNMP1	LC017801	528	919.31	389.54
OfurSNMP2	LC017802	523	1352.03	1317.27
OfurAOX1	LC017752	1275	134.36	161.16
OfurAOX2	LC017753	1279	189.54	297.3
OfurAOX3	LC017754	1280	19.36	15.92
OfurAOX4	LC017755	766	6.35	10.49
OfurAOX5	LC017756	593	3.36	6.88

OfurAOX6	LC017757	1268	13.28	13.98
OfurAOX7	LC017758	778	7.68	5.12
OfurAOX8	LC017759	378	2.44	4.03
OfurCXE1	LC017760	560	5.93	5.25
OfurCXE2	LC017761	541	135.42	116.35
OfurCXE3	LC017762	532	11.35	14.77
OfurCXE4	LC017763	559	66.51	34.96
OfurCXE5	LC017764	566	15.2	19.43
OfurCXE6	LC017765	511	13.59	15.4
OfurCXE7	LC017766	317	82.91	68.76
OfurCXE8	LC017767	566	354.93	237.89
OfurCXE9	LC017768	544	23.12	30.35
OfurCXE10	LC017769	542	23.89	28.67
OfurCXE11	LC017770	527	22.83	32.79
OfurCXE12	LC017771	519	17.16	30.97
OfurCXE13	LC017772	511	44.64	61.26
OfurCXE14	LC017773	562	175.11	229.03
OfurCXE15	LC017774	515	2.32	5.19
OfurAD1	LC017749	325	32.41	43.19
OfurAD2	LC017750	356	5.34	8.79
OfurAD3	LC017751	365	46.62	53.06

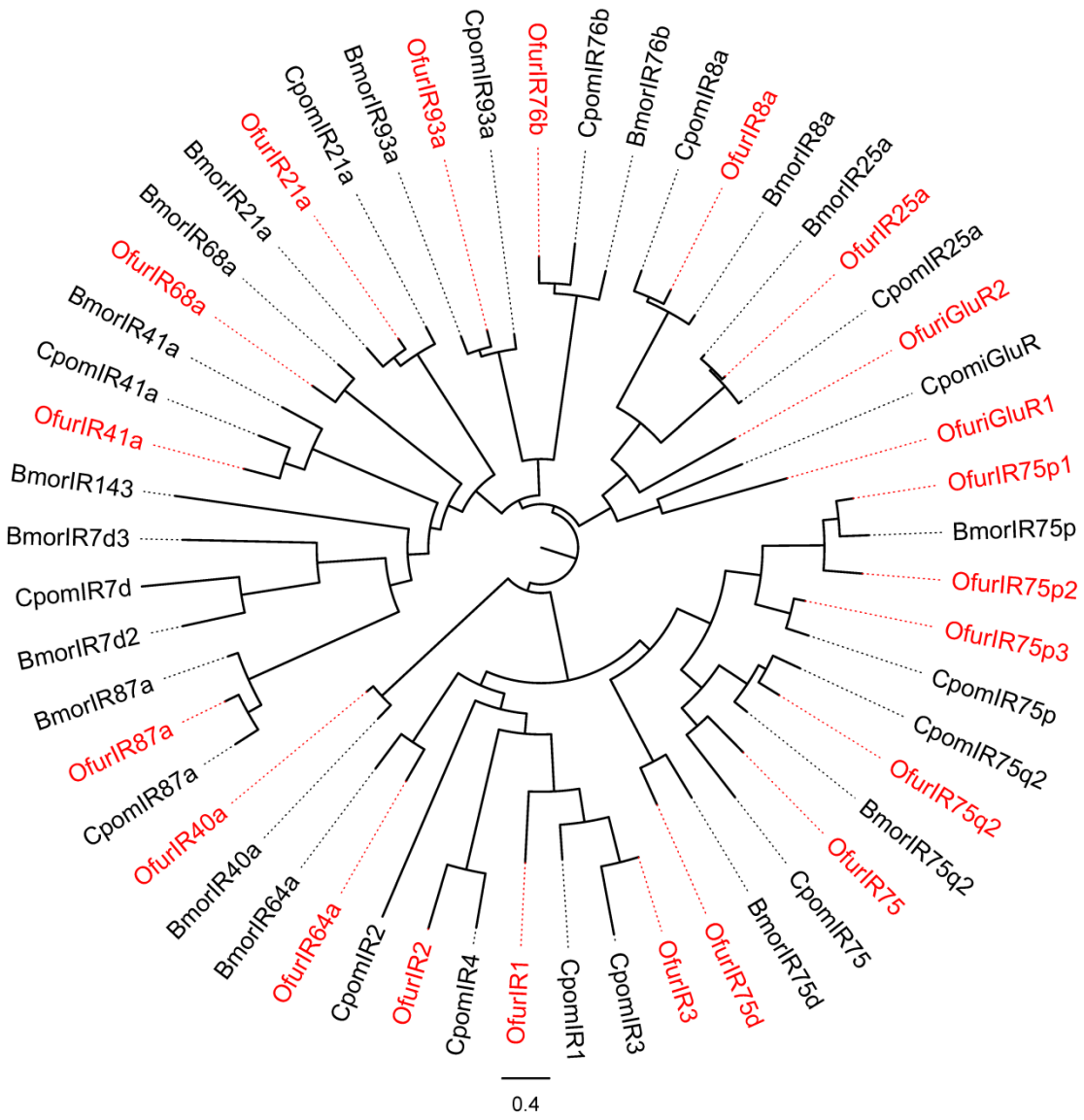


Figure 2-5. Phylogenetic relationship of *O. furnacalis* ionotropic receptors (*OfurIRs*) with those of *B. mori* (*BmorIRs*) and *C. pomonella* (*CpomIRs*). The tree was constructed by the maximum likelihood method using RAxML and visualized using FigTree. *OfurIRs* are indicated in red.

2.4 Discussion

2.4.1 Pheromone receptors

In the previous study, pheromone receptors in *O. furnacalis* were cloned by degenerate PCR (Miura et al, 2010). For this reason, the 5’ and 3’ terminal sequences of the ORFs were not known. In the present study, we identified complete ORF sequences for seven of the nine previously identified pheromone receptors. On the other hand, *OfurOR1* was not found in our RNA-seq analysis. It was also not detected in the independent qRT-PCR analysis, indicating that *OfurOR1* was not expressed in our sample. This might be due to intraspecies polymorphism because our samples and those used in the previous studies were derived from different localities in Japan (Miura et al, 2010). None of the 45 novel receptors found in this study showed male-biased expression as observed in the previously identified pheromone receptors. The previously identified pheromone receptors were structurally distinct from the other receptors; they formed a single clade in the phylogenetic analysis. Thus, it is likely that there are no additional pheromone receptors in *O. furnacalis* other than the already identified ones. However, the presence of some other receptors that incidentally respond to pheromone components was not excluded. Identification of novel odorant receptors in *O. furnacalis* provides the opportunity to experimentally examine this possibility.

2.4.2 Phylogenetic relationship of odorant receptors with sexually biased expression

Genome wide analysis of the expression pattern of odorant receptors has been carried out in several lepidopteran species including *B. mori* (Tanaka et al, 2009; Wanner et al, 2007), *Manduca sexta* (Grosse-Wilde et al, 2011), *Cydia pomonella*

(Bengtsson et al, 2012), *Helicoverpa armigera* (Liu et al, 2012a), and *Spodoptera littoralis* (Poivet et al, 2013). In each species, receptors with sex-specific expression have been identified. Some of these receptors are phylogenetically close to each other. The most significant example is the pheromone receptor group, which contains nine receptors from three species (*BmorOR1, 3, 4, 5, 6*; *HarmOR14, 15*; *SlitOR6, 13*) that were male specific (Tanaka et al, 2009; Liu et al, 2012a; Poivet et al, 2013). The previously identified pheromone receptors in *O. furnacalis* belonged to this group, and most of them were male specific (Miura et al, 2010). However, not all of the members were male biased. Seven receptors from three species (*CpomOR3, 5*; *HarmOR1, 2, 11*; *SlitOR11, 16*) were equally expressed in males and females (Bengtsson et al, 2012; Liu et al, 2012a; Poivet et al, 2013). Furthermore, *CpomOR15* was shown to be female specific (Bengtsson et al, 2012). Another example is a group of receptors including *OfurOR53* and *BmOR30*. Although these two receptors were specifically expressed in female antennae (Tanaka et al, 2009; Wanner et al, 2007; this study), orthologous receptors in other species (*CpomOR30* and *SlitOR30*) were also expressed in the male antennae (Bengtsson et al, 2012; Poivet et al, 2013). These examples indicated that sexually biased expression is under the influence of phylogenetic constraint to some extent, but it also evolves dynamically from sex-specific expression to sex-independent expression and vice versa. Nevertheless, it should be noted that in most of the previous studies, the expression levels were determined by non-quantitative methods, leaving the possibility that the difference between sexes was over- or under-estimated (Tanaka et al, 2009; Bengtsson et al, 2012; Grosse-Wilde et al, 2001; Liu et al, 2012a; Poivet et al, 2013). Quantitative analysis of the expression level is necessary to gain insight into the evolutionary pattern of sexually biased expression of odorant receptors.

2.4.3 Biological function of female-biased receptors in *O. furnacalis*

In this study, the expression levels of the all receptors were estimated quantitatively by two independent methods, which demonstrated that *OfurOR53*, *15*, and *39* had female-biased expression. Importantly, the latter two were the receptors with the highest expression level in female antennae next to *OfurOrco*. One possible function of these receptors is the perception of male sex pheromone, which was reported to be required for acceptance of mating by females in *O. nubilalis* (Leary et al, 2012). The *OfurOR7* is also a candidate for the male pheromone receptor. It belongs to the pheromone receptor group, and it was also expressed in the female antennae. Another possible function of the female-biased receptors is to recognize host-plant volatiles. Finding an appropriate host plant is crucial for reproduction in the herbivorous lepidopteran insects. Odorant receptors involved in host-plant detection would serve as a potential target for novel pest control techniques. In this regard, *OfurOR15* and *OfurOR39*, the receptors with the highest expression levels in female antennae, should be considered as the primary candidates for further characterization of their molecular function.

2.4.4 Repertoire of odorant receptors in *O. furnacalis*

Although an intensive analysis of the antennal transcriptome was conducted in this study, other tissues were not investigated. Therefore, odorant receptors not expressed in the antennae were not included in our analysis. Furthermore, receptors with extremely low expression levels may not have been identified. In fact, the ORF sequences appeared to be incomplete for a few receptors with low expression levels

(**Table 2-5**). Two receptors with a similar sequence, such as recently duplicated pairs, were indistinguishable in our analysis, as seen in the case of *OfurOR5a* and *OfurOR5b*. Finally, extremely divergent receptors that were not similar to any of the other insect odorant receptors may not be identified in our analysis, although the candidates excluded at the third screening (homology search against the NCBI nr database) were significantly similar to non-odorant-receptor proteins. These limitations mean that our method is conservative, and whole genome sequence analysis may identify additional odorant receptors in *O. furnacalis*. Nevertheless, our results provide a list of odorant receptors with significant expression in the antennae, thus they are considered to be biologically functional. Our present results will serve as a basis for studies to understand the evolution of the pheromone communication system, as well as for the development of novel control methods of agriculturally important pests.

Chapter 3

Targeted mutagenesis of *OfurOR4* and *OfurOrco* (*OR2*)

3.1 Introduction

Functional analysis of chemosensory genes is important for understanding the molecular mechanisms of chemosensation and the related behaviors in insects. However, the previous studies relied on ectopic expression systems such as *Xenopus* oocytes, leading to the lack of direct evidence for their *in vivo* function. Recently, several genome editing tools were developed for precise targeted mutagenesis, which made it possible to silence the target genes in non-model animals to analyze their function by comparisons of phenotypes between mutants and wild type. Currently, three tools were available for the targeted mutagenesis; zinc finger nucleases (ZFNs), transcription activator-like effector nuclease (TALEN), clustered regularly interspaced short palindromic repeats (CRISPR) and CRISPR-associated (CRISPR/Cas9) system. The principle for all these three tools is basically same. They consist of a DNA binding domain and an endonuclease domain, introducing double-stranded breaks at the target site resulting in indels by non-homologous-end-joining (NHEJ) that inactivate the target gene.

The CRISPR/Cas9 system is the newest tool for genome editing and it is comprised of an endonuclease Cas9 and the binding domain of a reconfigured single guided RNA (sgRNA) (Jinek et al, 2012; Makarova et al, 2011). The most advantageous point is that the endonuclease and binding domain was separated, and only the sgRNA need to be redesigned for different targets. However, off-target effect is not able to be pre-examined in most of the pest insects, for which the whole genome sequences are not available.

ZFN and TALEN use same endonuclease *Fok* I, which is a kind of

dimerisational enzyme that cleaves DNA only when it is paired, reducing off-target effect than Cas9. Disadvantages of ZFN and TALEN are that the *Fok* I was attached with a DNA binding domain adapted from zinc finger (ZF) and transcription activator-like effectors (TALEs), respectively, which makes the whole construct need to be redesigned for different targets (Urnov et al, 2010; Carroll, 2011; Schornack et al, 2006; Boch et al, 2009; Boch and Bonas, 2010). Because single unit of TALE recognizes and binds to a single nucleotide, the DNA binding domain of TALE can be freely assembled (Boch et al, 2009; Moscou and Bogdanove, 2009; Bogdanove and Voytas, 2011). In contrast, one ZF unit recognizes and binds to three nucleotides, and corresponding ZF units were not known for all combinations of nucleotides (Urnov et al, 2010; Carroll, 2011). Furthermore, it was shown that TALEN had higher efficiency to induce mutagenesis in species (Gupta et al, 2014; Ma et al, 2012).

Considering the smaller off-target effect and higher efficiency in moth species, TALEN would be a choice for the mutagenesis of chemosensory genes. To apply this tool in pest species, several issues remain to be addressed. First, the method of microinjection into early embryos is not established. For example, in the species that lay egg mass, individual embryos can not be separated from others. Therefore, specific strategy for preparation needs to be developed. Time schedule of egg collection also needs to be established according to the behavioral characteristics of the target species. Second, crossing and genetic management of many strains are required. Pairing with single individuals is necessary to isolate and establish mutant strains. This would be particularly difficult when pheromone receptors are silenced, because they are expected to be inactive for mating behavior. Finally, screening methods without genetic markers need to be established. Because the phenotype of the mutants for chemosensory genes

should be invisible, screening relies on efficient genotyping. Our experience coping with the above issues is limited, particularly in pest insect species. If TALEN can be applied in *O. furnacalis*, it would be a proof showing that this tool may be applied to other pest insects.

In this chapter, I established the application method of TALEN in *O. furnacalis*, using two odorant receptors, *OfurOR4* and *OfurOrco*, as target genes. They are the best characterized receptors in *O. furnacalis*. *OfurOR4* was suggested to have a major role in pheromone perception and to be responsible for evolution of the different pheromone preference between closely related species. Loss of *OfurOR4* gene function should result in specific defects in pheromone perception. *OfurOrco* is the co-receptor which is required for all the other odorant receptors. Loss of *OfurOrco* gene function should result in malfunction of olfactory system against broad range of odorants. Therefore, selecting these two target OR genes, I expected that the effectiveness of TALEN can be examined at phenotypic level once the mutant lines are successfully established. To achieve this end, various conditions and handling methods required for genome editing using TALEN was explored in this chapter.

3.2 Materials and Methods

3.2.1 Insect rearing

O. furnacalis were collected and cultured as described in chapter 2.

3.2.2 Design and RNA synthesis of TALEN

Odorant receptor sequences were obtained from Chapter 2. DNA binding domain of TALEN were designed considering the following issues: 1) 5' should have a T; 2) Spacer should be 11-18bp in length, in which 15 or 16bp is the best; 3) Sequences of binding area should have high specificity; 4) 5' should not contain more than 6consecutive A or T; 5) Less G in the sequence of binding domains; 6) Spacer should contain the restriction enzyme site. Due to the unspecific binding ability of TALEN unit NN (usually binding with G, but sometimes A), wobbles primers which G was taken place with R were used to check if there were other similar sequences in the *O. furnacalis* antennae transcriptome (**Table 3-1**). The pass through pairs of binding domains were used and designed as Truncated TAL Fok I from GeneArt® Custom Gene Synthesis (Life Technologies, Carlsbad, CA, USA). Two pairs of TALEN for *OfurOrco* and one pair of *OfurOR4* were provided in the entry vectors (**Fig. 3-1**). The RVD (repeat variable di-residue) of each repeat was designed following the general rule of TALEN recognition, which NG, HD, NI and NN recognize T, C, A and G respectively.

Table 3-1. Wobbles primers for TALEN.

Name	Primers
Orco-TAL-1-F	RATRACCAAARTRAAARC
Orco-TAL-1-R	RATRTTTRRCATCAAATC
Orco-TAL-2-F	CCRATTTRATRCCAAACA
Orco-TAL-2-R	RAATARRAAATRACCARC
OR4-TAL-1-F	CAAATATATTCRARRACA
OR4-TAL-1-R	CTCTRTRAATCCRAAARA

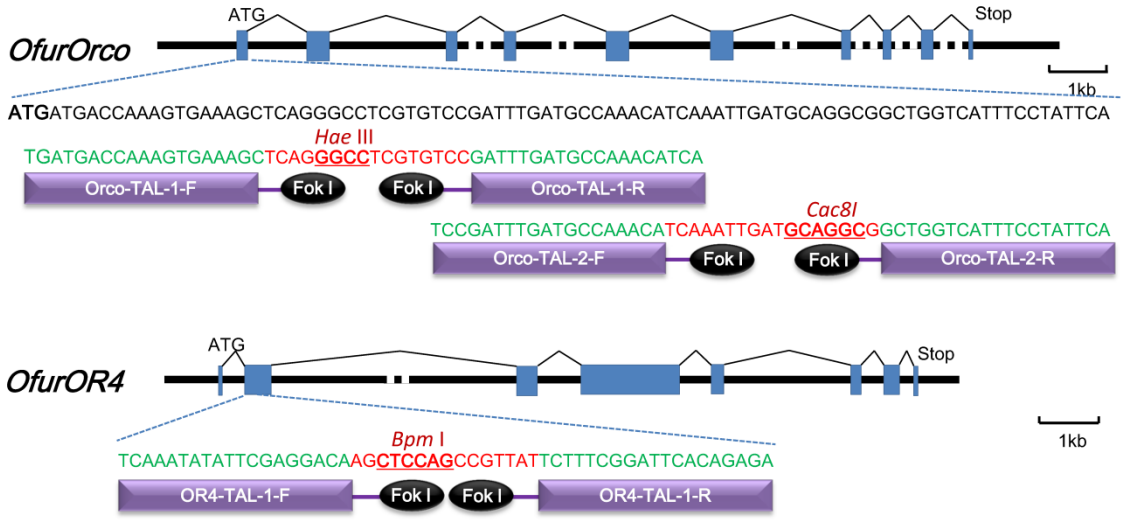


Figure 3-1. Design of TALEN from genomic sequences of *O. furnacalis*. Two pairs of TALEN for *OfurOrco* were designed in the first exon. One pair of TALEN for *OfurOR4* was designed in second exon. TALE binding domains were shown in green, spacers were in red. The underline indicates the restriction enzyme site for each pair of TALEN.

3.2.3 RNA syntheses for TALEN

The entry vectors were amplified using ECOS™ Competent *E. coli* DH5α （Nippon Gene, Fukuyama, Japan) and purified using QIAprep Spin Miniprep Kit (QIAGEN Inc., CA, USA) (**Fig. 3-2**). TALEN were recombined to a destination vector Gateway® pcDNA-DEST40 (Life Technologies, Carlsbad, CA, USA) using Gateway® LR Clonase® Enzyme mix (Life Technologies, Carlsbad, CA, USA). The destination vectors were amplified and purified using the same method described above and linearized by restriction enzyme *Pme*I (New England BioLabs Inc., Ipswich, MA, USA). The linearized DNA were purified with phenol/chloroform/isoamylalcohol (25:24:1) and transcribed into mRNA *in vitro* using mMESSAGE mMESSAGE® T7 ULTRA Transcription Kit (Ambion, Carlsbad, CA, USA). Synthesized mRNA was purified using MEGAclear™ Kit (Ambion, Carlsbad, CA, USA).

Table 3-2. Primers used in genotyping of *OfurOrco* and *OfurOR4.*

Name	Primers
OfurOrco-F	GTCAATCGATGGAGAGTCCAC
OfurOrco-R	GCGCAAGAGTGTTGTCATTCC
OfurOR4-F	AGGTCCGACAAAGTGATATGG
OfurOR4-R	TCACCATGGCCAAGAACGTTG

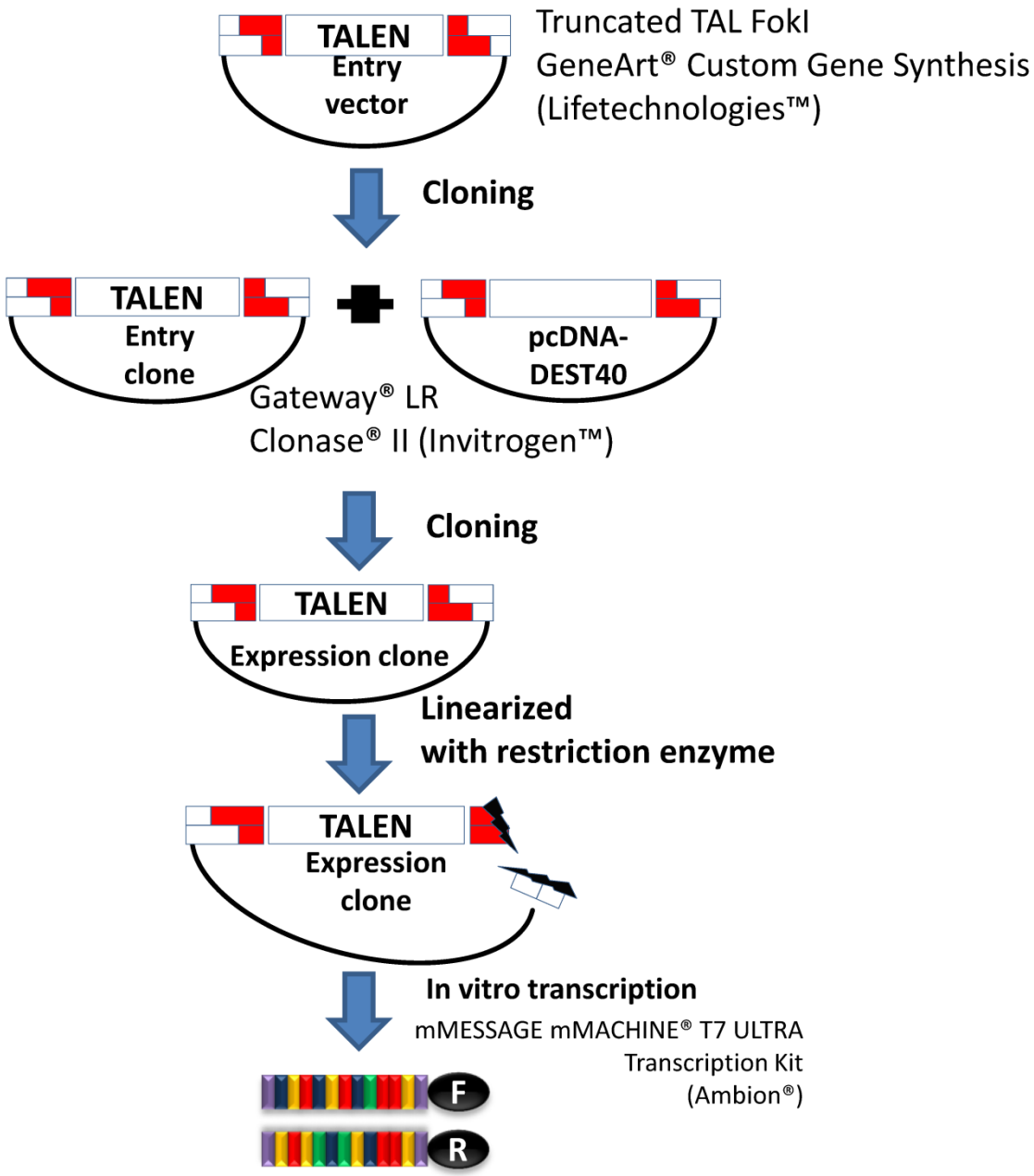


Figure 3-2. Schematic diagram of vector construction and *in vitro* transcription.

3.2.4 Preparation of eggs and setup for injection

Because adults of *O. furnacalis* lay eggs during night, the photoperiod was set reversed to ours (**Fig. 3-3A**). Dark period started from 6AM, and egg mass was collected at 10 AM. Injection was finished before 12AM to ensure that only the embryos at early stages of development were injected. The egg masses laid on a piece of plastic wrap were transferred to a cover slip that was immobilized with a small drop of distilled water on a SUPERFROST® micro glass slide (MATSUNAMI Inc., Osaka, Japan) (**Fig. 3-3B**). Only the eggs on the edge of egg mass were injected under an inverted microscope Primo Vert (ZEISS, Oberkochen, Germany), using microinjection needle Femtotips® II (Eppendorf, Hamburg, Germany). Needle was connected with a syringe using Grip head 0 for Universal Capillary holder, Adapter Femtotip and Pressure tube (CatNo.: 920007414, B2500887L, 920002081, Eppendorf, Hamburg, Germany). Solution filled in the needle was injected by manual application of air pressure using the connected syringe. Needle movement was controlled using Joystick Manipulator MN-151 attached on a manipulator mounting adaptor NZ-13 with a pipette holder (NARISHIGE, Tokyo, Japan). After injection, the cover slips were kept in a humid box with a paper infiltrated with distilled water and cultured at 23℃ as described below.

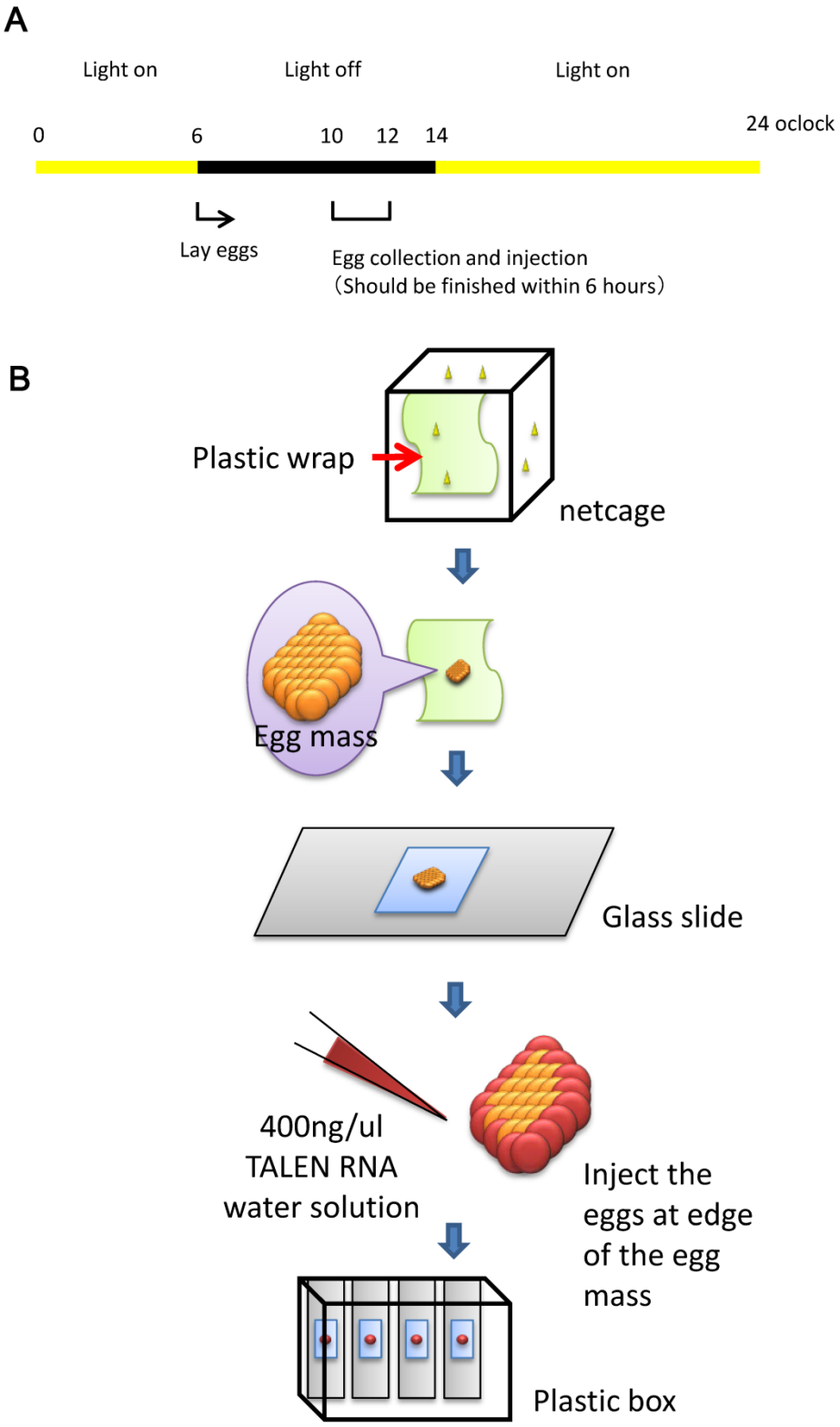


Figure 3-3. Time schedule (A) and workflow (B) for injection.

3.2.5 Selection of solvent to dissolve TALEN RNA

Distilled water and 0.2mM phosphate buffer (pH 7.5) were compared for better hatchability (**Fig. 3-4**). Distilled water and phosphate buffer were injected into the eggs in single egg masses. Internal eggs were left un-injected as control.

3.2.6 Microinjection with TALEN RNA against *OfurOrco* and *OfurOR4*

TALEN RNA was dissolved at a concentration of 400ng/ul (800ng/ul for a pair of TALENs) into distilled water, which showed better hatching rate than phosphate buffer in the experiment described above. Amount of injected solution was ca. 30-50nl per egg (**Fig. 3-5**). About 4 or 5 days after injection, a developing head of embryo was visible as a black spot. At this stage, non-injected embryos at the center of egg masses were killed using a needle. Injected eggs on a cover slip were transferred into a breeding jar with artificial diet. Egg masses were examined every day for their development, and the hatching rate was recorded. After pupation, collected pupae were sorted by sex.

3.2.7 Crossing scheme and screening strategy for *OfurOrco* mutants

Multiple males were required for crossing because of low copulation rate of the laboratory strain. Therefore, to trace the lineages of single injected individuals, only females of injected individuals were used for crossing (**Fig. 3-6A**). One injected female was paired with four wild type males in a plastic cup with wet cotton. After the females laid eggs or dead, whole body was stored at -20℃ for screening of somatic mutations. The G1 eggs laid by the females who contain the somatic mutations were cultured. At least eight individuals from each G1 families were examined for germline mutations. G1 adults from families with high germline mutation rates were single crossed within

family members. G1 parents of which laid eggs were genotyped for the mutation. If both of parents were heterozygous for the mutation, homozygotes were expected in the G2 generation.

3.2.8 Crossing scheme and screening strategy for *OfurOR4* mutants

Because *OfurOR4* is on the sex chromosome, all the G1 females were expected to be wild type (**Fig. 3-6B**). Homozygotes were expected in G3 generation.

3.2.9 Genotyping for mutations caused by TALEN

Whole body was homogenized with FastPrep 24 Instrument (MP Biomedicals, Santa Ana, California, USA), setting at model MP, speed 6 m/s for 60s, 3 times. Then the genomic DNA was isolated following the protocol of Quick Genomic DNA Prep for fly (**Fig. 3-7**). FastGene Gel/PCR Extraction Kit (Nippon Genetics, Tokyo, Japan) was use in the attached cleaning step. Primers designed from genomic sequence of *OfurOrco* and *OfurOR4* were used for PCR using TaKaRa Ex taq (TaKaRa, Shiga, Japan) (**Table 3-2**). The PCR products were digested using *Hae* III, *Cac8I* or *Bpm* I (New England BioLabs Inc., Ipswich, MA, USA) for *OfurOrco* or *OfurOR4*, respectively. Mutations were detected by electrophoresis on a 2% agarose gel in TAE buffer. DNA was visualized using ethidium bromide. Photographs were taken using ImageQuant 400 (GE healthcare, Buckinghamshire, United Kingdom).

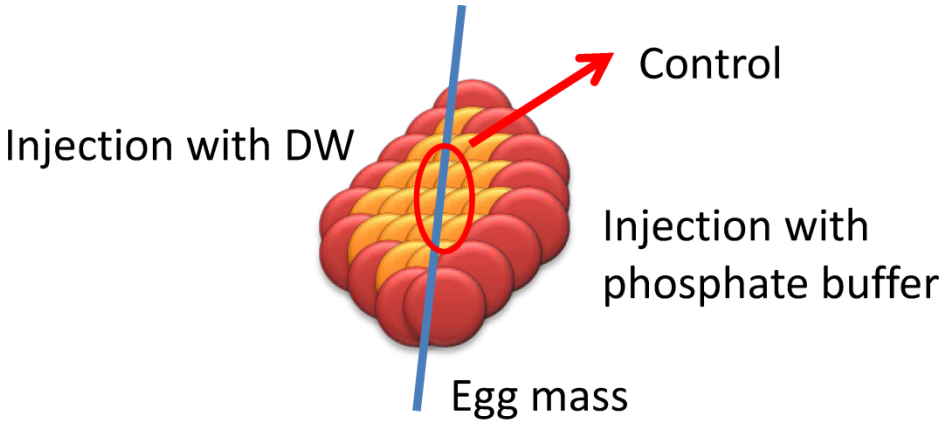


Figure 3-4. Strategy for the solvent selection experiment. Injected eggs were in red; intact eggs in center were in yellow.

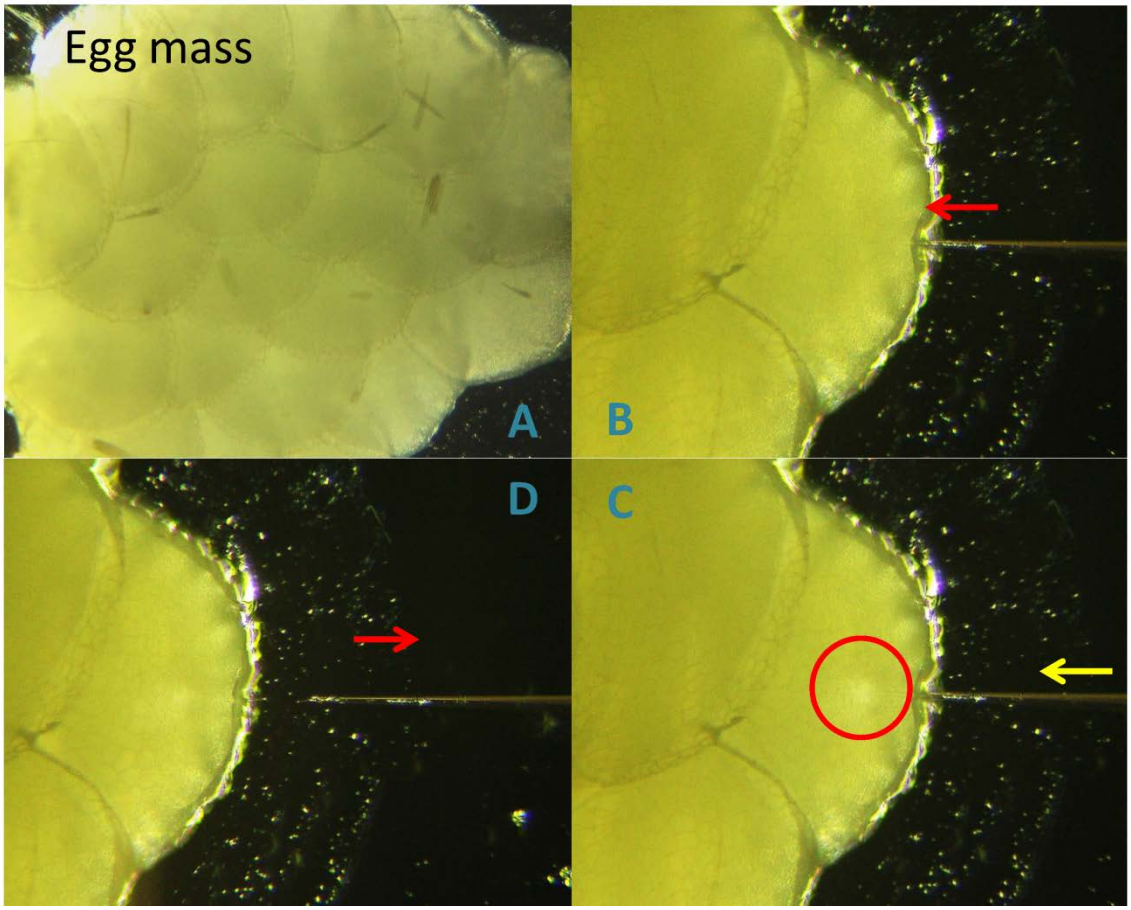


Figure 3-5. Egg mass and injection to the eggs. Needle for the injection was using Femtotips II (Eppendorf). A: egg mass; B: stab into an egg; C: inject with the TALEN RNA; D: pull out the needle.

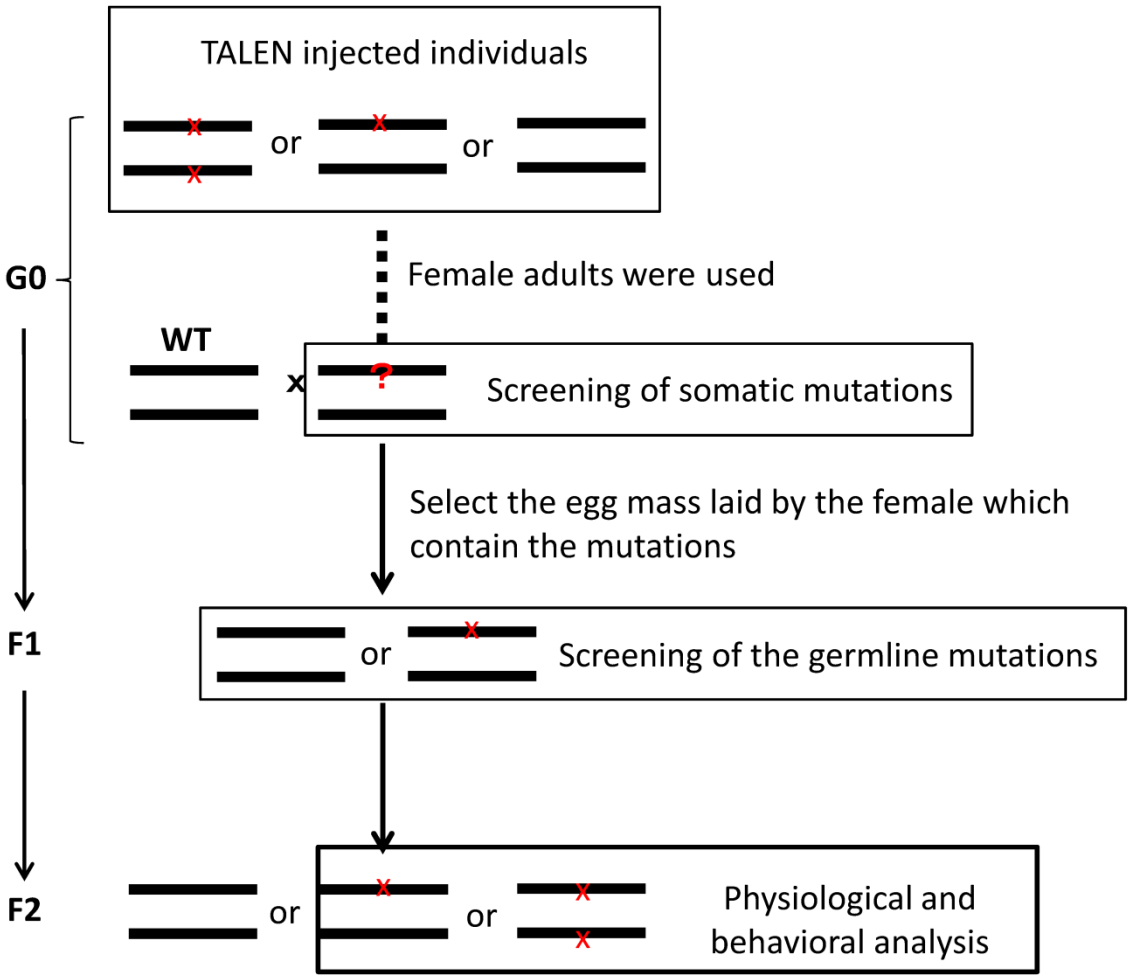


Figure 3-6A. Crossing scheme for mutagenesis of *OfurOrco*. Two black lines stand for the autosomes where *OfurOrco* located. Red crosses stand for mutation. One pair of autosomes indicates an individual.

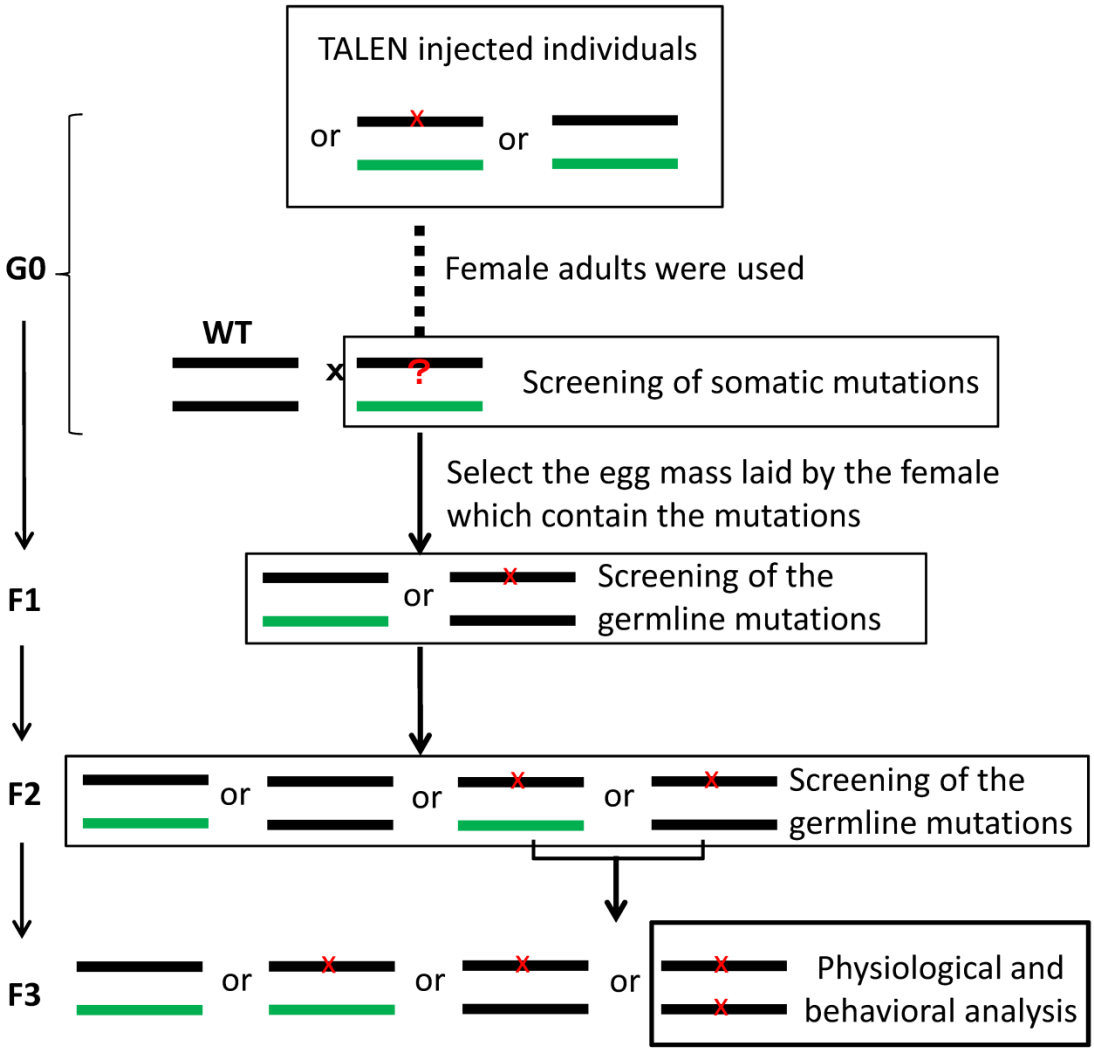


Figure 3-6B. Crossing scheme for mutagenesis of *OfurOR4*. Black and green line stands for Z and W chromosome, respectively. *OfurOR4* was located on Z chromosome. Red crosses stand for mutation.

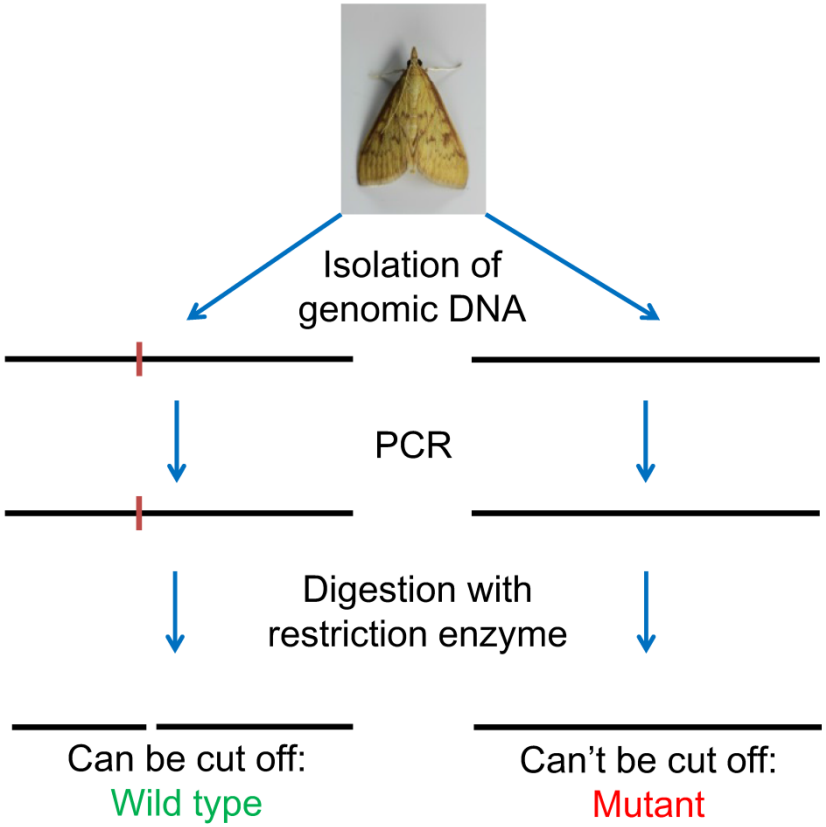


Figure 3-7. Principle of genotyping.

3.3 Results

3.3.1 Design of TALEN and preparation for injection

TALEN were designed close to the 5 prime end of ORF for a better silencing of the target (**Fig. 3-1**). For *OfurOrco*, TALEN were designed in first exon. For *OfurOR4*, because the ORF was only 12bp in the first exon, TALEN were designed at the second exon. Two pairs of TALEN for *OfurOrco* and one pair of *OfurOR4* were designed and obtained in the entry vectors. Recognition sites were generally 18bp in length, and spacer length was 15-17bp. N-terminus was 165aa in length and C-terminus was 57aa in length (**Table 3-3**). PCR on genomic DNA using wobbles primers at the target sites of Orco-TAL-1, Orco-TAL-2, and OR4-TAL-1 amplified the specific band only, suggesting that homologous sequences do not exist in the genome (**Table 3-4**). The vectors were cloned and recombined to a destination vector, amplified again and linearized for RNA syntheses by the methods described in **Figure 3-2**. The synthesized RNA were mixed in pairs and diluted to 400 ng/ul for each TALEN after checked by loading on a gel and stored at -80℃. Distilled water was used as solvent for RNA injection because there was no significant difference from phosphate buffer in hatchability (**Fig. 3-8**). Distilled water was also advantageous because of lower possibility of contamination that causes plugging of the needles.

Table 3-3. Summary of designed TALEN.

TALEN Pairs	Orco-TAL-1	Orco-TAL-2	OR4-TAL-1
N-terminus (aa)	165	165	165
C-terminus (aa)	57	57	57
Recognition Site Left (bp)	18	18	18
Recognition Site Right (bp)	18	18	18
Spacer (bp)	16	17	15
Restriction Enzyme Site	*Hae* III, *Sau96* I, *Mn* II, *Pho* I, *EcoO109I*	*HpyCH4V*, *Cac8* I	*BceA* I, *Bpm* I

Table 3-4. Target site sequences and corresponding designed TALEN.

			0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
Orco-TAL-1	**Left**	Target	T	G	A	T	G	A	C	C	A	A	A	G	T	G	A	A	A	G	C
		TALE		NN	NI	NG	NN	NI	HD	HD	NI	NI	NI	NN	NG	NN	NI	NI	NI	NN	HD
	Right	Target	T	G	A	T	G	T	T	T	G	G	C	A	T	C	A	A	A	T	C
		TALE		NN	NI	NG	NN	NG	NG	NG	NN	NN	HD	NI	NG	HD	NI	NI	NI	NG	HD
Orco-TAL-2	**Left**	Target	T	C	C	G	A	T	T	T	G	A	T	G	C	C	A	A	A	C	A
		TALE		HD	HD	NN	NI	NG	NG	NG	NN	NI	NG	NN	HD	HD	NI	NI	NI	HD	NI
	Right	Target	T	G	A	A	T	A	G	G	A	A	A	T	G	A	C	C	A	G	C
		TALE		NN	NI	NI	N	NI	NN	NN	NI	NI	NI	NG	NN	NI	HD	HD	NI	NN	HD
OR4-TAL-1	**Left**	Target	T	C	A	A	A	T	A	T	A	T	T	C	G	A	G	G	A	C	A
		TALE		HD	NI	NI	NI	NG	NI	NG	NI	NG	NG	HD	NN	NI	NN	NN	NI	HD	NI
	Right	Target	T	C	T	C	T	G	T	G	A	A	T	C	C	G	A	A	A	G	A
		TALE		HD	NG	HD	NG	NN	NG	NN	NI	NI	NG	HD	HD	NN	NI	NI	NI	NN	NI

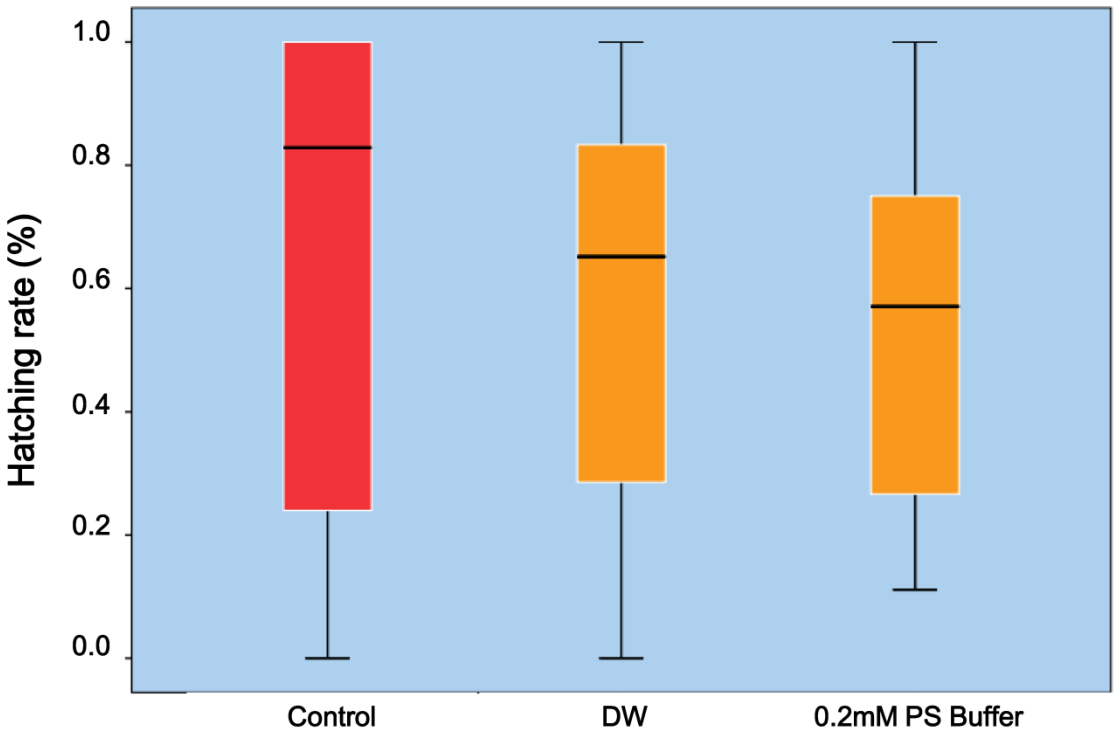


Figure 3-8. Result of the solvent selection experiment.

3.3.2 High rate of somatic mutations proved efficiency of TALEN

Microinjection may lead to the delayed development of some embryos and larvae. The injected individuals took about 40-45 days for development to pupae whereas the wild type took only 30-35 days. For *OfurOrco* targeted mutagenesis, 1408 embryos were injected with Orco-TAL-1 mRNAs, and 595 larvae hatched with a hatching rate at 42.3% (**Table 3-5**). Somatic mutation rate of G0 was 70.8% (75/106). Orco-TAL-2 was failed; no mutations were detected among G0 individuals. The difference in frequency of somatic mutation between the two TALEN suggests that the efficiency was heavily influenced by the selection of target site. For *OfurOR4* targeted mutagenesis, 1161 embryos were injected with OR4-TAL-1 mRNAs, and 593 larvae hatched with a hatching rate at 51.1%. Somatic mutation rate of G0 was 60.5% (69/114). Surprisingly, individuals without wild type bands were found in G0 (**Fig. 3-9**), suggesting that Orco-TAL-1 and OR4-TAL-1 could induce biallelic mutations at a high frequency.

3.3.3 Germline mutations induced by TALEN

To examine whether these mutations were heritable, G1 of injected females were screened for the targeted mutations. At least eight individuals of each family were genotyped. Mutation rates of G1 within single family were 0-100%. Among 35 egg masses laid by the females with somatic mutations caused by Orco-TAL-1, 22 families gained the heritable mutations with a germline mutation rate at 62.9% (**Table 3-5**). Among 28 egg masses laid by the females with somatic mutations caused by OR4-TAL-1, 21 families gained the heritable mutations with a germline mutation rate at 75.0%. The true germline mutation rate may be different from these results because not

all the offspring in one family were tested. Furthermore, G1 families from females without somatic mutations were not screened, which might also produce the germline mutants (Ma et al, 2012).

3.3.4 Analysis of the mutant sequences

To examine what types of mutations had been induced by TALEN, we extracted the genomic DNA from 9 *OfurOrco* targeted broods and 10 *OfurOR4* broods, respectively. Genomic DNA surrounding the target sites were amplified by PCR using corresponding primers (**Table 3-2**) and sequenced. Eighteen and 9 mutant sequences induced by Orco-TAL-1 and OR4-TAL-1, respectively, were identified. These mutations include deletions, insertions, and substitutions (**Fig. 3-10**). Most of the mutations occurred between the two TALEN binding domains. But in some case, large deletions were also detected. A 78bp and 70bp length deletions were found in *OfurOrco* mutants, and a large deletion of 583bp nucleotides were found in *OfurOR4* mutants.

Table 3-5. Somatic mutations and germline mutations.

		Orco-TAL-1	OR4-TAL-1
	Injected	1408	1161
	Hatched	595	593
	Hatching rate	42.3%	51.1%
G0	**Female adults**	129	184
(individual numbers)	**Screened female adults**	106	114
	Somatic mutation	75	69
	Somatic mutation rate	70.8%	60.5%
	Laid eggs	35 (57)	28 (58)
F1	**Germline mutation**	22	21
(Strains numbers)	**Germline mutation rate**	62.9% (38.6%)	75.0% (36.2%)

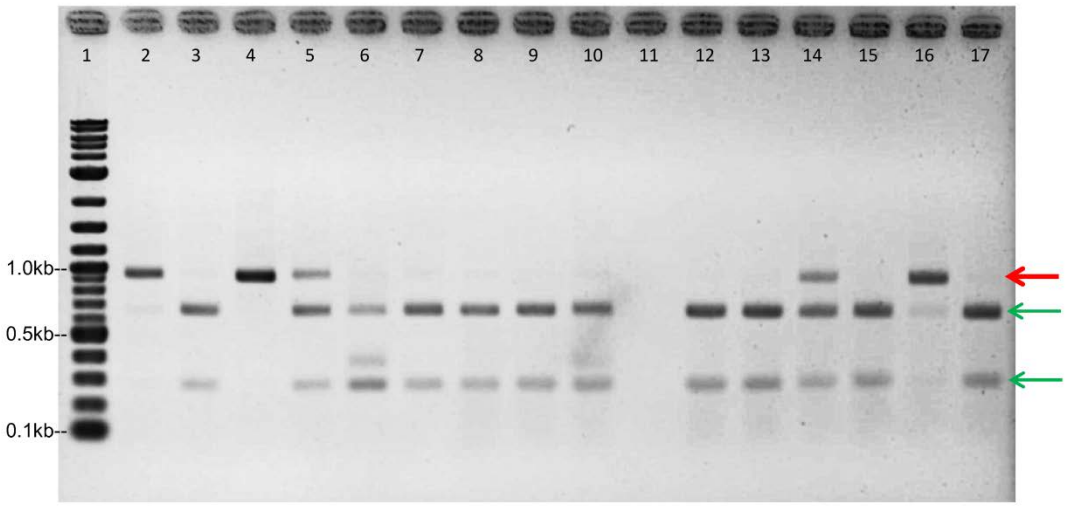


Figure 3-9. Genotyping for the targeted mutations in the Orco-TAL-1 injected G0 generation. Digested PCR fragments were examined by electrophoresis. Each lane represents an individual. Red arrow indicates the expected size of undigested fragments that may contain mutations. Green arrows indicate the expected size of digested fragments derived from wild type. The individual shown in lane 4 only had a mutant fragment, suggesting that all somatic cells contain biallelic mutations.

Figure 3-10. Mutations in targeted sites. Inserts were in green; deletions were shown as "-", large deletions were shown with numbers. (**A**) Mutations in *OfurOrco* sequences. (**B**) Mutations in *OfurOR4* sequences.

3.4 Discussion

TALEN is a powerful tool for targeted mutagenesis in non-model organisms. In this study, I have successfully induced the hereditable mutations for two odorant receptors by injecting TALEN mRNA into the embryos of *O. furnacalis*, showing that TALEN can be used in pest insects. But it should be noted that the efficiency of TALEN depends on the target sites. In this chapter, I designed three pairs of TALEN for two genes, but one pair was failed to induce mutation, whereas the others induced mutations at an extremely high rate. Such difference might depend on the binding efficiency of each TALEN, but is also possible that the off-target effect might underlie the failure of Orco-TAL-2.

There were several difficulties in a practical aspect. Screening TALEN-induced mutants is very time-consuming and costly, especially for genes with unknown function and phenotype (Ma et al, 2012; Sajwan et al, 2013). In our case, there was no visible phenotype for the mutant individuals. Therefore, we needed to genotype all the individuals after they produced offspring. Genotyping should be finished within a short period of time before the eggs hatch, otherwise it would be necessary to keep all of them, which also required additional labor and cost. Low survival rate and mating rate in *O. furnacalis* especially for the mutant individuals made it very difficult to get enough numbers of pairs to produce the homozygous offspring.

In conclusion, although it is time and labor consuming to get the homozygous individuals, TALEN system can be applied in pest insects to induce mutations in real target genes such as odorant receptors. Physiological and behavioral analyses of such mutants are expected to reveal the *in vivo* function of odorant receptors.

Chapter 4

Physiological and behavioral analysis of the *OfurOrco* (*OR2*) mutants

4.1 Introduction

In Chapter 3, two odorant receptors, *OfurOR4* and *OfurOrco*, were successfully mutagenized using TALEN in *O. furnacalis*. In this chapter, behavioral and physiological examinations of the *OfurOrco* mutants were carried out using wind tunnel and GC-EAD analysis. Although the involvement of odorant receptors in the pheromone perception was strongly supported by the experimental evidence, it has not been directly confirmed *in vivo*. *Orco* is the co-receptor subunit required for the function of all the other odorant receptors, and the *OfurOrco* mutant moths should have lost the function of all the odorant receptors including pheromone receptors. Therefore, the *OfurOrco* mutants offer an unique opportunity to confirm the *in vivo* function of odorant receptors in the pheromone perception.

4.2 Materials and Methods

4.2.1 Insect rearing

O. furnacalis were collected and cultured as described in chapter 2. To get homozygous mutants, heterozygotes from different strains were crossed to suppress background effect caused by possible mutations at other loci than *OfurOrco*.

4.2.2 Wind tunnel assay

F2 male adults from F1 parents that were heterozygous for the *OfurOrco* mutation were used. The males were first tested by wind tunnel assay (**Fig. 4-1**). The wind tunnel was 30cm in diameter and 200cm in length. Single adult was collected from netcage using a 25ml centrifuge tube, and then transferred to a handmade 6cm × 12cm wire-mesh tube. After 15min of incubation the wire-mesh tube containing a male was placed at the platform on the downwind side. Preparation of the pheromone source was carried out in another room. The pheromone components of *O. funacalis*, (*E*)- and (*Z*)-12-tetradecenyl acetate, were mixed (1:1) and applied onto a piece of filter paper (0.1cm × 0.3cm) at a total amount of 200 ng or 2000 ng. The filter paper was hanged at the center of upwind side of the wind tunnel. Whether the males reach to the pheromone source area were recorded.

4.2.3 Cup assay

Cup assay was used for simplified observation. The individuals were independently put in a plastic clean cup (60 ml, Risu Pack Co. Itd., Japan) with a 2 mm hole. Pheromones were added into the cap through the hole. Occurrences of courtship

behavior and wing vibration were recorded.

4.2.4 GC-EAD analysis

The individuals tested in wind tunnel or cup assay were then anesthetized by CO_2 and their antennae were dissected for GC-EAD analysis. (*E*)- and (*Z*)-12-tetradecenyl acetate were mixed (1:1) at a concentration of 300 ng/ul in hexane. 2ul was injected for each test. Program was set at a start temperature of 120℃, incubated for 2min, increased the temperature at a ramp of 12℃/min to 180℃, then change the ramp to 5℃/min, and incubated for 5 min when reach the final temperature 240℃ (**Fig. 4-2**).

4.2.5 Genotyping the tested individuals

All individuals examined in the above analyses were subjected for genotyping as described in chapter 3.

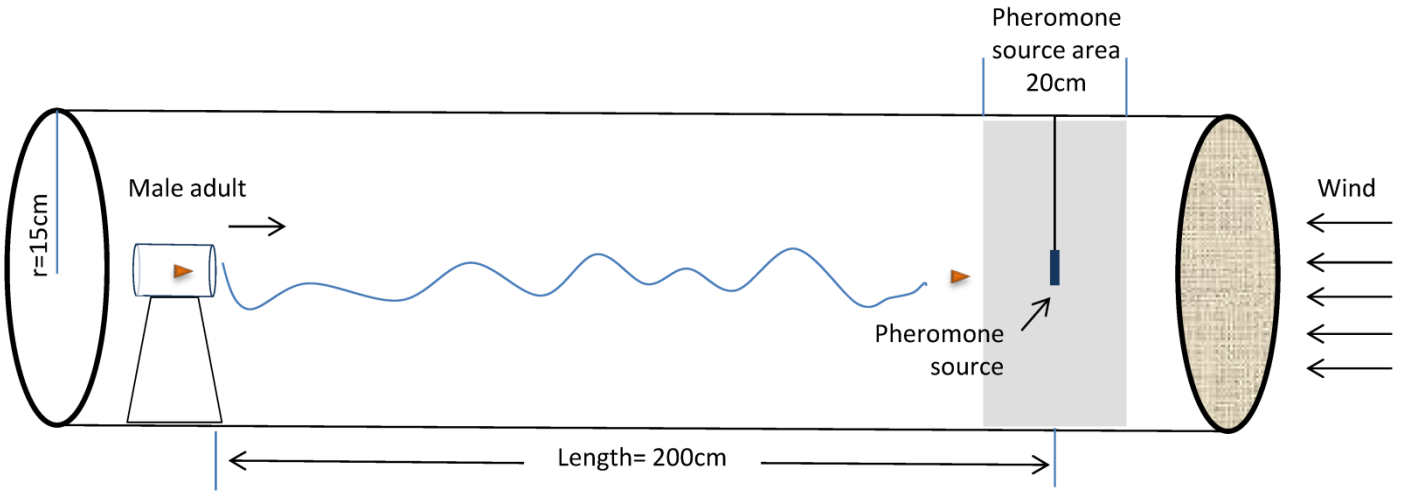


Figure 4-1. Setting of wind tunnel assay.

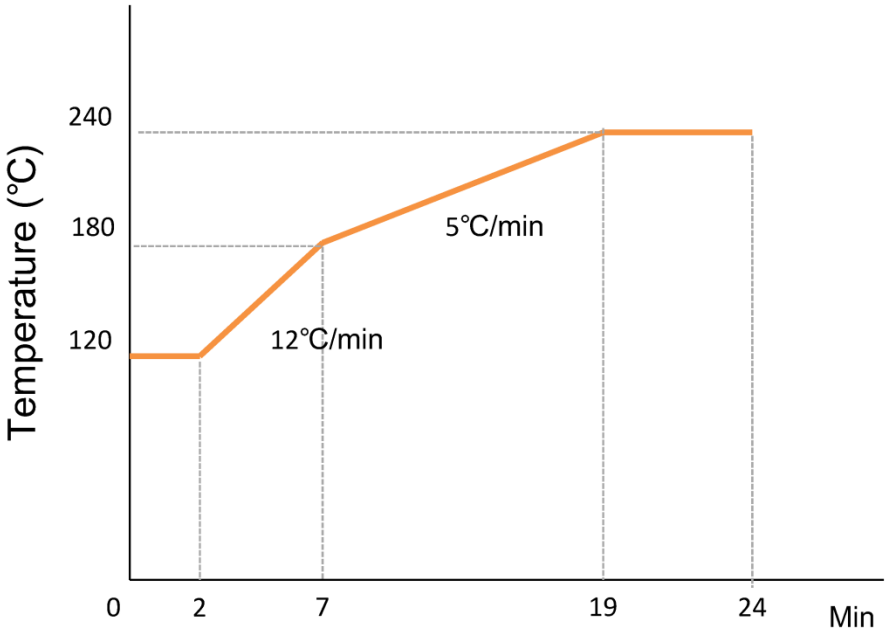


Figure 4-2. Temperature setting of GC-EAD.

4.3 Results:

4.3.1 The *OfurOrco* mutant moths lost the ability to detect the pheromone

Males homozygous for the *OfurOrco* mutation didn't show any response in wind tunnel (**Table 4-1**). Furthermore, the antennae from the homozygous males didn't respond to the pheromone components in GC-EAD analysis (**Fig. 4-3**).

4.3.2 Heterozygous males showed decreased sensitivity to the pheromone

Difference in the response to the pheromone was observed between the wild type and the *OfurOrco* heterozygous males. When the pheromone was presented at low amount (200ng), 60.0% of the wild type males reached to the pheromone source area, whereas only 33.3% of the heterozygous males did so. When the higher amount of the pheromone (2000ng) were presented, the result was reversed; only 38.5% of the wild type males reached to the pheromone source area, and 64.7% of the heterozygous males reached to the pheromone source area (**Table 4-1**). In GC-EAD results, by using 600 ng mixed pheromone components, antennae from heterozygotes had a smaller peak response to the pheromone than that from wild types, but they were not significant different (**Fig. 4-3**).

Table 4-1. Behavioral response to different amount of pheromones in the wind tunnel assay.

Genotpye	Fly to pheromone source area	
	200ng	**2000ng**
WT	60.0% (3/5)	38.5% (5/13)
Heterozygotes	33.3% (2/6)	64.7% (11/17)
Homozygotes	0 (0/4)	-

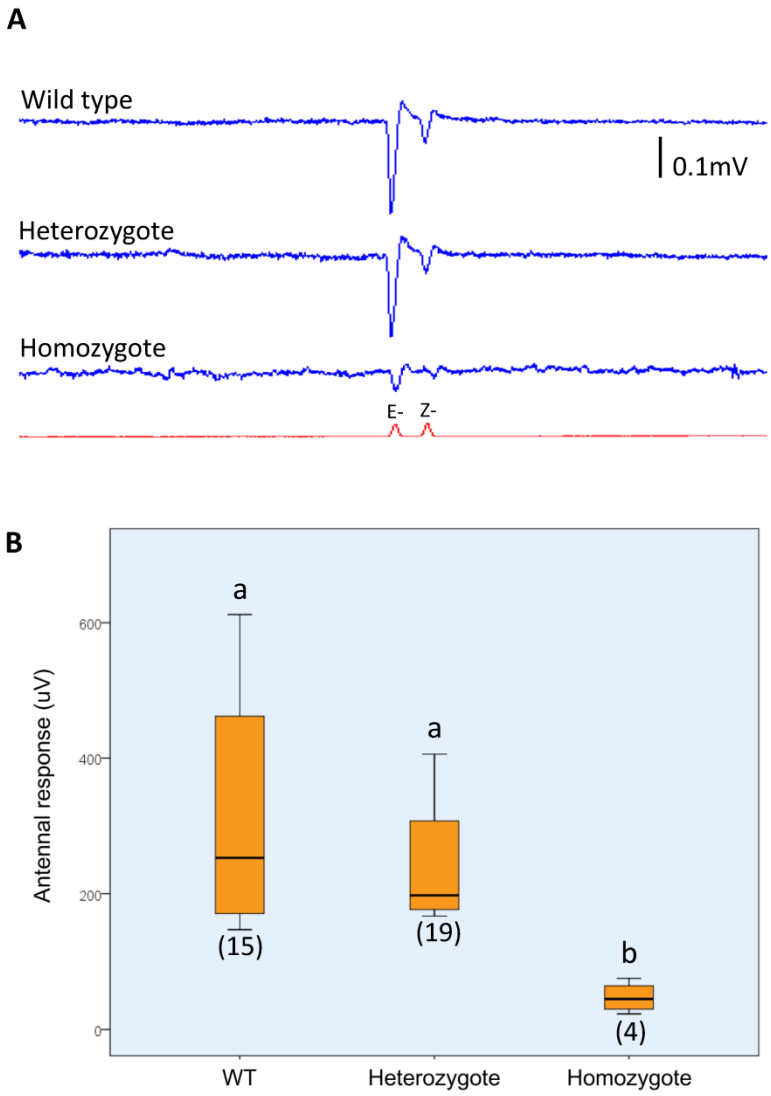


Figure 4-3. Results of GC-EAD. **A**: Typical antennal response of different genotype. FID outputs from gas chromatography were shown in red. Responses of antennae detected by EAD were shown in blue. Retention time of (*E*)- and (*Z*)-12-tetradecenyl acetate were 11.51 and 11.82 min. **B**: Antennal responses for each genotype. Sum of the antennal responses to (*E*)- and (*Z*)-12-tetradecenyl acetate were used. The letters above the boxes are shown the significant differences: same letters, not significant different at P >0.05; different letters, significant different at P <0.05. The P values were after Bonferroni correction by pairwise Wilcoxon rank sum tests. Numbers under the boxes are numbers of tested individuals.

4.4 Discussion:

In this chapter, I presented the direct evidence that odorant receptors are necessary for pheromone perception. Furthermore, it was suggested that the *OfurOrco* heterozygotes have decreased sensitivity to the pheromones. To make about 60% of individuals respond, 200ng of the pheromone was required for the wild type males but 2000ng for the *OfurOrco* heterozygotes. It means that the threshold of pheromone perception was higher in the *Orco* heterozygotes than in the wild type. In other words, gene dosage of *OfurOrco* affected the threshold of pheromone detection. The amount of active odorant receptors may influence the sensitivity of pheromone perception (**Fig. 4-4**). If the amount of *Orco* is reduced, the activity of other ORs will be also decreased. In such a case, the threshold of pheromone response becomes higher. On the other hand, decreased sensitivity may prevent the habituation in the central nervous system against the higher concentration of pheromones. In this sense, *Orco* should play an important role in regulation of sensitivity to pheromones.

General discussion

In this study, I took *O. furnacalis* as an example to apply a novel strategy combining two recently developed tools for identification and functional analysis of odorant receptors. In chapter 1, an RNA-seq analysis using the Roche 454 GS-Jr system was performed as a preliminary trial to see if low depth sequencing was enough to identify the repertoire of odorant receptors. As a result, 29 candidate odorant receptors were identified including eight of nine previously published receptors. Expression levels of the identified receptors were confirmed by qRT-PCR in various tissues. These results provided a basis for further analysis in chapter 2, where the Illumina MiSeq system was used to identify the entire repertoire of odorant receptors expressed in the antennae of *O. furnacalis*. As a result, 52 candidate odorant receptors were identified, among which three showed female-biased expression. In addition, candidate genes of 21 IRs, 5 GRs, 2 SNMPs, and 26 ODEs were identified. These results would be useful for studies about the molecular mechanisms of chemical perception in the *Ostrinia* species. In chapter 3, I took *OfurOR4* and *OfurORco* as a target of mutagenesis using TALEN. As a result, mutants for both genes were successfully obtained. In the *OfurOrco* targeted group, the somatic mutation rate at Generation-0 was 70.8% and the germline mutation rate observed in the next generation was 55.0%. In the *OfurOR4* targetd group, the somatic mutation rate was 60.5% and the germline mutation rate was 38.9%. This was the first study to knockout real target genes (not markers) using TALEN in pest insects. Finally, in the chapter 4, physiological and behavioral analysis of the *OfurOrco* mutants were performed to examine the *in vivo* function of the receptor. As a result, homozygous males were shown to have lost their response to the pheromone. In addition,

comparisons between wild type and heterozygous males suggested that *OfurOrco* was involved in the determination of sensitivity to the pheromone.

1. RNA-seq is a powerful tool to identify chemosensory genes in pest insects

In this study, RNA-seq was carried out using two types of sequencing platform, Roche 454 GS-Jr and Illumina MiSeq. Length of reads produced by Roche 454 systems is longer than that of Illumina, while Illumina systems have larger throughput than Roche 454 systems. Longer reads were expected to be advantageous to *de novo* assembly in species without whole genome sequences. For example, two or more adjacent repeats may be reduced to one in the assembly using short reads. However, as shown in this study, deeper sequencing was more important than longer read length for identification of odorant receptors because their expression level was low.

Results were also dependent on assemblers. In chapter 1, two assemblers, Trinity and Newbler, were compared. Trinity is inclined to dominating patterns, which means that if different nucleotides are observed for single sites, the one with highest frequency will be used. It is advantageous in reducing number of variants caused by sequencing errors. However, for genes with a similar sequence, for example the *OfurOR5a* and *OfurOR5b* in this study, trinity couldn't discriminate them. Newbler was better in this regard because it makes a large iso-group, which contains all the possibilities even with one single nucleotide difference. Nevertheless, it can be a weak point at the same time because it is usually difficult to tell which is correct or erroneous without reference genomes. In this study, trinity was superior in reconstructing longer transcripts.

2. TALEN is an adequate tool for efficient mutagenesis

Efficiency of TALEN has been confirmed in various insects. For example, in *Drosophila melanogaster*, inheritable modification at the *yellow* gene was detected from 31.2% of the injected F0 fertile individuals (Liu et al, 2012b). In the hemimetabolous insect *Gryllus bimaculatus*, 17% of founder animals transmit disrupted gene alleles at the targeted locus, *Gb'lac2* (Watanabe et al, 2012). In *Bombyx mori*, by injection of two pairs of TALEN for the *BmBlos2* gene, germline mutations were generated in 31% and 15% of injected individuals, respectively (Ma et al, 2012). In this study, 38.9% and 55.0% of G1 individuals contained inheritable mutations, showing that TALEN was effective in pest insects, too.

The CRISPR/Cas9 system is the newest tool for genome editing. Its efficiency seems to be more variable than TALEN according to the published data. High efficiency is necessary to isolate mutants in non-model insects because no markers are available. Preferably, germline mutation rate should be more than 10%, otherwise quite a lot of genotyping effort will be required. In this sense, suitability of the CRISPR/Cas9 system in mutagenesis of pest insects is still not clear. Nevertheless, once the application method of TALEN was established in *O. furnacalis*, CRISPR/Cas9 can be applied using the same method. It would be worth trying the CRISPR/Cas9 system in *O. furnacalis* in future studies.

Combination of the two tools opens a novel path in research of pest insects

Combination of the two powerful tools revealed a novel path to identification and functional analysis of any genes in any species. In this study, I identified the chemosensory genes in the Asian corn borer *O. furnacalis* using NGS, and obtained the

mutants for *OfurOR4* and *OfurORco* by genome editing. This strategy is expected to be applicable to any pest insects for elucidating molecular mechanisms necessary for their control.

Acknowledgements

Firstly, I express my deepest gratitude to my supervisor Associate Professor Takashi Matsuo. Without his patient teaching and constant supporting, this study could not be reached to the present state. I have been very fortunate in studying with a supervisor that strict in research and warm in the daily life. He taught me not only the knowledge in science but also the brain that how to think and behave as an independent researcher. Also, I would like to deeply appreciate Professor Yukio Ishikawa. He gave me the chance to study in a world famous university and helped a lot from the first day I came to Japan.

Furthermore, I would like to appreciate Assistant professor Takeshi Fujii, for there continuous guidance in my physiology study. Also I appreciate Dr. Masanobu Yamamoto for the experiment setting and data analysis with GC-EAD. I appreciate all the members in our laboratory for their helps in the daily life.

Last, I would like to greatly appreciate Dr. Katsuhisa Ozaki in JT Biohistory Research Hall, with his grate help in Illumina sequencing. I appreciate Yuji Yashikochi, for the genomic information supporting in *O. furnacalis*. I also appreciate Assistant professor Takashi Kiuchi in the Laboratory of Insect Genetics and Bioscience and Dr. Zhixin Luo in the University of Tsukuba, for their help in injection and wind tunnel.

References

Abuin L, Bargeton B, Ulbrich MH, Isacoff EY, Kellenberger S and Benton R. (2011) Functional architecture of olfactory ionotropic glutamate receptors. Neuron. 69: 44–60.

Ai M, Min S, Grosjean Y, Leblanc C, Bell R, Benton R and Suh GSB. (2010) Acid sensing by the *Drosophila* olfactory system. Nature. 468: 691-695

Allen JE and Wanner KW. (2011) Asian corn borer pheromone binding protein 3, a candidate for evolving specificity to the 12-tetradecenyl acetate sex pheromone. Insect Biochem Mol Biol. 41(3): 141-149

Andersson J, Borg-Karlson AK, Vongvanich N and Wiklund C. (2007) Male sex pheromone release and female mate choice in a butterfly. J Exp Biol 210: 964–970.

Baldwin IT and Preston CA. (1999) The eco-physiological complexity of plant responses to insect herbivores. Planta. 208: 137–145.

Bengtsson JM, Trona F, Montagné N, Anfora G, Ignell R et al. (2012) Putative chemosensory receptors of the codling moth, *Cydia pomonella*, identified by antennal transcriptome analysis. PLoS ONE. 7(2): e31620

Benton R (2007) Sensitivity and specificity in *Drosophila* pheromone perception. Trends Neurosci. 30: 512–519

Benton R, Vannice KS and Vosshall LB. (2007) An essential role for a *CD36*-related receptor in pheromone detection in *Drosophila*. Nature. 450: 289–293

Benton R, Vannice KS, Gomez-Diaz C and Vosshall LB. (2009) Variant Ionotropic Glutamate Receptors as Chemosensory Receptors in *Drosophila*. Cell. 136: 149–162.

Blomquist GJ and Vogt RG. (2003) Biochemistry of female moth sex pheromones. Insect Pheromone Biochemistry and Molecular Biology. London, UK: Elsevier

Academic Press.

Boch J and Bonas U. (2010) *Xanthomonas AvrBs3* family-type III effectors: discovery and function. Annu Rev Phytopathology. 48: 419-436

Boch J et al. (2009) Breaking the code of DNA binding specificity of TAL-type III effectors. Science. 326: 1509

Bogdanove AJ and Voytas DF. (2011) TAL effectors: customizable proteins for DNA targeting. Science. 333: 1843-1846

Bruyne de M and Baker TC. (2008) Odor Detection in Insects: Volatile Codes. J Chem Ecol. 34(7): 882-897

Bushra S and Tariq M. (2014) How aphid alarm pheromone can control aphids: a review. Archives of Phytopathology and Plant Protection. 47(13): 1563-1573

Caasi-Lit MT, Fernandez EC. (2006) Survey of alternate host plants of Asian corn borer (*Ostrinia furnacalis* (Guenée)) in major corn production areas of the Philippines. Asian Int J Life Sci. 15: 47–71.

Caasi-Lit MT, Sapin GD, Beltran AKM, de Leus EG, Mantala JP, Latiza SA. (2009) Larval survival and ovipositional preference of the Asian corn borer, *Ostrinia furnacalis* Guenée, for some alternate host plants at different growth stages. Philip Entomol. 23:184–185.

Calvo D, Gomez-Coronado D, Suarez Y, Lasuncion MA and Vega MA. (1998) Human *CD36* is a high affinity receptor for the native lipoproteins HDL, LDL, and VLDL. J Lipid Res. 39: 777–788

Carroll D. (2011) Genome engineering with zinc-finger nucleases. Genetics. 188: 773–782

Croset V, Rytz R, Cummins SF, Budd A, Brawand D, Kaessmann H, Gibson TJ and

Benton R. (2010) Ancient protostome origin of chemosensory ionotropic glutamate receptors and the evolution of insect taste and olfaction. PLoS Genet. 6: e1001064

Dicke M and van Loon JJA. (2000) Multitrophic effects of herbivore-induced plant volatiles in an evolutionary context. Entomol Exp Appl. 97: 237–249.

Ferkovich SM, van Essen F and Taylor TR. (1980) Hydrolysis of sex pheromone by antennal esterases of the cabbage looper, *Trichoplusia ni*. Chem Senses. 5: 33–46

Glover TJ, Perez N and Roelofs WL. (1989) Comparative analysis of sex-pheromone response antagonists in three races of European corn borer. J Chem Ecol. 15: 863-873

Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA et al. (2011) Full-length transcriptome assembly from RNA-seq data without a reference genome. Nat Biotechnol. 29(7): 644-652

Greg SBS, Allan MW, Anne CH, Jing WW, Anne FS, Seymour B, Richard A & David JA et al. (2004) A single population of olfactory sensory neurons mediates an innate avoidance behaviour in *Drosophila*. Nature. 431: 854–859

Grosse-Wilde E, Kuebler LS, Bucks S, Vogel H, Wicher D et al. (2011) Antennal transcriptome of *Manduca sexta*. Proc Natl Acad Sci USA. 108: 7449-7454

Gupta RM and Musunuru K. (2014) Expanding the genetic editing tool kit: ZFNs, TALENs, and CRISPR-Cas9. J Clincal Invstigation. 124(10): 4154-4161

Haas BJ, Papanicolaou A, Yassour M, Grabherr M, Blood PD et al. (2013) *De novo* transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. Nat Protocols. 8(8): 1494-1512

Halitschke R, Kealer A, Kahl J, Lorenz A, and Baldwin IT. (2000) Ecophysiological comparison of direct and indirect defenses in *Nicotiana attenuata*. Oecologia 124: 408–417.

Hallem EA, Dahanukar A and Carlson JR. (2006) Insect odor and taste receptors. Annu Rev Entomol. 51: 113–135

Hansson BS and Stensmyr MC. (2011) Evolution of insect olfaction. Neuron 72(5): 698-711

Hansson BS. (2014) From organism to molecule and back - Insect olfaction during 40 years. J Chem Ecol 40(5): 409-410

Hildebrand JG, Shepherd GM (1997). Mechanisms of olfactory discrimination: converging evidence for common principles across phyla. Annu Rev Neurosci. 20: 595–631.

Hirokawa T, Boon-Chieng S and Mitaku S. (1998) SOSUI: classification and secondary structure prediction system for membrane proteins. Bioinformatics 14(4): 378-379

Hoeppner MP, Lundquist A, Pirun M, Meadows JRA, Zamani N et al. (2014) An Improved canine cenome and a comprehensive catalogue of coding genes and non-coding transcripts. PLoS ONE. 9(3): e91172

Howlett N,Dauber KL, Shukla A, Morton B, Glendinning JI AND Brent E. (2012) Identification of chemosensory receptor genes in *Manduca sexta* and knockdown by RNA interference. BMC Genomics. 13:211

Huang YP, Honda H, Yoshiyasu Y, Hoshizaki S, Tatsuki S et al. (1998b) Sex pheromone of the butterbur borer, *Ostrinia zaguliaevi*. Entomol Exp Appl. 89: 281-287

Huang YP, Takanashi T, Hoshizaki S, Tatsuki S, Honda H et al. (1998) Geographic variation in sex pheromone of Asian corn borer, *Ostrinia furnacalis*, in Japan. J Chem Ecol. 24(12): 2079-2088.

Ishikawa Y, Takanashi T, Huang YP. (1999) Comparative studies on the sex pheromones of *Ostrinia spp.* in Japan: the burdock borer, *Ostrinia zealis*. Chemoecology. 9: 25-32

Ishikawa Y, Takanashi T, Kim CG, Hoshizaki S, Tatsuki S and Huang YP. (1999) *Ostrinia spp.* in Japan: their host plants and sex pheromones. Entomol Exp Appl. 91: 237-244

Jacquin-Joly E and Merlin C. (2004) Insect olfactory receptors: contributions of molecular biology to chemical ecology. J Chem Ecol. 30(12): 2359-2397

Ji Bao-Zhong, Liu Shu-Wen, Cao Dan-Dan et al. (2013) Research progress in trail pheromones in termites. Acta Entomologica Sinica. 56(9): 1038-1046

Jin X, Ha TS and Smith DP. (2008) SNMP is a signaling component required for pheromone sensitivity in *Drosophila*. Proc Natl Acad Sci USA. 105:10996-11001

Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA and Charpentier E. (2012) A Programmable Dual-RNA-Guided DNA Endonuclease in Adaptive Bacterial Immunity. Science. 337: 816–821

Johansson BG and Jones TM. (2007) The role of chemical communication in mate choice. Biol Rev. 82(2): 265-289

Jordan MD, Anderson A, Begum D, Carrher C, Authier A et al. (2009) Odorant Receptors from the Light brown Apple Moth (*Epiphyas postvittana*) Recognize Important volatile compounds produced by plants. Chem Senses. 34: 383-394

Joseph RM, Devineni AV, King IF and Heberlein U. (2009) Oviposition preference for and positional avoidance of acetic acid provide a model for competing behavioral drives in *Drosophila*. Proc Natl Acad Sci USA. 106: 11352–11357

Jurenka RA, Blomquist G and Vogt R. (2003) Biochemistry of female moth sex pheromones. Insect pheromone biochemistry and molecular biology. Elsevier, Amsterdam. pp. 53–80

Karlson P and Butenandt A. (1959) Pheromones (Ectohormones) in Insects. Annu Rev

Entomol. 4: 39-58

Katoh K and Toh H. (2010) Parallelization of the MAFFT multiple sequence alignment program. Bioinformatics. 26(15): 1899-1900

Kiely A, Authier A, Kralicek AV, Warr CG and Newcomb RD. (2007) Functional analysis of a Drosophila melanogaster olfactory receptor expressed in Sf9 cells. J Neutosci Methods. 159: 189-194

King JR, Christensen TA and Hildebrand JG. (2000) Response characteristics of an identified, sexually dimorphic olfactory glomerulus. J Neurosci. 20(6): 2391-2399

Knudsen JT, Eriksson R, Gershenzon J and Ståhl B. (2006) Diversity and distribution of floral scent. The Botanical Review. 72(1):1-120

Krogh A, Larsson B, von Heijne G and Sonnhammer EL. (2001) Predicting transmembrane protein topology with a hidden Markov model: Application to complete genomes. J Mol Biol. 305(3): 567-580

Kurtovic A, Widmer A, Dickson BJ (2007) A single class of olfactory neurons mediates behavioural responses to a *Drosophila* sex pheromone. Nature. 446: 542–546

Lassance JM and Lofstedt C. (2009) Concerted evolution of male and female display traits in the European corn borer, *Ostrinia nubilalis*. BMC Biology. 7: 10

Leal WS. (2005) Pheromone reception. Top Curr Chem. 240: 1–36

Leal WS. (2013) Odorant reception in insects: roles of receptors, binding proteins, and degrading enzymes. Annu Rev Entomol. 58: 373-391

Leary GP, Allen JE, Bunger PL, Luginbill JB, Linn CE et al. (2012) Single mutation to a sex pheromone receptor provides adaptive specificity between closely related moth species. Proc Natl Acad Sci USA. 109(35): 14081-14086.

Lemey P, Rambaut A, Drummond AJ and Suchard MA. (2009) Bayesian

phylogeography finds its roots. PLoS Comput Biol. 5(9): e1000520

Lewvanich A. (1973) Study of the identity of the corn stem borer in Thailand. Thai J Agric Sci. 7:103–109

Liu J, Li C, Yu Z, Huang P, Wu H, Wei C, Zhu N et al. (2012) Efficient and Specific Modifications of the Drosophila Genome by Means of an Easy TALEN Strategy. J Genetics and Genomics. 39(5): 209-215

Liu NY, Xu W, Papanicolaou A, Dong SL and Anderson A. (2014) Identification and characterization of three chemosensory receptor families in the cotton bollworm *Helicoverpa armigera*. BMC GENOMICS. 15: 597

Liu Y, Gu SH, Zhang YJ, Guo YY and Wang GR. (2012) Candidate olfaction genes identified within the *Helicoverpa armigera* antennal transcriptome. PLoS ONE. 7(10): e48260

Ma S, Zhang S, Wang F, Liu Y, Xu H, Liu C, Lin Y, Zhao P and Xia Q. (2012) Highly efficient and specific genome editing in silkworm using custom TALENs. PLoS ONE. 7(9): e45035

Makarova KS, Haft DH, Barrangou R, Brouns SJ, Charpentier E et al. (2011) Evolution and classification of the CRISPR–Cas systems. Nature. 9: 467-477

Miura N, Nakagawa T, Tatsuki S, Touhara K, Ishikawa Y. (2009) A male-specific odorant receptor conserved through the evolution of sex pheromones in *Ostrinia* moth species. Int J Biol Sci. 5: 319-330.

Miura N, Nakagawa T, Touhara K and Ishikawa Y. (2010) Broadly and narrowly tuned odorant receptors are involved in female sex pheromone reception in *Ostrinia* moths. Insect Biochem Mol Biol. 40(1): 64-73

Montell C. (2009) A taste of the *Drosophila* gustatory receptors. Curr Opin Neurobiol.

19(4): 345–353

Moraes de CM, Lewis WJ, Pare PW, Alborn HT, and Tumlinson JH. (1998) Hervibore-infested plants selectively attract parasitoids. Nature. 393: 570–573.

Mortazavi A, Williams BA, McCue K, Schaeffer L and Wold B. (2008) Mapping and quantifying mammalian transcriptomes by RNA-Seq. Nat Methods. 5(7): 621-628

Moscou MJ and Bogdanove AJ. (2009) A simple cipher governs DNA recognition by TAL effectors. Science. 326: 1501

Mutuura A and Munroe E. (1970) Taxonomy and distribution of the European corn borer and allied species: Genus *Ostrinia* (Lepidoptera: Pyralidae). Mem Entomol Soc Canada. 71:33-37.

Nakagawa T, Sakurai T, Nishioka T and Touhara K. (2005) Insect sex-pheromone signals mediated by specific combinations of olfactory receptors. Science. 307: 1638–1642

Obiero GFO, Mireji PO, Nyanjom SRG, Christoffels A, Robertson HM et al. (2014) Odorant and Gustatory Receptors in the Tsetse Fly *Glossina morsitans morsitans*. PLoS Neglected Tropical Diseases. 8(4): e2663

Ohno S. (2003) A new knotweed-boring species of the genus *Ostrinia Hübner* (Lepidoptera: Crambidae) from Japan. Entomol Sci. 6: 77-83

Orsi WD, Edgcomb VP, Christman GD and Biddle JF. (2013) Gene expression in the deep biosphere. Nature. doi:10.1038/nature12230

Paré PW and Tumlinson JH. (1999) Plant volatiles as a defense against insect herbivores. Plant Physiol. 121:325–331.

Pelosi P and Maida R. (1995) Odorant-binding proteins in insects. Comp Biochem Physiol B. 111:503-514

Poivet E, Gallot A, Montagne N, Glaser N, Legeai F et al. (2013) A comparison of the olfactory gene repertoires of adults and larvae in the noctuid moth *Spodoptera littoralis*. PLoS ONE 8(4): e60263

Reisenman CE, Riffell JA, Bernays EA and Hildebrand JG. (2010) Antagonistic effects of floral scent in an insect-plant interaction. Proc R Soc London, Ser B. 277:2371

Reisenman CE, Riffell JA, Duffy K, Pesque A, Mikles D and Goodwin B. (2012) Species-Specific Effects of Herbivory on the Oviposition Behavior of the *Moth Manduca*. Journal of Chemical Ecology. 10.1007/s10886-012-0228-1

Roberts A and Pachter L. (2012) Streaming fragment assignment for real-time analysis of sequencing experiments. Nat Methods. 10(1): 71-73

Roelofs WL, Du JW, Tang XH, Robbins PS and Eckenrode CJ. (1985) Three European corn borer populations in New York based on sex pheromones and voltinism. J Chem Ecol 11: 829-836

Rogers ME, Krieger J and Vogt RG. (2001) Antennal SNMPs (sensory neuron membrane proteins) of Lepidoptera define a unique family of invertebrate *CD36*-like proteins. J Neurobiol. 49 (1): 47-61

Rogers ME, Sun M, Lerner MR and Vogt RG. (1997) SNMP-1 a novel membrane protein of olfactory neurons of the silk moth *Antheraea Polyphemus* with homology to the *CD36* family of membrane proteins. J Biol Chem. 272: 14792-14799

Rutzler M and Zwiebel LJ. (2005) Molecular biology of insect olfaction: recent progress and conceptual models. J Comp Physiol A Neuroethol Sens Neural Behav Physiol. 191: 777–790

Sajwan S, Takasu Y, Tamura T, Uchino K, Sezutsu H and Zurovec M. (2013) Efficient disruption of endogenous *Bombyx* gene by TAL effector nucleases. Insect Biochem Mol

Biol. 43: 17–23

Sato K, Tanaka K and Touhara K. (2011) Sugar-regulated cation channel formed by an insect gustatory receptor. Proc Natl Acad Sci USA. 108(28): 11680-11685

Schnee C, Kollner TG, Held M, Turlings TCJ, Gershenzon J, and Degenhardt J. (2006) The products of a single maize sesquiterpene synthase form a volatile defense signal that attracts natural enemies of maize herbivores. Proc Natl Acad Sci USA. 103: 1129–1134.

Schornack S, Meyer A, Romer P, Jordan T and Lahaye T. (2006) Gene-for-gene-mediated recognition of nuclear-targeted *AvrBs3*-like bacterial effector proteins. J Plant Physiology. 163: 256 — 272

Shanbhag SR, Muller B, Steinbrech RA (1999). Atlas of olfactory organs of *Drosophila melanogaster* 1. Types, external organization, innervation and distribution of olfactory sensilla. Int J Insect Morphol Embryol. 28: 377–397.

Shirai Y. (1998) Laboratory evaluation of fight ability of the Oriental corn borer, Ostrinia furnacalis (Lepidoptera: Pyralidae). Bull Entomol Res. 88: 327–333.

Silbering AF and Benton R. (2010) Ionotropic and metabotropic mechanisms in chemoreception: ‘chance or design’? EMBo reports. 11(3): 173-179

Stamatakis A, Aberer AJ, Goll C, Smith SA, Berger SA et al. (2012) RAxML-light: a tool for computing terabyte phylogenies. Bioinformatics. 28(15): 2064-2066

Stensmyr MC,Dweck HK, Farhan A, Ibba I, Strutz A, Mukunda L, Linz J, Grabe V, Steck K, Lavista-Llanos S, Wicher D, Sachse S, Knaden M, Becher PG, Seki Y, Hansson BS. (2012) A conserved dedicated olfactory circuit for detecting harmful microbes in *Drosophila*. Cell. 151: 1345–1357

Sudhakar VS. Akella, William DJ Kirk, Yao-bin Lu, Tamotsu Murai,Keith FA Walters,

James GC Hamilton mail. (2007) Identification of the aggregation pheromone of the melon thrips, *Thrips palmi*. PLoS ONE. DOI: 10.1371/journal.pone.0103315

Takagi M, Hirose Y, Yamasaki M; Scriber JM, Tsubaki Y and Lederhouse RC. (1995) Antipredator defense in Papilio larvae: Effective or not? Scientific Publishers, Inc. {a}, Gainesville, Florida, USA.

Takanashi T et al. (2006) Unusual response characteristics of pheromone-specific olfactory receptor neurons in the Asian corn borer moth, *Ostrinia furnacalis*. J Exp Biol. 209: 4946–4956

Takanashi T, Ohno S, Huang YP, Tatsuki S, Honda H et al. (2000) A sex pheromone component novel to Ostrinia identified from *Ostrinia latipennis* (Lepidoptera: Crambidae). Chemoecology. 10: 143-147

Tanaka K, Uda Y, Ono Y, Nakagawa T, Suwa M et al. (2009) Highly selective tuning of a silkworm olfactory receptor to a key mulberry leaf volatile. Curr Biol. 19: 881-890

Turlings TCJ, Bernasconi ML, Bertossa R, Bigler F, Caloz G, and Dorn S. (1998). The induction of volatile emissions in maize by three herbivore species with different feeding habits: Possible consequences for their natural enemies. Biol Control. 11:122–129.

Turner SL and Ray A. (2009) Modification of CO2 avoidance behaviour in *Drosophila* by inhibitory odorants. Nature. 461: 277–281

Urnov FD, Rebar EJ, Holmes MC, Zhang HS and Gregory PD. (2010) Genome editing with engineered zinc finger nucleases. Nat Rev Genet. 11: 636–646

Vieira FG and Rozas J. (2011) Comparative genomics of the odorant-binding and chemosensory protein gene families across the Arthropoda: origin and evolutionary history of the chemosensory system. Genome Biol Evol. 3: 476-490

Vogt RG and Riddiford LM. (1981) Pheromone binding and inactivation by moth antennae. Nature. 293: 161–163

Vosshall LB and Stocker RF. (2007) Molecular architecture of smell and taste in *Drosophila*. Annu Rev Neurosci. 30: 505–533

Wang Z, Zhou D, Song Y, Li B, Zhang G, Gao S, Liu Y, Zheng L, Wang Y, Xie W et al. (1994) Studies on behavior of dispersal and possibility of migration in adult overwintering generation Asian corn borer by using release-and-recapture technique. Acta Phytophyl Sinica. 21:25–31.

Wanner KW and Robertson HM. (2008) The gustatory receptor family in the silkworm moth *Bombyx mori* is characterized by a large expansion of a single lineage of putative bitter receptors. Insect Mol Biol. 17(6): 621-629

Wanner KW, Anderson AR, Trowell SC, Theilmann DA, Robertson HM and Newcomb RD. (2007) Female-biased expression of odourant receptor genes in the adult antennae of the silkworm *Bombyx mori*. Insect Mol Biol. 16: 107–119

Wanner KW, Nichols AS, Allen JE, Bunger PL, Garczynski SF et al. (2010) Sex Pheromone Receptor Specificity in the European Corn Borer Moth, *Ostrinia nubilalis.* PLOS ONE. 5(1): 8685.

Watanabe T, Ochiai H, Sakuma T, Horch HW, Hamaguchi N et al. (2012) Non-transgenic genome modifications in a hemimetabolous insect using zinc-finger and TAL effector nucleases. Nature Communication. 3:1017

Witte V, Disney RHL, Weissflog A & Maschwitz U. (2010) Studies in European ant-decapitating flies (Diptera: Phoridae): ant alarm pheromone as host finding cue in Pseudacteon brevicauda, a parasite of Myrmica rubra (Formicidae: Myrmicinae). J Nat History. 44: 905–912

Yang Y and Smith SA. (2014) Orthology inference in nonmodel organisms using transcriptomes and low-coverage genomes: improving accuracy and matrix occupancy for phylogenomics. Mol Biol and Evo. 31(11): 3081-3092

Zhai B and Chen R. (1989) Flight capacity of Asian corn borer (*Ostrinia furnacalis* Guenée). J Jilin Agric Sci. 1:40–46.

Zhang J, Liu CC, Yan SW, Liu Y, Guo MB et al. (2013) An odorant receptor from the common cutworm (*Spodoptera litura*) exclusively tuned to the important plant volatile cis-3-Hexenyl acetate. Insect Mol Biol. 22(4): 424-432

Zhang SF, Zhang Z, Wang HB and Kong XB. (2014) Antennal transcriptome analysis and comparison of olfactory genes in two sympatric defoliators, *Dendrolimus houi* and *Dendrolimus kikuchii* (Lepidoptera: Lasiocampidae). Insect Biochem Mol Biol. 52: 69-81