Experimental Parasitology 129 (2011) 120–126

Contents lists available at ScienceDirect

Experimental Parasitology

journal homepage: www.elsevier.com/locate/yexpr

Construction of a cDNA library from female adult of *Toxocara canis*, and analysis of EST and immune-related genes expressions

Rongqiong Zhou [a,b], Qingyou Xia [b,*], Hancheng Huang [a], Min Lai [a], Zhenxin Wang [a]

[a] Department of Veterinary Medicine, Rongchang Campus, Southwest University, Chongqing 402460, PR China
[b] Institute of Sericulture and Systems Biology, Southwest University, Chongqing 400715, PR China

ARTICLE INFO

Article history:
Received 20 December 2010
Received in revised form 29 May 2011
Accepted 4 July 2011
Available online 13 July 2011

Keywords:
Expressed sequence tag (EST)
cDNA library
Toxocara canis
Female

ABSTRACT

Toxocara canis is a widespread intestinal nematode parasite of dogs, which can also cause disease in humans. We employed an expressed sequence tag (EST) strategy in order to study gene-expression including development, digestion and reproduction of *T. canis*. ESTs provided a rapid way to identify genes, particularly in organisms for which we have very little molecular information. In this study, a cDNA library was constructed from a female adult of *T. canis* and 215 high-quality ESTs from 5′-ends of the cDNA clones representing 79 unigenes were obtained. The titer of the primary cDNA library was 1.83×10^6 pfu/mL with a recombination rate of 99.33%. Most of the sequences ranged from 300 to 900 bp with an average length of 656 bp. Cluster analysis of these ESTs allowed identification of 79 unique sequences containing 28 contigs and 51 singletons. BLASTX searches revealed that 18 unigenes (22.78% of the total) or 70 ESTs (32.56% of the total) were novel genes that had no significant matches to any protein sequences in the public databases. The rest of the 61 unigenes (77.22% of the total) or 145 ESTs (67.44% of the total) were closely matched to the known genes or sequences deposited in the public databases. These genes were classified into seven groups based on their known or putative biological functions. We also confirmed the gene expression patterns of several immune-related genes using RT-PCR examination. This work will provide a valuable resource for the further investigations in the stage-, sex- and tissue-specific gene transcription or expression.

1. Introduction

The nematode *Toxocara canis* is a cosmopolitan parasite of canids and has a wide range of paratenic hosts, including humans. Human infection with *T. canis* is mostly asymptomatic (Taira et al., 2004). In some individuals, however, the immune system is unable to control larval migration at the liver. In these cases, severe disease with involvement of the central nervous system and/or the eye can occur. Three syndromes of *T. canis* infection have been described: visceral larva migrans (VLM), ocular larva migrans (OLM), and covert toxocariasis (CT) (Taylor and Holland, 2001). VLM is usually detected in young children (1–5 years old) and OLM mainly affects teenagers or young adults (Alba-Hurtado et al., 2009; Arias-Irigoyen and Senent-Sanchez, 1995). Clinical symptoms are more frequently observed in children than in adults. This can be explained by a relatively high infection rate at this age. This results from the tendency of many small children eating dirt food, thus ingesting eggs from a contaminated environment (Habluetzel et al., 2003; Overgaauw, 1997). A dog infected with adult worms of *T. canis* that are shedding thousands of eggs each day in its faeces can be found. The physical environment plays a crucial role in maintaining and distributing the infective eggs of *T. canis,* although this subject remains underappreciated. *T. canis* have a developed reproduction system, so we can study the genes which are related to reproduction.

Functional genomic strategies, such as the expressed sequence tag (EST) and microarray technology approaches are powerful tools for the identification of large numbers of genes. EST analysis is not only one of the more effective methods for gene discoveries, gene expression profiling and functional genome studies, but also one of the more efficient ways for the identification of differential genes (Zhao et al., 2009; Qi et al., 2008; Abernathy et al., 2007; Malde and Jonassen, 2008; Zhou et al., 2008; Wang et al., 2007; D'Agostino et al., 2007; Chakrabarti et al., 1994). The aims of this study are to create a cDNA library of female adults of *T. canis* which is suitable for the analysis of expressed sequence tags (ESTs) and to generate an EST resource to serve as a platform for its functional genomic studies including development, metabolism and reproduction of this parasite.

* Corresponding author. Fax: +86 23 68251128.
E-mail address: xiaqy@swu.edu.cn (Q. Xia).

0014-4894/$ - see front matter © 2011 Published by Elsevier Inc.
doi:10.1016/j.exppara.2011.07.005

2. Materials and methods

2.1. Materials and RNA isolation

Female adults of *T. canis* were collected from puppies in Rongchang County, Chongqing. The samples were stored in liquid nitrogen for RNA isolation. Total RNAs were extracted using TRIzol reagent from Invitrogen. The quality and concentration of RNA samples were examined by gel electrophoresis on formaldehyde denaturalization agarose.

2.2. Construction of a cDNA library

The cDNA library was constructed using a SMART™ cDNA Library Construction Kit (Clontech) according to the manufacturer's instructions. The first strand of cDNA synthesis was synthesized using SMARTScribe™ MMLV Reverse Transcriptase, the SMART IV oligonucleotide and the CDS III/3′ PCR primer provided in the kit. The double-stranded cDNA (ds-cDNA) was synthesized by long distance PCR (LD-PCR) with the 50 PCR primer and the CDS III/3′ PCR primer using the Advantage 2 PCR kit (Clontech, USA). The product was analyzed on 1.1% Agarose gel to determine the sizes and the amount of the cDNA products before proceeding to the next step. Then the LD-PCR products were used for treatment with proteinase K. After digestion of proteinase K at 45 °C for 20 min and *Sfi* I at 50 °C for 2 h, These *Sfi* products were separated by column CHROMA SPIN-400 and ligated to the λTriplEx2 vector. The ligation products were packaged with Gigapack III Gold Packaging Extract (Stratagene). Then the mixtures were transformed into *Escherichia coli* XL-1-Blue host cell culture to determine the titers of the unamplified and amplified libraries.

2.3. Converting λTriplEx2 to pTriplEx2

The λTriplEx2 multiple cloning site (MCS) is located within an embedded plasmid (pTriplEx2), which is flanked by *lox P* sites at the λ junctions. When the recombinant phage is transduced into *Escherichia coli* strain BM 25.8, Cre recombinase is expressed in the *E. coli* BM 25.8. The conversion of a λTriplEx2 phagemid to a pTriplEx2 plasmid involved a vivo excision and circularization of a complete plasmid from the recombinant phage. Briefly, 150 μl of the eluted positive plaque was combined overnight with 200 μl of BM 25.8 host cell culture and the mixture was incubated at 31 °C for 30 min without shaking. LB broth (400 μl) was added to the mixture, which was incubated at 31 °C for an additional 1 h with shaking (225 rpm). Finally, the infected cell suspension was spread on a Luria–Bertani (LB)/ampicillin plate to obtain isolated colonies. The insert fragment sizes of the positive recombinants were analyzed by PCR amplification using the vector-specific 5′- and 3′- sequencing primers.

2.4. EST sequencing and bioinformatic analysis

Plasmid DNA of the 218 randomly selected clones was extracted and single-pass sequenced from the 5′-end only using 5′-sequencing primer which was provided in the kit. The sequencing reaction was conducted on a PTC-225 Thermal Reactor (MJ-Research) and sequencing was performed using an ABI3730 DNA Sequencer (Applied Biosystems). All sequences were compared to the Gen-Bank database using BLASTX and BLASTN in a non-redundant (NR) program. The *E*-value scores lower than 1.0×10^{-5} were considered to be significant. The BLAST search results were used to obtain further information on function and relativity with other species through the database on GenBank.

2.5. RT-PCR detection of several immune-related genes

The expression patterns of several immune-related transcription units were tested and confirmed by RT-PCR. Total RNAs of male and female adults of *T. canis* were reverse-transcribed with M-MLV reverse transcriptase (Promega) at 42 °C, respectively. cDNA equivalent to 100 ng of total RNA was used for RT-PCR amplification in a volume of 25 μl with the primers shown in Table 1. RT-PCR was performed by preheating the samples at 94 °C for 4 min, followed by 30 cycles at 94 °C for 40 s, 58 °C for 40 s, and 72 °C for 1 min, followed by 10 min of extension at 72 °C. The PCR products were separated on 1.0% agarose gels and stained with EB.

3. Results

3.1. Construction of T. canis cDNA library

The integrity of the total RNA used to construct the cDNA library was detected by gel electrophoresis on formaldehyde denaturalization agarose, and the bands of 28S, 18S and 5S RNA could be observed at the corresponding positions (Fig. 1). The majority of the cDNAs produced from LD-PCR ranged from 200 to 4000 bp in size (Fig. 2), The titer of the primary cDNA library was 1.83×10^6 pfu/mL with a recombination rate of 99.33%. PCR amplification of randomly picked clones revealed that the inserted cDNA fragments ranged from 500 to 2000 bp which reflected that the size distribution of the first-strand cDNA was convergence (Fig. 3).

3.2. ESTs information

We sequenced the 218 ESTs isolated from the library. The raw sequences were edited to remove vector sequences and the terminal sequence of low reliability by using a cross-match program in combination with manual processing. The sequences representing rRNA, mitochondrial DNA and DNA of less than 150 bases were discarded. A total number of 215 high-quality ESTs were obtained from the library. The identities of the ESTs were determined by searching the GenBank databases. A large fraction of the sequences ranged from 300 to 900 bp, and the average length of a readable sequence was about 656 bp (Fig. 4). Cluster analysis of these ESTs identified 79 unigenes (unique sequences) containing 28 contigs and 51 singletons, suggesting that the overall redundancy of the library was 36.74%. There were only two unigenes (contig1, HO243918, and contig2, HO243919) which were assembled with more than 20 ESTs. Singletons (51) constituted 64.55% of all the

Table 1
Primers used in PCR amplification of several immune-related genes.

Gene name	5′ Primer sequence	3′ Primer sequence
TCF-1 (C-type lectin family member)	TCGCCTTTTGTCTGCTTCTC	AAAATCGCCTCCTAAAATCCTC
TCF-5 (immunosuppressive ovarian message protein)	ATGAGGTCGGGTATTATGCTGC	TGGTCTGTATCGTAACAAGGTCC
TCF-64 (surface-associated antigen 2)	TTTCGCACAAGGTCCTCAAG	TCCTATTAGATTCGGATTTTCAGC
TCF-66 (cathepsin Z1 preproprotein)	AGTCTCGTTTGCCCTCCTCA	CGTCTCCTCGGTATAGATCCCAC
18S	AATTGTTGGTCTTCAACGAGGA	AAAGGGCAGGGACgTAgTCAA