SHORT REPORT: ECHINOCOCCUS GRANULOSUS FROM XINJIANG, PR CHINA: cDNAS ENCODING THE EG95 VACCINE ANTIGEN ARE EXPRESSED IN DIFFERENT LIFE CYCLE STAGES AND ARE CONSERVED IN THE ONCOSPHERE

WENBAO ZHANG, JUN LI, HONG YOU, ZHUANGZHI ZHANG, GULINUL TURSON, ALEX LOUKAS, AND DONALD P. MCMANUS

Molecular Parasitology Laboratory, Australian Centre for International and Tropical Health and Nutrition, Queensland Institute of Medical Research and University of Queensland, Brisbane, Queensland, Australia; Veterinary Research Institute, Xinjiang Academy of Animal Science, Urumqi, Xinjiang, China.

Abstract. The EG95-based vaccine protects sheep from infection with the dog tapeworm Echinococcus granulosus. The EG95 encoding gene is a member of a multigene family, several members of which are expressed in the oncosphere, believed to be the target of immunity induced by the vaccine. E. granulosus exhibits extensive intraspecific (strain) variation, and variability of the eg95 gene in different isolates of E. granulosus may directly impact the effectiveness of the EG95-based vaccine. We analyzed the eg95 gene from E. granulosus collected in Xinjiang, in northwest China, where hydatid disease is hyperendemic. The gene is expressed in oncospheres, protoscoleces, and immature and mature adult worms, and the eg95 gene family was shown to comprise two basic sequence types. Very limited sequence variation was evident in the EG95 protein from oncospheres. This high degree of sequence conservation predicts that the vaccine will continue to be effective in China and elsewhere.

In protecting sheep from infection with the dog-tapeworm E. granulosus, EG95 prevents the development of cystic hydatid disease.\(^1\)\(^-\)\(^4\) Incomplete cDNA encoding EG95 was originally isolated by screening an oncospheral cDNA library (prepared from oncospheres of New Zealand sheep strain origin)\(^1\) with affinity-purified antibodies recognizing a 24.5 kDa native egg antigen.\(^5\) The eg95 cDNA encodes a protein of 153 residues with a molecular size of 16.5 kDa. Subsequent work has shown that the eg95 cDNA belongs to a gene family.\(^6\) Four eg95-related genes were predicted to express an identical EG95 protein, and all four were shown to be expressed in the oncosphere life-cycle stage,\(^6\) which is believed to be the target of host-protective immunity induced by the EG95 vaccine.\(^1\)

It is now well recognized that E. granulosus exhibits extensive intraspecific (strain) variation that may impact the epidemiology, pathology, and control of hydatid disease.\(^7\)\(^-\)\(^9\) with important implications also for the design and development of vaccines, diagnostic reagents, and drugs. Variability of the eg95 gene in different isolates of E. granulosus may directly impact the effectiveness of the EG95-based vaccine. Here, we present new data on the eg95 gene family showing, for the first time, that it is expressed in oncospheres, protoscoleces, and immature and mature adult worms of E. granulosus. The parasites used in the analysis were collected in Xinjiang/ Uygur Autonomous Region, in far northwest China, from sheep and experimentally infected dogs (Table 1). Hydatid disease is hyperendemic in this area.\(^10\)\(^-\)\(^12\)

Sheep hydatid cysts were collected from a slaughterhouse in Urumqi, the capital city of Xinjiang. Protoscoleces were aspirated from several individual cysts taken from a number of sheep, pooled, and washed with PBS. Adult worms were collected from dogs experimentally infected with a different pool of protoscoleces on day 35 (immature adult worms) and day 62 (mature adult worms with 37.5% harboring eggs) postinfection. Worms were first released from intestinal contents by soaking in 37°C PBS, treated with 1% (w/v) NaHCO\(_3\) solution to remove canine intestinal mucus, and then rinsed 10 times in PBS. Eggs were collected from mature adult worms. Oncospheres were hatched from eggs and activated\(^13\) and purified by density-gradient separation with 100% Percoll (Sigma, St Louis).\(^14\) All parasite materials were aliquoted into cryotubes immediately after a final PBS wash and stored in liquid nitrogen until used.

Total RNAs from the different stages of E. granulosus were isolated using Trizol Reagent (Gibco, BRL Life Technologies Inc., Gaithersburg, MA) according to the manufacturer’s instructions. Approximately 1 mL of parasite material was used for extracting total RNA with 15 mL of Trizol reagent. The total RNAs (3 µg were treated with RNase-free DNase (Promega), then reverse transcribed to cDNA using the SUPERSCRIPT\textsuperscript{TM} Preamplification System (Gibco, BRL Life Technologies Inc., Gaithersburg, MD) for first-strand cDNA synthesis in a total reaction volume of 20 µL. 1 µL of the reaction was used as template DNA for polymerase chain reaction (PCR) analysis. Primers were designed according to the parent eg95 sequence (GenBank accession number X90928), upstream primer: 5'-AACAGAGACTCGCTCC-GTAAAC-3' (position 77-100 of the eg95 sequence); downstream primer: 5'-AATGCAAGTGTTGGTGC-3' (position 511–534 of the eg95 sequence). An additional primer pair (5’-GGTGTGCTATGTCGCTCAGAC-T-3’ and 5’-CAAACAGAGATTTGGTHTCTC-3’) was also designed to amplify a fragment (379 bp) of actin I cDNA (GenBank accession number L07773) of E. granulosus\(^15\) to monitor the integrity of the RNA from each stage. PCR amplifications were carried out using a DNA Thermal cycler 2400 (Perkin Elmer Cetus, Emeryville, CA) in a 50-µL reaction mixture (94°C for 3 min, then 35 cycles of 94°C for 1 min, 55°C for 2 min, and 72°C for 2 min, 30 sec, and finally 72°C for 7 min) with Taq DNA polymerase (Promega). PCR-amplified DNA fragments were purified by a MinElute Gel Extraction Kit (Qiagen, Hilden, Germany) after separation by agarose gel electrophoresis, ligated into the pGEM-T Easy\textsuperscript{TM} plasmid (Promega) and sequenced using T7 and M13 reverse primers.

We amplified eg95 cDNAs by reverse-transcriptase (RT)-PCR from each of the four life cycle stages of E. granulosus.
TABLE 1
Details of the four life cycle stages of Echinococcus granulosus used for RNA isolation

<table>
<thead>
<tr>
<th>Stage</th>
<th>Number of organisms in 1 mL volume</th>
<th>Details of collection</th>
</tr>
</thead>
<tbody>
<tr>
<td>Protoscoleces</td>
<td>1.4 × 10^5</td>
<td>Protoscoleces were aspirated from a number of sheep hydatid cysts and pooled.</td>
</tr>
<tr>
<td>Immature adult worms</td>
<td>4100</td>
<td>Worms were collected on day 35 postinfection from two dogs experimentally infected with a different pool of protoscoleces. The collected worms were pooled for analysis.</td>
</tr>
<tr>
<td>Mature adult worms</td>
<td>3300</td>
<td>Worms were collected on day 62 postinfection from ten dogs experimentally infected with the same pool of protoscoleces. The worms were pooled for analysis.</td>
</tr>
<tr>
<td>Oncospheres</td>
<td>4 × 10^7</td>
<td>Oncospheres were obtained from the mature adult worms collected on day 62 postinfection. The collected oncospheres were pooled for analysis.</td>
</tr>
</tbody>
</table>

(Figure 1). The predicted size of the RT-PCR product was 458 bp. The amplified sequence covers the linear immunogenic regions of the EG95 protein^{16–17} that include the peptide sequences TETPLRKHFNLTPV (peptide 6), SLKAVNPSDPLVYKROTAKF (peptides 12 and 13), DIETPRAGKKESTVMTSGSA, (peptides 21 and 22), and SALTSAIAGFVFSC (peptide 24).^{16} The position of these sequences in the EG95 protein is highlighted in Figure 2.

RT-PCR using 35 cycles of PCR indicated that eg95 is expressed in all four life-cycle stages examined (Figure 1A, II): more-abundant expression was noted in hatched/activated oncospheres after 26 cycles of PCR (Figure 1A, I). However, although an abundant 0.8–0.9-kbp RNA species was identified in hatched and activated oncospheres on a Northern blot (Figure 1B), this is not unexpected as, although the precise function of EG95 is not known, it likely is involved in penetration of the epithelial border of the intestinal villi, with eg95 being expressed most abundantly in hatched and activated oncospheres that are infective to the intermediate host. The primers used in the RT-PCR analysis, designed according to the parent eg95 sequence, span an intron of 215 bp in the eg95 gene. But sequencing (see below) of all clones arising from RT-PCR products of the predicted size (458 bp; arrowed, Figure 1A) indicated that they contained cDNA sequences only, without the intron. Additional products were visible in the RT-PCR reactions (Figure 1A). Despite the fact that the original RNA preparations were treated with RNase-free DNase, these additional bands were probably generated from contaminating genomic DNA. Indeed, the band immediately above the eg95 cDNA band at 458 bp in the lane containing products obtained with protoscolex RNA was sequenced and was shown to have the eg95 sequence containing the intron. Further bands above this fragment evident in the immature adult worm and protoscolex stages may be non-specific bands due to the contaminating DNA or may be due to primer-dimerization of the eg95 primers we designed. As the EG95 host-protective antigen is expressed by a family of genes, some of the genes may have different intron lengths and/or some may have repetitive DNA inserted into introns, which also may account for the larger bands observed.

RT-PCR products banding at 458 bp from each of the four life-cycle stages were excised and ligated into T-ended plasmids; 10–15 clones containing inserts were randomly selected for sequencing for each life cycle stage RT-PCR product. The different sequences were compared with other eg95 sequences available in the EMBL, GenBank™, and DDBJ databases. Figure 2 shows an alignment of the deduced amino acid sequences. Although there are a number of amino acid substitutions in several of the individual sequences, the EG95
protein family comprises two basic sequence types. One is present in the protoscolex stage (the sequence designated as EG95-P7 in Figure 2), and the other occurs in oncospheres (sequences designated as EG95-O3, EG95-O4, and EG95-O6 in Figure 2). It should be emphasized that the immature and mature adult worms and oncospheres used in the study resulted from the same collection of pooled protoscoleces originally used to experimentally infect dogs. Protoscolex RNA was extracted from a different pool of parasites collected on a separate occasion. Only one sequence (EGP5-P7) was evident in the clones originating from the protoscolex stage. It is noteworthy that this sequence was identical to a sequence present in mature adult worms (EG95-M7 in Figure 2), indicating that the two different E.granulosus sources were very similar genotypically. Of 13 clones sequenced from the oncosphere RT-PCR product, the most highly represented sequence was EG95-O4 (9/13) followed by EG95-O6 (3/13) and EG95-O3 (1/13). Over 20% amino acid differences are evident between the two types, but the most significant difference is that the former has an insertion of seven amino acids in the C-terminal (Figure 2). Both sequence types also occur in immature adult worms (sequences represented as EG95-I4 and EG95-I7 in Figure 2) and mature adult worms (represented as EG95-M2, EG95-M5, and EG95-M7 in Figure 2).

In addition to the two basic sequence types, a number of amino acid substitutions were evident between members of the EG95 protein family in different life cycle stages. It is well recognized that PCR analysis may cause errors in generating new duplicated fragments. We used the same source of Taq polymerase (Promega) to amplify a number of other cDNAs sized 1.5–1.8 kb for subcloning into expression vectors. Subsequent sequencing analysis of the clones did not indicate any nucleotide substitutions. (Data not shown.) We are therefore confident that the polymorphisms we recorded in the eg95 sequences did not result from PCR or some other technical artifact, particularly as the same substitutions we provided in sequences from different life cycle stages. One example is EG95-O6, present in oncospheres, and EG95-M5 from mature adult worms, where there is a histidine/arginine substitution at the 20th position from the N-terminal end compared with the other EG95 sequences. Similarly, there is an aspartic acid/glutamic acid substitution at the 24th residue position from the N-terminal end in two oncosphere (EG95-O3/O6) and two mature worm (EG95-M2/M5) sequences compared with the other EG95 sequences (Figure 2). Neither of these substitutions occurs in linear immunogenic regions of the EG95 protein. With the exception of a single substitution (alanine/valine) in the end amino acid of peptide 21/22 in
EG95-O3, the oncospheral EG95 sequences share absolute identity to the parent EG95 protein sequence in the four immunogenic regions mapped previously. As we used samples of pooled parasites for each of the four life cycle stages, we do not know whether the variation we see in the eg95 cDNA sequences is due to genetic differences among the genes within an individual or among individuals in the sample population.

However, it is noteworthy that, overall, in the oncospheral sequences examined, very limited sequence variation (0–3 amino acid substitutions) was evident in the EG95 protein. Although protection against hydatid disease induced with the EG95 vaccine is now thought to be associated with conformational, rather than linear, epitopes, the high degree of sequence conservation, including the mapped immunodominant regions of the EG95 protein family, in oncospheres supports the widespread use of the EG95-based vaccine and predicts that it will continue to be effective in China and elsewhere. Indeed, the vaccine already has proven highly effective (96–100% protection) in challenge trials in China, Australia, New Zealand, and Argentina.

Received April 29, 2002. Accepted for publication September 11, 2002.

Financial support: This work was supported by the Wellcome Trust (WZ and DPM), the Australian National Health and Medical Research Council (DPM), and by Australian International Postgraduate Research (IPRS) and University of Queensland International Postgraduate Research Scholarships (UQIPRS) to WZ.

Authors’ addresses: Jun Li, Hong You, Alex Loukas, and Donald P. McManus. Molecular Parasitology Laboratory, Australian Centre for International and Tropical Health and Nutrition, The Queensland Institute of Medical Research and The University of Queensland, 300 Herston Road, Brisbane, Queensland 4006, Australia, Telephone: 61-7-3362 0401, Fax: 61-7-3362 0104, E-mail: donM@qimr.edu.au

Reprint requests: Professor Don McManus, Molecular Parasitology Laboratory, Queensland Institute of Medical Research, 300 Herston Road, Brisbane, Queensland 4006, Australia, Telephone: 61-7-3362 0401, Fax: 61-7-3362 0104, E-mail: donM@qimr.edu.au

REFERENCES