

GOPEN ACCESS

Citation: Ma L, Li Z-Q, Bian L, Cai X-M, Luo Z-X, Zhang Y-J, et al. (2016) Identification and Comparative Study of Chemosensory Genes Related to Host Selection by Legs Transcriptome Analysis in the Tea Geometrid *Ectropis obliqua*. PLoS ONE 11 (3): e0149591. doi:10.1371/journal.pone.0149591

Editor: J Joe Hull, USDA-ARS, UNITED STATES

Received: October 9, 2015

Accepted: February 1, 2016

Published: March 1, 2016

Copyright: © 2016 Ma et al. This is an open access article distributed under the terms of the <u>Creative</u> <u>Commons Attribution License</u>, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Funding: This work was supported by Modern Agricultural Industry Technology System (CARS-23), Open Foundation of State Key laboratory for Biology of Plant Diseases and Insect Pests (SKLOF201517) and Zhejiang Provincial Natural Science Foundation (LY16C140003). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interest exists.

RESEARCH ARTICLE

Identification and Comparative Study of Chemosensory Genes Related to Host Selection by Legs Transcriptome Analysis in the Tea Geometrid *Ectropis obliqua*

Long Ma¹°, Zhao-Qun Li¹°, Lei Bian¹, Xiao-Ming Cai¹, Zong-Xiu Luo¹, Yong-Jun Zhang²*, Zong-Mao Chen¹*

1 Key Laboratory of Tea Biology and Resources Utilization, Ministry of Agriculture, Tea Research Institute, Chinese Academy of Agricultural Sciences, Hangzhou, China, 2 State Key Laboratory for Biology of Plant Diseases and Insect Pests, Institute of Plant Protection, Chinese Academy of Agricultural Sciences, Beijing, China

• These authors contributed equally to this work.

* zmchen2006@163.com (ZMC); yjzhang@ippcaas.cn (YJZ)

Abstract

Host selection by female moths is fundamental to the survival of their larvae. Detecting and perceiving the non-volatile chemicals of the plant surface involved in gustatory detection determine the host preference. In many lepidopteran species, tarsal chemosensilla are sensitive to non-volatile chemicals and responsible for taste detection. The tea geometrid Ectropis oblique is one devastating chewing pest selectively feeding on limited plants, requiring the specialized sensors to forage certain host for oviposition. In present study, we revealed the distribution of chemosensilla in the ventral side of female fifth tarsomere in E. obligua. To investigate its molecular mechanism of gustatory perception, we performed HiSeg 2500 sequencing of the male- and female- legs transcriptome and identified 24 candidate odorant binding proteins (OBPs), 21 chemosensory proteins (CSPs), 2 sensory neuron membrane proteins (SNMPs), 3 gustatory receptors (GRs) and 4 odorant receptors (ORs). Several leg-specific or enriched chemosensory genes were screened by tissue expression analysis, and clustered with functionally validated genes from other moths, suggesting the potential involvement in taste sensation or other physiological processes. The RPKM value analysis revealed that 9 EobIOBPs showed sex discrepancy in the leg expression, 8 being up-regulated in female and only 1 being over expressed in male. These female-biased EobIOBPs indicated an ecological adaption related with host-seeking and oviposition behaviors. Our work will provide basic knowledge for further studies on the molecular mechanism of gustatory perception, and enlighten a host-selection-based control strategy of insect pests.

Introduction

Herbivorous insects must locate and identify host plants to meet their biological requirements and the demand of reproduction, while the process of host selection for feeding and oviposition involves foraging, landing, contact evaluation and final determination [1]. Olfaction and taste perception play crucial roles in chemical detection and discrimination of the host. As many lepidopteran species are designated to use a limited range of host plants, detecting and perceiving the semiochemicals from the host are particularly important. Generally, insects utilize their sensitive and selective antennae to detect air borne odorant molecules and guide the location from distance, while the subsequent contact evaluation of non-volatile chemicals involved in gustatory detection determines the host preference. Typically, insect contact chemoreceptors, derived from mechanosensory bristles and mainly scattered on the legs, the proboscis, the maxillary and the labial palps, are sensitive to phagostimulants (e.g. sugars, oviposition stimulants and amino acid) [2-5]. Many studies have confirmed the chemosensilla on legs play a principal role in perceiving phytochemical compounds after the insects land on the leaves and start drumming on the surface with the tarsi of prothoracic legs [6–9]. In Papilionidae (such as Papilio xuthus, Heliconius melpomene and Papilio polytes), female butterflies perceive oviposition stimulant by the chemosensilla located on the ventral side of their foreleg tarsus and further determine the suitable feeding plant for larvae [10-12]. In other lepidopteran species (such as *Mnesampela privata*, *Helicoverpa armigera* and *Heliothis virescens*) (Fig 1), tarsal chemosensilla of the prothoracic legs are sensitive to some salts, sugars and amino acids, which indicates a role in the assessment of food materials [13-15]. Legs of Drosphila, functioning as gustatory organ and being responsible for tastant recognition, contain several taste sensilla and make the initial contact with potential food resources [16].

Gustatory perception enables insects to efficiently detect the non-volatile chemosensory information, guiding the feeding and oviposition behaviors of insects. Gustatory stimuli are recognized by gustatory receptors (GRs), which share a common 7-transmembrane protein and plus C-terminal domain with olfactory receptors (ORs) but are more diverse [17–19]. According to the ligand profiles, GRs are classified into sugar GR genes [20–22], bitter taste receptors [23–25] and carbon dioxide receptors [26, 27]. Insect OBPs (odorant binding protein) are small, hydrophilic proteins, ferrying the hydrophobic semiochemicals across the sensilla lymph to olfactory receptors [28]. In Lepidoptera, two subfamilies of OBPs, general odorant-binding proteins (GOBPs) and pheromone binding proteins (PBPs), are responsible

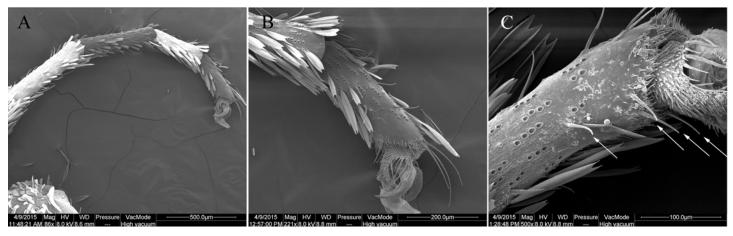


Fig 1. Scanning electron micrographs of foreleg tarsus and chemosensilla of the fifth tarsomere of adult female *E. obliqua*. (A) Foreleg tarsus. (B) Magnification of the fifth tarsomere by SEM. (C) Chemosensilla distributed in the ventral side of a female fifth tarsomere.

doi:10.1371/journal.pone.0149591.g001

for recognizing and transporting host odorants and pheromones, respectively [29, 30]. Although OBPs were identified originally in olfactory system, some OBPs appear in gustatory sensilla [31–34]. Recent studies have reported an unexpected role of OBP involved in gustatory perception. In *Drosphila*, *OBP49a* is indispensable for the suppression of sweet taste by bitter chemicals, and the loss of *OBP49a* will impair the inhibition [35]; two OBPs encoded by *OBP57e* and *OBP57d* are not only involved in taste perception but could also change the behavioral response to the host odors, and the mutation in these OBPs could shift the host preference [36]. CSPs (chemosensory protein) are also small soluble proteins enriched in the sensilla lymph [37], unlike OBPs which are considered antennae-specific, CSPs are much smaller (10–15 KDa) but widely expressed in many tissues, including antennae [38, 39], proboscises [40, 41], legs [42], wings [43], etc. In general, CSPs are believed to enhance the solubility of semiochemicals and deliver them to chemosensory receptors [44–47], however, the function remains unclear. SNMPs (sensory neuron membrane proteins), transmembrane domain-containing proteins, are expressed in pheromone-sensitive hair and proposed to participate in pheromone and general odorant reception [48–50].

The tea geometrid, *Ectropis obliqua* Prout (Lepidoptera: Geometridae) is one devastating chewing pest throughout the tea plantations in China. *E. obliqua* larvae, voracious worms that feed exclusively on tea leaves and tender buds, produce 6–7 generations throughout the growing season of tea, and eat the entire leaves of tea plants in several infestations, causing severe yield loss and deterioration in commercial tea quality [51]. As one member of Lepidoptera, the ability to locate suitable host plants is fundamental to the survival of their offspring, because the small larvae cannot easily forage for alternate hosts [52]. Field observation of the egg-laying behavior in *E. obliqua* reveals the female adults actively forage for the conspecific larvae-infested plants and preferentially oviposit on the splits of branches or the cracks between leaves and branches, during which the phytochemical compounds in leaves are the dominant attraction for oviposition [53, 54]. Thus, a deep insight into the molecular mechanisms of gustatory perception will largely contribute to the integrated management of *E. obliqua*.

Previous studies of gustatory system mainly focused on the electrophysiological recording of gustatory sensilla and behavioral response, and were confined to the limited butterflies and model species, such as *Bombyx mori* and *Drosophila melanogaster*, whose genomes were available. However, few studies were focused on agricultural pests and the underlying genetic mechanisms were poorly understood, primarily because of the difficulty in detecting the genes involved in taste sensation. Here we performed Illumina HiSeq 2500 sequencing of the transcriptome of adult male and female legs of this devastating economic pest, and reported the identification of 24 candidate OBPs (including 4 PBPs), 21 CSPs, 2 SNMP, 3GRs and 4 ORs genes. The transcripts analysis of RPKM metric was performed to highlight the most abundant genes and outline the comparative gene expression between samples. Through RT-PCR and real-time quantitative-PCR, we analyzed expression profiles of the chemosensory genes and their putative functions in chemoreception.

Materials and Methods

Insect material and RNA preparation

E. obliqua colony used in this study was originally collected from the experimental tea plantation of the Tea Research Institute of the Chinese Academy of Agricultural Sciences (Hangzhou, Zhejiang, China). Newly hatched larvae were transferred onto fresh tea shoots in enclosed nylon mesh cages ($70 \times 70 \times 70$ cm) and kept in a climate-controlled room at $25\pm1^{\circ}$ C and $70\pm5\%$ relative humidity under a photoperiod of 14:10 (light: dark). After pupation, male and female pupae were separated based on the body size and morphological characters [55], and kept in darkness until eclosion. After emergence, adult moths were given a 10% honey solution. For transcriptome sequencing, legs of unmated male and female individuals were collected 2–3 days after eclosion. For qPCR and RT-PCR, adult tissues were collected and divided into female antennae, male antennae, male legs, female legs, heads (without antennae), thoraxes, abdomens and wings. All tissues were immediately snap-frozen in liquid nitrogen, and stored at –80°C until extraction. Total RNA was extracted using Trizol reagent (Invitrogen, Shanghai, China). The integrity of RNA samples was detected by gel electrophoresis, and a NanoDrop 2000 spectrophotometer (NanoDrop, Wilmington, DE, USA) was used to determine RNA quantity.

Scanning electron microscopy (SEM)

Prothoracic tarsi of female moths were cut using a scalpel. The samples were first soaked in 70% ethanol for 3 h and then descaled in an ultrasonic bath for 10 s. After gradient elution in an ethanol series (80%, 90%, 95% and 100%), the samples were dried at 25°C overnight. The samples were mounted on stainless steel holders and coated with gold-palladium. Photomicrographs were viewed with a Quauta 200 FEG Environmental scanning electron microscopy (ESEM) (Feicompany, USA).

cDNA library construction and Illumina sequencing

cDNA library was constructed using 5 µg total RNA extracted from approximately 100 male or female legs. Oligo (dT) linked beads were used to isolate the mRNA from the total RNA (Illumina; San Diego, CA, USA). The isolated mRNA strands were digested into short fragments with Fragmentation Buffer. The fragmented mRNAs were used as templates to construct the cDNA libraries using a Truseq RNA Sample prep Kit (Illumina, San Diego, USA) following the manufacturer's instruction. Briefly, random hexamer-primers were used for first-strand cDNA, followed by second-strand cDNA synthesis using DNA Polymerase I and RNase H (Invitrogen, Carlsbad, CA). After end-repairing and ligation of adaptors, the products were amplified by PCR and purified with QIAquick PCR purification kit (Qiagen, Hilden, Germany) to construct a cDNA library. Then the two libraries created from the legs of male and female *E. obliqua* were sequenced on the Illumina HiSeq 2500 platform at Shanghai Majorbio Biopharm Technology Co., Ltd, generating 100 bp paired-end raw reads.

Transcriptome assembly and functional annotation

Clean-read datasets were obtained from the raw reads through the following procedures: first, remove the low quality reads (the bases with sequencing error rates more than 1% are over half in the read) and adaptor sequences; second, filter out the sequences with N (uncertain bases) exceeding 10%. These treatments were performed through SeqPrep (https://github.com/ jstjohn/SeqPrep) and Sickle (https://github.com/najoshi/sickl). The Q20, Q30, GC-content and sequence duplication level of the clean data were summarized simultaneously. Transcriptome assembly was performed through Trinity (trinityrnaseq-r2013-02-25). The trinity outputs made up two classes of unigenes: the consensus cluster sequences and singletons. To acquire the functional annotation, transcripts larger than 150 bp were submitted to NCBI BlastX homology search against a pooled database of non-redundant (nr) and SwissProt protein sequences with an E-value \leq 1.0E–5. We further imported the blast results into Blast2GO pipeline for Gene Ontology (GO) annotation. Then, the open reading frame (ORF) of genes was predicted by ORF finder (http://www.ncbi.nlm.nih.gov/gorf/gorf.html), and the SignalP 4.1 server (http://www.cbs.dtu.dk/services/SignalP/) was used to predict the signal peptide sites in the protein sequences. In order to explore the putative chemosensory receptors, the available GRs and ORs sequences from B. mori, D. melanogaster and Tribolium castaneum were

submitted as queries to run local homology search against the assembled transcripts using the BioEdit Sequence Alignment Editor 7.1.3.0, in avoidance of inaccurate gene annotation.

RACE-PCR and sequence verification

The retrieved unigenes did not always represent full-length transcripts and some contained only part. To confirm the reliability of output sequences and for better resolution of phylogenetic analysis, partial sequences of candidate chemosensory genes were extended using RACE-PCR, and subsequently followed by full-length assembly and cloning. Total RNA extracted from adult antennae or legs was used to synthesize 5' and 3' RACE cDNA templates through SMART RACE cDNA Amplification kit (Clontech). Primers were designed manually and listed in <u>S1 Table</u>. The RACE-PCR was operated in the means of touchdown following the manual of Advantage 2 PCR kit (Clontech, CA, USA). The PCR products were subcloned into pGEM-T (promega) and the inserts were sequenced using M13 primers. Afterwards, the 3' and 5' end sequences were aligned by BlastX against the GenBank to validate its correctness, and were sequence-matched to obtain the full length. Open reading frames (ORFs) of genes were predicted by ORF finder (<u>http://www.ncbi.nlm.nih.gov/gorf/gorf.html</u>). Then gene specific primers were designed using the Primer 5.0 software, and ORF sequences were amplified and verified by sequencing as mentioned before.

Phylogenetic analysis

Multiple alignments of the complete OBPs and CSPs amino acid sequences were performed by ClustalX 2.0 and further edited by GeneDoc 2.7. The phylogenetic trees were constructed by MEGA6.0 using the Neighbor-joining method with a p-distance model and a pairwise deletion of gaps. Bootstrap support was assessed by a boot strap procedure based on 1000 replicates. The data sets of OBPs and CSPs sequences which were chosen from other Lepidoptera species were listed in <u>S3</u> and <u>S4</u> Tables separately.

Comparative analysis of transcript abundance

To compare the differential expression of chemosensory genes between samples, the transcript expression abundances were calculated according to the metric RPKM (Reads per Kilobase per million mapped Reads) method, based on the formula: RPKM (A) = $(10^6 \times C \times 10^3)/(N \times L)$, where RPKM represents the expression of target gene A, C is the number of reads that are uniquely mapped to gene A, N is the whole number of reads that are uniquely aligned to all transcripts and L is the number of bases in gene A. RPKM metric is capable of eliminating the discrimination in gene lengths and sequencing discrepancies, which makes it possible to compare gene expression between samples [56]. Differentially expressed genes (DEGs) were identified by EdgeR (http://www.bioconductor.org/packages/2.12/bioc/html/edgeR.html) according to statistically significant differences with the threshold of false discovery rates (PDR)<0.05 and $|log_2Ratio \ge 1|$ (refer to Benjamini (2001) for details) in the manner of male transcriptome vs. female transcriptome. Subsequently, all DEGs were further annotated by GO and KEGG pathway enrichment analyses.

qRT-PCR analysis and RT-PCR confirmation

The tissue expression profiles of chemosensory genes in different tissues (male antennae, female antennae, male legs, female legs, heads without antennae, thoraxes, abdomens and wings) were measured by real-time qRT-PCR. After the digestion of residual genomic DNA from total RNA with DNase I (Promega), cDNAs were synthesized using 1 µg total RNA from

various tissues by the Fast Quant RT kit (TIANGEN, Beijing, China) following the instruction manual. qRT-PCR was conducted on an Bio-Rad CFX96 Touch Real-Time PCR Detection System. Specific primer pairs were designed by the Primer3 web program (http://primer3.ut.ee/) and listed in <u>S2 Table</u>. The reference gene β -actin (GenBank accession number KT860051) was used for normalization. To make sure that the amplification efficiencies of target genes and reference gene are approximately equal, the efficiency of each primer pair was analyzed by constructing a standard curve with three-fold cDNA dilution series. The qRT-PCR reaction contained 10 µl SuperReal PreMix Plus (TianGen, Beijing, China), 0.6 µl primer on each (10 μ M), 2 μ l sample cDNA (200 ng) and sterile H₂O up to 20 μ l. The qPCR procedure was 95°C for 15 min, followed by 40 cycles of 95°C for 10 s and 60°C for 30 s, melt curves stages at 95°C for 15 s, 60°C for 1 min, and 95°C for 15 s. A blank without template cDNA was included in each experiment serving as a negative control. To check reproducibility, each reaction included three biological replicates and was performed in triplicate (technical replicates). Relative transcript level in each tissue was calculated using the comparative $2^{-\Delta\Delta CT}$ method [57]. Data were first normalized to the endogenous β -actin levels from the same tissue, then the lowest-expression tissue was selected as the calibrator, and the relative expression level among different tissues was assessed by comparing the expression level of each target gene in other parts to that in the lowest one.

RT-PCR was implemented to verify the expression profiles of chemosensory genes. Specific primers were designed by Beacon Designer 7.7 (Premier Biosoft, Palo Alto, CA, USA) and listed in <u>S2 Table</u>. Each PCR reaction contained 200 ng cDNA template, performed by Taq Master Mix (CWBIO, Beijing, China) under general 3-step amplification by 34–36 cycles of 94°C for 20s, 58°C for 30s, 72°C for 40s. PCR products were checked by electrophoresis and further confirmed by sequencing. The β -actin gene served as an endogenous control. Each amplification was performed three times with different biological samples.

Statistical analysis

Data of relative expression levels in various tissue were subjected to one-way analysis of variance (ANOVA), followed by a least significant difference test (LSD) for mean comparison. Differences were considered significant at p<0.05. Data were analyzed using SAS 9.20 (SAS Institute, Cary, North Carolina, USA).

Results

Transcriptome overview

The transcriptome reads data were generated on an Illumina HiSeq 2500 platform using the paired end protocol. A total of 61.9 and 48.1 million raw reads were obtained from the *E. obliqua* female and male-leg libraries. After filtering the low quality and adaptor sequences, 59729104 and 46908620 clean reads were obtained, respectively, and assembled together into 83311 transcripts with an average length of 1402 bp. Of the clean reads, the Q20 percentage (proportion of sequences with a sequencing error rate less than 1%) for both libraries exceeded 98%. The clean reads obtained in this study are available at the NCBI/SRA database under the accession number SRX1502449. The length distribution of transcripts and unigenes was listed in <u>S1 Fig</u>.

The functional annotations of transcripts were performed using the sequence similarity searches against the Nr, SwissProt, KEGG, GO and COG databases with an E-value threshold of 10^{-5} . A percentage of 41.3%, 26.1% and 17.9% transcripts hit in Nr, SwissProt, and KEGG database, respectively. Among the annotated transcripts, 17390 (50.6%) of Nr-hit transcripts had a best match to *B. mori*, followed by 9282 (27.0%) to *Danaus plexippus* and 4924 (14.3%)

to *P. Xuthus* (Fig 2A). GO gene functional classification offers a strictly defined concept to depict the properties of genes and their products. Of the total transcripts, only 83311 (19.8%) could be annotated based on sequence homology, and the assembled transcripts were divided into 3 distinct subsets (Fig 2B). In the term of molecular function, the annotations were mostly enriched in binding (8539 transcripts accounting for 51.7% of the annotated) and catalytic

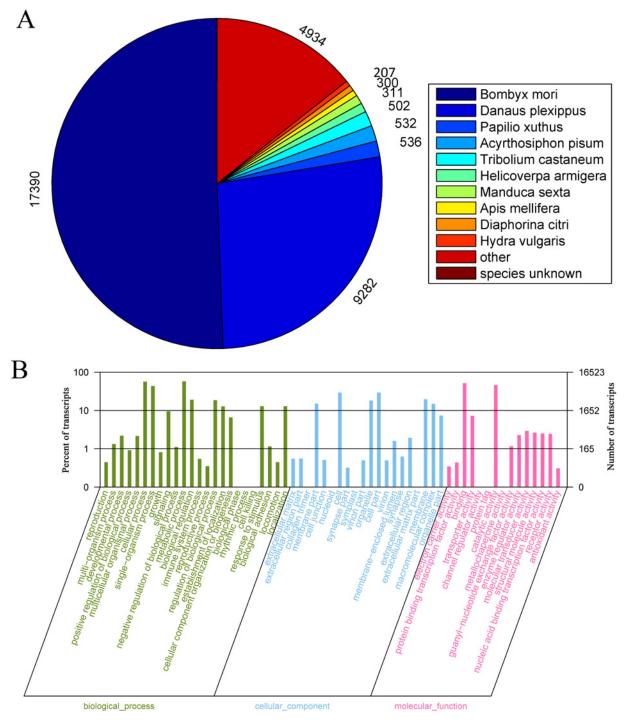


Fig 2. Summary for the annotation of *E. obliqua* legs transcripts. (A) Species distribution of the best Blastx hits. (B) Gene Ontology (GO) classifications of legs transcripts annotated at GO level 2 according to the involvement in biological processes, cellular component and molecular function.

doi:10.1371/journal.pone.0149591.g002

PLOS ONE

activity (7600 transcripts accounting for 46.0%). In biological process category, metabolic and cellular processes occurred most frequently; while the cell, cell part and organelle were most abundant in the category of cellular component.

Identification of putative chemosensory genes

Sequence annotation facilitated the identification of candidate chemosensory genes, generating a total of 24 OBPs (including 4 PBPs), 21 CSPs, 2 SNMP, 3GRs and 4 ORs genes (Table 1). Sequence predication revealed that 20 OBPs included the full-length open reading frame (ORF) with a predicted signal peptide, and the five incomplete OBPs were extended in either 5 or 3 directions by RACE-PCR. The ORF sequences of 24 EoblOBPs were verified by cloning and sequencing and further submitted to GenBank (Table 1). After acquiring the full length sequence, the 24 putative EoblOBPs shared a relative high similarity (34%-92%) to the known lepidopteran OBPs, among which EoblOBP20, EoblPBP1 and EoblPBP2 were matched to the orthologous gene from the Geometridae relative Ascotis selenaria. Based on the number and location of the conserved cysteines, 24 EoblOBPs can be divided into three subsets. EoblOBP8, EoblOBP16 and EoblOBP20 are classified as the Minus-C OBP family, which lack the conserved cysteines C2 and C5 (Fig 3A); while EoblOBP7 and EoblOBP10 belong to the Plus-C OBP family, which have extra 2 cysteines located behind the conserved C6. Moreover, the conserved C2 and C3 of these two Plus-C OBPs are separated by 4 amino acid residues rather than typical 3 in classic OBP, and the conserved C5 and C6 of EoblOBP7 are separated by 7 amino acid residues rather than usual 8 as in the Plus-C OBP (Fig 3A). The 19 EoblOBPs left are classic OBPs with the typical character of six conserved cysteines. The 24 EoblOBPs along with 153 OBPs from 7 other species (including B. mori, H. armigera, Agrotis ipsilon, Spodoptera exigua, Spodoptera litura, etc.) were chosen to construct a phylogenetic tree based on the amino acid sequences. The tree could be classified into several distinct branches: the PBP family, the GOBP family, the Plus-C OBP family and the Minus-C OBP family (Fig 4). On the whole, the identified EoblOBPs were clustered with different orthologous sequences in other species, except that EoblOBP3, EoblOBP6, EoblOBP18 and EoblOBP22 made up a homologous cluster. Besides, nine EoblOBPs were coupled with corresponding homologous OBPs from B. mori in one branch.

By homology analysis, we identified 21 transcripts encoding candidate CSPs in *E. obliqua*, among which 20 of the 21 EoblCSP genes had intact ORF with a signal peptide and four conserved cysteine residues, which represented the typical character of insect CSPs (Fig 3B). Sequence analysis revealed, relative to EoblOBPs, the 21 EoblCSPs were relatively conserved, the average identify of which was 64.8%. In phylogenetic analysis of CSPs from Lepidoptera species, EoblCSPs were spread across branches where they generally formed separate clusters with others. Only *EoblCSP2* and *EoblCSP13* formed one subgroup (Fig 5).

Two SNMPs were identified from our transcriptome and acquired the full length by RACE-PCR. Both *EoblSNMP1* and *EoblSNMP2* shared more than 60% (65% and 69%) identity with the corresponding SNMPs in *S. exigua*. The transcripts encoding chemosensory receptors were initially identified by the keyword search of functional annotation, and further confirmed by the local homology search. Four ORs and three GRs were easily identified with 7 transmembrane domains. From the annotation, *EoblGR2* shared 75.2% identity with *HarmGR9* which had been identified as a sugar receptor [58, 59].

Expression profiles of chemosensory genes

The expression profiles of chemosensory genes (OBPs, CSPs, SNMPs, ORs and GRs) were first examined by qRT-PCR and further confirmed by RT-PCR, illustrating that the majority of

nes in <i>E. obliqua</i> male- and female	Chemosensory genes in E. obliqua m
nes in <i>E. ol</i>	Chemosensory genes in E. of
	Chemosensory gei

					Best Blast P Match	ц	Iden-	RPKIV	RPKM value	log2FC	FDR	up-down-	signifi
	ÖN	(da)		(aa)		value	Â	Female	Male			regulation	-cant
Odorant Binc	Odorant Binding Protein (OBP)	BP)											
OBP3	KT327208	618	Yes	151	gb EHJ67766.1 antennal binding protein [Danaus plexippus]	1e-35	49%	16.49	6.07	1.5832	8.20E- 07	dn	yes
OBP4	KT327209	066	Yes	184	gb AKT26501.1 odorant binding protein 24 [Spodoptera exigua]	2e-123	92%	23.31	43.95	-0.8929	3E-04	down	оп
OBP5	KT327210	733	Yes	146	gb/All00976.1 odorant binding protein [Dendrolimus houi]	2e-40	56%	92.57	42.3	1.2555	2.55E- 07	dn	yes
OBP6	KT327211	660	Yes	153	gb AKT26503.1 odorant binding protein 26 [Spodoptera exigua]	4e-36	43%	4452.85	7846.6	-0.3940	0.16514	down	ou
OBP7	KT327212	591	Yes	145	gb AGK24580.1 odorant-binding protein 4 [Chilo suppressalis]	2e-67	68%	7.68	6.52	0.4119	0.44031	dn	ou
OBP8	KT327213	567	Yes	137	gb/All01007.1/odorant binding protein [Dendrolimus kikuchii]	5e-48	63%	2325.08	3066.73	-0.2268	0.49493	down	ou
ОВР9	KT327214	590	Yes	141	gb AFD34173.1 odorant binding protein 5 [Argyresthia conjugella]	2e-70	67%	67.24	183.35	-1.2827	7.15E- 08	down	yes
OBP10	KT327215	969	Yes	188	gb/All01009.1 odorant binding protein [Dendrolimus kikuchii]	5e-25	36%	1987.8	1775.41	0.2521	0.43405	dn	оп
OBP11	KT327216	560	Yes	142	gb AGM38605.1 odorant binding protein [Chilo suppressalis]	2e-61	64%	83.49	12.22	2.9454	7.87E- 29	dn	yes
OBP12	KT327217	693	Yes	191	gb CAS90131.1 odorant-binding protein 7 [Bombyx mori]	3e-32	53%	4.52	2.91	0.7661	0.14810	dn	оп
OBP13	KT327218	633	Yes	148	gb AAL60415.1 antennal binding protein 4 [Manduca sexta]	8e-75	78%	15.61	1.57	3.4075	4.52E- 17	dn	yes
OBP14	KT327219	669	Yes	151	gb AGP03457.1 SexiOBP11 [Spodoptera exigua]	8e-67	66%	6.84	1.47	2.1597	4.19E- 06	dn	yes
OBP15	KT327220	692	Yes	149	gb AEB54589.1 OBP8 [Helicoverpa armiger]	2e-72	79%	15.02	-	4.1939	5.36E- 20	dn	yes
OBP16	KT327221	1068	Yes	133	gb AAA85090.1 sericotropin [Galleria mellonella]	5e-82	91%	172.91	167.91	0.1214	0.77672	dn	оц
OBP17	KT327222	609	Yes	156	gb AGS36748.1 OBP6 [Sesamia inferens]	1e-32	53%	89.7	38.11	1.2966	3.33E- 08	dn	yes
OBP18	KT327223	1065	Yes	145	gb/All00968.1 odorant binding protein [Dendrolimus houi]	1e-27	40%	30.79	31.06	0.0671	0.93038	dn	ou
OBP19	KT327224	1 090	Yes	155	gb/AKI87962.1 odorant binding protein 1 [Spodoptera litura]	9e-63	71%	3.55	1.29	0.9815	0.02123	dn	ou
OBP20	KT327225	672	Yes	138	gb KOB71255.1 Odorant binding protein [Operophtera brumata]	1e-15	34%	0	0.52	noTest	noTest	noTest	noTest
OBP21	KT762010	597	Yes	147	gb/All00969.1 odorant binding protein [Dendrolimus houi]	1e-47	48%	8.14	3.5	1.3898	0.00131	dn	yes
OBP22	KT762011	592	Yes	150	gb AKT26503.1 odorant binding protein 26 [Spodoptera exigua]	4e-25	36%	0.83	0	noTest	noTest	noTest	noTest
PBP1	KT282990	856	Yes	169	dbj BAF63878.1 pheromone binding protein [Ascotis selenaria]	7e-91	78%	13.21	9.5	0.5794	0.09273	dn	ou
PBP2	KT282991	1178	Yes	164	dbj BAF64703.1 pheromone binding protein 2 [Ascotis selenaria]	2e-69	76%	22.98	27.6	-0.1930	0.62504	down	оп
PBP3	КТ282992	845	Yes	162	gb AIS72934.1 pheromone binding protein 3 [Spodoptera litura]	4e-55	57%	59.18	62.18	0.03240	0.99049	dn	оц

PLOS ONE | DOI:10.1371/journal.pone.0149591 March 1, 2016

GeneName	Accession	Length	CompleteORF	ORF (52)	Best Blast P Match	Ш	Iden-	RPKN	RPKM value	log2FC	FDR	up-down-	signifi
		(da)		(ad)		value	nıy	Female	Male			regulation	-Call
PBP4	KT282993	1115	Yes	170	gb AAF06142.1 pheromone binding protein [Synanthedon exitiosa]	2e-69	63%	15.14	12.09	0.4359	0.20968	dn	ou
chemosenso	Chemosensory Protein (CSP)	(H)											
CSP1	KT282970	526	Yes	124	gb AKT26488.1 chemosensory protein 12 [Spodoptera exigua]	5e-45	61%	1.6	1.13	0.7293	0.48061	dn	оц
CSP2	KT762012	592	Yes	131	gb AHC05672.1 chemosensory protein [Chilo suppressalis]	2e-24	42%	0.17	1.05	-2.2701	0.20298	down	ou
CSP3	KT282971	541	Yes	128	gb ABM67689.1 chemosensory protein 2 [Spodoptera exigua]	2e-46	61%	23512.11	12405.51	1.1052	1.78E- 06	dn	yes
CSP4	KT282972	861	Yes	128	gb AGR39576.1 chemosensory protein 6[Agrotis ipsilon]	4e-64	%92	12169.58	9274.41	0.4949	0.06465	dn	0 L
CSP5	KT282973	543	Yes	148	gb AGY49270.1 chemosensory protein [Sesamia inferens]	1e-52	78%	7.67	4.46	0.8364	0.01005	dn	С
CSP6	KT282974	551	Yes	121	gb ADO95154.1 chemosensory protein [Antheraea yamamai]	1e-40	58%	2771.92	1722.79	0.8649	3E-04	dn	ou
CSP7	KT282975	505	Yes	123	gb KOB75937.1 Chemosensory protein 15 [Operophtera brumata]	6e-56	69%	11.77	7.23	0.8994	0.02407	dn	ou
CSP8	KT282976	725	Yes	125	gb ABM67688.1 chemosensory protein 1 [Spodoptera exigua]	5e-51	%09	1013.44	345.55	1.5931	1.36E- 12	dn	yes
CSP9	KT282977	606	Yes	125	gb All01011.1 chemosensory protein [Dendrolimus houi]	2e-45	59%	0.15	7.66	-5.3451	1.02E- 15	down	yes
CSP10	KT282978	1184	Yes	121	gb KOB75936.1 Chemosensory protein [Operophtera brumata]	8e-67	79%	36.34	45.92	-0.2684	0.42525	down	ou
CSP11	KT282979	563	Yes	112	gb AJP61962.1 chemosensory protein [Phenacoccus solenopsis]	4e-07	41%	0.66	87.3	-6.9890	2.69E- 103	down	yes
CSP12	KT282980	822	Yes	122	gb EHJ73333.1 chemosensory protein 13 [Danaus plexippus]	4e-64	85%	3.77	7.61	-0.8258	0.02528	down	ou
CSP13	KT282981	590	Yes	118	ref[NP_001037068.1 chemosensory protein 7 [Bombyx mori]	1e-16	37%	12.94	3.62	1.9459	2.22E- 09	dn	yes
CSP14	KT282982	734	Yes	107	gb AIW65101.1 chemosensory protein [Helicoverpa armigera]	3e-61	85%	17.32	17.88	-0.0497	0.97152	down	ou
CSP15	KT282983	676	Yes	123	gb AKT26491.1 chemosensory protein 16 [Spodoptera exigua]	6e-37	69%	8.36	175.46	-4.9581	8.21E- 80	down	yes
CSP16	KT282984	612	Yes	123	ref NP_001037067.1 chemosensory protein 8 [Bombyx mori]	1e-53	64%	9.3	4.01	1.2743	3.60E- 05	dn	yes
CSP17	KT282985	1184	Yes	130	gb AHX37218.1 chemosensory protein 1 [Conogethes punctifieralis]	1e-71	%17	2389.76	2554.64	0.0175	-	dn	ou
CSP18	KT282986	648	Yes	126	gb AKI28976.1 chemosensory protein 2 [Bactrocera dorsalis]	7e-32	51%	1640.61	620.15	1.1747	3.23E- 07	dn	yes
CSP19	KT282987	1803	Yes	295	gb AIW65104.1 chemosensory protein [Helicoverpa armigera]	2e-108	65%	42.91	63.32	-0.5197	0.05756	down	ou
CSP20	KT282988	698	Yes	121	gb AEX07265.1 CSP2 [Helicoverpa armigera]	1e-52	78%	40399.5	23894.02	0.8912	2E-04	dn	ou
CSP21	KT282989	789	Yes	128	gb ABM67688.1 chemosensory protein 1 [Spodoptera exigua]	2e-58	65%	283.4	33.05	3.2140	9.14E- 42	dn	yes
Olfactory Receptor (OR)	ceptor (OR)												
ORco	KT373968	1693	Yes	473	gb BAG71418.1 olfactory receptor-2	0.0	85%	0	0.45	noTest	noTest	noTest	noTest

			Manal 4 0040
PLOS ONE	DOI:10.1371/	journal.pone.0149591	March 1, 2016

GeneName	Accession	Length	CompleteORF	ORF	Best Blast P Match	ш	Iden-	RPK	RPKM value	log2FC	FDR	-uwop-dn	signifi
	ÖN	(da)		(aa)		value	nuð	Female	Male			regulation	-cant
OR1	KT860045	2683	Yes	438	ref NP_001091791.1 candidate olfactory receptor [Bombyx mori]	4e-78	40%	1.64	0.54	1.6351	1.63511	dn	yes
OR2	KT860046	1931	Yes	391	gb All01092.1 odorant receptor [Dendrolimus kikuchii]	2e-146	59%	0.46	0.42	noTest	noTest	noTest	noTest
OR3	KT860047	1580	Yes	402	gb AJF23815.1 olfactory receptor OR59 [Planotortrix octo]	0.0	74%	0.74	0.4	noTest	noTest	noTest	noTest
Gustatory Re	Gustatory Receptor (GR)												
GR1	KT860048	1756	Yes	351	gb KDR12697.1 gustatory receptor 1 [Zootermopsis nevadensis]	8e-121	62%	1.21	1.38	-0.1306	0.96147	down	ou
GR2	KT860049	1676	Yes	466	gb AGA04648.1 gustatory receptor [Helicoverpa armigera]	0.0	75%	3.81	5.07	-0.4306	0.21840	down	ou
GR3	KT860050	1589	3' lost	423	gb AGK90023.1 gustatory receptor 1 [Helicoverpa assulta]	0.0	72%	0.37	0.32	noTest	noTest	noTest	noTest
Sensory Neu	Sensory Neuron Membrane Protein (SNMP)	Protein (S	(JMP)										
SNMP1	КТ282969	1908	Yes	524	gb AGN52676.1 sensory neuron membrane protein 1 [Spodoptera exigua]	0.0	65%	0	0.4	noTest	noTest	noTest	noTest
SNMP2	KP684219 *	1885	Yes	516	gb AKN78949.1 sensory neuron membrane protein 2 [Ectropis obliqua]	0.0	%66	92.38	530.35	-2.4820	2.55E- 28	uwop	yes

genes were identified according to statistically significant differences with the threshold of false discovery rates (PDR)<0.05 and |log2Ratio>1|. If no transcript representing the ٩م لام target gene was detected, RPKM value and the following analysis should not be taken into account. Log2FC: told change value after log2 transformat

* represents that the gene has been submitted already.

doi:10.1371/joumal.pone.0149591.t001

Δ		* 20	* 40 *	60	* 80	* 100	*
A		MAKL-FLFVVGVAVSLGIA MVRKFSGLLCCLCVFGISLSDSAIS					FFK 73 LLQCVYR 101
	Foblopp5 .	MINIICEPEUTOTEUUCUIC	AFKINDI KOFYDE	TT VECTAONDM	TAPDUPCTC-	KOKDTVNUNCTER	ALK 70
	EchlOBP6 .	MAKLOCEVELGVLAVESVAS	AATSEEDKNATKAEMLP	VLAE GKEHGV	TEKDVKEAK-	ESNN-VDAINPOFIA	FMK 75
							VFE 67
		MQASILFGLVFVAAGIN MTTATVPVFLALIAVAYCG					FFQ 72
	EoblOBP10:	MVGLGGLLLLVGLOIIASOEOG	PPHGPPPOWANHKCAGPPPAIKN	POKCCEIOOMF	TEEEMASCG-	-INKFEEENROGPPKPPD	NKO-ECLLK 91
	EoblOBP11:	MEK-SYWVVLITAIMIAG	DCDAMTKEOLRKTGKM	LRKOCLGKVGV	EEEKISOIE-	KGKFIEEKDVMCYIAC	VYQ 72
	EoblOBP12:	MNYTLLCLSLTITSVSLLVTCSKLSTET MWSSKINP-GAALAVCVFLLQML	TKVTATTESSNNMKPDDNN	KELDDSKPASD	TSRIGIDTKY	TSNDTNALYDEVMDVLTTC	NESFRIEIS 96 VYT 77
							VII //
	FobloBP15.	-MTNCTTECKYCUMETCETELTILT	T.SATMTMKOT.PSTCKM	MDKSCOPKNNU	FDEKTDRTA-	DOVETEEOEVKOVTAC	IMK 79
		MKTLIVLAVCFVAAQA MYKFGLVCFVLAASVVLKNDA					MLI 67 LYK 77
	EobloBP17:	MIRYUCVIAVILAGFMAAG	VTDEEKKOMHDYLVS	ATESCSKEFGI	PSEDFEKAK-	RNKE-LOSLDPOFVAC	ILK 72
	EoblOBP19:	MKDOISVINSTDYDYDGYG	SGSMGEKFVNSMPK	AADGRYYPYAN	GTNRTRRSE-	PLFSKPDNEOCLSOC	VFA 71
	EoblOBP20:	MFKILLFICAVSAVN	CEIVRTAITLPPEIAFDIAKA	IKEVCV	PEDRVPDIIR	MIREGETNNNTEFKKIIHC	VIK 74
	EODIOBP21: EobloBP22:	MARESEVAELGVAAVINIA	IATTEDEKNNTRINALP	ULTSCAOELGI	KMEDVVAAR-	DHKISDSPNAKCLMAC	IFK 68
	EoblPBP1 :	MARFTLSWRILALFAVFLAQIE	ERECSQEVMHKITKDFAY	VLEDCKKQENV	GDHIMQDIFN	IFWHEEYALVNPELGCVMLC	MAG 83
	EoblPBP2 :	MARFSFVAFLGVAAVINIA MARFILSWRILALFAVFLAQIE MTKLKELLLVLVISVIT MW-WKLVFVVVVGSAVV	RVQSSQDVMKSLTLNFGK	PMEVCKKELDL	PDAVTKEFLN	IFWREGYEVKNRLTGCAIIC	MSE 78
	EoblPBP3 : EoblPBP4 :	MW-WKLVFVVVVGSAVV MAKYHF-NKSLVLNVFLTVFLYFN	GTTEAADAMKLLASGFIS	VLEICQKELNI	EDGLISDLYH	IYWKLEFSMMQRDTGCALIC	MTK 77
	BODIEDEA .	Tantini Mobratti artisti		C	i boinbbi in	C C	101 00
		120 * 1	40 * 160	*	180 *	200 *	220
	EoblogP3 :	KIGFIDAKGLFDANVAL KVKAVDGYGFPTLEGI	EKNKKYFKAADDIA-KIEQMGKT VGLYSDGVNERGYFMAVLEASRE	LMRSHDHFSRTV	MD-NGRNODVSFDIFE	FLKEKANFT-PFESS	151
	EoblOBP5 :	KGGMMDDDGNLSVEGVF	KSAEAYLSDDPELLKKSELFTDA	KSVNDAPVSDGKR	(GCDRASLIFOC	SVEKAPSFOLF	146
	EoblOBP6 :	KRKIIDDEGKYAPEVAK	SEHAKYIHDAELVA-KLDEISDN	ASVNDOAVSDGAR	(GCERAKLLTAC	LTEHKDILTEIFKD	153
	EobloBP7 :	KSGFLIDKSTLNKETYK KAGIISPDGKLNVYK	AHLRQWAEKHKDWSAAVERAIED	VEKNLRQYLDIP-	TAYDVFT	TSIAMLKK PASSWK	145
	EoblOBP9 :	ESSLADEDGVVDYDMLL	SLIPEEYYDRTSKMILG	KHEDAPGKD	KOQSAFDVHK	SYQKDPDLYFLF	141
	EoblOBP10:	SKDCLNDDGSINHKA	VAEHLNNWASEEWKPAVEAAVAV	LGENEVPGPPHIC	CEANRLMFCIGGVIFSE	PTWQDNDDCKQLKEHINEC	KAAKFPPPN 188
	EobloBP11:	MTQIVKNNKLSYESALK YLVSLNETGSFPNETDKTPKCFLRCVLQ	QVDLMYPADMKA-SVKASIEN	KDVSKKYKD	PD-FKCHOFMAYNELK	LYDDNPKDFMFA	142
	EoblOBP13:	ITQVIKNNKLSYEAVIK	QVDMMFPPEMRT-AVKAAAEN	KEIAKKYKDD	ICEASYRTAKC	MYEYDAENFVFP	148
	EoblOBP14:	ITQVIKNNKLSYEAVIK DIGVMDGNGLIQKETAI	ELAKKVFEDAEELK-MIEDYLHS	AHVNTEPVSDGDF	(GCDRAIIAMKC	MIENASQFGFEL	151
	EoblOBP15:	MANAIKNGKLNYEAAIK KSELMTKDGKFKK	QADLLLPDEIKE-PAKESITV	RKVSDQYKD	ICEASFHTTK	IYNNNPAAFYFP	149
	FobloBP17.	KLGLMDDSGKISEKTAK	AATKKVFKEGDEMFTKVEELISP	THVNDAFTSDGDB	CGDRAKLAFE	FTEHAKELDLDVDL	156
	EoblOBP18:	GNGLIDDKGMFDPAKGT	SIAEKFIKSPDDIA-KVKKISDI	SSVNDEAVNDGDF	(GCDRAVLLLK	LMENKSLVV	145
	EoblOBP19: EoblOBP20:	GNGLIDDKGMFDPAKGT NLQVVDSRGIPREAEI EAKYMTADGKRINV	WNKIQSSVTSQQSRAALKDQTSA EKAASIFPNKVLMFKILSO	FQELQSEAEDN	GCSYSNKLER	LMLRFSDRKPSGTQTNNKQ FOENTPYRLSF	GTK 155
	EoblOBP21:	KADMMDDKGNYDLEKTN	IKWVETEFSDSATRLESARNLFNM	KKVNDEPVTDGEF	(GCERAYLLSKC	LVENSPKIGFATIE	147
	EoblOBP22:	KINVIDNDGLFVPAVAK	ANHQKYVHGADDLA-RLSASADT	TSVNDQAVTDGAN	IGCDRAKLLAR	FIDNHGVGPFSA	150
	EoblPBP1 :	KLDLMDGD-DMHHGNAH KLELLDEGLKLHHGNAK	DFAKKHGADDDLAK-QLVIMIHD	MESTPPNTD	PCMKTVDVAM	FKLKIHDLSWNPDPDLIIA	EVLAEA 163
	EoblPBP3 :	KLELLTDDGKFHHGVTK KLNLIDPEGKLHHGKAK	EFAMKNGADDNLAT-EMVSIIHS	ETKSEGLDD	ECLRALEVAK	FRVALHDLHWEPSPDVVIT	EVLGEM 162
	EoblPBP4 :	KLNLIDPEGKLHHGKAK	EFAMSHGADEGMAQ-QLIDIIHN	ENSTPQNED	GCLMVLAVAKC	FKVEIHKLNWTPSMDMVVG	EVLAES 170
				~	C		
		* 20	* 40 *	С	C C		
в		* 20 MKVLILLTLVVLASARP-D		60 NSRLLKAYAL	C C * 80 GRSK O TPEG-HDIKRWID	* 100 DEGTETR C AK C TPKQKVLVA	
В	EoblCSP2 :	MKVLILLTLVVLASARP-D MKSFYILFAFFAVCAAQTVAPATSTS	NHYDKKYDNFKVDELVS EAYYTSDENVDIEALVS	C 60 NSRLLKAYALCFLO NHDAMKQYVDCFTO	C C * 80 GRSK TPEG-HDIKRWII GKVE GPEA-GAAKGEFF	* 100 DEGTETRAK TPKQKVLVA PDALSDA AK TQVQKHTSK	VFF 95
В	EoblCSP2 : EoblCSP3 : EoblCSP4 :	MKVLILLTLVVLASARP-D MKSFYILFAFFAVCAAQTVAPATSTS MKGIVCVCLLAMAAAALARPDC MKLVYLITLAGVAAIAYGKPA-	NHYDKKYDNFKVDELVS EAYYTSDENVDIEALVS DHYTDKYDNINVDEILE QYTDKWDNINVDQILE	C 60 NSRLLKAYAL FLO NHDAMKQYVD FTO NSRLRDPYLS VLI SQRLLRGYVD	C C * 80 GRSKG TPEG-HDIKRWIE GRVEG GPEA-GAAKGEFF DVGKG APEA-KELKSHIK DRGRG TPDG-KALKETLF	* 100 DEGTETROAK TPKQKVLVA PDALSDACAK TQVQKHTSK CEALETHCLK TDIQKEQTK PDALEHECSK TAKQKSGSD	VFF 95 KVI 91 KVV 89
В	EoblCSP2 : EoblCSP3 : EoblCSP4 : EoblCSP5 :	MKVLILLTLVVLASARP-D MKSFYILFAFFAVCAAQTVAPATSTS MKGIVCVCLLAMAAAALARPDC MKLVYLITLAGVAAIAYGKPA- MLVLSLLSVFSVALTQCE-	NHYDKKYDNFKVDELVS 	C 60 NSRLLKAYALGFLC NHDAMKQYVDGFTC NSRLRDPYLSGVLI SQRLLRGYVDGLLI NERLLTGYVNGLLI	C C * 80 GRSKUTPEG-HDIKRWIL GRUEGPEA-GAAKGEFF DVGKCAPEA-KELKSHIK DRGRUTPDG-KALKETLE DLGPUTPDA-KELKKNLE	* 100 DEGTETROAK TPKOKVLVA POALSDAGAK TOVOKHTSK KEALETHILK TDIOKEOTK POALEHE SK TAKOKSGSD DOALEND EK TDRORDGAD	VFF 95 KVI 91 KVV 89 EVM 87
В	EoblCSP2 : EoblCSP3 : EoblCSP4 : EoblCSP5 : EoblCSP6 :	MKVLILLTLVVLASARP-D- MKSFILFAFFAVCAAQTVAPATSTS MKGIVCVCLAMAAAALARPDO -MKLVLITLAGVAAIAYGKPA 	NHYDKKYDNFKVDELVS EAYYTSDENVDIEALVS 	C 60 NSRLLKAYALCFLC NHDAMKQYVDCFTC NSRLRDPYLSOVLI SQRLLRGYVDCLLI NERLLTGYVNCLLI DLPTLQAFVGCFNI	C C 80 SRSKTPEG-HDIKRWII SKVELGPEA-GAAKGEFF DVGK APEA-KELKSHIF DRGR TPDG-KALKETLF DRGR TPDG-KALKKNLF DKVK DEKS-GDFKKDL7	* 100 DEGTETR AK TPKQKVLVA PDALSDALAK TQVQKHTSK (EALETH LK TDIQKEQTK PDALEHE SK TAKQKSGSD PDAIEND EK TDRQRDGAD LEALQQALAK TDAQKHIFK	VFF 95 KVI 91 KVV 89 EVM 87 VFL 86
В	EoblCSP2 : EoblCSP3 : EoblCSP4 : EoblCSP5 : EoblCSP6 : EoblCSP7 : EoblCSP8 :	MKVLILLTLVVLASARP-D MKSFYILFAFFAVCAAQTVAPATSTS MKGIVCVCLLAMAAAALARPDO MKLVLITLAGVAATAYGKPA WLVL5LLSVFSVALTQCE 	NHYDKYDNFKUDELVS; BAYTSDENVDIEALVS; DHYTDKYDNINVDEILE QYTDKWDNINVDQILE EIYTNKYDGVDLEEILA EIYTTENDDLNIEAVVA YDRRYDVFNIDYFVN YTSKYDGVDLDEILN;	C 60 NSRLLKAYAL FLO NHDAMKQYVD FTO NSRLRDPYLSVLI SQRLLRGYVD LLI NERLLTGYVN LLI DLPTLQAFVG FNI NPRLLKKYLN FLI NKRLLIPYVN CVL	C C * 80 SRSKTPEG-HDIRRWIT SKVE GPEA-GAAKGEFF VGK APEA-KELKSHIF ORGR TPDG-KELKSHIF DLGP TPDG-KELKKHIF DKVK DEKS-GPFKKDIF DCGP TPIG-RVFKQVLF DCGP TPIG-RVFKQVLF DCGR TADG-KELKSHIF	* 100 DEGTETR AK TPKQKVLVA PDALSDA AK TPKQKVLVA EGALETH LK TDIQKEQTK PDALEHE SK TAKQKSGSD PDAIEND EK TDRQRGAD BEAIQQA AK TDAQKHIFK EVIITA KK TPLQKRTY EGALETD AK TEAQQKGAE	VFF 95 KVI 91 KVV 89 EVM 87 VFL 86 KAF 88 KVF 91
В	EoblCSP2 : EoblCSP3 : EoblCSP4 : EoblCSP5 : EoblCSP7 : EoblCSP7 : EoblCSP8 : EoblCSP9 :	MKVLILLTLVVLASARP-D- MKSFYILFAFFAVCAAQTVAPATSTS MKGIVCVCLLMAAAALARPDO MKLVYLITLAGVAAIAYGKPA 	NHYDKKYDNFKVDELVS EAYYTSDENVDLEALVS QYTDKYDNINVDEILE QYTDKWDNINVDQILE EIYTNKUDGVDLEELLA YDRRYDYFNIDYFVN 	C 60 NSRLLKAYAL FLK NHDAMKQYVD FTC NSRLRGYVD LLI NERLLTGYVN LLI NERLLTGYVN FLI NERLLKYLN FLI NKRLLIPYVN VLI NKRLLAYYN VLI NKRLLAYYN VLI	C C * 80 GRSKUTPEG-HDIKRWII SKVEI GERA-GAAKGEFI VOGK APERA-KELKSHIFI DUGP TPDA-KELKSHIFI DOGF TPIG-RVFKQVLF EGGR TADG-KELKSHIFI VOGKUTPEG-KELKSHIFI VOGKUTPEG-KELKSHIFI	* 100 DEGTETR AK TPKQKVLVA PDALSDA AK TQVQKHTSK EALETH LK TDIQKEQTK PDALEHE SK TAKQKSGSD PDALEND EK TDAQKHIFK PEVITTA KK TPLQKRFTY EALEDT AK TPLQKRFTY EALETD AK TSPQRAGIR	VFF 95 KVI 91 KVV 89 EVM 87 VFL 86 KAF 88 KVF 91 KVI 88
В	EoblCSP2 : EoblCSP3 : EoblCSP4 : EoblCSP5 : EoblCSP6 : EoblCSP7 : EoblCSP9 : EoblCSP9 :	MKVLILLTLVVLASARP-D- MKS-VLFAFFAVCAAQTVAPATSTS MKGIVCVCLLAMAAAALARPDO MKLVILITLAGVAATAYGKPA WLVL5LLSVFSVALTQCE MKFFVVFALSLVVACSG MSVISVVSFVLIMTVAQEKY 	NHYDKKYDNFKVDELVS; BAYTSDENVDIEALVS; DHYTDKYDNINVDEILE 	C 60 NSRLLKAYAL FLG NBARKQYDC FTC NSRLRDPYLS VLI SQRLLRGYVN LLI NERLLKGYVN LLI NKRLLIPYVN VLI NKRLLIPYVN VLI NKRLLKAYIN FLI NDRLLNAYIN FLI	C C * 80 BRSK TPEG-HDIKKWII SKVE GPEA-GAAKGEFI VORK APEA-KELKSHI DRGR TPDG-KALKSTI DLGP TPDA-KELKSHI SKVK DEKS-GDFKKDIA DQGP TPIG-RVFKQVI EGGR TADG-KELKSHI VORK TPEG-ADFKKAL	* 100 DEGTETR AK TPKCKVLVA DALEDA AK TQVQKHTSK EALETH LK TDIQKEQTK DALEHE SK TAKQKSGSD DALEND EK TDRQRDGAD LEAIQQA AK TDAQKHIFK EVITTA KK TDAQKHFK DALEDT AK TEAQQKGAE DALQSG VK TGPQRAGTR DALGSG VK TGPQRAGTR	VFF 95 KVI 91 KVV 89 EVM 87 VFL 86 KAF 88 KVF 91 KVI 88 RVI 86
В	EoblCSP2 : EoblCSP3 : EoblCSP4 : EoblCSP6 : EoblCSP6 : EoblCSP7 : EoblCSP8 : EoblCSP10 : EoblCSP10 : EoblCSP11 :	MKVLILLTLVVLASARP-D- MKGIVCVCLLAMAAAALARPDC MKGIVCVCLLAMAAAALARPDC MKLVILITLAGVAATAYGKPA WLVISLSVFSVALTQCE MKFFVYFALSLVVACSG- MKVAVVLCILSFAAVAFSAPSC MKLLLIAVLITLASLHTLAQT MKVWLVFALMAVLASVTA		C 60 NSRLLKAYAL FLG NKRLRDYLS VLI SQRLLRGYVN LLI SQRLLRGYVN LLI DLPTLQAFVG FNI NRRLLKYLN FLI NKRLLKYLN VLI NKRLLKAYIN VLI NDRLLNAYVN FLI DEKWFQGFHD LMW ERILLAYVK VMI	C C * 80 SRSK TPEG-HDIKKWII SRVE GPEA-GAAKGEFF VIGK APEA-KELKSHF VIGK DEKS-GDFKKDIA JOGF TPIG-NELKSHIF VIGK TPEG-RVFKQVLE EEGR TADG-KELKSHIF VIGK TPEG-ADFKKALI TGPP PEEI-KKYTNII TGPP PEEI-KKYTNII	* 100 DEGTETR AK TPKOKVLVA DALSDA AK TQVQKHTSK (BALETH LK TDIQKEOTK 'DALENE SK TAKQKSGSD DAIEND EK TDRQRAGA EKAIQOA AK TDAQKHIFK VEVITTA KK TPLQKRFTY (EALETT GK SDKQKGNIR 'IYLKTA GP SPKEKEIFQ ETLTTA AK SPKOKOVVR	VFF 95 KVI 91 KVV 89 EVM 87 VFL 86 KAF 88 KVF 91 KVI 88 RVI 86 KHK 83 KML 86
В	EoblCSP2 : EoblCSP3 : EoblCSP4 : EoblCSP5 : EoblCSP6 : EoblCSP7 : EoblCSP9 : EoblCSP10: EoblCSP10: EoblCSP12: EoblCSP13:	MKVLILLTLVVLASARP-D MKSFYILFAFFAVCAAQTVAPATSTS MKGIVCVCLLAMAAAALARPDO MKLVILTLAGVAATAYGKPA WLVLSLLSVFSVALTQCE 	NHYDKYDNFKVDELVS; BAYTSDENVDIEALVS; DHYTDKYDNINVDEILE 	C 60 NSRLLKAYAL FLC NRDAWKQYDD FTC NSRLDRYLD VLI SQRLLRGYVD LLI NERLLRGYNW LLI NPRLLKYLM FLI NKRLLFYVN VLI NKRLLFYVN VLI NKRLLFYVN VLI DKRLLGYYK VM NERI LLAYYK VM NERI LLAYYK VM	C C * 80 GRSKE TPEG-HDIKRWII SKVEI GFEA-GAAKGEFF VOKG APEA-KELKSHIF DRGR TPDG-KALKETLF LOGF TPDA-KELKSHIF SKVK DEKS-GDFKKDIA DGP TFIG-RVFRQVLE SEGRT TADG-KELKSHIF VGG TPEG-ADFKKALF TGGP PEEI-KKYTNII SGGP TKDG-KNFKRVLE -GDS SOA-AALKTDFI -GD SSOA-AALKTDFI	* 100 DEGTETR AK TPKQKVLVA DALSDA AK TPKQKVLVA EGALETH LK TDIQKEQTK DALEND EK TDRQKEQTK DALEND EK TDRQKGRGAD AK TDAQKHIFK EVITTA KK TPLQKRTY EALETD AK TEAQQKGAE DALQQG VK TGPQRAGTR SDALQQG VK TGPQRAGTR SDALQQG VK TGPQRAGTR PALITT GK SDKQKGNIR 'IYLKTA GP SPKEKEIFQ DETLTTA AR SPKQKQVVR DETLTTA AR SPKQKQVVR	VFF 95 KVI 91 KVV 89 EVM 87 VFL 86 KAF 88 KVF 91 KVI 88 KVI 86 KHK 83 KML 86 LFI 83
В	EoblCSP2 : EoblCSP3 : EoblCSP5 : EoblCSP5 : EoblCSP6 : EoblCSP7 : EoblCSP8 : EoblCSP10: EoblCSP10: EoblCSP12: EoblCSP12: EoblCSP14:	MKVLILLTIVVLASARP-D- MKSFYILFAFFAVCAAQTVAPATSTS MKGIVCVCLLAMAAAALARPDO MKLVXLITLAGVAATAYGKPA- 	NHYDKYDNFKVDELVS; EAYTSDENVDIEALVS; 	C 60 NSRLLKAYAL FLC NNBARKQYVD FTC NSRLDPYLS VLI SQRLLRGYVD LLI SQRLLRGYVD LLI NERLLGYVN VLI NKRLLKAYIN VLI NKRLLKAYIN VLI NDRLLNAYVN FLM DEKWFQCHK DKU DKEAVQCVLK FMC DKFFIQRQLKALC	C C * 80 GRSKUTPEG-HDIKRWIL SKVEIGEA-GAAKGEFF DVGK APEA-KELKSHI DVGK TPDG-KALKETLF DLGP TFDA-KELKKNLF SKVK DEES-GOFKKDLA SUGK TPEG-KELKSHIF VGK TPEG-KELKSHIF DVGK TPEG-KELKSHIF DVGK TPEG-KELKKNLF SKGP TFDG-KELKSHI DVGK TPEG-KELKKLTLT SKGP TKDG-KNFKRVLF S-GD SSQA-AALTTDFF SAPD PDIG-KNLTTDFF	* 100 DEGTETR AK TPKQKVLVA PDALSDA AK TPKQKVLVA EXALETH LK TDIQKEOTK PDALEHE SK TAKQKSGSD PDAIEND EK TDAQKHIFK PEVITA KK TPLQKRTY EXALETD AK TEAQOKGAE PEALETT GK SDKQKGNIR PEALETT GK SPKCKUVVR PEILTTA AR SPKQKQVVR PLALGSA AE TPCQKHITK IVLRGA PO SPCETKCIQ	VFF 95 KVI 91 KVV 89 EVM 87 VFL 86 KVF 91 KVI 88 RVI 86 KMK 83 KML 86 LFI 83
В	EoblCSP2 : EoblCSP3 : EoblCSP4 : EoblCSP5 : EoblCSP6 : EoblCSP6 : EoblCSP8 : EoblCSP10 : EoblCSP11: EoblCSP12: EoblCSP13 : EoblCSP13 : EoblCSP15 :	MKVLILLTLVVLASARP-D- MKSIVCVCLLAMAAAALARPDO MKLVYLITLAGVAATAYGKPA WKLVYLITLAGVAATAYGKPA WKVWIVLSLSVFSVALTQCE MKVAVVLCILSFAAVAFSAPSO MKVAVVLCILSFAAVAFSAPSO MKVIULAVLITLASLHTLAQT MRSSMLLCAVALWAFVSGD MKSFCIVALGLVAACN MKSFCIVALGLVAACN MKSFCIVALGLVAACN MKALLSALIAVCIAQSQ MKA		C 60 NSRLKAYAL FLK NNDAMKQYVD FT NSRLRDPYLS VLI SQRLLRGYVD LLI NERLLTGYVN LLI DLPTLQAFVG FNI NRRLLFYVN VLI NKRLLFYVN VLI NKRLLFYVN VLI DEKWFQGFHD LMI DEKWFQGFHD LMI DEKYQKVLK FNO DKRFIQRQLK AL	C C * 80 BRSK TPEG-HDIKKWII SKK TPEG-HDIKKWII SKK TPEG-KALKSHI SKK TPEG-KALKSHI DGG TPDG-KALKSHI SKK TPEG-KELKSHI SKG TPEG-ADFKAL SKG TPEG-ADFKAL SKG TPEG-ADFKAL SGG TVERG-ADFKAL SGG TVERGADFKTLAF	* 100 DEGTETR AK TPKCKVLVA DALSDA AK TQVCKHTSK EALETH LK TDICKEOTK DALEHE SK TAKCKSGSD DAIEND EX TDRQRAGAD LEAIQOA AK TDACHIFK UEALETD AK TEAQCKGAE DDALSG VK TGPQRAGIR DALSG VK TGPQRAGIR PEAIETT GK SDKQKGVIR PIYLKTA GP SPKEKEIFQ PETLTTA AR SPKCKVVR LALGSA AE TPEQKHNTK LVLRGA PQ SPCKKUVQ EVIATS AK SPICKKUV	VFF 95 KVI 91 KVV 89 EVM 87 VFL 86 KVF 91 KVF 91 KVI 88 RVI 86 KHK 83 KHL 86 LFI 83 KTL 87 KTV 87
В	EoblCSP2 : EoblCSP3 : EoblCSP4 : EoblCSP5 : EoblCSP6 : EoblCSP6 : EoblCSP9 : EoblCSP10: EoblCSP10: EoblCSP12: EoblCSP13: EoblCSP14: EoblCSP15: EoblCSP17:	MKVLILLTIVVLASARP-D- MKGIVCVCLLAMAAAALARPDC MKGIVCVCLLAMAAAALARPDC MKLVILITLAGVAATAYGKPA WLVISLSVFSVALTQCE MKFFYVFALSLVVACSG- MKVAVVLCILSFAAVAFSAPSC MKLLILAVLITLASLHTLAQT MKSSMLLCAVLWAFVSGD MKSFCIVALGUVASCA MKSFCIVALGUVASCA MQIKLLLAIVCIAQSQ MKTLLLAIVTLAVELAQSQ MKTLLLAIVTLAVELAVC		C 60 NSRLLKAYAL FL NHDAMKQYDD FT NSRLRDPYLS VLI SQRLLRGYVD LLI DLPTLQAFVG FNI NRRLLKAYIN VLI NKRLLAYIN VLI NKRLLKAYIN VLI DEKWFQGFHD LM DERILLAYXK VM DKEFIQRQLK ALC DALFTKYID MLI NFRLLKAYAV FI NFRLLKAYAVD VM	C C * 80 SRSK TPEG-HDIKKWII SRVE GPEA-GAAKGEFF VIGK APEA-KELKSHIF VIGK DEKS-GDFKKDIA OGGT PTIG-RVENQUL EEGR TADG-KELKSHIF VIGK TPEG-RVENQUL EEGR TADG-KELKSHIF VIGK TPEG-ADFKKALI SGGT PTEG-ADFKKALI SGGP CONSCONTROL SGA PTEG-KELKSHIF VIGK TEGG-RVENQUL SGGP TEGG-RVENQUL SGGP TEGG-TDFKKVII SGGR SEGG-TEGG-TDFKKVII SGGR SEGG-TEGG-TDFKKVII SGGR SEGG-TEGK	* 100 DEGTETR AK TPKOKVLVA DALSDA AK TQVQKHTSK (EALETH LK TDIQKEOTK DALENE SK TAKQKSGD DAIEND EK TDRQRAGAD DAIEND EK TDRQRHIFK (EALETT AK TELQQKGAI PUITTA KK TPLQKRFTY (EALETT GK SDRQKGNI PIYLKTA GP SPKEKEIFQ PUITASA AK SPKQKVVR PLALGSA AE TPEQKHTK PULAGA AE TPEQKHTK PULAGA AK SPIQKKHVR PDAVKTK EK SPIQKKHVR DAVKTK EK SPIQKKHVR	VFF 95 KVI 91 KVV 89 EVM 87 VFL 86 KAF 88 KVF 91 KVI 86 KHK 83 KKV 86 LFI 83 KTL 87 KTV 87 IVV 88
В	EoblCSP2 : EoblCSP3 : EoblCSP4 : EoblCSP5 : EoblCSP6 : EoblCSP7 : EoblCSP10: EoblCSP10: EoblCSP12: EoblCSP12: EoblCSP15: EoblCSP15: EoblCSP16: EoblCSP16: EoblCSP18:	MKVLILLITJVVLASARP-D MKSFYILFAFFAVCAAQTVAPATSTS MKGIVCVCLLAMAAAALARPDO MKLVYLITLAGVAATAYGKPA WLVISLSVFSVALTQCE MKVAVVLCILSFAVAFSAPSG MKVLILAVLITLASLHTLAQT 	NHYDKYDNFKVDELVS; 	C 60 NSRLLKAYAL FLC NHDAHKQYDD FTC NSRLDPYLS VLI SQRLLRGYDD LLI NERLLTGYVD LLI NERLLTGYVD VLI NRRLLFYVN VLI NKRLLFYVN VLI DEKWFQGFHD LMI NERILLAYYN VM NERILLAYYN VM DREAVQKVLK FNC DREAVQKVLK FNC DRAFTGRQLK ALC NRRLLRAYDD VLI NRRLLKAYDD VLI NRRLLKAYDD VLI NRRLLKAYDD VLI NRRLLKAYDD VLI NRRLLKAYDD VLI NRRLLKAYDD VLI NRRLKAYDT LSS	C C * 80 BRSK TPEG-HDIKRWIL SKVEJ GEPA-GARKGEFF VOGK APEA-KELKSHIF DRGR TPDG-KALKETIF DRGR TPDG-KALKETIF DKVK DEKS-GDFKKDIA DGP TPIG-RVFKQVLF EGGR TADG-KELKSHIF VGK TPEG-KELKSHIF KTGP PEEI-KKYTNIIF KGP PEEI-KKYTNIIF EAPP DPIG-KRLKTLAF SGGP TSEG-TDFKKVIF SGGF TSEG-TDFKKVIF SGGF TSEG-TDFKKVIF SGGF VERADFKLEHLC SGGF TEGG-KELKEHLC SGGF VERADFKLEHLC SGGF VERADFKLEHLC	* 100 DEGTETR AK TPVGKVLVA DALSDA AK TQVGKHTSK GEALETH LK TDIQKEQTK DALEHE SK TAKQKSGSD DALEND EK TDRQRDGAD LEAIQQA AK TDAQKHIFK CEALETD AK TEAQQKGAE PIYLKTA GP SPKEKEIFQ DEALETT GK SDKQKGVIR PIYLKTA AR SPKCKVVR LALGSA AE TPEQKHNTK PUVLRG PQ SPKEKEIFQ DALANG AK TENNHMTAIR	VFF 95 KVV 89 EVM 87 VFL 86 KAF 88 KVF 91 KVI 88 RVI 86 KHK 83 KML 86 LFI 83 KTL 87 KTV 87 IVV 88 TVL 88 KVI 89
В	EoblCSP2 : EoblCSP3 : EoblCSP4 : EoblCSP5 : EoblCSP6 : EoblCSP6 : EoblCSP12: EoblCSP12: EoblCSP14: EoblCSP14: EoblCSP14: EoblCSP16: EoblCSP15: EoblCSP16: EoblCSP19: EoblCSP19:	MKVLILLTLVVLASARP-D- MKGIVCVCLLAMAAAALARPDC MKLVLUTLAGVAALAYGRPA MKLSLSVFSVALTQCE MKFFYVFALSLVVACSG- MKVAVVLCILSFAAVAFSAPSC MKVIVVFALJASLHTLAQT MKVIVVFACLVASVVTA MKVIVVFACLVASVVTA MKSSKLLGAVALWAFVSGD MKSTVLAVCALTVVVSSQQ MKSTVLAVCALTVVVSQQ MKSTVLAVCALTVVVSQQ MKS		C 60 NSRLLKAYAL FL NHDAMKQYVD FT NSRLRDYVLS VI SQRLLRGYVD LI DLPTLQAFVG FNI NRRLLKAYIN FL NKRLLKAYIN VI NKRLLKAYIN VI DKENGCFHD LM DREILLAYVK VM DREILLAYVK VM DREILLAYKY VM NRRLKAYVD VM TPRLLKNYAK FI NRRLKAYVD VM TRRLFHGYYLD LS	C C * 80 SRSK TPEG-HDIKKWII SRVE GPEA-GAAKGEFF VIGK APEA-KELKSHI DVGK DEKS-GDFKKDIA DGGT TDIG-NELKSHI DVGK TPEG-ADFKKALIA DVGK TPEG-ADFKKALIA SKGGT TKOG-KNFRKVLI SGGP TGG-KNFRKVLI SGGP TGG-RKHKTLAF SGGP TSGG-TDFKKVIF SGGF TSGG-TDFKKVIF SGGF TSGG-TDFKKVIF SGGF TSGG-TDFKKVIF SGGF TFGG-KLEKENMF SSVR TPEG-KALKENMF SVR TPGG-KALKRLENMF	* 100 DEGTETR AK TPKCKVLVA DALSDA AK TQVCKHTSK KEALETH LK TDIQKECTK DALENE SK TAKCKSGD DAIEND EK TDRQRADA VENTTA KK TPLQKRFTY KEALETT GK SDKQKGNIR VILALGS AK TGPQRAGIR VENTTA KK SDKQKGVR VILALGS A B TPECKHTK VLALGS A B TPECKHTK VLALGS A B TPECKHTK VLALGS A STCKVCV DALATS KK SDKQKUR PDAVKTK EK SDKQKELIR PDAVKTK EK SDKQKELIR VDALENG EK TEKQEKGSY DALANG AK TERQKRTAV	VFF 95 KVI 91 KVV 89 EVM 87 VFL 86 KAF 88 KVF 91 KVI 86 KVI 86 KML 86 LFI 83 KML 83 KTL 87 KTV 87 IVV 88 TVL 88 KVI 89 KVI 89
В	EoblCSP2 : EoblCSP3 : EoblCSP4 : EoblCSP5 : EoblCSP6 : EoblCSP6 : EoblCSP10: EoblCSP10: EoblCSP13: EoblCSP14: EoblCSP14: EoblCSP15: EoblCSP17: EoblCSP17: EoblCSP18: EoblCSP18: EoblCSP19:	MKVLILLITJVVLASARP-D MKSFYILFAFFAVCAAQTVAPATSTS MKGIVCVCLLAMAAAALARPDO MKLVYLITLAGVAATAYGKPA WLVISLSVFSVALTQCE MKVAVVLCILSFAVAFSAPSG MKVLILAVLITLASLHTLAQT 		C 60 NSRLIKAYAL FL NHDAMKQYVD FT NSRLRDPYLS VI SQRLIRGYVD LI NERLITGYVN LI DLPTLQAFVG FN NRRLIFYVN VL NKRLIFYVN VL NKRLIFYVN VL DKAFGGFHD LM DERILLAYYK VM DKEAVQKVLK FN DKAFIQRQLK AL NERILLAYYK VM NFRLIKNYAK FL NKRLIKAYVD VM TPRLFHGYYL LS NKRLVKIYVD T	C C * 80 SRSK TPEG-HDIKKWII SRVE GPEA-GAAKGEFF VIGK APEA-KELKSHI DVGK DEKS-GDFKKDIA DGGT TDIG-NELKSHI DVGK TPEG-ADFKKALIA DVGK TPEG-ADFKKALIA SKGGT TKOG-KNFRKVLI SGGP TGG-KNFRKVLI SGGP TGG-RKHKTLAF SGGP TSGG-TDFKKVIF SGGF TSGG-TDFKKVIF SGGF TSGG-TDFKKVIF SGGF TSGG-TDFKKVIF SGGF TFGG-KLEKENMF SSVR TPEG-KALKENMF SVR TPGG-KALKRLENMF	* 100 DEGTETR AK TPKCKVLVA DEDALSDA AK TQVCKHTSK (EALETH LK TDICKEQTK DALEHE SK TAKCKSGSD DAIEND EX TDRCRGAGA DAIEND EX TDRCRGAGA PEVITTA KK TPLOKRFTY (EALETD AK TEAQCKGAE DALOSG VK TPLOKRFTY DAUGAG VK SDKCKGNIR PILLAGA AC SPECKGIQ DETLTTA AR SPKCKUVW PLALAGA PC SPCKVUV PLALAGA AC SPOETKCIQ DAIETG EK TEKCKUVK DAVKK EK SDKCKUVR DAVKK EK SPCKKUV DALANG AK TENHMTAIR PEALETR AK TEKCKMAK	VFF 95 KVV 91 KVV 89 EVM 87 VFL 86 KAF 88 KVF 91 KVI 86 KHK 83 KML 86 KHK 83 KML 86 KHK 87 KVI 87 IVV 88 KVI 87 IVV 88 KVI 89 KVI 89 KVI 96 QLA 86
В	EoblCSP2 : EoblCSP3 : EoblCSP4 : EoblCSP5 : EoblCSP6 : EoblCSP6 : EoblCSP10: EoblCSP10: EoblCSP13: EoblCSP14: EoblCSP14: EoblCSP15: EoblCSP17: EoblCSP17: EoblCSP18: EoblCSP18: EoblCSP19:	MKVLILLTIVVLASARP-D- MKSIVCVCLLAMAAAALARPO MKGIVCVCLLAMAAAALARPO MKLVLITLAGVAATAYGKPA WLVISLSVFSVALTQCE MKYVVSFULMTVVAQEKY 	NHYDKKYDNFKVDELVS; 	C 60 NSRLLKAYAL FL NHDAMKQYDD FT NSRLRDPYLS VLI SQRLLRGYVDD LLI DLPTLQAFVG FNI NERLLTGYVN LLI NKRLLIPYVN VLI NKRLLKAYIN VLI NKRLLKAYNN FLI DEKWFQGFHD LMW NERLLKAYKV VM DKEFIQRQLK ALC DDALFFKYID MLI NFRLLKNYAK FLI NFRLLKNYAK FLI NKRLVKIVD LYT NFRLLNYAK LLI NKRLVKIVD LYT NFRLLNYAK LLI NKRLVKIVD LYT NFRLLNYAK LLI NKRLVKIVD LYT NKRLVKIVD LYT NFRLLNYAK LLI	C C * 80 BRSK TPEG-HDIKKWII SRVE GPEA-GAAKGEFF VVGK APEA-KELKSHIF VGK DEKS-GDFKKDIA OGGT PTIG-NENKUNI SKORT TPEG-RVFKQVLE EEGR TADG-KELKSHIF VVGK TPEG-ADFKKALI SKORT TPEG-ADFKKALI SKORT PEG-SADFKKALI SKORT SEG-SADFKKALI SKORT SEG-STADFKKVII SKORT SEG-STADFKKVII SKORT SEG-SKELKENME SVR TPEG-KELKENME SVR TPEG-KALKRLI SKORT PYG-KELKENME SVR TPEG-KALKRLI SKORT SPGG-KELKSHIF C	* 100 DEGTETR AK TPKOKVLVAK REALETH LK TDUQKHTSK REALETH LK TDUQKEDTK PDALEND SK TAKQKSGSD DDAIEND EK TDRQRHIFK VENUTTA KK TPLOKRFTY REALETT GK SDRQKGNIR PIYLKTA GP SPKEKEIFQ PIYLTA K SDRQKGNUR PLALGSA AE TPEQKHTK PLVLRGA PQ SPQETKQIQ PUXLRSA AK SPIQKKHVR PDAVKK EK SPIQKKHVR PDAVKK EK SPIQKKHVR PDAVKK EK SPIQKKHVR PDAVKK EK SPIQKKHVR PDALETG EK TEKQEKGSY PDALANG AK TENIMTAIR PEALETK YAK TEKQKQMGK REALETY AK TEKQKQMGK REALETY AK TEKQKQMGK	VFF 95 KVV 91 KVV 89 EVM 87 VFL 86 KAF 88 KVF 91 KVI 86 KHK 83 KML 86 KHK 83 KML 86 KHK 87 KVI 87 IVV 88 KVI 87 IVV 88 KVI 89 KVI 89 KVI 96 QLA 86
В	EoblCSP2 : EoblCSP3 : EoblCSP4 : EoblCSP5 : EoblCSP6 : EoblCSP6 : EoblCSP12: EoblCSP12: EoblCSP12: EoblCSP14: EoblCSP15: EoblCSP15: EoblCSP15: EoblCSP15: EoblCSP15: EoblCSP19: EoblCSP19: EoblCSP19: EoblCSP11 : EoblCSP11 :	MKVLILLTIVVLASARP-D MKSFYILFAFFAVCAAQTVAPATSTS MKGIVCVCLLAMAAAALARPDO MKLVLITLAGVAATAYGKPA- MLVLSLSVFSVALTQCE MKVVVSFVLMTVVAQEKY 		C 60 NSRLLKAYAL FL NHDAMKQYVD FT NSRLRDPYLS VLI SQRLLRGYVD LLI DLPTLQAFVG FNI NERLLTGYVN LLI NKRLLIPYVN VLI NKRLLIPYVN VLI NKRLLKAYIN VLI NKRLLKAYVN FI DEKWFQGFHD LMW DKEFIQRQLKALC DALFFKYID MLI NFRLLKNYAK FI NFRLLKNYAK FI NRRLLKAYVD VM NKRLVRIVVD LT NRRLLKYVD LT NRRLLKYVD LT NRRLLNYAK LLI NKRLVRIVUD LT NRRLLNYAK LLI NKRLVRIVUD LT NRRLLNYAK LLI NKRLVRIVUD LT NRRLLNYAK LLI NKRLVRIVN L	C C * 80 BRSK TPEG-HDIKKWII SRVE GPEA-GAAKGEFF VGK APEA-KELKSHF VGK APEA-KELKSHF DKVK DEKS-GDFKKDIA DGGT PIG-RVFKQVLE EEGR TADG-KELKSHIF VGK TPEG-RVFKQVLE EEGR TADG-KELKSHIF VGK TPEG-ADFKKALF SGGT PIG-KNFKRVLF SGGP KOE-KNFKRVLF SGGP KELKENMF SVGF TPGG-KELKENMF SVR TPEG-KALKRLF NGKS SPEG-KALKSLIF NGK SPEG-KALKSLIF NGK SPEG-KALKSLIF NGK SPEG-KALKSLIF NGK SPEG-KALKSLIF C * 180	* 100 DEGTETR AK TPKOKVLVAK TOPALSDA AK TQVQKHTSK TEALETH LK TDIQKEOTK TDALEND SK TAKQKSGD DAIEND EK TDRQKDGAD TDAQKHIFK TDRQKDGA AK TDAQKHIFK TEALETT GK SDRQKGNTR TYLKTA GP SPKEKEIFQ TYLKTA SFKEKEIFQ TYLKTA SFKEKEIFG	VFF 95 KVI 91 KVV 89 EVM 87 VFL 86 KAF 88 KVF 91 KVI 86 KVI 88 KML 86 LFI 83 KML 83 KML 87 KTV 87 IVV 88 TVL 88 KVI 89 KVI 91
В	EoblCSP2 : EoblCSP3 : EoblCSP4 : EoblCSP4 : EoblCSP6 : EoblCSP6 : EoblCSP8 : EoblCSP1 : EoblCSP13: EoblCSP14: EoblCSP14: EoblCSP14: EoblCSP15: EoblCSP15: EoblCSP16: EoblCSP17: EoblCSP20: EoblCSP21: EoblCSP21:	MKVLILLITJVVLASARP-D 		C 60 NSRLLKAYAL FLK NHDAMKQYDD FTY SQRLLRGYUD LLI SQRLLRGYUD LLI NERLLTGYUN LLI NERLLTGYUN LLI NERLLKYLN FLI NKRLLFYVN VLI NKRLLLAYYN VI DEKAVQKYLK FM NERILLAYYN VI NERLILAYYN VI NRELLAYYN VI NRRLLRAYDD VI NRRLLKAYDD VI NRRLLKAYDD VI NKRLLVRIYD LIS NKRLVRIYD VI C 160	C C * 80 BRSK TPEG-HDIKRWII SKVEJ GFEA-GARGEFF VORG APEA-KELKSHIF DRGR TPDG-KALKETIF DRGR TPDG-KALKETIF DRGR TPEG-KELKSHIF SKGF TPEG-KELKSHIF SKGF TPEG-ADFKKALF SKGF TPEG-ADFKKALF SKGF TPEG-ADFKKALF SKGF CFEA-ADFKKALF SKGF TFEG-ADFKKALF SKGF TFEG-KALKTLAF SKGF TFEG-KELKSHIF C C 180 * 180	* 100 DEGTETR AK TPKOKVLVA DALSDA AK TPKOKVLVA EGALETH LK TDIQKEOTK DALEHE SK TAKOKOSGD DAIEND EK TDRORDGAD LEAIQA AK TDAOKHIFK EDVITTA KK TEAQOKGAE DDALSOG VK TGPQRAGTR DEALETT GK SDKOKGNIR PIYLKTA GP SPKEKEIFQ DEALETT GK SDKOKGNIR PIYLKTA AR SPIQKHVTK LALGSA AE TPEQKHNTK LALGSA AE TPEQKHNTK UVLRGA PO SPOETKOIQ DUNKTK EK SDKOKELIR DDALETG EK TENGKKIVA PDAVKTK EK SDKOKELIR DDALSTG EK TENGKKIVA PDALKT AK TENGKOMGK (EALETY AK TENGKOKG	VFF 95 KVV 91 KVV 89 EVM 87 VFL 86 KAF 88 KVF 91 KVI 86 KVI 88 KHK 83 KHL 87 KTV 87 IVV 88 TVL 88 KTV 87 IVV 88 KVI 89 KVI 91 124 130
В	EoblCSP2 : EoblCSP3 : EoblCSP4 : EoblCSP5 : EoblCSP6 : EoblCSP6 : EoblCSP12: EoblCSP12: EoblCSP12: EoblCSP14: EoblCSP15: EoblCSP15: EoblCSP15: EoblCSP15: EoblCSP22: EoblCSP21: EoblCSP21:	MKVLILLTIVVLASARP-D MKSFYILFAFFAVCAAQTVAPATSTS MKGIVCVCLLAMAAAALARPDO MKLVLITLAGVAATAYGKPA WLVISLSVFSVALTQCE MKVVVSVLIMTVVAQEKY 		C 60 NSRLLKAYAL FL NHDAMKQYVD FT NSRLRDPYLS VLI SQRLLRGYVDP LLI DLPTLQAFVG FNI NERLLTGYVN LLI NKRLLIPYVN VLI NKRLLKAYIN VLI NKRLLKAYIN VLI DEKWFQGFHD LMN NERILLAYYK VMI DKEAVQKVLK FNU DKEFIQRQLKA DALFTKYID MLI NFRLLKNYAK FLS NKRLKAYVD VMI TPRLFHGYYL LSS NKRLKAYVD VMI NRRLLKYVD LTT NFRLLNYAK FLS NKRLKAYVD VMI C 160	C C * 80 BRSK TPEG-HDIKKWII SVUE GPEA-GAAKGEFF VOGK APEA-KELKSHIF DRGRTPDG-KALKETIF LIGFTPDA-KELKSHIF VOKK DEKS-GDFKKDIF DOGT PIG-RVFKQVIE EGGR TADG-KELKSHIF VOGK TFEG-RVFKQVIE SUGRTFEG-ADFKKALF SGRTPKG-ADFKKALF SGRTPKG-KNFKRVIF SGRTPKG-KNFKRVIF SGRT SEG-TDFKKVIF SGRK SPEG-KELKEHNF SVR TPEG-KALKKLLF RGGP TPEV-KQVKEKLF RGG TPEV-KQVKEKLF RGG SPDG-KELKSHIF C * 180	* 100 DEGTETR AK TPKOKVLVAK DEDALSDA AK TQVQKHTSK REALETH LK TDIQKEQTK PDALENE SK TAKQKSGSD DDAIEND EX TDRQKAGAR DEALETN AK TDAQKHIFK VEALETT AK TEAQOKGAE DDALOGS VK TEAQOKGAE DDALOGS VK TEAQOKGAE PLALGSA AK TEAQOKGAE PLALGSA AK TEAQOKOVW PLALGSA AK TEAQOKOVW PLALGSA AK SPIQKKHVW PLALGSA AK SPIQKKHVW PLALGSA AK SPIQKKHVW PDAVKK EK SPIQKKHVW PDALGSA AK TEKQKRAV PDALANG AK TEKQKRAV PEALETR AK TEKQKMAK C C * 200	VFF 95 KVI 91 KVV 89 EVM 87 VFL 86 KAF 88 KVF 88 RVI 86 KVI 88 KML 86 LFI 83 KML 86 LFI 83 KML 87 KTV 87 IVV 88 TVL 88 KVI 89 CQLA 86 RVI 91 124 124
В	EoblCSP2 : EoblCSP3 : EoblCSP5 : EoblCSP5 : EoblCSP6 : EoblCSP6 : EoblCSP6 : EoblCSP10 : EoblCSP12 : EoblCSP13 : EoblCSP14 : EoblCSP15 : EoblCSP15 : EoblCSP15 : EoblCSP1 : EoblCSP2 : EoblCSP2 : EoblCSP2 : EoblCSP3 : EoblCSP3 : EoblCSP3 : EoblCSP3 : EoblCSP3 : EoblCSP3 : EoblCSP3 :	MKVLILLTIVVLASARP-D MKGIVCVCLLAMAAAALARPDO MKLVLLITLAGVAATAYGKPA- WLVLSLLSVFSVALTQCE 		C 60 NSRLLKAYAL FLC NHDAMKQYUD FTC NSRLDPYLS VLI SQRLLRGYUD LLI NERLLTGYVN VLI DLPTLQAFVG FNI NERLLKYYN VLI NERLLKAYNN VLI DEKWFQGFHD LMI NERLLAYYK VMI DEKAVQKVKK FNC DKRAYQKVKK FNC DKRAYQKVKK FNC DKRAYQKVKK FNC DALFTKYID MLI NFRLLKAYVD VMI NERLLKAYVD VMI NERLLKAYVD VMI NERLLKAYVD VMI NERLLKAYVD VMI NERLLKAYVD VMI NERLLKAYVD VMI C 160 	C C * 80 GRSKE TPEG-HDIKRWIC SKVEJ GFEA-GAAKGEFF VORGA APEA-CAAKGEFF VORGA APEA-KELKSHIF DRGR TPDG-KALKETLE DKVK DEKS-GDFKKDIA DGF TPDA-KELKSHIF DKGR TPEG-KELKSHIF VGRG TPEG-KELKSHIF CGP SCA-ADFKKALL SGGP TSGG-TDFKKVLI SGGP TSGG-TDFKKVLI SGGP TSGG-TDFKKVLI SGGP TSGG-TDFKKVLI SGGP TSGG-TDFKKVLI SGGP TEGG-KALKRLLE SGK TPYG-KELKSHIF C * 180	* 100 DEGTETR AK TPKQKVLVA DALSDA AK TPKQKVLVA CEALETH LK TDIQKEQTK TDALEND EK TDRQREDA DALEND EK TDRQREDA DALEND EK TDRQREDA DALEND EK TDRQREDA DALEND EK TDRQREDA DALEND EK TEAQQKGAE PIYLKTA GP SPKEKEIFQ DALSGA AK TENENKHVR PIYLKTA GP SPKEKEIFQ ULIGAA PC SPQETKCIQ PETLITA AK SPIQKHVR PDAVKK EK SDKQKELIR DALANG AK TENENKTAIR PDALANG AK TENENKTAIR PEALETK VR TERQKTAK C C 200	VFF 95 KVV 89 EVM 87 VFL 86 KAF 88 KVF 91 KVI 88 RVI 86 KKH 83 KHL 86 LFI 83 KTL 87 KTV 87 IVV 88 KTI 87 KTV 87 IVV 88 KVI 89 KII 96 QLA 86 RVI 91
В	EoblCSP2 : EoblCSP3 : EoblCSP4 : EoblCSP4 : EoblCSP6 : EoblCSP6 : EoblCSP8 : EoblCSP10: EoblCSP13: EoblCSP14: EoblCSP14: EoblCSP15: EoblCSP15: EoblCSP16: EoblCSP17: EoblCSP17: EoblCSP12 : EoblCSP21 : EoblCSP2 : EoblCSP4 : EoblCSP4 : EoblCSP4 : EoblCSP4 : EoblCSP4 : EoblCSP4 : EoblCSP4 : EoblCSP4 : EoblCSP4 : EoblCSP5 : EoblCSP4 : EoblCSP4 : EoblCSP4 : EoblCSP5 : EoblCSP4 : EoblCSP5 : EoblCSP4 : EoblCSP5	MKVLILLTIVVLASARP-D MKSFYILFAFFAVCAAQTVAPATSTS MKGIVCVCLLAMAAAALARPDO MKLVLITLAGVAATAYGKPA WLVISLLSVFSVALTQCE MKV		C 60 NSRLLKAYAL FL NHDAMKQYVD FT NSRLRDPYLS VLI SQRLLRGYVDP LLI DLPTLQAFVG FNI NRRLLFYVN VLI NRRLLFYVN VLI NRRLLFYVN VLI NRRLLFYVN VLI DEKWFQGFHD LMI DEKWFQGFHD LMI DEKLKYYK VMI DERILLAYYK VMI DERILLAYYK VMI DKEAVQKVLK FNG DKAFIQRQLK ALI NRRLLKAYVD VMI TPRLFHGYVL VSS NKRLVRIYVD LTY PRLLNSYAK LLI NKRLVPYIN VLI 160	C C * 80 BRSK TPEG-HDIKKWII SKVE GPEA-GAAKGEFF VOGK APEA-KELKSHE DRGNTPDG-KALKETLE LIGF TPDA-KELKSHE DKVK DEKS-GDFKKDIA DQGF TFIG-RVFKQVIE DEGR TADG-KELKSHE DGGF TFIG-RVFKQVIE SGRT TPEG-ADFKKALE SGRT TPGG-KNFKRVLE SGRT SEG-DFKKVIE SGR SEG-DFKKVIE SGR SEG-REIKEHLE SGR TPGG-KELKENME SVR TPEG-KALKRLE NGF TPEG-KALKRLE NGF TPEG-KALKSHIF C * 180	* 100 DEGTETR AK TPKOKVLVA DDALSDA AK TQVQKHTSK KEALETH LK TDIQKEQTK PDALEND SK TAKQKSGSD DDAIEND EK TDRQKDGAD HEAIQOA AK TDAQKHIFK PEVITTA KK TELQQKGALE DDALOGS VK TEQQKGALE DDALOGS VK TEQQKGALE PIYLKTA GP SPKEKEIFQ VELTTA AK SPKQKVVR PLALGSA AE TPEQKHNTK PLALGSA AE TPEQKHNTK PLALGSA AK SPIQKKHVR PLALGSA AK SPIQKKHVR PDAVKK EK SDKQKELR PDALKTK VR TERQKRAV PEALETR AK TEKQKRAV VELLETR AK TEKQKQMGK C C * 200	VFF 95 KVV 99 EVM 87 VFL 86 KAF 88 KVF 91 KVI 86 KKK 83 KKL 86 KKK 83 KKL 86 KKT 87 KTV 87 IVV 88 KVI 89 KVI 89 KVI 89 RVI 91
В	EoblCSP2 : EoblCSP3 : EoblCSP4 : EoblCSP5 : EoblCSP6 : EoblCSP6 : EoblCSP6 : EoblCSP10 : EoblCSP12 : EoblCSP13 : EoblCSP14 : EoblCSP15 : EoblCSP15 : EoblCSP15 : EoblCSP1 : EoblCSP2 : EoblCSP2 : EoblCSP2 : EoblCSP3 : EoblCSP3 : EoblCSP3 : EoblCSP3 : EoblCSP3 : EoblCSP3 : EoblCSP3 : EoblCSP3 : EoblCSP3 : EoblCSP5 : Eobl	MKVLILLITJVVLASARP-D 		C 60 NSRLLKAYAL FLC NNEDAHKQYDD FTC SQRLLRGYDD FTC SQRLLRGYDD LLI NERLLTGYUND LLI NERLLTGYUND LLI NERLLKAYUN FLI NKRLL FYVNN VLI DEKWFQGFHD LMI NERLLAYYK VMI NERLLAYYK VMI NERLLAYYK VMI NERLLAYYK VMI NERLLAYYK VMI NRELLAYYK VMI NRELLAYYK VMI NRELLKAYVD VKI NKRLLKAYVD VMI C 160 	C C * 80 BRSK TPEG-HDIKRWIC SKVEJ GFEA-GARGEFT VORG APEA-KELKSHIF DRGR TPDG-KALKETIF DRGR TPDG-KALKETIF DRGR TPGG-KELKSHIF VGK DEKS-GDFKKDIA DGF TPDA-KELKSHIF C C * 180	* 100 DEGTETR AK TPKQKVLVA DALSDA AK TPKQKVLVA CEALETH LK TDIQKEQTK DALEHE SK TAKQKSGSD DALEND EK TDRQRDGAD LEALEYA AK TDAQKHIFK TDAQKAK TDAQKHFK EQUITTA KK TPLQKRTTY PLALEXT AK TEAQQKGAE PIYLKTA GP SPKEKEIFQ DALAGSA AK TENAMTAIR POAVKTK EK SDKQKHVR DALANG AK TENAMTAIR DALAKG AK TENAMTAIR DALAKG AK TENAMTAIR DALAKG AK TENAMTAIR PEALETR KK TEKQKKGVK KEALEYY AK TDVQRSGTR C C * 200	VFF 95 KVV 91 KVV 89 EVM 87 VFL 86 KAF 88 KVF 91 KVI 86 KVI 86 KHK 83 KTL 87 KTV 87 TVL 88 KTL 87 KTV 87 TVL 88 KTI 96 QLA 86 RVI 91
В	EoblCSP2 : EoblCSP3 : EoblCSP3 : EoblCSP4 : EoblCSP6 : EoblCSP6 : EoblCSP6 : EoblCSP10 : EoblCSP11: EoblCSP12: EoblCSP14: EoblCSP14: EoblCSP15: EoblCSP16: EoblCSP16: EoblCSP17: EoblCSP2 : EoblCSP2 : EoblCSP3 : EoblCSP3 : EoblCSP3 : EoblCSP3 : EoblCSP3 : EoblCSP3 : EoblCSP3 : EoblCSP3 : EoblCSP3 : EoblCSP5 :	MKVLILLITJVVLASARP-D 		C 60 NSRLLKAYAL FLK NNEDAMKQYUD FTI SQRLLRGYUD LLI SQRLLRGYUD LLI DLPTLQAFVG FNI NERLLTGYUN LLI NERLLKYLN FLI NKRLLAYVN VI DEKAVQKVLK FNI DEKAVQKVLK FNI DEKAVQKVLK FNI DREAVQKVLK FNI DREAVQKVLK FNI DREAVQKVLK FNI NRRLLAYYD UT TRLFHGYUL LSI NKRLLKAYDD WLI C 160 	C C * 80 BRSK TPEG-HDIKRWII SKVEJ GFEA-GARGEFF VOGK APEA-KELKSHIF DRGR TPDG-KALKETIF DRGR TPDG-KALKETIF DRGR TPGG-KELKSHIF SKGF TPEG-KELKSHIF SKGF TPEG-ADFKKALF SKGF TPEG-ADFKKALF SKGF CFEA-ADFKKALF SKGF CFEA-ADFKKALF SKGF CFEA-ADFKKALF SKGF CFEA-ADFKKALF SKGF CFEA-ADFKKALF SKGF CFEA-ADFKKALF SKGF CFEA-ADFKKALF SKGF CFEA-ADFKKALF SKGF SFG-KELKSHIF C 180	* 100 DEGTETR AK TPKQKVLVA DALSDA AK TPKQKVLVA EGALETH LK TDIQKEQTK DALEHE SK TAKQKSGSD DALEND EK TDRQRDGAD LEAIQQA AK TDAQKHIFK EDVITTA KK TPLQKRTTY EGALETD AK TEAQCKGAL PIYLKTA GP SPKEKEIFQ DALSGY KK TGAQCKGAL PIYLKTA GP SPKEKEIFQ LALGSA AE TPEQKHNTK LALGSA AE TPEQKHNTK LALGSA AE TPEQKHNTK UVLRGA PQ SPQETKQTQ DALANG AK SPIQKKHVR DALKT KK SDKQKELIR DALANG AK TENNHTAIR PALKTR KK TEKQKKGVK (EALETY AK TEKQCMGK (EALETY AK TEKQCMGK (EALETY AK TEKQCMGK) * 200	VFF 95 KVV 89 EVM 87 VFL 86 KAF 88 KVF 91 KVI 86 KVI 86 KVI 86 KKH 83 KML 86 KTL 87 KTV 87 IVV 88 KTL 87 KTV 87 IVV 88 KTI 96 QLA 86 RVI 91 124 128 128 123 125
В	EoblCSP2 : EoblCSP3 : EoblCSP3 : EoblCSP4 : EoblCSP6 : EoblCSP6 : EoblCSP6 : EoblCSP10 : EoblCSP11: EoblCSP12: EoblCSP14: EoblCSP14: EoblCSP15: EoblCSP16: EoblCSP16: EoblCSP17: EoblCSP2 : EoblCSP2 : EoblCSP3 : EoblCSP3 : EoblCSP3 : EoblCSP3 : EoblCSP3 : EoblCSP3 : EoblCSP3 : EoblCSP3 : EoblCSP3 : EoblCSP5 :	MKVLILLITJVVLASARP-D 		C 60 NSRLLKAYAL FLK NNEDAMKQYUD FTI SQRLLRGYUD LLI SQRLLRGYUD LLI DLPTLQAFVG FNI NERLLTGYUN LLI NERLLKYLN FLI NKRLLAYVN VI DEKAVQKVLK FNI DEKAVQKVLK FNI DEKAVQKVLK FNI DREAVQKVLK FNI DREAVQKVLK FNI DREAVQKVLK FNI NRRLLAYYD UT TRLFHGYUL LSI NKRLLKAYDD WLI C 160 	C C * 80 BRSK TPEG-HDIKRWII SKVEJ GFEA-GARGEFF VOGK APEA-KELKSHIF DRGR TPDG-KALKETIF DRGR TPDG-KALKETIF DRGR TPGG-KELKSHIF SKGF TPEG-KELKSHIF SKGF TPEG-ADFKKALF SKGF TPEG-ADFKKALF SKGF CFEA-ADFKKALF SKGF CFEA-ADFKKALF SKGF CFEA-ADFKKALF SKGF CFEA-ADFKKALF SKGF CFEA-ADFKKALF SKGF CFEA-ADFKKALF SKGF CFEA-ADFKKALF SKGF CFEA-ADFKKALF SKGF SFG-KELKSHIF C 180	* 100 DEGTETR AK TPKQKVLVA DALSDA AK TPKQKVLVA EGALETH LK TDIQKEQTK DALEHE SK TAKQKSGSD DALEND EK TDRQRDGAD LEAIQQA AK TDAQKHIFK EDVITTA KK TPLQKRTTY EGALETD AK TEAQCKGAL PIYLKTA GP SPKEKEIFQ DALSGY KK TGAQCKGAL PIYLKTA GP SPKEKEIFQ LALGSA AE TPEQKHNTK LALGSA AE TPEQKHNTK LALGSA AE TPEQKHNTK UVLRGA PQ SPQETKQTQ DALANG AK SPIQKKHVR DALKT KK SDKQKELIR DALANG AK TENNHTAIR PALKTR KK TEKQKKGVK (EALETY AK TEKQCMGK (EALETY AK TEKQCMGK (EALETY AK TEKQCMGK) * 200	VFF 95 KVV 89 EVM 87 VFL 86 KAF 88 KVF 91 KVI 86 KVI 86 KVI 86 KKH 83 KML 86 KTL 87 KTV 87 IVV 88 KTL 87 KTV 87 IVV 88 KTI 96 QLA 86 RVI 91 124 128 128 123 125
В	EoblCSP2 : EoblCSP3 : EoblCSP3 : EoblCSP4 : EoblCSP6 : EoblCSP6 : EoblCSP6 : EoblCSP10 : EoblCSP11: EoblCSP12: EoblCSP14: EoblCSP14: EoblCSP15: EoblCSP16: EoblCSP16: EoblCSP17: EoblCSP2 : EoblCSP2 : EoblCSP3 : EoblCSP3 : EoblCSP3 : EoblCSP3 : EoblCSP3 : EoblCSP3 : EoblCSP3 : EoblCSP3 : EoblCSP3 : EoblCSP5 :	MKVLILLITJVVLASARP-D 		C 60 NSRLLKAYAL FLK NNEDAMKQYUD FTI SQRLLRGYUD LLI SQRLLRGYUD LLI DLPTLQAFVG FNI NERLLTGYUN LLI NERLLKYLN FLI NKRLLAYVN VI DEKAVQKVLK FNI DEKAVQKVLK FNI DEKAVQKVLK FNI DREAVQKVLK FNI DREAVQKVLK FNI DREAVQKVLK FNI NRRLLAYYD UT TRLFHGYUL LSI NKRLLKAYDD WLI C 160 	C C * 80 BRSK TPEG-HDIKRWII SKVEJ GFEA-GARGEFF VOGK APEA-KELKSHIF DRGR TPDG-KALKETIF DRGR TPDG-KALKETIF DRGR TPGG-KELKSHIF SKGF TPEG-KELKSHIF SKGF TPEG-ADFKKALF SKGF TPEG-ADFKKALF SKGF CFEA-ADFKKALF SKGF CFEA-ADFKKALF SKGF CFEA-ADFKKALF SKGF CFEA-ADFKKALF SKGF CFEA-ADFKKALF SKGF CFEA-ADFKKALF SKGF CFEA-ADFKKALF SKGF CFEA-ADFKKALF SKGF SFG-KELKSHIF C 180	* 100 DEGTETR AK TPKQKVLVA DALSDA AK TPKQKVLVA EGALETH LK TDIQKEQTK DALEHE SK TAKQKSGSD DALEND EK TDRQRDGAD LEAIQQA AK TDAQKHIFK EDVITTA KK TPLQKRTTY EGALETD AK TEAQCKGAL PIYLKTA GP SPKEKEIFQ DALSGY KK TGAQCKGAL PIYLKTA GP SPKEKEIFQ LALGSA AE TPEQKHNTK LALGSA AE TPEQKHNTK LALGSA AE TPEQKHNTK UVLRGA PQ SPQETKQTQ DALANG AK SPIQKKHVR DALKT KK SDKQKELIR DALANG AK TENNHTAIR PALKTR KK TEKQKKGVK (EALETY AK TEKQCMGK (EALETY AK TEKQCMGK (EALETY AK TEKQCMGK) * 200	VFF 95 KVV 89 EVM 87 VFL 86 KAF 88 KVF 91 KVI 86 KVI 86 KVI 86 KKH 83 KML 86 KTL 87 KTV 87 IVV 88 KTL 87 KTV 87 IVV 88 KTI 96 QLA 86 RVI 91 124 128 128 123 125
В	EoblCSP2 : EoblCSP3 : EoblCSP3 : EoblCSP4 : EoblCSP6 : EoblCSP6 : EoblCSP6 : EoblCSP10 : EoblCSP11: EoblCSP12: EoblCSP14: EoblCSP14: EoblCSP15: EoblCSP16: EoblCSP16: EoblCSP17: EoblCSP2 : EoblCSP2 : EoblCSP3 : EoblCSP3 : EoblCSP3 : EoblCSP3 : EoblCSP3 : EoblCSP3 : EoblCSP3 : EoblCSP3 : EoblCSP3 : EoblCSP5 :	MKVLILLITJVVLASARP-D 		C 60 NSRLLKAYAL FLK NNEDAMKQYUD FTI SQRLLRGYUD LLI SQRLLRGYUD LLI DLPTLQAFVG FNI NERLLTGYUN LLI NERLLKYLN FLI NKRLLAYVN VI DEKAVQKVLK FNI DEKAVQKVLK FNI DEKAVQKVLK FNI DREAVQKVLK FNI DREAVQKVLK FNI DREAVQKVLK FNI NRRLLAYYD UT TRLFHGYUL LSI NKRLLKAYDD WLI C 160 	C C * 80 BRSK TPEG-HDIKRWII SKVEJ GFEA-GARGEFF VOGK APEA-KELKSHIF DRGR TPDG-KALKETIF DRGR TPDG-KALKETIF DRGR TPGG-KELKSHIF SKGF TPEG-KELKSHIF SKGF TPEG-ADFKKALF SKGF TPEG-ADFKKALF SKGF CFEA-ADFKKALF SKGF CFEA-ADFKKALF SKGF CFEA-ADFKKALF SKGF CFEA-ADFKKALF SKGF CFEA-ADFKKALF SKGF CFEA-ADFKKALF SKGF CFEA-ADFKKALF SKGF CFEA-ADFKKALF SKGF SFG-KELKSHIF C 180	* 100 DEGTETR AK TPKQKVLVA DALSDA AK TPKQKVLVA EGALETH LK TDIQKEQTK DALEHE SK TAKQKSGSD DALEND EK TDRQRDGAD LEAIQQA AK TDAQKHIFK EDVITTA KK TPLQKRTTY EGALETD AK TEAQCKGAL PIYLKTA GP SPKEKEIFQ DALSGY KK TGAQCKGAL PIYLKTA GP SPKEKEIFQ LALGSA AE TPEQKHNTK LALGSA AE TPEQKHNTK LALGSA AE TPEQKHNTK UVLRGA PQ SPQETKQTQ DALANG AK SPIQKKHVR DALKT KK SDKQKELIR DALANG AK TENNHTAIR PALKTR KK TEKQKKGVK (EALETY AK TEKQCMGK (EALETY AK TEKQCMGK (EALETY AK TEKQCMGK) * 200	VFF 95 KVV 89 EVM 87 VFL 86 KAF 88 KVF 91 KVI 86 KVI 86 KVI 86 KKH 83 KML 86 KTL 87 KTV 87 IVV 88 KTL 87 KTV 87 IVV 88 KTI 96 QLA 86 RVI 91 124 128 128 123 125
В	EoblCSP2 : EoblCSP3 : EoblCSP3 : EoblCSP4 : EoblCSP6 : EoblCSP6 : EoblCSP1 : EoblCSP1 : EoblCSP12 : EoblCSP13 : EoblCSP14 : EoblCSP14 : EoblCSP14 : EoblCSP15 : EoblCSP15 : EoblCSP1 : EoblCSP1 : EoblCSP2 : EoblCSP3 : Eobl			C 60 NSRLKAYAL FI NIDAMKQYDD FI SQRLLRGYVD LLI SQRLLRGYVD LLI DLPTLQAFVG FNI NRRLLFYVN VLI NRRLLFYVN VLI NRRLLFYVN VLI DEKWFQGFHD LMI DEKWFQGFHD LMI NERILLAYYK VMI DEKAVQKVLK FNO DKRF1QRQLKALK NKRLLKAYVD VMI NFRLHGYVL LSI NKRLLKNYAK FI NKRLLKNYAK FI NKRLLKNYAK FI NKRLLKNYAK FI 160	C C * 80 BRSK TPEG-HDIKKWII KVEZ GPEA-GAAKGEFI VOKK APEA-KELKSHI DAGR TPDG-KALKETLI LIGP TPDA-KELKSHI DAGR TPEG-KELKSHI VOKK TPEG-KELKSHI VOKK TPEG-ADFKKALI SOGP TKDG-KNFKRVLI -GO SSQA-AALKTDFI EAP DPIG-KRIKTLAFI SOGP TSEG-DFKKVII SOGP TSEG-TDFKKVII SOGF KELKSADFKKLLE SOGF SEG-KELKEHLC CGK TYG-KELKENHL C 180 * 180	* 100 DEGTETR AK TPKCKVLVA DALSDA AK TQVQKHTSK EALETH LK TDIQKEQTK DALEHE SK TAKQKSGSD DAIEND EX TDRQRDGAD LEAIQOA AK TDAQKHIFK UEALETD AK TEAQQKGAE DDALSG VK TGPQRAGIR DALGSG VK TGPQRAGIR PIYLKTA GP SPKEKEIFO PETLITA AR SPKQKUVO PETLITA AR SPKCKUVO PETLITA AK SPKCKUVO PETLITA AK SPKCKUVO PALANGA AE TPEQKHNTK LVLRGA PQ SPCKNUV DAUKTK EK SDKQKEUVO DALANG AK TENNHTAIR DALANG AK TENNHTAIR PEALETK VR TERQKRTAV DALANG AK TENQKOWCK (EALETY AK TEXQCMCK 200	VFF 95 KVV 89 EVM 87 VFL 86 KAF 88 KVF 91 KVI 86 KVI 88 RVI 86 KVI 86 KVI 87 IVV 88 KTL 87 KTV 87 IVV 88 KTV 87 IVV 88 KVI 99 CL24
В	EoblCSP2 : EoblCSP3 : EoblCSP4 : EoblCSP4 : EoblCSP5 : EoblCSP6 : EoblCSP6 : EoblCSP10 : EoblCSP10 : EoblCSP13 : EoblCSP14 : EoblCSP15 : EoblCSP15 : EoblCSP16 : EoblCSP17 : EoblCSP2 : EoblCSP4 : EoblCSP4 : EoblCSP4 : EoblCSP4 : EoblCSP4 : EoblCSP5 : EoblCSP5 : EoblCSP5 : EoblCSP6 : EoblCSP6 : EoblCSP6 : EoblCSP6 : EoblCSP6 : EoblCSP6 : EoblCSP13 : EoblCSP14 : EoblCSP13 : EoblCSP13 : EoblCSP13 : EoblCSP14 : EoblCSP13 : EoblCSP13 : EoblCSP13 : EoblCSP13 : EoblCSP13 : EoblCSP13 : EoblCSP14 : EoblCSP13 : EoblCSP13 : EoblCSP14 : EoblCSP13 : EoblCSP14 : EoblCSP15 : EoblCSP15 : EoblCSP15 : EoblCSP15 : EoblCSP15 : EoblCSP15 : EoblCSP15 : EoblCSP16 : Eobl	MKVLILLTIVVLASARP-D MKSIVCVCLLAMAAAALARPO MKGIVCVCLLAMAAAALARPO MKLVILIILAGVAATAYGKPA WLVISLUSYFSVALTQCE MKVVVSFULMTVVAQEKY 		C 60 NSRLLKAYAL FL NHDAMKQYDD FT NSRLRDPYLS VLI SQRLLRGYVDD LLI DLPTLQAFVG FNI NERLLTGYVN LLI NKRLLIPYVN VLI NKRLLIPYVN VLI NKRLLKAYIN VLI NKRLLKAYN FI DEKWFQGFHD LMN DKEFIQRQLKALC DALFFKYID MLI NFRLLKNYAK FL NKRLLKAYVD VM TPRLFHGYYL LSS NKRLVRYVD LT NKRLVRYVD CT NRRLLNYAK FL 160	C C * 80 BRSK TPEG-HDIKKWII SKVE GPEA-GAAKGEFI VOGK APEA-KELKSHIF DRGR TPDG-KALKETIFI LIGP TPDA-KELKSHIF DRGR TPGG-KELKSHIF VOGK TFEG-RVFKQVLE EEGR TADG-KELKSHIF VOGK TFEG-ADFKKALF SKGR TPEG-ADFKKALF -GD SSQA-AALKTDFF BEAP DPIG-KRIKTIAF NGCP TVGC-KNFKRVLI -GD SSQA-AALKTDFF BEAP DPIG-KRIKTIAF NGCP CHISADFKKLLE SKGR SPEG-KELKEHL NGCP TEGG-KELKEHL NGCP TPG-KELKSHIF C * 180	* 100 DEGTETR AK TPKOKVLVAK PDALSDA AK TQVQKHTSK EGALETH LK TDIQKEQTK PDALEND SK TAKQKSGSD DAIEND EK TDRQRHIK VEALETT AK TPLQKRFTY KEALETT AK TEAQOKGAE DDALSOG VK TEAQOKGKIR PIYLKTA GP SPKEKEIFQ PULTATA K SPKQKVVR PLALSSA AE TPEQKHTK PLVLRGA PQ SPQETKQIQ PULTAS AK SPIQKKHVR PLVLRGA PX SPCKKUVY PLALSSA AE TPEQKHTKY PALANG AK TENMMTAIR PDAVKT EK SPKQKGVR PDALANG AK TENMMTAIR PEALETR AK TEKQKQMGK KEALETY AK TEVQRSTR C C	VFF 95 KVV 89 EVM 87 VVF 86 KAF 88 KVF 91 KVI 86 KKV 83 KVI 86 KKK 83 KKL 86 KKK 83 KKT 87 KTV 87 IVV 88 KVI 89 KVI 89 CLA 86 RVI 91
В	EoblCSP1 : EoblCSP3 : EoblCSP3 : EoblCSP4 : EoblCSP6 : EoblCSP6 : EoblCSP7 : EoblCSP1 : EoblCSP13 : EoblCSP14 : EoblCSP14 : EoblCSP14 : EoblCSP15 : EoblCSP15 : EoblCSP15 : EoblCSP1 : EoblCSP1 : EoblCSP1 : EoblCSP2 : EoblCSP3 : EoblCSP3 : EoblCSP4 : EoblCSP4 : EoblCSP5 : EoblCSP5 : EoblCSP5 : EoblCSP5 : EoblCSP5 : EoblCSP5 : EoblCSP5 : EoblCSP5 : EoblCSP1 : Eobl	MKVLILLTIVVLASARP-D MKSFYILFAFFAVCAAQTVAPATSTS MKGIVCVCLLAMAAAALARPDO MKLVILIILAGVAATAYGKPA WLVISLLSVFSVALTQCE MKVAVVLCILSFAAVAFSAPSO MKVAVVLCILSFAAVAFSAPSO MKLLILAVLITLASLHTLAQT MKS		C 60 NSRLLKAYAL FLK NHDAMKQYVD FTY NSRLRDPYLS VLI SQRLLRGYVD LLI DLPTLQAFVG FNI NRRLLFYVN VLI NRRLLFYVN VLI NRRLLFYVN VLI DKWFQGFHD LMN DREINLJYVN FIL DEKWFQGFHD LMN DREINLYVN FIL DEKWFQGFHD LMN NRRLLKYVN FNI DRENLSYAK LLI NRRLVRYVD VM TPRLLNYAK FLS NKRLVPY IN VLI C 160	C C * 80 BRSK TPEG-HDIKKWII SKVE GPEA-GAAKGEFF VOKK APEA-KELKSHLF DRGNTPDG-KALKETLF LIGP TPDA-KELKSHLF SKUE DEKS-GDFKKDIA DQGF TPIG-RVFKQVLF EGGR TADG-KELKSHLF VGKGT FEG-ADFKKALF SGGD SSQA-AALKTDFF SEAP DPIG-KNEKKVLF SGGD SSQA-AALKTDFF SALKTON SGGT KEGK-CHEKKSHIF C * 180 * 180	* 100 DEGTETR AK TPKOKVLVA PDALSDA AK TQVQKHTSK EALETH LK TDIQKEQTK PDALENE SK TAKQKSGSD DDAIEND EX TDRQRAGAD HEALGOA AK TDAQKHIFK TEALETT AK TELQCKGAL PUITTA KK TPLQKRFY TEALETT AK TELQCKGAL PUITTA AK SPKQKVVR PLALGSA AE TPEQKHNTK PLALGSA AE TPEQKHNTK C C * 200	VFF 95 KVV 89 EVM 87 VFL 86 KAF 88 KVF 91 KVI 86 KVI 88 RVI 86 KKH 83 KKL 87 KTL 87 KTL 87 KTL 87 KTL 87 KTV 88 RVI 91
В	EoblCSP2 : EoblCSP3 : EoblCSP3 : EoblCSP4 : EoblCSP6 : EoblCSP6 : EoblCSP10: EoblCSP10: EoblCSP13: EoblCSP13: EoblCSP14: EoblCSP14: EoblCSP15: EoblCSP17: EoblCSP17: EoblCSP16: EoblCSP17 : EoblCSP20: EoblCSP4 : EoblCSP4 : EoblCSP4 : EoblCSP5 : EoblCSP5 : EoblCSP6 : EoblCSP17 : EoblCSP6 : EoblCSP6 : EoblCSP6 : EoblCSP12: EoblCSP13: EoblCSP13: EoblCSP13: EoblCSP14: EoblCSP14: EoblCSP14: EoblCSP14: EoblCSP14: EoblCSP14: EoblCSP14: EoblCSP14: EoblCSP14: EoblCSP14: EoblCSP14: EoblCSP14: EoblCSP15: EoblCSP14: EoblCSP14: EoblCSP15: EoblCSP14: EoblCSP15: EoblCSP15: EoblCSP15: EoblCSP16: EoblCSP17: EoblCSP16			C 60 NSRLLKAYAL FL NHDAMKQYDD FT NSRLRDPYLS VLI SQRLLRGYVD LLI DLPTLQAFVG FNI NERLLTGYVN LLI NKRLLIPYVN VLI NKRLLIPYVN VLI NKRLLAYIN VLI DEKWFQGFHD LMN DERILLAYYK VMI DERILLAYYK VMI DERILLAYYK VMI DERILLAYYK VMI DERILLAYYK VMI DERILLAYYK VMI NERILLAYYK VMI NERILLAYYK VMI NERILLAYYK VMI NERLLNYAK FLI NERLLNYAK FLI NERLLNYAK FLI NERLLNYAK LLI NERLLYYV D LT NERLLNYAK LLI NERLLYYV D LT 160	C C * 80 SRSK TPEG-HDIKKWII SKVE GPEA-GAAKGEFI VOGK APEA-KELKSHIF SRGR TPDG-KALKETIFI LIGPT PDA-KELKSHIF SKVK DEKS-GDFKKDIA DOGT PTIG-RVFKQVLE EEGR TADG-KELKSHIF VOGK TPEG-ADFKKALI SKGR TPEG-ADFKKALI SGGP TOG-KNFKRVLI SGGP TOG-KNFKRVLI SGGP TOG-KNFKRVLI SGGP CHSADFKKLIKLAFI SGGP CHSADFKKLIKLAFI SGGP CHSADFKKLIKLAFI SGGP CHSADFKKLIKLAFI SGGP TEGG-TDFKKVII SGGP SEG-CALKKLIKLAFI SGGP TEGG-RELKENMI SVR TPEG-KELKENMI SVR TPEG-KALKRLLI SVR TPEG-KALKRLI SVR TPEG-KALKRLIFI C * 180	* 100 DEGTETR AK TPKOKVLVAK PDALSDA AK TQVQKHTSK EGALETH LK TDIQKEOTK PDALEND SK TAKQKSGD DDAIEND EK TDRQRHTK VEALETT AK TPLQKRFTY VEALETT GK SDKQKGNIR IYLKTA GP SPKEKEIFQ PUVITA K SPKQKVVR PLALGSA AE TPEQKHTK PLVLRGA PQ SPQETKQIQ PUVLRGA PQ SPQETKQIQ PUVLRGA AK TENUMTAIR PLVLRGA PX SPKQKUVR PLALST AK SPIQKKHVR PDAVKT EK SPIQKKHVR PDAVKT EK TEKQKQKSTR C C * 200	VFF 95 KVV 99 KVV 89 EVM 87 VFL 86 KAF 88 KVF 91 KVI 86 KKF 83 KKL 86 KKK 83 KKL 86 KKT 87 KTV 87 IVV 88 KVI 89 KVI 89 CLA 86 RVI 91
В	EoblCSP2 : EoblCSP3 : EoblCSP3 : EoblCSP4 : EoblCSP6 : EoblCSP6 : EoblCSP1 : EoblCSP1 : EoblCSP13 : EoblCSP13 : EoblCSP14 : EoblCSP14 : EoblCSP14 : EoblCSP15 : EoblCSP15 : EoblCSP1 : EoblCSP2 : EoblCSP2 : EoblCSP3 : Eobl	MKVLILLITUVULASARP-D 		C 60 NSRLKAYAL FX NSLRDPYLS VI SQRLLRGYVD LL SQRLLRGYVD LL DLPTLQAFVG FNI NRRLLFYVN VL NRRLLFYVN VL NRRLLFYVN VL DEWFQGFHD LM NRRLLFYVN VL DEWFQGFHD LM NERILLAYK VM DEKAVQKVLK FN DEKAVQKVLK FN DRFILLAYK VM NRRLLAYK VM NRRLLKNYAK FL NRRLLKNYAK FL SC 160	C C * 80 BRSK TPEG-HDIKKWII KVEI GPEA-GAAKGEFI VOKK APEA-KELKSHIF DRGR TPDG-KALKETLI LIGP TPDA-KELKSHIF DRGR TPEG-KELKSHIF VOK TPEG-KELKSHIF VOK TPEG-ADFKALI SOGP TKDG-KNFKRVLI -GO SOGA-AALKTDFF EAP DPIG-KRLKTLAF VGK TPEG-KELKSHIF C 180 * 180	* 100 DEGTETR AK TPKCKVLVA DALSDA AK TQVQKHTSK EALETH LK TDIQKEQTK DALEHE SK TAKQKSGSD DAIEND EX TDRQRAGAD LEAIQOA AK TDAQKHIFK USALETD AK TEAQCKGAE DDALSG VK TGPQRAGIR PAILTT GK SDKQKGVIR PIYLKTA GP SPKEKEIFQ PETLITA AR SPKCKUVQ PETLITA AK SPKCKUVQ PETLITA AK SPKCKUVQ DALANGA AE TPEQKHNTK LVLRGA PQ SPCKNUV PDAVKTK EK SDKQKEUFQ DALANG AK TENHMTAIR PDALST AK SPLQKKUVQ PDAVKTK EK SDKQKEUFQ DALANG AK TENHMTAIR C C * 200	VFF 95 KVV 89 EVM 87 VFL 86 KAF 88 KVF 91 KVI 86 KVI 88 RVI 86 KVI 87 IVI 86 KTL 87 KTV 87 IVV 88 KTL 87 KTV 87 IVV 88 KTI 96 QLA 86 RVI 91
В	EoblCSP2 : EoblCSP3 : EoblCSP3 : EoblCSP4 : EoblCSP6 : EoblCSP6 : EoblCSP10 : EoblCSP10 : EoblCSP11 : EoblCSP13 : EoblCSP14 : EoblCSP15 : EoblCSP15 : EoblCSP16 : EoblCSP17 : EoblCSP17 : EoblCSP2 : EoblCSP4 : EoblCSP4 : EoblCSP4 : EoblCSP5 : EoblCSP5 : EoblCSP5 : EoblCSP5 : EoblCSP6 : EoblCSP12 : EoblCSP6 : EoblCSP13 : EoblCSP6 : EoblCSP13 : EoblCSP13 : EoblCSP13 : EoblCSP13 : EoblCSP13 : EoblCSP13 : EoblCSP13 : EoblCSP14 : EoblCSP14 : EoblCSP14 : EoblCSP14 : EoblCSP14 : EoblCSP14 : EoblCSP15 : EoblCSP14 : EoblCSP15 : EoblCSP15 : EoblCSP16 : EoblCSP16 : EoblCSP16 : EoblCSP16 : EoblCSP16 : EoblCSP16 : EoblCSP16 : EoblCSP16 : EoblCSP17 : EoblCSP16 : EoblCSP17 : EoblCSP16 : EoblCSP16 : EoblCSP16 : EoblCSP16 : EoblCSP17 : EoblCSP16 :	MKVLILLTIVVLASARP-D MKSIVCVCLLAMAAAALARPO MKGIVCVCLLAMAAAALARPO MKLVILIILAGVAATAYGKPA WLVISLUSYFSVALTQCE MKVVVSFULMTVVAQEKY 		C 60 NSRLLKAYAL FL NHDAMKQYDD FT NSRLRDPYLS VLI SQRLLRGYVDD LLI NERLLTGYVNN LLI DLPTLQAFVG FNI NRRLLAYIN VLI NKRLLIPYVN VLI NKRLLAYIN VLI NKRLLKAYIN VLI DEKWFQGFHD LMI DEKWFQGFHD LMI NKRLLKAYVN FI DEKWFQGFUC ALC DALFFKYID MLI NFRLLKNYAK FL NKRLKAYVD VM NKRLVRYVD VT NFRLLNYAK FL NKRLVRYVD VT C 160	C C * 80 SRSK TPEG-HDIKKWII SKVE GPEA-GAAKGEFF VOGK APEA-KELKSHIF DRGR TPDG-KALKETIF LIGGT PDDA-KELKSHIF SKORT TPEG-RVFKQVLE EEGR TADG-KELKSHIF VOGK TFEG-RVFKQVLE SKGR TPEG-ADFKKALF SKGR TPEG-ADFKKALF -GD SSQA-AALKTDFF SEAP DPIG-KRIKTLAF SKGF EVEHSADFKKLIF SGGP TSGG-TDFKKVII SKGF EVEHSADFKKLIF SGGP TEGG-REIKEHLAF SKGF EVEHSADFKKLIF SKGF SPIG-KELKENH SVR TPEG-KALKRLIF RGK SPEG-KALKRLIF SKGF 180 * 180 C * 180	* 100 DEGTETR AK TPKOKVLVAK PDALSDA AK TQVQKHTSK EGALETH LK TDIQKEOTK PDALEND SK TAKQKSGD DDAIEND EK TDRQRHTK VEALETT AK TPLQKRFTY VEALETT GK SDKQKGNIR IYLKTA GP SPKEKEIFQ PUVITA K SPKQKOVVR PLALGSA AE TPEQKHTK PLVLRGA PQ SPQETKQIQ PLALGSA AE TPEQKHTK PLVLRGA PX SPKCKIFY PLALATK K TEKQKQKG PDAIENG EK TEKQEKGSY PDALANG AK TENMMTAIR PDAVKT EK SPKQKUGK PDAUKTY AK TEKQKQMGK C C * 200	VFF 95 KVV 89 EVM 87 VVF 86 KAF 88 KVF 91 KVI 86 KKF 83 KKF 83 KKF 86 KKF 83 KKF 86 KKF 87 KTV 87 IVV 88 KVI 89 KVI 89 KVI 89 CLA 86 RVI 91
В	EoblCSP2 : EoblCSP3 : EoblCSP3 : EoblCSP4 : EoblCSP6 : EoblCSP6 : EoblCSP10 : EoblCSP10 : EoblCSP11 : EoblCSP13 : EoblCSP14 : EoblCSP15 : EoblCSP15 : EoblCSP16 : EoblCSP17 : EoblCSP17 : EoblCSP2 : EoblCSP4 : EoblCSP4 : EoblCSP4 : EoblCSP5 : EoblCSP5 : EoblCSP5 : EoblCSP5 : EoblCSP6 : EoblCSP12 : EoblCSP6 : EoblCSP13 : EoblCSP6 : EoblCSP13 : EoblCSP13 : EoblCSP13 : EoblCSP13 : EoblCSP13 : EoblCSP13 : EoblCSP13 : EoblCSP14 : EoblCSP14 : EoblCSP14 : EoblCSP14 : EoblCSP14 : EoblCSP14 : EoblCSP15 : EoblCSP14 : EoblCSP15 : EoblCSP15 : EoblCSP16 : EoblCSP16 : EoblCSP16 : EoblCSP16 : EoblCSP16 : EoblCSP16 : EoblCSP16 : EoblCSP16 : EoblCSP17 : EoblCSP16 : EoblCSP17 : EoblCSP16 :	MKVLILLTIVVLASARP-D 		C 60 NSRLKAYAL FL NHDAMKQYVD FT NSRLRDPYLS VL SQRLLRGYVD LL DLPTLQAFVG FN NRRLLRYIN VL NKRLLIPYVN VL NKRLLAYIN VL DDALFAYN FL NKRLKAYN FL DALFYGYNC VM DRALKAYN FL DALFYGYNC VM DRALFYGYNC VM DRALFYC VC DALFYYV VL DALFYGYYL LS NKRLKAYVD VM TPRLFHGYYL LS NKRLVRIVVD VM TPRLSYXK VL NKRLVRYVD VM TC 160 C	C C * 80 SRSK TPEG-HDIKKWII SKVE GPEA-GAAKGEFI VOGK APEA-KELKSHIF SRGR TPDG-KALKETIFI LIGPT PDA-KELKSHIF SKVE GEKS-GDFKKDIA OGGT PIIG-RVFKQVLE EEGR TADG-KELKSHIF VOGK TPEG-ADFKKALI SKGR TPEG-ADFKKALI SGGP TOG-KNFRKVLI SGGP TOG-KNFRKVLI SGGP CANFRKVLE SGGP SGG-ALKTDFF SGGP SGG-KELKEHLF SGGP SGG-KELKSHIF C * 180 SGGP CANFRKVLE SGGP SGGP CANFRKVLE SGGP S	* 100 DEGTETR AK TPKOKVLVAK CDALSDA AK TQVQKHTSK (EALETH LK TDIQKEOTK DDALEND EK TDRQRADA LEAIQAA AK TDAQKHIFK EVITTA KK TPLQKRTY (EALETD AK TEAQQKGAE DDALQG VK TGQQAGIR 11YLKTA GP SPKEKIFQ DDALGG VK TGQQAGK PULALGSA AE TPEQKHTK VLALGSA AE TPEQKHTK VLALGSA AK TENMMTAIR PLALGTK K KK SPIQKKHVR PDAVKK EK SPIQKKHVR DDAVKK EK SPIQKKOVOR EVIATS AK SPIQKKUV PDALETG EK TEKQEKGSY DDALETG EK TEKQEKGSY DDALETG EK TEKQEKGSY DDALETG EK TEKQEKGSY DDALETG K TEKQKTAV FEALETY AK TEKQKUMCK (EALETY AK TEKQKUMCK (EALETY AK TEKQKUMCK (EALETY AK TDVQRSGTR C C * 200	VFF 95 KVV 89 EVM 87 VVF 86 KAF 88 KVF 91 KVI 86 KKF 83 KKF 83 KKF 86 KKF 83 KKF 86 KKF 87 KTV 87 IVV 88 KVI 89 KVI 89 KVI 89 CLA 86 RVI 91
В	EoblCSP2 : EoblCSP3 : EoblCSP3 : EoblCSP4 : EoblCSP5 : EoblCSP6 : EoblCSP6 : EoblCSP1 : EoblCSP12 : EoblCSP13 : EoblCSP14 : EoblCSP14 : EoblCSP15 : EoblCSP15 : EoblCSP16 : EoblCSP16 : EoblCSP16 : EoblCSP17 : EoblCSP2 : EoblCSP2 : EoblCSP4 : EoblCSP4 : EoblCSP4 : EoblCSP5 : EoblCSP5 : EoblCSP5 : EoblCSP5 : EoblCSP6 : EoblCSP16 : EoblCSP17 : EoblCSP16 : EoblCSP16 : EoblCSP16 : EoblCSP16 : EoblCSP16 : EoblCSP16 : EoblCSP16 : EoblCSP16 : EoblCSP16 : EoblCSP17 : EoblCSP16 : EoblCSP16 : EoblCSP17 : EoblCSP17 : EoblCSP16 : EoblCSP16 : EoblCSP16 : EoblCSP16 : EoblCSP17 : EoblCSP16 : EoblCSP16 : EoblCSP16 : EoblCSP16 : EoblCSP17 : EoblCSP16 : EoblCSP17 : EoblCSP16 :	MKVLILLTIVVLASARP-D 		C 60 NSRLLKAYAL FLK NHDAMKQYVD FTY NSRLRDPYLS VLI SQRLLRGYVD LLI NERLLTGYVN LLI DLPTLQAFVG FNI NRRLLPYVN VLI NKRLLFYVN VLI DKENFQGFHD LMI DEKWFQGFHD LMI DEKWFQGFHD LMI NERILLAYYN VMI DKEAVQKVLK FNC DKAFIQRQLK ALL NKRLLKYYN DMI NERILLSYNC DMI NERILLSYNC DMI NERLLNYAK FLS NKRLVPY IN VLI TPLINSYAK LLI NKRLVPY IN VLI 160 EDSEEKDESE	C C * 80 BRSK TPEG-HDIKKWII SKVE GPEA-GAAKGEFF VOGK APEA-KELKSHLF DRGN TPDG-KALKSTLF LIGP TPDA-KELKSHLF SKUE DEKS-GDFKKDIA DQGF TPIG-RVFKQVLF EGGR TADG-KELKSHT SGGP SKOE-ADFKKALF SGGP SKOE-KNFKRVLI SGGP SKOE-ADFKKALF SGGP SKOE-ADFKKALF SGGP SKOE-KNFKRVLI SGGP SKOE-KNFKRVLI SGGP SKOE-KNFKRVLI SGGP SKOE-KNFKRVLI SVR TPEG-KALKRLIF NGGP TEVE-VQVEKLKSHTP C * 180 * 180 * 180	* 100 DEGTETR AK TPKOKVLVA PDALSDA AK TQVQKHTSK EALETH LK TDIQKEQTK PDALEND EX TDRQRAGAD LEALTDA LK TDIQKEQTK PDALEND EX TDRQRAGAD HEALGOA AK TDAQKHIFK EALETT AK TELQCKGAL PLALETD AK TELQCKGAL PLALETA AK TELQCKGAL PLALGSA AE TPEQKHNTK PLALGSA AE TPEQKHNTK C C * 200 * 200 * 200	<pre>VFF 95 KVV 89 EVM 87 VVFL 86 KAF 88 KVF 91 KVI 86 KVF 91 KVI 86 KVI 87 IVV 88 KIL 87 KTL 87 KVI 87 IVV 88 KVI 89 CVI 91</pre>

Fig 3. Alignment of amino acid sequences of candidate OBPs and CSPs from *E. obliqua*. (A) Amino acid alignment of the candidate OBPs. (B) Amino acid alignment of the candidate CSPs. Green boxes show the conserved cysteine residues. Accession numbers of the *E. obliqua* OBPs and CSPs are listed in Table 1.

doi:10.1371/journal.pone.0149591.g003

PLOS ONE

EoblOBPs were abundant in antennae (Fig 6). 16 of 24 total EoblOBPs (*OBP8-16*, *OBP18-19*, *OBP21* and *PBP1-4*) were uniquely or primarily expressed in the male and female antennae; while three EoblOBPs (*OBP3-4* and *OBP17*) were highly expressed in legs; *EoblOBP7* were

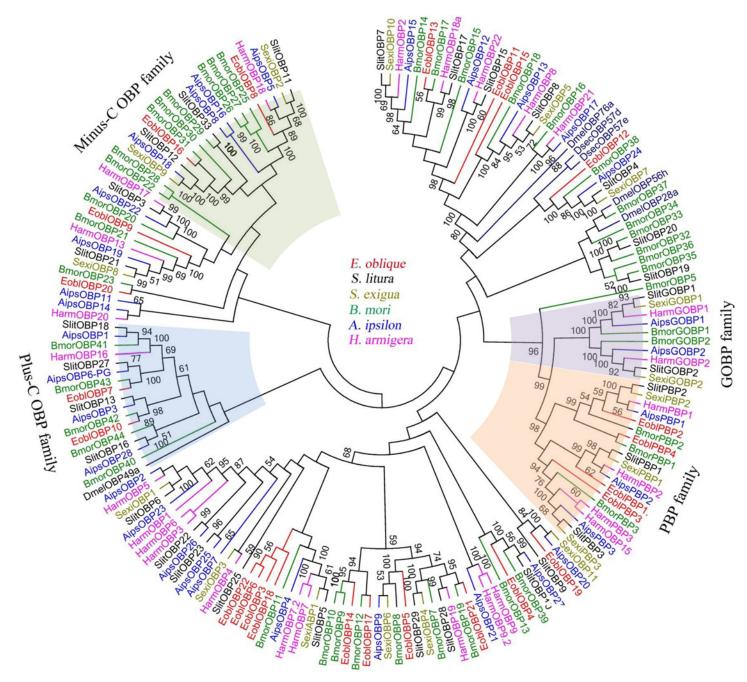


Fig 4. Neighbor-joining tree of candidate OBP proteins from Lepidoptera species. The protein names and sequences of the 178 OBPs used in this analysis are listed in <u>S3 Table</u>. Eobl, *E. obliqua*; Slit, *S. litura*; Sexi, *S. exigua*; Bmor, *B. mori*; Aips, *A. ipsilon*; Harm, *H. armigera*.

doi:10.1371/journal.pone.0149591.g004

primarily detected in the abdomen. However, the remaining 5 OBPs were distributed in varying tissues, among which *EoblOBP5*, *EoblOBP6* and *EoblOBP22* were primarily enriched in antennae and legs.

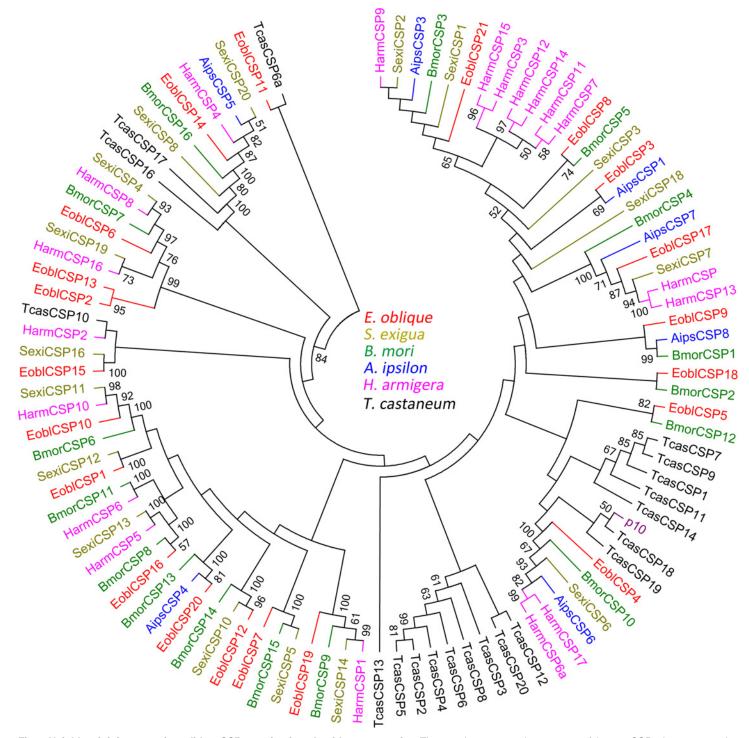


Fig 5. Neighbor-joining tree of candidate CSP proteins from Lepidoptera species. The protein names and sequences of the 101 CSPs that were used in this analysis are listed in <u>S4 Table</u>. Eobl, *E. obliqua*; Tcas, *T. castaneum*; Sexi, *S. exigua*; Bmor, *B. mori*; Aips, *A. ipsilon*; Harm, *H. armigera*. p10 is one CSP reported in the cockroach *Periplaneta Americana* (Kitabayashi et al., 1998).

doi:10.1371/journal.pone.0149591.g005

PLOS ONE

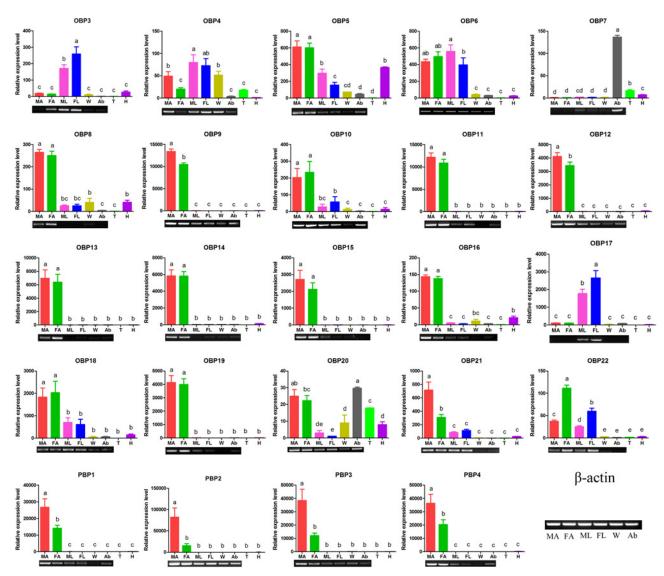


Fig 6. Tissue- and sex- specific expression profiles of *E. obliqua* OBP genes by qRT-PCR analysis and RT-PCR confirmation. FA: female antennae, MA: male antennae, FL: female legs, ML: male legs, H: heads, T: thoraxes, A: abdomens, L: legs, W: wings. In qPCR data were first normalized to endogenous β -actin levels from the same tissue, and the lowest-expression tissue was selected as the calibrator. The standard error is represented by the error bar, and the different letters above each bar represent significant differences (p<0.05). EobIOBPs expression of the former six tissues were confirmed by RT-PCR and arranged in the same order as that of qRT-PCR. β -actin was used as an internal reference gene to test the integrity of each cDNA template.

PLOS ONE

The expression pattern of 21 EoblCSPs showed diverse and wide expression (Fig_7). Six EoblCSPs (*CSP7*, *CSP8*, *CSP11*, *CSP15*, *CSP18* and *CSP21*) were dominantly expressed in legs, among which *EoblCSP11* and *EoblCSP15* were highly enriched in male legs. *EoblCSP2*, *EoblCSP10* and *EoblCSP16* were mostly distributed in abdomen, while *EoblCSP6* were uniquely expressed in antennae. The other EoblCSPs were ubiquitous in most tissues. In addition, *EoblCSP1*, *EoblCSP3*, *EoblCSP4*, *EoblCSP9*, *EoblCSP12*, *EoblCSP13*, *EoblCSP14*, *EoblCSP17*, *EoblCSP19* and *EoblCSP20* were abundant in legs at a relatively high level.

We also characterized the expression levels of ORs, GRs and SNMP in different tissues (Fig 8). The results indicated that the *EoblSNMP1* and *EoblSNMP2* were expressed significantly higher in antennae than in other tissues of both sexes. Four ORs were mainly expressed in the moth antennae. In addition, the transcript level of *EoblORco* exceeded 15000 fold changes

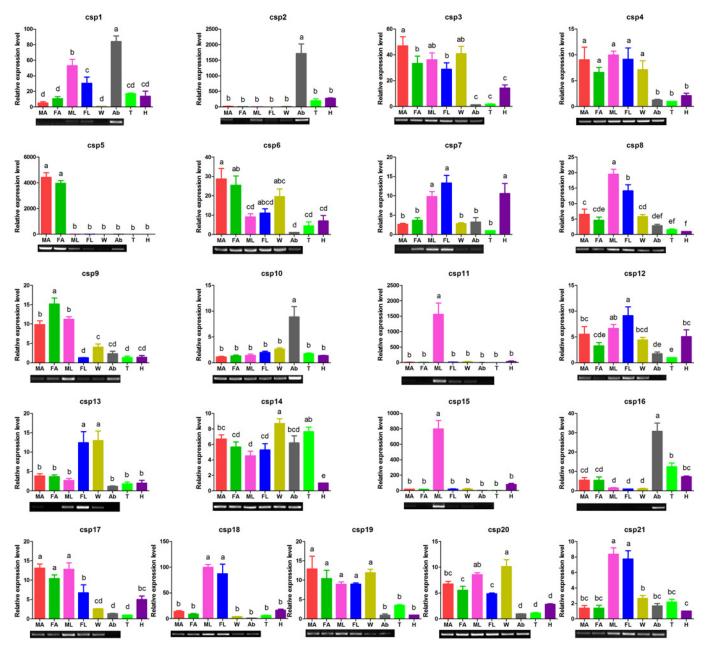


Fig 7. Tissue- and sex- specific expression profiles of *E. obliqua* CSPs genesc by qRT-PCR analysis and RT-PCR confirmation. The details were same as mentioned in Fig 6.

PLOS ONE

relative to the lowest expression in thoraxes. Among the three EoblGRs identified, *EoblGR1* were enriched in abdomen and *EoblGR2* had antennae-enriched expression, while *EoblGR3* were detected in both legs and wings.

Abundant analysis of chemosensory genes

RPKM metric was calculated to evaluate the comparative expression abundance in male and female legs. Of the 24 EoblOBPs, *EoblOBP6* showed the highest expression (7846.6 RPKM in male transcriptome), followed by *EoblOBP8* and *EoblOBP10*. Among the 21 EoblCSPs,



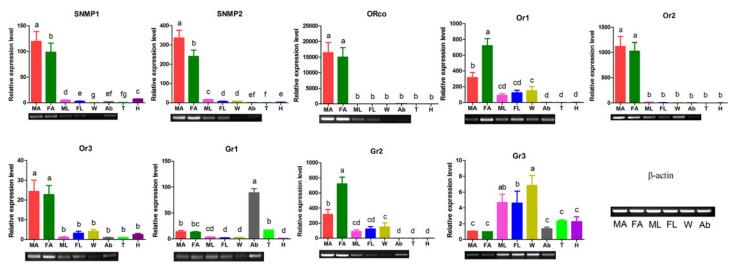


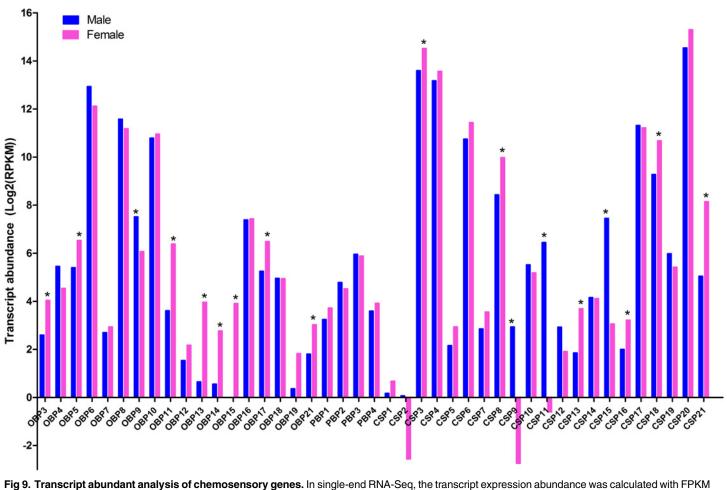
Fig 8. Tissue- and sex- specific expression profiles of *E. obliqua* chemoreceptor genes by qRT-PCR analysis and RT-PCR confirmation. The details were same as mentioned in Fig 6.

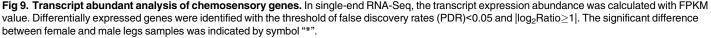
EoblCSP20 was the most abundant (40399.5 RPKM in female transcriptome), followed by *EoblCSP3*, *EoblCSP4*, *EoblCSP17* and *EoblCSP6* (Fig 9). Overall, the levels of expression of EoblOBPs in leg transcriptome were extremely variable, with RPKM values ranging from 1 to 7846; while the EoblCSP expressions were more diverse, from less than 10 to 40399. For chemosensory receptors, both ORs and GRs remained low expression in legs (<10 RPKM), of which *EoblGR2* had the highest expression in both sexes (3.81 RPKM in female and 5.07 RPKM in male). In addition, relative to other transmembrane proteins, *EoblSNMP2* showed an unexpectedly high expression (92.38 RPKM in female and 530.35 RPKM in male).

After summarizing the gene comparative expression between samples, a total of 1933 and 1985 up- and down-regulated unigenes, respectively, showed significantly altered expression (FDR \leq 0.05 and $|\log_2 \text{Ratio}| \geq 1$), as compared to the male transcriptome. The majority of the unigenes (91.6%), however, were expressed within a two-fold difference (Fig 10). For transporters, 9 EoblOBPs showed sex discrepancy in their levels of legs expression, 8 (*OBP3, OBP5, OBP11, OBP13, OBP14, OBP15, OBP17* and *OBP21*) being over expressed in female and only 1 (*OBP9*) over expressed in male; while 8 EoblCSPs presented sex differences in their expressions, 5 (*CSP3, CSP4, CSP13, CSP16* and *CSP18*) being more expressed in female and only 3 (*CSP9, CSP11* and *CSP 15*) up-regulated in male (Fig 9). Unexpectedly, relative to other chemosensory receptors, *EoblSNMP2* showed an abundant expression level and was 4.7-fold greater expressed in male legs. Go classification of the significantly regulated genes was implemented to identify the functional processes involved in sex differences (Fig 11). Overall, these regulated genes were mainly concentrated on cellular process, metabolic process, single-organism process, binding and catalytic activity. Besides, several subcategories were only involved in one-sex-regulation, such as reproduction, growth, multi-organism process and etc.

Discussion

Host plant selection by herbivorous insects is particularly important for reproduction, involved in searching, landing, contact evaluation, and final determination for oviposition [1]. Generally female adults avoid laying eggs on non-host plants in order to maximize the survival chances of their progenies. Monophagous herbivorous pests, such as *E. obliqua*, selectively utilize a limit of host plants, therefore requiring the specialized sensors to explore certain host for





PLOS ONE

oviposition. Olfaction and taste of insects are crucial in detecting and discriminating the chemical compounds from host. In spite that olfaction plays a primary role in perceiving plant volatile odorants from distance, taste is indispensable for non-volatile chemicals recognition after landing on the plant [4]. To ascertain host-plant identity, female butterflies and moths usually contact their forelegs on the leaves of host, which is the first perception of phytochemical compounds when landing on the surface. This initial contact presumably permits insects to taste phytochemical compounds [53]. Consistent with this action, butterflies including *P. xuthus*, *H.* melpomene and P. polytes possess groups of trichoid sensilla along with pairs of cuticular spines in female foretarsi, which get involved in the recognition of oviposition stimulants [12, 60, 61]; while 14 gustatory trichoid chemosensilla sensitive to either sugars or amino acids are found in prothoracic legs of moth H. armigera, M. privata and L. botrana [13, 15, 62]. In lepidopteran species, the tarsus is further divided into five tarsomeres, the fifth of which is the most distal part of the tarsus and bears more chemosensory sensilla than the other four tarsomeres. The arrangement of the gustatory sensilla in proximity to prominent tarsal spines is unique and could represent an adaptation which enables them to penetrate the wax layer and contact with metabolites present closer to the leaf surface [13]. Our microscopy of *E. obliqua* revealed the

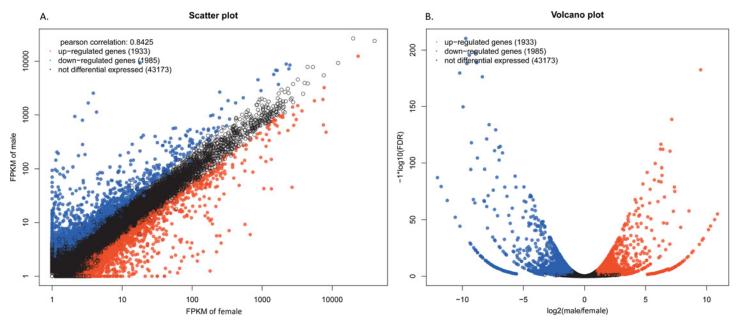


Fig 10. Analysis of differentially expressed genes exhibited in (A) Scatter plot and (B) Volcano plot. Genes are divided into three distinct subsets: blue genes represent the up-regulated ones in the female legs transcriptome vs. the male legs transcriptome, red genes are the down-regulated class compared in the same way, and black part represents the non-differentially expressed transcripts. Differentially expressed genes are identified according to statistically significant differences with the threshold of false discovery rates (PDR)<0.05 and $|log2Ratio \ge 1|$.

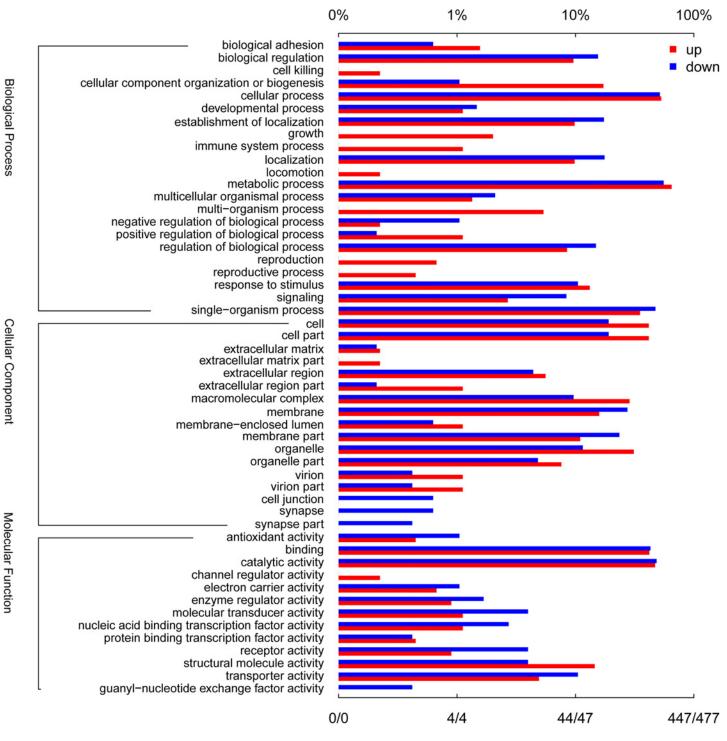
distribution of chemosensilla in the ventral side of a female fifth tarsomere (<u>Fig 1</u>), suggesting that the leg tarsi of *E. obliqua* were also responsible for taste detection.

Outside the limited butterfly species and model species whose genomes are available, rare studies are focused on gustatory system of other insect species. In fact, the remarkable selectivity and sensitivity of the chemosensory systems depend primarily on the performance of chemosensory neurons, which in turn rely ultimately on odorant receptors, gustatory receptors and selective transporters. So there is a special need to explore the candidate chemosensory genes. From our transcriptome analysis of *E. obliqua* legs, 24 OBPs (including 4 PBPs), 21 CSPs, 4 PBPs, 2 SNMP, 3 GRs and 4 ORs genes were identified.

Due to the low expression level of GR [17, 18], only three GR-encoding transcripts were identified from the legs transcriptome. Two of three EoblGRs were highly expressed in abdomen, among which *EoblGR2* shared 75.3% homology with *HarmGR4* that had been identified as a sugar receptor concentrated in larval foregut, female antennae and proleg tarsi of *H. armi-gera* [58, 59]. Thus, we can reasonably assume that *EoblGR2* is also a sugar receptor and could participate in the sugar detection and consumption. The abundance of *EoblGR2* in legs (the highest RPKM value among chemosensory receptors) is of great physiological significance, as most adult lepidopteran insects feed on floral nectar and honeydew, which contributes to female reproductive success [63]. Most ORs in insects are extensively distributed in antennae [64]. The tissue expression profiles of 4 EoblORs demonstrate the obviously antennal-abundance, however, these ORs are also distributed in other organs. The distribution of ORs in non-olfactory tissues suggests that they may participate in other physiological processes besides olfaction. For example, ORco expressed in the testes is involved in mediating activation of spermatozoa in *Anopheles gambiae* [65].

The majority of EoblOBPs (16 in total 24 OBPs) show relatively high expression in antenna, which corresponds to the commonly accepted concept that OBPs function as carriers of hydrophobic ligands to olfactory receptors in antenna [28], however, six EoblOBPs remain highly





Number of genes (Up/Down)

Fig 11. Gene Ontology (GO) enrichment analysis of all the differentially expressed genes. Horizontal axis in the top displays the percentage of significant genes in each column, while axis in the bottom is the number of significant genes. Vertical axis displays the detailed GO annotation corresponding to each functional type. Differentially expressed genes are compared in the manner of the female legs transcriptome vs. the male legs transcriptome.

doi:10.1371/journal.pone.0149591.g011

expressed or relative high in legs. A correlation of some OBPs reported (OBP49a, OBP57e and *OBP57d*) to host selection [35, 36, 66] and their unexpected distribution in taste organs, such as labellum, wing margins, tarsi, labial palps and etc. [67], raises the possibility that OBPs also participate in taste perception. In fact, non-volatile metabolites in plants are comparable to odorant in the way that they are both small poorly-water-soluble molecules, such as alkaloids and parts of terpenoids [68]. Previous studies have reported the binding of bitter compounds (berberine, denatonium and quinine) to OBP49a [35]. Consequently, it is reasonable to conclude that OBPs may act as transporters of hydrophobic compounds to gustatory receptors, which is similar to their performance in olfaction. RPKM metric facilitates the comparative study of expression between samples in mRNA-seq. Our comparative study revealed that 9 EoblOBPs showed sex differences in expression, 8 being up-regulated in female and only 1 being over expressed in male. Previous studies have reported the profound differences in the expression of OBP between sexes [69–71]. Considering that the adult history of male and female moths is guite similar in regard to the aim to fuel their body and the need to mate, the only difference is that females have to identify suitable host for oviposition. This sex difference may have ecological significance as females have to evaluate oviposition sites, so it stands to reason that OBPs with female-biased expression may participate in host selection, and that the female oviposition behavior drives the diversity of OBP expression between sexes.

The phylogenetic analysis reveals most EoblOBPs are clustered with different orthologous sequences from other species, suggesting that the Lepidoptera OBPs have differentiated into several different groups after a long time evolution. However, *EoblOBP3, EoblOBP6, EoblOBP18* and *EoblOBP22* share a high identity and are clustered in one branch, indicating recent gene duplication events. Besides, *EoblOBP21* shares 37.6% identity and similar expression profile with *HarmOBP10*, which was previously reported to bind one insect repellent 1-dodecene [72].

CSPs are soluble proteins and believed to play a role which is similar to that of OBPs in the perception of odorants [44–47]. The relative expression patterns of 21 EoblCSPs are diverse and widely distributed. Apart from *EoblCSP5* specially expressed in antenna, three CSPs (EoblCSP2, EoblCSP10 and EoblCSP16) are primarily present in abdomen, where they might transport semiochemicals in reproductive organs or sex glands, assisting their release into the environment [44, 73, 74]. Fortunately, HarmCSP6, sharing 45.2% homology and closely clustered with EoblCSP16, was reported to be highly transcribed in pheromone glands and display high binding affinity for pheromone components [75]. In addition, six EoblCSPs are dominantly expressed in legs, besides, 10 EoblCSPs are abundant in legs at the relatively high level. Among them, EoblCSP21 shared 59.4% identity with HarmCSP4 which was detected to be exclusively present in proboscis and could help solubilizing terpenoids present in flower nectar [76]; EoblCSP4 is exceptionally abundant in legs (9274 RPKM in male legs and 12169 RPKM in female legs), sharing 41.7% homology and closely clustered with Pamep10 which seemed to be involved in limb regeneration [77]. To our surprise, these functions mentioned above are completely unrelated to chemical communication. Actually, the compact structure of CSPs, their soluble nature and flexible polypeptide folding, permit this protein to bind a variety of ligands and therefore could undertake several tasks in the biological process [78].

In summary, a large number of chemosensory genes were identified in *E. obliqua*, and tissue distribution profiles were investigated. Several leg-specific or enriched genes were screened, and clustered with functionally validated genes from other moths, suggesting potential involvement in taste sensation or other physiological processes. The female-biased EoblOBPs indicated an ecological adaption related with host-seeking and oviposition behaviors. Our studies will provide the basic knowledge for further research on the molecular mechanism of gustatory perception, and enlighten a host-selection-based control strategy of insect pests.

Supporting Information

S1 Fig. (TIF) **S1 Table.** (DOC) **S2 Table.** (DOCX) **S3 Table.** (DOCX) **S4 Table.** (DOCX)

(DOCA)

Acknowledgments

We are grateful to Jianying Jin, Kunshan Yin and Meijun Tang for supplying the insects. This work was supported by Modern Agricultural Industry Technology System (CARS-23), Open Foundation of State Key laboratory for Biology of Plant Diseases and Insect Pests (SKLOF201517) and Zhejiang Provincial Natural Science Foundation (LY16C140003).

Author Contributions

Conceived and designed the experiments: LM ZMC YJZ. Performed the experiments: LM. Analyzed the data: LM. Contributed reagents/materials/analysis tools: LM LB XMC ZXL. Wrote the paper: LM ZMC ZQL. Polished the language of the manuscript: ZMC ZQL.

References

- Schoonhoven LM, Van Loon JJA, Dicke M. Insect-Plant Biology, 2nd ed. Oxford, UK: Oxford University Press; 2005.
- 2. Schneider D. Insect antennae. Annu Rev Entomol. 1964; 9: 103–122.
- Schoonhoven LM, Dethier VM. Sensory aspects of host-plant discrimination by lepidopterous larvae. Arch Neerl Zool. 1966; 16: 497–530.
- 4. Chun MW, Schoonhoven LM. Tarsal contact chemosensory hairs of the large white butterfly *Pieris* brassicae and their possible role in oviposition behaviour. Entom Exp Applicate. 1973; 16: 343–357.
- Chapman RF. Contact chemoreception in feeding by phytophagous insects. Annu Rev Entomol. 2003; 48 (1): 455–484.
- 6. Klijnstra J, Roessingh P. Perception of the oviposition deterring pheromone by tarsal and abdominal contact chemoreceptors in *Pieris brassicae*. Entom Exp Applicate. 1986; 40: 71–79.
- 7. Qiu YT, van Loon JJA, Roessingh P. Chemoreception of oviposition inhibiting terpenoids in the diamondback moth *Plutella xylostella*. Entomol Exp Appl. 1998; 87: 143–155.
- Maher N, Thiery D, Stadler E. Oviposition by *Lobesia botranais* stimulated by sugars detected by contact chemoreceptors. Physiol Entomol. 2006; 31 (1): 14–22.
- 9. Dethier VG. The hungry fly. Cambridge, MA: Harvard UP; 1976.
- Carter M, Sachdev-Gupta K, Feeny P. Tyramine from the leaves of wild parsnip: a stimulant and synergist for oviposition by the black swallowtail butterfly. Physiol Entomol. 1998; 23 (4): 303–312.
- 11. Nakayama T, Honda K, Omura H, Hayashi N. Oviposition stimulants for the tropical swallowtail butterfly, *Papilio polytes*, feeding on a rutaceous plant, *Toddalia asiatica*. J Chem Ecol. 2003; 29 (7): 1621– 1634. PMID: 12921440
- Briscoe AD, Macias-Muñoz A, Kozak KM, Walters JR, Yuan F, Jamie GA, et al. Female behaviour drives expression and evolution of gustatory receptors in butterflies. PLoS Genet. 2013; 9(7): e1003620. doi: <u>10.1371/journal.pgen.1003620</u> PMID: <u>23950722</u>

- Calas D, Marion-Poll F, Steinbauer MJ. Tarsal taste sensilla of the autumn gum moth, *Mnesampela privata*: morphology and electrophysiological activity. Entomol. Exp. Appl. 2009; 133(2): 186–192.
- Blaney W. Simmonds M. A behavioural and electrophysiological study of the role of tarsal chemoreceptors in feeding by adults of *Spodoptera*, *Heliothis virescens* and *Helicoverpa armigera*. J Insect Physiol. 1990; 36(10): 743–756.
- Zhang YF, van Loon JJA, Wang CZ. Tarsal taste neuron activity and proboscis extension reflex in response to sugars and amino acids in *Helicoverpa armigera* (Hubner). J Exp Biol. 2010; 213(16): 2889–2895.
- Ling F, Dahanukar A, Weiss LA, Kwon JY, Carlson JR. The molecular and cellular basis of taste coding in the legs of *Drosophila*. J Neurosci. 2014; 34(21): 7148–7164. doi: <u>10.1523/JNEUROSCI.0649-14</u>. <u>2014</u> PMID: <u>24849350</u>
- Clyne PJ, Warr CG, Carlson JR. Candidate taste receptors in *Drosophila*. Science. 2000; 287 (5459): 1830–1834. PMID: <u>10710312</u>
- Scott K, Brady R, Cravchik A, Morozov P, Rzhetsky A, Zuker C, et al. A chemosensory gene family encoding candidate gustatory and olfactory receptors in *Drosophila*. Cell. 2001; 104 (5): 661–673. PMID: <u>11257221</u>
- Dunipace L, Meister S, McNealy C, Amrein H. Spatially restricted expression of candidate taste receptors in the *Drosophila* gustatory system. Curr Biol. 2001; 11 (11): 822–835. PMID: <u>11516643</u>
- Dahanukar A, Foster K, Carlson JR. A Gr receptor is required for response to the sugar trehalose in taste neurons of *Drosophila*. Nat Neurosci. 2001; 4 (12): 1182–1186. PMID: <u>11704765</u>
- Miyamoto T, Chen Y, Slone J, Amrein H. Identification of a *Drosophila* glucose receptor using Ca²⁺ imaging of single chemosensory neurons. PLoS One. 2013; 8: e56304. doi: <u>10.1371/journal.pone.</u> <u>0056304</u> PMID: <u>23418550</u>
- Jiao Y, Moon SJ, Montell C. A *Drosophila* gustatory receptor required for the responses to sucrose, glucose, and maltose identified by mRNA tagging. Proc Natl Acad Sci USA. 2007; 104: 14110–14115. PMID: 17715294
- Lee Y, Moon SJ, Montell C. Multiple gustatory receptors required for the caffeine response in Drosophila. Proc Natl Acad Sci USA. 2009; 106: 4495–4500. doi: 10.1073/pnas.0811744106 PMID: 19246397
- Moon SJ, Kottgen M, Jiao Y, Xu H, Montell C. A taste receptor required for the caffeine response in vivo. Curr Biol. 2006; 16: 1812–1817. PMID: <u>16979558</u>
- Moon SJ, Lee Y, Jiao Y, Montell C. A *Drosophila* gustatory receptor essential for aversive taste and inhibiting male-to-male courtship. Curr Biol. 2009; 19: 1623–1627. doi: <u>10.1016/j.cub.2009.07.061</u> PMID: <u>19765987</u>
- Jones WD, Cayirlioglu P, Kadow IG, Vosshall LB. Two chemosensory receptors together mediate carbon dioxide detection in *Drosophila*. Nature. 2007; 445:86–90. PMID: 17167414
- Kwon JY, Dahanukar A, Weiss LA, Carlson JR. The molecular basis of CO2 reception in *Drosophila*. Proc Natl Acad Sci USA. 2007; 104: 3574–3578. PMID: <u>17360684</u>
- Leal WS. Odorant reception in insects: roles of receptors, binding proteins, and degrading enzymes. Annu Rev Entomol. 2013; 58: 373–391. doi: 10.1146/annurev-ento-120811-153635 PMID: 23020622
- Robertson HM, Martos R, Sears CR, Todres EZ, Walden KKO, Nardi JB. Diversity of odourant binding proteins revealed by an expressed sequence tag project on male *Manduca sexta* moth antennae. Insect Mol Biol. 1999; 8: 501–518. PMID: 10620045
- **30.** Blomquist GJ, Vogt RG. Insect pheromone biochemistry and molecular biology: The biosynthesis and detection of pheromones and plant volatiles. London: Elsevier Academic Press; 2003: 391–445.
- McKenna MP, Hekmat-Scafe DS, Gaines P, Carlson JR. Putative *Drosophila* pheromone-binding proteins expressed in a subregion of the olfactory system. J Biol Chem. 1994; 269: 16340–16347. PMID: <u>8206941</u>
- Pikielny CW, Hasan G, Rouyer F, Rosbash M. Members of a family of *Drosophila* putative odorant-binding proteins are expressed in different subsets of olfactory hairs. Neuron. 1994; 12: 35–49. PMID: 7545907
- Ozaki M, Morisaki K, Idei W, Ozaki K, Tokunaga F. A putative lipophilic stimulant carrier protein commonly found in the taste and olfactory systems. Eur J Biochem. 1995; 230: 298–308. PMID: <u>7601113</u>
- Yasukawa J, Tomioka S, Aigaki T, Matsuo T. Evolution of expression patterns of two odorant-binding protein genes, Obp57d and Obp57e, in *Drosophila*. Gene. 2010; 467: 25–34. doi: <u>10.1016/j.gene.</u> <u>2010.07.006</u> PMID: <u>20637846</u>
- Jeong YT, Shim J, Oh SR, Yoon HI, Kim CH, Moon SJ, et al. An odorant-binding protein required for suppression of sweet taste by bitter chemicals. Neuron. 2013; 79: 725–737. doi: <u>10.1016/j.neuron.</u> <u>2013.06.025</u> PMID: <u>23972598</u>

- **36.** Matsuo T, Sugaya S, Yasukawa J, Aigaki T, Fuyama Y. Odorant-binding proteins OBP57d and OBP57e affect taste perception and host-plant preference in *Drosophila sechellia*. PLoS Biol. 2007; 5: e118. PMID: <u>17456006</u>
- Gong D P, Zhang H, Zhao P, Lin Y, Xia QY, Xiang ZH. Identification and expression pattern of the chemosensory protein gene family in the silkworm, *Bombyx mori*. Insect Biochem Mol Biol. 2007; 37: 266– 277. PMID: <u>17296501</u>
- Angeli S, Ceron F, Scaloni A. Monti M, Monteforti G, Minnocci A, et al. Purification, structural characterization, cloning and immunocytochemical localization of chemoreception proteins from Schistocerca gregaria. Eur J Biochem. 1999; 262: 745–754. PMID: 10411636
- Jacobs SP, Liggins AP, Zhou JJ, Pickett JA, Jin X, Field LM. OS-D-like genes and their expression in aphids (Hemiptera: Aphididae). Insect Mol Biol. 2005; 14: 423–432. PMID: <u>16033435</u>
- Maleszka R, Stange G. Molecular cloning, by a novel approach, of a cDNA encoding a putative olfactory protein in the labial palps of the moth *Cactoblastis cactorum*. Gene. 1997; 202: 39–43. PMID: 9427543
- Jin X, Brandazza A, Navarrini A, Ban L, Zhang S, Steinbrecht RA, et al. Expression and immunolocalisation of odorant-binding and chemosensory proteins in locusts. Cell Mol Life Sci. 2005; 62: 1156– 1166. PMID: <u>15928808</u>
- Ban L, Zhang L, Yan Y, Pelosi P. Binding properties of a locust's chemosensory protein. Biochem Biophys Res Commun. 2002; 293: 50–54. PMID: <u>12054562</u>
- Zhou SH, Zhang J, Zhang SG, Zhang L. Expression of chemosensory proteins in hairs on wings of Locusta migratoria (Orthoptera: Acrididae). J Appl Entomol. 2008; 132: 439–450.
- Jacquin-Joly E, Vogt RG, François MC, Meillour PNL. Functional and expression pattern analysis of chemosensory proteins expressed in antennae and pheromonal gland of *Mamestra brassicae*. Chem senses. 2001; 26: 833–844. PMID: <u>11555479</u>
- Monteforti G, Angeli S, Petacchi R, Minnocci A. Ultrastructural characterization of antennal sensilla and immunocytochemical localization of a chemosensory protein in *Carausius morosus* Brünner (Phasmida: Phasmatidae). Arthropod Struct Dev. 2002; 30: 195–205. PMID: <u>18088955</u>
- Calvello M, Guerra N, Brandazza A, D'Ambrosio C, Scaloni A, Dani FR, et al. Soluble proteins of chemical communication in the social wasp *Polistes dominulus*. Cell Mol Life Sci. 2003; 60: 1933–1943. PMID: 14523553
- Ozaki M, Wada-Katsumata A, Fujikawa K, Iwasaki M, Yokohari F, Satoji Y, et al. Ant nestmate and nonnestmate discrimination by a chemosensory sensillum. Science. 2005; 309: 311–314. PMID: 15947139
- Benton R, Vannice KS, Vosshall LB, An essential role for a CD36-related receptor in pheromone detection in *Drosophila*. Nature. 2007; 450: 289–293. PMID: 17943085
- Vogt RG, Miller NE, Litvack R, Fandino RA, Sparks J, Staples J, et al. The insect SNMP gene family. Insect Biochem Mol Biol. 2009; 39: 448–456. doi: <u>10.1016/j.ibmb.2009.03.007</u> PMID: <u>19364529</u>
- 50. Gu SH, Yang RN, Guo MB, Wang GR, Wu KM, Guo YY, et al. Molecular identification and differential expression of sensory neuron membrane proteins in the antennae of the black cutworm moth Agrotis ipsilon. J Insect Physiol. 2013; 59: 430–443. doi: 10.1016/j.jinsphys.2013.02.003 PMID: 23454276
- Zhang ZQ, Sun XL, Luo ZX, Gao Y, Bian L and Chen ZM, Effect of odors from different aromatic plants and extracts on the behavior of the tea geometrid, *Ectropis obliqua* (Prout). Acta Phytophys Sinica. 2012; 39: 541–548.
- Ryuda M, Calas-List D, Yamada A, Marion-Poll F, Yoshikawa H, Tanimura T, et al. Gustatory sensing mechanism coding for multiple oviposition stimulants in the swallowtail butterfly, *Papilio xuthus*. J Neurosci. 2013; 33: 914–924. doi: 10.1523/JNEUROSCI.1405-12.2013 PMID: 23325231
- Feeny P, Rosenberry L, Carter M. Chemical aspects of oviposition behavior in butterflies. In: Herbivorous insects: host-seeking behavior and mechanisms (Ahmad S, edited). New York; 1983: 26–76.
- 54. Sun XL, Wang GC, Gao Y, Zhang XZ, Xin ZJ, Chen, ZM. Volatiles emitted from tea plants infested by *Ectropis obliqua* larvae are attractive to conspecific moths. J Chem Ecol. 2014; 40: 1080–1089. doi: <u>10.1007/s10886-014-0502-5</u> PMID: <u>25378120</u>
- 55. Miller TA, Highfill JW, Cooper WJ. Relationships between pupal size and sex in giant silkworm moths (Saturniidae). J Lepid Soc. 1982; 36: 207–216.
- Mortazavi A, Williams BA, McCue K, Schaeffer L, Wold B. Mapping and quantifying mammalian transcriptomes by RNA-Seq. Nat Methods. 2008; 5: 621–628. doi: <u>10.1038/nmeth.1226</u> PMID: <u>18516045</u>
- 57. Pfaffl MW. A new mathematical model for relative quantification in real-time RT-PCR. Nucleic Acids Res. 2001; 29: e45–e45. PMID: <u>11328886</u>

- Xu W, Zhang H J, Anderson A. A sugar gustatory receptor identified from the foregut of cotton bollworm Helicoverpa armigera. J Chem Ecol. 2012; 38: 1513–1520. doi: <u>10.1007/s10886-012-0221-8</u> PMID: <u>23224441</u>
- Jiang XJ, Ning C, Guo H, Jia YY, Huang LQ, Qu MJ, et al. A gustatory receptor tuned to D-fructose in antennal sensilla chaetica of *Helicoverpa armigera*. Insect Biochem Mol Biol. 2015; 60: 39–46. doi: <u>10.</u> <u>1016/j.ibmb.2015.03.002</u> PMID: <u>25784630</u>
- **60.** Renou M, Les re cepteurs gustatifs du tarse ante rieur de la femelle d'*Heliconius charitonius* (Lep.: Heliconiidae). Annls Soc Ent Fr (NS). 1983; 19: 101–106.
- Ozaki K, Ryuda M. Ozaki K, Ryuda M, Yamada A, Utoguchi A, et al. A gustatory receptor involved in host plant recognition for oviposition of a swallowtail butterfly. Nat Commun. 2011; 2: 542. doi: <u>10.</u> <u>1038/ncomms1548</u> PMID: <u>22086342</u>
- Calas D, Thiéry D, Marion-Poll F. 20-Hydroxyecdysone deters oviposition and larval feeding in the European grapevine moth, *Lobesia botrana*. J Chem Ecol. 2006; 32: 2443–2454. PMID: <u>17082989</u>
- Bauerfeind S. and Fischer K. Effects of adult-derived carbohydrates, amino acids and micronutrients on female reproduction in a fruit-feeding butterfly. J Insect Physiol. 2005; 51: 545–554. PMID: <u>15894001</u>
- Vosshall LB, Amrein H, Morozov PS, Rzhetsky A, Axel R. A spatial map of olfactory receptor expression in the *Drosophila* antenna. Cell. 1999; 96: 725–736. PMID: <u>10089887</u>
- Pitts RJ, Liu C, Zhou X, Malpartida JC, Zwiebel LJ. Odorant receptor-mediated sperm activation in disease vector mosquitoes. Proc Natl Acad Sci USA. 2014; 111: 2566–2571. doi: <u>10.1073/pnas.</u> <u>1322923111</u> PMID: 24550284
- 66. Matsuo T. Genes for host-plant selection in Drosophila. J Neurogen. 2008; 22: 195–210.
- Galindo K, Smith DP. A large family of divergent Drosophila odorant-binding proteins expressed in gustatory and olfactory sensilla. Genetics. 2001; 159: 1059–1072. PMID: <u>11729153</u>
- Swarup S, Morozova TV, Sridhar S, Nokes M, Anholt RR. Modulation of feeding behavior by odorantbinding proteins in *Drosophila melanogaster*. Chem Senses. 2014; 39: 125–132. doi: <u>10.1093/</u> chemse/bjt061 PMID: 24302688
- Zhou S, Stone EA, Mackay TFC, Anholt RRH. Plasticity of the chemoreceptor repertoire in *Drosophila* melanogaster. PLoS Genet. 2009; 5: e1000681. doi: <u>10.1371/journal.pgen.1000681</u> PMID: <u>19816562</u>
- 70. Gu SH, Sun L. Yang RN, Wu KM, Guo YY, Li XC, et al. Molecular characterization and differential expression of olfactory genes in the antennae of the black cutworm moth *Agrotis ipsilon*. PLoS One. 2014; 9: e103420. doi: 10.1371/journal.pone.0103420 PMID: 25083706
- Gu SH, Zhou JJ, Gao S, Wang DH, Li XC, Guo YY, Zhang YJ. Identification and comparative expression analysis of odorant binding protein genes in the tobacco cutworm *Spodoptera litura*. Sci Rep. 2015; 5: 13800. doi: <u>10.1038/srep13800</u> PMID: <u>26346731</u>
- Sun YL, Huang LQ, Pelosi P, Wang CZ. Expression in antennae and reproductive organs suggests a dual role of an odorant-binding protein in two sibling *Helicoverpa* species. PLoS One. 2012; 7: e30040. doi: <u>10.1371/journal.pone.0030040</u> PMID: <u>22291900</u>
- Iovinella I, Bozza F, Caputo B, della Torre A. Pelosi P. Ligand-binding study of Anopheles gambiae chemosensory proteins. Chem Senses. 2013; 38: 409–419. doi: <u>10.1093/chemse/bjt012</u> PMID: <u>23599217</u>
- Zhou XH, Ban LP, Iovinella I, Zhao LJ, Gao Q, Felicioli A, et al. Diversity, abundance, and sex-specific expression of chemosensory proteins in the reproductive organs of the locust *Locusta migratoria manilensis*. Biol Chem. 2013; 394: 43–54. doi: 10.1515/hsz-2012-0114 PMID: 23096575
- Li ZQ, Zhang S, Luo JY, Zhu J, Cui JJ, Dong SL. Expression Analysis and Binding Assays in the Chemosensory Protein Gene Family Indicate Multiple Roles in *Helicoverpa armigera*. J Chem Ecol. 2015; 41(5): 473–485. doi: <u>10.1007/s10886-015-0574-x</u> PMID: <u>25893790</u>
- Liu YL, Guo H, Huang LQ, Pelosi P, Wang CZ. Unique function of a chemosensory protein in the proboscis of two *Helicoverpa* species. J Exp Biol. 2014; 217(10): 1821–1826.
- 77. Kitabayashi AN, Arai T, Kubo T, Natori S. Molecular cloning of cDNA for p10, a novel protein that increases in the regenerating legs of *Periplaneta americana* (American cockroach). Insect Bioch Mol Biol. 1998; 28(10): 785–790.
- Lartigue A, Campanacci V, Roussel A, Larsson AM, Jones TA, Tegoni M, et al. X-ray structure and ligand binding study of a moth chemosensory protein. J Biol Chem. 2002; 277(35): 32094–32098. PMID: 12068017