Korean J Parasitol. Vol. 46, No. 2: 59-63, June 2008 DOI: 10.3347/kjp.2008.46.2.59

Expressed Sequence Tags of *Trichinella spiralis* Muscle Stage Larvae

Hae Kyung Park[1,3], Seong Won Chang[2], Se Won Kang[4], Min Kyoung Cho[2], Sun Hee Choi[2], Yeon Chul Hong[5], Yong Seok Lee[4], Hae Jin Jeong[2,3] and Hak Sun Yu[2,3,*]

[1]Department of Internal Medicine; [2]Department of Parasitology, School of Medicine, Pusan National University, Busan 609-735, Korea; [3]Pusan National University Hospital Medical Research Institute, Busan 602-739, Korea; [4]Department of Parasitology and Malariology, PICR, College of Medicine and Frontier Inje Research for Science and Technology, Inje University, Busan 614-735, Korea; [5]Department of Parasitology, School of Medicine, Kyungpook National University, Daegu 700-422, Korea

Abstract: In order to obtain greater insight into the relevant genomic expression patterns of *Trichinella spiralis*, 992 expressed sequence tags (ESTs) were collected from a cDNA library of *T. spiralis* muscle stage larvae and assembled into 60 clusters and 385 singletons. Of them, 445 (44.7%) ESTs were annotated to their homologous genes, and small fractions were matched to known genes of nematodes. The annotated ESTs were classified into 25 eukaryotic orthologous groups (KOG). Cytochrome C oxidase (34 clones) was found to be most frequent species.

Key words: *Trichinella spiralis*, Expressed sequence tags, muscle stage larva, eukaryotic orthologous groups (KOG)

INTRODUCTION

Genus *Trichinella*, which consists of several species, is found to infect mammals, birds, and reptiles. These nematodes alternate enteric and skeletal muscle stages during their life cycle. *T. spiralis* is a relatively small nematode, adult female being 1.4-4.0 mm, adult male 1.4-1.8 mm, and muscle stage larvae approximately 1 mm long [1,2]. The life cycle of this nematode begins with the enteral phase of infection, when a person or an animal eats contaminated meat containing first-stage muscle larvae. Digestive juices from the stomach (pepsin and hydrochloric acid) dissolve the capsule-like cyst and release the larvae, penetrating into epithelial cells of the small intestine [3]. Shortly thereafter, the larvae molt 4 times (10 through 28 hr post-oral ingestion [POI]), and mature into adults that mate (30-34 hr POI). Female worms can produce 500-1,500 newborn larvae (immature L1) during a lifespan, prior to expulsion by the host immune system. The migratory phase of infection begins when these newborn larvae are passed into tissues, enter the lymphatic system, and then enter the general circulation at the thoracic duct. These larvae are distributed widely throughout tissues by circulation, and eventually make their way through blood capillaries into muscle fibers, so initiating the muscle phase of infection. Once in muscle fibers, they encyst, undergo development, become infective within 15 days, and remain for months to years.

Trichinella spiralis evidences several biological differences that include host specificity, phenotypic characteristics induced in infected host cells, and interactions with the host immune system. Additionally, each of these parasites must migrate extracellularly through the tissues of the host in order to infect the various host cells they inhabit [4]. The mechanisms involved in this migration remain to be elucidated. Furthermore, the role of the parasite in regulating changes in these host cells remains unknown. A number of technical obstacles exist which impede progress in many of these issues, including a paucity of information regarding the genes that comprise *Trichinella* spp [5,6].

Reverse-genetic and/or post-genomic approaches predicated on nucleotide sequence information are vital for such investigations [7,8]. However, the availability of sequence information required for such approaches remains insufficient. Expressed sequence tags (ESTs) analysis enables the collection of nucleotide sequence information regarding protein-coding regions in a rapid and efficient manner, and should be helpful in detecting and identifying genes in genomes. It also provides gene expression information for a particular type of cultivation or for a particular growth phase. For these reasons, EST analysis is conducted as a component of many genome projects [9-11]. In addition to the genes that are candidates for pathogenic factors, many novel genes for stage-specific or cell cycle regulation, in addition to a

• Received 2 February 2008, accepted after revision 20 May, 2008.
* Corresponding author (hsyu@pusan.ac.kr)

great many cell signaling-associated genes, have been identified in these parasites via EST analysis [12-14].

ESTs analysis and the single-pass sequencing of randomly selected cDNA clones have been used to characterize the transcribed genes of a variety of organisms, including parasitic nematodes, such as B*rugia malayi* [15], *Toxocara canis* [12], *Strongyloides ratti* [16], and *Anisakis simplex* [17]. Recently, some ESTs were analyzed using a new database, referred to as eukaryotic Orthologous Groups (KOG) [18,19]. In order to gain greater insight into the relevant genomic expression patterns, we generated the ESTs of *T. spiralis* third-stage larva and investigated their functions using KOG database analysis. The characterization of the EST data is expected to provide valuable insights into both the metabolism and the pathogenicity of this parasite.

MATERIALS AND METHODS

Trichinella spiralis strain maintenance and isolation

Trichinella spiralis was kindly provided from Dr. Lee of Korea University, and maintained and recovered as described previously [20]. Albino rats (SD strain) were infected with 5,000 larvae and sacrificed 2 mo after infection.

Preparation of total RNA

Approximately 50,000 muscle stage larvae of *T. spiralis* were homogenized using a tissue homogenizer. Total RNA was extracted from homogenized tissue solution using Trizol reagent, according to the manufacturer's protocols (Invitrogen Life Technologies Inc., Carlsbad, California, USA). After lysing the larvae with 10 ml of Trizol reagent, 1/5 volume of chloroform was added and the sample was vigorously mixed. Aqueous layer was collected by centrifugation at 12,000 rpm for 15 min at 4℃. Total RNA was precipitated with an equal volume of isopropanol. The precipitated pellet was washed with ice-cold 75% ethanol and air-dried. The pellet was dissolved in DEPC-treated water and incubated for 10 min at 55℃. The quantity and quality of the RNA was assessed by measuring absorbance at 260 nm and 280 nm with a spectrophotometer. To assess prepared RNA quality, 3 μg of total RNA was run in 1% denaturing formaldehyde agarose gel.

Construction of cDNA library

With 5 μg mRNA purified with mRNA isolation kit (QIAGEN, Hilden, Germany) from the 1 mg total RNA, a cDNA library was constructed in Uni-ZAP™ XR expression vector according to the manufacturer's instructions (Stratagene, La Jolla, California, USA).

Sequencing and analysis of ESTs

The chromatogram files obtained from the sequencer were initially submitted to Phred [21,22] for base calling and quality assignment. The trace files were trimmed with trim-alt 0.05 (Phred score $\geq$ 20). Each vector sequence was identified and trimmed off using cross-match software (http://www.phrap.org). EST sequences shorter than 100 bp were discarded. The EST sequences were assembled and clustered using TGI Clustering Tools (TGICL) [23] package and cap3 software. Finally, the clusters and singletons were analyzed for annotation using a homology search engine, local BLAST [24,25] against Parablast database (http://blast.inje.ac.kr/parablast) (not released) and National center for biotechnology information (NCBI) nr database, NLM, USA. The significant matches were determined when expectation value was less than $1 \times e^{-5}$ with previously report genes. Function of ESTs was predicted through KOG analysis.

RESULTS

Summary of *T. spiralis* muscle stage larva ESTs

A total of 1,010 reads were collected from the *T. spiralis* muscle stage larva cDNA library. The length of the ESTs ranged from 78 to 801 bp with an average of 652 bp. Among the 992 ESTs submitted for BLASTx searches, 448 (45.2%) ESTs matched significantly with formerly identified gene transcripts, but 544 (55.8 %) ESTs were not. These ESTs were assembled into 60 clusters and 385singletons (Table 1).

Among the ESTs matched, 445 ESTs (44.7%) appeared homologous significantly with formerly identified gene transcripts and, therefore, annotated respective gene transcripts. Of the annotated ESTs, 58 ESTs (5.9%) were homologous with *Trichinella* gene transcripts and 34 ESTs (3.4%) matched with *C. elegans* genes. The annotated ESTs were classified into 25 KOG categories: translation, ribosomal structure, and biogenesis (14.4%); RNA pro-

Table 1. Statistics of ESTs of *T. spiralis*

ST category	No. of clones (%)
Total clones sequenced	1,010
ESTs submitted for BLAST search	992
No. of EST clusters	60
Singletons	385
Annotated ESTs	445 (44.7)
ESTs with homology to *Trichinella spiralis*	58 (5.9)
ESTs with homology to *Caenorhabditis elegans*	34 (3.4)

ESTs, expressed sequence tags; BLAST, Basic local alinment and search tool.

Table 2. Frequently identified ESTs in *T. spiralis* muscle stage larva ESTs

Cluster ID	Length	E-value	Annotation	No. of ESTs
TS0799-T3.ab1	748	$1.00e^{-123}$	Cytochrome c oxidase	34
TS1030-T3.ab1	490	$6.00e^{-44}$	Ts-abundant larval transcript-1	11
TS0416-T3.ab1	782	$2.00e^{-17}$	Ts-abundant larval transcript-2	11
TS0920-T3.ab1	504	$5.00e^{-55}$	Ts-abundant larval transcript-3	8
TS0813-T3.ab1	541	$2.00e^{-48}$	NADH dehydrogenase	8
TS0753-T3.ab1	646	$2.00e^{-69}$	Histone H3.3	8
TS0126-T3.ab1	772	$5.00e^{-17}$	Cytochrome *b* protein	6
TS0971-T3.ab1	765	$6.00e^{-93}$	Small heat shock protein	5
TS0153-T3.ab1	777	$4.00e^{-51}$	MLP/CRP family	5

ESTs, expressed sequence tags; NADH, nicotinamide adenine dinucleotide; MLP, muscle lim protein; CRP, cysteine-rich protein.

cessing and modification (3.0%); transcription (0.6%); replication, recombination, and repair (1.2%); chromatin structure and dynamics (3.3%); energy production and conversion (16.8%); carbohydrate transport and metabolism (4.2%); amino acid transport and metabolism (1.5%); nucleotide transport and metabolism (0.9%); coenzyme transport and metabolism (0.6%); lipid transport and metabolism (1.5%); inorganic ion transport and metabolism (0.9%); secondary metabolite biosynthesis, transport, and metabolism (0.3%); cell cycle control, cell division, and chromosome partitioning (0.9%); nuclear structure (0.6%); defense mechanisms (0.9%); signal transduction mechanisms (10.5%); cell wall/membrane/envelope biogenesis (0.0%); cell motility (0.9%); cytoskeleton (1.8%); extracellular structures (2.1%); intracellular trafficking, secretion, and vesicular transport (3.0%); posttranslational modification, protein turnover, and chaperones (12.6%); general function prediction only (9.6%); and function unknown (7.5%). Clusters associated with cell respiration were very frequently detected in this study. In particular, cytochrome c oxidase (34 ESTs) was identified the most frequently (Table 2).

DISCUSSION

Trichinella spp. is intracellular parasites of vertebrates, in which the entire life cycle is confined to the host. These parasites alternate between enteric and muscle phases of infection, and several species have been shown to induce mortality and morbidity in humans. The parasites establish remarkable intracellular interactions with different host cell types when adapting to these diverse habitats (intestinal mucosal epithelial cells and striated skeletal muscle cells) [26,27]. A wide range of changes are induced in host cells as the result of infection with *Trichinella* spp., but these have been poorly characterized with regard to the relevant mechanisms and their significance to parasite survival. Furthermore, the role of the parasite in regulating these host cell changes remains largely unknown.

The KOG system classifies genes on the basis of orthologous relationships between genes [28]. The KOG classification of ESTs from the *Trichinella* cDNA library was utilized in this study. The largest group of ESTs was genes belonging to the group C associated with energy production and conversion. The cytochrome-associated proteins, particularly cytochrome c oxidase, were most frequently identified in this study. Eukaryotic cytochrome *c* oxidase is the terminal multicomponent enzyme of the energy-transducing mitochondrial electron transport chain [29]. It is a member of the superfamily of hemecopper-containing terminal oxidases, characterized by the presence of histidine ligands to 2 heme groups and to a CuB copper ion [30]. Although currently it is not clear why this gene is highly transcribed once the infectious larval stage is reached, we believe that the infectious larva might require much more energy than in the muscle stage; therefore it is necessary for the genes associated with energy metabolism, including cytochromes, to be abundantly transcribed.

The translation, ribosomal structure, and biogenesis group (14.4%) was expressed at the second highest level. This means that many proteins were newly generated at the infectious larva stage. The group with the third-highest expression level was also associated with protein synthesis (posttranslational modification, protein turnover, chaperones group). Genes associated with carbohydrate transport and metabolisms were identified more frequently (4.2%) than were genes associated with amino acids, lipids, nucleic acid transport, and metabolism (1.5%, 1.5 %, and 0.9%, respectively).

Serine proteinase has been reported to perform functions in host tissue invasion or molting. This serine proteinase may be attributable to infection into host muscle cells. Previously, sev-

eral studies have been conducted concerning *Trichinella* serine protease. In a comparison of two trypsin-like domains, Trap et al. [31] previously suggested that serine protease was detected at both the larval and adult stages. Immune- localization analysis also indicated that the protease was located on the inner layer of the cuticle and esophagus of the parasite, thereby suggesting a potential role in its molting and/or digestive functions [31]. The protease of *Trichinella* was well characterized by Moczon and Wranicz [32]. Under in vitro conditions, the muscle larvae of *T. spiralis* secreted minute amounts of a cysteine proteinase into the outer environment from the stichosome. Although multiple proteinase activities were histochemically detected in the somatic muscles, stichosome, midgut, and genital primordium of the muscle larvae, none of these enzymes appeared to be the one that was secreted. Several histochemically demonstrable proteinases were also detected in the cells of 48- to 72-hr-old juvenile worms. One was localized within the esophageal lumen and at or around the anterior esophagus of the larvae, where the developing stichocytes are believed to occur.

Three elongation factor genes were identified in this study. This gene was found to be important in tRNA studies. EF-Tu delivers aminoacyl-tRNAs to ribosomes in the translation system. However, unusual truncations detected in some animal mitochondrial tRNAs appear to prevent recognition via a canonical EF-Tu [33]. Also, it was previously demonstrated that the chromadorean nematode harbors two distinct EF-Tus, one of which (EF-Tu1) binds only to T-armless aminoacyl-tRNAs, whereas the other (EF-Tu2) binds to D-armless Ser-tRNAs [33,34]. Neither of the EF-Tus can bind to canonical cloverleaf tRNAs. It has been recently reported that *Trichinella spiralis* mitochondria DNA harbors genes that encode for three distinct types of tRNAs: T-armless tRNAs, D-armless tRNAs, and cloverleaf tRNAs with a short T arm [35]. The translation system of *Trichinella* species may be an intermediate evolutionary state between the canonical system which utilizes uses only cloverleaf tRNAs and the unusual chromadorean system which utilizes only tRNAs that lack the T or D arm. Mitochondrial EF-Tu species, EF-Tu1 and EF-Tu2, were reported from *Trichinella britovi*. *T. britovi* EF-Tu2 was shown to bind only to D-armless Ser-tRNA, as does *Caenorhabditis elegans* EF-Tu2 [33]. In contrast to the case of *C. elegans* EF-Tu1, however, *T. britovi* EF-Tu1 bound to all 3 types of tRNA present in the mitochondria of *Trichinella*. These results indicate that the *Trichinella* mitochondrial translation system, and in particular the tRNA-binding specificity of EF-Tu1, could represent an intermediate state between the canonical system and the chromadorean nematode mitochondrial system [33].

In the present study, we verified that the *T. spiralis* mitochondria translation system does include at least one cloverleaf tRNA. We cloned 2 EF-Tu species from *T. britovi*, a close relative of *T. spiralis*, and assessed their aminoacyl-tRNA specificity.

According to this study, 45% of the *T. spiralis* EST clusters evidenced homology with proteins from other species, whereas 55% were located in the category of unknown proteins. Among these, 3 highly transcribed genes (8-11 transcripts of 992 ESTs) were identified. Although we could not determine their functions, these genes might be centrally involved in larval infection or larval life maintenance. More information regarding these genes is required for *Trichinella* studies.

ACKNOWLEDGEMENTS

This study was supported by a Pusan National University Medical Research Institute grant (2006-50).

REFERENCES

1. Clausen MR, Meyer CN, Krantz T, Moser C, Gomme G, Kayser L, Albrectsen J, Kapel CM, Bygbjerg IC. *Trichinella* infection and clinical disease. QJM 1996; 89: 631-636.
2. Pozio E, Darwin Murrell K. Systematics and epidemiology of *Trichinella*. Adv Parasitol 2006; 63: 367-439.
3. Katz M, Despommier DD, Gwadz RW. Parasitic Diseases. 2nd ed. New York, USA. Springer-Verlag. 1989.
4. Campbell WC. *Trichinella* and trichinosis. New York, USA. Plenum Press. 1983.
5. Mahida YR. Host-parasite interactions in rodent nematode infections. J Helminthol 2003; 77: 125-131.
6. Mitreva M, Jasmer DP. Biology and genome of *Trichinella spiralis*. WormBook 2006; 1-21.
7. Hudson DF, Morrison C, Ruchaud S, Earnshaw WC. Reverse genetics of essential genes in tissue-culture cells: 'dead cells talking'. Trends Cell Biol 2002; 12: 281-287.
8. Serruto D, Rappuoli R. Post-genomic vaccine development. FEBS lett 2006; 580: 2985-2992.
9. Wolfsberg TG, Landsman D. A comparison of expressed sequence tags (ESTs) to human genomic sequences. Nucleic Acids Res 1997; 25: 1626-1632.
10. de Koning AP, Tartar A, Boucias DG, Keeling PJ. Expressed sequence tag (EST) survey of the highly adapted green algal parasite, *Helicosporidium*. Protist 2005; 156: 181-190.
11. Ramirez M, Graham MA, Blanco-López L, Silvente S, Medrano-Soto A, Blair MW, Hernández G, Vance CP, Lara M. Sequencing and analysis of common bean ESTs. Building a foundation for functional genomics. Plant Physiol 2005; 137: 1211-1227.

12. Tetteh KK, Loukas A, Tripp C, Maizels RM. Identification of abundantly expressed novel and conserved genes from the infective larval stage of *Toxocara canis* by an expressed sequence tag strategy. Infect Immun 1999; 67: 4771-4779.
13. Maizels RM, Tetteh KK, Loukas A. *Toxocara canis*: genes expressed by the arrested infective larval stage of a parasitic nematode. Int J Parasitol 2000; 30: 495-508.
14. Ng ST, Sanusi Jangi M, Shirley MW, Tomley FM, Wan KL. Comparative EST analyses provide insights into gene expression in two asexual developmental stages of *Eimeria tenella.* Exp Parasitol 2002; 101: 168-173.
15. Blaxter M, Daub J, Guiliano D, Parkinson J, Whitton C, Filarial Genome Project. Pathogen genomes and human health. The *Brugia malayi* genome project: expressed sequence tags and gene discovery. Trans R Soc Trop Med Hyg 2002; 96: 7-17.
16. Thompson FJ, Mitreva M, Barker GL, Martin J, Waterston RH, McCarter JP, Viney ME. An expressed sequence tag analysis of the lifecycle of the parasitic nematode *Strongyloides ratti.* Mol Biochem Parasitol 2005; 142: 32-46.
17. Yu HS, Park SK, Lee KH, Lee SJ, Choi SH, Ock MS, Jeong HJ. *Anisakis simplex:* analysis of expressed sequence tags (ESTs) of third-stage larva. Exp Parasitol 2007; 117: 51-56.
18. Akao T, Sano M, Yamada O, Akeno T, Fujii K, Goto K, Ohashi-Kunihiro S, Takase K, Yasukawa-Watanabe M, Yamaguchi K, Kurihara Y, Maruyama J, Juvvadi PR, Tanaka A, Hata Y, Koyama Y, Yamaguchi S, Kitamoto N, Gomi K, Abe K, Takeuchi M, Kobayashi T, Horiuchi H, Kitamoto K, Kashiwagi Y, Machida M, Akita O. Analysis of expressed sequence tags from the fungus *Aspergillus oryzae* cultured under different conditions. DNA Res 2007; 14: 47-57.
19. Schwartz EF, Diego-Garcia E, Rodriguez de la Vega RC, Possani LD. Transcriptome analysis of the venom gland of the Mexican scorpion *Hadrurus gertschi* (Arachnida: Scorpiones). BMC genomics 2007; 8: 119.
20. Wakelin D, Lloyd M. Immunity to primary and challenge infections of *Trichinella spiralis* in mice: a re-examination of conventional parameters. Parasitology 1976; 72: 173-182.
21. Ewing B, Green P. Base-calling of automated sequencer traces using phred. II. Error probabilities. Genome Res 1998; 8: 186-194.
22. Ewing B, Hillier L, Wendl MC, Green P. Base-calling of automated sequencer traces using phred. I. Accuracy assessment. Genome Res 1998; 8: 175-185.
23. Pertea G, Huang X, Liang F, Antonescu V, Sultana R, Karamycheva S, Lee Y, White J, Cheung F, Parvizi B, Tsai J, Quackenbush J. TIGR Gene Indices clustering tools (TGICL): a software system for fast clustering of large EST datasets. Bioinformatics 2003; 19: 651-652.
24. Altschul MS. The circumcision controversy. Am Fam Physician 1990; 41: 817-820.
25. Altschul SF, Lipman DJ. Protein database searches for multiple alignments. Proc Natl Acad Sci USA 1990; 87: 5509-5513.
26. Capó VA, Despommier DD, Polvere RI. *Trichinella spiralis*: vascular endothelial growth factor is up-regulated within the nurse cell during the early phase of its formation. J Parasitol 1998; 84: 209-214.
27. Appleton JA, Romaris F. A pivotal role for glycans at the interface between *Trichinella spiralis* and its host. Vet Parasitol 2001; 101: 249-260.
28. Tatusov RL, Fedorova ND, Jackson JD, Jacobs AR, Kiryutin B, Koonin EV, Krylov DM, Mazumder R, Mekhedov SL, Nikolskaya AN, Rao BS, Smirnov S, Sverdlov AV, Vasudevan S, Wolf YI, Yin JJ, Natale DA. The COG database: an updated version includes eukaryotes. BMC Bioinformatics 2003; 4: 41.
29. Capaldi RA. Structure and assembly of cytochrome c oxidase. Arch Biochem Biophys 1990; 280: 252-262.
30. Michel H, Behr J, Harrenga A, Kannt A. Cytochrome c oxidase: structure and spectroscopy. Annu Rev Biophys Biomol Struct 1998; 27: 329-356.
31. Trap C, Fu B, Le Guerhier F, Liu M, Le Rhun D, Romand T, Perret C, Blaga R, Boireau P. Cloning and analysis of a cDNA encoding a putative serine protease comprising two trypsin-like domains of *Trichinella spiralis*. Parasitol Res 2006; 98: 288-294.
32. Moczon T, Wranicz M. *Trichinella spiralis*: proteinases in the larvae. Parasitol Res 1999; 85: 47-58.
33. Arita M, Suematsu T, Osanai A, Inaba T, Kamiya H, Kita K, Sisido M, Watanabe Y, Ohtsuki T. An evolutionary intermediate state of mitochondrial translation systems found in *Trichinella* species of parasitic nematodes: co-evolution of tRNA and EF-Tu. Nucleic Acids Res 2006; 34: 5291-5299.
34. Ohtsuki T, Sato A, Watanabe Y, Watanabe K. A unique serine-specific elongation factor Tu found in nematode mitochondria. Nat Struct Biol 2002; 9: 669-673.
35. Lavrov DV, Brown WM. *Trichinella spiralis* mtDNA: a nematode mitochondrial genome that encodes a putative ATP8 and normally structured tRNAS and has a gene arrangement relatable to those of coelomate metazoans. Genetics 2001; 157: 621-637.