Published in final edited form as:

Dev Comp Immunol. 2008 ; 32(10): 1119–1130. doi:10.1016/j.dci.2008.03.001.

Fibrinogen-bearing protein genes in the snail *Biomphalaria glabrata*: characterization of two novel genes and expression studies during ontogenesis and trematode infection

Si-Ming Zhang *

Center for Evolutionary and Theoretical Immunology, Department of Biology, MSC03 2020, University of New Mexico, Albuquerque, New Mexico 87131, USA

Hong Nian,

Center for Evolutionary and Theoretical Immunology, Department of Biology, MSC03 2020, University of New Mexico, Albuquerque, New Mexico 87131, USA

Yong Zeng, and

Center for Evolutionary and Theoretical Immunology, Department of Biology, MSC03 2020, University of New Mexico, Albuquerque, New Mexico 87131, USA

Randall J. DeJong

Biology Department, Calvin College, Grand Rapids, Michigan 49546, USA

Abstract

All fibrinogen (FBG)-bearing proteins documented to date in the freshwater snail *Biomphalaria glabrata*, the intermediate host of the human blood fluke *Schistosoma mansoni*, possess the same molecular structure; one or two immunoglobulin superfamily (IgSF) domains at the N-terminus and a FBG domain at the C-terminus (named as FBG-related protein (FREP)). Here we report two novel genes that encode FBG-bearing proteins from *B. glabrata*. Different from all known FREPs, the first gene encodes a protein (657 amino acids (aa)) composed of a long N-terminal region with no sequence homology to any known protein, a middle epidermal growth factor (EGF) repeat region and a C-terminal FBG domain, designated FBG-related molecule (FREm). A differential expression at 2 days post exposure (dpe) to the trematode *S. mansoni* or *Echinostoma paraensei* has been found in the *S. mansoni* susceptible M line and resistant BS-90 snail strains. The second gene is a new member of the FREP family, designated FREP14, which encodes a 399 aa putative secreted protein. FREP14 is different from known FREPs in that it is encoded by a single locus and is not upregulated in early or late stage *S. mansoni* exposure, but is upregulated in late stage *E. paraensei* infection. Furthermore, gene expression during the snail's ontogenesis and at a late stage of trematode-infection (52 dpe) has been investigated in the two newly-identified genes (FREm and FREP14) described in this paper and five representative members of known FREPs (FREPs 2, 3, 4, 12, and 13). A variety of expression patterns were observed, suggestive of functional diversity among the members of FBG-bearing proteins. Our findings further broaden our understanding of the diversity and function of the FBG-bearing protein encoded genes in *B. glabrata*.

*Address for Correspondence, Center for Evolutionary and Theoretical Immunology, Department of Biology, MSC03 2020, University of New Mexico, Albuquerque, New Mexico 97131, USA, Phone: 505 277 4589; Fax: 505 277 0304, Email: zhangsm@unm.edu.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Keywords

epidermal growth factor (EGF); fibrinogen-related protein (FREP); fibrinogen-related molecule (FREm); mollusc; innate immunity; *Biomphalaria glabrata*; *Schistosoma mansoni*; *Echinostoma paraense*

1. Introduction

Fibrinogen (FBG) proteins have well-characterized functions in blood clotting and in innate responses in vertebrate animals [1,2]. However, FBG-encoding sequences or molecules have been recently discovered in a variety of invertebrates [3,4]. An accumulating literature increasingly suggests that FBG-bearing proteins in a broad phylogenetic spectrum of animals play a role in defense against a range of pathogens, although specific details for how this occurs is often lacking. In the sponge *Suberites domuncula*, a representative of the most basal metazoan phylum, a protein containing a FBG-like domain binds carbohydrates and is upregulated by curdlan [5]. It has been shown that ascidian ficolins, the FBG-bearing proteins, bind to *N*-acetyl-D-glucosamine (GlcNAc) and *N*-acetyl-D-galactosamine (GalNAc) and activate immune responses [6]. In horseshoe crab, tachylections have antibacterial activities and also function in non-self recognition [7]. In mosquitoes of the genera *Anopheles* and *Armigeres*, FBG-like proteins have been implicated in immune responses to microbes or malaria parasites because their transcript levels are upregulated after infection [8–10]. A further study using microarrays and RNA interference (RNAi) revealed that some members of the *Anopheles* FBG-bearing proteins called FBG immunolectins (FBN) such as FBNs 8, 9 and 39 strongly influence *Plasmodium* development *in vivo*, likely playing a role in initial recognition of the pathogens [11]. With respect to human ficolins which are the best-studied FBG-bearing proteins in terms of structure and function, three types of ficolins (L-, H-, and M-ficolin) have been found to have lectin activity and can activate the lectin complement pathway [4].

It has been shown that the freshwater snail *Biomphalaria glabrata* (Gastropoda, Planorbidae) produce a number of FBG-bearing proteins, which were designated FBG-related proteins, termed FREPs [12]. *Biomphalaria glabrata* is an intermediate host for the parasite *Schistosoma mansoni*, one of the causative agents of schistosomiasis, affecting over 200 million people world wide [13]. FREPs have a unique molecular structure among all known proteins in combining one or two N-terminal immunoglobulin superfamily (IgSF) domains and a C-terminal FBG β/γ domain [14–16]. Earlier studies revealed that *FREPs* display considerable diversity at the messenger RNA (mRNA) and DNA levels, and 13 subfamilies have been documented [16,17]. Moreover, it has been found that diversification of *FREPs* is generated through gene conversion and point mutation [18]. Diversification of innate immune effectors has also been found in other invertebrates, leading to general reconsideration of invertebrate immune capabilities [19–23].

While a number of *FREP* genes have been identified from *B. glabrata*, many questions have been raised about the biological significance of their diversity, structure and function, which remain to be further elucidated. As described above, at present, we know that the *B. glabrata* genome contains multiple *FREP* subfamilies [16]. All FREPs thus far observed are associated with a single FBG-like domain and one or two IgSF domains [14–16], but whether some of the many FBG-like sequences in *B. glabrata* might be part of other domain combinations has not been determined. In a previous study [17], we found four partial FBG-like sequences (in two groups) that did not cluster with FBG-like sequences of known *FREPs* in a phylogenetic analysis. We therefore characterized the full-length complementary DNAs (cDNA) of two of these unusual partial FBG-like sequences (Bg80t and Bg106m; see Fig. 6 of [17]), each representing one group. Here we present two novel cDNAs, one encoding a protein with a

novel type of domain combination (epidermal growth factor (EGF) + FBG), and the other revealed to be a new member of the FREP family.

FREPs, present in the snail hemolymph, function as calcium-dependent lectins, and bind to the trematode *Echinostoma paraense* [12,24]. The functions of FREPs in the snail are not yet fully understood. This may be due to the complex nature of FREPs. For example, it has been shown that the FREP expression in response to different pathogens varies [24,25]. Like most earlier work performed *B. glabrata* internal defense, studies on the response of FREPs to the trematode parasite *S. mansoni* or *E. paraense* have focused on the juvenile snails or the juvenile snails at early stage of infection, normally at less than 10 days post-exposure [12,25]. It is well-known that trematode infection within a snail host is a long term process [26], but little is known about the expression of the individual *FREP* members in response to the late infection stages, when the parasite has become quite massive, and a large number of cercaria are being produced in a snail host. Likewise, the ontogenic expression of *FREPs* has not been studied. Studies from other animals suggested that some immune-related proteins also play a role in development, for example, Toll and Dscam in *Drosophila* [20,27]. Moreover, it has been noted that snail age affects the compatibility between snail and parasite [28] and that *FREPs* may have multiple functions [15]. Therefore, the second goal of this study was to investigate expression of five representative *FREPs* and the two newly-identified FBG-bearing genes during snail ontogenesis and at late stage of *S. mansoni* or *E. paraense* infection. The comparative expression studies enable us to gain insight into the possible function of the FBG-bearing genes in *B. glabrata*.

2. Materials and methods

2.1. Snails and parasites

Two strains (M line and BS-90) of the snail *B. glabrata* were used for cloning novel FBG-like genes and for performing Southern and Northern analyses. M line and BS-90 snails are susceptible and resistant to *S. mansoni*, respectively. The digenetic trematodes *S. mansoni* and *E. paraense* were maintained in the laboratory as described by [29] and [30], respectively. Both trematodes use the snail *B. glabrata* as an intermediate host for their life cycles.

2.2. Infection of snails with parasites

In the gene expression study, snails at two stages of trematode-infection, early stage (2 days post-exposure, dpe) and late stage (52 dpe) of infection, were investigated. Since the response of *FREP* genes to the early stage of trematode-infection including 2 dpe was reported [25], in this study gene expression at the early stage of infection was only measured in the two newly-identified genes described in this paper. At late stage infection and during snail ontogenesis, gene expression of the two new genes and five known *FREPs* was measured.

The number of miracidia for *E. paraense* and *S. mansoni* used to infect a snail (0.6–0.9 cm in shell diameter) was 20–30 and 10–20, respectively. In the early stage infection study, total RNA was collected from pooled snails (n=12) at 2 dpe to either *E. paraense* or *S. mansoni*, as well as non-exposed snails. In the case of *E. paraense* infection, only those snails that had been confirmed to be infected by microscopy were used for RNA extraction. In the case of *S. mansoni* infection, all snails that were exposed to the *S. mansoni* miracidia were used because it is difficult to identify the infection status at 2 dpe.

Late stage infection (52 dpe) M line snails exposed to *S. mansoni* (two groups of snails with two different sizes: 1.3–1.5 cm (n=4) and 1.6–1.7 cm (n=4)) were confirmed to be infected by monitoring for cercaria shedding. Infected and non-exposure control snails of similar sizes were conditioned overnight in separate tanks. For the M line snails exposed to *E. paraense*

(n=4; 1.3–1.5 cm), infection was verified by observing presence of sporocysts at 4 dpe. Those snails confirmed to be infected were maintained separately until they were used for RNA extraction (i. e., 52 dpe).

For ontogenic expression studies, M line snails at four different sizes (from smallest to largest, size classes 1–4: 0.2–0.4, 0.7–0.8, 1.4–1.6, and 2.0–2.3 cm, respectively) were used. Size classes 1 and 2 are immature with size 1 very recently hatched. Sizes 3 and 4 are mature, with size 4 the maximum size reached in our cultures. The number of the snails used from the size 1 to 4 is 40, 10, 5 and 5, respectively.

2.3. Extraction of total RNA and genomic DNA

Whole snails were used for RNA extraction throughout this study. Total RNA was extracted initially by TRIzol reagent (Invitrogen), and the resulting RNA was further purified using the NucleoSpin RNA II kit (Clontech). Once the RNA was obtained, cDNA synthesis was performed immediately using SuperScript III First-Strand Synthesis System for RT-PCR (Invitrogen). The cDNA and remained RNA were stored in –70 °C for quantitative polymerase chain reaction (qPCR) and Northern blot analyses, respectively.

For Southern blotting, genomic DNA was extracted from an M line snail, which was inbred for 20-generations [31], using CTAB method [32]. The concentration and purity of total RNA and genomic DNA were determined spectrophotometrically.

2.4. Generation of full-length cDNA

To generate the full-length cDNA of interest, the forward primer Bg80/13F1 (5'-gacgagctgagagctaccacaacg-3') and reverse primer Bg80/13R1 (5'-actctcagctcgtctgcaatagaatcg-3') were designed based on Bg80t (GenBank accession number, AY559491); the primers Bg106F1 (5'-tatcaaaaaacatgctccttagca-3') and Bg106R1 (5'-tagcaatagaaagcttgagtattgggc-3') were designed based on Bg106m (AY559485). GeneRacer (Invitrogen) was used for 5'- and 3'-RACE (rapid amplification of cDNA end). Additionally, some sequences were obtained and/or confirmed from a Bluescript cDNA library which had been generated from M line snails at 4 dpe to *E. paraensei*. Those sequences were generated using a gene-specific primer and the vector primer M13 reverse or T7 [14]. After a contiguous sequence was generated, the primers at 5'- and 3'- UTR (untranslated region) were designed and used for amplifying the cDNA to confirm the full-length cDNA sequence obtained.

2.5. Cloning, sequencing and sequence analyses

Amplicons were cloned into pCR2.1 TOPO vector (Invitrogen). The extension products resulting from cycle sequencing were analyzed on an ABI 377 sequencer. Nucleotide (nt) sequences were checked for accuracy using Chromas v 1.42 (www.technelysium.com.au/chromas) and then blasted using NCBI blast search program (www.ncbi.nlm.nih.gov/blast). Contiguous sequences were generated using Sequencher (version 4.2; Gene Codes Corporation). Signal peptide (SP) and conserved domains were predicted by SMART program (<http://smart.embl-heidelberg.de>) [33].

2.6. Southern and Northern blot analyses

Oligonucleotide probes were amplified by PCR from corresponding cDNA templates. The probe for FBG-related molecule gene (*FReM*) (1399 bp, at the N-terminal non-FBG region) was generated using the forward primer Bg80F7 (5'-cacgtagcagtggaaacttaac-3') and the reverse primer Bg80R6 (5'-caatgggagtcacttgacgtggtttcc-3'). The probe for *FREPI4* (449 bp, at the N-terminal IgSF region) was produced by the forward primer Bg106F4c (5'-gaagacttattgaagtttgagacactta-3') and the reverse primer Bg106R3 (5'-

tgcaacagcctgaccttagagagcgt-3'). Primers used to generate probes for *FREPs* 2, 3, 4, 12, and 13 were published [17].

All probes were radiolabelled with [α - 32 P]dATP using random priming (Ambion). For a given gene, the same probe was used for both Southern and Northern blotting. Southern and Northern analyses were conducted using a modified version of the SouthernMax and NorthernMax method (Ambion), and the details were provided [17]. For Northern blotting, an equal amount of RNA (20 or 25 μ g) from control snails and snails under different treatments described was loaded into each lane. Hybridization was conducted and re-probing was performed using the Strip-EZ method (Ambion).

2.7. qPCR

To validate Northern blot analyses, and to quantitate gene expression, qPCR was performed in the seven genes (*FReM*, *FREPs* 2, 3, 4, 12, 13, and 14). The qPCR primers for *FReM* are qBg80F1 (5'-cctgtaagttagatggaacatg-3') and qBg80R1 (5'-cacaagttttactacaggcgtgac-3'). The primers for *FREPI4* are qBg106F1 (5'-agcttcagagcaaaaggatcatg-3') and qBg106R1 (5'-tcccataattcttcttgatgataa-3'). The primers for *FREPI2* are as follow; the forward (5'-tggtcccttcagaatctggataataatt-3') and reverse one (5'-gccgtgttaaagtactctctgtttg-3'). The qPCR primers for *FREPs* 2, 3, 4, and 13 were published [34]. All primers are located at the highly specific region (non-FBG region) of each gene. qPCR was performed on a Sequence Detection System 7000 (Applied Biosystems) using a SYBR Green PCR Core Reagents kit (Applied Biosystems), as used previously to assess differential expression of *FREPs* [34].

Relative gene expression data were obtained using the $\Delta\Delta C_t$ method [35], with *18S* rDNA as the internal standard. For each gene tested, qPCR was performed in triplicate from a cDNA transcribed from the same RNA sample used for the Northern analysis. So, statistical analysis has not been performed because data were collected from the technical replicates, rather than different biological samples. The data shown in the corresponding figure are the average value related to those of the control snails and the standard error bars in each experiment.

3. Results

3.1. Characterization of a cDNA that encodes a FBG-bearing protein with a novel domain combination, designated FBG-related molecule (FReM)

Previous phylogenetic analysis showed that Bg80t, a partial FBG-like sequence (Genbank accession number: AY559491) uncovered from the *B. glabrata* genome, did not cluster to the FBG sequences of known *FREPs*, which was further confirmed by canonical residues analysis [17]. Determination of the entire cDNA sequence revealed that the deduced 657 amino acid (aa) protein encoded by the gene characterized from the Bg80t sequence has a novel domain structure (epidermal growth factors (EGFs) and FBG), which is quite different from the known *FREPs* (Fig. 1A and B; detailed description below). In order to distinguish this novel type of FBG-bearing protein gene from the family of proteins known as *FREP* genes, which bear both fibrinogen domains and IgSF domains, the gene product was designated FReM (FBG-related molecule) (EF467291).

NCBI blast search revealed that FReM contains an N-terminal region (304 aa; from 21 to 323 aa) that did not match to any known protein sequences, followed by two EGF-like repeats of 42 and 43 aa. EGF is a common module found in various organisms, and possesses 6 or 8 most conserved cysteines. In some cases, glycine is also conserved in some positions across EGF repeats [36–39]. In the FReM, in addition to 8 conserved cysteines found in the two repeats, the first and second EGF repeats contain 4 and 3 conserved glycines, respectively. The aa identity between the two EGFs is 39% (Fig. 2). Computational analysis using Blast search

indicated that the two EGF repeats are similar to the EGF repeats of several proteins, and most closely related to the EGF repeats located within the tyrosine-protein kinase receptor Tie-2 in vertebrates. The aa identity between the EGFs of the FReM and EGFs of tyrosine-protein kinase receptor Tie-2 from mouse, human and zebrafish is 34%, 33%, and 32%, respectively.

A well-conserved 211 aa FBG-like domain (from 447 to 657 aa) is located at the FReM C-terminus, which maintains 20 canonical residues [40] (Fig. 1A). When FBG domain of FReM was compared to FBG sequences of the known FREPs, it was revealed that the FBG-like domain of FReM is more similar to the FBG-like domain of FREPs 3 and 12 than to the FBG-like domain of other FREPs (Table 1).

To determine the approximate number of loci present in the snail genome, a gene-specific probe (1399 bp) containing the N-terminal region of FReM was applied to hybridize to the snail genomic DNA after digestion by three restriction enzymes. Two bands were revealed by Southern blotting (Fig. 3A). Since the source of DNA used for the Southern blotting was from a M line snail produced by 20 generations of self-fertilization, it is likely that most genes are homozygous. With that assumption, it is estimated that 2 loci are present in the snail genome, consistent with the data obtained using the probe derived from the FBG-like region (Bg80t probe, see [17]).

3.2. Expression of *FReM* after exposure to the trematode *E. paraensei* or *S. mansoni*

In order to get insight into the defense function of *FReM* in snails, Northern and qPCR analyses were used to determine if *FReM* is induced during early stage infection (2 dpe) with *E. paraensei* or *S. mansoni*. In the *S. mansoni* susceptible M line snails, the expression level was heightened after exposure to *S. mansoni*. However, downregulation was observed in the *E. paraensei*-infected snails. In the case of the *S. mansoni* resistant BS-90 snails, *E. paraensei* or *S. mansoni* was shown to downregulate *FReM* expression. Northern blot and qPCR data gave highly similar results (Fig. 4A and B).

3.3. *FREP14*, a new member of *FREPs* is encoded by a single locus, but not inducible by trematode-infection

Sequencing of the full-length cDNA associated with the unique FBG-like sequence Bg106m [17] revealed that the corresponding gene encodes a 399-aa protein. This protein is composed of a 19 aa SP, a 133 aa N-terminal IgSF domain (from 20 to 153 aa), a 43 aa interceding region (ICR) 154–197 aa), and a 201 aa C-terminal FBG domain (22 canonical residues; 198–399 aa) (Figs. 5A, B). Based on the domain structure, it was classified as a one-IgSF FREP. Since 13 *FREP* subfamilies were documented previously [16], the gene product is designated *FREP14* (EF467292).

Sequence analysis revealed that the IgSF domain of *FREP14* is more similar to IgSF of *FREP2* (49% aa identity) than to the IgSF2 of two-IgSF FREPs such as FREPs 3 and 7 (43–38%), but less similar to IgSF of *FREP4* (35%) (Table 1). The aa identity of the FBG domain of *FREP14* to the FBG of FREPs 2, 3, 4, 7, 12, and 13 is very comparable (61–67%). However, the identity between the FBG of *FREP14* and *FReM* is lower (54%) (Table 1). Southern blotting using a *FREP14* gene-specific probe (449 bp) derived from the N-terminal IgSF region clearly showed one band in three genomic DNA samples digested by one of three restriction enzymes (*EcoR* I, *Hind* III and *Pst* I), suggestive of one locus present in the genome (Fig. 3B).

To test whether expression of *FREP14* in the two snail strains is inducible after exposure to the trematode *S. mansoni* or *E. paraensei*, the same RNA samples that were used to test the expression of *FReM* were used to test *FREP14* gene expression. In contrast to the *FReM*

expression described, we did not observe the significant change of *FREP14* expression in snails after trematode-exposure (data not shown).

3.4. Ontogenic expression of *FReM* and *FREPs* 2, 3, 4, 12, 13, and 14

The ontogenic expression in *FREP* genes has not been investigated previously. Five representative *FREPs* (*FREPs* 2, 3, 4, 12, and 13) and the two novel genes described here, *FReM* and *FREP14*, were selected for ontogenetic expression studies. The expression study was performed on the M line snails of four different sizes, as shown in Fig. 6A (lane 1 to 4). Northern analysis showed that expression signal of *FReM* was very weak in the two different sizes of snails (lanes 1 and 2, in Fig. 6A). The expression level was heightened in mature snails (lane 3 and 4, in Fig. 6A), which was generally consistent with the data obtained by qPCR (Fig. 6B). In contrast, each individual *FREP* was detected by its respective probe in immature and mature size classes (Fig. 6A), although the relative expression level did vary ontogenetically (Fig. 6B). Among the *FREPs* tested, *FREP2* was found to be highly expressed in the two juvenile size classes relative to the adult size classes. Expression of *FREP4* was also higher in immature than mature snails, with a pronounced peak in size class 2, just before maturation. *FREPs* 3 and 12 were consistently expressed throughout all the sizes of the snails. In the case of *FREP14*, the lowest expression level was noted in the oldest snails (see Figs. 6A and B). All Northern blots indicated single transcripts except for *FREP13*, for which the probe detected two transcripts at an approximate size of 2.4 and 3.4 kb (Fig. 6A).

3.5. Expression of *FReM* and *FREPs* 2, 3, 4, 12, 13, and 14 at late stage trematode-infection

We investigated expression profiles in M line snails at late stage infection (52 dpe) for the two newly-identified genes, *FReM* and *FREP14*, and the same representative *FREPs* as above (2, 3, 4, 12, and 13). A wide variety of expression patterns were present among these genes (Fig. 7A and B). Expression of *FReM* was found to be downregulated by during late stage trematode infection. *FReM* expression was lower in both size classes of *S. mansoni*-infected snails, and in *E. paraensei*-infected snails, relative to respective controls. In contrast, *FREP14* expression was not higher in *S. mansoni*-infected snails relative to controls, but was elevated in *E. paraensei*-infected controls. *FREPs* 2, 3, and 4 were strongly elevated in *S. mansoni*-infected M line snails, but only *FREP4* was also elevated in *E. paraensei*-infected snails.

4. Discussion

Rapidly growing evidence suggests that FBG-bearing proteins are a group of proteins mainly functioning in vertebrate and invertebrate innate immunity. It has been shown that all FBG-bearing proteins contain a common C-terminal FBG domain, but differ in their N-terminal regions [41,42]. For example, ficolins possess a short N-terminal collagen-like domain whereas snail *FREPs* have an N-terminal IgSF domain. The N-terminal regions of some FBG-bearing proteins have no homology to known proteins, for instance, horseshoe crab tachylectin [7], and most FBG-like sequences in the fruitfly *Drosophila melanogaster* and the mosquito *Anopheles gambiae* [42]. Alternatively, the molecules contain only FBG-like domain(s), as in some mosquito FBG-like sequences [42] and the slug FBG-like lectins [43]. In the present study, in addition to identifying a new member of *FREP* (*FREP14*), more importantly, we report a gene, *FReM*, from *B. glabrata* encoding a FBG-bearing protein with a novel domain combination: an N-terminal region with no homology to any known protein, a middle EGF repeat region and a C-terminal FBG domain.

The FBG domain of *FReM* contains 20 of the 24 proposed canonical residues [40], slightly less than FBG domains of known *FREPs*, which normally have 22 canonical residues [14–16]. Unlike the FBG probe of *FREPs* that can cross-react to all or multiple *FREPs*, the FBG probe which was confirmed in this study was not able to hybridize to *FREP* sequences [17].

Taking into account the several lines of evidence available (phylogenetic, Southern and canonical residue analyses), it is suggested that FBG domain of *FReM* might have diverged from a precursor of FBGs, before the FBGs associated with *FREPs* had duplicated.

EGF-containing proteins have been found to play a role in extracellular protein-protein interactions such as adhesion, coagulation, and receptor-ligand interactions [38]. In addition to well-known functions in developmental processes and homeostasis [44], EGF modules have been recently described to function in innate immunity. In the beetle *Holotrichia diomphalia*, a secreted protein containing 6 EGF repeats has been found to act as a pattern recognition protein for lipopolysaccharide and agglutinated bacteria [45]. Eater and Nimrod are proteins composed of 32 and 10 EGF repeats respectively and both mediate phagocytosis of bacterial pathogens in *Drosophila* [38,39]. Similarly, a 120-kDa hemocyte-specific membrane protein containing 18 EGF repeats has been identified in the flesh fly *Sacrophaga* and was believed to be a scavenger receptor [46]. Moreover, both EGF- and FBG-like genes have been found in the sponge *Suberites domuncula* to respond to curdlan, suggesting EGF- and FBG-like domains may have been involved in the ancient immune response before the divergence of deuterostome and protostome [5]. The similar immune functions of both domains (EGF and FBG) may have maintained them in the two metazoan lineages.

The expression pattern of the *FReM* gene in response to trematodes is intriguing. A previous study showed that the two snail strains responded to *E. paraensei*, but only BS-90 snails had elevated *FREPs* 2 and 4 expression to *S. mansoni* [25]. Those patterns are in contrast to what we observed in *FReM* in the snails at 2 dpe to *E. paraensei* or *S. mansoni* (see Fig. 4). Moreover, ontogenic expression in the juveniles was found to be different between *FReM* and all *FREPs* tested (see Fig. 6). Together, difference in gene expression under different challenges and developmental stages detected in our study may reflect a functional difference between *FReM* and *FREPs*.

FREP14, a new *FREP* member identified in this study, is different from previously known *FREPs* in several aspects. It is encoded by a single locus, which is different from most *FREP* subfamilies encoded by multiple loci [17]. Expression studies revealed that it was not induced by the two trematode parasites, suggesting that it may not function in the context of parasite-induced immune responses. In mosquitoes, it has been shown that FBNs show varied responses to different challenges such as in the case of malaria parasite exposure or microbial infections [11]. Therefore, it is also possible that *FREP14* might be involved in snail's immune responses to non-trematode pathogens, such as microorganisms. *FREP14* might have diverged from other *FREPs* at a very early time, as suggested by previous phylogenetic analysis [17]. Recent gene duplication was likely not involved because there is a single locus present in the *B. glabrata* genome. Although gene conversion has been found to play a role in the *FREP* diversification [18], we do not know whether such mechanism was involved in the evolution of *FREP14*. Nevertheless, the finding of a new member of the *FREP* family provides further insight into diverse functions of *FREPs*.

In contrast to *FREPs* which were found to be expressed in all four stages of snails tested, our study shows that expression of *FReM* was very low in the nonmature snails (normally undetectable by Northern blotting). Moreover, within members of *FREPs* tested, the ontogenic expression pattern has also been found to vary among the *FREP* members. For example, it has been shown that expression of *FREPs* 2 and 4 was lower in mature snails than in immature ones, whereas expression of the other *FREPs* did not vary among different sizes. This implies that some FBG-bearing proteins may have functions in development. The different functional role of individual *FREP* members in the snail internal defense and the possible roles of *FREPs* in non-immune responses have been proposed previously [15,25].

Two *FREP13* transcripts were revealed by Northern blotting. The size of the smaller one (2.4 kb) is consistent with that of *FREP13* mRNA reported previously [16]. However, the larger transcript (3.4 kb) is larger than other *FREPs* and *FReM*. So, it seems unlikely that either of the transcripts resulted from the process of alternative splicing because the alternatively spliced transcripts should be smaller than the mRNA transcribed from the original gene. There are two types of *FREPs*, one-IgSF and two-IgSF *FREPs*, all containing a FBG-like domain. Our explanation is that the *FREP13*-like protein transcribed from the larger transcript may contain more than two IgSF domains and/or more than one FBG domain.

This study includes the first examination of *B. glabrata* FBG-bearing protein transcript expression during late stage trematode infection. Generally, snail's responses to the late stage of trematode-infection should be quite different from those at the early stage infection. After about one month exposure, the successful trematode *S. mansoni* or *E. paraensei* proliferates massively within snail host. As a consequence, a large space is occupied and nutrients are consumed by the parasites. Moreover, snails get wounded due to the shedding of large numbers of cercariae from them on a daily basis. Immunological and/or physiological responses by the snail may be necessary to cope with such stress. Among the genes tested, it seems that *FREPs* 2 and 4 are particularly interesting. *FREPs* 2 and 4 expression are elevated in late stage *S. mansoni*-infected snails, but *FREP4* is also elevated in late stage *E. paraensei*-infected snails, whereas *FREP2* is not. In addition, *FREP14* shows little response to late stage *S. mansoni*, but is elevated in late stage *E. paraensei*-infected snails. It is well-known that the snail immune system interacts with the two trematodes by different mechanisms, especially during early stages of infection [37]. Our findings substantiate this notion and suggest differences in interaction mechanisms may persist throughout the course of infection.

Snail genes which response to infection may be part of the defense response, or they be regulated by parasitic factors to promote parasite development and proliferation. Obviously, it would be interesting to know whether disruption of *FREPs* 2 and 4 at the late stage of infection will inhibit schistosomiasis transmission within the snail host, our previous attention was drawn to the snail responses at the early stage of infection, which is normally an immunological response. Inspired by the current findings, prevention of parasite development within snail host may be an alternative way to control schistosomiasis transmission. Conceivably, manipulation of molecules critical for parasite survive or development within snail host might be able to block transmission to human hosts even though snail infection was not prevented. This long-term goal requires considerable knowledge and technologies be developed for this snail-trematode model (for example, genetic knockout, transgenesis. etc).

In summary, data presented here provides valuable information regarding two newly-identified FBG-bearing protein encoded genes (*FReM* and *FREP14*) and comparative expression profiling of the two new genes and five *FREP* members during ontogenesis and a later stage of trematode-infection. This study further expands our understanding of diversity and function of FBG-like molecules in *B. glabrata*, the prominent intermediate snail host for the human parasite *S. mansoni*. Recent genome sequencing project from different animal species has revealed that FBG-bearing proteins encoded genes are abundant. For example, the ongoing genome project of the mollusc *Lottia gigantea* (Gastropoda, Lottiidae) shows that 70 FBG-bearing proteins were annotated based on the incomplete genome sequence, suggestive of at least 70 FBG-like genes present in this mollusc genome (http://supfam.cs.bris.ac.uk/SUPERFAMILY/cgi-bin/gen_list.cgi?genome=gy). A similar number of FBG sequences had been uncovered from mosquitoes [42]. It is expected that in the near future more FBG-bearing proteins or genes will be identified from various animal species, and that their functions will be elucidated. The increasing information about snail FBG-bearing

molecules makes them an ideal model for the study of invertebrate immunity and snail-trematode interactions.

Acknowledgments

This work was supported by National Institute of Health Grant RO1 AI067686 (to S-MZ) and NIH Grant Number RR-1P20RR18754 from the Institutional Development Award (IDeA) Program of the National Center for Research Resources. The authors would like to thank Dr. Sam Loker for his encouragement and generous support.

References

1. Doolittle RF. The evolution of vertebrate blood-coagulation – a case of Yin and Yang. *Thromb Haemost* 1993;70:24–28. [PubMed: 8236110]
2. Lu J, Teh C, Kishore U, Reid KBM. Collectins and ficolins: sugar pattern recognition molecules of the mammalian innate immune system. *Biochimica et Biophysica Acta-General Subjects* 2002;1572:387–400.
3. Holmskov U, Thiel S, Jensenius J. Collectins and ficolins: humoral lectins of the innate immunity defense. *Annu Rev Immunol* 2003;21:547–578. [PubMed: 12524383]
4. Fujita T, Matsushita M, Endo Y. The lectin-complement pathway-its role in innate immunity and evolution. *Immunological Rev* 2004;198:185–202.
5. Perović-Ottstadt S, Adell T, Proksch P, Wiens M, Korzhev M, Gamulin V, et al. A (1-3)- β -D-glucan recognition protein from the sponge *Suberites domuncula*: mediated activation of fibrinogen-like protein and epidermal growth factor gene expression. *Eur J Biochem* 2004;271:1924–1937. [PubMed: 15128302]
6. Kenjo A, Takahashi M, Matsushita M, Endo Y, Nakata M, Mizuochi T, et al. Cloning and characterization of novel ficolins from the solitary ascidian, *Halocynthia roretzi*. *J Biol Chem* 2001;276:19959–19965. [PubMed: 11259430]
7. Gokudan S, Muta T, Tsuda R, Koori K, Kawahara T, Seki N, et al. Horseshoe crab acetyl group-recognizing lectins involved in innate immunity are structurally related to fibrinogen. *Proc Natl Acad Sci USA* 1999;96:10086–10091. [PubMed: 10468566]
8. Christophides GK, Zdobnov E, Barillas-Mury C, Birney E, Blandin S, Blass C, et al. Immunity-related genes and gene families in *Anopheles gambiae*. *Science* 2002;298:159–165. [PubMed: 12364793]
9. Srinivasan P, Abraham EG, Ghosh AK, Valenzuela J, Ribeiro JMC, Dimopoulos G, et al. Analysis of the *Plasmodium* and *Anopheles* transcriptomes during oocyst differentiation. *J Biol Chem* 2004;279:5581–5587. [PubMed: 14627711]
10. Wang X, Rocheleau TA, Fuchs JF, Hillyer JF, Chen C-C, Christensen BM. A novel lectin with a fibrinogen-like domain and its potential involvement in the innate immune response of *Armigeres subalbatus* against bacteria. *Insect Mol Biol* 2004;13:273–282.
11. Dong YM, Aguila R, Xi Z, Warr E, Mongin E, Dimopoulos G. *Anopheles gambiae* immune response to human and rodent *Plasmodium* parasite species. *PLoS Pathogens* 2006a;26:e52.
12. Adema CM, Hertel LA, Miller RD, Loker ES. A family of fibrinogen-related proteins that precipitates parasite-derived molecules is produced by an invertebrate after infection. *Proc Natl Acad Sci USA* 1997;94:8691–8996. [PubMed: 9238039]
13. Pearce EJ, MacDonald AS. The immunobiology of schistosomiasis. *Nat Rev Immunol* 2002;2:499–511. [PubMed: 12094224]
14. Léonard PM, Adema CM, Zhang S-M, Loker ES. Structure of two *FREP* genes that combine IgSF and fibrinogen domains, with comments on diversity of the *FREP* gene family in the snail *Biomphalaria glabrata*. *Gene* 2001;269:155–165. [PubMed: 11376947]
15. Zhang S-M, Léonard PM, Adema CM, Loker ES. Parasite-responsive IgSF members in the snail *Biomphalaria glabrata*: characterization of novel genes with tandemly arranged IgSF domains and a fibrinogen domain. *Immunogenetics* 2001;53:684–694. [PubMed: 11797103]
16. Zhang S-M, Loker ES. The *FREP* gene family in the snail *Biomphalaria glabrata*: additional members, and evidence consistent with alternative splicing and *FREP* retrosequences. *Dev Comp Immunol* 2003;27:175–187. [PubMed: 12590969]

17. Zhang S-M, Loker ES. Representation of an immune responsive gene family encoding fibrinogen-related proteins (FREPs) in the freshwater mollusc *Biomphalaria glabrata*, an intermediate host for *Schistosoma mansoni*. *Gene* 2004;341:255–266. [PubMed: 15474308]
18. Zhang S-M, Adema CM, Kepler TB, Loker ES. Diversification of Ig superfamily genes in an invertebrate. *Science* 2004;305:251–254. [PubMed: 15247481]
19. Litman GW, Cannon JP, Dishaw LJ. Reconstructing immune phylogeny: New perspectives. *Nat Rev Immunol* 2005;5:866–879. [PubMed: 16261174]
20. Watson FL, Puttmann-Holgado R, Thomas F, Lamar DL, Hughes M, Kondo M, et al. Extensive diversity of Ig-superfamily proteins in the immune system of insects. *Science* 2005;309:1874–1878. [PubMed: 16109846]
21. Dong YM, Taylor HE, Dimopoulos G. AgDscam, a hypervariable immunoglobulin domain-containing receptor of the *Anopheles gambiae* innate immune system. *PLOS Biol* 2006b;4:1137–1146.
22. Terwilliger DP, Buckley KM, Mehta D, Moorjani PG, Smith LC. Unexpected diversity displayed in cDNA expressed in the immune cells of the purple sea urchin, *Strongylocentrotus purpuratus*. *Physiol Genomics* 2006;26:134–144. [PubMed: 16837652]
23. Terwilliger DP, Buckley KM, Brockton V, Ritter NJ, Smith LC. Distinctive expression patterns of 185/333 genes in the purple sea urchin, *Strongylocentrotus purpuratus*: an unexpectedly diverse family of transcripts in response to LPS, beta-1,3-glucan, and dsRNA. *BMC Mol Biol* 2007;8:16. [PubMed: 17331248]
24. Stout, BA.; Adema, CM.; Zhang, S-M.; Loker, ES. The biology of FREPs: diversified lectins with fibrinogen-related domains from the freshwater snail *B. glabrata*. In: Vasta, G.; Ahmed, H., editors. *Animal Lectins: A Functional View*. Taylor & Francis: CRC Press; 2008. in press
25. Hertel LA, Adema CM, Loker ES. Differential expression of *FREP* genes in two strains of *Biomphalaria glabrata* following exposure to the digenetic trematodes *Schistosoma mansoni* and *Echinostoma paraensei*. *Dev Comp Immunol* 2005;29:295–303. [PubMed: 15859234]
26. Galaktionov, KV.; Dobrovolskij, AA. The biology and evolution of trematodes. Dordrecht, The Netherlands: Kluwer Academic Publishers; 2003.
27. Hoffmann JA. The immune response of *Drosophila*. *Nature* 2003;426:33–38. [PubMed: 14603309]
28. Lewis FA, Patterson CN, Knight M, Richards CS. The relationship between *Schistosoma mansoni* and *Biomphalaria glabrata*: genetic and molecular approaches. *Parasitology* 2001;123:S169–S179. [PubMed: 11769281]
29. Stibbs HH, Owczarzak A, Bayne CJ, Dewan P. Schistosome sporocyst-killing amebas isolated from *Biomphalaria glabrata*. *J Invertebr Pathol* 1979;33:59–170.
30. Loker ES, Hertel LA. Alternations in *Biomphalaria glabrata* plasma induced by infection with the digenetic trematode *Echinostoma paraensei*. *J Parasitol* 1987;75:505–513.
31. Zhang S-M, Zeng Y, Loker ES. Characterization of immune genes from the schistosome host snail *Biomphalaria glabrata* that encode peptidoglycan recognition proteins and gram-negative bacteria binding protein. *Immunogenetics* 2007;59:883–898. [PubMed: 17805526]
32. Winnepeninck B, Backeljau T, DeWachter R. Extraction of high-molecular-weight DNA from mollusks. *Trends Genet* 1993;9:407–407. [PubMed: 8122306]
33. Schultz J, Milpetz F, Bork P, Ponting CP. SMART, a simple modular architecture research tool: Identification of signaling domains. *Proc Natl Acad Sci USA* 1998;95:5857–5864. [PubMed: 9600884]
34. Jiang Y, Loker ES, Zhang S-M. *In vivo* and *in vitro* knockdown of *FREP2* gene expression in the snail *Biomphalaria glabrata* using RNA interference. *Dev Comp Immunol* 2006;60:855–866. [PubMed: 16442620]
35. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta CT}$ method. *Methods* 2001;25:402–408. [PubMed: 11846609]
36. Nakayama M, Nakajima D, Nagase T, Nomura N, Seki N, Ohara O. Identification of high-molecular-weight proteins with multiple EGF-like motifs by motif-trap screening. *Genomics* 1998;51:27–34. [PubMed: 9693030]
37. Stein RA, Staros JV. Evolutionary analysis of the ErbB receptors and ligand families. *J Mol Evol* 2000;50:397–412. [PubMed: 10824084]

38. Kocks C, Cho JH, Nehme N, Ulvila J, Pearson AM, Meister M, et al. Eater, a transmembrane protein mediating phagocytosis of bacterial pathogens in *Drosophila*. *Cell* 2005;123:335–346. [PubMed: 16239149]
39. Kurucz É, Márkus R, Zsámboki J, Folkl-Medzihradsky K, Darula Z, Vilmos P, et al. Nimrod, a putative phagocytosis receptor with EGF repeats in *Drosophila* plasmatocytes. *Curr Biol* 2007;17:649–654. [PubMed: 17363253]
40. Doolittle RF. A detailed consideration of a principal domain of vertebrate fibrinogen and its relatives. *Prot Sci* 1992;1:1563–1577.
41. Lu J, Le Y. Ficolins and the fibrinogen-like domain. *Immunobiol* 1998;199:190–199.
42. Wang X, Zhao Q, Christensen BM. Identification and characterization of the fibrinogen-like domain of fibrinogen-related proteins in the mosquito, *Anopheles gambiae*, and the fruitfly, *Drosophila melanogaster*, genomes. *BMC Genomics* 2005;6:114. [PubMed: 16150145]
43. Kurachi S, Song Z, Takagaki M, Yang Q, Winter HC, Kurachi K, et al. Sialic-acid-binding lectin from the slug *Limax flavus*: cloning, expression of the polypeptide, and tissue location. *Eur J Biochem* 1998;254:217–222. [PubMed: 9660173]
44. Kojima T. The mechanism of *Drosophila* leg development along the proximodistal axis. *Dev Growth Diff* 2004;46:115–129.
45. Ju JS, Cho MH, Brade L, Kim JH, Park JW, Ha NC, et al. A novel 40-kDa protein containing six repeats of an epidermal growth factor-like domain functions as a pattern recognition protein for lipopolysaccharide. *J Immunol* 2006;177:1838–1845. [PubMed: 16849495]
46. Hori S, Kobayashi A, Natori S. A novel hemocyte-specific membrane protein of *Sarcophaga* (flesh fly). *Eur J Biochem* 2000;267:5397–5403. [PubMed: 10951197]
47. Adema, CM.; Loker, ES. Specificity and immunobiology of larval digenean-snail associations. In: Fried, B.; Graczyk, TK., editors. *Advance in Trematode Biology*. NY: CRC Press; 1997.

A

```

nt 1  cacgtagcagtggaacttcaacataaacttctctctgacagactgtgtcccatctgctgtATGAAATGCGTTGGTGGTGCTACTTTCTTTCTGGAAATG
aa 1  M K M R W W L L L S F L E M
106  GCCCAATAACTGTATTGCCAAGACATAATAGAGGAGTGCCCGTTGGTTGGTTCGGCACAGAATGTCAATATCAAAACCCACTGCCGGACCAAGTGTGAACAGAT
15  A H K L Y C Q D I I E E C P V G W F G T E C Q Y Q T H C R T K C E T D
211  GGGCACTGTATGGGCAACGCTGCAAGATAGGATGTTGGTTACAAATGTCAGTACCAGGATATGGTGTCTAGATTTGAAGTCCCGTATCCCAAAATCCACTG
50  G D C M G N V C K I G W F G Y K C Q Y Q D M V S R F E V P Y P K I P L
316  TCCAAATGGTCAAAATACCAATGTCTCTGAGTGTATAGAGGATGAGAGAATTGAAGAAGTCCATGTCTACTTAAAGTCGACAACTACTTTACCTGGTTAAGTTTA
85  S K W S N T N V S E C I E D E R I E E V H V Y L K S T I Y F T W L S L
421  ACAGTCAAGAAATGCGGATTTGTAGACCTTTTACCATTCAACTTGAGAACGGTGACAAGAAATGGAATGTGAAAGCAGCTATACACTATCAGAAATATAAT
120  T V K N A D L L D L F T I Q L E N G D K K W K C E K Q L Y T I R N N N
526  ACTGCGCAATAGAATGTCAAGGGAGAGAGCTTGTGAGCAAACTCACTCTTAAAGGCAAAGGAGTTACTTCCCTTTGTTCTATTACGTCAGCAAAGTCGCAAC
155  T A R I E C Q G R E L V S K L T L K G K G V T S L C S I H V S K G R N
631  TTGGCGATCAAAACAAAAGTGACGCTATCCTTAGGCATGGACACGACGCTTCCCTGATTGGTGGACGGAATAAGTGACACCCCAACTCAGTTCTCTGCACA
190  L A I K Q K V T L S L G M D T N A S P D L V D G I S D T T Q L S S C T
736  CCAATTACCGGTTACAAAATGCGCCGCGCAAAACATGGTCTCTAACTTTTGTGAACCTGTTGTGGCAAGTAGTATTGAAATTTATTACAAAGGCACTATTGGTA
225  P I T G S Q N A P R K T W S L T F A E P V V A S S I E I Y Y K G T M V
841  GACAAGAGCCCTTGTCTATTGAAACACTAGATGGAGACAAAGCTGTGGATTTTAATGAACTGACTCAAAGGATTATATAAATTTATCAACGTTGAATGGAGAT
260  D K S L L L I E T L D G D K A V D F N E T D S K D Y I N L F N V N G D
946  CCAATCACTCAGTTTAAAGTAACAGATGCGAAGAAAGAAAACGTGTCTAACCTGATTACCCATATGTGAGGTCAAAATCTTTGAGTGTGTGTACCAGGTAATGG
295  P I T Q F K V T D A K K E N V S N L I T L C E V K I F G V C V P G K W
1051  GGCTTAAACTGTAATAACACCTGTGACAGCAAATGTAATGCATCTGTATGAGATTGATGGAACATGCCCTGATGGTTGCCTTGGTTACAGCGGTGATCAGGCC
330  G L N C N N T C D S K C N A S C N E I D G T C P D G C L G Y S G D H A
1156  TGTAGTAAACTTTGTGACAGCTCAATGGGTGTGAAGTCAAGAAAGGATTGGCATAAGAAATGTGCCCTCCACGCTTGTGACGACAAAATGGTGAATGCATC
365  C S K T C E T A Q W G V N C K K D C D K K C A F H A C D D K N G E C I
1261  TTAGGCTGCAACGGCTTCTCGGACCCCGCAATGCTCTGAAGTTTGCAGTGACTCTTACGGCCCCAACTGTTTATGTACGAGCGGGAACCACTTTTGC
400  L G C N G F S D P P Q C S E V C S D D S Y G P N C L C T D G K Q P F C
1366  CCACCGGAAATACCAGGTCAAGTGACTCCCACTGAACCAAGCCTGAATCTTGCCAAGGCGTAAATAGCACCAAGAAAGGGTCGTGGTGAATTTGGCTCGGGGG
435  P P E I P G Q V T P I E P K P E S Q G V N S T Q E R V V V L A R G
1471  ATAAAGGTCATGTGTGACACTCAGACAGACGGCGGAGGCTGGATCGTCATTCAACGAAGATTCAACGGCACAGTTCGATTTCTCAGAGGCTGGTCTGAGTACCGT
470  I . K . X . M . D . T . Q . T . D . G . G . G . I . X . I . Q . R . F . N . T . Y . R . F . R . G . W . S . E . R .
1576  AAAGGCTTCGGAGACTACGACATGGGAGAGTTTACCTGGGCAACGAGAACCTCTTCAGCTGACGTCCACGGGGCAATACGAGCTCAAAATAGACCTGAAGTAT
505  K . G . G . D . Y . D . M . G . F . Q . D . N . E . N . H . F . Q . L . F . S . T . G . Q . Y . E . L . K . I . P . L . K . X
1681  AATGACAAAGAGTACAGCTTAACTACTCGAGCTTTCAAATTTGCAATGAGAGAGAGAAGTACAACTTCAGATCCAGACTTTCTTCAACGATTTCTATTGAGAC
540  N . D . K . E . Y . S . N . S . S . F . Q . I . S . E . R . E . K . Y . K . L . Q . I . Q . T . F . N . D . S . I . D .
1786  GAGCTGAGAGCTACCAACACGGCATGCCCTTCAGCACTTTTCGACACACACACGACGCGCAGCGTCAGCTCCAACCTGTGCCGAGATTCCGCTGAGCCTGGTGG
575  R . H . R . Y . X . H . N . G . M . R . D . R . X . D . N . D . N . D . R . S . Y . S . S . N . D . A . R . R . S . A . R . A . M . D
1891  TACAACAATTGCTATGAGGTCAATCTGAACGGCTTTTGAACGACACTTTATTGGCCATGGTTTGCATGGAGGTCAATAACTGGCTATGATGGTGGTGAACA
610  Y . N . N . Y . E . V . N . L . N . F . N . D . T . L . F . G . H . G . L . Q . W . R . S . I . T . G . Y . D . G . G . V . T
1996  TTTACCGCATCAAAATCAGAGCTGGGGAGCTTAATTTAACTggtagcagtgccctaaagaagtgaataacgaagagaacttctaactactaatcatgtta
645  F T E M K I R E V G S L I *
2101  gcagtaataatcttaacgatattactggttagttatagaattagaataacacttaataatgaaattagaataagtcacacaaagatgtactcaattcttagtacctat
2206  aactaactggtataatacatgttcgaattcagttgtaactctcacacttagggtataaaaaaaatgttattaaaaatttaattagatttcaaaaaaaagggga
2311  aaagattatattgaaatgaaaaaaatattttcaataaaaaaaactcttaaatattgcttacttctatcatgtcaaatgaaacataataaatgttgatca
2416  aatctgtgtacctatytatgttattgaactttatawccagtaatttgaatccaaagttcgataaaaaatcattgtaactttaagtttataaagaatgaagattaca
2521  cgtgaaaaattaaaaaaataaaaaaa

```

B **Fig. 1.**

A) Nucleotide (nt) and deduced aa sequences of FReM (EF467291). The signal peptide (SP) is underlined by a double solid line. Two EGF repeats and FBG domain are indicated by the single solid and dotted lines, respectively. The corresponding positions for the 24 canonical residues in the FBG domain are boxed (identical) or circled (non-identical) (applied to figure 5 below). Numbers in the left margin designate nt (Arabic) and aa (italicized). B) Schematic of FReM (not drawn to scale).

	*	●	*	*	*	*	*	●	*	*	●	●	*																																		
sEGF1	C	V	P	G	K	W	G	L	N	C	N	N	T	C	D	S	K	--	C	N	A	S	--	C	N	E	I	D	-	G	T	C	P	D	G	C	L	-	G	Y	S	G	D	H	A	C	
sEGF2	C	E	T	A	Q	W	G	V	N	C	K	K	D	C	D	K	K	--	C	A	F	H	A	-	C	D	D	K	N	-	G	E	C	I	L	G	C	N	-	G	F	S	D	P	P	Q	C
mEGF1	C	E	A	Q	K	W	G	P	D	C	S	R	P	C	T	T	---	C	K	N	N	G	V	C	H	E	D	T	-	G	E	C	I	--	C	P	P	G	F	M	G	-	R	T	C		
mEGF2	C	E	P	H	T	F	G	R	T	C	K	E	R	C	S	G	P	E	G	C	K	S	Y	V	F	C	L	P	D	P	Y	G	-	C	S	--	C	A	T	G	W	R	G	-	L	Q	C
mEGF3	C	P	S	G	Y	Y	G	P	D	C	K	L	R	C	H	---	C	T	N	E	E	I	C	D	R	F	Q	-	G	-	C	L	--	C	S	Q	G	W	Q	G	-	L	Q	C			
hEGF1	C	E	A	Q	K	W	G	P	E	C	N	H	L	C	T	A	---	C	M	N	N	G	V	C	H	E	D	T	-	G	E	C	I	--	C	P	P	G	F	M	G	-	R	T	C		
hEGF2	C	E	L	H	T	F	G	R	T	C	K	E	R	C	S	G	Q	E	G	C	K	S	Y	V	F	C	L	P	D	P	Y	G	-	C	S	--	C	A	T	G	W	K	G	-	L	Q	C
hEGF3	C	H	P	G	F	Y	G	P	D	C	K	L	R	C	S	---	C	N	N	G	E	M	C	D	R	F	Q	-	G	-	C	L	--	C	S	P	G	W	Q	G	-	L	Q	C			
zEGF1	C	R	A	G	F	W	G	P	N	C	T	E	S	C	P	R	---	C	A	N	G	G	V	C	D	D	T	T	-	G	E	C	V	--	C	P	P	G	F	R	G	-	H	T	C		
zEGF2	C	G	E	G	R	F	G	A	G	C	K	E	R	C	V	D	G	-	V	C	R	A	L	V	F	C	L	R	D	P	Y	G	-	C	S	--	C	A	S	G	W	R	G	-	L	S	C
zEGF3	C	P	D	G	Y	Y	G	A	G	C	T	Q	K	C	V	---	C	A	K	G	R	-	C	D	R	F	R	-	G	-	C	V	--	C	A	-	G	R	H	G	-	S	R	C			

Consensus CxxxxxGxxCxxxCxxxxxCxxxxxCxxxxxGxCxxxCxxGxxGxxxC

Fig. 2.

Alignment of the snail EGF (sEGF) repeats of FReM with EGF repeats from other species. The EGF repeats that are most closely related to sEGF repeats are located within the tyrosine-protein kinase receptor TIE-2 precursors of mouse (mEGFs) (Q02858), human (hEGFs) (Q02763) and zebrafish (zEGFs) (O73791). The conserved residues with >90% identity across all repeats are shaded. The conserved cysteine (C) and glycine (G) were indicated by asterisk and filled circle, respectively.

Fig. 3. Southern blot analyses of *FReM* (A) and *FREP14* (B). Three restriction enzymes, *EcoR* I (lane 1), *Hind* III (lane 2) and *Pst* I (lane 3) were used to digest genomic DNA. Size markers in kilobases (kb) are indicated on the right.

Fig. 4.

Expression of *FReM* in M line and BS-90 snails after exposure to *E. paraensei* or *S. mansoni* revealed by Northern blotting (**A**) and qPCR (**B**). In figure **A**, lanes 1 to 3 show the expression levels of *FReM* in M line snails that were not exposed (lane 1), or were at 2 dpe to *E. paraensei* (lane 2) or to *S. mansoni* (lane 3). Note that a relatively less total RNA sample suggested by the intensity of rRNA found in lane 3 may be due to slight degradation. However, the abundance of mRNA detected in lane 3 was still higher than that of lane 1, indicating that the expression level of *S. mansoni*-infected snails (lane 3) is higher than that of non-exposed control snails (lane 1). For BS-90 snails, lanes 1 to 3 show the mRNA abundance of *FReM* from snails that were not exposed (lane 1), or were at 2 dpe to *E. paraensei* (lane 2) or to *S. mansoni* (lane 3). Lower panel shows rRNA stained with ethidium bromide used as a loading control (applied to figure 6 and figure 7 below). In figure **B**, the expression level of *FReM* is relative to the non-exposure control (bar 1) in M line or BS-90 snails after exposure to *E. paraensei* (bar 2) or *S. mansoni* (bar 3). The expression value of non-exposure control snails was set to a value of 1. The numbers above the bars are the average value among technical replicates.

Fig. 5.

A) Nucleotide and deduced aa sequences of FREP14 (EF467292). SP, IgSF and FBG are underlined by doubled solid, dashed and dotted lines respectively. Numbers in the left margin are designated as those in Fig.1A. B) Schematic of FREP14 (not drawn to scale).

Fig. 6. Expression of *FReM* and multiple *FREP* members in M line snails during ontogenesis revealed by Northern blotting (**A**) and qPCR (**B**). **A**) The size of unexposed snails used in lanes 1 to 4 is 0.2–0.4, 0.7–0.8, 1.4–1.6, and 2.0–2.3 cm, respectively. A total of three Northern blots generated from same RNA sources were used for all expression studies presented in the figure A using the re-probing method. Expression profiling of *FReM*, *FREPs* 13 and 14 were obtained from one membrane. Similarly, *FREPs* 4 and 12 were from one membrane, and *FREPs* 2 and 3 were from another. **B**) Shows the relative expression of seven genes in the four size classes of snails. The RNA samples used to produce the cDNA assayed by qPCR were the same as those for Northern blots (left panel).

Fig. 7.

Expression of *FReM* and multiple *FREP* members in M line snails at 52 dpe to *S. mansoni* or *E. paraensei* revealed by Northern blotting (A) and qPCR (B and C). A) Lane 1 shows the mRNA extracted from the non-exposed snails of similar sizes as infected snails (1.4–1.6 cm). Lanes 2 and 3 show the mRNA from two different sizes of infected M line snails (lane 5: 1.3–1.5 cm; lane 6: 1.5–1.6 cm) at 52 dpe to *S. mansoni*. B) The same RNA used for Northern analysis was used to produce the cDNA assayed by qPCR. The expression level of the *S. mansoni*-infected snails (bar 2 and 3) is shown relative to the non-exposed snails (bar 1). C) Relative expression of non-exposure M line snails (bar 1) and the infected M line snails of

similar sizes at 52 dpe to *E. paraensei* (bar 2) based on qPCR analysis. Northern analysis was not performed due to the limited amount of RNA sample available.

Table 1
Amino acid (aa) identity (%) of FBG domain or IgSF domain between the FReM or FREP14 and the corresponding domain of different FREPs

FBG domain							
	FREP2	FREP4	FREP3	FREP7	FREP12	FREP13	FReM
FReM	54	53	57	54	57	51	
FREP14	66	61	65	62	67	66	54
IgSF domain							
	FREP2	FREP4	FREP3	FREP7	FREP12	FREP13	FReM
FREP14	49	35	43	42	38	40	

The aa identity between two sequences was performed using NCBI Blast 2 sequences program (<http://www.ncbi.nlm.nih.gov/blast/bl2seq/wblast2.cgi>). No significant similarity was found between IgSF of FREP14 and IgSF1 of any two-IgSF FREPs (i. e., FREPs 3, 7, 12, and 13). So, for the two-IgSF FREPs, the identity showed in the table 1 is the identity between IgSF of FREP14 and IgSF2 of a corresponding FREP. Exact position of each domain for the analysis is given in this paper and previous publications [14–16].