

Transcriptomics-Based Identification of Aquaporin Diversity in the House Dust Mite *Dermatophagoides farinae* (Acariformes: Pyroglyphidae)

Jianglong Peng,^{1,*} Ying Zhou,^{2,*} Haoyuan Jia,³ Ling Li,⁴ Jun Qian,⁴ Feifei Han,³ Hao Yin,³ and Yubao Cui^{3,5,✉}

¹School of Tropical Medicine and Laboratory Medicine, Hainan Medical University, Haikou 571101, China, ²Department of Pediatrics Laboratory, Wuxi Children's Hospital, Wuxi 214023, China, ³Department of Clinical Laboratory, Wuxi People's Hospital Affiliated to Nanjing Medical University, No. 299 at Qingyang Road, Wuxi 214023, Jiangsu Province, China, ⁴Department of Pediatrics Respiratory, Wuxi Children's Hospital, Wuxi 214023, China, and ⁵Corresponding author, e-mail: ybcui1975@hotmail.com; ybcui1975@njmu.edu.cn

*Equal contribution, co-first authors.

Subject Editor: Bill Bendena

Received 20 August 2018; Editorial decision 13 September 2018

Abstract

Aquaporin water channel proteins are highly conserved across many diverse species. Some evidence indicates that aquaporins in insects may contribute to insect-related mammalian diseases and inflammation, and thus these proteins may represent viable therapeutic targets. Here, we used RNA sequencing and bioinformatics to identify putative aquaporins from the house dust mite, *Dermatophagoides farinae*. Six putative aquaporins were identified based on sequence similarity with aquaporins from other species. These putative aquaporins, deposited in GenBank and named DerfAQP1–4 (KY231248, KY231249, KY231250, and KY231251, respectively), DerfAQP5.01, and DerfAQP5.02 (KY231252 and KY231253), were successfully cloned into a bacterial plasmid. The identification of full-length aquaporin sequences from *D. farinae* provides a foundation for future molecular and biochemical studies of these proteins in *D. farinae* and related species.

Key words: aquaporin, *Dermatophagoides farinae*, transcriptomics, RNA-Seq

Aquaporins (AQPs), also called water channels or major intrinsic proteins, facilitate the passive transport of water and a variety of low-molecular-weight solutes, such as glycerol, urea, boric acid, silicic acid, ammonia (NH₃), carbon dioxide (CO₂), and hydrogen peroxide (H₂O₂) (Maurel et al. 2008, Gomes et al. 2009). AQPs are highly conserved: they are found in diverse organisms from eubacteria to animals. Thirteen human AQPs have been identified to date (AQP0–AQP12). These AQPs can be grouped into three major categories according to their transport properties. The first category, classical aquaporins, comprises seven members (AQP0, -1, -2, -4, -5, -6, and -8) that are believed to transport only water, thus serving an essential role in trans-cellular water transport. However, AQPs -6 and -8 have been included in the classical category primarily based on sequence homology, since AQP6 is also permeated by anions, and AQP8 possibly by both water and urea (Fu and Lu 2007). The second category, aquaglyceroporins, comprises four members (AQP3, -7, -9, and -10) believed to transport small, uncharged molecules in addition to water. The third category, unorthodox aquaporins, comprises AQP11 and -12, whose functions are currently being explored. Some archaeobacterial AQPs may also belong to this third category, which lies somewhere between classical aquaporins and aquaglyceroporins (Kozono et al. 2002).

In the subclass Acari, a cDNA encoding a putative aquaporin was cloned from the ovaries of the American dog tick *Dermacentor variabilis* (Acari: Ixodidae) (Holmes et al. 2008). An aquaporin, RsAQP1, was identified in a salivary gland cDNA library in the brown dog tick, *Rhipicephalus sanguineus*. Subsequent functional characterization of an acarine AQP was performed in *Xenopus* oocytes. Oocytes expressing RsAQP1 became water permeable; however, RsAQP1 did not transport glycerol or urea (Ball et al. 2009). An aquaporin, IrAQP1, identified in the sheep tick *Ixodes ricinus* was present only in tissues involved in mass water flux (namely the gut and rectal sac, and especially abundant in the salivary glands) (Campbell et al. 2010). IrAQP1 thus played a pivotal role in the blood, and in meal and water handling through the gut and salivary glands. In transcriptomic studies on gut tissues dissected from fully engorged adult female cattle ticks (*Rhipicephalus microplus*), a cDNA encoding for an aquaporin, RmAQP1, had similarity to the human AQP7. Subsequently, a cDNA encoding a significant portion of RmAQP1 was expressed as a recombinant protein in *Pichia pastoris*, purified, and used to intramuscularly vaccinate cattle. This recombinant protein immunization provided 75 and 68% efficacy in reducing the numbers of adult female ticks in two cattle pen trials in Brazil. The

© The Author(s) 2018. Published by Oxford University Press on behalf of Entomological Society of America.

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact journals.permissions@oup.com

effectiveness of the vaccine indicates that this aquaporin antigen holds promise as an active ingredient in cattle vaccines to prevent infestations of *R. microplus* (Guerrero et al. 2014).

Dermatophagoides farinae is one of the major house dust mite species that induces allergic diseases, such as bronchial asthma, rhinitis, and atopic dermatitis, in many people throughout the world. Thus, identifying and understanding the potentially antigenic proteins produced by this species is critical to reducing allergy burdens. Here, in transcriptomic analyses of *D. farinae*, six putative aquaporins were identified. We describe the first successful cDNA cloning and consecutive molecular characterization of these putative aquaporins.

Materials and Methods

Total RNA of *Dermatophagoides farinae*

Dermatophagoides farinae mites were isolated and cultured according to our previously reported methods (Cui et al. 2008, 2010, 2012a,b, 2013, 2014, 2016b). Approximately 2,000 *D. farinae* mites were selected under a stereomicroscope before homogenization. Total RNA was extracted from homogenized mites according to the mini handbook in the RNeasy Mini Kit (Qiagen 74104, Hilden, Germany). To evaluate the quality of the total RNA, the concentrations, 28S/18S, and RNA integrity number (RIN) were detected on Agilent 2100 Bioanalyzer with Agilent RNA 6000 Nano Reagents Port 1. The RNA was of good purity (UV260/280 = 1.95, UV260/230 = 1.37). The RIN value was 6.4. Total RNA was dissolved in RNase-free water and stored at -80°C.

Transcriptome Library Construction, Sequencing, De Novo Assembly, and Annotation

As described in our previous study (Cui et al. 2016a), a transcriptome library was constructed with a SMART cDNA Library Construction Kit (CLONTECH Corporation, Mountain View, CA; Code No. 634901). In brief, 10 µg total RNA of *D. farinae* was taken to construct the transcriptome library using SMART cDNA Library Construction Kit (CLONTECH, Code No. 634901). Magnetic beads with Oligo (dT) were used to bind and purify mRNA. mRNA was randomly broken into short fragments, which were used as a template for reverse transcription with random primers. cDNA was purified and blunt end repair was performed, adding a base A and a linker to 3' end. Products were purified twice using Agencourt AMPure XP 60 ml Kit (Beckman coulter, Miami, FL), to thoroughly remove excess linker. Then 12 cycles of PCR amplification were performed; after library quality inspection, sequencing was performed using Illumina HiSeq 2500 Sequencer.

Sequencing data were pretreated by removing the reads with adapters, with >5% ratio of information on unidentified bases, or with low quality (number of bases with quality values ≤10 was >20 among the entire reads). Trinity, a short-reads assembling software, was used to assemble de novo sequence (<http://trinityrnaseq.sourceforge.net/>) (Grabherr et al. 2011). CD-HIT software was used to homogeneously cluster all the transcripts, using a sequence similarity of 95% as the threshold value. Unigene was compared separately with the following public databases: Swiss-Prot, NR, KEGG, COG, and Pfam (ftp://ftp.uniprot.org/pub/databases/uniprot/current_release/knowledgebase/complete/uniprot_sprot.fasta.gz; <ftp://ftp.ncbi.nlm.nih.gov/blast/db/FASTA/nr.gz>; <http://www.kegg.jp/kegg/download/>; <http://www.ncbi.nlm.nih.gov/COG/grace/shokog.cgi>; and <ftp://ftp.sanger.ac.uk/pub/databases/Pfam/release/Pfam27.0/Pfam-A.fasta.gz>). Functional annotation was carried out by the similarity of the genes: taking *E*-value ≤ 1E-5 annotated and combined genotypes to obtain all annotation details, then performing COG and KEGG classification for unigenes, as

well as GO (Gene Ontology) function annotation. Genescan software was used to predict the coding region of the Unigene sequence, and to obtain nucleotide and amino acid sequences of the coding region, with partial results displayed by CDS nucleotide sequence. The assembled transcripts (≥200 bp) were used as the library, and the pretreated sample database was mapped to the transcript using Bowtie software (<http://bowtie-bio.sourceforge.net/>). Single nucleotide polymorphism (SNP) analysis was carried out according to the mapping results, then variable sites with high possibility were screened by Samtools software (<http://samtools.sourceforge.net/>). The assembled genes (≥200 bp) were used as the library to calculate the expression abundance of each gene in each sample by utilizing the sequence similarity matching method and Bowtie software 0.12.8, with one reads mapped to multiple genes using the single-end mapping method. The gene expression quantity was represented by RPKM value (reads per kilobase of exon model per million mapped reads), calculated as follows:

$$\text{RPKM} = \frac{\text{Total exon reads}}{\text{mapped reads (millions)} * \text{exonlength (KB)}}$$

Cloning, Sequencing, and Bioinformatics Analysis for Aquaporin-Coding Genes of *Dermatophagoides farinae*

The aquaporin coding sequence was obtained by sequencing and functional annotation of the transcriptome, that is to say, the unigenes were aligned by BLASTX (*E*-value ≤ 1.0E-5; <http://blast.ncbi.nlm.nih.gov/>) to known AQP from other species in protein databases in the priority order NR, Swiss-Prot, KEGG, and COG. Primers were designed against these sequences, and potential aquaporins were amplified from total RNA by RT-PCR. Full-length genes were obtained by RACE (Cui et al. 2016b). The primers were designed according to the full-length gene sequence (as shown in Table 1); the total RNA of *D. farinae* was used as a template, and the RT-PCR amplification was performed to obtain six putative aquaporin-coding genes. These cDNA fragments were then inserted into pET28(a) for molecular cloning and sequencing validation.

Results

In total, 38,137 transcripts, with an average length of 653 bases, were obtained from total RNA of *D. farinae*.

The 22,023 unigenes identified from the *D. farinae* transcriptome were annotated by BLAST tools, including Swiss-prot (5,904; 26.81%), nr (7,952; 36.11%), Pfam (6,412; 29.12%), KEGG (7,402; 33.61%), and COG (6,081; 27.61%). The NR database queries revealed similarities in *D. farinae* sequence to *Ixodes scapularis* 11.9%, *Metaselulus occidentalis* 5.5%, *Daphnia pulex* 3.9%, *Tribolium castaneum* 3.3%, *Wallemia sebi* 3.2%, *Pediculus humanus* 2.4%, and others 69.8%. The function annotations from COG and GO are provided in Figs. 1 and 2. The most common functions identified were transcription/translation and ATP binding.

Unigene SNP Analysis

SNPs at the genomic level can affect gene functions. Thus, we assessed SNP sites of within coding sequences of transcripts. We identified 25,573 SNPs; 16,702 were transition (A-G, C-T) variants and 8,871 were transversions (A-C, A-T, C-G, G-T).

Cloning and Sequencing of a Full-Length Gene of an Aquaporin of *Dermatophagoides farinae*

Transcript sequences were input into the public databases Swiss-Prot, NR, KEGG, KOG, and Pfam for BLASTX. Nine aquaporin-coding

Table 1. Identification of aquaporins from *Dermatophagoides farinae*

Aquaporins names	GenBank accession number	cDNA length (bp)	Deduced no. of amino acids	Deduced molecular weight (Da)	Theoretical pI	Homolog (% similarity) ^a	Forward primers in PCR	Reverse primers in PCR	Enzyme site
DerfAQP1	KY231248	780	259	28,550.50	8.50	ABI53034 (41%)	AATGGGTGCGGGATCCATGG AAAAAACATTGAAACG	ATCTCAGTGGTGGTGGTGGTG CTCGAGTTATTTCCAATGGTTACCGA	<i>Bam</i> H I/ <i>Xho</i> I
DerfAQP2	KY231249	984	327	36,127.71	6.37	KPM02745 (57%)	AATGGGTGCGGGATCCATGC ATAATCTTTTCAAAGA	ATCTCAGTGGTGGTGGTGGTG CTCGAGTGGGGCCGC TTATTGTTGAAATTCAGTAT	<i>Bam</i> H I/ <i>Not</i> I
DerfAQP3	KY231250	1,866	621	71,039.57	8.87	KPM02970 (46%)	AATGGGTGCGGGATCCATGGAA CAAATCTGATTGA	ATCTCAGTGGTGGTGGTGGTG TGCTCGAGTGGGGCCCTAAT CATATTGAACACGTG	<i>Bam</i> H I/ <i>Not</i> I
DerfAQP4	KY231251	966	321	35,043.59	6.21	KPM07253 (70%)	AATGGGTGCGGGATCCATGGAA GCCGAAACAATAAC	ATCTCAGTGGTGGTGGTGGTG CTCGAGTCCGG CCGCTCAATGTTTTCATTATCCT	<i>Bam</i> H I/ <i>Not</i> I
DerfAQP5.01	KY231252	1,635	544	58,901.00	7.78	AMZ04828 (44%)	AATGGGTGCGGGATCCATGGCA GCATCAACATCACA	ATCTCAGTGGTGGTGGTGGTGGTG AGAAATTTGGTCCCATCA	<i>Bam</i> H I/ <i>Xho</i> I
DerfAQP5.02	KY231253	1,638	545	58,903.97	8.68	AMZ04828 (45%)	AATGGGTGCGGGATCCATGGCA GCATCAACATCACA	ATCTCAGTGGTGGTGGTGGTGGTG TCAGAAATTTGGTCCCATCA	<i>Bam</i> H I/ <i>Xho</i> I

^aBLASTp GenBank accession numbers and similarity are listed.

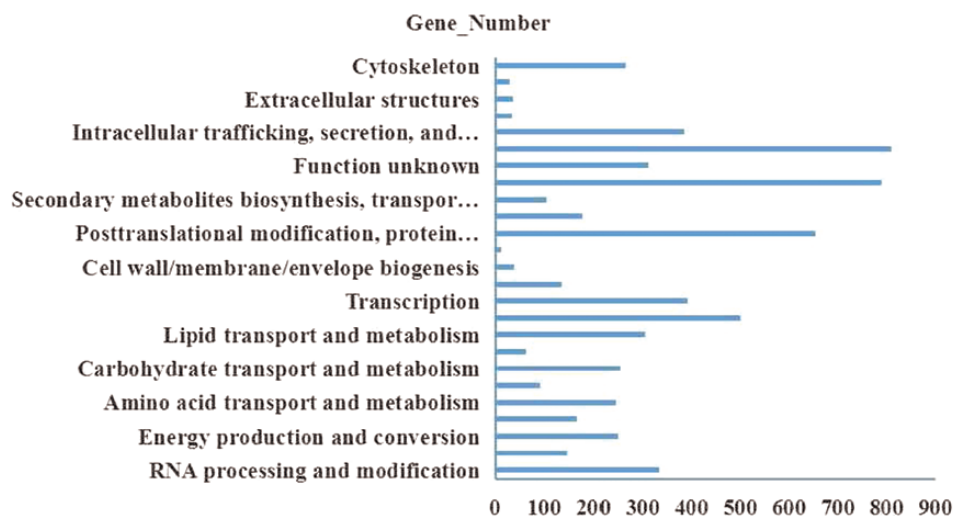


Fig. 1. COG function classification of unigenes from *Dermatophagoides farinae*.

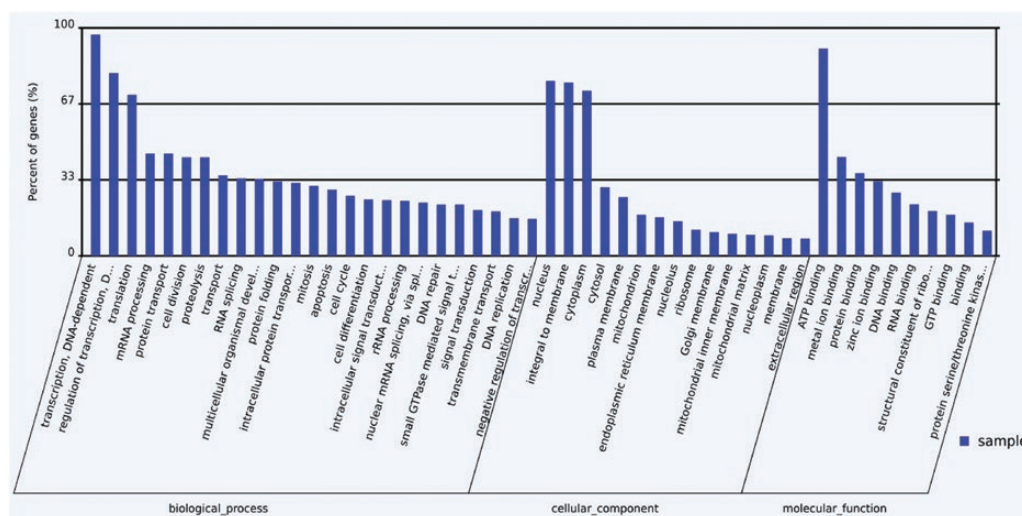


Fig. 2. GO function annotation of *Dermatophagoides farinae* unigenes.

genes were identified from BLAST alignments, two of which were complete CDS, and the remaining seven were fragments. Primers were designed using the total RNA of *D. farinae* as a template, utilizing the nine identified sequences, and RT-PCR amplification was performed to confirm these transcripts. Full-length gene sequences were obtained through RACE, and targeted primers amplified six putative aquaporin-coding genes, which were then inserted into pET28(a) for molecular cloning and sequencing validation. These full-length genes, named DerfAQP1, DerfAQP2, DerfAQP3, DerfAQP4, DerfAQP5.01, and DerfAQP5.02, were submitted to GenBank (Fig. 3A–C). DerfAQP5.01 and DerfAQP5.02 differed by just one amino acid (Fig. 3B).

Clustering analysis performed with Mega 6.0 resulted in three clusters (Fig. 4): DerfAQP1, DerfAQP2, DerfAQP4 and hAQP3, hAQP7, hAQP9, hAQP10 were clustered together; DerfAQP5.01, DerfAQP5.02 and hAQP0, hAQP1, hAQP2, hAQP4, hAQP5, hAQP6, hAQP8 were clustered; and DerfAQP3 and hAQP11, hAQP12 were clustered.

Sequence alignment was performed for DerfAQP1, DerfAQP2, DerfAQP4, and hAQP3, hAQP7, hAQP9, hAQP10 by Clustal Omega. Figure 3A also shows identification of the two NPA tag motifs, which are highly conserved motifs in aquaporins. Conserved amino acids K

(Lysine), Y (Tyrosine), Q (Glutamine) were detected at positions 15, 19, 23 of downstream of the first NPA; conserved motif 'IF-T' was detected at position 73 downstream of the first NPA. In addition, the conserved amino acids T (Threonine), D (Aspartic acid), and G (Glycine) were detected at positions 96, 108, 125 of downstream of the first NPA. According to the sequence alignment results, the second NPA motif can be extended to the tag sequence 'NPARD - PR'; at position 4 of its upstream conserved G (Glycine), and at amino acid positions 13 and 19 located at conserved G (Glycine) and F (Phenylalanine) residues.

Sequence alignment was carried out using Clustal Omega for DerfAQP5.01, DerfAQP5.02 and hAQP0, hAQP1, hAQP2, hAQP4, hAQP5, hAQP6, hAQP8 (Fig. 3B). The two NPA tag motifs of aquaporins were identified. The first NPA tag site can be extended to the conserved sequence 'SG-H-NPA', with an 'AEF' box at the -60 amino acid position, and with conserved Y (Tyrosine), G (Glycine), E (Glutamic acid), G (Glycine), L (Leucine) at downstream positions 19, 26, 68, 72, 100, 108. For the second NPA, it can be extended to sequence 'G--MNPARG', with conserved 'A' at its downstream sixth amino acid position, and conserved 'HW—W-GP' motif at its downstream 16th amino acid position. Identification of these conserved sequences provides tags for discovering similar aquaporins.

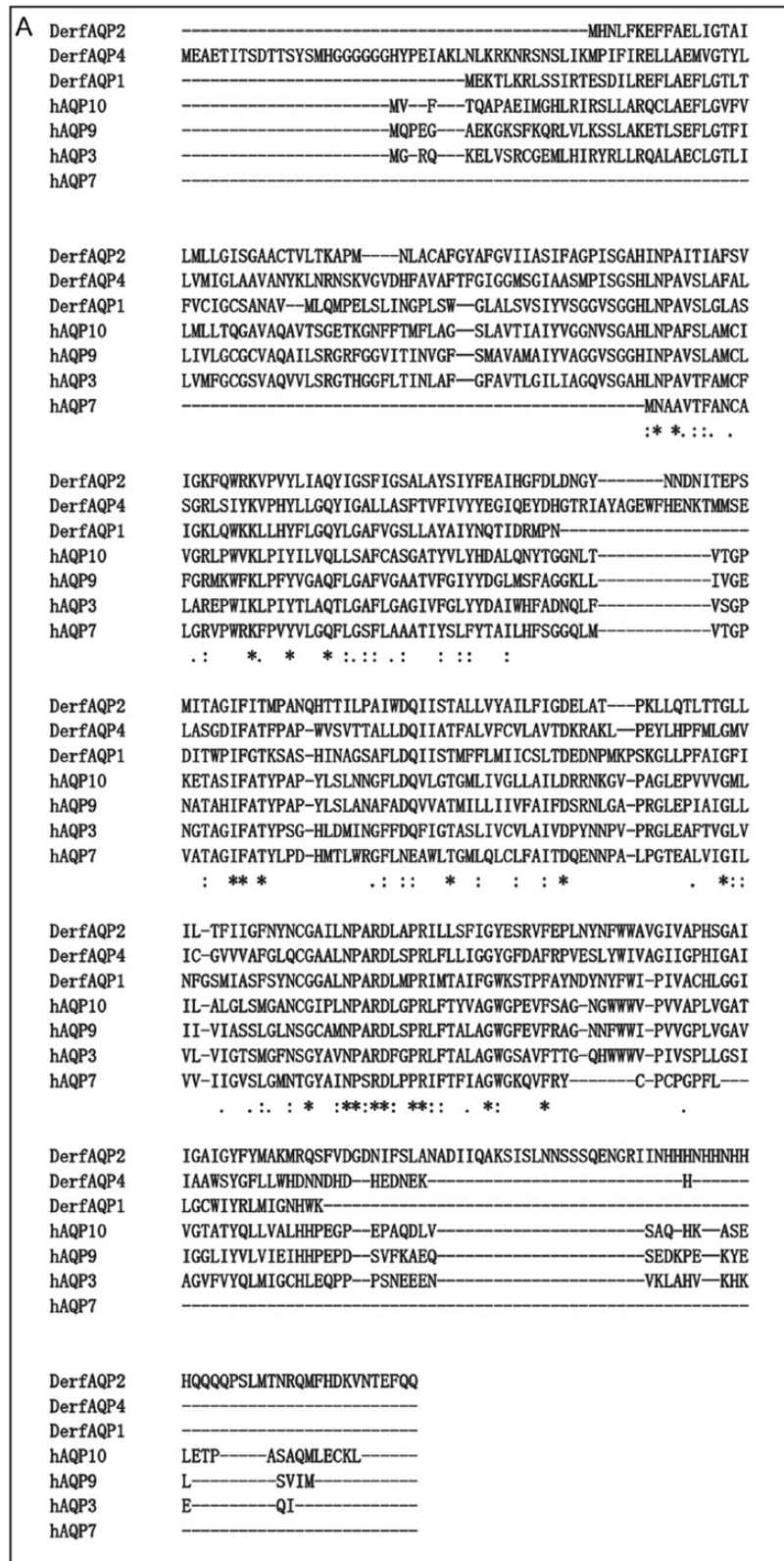


Fig. 3. Aquaporin amino acid sequence alignments. (A) Among DerfAQP1, DerfAQP2, DerfAQP4 and hAQP3, hAQP7, hAQP9, hAQP10; (B) Among DerfAQP5.01, DerfAQP5.02 and hAQP0, hAQP1, hAQP2, hAQP4, hAQP5, hAQP6, hAQP8; (C) Amino acid sequence alignment among DerfAQP3, hAQP11 and hAQP12. Alignments performed using Clustal Omega (<http://www.ebi.ac.uk/Tools/msa/clustalo/>). *Residues or nucleotides in that column are identical in all sequences in the alignment; ':' indicates conserved substitutions; ':' indicates semiconserved substitutions.

Discussion

RNA sequencing (RNA-Seq), also known as whole-transcriptome shotgun sequencing, provides a cost-effective method for performing

qualitative and quantitative analyses of gene transcripts in many nonmodel organisms, such as *D. farinae* mites, for which whole-genome sequences are not available. Through RNA-Seq we obtained

[illegible]

Fig 3. Continued

51,678,000 clean reads, and de novo transcript assembly produced 38,137 transcripts and 22,023 unigenes. These observations suggest that we have generated a representative transcriptome of *D. farinae*.

In annotation, nine aquaporin-like genes were predicted by BLASTX in the public databases SWISS-PROT, NR, KEGG, KOG, and Pfam. Of these, six full-length cDNA were produced and

C	DerfAQP3	MEQILIDDSFVSDLINHMGQKYPRILIQNDDIVKPALAPEHLCRDIFVERNDQNLVRL
	hAQP11	-----
	hAQP12	-----
	DerfAQP3	LRHYRTKGFISMLNFIKQSYHEYHRQSKQSDDNHSGKPTNSKNYGLLLYIFALIILQI
	hAQP11	-----
	hAQP12	-----
	DerfAQP3	YSIYLAIHMYFYPFDITYSDDLLNEFNLSGWKSLVRTFVWHPSSLRCALALVNDLIYV
	hAQP11	-----
	hAQP12	-----
	DerfAQP3	YSADHKRPLNHITCVKNVHQRISSMCWHPKLDNVLVACRDKILIWFESLSTNVVQSS
	hAQP11	-----
	hAQP12	-----
	DerfAQP3	PALLAQVKAESLSTRQPSLQIVNTHLQSPITSLAFNMNGNMLVSC-----
	hAQP11	-----MSPLL-GLRSELQDTCTSLGLMSVLLMGLARVVARQQLHRPVA
	hAQP12	-----MAGLNV-----
		. : :
	DerfAQP3	---SPRNASLSLRFDDKCQLAKIESIRMKQLSSLIHLPLQLTSLQWSPDGRRLALTTL
	hAQP11	HAFVLEFLATFQLCCCTHE-----
	hAQP12	--SLSFFFATFALCEARRASKALL-----
		*: * *
	DerfAQP3	RSKPICVFESIKWSYKFWTFPPEHNRQLCQTAVWLRPNGMIMLFVPKQSSKIFALTTFYDR
	hAQP11	-----LQLLS-----EQHPAHTWTL-----TLVYFFSL
	hAQP12	---PVG-----AYEVFAREAMRTLVELGPWAGDFGPDLLL-----TLLFLLFL
		: * : : :
	DerfAQP3	AEADDVGGSSDQCTLILDTRQ-----FFNQDGITSLPIIHQLC---WDIKCTK
	hAQP11	VHGLTLVGTSSNPCGVMQMMLGGMSPETGAVRLLAQLV---SALCSRYCTSAWLSGLTQ
	hAQP12	AHGVTLDGASANPTVSLQEFLEMAEQSLPGTLLKLAAGLGMQAACITMLRLCWAWELSDLH
		... : * *: : : : : : : * : : :
	DerfAQP3	LVVSFQD---PMET-----IAI---FST---SLNPGLNIQ-----
	hAQP11	YHVSERSFACKNPIRVDLLKAVITEAVCSFLFHSALLHFQEVRTKLRIHLLAALITFLVY
	hAQP12	LLQSLMAQSCSSALRTSVPHGALVEAACAFCHLTLLHLRHSPPAYSGPAVALLVTVTAY
		* :.. *
	DerfAQP3	---PFRILKPAFTSGKNVRCLSMNFHDNY---VDP---GCLITICWSNGILTHIPLE
	hAQP11	AGGSLTGAVFNPALALSLHFMCFDEAFQFFIVYWLAPSLGILLMILMFSFPLPWLHNNH
	hAQP12	TAGPFTSAFFNPALAASVTFACSGHTLLEYVQVYWLGPLTGMVLAVLLHQGRPLPHLQORN
		..*: : . . * . : : : * * : : . * : ..
	DerfAQP3	TDIDKNDRTPNESIRSLTSLCTSFDESSFVRRAGEFSLSRCSIASPLQHNGNGSIILSQT
	hAQP11	T-----INKE-----
	hAQP12	L-----FYGQKNKYRAPRG---KPAPASGDTQ---TP
		: :
	DerfAQP3	HTGKKSQRPEKTKSFGTSTPPSTPFRRSVLYSTRVQYD
	hAQP11	-----
	hAQP12	AKGSSVREPGRSGV---EGPHSS-----

Fig 3. Continued

confirmed in subsequent biological experiments that included RT-PCR and RACE, and then verified by nucleotide sequencing of the recombinant plasmids by linking the cDNA obtained with pET28(a) plasmid. These six full-length sequences received GenBank accession numbers KY231248, KY231249, KY231250, KY231251, KY231252, and KY231253, which we have named, respectively, DerfAQP1-4, DerfAQP5.01, and DerfAQP5.02. The only difference between KY231252 and KY231253 is one amino acid; thus,

it was presumed that these two reflect a polymorphism of the same aquaporin.

Phylogenetic analyses of aquaporins from *D. farinae* identified three subgroups: 1) aquaporins DerfAQP5.01 and DerfAQP5.02, which were clustered with hAQP0, -1, -2, -4, -5 -6, and -8; 2) aquaglyceroporins DerfAQP1, DerfAQP2, and DerfAQP4, which were clustered with hAQP3, -7, -9, and -10; and 3) unorthodox aquaporin DerfAQP3, which was clustered with hAQP11 and hAQP12.

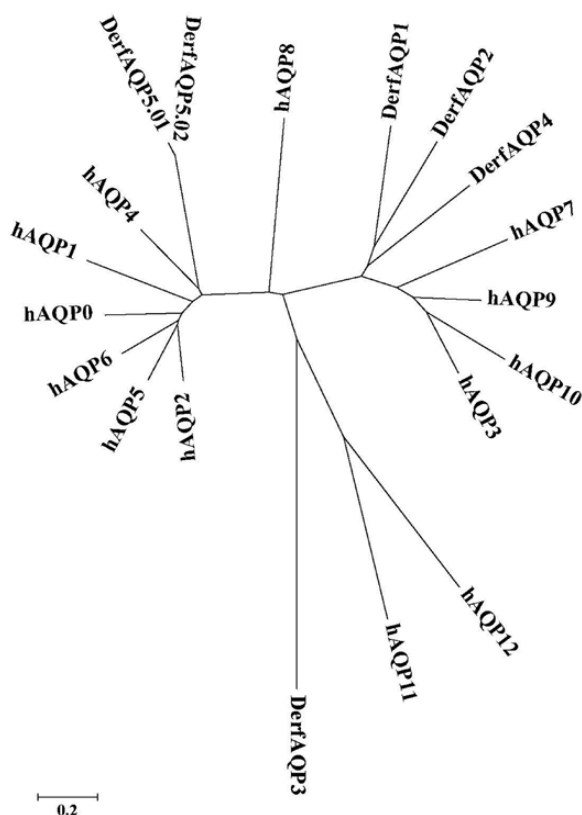


Fig. 4. Phylogenetic relationship of aquaporins from *Dermatophagoides farinae* and *Homo sapiens* (human). Tree was constructed from amino acid sequences using the neighbor-joining method of MEGA (6.0). Scale bar represents an estimate of the number of amino acid substitutions per site. Accession and database sequences identifiers are as follows: hAQP1 (AB451275), hAQP2 (AH007817), hAQP3 (BT007199), hAQP4 (BC022286), hAQP5 (AH006636), hAQP6 (NM_001652), hAQP7 (BC119672), hAQP8 (AF067797), hAQP9 (AB008775), AQP10 (BC069607), AQP11 (BC040443), AQP12 (AB040748), DerfAQP1 (KY231248), DerfAQP2 (KY231249), DerfAQP3 (KY231250), DerfAQP4 (KY231251), DerfAQP5.01 (KY231252), and DerfAQP5.02 (KY231253).

Great effort has been focused on determining the structures of AQPs that characterize their unique and specific permeability. Based on primary sequences, AQPs contain six membrane-spanning segments (TM1–6) with five connecting loops (A–E), and two highly conserved regions called NPA boxes (or motifs) with three amino acid residues (asparagine, proline, alanine: Asn-Pro-Ala). The NPA boxes have been called the ‘signature’ of AQPs. In our analyses, the two NPA motifs were found in DerfAQP1, DerfAQP2, DerfAQP4, and DerfAQP5.02. However, analysis of the MIP family database revealed considerable variation of NPA motifs, like we found in DerfAQP5.01 and DerfAQP3. Further sequence alignment revealed homologies among DerfAQP1, DerfAQP2, and DerfAQP4 and hAQP3, hAQP7, hAQP9, and hAQP10; and among DerfAQP5.01 and DerfAQP5.02 and hAQP0, hAQP1, hAQP2, hAQP4, hAQP5, hAQP6, and hAQP8. DerfAQP5.01 and DerfAQP5.02 differed by just one amino acid, which may be a polymorphic site. To demonstrate this possibility, further study is required.

By 2009, seven types of AQPs had been functionally expressed within the class Insecta, specifically BIB from *Drosophila melanogaster*, PvAQP1 and PvAQP2 from *Polypedium vanderplanki*, AQPcic from *Cicadella viridis*, AeaAQP from *Aedes aegypti*, RpMIP from *Rhodnius prolixus*, and DRIP from *D. melanogaster*

(Spring et al. 2009). Since then, AQP homologues have been found in insect genomes with the greatest coverage and fullest assembly. The genomes of each known *Drosophila* species include seven AQP homologues. In genomes of the mosquitoes *Ae. aegypti*, *Anopheles gambiae*, and those in the genus *Culex*, five have been identified; in genomes of the honey bee, *Apis mellifera*, the jewel wasp, *Naso vitripennis*, and the red flour beetle, *Tribolium castaneum*, six have been identified (Campbell et al. 2008). In 2014, Fabrick et al. cloned and characterized five new AQPs from the western tarnished plant bug, *Lygus hesperus*, a polyphagous insect pest of food and fiber crops throughout western North America (Fabrick et al. 2014). Recently, eight aquaporins (BtAQPs) were identified in the Middle East-Asia Minor 1 (MEAM1) whitefly, *Bemisia tabaci* (Gennadius). Seven of these belonged to the classical AQP4-related channels, including BIB, DRIP, Prip, and Egpls, and one belonged to the unorthodox group of AQP12-like channels (Ekert et al. 2016).

For the six gene fragments we cloned from the *D. farinae* transcriptome, we unfortunately failed to express the full-length cDNAs in *Escherichia coli* BL21 (DE3) T1R cells. Nonetheless, the identification of the full-length gene sequences confirms the presence of multiple aquaporin genes in this house dust mite species. Further, bioinformatics analyses allowed us to classify them into three paralogous groups, i.e., aquaporins, aquaglyceroporins, and superaquaporins. Some signature sequence will be helpful to explore the mite aquaporins in the future, and can lay the foundation for investigations of their relationship to human allergy.

Acknowledgments

We thank for Xiaowei Zhang, Ying Han, and Liwei Lin (TaKaRa Biotechnology Limited Company, Dalian, China) for helping us in molecular techniques; and Zhiyuan Huang (LC-Bio Co. Ltd, Hangzhou, China) for help in transcriptomic analyses. This study was supported by the National Natural Sciences Foundation of China (grant no. NSFC31272369 and NSFC81360456), the 333 project of Jiangsu Province in 2017 (BRA2017216), Medical Innovation Team of Jiangsu Province (grant no. CXTDB 2017016) and the Major Program of Wuxi health and Family Planning Commission (z201606 and Z201701), and Primary Research & Development Plan of Jiangsu Province (grant no. BE2018627).

References Cited

- Ball, A., E. M. Campbell, J. Jacob, S. Hoppler, and A. S. Bowman. 2009. Identification, functional characterization and expression patterns of a water-specific aquaporin in the brown dog tick, *Rhipicephalus sanguineus*. *Insect Biochem. Mol. Biol.* 39: 105–112.
- Campbell, E. M., A. Ball, S. Hoppler, and A. S. Bowman. 2008. Invertebrate aquaporins: a review. *J. Comp. Physiol. B.* 178: 935–955.
- Campbell, E. M., M. Burdin, S. Hoppler, and A. S. Bowman. 2010. Role of an aquaporin in the sheep tick *Ixodes ricinus*: assessment as a potential control target. *Int. J. Parasitol.* 40: 15–23.
- Cui, Y., P. Zhou, J. Peng, M. Peng, Y. Zhou, Y. Lin, and L. Liu. 2008. Cloning, sequence analysis, and expression of cDNA coding for the major house dust mite allergen, Der f 1, in *Escherichia coli*. *Braz. J. Med. Biol. Res.* 41: 380–388.
- Cui, Y. B., H. X. Cai, Y. Zhou, C. X. Gao, W. H. Shi, M. Yu, and L. Li. 2010. Cloning, expression, and characterization of Der f 7, an allergen of *Dermatophagoides farinae* from China. *J. Med. Entomol.* 47: 868–876.
- Cui, Y., Y. Zhou, W. Shi, G. Ma, L. Yang, Y. Wang, and L. Li. 2012a. Molecular cloning, expression, sequence analyses of dust mite allergen Der f 6 and its IgE-binding reactivity with mite allergic asthma patients in southeast China. *Mol. Biol. Rep.* 39: 961–968.
- Cui, Y., Y. Zhou, G. Ma, L. Yang, Y. Wang, and W. Shi. 2012b. Cloning, bioinformatics analysis, and expression of the dust mite allergen Der f 5 of *Dermatophagoides farinae*. *Braz. J. Med. Biol. Res.* 45: 746–752.

- Cui, Y., Y. Zhou, Y. Wang, G. Ma, and L. Yang. 2013. The group 10 allergen of *Dermatophagoides farinae* (Acari: Pyroglyphidae): cDNA cloning, sequence analysis, and expression in *Escherichia coli* BL21. *J. Med. Entomol.* 50: 205–208.
- Cui, Y. B., Y. Zhou, N. Wang, F. X. Teng, L. L. Yu, Y. H. Bian, J. X. Song, L. Yang, and C. B. Zhang. 2014. Expression, cloning, and IgE-binding of the full-length dust mite allergen Der f 8. *Immunol. Res.* 60: 60–68.
- Cui, Y., L. Yu, F. Teng, C. Zhang, N. Wang, L. Yang, and Y. Zhou. 2016a. Transcriptomic/proteomic identification of allergens in the mite *Tyrophagus putrescentiae*. *Allergy.* 71: 1635–1639.
- Cui, Y. B., L. L. Yu, F. X. Teng, N. Wang, Y. Zhou, L. Yang, and C. B. Zhang. 2016b. Dust mite allergen Der f 4: expression, characterization, and IgE binding in pediatric asthma. *Pediatr. Allergy Immunol.* 27: 391–397.
- Ekert, E. V., F. Chauvigné, R. N. Finn, L. G. Mathew, J. J. Hull, J. Cerdà, and J. A. Fabrick. 2016. Molecular and functional characterization of *Bemisia tabaci* aquaporins reveals the water channel diversity of hemipteran insects. *Insect Biochem. Mol. Biol.* 77: 39–51.
- Fabrick, J. A., J. Pei, J. J. Hull, and A. J. Yool. 2014. Molecular and functional characterization of multiple aquaporin water channel proteins from the western tarnished plant bug, *Lygus hesperus*. *Insect Biochem. Mol. Biol.* 45: 125–140.
- Fu, D., and M. Lu. 2007. The structural basis of water permeation and proton exclusion in aquaporins. *Mol. Membr. Biol.* 24: 366–374.
- Gomes, D., A. Agasse, P. Thiébaud, S. Delrot, H. Gerós, and F. Chaumont. 2009. Aquaporins are multifunctional water and solute transporters highly divergent in living organisms. *Biochim. Biophys. Acta.* 1788: 1213–1228.
- Grabherr, M. G., E. Mauceli, and L. J. Ma. 2011. Genome sequencing and assembly. *Methods Mol. Biol.* 722: 1–9.
- Guerrero, F. D., R. Andreotti, K. G. Bendele, R. C. Cunha, R. J. Miller, K. Yeater, and A. A. P. D. León. 2014. *Rhipicephalus (Boophilus) microplus* aquaporin as an effective vaccine antigen to protect against cattle tick infestations. *Parasit. Vectors.* 7: 1–12.
- Holmes, S. P., D. Li, S. M. Ceraul, and A. F. Azad. 2008. An aquaporin-like protein from the ovaries and gut of American dog tick (Acari: Ixodidae). *J. Med. Entomol.* 45: 68–74.
- Kozono, D., M. Yasui, L. S. King, and P. Agre. 2002. Aquaporin water channels: atomic structure molecular dynamics meet clinical medicine. *J. Clin. Invest.* 109: 1395–1399.
- Maurel, C., L. Verdoucq, D. T. Luu, and V. Santoni. 2008. Plant aquaporins: membrane channels with multiple integrated functions. *Annu. Rev. Plant Biol.* 59: 595–624.
- Spring, J. H., S. R. Robichaux, and J. A. Hamlin. 2009. The role of aquaporins in excretion in insects. *J. Exp. Biol.* 212: 358–362.