University of Nebraska - Lincoln

DigitalCommons@University of Nebraska - Lincoln

Faculty Publications from the Harold W. Manter Laboratory of Parasitology

Parasitology, Harold W. Manter Laboratory of

2002

Identification and Characterisation of a cDNA Sequence Encoding a Glutamic Acid-Rich Protein Specifically Transcribed in *Trichinella spiralis* Newborn Larvae and Recognised by Infected Swine Serum

Dante S. Zarlenga
United States Department of Agriculture

Patricia Boyd
United States Department of Agriculture

J. Ralph Lichtenfels
Animal Parasitic Disease Lab, ARS, United States Department of Agriculture, 2jrcgl@gmail.com

Dolores Hill
Animal Parasitic Disease Lab, ARS, United States Department of Agriculture

H. Ray Gamble
National Research Council, rgamble@nas.edu

Follow this and additional works at: https://digitalcommons.unl.edu/parasitologyfacpubs

Part of the Parasitology Commons

Identification and characterisation of a cDNA sequence encoding a glutamic acid-rich protein specifically transcribed in *Trichinella spiralis* newborn larvae and recognised by infected swine serum

Dante S. Zarlenga[a], Patricia Boyd[a], J. Ralph Lichtenfels[b], Dolores Hill[b], H. Ray Gamble[c]

[a]*Immunology and Disease Resistance Laboratory, USDA, ARS, ANRI, Beltsville, MD 20705, USA*
[b]*Parasite Biology, Epidemiology and Systematics Laboratory, USDA, ARS, ANRI, Beltsville, MD 20705, USA*
[c]*National Research Council, Washington, DC 20418, USA*

Abstract

Presently, little is known of the mechanism by which *Trichinella* penetrates and modulates reprogramming of muscle cells. In light of evidence demonstrating strong protective characteristics of antigens derived from this stage, understanding this process may shed light on potential targets for effective abatement of infection. To this end, a PCR-derived cDNA expression library was constructed using 0.5 μg of total RNA from *Trichinella spiralis* newborn larvae. The library consisted of >125,000 insert-containing clones. Approximately 40–50×10^3 clones were screened immunologically using sera from pigs experimentally infected with 7,000 *Trichinella* L1. Multiple clones reacting positively with the swine infection serum and encoding portions of a glutamic acid-rich protein were identified. Northern and Southern blots indicated at least two distinct genes that encoded the glutamic acid-rich proteins and that these genes were transcribed specifically in the newborn larvae stage. cDNA sequence data predicted open reading frames of 1,497 and 1,716 bp generating proteins of 498 amino acids and 571 amino acids, respectively. Both sequences consisted of approximately 39% glutamic acid and 16% serine residues, and differed by the presence of a 219 bp fragment present in the 1,716 bp sequence that was absent from the 1,497 bp sequence. PCR data indicated that additional isoforms exist within this gene family that are different in length from those described above. In addition, it was found that more than one isoform can exist within a single worm and that this pattern can vary between individual worms within a population. Mouse antibodies to recombinant antigen localised the glutamic acid-rich proteins to the periphery of the developing stichocyte cells within the newborn larvae consistent with the hypothesis that the newborn larval antigens are secreted. Published by Elsevier Science Ltd. on behalf of Australian Society for Parasitology Inc.

Keywords: *Trichinella*; Nematode; Newborn larvae; cDNA library; Cloning; Glutamic acid-rich

1. Introduction

Nematodes of the genus *Trichinella* form a unique class of intracellular parasites in that they complete their entire life cycle within a single host. Except for the short time following extrusion from the female, larvae are confined to the skeletal muscle cell. In general, successful infection of the muscle cell is paralleled by rapid changes in other cellular events that invoke a reprogramming of the cell, also referred to as basophilic transformation (Gabryel and Gustowska, 1970), host cell-cycle re-entry (Jasmer and Neary, 1994) and dedifferentiation (Vassilatis et al., 1996). Cell reprogramming is accompanied by DNA replication (Jasmer, 1993), enlargement and migration of infected cell nuclei (Despommier et al., 1991), reduction in myofibrillar proteins (Jasmer et al., 1990), termination of the host cell development in the G$_2$/M phase with cessa-

tion of muscle gene transcription (Jasmer, 1993), and formation of a collagen-based envelope or cyst wall, around the cell. It has been proposed that the purpose for the cyst around this newly developed 'nurse cell' is to serve as a protective and metabolic niche for the parasite which occurs in the absence of muscle cell death. However, two species of *Trichinella*, *Trichinella pseudospiralis* and *Trichinella papuae* (Pozio et al., 1999; La Rosa et al., 2001), that also induce cellular reprogramming, survive well within infected animals, but fail to invoke cyst formation as a consequence of infection, bring into question the primary role of encystment as a protective environment for the parasite.

To date, substantial immunological, biochemical and molecular studies have focused on the muscle larvae due in part to the ease with which this stage can be propagated and isolated. In addition, *Trichinella spiralis* muscle larvae have garnered interest from investigators because of the

strong host immune responses that develop as a consequence of antigens secreted from this stage. It had been postulated that the muscle larval ES antigens were responsible for reprogramming infected striated muscle cells; however, evidence has been advanced indicating that it is the newborn larvae that initiate these cellular events (Jasmer and Neary, 1994; Gabryel et al., 1995). Wranicz et al. (1998) showed further that the genesis of cellular reprogramming can occur within muscle cells that do not directly contact newborn larvae but are proximal to those recently infected with *Trichinella*, suggesting that secretions from the newborn larvae present within the local milieu surrounding the newly infected cell are capable of inducing changes in the adjacent cells as well. Changes in the morphology of neighbouring muscle cells persist for up to 30 days following parasite penetration of the muscle cell, after which a reversion to normal morphology occurs. Infectivity appears greatest during the first 9 h. following larval birth, then falls off precipitously thereafter, suggesting that the newborn larvae are likely born with and rapidly lose their transformation potential (Wranicz et al., 1998).

In contrast to the moderate protective responses elicited by muscle larval ES antigens in pig immunisation trials, pigs receiving crude extracts of newborn larval antigens were greater than 85% protected against challenge infections suggesting that surface proteins were among the most protective of this mix (Marti et al., 1987). These results argued that immunity to newborn larvae may confer greater resistance to challenge than does intestinal immunity. This work supported earlier studies showing that passively transferred rodent antibodies (Moloney and Denham, 1979; Perrudet-Badoux et al., 1981; Ortega-Pierres et al., 1984) and newborn larvae given i.v. (Despommier, 1971) were also protective.

The above information support the need for understanding the mechanisms by which newborn larvae penetrate and reprogram the muscle cells as well as confer protection to the infected host. To date, however, no work has been performed studying genes associated with this stage of *Trichinella* because of difficulties in obtaining sufficient material for genetic analysis. The research presented herein is the first such report demonstrating not only characterisation of a family of genes encoding glutamic acid and serine-rich proteins specific to this stage, but the identification of these genes from a PCR-derived expression library using swine infection sera as a screening agent.

2. Materials and methods

2.1. Parasite isolation

Trichinella spiralis (ISS 31) was maintained by serial passage in female Sprague–Dawley rats. Muscle larvae were recovered by a 3 h pepsin:HCl (1% v/v) digestion of infected rat muscle 30 days post infection (p.i.), and used to generate adult parasites and newborn larvae. For the production of adult parasites and newborn larvae, rats infected orally with 7,000 muscle larvae were killed on the 6th day p.i. Adults were recovered from washed intestines and cultured overnight for the production of newborn larvae as described (Marti et al., 1987). Newborn larvae were separated from adults by twice passing culture fluids through a 40 μm mesh filter. Preparations were confirmed microscopically to contain less than 5% adult parasites. Larvae were either frozen at −70°C, or immediately suspended in Trizol Reagent (Invitrogen) for RNA isolation. Adults remaining after culture and substantially devoid of embryonated eggs were used for isolating RNA. Muscle larvae not used to infect rats were frozen for the isolation of muscle larvae RNA and genomic DNA. Individual adult parasites and muscle larvae were morphologically distinguished as male and female by the presence of the caudal, bilobed, copulatory bursa in adult males (Lichtenfels et al., 1983). Male and female muscle larvae were separated based upon characteristics of the proximal end of the reproductive tracts near the junction of the intestine and stichocyte-esophagus, and by the relative lengths of the rectum (Kozek, 1975). Individual parasites were immediately processed for DNA isolation.

2.2. Nucleic acid isolation

Total RNA was isolated from newborn larvae by sonication in Trizol Reagent (Invitrogen) as described by the manufacturer. RNA from muscle larvae and adult parasites was generated by proteinase K:SDS digestion of parasite material followed by phenol and phenol:chloroform extraction and ethanol precipitation (Dame et al., 1987). All RNA samples were DNAse treated, extracted with phenol:chloroform, ethanol precipitated and dissolved in sterile water. For genomic DNA from muscle larvae, a proteinase K:SDS digest was RNAse A/T1 treated, re-extracted and ethanol precipitated. Total nucleic acids were isolated from microscopically selected individual worms in 10 μl of PCR buffer (10 mM Tris–HCl, pH 8.3, 50 mM KCl, 1.5 mM MgCl$_2$, 0.01% (w/v) gelatin) as described (Zarlenga et al., 1999). Samples were stored at −20°C prior to PCR. Southern and Northern blots were carried out using Nytran™ membranes as described (Southern, 1975). For Southern blots, approximately 10 μg of *T. spiralis* muscle larvae genomic DNA was used for each restriction enzyme digestion. Northern blots contained approximately 12 μg of total RNA per lane. Radiolabelled probes were synthesised from a partial clone, designated NBL4.1.4, in the presence of [^{32}P] dCTP and random primers as described by the manufacturer (Invitrogen). Blots were exposed overnight to Kodak Biomax-MS film.

2.3. Library construction and immunoscreening

A cDNA library was constructed in the *Eco* RI site of γgt11 using the SMART™ I PCR cDNA construction kit (BD Biosciences Clontech) and 0.5 μg of total newborn

larval RNA. PCR amplification was carried out for 22 cycles of 95°C for 15 s followed by 68°C for 5 min in a total volume of 100 μl using Advantage 2 Polymerase Mix (BD Biosciences Clontech) and supplied primers. Manufacturer recommendations were followed for library construction except for the ligation of hemi-phosphorylated *Eco* RI/ *Not* I adaptors (Invitrogen) and their purification from double-stranded cDNA (dscDNA). To this end, unincorporated adaptors were removed by two passages through a 1 cm Sephacryl-S400 HR spin column (Promega) using a TOMY swinging bucket rotor at $800 \times g$ for 5 min. The purified dscDNA was ethanol precipitated using 1 μg of phosphatased γgt11 arms as carrier and ligated overnight in the presence of T4 DNA ligase in a total volume of 5 μl.

The ligated dscDNA was packaged, used to infect *Escherichia coli* Y1090 cells, and immunologically screened with serum from pigs infected with 7,000 *T. spiralis* muscle larvae and bled 8 weeks p.i. Approximately 0.5×10^5 clones were transferred to isopropyl-1-thio-β-D-galactopyranoside (IPTG)-impregnated nitrocellulose filters and screened overnight with swine infection serum diluted 1:50 in $1\times$ blocking solution (Kirkegaard and Perry) followed by 1:500 dilution of peroxidase-conjugated, rabbit anti-swine in $1\times$ blocking solution as described (Zarlenga et al., 1994).

2.4. Polymerase chain reaction

Positive clones picked through immunoscreening the cDNA library were subcloned by PCR using the Advantage 2 Polymerase Mix (BD Biosciences Clontech). Reactions were performed in 50 μl containing 200 μM of each dNTP, 0.25 μM each of γgt11 primers #483 (5′ AAAGAATTCGACGACTCCTGGAGCCCG) and #484 (AAAAGTCGACACCAGACCAACTGGTAATG), and manufacturer provided buffers. Enzymatic amplification was performed for 30–35 cycles under the following conditions: 94°C for 1 min, 60°C for 1 min and 72°C for 2 min. PCR products were separated on 2% MetaPhor agarose gels, and visualised by ethidium bromide staining prior to gel purification and insertion either into the pCR 2.1 vector by TA cloning, or into the *Eco* RI/*Sal* I site of pMal-c2 (New England Biolabs) for protein expression. Positive clones were sequenced automatically on an ABI Sequencer model 3100. A full-length sequence was identified after further library screening. Multiple isoforms were cloned by PCR amplification directly from cDNA using forward primer #661 (5′ AGA*CATATG*TGGCTTTTCCGCTCTCC) and reverse primer #662 (AAA*GCGGCCGC*ACA-CACTTTCATCCGATTCTTCC), which bind at the start and stop sites of the 1,497 bp clone, and contain *Nde* I and *Not* I restriction sites (italicized), respectively. For cloning the multiple isoforms, PCR extension was increased to 3 min. Examination of individual parasites for isoform profiles was performed using PCR primers #599 (5′AGATACATCTGAGGAACAAGC) and #662, the

above described amplification conditions, and DNA (3 μl) from individual adult parasites or muscle larvae per reaction.

2.5. Protein expression and antibody production

Partial newborn larval clone, NBL4.1.4, which contained a portion of the glutamic acid-rich region, was subcloned into the *Eco* RI/*Sal* I site of pMal-c2. The cDNA was expressed as an N-terminal, maltose-binding fusion protein under the control of the lac repressor, and affinity purified over an amylose column as defined by the manufacturer (New England Biolabs). The approximate concentration of purified recombinant protein was determined spectrophotometrically (Stoschek, 1990). The NBL4.1.4 protein was assayed by Western blot as described (Zarlenga et al., 1994). Antiserum to the NBL4.1.4 protein was produced in Swiss–Webster female white mice by i.p. injection (3×50 μg protein/dose) at 2-week intervals using alum (400 μg/injection) as adjuvant. Sera were collected before inoculation and 1 week after the final inoculation.

2.6. Immunohistochemistry

Trichinella newborn larvae were fixed with 3% paraformaldehyde and 0.5% gluteraldehyde in 0.2 M sodium cacodylate buffer, pH 7.4, and processed for light microscopy (Hill et al., 1990). Fixed worms were dehydrated in ascending steps of ethanol and embedded in LR White resin (Polysciences). Thick sections (80 nm) were cut on a Porter-Blum Ultramicrotome, collected onto uncoated glass slides and processed for immunohistochemistry using mouse anti-NBL4.1.4 (1:100) as the first-step antibody, and biotinylated goat anti-mouse (1:1000) as the second-step antibody. Sections were incubated in streptavidin colloidal gold with 10-nm particles (Sigma) followed by silver amplification with IntenSE M (Amersham). All incubations were performed at room temperature. Sections were counterstained with Ladd's Multistain, mounted with Permount, and photographed using an Olympus AH-2 photomicroscope with bright-field optics. Control sections were treated identically as above except for the use of normal (pre-immunisation) mouse sera as the first-step antibody.

3. Results

The newborn larvae cDNA library consisted of approximately 1.25×10^5 insert containing clones. Immunological screening of 40–50,000 plaques with swine infection serum resulted in the identification of 14 positively reacting clones, the most strongly of which were encoded by portions of the same or similar genes. Southern blots screened with a radiolabelled probe NBL4.1.4, made from one of the subcloned, abundant sequences, generated a profile suggestive of a gene family showing multiple bands predominantly within the *Hind*III-digested DNA (Fig. 1A). Northern blots

containing total RNA from newborn larvae, muscle larvae and adult parasite exhibited two bands migrating at approximately 1,550 and 1,750 bases in the RNA from newborn larvae only (Fig. 1B). The radiolabelled probe spanned bases 202–765 (Fig. 2) and contained a single, internal *Hind* III site.

Sequence analysis of a full-length clone, designated NBL1500, indicated an open reading frame of 1,497 bp
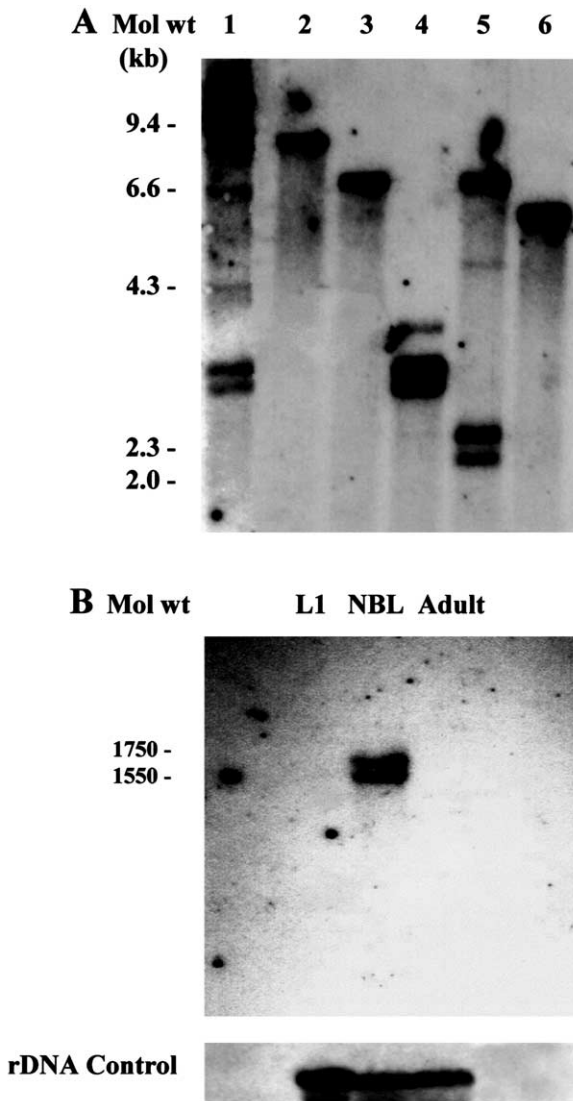

Fig. 1. A Southern blot of *Trichinella spiralis* genomic DNA. Single and double restriction enzyme digests were performed on 10 μg of genomic DNA as follows: (1) *Hind* III; (2) *Eco* RI; (3) *Eco* RV; (4) *Hind* III/*Eco* RI; (5) *Hind* III/*Eco* RV; and (6) *Eco* RI/ *Eco* RV. Lane 1 required over exposure to better reveal the DNA banding pattern. (B). Northern blot of *T. spiralis* RNA. Total RNA (12 μg) from *Trichinella* muscle larvae (L1), newborn larvae (NBL) and adult parasite were separated on a formaldehyde denaturing agarose gel. In both figures A and B, nucleic acids were transferred to Nytran™ and screened with [³²P]dCTP radiolabelled NBL 4.1.4, a partial cDNA clone of the NBL 1500 sequence which spans bases 202-765 (see Fig. 2). For control of RNA loading, the Northern blot was rescreened with a probe, designated pTsr 2.2, encoding the full-length small subunit rDNA from *T. spiralis* (Zarlenga and Murrell, 1989).

(Fig. 2) encoding a predicted protein sequence of 498 amino acids with an isoelectric point of 3.4. PCR amplification of cDNA using gene specific primers #661 and #662 that bind at the putative start and stop sites produced at least two dominant fragments; one coinciding with the 1,497 bp fragment and the other migrating at approximately 1,700 bp. Sequence data from the larger fragment, designated NBL1700, indicated an open reading frame of 1,716 bp encoding a protein of 571 amino acids (Fig. 2), and as with the 1,497 bp fragment, consisted of approximately 39% glutamic acid and 16% serine residues. Both the NBL1500 and NBL1700 sequences were essentially the same except for a 219 bp (73 amino acid) region in NBL1700 that was absent from the NBL1500. The preponderance of the glutamic acid in both sequences was nearest the C-terminus where the local concentration of both glutamic acid and total acidic residues increased to 48 and 53%, respectively. The pentapeptide repeat, ASEEX, appears frequently throughout both sequences where X is predominantly K, E or Q. A variation of the repeat, ASQEE, also appears, though it is substantially less abundant. Two additional, imperfect repeats 31 amino acids (Repeat A; Fig. 3A) and 40 amino acids (Repeat B; Fig. 3B) in length, are also present within the sequences from newborn larvae where Repeat B within NBL1500 exists in tandem and is interrupted by the 73 amino acid insert present in NBL1700. Within each of the two sequences, there exist seven putative N-glycosylation sites and a hydrophobic region within the first 25 amino acids proximal to the N-terminus. Analysis of multiple clones of both the NBL1500 and NBL1700 cDNAs did not reveal absolute sequence identity within the 1,000–1,200 bp closest the 3' terminus (data not shown). Artifacts derived from PCR could not explain all of the inconsistencies. We instead interpret this as variability among cDNA isoforms within and between individual parasites (discussed below). Thus, the NBL1500 and NBL1700 sequences illustrated in Fig. 2 should be regarded as representative of a family of gene sequences within the newborn larval stage.

Amplification of a portion of the newborn larval sequences spanning bases 519 (primer #599) and the 3'-terminus (primer #662) using DNA derived from individual adult parasite, produced multiple banding patterns. The predicted 979 and 1,198 bp fragments appeared in all adult female samples, but an additional weak band migrating at approximately 1,030 bp was evident in some worms (Fig. 4A). Both the 979 and the 1,198 bp fragments were found when males were assayed within the population; however, not all parasites contained both fragments. In addition, a fourth band appeared in one male parasite migrating at approximately 1,140 bp. Thus, at least four distinct banding patterns emerged from the adult male samples. Within the male muscle larvae, a fifth pattern was observed exhibiting the 979 bp fragment only (Fig. 4B). Except for one tested sample, PCR products amplified from female muscle larvae were similar to those generated from the adult female parasites.

```
     M  W  L  F  R  C  P  I  Y  F  V  L  L  Q  L  F  F  L  T  F        20
     atgtggcttttccgctgtccaatttatttcgtgctgctgcaactattttttttaactttt     60
     L  T  V  T  S  S  N  A  I  P  G  R  S  S  S  R  L  R  L  L
     ttaacagttacttccagcaatgccattcctggacgatcctctagtcggctgagattatta
     E  R  Y  D  S  L  P  S  L  R  S  H  S  E  D  R  Y  D  D  G
     gaaagatatgactcgttaccaagtttgcgaagtcattcggaagaccgatacgatgatgga
     V  D  R  K  W  K  K  R  E  G  N  S  D  D  I  C  T  E  D
     gtcgatcgtaaatggaaaaaacgagaaggaaattcagatgatatttgtacggaagatgaa
     T  T  V  I  E  K  E  S  E  N  G  V  D  K  E  K  P  T  S  K       100
     actactgttattgaaaaagaatctgaaatggagtagataaagaaaagccaacaagtaag    300
     E  E  S  G  E  K  T  S  Q  E  K  E  S  E  E  K  S  S  Q  E
     gaagagtcaggagaaaaaacatctcaagaaaaagaatcagaagaaaatcatcacaagaa
     K  D  E  D  K  S  E  S  E [A  S  E  E  K] D  V  S  Q  E  Q
     aaagatgaggat100aaatcagaaagtgaagcatctgaagaaaaagacgtatctcaggaacaa
     N  S  K  E  E  K  G [A  S  E  E  D] E  D  T  P  E  E  Q  N
     aattctaaagaagaaaaaggtgcatctgaagaagatgaagatacacctgaagaacaaaac
     S  K  E  E  N  G  S  S  E  D  D  E  D [A  S  E  E  Q] A
     tcaaaagaagaaaacggatcatctgaagaagatgatgaagatacatctgaggaacaagct
     S  N  E  E  K  E [A  S  E  E  K] N  T  V  S  E  E  R  K  G     200
     tcaaatgaggaaaaagaagcatctgaagaaaaaaatactgtttctgaggaaagaaaaggt    600
    [A  S  E  E  E] D  E  E  K  D  D  G  H  E  S  E  V  E  S  Q
     gcatctgaagaagaagatgaagaaaaagatgatggacatgaaagtgaagtagaaagtcaa
    [A  S  E  E  Q] T  T  E  E  G [A  S  E  E  E] D  E  E  S [A
     gcatctgaggaacaaaccactgaagaaggtgcatcagaagaggaagatgaagaaagtgca
     S  E  E  Q] T  S  E  G  E  K  G [A  S  Q  E  E] E  E  D
     tctgaagaacaaacatctgaaggagaagaaaaaggcgcatctcaggaagaagaggaagat
     E  G  N  E  Q  S  E  V  E  S  Q [A  S  E  E  Q] T  S  E
     gaaggaaatgaacaagaaagtgaagtagaaagtcaagcatctgaggaacaaacctctgaa
     E  E  S [A  S  E  E  E] D  E  E  N  E  S  K  E  Q  T  T       300
     gaagaagaaagtgcatcagaagaggaagatgaagaaatgaatctaaggaacaaaccact    900
     E  E  E  S [A  S  E  E  E] D  E  E  S  A  S  E  R  E  E
     gaagaagaagaaagtgcatcagaagaggaagatgaagaaagtgcatctgaaagagaagaa
     K  N [A  S  Q  E  E] E  E  D  E  G  N  E  S  K  E  Q  T  T
     aaaacgcatctcaggaagaagaggaagatgaaggaaatgaatctaaggaacaaaccact
     E  E  E  E  S [A  S  E  E  E] D  E  E  S  V  S  E  E  Q  T
     gaagaagaagaaagtgcatcagaagaggaagatgaagaaagtgtatctgaagaacaaaca
     S  E  G  E  E  K  G [A  S  Q  E  E] E  E  D  E  G  N  D  Q
     tctgaaggagaagaaaaaggcgcatctcaggaagaagaggaagatgagggaaatgaccaa
     E  S  E  V  E  S  Q [A  S  E  E  Q] T  S  E  E  E  G [A  S      400
     gaaagtgaagtagaaagtcaagcatctgaggaacaaacctctgaagaagaaggtgcatca   1200
    [E  E  E] D  E  E  N  E  S  E  E  Q  T  T  E  E  S [A  S
     gaagaggaagatgaagaaaatgaatctgaggaacaaaccactgaagaagaaagtgcatca
     E  E  E] D  E  E  S [A  S  E  G  E] K  N [A  S  Q  E  E]
     gaagaggaagatgaagaaagtgcatctgaaggagaagaaaaaaacgcatctcaggaagaa
     E  E  D  E  G  N  E  Q  S  E  V  E  S  Q [A  S  E  E  Q]
     gaggaagatgaaggaaatgaacaagaaagtgaagtagaaagccaagcatctgaggaacaa
     T  S  E  E  E  K  E  G [A  S  Q  E  E] D  E  E  N  E  S
     acctctgaagaagaagaaaaagaaggtgcatcacaagaggaagatgaagaaatgaatct
     E  E  Q  T  S  E  E  E  E  G [A  S  E  E  E] D  E  E  S        500
     gaggaacaaacctctgaagaagaggaagaaggtgcatctgaagaggaagatgaagaaagt   1500
     A  F  E  E  Q  T  S  E  E  E  E  K  G [A  S  Q  E  E] E
     gcatttgaggaacaaacatctgaagaagaagaagaaaaaggtgcatctcaggaagaagag
     E  D  E  E  N  E  Q  S  E  V  E  S  Q [A  S  E  E  Q] T
     gaagatgaagaaaatgaacaagaaagtgaagtagaaagtcaagcatctgaggaacaaacc
     S  E  E  G [A  S  E  E  G] Q  D [A  S  E  E  E] D  E  D
     tctgaagaagaaggtgcatcagaagaaggacaagatgcatctgaggaagaagacgaagat
     E  S  E  E  E  S  D  E  S  V  -                                571
     gaatctgaagaggaagaatcggatgaaagtgtgtga                          1716
```

Fig. 2. DNA and predicted amino acid sequence of the NBL1500 and NBL1700 cDNAs. Underlined sequences beginning at bases 1, 519 and 1,692 are the binding sites of primers 661, 599 and 662, respectively, where primer 662 is complementary to that shown. DNA and amino acid sequences highlighted in bold type reflect the region present in NBL1700 only. Boxed amino acid sequences represent randomly dispersed ASEEX and ASQEE pentapeptide repeats. The NBL1500 and NBL1700 sequences appearing in Fig. 2 have been assigned GenBank accession numbers AY092819 and AY092820, respectively.

**Location
(aa)**　　　　　　　　　　　　**Repeat A**

```
449    ESEVESQASEEQTSEEEEKEGASQ--EEDEENE
214    ..............T..G----..E--......
266    ...................--....E--.......
381    ...................---...E--.......
480    --------............-...E--.....
500    ------S.F.........EK....E........
528    ...................---...EE
```

**Location
(aa)**　　　　　　　　　　　　**Repeat B**

```
295    SKEQTTEEEESASEEEDEESASER-----EEKNASQEEEEDEGNE
335a   ..........................G-----...............
335b   ...................V..EQTSEG...G...........D
```

Fig. 3. Two long amino acid repeats within the predicted NBL1500 and NBL1700 sequences. The repeat number defines the amino acid location relative to Fig. 2. In Repeat B, 295 is present in both newborn larval sequences. Repeats 335a and 335b are present in NBL1500 and NBL1700, respectively, where the underlined amino acids in 335b define a portion of the 73 amino acid insert in NBL1700 beginning with residue 'V'. Amino acid identity is indicated by '·' and gaps are represented by '-'.

Western blots of NBL4.1.4 and factor Xa-digested NBL4.1.4 indicated that under the transfer conditions used, the incompletely expressed protein, unlinked to the maltose-binding protein did not bind to the transfer membrane (data not shown). Coomassie blue staining of SDS-PAGE gels following electrophoretic transfer showed no presence of the untagged, expressed protein left in the gel following electroelution. Consequently, mouse anti-NBL4.1.4 antibodies were unable to detect the natural protein(s) within newborn larval extracts by Western blot. The inability of the glutamic acid-rich proteins to adequately bind nitrocellulose or charged nylon membranes is likely due to their extraordinary composition. This phenomenon has been observed elsewhere with glutamic acid-rich proteins (Triglia et al., 1988).

Immunohistochemical staining of fixed newborn larvae showed antibody binding to the parasite stichosome (Fig. 5). No interaction was observed with parasite cuticle, muscle or other organs. Binding within the stichosome was generally localised to the periphery of the developing stichocyte cells (Fig. 5).

4. Discussion

This is the first report on the identification of genes encoding host recognised antigens associated with the newborn larval stage of *T. spiralis*. Using serum from pigs experimentally infected with *T. spiralis* to screen a newborn larval expression library, an immunodominant sequence, designated NBL1500, was identified that encodes a protein specifically transcribed in the newborn larval stage and highly enriched in glutamic acid (39%) and serine (16%) residues where the largest number of acidic amino acids (53%) resides nearest the C-terminus. PCR amplification of cDNA with primers that span the entire open reading frame of NBL1500 resulted in the cloning of at least one other gene, designated NBL1700, that contains a 219 bp insert/repeat relative to NBL1500, and the identification of additional isoforms that remain uncharacterised. Conformational analysis using Coils 2.1 (Lupas et al., 1991) suggests that the predicted NBL1500/1700 proteins assume an α-helical, coiled-coil conformation within the glutamic acid-rich region i.e. from amino acid 100 through the C-terminus, though the primary structures do not strictly conform to the amphipathic character of alternating hydrophobic and hydrophilic residues proposed by Lupas et al. (1991). Within this region, three separate domains/coils have been calculated that centre around amino acids 191, 299 and 521 where the values approach 1.0 and substantially exceed the threshold probability of 0.5 units. The small, hydrophobic region at the N-terminus is consistent with this functioning either as an inside–out transmembrane helix or as a secretory signal with a putative cleavage site at amino acid 28. The binding of anti-glutamic acid-rich antibodies to the periphery of the developing stichocyte cells is consistent
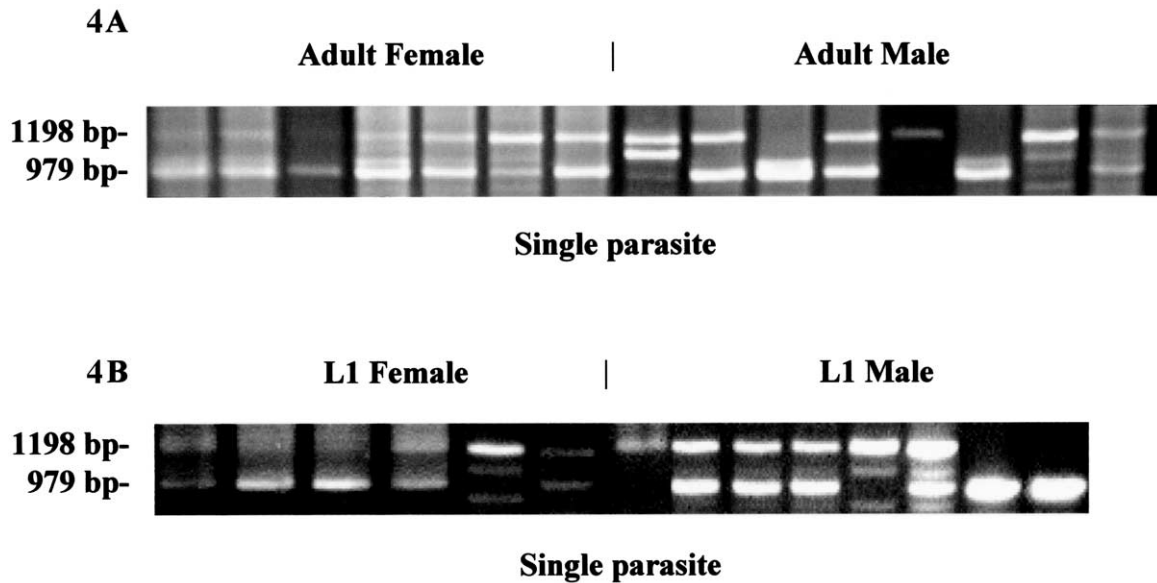
Fig. 4. Enzymatic amplification of genomic DNA from individual female and male adult worms (A) and muscle larvae (B) using primers #599 and #662. PCR primers span the 219 bp deletion in NBL1500 relative to NBL1700 and produce 1,198 and 979 bp truncated fragments characteristic of the NBL1700 and NBL1500 sequences, respectively. Non-uniform banding patterns were generated throughout the samples tested with the appearance of unique sized newborn larval fragments.

with these proteins being membrane associated; however, apart from the hydrophobic character of the N-terminus, the sequences from newborn larvae lack features of integral membrane proteins in that no additional hydrophobic character is present to provide a membrane spanning domain. Furthermore, there is no signal sequence at the C-terminus for membrane attachment through a glycosyl-phosphatidy-linositol anchor (Berger et al., 1988); consequently, the hydrophobic N-terminus may be more consistent with a secretory function. Stichocyte cells within the muscle larvae are well known for their secretory characteristics.

When DNAs from individual worms were PCR amplified, no fewer than two additional fragments appeared that migrated between the NBL1500 and NBL1700 sequences. No substantial difference was observed in the banding patterns between female muscle larvae and adult parasites, which potentially harbor embryonated eggs or unhatched, fully developed newborn larvae. These unique banding patterns are consistent with a multiple gene family; however, multiple alleles where sequence variation was observed among the isoforms cannot be ruled out. In addition, these data are among the first demonstrating genetic variation at the genome level between individual worms within a single *Trichinella* population, where the highest level of variability was observed among the males. It should be noted that the *T. spiralis* isolate used in these experiments is a well-characterised laboratory strain maintained in mice and pigs in excess of 25 years, indicating that the variability in this character has been retained with time.

In most animal–parasitic nematodes, gametes originate from a single cell (telegonic gonads); however, nematodes of the genus *Trichinella* have hologonic gonads in both the male and female where the germinal zone for gametes extends the entire length of the organ. It has been proposed (Murrell et al., 2000) that dispersion of the germinal zone may impart requisite genetic variation within a population that normally presents little possibility for natural variation by inbreeding between populations. Within *Plasmodium reichenowi*, Theisen et al. (2001) found that an excess of replacement polymorphisms within the glutamic acid-rich region of the PrGLURP protein among field isolates was consistent with positive diversifying selection which in turn would help the parasite avoid the host immune defenses. Though avoidance of host immunity cannot be ruled out, it is unlikely the sole impetus for the variation observed in the NBL1500/1700 sequences between or within individual worms given that: (1) transcription of the genes is limited to the NBL whose life is usually measured in hours prior to encysting within the muscle cell; (2) the window for peak, non-synchronous production of newborn larvae within the host lasts less than 48 h, and; (3) sufficient sequence identity exists within the glutamic acid-rich regions of the newborn larval isoforms that host immune responses are likely overlapping. The α-helical, coiled-coil domains predicted for the glutamic acid-rich region of NBL1500/1700 have been associated most often with functional proteins that bind other molecules such as myosins, intermediate filaments, transcriptional activators, and α- and β-tubulins (Lupas et al., 1991). Thus, the predicted secondary structure of the NBL1500/1700 would ascribe a regulatory role to the proteins thereby linking variability to functionality. A similar set of at least two secreted glutamic acid-rich isoforms, each transcribed by unique genes, was identified in rat where
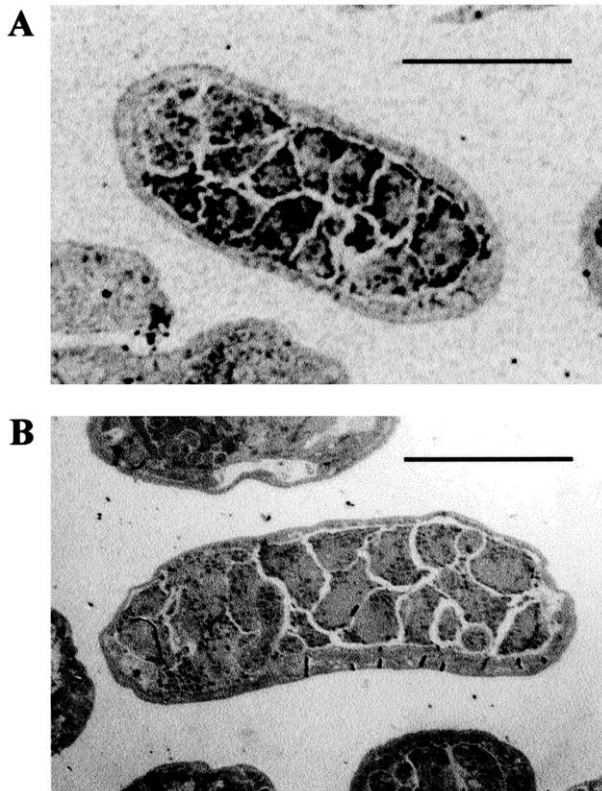
Fig. 5. Localisation of *Trichinella spiralis* newborn larval 1500/1700 proteins performed on tangential cross-sections through the parasite midbody using mouse anti-NBL4.1.4 serum (750×). (A) NBL1500/1700 and associated isoforms were localised to the stichosome. No binding was observed to cuticle, muscle, or other structures. (B) Control sections using normal mouse serum as 1° antibody. Bar = 37 μm.

differential expression occurred among the isoforms in response to β-adrenergic agonists (Cooper et al., 1991). Evidence from immunologically identified, overlapping sequences among smaller, incomplete clones of the NBL1500/1700 family (data not shown), support the conclusion that the glutamic acid-rich region contributes substantially to the immunogenicity of the NBL1500/1700 proteins. Antigenic, glutamic acid-rich proteins are not unique to parasites, though the newborn larval proteins described here are among the first reported in nematodes. Partially characterised, highly immunogenic, glutamic acid-rich proteins have been identified in several species of *Plasmodium* (Triglia et al., 1988; Giraldo et al., 1998) and *Trypanosoma* (Beecroft et al., 1993; Bayne et al., 1993). Within *Plasmodium*, some of the glutamic acid-rich proteins have been postulated to divert the host immune system; however, Giraldo et al. (1998), Foley et al. (1991) and Lustigman et al. (1990) have all shown that three independently identified and distinct glutamic acid-rich proteins from *Plasmodium falciparum* or *Plasmodium chabaudi* interact directly with the cytoplasmic face of the erythrocyte membrane during the infection process, and all are predicted to assume an α-helical conformation. Giraldo

et al. (1998) further suggested that the α-helical structure, glutamic-acid content, and internal repeat arrangement, similar to the larger repeats seen within *T. spiralis*, are attributes found within cytoskeletal proteins that if secreted from *Plasmodium*, may affect erythrocyte membrane function and structure. In support of this, others have putatively linked the *P. falciparum* glutamic acid-rich proteins designated Ag332 , Pf 11.1, Pf 155/RESA and D260 either to cytoadherence or to host cell invasion (Udomsangpetch et al., 1989; Mattei and Scherf, 1992; Wahlin et al., 1992). Each are abundant in glutamic acid, dipeptide repeats comparable with that found within the *T. spiralis* NBL1500/1700 sequences. The proteins D260 and Ag332 are particularly interesting in that, like NBL1500/1700, the molecular weight among the isoforms show substantial variability between isolates of *P. falciparum*. It is difficult to draw comparisons between protozoan and helminth parasites in general; however, nematodes of the genus *Trichinella* are uncharacteristically intracellular, have a predilection for striated muscle cells and therefore must have mechanisms for cell surface recognition, binding and penetration. In addition, newborn larvae of the genus *Trichinella* have the ability to alter the topography of the muscle cell prior to invasion (Gabryel et al., 1995; Wranicz et al., 1998). The NBL1500/1700 sequences presented here were expressed quite specifically in newborn larvae, which is the stage responsible for invasion. Work is underway to determine whether the newborn larval proteins have a propensity for interacting with the surface of muscle cells. Immunohistochemistry results are consistent with these proteins having a predilection for membrane surfaces within the stichosome; however, it has yet to be determined if these proteins are secreted from the parasite.

Given that *T. spiralis* is perceived as the most cosmopolitan of the *Trichinella* species and the one with the broadest documented host range, one can postulate that the ability of *T. spiralis* to non-discriminately penetrate and reprogram striated muscle cells may be facilitated by a family of regulatory proteins that can account for variation in muscle cell physiology among its innumerable potential hosts. Of particular interest is the large number of serine residues (16%) present within NBL1500/1700 sequences and the potential for phosphorylation at these sites. Phosphorylation is known to cause alterations in protein–protein interactions and in particular, proteins which can assume a coiled-coil form. The classic example for such changes was observed in skeletal muscle in the interaction between actin and phosphorylated light chain myosin, where phosphorylation-induced conformation changes in the protein from a folded form to a more linear, fibrous molecule (Craig et al., 1983). Thus the high number of non-systematically arrayed serine residues present within the ASEEX and ASQEE pentapeptides, in conjunction with the multiple isoforms of the NBL1500/1700 molecules could provide the individual parasite as well as the entire population, a mechanism by which numerous configurations of the protein could be

generated. Future work will focus on the role, if any, that these glutamic acid-rich proteins play in the infection process.

NBL1500 and NBL1700 sequences have been assigned GenBank accession numbers AY092819 and AY092820, respectively.

Acknowledgements

References

Bayne, R.A., Kilbride, E.A., Lainson, F.A., Tetley, L., Barry, J.D., 1993. A major surface antigen of procyclic stage *Trypanosoma congolense*. Mol. Biochem. Parasit. 61, 295–310.

Beecroft, R.P., Roditi, I., Pearson, T.W., 1993. Identification and characterization of an acidic major surface glycoprotein from procyclic stage *Trypanosoma congolense*. Mol. Biochem. Parasit. 61, 285–94.

Berger, J., Howard, A.D., Brink, L., Gerber, L., Hauber, J., Cullen, B.R., Udenfriend, S., 1988. COOH-terminal requirements for the correct processing of a phosphatidylinositol-glycan anchored membrane protein. J. Biol. Chem. 263, 10016–21.

Cooper, L.F., Elia, D.M., Tabak, L.A., 1991. Secretagogue-coupled changes in the expression of glutamine/glutamic acid-rich proteins (GRPs). J. Biol. Chem. 266, 3532–9.

Craig, R., Smith, R., Kendrick-Jones, J., 1983. Light-chain phosphorylation controls the conformation of vertebrate non-muscle and smooth muscle myosin molecules. Nature 302, 436–9.

Dame, J.B., Murrell, K.D., Worley, D.E., Schad, G.A., 1987. *Trichinella spiralis*: genetic evidence for synanthropic subspecies in sylvatic hosts. Exp. Parasit. 64, 195–203.

Despommier, D.D., 1971. Immunogenicity of the newborn larva of *Trichinella spiralis*. J. Parasit. 57, 531–5.

Despommier, D.D., Symmans, W.F., Dell, R., 1991. Changes in nurse cell nuclei during synchronous infection with *Trichinella spiralis*. J. Parasit. 77, 290–5.

Foley, M., Tilley, L., Sawyer, W.H., Anders, R.F., 1991. The ring-infected erythrocyte surface antigen of *Plasmodium falciparum* associates with spectrin in the erythrocyte membrane. Mol. Biochem. Parasit. 46, 137–48.

Gabryel, P., Leokadia, G., Blotna-Filipiak, M., 1995. The unique and specific transformation of muscle cell infected with *Trichinella spiralis*. Basic Appl. Myol. 5, 231–8.

Gabryel, P., Gustowska, L., 1970. Trichinellous basophilia of the muscle fibre. Acta Parasitol. Pol. 18, 1–6.

Giraldo, L.E., Grab, D.J., Wiser, M.F., 1998. Molecular characterization of a *Plasmodium chabaudi* erythrocyte membrane-associated protein with glutamate-rich tandem repeats. Euk. Microbiol. 45, 528–34.

Hill, D.E., Fetterer, R.H., Urban Jr, J.F., 1990. Biotin as a probe of the surface of *Ascaris suum* developmental stages. Mol. Biochem. Parasitol. 41, 45–52.

Jasmer, D.P., 1993. *Trichinella spiralis* infected skeletal muscle cells: arrest in G2/M is associated with the loss of muscle gene expression. J. Cell Biol. 121, 785–93.

Jasmer, D.P., Bohnet, S., Prieur, D.J., 1990. *Trichinella spp.*: differential expression of acid phosphatase and myofibrillar proteins in infected muscle cells. Exp. Parasit. 72, 321–31.

Jasmer, D.P., Neary, M.N., 1994. *Trichinella spiralis*: inhibition of muscle larva growth and development is associated with delay in expression of infected skeletal muscle characteristics. Exp. Parasit. 78, 317–325.

Kozek, W.J., 1975. *Trichinella spiralis*: morphological characteristics of male and female intestine-infecting larvae. Exp. Parasit. 37, 380–7.

La Rosa, G., Marucci, G., Zarlenga, D.S., Pozio, E., 2001. *Trichinella pseudospiralis* populations of the Palearctic region and their relationship with populations of the Nearctic and Australian regions. Int. J. Parasitol. 31, 297–305.

Lichtenfels, J.R., Murrell, K.D., Pilitt, P.A., 1983. Comparison of three subspecies of *Trichinella spiralis* by scanning electron microscopy. J. Parasitol. 69, 1131–40.

Lupas, A., Van Dyke, M., Stock, J., 1991. Predicting coiled-coils from protein sequences. Science 252, 1162–4.

Lustigman, S., Anders, R.F., Brown, G.V., Coppel, R.I., 1990. The mature parasite-infected erythrocyte surface antigen (MESA) associates with the erythrocyte membrane skeletal protein band 4.1. Mol. Biochem. Parasit. 38, 261–70.

Marti, H.P., Murrell, K.D., Gamble, H.R., 1987. *Trichinella spiralis*: immunization of pigs with newborn larval antigens. Exp. Parasit. 63, 68–73.

Mattei, D., Scherf, A., 1992. The Pf332 gene of *Plasmodium falciparum* codes for a giant protein that is translocated from the parasite to the membrane of infected erythrocytes. Gene 110, 71–79.

Moloney, A., Denham, D.A., 1979. Immunization of swine against *Trichinella spiralis*. Vet. Parasit. 15, 263–70.

Murrell, K.D., Lichtenfels, J.R., Zarlenga, D.S., Pozio, E., 2000. The systematics of the genus *Trichinella* with a key to species. Vet. Parasitol. 93, 293–307.

Ortega-Pierres, M.G., Mackenzie, C.D., Parkhouse, R.M.E., 1984. Protection against *Trichinella spiralis* induced by a monoclonal antibody that promotes killing of newborn larvae by granulocytes. Parasite Immunol. 6, 275–84.

Perrudet-Badoux, A., Binaghi, R.A., Boussac-Aron, Y., Ruitenberg, E.J., 1981. Antibody dependent mechanisms of immunity against migrating larvae of *Trichinella spiralis*. Vet. Parasit. 8, 89–94.

Pozio, E., Owen, I.L., La Rosa, G., Sacchi, L., Rossi, P., Corona, S., 1999. *Trichinella papuae* n.sp. (Nematoda), a new non-encapsulated species from domestic and sylvatic swine of Papua New Guinea. Int. J. Parasitol. 29, 1825–39.

Southern, E.M., 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. 98, 503–17.

Stoschek, C.M., 1990. Quantitation of proteins. Methods Enzymol. 182, 50–68.

Theisen, M., Thomas, A.W., Jepsen, S., 2001. Cloning, nucleotide sequencing and analysis of the gene encoding the glutamate-rich protein (GLURP) from *Plasmodium reichenowi*. Mol. Biochem. Parasit. 115, 269–73.

Triglia, T., Stahl, H-D., Crewther, P.E., Silva, A., Anders, R.F., Kemp, D.J., 1988. Structure of a *Plasmodium falciparum* gene that encodes a glutamic acid-rich protein (GARP). Mol. Biochem. Parasit. 31, 199–202.

Udomsangpetch, R., Carlsson, J., Wahlin, B., Holmquist, G., Ozaki, L.S., Scherf, A., Mattei, D., Mercereau-Puijalon, O., Uni, S., Aikawa, M., et al., 1989. Reactivity of the human monoclonal antibody 33G2 with repeated sequences of three distinct *Plasmodium falciparum* antigens. J. Immunol. 142, 3620–6.

Vassilatis, D.K., Polvere, R.I., Despommier, D.D., Gold, A.M., Van der Ploeg, L.H.T., 1996. Developmental expression of a 43-kDa secreted glycoprotein from *Trichinella spiralis*. Mol. Biochem. Parasit. 78, 13–23.

Wahlin, B., Sjolander, A., Ahlborg, N., Udomsangpetch, R., Scherf, A., Mattei, D., Berzins, K., Perlmann, P., 1992. Involvement of Pf155/RESA and cross-reactive antigens in *Plasmodium falciparum* merozoite invasion in vitro. Infect. Immun. 60, 443–9.

Wranicz, M.J., Gustowska, L., Gabryel, P., Kucharska, E., Cabaj, W., 1998. *Trichinella spiralis*: induction of the basophilic transformation

of muscle cells by synchronous newborn larvae. Parasitol. Res. 84, 403–7.

Zarlenga, D.S., Chute, M.B., Martin, A., Kapel, C.M.O., 1999. A single multiplex PCR for unequivocal differentiation of six encapsulated and three non-encapsulated genotypes of *Trichinella*. Int. J. Parasitol. 29, 141–9.

Zarlenga, D.S., Rhoads, M.L., Al-Yaman, F.M., 1994. A *Taenia crassiceps* cDNA sequence encoding a putative immunodiagnostic antigen for bovine cysticercosis. Mol. Biochem. Parasitol. 67, 215–23.

Zarlenga, D.S., Murrell, K.D., 1989. Molecular cloning of *Trichinella spiralis* ribosomal genes: application as genetic markers for isolate classification. Tanner, C.E. (Ed.), Trichinellosis, vol. 7. CSIC Press, Madrid, pp. 35–40.