Coordination of humoral immune factors dictates compatibility between *Schistosoma mansoni* and *Biomphalaria glabrata*

Hongyu Li[1,2], Jacob R. Hambrook[2], Emmanuel A. Pila[2], Abdullah A. Gharamah[2], Jing Fang[1,2], Xinzhong Wu[1,3,*], Patrick C. Hanington[2,*]

Affiliations

[a]Ocean College, Beibu Gulf University, Qinzhou, Guangxi 535011, China;

[b]School of Public Health, University of Alberta, Edmonton, AB T6G2G7, Canada;

[c]College of Animal Sciences, Zhejiang University, Hangzhou, Zhejiang 310058, China.

***To whom correspondence should be addressed.** Patrick C. Hanington, pch1@ualberta.ca; Xinzhong Wu, wuxzqinzhou@163.com.

Abstract

Immune factors in snails of the genus *Biomphalaria* are critical for combating *Schistosoma mansoni*, the predominant cause of human intestinal schistosomiasis. Independently, many of these factors are known to play an important role in, but not fully define, the compatibility between the model snail *B. glabrata,* and *S. mansoni*. Here, we demonstrate association between four, previously characterized humoral immune molecules; *Bg*FREP3, *Bg*TEP1, *Bg*FREP2 and Biomphalysin. We also identify unique immune determinants in the plasma of *S. mansoni*-resistant *B. glabrata* that explain the incompatible phenotype. These factors coordinate to initiate haemocyte-mediated deestruction of *S. mansoni* sporocysts via production of reactive oxygen species. The inclusion of *Bg*FREP2 in a *Bg*FREP3-initiated complex that also includes *Bg*TEP1 almost completely explains resistance to *S. mansoni* in this model. Our study unifies many

independent lines of investigation to provide a more comprehensive understanding of the snail immune system in the context of infection by this important human parasite.

Introduction

Schistosomiasis, a disease caused by parasitic trematodes of the genus *Schistosoma*, is the second-most socioeconomically devastating parasitic disease, with an estimated 252 million people infected worldwide in 2015 (Disease, Injury, & Prevalence, 2016; Mortality & Causes of Death, 2016). Larval digenean trematodes share a common feature in the use of snails (gastropod mollusks) as intermediate hosts for transmission to a vertebrate host (Esch, Barger, & Fellis, 2002). The snail *Biomphalaria glabrata* transmits several species of trematode, including *Schistosoma mansoni*, the predominant causal species of intestinal schistosomiasis (Pila, Gordy, et al., 2016). The molecular interactions between *B. glabrata* and *S. mansoni* have been studied extensively towards better understanding this essential life cycle stage of an important human parasite (Adema, 2015).

Various genetically determined resistance phenotypes exist with respect to *S. mansoni* infection in *B. glabrata* (Richards, 1975a, 1975b). Some *B. glabrata* snails (such as BS-90 strain) naturally resist *S. mansoni* infection. Susceptible *B. glabrata* (such as M-line strain) can develop transient levels of acquired resistance following previous exposure to incompatible or radiation-attenuated parasites with modest cross protection to related digenean species (Lie & Heyneman, 1975; Lie, Heyneman, & Lim, 1975; Sullivan, Richards, Joe, & Heyneman, 1982). Previous research has focused on identifying the mechanisms responsible for determining resistance profiles. Using in vivo

and in vitro models of *B. glabrata* snails, it has been demonstrated that the killing of *S. mansoni* larvae is associated with a haemocyte-mediated cytotoxic mechanism, and passive transfer of natural resistance to *S. mansoni* has been successfully accomplished when haemocytes from a susceptible *B. glabrata* strain are incubated in cell-free hemolymph (plasma) from a resistant strain (Bayne, Buckley, & DeWan, 1980a, 1980b; Granath & Yoshino, 1984; Loker & Bayne, 1982). Thus, haemocytes from resistant or susceptible *B. glabrata* strains do not appear to differ a priori in their cytotoxic capabilities, but their response requires activation by some humoral factor(s) (Bayne et al., 1980b; Granath & Yoshino, 1984; Vasquez & Sullivan, 2001), for proper recognition of *S. mansoni* and enhancing haemocyte cytotoxicity (Hahn, Bender, & Bayne, 2001).

Researchers have long sought immune determinants present in resistant *B. glabrata* plasma that specifically activate haemocytes to encapsulate and destroy *S. mansoni* sporocysts. Identified in *B. glabrata*, a large family of soluble lectins termed fibrinogen-related proteins (*Bg*FREPs) attracted widespread interest because of their unique structure (Adema, Hertel, Miller, & Loker, 1997; Gordy, Pila, & Hanington, 2015), capacity for somatic diversification (Zhang, Adema, Kepler, & Loker, 2004), and immune function (Hanington et al., 2010; Hanington & Zhang, 2011). *Bg*FREPs are composed of a C-terminal fibrinogen (FBG) domain and either one or two N-terminal immunoglobulin superfamily (IgSF) domains linked to the FBG domain by an interceding region (ICR) (Hanington et al., 2010; Pila, Li, Hambrook, Wu, & Hanington, 2017). Numerous studies have implicated *Bg*FREP3 as playing a central role in *B. glabrata* resistance to digenetic trematodes (Hanington et al., 2010; Hanington, Forys, & Loker, 2012; Hanington &

Zhang, 2011; Lockyer et al., 2012; Lockyer et al., 2008; Pila, Li, et al., 2017). *Bg*FREP3 contains two tandem IgSF domains whereas *Bg*FREP2 has one (Leonard, Adema, Zhang, & Loker, 2001). The expression of *Bg*FREP2 in BS-90 snails is dramatically up-regulated (57-fold) after exposure to *S. mansoni*, whereas M-line snails do not feature such a drastic upregulation, thereby suggesting a role for *Bg*FREP2 in resistance (Hertel, Adema, & Loker, 2005). Interestingly, the time course for heightened *Bg*FREP2 expression overlaps the interval within which sporocysts are encapsulated and killed, and only occurs in BS-90 haemocytes (Dinguirard et al., 2018; Hertel et al., 2005). Furthermore it was reported that *Bg*FREP2 is reactive with polymorphic glycoproteins found on the tegumental surface of larval *S. mansoni* (Mone et al., 2010).

In addition to *Bg*FREPs, other soluble immune effector factors which are involved in both the direct killing of sporocysts and preparation of haemocytes to mount a cell-mediated response have been characterized (Pila, Li, et al., 2017). These factors include *B. glabrata* thioester-containing proteins (*Bg*TEPs) (Bender, Fryer, & Bayne, 1992; Mone et al., 2010; Portet et al., 2018) and Biomphalysin (Galinier et al., 2013). TEPs are key components of invertebrate and vertebrat immune responses, aiding in melanization, opsonization, and killing of invading pathogens (Baxter et al., 2007; Blandin, Marois, & Levashina, 2008; Levashina et al., 2001; Pila, Li, et al., 2017; Portet et al., 2018; Povelones et al., 2013; Yassine & Osta, 2010). *Bg*TEP was first characterized as an alpha macroglobulin proteinase inhibitor, and has recently re-immerged as a molecule of interest due to its association with *Bg*FREP2 and *S. mansoni* polymorphic mucins (*Sm*PoMucs), and the capacity to bind to various invading pathogens (Bender et al., 1992;

Mone et al., 2010; Portet et al., 2018). Biomphalysin is a cytolytic protein in *B. glabrata* belonging to the β pore-forming toxin (β-PFT) superfamily (Galinier et al., 2013). Biomphalysin binds to the surface of *S. mansoni* sporocysts in the absence of plasma, while its cytolytic activity is drastically increased when plasma is present, suggesting that other factor(s) within the plasma may mediate the conversion of the oligomeric pre-pore to a functional pore (Galinier et al., 2013). Although the functional mechanisms of these factors are not thoroughly understood, studies suggest that these factors function as key determinants in the final outcome of *S. mansoni* challenge of *B. glabrata* (Galinier et al., 2013; Mone et al., 2010).

While studies have implicated *Bg*FREP3 in resistance to *S. mansoni*, the underlying mechanism of *Bg*FREP3 function; how it binds to *S. mansoni* sporocyst surfaces, and then how recognition is translated into haemocyte engagement, activation, and ultimately parasite encapsulation, is still unknown (Hanington et al., 2010). Here, we report an association between *Bg*FREP3, *Bg*TEP1, Biomphalysin (UniProtKB/TrEMBL: A0A182YTN9 and A0A182YTZ4) and *Bg*FREP2 (UniProtKB/TrEMBL: A0A2C9L9F5). *Bg*FREP3 associates with *Bg*TEP1 and Biomphalysin in both M-line and BS-90 strains, but only in BS-90 strain uniquely interacts with *Bg*FREP2 and other versions of *Bg*FREP3 (a variant of *Bg*FREP3.3), providing us with an important insight into why BS-90 strain is refractory to *S. mansoni* infection. In this study, we demonstrate that *Bg*FREP3 binds to *S. mansoni* sporocysts without any other soluble plasma factors, yet *Bg*FREP2 relies on *Bg*TEP1. However, *Bg*FREP3 still interacts with *Bg*TEP1 to form unique immune complexes, which significantly enhance the ability of haemocytes and

plasma from *S. mansoni*-susceptible *B. glabrata* (M-line) to kill *S. mansoni* sporocysts. A more striking finding is that the combination of *Bg*FREP3, *Bg*TEP1 and *Bg*FREP2 renders M-line haemocytes capable of killing *S. mansoni* sporocysts at nearly the same level as *S. mansoni*-resistant BS-90 haemocytes. Reactive oxygen species (ROS) is shown to play an important role during this haemocyte-mediated killing of *S. mansoni* sporocysts. These results provide insight into how the numerous previously characterized immune factors known to be important in the anti-*S. mansoni* immune response to *B. glabrata* are acting in concert to defend the snail host.

Results

*Bg*FREP3 interacts with *Bg*TEP1, Biomphalysin, *Bg*FREP2 and other *Bg*FREP3 variants in snail plasma.

To identify the factors in snail plasma which interact with *Bg*FREP3, we produced recombinant *Bg*FREP3 (r*Bg*FREP3, GenBank: AAK28656.1) using an insect expression system (Fig. 1A and Figure 1 – figure supplement 1) and performed a series of pull-down experiments. We initiated our investigations with *Bg*FREP3 because we observed that it is expressed at greater abundance in BS-90 compared with M-line snails. This observation was confirmed using Western-blot that showed that the hemolymph of BS-90 snails contain more *Bg*FREP3 than does M-line snail hemolymph (Fig. 1B). Among the proteins identified using liquid chromatography-tandem mass spectrometry (LC-MS/MS) in the eluent of these pull-down experiments was *Bg*TEP1 (Fig. 1C). We performed four r*Bg*FREP3 pull-down experiments, each yielded a band at approximately 200-kDa, which corresponded to the full-length *Bg*TEP1, which was identified from both M-line and BS-

90 strains (Fig. 1C). In three instances, we cut the bands and sent for mass spectrometry identification, two of which detected the *Bg*TEP1 pulled down by r*Bg*FREP3 in M-line plasma, while the third detected *Bg*TEP1 in both M-line and BS-90 strains (Fig. 1D). From the LC-MS/MS analysis we obtained a total of 17 unique peptides, 16 peptides were detected from M-line strain, 4 peptides were from BS-90 strain, and 3 peptides were present in both BS-90 and M-line strains (Table. S2 and Fig. 1E). While all peptides mapped to *Bg*TEP, there is uncertainty as to which specific *Bg*TEP proteins are present. Twelve peptides mapped to *Bg*TEP-ACL00841.1; 15 peptides mapped to *Bg*TEP1.1-ADE45332.1, *Bg*TEP1.2-ADE45339.1, *Bg*TEP1.3-ADE45340.1, *Bg*TEP1 1.4-ADE45341.1 and *Bg*TEP1.5-ADE45333.1, however, there was overlap between some of these peptides, which mapped to all known *Bg*TEP sequences (Figure 1 – figure supplement 2). The identified *Bg*TEP1 showed a relatively large difference from *Bg*TEP (ACL00841.1), but could not be distinguished from other *Bg*TEP1 sequences (Figure 1 – figure supplement 2), so we termed it *Bg*TEP1.

The full-length sequence of *Bg*TEP1 generally contains 1445 or 1446 amino acids (Figure 1 – figure supplement 2). The peptides identified by LC-MS/MS cover 16.3% of the full-length sequence (Figure 1 – figure supplement 2). The identified peptides are distributed over the full-length *Bg*TEP1, with the peptides identified from the BS-90 strain being concentrated in the N-terminal region (Fig. 1D). Background transcript abundance of *Bg*TEP1 detected by qRT-PCR in M-line and BS-90 snails, suggest that there is no significant difference ($P > 0.05$) in baseline *Bg*TEP1 expression between the two strains (Figure 1 – figure supplement 3). These results indicate that the interaction of *Bg*FREP3 with *Bg*TEP1 doesn't require the involvement of any parasite components.

158 Recombinant *Bg*TEP1 (r*Bg*TEP1, GenBank: ADE45332.1) was also used to perform
159 a series of pull-down experiments (Fig. 1A and E). Biomphalysin (UniProtKB/TrEMBL:
160 A0A182YTN9 and A0A182YTZ4) was identified in both M-line and BS-90 plasma from
161 r*Bg*TEP1 pull-down experiments and was also found to be present in the pull-down
162 studies using *Bg*FREP3 (Fig. 1C and E). Analysis suggests that this Biomphalysin
163 (A0A182YTZ4) was most similar (97% amino acid identity) to the Biomphalysin
164 previously reported by Galinier and Portela *et al.* (Galinier et al., 2013) (GenBank:
165 KC012466) (Figure 1 – figure supplement 4). The 13 peptides identified by LC-MS/MS
166 covered 28.3% of Biomphalysin (A0A182YTN9) amino acid sequence (Fig. 1F, Table S2
167 and Figure 1 – figure supplement 4). These peptides were evenly distributed throughout
168 the Biomphalysin protein (Fig. 1F). The theoretical molecular weight of Biomphalysin
169 (A0A182YTN9 and A0A182YTZ4) is ~64-kDa. The Biomphalysin we identified via
170 protein pull-down emerged at three different molecular weights, one at > 400-kDa (Fig.
171 1C), which likely represents the heptamer formed by Biomphalysin (Galinier et al.,
172 2013), the second site at roughly 60-kDa (Fig. 1C and E), which is in agreement with the
173 theoretical molecular weight, and the third at around 55-kDa (Fig. 1C), which may
174 represent the size after proteolysis (Galinier et al., 2013).

175 Comparing r*Bg*FREP3 pull-down results between M-line and BS-90 *B. glabrata*
176 identified 2 unique proteins in the BS-90 lane. One of these proteins is *Bg*FREP2
177 (UniProtKB/TrEMBL: A0A2C9L9F5) (Fig. 1C and E). Five peptides were identified by
178 LC-MS/MS analysis and covered 16.6% of *Bg*FREP2 (GenBank: AAK13550.1) amino
179 acid sequence and were located throughout the IgSF and FBG of the protein (Fig. 1G and
180 Table S2 and Figure 1 – figure supplement 5). Three of the five identified peptides

181 specifically belonged to *Bg*FREP2 (AAK13550.1), so the identified *Bg*FREP2 appears to
182 represent a specific *Bg*FREP2 variant that is distinguished from other *Bg*FREP members
183 that have been published (Figure 1 – figure supplement 5).

184 Only *Bg*FREP2 was identified, at ~55-kDa, from the BS-90 strain plasma when
185 r*Bg*FREP3 was used as the bait protein in pull-down experiments (Fig. 1C). This
186 contrasts with our other data in which *Bg*FREP2 was identified from plasma of both M-
187 line and BS-90 strains when r*Bg*TEP1was used as bait in pull-down experiments (Fig.
188 1E). *Bg*FREP2 was identified by LC-MS/MS from two distinct molecular weight bands
189 in the r*Bg*TEP1 study, one band, at ~60 kDa also included peptides that mapped to
190 Biomphalysin, the other band, identified at ~55-kDa was contained peptides for only
191 *Bg*FREP2 (Fig. 1C and E). Thus, all of the *Bg*FREP2 peptides were identified from
192 protein bands that were of greater molecular weight than the predicted *Bg*FREP2, which
193 is 43.8-kDa. However, this is consistent with the observed molecular weight of *Bg*FREP2
194 by Moné *et al.* (Mone et al., 2010), and could be explained by post-translational
195 modifications such as glycosylation (Adema et al., 1997; Mone et al., 2010; Zhang, Zeng,
196 & Loker, 2008).

197 Another unique protein identified from the BS-90 lane was a variant of *Bg*FREP3.
198 Of the 9 peptides identified by LC-MS/MS, 7 specifically matched *Bg*FREP3.3
199 (Genbank: AEO50747.1), the other two matched other members of *Bg*FREP3 family
200 (*Bg*FREP3.3-AAO59915.1, *Bg*FREP3.2-AEO50746.1, *Bg*FREP3.2-AAK28656.1 and
201 *Bg*MFREP3-AAK13548.1), suggesting that this protein is likely a FREP3 variant (Figure
202 1 – figure supplement 6 and Table S2). The 7 peptides that specifically belonged to
203 *Bg*FREP3.3 (AEO50747.1) covered 15.6% of the amino acid sequence length and mainly

204 concentrated in the interceding region (ICR) (5 peptides), one peptide in each of the
205 IgSF2 domain and FBG domain of *Bg*FREP3.3 were also identified (Figure 1 – figure
206 supplement 6).

207 **_Bg_FREP3 independently recognizes and binds to primary sporocysts of _S. mansoni_.**

208 Our data suggests a close and complex interaction between *Bg*FREP3, *Bg*TEP1,
209 Biomphalysin and *Bg*FREP2. These factors have each independently been reported to be
210 involved in the recognition of *S. mansoni* sporocysts (Adema et al., 1997; Galinier et al.,
211 2013; Hanington et al., 2010; Hanington et al., 2012; Mone et al., 2010; Portet et al.,
212 2018; X. J. Wu et al., 2017). However, only Biomphalysin is known to directly bind to
213 the surface of *S. mansoni* sporocysts without the aid of any other plasma factors (Galinier
214 et al., 2013). The mechanism by which *Bg*FREP3, *Bg*TEP1 and *Bg*FREP2 bind to
215 primary sporocysts of *S. mansoni* is still not clear. To explore this issue, we produced
216 r*Bg*FREP3, r*Bg*TEP1 and r*Bg*FREP2 (Fig. 1A) to observe whether they associate with
217 primary sporocysts of *S. mansoni*. Immunocytochemistry clearly shows that r*Bg*FREP3 is
218 able to interact with *S. mansoni* sporocysts independently, while r*Bg*TEP1 and r*Bg*FREP2
219 alone do not have such capabilities (Fig. 2). However, incubation of sporocysts with pre-
220 combined r*Bg*FREP2 and r*Bg*TEP1 yields a signal, suggesting that these two factors
221 require complex formation prior to being able to associate with *S. mansoni* sporocysts
222 (Fig. 2).

223 **_Bg_FREP3 and _Bg_TEP1 form an immune complex.**

224 To further investigate whether *Bg*FREP3 and *Bg*TEP1 form a complex, we carried
225 out a set of immunoblotting experiments with recombinant proteins. As expected,
226 r*Bg*FREP3 did form a complex with r*Bg*TEP1 (Fig. 3A). The complex appeared in a

region of high molecular weight (> 460-kDa), and the signal decreased gradually with the decrease of r*Bg*TEP1:r*Bg*FREP3 ratio in a dose-dependent manner (Fig. 3A). As shown in Fig. 3A, r*Bg*FREP3 exists as a high molecular weight multimer in a non-denatured state. Since the molecular weights of the r*Bg*FREP3 multimer and the r*Bg*FREP3-r*Bg*TEP1 complex were both > 460-kDa, it was difficult to distinguish them using Western bolt. To confirm the interaction between r*Bg*FREP3 and r*Bg*TEP1, a far-Western blotting approach, in which decreasing amount of r*Bg*TEP1 was imprinted on the membrane, and then detected by using biotinylated r*Bg*FREP3 as primary recognition agent was performed. A band was detected in the lane with the largest amount of r*Bg*TEP1, confirming the interaction between r*Bg*FREP3 and r*Bg*TEP1 (Fig. 3B).

*Bg*FREP2 and *Bg*TEP1 have been shown to form immune complexes with *Sm*PoMucs derived from *S. mansoni* (Mone et al., 2010), suggesting that *Bg*FREP2 and *Bg*TEP1 associate with each other in snail plasma. To test this assumption, we performed a set of immunoblot assays using r*Bg*FREP2 and r*Bg*TEP1. The results of Western blot under non-denaturing conditions showed that a protein at > 460 kDa after mixing r*Bg*FREP2 and r*Bg*TEP1. This protein decreased in intensity on the Western blot in a dose dependent manner along with the decrease of the r*Bg*TEP1:r*Bg*FREP2 ratio (Fig. 3C upper panel). However, r*Bg*FREP2 does not exist as a multimer under non-denaturing conditions (Fig. 3C upper panel), which is consistent with the findings of Zhang *et al.* that *Bg*FREP2 did not form a multimer in snail plasma (Zhang et al., 2008). In addition, we further verified the interaction between r*Bg*FREP2 and r*Bg*TEP1 by Far-Western blot (Fig. 3C lower panel).

Association of *Bg*FREP3 and *Bg*TEP1 facilitates recognition and killing of *S. mansoni* sporocysts.

Since *Bg*FREP3 does not rely on *Bg*TEP1 to recognize pathogens, but does associate with *Bg*TEP1, we hypothesize that *Bg*TEP1 plays a role in the downstream immune response triggered by *Bg*FREP3 recognition of *S. mansoni* sporocysts. We treated primary sporocysts with haemocytes or plasma from BS-90 and M-line *B. glabrata* snails with r*Bg*FREP3, r*Bg*TEP1 and a combination of the two in vitro.

Plasma (hemoglobin-low/free) from the resistant BS-90 strain of *B. glabrata* snails is naturally able to kill 67% (SEM 3.7%; n=10) of *S. mansoni* sporocysts by 48 h post incubation, while plasma from the susceptible M-line strain, which kills 20% (SEM 4.5%; n=10), is not significantly more effective than the medium control in which 13% (SEM 2.1%; n=10) were killed at 48 h (Fig. 4A). The addition of r*Bg*FREP3 (36% [SEM 5.0%; n=10]) and r*Bg*TEP1 (32% [SEM 3.3%; n=10]) independently did not significantly enhanced the ability of M-line plasma to kill *S. mansoni* sporocysts by 48 h (Fig. 4A). However, the combined addition of r*Bg*FREP3 and r*Bg*TEP1 did significant increase M-line-plasma-mediated killing of *S. mansoni* sporocysts by 48 h post incubation (56% [SEM 4.8%; n=10]) compared to the application of either recombinant protein alone. The combined addition of r*Bg*FREP3 and r*Bg*TEP1 rendered the capacity of M-line plasma able to kill *S. mansoni* sporocysts statistically insignificant from BS-90 plasma ($P > 0.05$) (Fig. 4A). Recombinant proteins, in the absence of *B. glabrata* plasma, did not display any capacity to kill *S. mansoni* sporocysts (Figure 4 – figure supplement 1A). This suggests that r*Bg*FREP3 and r*Bg*TEP1 activate a factor in plasma to confer the ability to kill sporocysts.

Haemocytes derived from BS-90 strain *B. glabrata* are innately capable of destroying *S. mansoni* sporocysts, killing 88% (SEM 3.6%; n=10) of sporocysts after 48 h of incubation in vitro (Fig. 4B). The basal ability of M-line haemocytes to kill sporocysts (13% [SEM 3%; n=10]) is not significantly different from the medium-alone control (13% [SEM 2.1%; n=10]), and is only slightly increased by adding r*Bg*FREP3 (24% [SEM 3.4%; n=10]) at 48 h (Fig. 4B). Addition of r*Bg*TEP1 did significantly increase sporocyst killing by M-line haemocytes to 40% (SEM 3.9%; n=10) (Fig. 4B), however, this killing response was significantly enhanced to 61% (SEM 5.5%; n=10) when r*Bg*FREP3 and r*Bg*TEP1 were incubated with M-line haemocytes together (Fig. 4B). However, the combination of r*Bg*FREP3 and r*Bg*TEP1 did not convey a killing capacity to M-line snails that paralleled BS-90 haemocytes (Fig. 4B). This suggests that there is yet another factor(s) affecting this process, causing haemocytes from BS-90 and M-line to differ in their ability to destroy sporocysts. In our r*Bg*FREP3 pull-down experiments, *Bg*FREP2 was one of the two proteins uniquely identified as associating with r*Bg*FREP3 in the plasma of BS-90 *B. glabrata* (Fig. 1C). Incubation of r*Bg*FREP2 in combination with r*Bg*FREP3 and r*Bg*TEP1 enabled M-line haemocytes to kill *S. mansoni* sporocysts (83% [SEM 4.8%; n=6]) to statistically indistinguishable ($p > 0.05$) levels as BS-90 haemocytes (88% [SEM 4.7%; n=6])) (Fig. 4C). To verify whether the r*Bg*FREP2-r*Bg*TEP1 complex dramatically enhanced the ability of M-line haemocytes to killing *S. mansoni* sporocysts above either factor independently, as does the r*Bg*FREP3-r*Bg*TEP1 complex, a similar experiment was performed, showing that r*Bg*FREP2-r*Bg*TEP1 did not play such a role (Fig. 4D). Independent treatments of r*Bg*FREP2 (21% [SEM 3.1%; n=10]) and r*Bg*TEP1 served controls (Figure 4 – figure supplement 1B),

295 demonstrating that *Bg*FREP2 on alone does not facilitate sporocyst killing above control,
296 whereas r*Bg*TEP1 killing (50% [SEM 3.9%; n=10]) is indistinguishable ($P > 0.05$) from
297 r*Bg*FREP2 and r*Bg*TEP1 combined (55% [SEM 3.7%; n=10]) (Fig. 4D).
298 These results made us speculate that *Bg*FREP3 was the determining factor
299 controlling sporocyst engagement and killing. Addition of three anti-r*Bg*FREP3
300 antibodies to the mixture of r*Bg*FREP3, r*Bg*TEP1 and r*Bg*FREP2 significantly abrogates
301 the increase in the ability of M-line haemocytes to kill sporocysts (51% [SEM 3.1%; n=6]
302 compared to 83% [SEM 4.8%; n=6]) (Fig. 4C). Pre-incubation of BS-90 haemocytes with
303 the anti-r*Bg*FREP3 antibodies also significantly reduced their ability to destroy
304 sporocysts (68% [SEM 3.1%; n=6] compared to 88% [SEM 4.7%; n=6]) (Fig. 4C).
305 Combined pre-immunized IgG purified from serum from the same rabbits from which the
306 antibodies were produced was used as a negative control, demonstrating that it did not
307 significantly affect the killing *S. mansoni* sporocysts by M-line (51% [SEM 3%; n=6])
308 and BS-90 haemocytes (90% [SEM 3.7%; n=6]) (Figure 4 – figure supplement 1C).
309 These results suggest that r*Bg*FREP3 plays a central role in initiating snail haemocyte-
310 mediated anti-*S. mansoni* immune responses.
311 Finally, we hypothesize that haemocyte-mediated killing likely involves the
312 production of ROS, which have been shown to be crucial to the haemocyte-mediated
313 clearance of *S. mansoni* sporocysts (Hahn, Bender, & Bayne, 2000; Hahn et al., 2001). To
314 investigate the role of ROS in the combined ability of r*Bg*FREP3 and r*Bg*TEP1 to
315 enhance the ability of M-line haemocytes to kill *S. mansoni* sporocysts, ROS inhibitors
316 N-acetyl-L-cysteine (NAC) and Diphenyleneiodonium (DPI) were applied to the
317 sporocyst killing assays. Co-incubation of BS-90 haemocytes with NAC and DPI

318 significantly reduced their ability to kill *S. mansoni* sporocysts (45% [SEM 5.6%; n=6]
319 and 57% [SEM 4.2%; n=6] respectively) compared to BS-90 haemocytes in the vehicle
320 control (89% [SEM 3.1%; n=6]) by 48 h post incubation (Fig. 4E). The ability of M-line
321 haemocytes primed by the r*Bg*FREP3-r*Bg*TEP1 complex to destroy sporocysts (67%
322 [SEM 4.9%; n=6]) was also significantly abrogated by the addition of NAC (40% [SEM
323 3.7%; n=6]) or DPI (52% [SEM 3.1%; n=6]) (Fig. 4F).
324

325 **Discussion**

326 In 1984, Yoshino *et al.* successfully achieved passive transfer of resistance to *S.*
327 *mansoni* from a refractory *B. glabrata* strain to susceptible snails via injection of cell-free
328 hemolymph (plasma) (Granath & Yoshino, 1984). Similar passive transfer of resistance
329 had previously been shown using an in vitro model (Bayne et al., 1980a, 1980b; Loker &
330 Bayne, 1982). Although the oxidation of haemoglobin, which is the main protein
331 component (97% of total protein) in *B. glabrata* plasma was later found to be toxic to *S.*
332 *mansoni* sporocysts in long-term cultures (Bender, Bixler, Lerner, & Bayne, 2002), and
333 thus likely explained the passive transfer study results, further studies suggest that
334 specific anti-parasitic factors must be present at or above a certain threshold level in *S.*
335 *mansoni*-resistant snail plasma to stimulate a haemocyte-mediated destruction of
336 sporocysts within 24–48 h post-infection (Dinguirard et al., 2018). Although many
337 important immune factors have been identified from *B. glabrata* plasma/haemocytes,
338 such as *Bg*FREPs (Adema et al., 1997), *Bg*TEP1 (Mone et al., 2010), Biomphalysin
339 (Galinier et al., 2013), Toll-like receptors (*Bg*TLRs) (Pila, Tarrabain, Kabore, &
340 Hanington, 2016), granulin (*Bg*GRN) (Pila, Gordy, et al., 2016) and macrophage

migration inhibitory factor (*Bg*MIF) (Baeza Garcia et al., 2010), it is still not known which factors ultimately underpin Biomphalaria immunity *S. mansoni*. Here, we provide insight into the broader interactions that take place between *B. glabrata* immune factors that leads to development of an *S. mansoni*-resistant phenotype.

Previous studies have shown that *Bg*FREPs are capable of binding to digenean trematode sporocysts and are up-regulated in response to larval infection (Adema et al., 1997; X. J. Wu et al., 2017; Zhang, Leonard, Adema, & Loker, 2001). *Bg*FREP3 has been shown to be a central part of the anti-*S. mansoni* immune response (Hanington et al., 2010). *Bg*FREP3 is up-regulated based on three characteristics (size, strain, and acquired resistance) of resistance of *B. glabrata* to infection with *S. mansoni* or *Echinostoma paraensei* (Hanington et al., 2010). Knock-down of *Bg*FREP3 in resistant snails using siRNA-mediated interference altered the phenotype of resistant *B. glabrata* by increasing susceptibility to *E. paraensei* infection (Adema, 2015; Hanington et al., 2010). Nevertheless, we have little knowledge of the underlying mechanism by which *Bg*FREP3 interacts with other snail plasma factors in order to signal and trigger an immune response after recognizing the pathogen. In order to explore this mechanism, we conducted a series of pull-down experiments that allowed us to discover the interaction between *Bg*TEP1, Biomphalysin and *Bg*FREP2 with *Bg*FREP3 (Fig. 1).

Proteomic analysis of *B. glabrata* plasma revealed that *Bg*FREP3, *Bg*TEP1 and *Bg*FREP2 display affinity for *S. mansoni* sporocyst tegumental membrane proteins (Portet et al., 2018; X. J. Wu et al., 2017). *Bg*TEP1 was previously reported to play a role in *Bg*FREP2 recognition of *S. mansoni* sporocyst *Sm*PoMucs, (Mone et al., 2010; Portet et al., 2018). Our immunofluorescence results illustrate that *Bg*FREP3 independently

recognizes *S. mansoni* sporocysts, whereas *Bg*FREP2 must be involved with the participation of *Bg*TEP1 (Fig. 2). We speculate that this may be due to their structural differences: *Bg*FREP3 possesses two IgSF domains, yet *Bg*FREP2 has only one; *Bg*FREP3 exists as a multimer in its natural state (Fig. 3), whereas *Bg*FREP2 does not (Zhang et al., 2008). Alternatively, it is possible that the targets they recognize are different, so the immunoaffinity reactions that occur are also fundamentally different. A recent report suggests that *Bg*FREPs are likely to form multimers through the coiled-coil region of the ICR. However, this has not been supported by experimental evidence and does not completely rule out the possibility that *Bg*FREP multimer formation is mediated by the FBG or IgSF domain (Gorbushin, 2019).

*Bg*TEP1 and *Bg*FREP3 form a complex without any *S. mansoni* molecules present (Fig. 3). Since *Bg*FREP3 does not require *Bg*TEP1 to recognize pathogens, *Bg*TEP1 is more likely to play a role in the downstream immune response initiated by *Bg*FREP3. Following the addition of r*Bg*FREP3 and r*Bg*TEP1, plasma demonstrated an increased capacity to kill sporocysts compared to controls (Fig. 4A). Because the *B. glabrata* plasma used in these tests had haemoglobin removed, the toxicity of hemoglobin oxidation to *S. mansoni* sporocysts was ruled out. These studies also demonstrated that recombinant *Bg*FREP3 and *Bg*TEP1 alone were not able to increase sporocyst killing, but showed that the addition of cell-free plasma with both recombinants was required, thereby implying the necessity of another factor. Based on our pull-down experiments, we think that this factor could be a Biomphalysin. Biomphalysin displays similarities to members of the β-PFT superfamily (Howard & Buckley, 1985; MacKenzie, Hirama, & Buckley, 1999; Wilmsen, Leonard, Tichelaar, Buckley, & Pattus, 1992; Xu, Wang, Yu, &

Sun, 2014), and has cytolytic activity mediated by a plasma factor(s) that remains to be identified (Galinier et al., 2013). Our results suggest a scenario in which *Bg*FREP3 and/or *Bg*TEP1 could be the factors associating with Biomphalysin to promote plasma-mediated sporocyst killing. *Bg*FREP3 and/or *Bg*TEP1 might activate Biomphalysin in an unknown manner, assist in the oligomerization of Biomphalysin while it forms its heptameric channel, or mediate the conversion of the oligomeric prepore to a functional pore. Further investigation is warranted to illustrate the role of *Bg*FREP3 and *Bg*TEP1 in the activation of Biomphalysin.

Within 2–3 h of *S. mansoni* entry into *B. glabrata* , haemocytes surround the developing sporocysts forming multi-layered cellular capsules that kill encapsulated larvae within 24–48 h post-infection (Dinguirard et al., 2018; Loker & Bayne, 1982). This occurs in both susceptible and resistant snail strains and has been replicated in our in vitro functional studies, where only resistant snails can eventually destroy the sporocysts without external assistance. Dinguirard *et al.* conducted a comparative proteomic analysis on the responses of haemocytes participating in *S. mansoni* sporocyst encapsulation reactions from susceptible and resistant *B. glabrata* strains, showing that striking differences in proteins expressed, with susceptible snail haemocytes exhibiting extensive downregulation of protein expression and a lower level of constitutively expressed proteins involved immunity (e.g., *Bg*FREP2) and oxidation-reduction compared to resistant snails (Dinguirard et al., 2018). This could explain why BS-90 haemocytes alone can kill *S. mansoni* sporocysts. After challenge, the transcript abundance of numerous *Bg*FREPs increases by 3-fold or greater in BS-90 when compared to M-line *B. glabrata* (Adema et al., 1997). Noteworthy in the context of our current study is that following

challenge BS-90 snails with *S. mansoni*, transcript abundance for *Bg*FREP2 increases over 50-fold by 1 day post-exposure (Dinguirard et al., 2018; Hertel et al., 2005), *Bg*FREP3 increases around 4-fold by 1 day post-exposure (Hanington et al., 2010), and *Bg*TEP1 which increases by 2- to 3.5-fold from 6 to 24 h post-exposure (Portet et al., 2018). These dramatic increases in the abundance of immune-relevant transcripts occurring over the first 12-24 h post exposure to *S. mansoni* sporocysts in BS-90 snails aligns closely with our results that demonstrate sporocyst killing by BS-90 haemocytes rapidly increases after 12 h post-exposure (Fig. 4).

Numerous studies have demonstrated the production of ROS by haemocytes in *B. glabrata*, especially H2O2, which plays a vital role in anti-schistosome defense (Galinier et al., 2013; Hahn et al., 2000, 2001). Haemocytes from *S. mansoni*-resistant snails produce significantly more ROS than susceptible snails (Bender, Broderick, Goodall, & Bayne, 2005; Bender, Goodall, Blouin, & Bayne, 2007; Goodall, Bender, Broderick, & Bayne, 2004). In our study, ROS was found to be critical in BS-90 haemocyte-mediated *S. mansoni* sporocyst killing, as this ability was significantly reduced after pre-incubation with the ROS inhibitors NAC and DPI (Fig. 4E). However, according to the phenomenon of passively transferred resistance, the cytotoxic potential of susceptible haemocytes may not fundamentally different from resistant snails (Granath & Yoshino, 1984). We infer that M-line plasma may lack some factors or the levels of these factors do not reach a certain threshold and fail to stimulate haemocytes to produce ROS, or M-line haemocytes are short of some receptors to receive signals that up-regulate ROS. Previously, our team identified a toll-like receptor (*Bg*TLR) from haemocyte surface of *B. glabrata* snails, which differed in abundance between M-line and BS-90 strains (Pila, Tarrabain, et al.,

433 2016). It has been repeatedly demonstrated in other organisms that activation of Toll-like
434 receptors is associated with increased intracellular ROS (Asehnoune, Strassheim, Mitra,
435 Kim, & Abraham, 2004; Tsung et al., 2007; Wong et al., 2009).

436 Our results indicate that the *Bg*FREP3-*Bg*TEP1 complex plays a role in raising ROS
437 levels in M-line haemocytes (Fig. 4F). We found that BS-90 haemolymph contains more
438 *Bg*FREP3 protein than M-line (Fig. 1B). Moreover, the BS-90 haemocytes significantly
439 increased the expression level of *Bg*FREP2 during encapsulating *S. mansoni* sporocysts
440 compared with susceptible snail haemocytes (Dinguirard et al., 2018). These suggest that
441 *Bg*FREP3 and *Bg*FREP2 may play a crucial role in inducing haemocytes which take part
442 in encapsulating sporocysts to generate ROS. We suppose that *Bg*TEP1 also plays an
443 opsonin role in the *B. glabrata* snail haemolymph, which interacts with molecules on the
444 surface of haemocytes, thereby promoting the phagocytosis of pathogens.

445 Finally, our data hints at fundamental differences in the *Bg*FREP3-mediated immune
446 response between BS-90 and M-line *B. glabrata*. Protein Pull-down experiments yielded
447 two obvious and unique protein bands that were present in BS-90 but not M-line pull-
448 downs. We identified the bands to be a variant of *Bg*FREP3.3 and *Bg*FREP2 (Fig. 1C),
449 showing different immune factor interactomes between different strains. This suggests
450 that *Bg*FREP3 variants and *Bg*FREP2 in BS-90 plasma are more prone to r*Bg*FREP3
451 binding, and the combination of different versions of *Bg*FREPs seems to play an
452 important and unknown role in snail resistance. Both *Bg*FREP2 and *Bg*FREP3 are
453 characterized by high diversity, our results imply that the versions of *Bg*FREP2 and
454 *Bg*FREP3 in the two snail strains are different. After exposure to *S. mansoni*, the
455 expression of *Bg*FREP2 in resistant snails is superior to that of susceptible snails

(Dinguirard et al., 2018; Gordy et al., 2015; Hertel et al., 2005). This again suggests that *Bg*FREP2 is important in determining the outcome of the compatibility of *B. glabrata* against *S. mansoni*. Our study implies that *Bg*FREP3 seems to play an "all or none" switching effect in determining B.glabrata-*S. mansoni* compatibility, while *Bg*FREP2 seems to play an additive role.

It has been decades since the discovery and initial characterization of *Bg*FREPs in the *B. glabrata* immune response to *S. mansoni*. While numerous studies since have implicated both *Bg*FREPs (*Bg*FREP3 and *Bg*FREP2) along with a suite of other immune factors, as being important elements of the snail immune response to *S. mansoni*, how the response is coordinated has remained elusive. Our study presents a model in which many of the known determinants of snail resistance to *S. mansoni* are found to work in concert to protect the snail host against infection. We map a mechanism for the interaction of *Bg*FREP3, *Bg*TEP1, *Bg*FREP2 and Biomphalysin with *S. mansoni* sporocysts in *B. glabrata* snail haemolymph that ultimately leads to sporocyst killing (Fig. 5). Central to this response is *Bg*FREP3, which is able to recruit unique representatives of both *Bg*FREP2 and *Bg*FREP3 in *S. mansoni*-resistant BS-90 *B. glabrata*. With this new-found understanding of the *B. glabrata* immune response comes the ability to now undertake comprehensive assessments of how modifications to both the snail host and the parasite influence the underpinning immunological drivers of compatibility in this important model for studying human schistosomes.

Materials and Methods

Live material and Experimental Treatments.

M-line and BS-90 strain *B. glabrata* snails were used in this study. The M-line strain is highly susceptible to the PR-1/ NMRI strain *S. mansoni* infection (Newton, 1955), while the BS-90 strain is resistant. M-line and BS-90 strain *B. glabrata* snails, and *S. mansoni* were maintained at the University of Alberta as described previously (Pila, Gordy, et al., 2016). NMRI strain of *S. mansoni* was obtained from infected Swiss-Webster mice provided by the NIH/NIAID Schistosomiasis Resource Center at the Biomedical Research Institute (Cody et al., 2016). All animal work observed ethical requirements and was approved by the Canadian Council of Animal Care and Use Committee (Biosciences) for the University of Alberta (AUP00000057).

Recombinant *Bg*FREP3, *Bg*TEP1 and *Bg*FREP2 Synthesis and Purification.

Recombinant *Bg*FREP3 (r*Bg*FREP3, derived from GenBank: AY028461.1), *Bg*TEP1 (r*Bg*TEP1, GenBank: HM003907.1) and *Bg*FREP2 (r*Bg*FREP2, GenBank: AY012700.1) were generated by using the Gateway cloning system according to the manufacturer's instructions (Life Technologies) as previously described (Hambrook, Kabore, Pila, & Hanington, 2018; Pila, Gordy, et al., 2016; Pila, Peck, & Hanington, 2017; Pila, Tarrabain, et al., 2016), with primer information as shown in Table. S1 and Sf9 cell codon-optimized *Bg*FREP3 Genscript as shown in Fig. S1. Briefly, the coding regions were amplified with Phusion high-fidelity DNA polymerase from targeted DNA templates in pUC57 plasmids (synthesized by GenScript). Blunt-end PCR products were cloned into the pENTR/D-TOPO vectors to generate entry clones. Plasmids from these entry clones were extracted and cloned into pIB/V5-His-DEST vectors in a Clonase recombination reaction to produce the expression clones. Recombinant plasmids were extracted from the expression clones and then transfected into Sf9 insect cells using

Cellfectin reagent (Life Technologies). Cells that survived the screening of the antibiotic blasticidin (Thermo Fisher Scientific) were collected and lysed. Protein expression of r*Bg*FREP3, r*Bg*TEP1 and r*Bg*FREP2 were detected by Western blot using monoclonal antibody against the V5 tag (Thermo Fisher Scientific) on the recombinant proteins.

Sf9 cells stably expressing recombinant r*Bg*FREP3, r*Bg*TEP1 and r*Bg*FREP2 were selected, some of them were frozen as a seed stock and the rest were maintained in a medium containing blasticidin. Cultures of Sf9 cells expressing recombinant proteins were scaled up, and the pellet (containing r*Bg*FREP3) or the medium (containing secreted r*Bg*TEP1 and r*Bg*FREP2) was collected. Recombinant proteins were purified from the cell pellet lysate or the culture medium with HisTrap FF (GE Healthcare) columns using fast protein liquid chromatography (ÄKTA Pure, GE Healthcare). Purified r*Bg*FREP3, r*Bg*TEP1 and r*Bg*FREP2 were dialyzed against PBS buffer twice for 2 h each and then once overnight using Slide-A-Lyzer dialysis kit (Thermo Scientific).

Preparation of *B. glabrata* plasma and pull-down experiment.

Hemolymph was obtained from M-line and BS-90 strain *B. glabrata* snails by the head-foot retraction method (Pila, Gordy, et al., 2016). Upon collection, hemolymph from ~5 snails of each strain (8–12 mm in diameter) was immediately placed in 1.5 mL tubes on ice, EDTA was added to a final concentration of 150 µg/mL. Samples were centrifuged twice at 20,000 × g for 15 min and the cell-free fraction was collected. The 4 M Imidazole Stock Solution was added to a final concentration of 10 mM imidazole. Plasma samples were used immediately for subsequent pull-down experiment.

Pull-down assays were conducted by using Pull-Down PolyHis Protein:Protein Interaction Kit (Thermo Scientific) according to the manufacturer's instructions with

525 some modifications. Briefly, first we prepared the "bait" proteins, r*Bg*FREP3 and
526 r*Bg*TEP1, from previously purified proteins or Sf9 cell lysate expressing r*Bg*FREP3 or
527 concentrated medium containing r*Bg*TEP1. We then loaded the HisPur Cobalt Resin
528 equilibrated for at least 30 minutes at room temperature onto the column according to the
529 kit instructions. The prepared polyhistidine-tagged r*Bg*FREP3 and r*Bg*TEP1 was
530 immobilized on the columns after five times washing with wash solution [1:1 of Tris-
531 buffered saline (TBS) solution:Pierce Lysis Buffer containing 10 mM imidazole]. For cell
532 lysates or concentrated medium containing r*Bg*FREP3 and r*Bg*TEP1, protein expression
533 was previously analyzed by SDS-PAGE and Western blot, and 800 µL was added; for
534 purified r*Bg*FREP3, at least 150 µg (about 500 µL) was added. After incubation for 2 h at
535 room temperature or overnight at 4°C, the columns were washed 5-8 times with wash
536 solution and then 800 µL of prepared plasma (prey) containing 10 mM imidazole was
537 added to the columns. The columns were incubated at 4°C for overnight with gentle
538 rocking motion every few h. After that, the columns were washed 5-8 times with the
539 wash solution, and then eluent was added and incubated at room temperature for 30 min
540 with gentle rocking on a rotating platform.

541 The eluent collected by centrifugation was used for SDS-PAGE analysis and silver
542 staining (Pierce™ Silver Stain for Mass Spectrometry, Thermo Scientific) was performed
543 according to the manufacturer's instructions. After silver staining, we excised the desired
544 bands for LC-MS/MS to identify the proteins.

545 **Mass spectrometry analysis.**

546 Mass spectrometry and protein identification was completed by Alberta Proteomics
547 and Mass Spectrometry Facility at University of Alberta. In-gel trypsin digestion was

548 performed on the samples. Briefly, the excised gel bands were reduced (10 mM
549 mercaptoethanol in 100 mM bicarbonate) and alkylated (55 mM iodoacetamide in 100
550 mM bicarbonate). After dehydration enough trypsin (6 ng/μL, Promega Sequencing
551 grade) was added to just cover the gel pieces and the digestion was allowed to proceed
552 overnight (~16 h) at room temperature. Tryptic peptides were first extracted from the gel
553 using 97% water/2% acetonitrile/1% formic acid followed by a second extraction using
554 50% of the first extraction buffer and 50% acetonitrile.

555 The digested samples containing tryptic peptides were resolved and ionized by using
556 nanoflow HPLC (Easy-nLC II, Thermo Scientific) coupled to an LTQ Orbitrap XL
557 hybrid mass spectrometer (Thermo Scientific). Nanoflow chromatography and
558 electrospray ionization were accomplished by using a PicoFrit fused silica capillary
559 column (ProteoPepII, C18) with 100 μm inner diameter (300 Å, 5 μm, New Objective).
560 The mass spectrometer was operated in data-dependent acquisition mode, recording high-
561 accuracy and high-resolution survey Orbitrap spectra using external mass calibration,
562 with a resolution of 30,000 and m/z range of 400–2,000. The fourteen most intense
563 multiply charged ions were sequentially fragmented by using collision induced
564 dissociation, and spectra of their fragments were recorded in the linear ion trap; after two
565 fragmentations all precursors selected for dissociation were dynamically excluded for 60
566 s. Data was processed using Proteome Discoverer 1.4 (Thermo Scientific) and a *B.*
567 *glabrata* proteome database (UniProt) was searched using SEQUEST (Thermo
568 Scientific). Search parameters included a precursor mass tolerance of 10 ppm and a
569 fragment mass tolerance of 0.8-Da. Peptides were searched with carbamidomethyl

570 cysteine as a static modification and oxidized methionine and deamidated glutamine and
571 asparagine as dynamic modifications.

572 **Sporocyst transformations of *S. mansoni* and immunofluorescence staining.**

573 Miracidia were obtained from eggs isolated from the livers of mice infected with *S.*
574 *mansoni* (NMRI) as described previously (Hambrook et al., 2018). Briefly, newly
575 hatched miracidia were collected for cultivation in Chernin's Balanced Salt Solution
576 (CBSS) (Chernin, 1963) containing 1 g/L each of glucose and trehalose as well as 1%
577 pen-strep antibiotics (Life Technologies). After cultivation for 24 h at 27°C under
578 normoxic conditions, most miracidia transformed to primary sporocysts. Primary
579 sporocysts were washed five times with snail PBS (sPBS) (Yoshino & Laursen, 1995)
580 and transferred to 15 mL centrifuge tubes (Corning). All in-tube washes were performed
581 at 4°C and sporocysts were pelleted by centrifugation for 2 min at 300 × g.

582 After overnight 2% paraformaldehyde/sPBS (pH 7.2) fixation, sporocysts were
583 washed five times with sPBS followed by 1 h incubation in 5% BSA/0.02% azide/sPBS
584 (blocking buffer) at room temperature. Sporocysts was grouped and incubated with
585 r*Bg*FREP3 (500 ng/µL), r*Bg*TEP1 (50 ng/µL) and r*Bg*FREP2 (150 ng/µL) in blocking
586 buffer overnight at 4°C; control group were incubated in blocking buffer without
587 recombinant proteins. Following five washes with 1% BSA/sPBS, sporocysts were
588 treated with anti-V5 antibody (Life Technologies) diluted 1:400 in blocking buffer.
589 Following primary antibody treatment, sporocysts were washed five times with 1%
590 BSA/sPBS and incubated with a solution containing 7.5 U/mL Alexa Fluor 488
591 phalloidin (Invitrogen) and 4 µg/mL Alexa Fluor 633 goat anti-mouse secondary
592 antibody (Invitrogen) in blocking buffer overnight at 4°C. Finally, sporocysts were

washed five times with 1% BSA/sPBS, dropped into a drop of DAPI staining solution, incubated for 5 min at room temperature, and transferred onto microscope slides. Preparations were then covered with a cover glass and sealed with nail polish to prevent desiccation. Slides were imaged using an Axio imager A2 microscope (Zeiss), and analyzed using Zen 2011 software (Zeiss) and Photoshop CS4 (Adobe Systems Inc., USA).

Anti- *Bg*FREP3 polyclonal antibody generation and validation.

The IgSF1 and FBG domains of *Bg*FREP3 were screened for regions of presumed antigenicity using the OptimumAntigen^{TM} design tool (GenScript). Three polyclonal antibodies were generated specific for different peptides (14-15 amino acids long, ~ 4.4-kDa) targeting each domain. The information of these peptides is: anti-IgSF #1: CSFKKDDLSDSKQRS; anti-IgSF #2: CHNKYSEGRIDKSSN; anti-IgSF #3: CNINKDLDFKEQNIT; anti-FBG #1: TTFDRDNDEYSYNC; anti-FBG #2: GWKEYRDGFGDYNIC and anti-FBG #3: CLNGKWGSSDFAKGV. Peptides were synthesized by GenScript and used to immunize rabbits. Rabbits received a primary immunization and two secondary boosts at 2 and 5 weeks. One week after the second boost, rabbits were exsanguinated and the serum IgG was affinity-purified using protein A/G affinity column. These were further purified using affinity resin conjugated with the immunization peptides. The anti-*Bg*FREP3 antibodies used in this study were effective for Western blot detections at a concentration of 1:1,000.

The purified antibodies were supplied by GenScript along with synthesized peptides and pre-immune serum. Specificity of the antibodies was tested against the respective peptides as well as snail plasma and haemocyte lysates. Dot blots were initially used to

616 determine whether the antibodies recognized their peptides as well as the specificity of
617 the recognition. Briefly, 2 µL of each supplied peptide was slowly pipetted in one spot on
618 a nitrocellulose membrane that was divided into a grid with a pencil. The deposited
619 peptides were dried at room temperature for about 10 min, creating a spot of
620 approximately 2–3 mm. The membrane was then incubated in blocking buffer and
621 processed using the same procedure described for Western blot following sample transfer
622 to nitrocellulose membranes. In order to determine that the peptides were not being
623 recognized by any pre-immune serum components, the above procedure was followed
624 exactly with the exception that the respective pre-immune sera were used for the primary
625 incubations in place of antibodies.

626 **Western and far-Western blot analysis.**

627 To test whether r*Bg*TEP1 complexes with r*Bg*FREP3 and r*Bg*FREP2, purified
628 r*Bg*TEP1 was incubated with r*Bg*FREP3 or r*Bg*FREP2 in the following ratios (wt/wt in
629 µg): 1:1, 1:0.75, 1:0.5, and 1:0.25 in 1.5 mL tubes for 2 h at room temperature on a
630 Labquake shaker (ThermoFisher Scientific) to allow potential complexes to form. After
631 the incubation, samples were subjected to western blot detection under denaturing
632 (Laemmli protein loading buffer with β-mercaptoethanol and heated at 95°C for 10 min)
633 and non-denaturing (no β-mercaptoethanol nor heating) conditions, respectively.

634 The Western blot experiment process was as described previously (Pila, Peck, et al.,
635 2017). Samples were loaded on 6-14% (vol/vol) SDS/PAGE gels and run on the Mini
636 PROTEAN Tetra system (Bio-Rad) at 150 V for 1.5-2 h. Samples were then transferred
637 for 1.5 h onto 0.45 µm supported nitrocellulose membranes (Bio-Rad). Blocking was
638 done for 1 h at room temperature in 5% (wt/vol) skimmed milk or bovine serum albumin

(BSA) prepared in TBS solution plus 0.1% Tween-20 (TBS-T buffer) before staining for 1 h in anti-V5 mouse primary antibody at a dilution of 1:5,000 in blocking buffer. Membranes were washed in TBS-T buffer for 10 min, then twice for 5 min each and once in TBS solution for 5 min. Membranes were then incubated for 1 h in HRP-conjugated rabbit anti-mouse antibody diluted 1:5,000 in blocking buffer followed by a wash step as described above. Finally, the blot was developed by incubating the membranes in SuperSignal West Dura Extended Duration substrate (Thermo Scientific). Chemiluminescent signals were acquired on the ImageQuant LAS 4000 machine (GE Healthcare).

To detect the *Bg*FREP3 background expression in the *B. glabrata* snail hemolymph, the hemolymph of four snails was obtained from each of M-line and BS-90 strain *B. glabrata* snails. The hemolymph from each snail was diluted twice with lysis buffer (50 mM Tris, pH 7.8; 150 mM NaCl; 1% Nonidet P-40 and 150 µg/mL EDTA). After incubation for 30 minutes on ice, samples were centrifuged at 20,000 × g for 15 min and the supernatant was collected. After quantification, samples (160 µg/well) were used for subsequent Western blot analysis. The primary antibodies were polyclonal antibodies raised in rabbits against the IgSF and FBG domains of r*Bg*FREP3 (1:500 dilution) and the secondary antibody was a HRP-conjugated goat anti-rabbit IgG (1:5000 dilution).

Far-Western blotting is used to probe a membrane containing transferred protein with another protein to detect specific protein-protein interactions (Y. Wu, Li, & Chen, 2007). In our study, we detected r*Bg*TEP1 (“prey” protein) separated by SDS-PAGE with the biotinylated r*Bg*FREP3 and r*Bg*FREP2 (“bait” or “probe” proteins) and a streptavidin-HRP chemiluminescent detection system. If r*Bg*TEP1 forms a complex with

r*Bg*FREP3 or r*Bg*FREP2, r*Bg*FREP3 would be detected on spots in the membrane where r*Bg*TEP1 was located. First, r*Bg*FREP3 and r*Bg*FREP2 was biotinylated with EZ-Link Sulfo-NHS-LC-Biotin (Thermo Scientific) according to the manufacturer's instructions. Recombinant *Bg*TEP1 was separated by SDS-PAGE, and transferred to a membrane, as in a standard western blot. Recombinant *Bg*TEP1 in the membrane was then denatured and renatured as described previously (Y. Wu et al., 2007). The membrane was then blocked and probed with biotinylated r*Bg*FREP3 or r*Bg*FREP2. After washing as described above in a standard western blot, the membrane was incubated with streptavidin-HRP diluted 1:100,000 in blocking buffer at room temperature for 1 h. Chemiluminescent signals were detected using the ImageQuant LAS 4000 machine (GE Healthcare).

Quantitative real-time PCR analysis of *Bg*TEP1 expression.

A TRIzol reagent was utilized to extract total RNA from 5 whole unchallenged snails from each strain independently, after which RNA was purified using a PureLink RNA mini kit (Life Technologies). A NanoVue spectrophotometer (ThermoFisher Scientific) was used to determine RNA concentration, and first-strand cDNA synthesis using the qScript cDNA synthesis kit (Quanta Biosciences) was performed using 1 µg of RNA template. The cDNA was diluted fivefold, and 5 µL was used as template in RT-PCR using primers (Table S1) specific for *Bg*TEP1 and *B. glabrata* β-actin (*Bg*Actin) in a SYBR Green detection system (PerfeCTa SYBR Green FastMix; Quanta Biosciences). All RT-PCRs were performed on a QuantStudio 3 PCR system (Applied Biosystems) using the following thermocycling conditions: initial hold at 95°C for 10 min, followed by 40 cycles of 95°C for 15 s and 60°C for 1 min, with data collection every cycle.

Specificity for the gene specific primers RT-PCR was confirmed by continuous melt curve analysis.

In vitro S. mansoni sporocyst killing assays.

Haemolymph from 25 M-line or BS-90 strain *B. glabrata,* sterilized as reported in Hahn *et al.*, 2001 (Hahn et al., 2001), was pooled into 1.5 mL tubes and kept on ice. Haemocytes were isolated from plasma by centrifugation at 70 × g for 10 min at 4°C followed by aspiration of the cell free plasma, which was kept for assessment of plasma-mediated killing of sporocysts. Haemocytes were resuspended in modified *Bg*e-cell medium (m*Bg*e) medium [22% Schneider's Drosophila medium (Thermo Scientific), 7 mM D-glucose, 24 mM NaCl, and 20 mg/mL gentamicin at pH 7.4] and washed three times following the same centrifugation protocol. Haemocytes were then resuspended to a final concentration of 50 cells/µL in m*Bg*e medium and 200 µL of this suspension was added to the wells of a 96-well tissue culture plate pre-treated with poly-L-lysine that contained 10 *S. mansoni* sporocysts that were transformed using previously published protocols (Hambrook et al., 2018). Treatment consisted of 200 pM of each recombinant protein (*Bg*FREP3, *Bg*TEP1, *Bg*FREP2). Three anti-*Bg*FREP3 antibodies (anti-IgSF #1 and #3, anti-FBG #1) that were generated to peptides within the IgSF and FBG domains were used as a cocktail (combined concentration of 2.5 mg/mL) were used the antibody blocking assays. Sporocyst viability was assessed using the protocol published by Hahn *et al.*, 2001 (Hahn et al., 2001). Briefly, Propidium iodide (Millipore Sigma), was added to each well to a final concentration of 10 µg/mL. Sporocyst viability was assessed using a fluorescent inverted microscope (Zeiss) starting at 3 h post incubation with haemocytes/plasma and then measured again at 6, 12, 24, 36 and 48 h post incubation.

Sporocyst death was determined by propidium iodide-staining of the nuclei of the sporocyst. Any sporocyst deemed dead, was left in the well until the end of the time-course so that the experimental conditions were not disturbed differentially in wells where sporocyst mortality was more frequent.

The pooled plasma from M-line and BS-90 snails. Cell-free plasma was ultracentrifuged at 30,000 rpm for 3 h at 4°C to remove haemoglobin, which is known to have a toxic effect on sporocysts (Bender et al., 2002). Ultracentrifuged plasma was mixed 50:50 with m*Bg*e medium at room temperature and then incubated with 10 *S. mansoni* sporocysts according to the protocol outlined above.

Assessment of reactive oxygen species production by haemocytes *in vitro*.

To assess the role that ROS had on sporocyst killing by haemocytes the above haemocyte *in vitro* killing assays were repeated with addition of two well characterized inhibitors of ROS; diphenyleneiodonium chloride (DPI) and N-acetyl-L-cysteine (NAC). Both ROS inhibitors (DPI at a final concentration of 150 nM and NAC at a final concentration of 1 mM) were incubated along with haemocytes and *S. mansoni* sporocysts as described above in the haemocyte-based sporocyst killing assay.

Bioinformatic and Statistical analysis

We used Clustal Omega (https://www.ebi.ac.uk/Tools/msa/clustalo/) to align the amino acid sequences for *Bg*TEPs, Biomphalysins and *Bg*FREPs which acquired from GenBank [National Center for Biotechnology Information (NCBI)]. The alignment files were downloaded and visually inspected for obvious inaccuracies, modified if need be, and then annotated based on information provided by NCBI. To show the distribution of the identified peptides over a full length sequence, the amino acid sequences of

731 *Bg*TEP1.5 (ADE45333.1), Biomphalysin (A0A182YTN9) and *Bg*MFREP2
732 (AAK13550.1) were imported into Vector NTI Advance 11.5 (Thermo Fisher Scientific)
733 and the identified peptides were annotated. Images produced by Vector NTI software
734 were saved by FSCapture and further processed with FSCapture and Adobe Photoshop
735 CS4. To determine significant differences in sporocysts killing assays, one-way ANOVA
736 with Tukey's post hoc tests were performed using GraphPad Prism version 6.0f for Mac
737 OS X (GraphPad; www.graphpad.com). Statistical significance threshold was set at $P \leq$
738 0.05.

739 **Acknowledgements**

740 *B. glabrata* snails provided by the NIAID Schistosomiasis Resource Center of the
741 Biomedical Research Institute (Rockville, MD) through NIH-NIAID Contract
742 HHSN272201700014I for distribution through BEI Resources. These studies were
743 supported by funds provided by the Natural Sciences and Engineering Research Council
744 of Canada #2018-05209 and 2018- 522661 (PCH), and National Natural Science
745 Foundation of China (NSFC 31272682), Guangxi 16 Natural Science Foundation (Key
746 Project, 2016JJD130059) Special Fund for Team Building, Beibu Gulf University (
747 Former Qinzhou University, 2015).

748
749 **Author Contributions**

750 H.L. and P.C.H. designed research; H.L., J.R.H., E.A.P., A.A.G., and J.F. performed
751 research; H.L., J.R.H., E.A.P., X.W. and P.C.H. analyzed data; and H.L., J.R.H., E.A.P.,
752 A.A.G., and P.C.H. wrote the paper.

753
754 **Declaration of interests:** The authors declare no competing interests.

References

Adema, C. M. (2015). Fibrinogen-Related Proteins (FREPs) in Mollusks. *Results Probl Cell Differ, 57*, 111-129. doi: 10.1007/978-3-319-20819-0_5

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*(16), 8691-8696.

Asehnoune, K., Strassheim, D., Mitra, S., Kim, J. Y., & Abraham, E. (2004). Involvement of reactive oxygen species in Toll-like receptor 4-dependent activation of NF-kappa B. *J Immunol, 172*(4), 2522-2529. doi: 10.4049/jimmunol.172.4.2522

Baeza Garcia, A., Pierce, R. J., Gourbal, B., Werkmeister, E., Colinet, D., Reichhart, J. M., . . . Coustau, C. (2010). Involvement of the cytokine MIF in the snail host immune response to the parasite Schistosoma mansoni. *PLoS Pathog, 6*(9), e1001115. doi: 10.1371/journal.ppat.1001115

Baxter, R. H., Chang, C. I., Chelliah, Y., Blandin, S., Levashina, E. A., & Deisenhofer, J. (2007). Structural basis for conserved complement factor-like function in the antimalarial protein TEP1. *Proc Natl Acad Sci U S A, 104*(28), 11615-11620. doi: 10.1073/pnas.0704967104

Bayne, C. J., Buckley, P. M., & DeWan, P. C. (1980a). Macrophagelike hemocytes of resistant Biomphalaria glabrata are cytotoxic for sporocysts of Schistosoma mansoni in vitro. *J Parasitol, 66*(3), 413-419.

Bayne, C. J., Buckley, P. M., & DeWan, P. C. (1980b). Schistosoma mansoni: cytotoxicity of hemocytes from susceptible snail hosts for sporocysts in plasma from resistant Biomphalaria glabrata. *Exp Parasitol, 50*(3), 409-416.

Bender, R. C., Bixler, L. M., Lerner, J. P., & Bayne, C. J. (2002). Schistosoma mansoni sporocysts in culture: host plasma hemoglobin contributes to in vitro oxidative stress. *J Parasitol, 88*(1), 14-18. doi: 10.1645/0022-3395(2002)088[0014:SMSICH]2.0.CO;2

Bender, R. C., Broderick, E. J., Goodall, C. P., & Bayne, C. J. (2005). Respiratory burst of Biomphalaria glabrata hemocytes: Schistosoma mansoni-resistant snails produce more extracellular H2O2 than susceptible snails. *J Parasitol, 91*(2), 275-279. doi: 10.1645/GE-415R

Bender, R. C., Fryer, S. E., & Bayne, C. J. (1992). Proteinase inhibitory activity in the plasma of a mollusc: evidence for the presence of alpha-macroglobulin in Biomphalaria glabrata. *Comp Biochem Physiol B, 102*(4), 821-824.

Bender, R. C., Goodall, C. P., Blouin, M. S., & Bayne, C. J. (2007). Variation in expression of Biomphalaria glabrata SOD1: a potential controlling factor in susceptibility/resistance to Schistosoma mansoni. *Dev Comp Immunol, 31*(9), 874-878. doi: 10.1016/j.dci.2006.12.005

Blandin, S. A., Marois, E., & Levashina, E. A. (2008). Antimalarial responses in Anopheles gambiae: from a complement-like protein to a

complement-like pathway. *Cell Host Microbe, 3*(6), 364-374. doi: 10.1016/j.chom.2008.05.007

Chernin, E. (1963). Observations on hearts explanted in vitro from the snail Australorbis glabratus. *J Parasitol, 49*, 353-364.

Cody, J. J., Ittiprasert, W., Miller, A. N., Henein, L., Mentink-Kane, M. M., & Hsieh, M. H. (2016). The NIH-NIAID Schistosomiasis Resource Center at the Biomedical Research Institute: Molecular Redux. *PLoS Negl Trop Dis, 10*(10), e0005022. doi: 10.1371/journal.pntd.0005022

Dinguirard, N., Cavalcanti, M. G. S., Wu, X. J., Bickham-Wright, U., Sabat, G., & Yoshino, T. P. (2018). Proteomic Analysis of Biomphalaria glabrata Hemocytes During in vitro Encapsulation of Schistosoma mansoni Sporocysts. *Front Immunol, 9*, 2773. doi: 10.3389/fimmu.2018.02773

Disease, G. B. D., Injury, I., & Prevalence, C. (2016). Global, regional, and national incidence, prevalence, and years lived with disability for 310 diseases and injuries, 1990-2015: a systematic analysis for the Global Burden of Disease Study 2015. *Lancet, 388*(10053), 1545-1602. doi: 10.1016/S0140-6736(16)31678-6

Esch, G. W., Barger, M. A., & Fellis, K. J. (2002). The transmission of digenetic trematodes: style, elegance, complexity. *Integr Comp Biol, 42*(2), 304-312. doi: 10.1093/icb/42.2.304

Galinier, R., Portela, J., Mone, Y., Allienne, J. F., Henri, H., Delbecq, S., . . . Duval, D. (2013). Biomphalysin, a new beta pore-forming toxin involved in Biomphalaria glabrata immune defense against Schistosoma mansoni. *PLoS Pathog, 9*(3), e1003216. doi: 10.1371/journal.ppat.1003216

Goodall, C. P., Bender, R. C., Broderick, E. J., & Bayne, C. J. (2004). Constitutive differences in Cu/Zn superoxide dismutase mRNA levels and activity in hemocytes of Biomphalaria glabrata (Mollusca) that are either susceptible or resistant to Schistosoma mansoni (Trematoda). *Mol Biochem Parasitol, 137*(2), 321-328. doi: 10.1016/j.molbiopara.2004.06.011

Gorbushin, A. M. (2019). Derivatives of the lectin complement pathway in Lophotrochozoa. *Dev Comp Immunol, 94*, 35-58. doi: 10.1016/j.dci.2019.01.010

Gordy, M. A., Pila, E. A., & Hanington, P. C. (2015). The role of fibrinogen-related proteins in the gastropod immune response. *Fish Shellfish Immunol, 46*(1), 39-49. doi: 10.1016/j.fsi.2015.03.005

Granath, W. O., Jr., & Yoshino, T. P. (1984). Schistosoma mansoni: passive transfer of resistance by serum in the vector snail, Biomphalaria glabrata. *Exp Parasitol, 58*(2), 188-193.

Hahn, U. K., Bender, R. C., & Bayne, C. J. (2000). Production of reactive oxygen species by hemocytes of Biomphalaria glabrata: carbohydrate-specific stimulation. *Dev Comp Immunol, 24*(6-7), 531-541.

Hahn, U. K., Bender, R. C., & Bayne, C. J. (2001). Killing of Schistosoma mansoni sporocysts by hemocytes from resistant Biomphalaria glabrata: role of reactive oxygen species. *J Parasitol, 87*(2), 292-299. doi: 10.1645/0022-3395(2001)087[0292:KOSMSB]2.0.CO;2

Hambrook, J. R., Kabore, A. L., Pila, E. A., & Hanington, P. C. (2018). A

metalloprotease produced by larval Schistosoma mansoni facilitates infection establishment and maintenance in the snail host by interfering with immune cell function. *PLoS Pathog, 14*(10), e1007393. doi: 10.1371/journal.ppat.1007393

Hanington, P. C., Forys, M. A., Dragoo, J. W., Zhang, S. M., Adema, C. M., & Loker, E. S. (2010). Role for a somatically diversified lectin in resistance of an invertebrate to parasite infection. *Proc Natl Acad Sci U S A, 107*(49), 21087-21092. doi: 10.1073/pnas.1011242107

Hanington, P. C., Forys, M. A., & Loker, E. S. (2012). A somatically diversified defense factor, FREP3, is a determinant of snail resistance to schistosome infection. *PLoS Negl Trop Dis, 6*(3), e1591. doi: 10.1371/journal.pntd.0001591

Hanington, P. C., & Zhang, S. M. (2011). The primary role of fibrinogen-related proteins in invertebrates is defense, not coagulation. *J Innate Immun, 3*(1), 17-27. doi: 10.1159/000321882

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*(4), 295-303.

Howard, S. P., & Buckley, J. T. (1985). Activation of the hole-forming toxin aerolysin by extracellular processing. *J Bacteriol, 163*(1), 336-340.

Leonard, P. M., Adema, C. M., Zhang, S. M., & Loker, E. S. (2001). 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, 269*(1-2), 155-165. doi: 10.1016/s0378-1119(01)00444-9

Levashina, E. A., Moita, L. F., Blandin, S., Vriend, G., Lagueux, M., & Kafatos, F. C. (2001). Conserved role of a complement-like protein in phagocytosis revealed by dsRNA knockout in cultured cells of the mosquito, Anopheles gambiae. *Cell, 104*(5), 709-718.

Lie, K. J., & Heyneman, D. (1975). Studies on resistance in snails: a specific tissue reaction to Echinostoma lindoense in Biomphalaria glabrata snails. *Int J Parasitol, 5*(6), 621-625.

Lie, K. J., Heyneman, D., & Lim, H. K. (1975). Studies on resistance in snails: specific resistance induced by irradiated miracidia of Echinostoma lindoense in Biomphalaria glabrata snails. *Int J Parasitol, 5*(6), 627-631.

Lockyer, A. E., Emery, A. M., Kane, R. A., Walker, A. J., Mayer, C. D., Mitta, G., . . . Jones, C. S. (2012). Early differential gene expression in haemocytes from resistant and susceptible Biomphalaria glabrata strains in response to Schistosoma mansoni. *PLoS One, 7*(12), e51102. doi: 10.1371/journal.pone.0051102

Lockyer, A. E., Spinks, J., Kane, R. A., Hoffmann, K. F., Fitzpatrick, J. M., Rollinson, D., . . . Jones, C. S. (2008). Biomphalaria glabrata transcriptome: cDNA microarray profiling identifies resistant- and susceptible-specific gene expression in haemocytes from snail strains exposed to Schistosoma mansoni. *BMC Genomics, 9*, 634. doi: 10.1186/1471-2164-9-634

Loker, E. S., & Bayne, C. J. (1982). In vitro encounters between Schistosoma

mansoni primary sporocysts and hemolymph components of susceptible and resistant strains of Biomphalaria glabrata. *Am J Trop Med Hyg, 31*(5), 999-1005. doi: 10.4269/ajtmh.1982.31.999

MacKenzie, C. R., Hirama, T., & Buckley, J. T. (1999). Analysis of receptor binding by the channel-forming toxin aerolysin using surface plasmon resonance. *J Biol Chem, 274*(32), 22604-22609. doi: 10.1074/jbc.274.32.22604

Mone, Y., Gourbal, B., Duval, D., Du Pasquier, L., Kieffer-Jaquinod, S., & Mitta, G. (2010). A large repertoire of parasite epitopes matched by a large repertoire of host immune receptors in an invertebrate host/parasite model. *PLoS Negl Trop Dis, 4*(9). doi: 10.1371/journal.pntd.0000813

Mortality, G. B. D., & Causes of Death, C. (2016). Global, regional, and national life expectancy, all-cause mortality, and cause-specific mortality for 249 causes of death, 1980-2015: a systematic analysis for the Global Burden of Disease Study 2015. *Lancet, 388*(10053), 1459-1544. doi: 10.1016/S0140-6736(16)31012-1

Newton, W. L. (1955). The establishment of a strain of Australorbis glabratus which combines albinism and high susceptibility to infection with Schistosoma mansoni. *J Parasitol, 41*(5), 526-528.

Pila, E. A., Gordy, M. A., Phillips, V. K., Kabore, A. L., Rudko, S. P., & Hanington, P. C. (2016). Endogenous growth factor stimulation of hemocyte proliferation induces resistance to Schistosoma mansoni challenge in the snail host. *Proc Natl Acad Sci U S A, 113*(19), 5305-5310. doi: 10.1073/pnas.1521239113

Pila, E. A., Li, H., Hambrook, J. R., Wu, X., & Hanington, P. C. (2017). Schistosomiasis from a Snail's Perspective: Advances in Snail Immunity. *Trends Parasitol, 33*(11), 845-857. doi: 10.1016/j.pt.2017.07.006

Pila, E. A., Peck, S. J., & Hanington, P. C. (2017). The protein pheromone temptin is an attractant of the gastropod Biomphalaria glabrata. *J Comp Physiol A Neuroethol Sens Neural Behav Physiol, 203*(10), 855-866. doi: 10.1007/s00359-017-1198-0

Pila, E. A., Tarrabain, M., Kabore, A. L., & Hanington, P. C. (2016). A Novel Toll-Like Receptor (TLR) Influences Compatibility between the Gastropod Biomphalaria glabrata, and the Digenean Trematode Schistosoma mansoni. *PLoS Pathog, 12*(3), e1005513. doi: 10.1371/journal.ppat.1005513

Portet, A., Galinier, R., Pinaud, S., Portela, J., Nowacki, F., Gourbal, B., & Duval, D. (2018). BgTEP: An Antiprotease Involved in Innate Immune Sensing in Biomphalaria glabrata. *Front Immunol, 9*, 1206. doi: 10.3389/fimmu.2018.01206

Povelones, M., Bhagavatula, L., Yassine, H., Tan, L. A., Upton, L. M., Osta, M. A., & Christophides, G. K. (2013). The CLIP-domain serine protease homolog SPCLIP1 regulates complement recruitment to microbial surfaces in the malaria mosquito Anopheles gambiae. *PLoS Pathog, 9*(9), e1003623. doi: 10.1371/journal.ppat.1003623

Richards, C. S. (1975a). Genetic factors in susceptibility of Biomphalaria glabrata for different strains of Schistosoma mansoni. *Parasitology, 70*(2), 231-241.

Richards, C. S. (1975b). Genetic studies of pathologic conditions and susceptibility to infection in Biomphalaria glabrata. *Ann N Y Acad Sci, 266*, 394-410.

Sullivan, J. T., Richards, C. S., Joe, L. K., & Heyneman, D. (1982). Ribeiroia marini: irradiated miracidia and induction of acquired resistance in Biomphalaria glabrata. *Exp Parasitol, 53*(1), 17-25.

Tsung, A., Klune, J. R., Zhang, X., Jeyabalan, G., Cao, Z., Peng, X., . . . Billiar, T. R. (2007). HMGB1 release induced by liver ischemia involves Toll-like receptor 4 dependent reactive oxygen species production and calcium-mediated signaling. *J Exp Med, 204*(12), 2913-2923. doi: 10.1084/jem.20070247

Vasquez, R. E., & Sullivan, J. T. (2001). Further characterization of passively transferred resistance to Schistosoma mansoni in the snail intermediate host Biomphalaria glabrata. *J Parasitol, 87*(6), 1360-1365. doi: 10.1645/0022-3395(2001)087[1360:FCOPTR]2.0.CO;2

Wilmsen, H. U., Leonard, K. R., Tichelaar, W., Buckley, J. T., & Pattus, F. (1992). The aerolysin membrane channel is formed by heptamerization of the monomer. *EMBO J, 11*(7), 2457-2463.

Wong, S. W., Kwon, M. J., Choi, A. M., Kim, H. P., Nakahira, K., & Hwang, D. H. (2009). Fatty acids modulate Toll-like receptor 4 activation through regulation of receptor dimerization and recruitment into lipid rafts in a reactive oxygen species-dependent manner. *J Biol Chem, 284*(40), 27384-27392. doi: 10.1074/jbc.M109.044065

Wu, X. J., Dinguirard, N., Sabat, G., Lui, H. D., Gonzalez, L., Gehring, M., . . . Yoshino, T. P. (2017). Proteomic analysis of Biomphalaria glabrata plasma proteins with binding affinity to those expressed by early developing larval Schistosoma mansoni. *PLoS Pathog, 13*(5), e1006081. doi: 10.1371/journal.ppat.1006081

Wu, Y., Li, Q., & Chen, X. Z. (2007). Detecting protein-protein interactions by Far western blotting. *Nat Protoc, 2*(12), 3278-3284. doi: 10.1038/nprot.2007.459

Xu, C., Wang, B. C., Yu, Z., & Sun, M. (2014). Structural insights into Bacillus thuringiensis Cry, Cyt and parasporin toxins. *Toxins (Basel), 6*(9), 2732-2770. doi: 10.3390/toxins6092732

Yassine, H., & Osta, M. A. (2010). Anopheles gambiae innate immunity. *Cell Microbiol, 12*(1), 1-9. doi: 10.1111/j.1462-5822.2009.01388.x

Yoshino, T. P., & Laursen, J. R. (1995). Production of Schistosoma mansoni daughter sporocysts from mother sporocysts maintained in synxenic culture with Biomphalaria glabrata embryonic (Bge) cells. *J Parasitol, 81*(5), 714-722.

Zhang, S. M., Adema, C. M., Kepler, T. B., & Loker, E. S. (2004). Diversification of Ig superfamily genes in an invertebrate. *Science, 305*(5681), 251-254. doi: 10.1126/science.1088069

Zhang, S. M., 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*(8), 684-694. doi: 10.1007/s00251-001-0386-8

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*(3), 175-189. doi: 10.1177/1753425908093800

983 **Figure titles and legends**

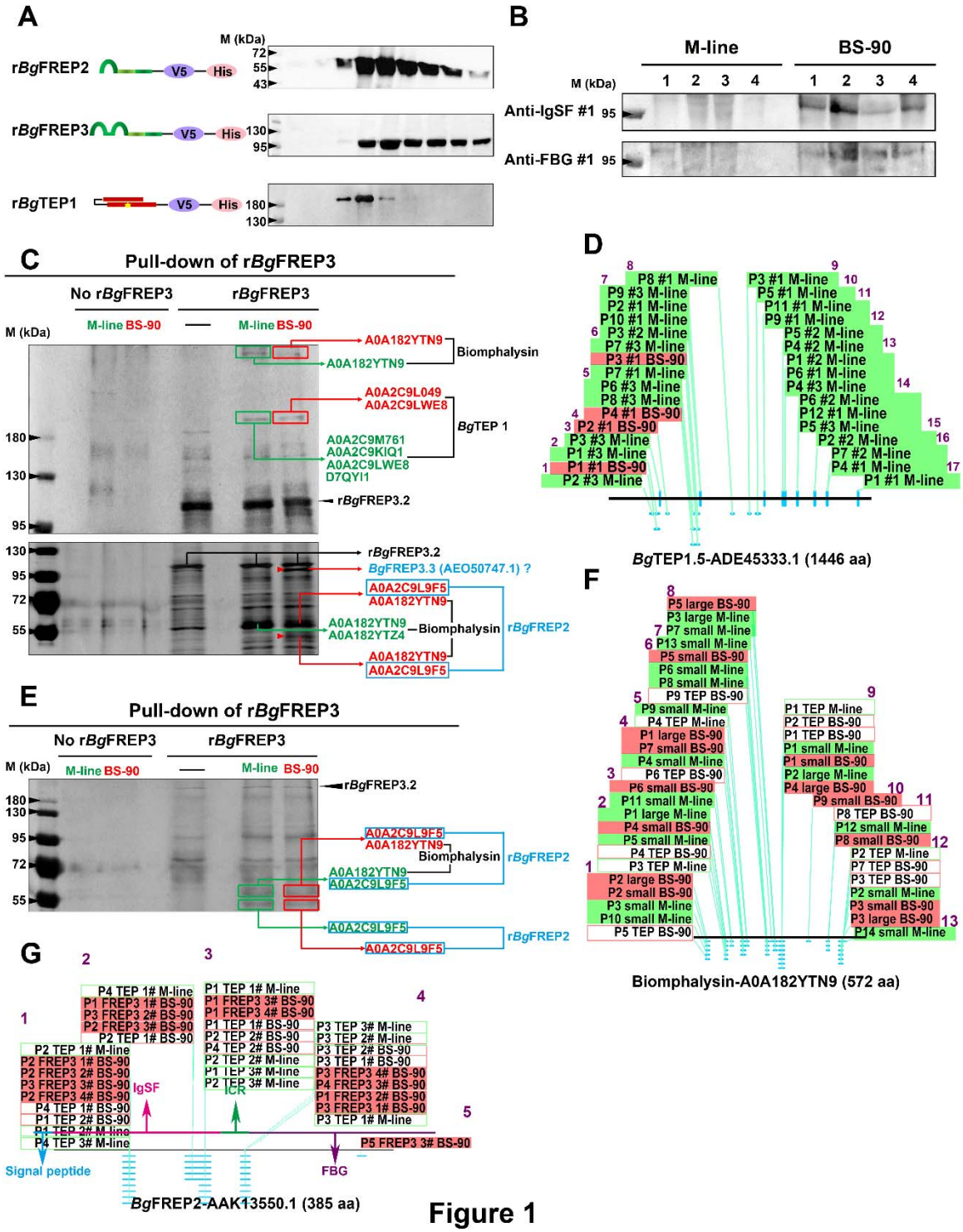

Figure 1

984
985

Figure 1. Pull-down experiments of r*Bg*FREP3 and r*Bg*TEP1. A: Schematic diagram and Western blot analysis of recombinant proteins, the ends of which are coupled to the tags V5 and ploy-His for protein immunoassay and purification operations. **B:** Western blot analysis of *Bg*FREP3 background expression in hemolymph of M-line and BS-90 snails. The hemolymph of four snails from each M-line and BS-90 strain was used for detection, and anti-IgSF #1 and anti-FBG #1 antibodies were used as primary antibodies. **C:** Pull-down experiments with r*Bg*FREP3. The eluate was collected and separated on 6% (upper panel) and 10% (lower panal) SDS-PAGE before silver staining. **E:** Pull-down experiments with r*Bg*TEP1. The eluate was separated on 10% SDS-PAGE. In the pull-down experiment, cell-free plasma from susceptible (green) or resistant (red) *B. glabrata* snails was incubated with Sf9 cell lysates expressing r*Bg*FREP3 or concentrated medium containing r*Bg*TEP1 together (interactome experiments, last two lanes) or alone (controls, first three lanes). Bands that differ between control and interactome experiments were cut, proteins submitted to tryptic digest and analyzed by LC-MS/MS for identification. In interactome experiments, bands that differ between M-line and BS-90 were marked with red arrows. The results of mass spectrometry analysis are shown in the figure. **D, F and G:** Distribution of peptides identified by LC-MS/MS on *Bg*TEP1.5 (ADE45333.1), Biomphalysin (A0A182YTN9) and *Bg*FREP2 (AAK13550.1). Green background: from M-line strain in the r*Bg*FREP3-pull-down; red background: from BS-90 strain in the r*Bg*FREP3-pull-down; green box: from M-line strain in the r*Bg*TEP1-pull-down; red box: from BS-90 strain in the r*Bg*TEP1-pull-down.

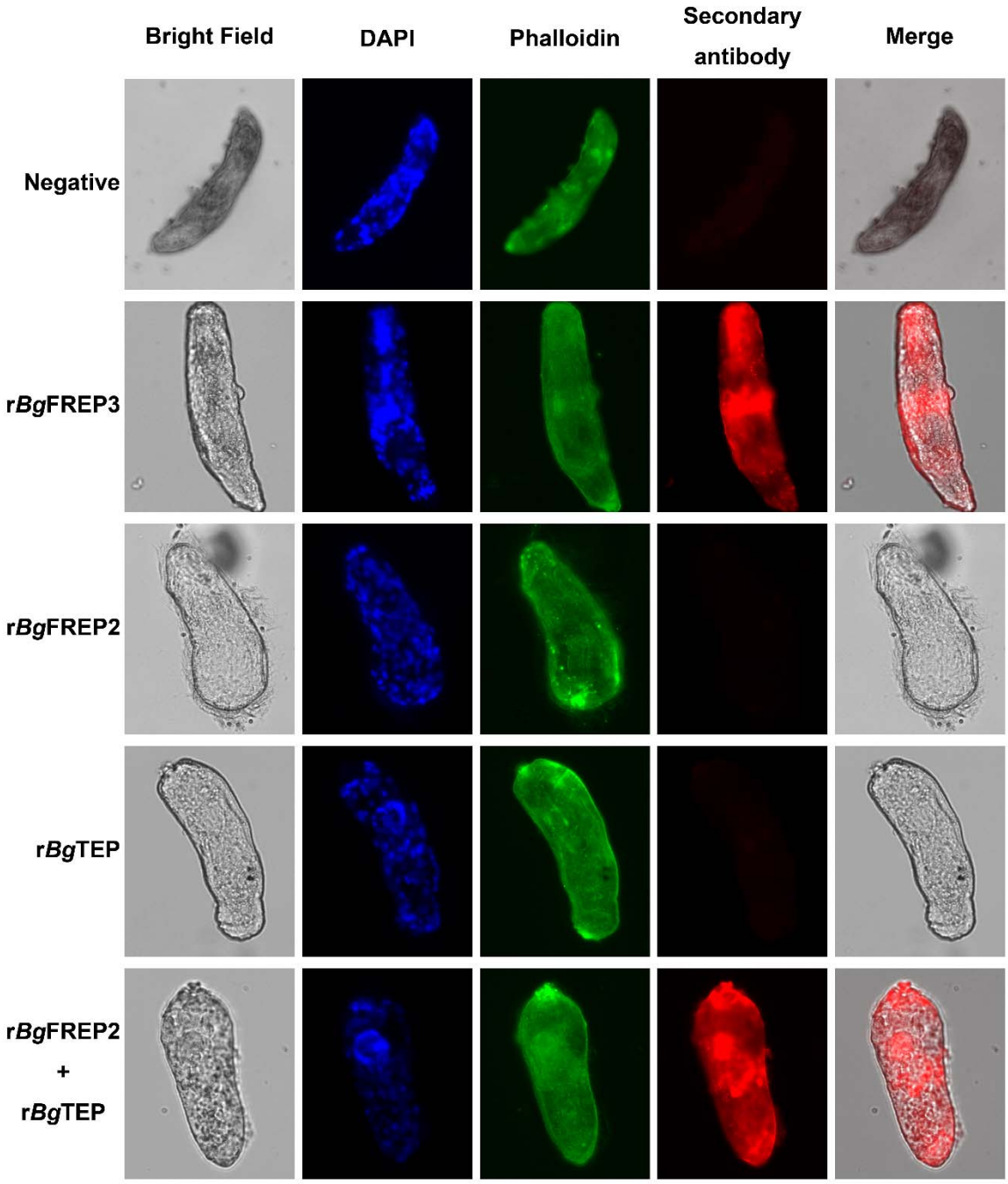

Figure 2

Figure 2. Immunofluorescence analysis of interaction between r*Bg*FREP3, r*Bg*FREP2, and r*Bg*TEP1 with *S. mansoni* sporocysts. Sporocysts were incubated with vehicle or recombinant proteins (r*Bg*FREP3, r*Bg*FREP2, and r*Bg*TEP1 alone) and

combination (r*Bg*FREP2 plus r*Bg*TEP1) and immunostained using anti-V5 primary IgG and Alexa Fluor 633 goat anti-mouse secondary antibody. The sporocysts were stained with DAPI (for nucleus, blue), phalloidin (for F-actin labeling, green) and fluorescent secondary antibody (for recombinant proteins, red).

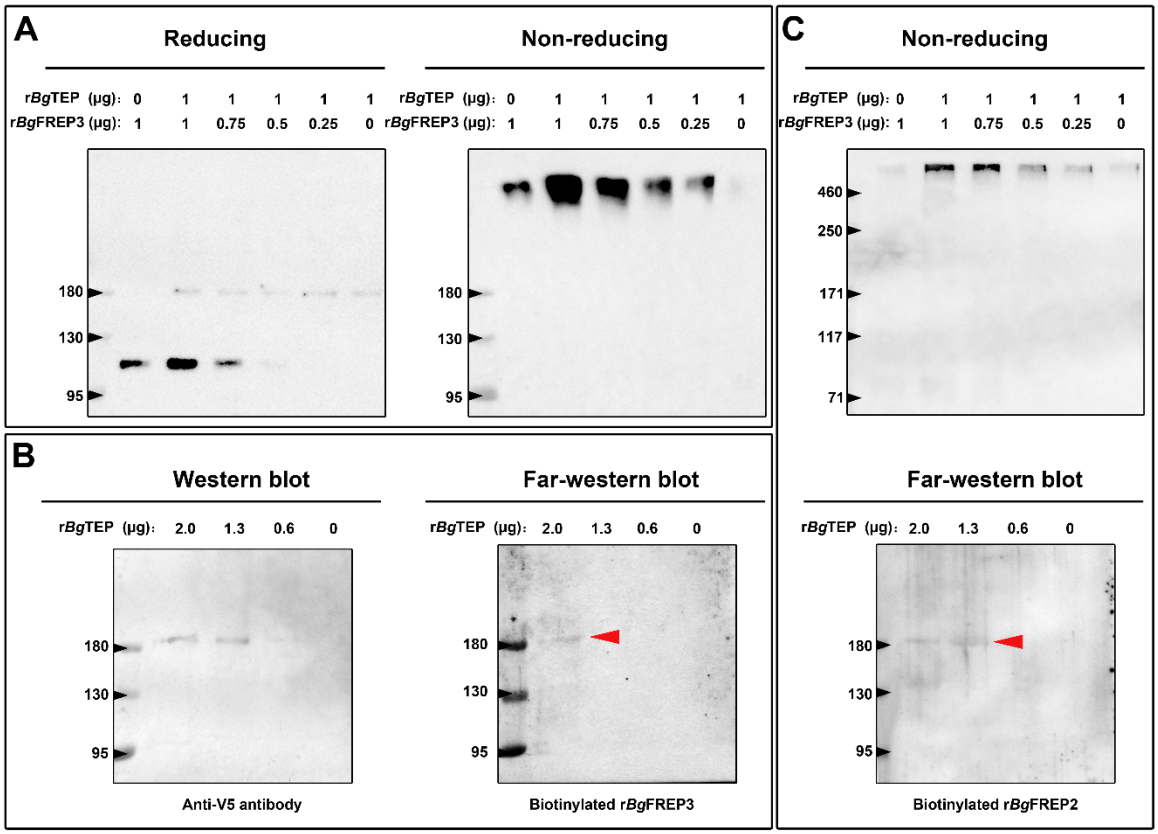

Figure 3

Figure 3. Immunoblot analysis of the interaction between r*Bg*TEP1 and r*Bg*FREP3 and r*Bg*FREP2. A: Purified r*Bg*TEP1 was incubated with r*Bg*FREP3 in the indicated ratios (wt/wt in μg) for 2 h at room temperature. After the incubation, Western blot analysis was carried out under denaturing (heating with reducing agent, the right) and non-denaturing (no heating, no reducing agent, the left) conditions, respectively. **B:** Decreasing amount of r*Bg*TEP1 was blotted onto the membrane and then subjected to standard Western blot (the left) and far-Western blot (the right) analysis. **C:** The interaction between r*Bg*FREP2 and r*Bg*TEP1. Western blot analysis under non-denaturing (the upper) and far-Western blot analysis (the below) was carried out,

1028 respectively. In the far-Western blot, the biotinylated r*Bg*FREP3 and r*Bg*FREP2 acts as a
1029 primary antibody and the secondary antibody is a streptavidin-HRP; the red arrow
1030 indicates the band that symbolizes the interaction between *Bg*TEP1 and r*Bg*FREP3 or
1031 r*Bg*FREP2.
1032
1033

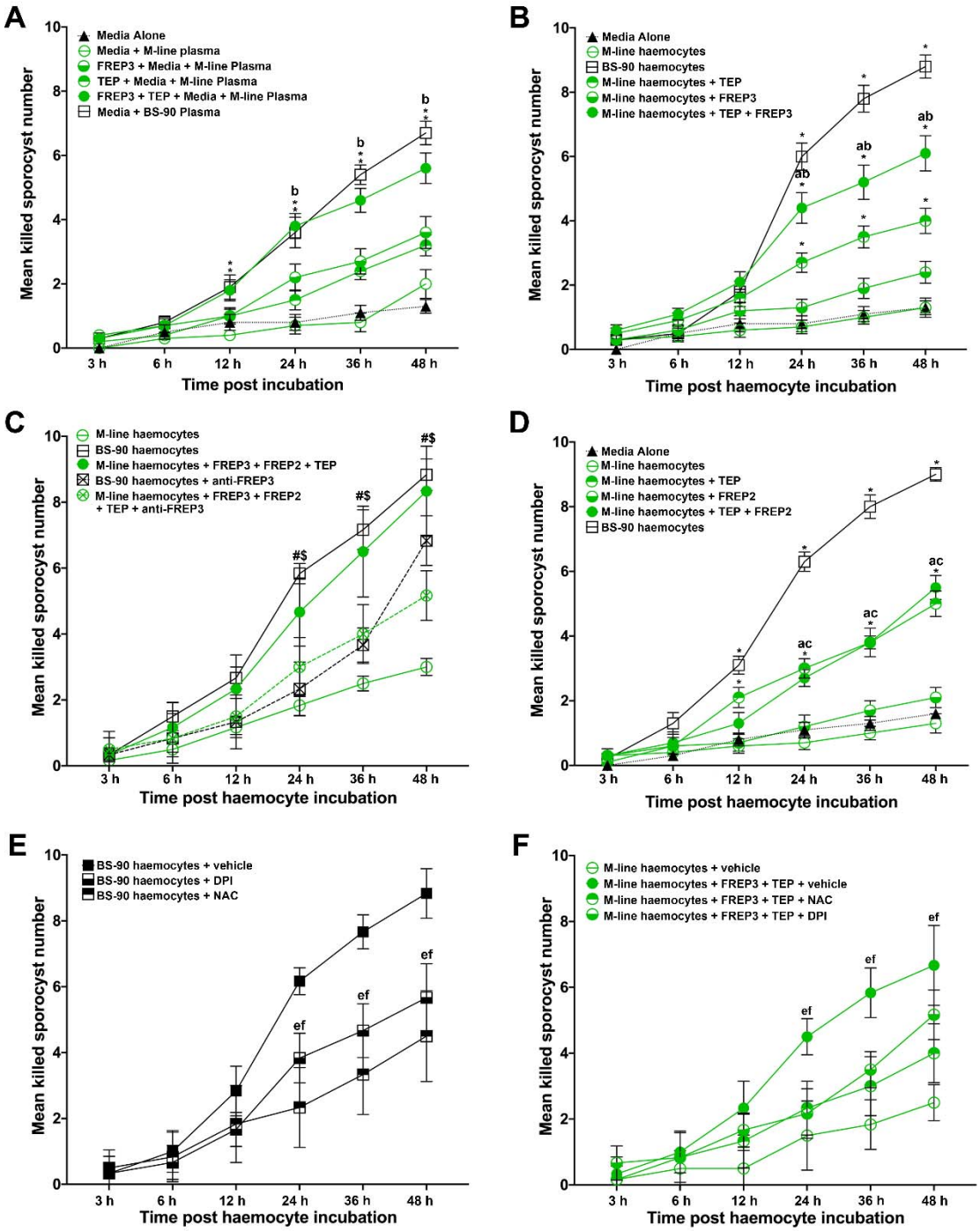

Figure 4

Figure 4. Association of r*Bg*FREP3, r*Bg*TEP1 and r*Bg*FREP2 facilitates M-line snails to kill *S. mansoni* sporocytes. Effects of r*Bg*FREP3, r*Bg*TEP1 and r*Bg*FREP3-

r*Bg*TEP1 complex on killing of *S. mansoni* sporocysts by M-line plasma (**A**) and haemocytes (**B**). **C:** Before being exposed to sporocysts, M-line haemocytes were pre-incubated with r*Bg*FREP3-r*Bg*TEP1-r*Bg*FREP2 combination in presence or in absence of anti- *Bg*FREP3 antibodies (abrogation treatment). BS-90 haemocytes were incubated with anti-*Bg*FREP3 antibodies. **D:** The effect of r*Bg*FREP2, r*Bg*TEP1 and r*Bg*FREP2-r*Bg*TEP1 complex on the destruction of sporocysts by M-line haemocytes. **E and F:** The role of ROS in killing of *S. mansoni* sporocysts by haemocytes from *B. glabrata* snails. **E:** The addition of ROS inhibitors NAC and DPI abolished the ability of BS-90 haemocytes to destroy sporocysts. **F:** Killing sporocysts of M-line haemocytes rendered by r*Bg*FREP3-r*Bg*TEP1 complex was annulled by pre-incubation with ROS inhibitors NAC and DPI. Statistically significant difference symbols in **A**, **B** and **D**: asterisk (*) represents comparison with M-line haemocytes or plasma, * $P < 0.05$, ** $P < 0.01$; a represents comparison with BS-90 haemocytes or plasma, $P < 0.05$; b represents comparison with both single treatment (r*Bg*TEP1/r*Bg*FREP3 or r*Bg*FREP2), $P < 0.05$; c represents comparison with only *Bg*FREPs treatment, $P < 0.05$; bars represent SEM, n=10. Statistically significant difference symbols in **C**: # represents comparison BS-90 haemocytes with BS-90 haemocytes+anti-r*Bg*FREP3 antibodies, $P < 0.05$; $ represents comparison M-line haemocytes+r*Bg*FREP3+r*Bg*FREP2+r*Bg*TEP1 with M-line haemocytes+r*Bg*FREP3+r*Bg*FREP2+r*Bg*TEP1+anti-r*Bg*FREP3 antibodies, $P < 0.05$; bars represent SEM, n=10. Statistically significant difference symbols in **E** and **F:** e represents comparison NAC treatment with BS-90 haemocytes or M-line haemocytes+r*Bg*FREP3+r*Bg*TEP1, $P < 0.05$; f represents comparison DPI treatment with

1059 BS-90 haemocytes or M-line haemocytes+r*Bg*FREP3+r*Bg*TEP1, $P < 0.05$; bars represent

1060 SEM, n=6.

1061

1062

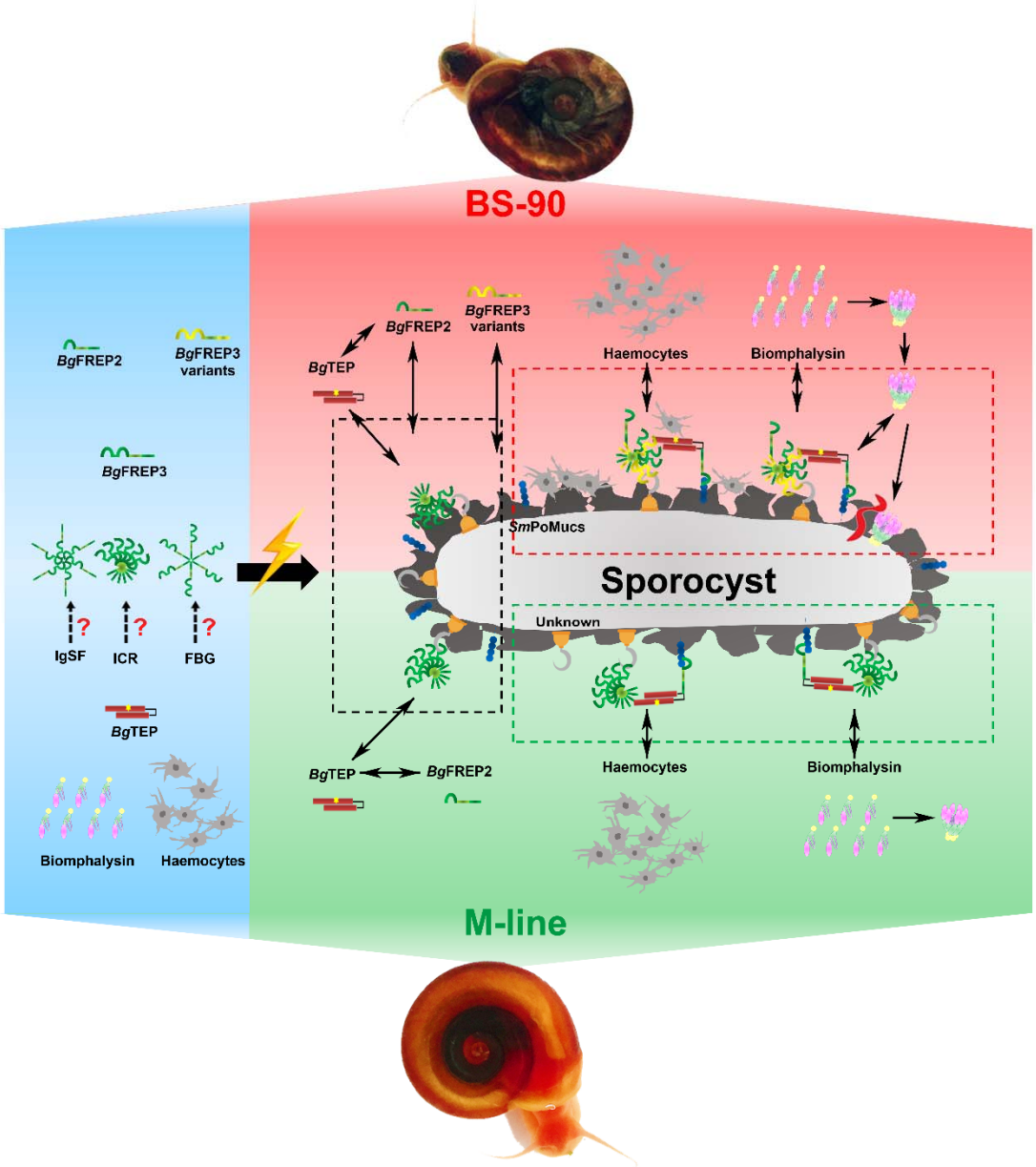

Figure 5

Figure 5. Interaction of *Bg*FREP3, *Bg*FREP2, *Bg*TEP1 and Biomphalysin in *B. glabrata* snail hemolymph and possible functional mechanisms against *S. mansoni* infection. This figure is a schematic drawing based on the results of previous studies and our work.

Supplemental figure titles and legends

>Derived from GenBank AY028461.1 *Biomphalaria glabrata* fibrinogen-related protein 3-2 precursor, gene, complete cds.

ATGGCTCGTCTCTTCTTGCTCTTCATTCTCTGCGTCTTCGTGGTGTCACTCGCTGGCTCGGA
ACTCGTCATTGATGTCCAACCCAACGTCATTTCCCCGGAAATCACCCCTCAGCTGGTCATC
AACTGCAGCATCACGAACAATGAGGTCCAAACACTGGACCTCATCAAGTCTTTGACCCTG
TCAAGGTACAACGAGACTATTCGCGAGTTCGATGAATTGATCGCCCTGGACAGCCTCACCT
TGAACCTGAAGCAGTTCGTGCGCTTCAAGTACTCGCAAATCAGTTTCGGCAACCGTTACAT
CACGCTCATTTTGCACAATCCCACACAGTTCGACGCCCGCATCTACAAGTGCAATGCTACG
GGCACAAACTCCGAGGGTGCCAATATCAGCCTGTTCGCCAAGAAGGCTGTGGAGTACGAA
ACCAACTCTACTGCTCTGATCGAGGAAATTCGCCGTATCAAGAAGGATGAGAACTACTGCA
GCTTCAAGAAGGACGATCTGTCTGACTCAAAGCAGCGTTCTAGGGTGTACTTCTCAGGTT
CCAGCGATATCATTAAGGAGCGTATCGAACCCCTGACGCTCAAGTGCACATTCCAGGTCCT
GAAGACCGACCAAAACGAAACTTCCAGGCTCCAGAGCTTGTACATTCTGCACGAGTCCAA
GGGAGTCATCGCCTACGTGAACAAGGATCAGCCAGTGGTCACCTCGCTGCAGGGCAGTAA
CATCCAAGACGTGGAGGGTGAAATTTACGATAATGCCATCAAGGACTCGTACTTGCAGGTC
ACTTGGAGTAACCTGAAGCACTCCGAGAGCGGAAAGTACTTCTGCGAAGCTCATAACAAG
TACTCGGAGGGCAGGATTGATAAGTCGAGTAATATGCTGACCATCACTGTCGAACGCCCCA
CCTTCGACGATCTCGTGGAGGCTATGCACAAGTTGTTCACTCAGGTGGACGGTGCCAAGG
AGTCCCTGAAGGCTATCAACCAAAACATCAAGAACATCAATAAGGACCTGGATTTCAAGG
AACAGAACATTACCTCGATCAAGCAAGAGGTCATCCGCAATCAGAACAATATCCAAATTCT
GTCCGAGGATCTCAACATTAAGGAACAGAATCTGACGTCGATCAAGGCCGACCTCAGTAC
CAAGCAGCAAACTTTCTTGAACATTAAGGAGGATGTGATCCTGAATCAGCAAATCATTCAT
AAGATCAAGCAGGACCTGAACACCTACCGCCACAATATGTCCAACATTGAGGAACATCTG
GAAGTGATCCTCACGAACTTGTCTACAGCTTCAATCAAGGTCAAGAATCAGACGGATGAG
GGATCCAAGATGAGCTACCCCCCACGCAAGTCTTGCCGTGATGTGAACTCAACAGACGAG
CGTGTGGTCGTGACGTTGACAAGCGGCCTGAAGGTCATGTGCGATACCAAGACTGACGGC
GGTGGATGGATCATTTTCCAGAGGCGCATCAACGGTAATGTGGACTTCTACCGTGGATGGA
AGGAGTACAGGGATGGCTTCGGTGACTACAACATTGGAGAGTTCTACCTCGGCAACGAAA
ATATCTACATGTTGACGTCGACAGGCCAGTACAACCTCAGGATCGATTTGAAGTACAAGAA
TAAGGCCTTCTTCGCTCAATACTCCGGTTTCAAGATTCTGAGCGAGAAGGAAAAGTACAA
GCTCAACATCGGAGCCTACTCGGGAAATAGTGGCGACAACTTCTCTTCACACAACAATGCT
TTCTTCACCACTTTCGACCGTGATAACGACGAGTACTCCTACAATTGCGCCGTGGATTACA
CCGGCGCTTGGTGGTACCATTCCAGCTGCCTGAACTGCAATCTCAACGGCAAGTGGGGAT
CGAGTGACTTCGCCAAGGGTGTCAACTGGTACGACCTGAGCCGTTTCGATTCGTCCGTGTC
GTTCACTGAGATGAAGATTAGGGAGATTTAA

Figure 1—figure supplement 1. Nucleotide sequence for cloning r*Bg*FREP3. This sequence is derived from GenBank AY028461.1. There are differences in nucleotide sequence but the amino acids match 100%. The differences are due to codon-optimization for expression in insect Sf9 cells. GenBank annotations are used to locate the individual *Bg*FREP3.2 domains: underline, signal peptide; gray shading, IgSF1 domain; double

1111 underline, small connecting region; gray shading plus frame, IgSF2 domain; wavy

1112 underline, ICR; bold underline, FBG domain; delete line, termination codon.

1113
1114

```
CLUSTAL O (1.2.4) multiple sequence alignment

BgTEP           MRMKLNLILFVFYLVFQECQGGKYFISAPRNVVPGTAYDISVDILKQDIGNVTVEAILQD    60
BgTEP1.1        MRMKLNLILFVFYLVFQECQGGKYFISAPRNVVPGTAYDISVDILKQDIDNVTVEAILQD    60
BgTEP1.2        MRMKLNLILFVFYLVFQECQGGKYFISAPRNVVPGTAYDISVDILKQDIDNVTVEAILQD    60
BgTEP1.3        MRMKLNLILFVFYLVFQECQGGKYFISAPRNVVPGTAYDISVDILKQDIDNVTVGAILQD    60
BgTEP1.4        MRMKLNLILFVFYLVFQECQGGKYFISAPRNVVPGTAYDISVDILKQDIDNVTVEAILQD    60
BgTEP1.5        MRMKLNLILFVFYLVFQECQGGKYFISAPRNVVPGTAYDISVDILKQDIGNVTVEAILQD    60
                ************************************************.**** *****

BgTEP           YSFSIPEGPKSLLTANGTFSPGVRGTLSMPIDFNLHCSYCRILLKGYNPLQFEQDIFIQI    120
BgTEP1.1        YSFSIPEGPKSLLTANGTFSPGVRGTLSMPIDFNLHCSYCRILLKGYNPLQFEQDIFIQI    120
BgTEP1.2        YSFSIPEGPRSLLTANGTFSPGVRGTLSMPIDFNLHCSYCRILLKGYNPLQFEQDIFIQI    120
BgTEP1.3        YSFSIPEGPKSLLTANGTFSPGVRGTLSMPIDFNLHCSYCRILLKGYNPLQFEQDIFIQI    120
BgTEP1.4        YSFSIPEGPKSLLTANGTFSPGVRGTLSMPIDFNLHCSYCRILLKGYNPLQFEQDIFIQI    120
BgTEP1.5        YSFSIPEGPKSLLTANGTFSPGVRGTLSMPIDFNLHCSYCRILLKGYNPLQFEQDIFIQI    120
                *********:**************************************************

BgTEP           SSDILSILIQTDKAIYKPKERVNFRILAAYYNLQLYTGTFHYEILDPYDNKINVLSGVSG    180
BgTEP1.1        SSDILSILIQTDKAIYKPKERVNFRILAAYYNLQLYTGTFHYEILDPYDNKINVLSGVSG    180
BgTEP1.2        SSDILSILIQTDKAIYKPKERVNFRILAAYYNLQLYTGTFHYEILDPYDNKINVLSGVSG    180
BgTEP1.3        SSDILSILIQTDKAIYKPKERVNFRILAAYYNLQLYTGTFHYEILDPYDNKINVLSGVSG    180
BgTEP1.4        SSDILSILIQTDKAIYKPKERVNFRILAAYYNLQLYTGTFHYEILDPYDNKINVLSGVSG    180
BgTEP1.5        SSDILSILIQTDKAIYKPKERVNFRILAVYYNLQLYTGTFHYEILDPYDNKINVLSGVSG    180
                ****************************.**************************------

BgTEP           TFGVVEGFFDLSDQPSFGTWKINVRTETVSGEESQLFEVAEYDLPRFQVDVGLPPFALLS    240
BgTEP1.1        TFGVVEGFFDLSDQPSFGTWKINVRTETVSGAESQFFEVAEYDLPRFQVDVGLPPFALLS    240
BgTEP1.2        TFGVVEGFFDLSDQPSFGTWKINVRTETVSGAESQFFEVAEYDLPRFQVDVGLPPFALLS    240
BgTEP1.3        TFGVVEGFFDLSDQPSFGTWKINVRTETVSGAESQFFEVAEYDLPRFQVDVGLPPFALLS    240
BgTEP1.4        TFGVVEGFFDLSDQPSFGTWKINVRTETVSGAESQFFEVAEYDLPRFQVDVGLPPFALLS    240
BgTEP1.5        TFGVVEGFFDLSDQPSFGTWKINVRTETVSGAESQFFEVAEYDLPRFQVDVGLPPFALLS    240
                ******************************* ***:************************

BgTEP           DTTLSGSVEAKYTFGQPVYGLVLLQIGENVDTIDKCNVNRKVTEISFEIKGKENFSVPLE    300
BgTEP1.1        DTTLSGSVEAKYTFGQPVYGLVLLQIGENVDTIDKCNVNRKVTEISFEIKGKGNFSVPLE    300
BgTEP1.2        DTTLSGSVEAKYTFGQPVYGLVLLQIGENVDTIDKCNVNRKVTEISFEIKGKGNFSVPLE    300
BgTEP1.3        DTTLSGSVEAKYTFGQPVYGLVLLQIGENVDTIDKCNVNRKVTEISFEIKGKGNFSVPLE    300
BgTEP1.4        DTTLSGSVEAKYTFGQPVYGLVLLQIGENVDTIDKCNVNRKVTEISFEIKGKGNFSVPLE    300
BgTEP1.5        DTTLSGSVEAKYTFGQPVYGLVLLQIGENVDTIDKCNVNRKVTEISFEIKGKGNFSVPLE    300
                ************************************************* *******

BgTEP           DIRRSVHLNEKKKIKITAFVTEASTGIKLNGSSVITYYGNRYQIKFLEMTPAVFKPGLQY    360
BgTEP1.1        DIRRSVHLNEKKKIKITAFVTEASTGIKLNGSSVITYYGNRYQIKFLEMTPAVFKPGLQY    360
BgTEP1.2        DIRRSVHLNEKKKIKITAFVTEASTGIKLNGSSVITYYGNRYQIKFLEMTPAVFKPGLQY    360
BgTEP1.3        DIRRSVHLNEKKKIKITAFVTEASTGIKLNGSSVITYYGNRYQIKFLEMTPAVFKPGLQY    360
BgTEP1.4        DIRRSVHLNEKKKIKITAFVTEASTGIKLNGSSVITYYGNRYQIKFLEMTPAVFKPGLQY    360
BgTEP1.5        DIRRSVHLNEKKKIKITAFVIEASTGIKLNGSSVITYYGNRYQIKFLEMTPAVFKPGLQY    360
                ******************* ************************************

BgTEP           TAYVQVTTPDGLPPTDSNLSLSVYTSVTYQMTVPDQELYSPSSFSGSYPLPGQNMSLPAN    420
BgTEP1.1        TAYVQVTTPDGLPPTDSNLSLSVYTSVTYQMTVPDQELYSPSSFSGSYPLPGQNMSLPAN    420
BgTEP1.2        TAYVQVTTPDGLPPTDSNLSLSVYTSVTYQMTVPDQELYSPSSFSGSYPLPGQNMSLPAN    420
BgTEP1.3        TAYVQVTTPDGLPPTDSNLSLSVYTSVTYQMTVPDQELYSPSSFSGSYPLPGQNMSLPAN    420
BgTEP1.4        TAYVQVTTPDGLPPTDSNLSLSVYTSVTYQMTVPDQELYSPSSFSGSYPLPGQNMSLPAN    420
BgTEP1.5        TAYVQVTTPDGLPPTESNLSLSVYTSVTYQMTVPDQELYSPSSFSGSYPLPGQNMSLPAN    420
                ***************:************************************************

BgTEP           GILSIDIDIPLNATSIDIKVSLNKQTTAEKRISKSYSMSNNYLQLSLLSKLVKAESDVLI    480
BgTEP1.1        GILSIDIDIPLNATSIDIKVSLNKETTAEKRISKSYSMSNNYLQLSLLSKLVKAESDVLI    480
BgTEP1.2        GILSIDIDIPLNATSIDIKVSLNKETTAEKRISKSYSMSNNYLQLSLLSKLVKAESDVLI    480
BgTEP1.3        GILSIDIDIPLNATSIDIKVSLNKETTAEKRISKSYSMSNNYLQLSLLSKLVKAESDVLI    480
BgTEP1.4        GILSIDIDIPLNATSIDIKVSLNKETTAEKRISKSYSMSNNYLQLSLLSKLVKAESDVLI    480
BgTEP1.5        GILSIDIDIPLNATSIDIKVSLNKQTTAEKRISKSYSMSNNYLQLSLLSKLVKAESDVLI    480
                ************************:************************************

BgTEP           KITSTEAIDSLAYEIRSRSDRVKSGVLELNGQREFNATFKVEPSWAPIAQLLMYYIRRDS    540
BgTEP1.1        KITSTEAIDSLAYEIRSRSDHVKSGVLELSGQREFNATFKVEPSWAPIAQLLMYYIRRDS    540
```

```
1184  BgTEP1.2    KITSTEAIDSLAYEIRSRSDHVKSGVLELSGQREFNATFKVEPSWAPIAQLLMYYIRRDS  540
1185  BgTEP1.3    KITSTEAIDSLAYEIRSRSDHVKSGVLELSGQREFNATFKVEPSWAPIAQLLMYYIRRDS  540
1186  BgTEP1.4    KITSTEAIDSLAYEIRSRSDHVKSGVLELSGQREFNATFKVEPSWAPIAQLLMYYIRRDS  540
1187  BgTEP1.5    KITSTEAIDSLAYEIRSRSDHVKSGVLELNGQREFNATFKVGPSWAPIAQLLMYYIRRDS  540
1188              ********************:********.********** ******************
1189
1190  BgTEP       NEVVTDSLAFNVEGMFKNKVNVAFKENETDINKNVSLELPADSDSQIYVLAVDXSVLLLK  600
1191  BgTEP1.1    NEVVTDSLAFNVEGMFKNKVNVAFKENETDINKNVSLELSADSDSQIYVLAVDQSVLLLK  600
1192  BgTEP1.2    NEVVTDSLAFNVEGMFKNKVNVAFKENETDINKNVSLELSADSDSQIYVLAVDQSVLLLK  600
1193  BgTEP1.3    NEVVTDSLAFNVEGMFKNKVNVAFKENETDINKNVSLELSADSDSQIYVLAVDQSVLLLK  600
1194  BgTEP1.4    NEVVTDSLAFNVEGMFKNKVNVAFKENETDINKNVSLELSADSDSQIYVLAVDQSVLLLK  600
1195  BgTEP1.5    NEIVTDSLAFNVEGMFENKVNVAFKENETDINKNVSLELSADSDSQIYVLAVDQSVLLLK  600
1196              **:*************:********************** ************ ******
1197
1198  BgTEP       TGNDXTPXKVKDSFXSXFHKGAIPTDSNFALSYSGSSIXEVFSNMGLVIATDLNIFAPFR  660
1199  BgTEP1.1    TGNDLTPNKVKDSFISKFHKGAIPTDSNFALSYSGSSINEVFSNMGLVIATDLNIFAPFR  660
1200  BgTEP1.2    TGNDLTPNKVKDSFISKFHKGAIPTDSNFALSYSGSSINEVFSNMGLVIATDLNIFAPFR  660
1201  BgTEP1.3    TGNDLTPNKVKDSFISKFHKGAIPTDSNFALSYSGSSINEVFSNMGLVIATDLNIFAPFR  660
1202  BgTEP1.4    TGNDLTPNKVKDSFISKFHKGAIPTDSNFALSYSGSSINEVFSNMGLVIATDLNIFAPFR  660
1203  BgTEP1.5    TGNDLTPNKVKDSFISKFHKGEVPTDSNFALSYSGSSINEVFSNMGLVIATDLNIFAPFR  660
1204              **** ** ****** * **** :*************** ********************
1205
1206  BgTEP       PIALGRFPSSGFDRQXMMGAPXAMSFXDDXAMXSASFEMDVTTSTKPVERVRSFFPESWL  720
1207  BgTEP1.1    PIALGRFPSSGFDRQGMMGAPMAMSFRDDNAMESASFEMDVATSTKPVERVRSFFPESWL  720
1208  BgTEP1.2    PIALGRFPSSGFDRQGMMGAPMAMSFRDDNAMESASFEMDVATSTKPVERVRSFFPESWL  720
1209  BgTEP1.3    PIALGRFPSSGFDRQGMMGAPMAMSFRDDNAMESASFEMDVATSTKPVERVRSFFPESWL  720
1210  BgTEP1.4    PIALGRFPSSGFDRQGMMGAPMAMSFRDDNAMESASFEMDVATSTKPVERVRSFFPESWL  720
1211  BgTEP1.5    PIALGRFPSSGFDRQGMMGAPMAMSFRDDNAMESASFEMDVATSTKPVERVRSFFPESWL  720
1212              ********************* ***** **** ** ** ********:**************
1213
1214  BgTEP       WTSVKSINGHATLTTTVPDTITSWIGSAFATNSDTGLGVAPTTSKLPGFRPFFGSLTYPR  780
1215  BgTEP1.1    WTSVKSINGHATLTTTVPDTITSWIVSAFATNSDTGLGVAPTTSKLRVFRPFFVSLTYPR  780
1216  BgTEP1.2    WTSVKSINGHATLTTTVPDTITSWIVSAFATNSDTGLGVAPTTSKLRVFRPFFVSLTYPR  780
1217  BgTEP1.3    WTSVKSINGHATLTTTVPDTITSWIVSAFATNSDTGLGVAPTTSKLRVFRPFFVSLTYPR  780
1218  BgTEP1.4    WTSVKSINGHATLTTTVPDTITSWIVSAFATNSDTGLGVAPTTSKLRVFRPFFVSLTYPR  780
1219  BgTEP1.5    WTSVKSINGHATLTTTVPDTITSWIVSAFATNSDTGLGVAPTTSKLRVFRPFFVSLTYPR  780
1220              ************************ ********************  ***** ******
1221
1222  BgTEP       SVTRNEQFIVQATVFNYLPVDLMVTVSLKENPFLTPVTPGPGNQASNIQVRANEQRTVYF  840
1223  BgTEP1.1    SVTRNEQFIVQATVFNYLPVDLMVTVSLKENPFLTPITPGPGNQASNIQVRANEQGIVYF  840
1224  BgTEP1.2    SVTRNEQFIVQATVFNYLPVDLMVTVSLKENPFLTPITPGPGNQASNIQVRANEQGIVYF  840
1225  BgTEP1.3    SVTRNEQFIVQATVFNYLPVDLMVTVSLKENPFLTPITPGPGNQASNIQVRANEQGIVYF  840
1226  BgTEP1.4    SVTRNEQFIVQATVFNYLPVDLMVTVSLKENPFLTPITPGPGNQASNIQVRANEQGIVYF  840
1227  BgTEP1.5    SVTRNEQFIVQATVFNYLPVDLMVTVSLKENPFLTPITPGPGNQASNIQVRANEQGIVYF  840
1228              ************************************:******************  ***
1229
1230  BgTEP       SLSALIVGALDIEVSARSNMAADATVRQILVRHEGAPVVYNNPILISLSNNQSTFEKNIA  900
1231  BgTEP1.1    SLSALTVGSLDIEVSARSNMAADATVRQILIKHEGAPVVYNNPILINLSNNQSTFEKNIA  900
1232  BgTEP1.2    SLSALTVGSLDIEVSARSNMAADATVRQILIKHEGAPVVYNNPILINLSNNQSTFEKNIA  900
1233  BgTEP1.3    SLSALTVGSLDIEVSARSNMAADATVRQILIKHEGAPVVYNNPILINLSNNQSTFEKNIA  900
1234  BgTEP1.4    SLSAPTVGSLDIEVSARSNMAADATVRQILIKHEGAPVVYNNPILINLSNNQSTFEKNIA  900
1235  BgTEP1.5    SLSALTVGSLDTEVSARSNMAADAIVRQILIKHEGAPVVYNNPILINLSNNQSTFEKNIA  900
1236              ****  **:** ******************::***************.**************
1237
1238  BgTEP       FTLPDSLVPESHRIRVKVTGDLIGSTVQSLTSLLTLPTGCGEQSLVKFTPNIHIGRYLKA  960
1239  BgTEP1.1    FTLPDSLVPESQRIRVKVTGDLIGSTVQSLTSLLTLPTGCGEQSLVKFTPNIHIGRYLKA  960
1240  BgTEP1.2    FTLPDSLVPESQRIRVKVTGDLIGSTVQSLTSLLTLPTGCGEQSLVKFTPNIHIGRYLKA  960
1241  BgTEP1.3    FTLPDSLVPESQRIRVKVTGDLIGSTVQSLTSLLTLPTGCGEQSLVKFTPNIHIGRYLKA  960
1242  BgTEP1.4    FTLPDSLVPESQRIRVKVTGDLIGSTVQSLTSLLTLPTGCGEQSLVKFTPNIHIGRYLKA  960
1243  BgTEP1.5    FTLPDSLVPESQRIRVKVTGDLIGSTVQSLTSLLTLPTGCGEQSLVKFTPNIHIGRYLKA  960
1244              ***********:************************************************
1245
1246  BgTEP       TNQLSKELNKKIIDLLNNGYQRQLTYKRYDNGFSAFGNYDLSSSTWLTALVVTSFAEAQE  1020
1247  BgTEP1.1    TNQLSEELNKKIIDLLNDGYQRQLTYKRYDNGFSAFGNYDISSSTWLTALVVTSFAEAQE  1020
1248  BgTEP1.2    TNQLSEELNKKIIDLLNDGYQRQLTYKRYDNGFSAFGNYDISSSTWLTALVVTSFAEAQE  1020
1249  BgTEP1.3    TNQLSEELNKKIIDLLNDGYQRQLTYKRYDNGFSAFGNYDISSSTWLTALVVTSFAEAQE  1020
1250  BgTEP1.4    TNQLSEELNKKIIDLLNDGYQRQLTYKRYDNGFSAFGNYDISSSTWLTALVVTSFAEAQE  1020
1251  BgTEP1.5    TNQLSEELNKKIIDLLNDGYQRQLTYKRYDNGFSAFGNYDISSSTWLTALVVTSFAEAQE  1020
1252              *****:***********:**********************:********************
```

```
1253
1254  BgTEP          FIFVDKEIILKASMLLIDRQNLDGSFNEFGKVLDRNTQGT-TAGPALTAFVLVALLKAKE   1079
1255  BgTEP1.1       FIFVDKEIILKASMLLIDRQNIDGSFNEFGKVLDRNTQGT-TAGPALTAFVLVALLKAKE   1079
1256  BgTEP1.2       FIFVDKEIILKASMLLIDRQNIDGSFNEFGKVLDRNTQGT-TAGPALTAFVLVALLKAKE   1079
1257  BgTEP1.3       FIFVDKEIILKASMLLIDRQNIDGSFNEFGKVLDRNTQGT-TAGPALTAFVLVALLKAKE   1079
1258  BgTEP1.4       FIFVDKEIILKASMLLIDRQNIDGSFNEFGKVLDRNTQGT-TAGPALTAFVLVALLKAKE   1079
1259  BgTEP1.5       FIFVDKEIILKASMLLIDRQNIDGSFNEFGKVLDRNTQGTTTAGPALTAFVLVALLKAKE   1080
1260                 ********************:******************* ******************
1261
1262  BgTEP          LADVQYCKNNNKCRYYLLGNATLNATRNLERLMLADSIDDQFSLAVTSYALAEAKSQLAQ   1139
1263  BgTEP1.1       LADVQDCKNNNKCRYYLLGNATLNATRNLERLMLADSIDDQFSLAVASYAFAEAKSQLAQ   1139
1264  BgTEP1.2       LADVQDCKNNNKCRYYLLGNATLNATRNLERLMLADSIDDQFSLAVASYAFAEAKSQLAQ   1139
1265  BgTEP1.3       LADVQDCKNNNKCRYYLLGNATLNATRNLERLMLADSIDDQFSLAVASYAFAEAKSQLAQ   1139
1266  BgTEP1.4       LADVQDCKNNNKCRYYLLGNATLNATRNLERLMLADSIDDQFSLAVASYAFAEAKSQLAQ   1139
1267  BgTEP1.5       LADVQDCKNNNKCRYYLLGNATLNATRNLERLMLADSIDDQFSLAVASYALAEAKSQLAQ   1140
1268                 ***** ****************************************:***:*********
1269
1270  BgTEP          STFEKLLTFVKQEGGLEYRSANSTVNNEELNRFINWRPPRLQARPIDILITSYAILTYSP   1199
1271  BgTEP1.1       STFEKLLTFVKQEGGLEYWSANSTVNNEELNRFINWRPPRLQARPIDILITSYAILTYSS   1199
1272  BgTEP1.2       STFEKLLTFVKQEGGLEYWSANSTVNNEELNRFINWRPPRLQARPIDILITSYAILTYSS   1199
1273  BgTEP1.3       STFEKLLTFVKQEGGLEYWSANSTVNNEELNRFINWRPPRLQARPIDILITSYAILTYSS   1199
1274  BgTEP1.4       STFEKLLTFVKQEGGLEYWSANSTVNNEELNRFINWRPPRLQARPIDILITSYAILTYSS   1199
1275  BgTEP1.5       STFEKLLTFVKQEGGLEYWSANSTVNNEELNRFINWRPPRLQARPIDILITSYAILTYSS   1200
1276                 ****************** ****************************************
1277
1278  BgTEP          LGRLDEALPSVRWLTLQKNAQGGFVSTQDTVVGLQALSTYGSKSFRPDTNITIYVSDMNT   1259
1279  BgTEP1.1       LGRLDEALPSVRWLTLQKNAQGGFVSTQDTVVGLQALSSYGSKSFRPDTNITIYVSDMNT   1259
1280  BgTEP1.2       LGRLDEALPSVRWLTLQKNAQGGFVSTQDTVVGLQALSSYGSKSFRPDTNITIYVSDMNT   1259
1281  BgTEP1.3       LGRLDEALPSVRWLTLQKNAQGGFVSTQDTVVGLQALSSYGSKSFRPDTNITIYVSDMNT   1259
1282  BgTEP1.4       LGRLDEALPSVRWLTLQKNAQGGFVSTQDTVVGLQALSSYGSKSFRPDTNITIYVSDMNT   1259
1283  BgTEP1.5       LGRLDEALPSVRWLTLQKNAQGGFVSTQDTVVGLQALSFYGSKSFRPDTNITIYVSDMNT   1260
1284                 ************************************ ********************
1285
1286  BgTEP          HLTMNVKSDNALSLQIQEIQSNSQDFSITASGSGLALLDIEYSFNVLKELSKPVFDVNTV   1319
1287  BgTEP1.1       HLTMNVNSENALSLQIQEIQSNSQDFSITASGSGLALLDIEYSFNVLKELSKPVFDVNTV   1319
1288  BgTEP1.2       HLTMNVNSENALSLQIQEIQSNSQDFSITASGSGLALLDIEYSFNVLKELSKPVFDVNTV   1319
1289  BgTEP1.3       HLTMNVNSENALSLQIQEIQSNSQDFSITASGSGLALLDIEYSFNVLKELSKPVFDVNTV   1319
1290  BgTEP1.4       HLTMNVNSENALSLQIQEIQSNSQDFSITASGSGLALLDIEYSFNVLKELSKPVFDVNTV   1319
1291  BgTEP1.5       HLTMNVNSENALSLQIQEIQSNSQDFSITASGSGLALLDIEYSFNVLKELSKPVFDANTV   1320
1292                 ******:*:************************************************.***
1293
1294  BgTEP          LLDDKLDSFNIMVCTKFLLKHDTGMVVQEVSIPSGFVPDLSTLGQVAGVKRSERKGSIVA   1379
1295  BgTEP1.1       LLDDKLDSFNIMVCTKFLMKHDTGMVVQELSIPSGFVPDLSTLGQVAGVKRSERKGSIVA   1379
1296  BgTEP1.2       LLDDKLDSFNIMVCTKFLMKHDTGMVVQELSIPSGFVPDLSTLGQVAGVKRSERKGSIVA   1379
1297  BgTEP1.3       LLDDKLDSFNIMVCTKFLMKHDTGMVVQELSIPSGFVPDLSTLGQVAGVKRSERKGSIVA   1379
1298  BgTEP1.4       LLDDKLDSFNIMVCTKFLMKHDTGMVVQELSIPSGFVPDLSTLGQVAGVKRSERKGSIVA   1379
1299  BgTEP1.5       LLDDKLDSFNIMVCTKFLMKHDTGMVVQELSIPSGFVPDLSTLGQVAGVKRSERKGSIVA   1380
1300                 ******************:**********:*******************************
1301
1302  BgTEP          IYFDKISGSSLCYSIVMTREAKVAKSQKSYVRTYDYYEPANQATVFYQPRTLRDSTVCDV   1439
1303  BgTEP1.1       IYFDKISGSSLCYSIVMTREAKVAKSQKSYVRTYDYYEPANQATVFYQPRTLRDSTVCDV   1439
1304  BgTEP1.2       IYFDKISGSSLCYSIVMTREAKVAKSQKSYVRTYDYYEPANQATVFYQPRTLRDSTVCDV   1439
1305  BgTEP1.3       IYFDKISGSSLCYSIVMTREAKVAKSQKSYVRTYDYYEPANQATVFYQPRTLRDSTVCDV   1439
1306  BgTEP1.4       IYFDKISGSSLCYSIVMTREAKVAKSQKSYVRTYDYYEPANQATVFYQPRTLRDSTVCDV   1439
1307  BgTEP1.5       IYFDKISGSSLCYSIVMTREAKVAKSQKSYVRTYDYYEPANQATVFYQPRTLRDSTVCDV   1440
1308                 ************************************************************
1309
1310  BgTEP          CLNCCP  1445
1311  BgTEP1.1       CLNCCP  1445
1312  BgTEP1.2       CLNCCP  1445
1313  BgTEP1.3       CLNCCP  1445
1314  BgTEP1.4       CLNCCP  1445
1315  BgTEP1.5       CLNCCP  1446
1316                 ******
1317
```

Figure 1—figure supplement 2. Alignment of multiple *Bg*TEP amino acid sequences and distribution of identified peptides. Peptides identified by LC-MS/MS from r*Bg*FREP3 pull-down experiments are highlighted in gray. The GenBank accession numbers of each entry are: *Bg*TEP, ACL00841.1; *Bg*TEP1.1, ADE45332.1; *Bg*TEP1.2, ADE45339.1; *Bg*TEP1.3, ADE45340.1; *Bg*TEP1.4, ADE45341.1 and *Bg*TEP1.5, ADE45333.1.

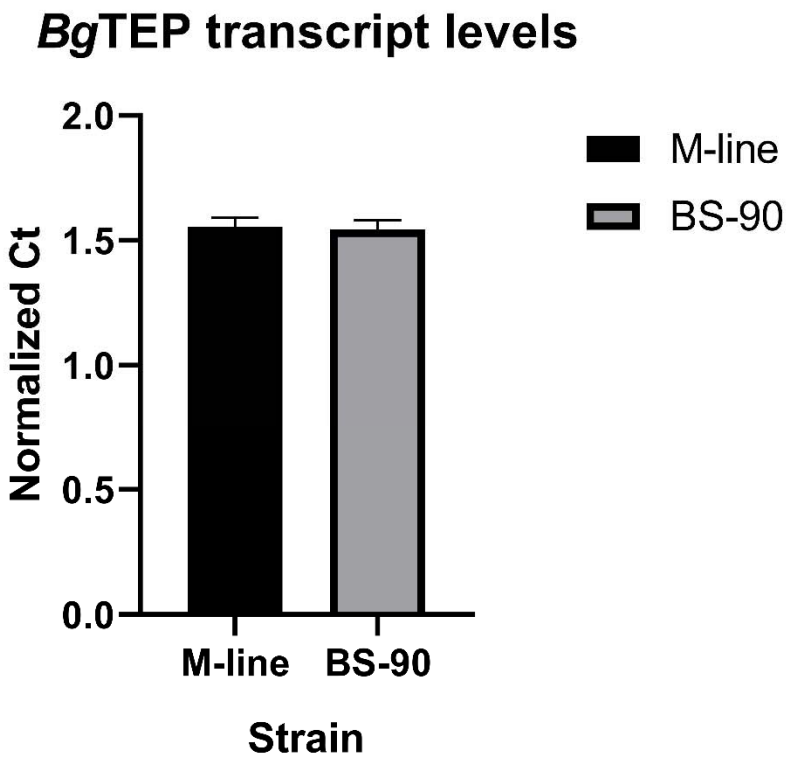

Figure 1—figure supplement 3.

Figure 1—figure supplement 3. Normalized transcript abundance of *Bg*TEP1 in unchallenged M-line (black bar, n=5) and BS-90 (grey bar, n=5) snails indicating a lack of significant difference ($p > 0.05$) between the two strains.

Score	Expect	Method	Identities	Positives	Gaps
1143 bits(2957)	0.0	Compositional matrix adjust.	538/572(94%)	555/572(97%)	0/572(0%)

```
CLUSTAL O (1.2.4) multiple sequence alignment

A0A182YTN9          MLVQFLFAATLLQYVSSQCTYSSWWYSFDTPGQSKCNEINSYINALDRNDVNWADDALSN  60
A0A182YTZ4          MFLQIFVAVTLVQYVSSQCTYSSWWYSFDTPGQSKCNEINSYINALDRNDVNWADDALSN  60
                    *::*::.*.**:**************************************************

A0A182YTN9          LEGVQCCRPPAPWNNVEQQVVYEDWTATLDSDYTWAFCRVGYFLQGLYRSDTGWPRFKGY  120
A0A182YTZ4          LEGVQCCRPPAPWNNVEQQVVYEDWTATLDSDYTWAFCRVGYFLQGLYRSDTGWPRFKGY  120
                    ************************************************************

A0A182YTN9          LFNLESARCTKPANHPLNYGNCQDIDVSSCMGRKGQCSCPGGYFLTGLYRADGDDLYFLK  180
A0A182YTZ4          LFNLESARCTKPANHPLNYGTCQDIDVSSCMGRKGQCSCPGGYFLTGLYRADGDDLYFLK  180
                    **********************.*************************************

A0A182YTN9          KIRCCTPAAKPLEMDEKSKIQTRIMDTTLWNMATLAHYLGYGWCYGCRGVAVGEDFTRNG  240
A0A182YTZ4          KIRCCTPAAKPLEMDEKSKIQTRIMDTTLWNMATLAHYMGYGWCYGCHGLAVGEDFTRNG  240
                    ****************************************:********:*:**********

A0A182YTN9          FTWAADTRSFWGKWCEGDKNGERLNLVFGDWGFAVKEIIYGKSVIEDLQAESVDSGVLYN  300
A0A182YTZ4          FTWAADTRTFWGKWCEGDKNGERLNLVFGDWGFAVKEIIYGKSVIEDLQAESVDSGVLYN  300
                    ********:****************************************************

A0A182YTN9          RASSPVTESIDRTKTIEETVTHSTTSTFTNSHELGIELNFEIASVSGKASYTTKFEYSKA  360
A0A182YTZ4          RASSPVTESIERSKTIQETITHSTTSTFTNSHGLGVELEFEIASVKGKASYKTRFEYSTS  360
                    **********:*:***:**:************ **:**:******.*****.*:****.:

A0A182YTN9          TTNEKSISQTAGFTKKSSITLGPMEGAKYEIIMSKSRTTVPYTAIITTKFSTEMKGFLRW  420
A0A182YTZ4          TTNSKSISETQGFTKQSSITLGPMEGAKYEVIMSKSRTTVPYTAIITTKFSTEMKGFLRW  420
                    ***.****:* ****:***************:******************************

A0A182YTN9          EDGNGNFHQDYRTNSGRPTYNYRFGDSSVPFYKALKKQSDNNEGVWMWGMLFQKFPDARR  480
A0A182YTZ4          EDGNGNFHQDYRTNSGRPTFNYRFGDSSVPFYKALKKQSDNNEGVWMWGMLFQKFPDARR  480
                    ********************:*******************************************

A0A182YTN9          VINRLTDETQYQFTLTGKLEKVEGTSVNVKWEKIKLNRRDVSGNDAPGSNITTYIAASGP  540
A0A182YTZ4          VTNRLTDETQYQFTLAGKLEKVEGTSVNVKWEKMKLNRRDVSGNDEPGSNITTYIAASGP  540
                    * ************:*****************:*********** **************

A0A182YTN9          ADKPAVVEYPKVNLNNKEPFKPIEIPVTEVKV 572
A0A182YTZ4          ADKPAVVEYPKVNLNNKEPFKPIEIPVTEVKV 572
                    ********************************
```

Figure 1—figure supplement 4. The alignments of two Biomphalysin variants (UniProtKB/TrEMBL: A0A182YTN9 and A0A182YTZ4) identified by LC-MS/MS. Peptides identified from A0A182YTN9 and A0A182YTZ4 are highlighted in gray and underline, respectively. Two red oval boxes representing differences in peptide sequences that distinguish two Biomphalysin variants.

```
CLUSTAL O (1.2.4) multiple sequence alignment

BgMFREP1 PARTIAL ------------------------------------------------------------  0
BgMFREP2 PRE     ------------------------------------------------------------  0
BgFREP3-2 PRE    MARLFLLFILCVFVVSLAGSELVIDVQPNVISPEITPQLVINCSITNNEVQTLDLIKSLT  60
BgMFREP4 PRE     ------------------------------------------------------------  0
BgMFREP5 PARTIAL ------------------------------------------------------------  0
BgMFREP6 PARTIAL ------------------------------------------------------------  0
BgFREP7.1 PRE    MTNLLLRLVFFQSLLPLLSSELVIDVQPDIISAEITAQLVINCSVTNNQVQHLDVIRSLT  60
BgMFREP8 PARTIAL ------------------------------------------------------------  0
BgMFREP9 PARTIAL ------------------------------------------------------------  0
BgMFREP10PARTIAL ------------------------------------------------------------  0
BgMFREP11PARTIAL ------------------------------------------------------------  0
BgFREP12.1 PRE   -MTLLQHFLLFVSLFLFSSSELVIDVQPNVISAEITAQLVINCSITNNQVQHLDVIRSLT  59
BgFREP12-FBG2    ------------------------------------------------------------  0
BgFREP13.1 PRE   -MATFLYVILVVALVTLSTSELVIDVQPNVISPEITAQLAINCSVTNNQAQNIDVIKSLT  59
BgFREP14         ------------------------------------------------------------  0

BgMFREP1 PARTIAL ------------------------------------------------------------  0
BgMFREP2 PRE     ------------------------------------------------------------  0
BgFREP3-2 PRE    LSRYNETIREFDELIALDSLTLNLKQFVRFKYSQISFGNRYITLILHNPTQFDARIYKCN  120
BgMFREP4 PRE     ------------------------------------------------------------  0
BgMFREP5 PARTIAL ------------------------------------------------------------  0
BgMFREP6 PARTIAL ------------------------------------------------------------  0
BgFREP7.1 PRE    LSRYNQTLRDFEDITALDLLTLNLKQLVKFKHSHISFGNVFISLTLLYPTQFDANVYRCS  120
BgMFREP8 PARTIAL ------------------------------------------------------------  0
BgMFREP9 PARTIAL ------------------------------------------------------------  0
BgMFREP10PARTIAL ------------------------------------------------------------  0
BgMFREP11PARTIAL ------------------------------------------------------------  0
BgFREP12.1 PRE   LSRYDETLKEFLDLITLEAKTLNLSQLVQLHHAQISFGNLSISLTLHNPTQFDAKVYRCK  119
BgFREP12-FBG2    ------------------------------------------------------------  0
BgFREP13.1 PRE   LSRYNETIRDFEVMIXLDLLTLNLKQLVQFNYSLISFGNVFISLTILHPTKSDAKVYRCS  119
BgFREP14         ------------------------------------------------------------  0

BgMFREP1 PARTIAL ------------------------------------------------------------  0
BgMFREP2 PRE     -------------------------MASLPLR---------LVLLV----SMATLIRSSSWL  24
BgFREP3-2 PRE    ATGTNSEGANISLFAKKAVEYETNSTALIEEIRRIKK--DENYCSFKKDDLSDSKQRSRV  178
BgMFREP4 PRE     ---------------------------------------MKNLLLCLF----LVSATLGSRL  19
BgMFREP5 PARTIAL ------------------------------------------------------------  0
BgMFREP6 PARTIAL ------------------------------------------------------------  0
BgFREP7.1 PRE    VKGGDPNNKDMSLFSKKTVEYETNSTALVEEIRRYKIDENKCLCSLASNDRTSTNKRLRV  180
BgMFREP8 PARTIAL ------------------------------------------------------------  0
BgMFREP9 PARTIAL ------------------------------------------------------------  0
BgMFREP10PARTIAL ------------------------------------------------------------  0
BgMFREP11PARTIAL ------------------------------------------------------------  0
BgFREP12.1 PRE   VEGDKTNAASSSIVAKKEVEYRTNMTALIEEIRRLKVVEKNDQCSLKXEELTSYHQKTKL  179
BgFREP12-FBG2    ------------------------------------------------------------  0
BgFREP13.1 PRE   VKGDNSNERNISLFAKKAVEYETNSTALIEEIQRCKKVKSKCQCSLENNDSSSNYKRSRV  179
BgFREP14         -------------------------MGALIL---------HVILCAS---IVPVISSRPKL  24

BgMFREP1 PARTIAL ------------------------------------------------------------  0
BgMFREP2 PRE     NFTGNSETIRELIQPLKLTCTFQISKNDSDNDSQVLFMSIYHETKRVIASISKYQPVATS  84
BgMFREP3-2 PRE   YFSGSSDIIKERIEPLTLKCTFQVLKTDQNETSRLQSLYILHESKGVIAYVNKDQPVVTS  238
BgMFREP4 PRE     SFNANVEKINEVIXPLMLTCSFEVSRNDSWQNTKVQLMYIMHETKGFVATITKDQNITG-  78
BgMFREP5 PARTIAL ---------------------------------------------SNRIIANINKDRQVLTP  17
BgMFREP6 PARTIAL ------------------------------------------------------------  0
BgFREP7.1 PRE    NFSGNSEIIKERVETLTLNCTFQVLNQDQNETSSLQSLYILHETNGVIANINKGQPVLK-  239
BgMFREP8 PARTIAL ------------------------------------------------------------  0
BgMFREP9 PARTIAL ------------------------------------------------------------  0
BgMFREP10PARTIAL ------------------------------------------------------------  0
BgMFREP11PARTIAL ------------------------------------------------------------  0
BgFREP12.1 PRE   HFVGSSKVIKELIEPLTLTCSLQNLDN----NSTVQFMYILHESNGVIATINKDQPVVTT  235
BgFREP12-FBG2    ------------------------------------------------------------  0
BgFREP13.1 PRE   HFSGSSEIIKERIETVTLNCTYQALKHQRNE-TSLQSLYILHEANGVIANINKGQPVVQ-  237
BgFREP14         KFSGNAEILKELIEPLILSCSLKSFNNETNEHFKVHIMFIQHETNGVISTLSKDQAVAVS  84
```

```
1451
1452  BgMFREP1 PARTIAL ------------------------------------------------------------  0
1453  BgMFREP2 PRE     LYPSVTKVQGHIYHSNESKDSYLQVTWTHPKLSESGKYFCLAHAWNSTSQNSVFDADITV  144
1454  BgFREP3-2 PRE    LQGSNIQDVEGEIYDNAIKDSYLQVTWSNLKHSESGKYFCEAHNKYSEGRIDKSSNMLTI  298
1455  BgMFREP4 PRE     NADMTFSEGQGTLNNEIDNTSFXQVTWKNASNELSGKYICVVHATNAEGKVEFLSASLKV  138
1456  BgMFREP5 PARTIAL SQRDNSKNIQGEIHDDGSKDSYLQVTWSNVKSSESGKYFCEANVKHSDGRAERLSEMLII  77
1457  BgMFREP6 PARTIAL ------------------------------------------------------------  0
1458  BgFREP7.1 PRE    --GSHLKDAQGNIFDSGLKDSYLQVTWSNVKLSDSGKYFCEANVKHSDGRAERLSEMLIL  297
1459  BgMFREP8 PARTIAL ------------------------------------------------------------  0
1460  BgMFREP9 PARTIAL ------------------------------------------------------------  0
1461  BgMFREP10PARTIAL ------------------------------------------------------------  0
1462  BgMFREP11PARTIAL ------------------------------------------------------------  0
1463  BgFREP12.1 PRE   KQESNFNTAKGVLSDTKSKASFIEVSWSYIKSSESGNYFCGAHVMGPDGRSERLNEMLAI  295
1464  BgFREP12-FBG2    ------------------------------------------------------------  0
1465  BgFREP13.1 PRE   --GSNLKNAEGEIFHNESKDSYLQVTWSNLKFSESGKYFCEANVKHSDGRAERLSEMLII  295
1466  BgFREP14         A-DQSSTHAHGKIYNKDLQDSYLQVILKNPKISESGKYFCLAYAKNSTGQDSVFQSTVTI  143
1467
1468
1469  BgMFREP1 PARTIAL ------------------------------------------------------------  0
1470  BgMFREP2 PRE     NVIKSSTDDLAVALSYIQDRLDKD----------------------------------  168
1471  BgFREP3-2 PRE    TVERPTFDDLVEAMHKLFTQVDGAKESLKAINQNIKNINKDLDFKEQNITSIKQEVIRNQ  358
1472  BgMFREP4 PRE     QVQKLEIADLAQYVVDLTARVKESDDKIQNYTR---------------NVTSI-------  176
1473  BgMFREP5 PARTIAL TVVRPTFDDLVKVVEKLLEQFNEYKTHLQENV---EKNRAMLDRNKQTILLIKKDVLANQ  134
1474  BgMFREP6 PARTIAL ------------------------------------------------------------  0
1475  BgFREP7.1 PRE    TVVSPTVDDLVKVIEKLLGQVDEDTKHIQENKQNIKNIKEELKTKEQNILSNTADLNSTQ  357
1476  BgMFREP8 PARTIAL ------------------------------------------------------------  0
1477  BgMFREP9 PARTIAL ------------------------------------------------------------  0
1478  BgMFREP10PARTIAL ------------------------------------------------------------  0
1479  BgMFREP11PARTIAL ------------------------------------------------------------  0
1480  BgFREP12.1 PRE   TVSNPTFDDVVKVIPKLLRQADIEKENILENKQNIYDIKEYFKSKQQNIISIKDGLNTNR  355
1481  BgFREP12-FBG2    ------------------------------------------------------------  0
1482  BgFREP13.1 PRE   TVVSPTFDDLVKVIEKLLGQVDGDTRHIQENNQSIKNIKEELKIKEQNIISITADLNSTQ  355
1483  BgFREP14         KVLKPKADDLVQVLGQLLKRVDTLEQLLKGNETNFGGV----------------------  181
1484
1485
1486  BgMFREP1 PARTIAL ------------------------------------------------------------  0
1487  BgMFREP2 PRE     ------------------------------------------------------------  168
1488  BgFREP3-2 PRE    NNIQILSEDLNIKEQNLTSIKADLSTKQQTFLNIKED----------------------  395
1489  BgMFREP4 PRE     ------------------------------------------------------------  176
1490  BgMFREP5 PARTIAL QSLQNMKEDWNSNQT---------------------------------------------  149
1491  BgMFREP6 PARTIAL ------------------------------------------------------------  0
1492  BgFREP7.1 PRE    QTIRSMKEDIAINQHNMSSLKEFVDANLESLQSIKEDLNIQQRNIIS-------------  404
1493  BgMFREP8 PARTIAL ------------------------------------------------------------  0
1494  BgMFREP9 PARTIAL ------------------------------------------------------------  0
1495  BgMFREP10PARTIAL ------------------------------------------------------------  0
1496  BgMFREP11PARTIAL ------------------------------------------------------------  0
1497  BgFREP12.1 PRE   HNIKSIADDLNVNKENIASHNDEINTLRQMVNNVQSDLSICKKSIHNFSNDLDTKKQSIA  415
1498  BgFREP12-FBG2    ------------------------------------------------------------  0
1499  BgFREP13.1 PRE   QIISIIKEDITQNQQNISSMKEDLIINQEN-----------------------------  385
1500  BgFREP14         ------------------------------------------------------------  181
1501
1502
1503  BgMFREP1 PARTIAL ------------------------------------------------------------  0
1504  BgMFREP2 PRE     ------------------------------------------------------------  168
1505  BgFREP3-2 PRE    ----------------------------------------------VILNQQIIHKIKQ  408
1506  BgMFREP4 PRE     ------------------------------------------------------KE  178
1507  BgMFREP5 PARTIAL ----------------------------------------------------NIISIKE  156
1508  BgMFREP6 PARTIAL ------------------------------------------------------------  0
1509  BgFREP7.1 PRE    ---------------VKEDIAINQQNISSIKTDVAVNEENLINLQKDFNIQQRNISSIKE  449
1510  BgMFREP8 PARTIAL ------------------------------------------------------------  0
1511  BgMFREP9 PARTIAL ------------------------------------------------------------  0
1512  BgMFREP10PARTIAL ------------------------------------------------------------  0
1513  BgMFREP11PARTIAL ------------------------------------------------------------  0
1514  BgFREP12.1 PRE   SHKDELNSFGQIVNSLKDDLRTNKQNLQSVTDEENTNKENINQLKEDFKSNKQKIQNITQ  475
1515  BgFREP12-FBG2    ------------------------------------------------------------  0
1516  BgFREP13.1 PRE   ----------------------------------------LKNVKEDFNIQQRNILSLEK  405
1517  BgFREP14         ------------------------------------------------------------  181
1518
1519
```

```
1520  BgMFREP1 PARTIAL ------------------------------------------DVRGYNVITNRQPVSCR  17
1521  BgMFREP2 PRE     ------------------------------------------GVDSIQISRASPTLPESCR  187
1522  BgFREP3-2 PRE    DLNTYRHN----------MSNIEEHLEVILTNLSTASIKVKNQTDEGSKMSYPPRKSCR  457
1523  BgMFREP4 PRE     ELNALKENHLAA------------------LRS--LDIIKKVNKNLQLSCECLAKPTSCR  218
1524  BgMFREP5 PARTIAL ELQTHRQN----------MSTLKENFETVFSNFSTALIDIKNQIVK-ERSGFKQVLSCR  204
1525  BgMFREP6 PARTIAL ------------------------------------------GVDSIQISRASPTLPESCR  19
1526  BgFREP7.1 PRE    DFKTFYKNT--------------SSFQEIILANLSATLEKVKN------ETELLHPTSCR  489
1527  BgMFREP8 PARTIAL ------------------------------------------------------------  0
1528  BgMFREP9 PARTIAL ------------------------------------------------------------  0
1529  BgMFREP10PARTIAL ------------------------------------------------------------  0
1530  BgMFREP11PARTIAL ------------------------------------------------------------  0
1531  BgFREP12.1 PRE   EVNANRQNIINLNNTSQQSLSKLEADLETQLTNLSTALTQIN------EKIKKGLPSSCR  529
1532  BgFREP12-FBG2    ------------------------------------------------------------  0
1533  BgFREP13.1 PRE   DFHTHQQN----------ISNFQENLEIVLSNFSTALMEVKNQTDK-ERKGDDQITSCR  453
1534  BgFREP14         ---------------------------------TA-------DDILISTTHKTPFTSCR  200
1535
1536  BgMFREP1 PARTIAL DVNSTEDRMVVTLASGLKVMCDTKTDGGGWIIFQRRINGKVDFYRGWKEYRDGFGDYNIG  77
1537  BgMFREP2 PRE     DVISSEDRVVVTLASGLKVMCDTKTDGGGWIIFQRRINGYVDFYRGWKEYRDGFGDYDIG  247
1538  BgFREP3-2 PRE    DVNSTDERVVVTLTSGLKVMCDTKTDGGGWIIFQRRINGNVDFYRGWKEYRDGFGDYNIG  517
1539  BgMFREP4 PRE     DVISTEDRVVVTLASGLEVMCDTTTDGGGWTIFQRRFNGSIDFYRDWKEYRDGFGDYNIG  278
1540  BgMFREP5 PARTIAL DVRSIADRLVVFLISGLKVMCDTKTDGGGWLIFQRRINGKVDFYRGWKEYRDGFGDYDIG  264
1541  BgMFREP6 PARTIAL DVISSEDRVVVTLASGSKVMCDTKTDGGGWIIFRRRINGNVDFYRGWKEYRDGFGSFSIG  79
1542  BgFREP7.1 PRE    KVIYKEDRAIVTLASGLKVMCDTKTDGGGWIIFQRRVNGSVDFYRGWQEYRDGFGDYNIG  549
1543  BgMFREP8 PARTIAL ----------------------------GWTMFQRRITGNLSFYRGWEEYKYGFGDVTTG  32
1544  BgMFREP9 PARTIAL ----------------------------GWIIFQRRINGSVDFYRGWKEYRDGFGDYNIG  32
1545  BgMFREP10PARTIAL ----------------------------GWTIFQRRINGKVDFYRGWKEYRDGFGDYNIG  32
1546  BgMFREP11PARTIAL ----------------------------GWIIFQRRINGKVDFYRGWKEYRDGFGDYTIG  32
1547  BgFREP12.1 PRE   EINSFQERVIVTLTSGLKVMCDTKTDGGGWIIFQRRINGKVDFYRGWKEYRDGFGDYDIG  589
1548  BgFREP12-FBG2    ----------------------TKTDGGGWIIFQRRINGKVNFYRGWKEYRDGFGDYDIG  38
1549  BgFREP13.1 PRE   DVTSKDDRVVVTLASGLKVMCDTKTDGGGWIIFQRRINGKVDFYRNWQVYRDGFGDYDIG  513
1550  BgFREP14         DVNSSDERVVVTLASEQKVMCDTKTDGGGWIIIQRRIMGYVDFYRGWKEYRDGFGDYNIG  260
1551                                          ** :::**. * :.***.*: *: ***.   *
1552
1553  BgMFREP1 PARTIAL EFYLGNENIFMLTSGRQYELRIDMEFKTKKYFAKYSRFKVLSEANNYQLKIESFSGNAGD  137
1554  BgMFREP2 PRE     EFYLGNENIFKLTSSKKYDLRIDLEFNNTKYFAFYTRFEILGEQDYYKLQIGGYSGNAGD  307
1555  BgFREP3-2 PRE    EFYLGNENIYMLTSTGQYNLRIDLKYKNKAFFAQYSGFKILSEKEKYKLNIGAYSGNSGD  577
1556  BgMFREP4 PRE     EFYLGNENIFNLTSSRKYELRFDLEYENKKYFAHYSDFKLLDENNKYKLIIGSYSGNAGD  338
1557  BgMFREP5 PARTIAL EFYLGNENIFNLTSTGKYDLRIDLKYND-FYHAQYSDFQILSEKEKYKLKIGAYSGNAGD  323
1558  BgMFREP6 PARTIAL EFYLGNENIFKLTSSKKYDLRIDLEFNNTKYFALYTSFEILXXQDYYTLQIGGYSGNAGD  139
1559  BgFREP7.1 PRE    EFYLGNENIFKLTSTGKYNLRIDLVLKNQPYYAQYSIFQILSEKNHYKINVGGYSGNAGN  609
1560  BgMFREP8 PARTIAL EFYLGNENMHQITTKRRHELQIDLTFNQAKFSATYSRFSLYGEPEKYRLKVSGYSGNAGD  92
1561  BgMFREP9 PARTIAL EFYLGNENIFNLTSTGQYDLRIDLEYNDKKYFAQYENFKVLSETEKYKLKIRKYSGNAGN  92
1562  BgMFREP10PARTIAL EFYLGNEKVYQLTSEKENELRVDLETNITSYFAYYSEFYLLNETENYKLKVGGFNGTMSD  92
1563  BgMFREP11PARTIAL EFYLGNEYISKLTSTRNFDLRIDFKFNHKTYFVEFSDFRILNETNNYQLKIGKYKGNASD  92
1564  BgFREP12.1 PRE   EFYLGNENIFGLISTGQYDLRIDLEFNNTNYFAQYENFKVLSETEKYKLQTGTYSGNAGD  649
1565  BgFREP12-FBG2    EFYLGNENIFKLTSTGQYDLRIDLEFKNTKYFAQYEDFKVLSETEKYKLQIGDYLGNAGD  98
1566  BgFREP13.1 PRE   EFYLGNENVYNLTTSGNFDLRIDLKFNYQSYFAQYTHFKLLSEKEFYRLKFGDYSGNAGD  573
1567  BgFREP14         EFYLGNENIHQLTSKKPHELRIDLEFDNENYFAQYSSFLLLGEEEFYKLQIGDYSGNAGD  320
1568                   ******* :  : :    :*:.*:  .   : . :  * :    : * :    : *. .:
1569
1570  BgMFREP1 PARTIAL SLT-YHNDQFFSTF------------------------------------------------  150
1571  BgMFREP2 PRE     ALTNIHNDKFFSTYDKDNDLGTSGSCAVTHKGAWWYR-DCYDSNLNGKWGSDR----VNW  362
1572  BgFREP3-2 PRE    NFS-SHNNAFFTTFDRDNDEYS-YNCAVDYTGAWWYHSSCLNCNLNGKWGSSDFAKGVNW  635
1573  BgMFREP4 PRE     SMR-RHVNKFFTTFDKDNDDSPNDNCAIIRRGAWWYQ-NCADVNLNGNWGRGE-PDGVFW  395
1574  BgMFREP5 PARTIAL SLS-YHNDAHFSTYDKDNDNSS-INCASTVSGAWWYK-SCHHVNLNGRWRSKESGKSVIW  380
1575  BgMFREP6 PARTIAL ALTNIHNDKFFSTYDKDNDMDTSGSCAVDYKGAWWYR-SCYDSNLNGKWGSDW----VNW  194
1576  BgFREP7.1 PRE    SLS-YHNDMYFSTFDNDNDAYSGVNCAVKYHGAWWYK-TCYESNLNGKWGSSERDKDLNW  667
1577  BgMFREP8 PARTIAL NLRKFDNVKFTT------------------------------------------------  104
1578  BgMFREP9 PARTIAL GLF-YHNNMFFTT-----------------------------------------------  104
1579  BgMFREP10PARTIAL DLH-HNNNKAFST-----------------------------------------------  104
1580  BgMFREP11PARTIAL DFS-YHNNMQFST-----------------------------------------------  104
1581  BgFREP12.1 PRE   DLS-PHNNKFFSTFDRDNDISSSTNCAEYNSGAWWYE-SCHHSNLNGQWGRTS-NKGMNW  706
1582  BgFREP12-FBG2    DLS-PHNNMFFSTFDRDNDVDSHLNCAEYCS----------------------------  128
1583  BgFREP13.1 PRE   SLS-YHKDMYFSTFDKDNNIDS-RNCATEYWGAWWYR-SCHYSHLNGVWGGKD-PKGLIW  629
1584  BgFREP14         SLA-YQKNMSFSTYDRDNDNASGENCAVSYSGAWWYN-ACHMSNLNGDWGNDAYGKGVNW  378
1585                   :   .     :
```

*Bg*MFREP1 PARTIAL	----------	150
*Bg*MFREP2 PRE	SKLTG-ITKSVTFTEMKIREIELN----	385
*Bg*FREP3-2 PRE	YDLSR-FDSSVSFTEMKIREI-------	655
*Bg*MFREP4 PRE	DNITV--WESVSFSEIKIREIDKEKNKS	421
*Bg*MFREP5 PARTIAL	RALTG-EQESVLYSEMKIREKD------	401
*Bg*MFREP6 PARTIAL	SKLTG-ITKSVTFSEMKIREIELD----	217
*Bg*FREP7.1 PRE	NTLLQTAHVGVSFTEMKIRERE------	689
*Bg*MFREP8 PARTIAL	----------	104
*Bg*MFREP9 PARTIAL	----------	104
*Bg*MFREP10PARTIAL	----------	104
*Bg*MFREP11PARTIAL	----------	104
*Bg*FREP12.1 PRE	FKLTR-GSNSVSFTEMKIREREKNYFH-	732
*Bg*FREP12-FBG2	----------	128
*Bg*FREP13.1 PRE	DKVTN-YEASVSFTEMKIREKS------	650
*Bg*FREP14	DGVTG-LHDSVVFSEMKLRELD------	399

Figure 1—figure supplement 5. Alignment of multiple *Bg*FREP amino acid sequences and distribution of identified peptides. Peptides identified by LC-MS/MS are highlighted in gray. The GenBank accession numbers of each entry are: *Bg*MFREP1 partial, AAK13549; *Bg*MFREP2 precursor, AAK13550; *Bg*FREP3-2 precursor, AAK28656; *Bg*MFREP4 precursor, AAK13551; *Bg*MFREP5 partial, AAK13546; *Bg*MFREP6 partial, AAK13552; *Bg*FREP7.1 precursor, AAK28657; *Bg*MFREP8 partial, AAK13553; *Bg*MFREP9 partial, AAK13554; *Bg*MFREP10 partial, AAK13555; *Bg*MFREP11 partial, AAK13556; *Bg*FREP12.1 precursor, AAO59918; *Bg*FREP12-FBG2, AAT58639; *Bg*FREP13.1 precursor, AAO59922 and *Bg*FREP14, ABO61860.

```
1614    CLUSTAL O (1.2.4) multiple sequence alignment
1615
1616
1617    BgFREP3.3 AEO50747.1      --MLLSLELIRKFPTDTRSSHELVIDAQPEVISLELTPQLVVNCSITDSHVPGLDTINSL  58
1618    BgFREP3.3 AAO59915.1      MARLFLLFILC-VFVVSLAGSELVIDVQPNVISPEITPQLVINCSITNNEVQQLDLIKSL  59
1619    BgFREP3.2 AEO50746.1      MERLFLLLVLC-VFVVSLAGSELVIDVQPNVISPEITPQLVINCSITNNKVQQLDLIKSL  59
1620    BgFREP3.2 AAK28656.1      MARLFLLFILC-VFVVSLAGSELVIDVQPNVISPEITPQLVINCSITNNEVQTLDLIKSL  59
1621    BgFREP3.1 AEO50745.1      MARLFPRLVLC-VFIVPLAGSELVIDVQPNVISPEITPQLVINCSVTNNQVQHLEVIKSL  59
1622    BgMFREP3  AAK13548.1      ---LFLLFILC-VFVVSLAGSELVIDVQPNVISPEITPQLVINCSITNNEVQTLDLIKSL  56
1623                                 *:   ::  .     :. *****.**:*** *:*****:***:*:..*  *: *:**
1624
1625    BgFREP3.3 AEO50747.1      SLSRYNETKKEFDVLLSLDTHTLSLQQLVQFRHAQISFGNLYVTLTLRNPTQSDAKVYRC  118
1626    BgFREP3.3 AAO59915.1      TLSRYNETIREFDDLIALDSLTLNLKQFVRFKYSQISFGNLYITLTLPNPTQFDARIYKC  119
1627    BgFREP3.2 AEO50746.1      TLSRYNETIRDFDELIALDSLTLNLKQFVRFKYSQISFGNRYITLILHNPTQFDARIYKC  119
1628    BgFREP3.2 AAK28656.1      TLSRYNETIREFDELIALDSLTLNLKQFVRFKYSQISFGNRYITLILHNPTQFDARIYKC  119
1629    BgFREP3.1 AEO50745.1      TLSRYNEIIREFDELIALDSLTQNLKQFVRFKYSQISFGNLYITLTLPNPTQFDARIYRC  119
1630    BgMFREP3  AAK13548.1      TLSRYNETIREFDELIALDSLTLNLKQFVRFKYSQISFGNRYITLILHNPTQFDARIYKC  116
1631                              :******  ::** *::**: * .*:*:*:*:::****** *:** * **** **::*:*
1632
1633    BgFREP3.3 AEO50747.1      NVSGDDSLWKNITRVFKKEIKYETNLTVLLEEIRRLREEKDRDQLSCQKEKLN----DSK  174
1634    BgFREP3.3 AAO59915.1      NATGANSDGTNISLFAKKAVEYETNSTALIEEIRRIKKDENY--CSFKKDDLADIKQRSR  177
1635    BgFREP3.2 AEO50746.1      NATGDNSEGANISLFAKKGVEYETNSTALIEEIRRIKKDENY--CSFKKDDLSDSKQRSR  177
1636    BgFREP3.2 AAK28656.1      NATGTNSEGANISLFAKKAVEYETNSTALIEEIRRIKKDENY--CSFKKDDLSDSKQRSR  177
1637    BgFREP3.1 AEO50745.1      NADGANSEGTNISLFTKKAVEYETNSTALIEEIRRIKKDENK--CSLKKDDLSDIKQRWR  177
1638    BgMFREP3  AAK13548.1      NATGTNSEGANISLFAKKAVEYETNSTALIEEIRRIKKDENY--CSFKKDDLSDXKQRSR  174
1639                              *. * :*   **: . ** ::**** *.*:*****:::::    * :*:.*       :
1640
1641    BgFREP3.3 AEO50747.1      LHFVGNSKVVKELFDPLTLTCSIQDLMNDRNETSTVQSIYILHEANGIIATISKDQPVVT  234
1642    BgFREP3.3 AAO59915.1      VYFSGSSDIIKERIEPLTLKCTFQVLKTDQNDTSRLQSLYILHESKGVIAYVNKDQPVVT  237
1643    BgFREP3.2 AEO50746.1      VYFSGSSDIIKERIEPLTLKCTFQVLKTDQNEISRLQSLYILHETKGVIAYVNKDQPVVT  237
1644    BgFREP3.2 AAK28656.1      VYFSGSSDIIKERIEPLTLKCTFQVLKTDQNETSRLQSLYILHESKGVIAYVNKDQPVVT  237
1645    BgFREP3.1 AEO50745.1      VYFSESSKIIKERIEPLTLKCTFQILTPDENETSRLQSLYILHESNGVVANINKDQAVIT  237
1646    BgMFREP3  AAK13548.1      VYFSGSSDIIKERIEPLTLKCTFQVLKTDQNETSRLQSLYILHESKGVIAYVNKDQPVVT  234
1647                              ::*  .*.::** ::****.*::* *  *.*: * :**:*****::*::* :.*** *:*
1648
1649    BgFREP3.3 AEO50747.1      TNQDVNLLAVIGKLHDDSSKNSYLQVTWSNPKFSESGKYFCGAHANNAQGRNEHWNEMLT  294
1650    BgFREP3.3 AAO59915.1      SLQGSNIQDVEGEIYDNAIKDSYLQVTWSNLKHSESGKYFCEAHNKYSEGRIDKSSNMLT  297
1651    BgFREP3.2 AEO50746.1      SLQGSHIQDVEGEIYDNAIKDSYLQVTWSNLKHTESGKYFCEAHNQYSEGRIDKTSNMLT  297
1652    BgFREP3.2 AAK28656.1      SLQGSNIQDVEGEIYDNAIKDSYLQVTWSNLKHSESGKYFCEAHNKYSEGRIDKSSNMLT  297
1653    BgFREP3.1 AEO50745.1      TIQGGNFENAQGEISGDQSKESYLQVTWSNLKHSDSGKYFCEAHVKHSKGKAERLNEMLT  297
1654    BgMFREP3  AAK13548.1      SLQGSNIQDVEGEIYDNAIKDSYLQVTWSNLKHSESGKYFCEAHNXYSEGRIDXSSNMLT  294
1655                              : *. ::  . *:: .:  *:********* *.::****** **   ::*: :  .:***
1656
1657    BgFREP3.3 AEO50747.1      ITVERLQFDDIVKVMYDIQRQVDEDKKRLQTFHENLTNNFIILNTNLQSIENVRRDVRTN  354
1658    BgFREP3.3 AAO59915.1      ITVERPTFDDLVEAMHKLFTQVDGAKESLKAINQNIKN----------------------  335
1659    BgFREP3.2 AEO50746.1      ITVERPTFDDLVEAMHKLFTQVDGAKESLKAINQNIKN----------------------  335
1660    BgFREP3.2 AAK28656.1      ITVERPTFDDLVEAMHKLFTQVDGAKESLKAINQNIKN----------------------  335
1661    BgFREP3.1 AEO50745.1      IEVISPTIDDLMEVIQKLVTQVDGDKESLQDVKQNIMN----------------------  335
1662    BgMFREP3  AAK13548.1      ITVERPTFDDLVEAXXKLFTQVDGAKESLKAINQNIKN----------------------  332
1663                              * *    :**:::.  .:  ***  *: *: .::*: *
1664
1665    BgFREP3.3 AEO50747.1      QESINGIKDELLSNKQNIVNNKKDINTAKESINVIKEELLSNKQNIVNNKRDIDTMEESI  414
1666    BgFREP3.3 AAO59915.1      ----------------------INKDLDFKEQNITSIKQEVIRNQNNIQILSEDLNIKE---  372
1667    BgFREP3.2 AEO50746.1      ----------------------INKDLDFKEQNITSIKEEVIRNQNNIQILSEDSNIKE---  372
1668    BgFREP3.2 AAK28656.1      ----------------------INKDLDFKEQNITSIKQEVIRNQNNIQILSEDLNIKE---  372
1669    BgFREP3.1 AEO50745.1      ----------------------IKEDLNTKEQNIISIKEDLNTKQQSIISIKEEFKTKQ---  372
1670    BgMFREP3  AAK13548.1      ----------------------INKDLDFKEQNITSIKQEVIRNQNNIQILSEDLNIKE---  369
1671                                                    ::*::  ::.*  **:::  :::.*   ..: .  :
1672
1673    BgFREP3.3 AEO50747.1      NVIRHELLSNKQNIVNNKKDIHTTQESIIGILEELLRNQQNIVNNKKDINTTQDSIREIQ  474
1674    BgFREP3.3 AAO59915.1      ------------------------------------QNLTSIKADLSTKQQTFLNIK  393
1675    BgFREP3.2 AEO50746.1      ------------------------------------QNMTSIREDLSTKQQTFLNIK  393
1676    BgFREP3.2 AAK28656.1      ------------------------------------QNLTSIKADLSTKQQTFLNIK  393
1677    BgFREP3.1 AEO50745.1      ------------------------------------EN---IQKDVTINQQNIQKIK  390
1678    BgMFREP3  AAK13548.1      ------------------------------------QNLTSIKADLSTKQQTFLNIK  390
1679                                                                  :*    : *:. .*:.: :*:
```

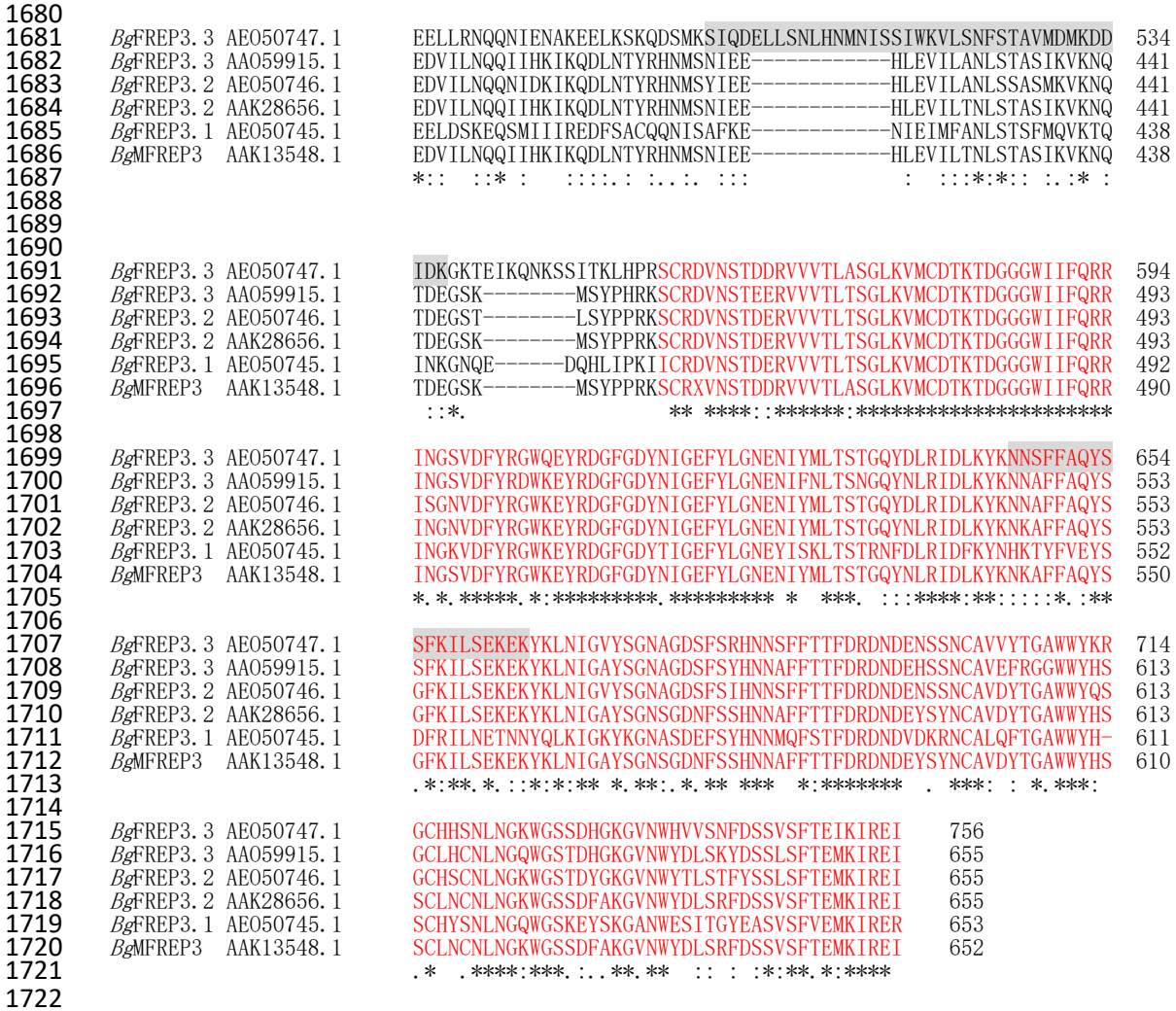

Figure 1—figure supplement 6. Alignment of multiple *Bg*FREP3 amino acid sequences and distribution of identified peptides. Peptides identified by LC-MS/MS are highlighted in gray, the overlapping region of peptides are bold underline. GenBank annotations of *Bg*FREP3.2 (AAK28656.1) are used for locate the individual *Bg*FREP3 domains: orange, signal peptide; green, IgSF1 domain; blue, small connecting region; pink, IgSF2 domain; black, ICR; red, FBG domain.

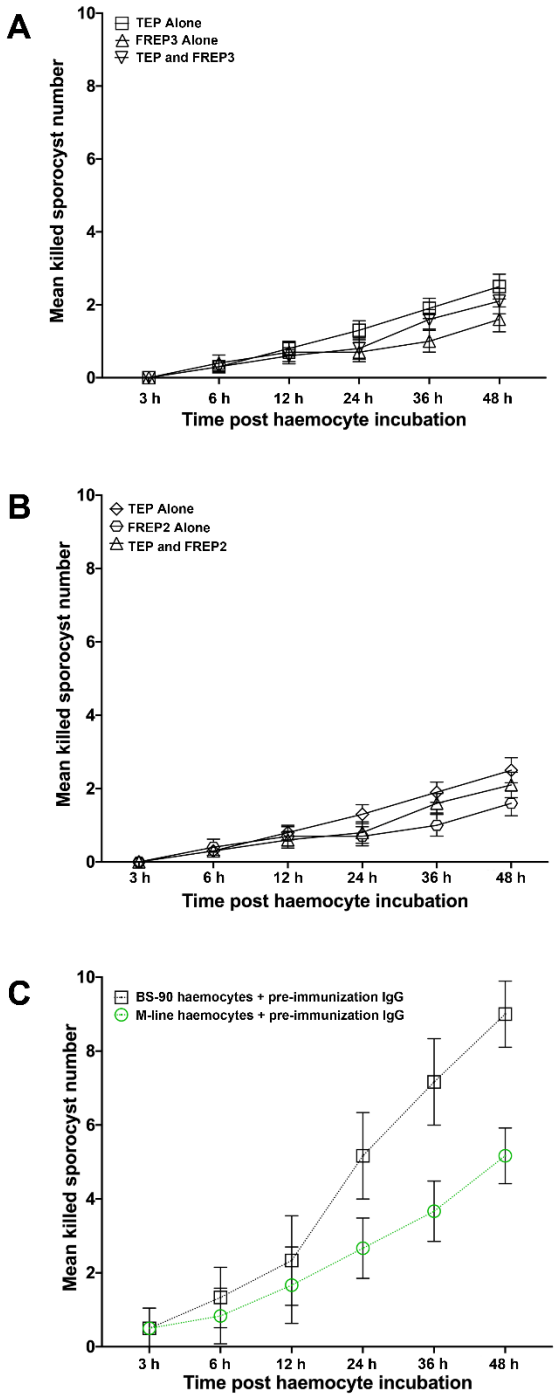

Figure 4—figure supplement 1.

Figure 4—figure supplement 1. Controls of functional verification experiments. A: Only the media and recombinant proteins r*Bg*FREP3 and r*Bg*TEP1 were present, without

snail plasma or haemocytes, to control the sporocyst killing assay of r*Bg*FREP3, r*Bg*TEP1 and the combination of both. **B:** The controls of the sporocyst killing assay of r*Bg*FREP2, r*Bg*TEP1 and r*Bg*FREP2-r*Bg*TEP1 complex. **C:** Pre-immune serum from the same rabbit which raised the anti-r*Bg*FREP3 antibodies were incubated with BS-90 and M-line haemocytes for control.

1741 **Table S1** Primer list for cloning and quantitative RT-PCR.
1742

Name	Protein	Primer sequence
For cloning:		
*Bg*MFREP2 precursor (AY012700.1)	AAK13550.1	Forward:5'- CAC CAG GAG GAT CTT AGT AAT GGC GTC GC -3' Reverse: 5'- GAC CCT TGG CGC GTT TAG CTC TAT TTC TCT -3'
*Bg*FREP3.2 precursor (Derived from AY028461.1)	AAK28656.1	Forward: 5'- CAC CAG GAG AAG AAA CGA AAT GGC TCG TCT CTT CTT GCT CTT C -3' Reverse: 5'- CAT CTC AGT GAA CGA CAC GGA -3'
*Bg*TEP1.1 (HM003907.1)	ADE45332.1	Forward: 5'- CAC CAT GAG AAT GAA GCT GAA T -3' Reverse: 5'- TGG GCA ACA GTT GAG GCA -3'
For q-PCR:		
*Bg*TEP1.1		Forward: 5'-CACCAATGGCGATGTCATTTAG-3' Reverse: 5'-CTGACTCTCTCAACAGGCTTAG-3'
*Bg*Actin		Forward: 5'-GCTTCCACCTCTTCATCTCTT - 3' Reverse: 5' -GAACGTAGCTTCTGGACATCTG-3'

1743

1744 **Table S2** The identified peptides of *Bg*TEP1, Biomphalysin, *Bg*FREP2 and *Bg*FREP3.3
1745 by LC-MS/MS.
1746
1747

	Sequence	# PSMs	# Protein groups	Protein Group Accessions	# times identified
***Bg*TEP1:**					
P1	FQVDVGLPPFALLSDTTLSGSVEAK	2	1	A0A2C9M761	1
P2	YTFGQPVYGLVLLQIGENVDTIDK	4	2	A0A2C9L049; A0A2C9M761	2
P3	KVTEISFEIK	1	1	A0A2C9M761	1
P4	ITAFVTEASTGIK	1	1	A0A2C9L049	1
P5	SYSMSNNYLQLSLLSK	3	1	A0A2C9LWE8	3
P6	ITSTEAIDSLAYEIR	3	1	A0A2C9LWE8	3
P7	SGVLELNGQR	7	2	A0A2C9LWE8; D7QYI1	4
P8	QGMMGAPMAMSFR	1	1	D7QYI1	1
P9	VFRPFFVSLTYPR	1	1	D7QYI1	1
P10	ENPFLTPITPGPGNQASNIQVR	1	1	D7QYI1	1
P11	SNMAADAIVR	2	1	D7QYI1	2
P12	ATNQLSEELNK	2	1	D7QYI1	2
P13	IIDLLNNGYQR	11	2	A0A2C9KIQ1; D7QYI1	3
P14	QNIDGSFNEFGK	3	2	A0A2C9KIQ1; D7QYI1	3
P15	SQLAQSTFEK	2	1	D7QYI1	1
P16	LDEALPSVR	3	1	D7QYI1	2
P17	KGSIVAIYFDK	1	1	D7QYI1	1
Biomphalysin:					
P1	CNEINSYINALDR	6	2	A0A182YTN9; A0A182YTZ4	5
P2	VGYFLQGLYR	9	2	A0A182YTN9; A0A182YTZ4	5
P3	GYLFNLESAR	2	2	A0A182YTN9; A0A182YTZ4	2
P4	GQCSCPGGYFLTGLYR	5	2	A0A182YTN9; A0A182YTZ4	4
P5	ADGDDLYFLK	2	2	A0A182YTN9; A0A182YTZ4	2
P6	NGFTWAADTR	7	2	A0A182YTN9; A0A182YTZ4	4
P7	WCEGDKNGER	1	2	A0A182YTN9; A0A182YTZ4	1
P8	LNLVFGDWGFAVK	2	2	A0A182YTN9; A0A182YTZ4	3
P9	SVIEDLQAESVDSGVLYNR	9	2	A0A182YTN9; A0A182YTZ4	6
P10	SSITLGPMEGAK	1	2	A0A182YTN9; A0A182YTZ4	1
P10/2	QSSITLGPMEGAK	1	1	A0A182YTZ4	1
P11	FGDSSVPFYK	3	2	A0A182YTN9; A0A182YTZ4	3
P12	LTDETQYQFTLTGK	7	1	A0A182YTN9;	6
P13	LEKVEGTSVNVK	2	2	A0A182YTN9; A0A182YTZ4	1
***Bg*FREP2:**					
P1	YQPVATSLYPSVTK	16	1	A0A2C9L9F5	9
P2	SSTDDLAVALSYIQDR	5	1	A0A2C9L9F5	5
P3	LDKDGVDSIQISR	22	1	A0A2C9L9F5	9
P4	TDGGGWIIFQR	13	1	A0A2C9L9F5	9
P5	DCYDSNLNGK	1	1	A0A2C9L9F5	1
***Bg*FREP3.3:**					
P1	SIQDELLSNLHNMNISSIWKVLSNFST AVMDMKDDIDK	2	1	AEO50747.1	1
P2	DIDTMEESINVIRHELLSNK	1	1	AEO50747.1	1
P3	NNSFFAQYSSFKILSEKEK	3	1	AEO50747.1	1
P4	NEHWNEMLTITVERLQFDDIVK	1	1	AEO50747.1	1
P5	KDINTTQDSIR	2	1	AEO50747.1	1
P6	QNIVNNKR	2	1	AEO50747.1;	1
P7	LQFDDIVK	1	2	AEO50747.1; XP_013062145.1	1
P8	ESLKAINQNIKNINK	8	8	AAO59915.1; AAK28656.1; AEO50746.1; AAK13548.1; AAC47701.1; AQX34544.1; AQX34587.1; AQX34558.1.	1

P9	AINQNIKNINKDLDFK	2	8	AAO59915.1; AEO50746.1; AAC47701.1; AQX34587.1;	AAK28656.1; AAK13548.1; AQX34544.1; AQX34558.1.	1

PSMs: peptide-to-spectrum matches.

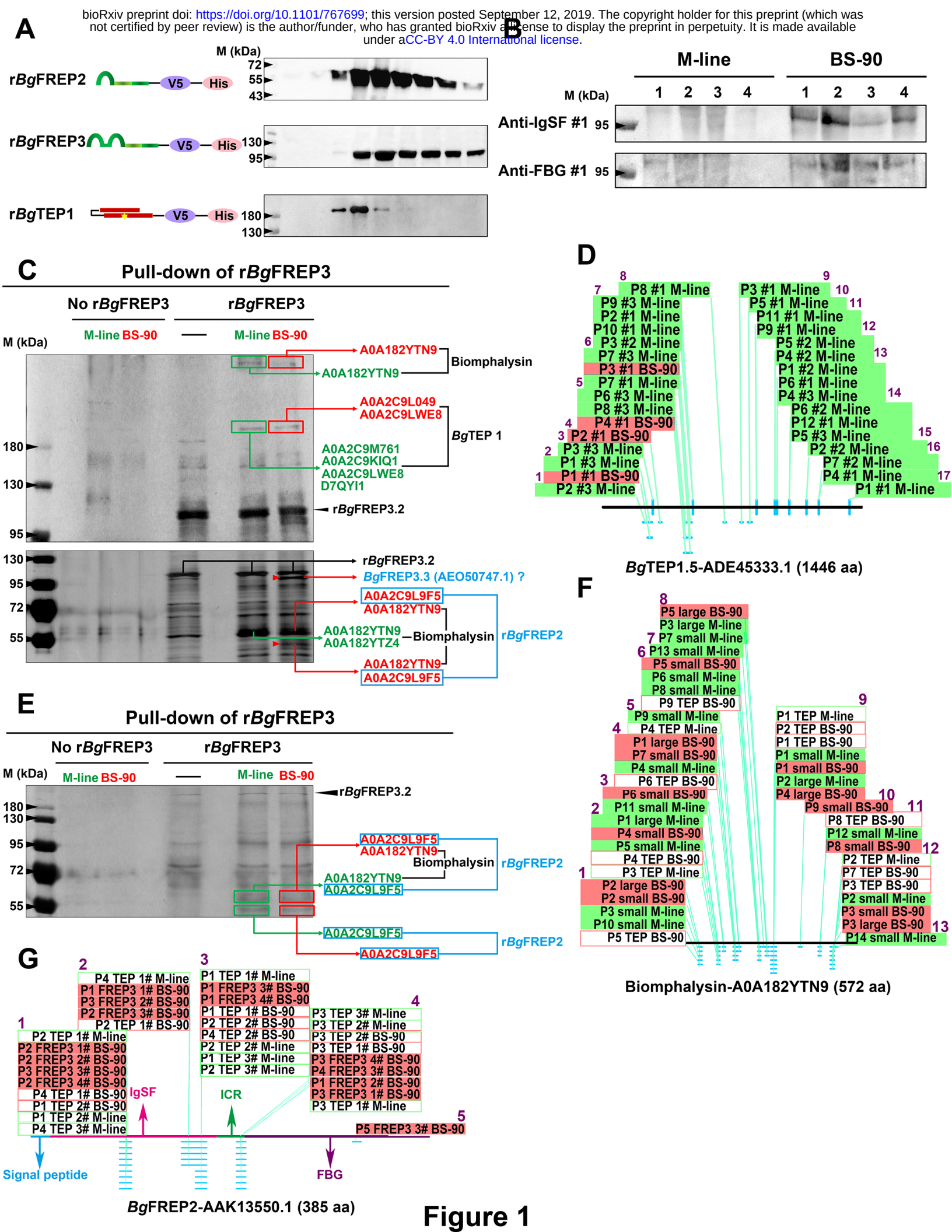

Figure 1

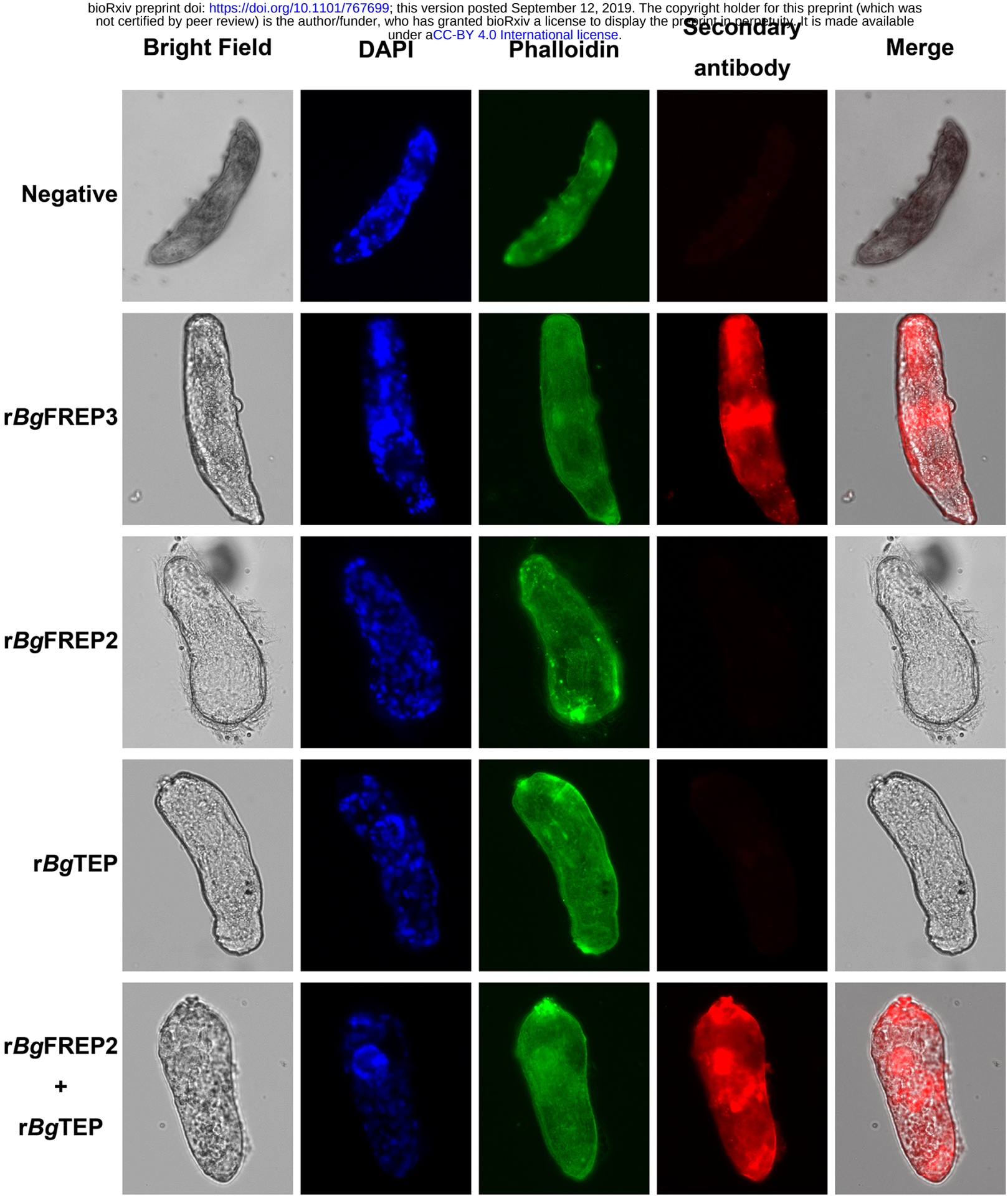

Figure 2

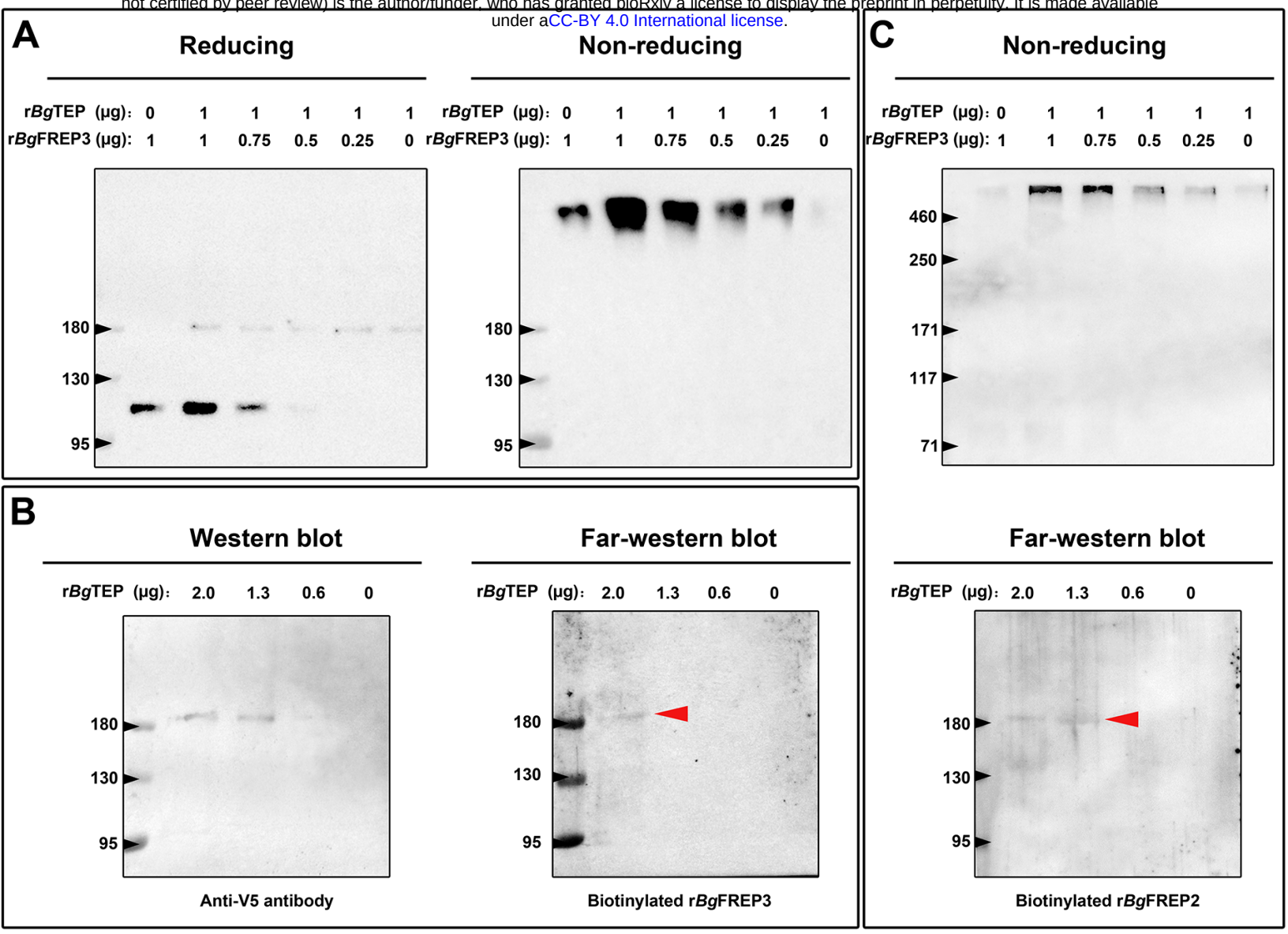

Figure 3

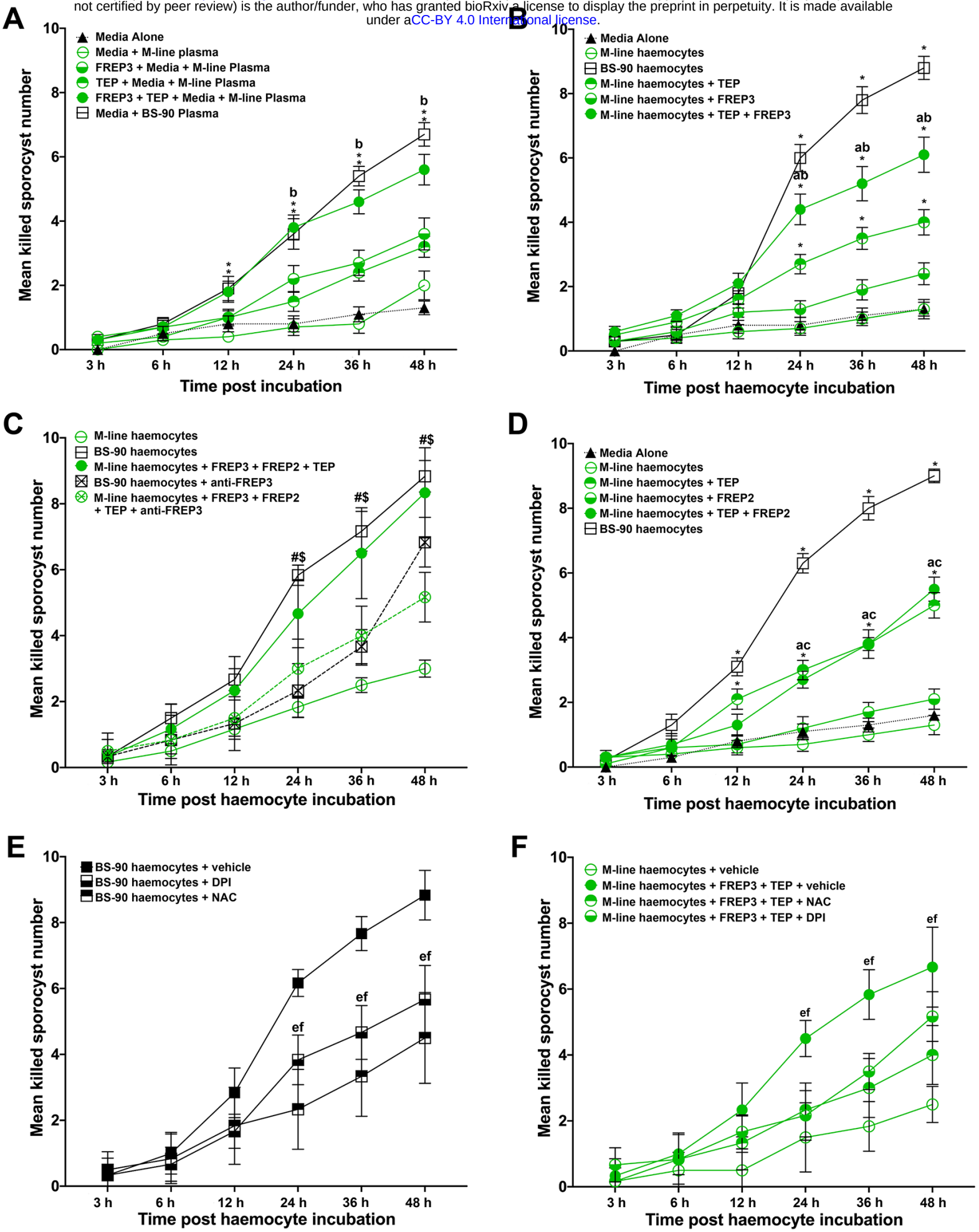

Figure 4

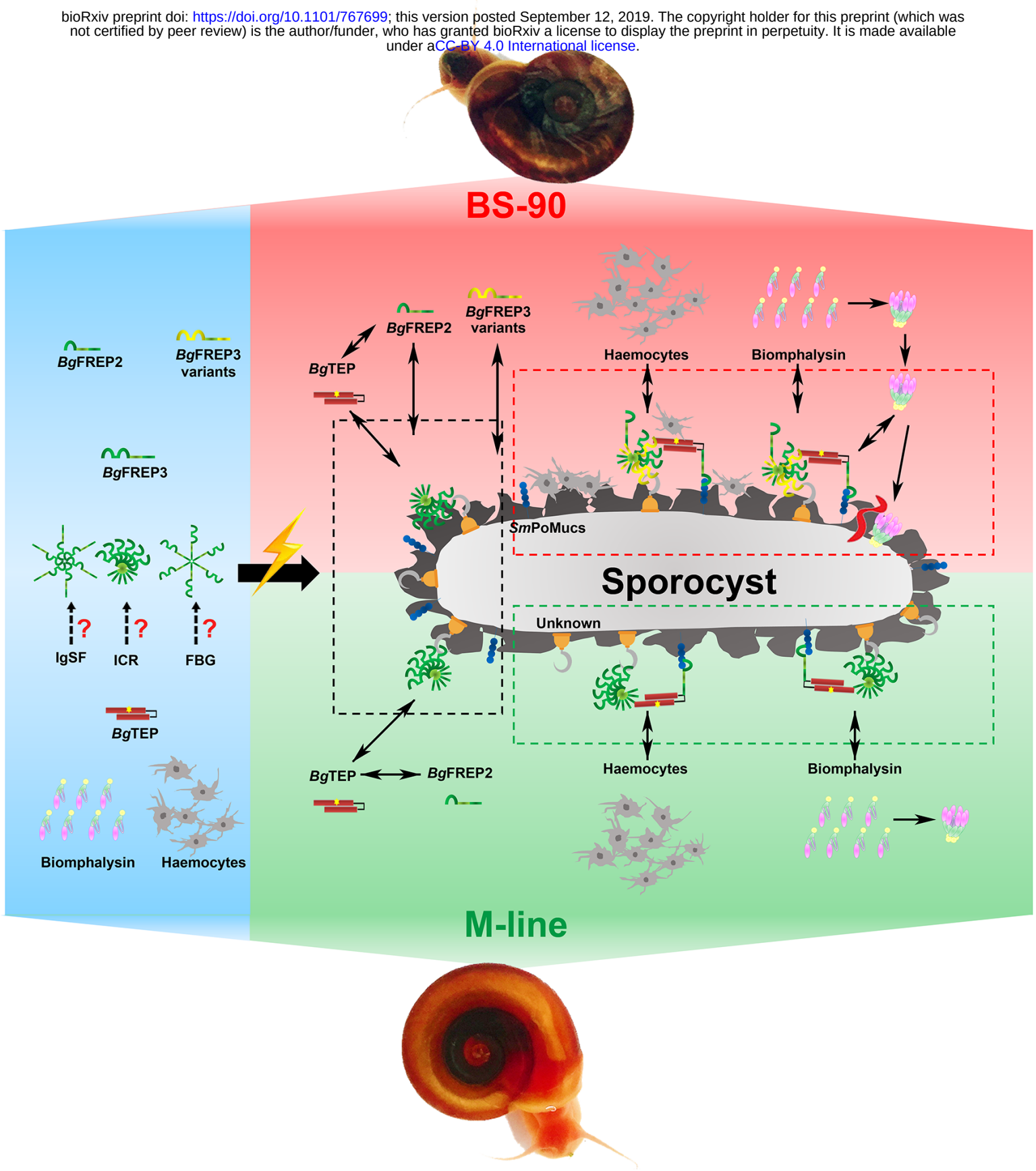

Figure 5

>Derived from GenBank AY028461.1 *Biomphalaria glabrata* fibrinogen-related protein 3-2 precursor, gene, complete cds.

ATGGCTCGTCTCTTCTTGCTCTTCATTCTCTGCGTCTTCGTGGTGTCACTCGCTGGCTCGGA
ACTCGTCATTGATGTCCAACCCAACGTCATTTCCCCGGAAATCACCCCTCAGCTGGTCATC
AACTGCAGCATCACGAACAATGAGGTCCAAACACTGGACCTCATCAAGTCTTTGACCCTG
TCAAGGTACAACGAGACTATTCGCGAGTTCGATGAATTGATCGCCCTGGACAGCCTCACCT
TGAACCTGAAGCAGTTCGTGCGCTTCAAGTACTCGCAAATCAGTTTCGGCAACCGTTACAT
CACGCTCATTTTGCACAATCCCACACAGTTCGACGCCCGCATCTACAAGTGCAATGCTACG
GGCACAAACTCCGAGGGTGCCAATATCAGCCTGTTCGCCAAGAAGGCTGTGGAGTACGAA
ACCAACTCTACTGCTCTGATCGAGGAAATTCGCCGTATCAAGAAGGATGAGAACTACTGCA
GCTTCAAGAAGGACGATCTGTCTGACTCAAAGCAGCGTTCTAGGGTGTACTTCTCAGGTT
CCAGCGATATCATTAAGGAGCGTATCGAACCCCTGACGCTCAAGTGCACATTCCAGGTCCT
GAAGACCGACCAAAACGAAACTTCCAGGCTCCAGAGCTTGTACATTCTGCACGAGTCCAA
GGGAGTCATCGCCTACGTGAACAAGGATCAGCCAGTGGTCACCTCGCTGCAGGGCAGTAA
CATCCAAGACGTGGAGGGTGAAATTTACGATAATGCCATCAAGGACTCGTACTTGCAGGTC
ACTTGGAGTAACCTGAAGCACTCCGAGAGCGGAAAGTACTTCTGCGAAGCTCATAACAAG
TACTCGGAGGGCAGGATTGATAAGTCGAGTAATATGCTGACCATCACTGTCGAACGCCCCA
CCTTCGACGATCTCGTGGAGGCTATGCACAAGTTGTTCACTCAGGTGGACGGTGCCAAGG
AGTCCCTGAAGGCTATCAACCAAAACATCAAGAACATCAATAAGGACCTGGATTTCAAGG
AACAGAACATTACCTCGATCAAGCAAGAGGTCATCCGCAATCAGAACAATATCCAAATTCT
GTCCGAGGATCTCAACATTAAGGAACAGAATCTGACGTCGATCAAGGCCGACCTCAGTAC
CAAGCAGCAAACTTTCTTGAACATTAAGGAGGATGTGATCCTGAATCAGCAAATCATTCAT
AAGATCAAGCAGGACCTGAACACCTACCGCCACAATATGTCCAACATTGAGGAACATCTG
GAAGTGATCCTCACGAACTTGTCTACAGCTTCAATCAAGGTCAAGAATCAGACGGATGAG
GGATCCAAGATGAGCTACCCCCCACGCAAGTCTTGCCGTGATGTGAACTCAACAGACGAG
CGTGTGGTCGTGACGTTGACAAGCGGCCTGAAGGTCATGTGCGATACCAAGACTGACGGC
GGTGGATGGATCATTTTCCAGAGGCGCATCAACGGTAATGTGGACTTCTACCGTGGATGGA
AGGAGTACAGGGATGGCTTCGGTGACTACAACATTGGAGAGTTCTACCTCGGCAACGAAA
ATATCTACATGTTGACGTCGACAGGCCAGTACAACCTCAGGATCGATTTGAAGTACAAGAA
TAAGGCCTTCTTCGCTCAATACTCCGGTTTCAAGATTCTGAGCGAGAAGGAAAAGTACAA
GCTCAACATCGGAGCCTACTCGGGAAATAGTGGCGACAACTTCTCTTCACACAACAATGCT
TTCTTCACCACTTTCGACCGTGATAACGACGAGTACTCCTACAATTGCGCCGTGGATTACA
CCGGCGCTTGGTGGTACCATTCCAGCTGCCTGAACTGCAATCTCAACGGCAAGTGGGGAT
CGAGTGACTTCGCCAAGGGTGTCAACTGGTACGACCTGAGCCGTTTCGATTCGTCCGTGTC
GTTCACTGAGATGAAGATTAGGGAGATT~~TAA~~

Figure 1—figure supplement 1. Nucleotide sequence for cloning r*Bg*FREP3. This sequence is derived from GenBank AY028461.1. There are differences in nucleotide sequence but the amino acids match 100%. The differences are due to codon-optimization for expression in insect Sf9 cells. GenBank annotations are used to locate the individual *Bg*FREP3.2 domains: underline, signal peptide; gray shading, IgSF1 domain; double underline, small connecting region; gray shading plus frame, IgSF2 domain; wavy underline, ICR; bold underline, FBG domain; delete line, termination codon.

44

```
CLUSTAL O (1.2.4) multiple sequence alignment

BgTEP           MRMKLNLILFVFYLVFQECQGGKYFISAPRNVVPGTAYDISVDILKQDIGNVTVEAILQD   60
BgTEP1.1        MRMKLNLILFVFYLVFQECQGGKYFISAPRNVVPGTAYDISVDILKQDIDNVTVEAILQD   60
BgTEP1.2        MRMKLNLILFVFYLVFQECQGGKYFISAPRNVVPGTAYDISVDILKQDIDNVTVEAILQD   60
BgTEP1.3        MRMKLNLILFVFYLVFQECQGGKYFISAPRNVVPGTAYDISVDILKQDIDNVTVGAILQD   60
BgTEP1.4        MRMKLNLILFVFYLVFQECQGGKYFISAPRNVVPGTAYDISVDILKQDIDNVTVEAILQD   60
BgTEP1.5        MRMKLNLILFVFYLVFQECQGGKYFISAPRNVVPGTAYDISVDILKQDIGNVTVEAILQD   60
                ************************************************.**** *****

BgTEP           YSFSIPEGPKSLLTANGTFSPGVRGTLSMPIDFNLHCSYCRILLKGYNPLQFEQDIFIQI   120
BgTEP1.1        YSFSIPEGPKSLLTANGTFSPGVRGTLSMPIDFNLHCSYCRILLKGYNPLQFEQDIFIQI   120
BgTEP1.2        YSFSIPEGPRSLLTANGTFSPGVRGTLSMPIDFNLHCSYCRILLKGYNPLQFEQDIFIQI   120
BgTEP1.3        YSFSIPEGPKSLLTANGTFSPGVRGTLSMPIDFNLHCSYCRILLKGYNPLQFEQDIFIQI   120
BgTEP1.4        YSFSIPEGPKSLLTANGTFSPGVRGTLSMPIDFNLHCSYCRILLKGYNPLQFEQDIFIQI   120
BgTEP1.5        YSFSIPEGPKSLLTANGTFSPGVRGTLSMPIDFNLHCSYCRILLKGYNPLQFEQDIFIQI   120
                *********:**************************************************

BgTEP           SSDILSILIQTDKAIYKPKERVNFRILAAYYNLQLYTGTFHYEILDPYDNKINVLSGVSG   180
BgTEP1.1        SSDILSILIQTDKAIYKPKERVNFRILAAYYNLQLYTGTFHYEILDPYDNKINVLSGVSG   180
BgTEP1.2        SSDILSILIQTDKAIYKPKERVNFRILAAYYNLQLYTGTFHYEILDPYDNKINVLSGVSG   180
BgTEP1.3        SSDILSILIQTDKAIYKPKERVNFRILAAYYNLQLYTGTFHYEILDPYDNKINVLSGVSG   180
BgTEP1.4        SSDILSILIQTDKAIYKPKERVNFRILAAYYNLQLYTGTFHYEILDPYDNKINVLSGVSG   180
BgTEP1.5        SSDILSILIQTDKAIYKPKERVNFRILAVYYNLQLYTGTFHYEILDPYDNKINVLSGVSG   180
                ****************************.**************************************

BgTEP           TFGVVEGFFDLSDQPSFGTWKINVRTETVSGEESQLFEVAEYDLPRFQVDVGLPPFALLS   240
BgTEP1.1        TFGVVEGFFDLSDQPSFGTWKINVRTETVSGAESQFFEVAEYDLPRFQVDVGLPPFALLS   240
BgTEP1.2        TFGVVEGFFDLSDQPSFGTWKINVRTETVSGAESQFFEVAEYDLPRFQVDVGLPPFALLS   240
BgTEP1.3        TFGVVEGFFDLSDQPSFGTWKINVRTETVSGAESQFFEVAEYDLPRFQVDVGLPPFALLS   240
BgTEP1.4        TFGVVEGFFDLSDQPSFGTWKINVRTETVSGAESQFFEVAEYDLPRFQVDVGLPPFALLS   240
BgTEP1.5        TFGVVEGFFDLSDQPSFGTWKINVRTETVSGAESQFFEVAEYDLPRFQVDVGLPPFALLS   240
                ******************************** ***:************************

BgTEP           DTTLSGSVEAKYTFGQPVYGLVLLQIGENVDTTDKCNVNRKVTEISFEIKGKENFSVPLE   300
BgTEP1.1        DTTLSGSVEAKYTFGQPVYGLVLLQIGENVDTTDKCNVNRKVTEISFEIKGKGNFSVPLE   300
BgTEP1.2        DTTLSGSVEAKYTFGQPVYGLVLLQIGENVDTTDKCNVNRKVTEISFEIKGKGNFSVPLE   300
BgTEP1.3        DTTLSGSVEAKYTFGQPVYGLVLLQIGENVDTTDKCNVNRKVTEISFEIKGKGNFSVPLE   300
BgTEP1.4        DTTLSGSVEAKYTFGQPVYGLVLLQIGENVDTTDKCNVNRKVTEISFEIKGKGNFSVPLE   300
BgTEP1.5        DTTLSGSVEAKYTFGQPVYGLVLLQIGENVDTTDKCNVNRKVTEISFEIKGKGNFSVPLE   300
                ************************************************** *******

BgTEP           DIRRSVHLNEKKKIKITAFVTEASTGIKLNGSSVITYYGNRYQIKFLEMTPAVFKPGLQY   360
BgTEP1.1        DIRRSVHLNEKKKIKITAFVTEASTGIKLNGSSVITYYGNRYQIKFLEMTPAVFKPGLQY   360
BgTEP1.2        DIRRSVHLNEKKKIKITAFVTEASTGIKLNGSSVITYYGNRYQIKFLEMTPAVFKPGLQY   360
BgTEP1.3        DIRRSVHLNEKKKIKITAFVTEASTGIKLNGSSVITYYGNRYQIKFLEMTPAVFKPGLQY   360
BgTEP1.4        DIRRSVHLNEKKKIKITAFVTEASTGIKLNGSSVITYYGNRYQIKFLEMTPAVFKPGLQY   360
BgTEP1.5        DIRRSVHLNEKKKIKITAFVIEASTGIKLNGSSVITYYGNRYQIKFLEMTPAVFKPGLQY   360
                ******************** ****************************************

BgTEP           TAYVQVTTPDGLPPTDSNLSLSVYTSVTYQMTVPDQELYSPSSFSGSYPLPGQNMSLPAN   420
BgTEP1.1        TAYVQVTTPDGLPPTDSNLSLSVYTSVTYQMTVPDQELYSPSSFSGSYPLPGQNMSLPAN   420
BgTEP1.2        TAYVQVTTPDGLPPTDSNLSLSVYTSVTYQMTVPDQELYSPSSFSGSYPLPGQNMSLPAN   420
BgTEP1.3        TAYVQVTTPDGLPPTDSNLSLSVYTSVTYQMTVPDQELYSPSSFSGSYPLPGQNMSLPAN   420
BgTEP1.4        TAYVQVTTPDGLPPTDSNLSLSVYTSVTYQMTVPDQELYSPSSFSGSYPLPGQNMSLPAN   420
BgTEP1.5        TAYVQVTTPDGLPPTESNLSLSVYTSVTYQMTVPDQELYSPSSFSGSYPLPGQNMSLPAN   420
                ***************:************************************************

BgTEP           GILSIDIDIPLNATSIDIKVSLNKQTTAEKRISKSYSMSNNYLQLSLLSKLVKAESDVLI   480
BgTEP1.1        GILSIDIDIPLNATSIDIKVSLNKETTAEKRISKSYSMSNNYLQLSLLSKLVKAESDVLI   480
BgTEP1.2        GILSIDIDIPLNATSIDIKVSLNKETTAEKRISKSYSMSNNYLQLSLLSKLVKAESDVLI   480
BgTEP1.3        GILSIDIDIPLNATSIDIKVSLNKETTAEKRISKSYSMSNNYLQLSLLSKLVKAESDVLI   480
BgTEP1.4        GILSIDIDIPLNATSIDIKVSLNKETTAEKRISKSYSMSNNYLQLSLLSKLVKAESDVLI   480
BgTEP1.5        GILSIDIDIPLNATSIDIKVSLNKQTTAEKRISKSYSMSNNYLQLSLLSKLVKAESDVLI   480
                ************************:************************************

BgTEP           KITSTEAIDSLAYEIRSRSDRVKSGVLELNGQREFNATFKVEPSWAPIAQLLMYYIRRDS   540
BgTEP1.1        KITSTEAIDSLAYEIRSRSDHVKSGVLELSGQREFNATFKVEPSWAPIAQLLMYYIRRDS   540
BgTEP1.2        KITSTEAIDSLAYEIRSRSDHVKSGVLELSGQREFNATFKVEPSWAPIAQLLMYYIRRDS   540
```

```
70   BgTEP1.3      KITSTEAIDSLAYEIRSRSDHVKSGVLELSGQREFNATFKVEPSWAPIAQLLMYYIRRDS   540
71   BgTEP1.4      KITSTEAIDSLAYEIRSRSDHVKSGVLELSGQREFNATFKVEPSWAPIAQLLMYYIRRDS   540
72   BgTEP1.5      KITSTEAIDSLAYEIRSRSDHVKSGVLELNGQREFNATFKVGPSWAPIAQLLMYYIRRDS   540
73                 ********************:********.********** ******************
74
75   BgTEP         NEVVTDSLAFNVEGMFKNKVNVAFKENETDINKNVSLELPADSDSQIYVLAVDXSVLLLK   600
76   BgTEP1.1      NEVVTDSLAFNVEGMFKNKVNVAFKENETDINKNVSLELSADSDSQIYVLAVDQSVLLLK   600
77   BgTEP1.2      NEVVTDSLAFNVEGMFKNKVNVAFKENETDINKNVSLELSADSDSQIYVLAVDQSVLLLK   600
78   BgTEP1.3      NEVVTDSLAFNVEGMFKNKVNVAFKENETDINKNVSLELSADSDSQIYVLAVDQSVLLLK   600
79   BgTEP1.4      NEVVTDSLAFNVEGMFKNKVNVAFKENETDINKNVSLELSADSDSQIYVLAVDQSVLLLK   600
80   BgTEP1.5      NEIVTDSLAFNVEGMFENKVNVAFKENETDINKNVSLELSADSDSQIYVLAVDQSVLLLK   600
81                 **:**************:********************** ************* ******
82
83   BgTEP         TGNDXTPXKVKDSFXSXFHKGAIPTDSNFALSYSGSSIXEVFSNMGLVIATDLNIFAPFR   660
84   BgTEP1.1      TGNDLTPNKVKDSFISKFHKGAIPTDSNFALSYSGSSINEVFSNMGLVIATDLNIFAPFR   660
85   BgTEP1.2      TGNDLTPNKVKDSFISKFHKGAIPTDSNFALSYSGSSINEVFSNMGLVIATDLNIFAPFR   660
86   BgTEP1.3      TGNDLTPNKVKDSFISKFHKGAIPTDSNFALSYSGSSINEVFSNMGLVIATDLNIFAPFR   660
87   BgTEP1.4      TGNDLTPNKVKDSFISKFHKGAIPTDSNFALSYSGSSINEVFSNMGLVIATDLNIFAPFR   660
88   BgTEP1.5      TGNDLTPNKVKDSFISKFHKGEVPTDSNFALSYSGSSINEVFSNMGLVIATDLNIFAPFR   660
89                 **** ** ****** * **** :*************** ********************
90
91   BgTEP         PIALGRFPSSGFDRQXMMGAPXAMSFXDDXAMXSASFEMDVTTSTKPVERVRSFFPESWL   720
92   BgTEP1.1      PIALGRFPSSGFDRQGMMGAPMAMSFRDDNAMESASFEMDVATSTKPVERVRSFFPESWL   720
93   BgTEP1.2      PIALGRFPSSGFDRQGMMGAPMAMSFRDDNAMESASFEMDVATSTKPVERVRSFFPESWL   720
94   BgTEP1.3      PIALGRFPSSGFDRQGMMGAPMAMSFRDDNAMESASFEMDVATSTKPVERVRSFFPESWL   720
95   BgTEP1.4      PIALGRFPSSGFDRQGMMGAPMAMSFRDDNAMESASFEMDVATSTKPVERVRSFFPESWL   720
96   BgTEP1.5      PIALGRFPSSGFDRQGMMGAPMAMSFRDDNAMESASFEMDVATSTKPVERVRSFFPESWL   720
97                 *************** ***** **** ** ** ********:******************
98
99   BgTEP         WTSVKSINGHATLTTTVPDTITSWIGSAFATNSDTGLGVAPTTSKLPGFRPFFGSLTYPR   780
100  BgTEP1.1      WTSVKSINGHATLTTTVPDTITSWIVSAFATNSDTGLGVAPTTSKLRVFRPFFVSLTYPR   780
101  BgTEP1.2      WTSVKSINGHATLTTTVPDTITSWIVSAFATNSDTGLGVAPTTSKLRVFRPFFVSLTYPR   780
102  BgTEP1.3      WTSVKSINGHATLTTTVPDTITSWIVSAFATNSDTGLGVAPTTSKLRVFRPFFVSLTYPR   780
103  BgTEP1.4      WTSVKSINGHATLTTTVPDTITSWIVSAFATNSDTGLGVAPTTSKLRVFRPFFVSLTYPR   780
104  BgTEP1.5      WTSVKSINGHATLTTTVPDTITSWIVSAFATNSDTGLGVAPTTSKLRVFRPFFVSLTYPR   780
105                ************************ ******************  ***** ******
106
107  BgTEP         SVTRNEQFIVQATVFNYLPVDLMVTVSLKENPFLTPVTPGPGNQASNIQVRANEQRTVYF   840
108  BgTEP1.1      SVTRNEQFIVQATVFNYLPVDLMVTVSLKENPFLTPITPGPGNQASNIQVRANEQGIVYF   840
109  BgTEP1.2      SVTRNEQFIVQATVFNYLPVDLMVTVSLKENPFLTPITPGPGNQASNIQVRANEQGIVYF   840
110  BgTEP1.3      SVTRNEQFIVQATVFNYLPVDLMVTVSLKENPFLTPITPGPGNQASNIQVRANEQGIVYF   840
111  BgTEP1.4      SVTRNEQFIVQATVFNYLPVDLMVTVSLKENPFLTPITPGPGNQASNIQVRANEQGIVYF   840
112  BgTEP1.5      SVTRNEQFIVQATVFNYLPVDLMVTVSLKENPFLTPITPGPGNQASNIQVRANEQGIVYF   840
113                ************************************:******************  ***
114
115  BgTEP         SLSALIVGALDIEVSARSNMAADATVRQILVRHEGAPVVYNNPILISLSNNQSTFEKNIA   900
116  BgTEP1.1      SLSALTVGSLDIEVSARSNMAADATVRQILIKHEGAPVVYNNPILINLSNNQSTFEKNIA   900
117  BgTEP1.2      SLSALTVGSLDIEVSARSNMAADATVRQILIKHEGAPVVYNNPILINLSNNQSTFEKNIA   900
118  BgTEP1.3      SLSALTVGSLDIEVSARSNMAADATVRQILIKHEGAPVVYNNPILINLSNNQSTFEKNIA   900
119  BgTEP1.4      SLSAPTVGSLDIEVSARSNMAADATVRQILIKHEGAPVVYNNPILINLSNNQSTFEKNIA   900
120  BgTEP1.5      SLSALTVGSLDTEVSARSNMAADAIVRQILIKHEGAPVVYNNPILINLSNNQSTFEKNIA   900
121                ****  **:** ******************::**************.**************
122
123  BgTEP         FTLPDSLVPESHRIRVKVTGDLIGSTVQSLTSLLTLPTGCGEQSLVKFTPNIHIGRYLKA   960
124  BgTEP1.1      FTLPDSLVPESQRIRVKVTGDLIGSTVQSLTSLLTLPTGCGEQSLVKFTPNIHIGRYLKA   960
125  BgTEP1.2      FTLPDSLVPESQRIRVKVTGDLIGSTVQSLTSLLTLPTGCGEQSLVKFTPNIHIGRYLKA   960
126  BgTEP1.3      FTLPDSLVPESQRIRVKVTGDLIGSTVQSLTSLLTLPTGCGEQSLVKFTPNIHIGRYLKA   960
127  BgTEP1.4      FTLPDSLVPESQRIRVKVTGDLIGSTVQSLTSLLTLPTGCGEQSLVKFTPNIHIGRYLKA   960
128  BgTEP1.5      FTLPDSLVPESQRIRVKVTGDLIGSTVQSLTSLLTLPTGCGEQSLVKFTPNIHIGRYLKA   960
129                ***********:************************************************
130
131  BgTEP         TNQLSKELNKKIIDLLNNGYQRQLTYKRYDNGFSAFGNYDLSSSTWLTALVVTSFAEAQE  1020
132  BgTEP1.1      TNQLSEELNKKIIDLLNDGYQRQLTYKRYDNGFSAFGNYDISSSTWLTALVVTSFAEAQE  1020
133  BgTEP1.2      TNQLSEELNKKIIDLLNDGYQRQLTYKRYDNGFSAFGNYDISSSTWLTALVVTSFAEAQE  1020
134  BgTEP1.3      TNQLSEELNKKIIDLLNDGYQRQLTYKRYDNGFSAFGNYDISSSTWLTALVVTSFAEAQE  1020
135  BgTEP1.4      TNQLSEELNKKIIDLLNDGYQRQLTYKRYDNGFSAFGNYDISSSTWLTALVVTSFAEAQE  1020
136  BgTEP1.5      TNQLSEELNKKIIDLLNDGYQRQLTYKRYDNGFSAFGNYDISSSTWLTALVVTSFAEAQE  1020
137                *****:***********:**********************:********************
138
```

```
139  BgTEP          FIFVDKEIILKASMLLIDRQNLDGSFNEFGKVLDRNTQGT-TAGPALTAFVLVALLKAKE   1079
140  BgTEP1.1       FIFVDKEIILKASMLLIDRQNIDGSFNEFGKVLDRNTQGT-TAGPALTAFVLVALLKAKE   1079
141  BgTEP1.2       FIFVDKEIILKASMLLIDRQNIDGSFNEFGKVLDRNTQGT-TAGPALTAFVLVALLKAKE   1079
142  BgTEP1.3       FIFVDKEIILKASMLLIDRQNIDGSFNEFGKVLDRNTQGT-TAGPALTAFVLVALLKAKE   1079
143  BgTEP1.4       FIFVDKEIILKASMLLIDRQNIDGSFNEFGKVLDRNTQGT-TAGPALTAFVLVALLKAKE   1079
144  BgTEP1.5       FIFVDKEIILKASMLLIDRQNIDGSFNEFGKVLDRNTQGTTTAGPALTAFVLVALLKAKE   1080
145                 ********************:****************** ******************
146
147  BgTEP          LADVQYCKNNNKCRYYLLGNATLNATRNLERLMLADSIDDQFSLAVTSYALAEAKSQLAQ   1139
148  BgTEP1.1       LADVQDCKNNNKCRYYLLGNATLNATRNLERLMLADSIDDQFSLAVASYAFAEAKSQLAQ   1139
149  BgTEP1.2       LADVQDCKNNNKCRYYLLGNATLNATRNLERLMLADSIDDQFSLAVASYAFAEAKSQLAQ   1139
150  BgTEP1.3       LADVQDCKNNNKCRYYLLGNATLNATRNLERLMLADSIDDQFSLAVASYAFAEAKSQLAQ   1139
151  BgTEP1.4       LADVQDCKNNNKCRYYLLGNATLNATRNLERLMLADSIDDQFSLAVASYAFAEAKSQLAQ   1139
152  BgTEP1.5       LADVQDCKNNNKCRYYLLGNATLNATRNLERLMLADSIDDQFSLAVASYALAEAKSQLAQ   1140
153                 ***** ****************************************:***:*********
154
155  BgTEP          STFEKLLTFVKQEGGLEYRSANSTVNNEELNRFINWRPPRLQARPIDILITSYAILTYSP   1199
156  BgTEP1.1       STFEKLLTFVKQEGGLEYWSANSTVNNEELNRFINWRPPRLQARPIDILITSYAILTYSS   1199
157  BgTEP1.2       STFEKLLTFVKQEGGLEYWSANSTVNNEELNRFINWRPPRLQARPIDILITSYAILTYSS   1199
158  BgTEP1.3       STFEKLLTFVKQEGGLEYWSANSTVNNEELNRFINWRPPRLQARPIDILITSYAILTYSS   1199
159  BgTEP1.4       STFEKLLTFVKQEGGLEYWSANSTVNNEELNRFINWRPPRLQARPIDILITSYAILTYSS   1199
160  BgTEP1.5       STFEKLLTFVKQEGGLEYWSANSTVNNEELNRFINWRPPRLQARPIDILITSYAILTYSS   1200
161                 ****************** ****************************************
162
163  BgTEP          LGRLDEALPSVRWLTLQKNAQGGFVSTQDTVVGLQALSTYGSKSFRPDTNITIYVSDMNT   1259
164  BgTEP1.1       LGRLDEALPSVRWLTLQKNAQGGFVSTQDTVVGLQALSSYGSKSFRPDTNITIYVSDMNT   1259
165  BgTEP1.2       LGRLDEALPSVRWLTLQKNAQGGFVSTQDTVVGLQALSSYGSKSFRPDTNITIYVSDMNT   1259
166  BgTEP1.3       LGRLDEALPSVRWLTLQKNAQGGFVSTQDTVVGLQALSSYGSKSFRPDTNITIYVSDMNT   1259
167  BgTEP1.4       LGRLDEALPSVRWLTLQKNAQGGFVSTQDTVVGLQALSSYGSKSFRPDTNITIYVSDMNT   1259
168  BgTEP1.5       LGRLDEALPSVRWLTLQKNAQGGFVSTQDTVVGLQALSFYGSKSFRPDTNITIYVSDMNT   1260
169                 ************************************ ********************
170
171  BgTEP          HLTMNVKSDNALSLQIQEIQSNSQDFSITASGSGLALLDIEYSFNVLKELSKPVFDVNTV   1319
172  BgTEP1.1       HLTMNVNSENALSLQIQEIQSNSQDFSITASGSGLALLDIEYSFNVLKELSKPVFDVNTV   1319
173  BgTEP1.2       HLTMNVNSENALSLQIQEIQSNSQDFSITASGSGLALLDIEYSFNVLKELSKPVFDVNTV   1319
174  BgTEP1.3       HLTMNVNSENALSLQIQEIQSNSQDFSITASGSGLALLDIEYSFNVLKELSKPVFDVNTV   1319
175  BgTEP1.4       HLTMNVNSENALSLQIQEIQSNSQDFSITASGSGLALLDIEYSFNVLKELSKPVFDVNTV   1319
176  BgTEP1.5       HLTMNVNSENALSLQIQEIQSNSQDFSITASGSGLALLDIEYSFNVLKELSKPVFDANTV   1320
177                 ******:*:**************************************************. ***
178
179  BgTEP          LLDDKLDSFNIMVCTKFLLKHDTGMVVQEVSIPSGFVPDLSTLGQVAGVKRSERKGSIVA   1379
180  BgTEP1.1       LLDDKLDSFNIMVCTKFLMKHDTGMVVQELSIPSGFVPDLSTLGQVAGVKRSERKGSIVA   1379
181  BgTEP1.2       LLDDKLDSFNIMVCTKFLMKHDTGMVVQELSIPSGFVPDLSTLGQVAGVKRSERKGSIVA   1379
182  BgTEP1.3       LLDDKLDSFNIMVCTKFLMKHDTGMVVQELSIPSGFVPDLSTLGQVAGVKRSERKGSIVA   1379
183  BgTEP1.4       LLDDKLDSFNIMVCTKFLMKHDTGMVVQELSIPSGFVPDLSTLGQVAGVKRSERKGSIVA   1379
184  BgTEP1.5       LLDDKLDSFNIMVCTKFLMKHDTGMVVQELSIPSGFVPDLSTLGQVAGVKRSERKGSIVA   1380
185                 ******************:**********:******************************
186
187  BgTEP          IYFDKISGSSLCYSIVMTREAKVAKSQKSYVRTYDYYEPANQATVFYQPRTLRDSTVCDV   1439
188  BgTEP1.1       IYFDKISGSSLCYSIVMTREAKVAKSQKSYVRTYDYYEPANQATVFYQPRTLRDSTVCDV   1439
189  BgTEP1.2       IYFDKISGSSLCYSIVMTREAKVAKSQKSYVRTYDYYEPANQATVFYQPRTLRDSTVCDV   1439
190  BgTEP1.3       IYFDKISGSSLCYSIVMTREAKVAKSQKSYVRTYDYYEPANQATVFYQPRTLRDSTVCDV   1439
191  BgTEP1.4       IYFDKISGSSLCYSIVMTREAKVAKSQKSYVRTYDYYEPANQATVFYQPRTLRDSTVCDV   1439
192  BgTEP1.5       IYFDKISGSSLCYSIVMTREAKVAKSQKSYVRTYDYYEPANQATVFYQPRTLRDSTVCDV   1440
193                 ************************************************************
194
195  BgTEP          CLNCCP  1445
196  BgTEP1.1       CLNCCP  1445
197  BgTEP1.2       CLNCCP  1445
198  BgTEP1.3       CLNCCP  1445
199  BgTEP1.4       CLNCCP  1445
200  BgTEP1.5       CLNCCP  1446
201                 ******
202
```

Figure 1—figure supplement 2. Alignment of multiple *Bg*TEP amino acid sequences and distribution of identified peptides. Peptides identified by LC-MS/MS from

205 r*Bg*FREP3 pull-down experiments are highlighted in gray. The GenBank accession

206 numbers of each entry are: *Bg*TEP, ACL00841.1; *Bg*TEP1.1, ADE45332.1; *Bg*TEP1.2,

207 ADE45339.1; *Bg*TEP1.3, ADE45340.1; *Bg*TEP1.4, ADE45341.1 and *Bg*TEP1.5,

208 ADE45333.1.

209
210

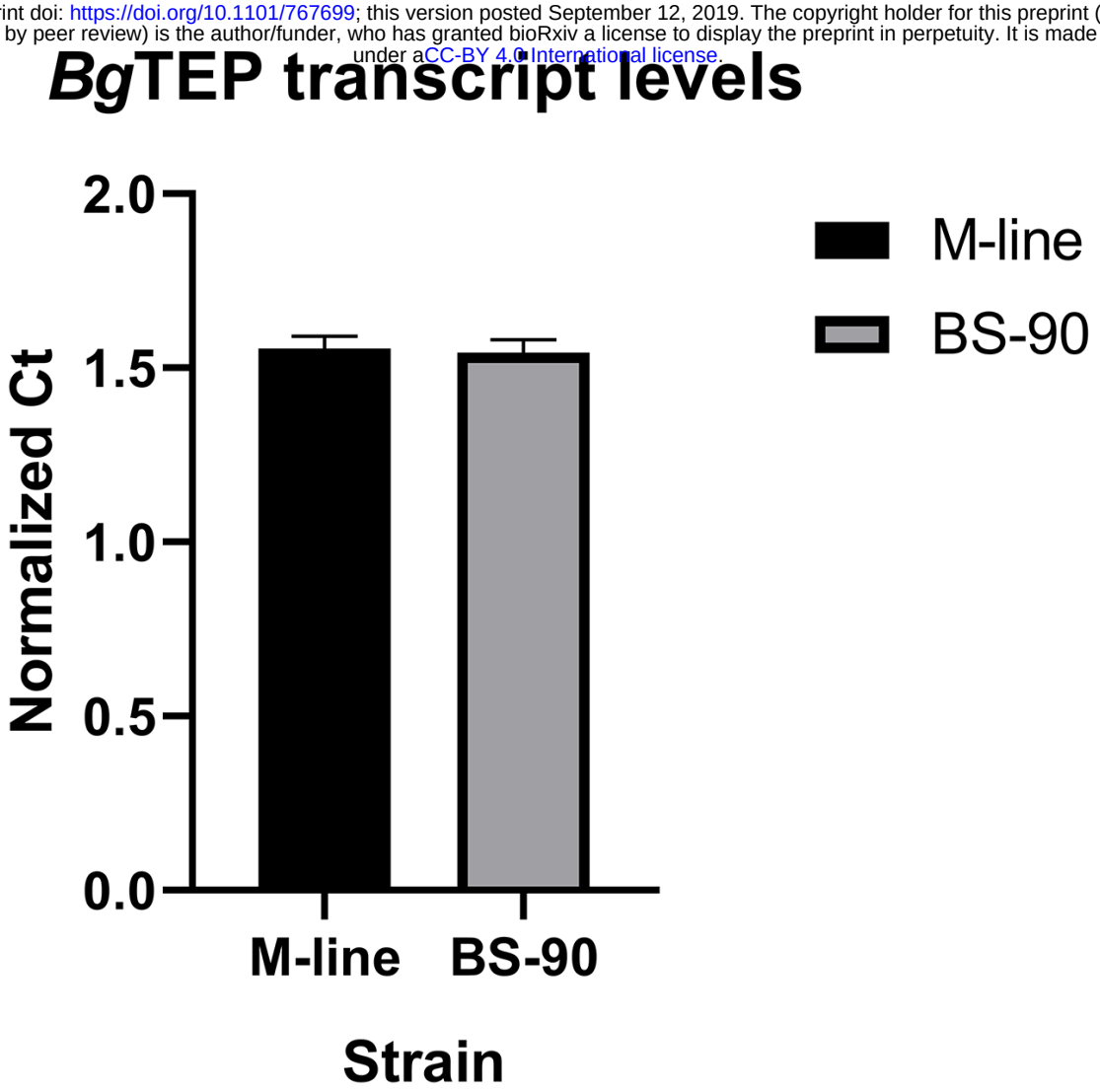

Figure 1—figure supplement 3.

Score	Expect	Method	Identities	Positives	Gaps
1143 bits(2957)	0.0	Compositional matrix adjust.	538/572(94%)	555/572(97%)	0/572(0%)

```
CLUSTAL O (1.2.4) multiple sequence alignment

A0A182YTN9          MLVQFLFAATLLQYVSSQCTYSSWWYSFDTPGQSKCNEINSYINALDRNDVNWADDALSN   60
A0A182YTZ4          MFLQIFVAVTLVQYVSSQCTYSSWWYSFDTPGQSKCNEINSYINALDRNDVNWADDALSN   60
                    *::*::.*.**:**************************************************

A0A182YTN9          LEGVQCCRPPAPWNNVEQQVVYEDWTATLDSDYTWAFCRVGYFLQGLYRSDTGWPRFKGY   120
A0A182YTZ4          LEGVQCCRPPAPWNNVEQQVVYEDWTATLDSDYTWAFCRVGYFLQGLYRSDTGWPRFKGY   120
                    ************************************************************

A0A182YTN9          LFNLESARCTKPANHPLNYGNCQDIDVSSCMGRKGQCSCPGGYFLTGLYRADGDDLYFLK   180
A0A182YTZ4          LFNLESARCTKPANHPLNYGTCQDIDVSSCMGRKGQCSCPGGYFLTGLYRADGDDLYFLK   180
                    ********************.***************************************

A0A182YTN9          KIRCCTPAAKPLEMDEKSKIQTRIMDTTLWNMATLAHYLGYGWCYGCRGVAVGEDFTRNG   240
A0A182YTZ4          KIRCCTPAAKPLEMDEKSKIQTRIMDTTLWNMATLAHYMGYGWCYGCHGLAVGEDFTRNG   240
                    ****************************************:*********:*:**********

A0A182YTN9          FTWAADTRSFWGKWCEGDKNGERLNLVFGDWGFAVKEIIYGKSVIEDLQAESVDSGVLYN   300
A0A182YTZ4          FTWAADTRTFWGKWCEGDKNGERLNLVFGDWGFAVKEIIYGKSVIEDLQAESVDSGVLYN   300
                    ********:************************************************************

A0A182YTN9          RASSPVTESIDRTKTIEETVTHSTTSTFTNSHELGIELNFEIASVSGKASYTTKFEYSKA   360
A0A182YTZ4          RASSPVTESIERSKTIQETITHSTTSTFTNSHGLGVELEFEIASVKGKASYKTRFEYSTS   360
                    **********:*:***:**:************ **:**:******.*****.*:****.:

A0A182YTN9          TTNEKSISQTAGFTKKSSITLGPMEGAKYEIIMSKSRTTVPYTAIITTKFSTEMKGFLRW   420
A0A182YTZ4          TTNSKSISETQGFTKQSSITLGPMEGAKYEVIMSKSRTTVPYTAIITTKFSTEMKGFLRW   420
                    ***.****:* ****:***************:******************************

A0A182YTN9          EDGNGNFHQDYRTNSGRPTYNYRFGDSSVPFYKALKKQSDNNEGVWMWGMLFQKFPDARR   480
A0A182YTZ4          EDGNGNFHQDYRTNSGRPTFNYRFGDSSVPFYKALKKQSDNNEGVWMWGMLFQKFPDARR   480
                    ********************:*********************************************

A0A182YTN9          VINRLTDETQYQFTLTGKLEKVEGTSVNVKWEKIKLNRRDVSGNDAPGSNITTYIAASGP   540
A0A182YTZ4          VTNRLTDETQYQFTLAGKLEKVEGTSVNVKWEKMKLNRRDVSGNDEPGSNITTYIAASGP   540
                    * ************:*****************:*********** **************

A0A182YTN9          ADKPAVVEYPKVNLNNKEPFKPIEIPVTEVKV 572
A0A182YTZ4          ADKPAVVEYPKVNLNNKEPFKPIEIPVTEVKV 572
                    ********************************
```

Figure 1—figure supplement 4. The alignments of two Biomphalysin variants (UniProtKB/TrEMBL: A0A182YTN9 and A0A182YTZ4) identified by LC-MS/MS. Peptides identified from A0A182YTN9 and A0A182YTZ4 are highlighted in gray and underline, respectively. Two red oval boxes representing differences in peptide sequences that distinguish two Biomphalysin variants.

```
CLUSTAL O (1.2.4) multiple sequence alignment

BgMFREP1 PARTIAL ------------------------------------------------------------  0
BgMFREP2 PRE     ------------------------------------------------------------  0
BgFREP3-2 PRE    MARLFLLFILCVFVVSLAGSELVIDVQPNVISPEITPQLVINCSITNNEVQTLDLIKSLT  60
BgMFREP4 PRE     ------------------------------------------------------------  0
BgMFREP5 PARTIAL ------------------------------------------------------------  0
BgMFREP6 PARTIAL ------------------------------------------------------------  0
BgFREP7.1 PRE    MTNLLLRLVFFQSLLPLLSSELVIDVQPDIISAEITAQLVINCSVTNNQVQHLDVIRSLT  60
BgMFREP8 PARTIAL ------------------------------------------------------------  0
BgMFREP9 PARTIAL ------------------------------------------------------------  0
BgMFREP10PARTIAL ------------------------------------------------------------  0
BgMFREP11PARTIAL ------------------------------------------------------------  0
BgFREP12.1 PRE   -MTLLQHFLLFVSLFLFSSSELVIDVQPNVISAEITAQLVINCSITNNQVQHLDVIRSLT  59
BgFREP12-FBG2    ------------------------------------------------------------  0
BgFREP13.1 PRE   -MATFLYVILVVALVTLSTSELVIDVQPNVISPEITAQLAINCSVTNNQAQNIDVIKSLT  59
BgFREP14         ------------------------------------------------------------  0


BgMFREP1 PARTIAL ------------------------------------------------------------  0
BgMFREP2 PRE     ------------------------------------------------------------  0
BgFREP3-2 PRE    LSRYNETIREFDELIALDSLTLNLKQFVRFKYSQISFGNRYITLILHNPTQFDARIYKCN  120
BgMFREP4 PRE     ------------------------------------------------------------  0
BgMFREP5 PARTIAL ------------------------------------------------------------  0
BgMFREP6 PARTIAL ------------------------------------------------------------  0
BgFREP7.1 PRE    LSRYNQTLRDFEDITALDLLTLNLKQLVKFKHSHISFGNVFISLTLLYPTQFDANVYRCS  120
BgMFREP8 PARTIAL ------------------------------------------------------------  0
BgMFREP9 PARTIAL ------------------------------------------------------------  0
BgMFREP10PARTIAL ------------------------------------------------------------  0
BgMFREP11PARTIAL ------------------------------------------------------------  0
BgFREP12.1 PRE   LSRYDETLKEFLDLITLEAKTLNLSQLVQLHHAQISFGNLSISLTLHNPTQFDAKVYRCK  119
BgFREP12-FBG2    ------------------------------------------------------------  0
BgFREP13.1 PRE   LSRYNETIRDFEVMIXLDLLTLNLKQLVQFNYSLISFGNVFISLTILHPTKSDAKVYRCS  119
BgFREP14         ------------------------------------------------------------  0


BgMFREP1 PARTIAL ------------------------------------------------------------  0
BgMFREP2 PRE     -------------------------MASLPLR---------LVLLV----SMATLIRSSSWL  24
BgFREP3-2 PRE    ATGTNSEGANISLFAKKAVEYETNSTALIEEIRRIKK--DENYCSFKKDDLSDSKQRSRV  178
BgMFREP4 PRE     ---------------------------------------MKNLLLCLF----LVSATLGSRL  19
BgMFREP5 PARTIAL ------------------------------------------------------------  0
BgMFREP6 PARTIAL ------------------------------------------------------------  0
BgFREP7.1 PRE    VKGGDPNNKDMSLFSKKTVEYETNSTALVEEIRRYKIDENKCLCSLASNDRTSTNKRLRV  180
BgMFREP8 PARTIAL ------------------------------------------------------------  0
BgMFREP9 PARTIAL ------------------------------------------------------------  0
BgMFREP10PARTIAL ------------------------------------------------------------  0
BgMFREP11PARTIAL ------------------------------------------------------------  0
BgFREP12.1 PRE   VEGDKTNAASSSIVAKKEVEYRTNMTALIEEIRRLKVVEKNDQCSLKXEELTSYHQKTKL  179
BgFREP12-FBG2    ------------------------------------------------------------  0
BgFREP13.1 PRE   VKGDNSNERNISLFAKKAVEYETNSTALIEEIQRCKKVKSKCQCSLENNDSSSNYKRSRV  179
BgFREP14         -------------------------MGALIL---------HVILCAS---IVPVISSRPKL  24


BgMFREP1 PARTIAL ------------------------------------------------------------  0
BgMFREP2 PRE     NFTGNSETIRELIQPLKLTCTFQISKNDSDNDSQVLFMSIYHETKRVIASISKYQPVATS  84
BgFREP3-2 PRE    YFSGSSDIIKERIEPLTLKCTFQVLKTDQNETSRLQSLYILHESKGVIAYVNKDQPVVTS  238
BgMFREP4 PRE     SFNANVEKINEVIXPLMLTCSFEVSRNDSWQNTKVQLMYIMHETKGFVATITKDQNITG-  78
BgMFREP5 PARTIAL ---------------------------------------------SNRIIANINKDRQVLTP  17
BgMFREP6 PARTIAL ------------------------------------------------------------  0
BgFREP7.1 PRE    NFSGNSEIIKERVETLTLNCTFQVLNQDQNETSSLQSLYILHETNGVIANINKGQPVLK-  239
BgMFREP8 PARTIAL ------------------------------------------------------------  0
BgMFREP9 PARTIAL ------------------------------------------------------------  0
BgMFREP10PARTIAL ------------------------------------------------------------  0
BgMFREP11PARTIAL ------------------------------------------------------------  0
BgFREP12.1 PRE   HFVGSSKVIKELIEPLTLTCSLQNLDN----NSTVQFMYILHESNGVIATINKDQPVVTT  235
BgFREP12-FBG2    ------------------------------------------------------------  0
BgFREP13.1 PRE   HFSGSSEIIKERIETVTLNCTYQALKHQRNE-TSLQSLYILHEANGVIANINKGQPVVQ-  237
BgFREP14         KFSGNAEILKELIEPLILSCSLKSFNNETNEHFKVHIMFIQHETNGVISTLSKDQAVAVS  84


```

```
70
71   BgMFREP1 PARTIAL ------------------------------------------------------------  0
72   BgMFREP2 PRE     LYPSVTKVQGHIYHSNESKDSYLQVTWTHPKLSESGKYFCLAHAWNSTSQNSVFDADITV  144
73   BgFREP3-2 PRE    LQGSNIQDVEGEIYDNAIKDSYLQVTWSNLKHSESGKYFCEAHNKYSEGRIDKSSNMLTI  298
74   BgMFREP4 PRE     NADMTFSEGQGTLNNEIDNTSFXQVTWKNASNELSGKYICVVHATNAEGKVEFLSASLKV  138
75   BgMFREP5 PARTIAL SQRDNSKNIQGEIHDDGSKDSYLQVTWSNVKSSESGKYFCEANVKHSDGRAERLSEMLII  77
76   BgMFREP6 PARTIAL ------------------------------------------------------------  0
77   BgFREP7.1 PRE    --GSHLKDAQGNIFDSGLKDSYLQVTWSNVKLSDSGKYFCEANVKHSDGRAERLSEMLIL  297
78   BgMFREP8 PARTIAL ------------------------------------------------------------  0
79   BgMFREP9 PARTIAL ------------------------------------------------------------  0
80   BgMFREP10PARTIAL ------------------------------------------------------------  0
81   BgMFREP11PARTIAL ------------------------------------------------------------  0
82   BgFREP12.1 PRE   KQESNFNTAKGVLSDTKSKASFIEVSWSYIKSSESGNYFCGAHVMGPDGRSERLNEMLAI  295
83   BgFREP12-FBG2    ------------------------------------------------------------  0
84   BgFREP13.1 PRE   --GSNLKNAEGEIFHNESKDSYLQVTWSNLKFSESGKYFCEANVKHSDGRAERLSEMLII  295
85   BgFREP14         A-DQSSTHAHGKIYNKDLQDSYLQVILKNPKISESGKYFCLAYAKNSTGQDSVFQSTVTI  143
86
87
88   BgMFREP1 PARTIAL ------------------------------------------------------------  0
89   BgMFREP2 PRE     NVIKSSTDDLAVALSYIQDRLDKD----------------------------------  168
90   BgFREP3-2 PRE    TVERPTFDDLVEAMHKLFTQVDGAKESLKAINQNIKNINKDLDFKEQNITSIKQEVIRNQ  358
91   BgMFREP4 PRE     QVQKLEIADLAQYVVDLTARVKESDDKIQNYTR---------------NVTSI-------  176
92   BgMFREP5 PARTIAL TVVRPTFDDLVKVVEKLLEQFNEYKTHLQENV---EKNRAMLDRNKQTILLIKKDVLANQ  134
93   BgMFREP6 PARTIAL ------------------------------------------------------------  0
94   BgFREP7.1 PRE    TVVSPTVDDLVKVIEKLLGQVDEDTKHIQENKQNIKNIKEELKTKEQNILSNTADLNSTQ  357
95   BgMFREP8 PARTIAL ------------------------------------------------------------  0
96   BgMFREP9 PARTIAL ------------------------------------------------------------  0
97   BgMFREP10PARTIAL ------------------------------------------------------------  0
98   BgMFREP11PARTIAL ------------------------------------------------------------  0
99   BgFREP12.1 PRE   TVSNPTFDDVVKVIPKLLRQADIEKENILENKQNIYDIKEYFKSKQQNIISIKDGLNTNR  355
100  BgFREP12-FBG2    ------------------------------------------------------------  0
101  BgFREP13.1 PRE   TVVSPTFDDLVKVIEKLLGQVDGDTRHIQENNQSIKNIKEELKIKEQNIISITADLNSTQ  355
102  BgFREP14         KVLKPKADDLVQVLGQLLKRVDTLEQLLKGNETNFGGV-----------------------  181
103
104
105  BgMFREP1 PARTIAL ------------------------------------------------------------  0
106  BgMFREP2 PRE     ------------------------------------------------------------  168
107  BgFREP3-2 PRE    NNIQILSEDLNIKEQNLTSIKADLSTKQQTFLNIKED----------------------  395
108  BgMFREP4 PRE     ------------------------------------------------------------  176
109  BgMFREP5 PARTIAL QSLQNMKEDWNSNQT---------------------------------------------  149
110  BgMFREP6 PARTIAL ------------------------------------------------------------  0
111  BgFREP7.1 PRE    QTIRSMKEDIAINQHNMSSLKEFVDANLESLQSIKEDLNIQQRNIIS--------------  404
112  BgMFREP8 PARTIAL ------------------------------------------------------------  0
113  BgMFREP9 PARTIAL ------------------------------------------------------------  0
114  BgMFREP10PARTIAL ------------------------------------------------------------  0
115  BgMFREP11PARTIAL ------------------------------------------------------------  0
116  BgFREP12.1 PRE   HNIKSIADDLNVNKENIASHNDEINTLRQMVNNVQSDLSICKKSIHNFSNDLDTKKQSIA  415
117  BgFREP12-FBG2    ------------------------------------------------------------  0
118  BgFREP13.1 PRE   QIISIIKEDITQNQQNISSMKEDLIINQEN-----------------------------  385
119  BgFREP14         ------------------------------------------------------------  181
120
121
122  BgMFREP1 PARTIAL ------------------------------------------------------------  0
123  BgMFREP2 PRE     ------------------------------------------------------------  168
124  BgFREP3-2 PRE    ---------------------------------------------VILNQQIIHKIKQ  408
125  BgMFREP4 PRE     ------------------------------------------------------KE  178
126  BgMFREP5 PARTIAL ---------------------------------------------NIISIKE  156
127  BgMFREP6 PARTIAL ------------------------------------------------------------  0
128  BgFREP7.1 PRE    ---------------VKEDIAINQQNISSIKTDVAVNEENLINLQKDFNIQQRNISSIKE  449
129  BgMFREP8 PARTIAL ------------------------------------------------------------  0
130  BgMFREP9 PARTIAL ------------------------------------------------------------  0
131  BgMFREP10PARTIAL ------------------------------------------------------------  0
132  BgMFREP11PARTIAL ------------------------------------------------------------  0
133  BgFREP12.1 PRE   SHKDELNSFGQIVNSLKDDLRTNKQNLQSVTDEENTNKENINQLKEDFKSNKQKIQNITQ  475
134  BgFREP12-FBG2    ------------------------------------------------------------  0
135  BgFREP13.1 PRE   ----------------------------------------LKNVKEDFNIQQRNILSLEK  405
136  BgFREP14         ------------------------------------------------------------  181
137
138
```

```
139  BgMFREP1 PARTIAL ----------------------------------------------DVRGYNVITNRQPVSCR  17
140  BgMFREP2 PRE     ----------------------------------------------GVDSIQISRASPTLPESCR  187
141  BgFREP3-2 PRE    DLNTYRHN-----------MSNIEEHLEVILTNLSTASIKVKNQTDEGSKMSYPPRKSCR  457
142  BgMFREP4 PRE     ELNALKENHLAA-------------------LRS--LDIIKKVNKNLQLSCECLAKPTSCR  218
143  BgMFREP5 PARTIAL ELQTHRQN-----------MSTLKENFETVFSNFSTALIDIKNQIVK-ERSGFKQVLSCR  204
144  BgMFREP6 PARTIAL ----------------------------------------------GVDSIQISRASPTLPESCR  19
145  BgFREP7.1 PRE    DFKTFYKNT----------SSFQEIILANLSATLEKVKN------ETELLHPTSCR  489
146  BgMFREP8 PARTIAL ------------------------------------------------------------  0
147  BgMFREP9 PARTIAL ------------------------------------------------------------  0
148  BgMFREP10PARTIAL ------------------------------------------------------------  0
149  BgMFREP11PARTIAL ------------------------------------------------------------  0
150  BgFREP12.1 PRE   EVNANRQNIINLNNTSQQSLSKLEADLETQLTNLSTALTQIN------EKIKKGLPSSCR  529
151  BgFREP12-FBG2    ------------------------------------------------------------  0
152  BgFREP13.1 PRE   DFHTHQQN-----------ISNFQENLEIVLSNFSTALMEVKNQTDK-ERKGDDQITSCR  453
153  BgFREP14         ---------------------------------TA-------DDILISTTHKTPFTSCR  200
154
155  BgMFREP1 PARTIAL DVNSTEDRMVVTLASGLKVMCDTKTDGGGWIIFQRRINGKVDFYRGWKEYRDGFGDYNIG  77
156  BgMFREP2 PRE     DVISSEDRVVVTLASGLKVMCDTKTDGGGWIIFQRRINGYVDFYRGWKEYRDGFGDYDIG  247
157  BgFREP3-2 PRE    DVNSTDERVVVTLTSGLKVMCDTKTDGGGWIIFQRRINGNVDFYRGWKEYRDGFGDYNIG  517
158  BgMFREP4 PRE     DVISTEDRVVVTLASGLEVMCDTTTDGGGWTIFQRRFNGSIDFYRDWKEYRDGFGDYNIG  278
159  BgMFREP5 PARTIAL DVRSIADRLVVFLISGLKVMCDTKTDGGGWLIFQRRINGKVDFYRGWKEYRDGFGDYDIG  264
160  BgMFREP6 PARTIAL DVISSEDRVVVTLASGSKVMCDTKTDGGGWIIFRRRINGNVDFYRGWKEYRDGFGSFSIG  79
161  BgFREP7.1 PRE    KVIYKEDRAIVTLASGLKVMCDTKTDGGGWIIFQRRVNGSVDFYRGWQEYRDGFGDYNIG  549
162  BgMFREP8 PARTIAL ------------------------------GWTMFQRRITGNLSFYRGWEEYKYGFGDVTTG  32
163  BgMFREP9 PARTIAL ------------------------------GWIIFQRRINGSVDFYRGWKEYRDGFGDYNIG  32
164  BgMFREP10PARTIAL ------------------------------GWTIFQRRINGKVDFYRGWKEYRDGFGDYNIG  32
165  BgMFREP11PARTIAL ------------------------------GWIIFQRRINGKVDFYRGWKEYRDGFGDYTIG  32
166  BgFREP12.1 PRE   EINSFQERVIVTLTSGLKVMCDTKTDGGGWIIFQRRINGKVDFYRGWKEYRDGFGDYDIG  589
167  BgFREP12-FBG2    ----------------------TKTDGGGWIIFQRRINGKVNFYRGWKEYRDGFGDYDIG  38
168  BgFREP13.1 PRE   DVTSKDDRVVVTLASGLKVMCDTKTDGGGWIIFQRRINGKVDFYRNWQVYRDGFGDYDIG  513
169  BgFREP14         DVNSSDERVVVTLASEQKVMCDTKTDGGGWIIIQRRIMGYVDFYRGWKEYRDGFGDYNIG  260
170                                              ** :::**. * :.***.*: *: ***.   *
171
172  BgMFREP1 PARTIAL EFYLGNENIFMLTSGRQYELRIDMEFKTKKYFAKYSRFKVLSEANNYQLKIESFSGNAGD  137
173  BgMFREP2 PRE     EFYLGNENIFKLTSSKKYDLRIDLEFNNTKYFAFYTRFEILGEQDYYKLQIGGYSGNAGD  307
174  BgFREP3-2 PRE    EFYLGNENIYMLTSTGQYNLRIDLKYKNKAFFAQYSGFKILSEKEKYKLNIGAYSGNSGD  577
175  BgMFREP4 PRE     EFYLGNENIFNLTSSRKYELRFDLEYENKKYFAHYSDFKLLDENNKYKLIIGSYSGNAGD  338
176  BgMFREP5 PARTIAL EFYLGNENIFNLTSTGKYDLRIDLKYND-FYHAQYSDFQILSEKEKYKLKIGAYSGNAGD  323
177  BgMFREP6 PARTIAL EFYLGNENIFKLTSSKKYDLRIDLEFNNTKYFALYTSFEILXXQDYYTLQIGGYSGNAGD  139
178  BgFREP7.1 PRE    EFYLGNENIFKLTSTGKYNLRIDLVLKNQPYYAQYSIFQILSEKNHYKINVGGYSGNAGN  609
179  BgMFREP8 PARTIAL EFYLGNENMHQITTKRRHELQIDLTFNQAKFSATYSRFSLYGEPEKYRLKVSGYSGNAGD  92
180  BgMFREP9 PARTIAL EFYLGNENIFNLTSTGQYDLRIDLEYNDKKYFAQYENFKVLSETEKYKLKIRKYSGNAGN  92
181  BgMFREP10PARTIAL EFYLGNEKVYQLTSEKENELRVDLETNITSYFAYYSEFYLLNETENYKLKVGGFNGTMSD  92
182  BgMFREP11PARTIAL EFYLGNEYISKLTSTRNFDLRIDFKFNHKTYFVEFSDFRILNETNNYQLKIGKYKGNASD  92
183  BgFREP12.1 PRE   EFYLGNENIFGLISTGQYDLRIDLEFNNTNYFAQYENFKVLSETEKYKLQTGTYSGNAGD  649
184  BgFREP12-FBG2    EFYLGNENIFKLTSTGQYDLRIDLEFKNTKYFAQYEDFKVLSETEKYKLQIGDYLGNAGD  98
185  BgFREP13.1 PRE   EFYLGNENVYNLTTSGNFDLRIDLKFNYQSYFAQYTHFKLLSEKEFYRLKFGDYSGNAGD  573
186  BgFREP14         EFYLGNENIHQLTSKKPHELRIDLEFDNENYFAQYSSFLLLGEEEFYKLQIGDYSGNAGD  320
187                   ******* :  : :    :*:.*:  .   : . :  * :    : * :    : *. .:
188
189  BgMFREP1 PARTIAL SLT-YHNDQFFSTF----------------------------------------------  150
190  BgMFREP2 PRE     ALTNIHNDKFFSTYDKDNDLGTSGSCAVTHKGAWWYR-DCYDSNLNGKWGSDR----VNW  362
191  BgFREP3-2 PRE    NFS-SHNNAFFTTFDRDNDEYS-YNCAVDYTGAWWYHSSCLNCNLNGKWGSSDFAKGVNW  635
192  BgMFREP4 PRE     SMR-RHVNKFFTTFDKDNDDSPNDNCAIIRRGAWWYQ-NCADVNLNGNWGRGE-PDGVFW  395
193  BgMFREP5 PARTIAL SLS-YHNDAHFSTYDKDNDNSS-INCASTVSGAWWYK-SCHHVNLNGRWRSKESGKSVIW  380
194  BgMFREP6 PARTIAL ALTNIHNDKFFSTYDKDNDMDTSGSCAVDYKGAWWYR-SCYDSNLNGKWGSDW----VNW  194
195  BgFREP7.1 PRE    SLS-YHNDMYFSTFDNDNDAYSGVNCAVKYHGAWWYK-TCYESNLNGKWGSSERDKDLNW  667
196  BgMFREP8 PARTIAL NLRKFDNVKFTT----------------------------------------------  104
197  BgMFREP9 PARTIAL GLF-YHNNMFFTT----------------------------------------------  104
198  BgMFREP10PARTIAL DLH-HNNNKAFST----------------------------------------------  104
199  BgMFREP11PARTIAL DFS-YHNNMQFST----------------------------------------------  104
200  BgFREP12.1 PRE   DLS-PHNNKFFSTFDRDNDISSSTNCAEYNSGAWWYE-SCHHSNLNGQWGRTS-NKGMNW  706
201  BgFREP12-FBG2    DLS-PHNNMFFSTFDRDNDVDSHLNCAEYCS----------------------------  128
202  BgFREP13.1 PRE   SLS-YHKDMYFSTFDKDNNIDS-RNCATEYWGAWWYR-SCHYSHLNGVWGGKD-PKGLIW  629
203  BgFREP14         SLA-YQKNMSFSTYDRDNDNASGENCAVSYSGAWWYN-ACHMSNLNGDWGNDAYGKGVNW  378
204                   :   .     :
```

```
BgMFREP1 PARTIAL ----------------------------      150
BgMFREP2 PRE     SKLTG-ITKSVTFTEMKIREIELN----      385
BgFREP3-2 PRE    YDLSR-FDSSVSFTEMKIREI-------      655
BgMFREP4 PRE     DNITV--WESVSFSEIKIREIDKEKNKS     421
BgMFREP5 PARTIAL RALTG-EQESVLYSEMKIREKD------      401
BgMFREP6 PARTIAL SKLTG-ITKSVTFSEMKIREIELD----      217
BgFREP7.1 PRE    NTLLQTAHVGVSFTEMKIRERE------      689
BgMFREP8 PARTIAL ----------------------------      104
BgMFREP9 PARTIAL ----------------------------      104
BgMFREP10PARTIAL ----------------------------      104
BgMFREP11PARTIAL ----------------------------      104
BgFREP12.1 PRE   FKLTR-GSNSVSFTEMKIREREKNYFH-     732
BgFREP12-FBG2    ----------------------------      128
BgFREP13.1 PRE   DKVTN-YEASVSFTEMKIREKS------      650
BgFREP14         DGVTG-LHDSVVFSEMKLRELD------      399
```

Figure 1—figure supplement 5. Alignment of multiple *Bg*FREP amino acid sequences and distribution of identified peptides. Peptides identified by LC-MS/MS are highlighted in gray. The GenBank accession numbers of each entry are: *Bg*MFREP1 partial, AAK13549; *Bg*MFREP2 precursor, AAK13550; *Bg*FREP3-2 precursor, AAK28656; *Bg*MFREP4 precursor, AAK13551; *Bg*MFREP5 partial, AAK13546; *Bg*MFREP6 partial, AAK13552; *Bg*FREP7.1 precursor, AAK28657; *Bg*MFREP8 partial, AAK13553; *Bg*MFREP9 partial, AAK13554; *Bg*MFREP10 partial, AAK13555; *Bg*MFREP11 partial, AAK13556; *Bg*FREP12.1 precursor, AAO59918; *Bg*FREP12-FBG2, AAT58639; *Bg*FREP13.1 precursor, AAO59922 and *Bg*FREP14, ABO61860.

```
CLUSTAL O (1.2.4) multiple sequence alignment


BgFREP3.3 AEO50747.1      --MLLSLELIRKFPTDTRSSHELVIDAQPEVISLELTPQLVVNCSITDSHVPGLDTINSL  58
BgFREP3.3 AAO59915.1      MARLFLLFILC-VFVVSLAGSELVIDVQPNVISPEITPQLVINCSITNNEVQQLDLIKSL  59
BgFREP3.2 AEO50746.1      MERLFLLLVLC-VFVVSLAGSELVIDVQPNVISPEITPQLVINCSITNNKVQQLDLIKSL  59
BgFREP3.2 AAK28656.1      MARLFLLFILC-VFVVSLAGSELVIDVQPNVISPEITPQLVINCSITNNEVQTLDLIKSL  59
BgFREP3.1 AEO50745.1      MARLFPRLVLC-VFIVPLAGSELVIDVQPNVISPEITPQLVINCSVTNNQVQHLEVIKSL  59
BgMFREP3  AAK13548.1      ---LFLLFILC-VFVVSLAGSELVIDVQPNVISPEITPQLVINCSITNNEVQTLDLIKSL  56
                             *:   ::  .     :. *****.**:*** *:*****:***:*:..*  *: *:**

BgFREP3.3 AEO50747.1      SLSRYNETKKEFDVLLSLDTHTLSLQQLVQFRHAQISFGNLYVTLTLRNPTQSDAKVYRC  118
BgFREP3.3 AAO59915.1      TLSRYNETIREFDDLIALDSLTLNLKQFVRFKYSQISFGNLYITLTLPNPTQFDARIYKC  119
BgFREP3.2 AEO50746.1      TLSRYNETIRDFDELIALDSLTLNLKQFVRFKYSQISFGNRYITLILHNPTQFDARIYKC  119
BgFREP3.2 AAK28656.1      TLSRYNETIREFDELIALDSLTLNLKQFVRFKYSQISFGNRYITLILHNPTQFDARIYKC  119
BgFREP3.1 AEO50745.1      TLSRYNEIIREFDELIALDSLTQNLKQFVRFKYSQISFGNLYITLTLPNPTQFDARIYRC  119
BgMFREP3  AAK13548.1      TLSRYNETIREFDELIALDSLTLNLKQFVRFKYSQISFGNRYITLILHNPTQFDARIYKC  116
                          :******  ::** *::**: * .*:*:*:*:::****** *:** * **** **::*:*

BgFREP3.3 AEO50747.1      NVSGDDSLWKNITRVFKKEIKYETNLTVLLEEIRRLREEKDRDQLSCQKEKLN----DSK  174
BgFREP3.3 AAO59915.1      NATGANSDGTNISLFAKKAVEYETNSTALIEEIRRIKKDENY--CSFKKDDLADIKQRSR  177
BgFREP3.2 AEO50746.1      NATGDNSEGANISLFAKKGVEYETNSTALIEEIRRIKKDENY--CSFKKDDLSDSKQRSR  177
BgFREP3.2 AAK28656.1      NATGTNSEGANISLFAKKAVEYETNSTALIEEIRRIKKDENY--CSFKKDDLSDSKQRSR  177
BgFREP3.1 AEO50745.1      NADGANSEGTNISLFTKKAVEYETNSTALIEEIRRIKKDENK--CSLKKDDLSDIKQRWR  177
BgMFREP3  AAK13548.1      NATGTNSEGANISLFAKKAVEYETNSTALIEEIRRIKKDENY--CSFKKDDLSDXKQRSR  174
                          *. * :*   **: . ** ::**** *.*:*****:::::    * :*:.*       :

BgFREP3.3 AEO50747.1      LHFVGNSKVVKELFDPLTLTCSIQDLMNDRNETSTVQSIYILHEANGIIATISKDQPVVT  234
BgFREP3.3 AAO59915.1      VYFSGSSDIIKERIEPLTLKCTFQVLKTDQNDTSRLQSLYILHESKGVIAYVNKDQPVVT  237
BgFREP3.2 AEO50746.1      VYFSGSSDIIKERIEPLTLKCTFQVLKTDQNEISRLQSLYILHETKGVIAYVNKDQPVVT  237
BgFREP3.2 AAK28656.1      VYFSGSSDIIKERIEPLTLKCTFQVLKTDQNETSRLQSLYILHESKGVIAYVNKDQPVVT  237
BgFREP3.1 AEO50745.1      VYFSESSKIIKERIEPLTLKCTFQILTPDENETSRLQSLYILHESNGVVANINKDQAVIT  237
BgMFREP3  AAK13548.1      VYFSGSSDIIKERIEPLTLKCTFQVLKTDQNETSRLQSLYILHESKGVIAYVNKDQPVVT  234
                          ::*  .*.::** ::****.*::* *  *.*: * :**:*****::*::* :.*** *:*

BgFREP3.3 AEO50747.1      TNQDVNLLAVIGKLHDDSSKNSYLQVTWSNPKFSESGKYFCGAHANNAQGRNEHWNEMLT  294
BgFREP3.3 AAO59915.1      SLQGSNIQDVEGEIYDNAIKDSYLQVTWSNLKHSESGKYFCEAHNKYSEGRIDKSSNMLT  297
BgFREP3.2 AEO50746.1      SLQGSHIQDVEGEIYDNAIKDSYLQVTWSNLKHTESGKYFCEAHNQYSEGRIDKTSNMLT  297
BgFREP3.2 AAK28656.1      SLQGSNIQDVEGEIYDNAIKDSYLQVTWSNLKHSESGKYFCEAHNKYSEGRIDKSSNMLT  297
BgFREP3.1 AEO50745.1      TIQGGNFENAQGEISGDQSKESYLQVTWSNLKHSDSGKYFCEAHVKHSKGKAERLNEMLT  297
BgMFREP3  AAK13548.1      SLQGSNIQDVEGEIYDNAIKDSYLQVTWSNLKHSESGKYFCEAHNXYSEGRIDXSSNMLT  294
                          : *. ::  . *:: .:  *:********* *.::****** **   ::*: :  .:***

BgFREP3.3 AEO50747.1      ITVERLQFDDIVKVMYDIQRQVDEDKKRLQTFHENLTNNFIILNTNLQSIENVRRDVRTN  354
BgFREP3.3 AAO59915.1      ITVERPTFDDLVEAMHKLFTQVDGAKESLKAINQNIKN----------------------  335
BgFREP3.2 AEO50746.1      ITVERPTFDDLVEAMHKLFTQVDGAKESLKAINQNIKN----------------------  335
BgFREP3.2 AAK28656.1      ITVERPTFDDLVEAMHKLFTQVDGAKESLKAINQNIKN----------------------  335
BgFREP3.1 AEO50745.1      IEVISPTIDDLMEVIQKLVTQVDGDKESLQDVKQNIMN----------------------  335
BgMFREP3  AAK13548.1      ITVERPTFDDLVEAXXKLFTQVDGAKESLKAINQNIKN----------------------  332
                          * *    :**:::.  .:  ***  *: *: .::*: *

BgFREP3.3 AEO50747.1      QESINGIKDELLSNKQNIVNNKKDINTAKESINVIKEELLSNKQNIVNNKRDIDTMEESI  414
BgFREP3.3 AAO59915.1      ----------------------INKDLDFKEQNITSIKQEVIRNQNNIQILSEDLNIKE---  372
BgFREP3.2 AEO50746.1      ----------------------INKDLDFKEQNITSIKEEVIRNQNNIQILSEDSNIKE---  372
BgFREP3.2 AAK28656.1      ----------------------INKDLDFKEQNITSIKQEVIRNQNNIQILSEDLNIKE---  372
BgFREP3.1 AEO50745.1      ----------------------IKEDLNTKEQNIISIKEDLNTKQQSIISIKEEFKTKQ---  372
BgMFREP3  AAK13548.1      ----------------------INKDLDFKEQNITSIKQEVIRNQNNIQILSEDLNIKE---  369
                                                 ::*::  ::.*  **:::  :::.*   ..: .  :

BgFREP3.3 AEO50747.1      NVIRHELLSNKQNIVNNKKDIHTTQESIIGILEELLRNQQNIVNNKKDINTTQDSIREIQ  474
BgFREP3.3 AAO59915.1      ------------------------------------QNLTSIKADLSTKQQTFLNIK  393
BgFREP3.2 AEO50746.1      ------------------------------------QNMTSIREDLSTKQQTFLNIK  393
BgFREP3.2 AAK28656.1      ------------------------------------QNLTSIKADLSTKQQTFLNIK  393
BgFREP3.1 AEO50745.1      ------------------------------------EN---IQKDVTINQQNIQKIK  390
BgMFREP3  AAK13548.1      ------------------------------------QNLTSIKADLSTKQQTFLNIK  390
                                                             :*    : *:. .*:.: :*:
```

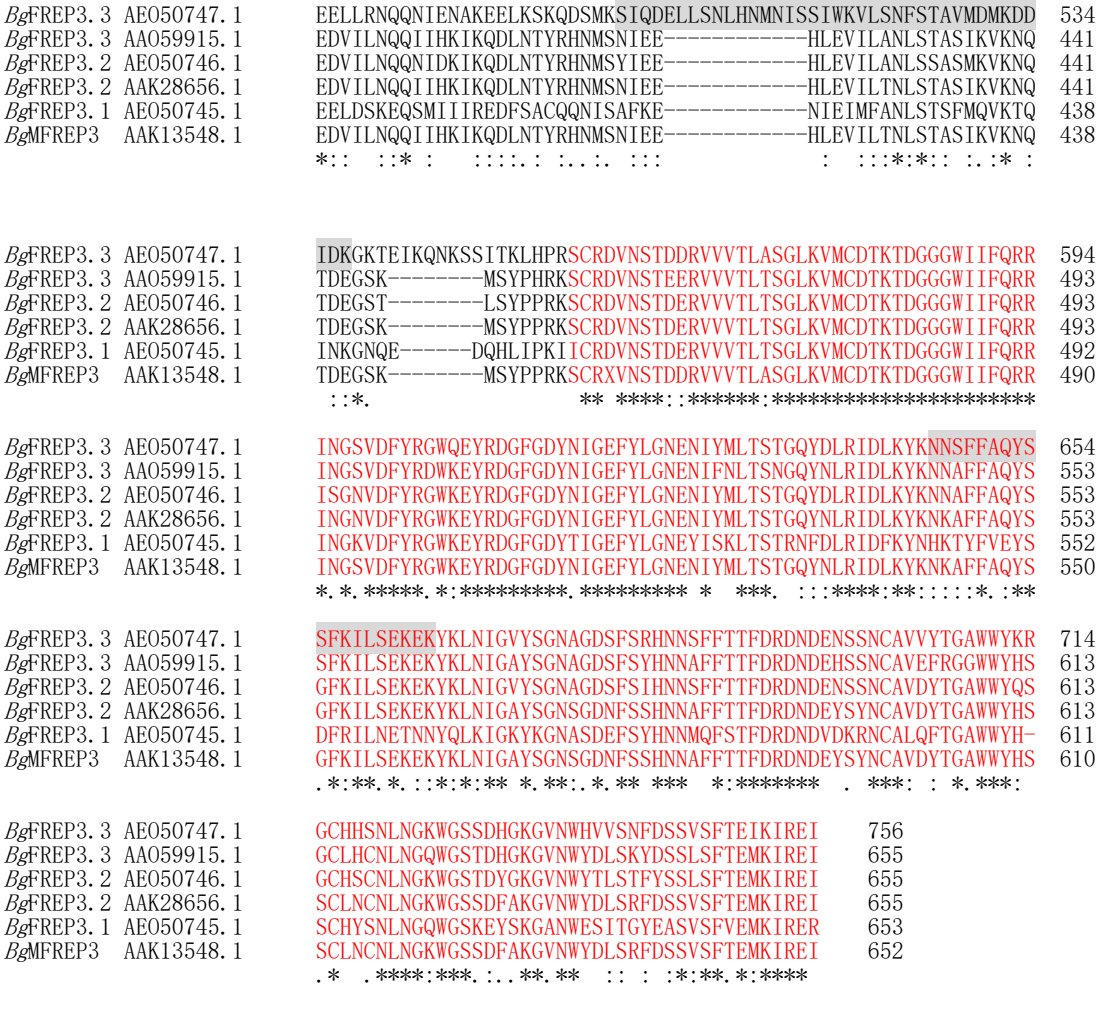

Figure 1—figure supplement 6. Alignment of multiple *Bg*FREP3 amino acid sequences and distribution of identified peptides. Peptides identified by LC-MS/MS are highlighted in gray, the overlapping region of peptides are bold underline. GenBank annotations of *Bg*FREP3.2 (AAK28656.1) are used for locate the individual *Bg*FREP3 domains: orange, signal peptide; green, IgSF1 domain; blue, small connecting region; pink, IgSF2 domain; black, ICR; red, FBG domain.

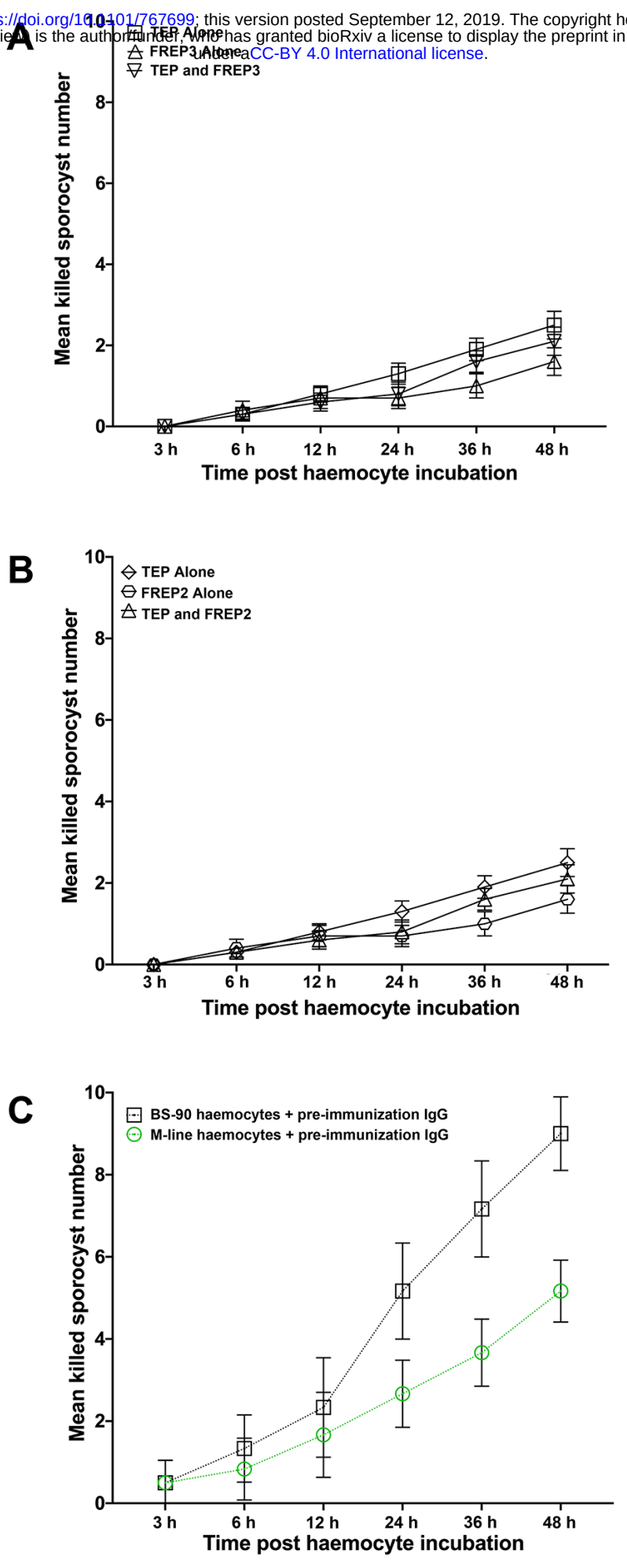

Figure 4—figure supplement 1.