

Expression and functionality of allergenic genes regulated by simulated gastric juice in *Anisakis pegreffii*

Xu Wang, Haiyan Jia, Haiyan Gong, Yan Zhang, Rongsheng Mi, Yehua Zhang, Xianghan Han, Zhaoguo Chen^{*}

Key Laboratory of Animal Parasitology of Ministry of Agriculture, Laboratory of Quality and Safety Risk Assessment for Animal Products on Biohazards (Shanghai) of Ministry of Agriculture, Shanghai Veterinary Research Institute, Chinese Academy of Agricultural Sciences, Shanghai, China

ARTICLE INFO

Keywords:

Anisakis pegreffii
Allergenic genes
Transcriptomic analysis
Gastric juice

ABSTRACT

To better understand the molecular mechanisms underlying allergens and parasite immunity and discover the stage-enriched gene expression of fish-borne zoonotic nematodes in the stomach, we used RNA-seq to study the transcriptome profiles of *Anisakis pegreffii* (Nematoda: Anisakidae, AP) in simulated gastric juice. Mobile L3 larvae were incubated in simulated medium at 37 °C in 5% CO₂ (AP-GJ) and the control group larvae were collected in PBS under the same conditions (AP-PBS). We found that the sequences of *A. pegreffii* were highly similar to *Toxocara canis* sequences. Among the transcripts, there would be 138 up-regulated putative genes and 251 down-regulated putative genes in AP-GJ group. Several lipid binding-related genes were more highly expressed in AP-GJ larvae. Moreover, 17 allergen genes were up-regulated and 29 were down-regulated in AP-GJ larvae. Eleven allergen genes belonged to one or more of the following three categories: biological process, cellular component, and molecular function. According to KEGG analysis, the main pathways that were represented included protein processing in transcription, immune system, cancer, and infectious disease. In particular, the most significant changes in the expression of parasite-derived allergen products occurred in AP-GJ larvae. This study helps us to extend our understanding of the biology of the fish-borne zoonotic parasite *A. pegreffii* and could be helpful for more precise risk assessment and providing guidelines for allergic consumers.

1. Introduction

Food safety is an important issue in every food industry process. In the last decade, many injuries and deaths were caused by food safety accidents with agricultural products, like seafood, husbandry products, and vegetables [1]. Foodborne diseases caused by parasites transmitted by fish and shellfish products pose a major public health problem [2]. Some of these parasites, such as *Anisakis* spp., are highly pathogenic and human infection is a result of the consumption of raw or undercooked parasite-infected fish [3]. Anisakiasis is caused by *Anisakis* spp. and is emerging as an important epidemiological problem. More than 20,000 cases of anisakiasis have been reported worldwide since 1960 [4].

Anisakis is a genus of nematodes that parasitizes marine mammals, fishes, shellfish, mollusks, and crustaceans. Free-swimming larvae are devoured by crustaceans and they mature into infective third-stage larvae (L3). At last, the infected crustaceans are eaten by fish and squid and the infective L3 are retained in the muscle tissues of fish [5].

Importantly, *Anisakis* spp. infection is well recognized by the Official Journal of the European Union since humans can be infected by third-stage *Anisakis* larvae through consumption of fish that were parasitized with live L3 [6]. Importantly, the penetrative ability of the L3 could play an important role in their pathogenicity through invasion of the gastrointestinal walls of patients. Gastric juice acid has generally been accepted as one of the drivers that influences marine parasites' biological functions and rates of adaptation [7,8]. The nematode life cycle is also affected indirectly by pH via other components of the parasitic environment [9,10]. Additionally, in experimental terrestrial mammals, decreased gastric secretion accelerates *Anisakis* larvae's invasion of the gastric mucosa of *Anisakis* larvae and hypoacidity has been reported in human cases of anisakiasis in Japan [11]. Previous studies have shown that *A. simplex* and *A. pegreffii* could survive in gastric juice for 5 days and described some phenotypic changes in such an environment [12]. This means that an alternative mechanism is utilized in *Anisakis* larvae that are stimulated by gastric juice.

^{*} Corresponding author.

E-mail address: zhaoguochen@shvri.ac.cn (Z. Chen).

<https://doi.org/10.1016/j.parint.2020.102223>

Received 17 March 2020; Received in revised form 31 August 2020; Accepted 13 September 2020

Available online 30 October 2020

1383-5769/© 2020 Elsevier B.V. All rights reserved.

The potential harm of *Anisakis* spp. to humans lies not only in the mechanical injury to the stomach and intestines, but also in the threat of allergens. Individuals infected by *Anisakis* spp. are at an increased risk of becoming sensitized to parasite allergens, leading to the onset of allergic anisakiasis. Additionally, nematodes present additional allergens (which were identified by proteomics) that pose an additional hazard to humans. Previous studies only generated lists of putative parasite allergens based on comparative analyses of next-generation sequencing (NGS) and bioinformatics. To date, a total of 14 *A. simplex* allergens have been identified [13,14]. Most *Anisakis* allergens have been detected in the excretory/secretory (ES) products of the parasite, such as Ani s 1, Ani s 7, and Ani s 9 [15,16]. These allergens, as targets of IgE antibodies, are a useful tool for diagnosing hypersensitivity. Meanwhile, it is generally thought that mRNA transcripts may potentially signal a potential hazard posed by an organism and these may change due to the local gastric juice conditions [17,18]. Thus, the genes involved in adaptation to putative parasite allergens, parasite development, and host immune response can be differentially expressed under distinct gastric juice conditions. However, thus far, the expression profiles of the genes that encode potential allergens in third-stage *Anisakis* larvae in response to gastric juice acid have not been investigated.

In this study, we completed transcriptomic analyses of *A. pegreffii*, the most commonly distributed *Anisakis* species worldwide and in China. The transcriptomes of third-stage *A. pegreffii* larvae induced in simulated gastric juice and PBS (as a control) conditions were obtained by NGS and de novo assembly was conducted. Comparative analyses were conducted to provide detailed information on the different levels of gene expression induced by gastric juice. The implications of these findings for understanding the allergens of this nematode are also discussed in this paper.

2. Materials and methods

2.1. Parasite preparation

L3 of *Anisakis* were collected between June and August 2019 in Shanghai, China. *Pseudosciaena polyactis* were purchased from aquatic product consume markets and were sent to the laboratory; L3 were extracted from all these fish following published procedures [19,20]. L3 larvae were removed from the *P. polyactis* with forceps and scissors, cleaned by successive washings with water or PBS.

2.2. Stimulation in artificial gastric juice

Stimulation of *Anisakis* L3 in artificial gastric juice (Group: GJ) was prepared as described by [12]. The medium consisted of 0.1% pepsin (proteolytic activity 1:10,000 NF (Lianshuo Co., Shanghai, China), 0.1% porcine stomach mucin (Sigma-Aldrich, St. Louis, USA), 0.02% KCl, 0.12% NaCl and 2% HCl. The pH was adjusted to 1.8 with 5 M HCl. Mobile larvae after washing (10 L3/culture dish, 6 dishes) were incubated in 20 mL of this medium at 37 °C in 5% CO₂ for 24 h. Apart from the surviving *Anisakis* L3, 60 treated larvae under PBS at the same condition were used as a control (Group: PBS).

2.3. Taxonomic identification

After acidic stress stimulation, individual *Anisakis* spp. L3 stage was placed in one Eppendorf tube after previously having a one-third part of the caudal region removed to allow molecular identification of each individual [21]. Immediately, other two-third part of each larvae was stored to ultra-low temperature freezer at -80 °C for RNA extraction until required. Genomic DNA identification of each one-third part was performed employing Marine Animal Genomic DNA Purification Kit (Shanghai Jinke, China), according to manufacturer's instructions. DNA quality and quantity was checked in a spectrophotometer Nanodrop® ND-2000 (Thermo Scientific, USA). The entire ITS1 region (ITS-1, 5.8S

rDNA gene and ITS-2) was amplified using the forward primer 5'-GTCAATTCGTAGGTGAACCTGCCGAAG GATCA-3' and reverse primer 5'-GCCGGATCC GAATCCTGGTTAGTTCTTTCT -3 [22]. PCR reactions were performed in a total volume of 50 µL containing 2 µL of genomic DNA (200 ng), 0.6 M primers, 25 µL Premix Taq™ Version 2.0 (Takara, Japan). PCRs were carried out in a T100 Thermal Cycler (Bio-rad, USA). The cycling protocol was 10 min at 95 °C, 30 cycles of 30 s at 95 °C, 30 s at 55 °C, and 75 s at 72 °C, followed by 7 min at 72 °C. Subsequently, molecular identification was carried out for each individual using PCR-RFLP [21]. Amplified DNA fragments were digested with the restriction enzymes *HhaI* and *Hinf I* (TaKaRa, Japan) following manufacturer's instructions. Restriction fragments were separated by electrophoresis in Tris-Acetate-EDTA (TAE) buffered 1.2% agarose gel, stained with GelGreen and visualized with UV illumination.

2.4. Library construction and Illumina HiSeq4000 sequencing

RNA was extracted from each *Anisakis* (120 L3s all sample) and using the Trizol reagent (Invitrogen, Life Technologies, USA), and DNA setreated using Turbo DNA-free (Ambion, USA) according to the manufacturer's instructions. To sequence each sample, a minimal of 5 µg total RNA was used for sequencing library construction. Strand-specific RNA-seq transcriptome libraries were prepared using TruSeq Stranded mRNA Sample Prep Kit from Illumina (San Diego, CA, USA). rRNA was removed using Ribo-Zero Magnetic kit (Epicentre, Madison, WI). The ribosomal-free RNA was fragmented using fragmentation buffer. First strand cDNA synthesis and second strand cDNA synthesis were performed according to Illumina's TruSeq RNAseq kit protocol. When synthesizing second strand cDNA, dUTP was incorporated in place of dTTP to generate ds cDNA [23]. DNA that has adapter molecules on both ends was selectively enriched by PCR. The quality and quantity of DNA library templates were analyzed by Agilent 2100 Bioanalyser and Illumina HiSeq4000 sequencing (100 bp*2), respectively.

2.5. Reference mapping

The paired-end raw reads were trimmed and quality controlled by SeqPrep (<https://github.com/jstjohn/SeqPrep>) and Sickle (<https://github.com/najoshi/sickle>) using default parameters. Then clean data from the samples (*Anisakis*) were used to do de novo assembly with Trinity (<http://trinityrnaseq.sourceforge.net/>) [24]. All the assembled transcripts were searched against the NCBI protein nonredundant (NR) using BLASTX to identify the proteins that had the highest sequence similarity with the given transcripts to retrieve their function annotations and a typical cut-off *E*-values less than 1.0×10^{-5} was set.

2.6. Differential expression analysis and gene network analysis of DEGs

RSEM (<http://deweylab.biostat.wisc.edu/rsem/>) [25] was used to quantify gene and isoform abundances. The trimmed mean of M-values (TMM) method was selected to compute normalization factors and the differentially expressed genes (DEGs) between two samples using the following criteria: 1) the logarithmic of fold change was greater than 2 and 2) the FDR should be less than 0.05. To understand the functions of the differentially expressed genes, Gene Ontology (GO) functional enrichment and Kyoto encyclopedia of genes genomes (KEGG) pathway analysis were carried out by Goatools (<https://github.com/tanghaibao/Goatools>) and KOBAS (<http://kobas.cbi.pku.edu.cn/home.do>) [26]. Differentially expressed genes (DEGs) were significantly enriched in GO terms and metabolic pathways when their Bonferroni-corrected *P*-value was less than 0.05. Putative homologues in other representatives of Clade III nematodes (i.e. *Toxocara canis* and *Brugia malayi*) [27] [28], other nematodes and organisms were also identified (e-value cut-off: 1e-05). Gene network analysis for DEGs of AP-GJ and AP-PBS was conducted based on the available of *Toxocara canis* in the GeneMANIA cytoscape plugin [29]. Comprehensive information on the interactions

of these DEGs was then downloaded for constructing the gene networks using the Cytoscape software (version 3.6.1) [30]. Sequence trimming, assembly and functional annotation analysis were performed using the free online platform of Majorbio I-Sanger Cloud Platform (www.i-sanger.com).

2.7. Identification and alignment of putative allergens

With the aim to also assess differential expression of genes encoding for putative allergens read counts of transcript clusters with annotation of GO term and KEGG pathway were exported to a new count file through online analyzing. Transcripts with at least a two-fold difference in expression ($|\log_2 \text{fold change}| > 1$) and a FDR < 0.05 were considered to be differentially expressed between two groups by DESeq. [31]. All above putative allergens from predicted peptides from each of these species were compared with sequence data currently available in the AllergenOnline database (<http://www.allergenonline.com/>) using the BLASTp algorithm (e-value cut-off: 1e-05; 70% identity match).

2.8. Quantitative RT-PCR (qRT-PCR) on gene expressions

Total RNAs from *Anisakis* without treatment (PBS) and from those treated with artificial gastric juice (GJ) overnight were used for qRT-PCR analysis. 5 µg of total RNA was reverse transcribed using PrimeScript™ II 1st Strand cDNA Synthesis Kit (Takara, Japan) and quantified with a TB Green™ Premix Ex Taq™ II (Takara, Japan). GAPDH as a reference was developed using primers (F: 5'-CCCCTTCATCAACATC-GACT-3') and (R: 5'-TCAGCTCCCATTTGATTTC-3') (Chen et al. 2014). The primers of putative allergens genes were shown in Supplementary Table S1. The relative expression of each gene was calculated and normalized using the $2^{-\Delta\Delta Ct}$ method. Calculation of log fold-change (logFC) for all genes was conducted. Each assay was performed in triplicate.

3. Results

3.1. Collection of *Anisakis nematodes* from *P. polyactis*

L3 were collected from the fish skin and viscera of *P. polyactis*. Based on the morphological examination and molecular identification, all L3 were assigned to *A. pegreffii* as described by Amelio [21]. The ITS-1 and ITS-2 sequence data and restriction fragment length polymorphism results confirmed that all these specimens were *A. pegreffii* (data not shown).

3.2. Summary of RNA-seq datasets

To obtain an overview of the transcriptome of *A. pegreffii* induced by simulated gastric juice (AP-GJ group) and PBS (AP-PBS group) conditions, the RNA samples were prepared and mRNA samples were subjected to high-throughput Illumina HiSeq 4000 sequencing. A total of 45,342,740 and 63,218,390 raw reads were obtained from the AP-GJ and AP-PBS groups, respectively. After a check for read quality and contamination removal, a reasonable proportion (70%) of high-quality clean reads (44,839,308 and 62,499,704 for the test and control groups, respectively) remained for assembly and further analysis (Table 1). The high-quality reads were submitted to the NCBI database (AP-GJ: SRX 7132391, AP-PBS: SRX 7132390).

The unigenes were taken and subjected to sequence splicing. Sequence clustering software was used to acquire non-redundant unigenes of maximal length. With paired-end reads, we produced 38,295 unigenes with an average length of 957.3 bp in the AP-PBS control group (Table 1). The length distribution of all the unigenes is presented in Fig. 1. About 57% of these unigenes were longer than 500 bp. To annotate these unigenes, we first searched the reference sequences using BLASTX against the non-redundant (NR) NCBI protein database with a

Table 1

Transcriptome sequencing and de novo assembly.

	AP_GJ	AP_PBS
Sequencing output		
No. of raw reads	45,342,740	63,218,390
No. of clean reads	44,839,308	62,499,704
GC content (%)	43.93	43.53
Mapped ratio	0.7555	0.7297
Transcript Assembly		
Total transcripts num	55,878	
Total unigenes num	38,295	
Total sequence base	53,491,985	
Largest	14,863	
Smallest	201	
Average length	957.3	
N50	1730	
E90N50	2210	
GC percent	39.9	
Mean mapped reads	830.4667456	
Annotation (percent)		
No. of predicted peptides with matches in Nr	13,234(0.3456)	
Swiss-Prot	9748(0.2546)	
Pfam	9863(0.2576)	
COG	2220(0.058)	
GO	10,527(0.2749)	
KEGG	7440(0.1943)	
Total no. of annotated predicted peptides	13,971(0.3648)	
Total	38,295(1)	

cut-off *E*-value of 10^{-5} . After functional annotations by BLASTX and Pfam, the putative functions of 11,697 sequences (88.4%) of 13,234 unigene sequences were annotated and 1488 sequences (11.6%) were found to be non-homologous sequences. The species distribution of the best-match result for each sequence is shown in Fig. 2. The *A. pegreffii* sequences showed 72.15% similarity with *Toxocara canis* sequences, followed by *Brugia malayi* (4.10%), *Loa loa* (3.91%), *Onchocerca flexuosa* (1.64%), and *Diploscapter pachys* (1.21%).

3.3. Putative novel *Anisakis* allergens

Comparative analyses of the predicted AP peptides with sequence data are available in the AllergenOnline Database (<http://www.allergenonline.com>) (Supplementary Table S2). Our results show a total number of 46 matches, including 28 sequences that matched previously known *Anisakis* allergens (Supplementary Table S2). Eighteen sequences encoding UA3-recognized allergens matched homologous allergenic proteins in *A. simplex* (AS). Other comparisons between sets of putative allergens revealed that *Ancylostoma* secreted protein, gut esterase, Ani s 1, and DVA-1 polypeptide matched sequences that were previously characterized in *Toxocara canis*. Two sequences, which encoded the Ani s 12 and Ani s 2 allergens, were unique to AP in the AllergenOnline database. Other putative novel allergens matched homologous allergenic proteins in arthropods (AP = 5) and helminthes (AP = 4). In the present study, gut esterase and carboxylesterase in AP were identified in *Anisakis* for the first time. An overview of the GO annotation and KEGG pathway analysis information for the whole transcriptomes of the putative novel allergens are shown in (Supplementary Fig. S1, S2).

3.4. Differences in gene expression and associated functions between AP-GJ and AP-PBS revealed by DEG analysis

DEG analysis was conducted to analyze differences in the gene expression profiles of the AP-GJ and AP-PBS groups. When DEGs were manually extracted under the categories mentioned above and filtered by *p*-value (*p*-value < 0.05), 138 putative genes and 251 putative genes were significantly up-regulated and down-regulated in AP-GJ, respectively. (Supplementary Table S3). The up-regulated genes that encoded putative allergens in the AP-GJ and AP-PBS groups are listed in Table 2.

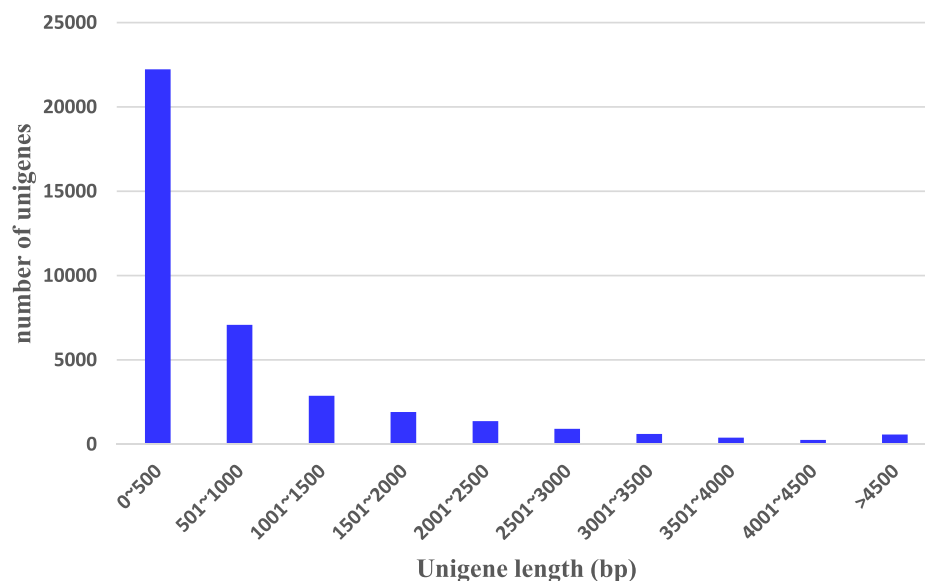


Fig. 1. Sequence length distribution of unigenes.

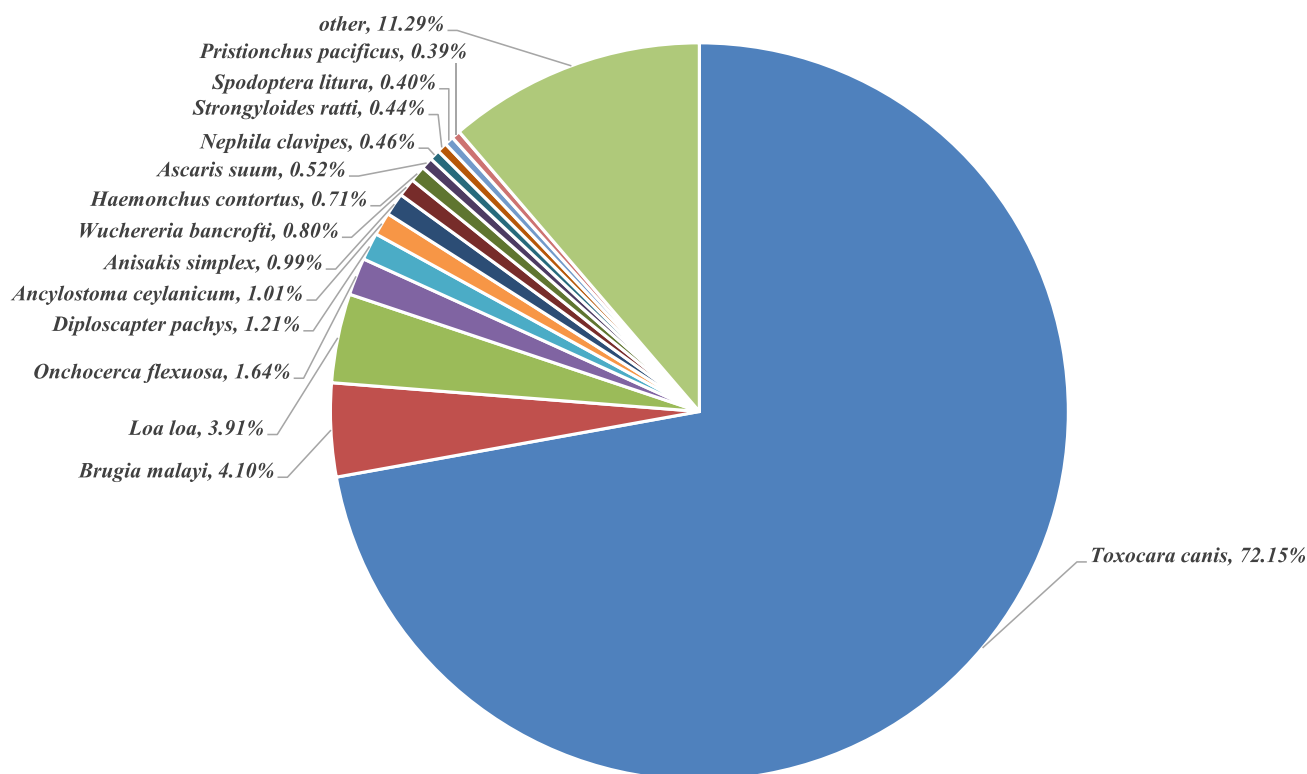


Fig. 2. Pie Graph showing the number of homologues among *A. pegreffii* and other similar species.

This figure shows the species distribution of unigene BLASTX results against the NR protein database with a cutoff E value $<10^{-5}$ and the proportions of each species. Different colors represent different species. Species with proportions of more than 1% are shown.

Two secretory-related genes and one allergen-related gene were highly expressed in AP-GJ larvae, whereas another seven genes were highly expressed in AP-PBS larvae. We found that there was one new gene (TRINITY_DN20534) that had no match in NR.

3.5. Gene network analysis

The networks of DEGs in the AP-PBS and AP-GJ groups were constructed using GeneMANIA Cytoscape (Fig. 3). Overall, 389 DEGs (138

up-regulated putative genes in the AP-GJ group and 251 down-regulated putative genes in AP-GJ, p -value <0.05) were used and 72 DEGs were mapped in GeneMANIA. The co-expression network can be loaded in Cytoscape, which is an open-source software program for integrating biomolecular interaction networks. Nep-2 and unc-52 had the highest number of links with other genes, followed by f32a5.3, asp, col-155, flna, and col-12. Table S4 shows the top 10 genes with the highest number of links and the NR description with the corresponding network. The GO analyses of these top 10 DEGs were enriched in several

Table 2
The significant regulated genes encoded putative allergens between AP-GJ and AP-PBS.

Gene_id	FC(AP_GJ/AP_PBS)	Log2FC(AP_GJ/AP_PBS)	Significant	Regulate	Closest match in Nr
DN10201_c0_g1	2.77	1.47	yes	up	Ancylostoma secreted protein [Toxocara canis]
DN21219_c0_g2	0.31	-1.72	yes	down	ani s 4 allergen [Anisakis simplex]
DN15482_c0_g1	0.34	-1.56	yes	down	Nematode polypeptide allergen ABA-1
DN19722_c2_g2	9.63	3.27	yes	up	vespid allergen antigen homolog [Wuchereria bancrofti]
DN15299_c0_g2	0.38	-1.4	yes	down	Ani s 12 allergen [Anisakis pegreffii]
DN20534_c0_g2	0.4	-1.33	yes	down	None
DN10961_c0_g1	0.29	-1.77	yes	down	DVA-1 polypeptide [Toxocara canis]
DN11409_c0_g1	18.85	4.23	yes	up	Ani s 14 allergen [Anisakis simplex]

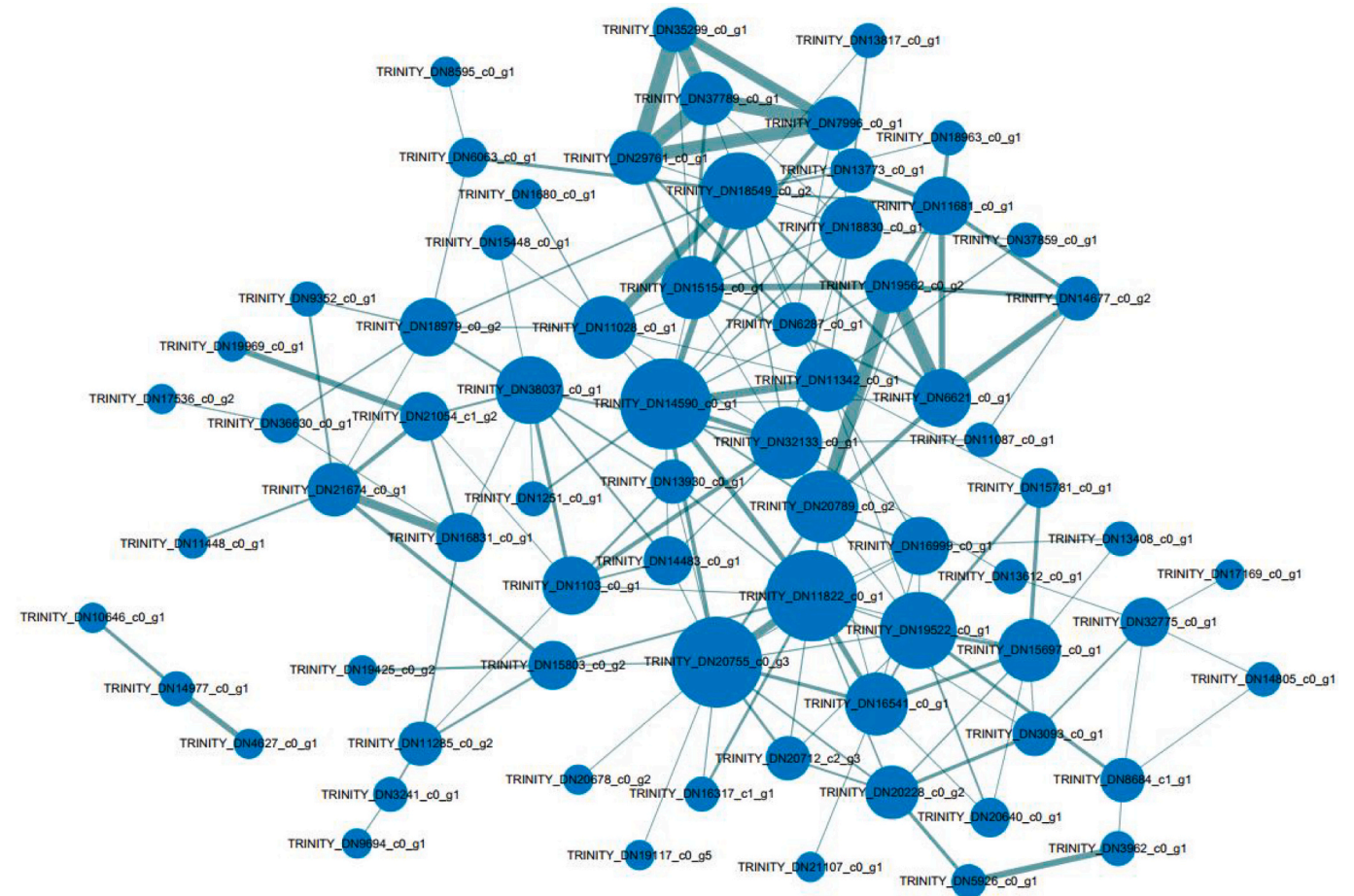


Fig. 3. Gene networks of DEGs between AP-GJ and AP-PBS significant transcriptomes. Network analysis of all the interactions connected to the identified DEGs by GeneMANIA. Node presents a gene and the different edge color presents the different relationship between the genes.

processes, such as being integral membrane components and metalloprotease activity (Supplementary Table S4). Most of these cross-linking proteins were similar to those of *Toxocara canis*.

3.6. Validation of DEGs by qRT-PCR

To validate the results of transcriptome sequencing and provide a better understanding of the nematodes' stress responses under gastric juice, we selected seven genes that encode allergens (Ani s 4 allergen (DN21219c0g2), Ani s 12 allergen (DN15299c0g2), DVA-1 polypeptide (DN10961c0g1), and Ani s 14 allergen (DN11409c0g1), ASP (DN10201c0g1), VAH (DN19722c2g2), ABA-1 (DN15482c0g1)) to examine differences in gene expression levels using qRT-PCR (Fig. 4). The results showed that except for Ani s 12 (DN15299c0g2), their expression levels were consistent with those of the transcriptome analysis. In particular, the expression levels of Ani s 4 allergen and DVA-1

polypeptide were higher in the AP-PBS group. In contrast, in the transcriptome analysis, the expression levels of Ani s 14 allergen (DN11409c0g1), Vespida allergen antigen (DN19722c2g2), and *Ancylostoma* secreted protein (TRINITY_DN10201_c0_g1) were higher in the AP-GJ group. These expression trends were in accord with those obtained by RNA-seq, suggesting that the RNA-seq data reliably reflected the gene expression alterations.

4. Discussion

Here we describe the first report of comparative de novo transcriptomic data of *A. pegreffii* in gastric juice and PBS. We obtained an average of 44,839,308 and 62,499,704 clean reads in the AP-GJ and AP-PBS groups, respectively. Additionally, the species distribution of the unigenes annotated in NR showed that *A. pegreffii* genes had the greatest match percentages with genes of *Toxocara canis*. A total number of 46

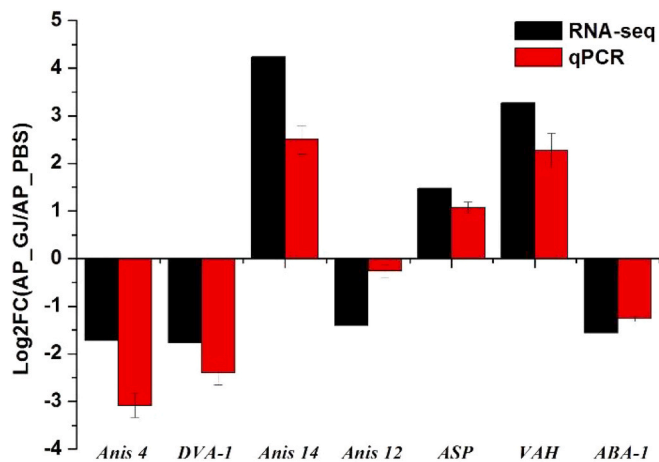


Fig. 4. Expression profiles of four candidate allergen genes as revealed by qPCR (black) and RNA-Seq (red) between AP-PBS and AP-GJ; the mean + s.e. represents three biological replicates. The relative expression implied the transcriptome data Log2FC (AP_GJ/AP_PBS). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

unigenes matched previously known *Anisakis* allergens. The DEG results indicated that 138 putative genes and 251 putative genes were significantly up-regulated and down-regulated in AP-GJ, respectively. The expression levels of three up-regulated allergens of *A. pegreffii* were significantly different between the AP-GJ and AP-PBS groups ($p < 0.05$). Remarkably, Ani s 14 allergen expression was highest in the AP-GJ group. The qRT-PCR results showed that the transcriptome sequence data of significant allergens were reliable in the gastric juice-stimulated *A. pegreffii*. *A. pegreffii* has already been reported in marine fish and as the most widespread species in the East China Sea [32,33]. Like previous studies, we found that *A. pegreffii* were widely distributed in *Pseudosciaena polyactis* in a Shanghai aquatic market (ISFH, 2018). Moreover, a previous study said that *A. pegreffii*, which causes gastric anisakiasis, is considered a major biohazard in coastal cities [34,35].

In the last few years, some studies focused on the transcriptomics and genetic data of AS, which is the same threatening nematode found in marine fish [36]. In our research, the transcriptome sequences were annotated using NCBI tools with an E -value of $10e^{-5}$ as cutoff threshold and the NCBI non-redundant protein (NR). We obtained 38,295 unigenes with an average length of 957.3 bp (Table 1). Although this number of unigenes is lower than those of previous studies, the average length and number of transcripts (957.3 bp, 55,878 transcripts) were better than the average length and number of transcripts reported by Baird et al. (100 bp, 33,737 transcripts). The number of predicted peptides with matches in NR obtained in our study (13,234) was greater than described in previous work [14]. In addition, we showed that the transcripts of *A. pegreffii* could be annotated via BLAST searches against the Nr (75.5% and 72.9%) It was consistent with the finding of Baird et al. (71.7%) [36] and Cavallero et al. (75%) [37]. We indicated that the observed sequences of *A. pegreffii* were most similar to those of *Toxocara canis* (72.15%), as well as the similar results in Baird et al. (75.0%). This data suggests that *A. pegreffii* and *T. canis* could produce similar levels of harm, which is supported by the fact that *A. pegreffii* and *T. canis* infection can lead to urticaria and silent parasitism [38,39]. It needs to be mentioned that the observed sequences of *A. pegreffii* have only 0.99% match for that of *A. simplex* in our results (Fig. 2). More specifically, all unigenes assembled were compared with the NCBI non-redundant (NR) protein database using BLASTX with an E -value cutoff of 1.0×10^{-5} . The identified proteins of *A. pegreffii* had the highest sequence similarities with the given transcripts to retrieve their function annotations in NCBI non-redundant (NR) protein database. There may indeed be tendencies when comparing annotations, such as prioritizing

comparisons to closely related species or patterns. The pattern species corresponds to more information in the NR database, so there are more corresponding results extracted. The prevalence of *Toxocara* and *Anisakis* infestations was evaluated by measuring the levels of specific IgG and IgE antibodies from whole larvae of *Anisakis simplex* (Ani s 1, Ani s 3 and Ani s 7) and *Toxocara canis* [40]. In addition, a number of third-stage-enriched transcripts were identified, which provide potential clues for understanding the molecular mechanisms of parasite biology.

It is well known that mechanical injuries and allergic hazards are two potential risks to human food safety, especially in raw marine products. In previous studies, heat shock proteins 70 (HSPs 70) and enolases were identified in both the AS and AP datasets via immune-proteomic approaches [41,42]. HSPs 70 from mites, black flies, midges and cockroaches are known allergens [43]. In particular, the expression of an HSP 70 from *A. pegreffii* was increased in L4s compared with L3s. It could be linked to a response of the parasite to the heat stress that follows infection [44]. In our research, the expression of HSP 70 was not changed in AP-GJ. We speculated that it may be related to this heat stress response. Subsequently, cyclophilins and ABA-1 domain-containing proteins were identified in AP, which were inferred to be novel putative allergens. ABA-1, which is the most common repeating unit of the nematode polypeptide allergens of *Ascaris suum*, may be transported within parasitic nematodes or via their interactions with host tissues. Cyclophilins (CyP) have also been identified in a collection of parasitic nematode species. Sabine described the IgE-binding capacity and cross-reactivity of CyP from different sources, showing that CyP can be considered a novel pan-allergen family. Of all the molecules inferred as novel putative allergens, 8 sequences that encoded related proteins from other species were identified in AP (Table 2). Tropomyosin is the first allergen that has been subsequently identified in over 150 invertebrate species (including shellfish and mites) [45,46] and is responsible for exacerbated immune responses in over 30% of the world's population [47,48]. UA3-recognized O-deglycosylated specific allergen was identified as a general allergen in *Anisakis app.* by antigen-capture ELISA, according to Lorenzo et al. [49]. The antibodies recognized by UA3 are good candidates for the serodiagnosis of Anisakiasis. Other than these two general allergens, five Ani s allergens (Ani s 1, Ani s 2, Ani s 3, Ani s 5 and Ani s 14) have been the investigated to detect the allergens for *A. simplex* and *A. pegreffii* L3 [50–53]. However, a new *Anisakis* allergen in our transcriptome data, Ani s 12 (TRINITY_DN15299 and TRINITY_DN15942), has been rarely mentioned in previous studies. Ani s 12 exhibits a lower rate of reactivity with patients' sera than Ani s 1, 2, 7 and 13, but it is also one major allergen that is recognized by more than 50% of *Anisakis*-allergic patients [53]. This provides us with new ideas for the diagnosis of allergens in anisakiasis. Surprisingly, several predictive allergens, like pancreatic lipase-related protein 2-like, GloA, carboxylesterase, and minor allergen Alt a 7-like, were first found in this research. Among the putative allergens characterized in the present study, some were expressed at significantly different levels in the AP-PBS and AP-GJ groups.

Recently, *Anisakis* homologous allergen haemoglobin (Ani s 13) was recognized by a large number (64.3%) of sensitized patients and in up to 80.9% of patients with GAA (gastro-allergic anisakiasis). Importantly, anti-*A. simplex* haemoglobin IgG4 and IgE levels immunized by Ani s 13 were higher in gastro-allergic anisakiasis than chronic urticaria patients [61]. Therefore, we suggested that Ani s 13 may be the major allergen that causes gastro-allergic anisakiasis. It is plausible that the acidic environment also contributes to the onset of allergic anisakiasis. Additionally, this study is the first to report variations in predicted key allergens under artificial gastric juice conditions. Many studies demonstrated that allergens are resistant to digestion by pepsin in simulated gastric fluid and are surprisingly stable [54,55]; this means some allergens exert more risk under stimulation by gastric juice in individuals who may be susceptible to those allergens. In our study, Ani s 14 allergen, Vespider allergen antigen, and *Ancylostoma* secreted protein were expressed at different levels between the AP-GJ and AP-PBS groups

(Table 3, Fig. 3). Vespidae allergen antigen and *Ancylostoma* secreted protein were identified to be secretory protein [56,57]. Secreted proteins are often thought to be involved in invasion or host modulation [58]. In addition, the recombinant vespidae allergens showed antigenic cross-reactivity of their homologous proteins at B-cell levels. Then, B-cell surface receptor CD22-mediated adhesion is dependent upon the presence of sialic acids on ligands. Thus, we suggested that vespidae allergen antigen could better bind to receptor proteins in acidic environment than for example in neutral pH [59,60]. Like Ani s 12, Ani s 14 exhibits a lower level of reactivity with patients' sera than Ani s 1, 2, 7 and 13. But Ani s 14 was major allergen recognized by more than 50% of Anisakis-allergic patients, and it was the sixth major *A. simplex* allergen. In our results, higher expression levels of Ani s 14 increased the risk of allergic symptoms under gastric juice. Recombinant Ani s 14 was verified to be IgE reactive and hence could be useful as a diagnostic tool for identifying *A. simplex* allergies [53]. However, diversity of immune responses of the Ani s 14 under gastric juice was not mentioned in previous study. We will continue this work in next plan. In conclusion, these three predicted key allergens were considered for constructing a diagnostic reagent for detection of the immune responses.

5. Conclusions

In the present study, we provided a detailed overview of the transcriptomes of *A. pegreffii* larvae under simulated gastric juice and PBS conditions in vitro. Our data provided comprehensive coverage of the *A. pegreffii* transcriptome and revealed that specific genes were differentially expressed between the AP-PBS and AP-GJ groups. Furthermore, we found the sequences of *A. pegreffii* showed the highest match percentages with *Toxocara canis* sequences. Several genes related to integral components of membranes and lipid binding were highly expressed in the AP-GJ group. Particularly, the parasite-derived allergen product Ani s 14 had the most significant expression change in the AP-GJ group. This study helps us to extend our understanding of the fish-borne zoonotic parasite *A. pegreffii* and increase the valuable source of transcriptomes under gastric juice to establish control strategies for anaphylaxis due to parasitic nematodes.

Declaration of Competing Interest

The authors declare that they have no conflicts of interest.

Acknowledgements

This research was supported by the Shanghai Agriculture Applied Technology Development Program, China (G20190303), Central Public-interest Scientific Institution Basal Research Fund (2018JB07).

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.parint.2020.102223>.

References

- [1] B.A. Alimi, Risk factors in street food practices in developing countries: a review, *Food Sci. Hum. Wellness* 5 (3) (2016) 141–148.
- [2] E. Abollo, C. Gestal, A. LãPez, A. Fgonzã Lez, A. Guerra, S. Pascual, Squid as trophic bridges for parasite flow within marine ecosystems: the case of *Anisakis simplex* (Nematoda: Anisakidae), or when the wrong way can be right, *S. Afr. J. Marine Sci.* 20 (1) (1998) 223–232.
- [3] C.R. Sterling, Food-Borne Nematode Infections, *Foodborne Parasites*, Springer, 2006, pp. 135–160.
- [4] H. Ruiz, D. Milanes, M. Frosini, A. Petrolini, E. Furfaro, R. Cardinale, B. Hamilton, K. De Bruyn, E. Hicks, J. Wang, Acute small-bowel obstruction from intestinal anisakiasis after the ingestion of raw clams; documenting a new method of marine-to-human parasitic transmission, *Oxford Univ. Press* 1 (2) (2014).
- [5] P.H.V. Thiel, F.C. Kuipers, R.T. Roskam, A nematode parasitic to herring, causing acute abdominal syndromes in man, *Trop. Geogr. Med.* 12 (1960) 97–113.
- [6] M. Bao, G.J. Pierce, N.J. Strachan, S. Pascual, M. González-Muñoz, A. Levsen, Human health, legislative and socioeconomic issues caused by the fish-borne zoonotic parasite Anisakis: challenges in risk assessment, *Trends Food Sci. Technol.* 86 (2019) 298–310.
- [7] R.A. Giannella, A.B. Selwyn, Z. Norman, Influence of gastric acidity on bacterial and parasitic enteric infections: a perspective, *Ann. Intern. Med.* 78 (2) (1973) 271–276.
- [8] I. Moneo, M.L. Caballero, *Anisakis simplex* larvae release allergens during a short incubation in diluted acid which can be used for clinical diagnosis, *Alergología e Inmunología Clínica* 17 (2002) 201–207.
- [9] N.F. Gray, Ecology of nematophagous fungi: Effect of soil moisture, organic matter, pH and nematode density on distribution, *Soil Biol. Biochem.* 17 (4) (2020) 499–507.
- [10] M.R. Samoiloff, S. Balakanich, M. Petrovich, Evidence for the two-state model of nematode behaviour, *Nature* 247 (5435) (2020) 73–74.
- [11] B.E. Matthews, Behaviour and enzyme release by *Anisakis* sp. larvae (Nematoda: Ascarididae), *J. Helminthol.* 56 (3) (1982) 177–183.
- [12] N. Arizono, M. Yamada, T. Tegoshi, M. Yoshikawa, *Anisakis simplex sensu stricto* and *Anisakis pegreffii*: biological characteristics and pathogenetic potential in human Anisakiasis, *Foodborne Pathog. Dis.* 9 (6) (2012) 517–521.
- [13] F.J. Baird, R.B. Gasser, A. Jabbar, A.L. Lopata, Foodborne anisakiasis and allergy, *Mol. Cell. Probe.* 28 (4) (2020) 167–174.
- [14] C. Llorens, S.C. Arcos, L. Robertson, R. Ramos, R. Futami, B. Soriano, S. Ciordia, M. Careche, M. González-Muñoz, Y. Jiménez-Ruiz, Functional insights into the infective larval stage of *Anisakis simplex* s.s., *Anisakis pegreffii* and their hybrids based on gene expression patterns, *BMC Genomics* 19 (1) (2018) 592.
- [15] A.M. Anadón, F. Romarís, M. Escalante, E. Rodríguez, F.M. Ubeira, The *Anisakis simplex* Ani s 7 major allergen as an indicator of true *Anisakis* infections, *Clin. Exp. Immunol.* 156 (3) (2009) 471–478.
- [16] R. Rodríguez-Pérez, I. Moneo, A. Rodríguez-Mahillo, M.L. Caballero, Cloning and expression of Ani s 9, a new *Anisakis simplex* allergen, *Mol. Biochem. Parasit.* 159 (2) (2020) 92–97.
- [17] S. Hu, Y. Yu, D. Zhou, R. Li, X. Xiao, H. Wu, Global transcriptomic acid tolerance response in *Salmonella enteritidis*, *LWT* 92 (2020) 330–338.
- [18] R.W. Li, R.K. Choudhary, A.V. Capuco, J.F. Urban Jr., Exploring the host transcriptome for mechanisms underlying protective immunity and resistance to nematode infections in ruminants, *Vet. Parasitol.* 190 (1–2) (2012) 1–11.
- [19] I. Iglesias, A. Valero, F.J. Adroher, Some factors which influence the in vitro maintenance of *Anisakis simplex* (Nematoda), *Folia Parasit.* 44 (4) (1997) 297–301.
- [20] A. Valero, M.D.M. López-Cuello, R. Benítez, F.J. Adroher, *Anisakis* spp. in European hake, *Merluccius merluccius* (L.) from the Atlantic off north-west Africa and the Mediterranean off southern Spain, *Acta. Parasitol.* 51 (3) (2020) 209–212.
- [21] S. D'Amelio, K.D. Mathiopoulos, O. Brandonisio, G. Lucarelli, L. Paggi, Diagnosis of a case of gastric anisakidosis by PCR-based restriction fragment length polymorphism analysis, *Parassitologia* 41 (4) (1999) 591–593.
- [22] J.P. Bachellerie, L.H. Qu, Ribosomal RNA Probes for Detection and Identification of Species, *Protocols in Molecular Parasitology* 21, Humana Press, 1993, pp. 249–263.
- [23] D. Parkhomchuk, T. Borodina, V. Amstislavskiy, M. Banaru, L. Hallen, S. Krobitsch, H. Lehrach, A. Soldatov, Transcriptome analysis by strand-specific sequencing of complementary DNA, *Nucleic Acids Res.* 37 (18) (2009) e123.
- [24] M.G. Grabherr, B.J. Haas, M. Yassour, J.Z. Levin, D.A. Thompson, I. Amit, X. Adiconis, L. Fan, R. Raychowdhury, Q. Zeng, Full-length transcriptome assembly from RNA-Seq data without a reference genome, *Nat. Biotechnol.* 29 (7) (2011) 644–652.
- [25] B. Li, C.N. Dewey, RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome, *BMC bioinformatics* 12 (1) (2011) 323.
- [26] C. Xie, X. Mao, J. Huang, Y. Ding, J. Wu, S. Dong, L. Kong, G. Gao, C.Y. Li, L. Wei, KOBAS 2.0: a web server for annotation and identification of enriched pathways and diseases, *Nucleic Acids Res.* 39 (Web Server issue) (2011) W316–22.
- [27] E. Ghedin, S. Wang, D. Spiro, E. Caler, Q. Zhao, J. Crabtree, J.E. Allen, A. L. Delcher, D.B. Guilian, D. Miranda-Saavedra, Draft genome of the filarial nematode parasite *Brugia malayi*, *Science* 317 (5845) (2007) 1756–1760.
- [28] X.-Q. Zhu, P.K. Korhonen, H. Cai, N.D. Young, P. Nejsun, G. von Samson-Himmelstjerna, P.R. Boag, P. Tan, Q. Li, J. Min, Genetic blueprint of the zoonotic pathogen *Toxocara canis*, *Nat. Commun.* 6 (1) (2015) 1–8.
- [29] J. Montojo, K. Zuberi, H. Rodriguez, F. Kazi, G. Wright, S.L. Donaldson, Q. Morris, G.D. Bader, GeneMANIA Cytoscape plugin: fast gene function predictions on the desktop, *Bioinformatics* 26 (22, 2010) 2927.
- [30] P. Shannon, Cytoscape: a software environment for integrated models of biomolecular interaction networks, *Genome Res.* 13 (11) (2020) 2498–2504.
- [31] S. Audic, J.-M. Claverie, The significance of digital gene expression profiles, *Genome Res.* 7 (10) (1997) 986–995.
- [32] W. Fang, F. Liu, S. Zhang, J. Lin, S. Xu, D. Luo, *Anisakis pegreffii*: a quantitative fluorescence PCR assay for detection in situ, *Exp. Parasitol.* 127 (2) (2011) 587–592.
- [33] D. Luo, W. Fang, Study on infestation and dynamic of *Anisakis* pathogen in marine fishes from Taiwan Strait, *J. Oceanogr. Taiwan Strait* 19 (2) (2000) 218–221.
- [34] S. Mattiucci, P. Fazii, A. De Rosa, M. Paoletti, A.S. Megna, A. Glielmo, M. De Angelis, A. Costa, C. Meucci, V. Calvaruso, Anisakiasis and gastroallergic reactions associated with *Anisakis pegreffii* infection, Italy, *Emerg. Infect. Dis.* 19 (3) (2013) 496.
- [35] A. Umehara, Y. Kawakami, J. Araki, A. Uchida, Multiplex PCR for the identification of *Anisakis simplex sensu stricto*, *Anisakis pegreffii* and the other anisakid nematodes, *Parasitol. Int.* 57 (1) (2008) 49–53.

- [36] F.J. Baird, S. Xiaopei, A. Ibukun, N.M.J. S. Hiromu, O. Domenico, L.A.L. C. Cinzia, F.R. Toshio, The *Anisakis* transcriptome provides a resource for fundamental and applied studies on allergy-causing parasites, *Plos Negl. Trop. Dis.* 10 (7) (2016) e0004845.
- [37] S. Cavallero, F. Lombardo, X. Su, M. Salvemini, C. Cantacessi, S. D'Amelio, Tissue-specific transcriptomes of *Anisakis simplex* (sensu stricto) and *Anisakis pegreffii* reveal potential molecular mechanisms involved in pathogenicity, *Parasite. Vector.* 11 (1) (2018) 1–13.
- [38] Z. Jin, N. Akao, N. Ohta, Prolactin evokes lactational transmission of larvae in mice infected with *Toxocara canis*, *Parasitol. Int.* 57 (4) (2008) 495–498.
- [39] P.L. Minciullo, A. Cascio, S. Gangemi, Association between urticaria and nematode infections, *Allergy Asthma Proc.* 39 (2) (2018).
- [40] M. Viñas, I. Postigo, E. Suñén, J. Martínez, Urticaria and silent parasitism by Ascaridoidea: component-resolved diagnosis reinforces the significance of this association, *PLoS Neglect. Trop. D.* 14 (4) (2020), e0008177.
- [41] C.K. Faeste, K.R. Jonscher, M.M. Dooper, W. Egge-Jacobsen, A. Moen, A. Daschner, E. Egaas, U. Christians, Characterisation of potential novel allergens in the fish parasite *Anisakis simplex*, *EuPA Open Proteomics* 4 (2014) 140–155.
- [42] S.C. Arcos, S. Ciordia, L. Roberston, I. Zapico, Y. Jiménez-Ruiz, M. Gonzalez-Muñoz, I. Moneo, N. Carballeda-Sangiao, A. Rodríguez-Mahillo, J.P. Albar, Proteomic profiling and characterization of differential allergens in the nematodes *Anisakis simplex* sensu stricto and *A. pegreffii*, *Proteomics* 14 (12) (2014) 1547–1568.
- [43] N. Yusuf, T.H. Nasti, C.-M. Huang, B.S. Huber, T. Jaleel, H.-Y. Lin, H. Xu, C. A. Elmetts, Heat shock proteins HSP27 and HSP70 are present in the skin and are important mediators of allergic contact hypersensitivity, *J. Immunol.* 182 (1) (2009) 675–683.
- [44] H.-Y. Chen, Y.-S. Cheng, H.-H. Shih, Expression patterns and structural modelling of Hsp70 and Hsp90 in a fish-borne zoonotic nematode *Anisakis pegreffii*, *Vet. Parasitol.* 212 (3–4) (2015) 281–291.
- [45] A.L. Lopata, S.B. Lehrer, New insights into seafood allergy, *Curr. Opin. Allergy Clin. I* 9 (3) (2009) 270–277.
- [46] K.N. Shanti, B.M. Martin, S. Nagpal, D.D. Metcalfe, P. Rao, Identification of tropomyosin as the major shrimp allergen and characterization of its IgE-binding epitopes, *J. Immunol.* 151 (10, 1993) 5354–5363.
- [47] R. Bernardini, G. Mistrello, E. Novembre, D. Roncarolo, S. Zanotta, E. Lombardi, A. Cianferoni, N. Pucci, M. De Martino, A. Vierucci, Cross-reactivity between IgE-binding proteins from *Anisakis simplex* and *Dermatophagoides pteronyssinus*, *Int. J. Immunopath. Ph.* 18 (4) (2005) 671–675.
- [48] E. Johansson, M. Aponno, M. Lundberg, M. Van Hage-Hamsten, Allergenic cross-reactivity between the nematode *Anisakis simplex* and the dust mites *Acarus siro*, *Lepidoglyphus destructor*, *Tyrophagus putrescentiae*, and *Dermatophagoides pteronyssinus*, *Allergy* 56 (7) (2001) 660–666.
- [49] S. Lorenzo, R. Iglesias, J. Leiro, F.M. Ubeira, I. Ansotegui, M. García, L.F.D. Corres, Usefulness of currently available methods for the diagnosis of *Anisakis simplex* allergy, *Allergy* 55 (7) (2000) 627–633.
- [50] F. Guarneri, C. Guarneri, S. Benvenga, Cross-reactivity of *Anisakis simplex*: possible role of Ani s 2 and Ani s 3, *Int. J. Dermatol.* 46 (2) (2007) 146–150.
- [51] M.L. Caballero, I. Moneo, Specific IgE determination to Ani s 1, a major allergen from *Anisakis simplex*, is a useful tool for diagnosis, *Ann. Allergy Asthma Im.* 89 (1) (2002) 74–77.
- [52] M.L. Caballero, I. Moneo, F. Gómez-Aguado, M.T. Corcuera, I. Casado, R. Rodríguez-Pérez, Isolation of Ani s 5, an excretory–secretory and highly heat-resistant allergen useful for the diagnosis of *Anisakis* larvae sensitization, *Parasitol. Res.* 103 (5) (2008) 1231–1233.
- [53] Y. Kobayashi, S. Kakemoto, K. Shimakura, K. Shiomi, Molecular cloning and expression of a new major allergen, Ani s 14, from *Anisakis simplex*, *J. Food Hyg. Soc. Jpn.* 56 (5) (2015) 194–199.
- [54] J. Spirc, S.J. Koppelman, A. Knulst, J.A. Nordlee, S.L. Taylor, J.L. Baumert, *In vitro* digestion and characterisation of 2S albumin and digestion-resistant peptides in pecan, *Int. J. Food Sci. Tech.* 53 (6) (2018) 1566–1578.
- [55] I. Ahmed, J.J. Ma, Z.X. Li, H. Lin, L.L. Xu, L.R. Sun, S.L. Tian, Effect of tyrosinase and caffeic acid crosslinking of turbot parvalbumin on the digestibility, and release of mediators and cytokines from activated RBL-2H3 cells, *Food Chem.* 300 (2019) 125209.
- [56] J. Murray, W.F. Gregory, N. Gomez-Escobar, A.K. Atmadja, R.M. Maizels, Expression and immune recognition of *Brugia malayi* VAL-1, a homologue of vespid venom allergens and *Ancylostoma* secreted proteins, *Mol. Biochem. Parasit.* 118 (1) (2001) 89–96.
- [57] J.M. Hawdon, S. Narasimhan, P.J. Hotez, *Ancylostoma* secreted protein 2: cloning and characterization of a second member of a family of nematode secreted proteins from *Ancylostoma caninum*, *Mol. Biochem. Parasit.* 99 (2) (1999) 149–165.
- [58] J.O. Park, J. Pan, F. Möhrlen, M.O. Schupp, H. Hutter, Characterization of the astacin family of metalloproteases in *C. elegans*, *Bmc Dev. Biol.* 10 (1) (2010) 14.
- [59] T.P. King, G. Lu, M. Gonzalez, N. Qian, L. Soldatova, Yellow jacket venom allergens, hyaluronidase and phospholipase: sequence similarity and antigenic cross-reactivity with their hornet and wasp homologs and possible implications for clinical allergy, *J. Allergy Clin. Immunol.* 98 (3) (1996) 588–600.
- [60] D. Sgroi, A. Varki, S. Braesch-Andersen, I. Stamenkovic, CD22, a B cell-specific immunoglobulin superfamily member, is a Sialic acid-binding Lectin, *J. Biol. Chem.* 268 (10) (1993) 7011–7018.
- [61] J. González-Fernández, A. Daschner, N.E. Nieuwenhuizen, A.L. Lopata, C.D. Frutos, A. Valls, C. Cuéllar, Haemoglobin, a new major allergen of *Anisakis simplex*, *Int. J. Parasitol.* 45 (6) (2020) 399–407.

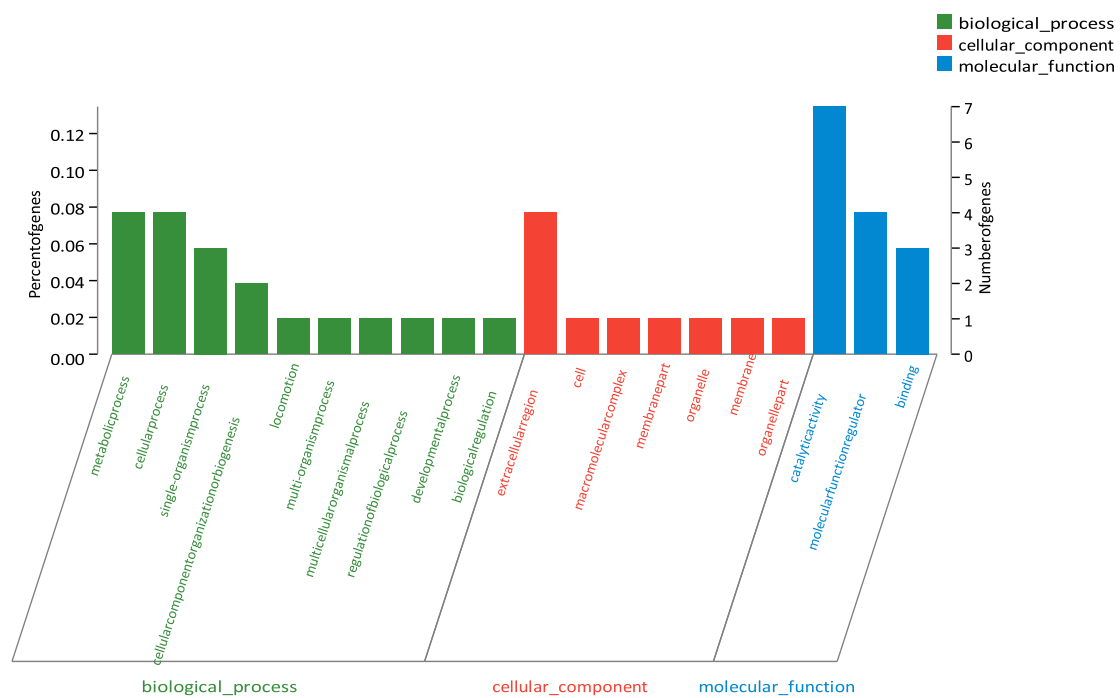


Figure S1. GO annotation analysis information for the whole transcriptomes of putative novel allergens

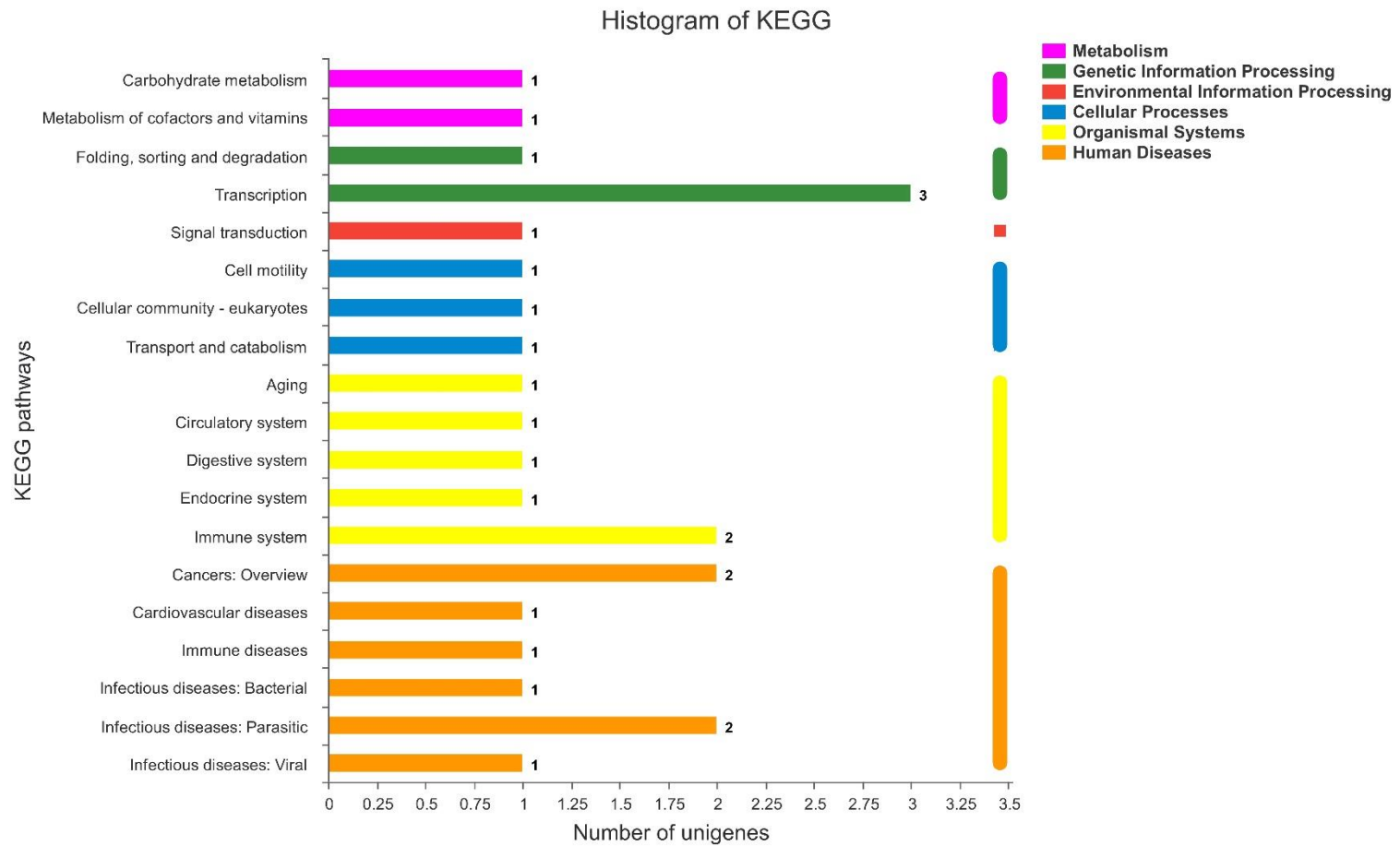


Figure S2. KEGG pathway analysis information for the whole transcriptomes of putative novel allergens

Table S1 Primers designed for putative allergens gene cloning and real time RT-PCR analysis.

Gene name	Primer Sequence (5'-3')	Amplicon length (bp)
GAPDH_RT	F: CCCCTTCATCAACATCGACT R: TCAGCTCCCCATTGATTTC	152
Ani s 14	F: AATCTGGTCTCGTGGCTCTGAAG R: AGGTACGGGATGGCTGTGGTG	115
Ani s 12	F: CCGCCATCAGTACGAGAACAATG R: GCCAATAGTAACAGGAGGTGCTC	138
Ani s 4	F:TGAACCTCGTCAAAGTGGTGTCTGC R:TGCCAAATGGTAACGGTACACAGG	117
DVA-1	F:TCTTCACCACGCCAGGACGAG R:GTCTGTCATCCAAGCGATCCACTC	132
ABA-1	F:TGTTGATGACTTCGGACCAGCATG R:ATCAGCATTGAGCCACTTCAGGTG	123
ASP	F:ACATCTTCGCTTCCTCAACAACCG R:ACTCGGACCACCAAGCCTCTG	84
VAH	F:TGTGGCTGCTGGCAATCAACTC R:ACCCTTGCGGAACCTCTCTGTCTG	139

Table S2 comparative analyses of *A. pegreffii* predicted peptides with sequence

Gene_id	Length	NR_hit-name	NR_description	Closest match in AllergenOnline
DN10201_c0_g1	1050	KHN73598.1;	Ancylostoma secreted protein [Toxocara canis];	Allergen Venom allergen 5
DN10961_c0_g1	849	KHN85858.1;	DVA-1 polyprotein [Toxocara canis];	ABA-1 allergen [Ascaris lum
DN11409_c0_g1	370	BAT62430.1;	Ani s 14 allergen [Anisakis simplex];	Ani s 14 allergen
DN13926_c0_g1	1129	ABL77410.1;	UA3-recognized allergen, partial [Anisakis simplex];	UA3-recognized allergen
DN14842_c0_g1	1574	ACZ95445.1;	Ani s 10 allergen precursor [Anisakis simplex];	Ani s 10 allergen precursor
DN15299_c0_g2	888	AGC60026.1;	Ani s 12 allergen [Anisakis pegreffii];	Ani s 12 allergen [Anisa
DN15482_c0_g1	1967	Q06811.2;	RecName: Full=Polyprotein ABA	ABA-1 allergen [Ascaris lum
DN15664_c0_g1	1415	ABL77410.1;	UA3-recognized allergen, partial [Anisakis simplex];	UA3-recognized allergen
DN15749_c0_g1	509	ACZ95445.1;	Ani s 10 allergen precursor [Anisakis simplex];	Ani s 10 allergen precur
DN15942_c0_g1	1377	BAJ78223.1;	Ani s 12 allergen precursor [Anisakis simplex];	Ani s 12 allergen [Anisa
DN16382_c1_g1	689	ABL77410.1;	UA3-recognized allergen, partial [Anisakis simplex];	UA3-recognized allergen
DN18255_c0_g1	1743	ABL77410.1;	UA3-recognized allergen, partial [Anisakis simplex];	UA3-recognized allergen
DN19246_c0_g2	4099	ABL77410.1;	UA3-recognized allergen, partial [Anisakis simplex];	UA3-recognized allergen, Ani s 14 allergen [Anisa
DN19246_c0_g3	2393	ABL77410.1;	UA3-recognized allergen, partial [Anisakis simplex];	UA3-recognized allergen Ani s 14 allergen [Anisa
DN19521_c0_g1	4468	ABL77410.1;	UA3-recognized allergen, partial [Anisakis simplex];	UA3-recognized allergen
DN19521_c0_g2	764	ABL77410.1;	UA3-recognized allergen, partial [Anisakis simplex];	UA3-recognized allergen
DN19722_c2_g2	2356	AAD16985.1;	vespid allergen antigen homolog [Wuchereria bancrofti];	Allergen allergen Pol d 5 precursor
DN20243_c0_g1	1472	ABV55106.1;	Ani s 9 allergen precursor [Anisakis simplex];	Ani s 9 allergen precursor

DN20642_c0_g1	2176	ABL77410.1;	UA3-recognized allergen, partial [Anisakis simplex];	Ani s 14 allergen [Anisa
DN21219_c0_g1	730	CAK50389.1;	ani s 4 allergen [Anisakis simplex];	Allergen major allergen Alt a 1 subunit
DN21219_c0_g2	532	CAK50389.1;	ani s 4 allergen [Anisakis simplex];	Allergen ani s 4 allergen [Anisakis
DN21375_c0_g1	1123	ABL77410.1;	UA3-recognized allergen, partial [Anisakis simplex];	Ani s 14 allergen [Anisa
DN21469_c0_g2	1835	KHN76256.1;	Major allergen Ani s 1 [Toxocara canis];	Allergen Major allergen Ani s 1 pre
DN21679_c0_g1	802	ABL77410.1;	UA3-recognized allergen, partial [Anisakis simplex];	UA3-recognized allergen
DN21762_c0_g3	1440	KHN81676.1;	Gut esterase 1 [Toxocara canis];	triosephosphat-isomerase
DN21808_c0_g1	1218	OZC07933.1;	hypothetical protein X798_05040 [Onchocerca flexuosa];	ABA-1 polyprotein precurs
DN21827_c0_g2	3074	ABL77410.1;	UA3-recognized allergen, partial [Anisakis simplex];	UA3-recognized allergen
DN21922_c1_g1	4828	AGC60020.1;	Ani s 2 allergen [Anisakis pegreffii];	Ani s 2 allergen [Anisakis
DN22223_c2_g2	2971	ACN32322.1;	tropomyosin [Ascaris lumbricoides];	tropomyosin [Ascaris lum
DN22574_c0_g1	628	KHN88507.1;	U4/U6.U5 tri-snRNP-associated protein 1 [Toxocara canis];	RNA dependent RNA polym
DN22574_c0_g5	695	XP_001893654.1;	SART-1 family protein [Brugia malayi];	Allergen tropomyosin [Dermatophagoid
DN22728_c0_g1	2317	ABL77410.1;	UA3-recognized allergen, partial [Anisakis simplex];	Ani s 14 allergen [Anisa
DN22751_c1_g1	1264	BAT62430.1;	Ani s 14 allergen [Anisakis simplex];	Ani s 14 allergen
DN22751_c2_g1	2254	ABL77410.1;	UA3-recognized allergen, partial [Anisakis simplex];	UA3-recognized allergen
DN24663_c0_g1	335	XP_013183788.1;	PREDICTED: pancreatic lipase-related protein 2-like [Amyeloidis transitella];	venom phospholipase A1 2 p
DN26355_c0_g1	519	PRD39420.1;	gloA [Nephila clavipes];	Lactoylglutathione lyase
DN2811_c0_g1	383	-----	-----	Fag s 1 pollen allergen
DN2994_c0_g1	1646	ABL77410.1;	UA3-recognized allergen, partial [Anisakis simplex];	Ani s 12 allergen [Anisa

DN31078_c0_g1	219	ABL77410.1;	UA3-recognized allergen, partial [Anisakis simplex];	UA3-recognized allergen	Ani s 14
DN35831_c0_g1	735	KHN85858.1;	DVA-1 polyprotein [Toxocara canis];	allergen [Anisa	
DN37518_c0_g1	579	AFH01340.1;	carboxylesterase [Nilaparvata lugens];	ABA-1 allergen [Ascaris lum	
DN38104_c0_g1	607	XP_022204320.1;	minor allergen Alt a 7-like [Nilaparvata lugens];	carboxylesterase	
DN38589_c0_g1	1066	ABL77410.1;	UA3-recognized allergen, partial [Anisakis simplex];	Allergen Alt a 7 [Alterna	
DN39207_c0_g1	350	AVA17421.1;	putative Per a allergen [Periplaneta americana];	Ani s 14 allergen [Anisa	
DN6357_c0_g1	293	BAT62430.1;	Ani s 14 allergen [Anisakis simplex];	heat shock-like protein	
DN9406_c0_g1	1249	ABL77410.1;	UA3-recognized allergen, partial [Anisakis simplex];	Ani s 14 allergen	
				Ani s 14 allergen [Anisa	

Table S3 Top 20 most differentially up-regulated expressed genes in AP-GJ.

No	Name	NR description	LogFC	Pvalue	Padjust
1	FKTN	KHN75690.1(Uncharacterized protein T07A5.1 [Toxocara canis])	8.9	2.07E-41	1.85E-40
2	DUF273	KHN79472.1(hypothetical protein Tcan_04706 [Toxocara canis])	7.1	3.70E-15	1.51E-14
3	GCNT2	KHN72286.1(N-acetyllactosaminide beta-1,6-N-acetylglucosaminyl-transferase, isoform C [Toxocara canis])	6.6	1.16E-11	3.83E-11
4	Tcan_18584	KHN80436.1(hypothetical protein Tcan_18584 [Toxocara canis])	6.4	4.32E-10	1.28E-09
5	hypothetical protein	ELU18102.1(hypothetical protein CAPTEDRAFT_143611, partial [Capitella teleta])	6.3	5.38E-51	5.68E-50
6	glpK	KHN77667.1(putative glycerol kinase [Toxocara canis])	6.3	5.42E-308	2.72E-306
7	Gp-FAR-1	KHN88406.1(Fatty-acid and retinol-binding protein 1 [Toxocara canis])	6.1	1.48E-201	4.90E-200
8	hypothetical protein	KRX86393.1(hypothetical protein T12_13365, partial [Trichinella patagoniensis])	6.1	2.18E-29	1.50E-28
9	MCFD2	XP_024502182.1(Multiple coagulation factor deficiency protein 2 [Strongyloides ratti])	6.0	6.58E-08	1.61E-07
10	SLC24A2	KHN84113.1(Sodium/potassium/calcium exchanger 2 [Toxocara canis])	5.6	1.76E-06	3.74E-06
11	AQP10	KHN76701.1(Aquaporin-10 [Toxocara canis])	5.6	1.64E-40	1.45E-39
12	LRR1	KHN87939.1(Leucine-rich repeat protein 1 [Toxocara canis])	5.3	1.35E-05	2.58E-05
13	MBF2	XP_021197663.1(uncharacterized protein LOC110381636 isoform X1 [Helicoverpa armigera])	5.1	1.14E-08	3.00E-08
14	DUF1759	CAA82797.1(gag protein [Ascaris lumbricoides])	4.9	0.000224714	0.000353224
15	RP-S23e, RPS23	NP_001298622.1(40S ribosomal protein S23 [Papilio polytes])	4.9	0.000224714	0.000353224

16	ART2	KZC15096.1(Putative uncharacterized protein ART2 [Dufourea novaeangliae])	4.8	8.35E-40	7.26E-39
17	hypothetical protein	XP_003370460.1(conserved hypothetical protein [Trichinella spiralis])	4.6	3.70E-56	4.25E-55
18	secreted protein	ABI52743.1(10 kDa putative secreted protein [Argas monolakensis])	4.6	9.54E-40	8.26E-39
19	nas-13	KHN70807.1(Zinc metalloproteinase nas-13 [Toxocara canis])	4.6	8.76E-267	3.84E-265
20	cpg-3	KHN73077.1(Chondroitin proteoglycan 3 [Toxocara canis])	4.4	2.52E-18	1.19E-17

Table S4 The top 10 genes having the highest number of links and their corresponding GO terms, their interacting genes and the networks by Gene MANIA network analysis.

The number of links	Gene	GO term	NR
14	nep-2	integral component of membrane(GO:0016021),metalloendopeptidase activity(GO:0004222)	Neprilysin-2 [Toxocara canis]
14	unc-52	calcium ion binding(GO:0005509)	Basement membrane proteoglycan [Toxocara canis]
14	F32A5.3	proteolysis involved in cellular protein catabolic process(GO:0051603),serine-type carboxypeptidase activity (GO:0004185)	putative serine carboxypeptidase F32A5.3 [Toxocara canis]
11	ASP	extracellular region(GO:0005576)	Ancylostoma secreted protein [Toxocara canis]
11	col-155	collagen trimer(GO:0005581),structural constituent of cuticle(GO:0042302)	Putative cuticle collagen [Toxocara canis]
10	Flna	Not Found	Filamin-A [Toxocara canis]
10	col-12	collagen trimer(GO:0005581),structural constituent of cuticle(GO:0042302)	Cuticle collagen 12 [Toxocara canis]
9	4CL1	metabolic process(GO:0008152),catalytic activity(GO:0003824)	4-coumarate--CoA ligase 1 [Toxocara canis]
8	SRAE_2000085100	intracellular(GO:0005622),MAP kinase kinase kinase activity(GO:0004709)	Mitogen-activated protein kinase MLK1 [Toxocara canis]
8	SdaC	integral component of membrane(GO:0016021)	Lysine histidine transporter-like 4 [Toxocara canis]

The number of links	Gene	GO term	NR
14	nep-2	integral component of membrane(GO:0016021),metalloendopeptidase activity(GO:0004222)	Neprilysin-2 [Toxocara canis]
14	unc-52	calcium ion binding(GO:0005509)	Basement membrane protein 52 [Toxocara canis]
14	F32A5.3	proteolysis involved in cellular protein catabolic process(GO:0051603),serine-type carboxypeptidase activity (GO:0004185)	putative serine carboxypeptidase [Toxocara canis]
11	ASP	extracellular region(GO:0005576)	Ancylostoma secreted protein
11	col-155	collagen trimer(GO:0005581),structural constituent of cuticle(GO:0042302)	Putative cuticle collagen 155 [Toxocara canis]
10	Flna	Not Found	Filamin-A [Toxocara canis]
10	col-12	collagen trimer(GO:0005581),structural constituent of cuticle(GO:0042302)	Cuticle collagen 12 [Toxocara canis]
9	4CL1	metabolic process(GO:0008152),catalytic activity(GO:0003824)	4-coumarate--CoA ligase [Toxocara canis]
8	SRAE_2000085100	intracellular(GO:0005622),MAP kinase kinase kinase activity(GO:0004709)	Mitogen-activated protein kinase [Toxocara canis]
8	SdaC	integral component of membrane(GO:0016021)	Lysine histidine transporter [Toxocara canis]