

HELMINTHOLOGIA, 57, 3: 196 - 210, 2020

Differentially expressed homologous genes reveal interspecies differences of *Paragonimus proliferus* based on transcriptome analysis

S. H. LI^{1,#}, S. D. LI^{1,#}, H. J. LI^{2,#}, J. Y. LI², J. J. XU², G. J. CHANG², L. J. YANG¹, W. Q. WANG¹, Y. L. ZHANG², Z. Q. MA¹, S. M. HE¹, W. L. WANG^{1,*}, H. L. HUANG^{2,*}

¹School of Basic Medicine, Kunming Medical University, Kunming 650504, China, *E-mail: *wenlinwang331@163.com*; ²Department of (hepatology, oncology, infectious disease), the Third People's Hospital of Kunming, Kunming 650043, China, *E-mail: 1602062713@qq.com

Article info

Summary

Received November 8, 2019 Paragonimus proliferus (P. proliferus), one of 46 Paragonimus species registered in the National Accepted April 1, 2020 Center for Biotechnology Information database, may be much more widely distributed in Southeast Asia than previously thought, as its reported natural foci have increased in the past decades. However, very little is known about its molecular biology, especially at the transcriptome level. For the first time, the transcriptome of this species was sequenced and compared with four other common Paragonimus species, namely Paragonimus skrjabini, Paragonimus kellicotti, Paragonimus miyazakii, and Paragonimus westermani, to predict homologous genes and differentially expressed homologous genes to explore interspecies differences of Paragonimus proliferus. A total of 7393 genes were found to be significantly differentially expressed. Of these, 49 were considered to be core genes because they were differentially expressed in all four comparison groups. Annotations revealed that these genes were related mainly to "duplication, transcription, or translation", energy or nutrient metabolism, and parasitic growth, proliferation, motility, invasion, adaptation to the host, or virulence. Interestingly, a majority (5601/7393) of the identified genes, and in particular the core genes (48/49), were expressed at lower levels in P. proliferus. The identified genes may play essential roles in the biological differences between Paragonimus species. This work provides fundamental background information for further research into the molecular biology of *P. proliferus*. Keywords: Paragonimus proliferus; transcriptome; homologous genes; interspecies differences

Introduction

Paragonimiasis, which is caused by lung flukes (also called *Paragonimus*), is considered a neglected tropical disease by the World Health Organization (Cumberlidge *et al.*, 2018) and contributes to serious disease burdens all over the world. Worldwide, approximately 23 million people suffer from paragonimiasis (Blair, 2014), and an estimated 293.8 million people are at risk (Keiser & Utzinger, 2005; Silachamroon & Wattanagoon, 2020). At least 50 *Paragonimus* species, 46 of which are registered in the National

Center for Biotechnology Information (NCBI) database, have been discovered in the past decades (Kong *et al.*, 2015). Increasingly, *Paragonimus* species, including *Paragonimus* westermani, *Paragonimus* africanus, *Paragonimus* heterotremus, *Paragonimus* kellicotti, *Paragonimus* mexicanus, *Paragonimus* siamensis, *Paragonimus* skrjabini, *Paragonimus* skrjabini miyazakii, and *Paragonimus* uterobilateralis, are reported to cause human infections (Chai, 2013). Therefore, greater attention should be paid to seemingly rare parasite species such as these.

P. proliferus was first isolated in 1964 from freshwater crabs in

^{* -} corresponding author

^{# -} these authors contributed equally to this work

Yunnan Province, China (Chung *et al.*, 1964). The lineage of *P. proliferus* is cellular organisms, Eukaryota, Opisthokonta, Metazoa, Eumetazoa, Bilateria, Platyhelminthes, Trematoda, Digenea, Plagiorchiida, Troglotremata, Troglotrematidae, *Paragonimus*. The reported natural foci of *P. proliferus* include several counties in Yunnan Province in southern China, including Xishuangbanna Dai Autonomous Prefecture (formerly known as Jinghong County) (Chung *et al.*, 1964), and Lvchun County (Yang *et al.*, 2007), as well as Lai Chau Province in northern Vietnam and Quang Binh Province in central Vietnam (Doanh *et al.*, 2013).

P. proliferus can be distinguished from other Paragonimus species

by a number of biological differences. Morphologically, *P. proliferus* metacercaria and the uterus of the adult worm, which is located near the ventral sucker, are very large. However, an encysted metacercaria is difficult to detect because the cyst wall covering is so thin and fragile that the metacercaria can rapidly excyst. Another significant difference between the species is host preference. The natural definitive or paratenic hosts of *P. proliferus* include rats, but not monkeys, dogs, or cats (Ben-jiang, 2004). In contrast, dogs and cats are suitable hosts for other *Paragonimus* species, such as *P. kellicotti* (Harrus, Nyska, Colorni, & Markovics, 1997; Peregrine, Nykamp, Carey, & Kruth, 2014), *P. skrjabini* (Doanh et



Fig. 1. Metacercaria and adult worms stained magenta with hydrochloric acid. A metacercaria isolated from a crab (a), and an adult worm recovered from the lung tissue of a rat (b).

al., 2016), Paragonimus miyazakii (Madarame et al., 2009), and *P. westermani* (Doanh et al., 2016; Irie et al., 2017). The distinct host permissiveness of different *Paragonimus* species strongly suggests that their virulence and pathogenicity may also differ considerably. Therefore, it is necessary to explore the interspecies differences between *P. proliferus* and other *Paragonimus* species at the molecular level.

Transcriptome analysis is an efficient method to assess the molecular and biological features of a causative agent. However, according to the NCBI database, of 46 known *Paragonimus* species, the transcriptomes of only four have been sequenced, including *P. skrjabini, P. kellicotti, P. miyazakii*, and *P. westermani*, whereas the transcriptomes of the remaining species, including *P. proliferus*, have not. However, the transcriptomic datasets of other *Paragonimus* species do not represent the complete biological and molecular characteristics of *P. proliferus*, as it is difficult to predict interspecific and intraspecies differences induced by genetic variation during evolution.

Given that little is known about the molecular biology of *P. pro-liferus*, analysis of high-throughput transcriptome sequences is a



Fig. 2. Sequencing, pretreatment, and annotation of *Paragonimus proliferus* transcriptome. The transcriptome of adult *P. proliferus* was sequenced using RNA-Seq (Hi-Seq Illumina X 4000). A total of 119.20 Mb raw reads were obtained, and 89.34 clean reads remained after removing low-quality and adaptor-polluted reads and reads with a high N content. The total number of clean bases, Q20, Q30, and clean reads ratio were 13.40 Gb, 96.85%, 91.23%, and 74.95, respectively. Trinity assembly of the clean reads generated 47,959 transcripts (total length: 22,204,462 bp, mean length: 462 bp, N50: 632 bp, N70: 339 bp, N90: 209 bp, GC content: 47.83%). A total of 29,967 unigenes (total length, 17,089,849 bp; mean length, 570 bp; N50, 826 bp; N70, 446 bp; N90, 249 bp; GC content, 47.36%) were further obtained after clustering and removal of redundant sequences using Tgicl. Finally, a total of 20,669 (68.97%) unigenes were annotated using seven functional databases: NR, 19,994 (66.72%); NT, 11,458 (38.24%); Swiss-Prot, 13,794 (46.03%); KEGG, 14,131 (47.16%); COG, 6,766 (22.58%); Interpro, 12,626 (42.13%); and GO, 9,659 (32.23%).

fundamental approach to determine the mechanism of development of the unique features of *P. proliferus* as well as its virulence and pathogenicity. In this study, an Illumina Hi-Seq 4000 platform was used to sequence the transcriptome of *P. proliferus* adult worms. Homologous genes were predicted by comparison with four other *Paragonimus* species (*P. skrjabini, P. kellicotti, P. miyazakii*, and *P. westermani*), the transcriptomes of which are available in the NCBI database. The identification of these genes will provide a foundation to explore interspecies differences between *P. proliferus* and other *Paragonimus* species at the transcriptome level.

Material and Methods

Preparation of P. proliferus samples

Freshwater crabs were collected from a natural stream in Mengla County, Xishuangbanna Dai Autonomous Prefecture, Yunnan Province, China. The shells of the crabs were removed, and the limbs, muscles, viscera, and other soft tissues were torn or triturated into tiny pieces, which were then repeatedly ground in a wire mesh. Samples were then rinsed by passing a large amount of water through the wire mesh into a 1000 ml tapered cylinder. The turbid liquid was allowed to settle for 30 minutes and the supernatant was carefully discarded, leaving approximately one third of the cylinder filled with sediment. After three to five washes, the remaining turbid liquid was shaken and poured into several petri dishes on a black backdrop. Enough suspension was poured into the dishes for tiny worms to easily be detected by visual inspection or under a light microscope. Using a pipette, the detected metacercariae were moved into normal saline and kept at 4°C until further use. The obtained metacercariae were used to infect Sprague-Dawley rats (eight metacercariae per rat) via subcutaneous injection (Lin & Xueming, 2001) into the abdominal wall. The rats were maintained at the Animal Lab Center of Kunming Medical University and provided with food and water ad libitum. Eight weeks post-infection (PI), the rats were euthanized under anesthesia by intraperitoneal injection of pentobarbital. Adult *P. proliferus* worms were isolated from the lungs of the rats, and a sample of three adult worms, whose viscera were removed under a light microscope, was stored at -80 °C until further use.

Transcriptome data of P. proliferus and other Paragonimus species Three P. proliferus adults were homogenized in 1 ml TRIzol reagent using a microcentrifuge pestle, and total RNA was extracted from the homogenate using a TRIzol Plus RNA Purification Kit (Thermo Fisher Scientific). RNA was treated with DNase I, and oligo(dT) was used to isolate mRNA. mRNA was fragmented by mixing with fragmentation buffer, and cDNA was synthesized using the mRNA fragments as templates. Short fragments were purified and resolved with elution buffer for end reparation and single nucleotide A (adenine) addition, followed by the addition of adapters. Suitable fragments were then selected for PCR amplification. An Agilent



Fig. 3. Species distribution of annotated unigenes in the Paragonimus proliferus transcriptome. A total of 26.25%, 22.36%, 20.83%, and 3.7% of the annotated unigenes belonged to the genera and species Opisthorchis viverrini, Culex quinquefasciatus, Clonorchis sinensis, and Halyomorpha halys, respectively, whereas 28.86% belonged to other species.



Fig. 4. Venn diagram showing homologous genes among comparisons of Paragonimus proliferus vs. Paragonimus skrjabini, Paragonimus kellicotti, Paragonimus miyazakii, or Paragonimus westermani.

2100 Bioanaylzer and an ABI StepOnePlus Real-Time PCR System (Applied Biosystems) were used to assess the quantity and quality of the sample library. The library was then sequenced using an Illumina Hi-Seq 4000 RNA-Seq platform with a depth of 12 Gb. The four available *Paragonimus* transcriptomic datasets were downloaded from the following NCBI database websites: *P. skrjabini* (https://www.ncbi.nlm.nih.gov/sra/SRX1507709), *P. kellicotti* (https://www.ncbi.nlm.nih.gov/sra/SRX150776), *P. miyazakii* (https://www.ncbi.nlm.nih.gov/sra/SRX1100074), and *P. westermani* (https://www.ncbi.nlm.nih.gov/sra/SRX1507710).

Pretreatment of raw reads and de novo assembly

After obtaining raw reads of the five *Paragonimus* species, namely *P. proliferus*, *P. skrjabini*, *P. kellicotti*, *P. miyazakii*, and *P. westermani*, low-quality and adaptor-polluted reads and reads with a high content of unknown bases (N) were removed to obtain clean reads, which were then used to perform de novo assembly of unigenes using Trinity (version 2.0.6).

Transcript annotation: We used BLAST (version 2.2.23) to align unigenes to Non-Redundant Protein Sequence (NR), Nucleotide Sequence (NT), Clusters of Orthologous Groups (COG), Kyoto Encyclopedia of Genes and Genomes (KEGG) and Swiss-Prot annotations. Blast2GO (version 2.5.0) was used to obtain NR and Gene Ontology (GO) annotations, and InterProScan5 (version 5.11-51.0) was used to obtain InterPro annotations.

Prediction of differentially expressed homologous P. proliferus genes

Four bilateral comparisons of unigenes were performed among the five species (*P. proliferus* versus *P. skrjabini, P. kellicotti, P. miyazakii*, and *P. westermani*) using BLAST. Transcripts that were the best match between two species (e-value=10e-5) were defined as homologous genes. Pairwise comparisons were also performed using NOISeq to identify differentially expressed homologous genes with |log2FC|≥1 and probability > 0.8. Mutually differentially expressed genes among the four pairs of comparisons were considered core genes.

Ethical Approval and/or Informed Consent

This study was approved by the Ethics Committee of Kunming Medical University. The methods were carried out in accordance with approved guidelines.

Results and Discussion

Morphological identification of P. proliferus

Their extremely thin and fragile cystic wall, spindle or scaphoid shape, and huge size of 2.253 ± 0.364 mm × 0.668 ± 0.071 mm allowed metacercariae to easily and rapidly excyst after isolation



Fig. 5. Changes in expression of differentially expressed homologous genes between *Paragonimus proliferus* and four other common *Paragonimus* species. A total of 3950/5622 (70.26 %), 1049/1084 (96.77 %), 388/473 (82.03 %), and 189/214 (88.32 %) genes were expressed at lower levels in *P. proliferus* compared to *Paragonimus* kellicotti, *Paragonimus skrjabini*, *Paragonimus westermani*, and *Paragonimus miyazakii*, respectively.

from crabs and exposure to the environment. The ventral sucker, which was approximately three times the size of the oral sucker, was located in the anterior third end of the body. The two intestinal branches that extended alongside the body wall and towards the tail end were thin and curved before the ventral sucker level, and became thick and smooth thereafter (Fig. 1a).

Eight weeks PI, adult worms (identified by mature reproductive organs that could be clearly visualized under a light microscope) of 7.238 ± 0.704 mm × 3.571 ± 0.655 mm in size were detected in the lung tissue or thorax of rats. An abundance of eggs filled an eiloid uterus, which was located right at the ventral sucker level. One ovary was located at the opposite side of the uterus, and the two testicles, located in the middle-posterior part of the body, were flamboyancy with 4 - 6 lobules. The uterus and testes were also extremely huge (Fig. 1b).

The morphological features of both metacercariae and adults were identical to those of *P. proliferus* described by Ben-jiang (2004) and Doanh *et al.* (2008; 2013).

Overall description of P. proliferus transcriptome

As shown in Fig. 2, RNA-Seq was used to sequence the transcriptome of adult *P. proliferus* worms. The sequences were then assembled de novo and filtered to include only high-confidence transcript sequences. Of 89.34 Mb clean reads obtained from 119.20 Mb raw reads, a total of 47,959 transcripts corresponding to 29,967 unigenes were generated after removal of low-quality and adaptor-polluted reads and reads with a high content of unknown bases (N). The total length, mean length, N50, and GC content were 17,089,849 bp, 570 bp, 826 bp, and 47.36 %, respectively. Overall, a total of 20,669 (68.97 %) unigenes were annotated to seven functional databases (NR, NT, GO, COG, KEGG, Swiss-Prot, and InterPro). Fig. 3 shows that the annotated unigenes belonged to *Opisthorchis viverrini* (26.25 %), *Culex quinquefasciatus* (22.36 %), *Clonorchis sinensis* (20.83 %), *Halyomorpha halys* (3.7 %), and other species (28.86 %).

Homologous genes and differential expression

As shown in the Venn diagram in Fig. 4, a total of 10,629 (19.13 % of *P. proliferus* genes), 12,074 (21.74 %), 13,558 (24.41 %), and 8,051 (14.49 %) homologous genes were identified between *P. proliferus* and *P. skrjabini, P. kellicotti, P. miyazakii,* and *P. west-ermani,* respectively.

To characterize the interspecies differences in homologous gene expression, pairwise comparisons were performed using NOISeq, which identified 8192 differentially expressed homologous genes among the five species, 7393 of which had $|log2FC| \ge 1$ and probability > 0.8. Surprisingly, as shown in Fig. 5, 3950/5622 (70.26 %), 1049/1084 (96.77 %), 388/473 (82.03 %), and 189/214 (88.32 %) genes were expressed at lower levels in *P. proliferus* compared to *P. kellicotti*, *P. skrjabini*, *P. westermani*, and *P. miyazakii*, respectively. The biological characteristics of *P. proliferus* related to lower expression of these genes requires further exploration, but there





Fig. 7. Venn diagram showing 49 core genes identified from the 8192 differentially expressed homologous genes.

is indeed a possibility that these genes influence growth, development, invasion, reproduction, virulence, and motility.

GO annotations of differentially expressed homologous genes

There were 527, 693, and 757 genes annotated as cellular components (CC; 207 terms, 18 of which were significantly enriched with P < 0.05), biological processes (BP; 930 terms, 10 with P <0.05), and molecular functions (MF; 756 terms, 14 with P < 0.05), respectively. The top 42 enriched GO terms (P < 0.05) are shown in Fig. 6.

Of the 10 significantly enriched GO terms belonging to BP, four participate in "genetic central dogma", such as DNA replication/ synthesis of RNA primers, regulation of translational initiation, rRNA modification, and DNA-dependent DNA replication. The remaining six terms (monovalent inorganic cation transport, potassium ion transport, hydrogen ion transmembrane transport, ornithine metabolic process, and fructose metabolic process) are related to energy metabolism. These findings are in line with the enrichment of MF; at least five of the 14 terms (P < 0.05) take part in "duplication, transcription, or translation", such as endonuclease activity, DNA primase activity, ribonuclease activity, and transcriptional repressor activity. The majority of the remaining terms were related to the oxidation respiratory chain and the activity of some

energy production-related enzymes. Therefore, we inferred that changes in the processes of "duplication, transcription, or translation" and "energy metabolism" in *P. proliferus* may lead to key interspecies differences in development or biological behavior. This was further verified by the significantly enriched GO terms belonging to CC; seven of 18 terms (P < 0.05) belonged to chondriosome-related cellular components, four of 18 to ATPase, and three of 18 to ribosome. As is well known, the chondriosome is the site of the oxidation respiratory chain, and the ribosome is critical for protein synthesis.

KEGG annotations of differentially expressed homologous genes A total of 1412 genes were annotated to 290 KEGG pathways (as shown in Table 1, 13 pathways were significantly enriched with P < 0.05). Of the 13 pathways, four were assigned to signal transduction, including forkhead box, class O (FoxO), transforming growth factor (TGF) β, vascular endothelial growth factor (VEGF), and mechanistic target of rapamycin (mTOR) signaling, all of which participate in important biological processes; two were related to metabolism, mainly pyruvate metabolism and terpenoid backbone biosynthesis; and two (aldosterone-regulated sodium reabsorption and bile secretion) were assigned to the excretory and digestive systems. However, these genes also intrinsically belong to ion or





Pathways	Involving genes	P value	Pathway ID	Level 2
Bladder cancer	8	0.006593	ko05219	Cancers: Specific types
Melanoma	9	0.02046	ko05218	Cancers: Specific types
Thyroid cancer	9	0.03693	ko05216	Cancers: Specific types
Pyruvate metabolism	20	0.03104	ko00620	Carbohydrate metabolism
Bile secretion	14	0.04864	ko04976	Digestive system
Aldosterone-regulated sodium reabsorption	12	0.002475	ko04960	Excretory system
Primary immunodeficiency	6	0.02711	ko05340	Immune diseases
Terpenoid backbone biosynthesis	10	0.03597	ko00900	Metabolism of terpenoids and polyketides
DNA replication	17	0.01765	ko03030	Replication and repair
FoxO signaling pathway	24	0.0109	ko04068	Signal transduction
TGF-beta signaling pathway	15	0.01839	ko04350	Signal transduction
VEGF signaling pathway	11	0.03018	ko04370	Signal transduction
mTOR signaling pathway	13	0.03336	ko04150	Signal transduction

Table 1. Top 13 Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways significantly enriched by differentially expressed homologous genes (P < 0.05).

lipid metabolism as well as pathways related to DNA replication and repair.

In mammals, FoxO family members are involved in cell metabolism, growth, differentiation, oxidative stress, senescence, autophagy, and aging (Lee & Dong, 2017), and we suspect that they have similar biological functions in *Paragonimus*. TGF- β or its parasite mimics may regulate the immune response of the host (Musah-Eroje & Flynn, 2018), and mTOR signaling is related to *Leishmania* proliferation (Jaramillo *et al.*, 2011). Energy metabolism is closely related to parasitic growth or adaptation to the host environment (Caddigan *et al.*, 2017), as well as replication and virulence (Mancio-Silva *et al.*, 2017). Recently, it was reported that pyruvate homeostasis is a determinant of parasite growth and metabolic plasticity in *Toxoplasma gondii* (Xia *et al.*, 2019).

The processes of signal transduction, metabolism, and DNA replication and repair may affect the biological characteristics of *P. proliferus*; further functional verifications should be performed to determine the effect of genes on the key targets of these signaling pathways.

Core differentially expressed homologous genes

As shown in the Venn diagram in Fig. 7, of the 8192 differentially expressed homologous genes, 49 mutually differentially expressed genes were identified among the four pairs of comparisons and defined as core genes that may play key roles in the biological differences between *P. proliferus* and other *Paragonimus* species. Interestingly, a majority (48/49) of these core genes were expressed at lower levels in *P. proliferus* than in the four other species. A total of 16 of the 49 core genes were annotated by functional databases (partial annotation information is listed in Table 2), whereas the remaining 33 were not.

GO and KEGG annotations of core genes

Six core genes were annotated to 11, 21, and 26 GO terms belonging to CC (DNAH1, NOP56, and PI3K), BP (SSP7, PFK-2/ FBPase2, DNAH1, Adk, PI3K, and NOP56), and MF (SSP7, PFK-2/FBPase2, DNAH1, Adk, PI3K, and NOP56), respectively, as shown in Fig. 8. Of the 21 terms belonging to BP, 17 were involved in metabolism, and at least 12 of the 26 nucleotide or ribonucleotide-related terms belonging to MF may be involved in DNA replication and repair. Of all the enriched GO terms, four belonging to BP (phosphate-containing compound metabolic process, organophosphate metabolic process, phosphorus metabolic process, and carbohydrate derivative metabolic process) and three belonging to MF (phosphotransferase activity/alcohol group as acceptor, kinase activity, and transferase activity/transferring phosphorus-containing groups) were significantly enriched with P < 0.05, which seems to indicate that phosphorus and carbohydrate metabolism may play an important role in the interspecies differences between P. proliferus and other Paragonimus species.

The six genes annotated to the GO terms mentioned above, as well as four others annotated to the KEGG Orthology (*MRP*, *Tuba*, *Wasf2*, and *Ppp1r7*), were also annotated to 87 KEGG pathways, 28 of which had a *P* value < 0.05 (Table 3) and all of which were enriched with four genes (*PI3K*, *WASF2*, *MRP*, and *PFK-2/FB-Pase2*). Of the 28 pathways, four were assigned to signal transduction (including 5' adenosine monophosphate-activated protein kinase [AMPK], Janus kinase/signal transduces and activators of transcription [JAK-STAT], VEGF, and mTOR signaling), four participated in immune-related signaling (Fc gamma R-mediated phagocytosis, and Toll-like receptor, Fc epsilon RI, and B cell receptor signaling pathways), and seven were related to digestive or endocrine systems, as well as carbohydrate metabolism, and

Gene-ID	Database	Annotated ID	Description		
TR11359 c0_g1	KEGG	smm:Smp_137080	multidrug resistance protein; K05658 ATP-binding cassette, subfamily B (MDR/TAP), member 1 [EC:3.6.3.44]		
TR12766 c0_g1	KEGG	oas:101108295	tubulin alpha-3 chain; K07374 tubulin alpha		
TR16634 c0_g1	KEGG	smm:Smp_164960	phosphatidylinositol-45-bisphosphate 3-kinase catalytic subunit alpha PI3K; K00922 phosphatidylinositol-4,5- bisphosphate 3-kinase [EC:2.7.1.153]		
TR17957 c0_g1	KEGG	fab:101816860	WASF2; WAS protein family, member 2; K05748 WAS protein family, member 2		
TR89500 c0_g1	KEGG	smm:Smp_159120	family C48 unassigned peptidase (C48 family); K08596 sentrin-specific protease 7 [EC:3.4.22.68]		
TR10230 c0_g1	NR	gi 358340450 dbj GAA48338.1	retrovirus-related Pol polyprotein from transposon opus [Clonorchis sinensis]		
TR11281 c0_g1	NR	gi 358253292 dbj GAA52762.1	serine/threonine-protein phosphatase 2A regulatory subunit B" subunit alpha [Clonorchis sinensis]		
TR15039 c0_g2	NR	gi 684389238 ref XP_009169318.1	hypothetical protein T265_05919 [Opisthorchis viverrini] >gi 663050934 gb KER26939.1 hypothetical protein T265_05919 [Opisthorchis viverrini]		
TR18101 c0_g1	NR	gi 684372571 ref XP_009164145.1	hypothetical protein T265_01791 [Opisthorchis viverrini] >gi 663056274 gb KER32181.1 hypothetical protein T265_01791 [Opisthorchis viverrini]		
TR22019 c0_g1	NR	gi 358337500 dbj GAA32515.2	cell wall protein Awa1p [Clonorchis sinensis]		
TR18820 c0_g1	SwissProt	sp Q64640 ADK_RAT	Adenosine kinase OS=Rattus norvegicus GN=Adk PE=1 SV=3		
TR18976 c0_g1	SwissProt	sp Q9D6Z1 NOP56_MOUSE	Nucleolar protein 56 OS=Mus musculus GN=Nop56 PE=1 SV=2		
TR20969 c0_g1	SwissProt	sp Q5EB30 ODF3A_XENTR	Outer dense fiber protein 3 OS=Xenopus tropicalis GN=odf3 PE=2 SV=1		
TR21606 c0_g1	SwissProt	spjQ9P2D7jDYH1_HUMAN	Dynein heavy chain 1, axonemal OS=Homo sapiens GN=DNAH1 PE=2 SV=4		
TR275 c0_g1	SwissProt	sp Q6DTY7 F264_MOUSE	6-phosphofructo-2-kinase/fructose-2,6-bisphosphatase 4 OS=Mus musculus GN=Pfkfb4 PE=2 SV=4		
TR50927 c0_g1	SwissProt	sp Q3UM45 PP1R7_MOUSE	Protein phosphatase 1 regulatory subunit 7 OS=Mus musculus GN=Ppp1r7 PE=1 SV=2		

Table 2. Annotation information for the 16 core genes.

Pathway	Pathway ID	Level 2	Pvalue	Involving DEHGs
Choline metabolism in cancer	ko05231	Cancers: Overview	0.004227	PI3K;WASF2
MicroRNAs in cancer	ko05206	Cancers: Overview	0.01585	PI3K;MRP
Melanoma	ko05218	Cancers: Specific types	0.02688	PI3K
Acute myeloid leukemia	ko05221	Cancers: Specific types	0.0292	PI3K
Non-small cell lung cancer	ko05223	Cancers: Specific types	0.03013	PI3K
Pancreatic cancer	ko05212	Cancers: Specific types	0.03338	PI3K
Endometrial cancer	ko05213	Cancers: Specific types	0.04168	PI3K
Chronic myeloid leukemia	ko05220	Cancers: Specific types	0.04397	PI3K
Colorectal cancer	ko05210	Cancers: Specific types	0.04718	PI3K
Fructose and mannose metabolism	ko00051	Carbohydrate metabolism	0.04352	PFK-2/FBPase2
Apoptosis	ko04210	Cell growth and death	0.03661	PI3K
Carbohydrate digestion and absorption	ko04973	Digestive system	0.0292	PI3K
Type II diabetes mellitus	ko04930	Endocrine and metabolic diseases	0.02781	PI3K
Regulation of lipolysis in adipocytes	ko04923	Endocrine system	0.03292	PI3K
Thyroid hormone signaling pathway	ko04919	Endocrine system	0.03431	PI3K;PFK-2/FBPase2
Prolactin signaling pathway	ko04917	Endocrine system	0.03754	PI3K
Aldosterone-regulated sodium reabsorption	ko04960	Excretory system	0.03013	PI3K
Fc gamma R-mediated phagocytosis	ko04666	Immune system	0.005578	PI3K;WASF2
Toll-like receptor signaling pathway	ko04620	Immune system	0.02641	PI3K
Fc epsilon RI signaling pathway	ko04664	Immune system	0.03477	PI3K
B cell receptor signaling pathway	ko04662	Immune system	0.03754	PI3K
Bacterial invasion of epithelial cells	ko05100	Infectious diseases: Bacterial	0.02904	PI3K;WASF2
Chagas disease (American trypanosomiasis)	ko05142	Infectious diseases: Parasitic	0.04214	PI3K
ABC transporters	ko02010	Membrane transport	0.0426	MRP
AMPK signaling pathway	ko04152	Signal transduction	0.007408	PI3K;PFK-2/FBPase2
Jak-STAT signaling pathway	ko04630	Signal transduction	0.03615	PI3K
VEGF signaling pathway	ko04370	Signal transduction	0.03846	PI3K
mTOR signaling pathway	ko04150	Signal transduction	0.04901	PI3K

Table 3. Top 28 Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways significantly enriched by core genes (P < 0.05).

signaling pathways that are always involved directly or indirectly in metabolism. It is worth noting that there were also nine pathways related to level 2 cancers according to KEGG enrichment, which suggests that parasite infections may be a risk factor for tumorigenesis, a view that is supported by other reports (Feng & Cheng, 2017; van Tong *et al.*, 2017).

Phosphoinositide 3-kinase (PI3K), which is critical for parasite virulence, was involved in 26 significantly enriched pathways. In *Entamoeba histolytica*, PI3K signaling plays an important role in motility, phagocytosis, host cell adhesion, and proteolysis of the extracellular matrix (Koushik *et al.*, 2014). In *Trypanosoma cruzi*, PI3K-like activity plays an important role in host cell invasion, and parasite entry is impaired in trypomastigotes treated with a

PI3K inhibitor prior to infection (Wilkowsky *et al.*, 2001). Recently, PIKs have been proposed to be targets for the treatment of *T. cruzi* infection (Gimenez *et al.*, 2015). Phosphofructokinase (PFK-2)/ fructose 1,6-bisphospatase (FBPase), which is also found in *Trypanosoma brucei, Leishmania major,* and *T. cruzi* (Chevalier *et al.*, 2005), was identified as being involved in fructose and mannose metabolism and thyroid hormone signaling in our study. According to KEGG enrichment of the analyzed transcriptome, Wiskott-Aldrich syndrome protein family member 2 (WASF2), which was annotated to signaling of choline metabolism in cancer, Fc gamma R-mediated phagocytosis, and bacterial invasion of epithelial cells, may be involved in parasite metabolism, invasion,

and host immunity. Mitochondrial RNA processing (MRP) belongs

to signaling of microRNAs in cancer and ATP-binding cassette (ABC) transporter classes. Reports have revealed that disruption of MRP in *Plasmodium falciparum* was related to a reduction in parasitemia, removal of toxic metabolites, and transport of antimalarial drugs (Raj *et al.*, 2009). Additionally, drug efflux transporters of the ABC gene superfamily can confer drug resistance to pathogens (Carmona-Antonanzas *et al.*, 2015).

Tubulin alpha-8 chain (TUBA) is also essential for host cell invasion and parasite survival in T. gondii (Morrissette, 2015). Adenosine kinase, encoded by the Adk gene, is a key enzyme in the purine-salvage pathway, adenosine phosphorylation, and activation of cytotoxic analogues (Romanello et al., 2013; Timm et al., 2014). In T. brucei, inhibition or downregulation of adenosine kinase results in resistance to cordycepin, demonstrating its role in the activation of adenosine antimetabolites (Luscher et al., 2007) and silencing of nucleolar protein (NOP) 1, which contributes to defects in rRNA-processing and causes lethal modifications (Barth et al., 2008). However, it is unclear whether NOP56 has a similar role. Protein phosphatases (PPs) are also related to reproduction and development. For example, in T. cruzi protein phosphatase 2A is important for the complete transformation of trypomastigotes into amastigotes (Gonzalez et al., 2003), and recent research in Toxocara canis has shown that PP1 regulates kinetochore-microtubule interactions during spermatogenesis via PP1ca-PP1r7 mechanisms (Ma et al., 2015). The functions of dynein axonemal heavy chain 1 (DNAH1) and Ssp7 are still unknown in parasites, especially in *P. proliferus*, and further research on the molecular biology of these genes in parasites is necessary.

Serine/threonine-protein phosphatase 2A regulatory subunit B (STP) may play a functional role in parasite reproduction (Boag et al., 2003; Ma et al., 2014). In T. canis, STP is not detected in adult females, but is expressed at high levels in the spermary, vas deferens, and musculature of the adult male, suggesting that STP may be involved in spermatogenesis and mating behavior and may represent a potential molecular target for controlling the reproduction of T. canis (Ma et al., 2014). The outer dense fiber (ODF), a coiled-coil protein, is the main cytoskeletal structural component of the sperm tail in animals and is potentially involved in building the cellular cytoskeleton (Petersen et al., 2002). Some ODF genes were found to be aberrantly expressed in tumors, such as prostate adenocarcinomas (Ghafouri-Fard et al., 2012), but the functions of parasite ODF, retrovirus-related Pol polyprotein from transposon opus, STP, and the cell wall protein Awa1p, require further research.

Conclusions

This study is the first to analyze and characterize the transcriptome of *P. proliferus*. Bioinformatics analysis and comparison with other *Paragonimus* species revealed that differentially expressed homologous genes were mainly related to "duplication, transcription, or translation" and energy or nutrient metabolism, as well as parasite growth, proliferation, motility, invasion, adaptation to the host, or virulence. These differences may be the key to reported differences in morphology or host preference. More importantly, the identified differentially expressed homologous genes related to virulence, motility, or invasion may explain differences in the pathogenicity of *P. proliferus*. It will be valuable to further explore the functions of these genes in *P. proliferus*.

Conflict of Interest

The authors declare that they have no competing interests. All authors read, edited, and approved the final manuscript, and consent to its publication.

Acknowledgments

We acknowledge the residents of Mengla County, Yunnan Province, China for their assistance in collecting local freshwater crabs. We also thank International Science Editing (http://www.internationalscienceediting.com) for editing this manuscript.

Financial Support

This study was supported by Yunnan Local Colleges Applied Basic Research Projects (Grant No. 2018FH001-087), Projects Of Kunming Assurance Plan On Science And Technology Of People's Livelihood (Grant No. Kkjz2019-1-s-2531800000808), the National Natural Science Foundation of China (Grant No. 81360252), the Association Foundation Program of Yunnan Provincial Science and Technology Department and Kunming Medical University (Grant No. 201801PH00065), and the Scientific Research Foundation of Yunnan Education (Grant No. 2018JS157).

References

BARTH, S., SHALEM, B., HURY, A., TKACZ, I.D., LIANG, X.H., ULIEL, S., MYSLYUK, I., DONIGER, T., SALMON-DIVON, M., UNGER, R., MICHAELI, S. (2008): Elucidating the role of C/D snoRNA in rRNA processing and modification in *Trypanosoma brucei*. *Eukaryot Cell*, 7: 86 – 101. DOI: 10.1128/EC.00215-07

BEN-JIANG, Z. (2004): [Studies on Paragonimus Proliferus]. *Zhongguo Ji Sheng Chong Xue Yu Ji Sheng Chong Bing Za Zhi*, 22: 109 – 112 (In Chinese)

BLAIR, D. (2014): Paragonimiasis. *Adv Exp Med Biol*, 766: 115 – 152. DOI: 10.1007/978-1-4939-0915-5_5

BOAG, P.R., REN, P., NEWTON, S.E., GASSER, R.B. (2003): Molecular characterisation of a male-specific serine/threonine phosphatase from *Oesophagostomum dentatum* (Nematoda: Strongylida), and functional analysis of homologues in *Caenorhabditis elegans. Int J Parasitol*, 33: 313 – 325. DOI: 10.1016/s0020-7519(02)00263-1 CADDIGAN, S.C., PFENNING, A.C., SPARKES, T.C. (2017): Competitive growth, energy allocation, and host modification in the acanthoce-

phalan *Acanthocephalus dirus*: field data. *Parasitol Res*, 116: 199 – 206. DOI: 10.1007/s00436-016-5279-8

CARMONA-ANTONANZAS, G., CARMICHAEL, S.N., HEUMANN, J., TAGGART, J. B., GHARBI, K., BRON, J.E., BEKAERT, M., STURM, A. (2015): A survey of the ATP-binding cassette (ABC) gene superfamily in the salmon louse (*Lepeophtheirus salmonis*). *PLoS One*, 10: e0137394. DOI: 10.1371/journal.pone.0137394

CHAI, J.Y. (2013): Paragonimiasis. *Handb Clin Neurol*, 114: 283 – 296. DOI: 10.1016/b978-0-444-53490-3.00023-6

CHEVALIER, N., BERTRAND, L., RIDER, M.H., OPPERDOES, F.R., RIGDEN, D.J., MICHELS, P.A. (2005): 6-Phosphofructo-2-kinase and fructose-2,6-bisphosphatase in Trypanosomatidae. Molecular characterization, database searches, modelling studies and evolutionary analysis. *Febs J*, 272: 3542 – 3560. DOI: 10.1111/j.1742-4658.2005.04774.x

CHUNG, H.L., HO, L.Y., CHENG, L.T., TSAO, W.C. (1964): The discovery in Yunnan province of 2 new species of lung flukes—*Paragonimus tuanshanensis* sp. nov. and *Paragonimus menglaensis* sp. nov. Part I. Study on morphology and life history with discussion on possible pathogenicity to man. *Chin Med J*, 83: 641 – 659 CUMBERLIDGE, N., ROLLINSON, D., VERCRUYSSE, J., TCHUEM TCHUENTE, L.A., WEBSTER, B., CLARK P.F. (2018): *Paragonimus* and paragonimiasis in West and Central Africa: unresolved questions. *Parasitology*, 145: 1748 – 1757. DOI: 10.1017/s0031182018001439

DOANH, P.N., HIEN, H.V., NONAKA, N., HORII, Y. NAWA, Y. (2013): Genetically variant populations of Paragonimus proliferus Hsia & Chen, 1964 from central Vietnam. *J Helminthol*, 87: 141 – 146. DOI: 10.1017/S0022149X12000090

DOANH, P.N., HIEN, H.V., TU, L.A., NONAKA, N., HORII, Y., NAWA, Y. (2016): Molecular identification of the trematode *Paragonimus* in faecal samples from the wild cat *Prionailurus bengalensis* in the Da Krong Nature Reserve, Vietnam. *J Helminthol*, 90: 658 – 662. DOI: 10.1017/s0022149x15000838

DOANH, P.N., SHINOHARA, A., HORII, Y., HABE, S., NAWA, Y., LE, N.T. (2008): Discovery of *Paragonimus proliferus* in Northern Vietnam and their molecular phylogenetic status among genus *Paragonimus*. *Parasitol Res*, 102: 677 – 683. DOI: 10.1007/s00436-007-0811-5

FENG, M., CHENG, X. (2017): Parasite-associated cancers (blood flukes/liver flukes). *Adv Exp Med Biol*, 1018: 193 – 205. DOI: 10.1007/978-981-10-5765-6_12

GHAFOURI-FARD, S., GHAFOURI-FARD, S., MODARRESSI, M.H. (2012): Expression of splice variants of cancer-testis genes ODF3 and ODF4 in the testis of a prostate cancer patient. *Genet Mol Res*, 11: 3642 – 3648. DOI: 10.4238/2012.October.4.11

GIMENEZ, A.M., GESUMARIA, M.C., SCHOIJET, A.C., ALONSO, G.D., FLAWIA, M.M., RACAGNI, G.E., MACHADO, E.E. (2015): Phosphatidylinositol kinase activities in *Trypanosoma cruzi* epimastigotes. *Mol Biochem Parasitol*, 203: 14 – 24. DOI: 10.1016/j.molbiopara.2015.10.002

Gonzalez, J., Cornejo, A., Santos, M.R., Cordero, E.M., Gutierrez, B., Porcile, P., Mortara, R.A., Sagua, H., Da Silveira, J.F., Araya,

J.E. (2003): A novel protein phosphatase 2A (PP2A) is involved in the transformation of human protozoan parasite *Trypanosoma cruzi*. *Biochem J*, 374: 647 – 656. DOI: 10.1042/bj20030215

HARRUS, S., NYSKA, A., COLORNI, A., MARKOVICS, A. (1997) Sudden death due to *Paragonimus kellicotti* infection in a dog. *Vet Parasitol*, 71: 59 – 63. DOI: 10.1016/s0304-4017(97)00007-1

IRIE, T., YAMAGUCHI, Y., DOANH, P.N. GUO, Z.H., HABE, S., HORII, Y., NONAKA, N. (2017): Infection with *Paragonimus westermani* of boar-hunting dogs in Western Japan maintained via artificial feeding with wild boar meat by hunters. *J Vet Med Sci*, 79: 1419 – 1425. DOI: 10.1292/jvms.17-0149

JARAMILLO, M., GOMEZ, M.A., LARSSON, O., SHIO, M.T., TOPISIROVIC, I., CONTRERAS, I., LUXENBURG, R., ROSENFELD, A., COLINA, R., MCMASTER, R.W., OLIVIER, M., COSTA-MATTIOLI, M., SONENBERG, N. (2011): *Leishmania* repression of host translation through mTOR cleavage is required for parasite survival and infection. *Cell Host Microbe*, 9: 331 – 341. DOI: 10.1016/j.chom.2011.03.008

KEISER, J., UTZINGER, J. (2005): Emerging foodborne trematodiasis. *Emerg Infect Dis*, 11: 1507 – 1514. DOI: 10.3201/eid1110.050614 KONG, Y., NGOC DOANH, P., NAWA, Y. (2015): *Paragonimus*. In: XIAO, L., RYAN, U., FENG, Y. (Eds) *Biology of foodborne parasites*. 1st ed. Boca Raton: CRC Press; pp. 445 – 462. DOI: 10.1201/b18317

KOUSHIK, A.B., WELTER, B.H., ROCK, M.L., TEMESVARI, L.A. (2014): A genomewide overexpression screen identifies genes involved in the phosphatidylinositol 3-kinase pathway in the human protozoan parasite *Entamoeba histolytica*. *Eukaryot Cell*, 13: 401 – 411. DOI: 10.1128/ec.00329-13

LEE, S., DONG, H.H. (2017): FoxO integration of insulin signaling with glucose and lipid metabolism. *J Endocrinol*, 233: R67 – R79. DOI: 10.1530/joe-17-0002

LIN, R., XUEMING, L.I. (2001): ;Experiment infection of rats with encysted metacercariae of *Paragonimus heterotremus* through hypodermic and muscle inoculation. *Guangxi J. Prev. Med.*, 7: 201 – 203

LUSCHER, A., ONAL, P., SCHWEINGRUBER, A.M., MASER, P. (2007): Adenosine kinase of *Trypanosoma brucei* and its role in susceptibility to adenosine antimetabolites. *Antimicrob Agents Chemother*, 51: 3895 – 3901. DOI: 10.1128/aac.00458-07

Ma, G.X., ZHOU, R.Q., HU, S.J., HUANG, H.C., ZHU, T., XIA, Q.Y. (2014): Molecular characterization and functional analysis of serine/threonine protein phosphatase of *Toxocara canis*. *Exp Parasitol*, 141: 55 – 61. DOI: 10.1016/j.exppara.2014.03.019

Ma, G.X., ZHOU, R.Q., SONG, Z.H., ZHU, H.H., ZHOU, Z.Y., ZENG, Y.Q. (2015): Molecular mechanism of serine/threonine protein phosphatase 1 (PP1calpha-PP1r7) in spermatogenesis of *Toxocara canis. Acta Trop*, 149: 148 – 154. DOI: 10.1016/j.actatropica.2015.05.026

MADARAME, H., SUZUKI, H., SAITOH, Y., TACHIBANA, M., HABE, S., UCHI-DA, A., SUGIYAMA, H. (2009): Ectopic (subcutaneous) *Paragonimus miyazakii* infection in a dog. *Vet Pathol*, 46: 945 – 948. DOI: 10.1354/vp.08-VP-0237-M-CR

MANCIO-SILVA, L., SLAVIC, K., GRILO RUIVO, M.T., GROSSO, A.R.,

MODRZYNSKA, K.K., VERA, I.M., SALES-DIAS, J., GOMES, A.R., MACPHER-SON, C.R., CROZET, P., ADAMO, M., BAENA-GONZALEZ, E., TEWARI, R., LLINAS, M., BILLKER, O., MOTA, M.M. (2017): Nutrient sensing modulates malaria parasite virulence. *Nature*, 547: 213 – 216. DOI: 10.1038/nature23009

MORRISSETTE, N. (2015): Targeting *Toxoplasma* tubules: tubulin, microtubules, and associated proteins in a human pathogen. *Eukaryot Cell*, 14: 2 – 12. DOI: 10.1128/ec.00225-14

MUSAH-EROJE, M., FLYNN, R.J. (2018): Fasciola hepatica, Tgf- β and host mimicry: the enemy within. Curr Opin Microbiol, 46: 80 – 85. DOI: 10.1016/j.mib.2018.09.002

PEREGRINE, A.S., NYKAMP, S.G., CAREY, H., KRUTH, S. (2014): Paragonimosis in a cat and the temporal progression of pulmonary radiographic lesions following treatment. *J Am Anim Hosp Assoc*, 50: 356 – 360. DOI: 10.5326/jaaha-ms-6053

PETERSEN, C., AUMULLER, G., BAHRAMI, M., HOYER-FENDER, S. (2002): Molecular cloning of *Odf3* encoding a novel coiled-coil protein of sperm tail outer dense fibers. *Mol Reprod Dev*, 61: 102 – 112. DOI: 10.1002/mrd.1136

RAJ, D.K., MU, J., JIANG, H., KABAT, J., SINGH, S., SULLIVAN, M., FAY, M.P., McCutchan, T.F., SU, X.Z. (2009): Disruption of a *Plasmodium falciparum* multidrug resistance-associated protein (PfMRP) alters its fitness and transport of antimalarial drugs and glutathione. *J Biol Chem*, 284: 7687 – 7696. DOI: 10.1074/jbc.M806944200

ROMANELLO, L., BACHEGA, J.F., CASSAGO, A., BRANDAO-NETO, J., DE-MARCO, R., GARRATT, R.C., PEREIRA, H.D. (2013): Adenosine kinase from *Schistosoma mansoni*: structural basis for the differential incorporation of nucleoside analogues. *Acta Crystallogr D Biol Crystallogr*, 69: 126 – 136. DOI: 10.1107/s0907444912044800

SILACHAMROON, U., WATTANAGOON, Y. (2020): Paragonimiasis. In: RYAN, E.T., HILL, D.R., SOLOMON, T., ARONSON, N.E., ENDY, T.P. (Eds) *Hunter's tropical medicine and emerging infectious diseases*. 10th ed.: pp. 928 – 931. DOI: 10.1016/B978-0-323-55512-8.00129-0

TIMM, J., GONZALEZ-PACANOWSKA, D., WILSON, K.S. (2014): Structures of adenosine kinase from *Trypanosoma brucei brucei*. Acta *Crystallogr F Struct Biol Commun*, 70: 34 – 39. DOI: 10.1107/ s2053230x13033621

VAN TONG, H., BRINDLEY, P.J., MEYER, C.G., VELAVAN, T.P. (2017): Parasite Infection, Carcinogenesis and Human Malignancy. *EBioMedicine*, 15: 12 – 23. DOI: 10.1016/j.ebiom.2016.11.034

WILKOWSKY, S.E., BARBIERI, M.A., STAHL, P., ISOLA, E.L. (2001): *Trypanosoma cruzi*: phosphatidylinositol 3-kinase and protein kinase B activation is associated with parasite invasion. *Exp Cell Res*, 264: 211–218. DOI: 10.1006/excr.2000.5123

XIA, N., YE, S., LIANG, X., CHEN, P., ZHOU, Y., FANG, R., ZHAO, J., GUPTA, N., YANG, S., YUAN, J., SHEN, B. (2019): Pyruvate homeostasis as a determinant of parasite growth and metabolic plasticity in *Toxoplasma gondii*. *MBio*, 10: e00898-19. DOI: 10.1128/ mBio.00898-19

YANG, B.B., ZHOU, B.J., LI, R.Q., BAI, Z.W., WU O.B., GAO, X.F. (2007): [Preliminary investigation on *Paragonimus* in Lvchun County of Yunnan Province]. *Zhongguo Ji Sheng Chong Xue Yu Ji Sheng Chong Bing Za Zhi*, 25: 518 – 519 (In Chinese)