



OPEN

Identification and comparative expression analysis of odorant-binding proteins in the reproductive system and antennae of *Athetis dissimilis*

Yue-Qin Song, Zhi-Yu Song, Jun-Feng Dong, Qi-Hui Lv, Qing-Xiao Chen & Hui-Zhong Sun[✉]

Odorant-binding proteins (OBPs) are prevalent in the antennal transcriptomes of different orders of insects. Studies on OBPs have focused on their role in the insect chemosensory system, but knowledge of their functions in the insect testis is limited. We sequenced the transcriptomes of the *Athetis dissimilis* reproductive organs and analyzed the expression of *AdisOBP* genes in different tissues. We identified 23 OBPs in the testis and ovaries and 31 OBPs in antennal transcriptomes. The results of real-time quantitative PCR revealed that 23 of the 54 OBP genes were highly expressed in both female and male antennae, including three that exhibited male-biased expression and 15 that exhibited female-biased expression. A total of 24 OBPs were highly expressed in the testis of *A. dissimilis*, while expression of OBPs in the ovaries was very low. These findings highlight the functional diversity of OBPs in insects and can facilitate further studies on the OBPs in *A. dissimilis* and lepidopteran species.

The olfactory system in insects regulates their intersex communication, host-plant interactions, oviposition, foraging, escape from predators and reproduction^{1–5}. Insects have a complex chemosensory system in which pheromones and plant odors are initially recognized by odorant-binding proteins (OBPs) expressed in the antennal sensilla lymph that transfer the odorants to membrane-bound olfactory receptors (ORs) to activate olfactory receptor neurons (ORNs) and stimulate behavioral responses^{6–11}.

Understanding the molecular mechanisms of olfaction is essential for better using olfactory-based pest management strategies and the development of novel strategies. OBPs are more accessible targets for research, considering they are small, soluble, stable and easier to manipulate and modify. OBPs are small water soluble proteins that have six positionally conserved cysteines to form three interlocking disulphide bridges that stabilize the protein's three-dimensional structure^{12–19}. OBPs were first discovered in the antenna of *Antheraea polyphemus*, where they distinguish and bind to lipophilic odorant compounds^{20–25}. However, emerging data suggests that OBPs are not restricted to the sensory organs of insect and show expression in non-sensory organs including reproductive organs^{26,27}. Li et al. showed that *AaegOBP22* was highly expressed in the male reproductive organs of *Aedes aegypti* and transfers to females during mating. This suggests an additional function for this protein as pheromone carrier, analogously to vertebrates' urinary and salivary proteins as well as some insect chemosensory proteins²⁶. Sun et al. also found that *HarmOBP10* and *HassOBP10* is highly abundant in seminal fluid of *Helicoverpa armigera* and *H. assulta* and transfers to female during mating. *HarmOBP10* and *HassOBP10* also bind 1-dodecene, a known insect repellent²⁷.

Athetis dissimilis Hampson (Lepidoptera: Noctuidae) is an important agricultural pest and mainly distributed in Asian countries including China, Japan, Philippines, Korea, Indonesia and India, causing serious damages to maize, wheat, peanut, soybean and sweet potato^{28–30}. Because of the fact that larvae of *A. dissimilis* live under plant residues, it is difficult to control the spread of the pest with chemical pesticides. Therefore, novel control managements are urgently needed to mitigate crop damage. We first sequenced the antennal transcriptomes of *A. dissimilis*³¹ and characterized 5 OBPs that showed tissue-specific expression patterns³². Of note, *AdisOBP6* was highly expressed in the testes of *A. dissimilis*³². We reasoned that the testis of insects possess a defined set of OBPs in a manner comparable to the antenna. In this study, we reanalyzed the previous antennal transcriptome

College of Horticulture and Plant Protection, Henan University of Science and Technology, Luoyang 471000, China. ✉email: huizhong66@163.com

Sample name	Clean reads	Clean bases	GC content (%)	Q30 (%)
Ovaries				
Repeat 1	34,565,866	10.35 G	48.00	94.03
Repeat 2	32,154,799	9.63 G	48.35	94.36
Repeat 3	26,952,526	8.07 G	48.27	94.21
Testis				
Repeat 1	27,752,168	8.29 G	48.85	94.42
Repeat 2	28,900,040	8.65 G	47.20	94.27
Repeat 3	30,838,686	9.23 G	46.65	94.01

Table 1. Summary of the sequence assemblies according to the RNA-seq data of the *A. dissimilis*.

Length range	Transcript	Rate%	Unigene	Rate%
< 300	0	0	0	0
300–500	83,670	37.85	37,104	45.24
500–1000	70,088	31.70	24,792	30.23
1000–2000	44,935	20.33	12,864	15.68
> 2000	22,381	10.12	7256	8.85
Total number	221,074		82,016	
Total length	216,261,287		73,549,396	
N50 length	1350		1243	
Mean length	978.23		896.77	

Table 2. Summary of de novo assembly of the *A. dissimilis* transcriptomes.

data and identified 31 OBP genes. We also sequenced the transcriptomes of the *A. dissimilis* reproductive organs, and studied the expression of the OBPs in the antennae, testis and ovaries. Our study provides a new reference for studying the function of OBP genes.

Results

Illumina sequencing and assembly. A total of 34,565,866, 32,154,799, and 26,952,526 clean reads containing 10.35, 9.63, and 8.07 giga base (Gb) pairs of clean nucleotides respectively, were obtained from the three replicates of the *A. dissimilis* ovaries. A total of 27,752,168, 28,900,040, and 30,838,686 clean reads containing 8.29, 8.65 and 9.23 giga base (Gb) pairs of clean nucleotides respectively, were obtained from the three replicates of *A. dissimilis* testes. The quality of the transcriptome sequences was high, with Q30 percentages of 94.03%, 94.36%, 94.21%, 94.42%, 94.27% and 94.01% for the three replicates of *A. dissimilis* ovaries and testes, with a GC content of ~ 50% (Table 1). Then 221,074 transcripts and 82,016 unigenes with N50 length of 1350 and 1243 were obtained from assembled using Trinity (Table 2).

Functional annotation. Significant matches of 33,587 unigenes (96.91%) in the NR; 29,936 (86.38%) in the eggnog; 20,134 (58.09%) in the Pfam; 15,174 (43.78%) in the Swissprot database; 14,775 (42.63%) in the KEGG; 7797 (22.50%) in the GO; and 6712 (19.37%) in the COG were observed. As a result, up to 34,658 putative coding sequences were identified (Table 3). NR database queries revealed a high percentage of *A. dissimilis* sequences that closely matched to sequences of *H. armigera* (19,072, 56.87%), *Amyelois transitella* (1936, 5.77%), *Bombyx mori* (1543, 4.60%), *Papilio machaon* (1155, 3.44%), *Papilio xuthus* (868, 2.59%), *Plutella xylostella* (844, 2.52%), *Danaus plexippus* (634, 1.89%), *Branchiostoma belcheri* (473, 1.41%), and *Papilio polytes* (368, 1.10%) (Fig. 1).

For GO analysis, 7797 unigenes (22.50%) could be assigned to three GO terms including: cellular components, molecular functions and biological process (Fig. 2). For the “molecular functions” ontology, catalytic activity (4227, 42.19%) and binding (3972, 39.64%) were most prevalent.

Identification of putative odorant-binding proteins. In the *A. dissimilis* antennal and reproductive organ transcriptome, we identified 54 candidate OBPs (Genbank accession number: KR780027–KR780030, MH900289–MH900338), 31 of which were from the antennae (through the analysis of previous *A. dissimilis* antennal transcriptomes) and 23 from the testis and ovaries transcriptomes of *A. dissimilis* (Table 4). A total of 44 *Adis*OBPs sequences had full-length ORFs. Their cDNAs encoded protein of 131–293 amino acids with molecular weights of 11.6–33.2 kDa and isoelectric points of 4.44–9.74. Excluding 7 *Adis*OBPs (*Adis*OBP28, 30, 31, 35, 36, 41, 42, 52, 53 and 54) signal peptides were predicted at the N-terminus. *Adis*OBPs had 39–99% sequence homology with previously identified OBPs from other insect species, displaying a high level of sequence similar-

Database	Number	Rate (%)	$300 \leq \text{length} < 1000$	Length ≥ 1000
COG	6712	19.37	2638	4074
GO	7797	22.50	4453	3344
KEGG	14,775	42.63	8205	6570
Pfam	20,134	58.09	8577	11,557
Swissprot	15,174	43.78	6987	8187
eggNOG	29,936	86.38	16,283	13,653
NR	33,587	96.91	18,939	14,648
All	34,658		19,914	14,744

Table 3. Functional annotation of the *A. dissimilis* transcriptomes.

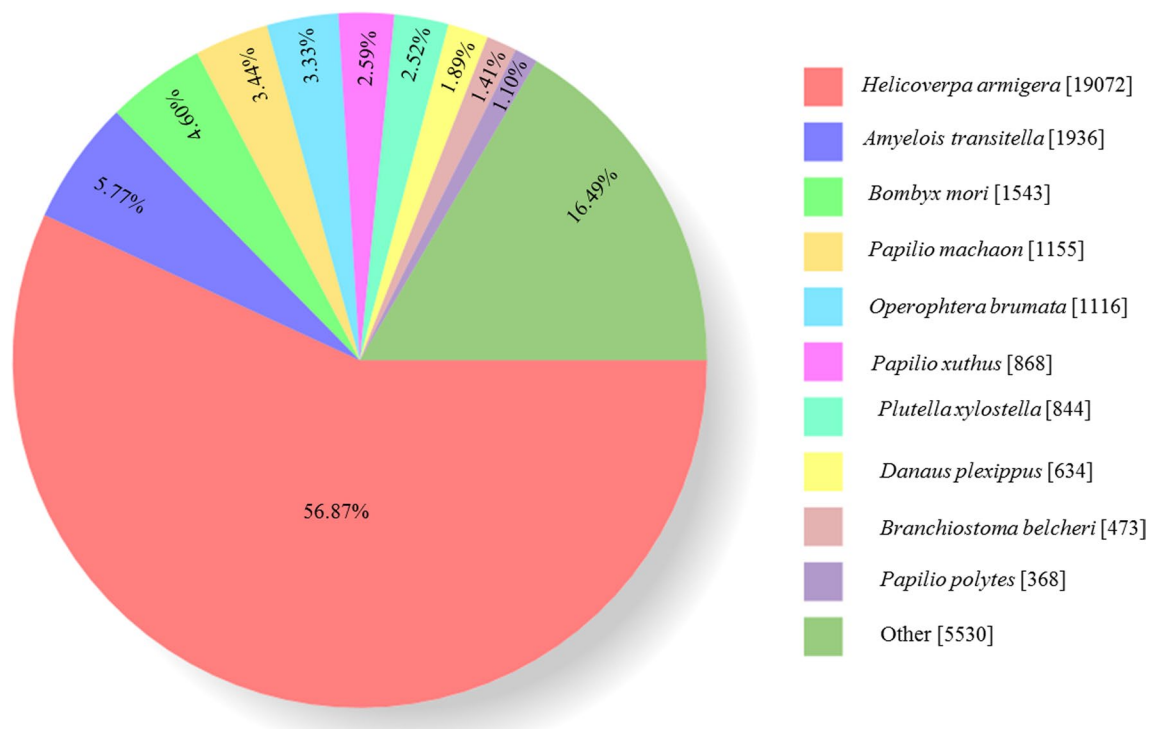


Figure 1. The Blastx results of *Athetis dissimilis* reproductive organs unigenes in NR database.

ity. For example, *AdisOBP13* had a 95% identity with *Spodoptera exigua* OBP9 (Table 4). There was 11.87% the lowest identity level in a pairwise comparison of *AdisOBPs*.

Multiple sequence alignments of the *A. dissimilis* OBPs revealed the presence of expected conserved cysteines (Fig. 3). The phylogenetic tree of *A. dissimilis* and other lepidopteran OBPs constructed using the neighbor-joining method, indicated five clades that contained four possible subclass OBPs (Fig. 4). In addition, the tree showed low levels of clustering highlighting the diversity of the lepidopteran OBPs. Five *AdisOBPs* (*AdisPBP1-3*, *GOBP1-2*) belonged to PBP/GOBP. A total of 30 OBPs (*AdisOBP2-3*, 9, 11, 20–24, 26–32, 34–35, 37, 39, 42, 45–48, 50–54) were ‘Classic’ OBPs that contained six positionally-conserved cysteine residues. Seven OBPs (*AdisOBP14-16*, 18, 33, 36 and 41) belonged to ‘Plus-C’ subclass OBP genes with more cysteines in addition to those of the conserved motif. Nine OBPs belonged to ‘Minus-C’ subclass OBP genes with only four cysteines. Interestingly, *AdisOBP1*, *AdisOBP17* and *AdisOBP40* did not belong to any of the four subclass OBPs (Fig. 4). However, according to BLAST results these three genes were homologous with OBP genes of *Bombyx mori*, *Spodoptera exigua* and *Dendrolimus punctatus* (Table 4). The transcription abundance of *A. dissimilis* OBPs in antennae of female and males, ovary and testis are profiled in Fig. 5.

Expression of the OBPs in the antennae, ovaries and testis of *A. dissimilis*. Next, we measured the relative expression levels of the identified OBPs in different tissues of *A. dissimilis* via fluorescence qRT-PCR (Fig. 6). A total of 23 OBPs (*AdisGOBP1-2*, *PBP1-3*, *OBP1-2*, 8–9, 11, 17, 20–22, 24, 26–31, 50 and 54) were highly expressed in the antennae compared to the reproductive organs, including three OBPs (*AdisPBP1*, *OBP17* and *OBP26*) that exhibited male-biased expression, 15 OBPs (*AdisGOBP2*, *PBP2-3*, *OBP1-2*, 11, 20–22, 27–28,

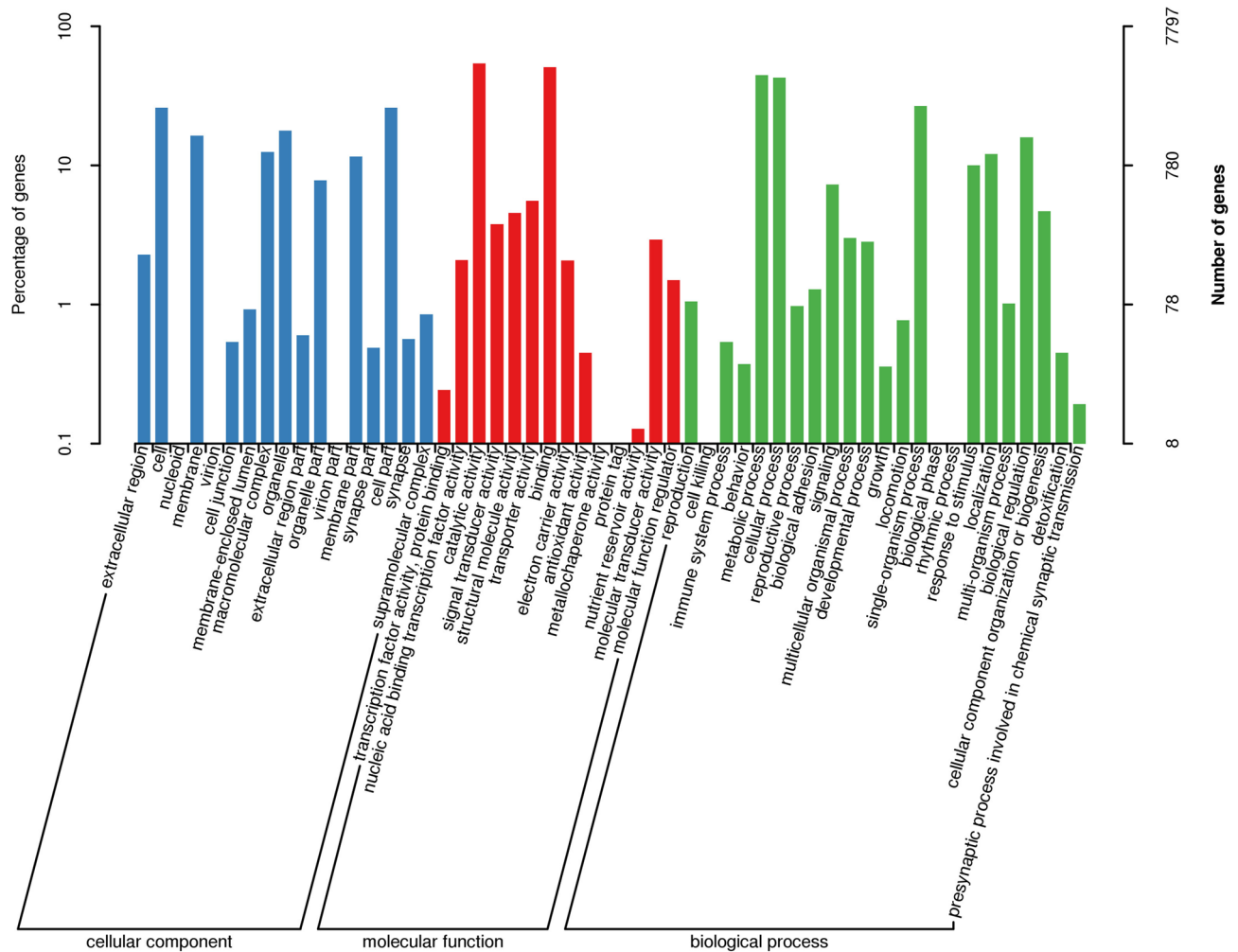


Figure 2. Gene Ontology (GO) classifications of *Athetis dissimilis* reproductive organs unigenes according to their involvement in biological processes, cellular component and molecular function.

30–31, 50 and 54) that exhibited female-biased expression, and five OBPs (*Adis GOBP1*, *OBP8-9*, 24 and 29) that showed comparable expression in the male and female antennae of *A. dissimilis*.

A total of 24 OBPs (*AdisOBP3*, 5, 15, 18–19, 23, 25, 33–41, 44–45, 47–49 and 51–53) were highly expressed in the testis of *A. dissimilis* compared to other tissues. The expression of the OBPs was low in the ovaries of *A. dissimilis*.

Discussion

In this study, we identified 31 novel OBPs through the analysis of *A. dissimilis* antennal transcriptomes, except for 5 *AdisOBP* genes identified in a previous study³². The number of OBPs in *A. dissimilis* antennae were similar to those in the antennal transcriptomes of *S. litura* (33)¹⁷ and *S. littoralis* (36)³³ but more abundant than *S. exigua* (11)³⁴, *M. sexta* (18)³⁵ and *H. armigera* (26)³⁶. We additionally sequenced the transcriptomes of *A. dissimilis* ovaries and testis. The alignments against the Nr database showed that 56.87% of the *A. dissimilis* unigenes were comparable to *H. armigera* sequences. A total of 23 OBPs were identified in the transcriptomes of *A. dissimilis* reproduction organs.

Based on sequence alignments and the cluster analysis of the phylogenetic trees, five PBP/GOBP genes, 35 Classic genes, 7 Plus-C genes and 9 Minus-C genes were obtained from the *A. dissimilis* antennal library. These results were similar to the classifications of most insect OBPs^{17,27,37}. Interestingly, *AdisOBP1*, *AdisOBP17* and *AdisOBP40* could not be clustered into any subfamilies, and multiple sequence alignments of all *AdisOBP* genes revealed that the three OBPs contain no conserved cysteines. The phylogenetic tree supports a highly dynamic evolutionary process for the lepidopteran OBP family and a high degree of OBP sequence divergence. The diversification of OBPs might be the result of multiple and late independent gene duplications. In addition, they might be derived from a common ancestor and later diverged into different subfamilies by different selection pressures, which has been evidenced by evolutionary selection analysis in several insect species^{38–40}.

OBPs are expressed specifically in the antennae and other parts associated with olfactory organs^{15,19,31,41–43}. Our comprehensive expression analysis revealed that 23 *AdisOBPs* were found to be restricted to the antenna. It is worth noting that only 3 *AdisOBPs* had male-biased expression pattern in the antennae, suggesting that females

Order	Gene name	GenBank accession no.	ORF (aa)	Molecular weight (kD)	Isoelectric point	Signal peptide	Full length	Homology search with the known proteins					
								Gene annotation	Species	Protein ID	Score	E-value	Identity (%)
c69042	AdisPBP1	KR780029	166	17.32	5.19	Yes	Yes	PBP1	<i>Mamestra brassicae</i>	AAC05702	266	3e-89	79
c65047	AdisPBP2	KR780030	162	18.08	5.30	Yes	Yes	PBP2	<i>Mamestra brassicae</i>	AAC05701	281	4e-95	81
c65143	AdisPBP3	MH900289	164	18.71	5.25	Yes	Yes	PBP3	<i>Agrotis ipsilon</i>	AFM36758	292	1e-99	82
c47645	Adis-GOBP1	KR780027	163	18.89	5.19	Yes	Yes	GOBP1	<i>Sesamia inferens</i>	AGS36742	289	3e-98	99
c60029	Adis-GOBP2	KR780028	161	18.09	5.09	Yes	Yes	GOBP2	<i>Agrotis ipsilon</i>	AFM36760	297	2e-101	88
c68783	AdisOBP1	MH900290	293	33.20	5.76	Yes	Yes	OBP	<i>Bombyx mori</i>	NP_001153663	264	1e-84	51
c69959	AdisOBP2	MH900291	246	27.36	5.40	Yes	Yes	OBP10	<i>Ostrinia furnacalis</i>	BAV56797	310	4e-104	66
c60098	AdisOBP3	MH900292	145	16.22	8.37	Yes	Yes	OBP	<i>Spodoptera exigua</i>	ADY17886	251	5e-84	79
c65852	AdisOBP5	MH900293	242	26.78	6.33	Yes	Yes	OBP35	<i>Dendrolimus punctatus</i>	ARO70194	215	2e-66	46
c72710	AdisOBP8	MH900294	240	27.01	6.53	Yes	Yes	OBP25	<i>Spodoptera exigua</i>	AKT26502	305	3e-102	63
c61153	AdisOBP9	MH900295	167	18.50	4.51	Yes	Yes	OBP10	<i>Sesamia inferens</i>	AGS36751	233	3e-76	79
c60049	Adis-OBP11	MH900296	141	16.38	4.47	Yes	Yes	OBP8	<i>Spodoptera exigua</i>	AGH70104	232	2e-76	86
c65401	Adis-OBP13	MH900297	133	15.14	9.01	Yes	Yes	OBP9	<i>Spodoptera exigua</i>	AGH70105	261	6e-88	95
c58306	Adis-OBP14	MH900298	185	20.13	6.04	Yes	Yes	OBP1	<i>Agrotis ipsilon</i>	AGR39564	279	1e-93	74
c64058	Adis-OBP15	MH900299	146	16.43	6.29	Yes	Yes	OBP6	<i>Agrotis ipsilon</i>	AGR39569	238	4e-79	88
c53621	Adis-OBP16	MH900300	118	-	-	-	Internal	OBP18	<i>Spodoptera exigua</i>	AKT26496	124	2e-33	48
c68160	Adis-OBP17	MH900301	252	28.95	6.19	Yes	Yes	OBP23	<i>Spodoptera exigua</i>	AKT26500	442	7e-156	81
c67912	Adis-OBP18	MH900302	203	22.50	5.69	Yes	Yes	OBP19	<i>Heliocoverpa assulta</i>	AGC92793	245	9e-80	62
c60881	Adis-OBP19	MH900303	139	14.55	8.58	Yes	Yes	OBP5	<i>Agrotis ipsilon</i>	AGR39568	168	4e-51	62
c71719	Adis-OBP20	MH900304	139	15.69	7.52	Yes	Yes	OBP8	<i>Spodoptera litura</i>	AKI87969	257	2e-86	87
c65033	Adis-OBP21	MH900305	147	15.65	4.90	Yes	Yes	OBP5	<i>Heliocoverpa armigera</i>	AEB54581	221	6e-72	75
c63129	Adis-OBP22	MH900306	146	15.92	7.53	Yes	Yes	OBP23	<i>Spodoptera litura</i>	XP_022826767	238	2e-78	77
c57331	Adis-OBP23	MH900307	149	15.96	5.03	Yes	Yes	OBP26	<i>Spodoptera exigua</i>	AKT26503	233	1e-76	76
c64709	Adis-OBP24	MH900308	148	16.77	5.45	Yes	Yes	OBP7	<i>Heliocoverpa armigera</i>	AEB54591	187	5e-57	57
c81048	Adis-OBP25	MH900309	71	-	-	-	Internal	OBP22	<i>Spodoptera exigua</i>	AKT26499	130	3e-37	87
c53707	Adis-OBP26	MH900310	134	14.28	4.51	Yes	Yes	OBP34	<i>Heliocoverpa assulta</i>	ASA40070	225	9e-74	86
c28876	Adis-OBP27	MH900311	124	-	-	-	Internal	OBP11	<i>Spodoptera exigua</i>	AGP03457.1	219	3e-71	81
c67118	Adis-OBP28	MH900312	236	27.80	4.90	No	Yes	OBP9	<i>Spodoptera litura</i>	ALD65883	383	1e-131	82
c57589	Adis-OBP29	MH900313	129	-	-	-	5' lose	OBP33	<i>Heliocoverpa assulta</i>	ASA40072	208	5e-67	76
c62521	Adis-OBP30	MH900314	180	20.26	4.84	No	Yes	OBP9	<i>Heliocoverpa armigera</i>	AEB54592	167	3e-50	54

Continued

Order	Gene name	GenBank accession no.	ORF (aa)	Molecular weight (kD)	Isoelectric point	Signal peptide	Full length	Homology search with the known proteins					
								Gene annotation	Species	Protein ID	Score	E-value	Identity (%)
c63839	Adis-OBP31	MH900315	116	12.77	6.12	No	Yes	OBP14	<i>Spodoptera exigua</i>	AGP03460	199	7e-64	83
Gene.53346	Adis-OBP32	MH900316	184	20.65	6.32	Yes	Yes	GOBP70	<i>Helicoverpa armigera</i>	XP_021188671	375	1e-131	98
Gene.77161	Adis-OBP33	MH900317	207	23.94	9.19	Yes	Yes	OBP19	<i>Helicoverpa assulta</i>	AGC92793	151	1e-42	39
Gene.60926	Adis-OBP34	MH900318	193	22.42	5.48	Yes	Yes	OBP9	<i>Cnaphalocrocis medinalis</i>	ALT31639	289	5e-97	70
Gene.32069	Adis-OBP35	MH900319	137	15.34	8.85	No	Yes	OBP	<i>Helicoverpa armigera</i>	AEX07279	238	7e-79	88
Gene.44893	Adis-OBP36	MH900320	143	15.92	5.57	No	Yes	OBP19	<i>Helicoverpa assulta</i>	AGC92793	187	1e-57	66
Gene.35132	Adis-OBP37	MH900321	102	-	-	-	5' lose	OBP24	<i>Cnaphalocrocis medinalis</i>	ALT31654	182	7e-58	86
Gene.54044	Adis-OBP38	MH900322	141	15.05	8.77	Yes	Yes	OBP5	<i>Agrotis ipsilon</i>	AGR39568	155	6e-46	57
Gene.7082	Adis-OBP39	MH900323	156	17.94	4.86	Yes	Yes	PBP1	<i>Helicoverpa armigera</i>	XP_021192649	129	1e-34	39
Gene.113597	Adis-OBP40	MH900324	166	19.09	8.61	Yes	Yes	OBP38	<i>Dendrolimus punctatus</i>	ARO70197	157	7e-46	63
Gene.77158	Adis-OBP41	MH900325	141	16.29	9.12	No	Yes	OBP19	<i>Helicoverpa assulta</i>	AGC92793	115	2e-29	44
Gene.14505	Adis-OBP42	MH900326	102	11.15	5.44	No	Yes	OBP23	<i>Spodoptera litura</i>	ALD65897	98.6	3e-24	49
Gene.54039	Adis-OBP43	MH900327	76	-	-	-	5' lose	OBP	<i>Helicoverpa armigera</i>	AEX07280	87.8	1e-20	59
Gene.58201	Adis-OBP44	MH900328	76	-	-	-	5' lose	OBP23	<i>Spodoptera litura</i>	ALD65897	71.6	6e-14	48
Gene.32531	Adis-OBP45	MH900329	150	16.43	4.77	Yes	Yes	OBP2	<i>Agrotis ipsilon</i>	AGR39565	119	1e-31	42
Gene.5319	Adis-OBP46	MH900330	70	-	-	-	5' lose	OBP14	<i>Spodoptera exigua</i>	AGP03460	117	2e-32	81
Gene.86678	Adis-OBP47	MH900331	120	-	-	-	5' lose	OBP13	<i>Sesamia inferens</i>	AGS36753	137	8e-39	53
Gene.141496	Adis-OBP48	MH900332	106	12.10	6.95	No	Yes	OBP39	<i>Dendrolimus punctatus</i>	ARO70198	183	4e-57	82
Gene.142856	Adis-OBP49	MH900333	157	17.96	9.74	Yes	Yes	OBP18	<i>Dendrolimus punctatus</i>	ARO70177	119	3e-31	51
Gene.17592	Adis-OBP50	MH900334	144	16.21	4.44	Yes	Yes	OBP9	<i>Helicoverpa armigera</i>	AEB54592	163	5e-49	54
Gene.54647	Adis-OBP51	MH900335	84	-	-	-	5' lose	OBP39	<i>Dendrolimus punctatus</i>	ARO70198	140	1e-40	86
Gene.76032	Adis-OBP52	MH900336	105	11.60	4.71	No	Yes	OBP	<i>Spodoptera litura</i>	ALD65897	111	4e-29	52
Gene.111996	Adis-OBP53	MH900337	105	12.28	8.21	No	Yes	OBP	<i>Operophtera brumata</i>	KOB73304	194	1e-61	88
Gene.158529	Adis-OBP54	MH900338	131	14.34	4.86	No	Yes	OBP11	<i>Spodoptera exigua</i>	AGP03457	226	3e-74	79

Table 4. The characteristic of candidate OBP genes in the antennae and reproductive organs of *A. dissimilis*. Genes beginning with the lowercase letter “c” came from the identification of antenna transcriptome, and genes beginning with “Gene” came from testis and ovary transcriptome identification.

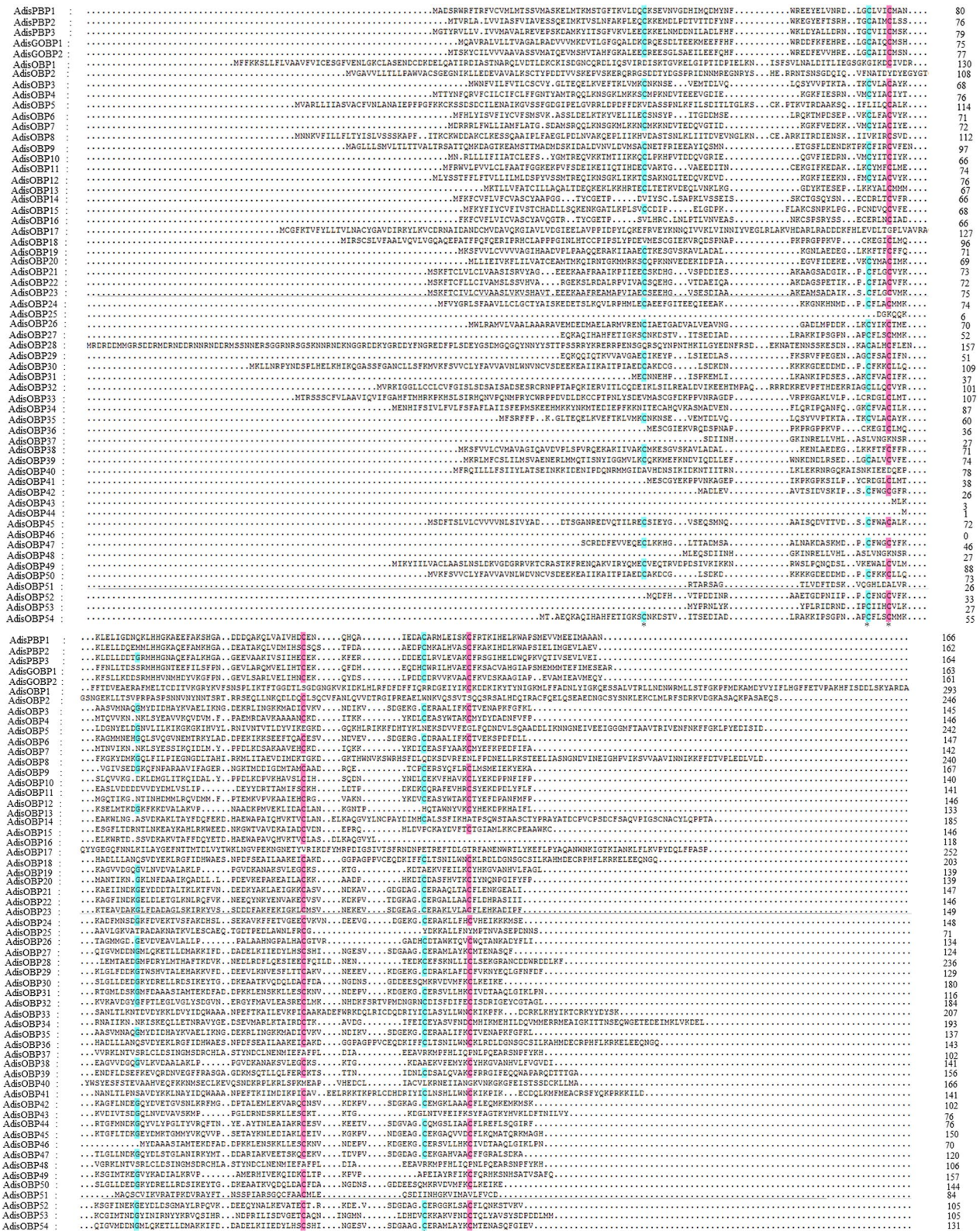


Figure 3. Sequence alignments of *Aethis dissimilis* OBPs. The six conserved cysteine residues are indicated with the asterisks under the sequence.

require more abundant OBPs for spawning. It is interesting to note 24 *AdisOBPs* showed significant expression in the testis of *A. dissimilis* compared to other tissues, but the expression of *AdisOBPs* in the ovaries was low. The

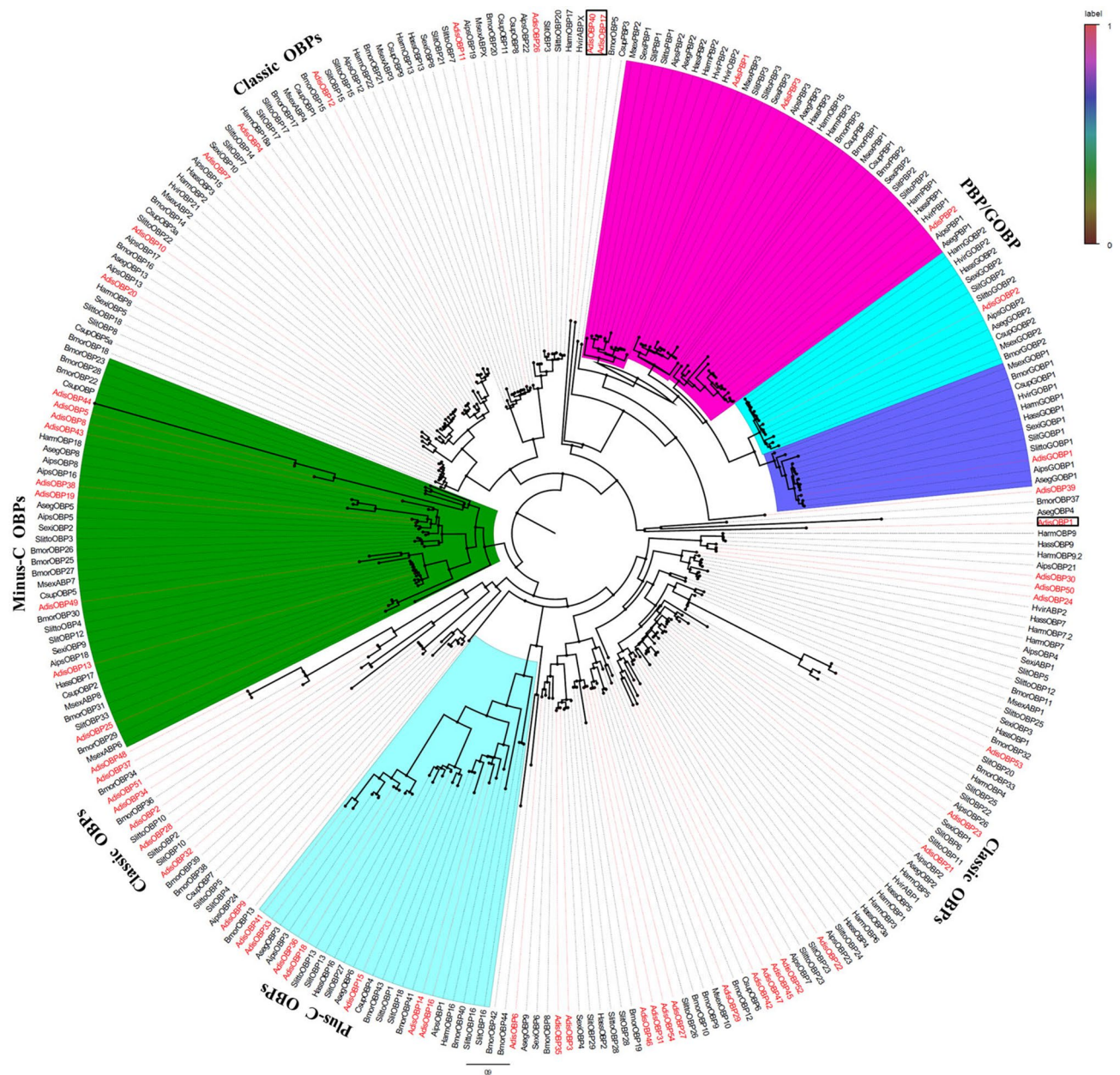


Figure 4. Phylogenetic relationships of candidate OBP proteins (including 5 OBPs identified in a previous study) from *Athetis dissimilis* and 33 Lepidoptera species.

expression of OBPs in reproduction has also been reported in some literature^{44–46}. It was previously speculated that OBPs expressed in the testis deliver compounds to the females during mating^{26,27}. Hence, it is understandable to presume that such stable proteins could be used in the testis of insect where there is need for transportation of hydrophobic molecules in aqueous media or protection of chemicals from degradation, as well as to assure a gradual release of semiochemicals in the environment. So these proteins have been named “encapsulins”, to imply the common role of encapsulating small ligands⁴⁷. qRT-PCR was conducted on 53 candidate genes, and the expression level of most genes were consistent with the variation of RPKM values.

Like the OBP families of insect antennae, insect testes contain a large number of OBP genes. The functions of these genes is unclear, and they need us to further study. Our results provide a reference for the study of these genes.

Materials and methods

Insect rearing and sample preparation. The *A. dissimilis* strain was collected from Luoyang (province of Henan, China) corn fields (112° 26' E, 34° 43' N) in 2014 and maintained at the Henan Science and Technology University. Colonies were reared on an artificial diet at 25 ± 1 °C, 80 ± 5% relative humidity and a 16-h/8-h light/dark cycle.

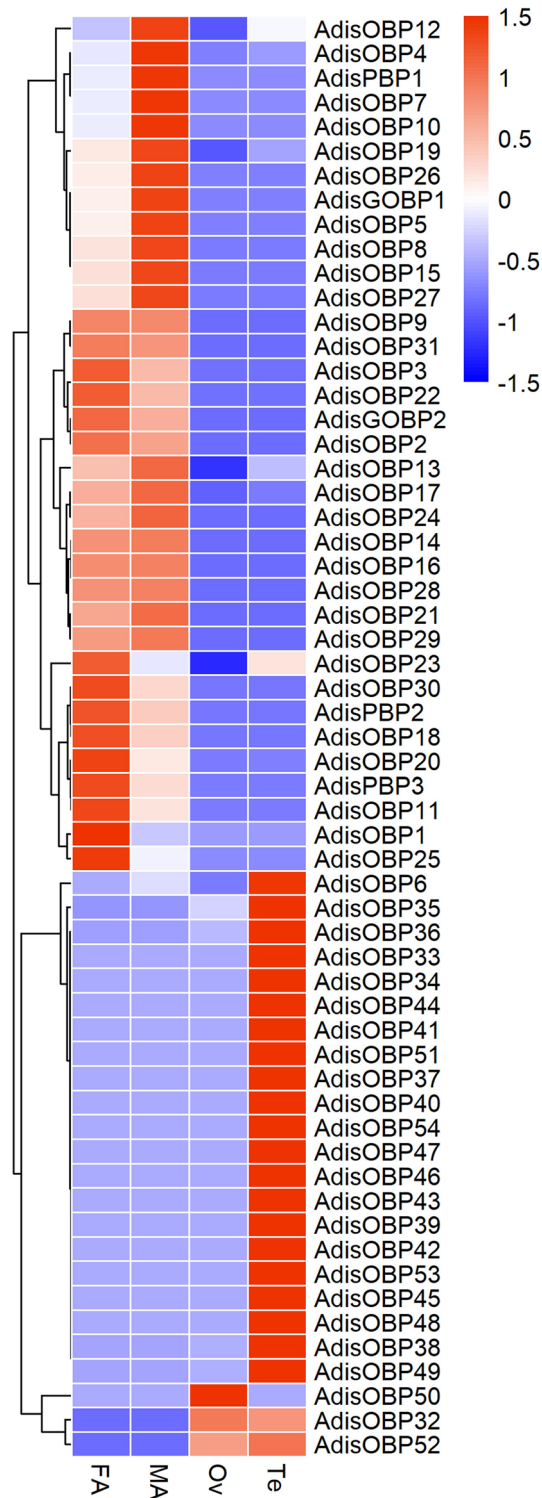


Figure 5. Heat map showing the abundance of unigenes encoding OBPs (including 5 OBPs identified in a previous study) in the *Athetis dissimilis* different tissues transcriptomes presented as normalized reads in reads per kilobase per million mapped reads (RPKM). In the figure each column represents 1 samples; each line represents 1 OBP gene. The color depth represents the number of reads contained in OBPs; red means more; blue means less. FA female antennae, MA male antennae, Ov ovaries, Te testis.

Based on preliminary data, we found that the *A. dissimilis* sperm and eggs began to mature 3 days after emergence. We respectively collected the ovaries and testes of 3-day old virgin females and male adults (n = 40 per

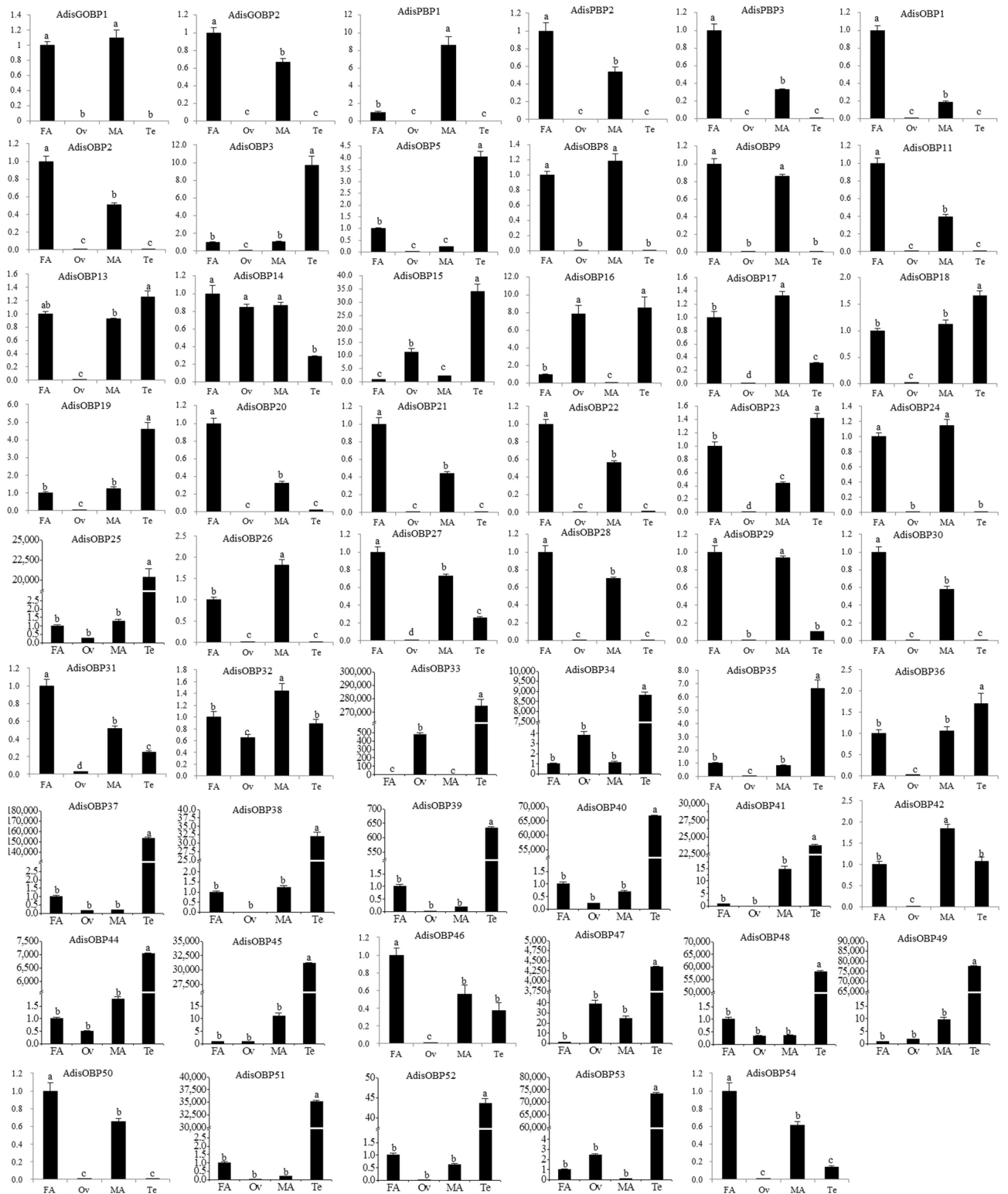


Figure 6. Expression profiles of the candidate OBPs in different tissues of *Athetis dissimilis*. FA female antennae, MA male antennae, Ov ovaries, Te testis. The standard errors are represented by the error bars; different lowercase letters (a–c) above the bars denote significant differences at $p < 0.05$.

treatment) from three biological replications. Dissections were performed in sterile PBS-DEPC and immediately frozen in liquid nitrogen until RNA isolation.

cDNA library preparation and sequencing. Total RNA from the *A. dissimilis* ovaries and testis tissues were extracted using RNAiso Plus kit (TaKaRa, Dalian, China) and treated with DNase I (TaKaRa, Dalian, China) as per the manufacturer's protocols. RNA was assessed through 1% agarose gel electrophoresis and Nanodrop 2000 (Thermo Scientific, Waltham, MA, USA), Qubit 2.0 (Life Technologies, Carlsbad, CA, USA) and Agilent 2100 (Agilent, Santa Clara, CA, USA) analysis.

Following the TruSeq RNA Sample Preparation Guide v2 (Illumina, San Diego, CA, USA), mRNA was enriched using magnetic beads crosslinked with Oligo (dT). Enriched RNA was then fragmented using fragmentation buffer and first-strand cDNA synthesis was used to produce small mRNA fragments, random primers, reverse transcriptase, and second-strand cDNA synthesis through the addition of dNTPs, DNA polymerase I, and RNase H. Double-stranded cDNA was purified with AMPure XP beads (Beckman Coulter, Brea, CA, USA) and treated to repair ends, remove poly(-A) tails, and link sequencing adapters. Fragment sizes were selected using AMPure XP beads and cDNA libraries were constructed through PCR amplification (Veriti 96-Well Thermal Cycle, Applied Biosystems, Foster City, USA). The concentration and insert size of the cDNA libraries were detected using Qubit 2.0 and Agilent 2100 and quantified via q-PCR (CFX-96, Bio-Rad, Hercules, CA, USA).

Finally, sequencing was performed using the Illumina HiSeq 4000 platform to generate 150-bp paired-end reads. Sequencing analyses were performed by the Genomics Services of the Beijing Biomarker Technologies Co., Ltd. (Beijing, China). Raw data processing and base calling were performed using Illumina software.

Assembly and functional annotation. Raw data (raw reads) in the FASTQ format were first modified into clean data (clean reads) through Perl scripts. This was performed through the removal of reads containing adapter sequences, >10% unknown nucleotides and quality values ≤ 20 . The Q20, Q30, and GC content were then calculated using high-quality data.

Transcriptomes were assembled using Trinity (version trinityrnaseq_r20131110) with default settings, except for min_kmer_cov set to 2⁴⁸. Unigene functions were annotated based on NCBI non-redundant protein sequences (NR, NCBI blast 2.2.28+, e-value = 1e-5), NCBI nucleotide sequences (NT, NCBI blast 2.2.28+, e-value = 1e-5), Protein family (Pfam, HMMER 3.0 package, hmmscan, e-value = 0.01), eukaryotic Ortholog Groups (KOG, NCBI blast 2.2.28+, e-value = 1e-3), SwissProt (NCBI blast 2.2.28+, e-value = 1e-5), the Kyoto Encyclopedia of Genes and Genomes (KEGG; KEGG Automatic Annotation Server [KASS], e-value = 1e-10) and Gene Ontology (GO, Blast2GO v2.5, e-value = 1e-6). Coding sequences (CDS) were predicted through aligning transcriptome sequences to the Nr and Swiss-Prot database or using estscan 3.0.3⁴⁹. The read count for each gene was obtained by mapping clean reads to the assembled transcriptome using RSEM (bowtie2 parameters: mismatch 0). The final read count was calculated as Fragments Per Kilobase of transcript per Million mapped reads (FPKM)⁵⁰.

Sequence and phylogenetic analysis. Sequence similarities were assessed using the NCBI-Blast network server (<http://blast.ncbi.nlm.nih.gov/>). The signal peptides of OBPs were predicted using SignalP 4.1 (<http://www.cbs.dtu.dk/services/SignalP/>)⁵¹. Multiple sequence alignments were assessed using DNAMAN 6.0. Sequence alignments of the candidate OBPs were performed using ClustalX 2.1⁵² and used to construct phylogenetic trees with PhyML in Seaview v.4 based on the Jones–Taylor–Thornton (JTT) model with nearest-neighbor interchanges. Trees were viewed and edited using FigTree v.1.3.1. Amino acid sequences of OBPs in phylogenetic tree were listed in Supplementary File 1.

Expression analysis through quantitative real-time polymerase chain reaction. Male antennae (100), female antennae (100), ovaries (80) and testes (150) tissue from adults at 3 post-eclosion were excised and frozen in liquid nitrogen. Total RNA was extracted using RNAiso Plus kits (TaKaRa, Dalian, China) and isolated RNA was transcribed to first-strand cDNA using PrimeScript RT reagent with gDNA Eraser (TaKaRa, Dalian, China) following the manufacturer's protocols. Real-time quantitative PCR (RT-qPCR) was performed with SYBR[®] Premix Ex Taq II (TaKaRa). The *A. dissimilis* GAPDH gene was used as an endogenous control to correct for sample-to-sample variations. A 200 ng/ μ L cDNA sample was used for per tissue. Primers were designed using Primer Premier 5.0 software and are listed in Supplementary File 2. RT-qPCR reactions contained: 10 μ L of SYBR Premix Ex Taq II, 20 ng of cDNA template, 0.2 μ M of each primer and nuclease-free water. The cycling conditions were 1 cycle of 95 °C for 5 min, followed by 40 cycles of 95 °C for 5 s and 55 °C for 30 s. Melt curve conditions were 95 °C for 10 s and 65 °C for 30 s. No-template controls (NTC) were included to detect possible contamination. Three biological replicates were analyzed and the relative expression of the OBP genes was measured using the 2^{- $\Delta\Delta$ CT} method⁵³. Expression was calculated relative to levels in the female antennae, which were arbitrarily set to 1. Differences in the expression of *Adis*OBP genes between the different tissues were compared using a one-way nested analysis of variance (ANOVA), followed by a Tukey's honestly significance difference (HSD) test using SPSS (SPSS Institute 17.0, SPSS Inc, Chicago, IL, USA).

Received: 3 February 2021; Accepted: 22 June 2021

Published online: 06 July 2021

References

- Field, L. M., Pickett, J. A. & Wadhams, L. J. Molecular studies in insect olfaction. *Insect Mol. Biol.* **9**, 545–551 (2000).
- Zhan, S., Merlin, C., Boore, J. L. & Reppert, S. M. The monarch butterfly genome yields insights into long-distance migration. *Cell* **147**, 1171–1185 (2011).
- Suh, E., Bohbot, J. & Zwiebel, L. J. Peripheral olfactory signaling in insects. *Curr. Opin. Insect Sci.* **6**, 86–92 (2014).

4. Sun, L. *et al.* Perception of potential sex pheromones and host-associated volatiles in the cotton plant bug, *Adelphocoris fasciaticollis* (Hemiptera: Miridae): Morphology and electrophysiology. *Appl. Entomol. Zool.* **49**, 43–57 (2014).
5. Zhang, J., Walker, W. B. & Wang, G. Pheromone reception in moths: From molecules to behaviors. *Prog. Mol. Biol. Transl. Sci.* **130**, 109–128 (2015).
6. Vogt, R. G. & Riddiford, L. M. Pheromone binding and inactivation by moth antennae. *Nature* **293**, 161–163 (1981).
7. Pelosi, P. & Maida, R. Odorant-binding proteins in insects. *Comp. Biochem. Phys. B* **111**, 503–514 (1995).
8. Angeli, S. *et al.* Purification, structural characterization, cloning and immunocytochemical localization of chemoreception proteins from *Schistocerca gregaria*. *Eur. J. Biochem.* **262**, 745–754 (1999).
9. Leal, W. S., Nikonova, L. & Peng, G. Disulfide structure of the pheromone binding protein from the silkworm moth, *Bombyx mori*. *FEBS Lett.* **464**, 85–90 (1999).
10. Zhou, J. J. Odorant-binding proteins in insects. *Vitamins Hormones* **83**, 241–272 (2010).
11. Lagarde, A. *et al.* The crystal structure of odorant binding protein 7 from *Anopheles gambiae* exhibits an outstanding adaptability of its binding site. *J. Mol. Biol.* **414**, 401–412 (2011).
12. Pelosi, P., Zhou, J. J., Ban, L. & Calvello, M. Soluble proteins in insect chemical communication. *Cell. Mol. Life Sci.* **63**, 1658–1676 (2006).
13. Leal, W. S. Odorant reception in insects: Roles of receptors, binding proteins, and degrading enzymes. *Annu. Rev. Entomol.* **58**, 373–391 (2013).
14. Liu, Z., Smagghe, G., Lei, Z. R. & Wang, J. J. Identification of male-and female-specific olfaction genes in antennae of the oriental fruit fly (*Bactrocera dorsalis*). *PLoS ONE* **11**, e0147783 (2016).
15. Qiao, H. *et al.* Discrimination of alarm pheromone (E)-beta-farnesene by aphid odorant-binding proteins. *Insect Biochem. Mol.* **39**, 414–419 (2009).
16. Zhou, J. J. *et al.* Characterisation of *Bombyx mori* Odorant-binding proteins reveals that a general odorant-binding protein discriminates between sex pheromone components. *J. Mol. Biol.* **389**, 529–545 (2009).
17. Gu, S. H. *et al.* Identification and comparative expression analysis of odorant binding protein genes in the tobacco cutworm *Spodoptera litura*. *Sci. Rep.* **5**, 13800 (2015).
18. Larter, N. K., Sun, J. S. & Carlson, J. R. Organization and function of *Drosophila* odorant binding proteins. *Elife* **5**, e20242 (2016).
19. Jia, X., Zhang, X., Liu, H., Wang, R. & Zhang, T. Identification of chemosensory genes from the antennal transcriptome of Indian meal moth *Plodia interpunctella*. *PLoS ONE* **13**, e0189889 (2018).
20. Cao, D. P. *et al.* Identification of candidate olfactory genes in *Chilo suppressalis* by antennal transcriptome analysis. *Int. J. Biol. Sci.* **10**, 846 (2014).
21. Zhang, S. F., Zhang, Z., Wang, H. B. & Kong, X. B. Antennal transcriptome analysis and comparison of olfactory genes in two sympatric defoliators, *Dendrolimus houi* and *Dendrolimus kikuchii* (Lepidoptera: Lasiocampidae). *Insect Biochem. Mol.* **52**, 69–81 (2014).
22. Jia, X. J. *et al.* Antennal transcriptome and differential expression of olfactory genes in the yellow peach moth, *Conogethes punctiferalis* (Lepidoptera: Crambidae). *Sci. Rep.* **6**, 29067 (2016).
23. Cheng, W. N., Zhang, Y. D., Liu, W., Li, G. W. & Zhu-Salzman, K. Molecular and functional characterization of three odorant-binding proteins from the wheat blossom midge, *Sitodiplosis mosellana*. *Insect Sci.* **98**, e21456 (2019).
24. Pregitzer, P. *et al.* Expression of odorant-binding proteins in mouthpart palps of the desert locust *Schistocerca gregaria*. *Insect Mol. Biol.* **28**, 264–276 (2019).
25. Yin, J. *et al.* Functional characterization of odorant-binding proteins from the scarab beetle *Holotrichia oblita* based on semiochemical-induced expression alteration and gene silencing. *Insect Biochem. Mol.* **104**, 11–19 (2019).
26. Li, S. *et al.* Multiple functions of an odorant-binding protein in the mosquito *Aedes aegypti*. *Biochem. Biophys. Res. Co.* **372**, 464–468 (2008).
27. Sun, Y. L., Huang, L. Q., Pelosi, P. & Wang, C. Z. Expression in antennae and reproductive organs suggests a dual role of an odorant-binding protein in two sibling *Helicoverpa* species. *PLoS ONE* **7**, e30040 (2012).
28. Li, J. W. *et al.* Morphologically a like species of *Aethis lepigone* (Möschler)—*A. dissimilis* (Hampson) found in Shandong Province of China. *Plant Prot.* **40**, 193–195 (2014).
29. Takahashi, M. *Aethis dissimilis* Hampson, a new nuisance? *Jpn. Soc. Med. Entomol. Zool.* **26**, 66 (1975).
30. Li, N. G. Relationships between cold hardiness and ice nucleating activity, glycerol and protein contents in the hemolymph of caterpillars, *Aporia crataegi* L.. *CryoLetters* **33**, 134–142 (2012).
31. Dong, J., Song, Y., Li, W., Shi, J. & Wang, Z. Identification of putative chemosensory receptor genes from the *Aethis dissimilis* antennal transcriptome. *PLoS ONE* **11**, e0147768 (2016).
32. Sun, H. Z., Song, Y. Q., Du, J., Wang, X. D. & Cheng, Z. J. Identification and tissue distribution of chemosensory protein and odorant binding protein genes in *Aethis dissimilis* (Lepidoptera: Noctuidae). *Appl. Entomol. Zool.* **51**, 409–420 (2016).
33. Poivet, E. *et al.* A comparison of the olfactory gene repertoires of adults and larvae in the noctuid moth *Spodoptera littoralis*. *PLoS ONE* **8**, e60263 (2013).
34. Zhu, J. Y., Zhang, L. F., Ze, S. Z., Wang, D. W. & Yang, B. Identification and tissue distribution of odorant binding protein genes in the beet armyworm, *Spodoptera exigua*. *J. Insect Physiol.* **59**, 722–728 (2013).
35. Grosse-Wilde, E. *et al.* Antennal transcriptome of *Manduca sexta*. *Proc. Natl. Acad. Sci.* **108**, 7449–7454 (2011).
36. Liu, Y., Gu, S. H., Zhang, Y. J., Guo, Y. Y. & Wang, G. R. Candidate olfaction genes identified within the *Helicoverpa armigera* antennal transcriptome. *PLoS ONE* **7**, e48260 (2012).
37. Chang, H. *et al.* Candidate odorant binding proteins and chemosensory proteins in the larval chemosensory tissues of two closely related noctuid moths, *Helicoverpa armigera* and *H. assulta*. *PLoS ONE* **12**, e0179243 (2017).
38. Forêt, S. & Maleszka, R. Function and evolution of a gene family encoding odorant binding-like proteins in a social insect, the honey bee (*Apis mellifera*). *Genome Res.* **16**, 1404–1413 (2006).
39. Zhou, J. J. *et al.* Genome annotation and comparative analyses of the odorant-binding proteins and chemosensory proteins in the pea aphid *Acyrtosiphon pisum*. *Insect Mol. Biol.* **19**, 113–122 (2010).
40. Zhang, G. H., Li, Y. P., Xu, X. L., Chen, H. & Wu, J. X. Identification and characterization of two general odorant binding protein genes from the oriental fruit moth, *Grapholita molesta* (Busck). *J. Chem. Ecol.* **38**, 427–436 (2012).
41. Feng, B., Guo, Q. S., Zheng, K. D., Qin, Y. X. & Du, Y. J. Antennal transcriptome analysis of the piercing moth *Oraesia emarginata* (Lepidoptera: Noctuidae). *PLoS ONE* **12**, e0179433 (2017).
42. Wei, H. S., Li, K. B., Zhang, S., Cao, Y. Z. & Yin, J. Identification of candidate chemosensory genes by transcriptome analysis in *Loxostege sticticalis* Linnaeus. *PLoS ONE* **12**, e0174036 (2017).
43. Tian, Z. Q. *et al.* Antennal transcriptome analysis of the chemosensory gene families in *Carposina sasakii* (Lepidoptera: Carposinidae). *BMC Genomics* **19**, 544 (2018).
44. Zhang, T. T. *et al.* Male-and female-biased gene expression of olfactory-related genes in the antennae of Asian corn borer, *Ostrinia furnacalis* (Guenee) (Lepidoptera: Crambidae). *PLoS ONE* **10**, e0128550 (2015).
45. Song, L. M. *et al.* Male tarsi specific odorant-binding proteins in the diving beetle *Cybister japonicus* sharp. *Sci. Rep.* **6**, 31848 (2016).
46. Chen, X. F. *et al.* Genome-wide identification and expression profiling of odorant-binding proteins in the oriental fruit fly, *Bactrocera dorsalis*. *Comp. Biochem. Phys. D* **31**, 100605 (2019).

47. Leal, W. S. Molecular-based chemical prospecting of mosquito attractants and repellents. In *Insect Repellents: Principles, Methods, and Uses* (eds Debboun, M. F. & Strickman, D.) 229–244 (CRC Press, 2006).
48. Grabherr, M. G. *et al.* Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat. Biotechnol.* **29**, 644–652 (2011).
49. Iseli, C., Jongeneel, C. V. & Bucher, P. ESTScan: A program for detecting, evaluating, and reconstructing potential coding regions in EST sequences. *Proc. Int. Conf. Intell. Syst. Mol. Biol.* **7**, 138–148 (1999).
50. Mortazavi, A., Williams, B. A., McCue, K., Schaeffer, L. & Wold, B. Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nat. Methods* **5**, 621–628 (2008).
51. Petersen, T. N., Brunak, S., von Heijne, G. & Nielsen, H. SignalP 4.0: Discriminating signal peptides from transmembrane regions. *Nat. Methods* **8**, 785–786 (2011).
52. Larkin, M. *et al.* Clustal W and Clustal X version 2.0. *Bioinformatics* **23**, 2947–2948 (2007).
53. Livak, K. J. & Schmittgen, T. D. Analysis of relative gene expression data using real-time quantitative PCR and the 2(T) (- $\Delta\Delta C$) method. *Methods* **25**, 402–408 (2001).

Acknowledgements

This study is supported by the National Natural Science Foundation of China (31701788) and Science and Technology Project in Henan Province (202102110069).

Author contributions

H.Z. and Y.Q. conceived this project and analyzed the data. Z.Y., Q.H., Q.X. and J.F. assisted the preparation of samples and operated experiment. Y.Q. wrote the main manuscript text. H.Z. edited the manuscript. All the authors reviewed the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-021-93423-1>.

Correspondence and requests for materials should be addressed to H.-Z.S.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021