Comparative Biochemistry and Physiology - Part D 25 (2018) 73–85

Contents lists available at ScienceDirect

Comparative Biochemistry and Physiology - Part D

journal homepage: www.elsevier.com/locate/cbpd

Identification and characterization of chemosensory gene families in the bark beetle, *Tomicus yunnanensis*

Nai-Yong Liu[a], Zong-Bo Li[a], Ning Zhao[a], Qi-Sheng Song[b], Jia-Ying Zhu[a,*], Bin Yang[a,*]

[a] Key Laboratory of Forest Disaster Warning and Control of Yunnan Province, Southwest Forestry University, Kunming 650224, China
[b] Division of Plant Sciences, University of Missouri, 1-31 Agriculture Building, Columbia, MO 65211, USA

ARTICLE INFO

Keywords:
Tomicus yunnanensis
Developmental transcriptome
Digital gene expression
Chemosensory gene
Expression profile

ABSTRACT

The bark beetle, *Tomicus yunnanensis* (Coleoptera: Scolytinae), is a seriously destructive pest of Yunnan pine (*Pinus yunnanensis*) and is distributed solely in Southwestern China. It has been a challenge to control this pest owing to its resistance to chemical pesticides, which have been used as the main control strategy of this species in recent years. Since this approach will continue until an alternative mitigation strategy is implemented, it is essential to develop novel or improved biocontrol approaches. In the current study, we aimed to identify most, if not all, of the bark beetle's chemosensory genes, and to address their respective phylogenetic relationships and expression characteristics. Digital gene expression (DGE) profiling and a comparison of the profiles at three developmental stages yielded 40,287,265 clean reads and a large number of differentially expressed genes (DEGs), with 21 up- and 20 down-regulated DEGs involved in chemoreception. Transcriptome of the three mixed stages revealed a total of 80 transcripts encoding chemosensory-related proteins comprising 45 odorant-binding proteins (OBPs), 12 chemosensory proteins (CSPs), 20 receptor proteins [9 odorant receptors (ORs), 8 gustatory receptors (GRs) and 3 ionotropic receptors (IRs)] and 3 sensory neuron membrane proteins (SNMPs). As many as 38 full-length sequences were acquired with a combination of transcriptomic analysis and rapid amplification of cDNA ends (RACE) strategy. Phylogenetic analysis showed that *T. yunnanensis* OBPs were clustered into four subgroups: 27 Minus-C OBPs, 5 antennal binding proteins (ABPIIs), 10 Classic OBPs and one Plus-C OBP; meanwhile, the ORs were grouped into four clades (1, 2, 7b and Orco). Expression profiles revealed that 66 of 80 genes were detected in the three DGE libraries, and 15 soluble olfactory proteins were antennae-predominant, possibly guiding olfactory-associated behaviors of this beetle. Taken together, our study has provided valuable data for further functional studies of this beetle and will facilitate the identification of potential molecular targets associated with chemosensory reception for use in biocontrol strategies.

1. Introduction

The bark beetle, *Tomicus yunnanensis*, belongs to the family Scolytinae and is a secondary pine pest in Southwestern China. The larvae and adults of this beetle generally live in the tree trunks or branches of *Pinus yunnanensis*, primarily feeding on host living shoots and ultimately leading to the death of the whole tree (Kirkendall et al., 2008; Lu et al., 2010). Over the past three decades, pine forests have been damaged greatly by this beetle, despite attempts to control its spread using chemical pesticides (Wang et al., 2015). In Yunnan Province alone, > 200,000 ha of *P. yunnanensis* forests have been infested with this pest to date; this has consequently led to withered pines over 93,000 ha (Liu et al., 2010; Lu et al., 2014). Thus, integrated pest management (IPM) strategies are urgently needed to effectively control this pest and avoid problems, such as pesticide resistance, non-target effects and environmental issues, caused by the chemical control method.

Semiochemicals, including plant volatiles and pheromones, are essential for mediating insect behaviors by serving as attractants or repellents. They could be used as effective biological agents to reduce the use of pesticides in IPM strategies by directly or indirectly manipulating communication between insects (Hansson and Stensmyr, 2011; Kergunteuil et al., 2015; Li et al., 2015; Liu et al., 2015a; Liu et al., 2015b). In addition to the direct use and large-scale screening of semiochemicals, understanding the molecular basis underlying chemosensory perception and recognition will undoubtedly facilitate the management of this pest. Such communication is associated with multiple chemosensory gene families residing in olfactory or non-olfactory organs, comprising soluble olfactory proteins, i.e. odorant binding proteins (OBPs) and chemosensory proteins (CSPs), and

* Corresponding authors.
E-mail addresses: jyzhu001@gmail.com (J.-Y. Zhu), yangbin48053@163.com (B. Yang).

https://doi.org/10.1016/j.cbd.2017.11.003
Received 17 August 2017; Received in revised form 6 November 2017; Accepted 7 November 2017

chemosensory membrane proteins, i.e. odorant receptors (ORs), gustatory receptors (GRs), ionotropic receptors (IRs), and sensory neuron membrane proteins (SNMPs) (Sanchez-Gracia et al., 2009; Vogt et al., 2009; Vieira and Rozas, 2011). These gene families could be used to screen potential behaviorally active compounds by reverse and conventional chemical ecological approaches (Leal et al., 2008).

The OBPs and CSPs are small (10–30 kDa), water-soluble proteins and are present primarily in olfactory- or gustatory-related organs guiding insect behaviors (Vieira and Rozas, 2011). The presence of signal peptides and conserved cysteines is generally considered as typical characteristics of these two families, and is used as the main criteria for gene identification (Pelosi et al., 2014; Pelosi et al., 2017). In Coleopteran CSPs, a strict residue-motif pattern with four conserved cysteines (C1-X_6-C2-X_{18}-C3-X_2-C4) is present (Xu et al., 2009; Dippel et al., 2014). Similarly, the OBP family in Coleoptera also has distinctively conserved numbers of cysteine residues. Accordingly, this family is classified into four sub-groups: Minus-C, Classic, Plus-C OBPs and antennal binding proteins (ABPIIs) (Dippel et al., 2014). In insects, many studies have indicated that OBPs and CSPs are distributed broadly among various tissues, and that they have diverse functional roles, including olfactory and non-olfactory functions (Pelosi et al., 2017).

In addition to the above two families of chemosensory genes, three receptor super-families of ORs, GRs and IRs are also responsible for effective communication of insects (Clyne et al., 1999; Clyne et al., 2000; Benton et al., 2009). Among these, ORs are seven transmembrane proteins and respond to pheromones and general odors by coexpression with a conserved co-receptor, Orco (Clyne et al., 1999; Chang et al., 2017; de Fouchier et al., 2017). In comparison with the OR family, the GRs are mainly responsible for the detection of sugars, bitter compounds and carbon dioxide (Jones et al., 2007; Slone et al., 2007; Xu et al., 2016). The IRs, which belong to the ionotropic glutamate receptor (iGluR) family, are three transmembrane receptors and sense amines, acids, general odors and sex pheromones (Benton et al., 2009; Ai et al., 2010; Silbering et al., 2011; Koh et al., 2014); they are also involved in gustation, thermosensation and hygrosensation (Knecht et al., 2016; Ni et al., 2016). Last, the SNMPs, members of the CD36 family, comprise two orthologs of SNMP1 and SNMP2, with the presence of several copy numbers for each ortholog in Coleoptera (Nichols and Vogt, 2008). This type of protein is essential for the sensation and recognition of sex pheromones as demonstrated in *Drosophila melanogaster* (Jin et al., 2008) and *Heliothis virescens* (Pregitzer et al., 2014).

Coleoptera is the largest order of insects containing forest, agricultural and stored-grain pests. With the increased availability of genome sequences, including those from *Tribolium castaneum* (*Tribolium* Genome Sequencing Consortium, 2008) and *Dendroctonus ponderosae* (Keeling et al., 2013), and a number of transcriptomes from coleopterans, growing numbers of genes associated with chemoreception have been identified, with a small number being functionally characterized (Engsontia et al., 2008; Andersson et al., 2013; Wang et al., 2014; Dippel et al., 2016; Hu et al., 2016). However, little is known about the chemosensory components of *T. yunnanensis*.

In a previous study, complete or partial sequences of 11 OBPs, 8 CSPs, 18 ORs and 8 GRs were retrieved from the head transcriptome of *T. yunnanensis* (Zhu et al., 2012). In this study, using integrated transcriptome and digital gene expression (DGE) sequencing, we extended and characterized several key chemosensory gene families of *T. yunnanensis* as potential molecular targets for pest control. The results of this study provide a basis to understand chemosensory recognition in the bark beetle and to highlight new potential molecular candidates for the control of this beetle.

2. Materials and methods

2.1. Insects and tissue collection

Adults of *T. yunnanensis* were collected from the host *P. yunnanensis* in Qujing city, Yunnan Province, China. Then, they were maintained in glass cages and fed with fresh host plants in the laboratory at room temperature and a photoperiod of 12:12 h light/dark.

For rapid amplification of cDNA ends (RACE), about 200 antennae with a sex ratio of 1:1 (male:female) were collected. For quantitative real-time PCR (qPCR), various tissues including antennae, heads (without antennae), thorax, abdomens, legs and wings from female and male adults were collected and immediately immersed in liquid nitrogen.

Table 1
Summary of digital gene expression data from different samples of *T. yunnanensis*.

Tissue category	Larva	Pupa	Adult
Raw read	13,021,670	13,429,571	13,890,736
Clean read	12,996,403	13,417,026	13,873,886
GC content (%)	45.60	42.15	42.97
Q20 (%)	97.96	98.07	98.04
Q30 (%)	94.20	94.43	94.35
Containing N	796 (0.01%)	794 (0.01%)	1,000 (0.01%)
Low quality	20,020 (0.15%)	9,617 (0.07%)	10,830 (0.08%)
Adapter related	4,451 (0.03%)	2,134 (0.02%)	5,020 (0.04%)

2.2. Total RNA extraction and first-strand cDNA synthesis

Total RNA was extracted from collected tissues using TRIzol Reagent (Ambion, Life Technologies, Carlsbad, CA, USA) following the manufacturer's instructions. RNA quality and integrity were measured by using NanoDrop™ 1000 Spectrophotometer (Thermo Fisher Scientific, USA) and Agilent 2100 Bioanalyzer System (Agilent Technologies, USA), respectively. Genomic DNA was digested by treatment with DNase I (Fermentas, Thermo Fisher Scientific, USA). First-strand cDNA was synthesized with a First Strand cDNA Synthesis Kit (TaKaRa, Dalian, Liaoning, China). RACE template was synthesized by using a SMART RACE cDNA Amplification Kit (Clontech, Mountain View, CA, USA) according to the manufacturer's protocols. The synthesized cDNA templates were stored at − 20 °C.

2.3. Library construction and sequencing

mRNAs, in the total RNA of larvae, pupae and adults, were enriched using magnetic bead with Oligo (dT). Then, they were cut into short fragments with fragmentation buffer, following the protocol in the NEBNext® Ultra™ RNA Library Prep Kit for Illumina® (NEB, USA). Using the fragments as templates first-strand cDNA was synthesized, which was then used to synthesize the second-strand cDNA with DNA polymerase I and RNase H. After purification with AMPure XP beads, the second-strand cDNA was end-repaired, a-tailed and ligated with indexed adapters. Suitable size products were selected and amplified by PCR and purified with AMPure XP beads for the creation of digital gene expression (DGE) libraries.

In addition, transcriptomic library was constructed as described above using a mix of equal amounts of mRNA from larvae, pupae and adults. The library was sequenced using Illumina HiSeq 2000 platform and the resulting raw data of DGE and transcriptomic libraries were deposited in the National Center for Biotechnology Information (NCBI) Short Read Archive (SRA) under the accession numbers of SRX3070092 and SRX2518383, respectively. The raw reads were processed through a stringent filtering process to discard low quality reads. The resulting clean reads of transcriptomic library were assembled into unigenes using Trinity v2012-10-05 (Grabherr et al., 2011). For DGE analysis, the resulting clean reads were then mapped to the transcriptomic unigenes using the RSEM software with default parameters (Li and Dewey, 2011). The number of read counts normalized by TMM for each mapped gene was used to calculate gene expression levels following the FPKM (fragments per kilobase of transcript sequence per millions base