Individual sequence variability and functional activities of fibrinogen-related proteins (FREPs) in the Mediterranean mussel (*Mytilus galloprovincialis*) suggest ancient and complex immune recognition models in invertebrates.

Alejandro Romero*, Sonia Dios*, Laura Poisa-Beiro*, Maria del Mar Costa*, David Posada†, Antonio Figueras*, Beatriz Novoa*[1].

* Instituto de Investigaciones Marinas. Consejo Superior de Investigaciones Científicas (CSIC). Eduardo Cabello 6, 36208, Vigo (Spain).

† Department of Biochemistry, Genetics and Immunology, Universidad de Vigo, Spain.

The sequences presented in this article have been submitted to GenBank under accession numbers as follows (www.ncbi.nlm.nih.gov/Genbank): HQ236391-HQ236407 and HO666697-HO666738.

[1]. Address correspondence and reprint requests to Dr. Beatriz Novoa, Instituto de Investigaciones Marinas, Consejo Superior de Investigaciones Científicas (CSIC), Eduardo Cabello, 6. 36208 Vigo, Spain. E-mail address: virus@iim.csic.es. Tlf: 34 986 21 44 63. Fax: 34 986 29 27 62

Abbreviations used in this paper: PAMP, pathogen-associated molecular patterns; MBL, mannose-binding lectin; MASPs, MBL-associated serine proteases; MAC, membrane attack complex; EST, expressed sequence tag; LTA, lipoteichoic acid ; FREP, fibrinogen-related protein; ORF, open reading frame; PRRs, Pattern Recognition Receptors; TLRs, Toll-like receptors.

Abstract

In this paper, we describe sequences of fibrinogen-related proteins (FREPs) in the Mediterranean mussel *Mytilus galloprovincialis* (MuFREPs) with the fibrinogen domain probably involved in the antigen recognition, but without the additional collagen-like domain of ficolins, molecules responsible for complement activation by the lectin pathway. Although they do not seem to be true or primive ficolins since the phylogenetic analysis are not conclusive enough, their expression is increased after bacterial infection or PAMPs treatment and they present opsonic activities similar to mammalian ficolins. The most remarkable aspect of these sequences was the existence of a very diverse set of FREP sequences among and within individuals (different mussels do not share any identical sequence) which parallels the extraordinary complexity of the immune system, suggesting the existence of a primitive system with a potential capacity to recognize and eliminate different kind of pathogens.

1. Introduction

Although in recent years there have been important advances in invertebrate innate immunity, there is not a comprehensive view of the immune mechanisms utilized across the broad spectrum of invertebrate phyla (Zhang et al., 2004).

Fibrinogen-related proteins (FREPs) are a family of glycoproteins that contain in the C-terminal portion a fibrinogen-like (FBG) domain but differ in the N-terminal region. This family comprises diverse proteins such us tenascins, tachylectins, ficolins, angiopoietins, ixoderins or fibrinogen β and γ chains (Adema et al., 1997; Lu and Le, 1998; Zhang et al., 2001; Gorbushin et al., 2010). FREPs are universally found in vertebrates (Doolittle,1992; Matsushita et al., 1996), urochordates (Kenjo et al., 2001) and invertebrates such us echinoderms (Xu and Doolittle, 1990), molluscs (Knibbs et al., 1993; Adema et al., 1997), and arthropods (Baker et al., 1990; Gokudan et al., 1999; Rego et al., 2005; Dixit et al., 2008), suggesting that a fibrinogen-related domain must have existed before the divergence of protostomes and deuterostomes (Gorbushin et al., 2010).

FREPs play an important role in the innate immune response of invertebrates against pathogens. They are up-regulated following relatively specific immunostimulation (Adema et al., 1999) and they bind to pathogens and precipitate parasites antigens (Adema et al., 1997; Zhang et al., 2008). For example, hemocytes from the snail *Biomphalaria glabrata* produced high amount of FREPs following challenge with trematode parasites and have lectin-like properties allowing them to precipitate soluble antigens derived from trematodes (Adema et al., 1997); FREPs from the mosquito *Armigeres subalbatus* (also called aslectin) are up-regulated by bacterial challenge and are able to bind bacteria *Escherichia coli* and *Micrococcus luteus* (Wang

et al., 2004). FREPs are also frequently implicated as mediators of non self-recognition in invertebrates (Richards and Renwrantz, 1991, Drickamer, 1995; Zhang et al., 2004; Hertel et al., 2005). In *Biomphalaria,* FREPs show a relatively high diversity (Zhang and Loker, 2003; Zhang et al., 2004). Likewise, vertebrate FREPs are also involved in immune reaction. It has been shown that human FREP1 is an acute phase reactant (Liu and Ukomadu, 2008), and mouse FREP2 contributes to CD4+CD25+ regulatory T cell activity (Shalev et al., 2008).

One of the most important FREPs is the ficolin group since they are important components of the lectin complement pathway. In mammals, ficolins are oligomeric proteins characterized by a short N-terminal segment, a collagen-like domain and a C-terminal fibrinogen-like domain (Matsushita et al., 1996). One of the main functions of ficolins is the recognition of sugars present in microorganisms (Tanio et al., 2007) playing an important role in innate immunity because they are able to enhance phagocytosis and activate the complement pathway (Turner, 1996; Lu and Le, 1998; Matsushita et al., 2000; Zhang and Ali, 2008).

The importance of ficolins in the recognition of pathogens is unquestionable in mammals. However, these molecules have not been found so far in fish, birds or reptiles (Garred et al., 2010), and there is only one case of ficolin-like proteins reported in ascidians (Kenjo et al., 2001).

In the Mediterranean mussel, the major immunological defense system is carried out by hemocytes present in the hemolymph (Allam et al., 2001; García-García et al., 2008; Pipe, 1990; Pipe et al., 1997). However, little is known about how this organism responds against pathogens. Despite the practice of culturing these animals at high densities, no important mortalities associated with pathogens have been found in these molluscs so far, suggesting the existence of an efficient defense mechanism.

Interestingly, clams and oysters cultured nearby have experienced mass mortalities associated with pathogens that have been also found in mussels. Accordingly, we have recently shown the high individual sequence variability for myticin C, a novel class of antimicrobial peptide (Pallavicini et al., 2008), suggesting that this wide repertoire of sequences could be related to the high degree of disease resistance observed (Costa et al., 2009). Another relevant aspect of bivalve molluscs is their important filtering activity: one adult mussel can filter roughly 8 liters of water in one hour (Meyhöfer, 1985; Norman, 1988; Hugh et al., 1992), which implies that they are in intimate contact with microorganisms.

The present work constitutes another astonishing example of highly diverse immune molecules in these animals and provides the first evidence of their possible involvement in immune defense.

2. Materials and Methods

2.1. Animals

Mediterranean mussels (*Mytilus galloprovincialis*) with a maximum shell length of 6 cm were obtained from a commercial shellfish farm from the Ría de Vigo (NW of Spain) during the summer season. Animals were maintained in open circuit filtered seawater tanks at 15 ºC with aeration. They were fed daily with *Isochrysis galbana*, *Tetraselmis suecica* and *Skeletonema costatum*. Bivalves were acclimatized for 1 week before the experimental work. All animal experiments were reviewed and approved by the CSIC National Committee on Bioethics.

2.2. Phylogenetic analysis

Fourteen mussel sequences putatively homologous to FREPs (GeneBank accession numbers from HQ236392 to HQ236405) were selected from two cDNA libraries previously constructed (Venier et al., 2009). Additional 62 ficolin and FREPs sequences from different animals were downloaded from GenBank. Protein sequences were aligned with MAFFT (Katoh et al., 2005) using the E-INS-i algorithm for multiple conserved domains and long gaps. Ambiguous columns in the alignment were filtered out with Gblocks (Castresana, 2000), with a minimum of 40 sequences for conserved/flanking positions, a maximum of 4 contiguous non-conserved positions, a minimum length of 4 amino acids in every block, and allowing gap positions. The best-fit model of amino acid replacement was selected according to the Akaike Information Criterion (AIC) (Akaike, 1974) using ProtTest (Abascal et al., 2005) and Phyml 3.0 (Guindon et al., 2010). This model was used to estimate a maximum likelihood (ML)

phylogenetic tree with RaxML 7.2.6 (Stamatakis, 2006). Nodal support was estimated using the same program with 1000 bootstrap replicates (Felsenstein, 1985).

2.3. Amplification of groups of FREPs by Rapid Amplification of cDNA Ends (RACE)

Total RNA was isolated from hemolymph using Trizol reagent according to the manufacturer's instructions. The SMART RACE cDNA Amplification Kit (BD Clontech) was used to complete the 5'- and 3'-cDNA ends of two representative types of FREPs found in the alignments (groups 2 and 3) (GeneBank accession numbers HQ236406 and HQ236407, respectively) using specific primers designed with the Primer3 (v. 0.4.0) software (Group 2: CAA ATC GTT GCT GTA TCA CCG TCC, Group 3: GAA GCC ATC GAA AGA GCC TGT CGG G). PCR products were purified from 1.2% agarose gel, subcloned using the Original TOPO T/A Cloning Kit, and sequenced.

2.4. Structural analysis

A consensus sequence for each group of FREPs was obtained by using the CAP3 Sequence Assembly Program and aligned with sequences from *Tachypleus tridentatus, Halocynthia roretzi, Xenopus laevis, Rattus norvegicus, Sus scrofa* and *Homo sapiens* available in the GenBank database using ClustalW2. The topology prediction of mussel FREPs was done using pTARGET software (Guda, 2006). The prediction of signal peptide, cleavage sites and the trans-membrane regions were carried out using different available software from the ExPASy Proteomics Server (SignalIP 3.0, SOSUI and SMART software, respectively).

2.5. Constitutive expression of FREPs

The expression of the three representative groups of FREPs was analyzed by quantitative PCR in different tissues. Mussel hemocytes, muscle, mantle, gills, gonads, foot and gland were extracted and preserved in Trizol Reagent (Invitrogen) until use. Total RNA was extracted using Trizol Reagent (Invitrogen) following the manufacturer's instructions. RNA (5 μg per sample) was treated with DNase I (Ambion) to remove contaminating DNA, and first-strand cDNA was synthesized with SuperScript™ II Reverse Transcriptase (Invitrogen). Quantitative PCR was performed using the following set of primers (Group 1; G1F: CCT GAC AAA TGC AAC AGT GG, G1R: TGG CCG TTG TGA TGT TCT AA. Group 2; G2F: GTG ATG CAT TCA GCG GAC TA and G2R: CCC CAA TTG ATA CCA GAT GC. Group 3; G3F: CAA CGT TGG TGA CTC ATT GG and G3R: CCG CCA AGA TAC TGT CCA TT). A total of 0.5 μl of each primer (10 μM) was mixed with 10.5 μl of SYBR green PCR master mix (Applied Biosystems) in a final volume of 21 μl. The standard cycling conditions were 95 ºC for 10 min, followed by 40 cycles of 95 ºC for 15 s and 60 ºC for 1 min. The comparative CT method (2-ΔΔCT method) was used to determine the expression level of analyzed genes (Livak and Schmittgen, 2001). The expression of candidate genes was normalized using the β-actin as housekeeping gene (ActinF: AAC CGC CGC TTC TTC ATC TTC and ActinR: CCG TCT TGT CTG GTG GTA). Fold units were calculated by dividing the normalized expression values of infected tissue by the normalized expression values of the controls. Data were analyzed using Student's t-test. The results were expressed as the mean ± standard deviation and differences were considered statistically significant at $p<0.05$.

2.6. Genomic organization

Genomic DNA from adult mussels was extracted from hemolymph with DNAzol (Invitrogen) following the manufacturer's instructions. The genomic sequence of FREP from group 1 (GeneBank accession number HQ236391) was then analyzed by PCR after consecutive amplification with specific primers designed using Primer3 (Table I). The PCR reaction was performed with a high fidelity Taq polymerase (TaKaRa ExTaq™ Hot Start Version; TaKaRa Bio Inc., Otsu, Siga, Japan), and the cycling protocol was 94 ºC for 5 min, 40 cycles of 94 ºC for 30 s, 50 ºC for 1 min and 72 ºC for 1 min followed by a final extension of 72 ºC for 7 min. The predicted intron-exon structure of the genomic sequence was obtained using Wise2 software.

2.7. Individual variability

Total RNA from three mussels was individually extracted, and the individual variability of the cDNA sequences of FREPs from group 1 was then analyzed by PCR using the specific primers previously described (G1F and G1R). The PCR reaction was done with a high fidelity Taq polymerase (TaKaRa ExTaq™ Hot Start Version; TaKaRa Bio Inc., Otsu, Siga, Japan), and the cycling protocol was 94 ºC for 5 min followed by 40 cycles of 94 ºC for 1 min, 55 ºC for 1 min and 72 ºC for 1 min with a final extension step of 72 ºC for 10 min. Products were analyzed as described above, and the PCR products matching the predicted product size (389 bp) were cloned into the pCR2.1-TOPO vector (Invitrogen) using DH5α™ Competent Cells (Invitrogen). Thirty clones from each mussel were sequenced (GeneBank accession numbers from HO666697 to HO666738), and the identity between sequences was analyzed using ClustalW2 software, in terms of the number of different nucleotides detected in the alignment.

2.8. Enrichment of FREPs in plasma fractions and functional activity

Hemolymph from 50 mussels was collected and centrifuged. FREPs were isolated from serum by affinity chromatography on a GlcNac-Sepharose 6B column at 4 ºC. Briefly, the column was washed with 7 M guanidine hydrochloride and three times with buffered solution (50 mM Tris-HCl, 1 M NaCl, 50 mM $CaCl_2$). The GlcNac-Sepharose was dispensed in sample tubes, which were kept overnight at 4 ºC with gentle mixing. The sample was loaded onto the column and unbound proteins were washed out using the same buffer. The bound proteins were eluted with 150 mM GlcNac in 50 mM Tris-HCl, 1 M NaCl and 50 mM $CaCl_2$. The presence of proteins in the bound fraction was assessed by SDS-PAGE on a 15% acrylamide:bisacrylamide gel using a Mini-PROTEAN electrophoresis system (Bio-Rad). Bands of SDS-PAGE gel were analyzed using a PMF system (Proteomic unit, Universidad Autónoma, Madrid).

The functional activity of the purified FREPs was assayed with phagocytosis experiments. One milliliter of hemolymph from each of 16 mussels was placed into individual wells of tissue-culture 24-well plates. The number of hemocytes in each sample was estimated by counting cells with the aid of a hemocytometer chamber. Hemocytes were allowed to adhere for 30 min in the dark at 15 °C. Hemolymph supernatant was then removed and cells were washed with filtered sea water. Fluorescein-labeled 1 μm latex beads were added at a 10:1 target:hemocyte ratio. Purified FREPs were also added to the plates at a final concentration of 2.27 μg/ml. Hemocytes with latex beads not treated with FREPs and hemocytes without FREPs or latex beads were included as controls. After 2 h of incubation at 15 °C in the dark, beads not internalized were removed by gently washing wells twice with 500 μl of PBS. Cells were finally resuspended in 150 μl of PBS by carefully detaching them from the bottom of the well using a rubber cell scraper. Fifty microliters of 0.8% trypan blue in PBS

were added to each sample to quench external fluorescence. Phagocytosis was evaluated in the four different hemocyte populations (with different morphology and functional activities) previously described in our laboratory (García-García et al., 2008). Ten thousand cells were then analyzed by flow cytometry. Cell fluorescence was analyzed in the FL-1 channel set to a log scale. FL-1 voltage was adjusted for positive cells (cells that internalized at least one particle) falling within the same fluorescence range. Phagocytosis was expressed as the percentage of cells that internalized at least one fluorescent particle (positive cells).

2.9. Expression of FREPs under different stimuli

To analyze the effect of external stimuli on the expression of the FREPs in hemocytes, several PAMPs (Pathogens-Associated Molecular Patterns) and live bacteria were administered intramuscularly (im) (*in vivo* experiments) or added to hemocyte primary cultures (*in vitro* experiments). For the PAMPs, solutions of 1 mg/ml of poly I:C, zymosan, lipopolysaccharide (LPS) or lipoteichoic acid (LTA) were prepared from a commercial stock (Sigma). To prepare a bacterial CpG solution, bacteria (*V. anguillarum*) were grown in TSA supplemented with 1% NaCl at room temperature over several days. Bacterial DNA was then isolated using phenol-chloroform (Maniatis et al., 1982), and the concentration of CpG was adjusted to 1 mg/ml. For live bacteria, *M. lysodeikticus* and *V. anguillarum* were used as Gram-positive and Gram-negative stimuli, respectively. *M. lysodeikticus* was grown in LB medium at 37 ºC, and *V. anguillarum* was cultured as previously described. Bacteria were resuspended in sterile-filtered seawater to obtain an OD_{620} of 0.033 ($1.6x10^7$ cfu *Vibrio*/ml and $1.2x10^6$ cfu *Micrococcus*/ml).

2.9.1. In vivo stimulation

Eight groups of 48 naive animals were injected on the posterior adductor muscle with 100 µl of PAMP solution (either poly I:C, zymosan, LPS, LTA or CpG) or live bacteria (*V. anguillarum* or *M. lysodeikticus*). The control group was injected with filtered seawater (FSW). All individuals were maintained out of the water for 20-30 min before and after the injection. Each treatment group was individually maintained in tanks with aeration until sampling. After 3, 6 and 24 h post inoculation, hemolymph was collected from the adductor muscle, pooled and adjusted to $2x10^5$ cells/ml. Pooled hemolymph from 4 individuals was used per each sampling point and treatment.

2.9.2. In vitro stimulation

A total of 20 hemocyte primary cultures were obtained from naive mussels. Hemolymph was collected from the adductor muscle of 5 individuals using the methodology previously described. For each experiment, pooled hemolymph (2-5 ml of hemolymph per individual) from 5 animals was used. Hemocytes were then incubated with the PAMPs solution (final concentration of 50 µg/ml) or live bacteria (OD_{620} 0.033). Samplings were performed after 1, 3 and 6 h post-inoculation. All the experiments were performed at 15 ºC and replicated at least twice.

2.9.3. Expression studies by Q-PCR

To determine and quantify the FREPs expression pattern, real time PCR was performed on hemocytes exposed to the different treatments using the cDNAs previously generated. The different treatments analyzed include bacterially infected samples, samples stimulated with each of the various PAMPs (described above) and

untreated control samples. Amplification was carried out using the same protocol previously described. The expression of candidate genes was normalized to the 18S ribosomal RNA as a housekeeping gene (Mussel-18S-F: GTA CAA AGG GCA GGG ACG TA and Mussel-18S-R: CTC CTT CGT GCT AGG GAT TG).

3. **Results**

3.1. Phylogenetic analysis

The final alignment was 366 residues long and was reduced to 166 residues once ambiguous columns were removed. The selected model of amino acid replacement was WAG (Whelan and Goldman, 2001), with a proportion of invariable sites and gamma distributed rate variation among sites (Yang, 1996) (i.e., WAG+I+G). The maximum likelihood estimate of the phylogeny (Figure 1) showed well-supported specific clades, although the relationships among these clades could not be resolved with confidence. The deepest split in the tree defined two lineages, although with very low support. The first lineage included fibrinogens and fibrinogen-like proteins, whereas the second was comprised of ficolins and ficolin-like proteins, tenascins, angiopoietins and tachylectins. Mussel FREPs (MuFREPs) clustered into 3 main groups with high confidence (99% of bootstrap values; Figure 1A). MuFREPs included in groups 1 and 3 were most closely related to each other forming, together with the *Argopecten* FREP, a sister group to the vertebrate ficolins. On the other hand, mussels sequences included in group 2 seemed to conform to a different lineage separated from other invertebrate and mammalian ficolins.

The position of the MuFREPs in the phylogenetic tree seems related to their domain structure (Figure 1B). Mussel sequences only possessed the fibrinogen domain, as observed in other invertebrate proteins such as tachylectins, ixoderins, carcinolectins, aslectins and several FREPs from *Anopheles* and *Branquiostoma*. The fibrinogen domain associated with other domains to constitute not only vertebrate ficolins, in combination with the collagen domain, but also other more complex proteins in combination with EGF, IG or fibronectin domains.

3.2. Structural analysis

The consensus sequence for group 1, with 789 bp, included the complete open reading frame (ORF), encoding a protein of 262 amino acids. Two sequences of 806 and 929 bp in length (from group 2 and 3, respectively) were finished by the RACE technique. The translated nucleotide sequences encoded two proteins of 230 and 249 amino acids, respectively.

The alignment of the MuFREPs sequences with other FREPs available in GenBank revealed that only the fibrinogen-like domain was present (Figure 2). A small leader 5´ signal peptide followed by a specific cleavage site was present in all groups of mussel FREPs. The cleavage site was TTQ-EP for group 1, ANA-EL for group 2 and VNS-TS for group 3. The fibrinogen-like domain consisted of 206-210 residues and was characterized by the presence of mostly hydrophobic amino acids, including four conserved cysteine residues. There were three potential N-glycosylation sites at residues 240, 300 and 305 and two potential calcium-binding sites in the latter half of the COOH-terminal. According to the topology prediction, all of the MuFREPs were classified as extracellular/secreted proteins. The hydrophobic signal peptides, which mediated protein translocation into the endoplasmic reticulum, and the cleavage sites were found in all mussel FREPs.

3.3. Constitutive expression of MuFREPs

To analyze the relative expression of the three groups of mussel FREPs, a quantitative PCR was conducted in several tissues (hemocytes, muscle, mantle, gills, gonads, foot and gland). The most highly expressed sequences were those corresponding to FREPs from group 1, which was especially evident in hemocytes and to a lesser extent in gills and digestive gland. However, the expression of FREPs from

groups 2 and 3 was almost undetectable (Figure 3). No differences were detected among male and female mussels (data not shown).

Because the FREP from group 1 was the highest expressed sequence, we focused in different aspects of this molecule, including genomic structure, expression in response to stimuli and individual variability.

3.4. Genomic organization

The sequence of the FREP gene from group 1 was 5339 nucleotides long. This sequence was divided into six exons, ranging in length from 72 to 247 bp. Their splice acceptor and donor sequences were consistent with the AG/GT rule. The first and second exons were 73 bp in length and encoded a signal peptide of 25 amino acids, a cleavage site and the N-terminal region. The next three exons encoded the upstream portion of the fibrinogen-like domain, and the last exon contained 247 bp encoding the remaining fibrinogen-like domain (Figure 4).

3.5. Individual variability

An assessment of polymorphisms and individual variability in mussel FREP sequences from group 1 was performed on 66 different clones from three different mussels (27, 17 and 22 clones from mussels 1, 2 and 3, respectively). The alignments showed that each sequence was unique with at least four different nucleotides per sequence. Moreover, the three mussels did not present any nucleotide or amino acid sequence in common. In total, 33, 26 and 36 of 389 nucleotides were variable in mussels 1, 2 and 3, respectively, indicating a percentage of variability ranging from 6.6 to 9.2 (Figure 5A).

Fifty-nine percent of all analyzed sequences were unique (18, 15 and 9 nucleotide sequences and 16, 9 and 9 amino acid sequences from mussels 1, 2 and 3, respectively). Furthermore, one specific sequence from each mussel was highly repeated (sequence A: 11 times in mussel 1; sequence B: 5 times in mussel 2 and sequence C: 15 times in mussel 3) (Figure 5B).

3.6. Functional activity of purified FREPs

The elution of mussel serum into the GlcNAc-Sepharose 6B column allowed us to identify a predominant protein of 45 kDa (Figure 6A) and two unidentified and minority proteins of 35 and 20 kDa. The peptide mass fingerprinting of the 45 kDa protein identified three peptides: one peptide as ficolin-2 precursor (peptide 1070.49 ficolin-2 precursor) with an amino acid sequence (QDGSVDFFR), which was present in the three groups of FREPs (highlighted in Figure 2); the other peptides (1163.62 and 1741.82) had similarity with the sequences available in the database (fibronectin and HC immunoglobulin, respectively).

The biological activity of the partially purified FREPs was measured by phagocytosis assay. The incubation of hemocytes with the purified FREPs induced an increase in the phagocytosis of fluorescent beads. Only 2 out of 16 mussels analyzed did not respond to this stimulation (Figure 6B). The phagocytic rate recorded in the different populations of hemocytes was different among mussels. Half of the analyzed mussels increased their phagocytic rate by less than 20% and 37.5% of the remaining mussels increased their activity by more than 20%, reaching up to 60% in some cases (Figure 6C). Moreover, the hemocyte subpopulations presented different levels of phagocytosis. The phagocytic activity in R1 and R2 cells was incremented up to 70% and 50%, respectively. R3 and R4 hemocytes presented less activity and the FREP

treatment only incrementally enhanced the phagocytic activity up to 30% and 10%, respectively (Figure 6D).

3.7. Expression of FREPs under different stimuli

After *in vitro* incubation of hemocytes with several PAMPs, the maximum expression of FREPs was obtained after 3 h post treatment in all cases. However, the results obtained after 1 and 6 h post stimulation did not show significant differences. LTA (50-fold induction) and LPS (34-fold induction) induced the highest FREP expression levels among the PAMPs administered (Figure 7A). Even though the expression levels obtained after poly I:C and zymosan stimulation were mild (4.9 and 7.7 fold increase, respectively), both PAMPs showed statistically significant increments when compared to controls from the same sampling point. The *in vivo* stimulation of mussels with the PAMPs showed that LTA was able to induce the maximum response after 3 h post stimulation (8.8 fold increase) (Figure 7B). Despite a reduction in expression values after 6 and 24 h post inoculation, significantly increased expression values were still observed at 24 h following LTA stimulation (3.9 fold induction). LPS also provoked a high FREP expression (8.1 fold increase) at 24 h post stimulation. However, the remaining PAMPs did not induce significant expression differences when compared to control samples.

The Gram-positive bacteria *M. lysodeikticus* was able to induce a significant increase in expression after 3 h post *in vitro* treatment (7.6 fold induction). No significant differences were found for the other sampling points following either *M. lysodeikticus* or *V. anguillarum* challenge (Figure 7C). In contrast, *in vivo* stimulation by Gram-negative or Gram-positive bacteria provoked a significant increase in the

expression at 24 h post stimulation (1451.2 and 9286.4 fold increase, respectively). No significant differences were found in the remaining sampling points (Figure 7D).

4. Discussion

It is well known that ficolins and MBL are the molecules responsible for complement activation by the lectin pathway (Endo et al., 2007; Runza et al., 2008). Ficolins have been identified in several species of amphibians (Kakinuma et al., 2003) and mammals (Fujimori et al., 1998; Ichijo et al., 1993; Matsushita et al., 1996; Ohashi and Erickson, 1998; Omori-Satoh et al., 2000). However, little is known about the presence of ficolins in fish, birds or reptiles. Ficolins from invertebrate animals have been only reported in the solitary ascidian *Halocynthia roretzi* (Kenjo et al., 2001) and as predicted proteins in *Ciona intestinalis*, *Strongylocentrotus purpuratus* and *Culex quinquefasciatus*. However, to date there has not been any description of these proteins in the protostome lineage. Because of this lack of information, the phylogenetic relationships between ficolin families remain unclear. Other proteins with fibrinogen-like domains, such us tachylectins and tenascins present in invertebrate and vertebrate animals (Gokudna et al., 1999; Kawabata and Tsuda, 2002; Mali et al., 2006; Ju et al., 2009), have been used in the past to understand the evolution of the ficolin genes (Endo et al., 2006). We cannot determine with certainty if the mussel FREPs sequences that we report in this study, resemble primitive structures from which the different FREPs (angiopoietins, tachylectins, carcinolectins and ficolins) evolved.

The fibrinogen-like domains of mussel FREPs conserve the Ca^{+2}-binding sites and likely bind carbohydrates in a calcium-dependent manner, as it has been described before for human ficolins and other FREPs (Matsushita et al., 2001; Kawabata and Tsuda, 2002; Zhang et al., 2009; Garlatti et al., 2010). The four conserved cysteines that are present in the C-terminal end could be involved in inter-chain disulfide bonds, as predicted according to the similar bonds found in human fibrinogen. The cDNA-derived

amino acid sequence for mussel FREP does not suggest the presence of typical transmembrane domains. This characteristic seems to be shared with other proteins of similar structures, such as C1q, collectins or tachylectins, which are soluble proteins that can be secreted into the circulatory system.

The exon-intron organization of FREPs in mussel is quite similar to the structure already described for mammalian ficolins (Runza et al., 2008; Garred et al., 2010). The fibrinogen-like domain is codified in both cases by 4 exons, from exons 5 to 8 in mammals and from exons 3 to 6 in mussel. The first and second exons encode the 5´UTR, the leader peptide and a short link region in both species.

Certain fibrinogen-like domains are involved in the recognition of microorganisms by lectins, including ficolin/P35 and the horseshoe crab lectins (Endo et al., 1996, 2005). This suggests that the mussel fibrinogen-like domain might play an important role in the first line of immune defense. In this sense, the high variability observed within mussels could explain the role of these molecules as a starting point for the activation of the lectin complement pathway (Zhu et al., 2005). Moreover, the high variability of MuFREPs, proposed here as an innate mechanism to fight pathogens (individual mussels do not share any identical sequence) has been described previously for other related invertebrate genes with allo-recognition. This is the case for FREPs of the freshwater snail *Biomphalaria glabrata*, which contain amino terminal immunoglobulin domains (Zhang et al., 2004); the highly polymorphic FuHc gene from *Botryllus schlosseri*; the self-sterility receptors of *Halocynthia* or the vCRL1 gene from *Ciona intestinalis* (Khalturin and Bosch, 2007). Also, we have already reported high individual variability, generated by a mechanism not yet determined, in the antimicrobial peptide myticin C (Costa et al., 2009; Pallavicini et al., 2008) and C1q-containing proteins (Gestal et al., 2010).

We could detect a constitutive mRNA expression of mussel FREP in different tissues (hemocytes, muscle, mantle, gills, male and female gonads, foot and digestive gland). Human ficolin mRNA has been detected by northern-blot techniques in peripheral blood leukocytes (Lu et al., 1996) and, to a lesser extent, in other tissues such as spleen, lung, thymus, placenta and skeletal muscle (Ichijo et al., 1993; Liu et al., 2005). This pattern of expression suggests that ficolins are mainly produced by peripheral blood leukocytes or tissue macrophages. The detection of mussel FREP mRNA in a wide range of tissues (especially in hemocytes) does not completely rule out the possibility that this cell type is the major producer of FREPs.

To analyze the functional activity of purified mussel FREPs, hemolymph samples were loaded onto a GlcNAc-Sepharose 6B column. This methodology has been also applied to the purification of ficolins in other species, such us *Xenopus laevis* and *Halocynthia roretzi*, with good results (Kenjo et al., 2001; Kakinuma et al., 2003). Peptide mass fingerprinting analyses revealed the presence of a ~45 kDa protein in serum identified as "ficolin-2 precursor". Although the molecular weight did not match with the one predicted, probably due to glycosylation, a peptide from the majoritary protein was present in the three groups of MuFREPs. The concentration used in these experiments (2.27 μg/ml) was quite similar to the concentration present in the serum of different vertebrate animals, which ranges from 3.0 to 5.0 μg/ml (Kilpatrick et al., 2003; Le et al., 1998). The incubation of hemocytes with enriched fractions of FREPs induced an increase in the phagocytic capacity. The phagocytic rate recorded in the overall hemocyte populations changed from one mussel to another ranging from 0% up to 60% and reflected heterogeneity in the immune response between individuals. As previously described (García-García et al., 2008), hemocyte groups R1 and R2 presented a high phagocytic capacity and a higher opsonization activity.

The invertebrate immune system is able to recognize molecular patterns (PAMPs) present on pathogen surfaces by means of several receptors (PRRs, Pattern Recognition Receptor) located on the host defense cells (Medzhitov and Janeway, 1997). Among the PRRs the Toll-like receptors (TLRs) are the best-known family of membrane receptors, which recognize several PAMPs. Among them, TLR-4 has been described as the main receptor for LPS, but can also recognize other PAMPs including zymosan and peptidoglycan. No Toll-like receptors have been characterized in *Mytilus galloprovincialis* so far. However, several ESTs with homology to TLRs have been identified in other bivalve molluscs, including the Japanese oyster, *Crassostrea gigas* (Tanguy et al., 2004) and the scallops *Argopencten irradians* (Song et al., 2006) and *Chlamys farreri* (Qiu et al., 2007). Other PRRs, such as LPS-binding proteins or glucan-binding proteins have been detected in *Crassostrea gigas* (González et al., 2007) and in *Mytilus galloprovincialis* (Venier et al., 2003), respectively. The existence of ESTs with homology to different PRRs in bivalves led us to consider that similar recognition mechanisms may occur in these organisms. Indeed, after both *in vitro* and *in vivo* stimulation with several PAMPs, the FREPs expression pattern on mussels seemed to be stimuli-dependent, suggesting that mussel hemocytes were able to discriminate between the different stimuli. LPS and LTA produced the highest FREP gene expression. Live bacteria challenges have also shown an increase in FREP expression. The highest level of FREP expression was found after challenge with the Gram-positive bacteria *M. lysodeikticus*, suggesting that some component of the bacteria would be responsible for this increase over the control. In humans, ficolins can specifically recognize LTA and activate the lectin pathway (Lynch et al., 2004). Thus, the contact of hemocytes with whole bacteria (dead or alive), or with a part of its structure (LPS or LTA), may be enough to up-regulate FREPs expression.

In conclusion, we have described several FREPs in mussels that could be related to mammalian ficolins because a) they cluster together with ficolins in a phylogenetic tree, b) their expression is increased after bacterial infection or PAMPs treatment and c) they have opsonic activities similar to mammalian ficolins. However, the lack of a collagen domain, the low confident phylogenetic position and the surprising lack of ficolins in some animal groups confound the identification of these simple molecules as primitive forms of more evolved ficolins. On the other hand, the high FREP variability observed within and among individuals in *M. galloprovincialis* seems to indicate an extraordinary innate complexity and a potential mechanism to fight pathogens. Also, the high variability described in FREPs and other innate molecules, such as myticin C, could be hypothesized as one of the main reasons why no mortalities have been associated with the Mediterranean mussel when kept at high densities. Further studies are needed to elucidate the function of each group of FREPs and to establish the functional relationship between them. Such studies will help to determine how the high sequence diversity found in bivalve genes with putative immune role, can be a key to pathogen resistance. These aspects would certainly add new insights to the origin and evolution of this important innate immune family of proteins.

Acknowledgements

This work has been funded by the European Union Integrated Project Imaquanim CT-2005-007103 and by the project AGL2008-05111/ACU from the Spanish Ministerio de Ciencia e Innovación. The authors have declared that no competing interests exist.

References

Abascal, F., Zardoya, R., Posada, D., 2005. ProtTest: selection of best-fit models of protein evolution. Bioinformatics 21, 2104-2105.

Adema, C.M., Hertel, L.A., Loker, E.S., 1999. Evidence from two planorbid snails of a complex and dedicated response to digenean (echinostome) infection. Parasitol. 119, 395–404.

Adema, C.M., Hertel, L.A., Miller, R.D., Loker, E.S., 1997. A family of fibrinogen-related proteins that precipitates parasite-derived molecules is produced by an invertebrate after infection. Proc. Natl. Acad. Sci. U.S.A. 94, 8691–8696.

Akaike, H., 1974. A new look at the statistical model identification. IEEE Trans Aut Control 19, 716-723.

Allam, B., Ashton-Alcox, K.A., Ford, S.E., 2001. Haemocyte parameters associated with resistance to brown ring disease in Ruditapes spp. clams. Dev. Comp. Immunol. 25, 365-375.

Baker, N.E., Mlodzik, M., Rubin, G.M., 1990. Spacing differentiation in the developing Drosophila eye: a fibrinogen-related lateral inhibitor encoded by scabrous. Science 250, 1370–1377.

Castresana, J., 2000. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. Mol. Biol. Evol. 17, 540-552.

Costa, M.M., Dios, S., Alonso-Gutierrez, J., Romero, A., Novoa, B., Figueras, A., 2009. Evidence of high individual diversity on myticin C in mussel (*Mytilus galloprovincialis*). Dev. Comp. Immunol. 33, 162-170.

Dixit, R., Roy, U., Patole, M.S., Shouche, Y.S., 2008. Molecular and phylogenetic analysis of a novel family of fibrinogen-related proteins from mosquito *Aedes albopictus* cell line. Comput. Biol. Chem. 32, 382–386.

Doolittle, R.F., 1992. A detailed consideration of a principal domain of vertebrate fibrinogen and its relatives. Protein Sci. 1, 1563–1577.

Drickamer, K., 1995. Increasing diversity of animal lectin structures. Curr. Opin. Struct. Biol. 5, 612–616.

Endo, Y., Liu, Y., Fujita, T., 2006. Structure and function of ficolins. Adv. Exp. Med. Biol. 586, 265-279.

Endo, Y., Matsushita, M.,Fujita, T., 2007. Role of ficolin in innate immunity and its molecular basis. Immunobiol. 212, 371-379.

Endo, Y., Nakazawa, N., Liu, Y., Iwaki, D., Takahashi, M., Fujita, T., Nakata, M., Matsushita, M., 2005. Carbohydrate-bindingspecifities of mouse ficolin A, a splicing variant of ficolin A and ficolin B and their complex formation with MASP-2 and aMAP. Immunogenetics 57, 837-844.

Endo, Y., Sato, Y., Matsushita, M., Fujita, T., 1996. Cloning and characterization of the human lectin P35 gene and its related gene. Genomics 36, 515–521.

Endo, Y., Takahashi, M., Fujita, T., 2006. Lectin complement system and pattern recognition. Immunobiol. 211, 283-293.

Felsenstein, J., 1985. Confidence limits on phylogenies: an approach using the bootstrap. Evolution 39, 783-791.

Fujimori, Y., Arrumilla, S., Fukumoto, Y., Miura, Y., Yagasaki, Y., Tachikawa, H., Fujimoto, D., 1998. Molecular cloning and characterization of mouse ficolin-A. Biochem. Biophys. Res. Commun. 244, 796–800.

García-García, E., Prado-Álvarez, M., Novoa, B., Figueras, A., Rosales, C., 2008. Immune responses of mussel hemocyte subpopulations are differentially regulated by enzymes of the PI 3-K, PKC, and ERK kinase families. Dev. Comparative Immunol. 32, 637-653.

Garlatti, V., Martin, L., Lacroix, M., Gout, E., Arlaud, G., Thielens, N., Gaboriaud, C., 2010. Structural insights into the recognition properties of human ficolins. J. Innate Immun. 2, 17-23.

Garred, P., Honoré, C., Ma, Y.J., Rorvig, S., Cowland, J., Borregaard, N., Hummelshoj, T., 2010. The genetics of ficolins. J. Innate Immun. 2, 3-16.

Gestal, C., Pallavicini, A., Venier, P., Novoa, B., Figueras, A., 2010. MgC1q, a novel C1q-domain-containing protein involved in the immune response of *Mytilus galloprovincialis*. Dev. Comp. Immunol. 34, 926-934.

Gokudna, S., Muta, T., Tsuda, R., Koori, K., Kawahara, T., Seki, N., Mizunoe, Y., Wai, S.N., Iwanaga, S., Kawabata S., 1999. Horseshoe crab acetyl group-recognizing lectins involved in innate immunity are structurally related to fibrinogen. Proc. Nat.l Acad. Sci. USA. 96, 10086-10091.

González, M., Gueguen, Y., Desserre, G., de Lorgeril, J., Romestand, B., Bachère, E., 2007. Molecular characterization of two isoforms of defensin from hemocytes of the oyster *Crassostrea gigas*. Dev. Comp. Immunol. 31, 332-339.

Gorbushin, A.M., Panchin, Y.V., Iakoleva, N.V., 2010. In search of the origin of FREPs: characterization of *Aplysia californica* fibrinogen-related proteins. Dev. Comp. Immunol. 34, 465-73.

Guda, C., 2006. pTARGET: A web server for predicting protein subcellular localization. Nucleic Acids Res. 35, 210-213.

Guindon, S., Dufayard, J. F., Lefort, V., Anisimova, M., Hordijk, W., Gascuel, O., 2010. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. Syst. Biol. 59, 307-321.

Hertel, L.A., Adema, C.M., Loker, E.S., 2005. 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. 29, 295–303.

Hugh, D., Owen, R., Tracy, A., 1992. Southern Gill dimensions, water pumping rate and body size in the mussel *Mytilus edulis* L. J. Exp. Ma.r Biol. Ecol. 155, 213-237.

Ichijo, H., Hellman, U., Wernstedt, C., Gonez, J.L., Claesson-Welsh, L., Heldin, C.H., Miyazono K., 1993. Molecular cloning and characterization of ficolin, a multimeric protein with fibrinogen- and collagen-like domains. J. Biol. Chem. 268, 14505–14513.

Ju, L., Zhang, S., Liang, Y., Sun, X., 2009. Identification, expression and antibacterial activity of a tachylectin-related homolog in amphioxus *Branchiostoma belcheri* with implications for involvement of the digestive system in acute phase response. Fish Shellfish Immunol. 26, 235-242.

Kakinuma, Y., Endo, Y., Takahashi, M., Nakata, M., Matsushita, M., Takenoshita, M. Fujita, T., 2003. Molecular cloning and characterization of novel ficolins from *Xenopus laevis*. Immunogenetics 55, 29–37.

Katoh, K., Kuma, K., Toh, H., Miyata, T., 2005. MAFFT version 5: improvement in accuracy of multiple sequence alignment. Nucleic Acids Res. 33, 511-518.

Kawabata, S., Tsuda, R., 2002. Molecular basis of non-self recognition by the horseshoe crab tachylectins. Biochim. Biophys. Acta 1572, 414-421.

Kenjo, A., Takahashi, M., Matsushita, M., Endo, Y., Nakata, M., Mizuochi, T., Fujita, T., 2001. Cloning and characterization of novel ficolins from the solitary ascidian, *Halocynthia roretzi*. J. Biol. Chem. 276, 19959–19965.

Khalturin, K., Bosch, T.C.G., 2007. Self/nonself discrimination at the basis of chordate evolution: limits on molecular conservation. Current Op. Immunol. 19, 4-9.

Kilpatrick, D.C., McLintock, L.A., Allan, E.K., Copland, M., Fujita, T., Jordanides, N.E., Koch, C., Matsushita, M., Shiraki, H., Stewart, K., Tsujimura, M., Turner, M.L., Franklin, I.M., Holyoake, T.L., 2003. No strong relationship between mannan binding lectin or plasma ficolins and chemotherapy-related infections. Clin. Exp. Immunol. 134, 279-284.

Knibbs, R.N., Osborne, S.E., Glick, G.D., Goldstein, I.J., 1993. Binding determinants of the sialic acid-specific lectin from the slug *Limax flavus*. J. Biol. Chem. 268, 18524–18531.

Le, Y., Lee, S.H., Kon, O.L., Lu, J., 1998. Human L-ficolin: plasma levels, sugar specifity and assignment of its lectin activity to the fibrinogen-like (FBG) domain. FEBS Lett. 425, 367-370.

Liu, Y., Endo, Y., Homma, S., Kanno, K., Yaginuma, H., Fujita, T., 2005. Ficolin A and Ficolin B are expressed in distinct ontogenic patterns and cell types in the mouse. Mol. Immunol. 42, 1265-1273.

Liu, Z., Ukomadu, C., 2008. Fibrinogen-like protein 1, a hepatocyte derived protein is an acute phase reactant. Biochem. Biophys. Res. Commun. 365, 729–734.

Livak, K., Schmittgen, T.D., 2001. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) method. Methods 25, 402–408.

Lu, J., Le, Y., 1998. Ficolins and the fibrinogen-like domain. Immunobiol. 199,190–199.

Lu, J., Tay, P.N., Kon, O.L., Reid, K.B., 1996. Human ficolin: cDNA cloning, demonstration of peripheral blood leucocytes as the major site of synthesis and assignment of the gene to chromosome 9. Biochem. J. 313, 473-478.

Lynch, N.J., Roscher, S., Hartung, T., Morath, S., Matsushita, M., Maennel, D.N., Kuraya, M., Fujita, T., Schwaeble, W.J., 2004. L-ficolin specifically binds to lipoteichoic acid, a cell wall constituent of Gram-positive bacteria, and activates the lectin pathway of complement. J. Immunol. 172, 1198-1202.

Mali, B., Soza-Ried, J., Frohme, M., Frank, U., 2006. Structural but not functional conservation of an immune molecule: a tachylectin-like gene in Hydractinia. Dev. Comparative Immunol. 30, 275-281.

Maniatis, T., Fritsch, E.F., Sambrook, J., 1982. Molecular Cloning. A Laboratory Manual. Cold Spring Harbor Laboratory Publication, New York.

Matsushita, M., Endo, Hamasaki, N., Fujita, T., 2001. Activation of the lectin complement pathway by ficolins. Int. Immunopharmacol. 1, 359–363.

Matsushita, M., Endo, Y., Fujita, T., 2000. Cutting edge: complement-activating complex of ficolin and mannose-binding lectin-associated serine protease. J. Immunol. 164, 2281-2284.

Matsushita, M., Endo, Y., Taira, S., Sato, Y., Fujita, T., Ichikawa, N., Nakata, N., Mizuochi, T., 1996. A novel human serum lectin with collagen- and fibrinogen-like domains that functions as an opsonin. J. Biol. Chem. 271, 2448–2454.

Medzhitov, R., Janeway, C.A., 1997. Innate immunity: the virtues of a non-clonal system of recognition. Cell 91, 295-298.

Meyhöfer, E., 1985. Comparative pumping rates in suspension-feeding bivalves. Marine Biol. 85, 137-142.

Norman, R., 1988. Hydrodynamics of flow in Mytilus gills. J. Exp. Mar. Bio. Ecol. 120, 171-182.

Ohashi, T., Erickson, H.P., 1998. Oligomeric structure and tissue distribution of ficolins from mouse, pig and human. Arch. Biochem. Biophys. 360, 223-232.

Omori-Satoh, T., Yamakawa, Y., Mebs, D., 2000. The antihemorrhagic factor, erinacin, from the European hedgehog (*Erinaceus europaeus*), a metalloprotease inhibitor of large molecular size possessing ficolin/opsonin P35 lectin domains. Toxicon 38, 1561–1580.

Pallavicini, A., Costa, M.M., Gestal, C., Dreos, R., Figueras, A., Venier, P., Novoa, B., 2008. High sequence variability of myticin transcripts in haemocytes of immune-stimulated mussels suggests ancient host-pathogen interactions. Dev. Comp. Immunol. 32, 213-226.

Pipe, R. 1990. Hydrolytic enzymes associated with the granular haemocytes of the marine mussel *Mytilus edulis*. Histochem. J. 22, 595-603.

Pipe, R., Farley, S., Coles, J., 1997. The separation and characterisation of haemocytes from the mussel *Mytilus edulis*. Cell Tissue Res. 289, 537-545.

Qiu, L., Song, L., Xu, W., Ni, D., Yu, Y., 2007. Molecular cloning and expression of a Toll receptor gene homologue from Zhikong Scallop, *Chlamys farreri*. Fish Shellfish Immunol. 22, 451-466.

Rego, R.O.M., Hajdusek, O., Kovar, V., Kopacek, P., Grubhoffer, L., Hypsa, V., 2005. Molecular cloning and comparative analysis of fibrinogen-related proteins from the soft tick *Ornithodoros moubata* and the hard tick *Ixodes ricinus*. Insect Biochem. Mol. Biol. 35, 991–1004.

Richards, E.H., Renwrantz, L.R., 1991. Two lectins on the surface of *Helix pomatia* haemocytes: a Ca^{2+}-dependent, GalNac-specific lectin and a Ca^{2+}-independent, mannose 6-phosphate-specific lectin which recognises activated homologous opsonins. J. Comp. Physiol. 161, 43–54.

Runza, V.L., Schwaeble, W., Männel, D.N., 2008. Ficolins: Novel pattern recognition molecules of the innate immune response. Immunobiol. 231, 297-306.

Shalev, I., Liu, H., Koscik, C., Bartczak, A., Javadi, M., Wong, K.M., Maknojia, A., He, W., Liu, M.F., Diao, J., Winter, E., Manuel, J., McCarthy, D., Cattral, M., Gommerman, J., Clark, D.A., Phillips, M.J., Gorczynski, R.R., Zhang, L., Downey, G., Grant, D., Cybulsky, M.I., Levy, G., 2008. Targeted deletion of fgl2 leads to impaired regulatory T cell activity and development of autoimmune glomerulonephritis. J. Immunol. 180, 249–260.

Song, L., Xu, W., Li, C., Li, H., Wu, L., Xiang, J., Guo, X., 2006. Development of expressed sequence tags from the bay scallop *Argopecten irradians irradians*. Mar. Biotechnol. 8, 161-169.

Stamatakis, A., 2006. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics. 22, 2688-2690.

Tanguy, A., Guo, X., Ford, S.E., 2004. Discovery of genes expressed in response to *Perkinsus marinus* challenge in Eastern (*Crassostrea virginica*) and Pacific (*C. gigas*) oysters. Gene 338, 121-131.

Tanio, M., Kondo, S., Sugio, S., Kohno, T., 2007. Trivalent recognition unit of innate immunity system: crystal structure of trimeric human M-ficolin fibrinogen-like domain. J. Biol. Chem. 282, 3889-3895.

Turner, M.W., 1996. Mannose-binding lectin: the pluripotent molecule of the innate immune system. Immunol. Today 17, 532-540.

Venier, P., De Pittà, C., Bernante, F., Varotto, L., De Nardi, B., Bovo, G., Roch, P., Novoa, B., Figueras, A., Pallavicini, A., Lanfranchi, G., 2009. MytiBase: a knowledgebase of mussel (*M. galloprovincialis*) transcribed sequences. BMC Genomics 10,72.

Venier, P., Pallavicini, A., De Nardi, B., Lanfranchi, G., 2003. Towards a catalogue of genes transcribed in multiples tissues of *Mytilus galloprovincialis*. Gene 314, 29-40.

Wang, X., Rocheleau, T.A., Fuchs, J.F., Hillyer, J.F., Chen, C., Christensen, B.M., 2004. 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. 13, 273–282.

Whelan, S., Goldman, N., 2001. A general empirical model of protein evolution derived from multiple protein families using a maximum-likelihood approach. Mol. Biol. Evol. 18, 691-699.

Xu, X., Doolittle, R.F., 1990. Presence of a vertebrate fibrinogen-like sequence in an echinoderm. Proc. Natl. Acad. Sci. U.S.A. 87, 2097–2101.

Yang, Z., 1996. Among-site rate variation and its impact on phylogenetic analysis. Trends Ecol. Evol. 11, 367-372.

Zhang, H., Wang, L., Song, L., Song, X., Wang, B., Mu, C., Zhang, Y., 2009. A fibrinogen-related protein from bay scallop *Argopecten irradians* involved in innate immunity as pattern recognition receptor. Fish Shellfish Immunol. 26, 56-64.

Zhang, S., Adema, C.M., Kepler, T.B., Loker, E.S., 2004. Diversification of Ig superfamily genes in an invertebrate. Science 305, 251–254.

Zhang, S., Leonard, P.M., Adema, C.M., Loker, E.S., 2001. Parasite-responsive IgSF members in the snail *Biomphalaria glabrata*: characterization of novel genes with tandemly arranged IgSF domains and a fibrinogen domain. Immunogenetics 53, 684–694.

Zhang, S.M., Loker, E.S., 2003. The FREP gene family in the snail *Biomphalaria glabrata*: additional members, and evidence consistent with alternative splicing and FREP retrosequences. Fibrinogen-related proteins. Dev. Comp. Immunol. 27, 175–187.

Zhang, S.M., Zeng, Y., Loker, E.S., 2008. Expression profiling and binding properties of fibrinogen-related proteins (FREPs), plasma proteins from the schistosome snail host *Biomphalaria glabrata.* Innate Immun. 14, 175–189.

Zhang, X., Ali, M., 2008. Ficolins: Structure, function and associated diseases, in Lambris, J.D. (Eds.), Current topics in complement II. Springer Science, New York.

Zhu, Y., Thangamani, S., Ho, B., Ding, J.L., 2005. The ancient origin of the complement system. EMBO J. 24, 382-394.

Figure legends

Figure 1. (A) Maximum likelihood phylogenetic tree for FREPs. Numbers on branches are bootstrap percentages. FIB: Fibrinogen. FIBL: Fibrinogen-like. ANGP: Angiopoietin. TL: Tachylectin. Fico: Ficolin. TENA: Tenascin. G1, G2 and G3 correspond to the names of the three FREPs found in mussel. (**B)** Domain structure of ficolin proteins and fibrinogen-containing proteins. The signal peptide and the low complexity regions are marked in red and pink, respectively. FBG: fibrinogen related domain. Small black box: Pfam collagen domain. Big black box: Pfam fib-alpha domain. IG: immunoglobulin domain. EGF: epidermal growth factor domain. FN3: fibronectin type 3 domain. G1, G2 and G3 correspond to the names of the three FREPs found in mussel.

Figure 2. Alignment of amino acid sequences of mussel FREPs and ficolins from *Halocynthia roretzi, Xenopus laevis, Rattus norvegicus, Sus scrofa* and *Homo sapiens*. The putative signal sequence is underlined and the cleavage sites are designated in bold and italics. The N-terminal region is highlighted in green and the conserved cysteine residues are highlighted in pink. The collagen-like domain, characterized by Gly-X-Y triplets, is highlighted in dark and light blue. The neck domain (only present in mammalian ficolins) is highlighted in yellow. Potential N-glycosylation sites are highlighted in red and the Ca^{+2}-binding sites are designated by two bold sequences in the COOH-terminal. The ficolin-2 precursor (QDGSVDFFR) purified from mussel hemolymph by affinity chromatography is highlighted in red.

Figure 3. Quantitative PCR of FREPs groups in different tissues from mussels. F gonad: female gonad, M gonad: male gonad. The results are presented as the mean ± SD. Bars represent the relative expression of FREPs normalized to β-Actin expression levels.

Figure 4. Exon-intron organization of mussel FREPs from group 1.

Figure 5. Sequence variability of FREPs mussel transcripts in group 1. **(A)** Numbers in bold indicate variable positions in 3 mussels. Numbers in italics indicate variable positions in 2 mussels and the asterisk (*) indicates the variable position in only 1 mussel. Transversal short bars represent point mutations in each mussel. Transition and transversion mutations are noted with letters. **(B)** Diagram summarizing the number of clones and the number of different nucleotide and amino acid sequences analyzed. Genetic identity was calculated according to the number of variable nucleotides between individuals.

Figure 6. Functional activity of purified mussel FREPs by phagocytosis assay. **(A)** Acrylamide:bisacrylamide gel stained with Coomassie blue showing the predominant protein of 45 kDa. (1) Molecular weight marker. (2) Hemolymph before passing though the column. (3) Hemolymph after passing though the column. (4) Washed unbound proteins. (5) Fraction of bound proteins. **(B)** Percent increase in phagocytosis (Y axis)

of 16 individual mussels. **(C)** Representative result of the phagocytosis experiments obtained with hemocytes isolated from mussel number 3 and treated with purified FREPs. The X axis represents the fluorescence intensity emitted by FITC-conjugated beads, as measured by flow cytometry. The Y axis represents the relative number of fluorescent events (phagocytic hemocytes) after treatment with purified FREPs. **(D)** Phagocytic activity of different populations of hemocytes isolated from the mussel number 13 after treatment with purified FREPs.

Figure 7. Quantitative expression of FREPs in hemocyte samples of mussels *M. galloprovincialis* after an *in vitro* (A) or *in vivo* (B) stimulation with PAMPs and following an *in vitro* (C) and *in vivo* (D) stimulation with bacteria during 1, 3 and 6 h or 3, 6 and 24 h post-inoculation, respectively. The results are presented as the mean ± SD. Lines represent the relative expression of FREPs as compared to controls, previously normalized to endogenous 18S expression levels. Significant values are represented with an asterisk (*) ($p<0.05$).

Table 1

Table I: Sequences of primers designed to analyze the genomic structure of FREPs from group 1.

Name	Sequence	Name	Sequence
FREPG1-Gen-For1	ACAAACGAAAAGGACAAATG	FREPG1-Gen-Rev2	GAGTGCAAATTATCGTTGCCTA
FREPG1-Gen-For2	TGTTTGTCATTTGTCCTTTTCG	FREPG1-Gen-Rev3	CGTTTGATTGACACGAGGTA
FREPG1-Gen-For3	TGTTTCTTGACGTCACTTCG	FREPG1-Gen-Rev4	AAGTTAAAACCCTCGGGAAG
FREPG1-Gen-For4	TCCGTTAGGTCCAGTACATCC	FREPG1-Gen-Rev5	CACCAGGCTCCGAGATAGTT
FREPG1-Gen-For5	ACTTTGTGCAAAATGGTCCAG	FREPG1-Gen-Rev6	CCTTTCCAAGTACGCCAGTA
FREPG1-Gen-For6	CAAGTTTTCCCCAATCCACA	FREPG1-Gen-Rev7	CGTCCAGTGTCCTTCATCTA
FREPG1-Gen-For7	CATTGGCACTCACACAACATC	FREPG1-Gen-Rev8	TGTTGTGTGAGTGCCAATGA
FREPG1-Gen-For8	GAGAAGAAGCATGGACACAGG	FREPG1-Gen-Rev9	GTGCCAATGAGACAACTCTCC
FREPG1-Gen-For9	GTGTTTGCAGACGAAACGTG	FREPG1-Gen-Rev10	TGTTTTCTCGTCAAGCACCA
FREPG1-Gen-Rev1	CTGAACTTGTTTCCGTTAGG	FREPG1-Gen-Rev11	AAATCATCATTTCCCGAGGAT

Figure 1

Click here to download high resolution image

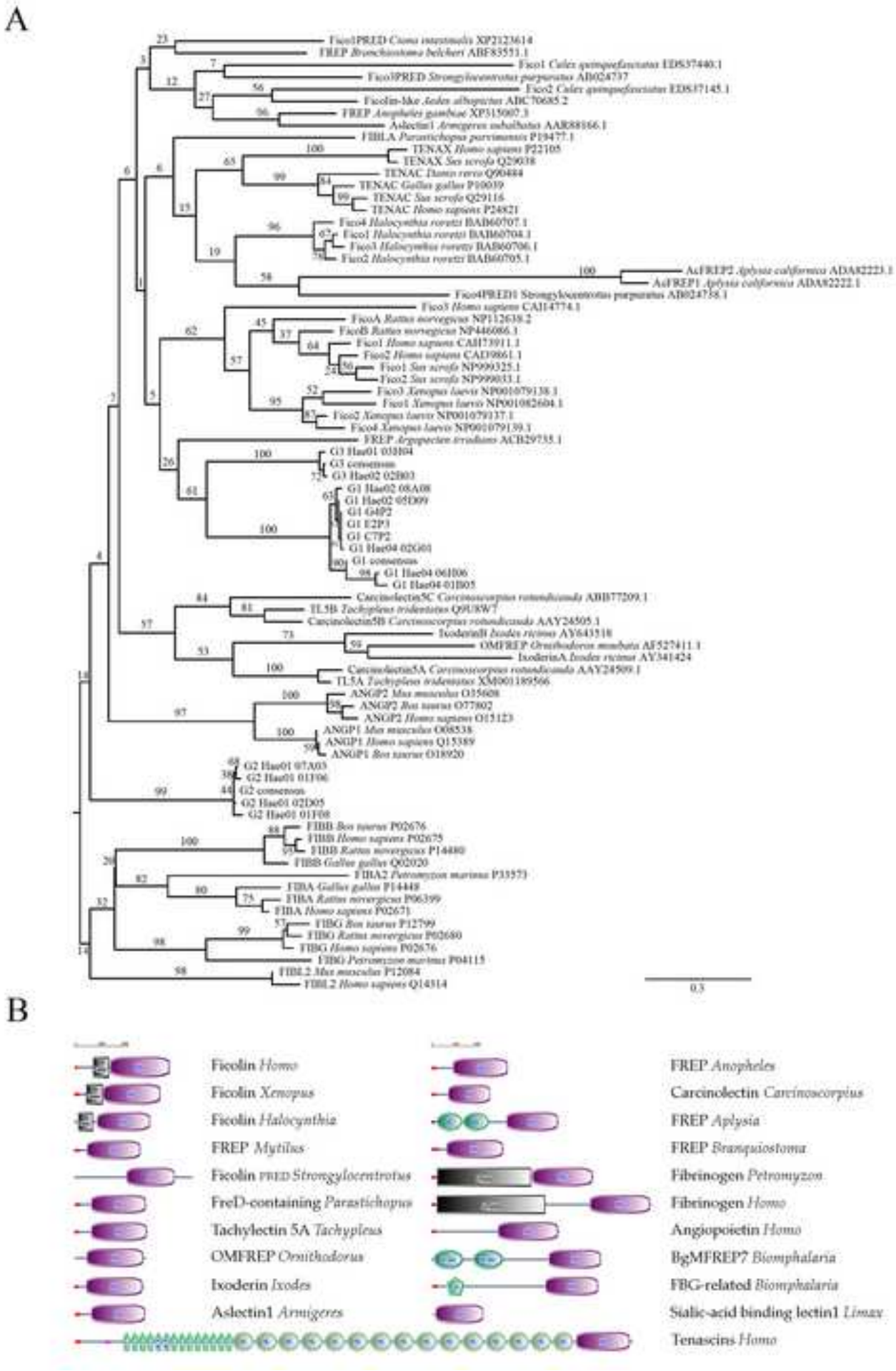
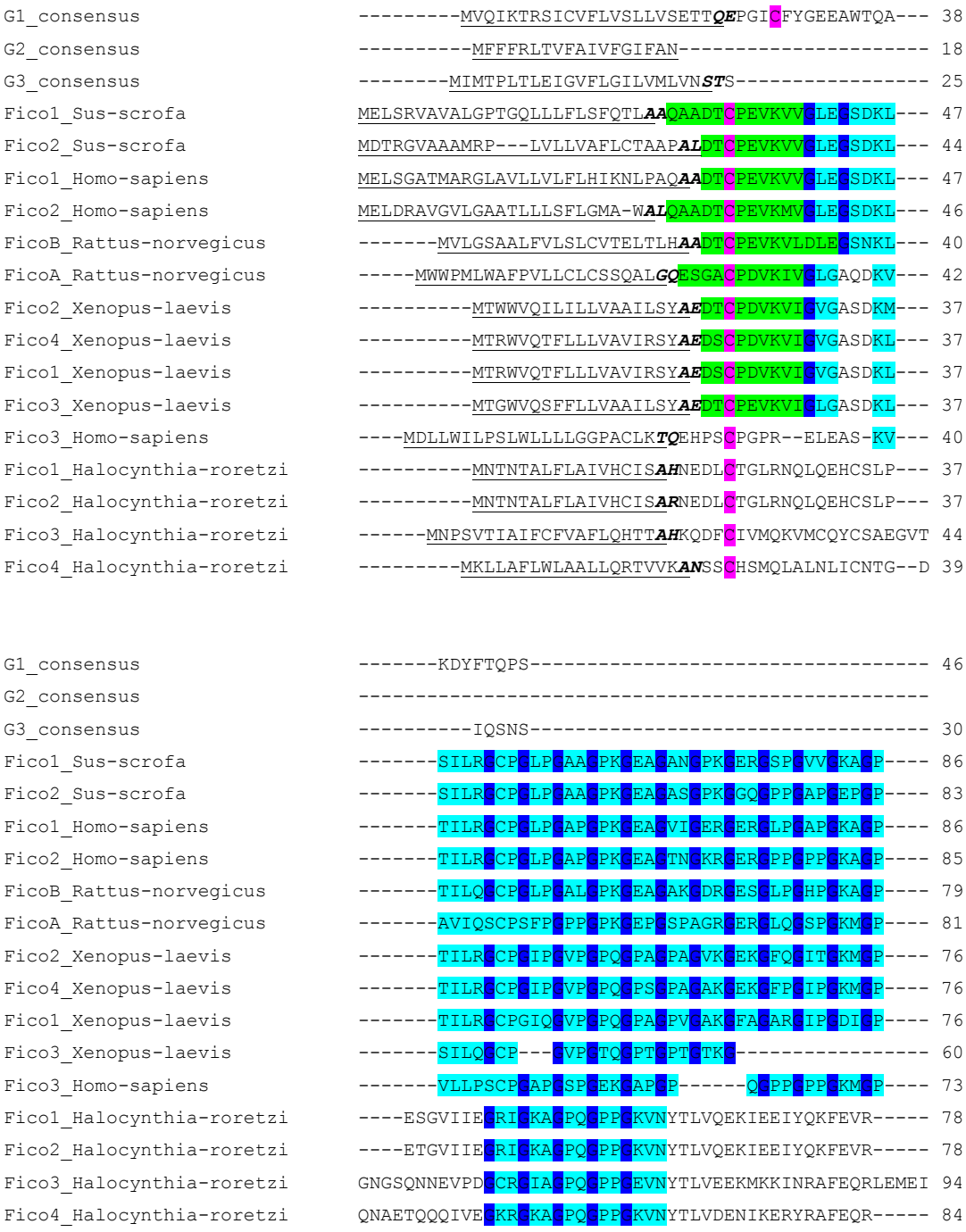


Figure 2


G1_consensus ---------MVQIKTRSICVFLVSLLVSETTQEPGICFYGEEAWTQA--- 38
G2_consensus ----------MFFFRLTVFAIVFGIFAN---------- 18
G3_consensus --------MIMTPLTLEIGVFLGILVMLVNSTS---------- 25
Fico1_Sus-scrofa MELSRVAVALGPTGQLLLFLSFQTLAAQAADTCPEVKVVGLEGSDKL--- 47
Fico2_Sus-scrofa MDTRGVAAAMRP---LVLLVAFLCTAAPALDTCPEVKVVGLEGSDKL--- 44
Fico1_Homo-sapiens MELSGATMARGLAVLLVLFLHIKNLPAQAADTCPEVKVVGLEGSDKL--- 47
Fico2_Homo-sapiens MELDRAVGVLGAATLLLSFLGMA-WALQAADTCPEVKMVGLEGSDKL--- 46
FicoB_Rattus-norvegicus -------MVLGSAALFVLSLCVTELTLHAADTCPEVKVLDLEGSNKL--- 40
FicoA_Rattus-norvegicus -----MWWPMLWAFPVLLCLCSSQALGQESGACPDVKIVGLGAQDKV--- 42
Fico2_Xenopus-laevis ----------MTWWVQILILLVAAILSYAEDTCPDVKVIGVGASDKM--- 37
Fico4_Xenopus-laevis ----------MTRWVQTFLLLVAVIRSYAEDSCPDVKVIGVGASDKL--- 37
Fico1_Xenopus-laevis ----------MTRWVQTFLLLVAVIRSYAEDSCPDVKVIGVGASDKL--- 37
Fico3_Xenopus-laevis ----------MTGWVQSFFLLVAAILSYAEDTCPEVKVIGLGASDKL--- 37
Fico3_Homo-sapiens ----MDLLWILPSLWLLLLGGPACLKTQEHPSCPGPR--ELEAS-KV--- 40
Fico1_Halocynthia-roretzi ----------MNTNTALFLAIVHCISAHNEDLCTGLRNQLQEHCSLP--- 37
Fico2_Halocynthia-roretzi ----------MNTNTALFLAIVHCISARNEDLCTGLRNQLQEHCSLP--- 37
Fico3_Halocynthia-roretzi ------MNPSVTIAIFCFVAFLQHTTAHKQDFCIVMQKVMCQYCSAEGVT 44
Fico4_Halocynthia-roretzi ---------MKLLAFLWLAALLQRTVVKANSSCHSMQLALNLICNTG--D 39
G1_consensus -------KDYFTQPS---------- 46
G2_consensus ----------
G3_consensus ----------IQSNS---------- 30
Fico1_Sus-scrofa -------SILRCCPGLPGAAGPKGEACANGPKGERGSPGVVGKAGP---- 86
Fico2_Sus-scrofa -------SILRCCPGLPGAAGPKGEACASGPKGGQGPPGAPGEPGP---- 83
Fico1_Homo-sapiens -------TILRCCPGLPGAPGPKGEACVIGERGERGLPGAPGKAGP---- 86
Fico2_Homo-sapiens -------TILRCCPGLPGAPGPKGEACTNGKRGERGPPGPPGKAGP---- 85
FicoB_Rattus-norvegicus -------TILQCCPGLPGALGPKGEACAKGDRGESGLPGHPGKAGP---- 79
FicoA_Rattus-norvegicus -------AVIQSCPSFPGPPGPKGEPGSPAGRGERGLQGSPGKMGP---- 81
Fico2_Xenopus-laevis -------TILRCCPGIPGVPGPQGPACPAGVKGEKGFQGITGKMGP---- 76
Fico4_Xenopus-laevis -------TILRCCPGIPGVPGPQGPSGPAGAKGEKGFPGIPGKMGP---- 76
Fico1_Xenopus-laevis -------TILRCCPGIQGVPGPQGPACPVGAKGFAGARGIPGDIGP---- 76
Fico3_Xenopus-laevis -------SILQCCP---GVPGTQGPTGPTGTKG---------- 60
Fico3_Homo-sapiens -------VLLPSCPGAPGSPGEKGAPGP------QGPPGPPGKMGP---- 73
Fico1_Halocynthia-roretzi ----ESGVIIEGRIGKAGPQGPPGKVNYTLVQEKIEEIYQKFEVR----- 78
Fico2_Halocynthia-roretzi ----ETGVIIEGRIGKAGPQGPPGKVNYTLVQEKIEEIYQKFEVR----- 78
Fico3_Halocynthia-roretzi GNGSQNNEVPDGCRGIAGPQGPPGEVNYTLVEEKMKKINRAFEQRLEMEI 94
Fico4_Halocynthia-roretzi QNAETQQQIVEGKRGKAGPQGPPGKVNYTLVDENIKERYRAFEQR----- 84

```
G1_consensus                  ----------------------------------------TCHEGRPID 55
G2_consensus                  ----------------------------------------AELPRE 24
G3_consensus                  ----------------------------------------GSYTRIPID 39
Fico1_Sus-scrofa              ------AGPKGDRGEKGARGEKGEPGQ------------LQSCATGPRT 117
Fico2_Sus-scrofa              ------PGPKGDRGEKGEPGPKGESWE------------TEQCLTGPRT 114
Fico1_Homo-sapiens            ------VGPKGDRGEKGMRGEKGDAGQ------------SQSCATGPRN 117
Fico2_Homo-sapiens            ------PGPNGAPGE-----------------------PQPCLTGPRT 104
FicoB_Rattus-norvegicus       ------TGPKGDRGEKGVRGEKGDTGP------------SQSCATGPRT 110
FicoA_Rattus-norvegicus       ------AGSKGEPGTMGPPGVKGEKGERGTASPLGQKELGDALCRRGPRS 125
Fico2_Xenopus-laevis          ------AGLKGERGISGPKGQKGDKGDP------------GIPAAGTAQN 108
Fico4_Xenopus-laevis          ------TGLKGERGISGPKGQKGDKGDP------------GIPVVGMAQN 108
Fico1_Xenopus-laevis          ------TGLKGEQGYPGARGLKGEKGDP------------GVPVPGTAQN 108
Fico3_Xenopus-laevis          ----------------------DKCDP------------GVPIPGTAKN 75
Fico3_Homo-sapiens            ------KGEPGDPVN-------------------------LLRCQEGPRN 92
Fico1_Halocynthia-roretzi     ------MDNRVDQKTETCSSQIKLLE-----------KRFNSLLTGCEKV 111
Fico2_Halocynthia-roretzi     ------MDNRVDQKTETCSSQIKLLE-----------KRFNSLLTGCEKV 111
Fico3_Halocynthia-roretzi     EKKFKIFSIKSERQIEMHSTEIKLLENKITELESRWHKRINS--TGCEKV 142
Fico4_Halocynthia-roretzi     ------FSLESERQIGMHSTEIKLLKNKITELESRWQKRFNSLLTGCEKV 128

G1_consensus                  CNDIPDK--CKSGVYKVFPK-QTQGFDVYCKMNLD--EGHWTVFQKRENG 100
G2_consensus                  CAELAITSCGVYKIYPFAKL-QP-GVSVYCKIDTS--GHIWTVIQQRFDG 70
G3_consensus                  CGDIDIK--RGSGVYMIYPTGSFDGFNVYCNMKVDNVGGGWTVFQRRLNG 87
Fico1_Sus-scrofa              CKELLTRGHFLSGWHTIYLP-DCQPLTVLCDMDTD--GGGWTVFQRRSDG 164
Fico2_Sus-scrofa              CKELLTRGHILSGWHTIYLP-DCQPLTVLCDMDTD--GGGWTVFQRRSDG 161
Fico1_Homo-sapiens            CKDLLDRGYFLSGWHTIYLP-DCRPLTVLCDMDTD--GGGWTVFQRRMDG 164
Fico2_Homo-sapiens            CKDLLDRGHFLSGWHTIYLP-DCRPLTVLCDMDTD--GGGWTVFQRRVDG 151
FicoB_Rattus-norvegicus       CKELLTRGYFLTGWYTIYLP-DCRPLTVLCDMDTD--GGGWTVFQRRIDG 157
FicoA_Rattus-norvegicus       CKDLLTRGIFLTGWYTIYLP-DCRPLTVLCDMDVD--GGGWTVFQRRVDG 172
Fico2_Xenopus-laevis          CKEWLDQGASISGWYTIYTP-NGLPLSVLCDMETD--GGGWIVFQRRMDG 155
Fico4_Xenopus-laevis          CKEWLDQGASISGWYTIYTT-NGLSLTVLCDMETD--GGGWIVFQRRMDG 155
Fico1_Xenopus-laevis          CKEWLDQGVTISGWYTIYTP-NGLTLSVLCDMETD--GGGWIVFQRRADG 155
Fico3_Xenopus-laevis          CKDWLDQGASITGWYTIYTS-TGRRLRVLCDMETD--GGGWTVFQRRSDG 122
Fico3_Homo-sapiens            CRELLSQGATLSGWYHLCLP-EGRALPVFCDMDTE--GGGWLVFQRRQDG 139
Fico1_Halocynthia-roretzi     SKYGALSWNGTGGIFNIYPDNPQQSIEVYCDLTSD--GGGWTVFQRRMDG 159
Fico2_Halocynthia-roretzi     SKYGALSWNGTGGIFNIYPDNPQQSIEVYCDLTSG--GGGWTVFQRRMDG 159
Fico3_Halocynthia-roretzi     SKYGAISWKGTGGIFNIYPDNPQESIEVYCDLTSD--GGGWTVFQRRMDG 190
Fico4_Halocynthia-roretzi     SKYGALSWNGTGGIFNIYPDNPQESIEVYCDLTSD--GGGWIVFQRRMDG 176
                              .             . .        . * *.:        * *:*:* :*
```

```
G1_consensus                YVDFYRGWNDYKSGFGNPKHEFWLGNENLHPLTSQHNYEMRIDLTDFEGN 150
G2_consensus                SVNFFRKWQNYKTGFGQPFGEYWLGNDVIHELTTGANHALRIEVEDFNGT 120
G3_consensus                AVGFYRGWDDYKAGFGTLEEEHWLGNENLHILTSQAEYQLLITLQDFANH 137
Fico1_Sus-scrofa            SVDFYRDWAAYKRGFGSQLGEFWLGNDHIHALTAQGTSELRVDLVDFEGN 214
Fico2_Sus-scrofa            SVDFYRDWAAYKRGFGSQLGEFWLGNDHIHALTAQGTNELRVDLVDFEGN 211
Fico1_Homo-sapiens          SVDFYRDWAAYKQGFGSQLGEFWLGNDNIHALTAQGSSELRVDLVDFEGN 214
Fico2_Homo-sapiens          SVDFYRDWATYKQGFGSRLGEFWLGNDNIHALTAQGTSELRVDLVDFEDN 201
FicoB_Rattus-norvegicus     TVDFFRDWTSYKQGFGSQLGEFWLGNDNIHALTTQGTNELRVDLADFDGN 207
FicoA_Rattus-norvegicus     SINFYRDWDSYKRGFGNLGTEFWLGNDYLHLLTANGNQELRVDLREFQGQ 222
Fico2_Xenopus-laevis        SVDFFRDWNSYKKGFGRQDSEFWLGNDNLHLLTATGNFQLRVDLTDFDKN 205
Fico4_Xenopus-laevis        SVDFFQDWISYKRGFGRQDSEFWLGNNNLHLLTVTGSFQLRVDLTDFGNN 205
Fico1_Xenopus-laevis        SVDFNRDWNSYKRGFGRKDSEFWLGNDNLHLLTATGNFQLRVDLTDFSDK 205
Fico3_Xenopus-laevis        SVDFFRDWDSYKRGFGLQQSEFWLGNENIHLLTSTGYFQLRIDLTDFEKK 172
Fico3_Homo-sapiens          SVDFFRSWSSYRAGFGNQESEFWLGNENLHQLTLQGNWELRVELEDFNGN 189
Fico1_Halocynthia-roretzi   SVDFYRGWNEYVNGFGEKNKEFWLGLETIHQLTKNGNYELRVDIGNWEGE 209
Fico2_Halocynthia-roretzi   SVDFYRGWDEYVNGFGEKDKEFWLGLETIHQLTKNGNYELRVDIGNWEGE 209
Fico3_Halocynthia-roretzi   SVDFYRGWNEYVNGFGEKDKEFWLGLETIHQLTKNGSYELRVDIGDWEGE 240
Fico4_Halocynthia-roretzi   SVDFYRGWNEYVNGFGENDKEFWLGLETIHQLTKNGNYELRVDIGDWEGE 226
                             :.* : *  *  ***    *.*** : :* **      : : : ::

G1_consensus                TAFAKYKVFAIGDESSKFKLTANGYHG-TAGNSIEHHNGHRFSTKDRDND 199
G2_consensus                SKYAEYENFSVSSEPNKYRLLVNGYSG-NAGDAFSGLNGQSFSTYDQDND 169
G3_consensus                TGYAKYANFNIANEAAKYKMTCSSYKG-NVGDSLARSIGQNFTTKDQDND 186
Fico1_Sus-scrofa            HQFAKYRSFQVAGEAEKYKLVLGGFLEGNAGDSLSSHRDQFFSTKDQDND 264
Fico2_Sus-scrofa            HQFAKYRSFQVADEAEKYMLVLGAFVEGNAGDSLTSHNNSLFTTKDQDND 261
Fico1_Homo-sapiens          HQFAKYKSFKVADEAEKYKLVLGAFVGGSAGNSLTGHNNNFFSTKDQDND 264
Fico2_Homo-sapiens          YQFAKYRSFKVADEAEKYNLVLGAFVEGSAGDSLTFHNNQSFSTKDQDND 251
FicoB_Rattus-norvegicus     HDFAKYSSFQIQGEAEKYKLILGNFLGGGAGDSLTSQNNMLFSTKDQDND 257
FicoA_Rattus-norvegicus     TSFAKYSSFQVSGEQEKYKLTLGQFLEGTAGDSLTKHNNMAFSTHDQDND 272
Fico2_Xenopus-laevis        HTSASYSNFRIAGESRNYTLSLGTFTGGDAGDSLSGHKNKGFSTKDRDND 255
Fico4_Xenopus-laevis        RTSASYSDFRIAAEAQNYTLSLGTFTGGDAGDSLYGHKNKGFSTKDRDND 255
Fico1_Xenopus-laevis        STYASYSNFSIAEESQSYTLSLRSFMGGDAGDSLSGHKNFSFSTKDRDNK 255
Fico3_Xenopus-laevis        HTYAAYSGFSITGDSNNYALRLGTFIGGDAGDSLSIHNNMAFSTKDRDND 222
Fico3_Homo-sapiens          RTFAHYATFRLLGEVDHYQLALGKFSEGTAGDSLSLHSGRPFTTYDADHD 239
Fico1_Halocynthia-roretzi   RRYAQYGTFSIAGSNDNYRLTVGDYSGTAGDSMTPRSNGQQFTTKDRDND 259
Fico2_Halocynthia-roretzi   RRYAQYGTFSIAGSNDNYRLTVGEYSGTAGDSLIANHNGKQFSTKDRDND 259
Fico3_Halocynthia-roretzi   RRYAQYGSFSIAGSNDNYRLTVGEYSGTAGDSMTPRSNGQQFSTKDRDND 290
Fico4_Halocynthia-roretzi   RRYAQYGTFSISGSNDNYRLTVGDYSGTAGDSLIGHHNGQQFSTKDQDND 276
                               * *  * :  .   : :    :     ..      .  *:* * *:.
```

```
G1_consensus                NNNS-KCATNYRGAWWFNKCVKADLNGQYFLKTPDAEHRGVYWRTW-KGA 247
G2_consensus                IWPS-NCAEKFKGAWWYSKCHSSNLNGLYWGGAHTEYASGINWGS--WGY 216
G3_consensus                KFPQ-NCAVSFKGAWWYKECHDSNLNGQYLGGTHTSFADGVNWKAW-KGY 234
Fico1_Sus-scrofa            NHSG-NCAEQYHGAWWYNACHSSNLNGRYLRGLHTSYANGVNWRSG-RGY 312
Fico2_Sus-scrofa            QYAS-NCAVLYQGAWWYNSCHVSNLNGRYLGGSHGSFANGVNWSSG-KGY 309
Fico1_Homo-sapiens          VSSS-NCAEKFQGAWWYADCHASNLNGLYLMGPHESYANGINWSAA-KGY 312
Fico2_Homo-sapiens          LNTG-NCAVMFQGAWWYKNCHVSNLNGRYLRGTHGSFANGINWKSG-KGY 299
FicoB_Rattus-norvegicus     QGSS-NCAVRYHGAWWYSDCHTSNLNGLYLRGLHKSYANGVNWKSW-KGY 305
FicoA_Rattus-norvegicus     TNGGKNCAALFHGAWWYHDCHQSNLNGRYLPGSHESYADGINWLSG-RGH 321
Fico2_Xenopus-laevis        SSPN-SCAERYKGAWWYTSCHVSHLNGLYLGGKHSSSANGVNWRSG-RGF 303
Fico4_Xenopus-laevis        SSPA-SCAERYRGAWWYTSCHSSNLNGLYLRGNHSSFANGVNWKSG-RGY 303
Fico1_Xenopus-laevis        SN----CAHTFKGGWWYETCHYSNLNGLYLHGNHTSYANGVNWSTG-RGY 300
Fico3_Xenopus-laevis        AHMAGNCAQNYKGAWWYESCHSSNLNGLYQQGEHSSSINGINWRTG-RGY 271
Fico3_Homo-sapiens          SSNS-NCAVIVHGAWWYASCYRSNLNGRYAVSEAAAHKYGIDWASG-RGV 287
Fico1_Halocynthia-roretzi   G-SGGNCAVEWSGAWWYEKCHVSNLNGIYLVGGTGATSKNVAWYHWGNNH 308
Fico2_Halocynthia-roretzi   E-YGSNCAVQWSGAWWYKSCHYSNLNGIYLVRGTGATAKNVAWYHWGNNY 308
Fico3_Halocynthia-roretzi   GWAAGHCAIDWSGAWWYGICHYSNLNGIYLVGGTGATPKNVAWYHWGNNH 340
Fico4_Halocynthia-roretzi   G-NSGNCAVSYTGAWWYQSCYNSNLNGVYHVGGTGANDKNIAWWQWKNTH 325
                                  **    *.**:  *  :.*** *          .: *

G1_consensus                NYSLKGSLMMMRRI-- 261
G2_consensus                HYSLQATTMMIRAT-- 230
G3_consensus                HYSLKATMMMIRRKK- 249
Fico1_Sus-scrofa            NYSYQVSEMKVRLT-- 326
Fico2_Sus-scrofa            NYSYKVSEMKFRAT-- 323
Fico1_Homo-sapiens          KYSYKVSEMKVRPA-- 326
Fico2_Homo-sapiens          NYSYKVSEMKVRPA-- 313
FicoB_Rattus-norvegicus     NYSYKVSEMKVRLI-- 319
FicoA_Rattus-norvegicus     RYSYKVAEMKIRAS-- 335
Fico2_Xenopus-laevis        NYSYKVSEMKFRPQS- 318
Fico4_Xenopus-laevis        KYSYEVSEIKFRPQP- 318
Fico1_Xenopus-laevis        ITHTRCPK-------- 308
Fico3_Xenopus-laevis        STLTRCQK-------- 279
Fico3_Homo-sapiens          GHPYRRVRMMLR---- 299
Fico1_Halocynthia-roretzi   VYSFKFTEIKFRRKQN 324
Fico2_Halocynthia-roretzi   VYSFKFTEIKFRKKQN 324
Fico3_Halocynthia-roretzi   VYSFKFTEIKFRKKQK 356
Fico4_Halocynthia-roretzi   NYSYKFTEIKFRKKQN 341
                                 .
```

Figure 3

Click here to download high resolution image

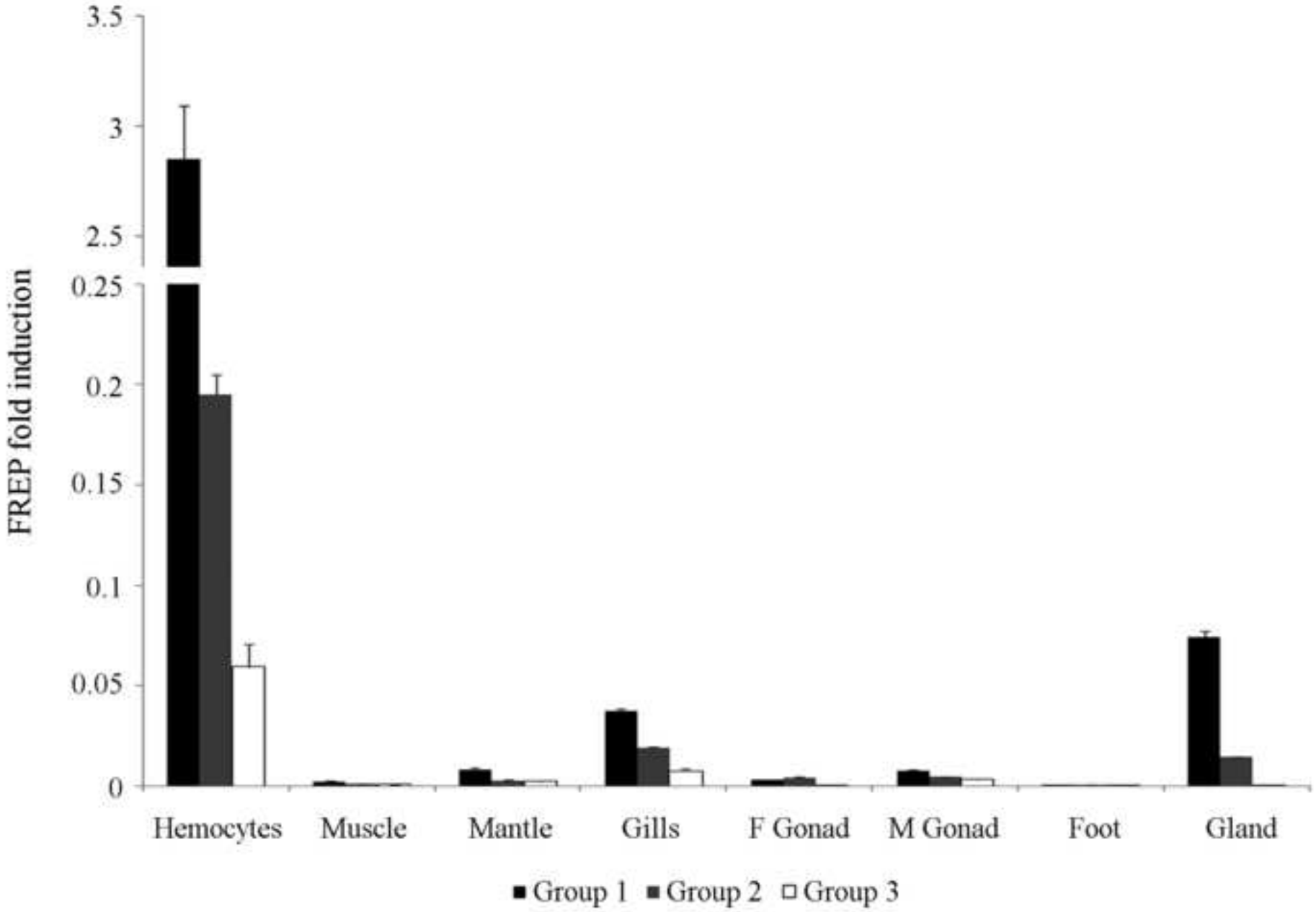


Figure 4

Click here to download high resolution image

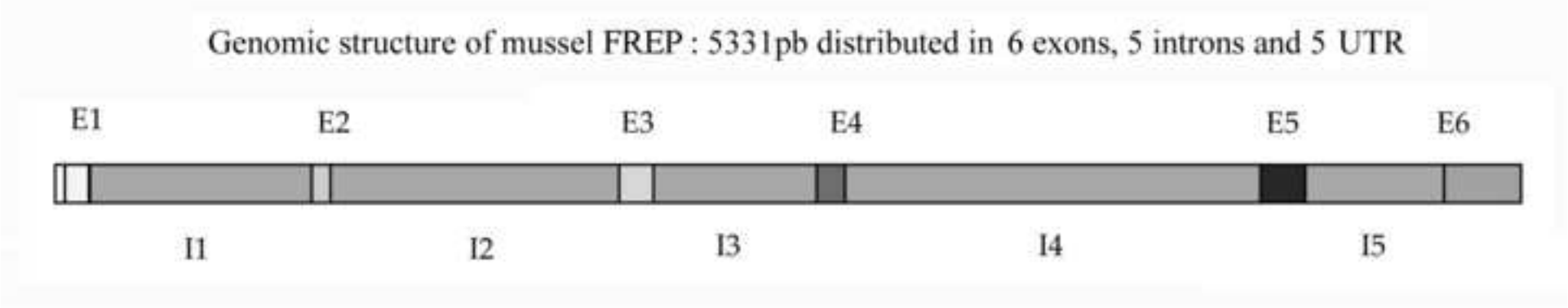


Figure 5

Click here to download high resolution image

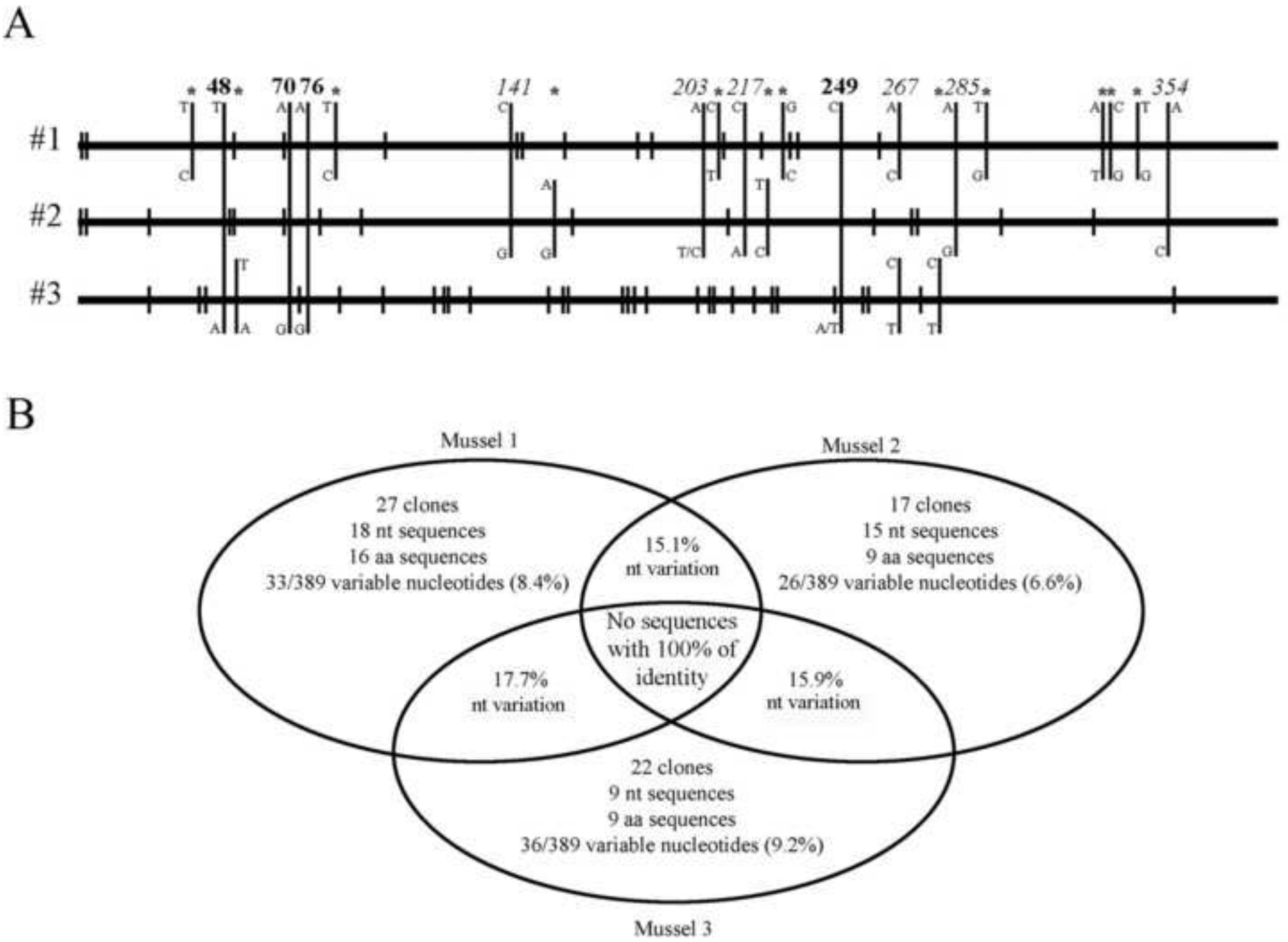


Figure 6

Click here to download high resolution image

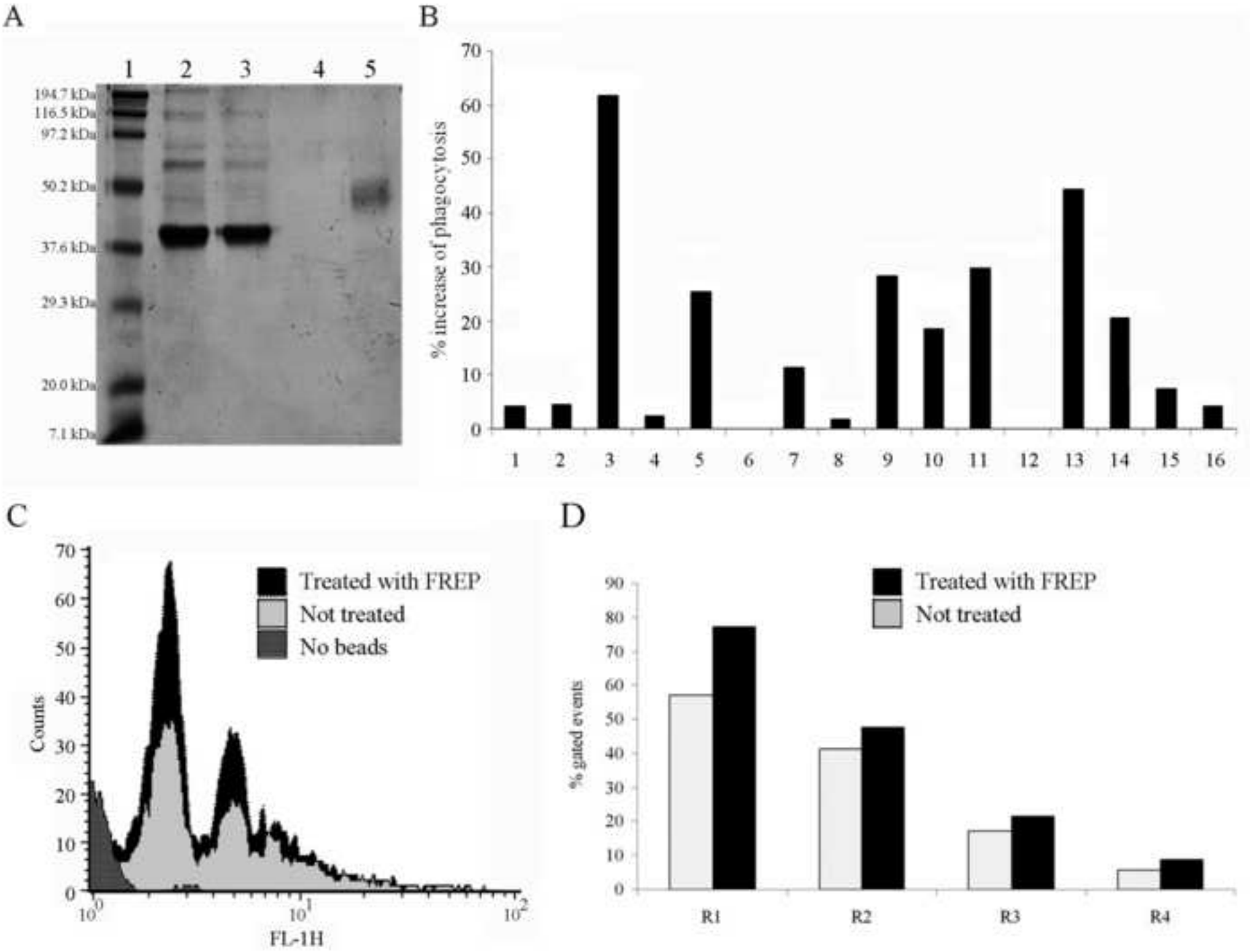


Figure 7

Click here to download high resolution image