6-1-2016

Characterization of nuclear factor-kappaB binding sites in the freshwater snail, Biomphalaria glabrata

Laura E. Deneckere
Lawrence University

Characterization of nuclear factor-kappaB binding sites in the freshwater snail, *Biomphalaria glabrata*

Laura Elizabeth Yu-Li Deneckere

Faculty Advisor: Judith Humphries

Biology Department

Lawrence University

Appleton, WI 54911

Monday, May 9, 2016

This project is presented for Honors at Graduation in Independent Study at Lawrence University, Appleton, WI

I hereby reaffirm the Lawrence University Honor Code

_____

ACKNOWLEDGEMENTS

I would like to express special appreciation to my research advisor, Dr. Judith Humphries, for her guidance and support throughout my undergraduate career. I am exceptionally fortunate to have worked in her research laboratory. Dr. Humphries's compassion and dedication extend beyond the school day, well into her own time and I am grateful for all that she has taught me.

I would like to thank the stockroom supervisors, JoAnn Stamm and Wayne Krueger, for their kind and generous help. Without their endless time and support, this project would not have been possible.

I am extremely appreciative to Briana Harter for her tremendous support throughout this entire process. She enthusiastically taught me important bench techniques and was a positive sounding board while troubleshooting the electrophoretic mobility shift assay protocol. Additionally, Briana meticulously read through countless drafts of this project.

I thank Zechariah Meunier for his thoughtful edits and support throughout this process. I would also like to acknowledge Savannah Vogel, Emily Stanislawski, Erin Maclaughlin, and Nicole Mitchell for editing drafts of this project.

I thank Henry Ward for collecting preliminary data.

I would like to extend my sincerest appreciation to Dr. Antoinette Powell for overseeing the honors committee. Lastly, I am grateful to my defense committee members, Dr. Judith Humphries, Dr. Brenda Jenike, Dr. Brian Piasecki, and Dr. Arnold Shober.

TABLE OF CONTENTS

TABLE OF FIGURES

ABBREVIATIONS

AP-1	Activator protein 1
APS	Ammonium persulfate
BLAST	Basic local alignment search tool
cAMP	Cyclic adenosine monophosphate
CREB	Cyclic adenosine monophosphate response element-binding protein
DTT	Dithiothreitol
ECL	Enhanced chemiluminescence
EDTA	Ethylenediaminetetraacetic acid
EMSA	Electrophoretic mobility shift assay
FASTA	Fast alignment
FREP	Fibrinogen-related protein
HPLC	High performance liquid chromatography
H_2O_2	Hydrogen peroxide
IκB	Inhibitory kappa-binding protein
IRAK	Interleukin-1 receptor-associated kinase
IRF	Interferon regulatory factor
κB site	Kappa-binding site
LASAGNA 2.0	Length-aware site alignment guided by nucleotide association 2.0
LPS	Lipopolysaccharide
LITAF	Lipopolysaccharide-induced tumor necrosis factor-α factor
MAPK	Mitogen-activated protein kinase
MyD88	Myeloid differentiation factor 88
NF-κB	Nuclear factor-kappaB
NIK	Nuclear factor-kappaB inducing enzyme
NLS	Nuclear localization sequence
NO	Nitric oxide

ORegAnno	Open regulatory annotation database
PAMP	Pathogen-associated molecular pattern
PGN	Peptidoglycan
RHD	Rel homology domain
RPM	Revolutions per minute
SA-HRP	Streptavidin horseradish peroxidase
STAT	Signal transducer and activator of transcription
TAK	Transforming growth factor β-associated kinase
TAB	Transforming growth factor β-associated kinase binding protein
TBE	Tris borate ethylenediaminetetraacetic acid
TdT	Terminal deoxynucleotidyl transferase
TEMED	Tetramethylethylenediamine
TIFF	Tagged image file format
TRANSFAC	Transcription factor database
TIR	Toll/interleukin-1 receptor
TLR	Toll-like receptor
TNF-α	Tumor necrosis factor-α
TRAF	Tumor necrosis factor receptor-associated factor
Uev1a	Ubiquitin-conjugating enzyme E2 variant 1A

ABSTRACT

Biomphalaria glabrata is an intermediate snail host for the digenean trematode, *Schistosoma mansoni,* which causes the human disease schistosomiasis. A lot of research has focused on the snail-schistosome interaction, especially in regards to the immune response of the snail. The nuclear factor-kappaB (NF-κB) pathway, which is involved in regulating the immune response, can be triggered by the Toll-like receptor (TLR) signaling pathway. However, not much is known about the specific molecular mechanisms regulating these responses. Both NF-κB and TLR homologues have recently been reported in *B. glabrata* so it is of great interest to determine if *Bg*NF-κB can regulate components of the pathway. We have used bioinformatics to identify putative κB sites upstream of 16 genes coding for members of the TLR pathway in *B. glabrata.* In order to determine if the snail's NF-κB p65 protein can recognize the κB sites of interest, electrophoretic mobility shift assays (EMSAs) were carried out. The functional analyses of putative κB sites are essential for advancing our understanding of the TLR mediated NF-κB pathway in *B. glabrata.* In the present study, six novel κB sites were characterized by showing that the Rel homology domain (RHD) of the *B. glabrata* NF-κB p65 protein could bind to these κB sites. Our findings suggest that *Bg*NF-κB can regulate the expression of these genes.

INTRODUCTION

1.1 Invertebrate immunology

Until recently, the field of immunology has largely focused on vertebrate systems (Rast & Messier-Solek, 2008). Vertebrates have both innate and adaptive immune systems, which allow organisms to respond to a vast array of foreign substances. The innate immune system serves as a first line of defense against pathogens, whereas the adaptive immune system can work to eliminate these foreign organisms during repeated infections (Owen *et al.* 2013). In contrast to vertebrate models, invertebrates rely solely on an innate immune system (Nyholm & Graf, 2012). However, invertebrates have highly diverse recognition receptors and effectors, as well as associated cell signaling pathways that can function through receptors. These various mechanisms allow invertebrates to recognize a variety of bacterial, viral, and fungal components on the surface of pathogens. Initial studies have revealed that invertebrate innate immune systems share many components that have also been reported in vertebrates (Schulenberg *et al.* 2007). In 1996, Jules Hoffman, Bruno Lematire, and colleagues discovered that *Toll* was shown to be involved in the production of an antifungal peptide called drosomycin. This finding led to the discovery of mammalian Toll-like receptors, which has transformed our understanding of innate immunity in humans (Lemaitre *et al.* 1996). Their findings suggest that both invertebrate and vertebrate animals share common defense mechanisms, giving us further insight into our own innate immune systems.

Despite their lack of antibody-based immune responses, invertebrates have been shown to exhibit highly effective immune responses (Table 1). Invertebrate animals have a vast array of antimicrobial peptides, which provide a rapid response to foreign microorganisms (Bulet *et al.*

2004). Other invertebrate defense mechanisms may include hemolymph coagulation and lectin-mediated complement activation (Iwanaga & Lee, 2005). Additionally, many invertebrates have been shown to exhibit immunological memory, which allows organisms to respond more effectively to a pathogen upon repeated exposure. To illustrate, Kurtz and Franz (2003) found that a cyclopoid copepod (*Macrocyclops albidus*) could show increased immunity upon repeated exposure to a tapeworm (*Schistocephalus solidus*). Similarly, acquired immunological priming has been reported in the water flea (*Daphnia magna*), as well as a species of bumblebee (*Bombus terrestrius*) (Little *et al.* 2005; Sadd & Schmid-Hempel, 2006). These characteristics were previously thought to be unique to vertebrates with adaptive immunity, which suggests that invertebrates can indeed exhibit highly elaborate immune responses.

Table 1. Major host defense systems in invertebrates. Table courtesy of Iwanaga and Lee (2005).

Major host defense mechanisms
Hemolymph coagulation system
Pro-phenoloxidase (pro-PO) activating system
Lectin-complement system
Agglutinin-lectin system
Antibacterial, antifungal, and antiviral systems mediated by Toll-like receptors (TLRs) and peptidoglycan-binding protein (PGBP)
Reactive oxygen-producing system
Phagocytic system

1.2 Molluscan immunology

Mollusks are considered as one of the most diverse phyla on the planet (Bunje, 2003). These animals display a wide array of defense mechanisms to protect themselves against pathogens and environmental stressors (Song *et al.* 2010). Most mollusks have physical (e.g., shells) and chemical (e.g., mucus) barriers to cover their soft bodies and defend against pathogens (Glinski & Jarosz, 1997). Similar to other invertebrates, mollusks rely exclusively on

an innate immune system that consists of two branches: the cellular response and the humoral response (Rowley & Powell, 2007). The former involves a variety of immune cells with different properties, such as the ability to engulf and isolate a respective pathogen. The humoral response consists of dissolved compounds (e.g., clotting factors, complement proteins, lectins) that circulate around the body and assist with pathogen elimination (Owen *et al.* 2013). These various components have contributed to the success of the phylum Mollusca.

Several molluscan species have already been widely researched, due to their role as intermediate hosts for parasites of medical and veterinary importance. For example, a number of *Lymnaea* species have been studied extensively due to their role as an intermediate host for a variety of trematode parasites, including the liver fluke (*Fasciola hepatica*), which causes serious infections in livestock (Roberts *et al.* 2012). Likewise, three genera of freshwater snails, *Biomphalaria, Bulinus,* and *Oncomelania*, can serve as the intermediate host for *Schistosoma* trematode parasites (World Health Organization, 1993). Schistosome infections affect over 261 million people in 78 countries worldwide (Fig. 2; WHO, 2015). For these reasons, several molluscan species have been widely studied, particularly within the field of tropical medicine.

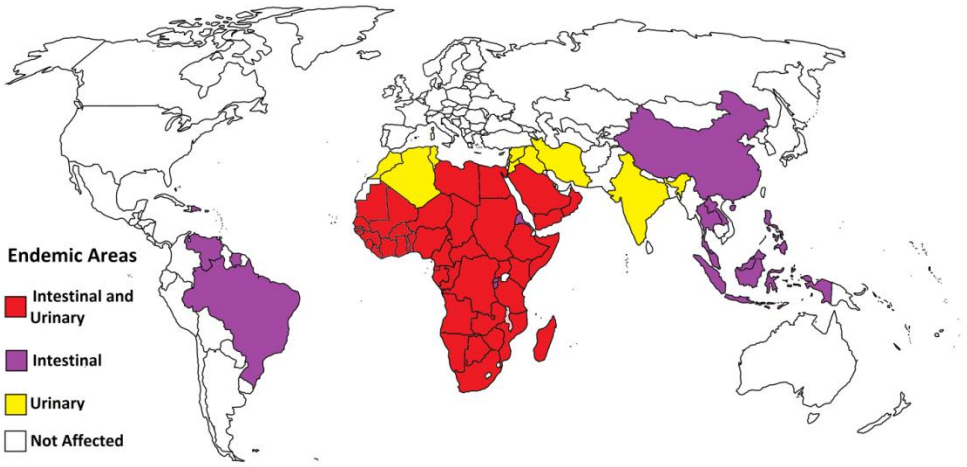

Figure 1. Worldwide distribution of schistosomiasis. Countries with reported cases of schistosomiasis are displayed above. Red regions are where both intestinal and urinary schistosomiasis can be reported. Purple regions indicate the presence of intestinal schistosomiasis only. Yellow regions show cases of urinary schistosomiasis only. Map was generated using data from the Centers for Disease Control and Prevention (2012).

1.3 Schistosomiasis overview

As an intermediate host, the snail plays a critical role in the schistosome life cycle, as it allows the parasite to develop into the stage that is infectious to humans (Fig. 3; Boros, 1989). When free-swimming, ciliated larval stages called miracidia infect their compatible snail hosts, they shed their ciliary plates to transform into sporocysts. Then, they undergo multiple rounds of asexual reproduction, generating numerous mother and daughter sporocyst stages for four to six weeks (Colley *et al.* 2014). The sporocysts give rise to free-swimming, fork-tailed cercariae that are released from the snail into the water (Boros, 1989). Within freshwater, the cercariae can remain infective to their human hosts for one to three days (Colley *et al.* 2014). Upon contact with a human host, cercariae penetrate the skin, shedding their tails upon entry (CDC, 2012). Cercariae mature and transform into juvenile forms called schistosomulae, which migrate through the blood vessels from the lungs to the circulatory system (Boros, 1989). The schistosomulae reach maturity in the liver, where they develop into sexually dimorphic male and

female blood flukes. Here, the adults pair up and live *in copula*, where the slender female resides in the gynaecophoric canal of the broader male to form a "hot dog in a bun" arrangement (CDC, 2012). Depending on the species of worm, the pair may relocate to the blood vessels of either the intestines or the bladder. Adults can live in the human host for an average of three to ten years, feeding on erythrocytes to harvest energy (Colley *et al.* 2014). Female worms can lay as many as 300 eggs into the bloodstream per day (Boros, 1989). While most of these eggs are excreted into the surrounding freshwater through feces or urine to continue the life cycle, some of these eggs can become lodged inside the intestines, liver, or bladder, causing serious inflammatory responses (Boros, 1989; Colley *et al.* 2014). Eggs passed into the surrounding freshwater hatch into miracidia, which are the infectious stage to snails. The miracidia then infect a compatible snail host and the life cycle repeats.

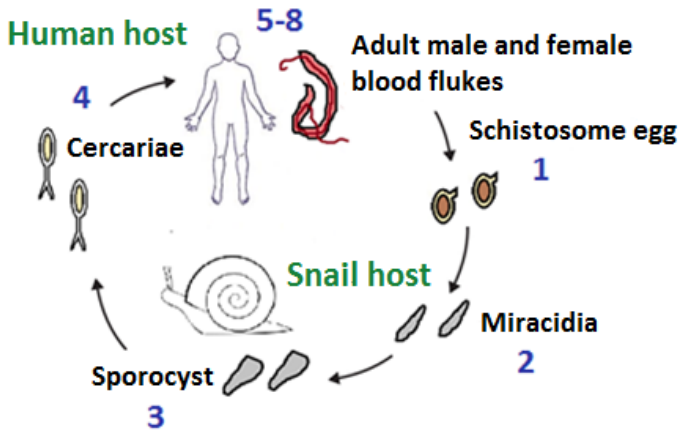

Figure 2. *Schistosoma mansoni* life cycle. 1) Eggs leave the human host via feces or urine. **2)** Miracidia hatch in freshwater and enter the snail. **3)** Sporocysts undergo multiple rounds of asexual reproduction and transform into cercariae. **4)** Cercariae shed from the snail and pass into surrounding freshwater. **5)** Cercariae penetrate human skin. **6)** Juvenile worms (schistosomulae) migrate around the body in blood vessels. **7)** Schistosomulae mature into adult male and female worms. **8)** Adult worms pair up and reproduce; females lay eggs within the blood vessels.

Understanding the *S. mansoni* life cycle is imperative for the successful control and eradication of this tropical disease. In previous years, government officials have taken preventative measures to control disease transmission, mainly through mass drug administration

programs, molluscicides, and biological control (Gray, 2011). Anthelminthic drugs, e.g. praziquantel, are used to treat schistosome infections. However, infection rates remain relatively unchanged because the infectious snails are still present within areas where people can bathe and collect drinking water (Colley *et al.* 2014). In addition, people are susceptible to reinfection with schistosomes following drug treatment (Gryseels, 2006).

1.4 *Biomphalaria glabrata* immune system

In the field of parasitology, *B. glabrata* is one of the most extensively researched organisms, largely due to its critical role in the survival of *S. mansoni* (Fig. 4; Bayne *et al.* 1980; Matricon-Gondran & Letocart, 1999; Hahn *et al.* 2001; Guillou *et al.* 2007; Humphries & Yoshino, 2008; Adema *et al.* 2009). Strains within the species differ in their susceptibility to *S. mansoni* infection and are characterized as either resistant or susceptible (Yoshino & Cousteau, 2011). Similar to other mollusks, the immune system of *B. glabrata* is comprised of both cellular and humoral components (Loker, 2004). Hemocytes (blood cells) circulate in the hemolymph, functioning as primary immune effector cells. These immune cells are released from the snail's hematopoietic organ to kill foreign pathogens via phagocytosis or cellular encapsulation (Owen *et al.* 2013). Hemolymph serves as a vehicle for carrying humoral factors around the body. Humoral components may include lectins, clotting factors, free radicals, and a variety of fibrinogen-related proteins (FREPs) (Zhang *et al.* 2008).

Figure 3. ***Biomphalaria glabrata.*** Image courtesy of www.vectorbase.org.

Many studies have emphasized the importance of hemocytes and hemolymph for their ability to generate anti-schistosome responses (Hahn *et al.* 2001; Humphries & Yoshino *et al.* 2008). For example, Hahn *et al.* (2001) examined the cytotoxic properties of hemocytes within susceptible and naturally resistant snail hosts. The authors in this study found that hemocytes in resistant snails can release free radicals, e.g. hydrogen peroxide (H_2O_2) and nitric oxide (NO), to kill sporocysts (Hahn *et al.* 2001). In contrast, susceptible snails were unable to use these defense strategies to kill the parasites and infection developed soon after parasite exposure. Other studies have been carried out to examine the protein profile of hemolymph before and after parasite infection (Adema *et al.* 1997; Zhang *et al.* 2008).

Recent studies using reverse transcriptase-polymerase chain reaction (RT-PCR), suppressive subtractive hybridization (SSH), and gene profiling via microarrays have identified many differentially expressed transcripts following exposure to trematode parasites (Lockyer *et al.* 2002; Nowak *et al.* 2004; Mitta *et al.* 2005; Adema *et al.* 2009). In fact significantly altered expression levels of various transcripts (e.g., defense-related lectins, sequences encoding histones) were reported in *B. glabrata* following *S. mansoni* exposure (Adema *et al.* 2009). Transcripts that show elevated expression levels following exposure to parasites are therefore hypothesized to play a role in defense.

It is expected that such changes in gene expression within *B. glabrata* are regulated by an underlying mechanism; however, very little is currently known regarding gene regulation in the snail. In other species, it is known that gene expression can be regulated in part by transcription factors, which are DNA-binding proteins. Although many transcription factors are involved in generating immune responses within vertebrates, transcription factor homologues have only recently been identified in *B. glabrata* (Zhang & Coultas, 2011). Recently characterized transcription factors include signal transducer and activator of transcription (STAT), cyclic adenosine monophosphate response element-binding proteins (CREB), activator protein 1 (AP-1), and lipopolysaccharide-induced tumor necrosis factor-α factor (LITAF) (Zhang & Coultas, 2011; Adema *et al.* 2016). In addition, two members of the nuclear factor-kappaB (NF-κB) transcription factor family, Rel (GenBank Accession No: FJ804761) and Relish (GenBank Accession No: FJ804762), were recently identified in *B. glabrata* (Zhang & Coultas, 2011).

1.5 Nuclear factor-kappaB pathway

One of the most important signaling pathways responsible for regulating immune responses in vertebrates is the NF-κB pathway (Moynagh, 2005). NF-κB proteins are well characterized in many animal phyla, including chordates (Gilmore & Wolenski, 2012). Homologues have been recently identified in molluscan species, including the Pacific oyster (*Crassostrea gigas*), the bobtail squid (*Euprymna scolopes*) and a species of abalone (*Haliotis diversicolor* supertexta) (Montagnani *et al.* 2004; Goodson *et al.* 2005; Jiang & Wu, 2007). However, little is known about molluscan NF-κB homologues regarding their functional role as transcription factors. The functional studies that have been carried out have mostly entailed expression analyses of the transcription factors themselves following exposure to pathogens (Adema *et al.* 2009; Zhang & Coultas, 2011).

In general, NF-κB proteins either form homo- or heterodimers that vary in their relative affinities to a particular DNA binding site (Chen & Ghosh, 1999). NF-κB typically functions as a heterodimer composed of p50 and p65 proteins. Interestingly, the two genes that were recently identified by Zhang and Coultas (2011) in *B. glabrata* show most similarly to p50 and p65 proteins in other species. All NF-κB proteins have a highly conserved amino-terminal domain called the Rel homology domain (RHD) (Fig. 4-5). The RHD is approximately 300 amino acids in length and contains regions that are involved in dimerization, sequence-specific DNA binding, inhibitory IκB binding, and nuclear localization (Chen & Ghosh, 1999). The carboxyl-terminus of the RHD has a nuclear localization sequence (NLS), which is responsible for "tagging" a protein for nuclear translocation (Liang *et al.* 2004; Franzenburg, 2012).

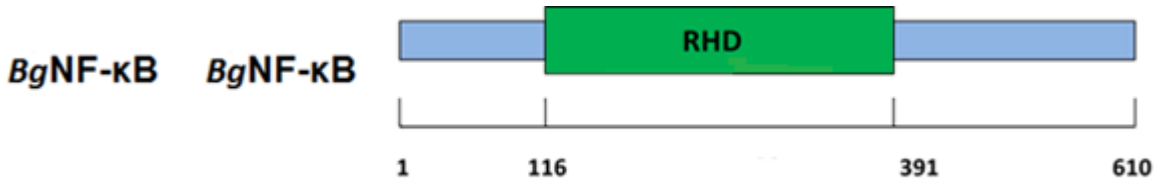

Figure 4. *Bg*NF-κB p65 protein indicating the Rel homology domain. This domain contains a nuclear localization sequence (NLS) and assists with highly specific DNA-protein interactions.

MMAMSFSDADLDLGDLQNMNSSELMRFVQNMNDDTAGQLINQMQQIHSPPQNITNSHPVHPV

QDSHHHAMQQYSNNPAVYAAPVYYPMQPPQQPQQQQPQPQLPPHVQTAASTPYVEILEQPKS

RGLRFRYECEGRSAGSVPGENSTNDHRTYPTIKIHNYNGPAIIVVSCVTKENPPHCKPHPHAIV

GRDCKKGVCTLRVKDTSDKIVFPQIGIQCAKKKDVESSLRLRKEINVDPYQTGFDHAQSNIDLNV

VRLCFQVFLPNEQGKVTRVVPPVCSHPIHDKKSSKDLVICRVDKSSGKARGGDEVFLLCDKVN

KEDIKVRFYEENEQGMVVWEDLGDFGQGDVHRQYAIVFRTPSYHNTEITRPAEVLMQLQRPSD

GETSEPIPFTYMPEDPDPDRIAEKRKRKAQRFLEYWNSTGVADYKRGASGASVRGRLHTMLKA

TRRIKKDPDVIPGLPGFMDTSGVDGVNGVGAVGGADATMSHVTADNSSMGSVTVSSFSAGSL

DSNVIAELSGLSSSDLSIVSTENGQLVISVSGGEGSGLSVDSINLPPNILSGELMNQIDIQQLNAYL

QDQGSGSEILLGNDPNQPIHILDHDPETEAALRGLNSQI

Figure 5. *Bg*NF-κB p65 protein sequence. The portion highlighted in red represents the Rel homology domain (RHD). Image adapted from J. Humphries (2015).

The NF-κB signaling pathway can be activated by a variety of transmembrane receptors, e.g. TLRs, that can bind to shared components on the surface of pathogens (Fig. 6; Franzenburg *et al.* 2012). In unstimulated cells, NF-κB is held within the cytoplasm in an inactive form by a family of inhibitory kappa-binding (IκB) proteins, such as IκBα and IκBβ (Jiang & Wu, 2007). These proteins bind to NF-κB via multiple ankyrin-repeats, which mask the NLS to inactivate NF-κB (Moynagh, 2005; Jiang & Wu, 2007). Upon stimulation, IκB kinases (IKKs) signal IκB to phosphorylate, which is immediately followed by the binding of a ligase complex that adds ubiquitin to IκB, targeting it for degradation by the 26S proteasome (Jiang & Wu, 2007; Franzenburg *et al.* 2012). This process exposes the NLS, thereby freeing the NF-κB and allowing it to translocate to the nucleus. Inside the nucleus, NF-κB binds to highly specific DNA sequences called response elements to activate target genes (e.g., cytokines, adhesion molecules, acute phase proteins) for gene transcription (Moynagh, 2005).

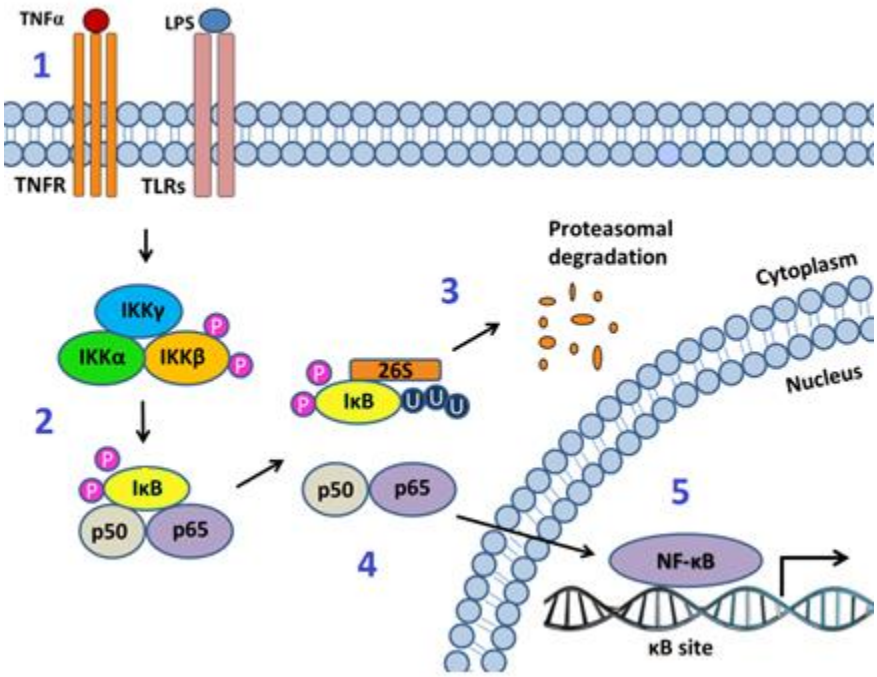

Figure 6. Nuclear factor-kappaB pathway. 1) NF-κB is activated by proinflammatory cytokines (e.g., tumor necrosis factor-α) or pathogen-associated molecular patterns (e.g., lipopolysaccharide). **2)** IκB kinases (IKKs) signal IκB to phosphorylate. **3)** The IκB inhibitor is ubiquitinated, leading to degradation by the 26S proteasome. **4)** NF-κB dimers (p50 and p65) then undergo nuclear translocation. **5)** Once inside the nucleus, transcription is initiated and target genes are activated.

NF-κB can be activated by TLRs, which are type-1 transmembrane receptors that can recognize components common to multiple infectious agents, e.g., peptidoglycan (PGN) (Fig. 7; Zhang & Ghosh, 2001). To illustrate, TLRs bind to their respective ligand, e.g. PGN, to initiate a signaling pathway that can ultimately activate NF-κB. The TLR–NF-κB pathway is thought to be an ancient host defense mechanism, as members of this pathway have been identified in cnidarians and sponges (Putnam *et al.* 2007; Gauthier & Degnan, 2008). Additionally, components of this pathway have also been identified in several molluscan species (Xiong *et al.* 2008; He *et al.* 2013; Toubiana *et al.* 2013; Valenzuela-Munoz, 2014).

TLR-ligand binding interactions can trigger a cascade of events that lead to NF-κB activation. A cascade can be considered as "a series of events in which one event triggers the

next in a linear fashion" (Hinkle & Bowman, 2003). Intracellular signaling is initiated by Toll/interleukin-1 receptor (TIR) domain containing cytoplasmic adaptor proteins, such as myeloid differentiation factor 88 (MyD88) (Kawai & Akira, 2006). Following MyD88 activation, a number of other downstream signaling proteins are activated, such as interleukin-1 receptor-associated kinase (IRAK), tumor necrosis factor receptor-associated factor (TRAF), transforming growth factor β-associated kinase (TAK1), and IκB kinase (IKK) (Zhang & Ghosh, 2001; Cornet *et al.* 2015).The activation of these downstream components ultimately leads to the nuclear translocation of NF-κB proteins and initiation of gene transcription (Franzenburg *et al.* 2012; Cornet *et al.* 2015).

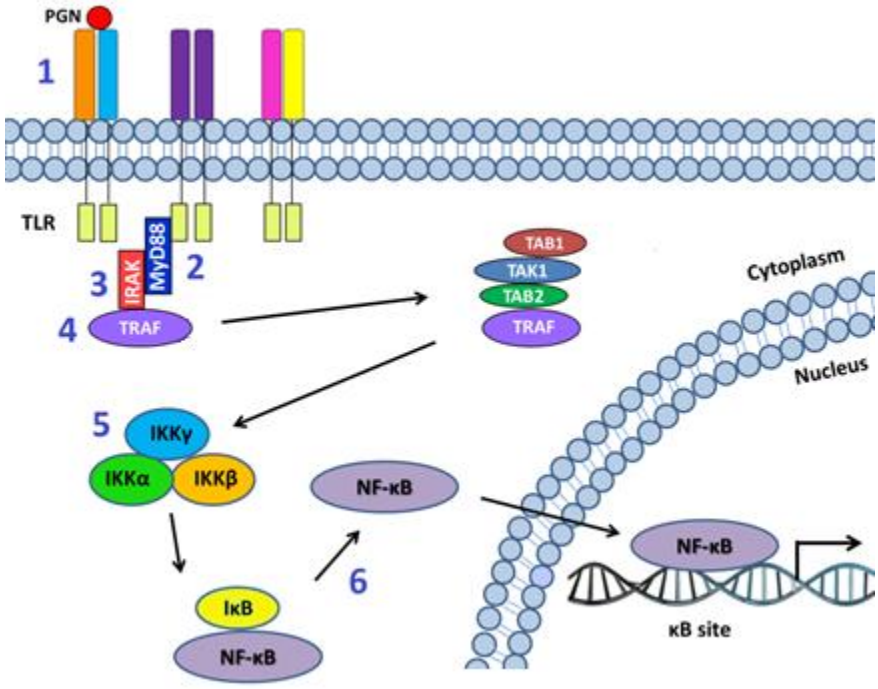

Figure 7. Toll-like receptor mediated nuclear factor-kappaB pathway. 1) The TLR binds to a pathogen-associated molecular pattern (PAMP), such as peptidoglycan (PGN). **2)** Intracellular signaling is initiated by adaptor proteins, e.g. myeloid differentiation factor 88 (MyD88). **3)** MyD88 associates with the cytoplasmic portion of TLR and recruits a downstream interleukin-1 receptor-associated kinase (IRAK) to the receptor complex. **4)** IRAK becomes phosphorylated and dissociates from the complex to recruit tumor necrosis factor receptor-associated factor (TRAF). **5)** This allows downstream kinases, e.g. TAK1, to stimulate an IκB kinase (IKK). **6)** After IKK is activated, NF-κB can translocate to the nucleus to initiate gene transcription.

1.6 Nuclear factor-kappaB binding site prediction

One way to study which genes are regulated by NF-κB transcription factors is to identify NF-κB binding sites, commonly referred to as kappa-binding (κB) sites (Fig. 8). These transcription factor binding sites, which are short genomic sequences (around 5-15 bp), can be found upstream, downstream, or within introns of the genes that they regulate (Lodish *et al.* 2000). In addition, they can be close or far in proximity to the beginning of the open reading frame of the target gene, which is the DNA sequence that codes for a protein. If the putative binding site is within the vicinity of the gene, then it is thought that this sequence might be able to bind to a protein of interest to assist with gene regulation. Transcription factor binding sites can be predicted by using databases, such as the Transcription Factor database (TRANSFAC) and the Open Regulatory Annotation database (ORegAnno) (Wingender *et al.* 2000; Montgomery *et al.* 2005). More recently, Lee and Huang (2014) developed a new transcription factor binding site search algorithm called Length-Aware Site Alignment Guided by Nucleotide Association 2.0 (LASAGNA 2.0), which is a specific algorithm based on the idea that binding sites share a common core with a few varying nucleotides on one or both ends of the sequence (Lee & Huang, 2014). The programs mentioned previously can find sequences that resemble other known binding sites, but LASAGNA 2.0 has been shown to be a faster and more accurate tool for predicting transcript expression (Lee & Huang, 2014).

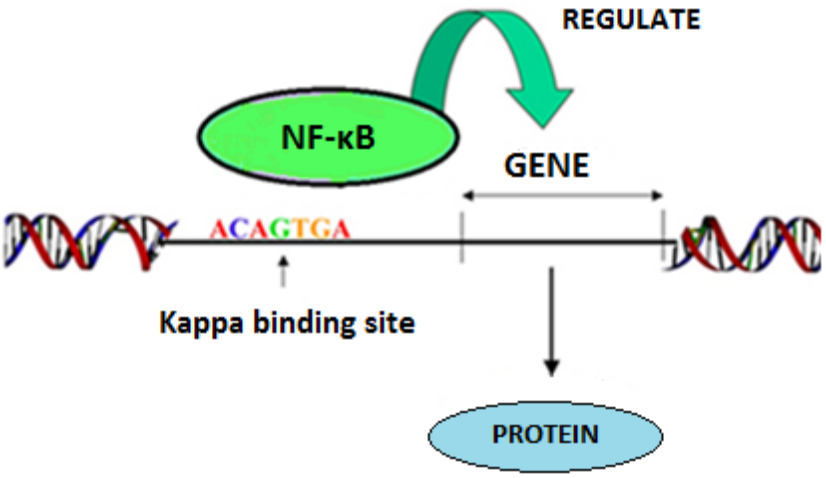

Figure 8. A nuclear factor-kappaB transcription factor interacting with a kappa-binding site to regulate gene expression. Image adapted and edited from www.igem.org.

It is especially important to perform functional assays because we have not yet confirmed if these predicted binding sites, which resemble known vertebrate sequences, may function similarly within invertebrates, such as *B. glabrata.* Transcription factor binding sites can be functionally characterized using a variety of approaches, including a common technique called an electrophoretic mobility shift assay (EMSA) (Fig. 9; Geertz & Maerkl, 2010). EMSAs are used to determine if transcription factor proteins can interact with and bind to nucleic acids. In this assay, various combinations of protein and DNA are incubated and then subjected to gel electrophoresis on a non-denaturing polyacrylamide gel (Hellman & Fried, 2007). DNA probes are labeled with radioisotopes, fluorophores, or biotin, which can then be detected via an array of imaging techniques (Rowe & O'Gara, 2016). This assay is based on the principle that a protein-DNA complex is higher in molecular weight than the labeled DNA probe alone. That is, the protein-DNA complex will migrate down the gel much more slowly compared to a free probe, which generates a characteristic "mobility shift" to compare the location of a protein-DNA complex to DNA alone (Fig. 9; Rowe & O'Gara, 2016). This procedure can also be used to

examine the specificity of a protein-DNA interaction. Mutant probes can be designed to examine how the altered nucleotides may impact the binding affinity of the DNA probe to the recombinant protein (Hellman & Fried, 2007). We can also subject various mixtures of protein and DNA to competition assays. By adding a competing unlabeled DNA probe to the mixture, we can determine if the protein of interest binds to the target probe with greater affinity than the competing probe (Rowe & O'Gara, 2016).

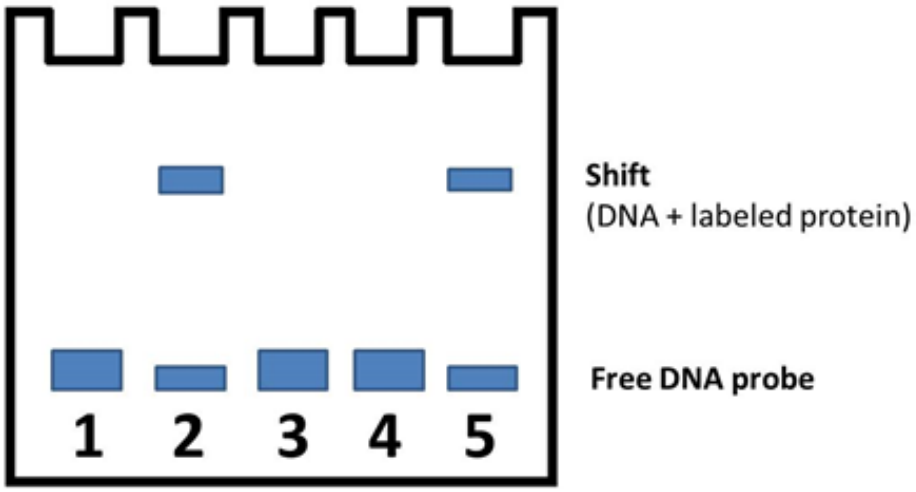

Figure 9. Overview of the electrophoretic mobility shift assay. The assay tests the functionality of a transcription factor by examining protein-DNA binding interactions. Lane 1 represents free DNA. Lane 2 shows a "shift", which indicates that the protein (transcription factor) is interacting with the DNA (binding site). Lane 3 examines the specificity of the interaction; upon altering the DNA sequence, no shift is observed because protein binding is inhibited. Competition assays can be performed. Lane 4 indicates a lack of DNA-protein binding when adding a higher fold excess of unlabeled wild-type probe. Lane 5 reveals that binding can still occur when adding the same fold excess of an unlabeled mutant probe.

EMSAs are sensitive to a variety of factors (e.g., pH, buffer composition, salinity) and therefore require much optimization (Hellman & Fried 2007). In addition, protein-DNA complexes are not at chemical equilibrium when run down a gel, so rapid dissociation during electrophoresis could prevent detection of the complex (Rowe & O'Gara, 2016). Despite these issues, the assay presents several advantages, including the use of small volumes (20 µL or less) as well as the ability to use a wide range of DNA probe lengths (Hellman & Fried, 2007).

1.7 Identification of nuclear factor-kappaB gene targets

More generally, EMSAs have been used to examine *in vitro* DNA-protein interactions. To date, DNA-NF-κB binding has been examined in just two molluscan species, *C. gigas* and *B. glabrata* (Jiang & Wu, 2007; Humphries & Harter, 2015). Jiang and Wu (2007) showed that the recombinant *Ab*-Rel RHD from *H. diversicolor* supertexta can bind to a vertebrate κB site. Likewise, Humphries and Harter (2015) reported that the *B. glabrata* RHD could bind to a vertebrate consensus probe. Humphries and Harter (2015) were also the first to use EMSAs to functionally characterize molluscan κB sites by showing specific binding between a molluscan RHD and a molluscan putative binding site. The molluscan κB sites that were characterized in their study were reported upstream of two immune genes (p38 mitogen-activating protein kinase (MAPK) and IκBα) in *B. glabrata* (Humphries & Harter, 2015). Previous studies suggest that p38 MAPK plays a role in generating immune responses within *B. glabrata.* More specifically, p38 MAPK has been shown to participate in *in vitro* H_2O_2 production, which is an anti-schistosome defense mechanism used by resistant strains of *B. glabrata* (Goodall *et al.* 2004; Yoshino & Humphries *et al.* 2008). Overall, Humphries and Harter (2015) showed that the *B. glabrata* RHD could interact with the κB sites discovered upstream of both *Bg*p38 MAPK and *Bg*IκBα, which implies that the expression of both of these genes may be regulated by *Bg*NF-κB. The findings in this study led us to wonder if *Bg*NF-κB may regulate other immune genes, such as components of the TLR pathway. Previous studies suggest that NF-κB regulates TLRs in vertebrate models, so it was of great interest to determine if a *B. glabrata* NF-κB p65 protein might be able to function similarly (Takeshita *et al.* 2004; Johnson & Tapping, 2007).

The aim of this research project was to determine if NF-κB can regulate other immune genes in the TLR–NF-κB pathway within *B. glabrata*. If NF-κB does indeed regulate such

immune genes, then putative κB sites should be found within the vicinity of the protein coding region of interest. In order to address this question, bioinformatics analyses were used to identify putative κB sites near various immune genes. Following, EMSAs were used to functionally assess the predicted κB sites.

MATERIALS AND METHODS

2.1 Identification of putative kappa-binding sites

Prediction of putative kappa-binding sites

In order to examine whether NF-κB might regulate gene expression of immune genes in the TLR pathway, we searched for putative kappa-binding (κB) sites within the vicinity of the genes of interest. The genomic DNA sequences upstream of protein coding regions for 20 genes representing the TLR mediated NF-κB pathway were acquired from VectorBase (http://www.vectorbase.com), which houses the genome for *B. glabrata*. The amino acid sequence for each gene was aligned to sequences within the snail genome using the Basic Local Alignment Search Tool (BLAST) (Altschul *et al.* 1990). BLAST allows one to search nucleotide and protein databases for sequences that are similar to the sequence of interest. The program, tblastn (peptide vs. translated nucleotide), was used to align the input protein sequence to nucleotides within the snail genome that were translated to hypothetical amino acid sequences. The section of the genomic DNA (scaffold) showing the highest level of identity to the input sequence was selected for further analysis. The upstream region (2000 bp) for each gene was then retrieved and exported as a Fast Alignment (FASTA) sequence.

The upstream region was then inserted into the search engine, Length-Aware Site Alignment Guided by Nucleotide Association 2.0 (LASAGNA 2.0; http://biogrid-lasagna.engr.uconn.edu/lasagna_search/), which is an algorithm that scans DNA sequences of interest for potential transcription factor binding sites (Lee & Huang, 2013). Certain Transcription Factor database (TRANSFAC) nucleotide matrices were selected to search for putative binding sites within the upstream regions of each gene (Wingender *et al.* 2000). Five

NF-κB matrices were used to find potential κB sites (NF-κB p65 (M00052), NF-κB p50 (M00051), NF-κB (M00208), NF-κB (M00194) and NF-κB (M00054)). Statistically significant results were accepted at $P<0.001$.

In addition to searching upstream of the protein coding region, we also looked within introns of several members of the TLR mediated NF-κB pathway. Methods for analyses were the same as above except BLAST was used to find the genomic DNA sequence of the protein coding region for each respective gene. These genomic sequences contained exons that code for the protein, interspersed with introns that do not code for the protein, but can play a regulatory role (e.g., by containing transcription factor binding sites). The sequence analysis software package program, Vector NTI® (Life Technologies, Grand Island, NY, USA), was then used to locate and acquire all the introns within each gene of interest.

Selection of putative kappa-binding sites for electrophoretic mobility shift assays

Members of the TLR-mediated NF-κB pathway were selected for functional analyses based on their fundamental role in the pathway. When looking for putative binding sites both upstream and within introns, some sequences contained many successful hits. The criteria for selection of predicted binding sites is as follows: (1) the sequence of interest was required to appear on the sense strand; (2) overlapping sequences generated from a single matrix were discarded and the hit with the best score was selected; and (3) hits for more than one matrix localized in the same site were discarded to avoid redundancy. If these criteria were met, the sequence that corresponded to the highest ranked matrix was considered as a positive hit. Out of the five matrices utilized in this study, putative binding sites that were predicted based on similarity to the NF-κB p65 matrix (M00052) were ranked the highest because the *B. glabrata* RHD used in EMSAs represents the RHD from *Bg*NF-κB p65 (NCBI Accession No.

ACZ25559.1). NF-κB p65 proteins often function as a heterodimer with NF-κB p50 proteins so additional matrices were selected because they represent binding sites recognized by vertebrate NF-κB p50 proteins.

Nine predicted κB sites were selected for further functional analyses using EMSAs. These sequences were selected based on their score, i.e. the predicted site with the highest score, lowest P-value, and lowest E-value (for any one gene). Oligonucleotide probes containing the predicted binding site were required to see if they interacted with the RHD protein via EMSAs. Wild-type oligonucleotide probes (30 bp in length) were designed and comprised of a 10 bp putative kappa-binding site with an additional 10 bp on both the 5' and 3' flanking regions. In order to generate mutant probes, the first half-site of the kappa-binding site was altered to 5'-CTCAA-3'. This five bp sequence was selected because it was shown to inhibit binding in a previous study between the *Bg*RHD and its associated *B. glabrata* κB site (Humphries & Harter, 2015). Both forward and reverse strands for these probes were synthesized by the Midland Certified Reagent Company (Midland, Texas, USA). Following synthesis, probes were purified using high performance liquid chromatography (HPLC).

2.2 Biotin labeling

The oligonucleotide probes were labeled with biotin, which is a stable, high-affinity protein, allowing the probe to be visualized upon later detection. Specifically, the probes were labeled with 1,3 biotinylated ribonucleotides onto the 3'-OH end of the DNA strands, according to the manufacturer's instructions (Biotin 3' End DNA Labeling Kit, Pierce ™, Rockford, IL, USA). Biotin labeling reactions (25 µL) were prepared by adding and mixing the following components: 12.5 µL Milli Q water, 5 µL 5 X TdT reaction buffer, 2.5 µL unlabeled oligonucleotide (1 µM), 2.5 µL biotin-11-UTP (5 µM), and 2.5 µL diluted terminal

deoxynucleotidyl transferase (TdT) (1.5 U/µL). The TdT enzyme catalyzed non-template-direct deoxynucleotide incorporation onto the 3'-OH end of each strand for each probe. Reactions were incubated at 37ºC for 30 min. Following incubation, 2.5 µL of 0.2 M ethylenediaminetetraacetic acid (EDTA) were added to each sample in order to stop the reaction. Fifty microliters of chloroform:isoamyl alcohol (24:1) were added to each reaction to extract the TdT. The mixture was then vortexed briefly and centrifuged at room temperature (2 min at 13.2 revolutions per min (RPM)) (Eppendorf 5415R, Hamburg, Germany) to separate the phases. Forward and reverse probes were incubated for 1 hr at room temperature. For unlabeled probes, equal parts of forward and reverse probes were combined in one tube and were given 1 hr to anneal at room temperature. All samples were stored at -20ºC for future use.

2.3 Preparation of Samples

The sequence representing the RHD from the *B. glabrata* NF-κB p65 (amino acids 83-392; NCBI Accession No. ACZ25559.1) was inserted into a vector (pET-32a). The recombinant *Bg*RHD was then expressed in *Escherichia coli* by the biotechnology company, GenScript (Piscataway, NJ, USA). Double stranded biotin-labeled probes (20 fmol) were incubated with 0.3 mg/mL *Bg*RHD, 0.5 µg/µL Poly I:C (ThermoScientific, Waltham, MA, USA), in addition to 2 µL 50% glycerol, and 2 µL 10 X binding buffer (10 mM Tris, 50 mM potassium chloride (KCl), 1 mM Dithiothreitol (DTT), pH 9.0). Samples were then incubated at room temperature for 30 min. Competition assays were also completed to examine the specificity of the interaction between the *Bg*RHD and the labeled wild-type oligonucleotide probes. These assays were used to determine if the *Bg*RHD could still bind to the labeled wild-type probe when in the presence of an excess unlabeled mutant probe. Additionally, the competition assays were performed to examine if the *Bg*RHD may bind with less affinity to the labeled wild-type probe when in the

presence of an excess unlabeled wild-type probe. Both unlabeled wild-type and unlabeled mutant probes were added at a 500-fold excess (4 pmol) for competition assays.

2.4 Electrophoretic mobility shift assay

Gel electrophoresis

A non-denaturing, 6% polyacrylamide gel in 0.25 X Tris borate EDTA buffer (EDTA pH 8.0; all adjusted to pH 9.0) was prepared and degassed for 15 min before adding tetramethylethylenediamine (TEMED) and ammonium persulfate (APS). The gel was then poured and was allowed to polymerize using the Bio-Rad Mini-PROTEAN® tetra cell hand casting module (Berkeley, CA). Acrylamide gels were given 2 hr to polymerize prior to use. Gels, which were wrapped in damp paper towels and plastic wrap, were stored at 4ºC until used within 48 hr of polymerization.

Gel electrophoresis was performed using a Bio-Rad MINI-PROTEAN® tetra vertical electrophoresis cell system (Berkeley, CA). After assembling the apparatus, each well was flushed out twice with 200 µL 0.25 X TBE buffer. The acrylamide gel was pre-run at 100 V in 0.25 X TBE buffer for 30 min to remove excess traces of APS. Afterwards, the wells were rinsed again twice with 0.25 X TBE buffer. Samples (20 µL) were then loaded into each well. Note that an aliquot (2.5 µL) of loading dye (ThermoScientific, Waltham, MA, USA) was added to the lane containing solely the labeled oligonucleotide probe in order to track the samples as they ran down the gel. The voltage was then reduced to 70 V and samples were allowed to migrate down the gel for 20-30 min.

Transferring the gel

A Mini Trans-Blot® tetra cell tank blotting system (Bio-Rad, Berkeley, CA) was used to transfer the DNA and protein contents from the gel onto a nylon membrane (type B positive, Fluka Analytical, St. Louis, MO, USA). The nylon membrane was presoaked in 0.25 X TBE buffer for 15 min prior to use. The membrane sandwich was prepared according to the manufacturer's instructions with sponges, filter paper (Extra Thick Blot Filter Paper, Bio-Rad, Berkeley, CA), the polyacrylamide gel, and the pre-soaked nylon membrane. The gel was then given 1-2 hr to transfer at 33 V in 0.25 X TBE and an ice pack was added during the transfer to prevent overheating of the gel. Afterwards, the free DNA, free protein, and protein-DNA complexes were all cross-linked onto the membrane by using a UVP ultraviolet crosslinker (CL-1000, Upland, CA).

Chemiluminescent detection

Chemiluminescence was used to detect the presence of biotin-labeled probes on the nylon membrane, according to the manufacturer's instructions (Chemiluminescent Nucleic Acid Detection Module, ThermoScientific, Waltham, MA, USA). The washing buffer and blocking buffer from the detection module were incubated in a water bath at 40ºC for approximately 10 min to dissolve all particulate matter. The transferred nylon membrane was then soaked in blocking buffer for 15 min. The membrane was incubated for an additional 15 min in blocking buffer that contained stabilized, streptavidin-horseradish peroxidase (SA-HRP) conjugate (1:300 dilution). This peroxide-conjugated streptavidin binding protein interacts very strongly with the biotin-labeled probes and provides signal amplification during visualization. The nylon membrane was then rinsed several times in 1 X wash buffer (4 x 5 min). Afterwards, the membrane was incubated in equilibration buffer for 5 min. Immediately following this step, the

membrane was developed with compatible enhanced chemiluminescence (ECL) reagents (Chemiluminescent Nucleic Acid Detection Module, ThermoScientific, Waltham, MA, USA). Equal parts (0.5 mL each) of stable peroxide solution and luminol enhancer were applied onto a Kodak imager (Image Station 4000 mm Pro, Rochester, NY) and the membrane was then placed protein side down onto the ECL solution. The membrane was exposed for 5 min. Images were collected and stored as tagged image file format (TIFF) files on the computer. For each predicted κB site that was tested, EMSAs were repeated for a total of three replicates.

RESULTS

3.1 Identification of putative kappa-binding sites associated with the Toll-like receptor mediated nuclear factor-kappaB pathway in *B. glabrata*

Upstream analysis

Approximately 2000 bp of genomic DNA upstream of the protein coding region for 20 members of the TLR pathway were surveyed for potential κB sites (Table 1). Sequences upstream of 16 of these genes (*Bg*TLR1, *Bg*TLR2, *Bg*TLR4, *Bg*MyD88, *Bg*SARM, *Bg*IRAK, *Bg*TAB, *Bg*TAK, *Bg*IKKα, *Bg*IKKγ, *Bg*Uev1a, *Bg*IκBβ, *Bg*NF-κB p65, *Bg*NF-κB p105, *Bg*LITAF, and *Bg*AP-1) contained putative κB sites. Often, more than one putative binding site was predicted upstream of each of these genes. Typically, there were between one and four reported hits for the 16 upstream regions described above. Out of the 20 DNA sequences subjected to analysis, only four (*Bg*TLR3, *Bg*TRAF, *Bg*IKKγ, and *Bg*IRF) did not contain putative κB sites.

Table 2. Identification of putative kappa-binding sites within 2000 bp upstream of 16 members of the TLR–NF-κB pathway in *B. glabrata*. The genomic DNA upstream of the protein coding region was submitted to the transcription binding site search engine, Length-Aware Site Alignment Guided by Nucleotide Association 2.0 (LASAGNA 2.0). Data highlighted in blue represent sequences that were selected for further analyses and their respective genes. Position number indicates number of base pairs upstream of the A nucleotide of the start codon (ATG).

Control Region	Sequence	Position	Number of κB sites
***Bg*TLR1**	agtgattccc	-274	1
***Bg*TLR2**	ggaagtttc gggaagtttc	-25 -1928	2
***Bg*TLR3**	-	-	None
***Bg*TLR4**	ggaatcttcc	-1829	1
***Bg*TLR5**	gggattttac	-1466	1
***Bg*MyD88**	ggacgttctt	-1639	1
***Bg*SARM**	ggaattcact gggaatgtct	-665 -512	2
***Bg*TRAF**	-	-	None
***Bg*IRAK**	ggtattttac	-175	1
***Bg*TAB**	ggaaactcca	-457	1
***Bg*TAK**	-	-	None
***Bg*IKKα**	ggggatctca	-862	1
***Bg*IKKγ**	tgtactttcc	-1464	1
***Bg*Uev1a**	ggaacttacc	-1157	1
***Bg*IκBβ**	tggaaattcc ggggaatccc tggaacttcc tcgaatttcc	-159 -290 -1742 -1774	4
***Bg*NF-κB p105**	gggaaagtcc	-1257	1
***Bg*NF-κB p65**	ggggattttc gggattttct	-820	2
***Bg*LITAF**	ggtttttcc gggtttctcc	-250 -267	2
***Bg*IRF**	-	-	None
***Bg*AP-1**	gggaaatcct	-356	1

In order to test our methodology, we surveyed the upstream region of two additional genes (*Bg*IκBα and *Bg*p38 MAPK) that were previously analyzed by Humphries and Harter (2015). Four hits were displayed for each gene (Table 6). Sequences exhibiting the highest scores from our bioinformatics analyses were identical to the putative κB sites that were selected for functional analyses by Humphries and Harter (2015).

Intron analysis

We also surveyed the introns for potential putative κB sites from the 20 members of the TLR–NF-κB pathway (Table 3). Twelve genes (*Bg*TLR1, *Bg*MyD88, *Bg*SARM, *Bg*TRAF, *Bg*IRAK, *Bg*TAK, *Bg*TAB, *Bg*Uev1a, *Bg*IKKγ, *Bg*NF-κB p105, *Bg*NF-κB p65 and *Bg*LITAF) contained introns for analyses. There were between one and eight predicted putative κB sites for these genes. In contrast, no putative binding sites were identified within the introns of *Bg*IκBβ. It should be noted that the genes for *Bg*TLR2, *Bg*TLR3, *Bg*TLR4 and *Bg*AP-1 did not contain introns so they were not included in this particular analysis. Additionally, it was not possible to analyze the introns within *Bg*IKKα because the gene was split between two scaffolds in the current draft of the genome.

Table 3. Identification of putative kappa-binding sites within introns of the genes coding for 12 members of the TLR–NF-κB pathway. The genomic DNA within introns of the protein coding region was submitted to the transcription factor binding site search engine, Length-Aware Site Alignment Guided by Nucleotide Association 2.0 (LASAGNA 2.0). Position number indicates number of base pairs downstream of the A nucleotide of the start codon (ATG).

Control Region	Sequence	Position	Number of κB sites
***Bg*TLR1**	gggacataac gggaagtttc	+1,046 +273	2
***Bg*MyD88**	cggagtttcc gggaaactcc ggaaaattac ggaactttac ttggaaacttctga tggaatcccc ggacttcaca	+3,004 +2,438 +2,881 +2,422 +3,714 +5,716 +7,608	7
***Bg*SARM**	tggagattcc aaggcacattcctc tttgaaatttctct aatggaattactat gggaataacc gggaagttac attgaactttacta	+6,496 +4,472 +10,659 +11,642 +14,462 +17,830 +20,316	7
***Bg*TRAF**	tggggattttc ttgaatttcc ggggatttct gggaattcac gggaaattat gggactctgct aatgaaatttacta gatggaacttctcc	+11,633 +23,355 +23,383 +16,030 +20,510 +17,758 +16,376 +22,828	8
***Bg*IRAK**	ggatactcca ggagatctcc	+10,219 +10,821	2
***Bg*TAB**	atggaactttcttt	+585	1
***Bg*TAK**	tgtggacataccac aaggggacattatat gggaaaatcc	+3,352 +5,274 +14,373	3
***Bg*IKKγ**	tgtactttcc tggaattcc	+1,704 +6,120	2
***Bg*Uev1a**	tgaaaattcc ggacactcca aatggatttatat cgggtttccc agtggacttactga	+5,010 +1,666 +549 +6,188 +6,297	5
***Bg*IκBβ**	-	-	None

***Bg*NF-κB p105**	tttggatgttcctt agagaaatttactt tgaaaattcc ttgaattccc ggaattccc gggacagccc gggatttttaa	+289 +2,148 +5,912 +11,603 +17,037 +18,145 +18,524	7
***Bg*NF-κB p65**	ggggattttc tgggaaattctca gtgattttcaa ttgtgaccttccat	+1,340 +3,093 +8,734 +12,722	4
***Bg*LITAF**	ggggttttct taggtactttcttc	+2,449 +3,919	2

Selection of genes for functional analyses using electrophoretic mobility shift assays

Nine genes (*Bg*TLR1, *Bg*TLR2, *Bg*MyD88, *Bg*IRAK, *Bg*IKKα, *Bg*IκBβ, *Bg*NF-κB p105, *Bg*NF-κB p65 and *Bg*LITAF) were selected for functional analyses. When more than one κB site was predicted upstream of the coding region for a gene of interest, the predicted site with the highest score and lowest respective E-value and P-value was selected for EMSA assays (Table 4). Putative κB sites were predicted upstream of the protein coding regions for *Bg*TLR1 (5'-AGTGATTCCC-3') and *Bg*TLR2 (5'-GGGAAGTTTC-3') at positions -274 bp and -25 bp from the start codon, respectively. Additionally, putative κB sites were reported upstream of the protein coding regions for *Bg*MyD88 (5'-GGACGTTCTT-3'), and *Bg*IRAK (5'-GGTATTTTAC-3') at -1639 bp and -175 bp from the start codon, respectively. Putative κB sites were also shown for *Bg*IKKα (5'-GGGGATCTCA-3'), *Bg*IκBβ (5'-TGGAAATTCC-3'), *Bg*NF-κB p105 (5'-GGGGATTTTC-3'), and *Bg*NF-κB p65 (5'-GGGAAAGTCC-3'). The putative κB sites were predicted upstream of the coding region for these genes, -862 bp, -159 bp, -1257 bp, and -820 bp from the start codon, respectively. Additionally, a LITAF homologue was selected for functional assays to determine if *Bg*NF-κB can potentially play a regulatory role in the expression of another transcription factor. The putative binding site for *Bg*LITAF

(5’-GGGTTTTTCC-3’) was predicted -250 bp upstream the start codon. The 10 bp sequences for all of the genes listed above were used to create oligonucleotide probes (Table 4).

Table 4. Putative kappa-binding site predictions in *B. glabrata*. The predicted sequences were found using the transcription factor binding site search engine, Length-Aware Site Alignment Guided by Nucleotide Association 2.0 (LASAGNA 2.0). The position is indicative of how close binding site is from the first nucleotide (A) of the start codon (ATG). The score represents the sequence similarity of the predicted binding site to the respective matrix. The P-value is the probability that an event occurs by chance ($P<0.05$), whereas the E-value is the expected number of times a hit of the same or higher score is found in the promoter sequence by chance.

Name	Sequence	Position	Score	P-value	E-value
***Bg*TLR1**	AGTGATTCCC	-274	8.83	0.0006	1.31
***Bg*TLR2**	GGCCATTCCC	-25	8.09	0.001	2.09
***Bg*MyD88**	GGACGTTCTT	-1639	6.89	0.001	2.03
***Bg*IRAK**	GGTATTTTAC	-175	9.91	0.0003	0.63
***Bg*IKKα**	GGGGATCTCA	-862	7.25	0.0003	0.58
***Bg*IκBβ**	TGGAAATTCC	-159	13.78	0	0
***Bg*NF-κB p105**	GGGAAAGTCC	-1257	9.11	0.0002	0.42
***Bg*NF-κB p65**	GGGAACTTCC	-820	10.96	0.0001	0.15
***Bg*LITAF**	GGGTTTTTCC	-250	12.49	0.00008	0.149

Table 5. Oligonucleotide probes used in electrophoretic mobility shift assays. All probes (30 bp long) consisted of a central predicted κB site (in purple), in addition to both 5' and 3' flanking regions. Mutant probes were generated by changing the first half-site of the κB site to 5'-CTCAA-3' (in red). Only the forward primer is displayed for both wild-type and mutant probes, but both forward and reverse strands were used in electrophoretic mobility shift assays.

Probe	Nucleotide sequence
*Bg*TLR1 wild-type	**ttcttttacgagtgattcccc aggcggcta**
*Bg*TLR1 mutant	**ttcttttacgctcaattcccc aggcggcta**
*Bg*TLR2 wild-type	**tgtgcaactgggccattcccc catttatat**
*Bg*TLR2 mutant	**tgtgcaactgctcaattcccc catttatat**
*Bg*MyD88 wild-type	**tttcttccttggacgttctt cagtgccagc**
*Bg*MyD88 mutant	**tttcttccttctcaattctt cagtgccagc**
*Bg*IRAK wild-type	**atgtgtcattggtattttac attcataaaa**
*Bg*IRAK mutant	**atgtgtcattctcaatttac attcataaaa**
*Bg*IKKα wild-type	**tgccagccctggggatctca ttatatgtct**
*Bg*IKKα mutant	**tgccagccctctcaatctca ttatatgtct**
Bg IκBβ wild-type	**tgttgtcgcatggaaattcc accgaaaaat**
*Bg*IκBβ mutant	**tgttgtcgcactcaaattcc accgaaaaat**
*Bg*NF-κB p105 wild-type	**accttaactagggaaagtcc ctacaatatt**
*Bg*NF-κB p105 mutant	**accttaactactcaaagtcc ctacaatatt**
*Bg*NF-κB p65 wild-type	**gggtccgggggggaacttcc ccccccccc**
*Bg*NF-κB p65 mutant	**gggtccgggg ctcaacttcc ccccccccc**
*Bg*LITAF wild-type	**tccagctcgtgggtttttcc agcgtgtgaa**
*Bg*LITAF mutant	**tccagctcgtctcaatttcc agcgtgtgaa**

3.2 Electrophoretic mobility shift assays

EMSAs were performed to determine whether the *Bg*RHD could recognize and bind to several putative κB sites that were identified upstream of the protein coding region for members of the TLR mediated NF-κB pathway. EMSA analyses confirmed that *Bg*RHD could recognize

and bind to six wild-type, biotin-labeled probes: *Bg*TLR2, *Bg*IRAK, *Bg*IκBβ, *Bg*NF-κB p105, *Bg*NF-κB p65 and *Bg*LITAF. In comparison, *Bg*RHD binding was substantially reduced or inhibited when incubated with mutant, biotin-labeled probes. Competition assays were performed to determine if the interaction between the *Bg*RHD and probe complex was specific. Binding between the *Bg*RHD and wild type, biotin-labeled probes was decreased and in some occasions eliminated by the addition of a 500-fold excess of the equivalently unlabeled wild type probe. In contrast, binding between the *Bg*RHD and labeled wild-type probe remained unchanged with the addition of a 500-fold excess of equivalently unlabeled mutant probe. This indicates that the interactions between both *Bg*RHD and six wild-type probes are specific (Fig. 10-15).

However, our findings vary from κB site to κB site in several ways. Binding interactions between the *Bg*RHD and altered mutant probes (*Bg*TLR2 and *Bg*NF-κB p65) were not completely inhibited in some cases (Fig. 10 and 14). Additionally, the findings from our competition assays varied for some of the probes, where a reduction in specific binding was displayed between the *Bg*RHD and the labeled wild-type probe (*Bg*TLR2, *Bg*IRAK, *Bg*NF-κB p65, and *Bg*LITAF) when incubated with a higher fold excess of unlabeled equivalent probe (Fig. 10-11, and 14-15, respectively; lane 4). In contrast, binding between the *Bg*RHD and labeled wild-type probe for both *Bg*IκBβ and *Bg*NF-κB p105 was completely outcompeted when the specific DNA-protein complex was incubated with the addition of a higher fold excess of unlabeled wild-type probe (Fig. 12-13; lane 4).

EMSAs were also carried out for the putative κB sites identified upstream of *Bg*TLR1, *Bg*MyD88, and *Bg*IKKα. When examining *Bg*TLR1, *Bg*RHD bound to the labeled wild-type probe but *Bg*RHD also bound to the labeled mutant probe. This implies that the interaction

between the *Bg*RHD and labeled wild-type was non-specific. In addition, *Bg*RHD did not bind to labeled wild-type probes for either *Bg*MyD88 or *Bg*IKKα.

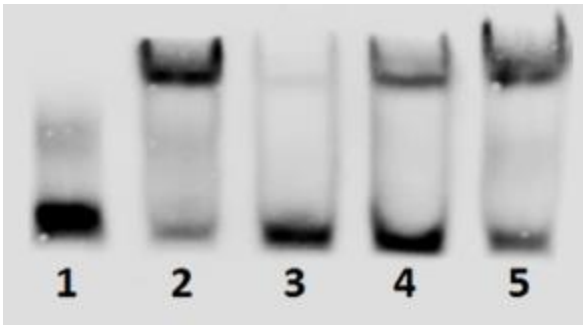

Figure 10. Electrophoretic mobility shift assay for *Bg*TLR2. The labeled wild-type probe is displayed in lane 1. The *Bg*RHD was incubated with either the labeled wild-type or labeled mutant probe (lanes 2 and 3, respectively). Competition assays assess the specificity of the interaction by incubating the *Bg*RHD and labeled wild-type probe with either excess unlabeled wild-type or unlabeled mutant probes (lanes 4 and 5, respectively) (n = 3).

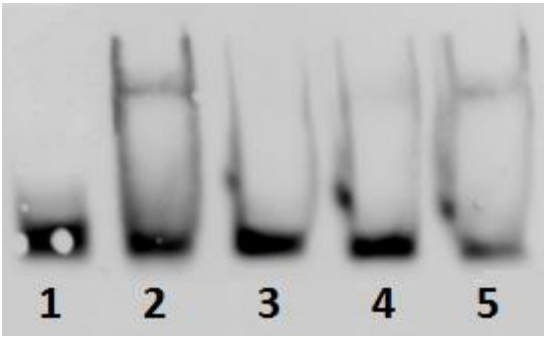

Figure 11. Electrophoretic mobility shift assays for *Bg*IRAK. The labeled wild-type probe is displayed in lane 1. The *Bg*RHD was incubated with either the labeled wild-type or labeled mutant probe (lanes 2 and 3, respectively). Competition assays assess the specificity of the interaction by incubating the *Bg*RHD and labeled wild-type probe with either excess unlabeled wild-type or unlabeled mutant probes (lanes 4 and 5, respectively) (n = 3).

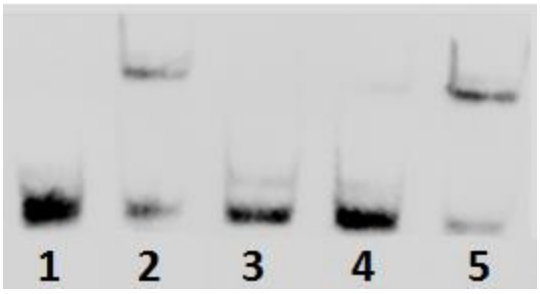

Figure 12. Electrophoretic mobility shift assays for *Bg*IκBβ. The labeled wild-type probe is displayed in lane 1. The *Bg*RHD was incubated with either the labeled wild-type or labeled mutant probe (lanes 2 and 3, respectively). Competition assays assess the specificity of the interaction by incubating the *Bg*RHD and labeled wild-type probe with either excess unlabeled wild-type or unlabeled mutant probes (lanes 4 and 5, respectively) (n = 3).

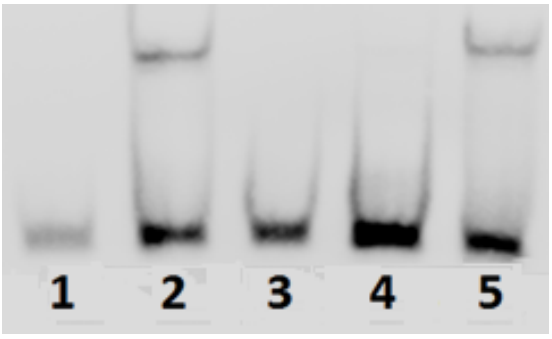

Figure 13. Electrophoretic mobility shift assays for *Bg*NF-κB p105. The labeled wild-type probe is displayed in lane 1. The *Bg*RHD was incubated with either the labeled wild-type or labeled mutant probe (lanes 2 and 3, respectively). Competition assays assess the specificity of the interaction by incubating the *Bg*RHD and labeled wild-type probe with either excess unlabeled wild-type or unlabeled mutant probes (lanes 4 and 5, respectively) (n = 3).

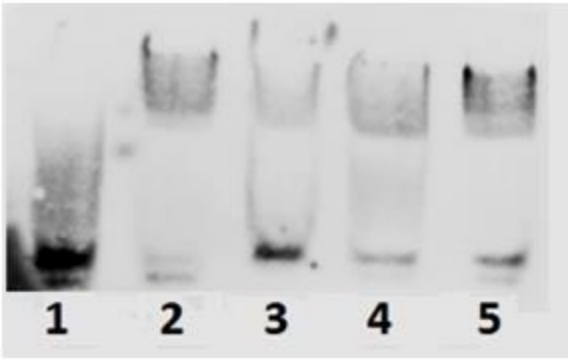

Figure 14. Electrophoretic mobility shift assays for *Bg*NF-κB p65. The labeled wild-type probe is displayed in lane 1. The *Bg*RHD was incubated with either the labeled wild-type or labeled mutant probe (lanes 2 and 3, respectively). Competition assays assess the specificity of the interaction by incubating the *Bg*RHD and labeled wild-type probe with either excess unlabeled wild-type or unlabeled mutant probes (lanes 4 and 5, respectively) (n = 3).

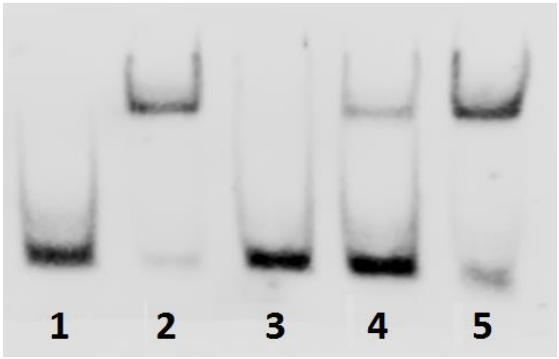

Figure 15. Electrophoretic mobility shift assays for *Bg*LITAF. The labeled wild-type probe is displayed in lane 1. The *Bg*RHD was incubated with either the labeled wild-type or labeled mutant probe (lanes 2 and 3, respectively). Competition assays assess the specificity of the interaction by incubating the *Bg*RHD and labeled wild-type probe with either excess unlabeled wild-type or unlabeled mutant probes (lanes 4 and 5, respectively) (n = 3).

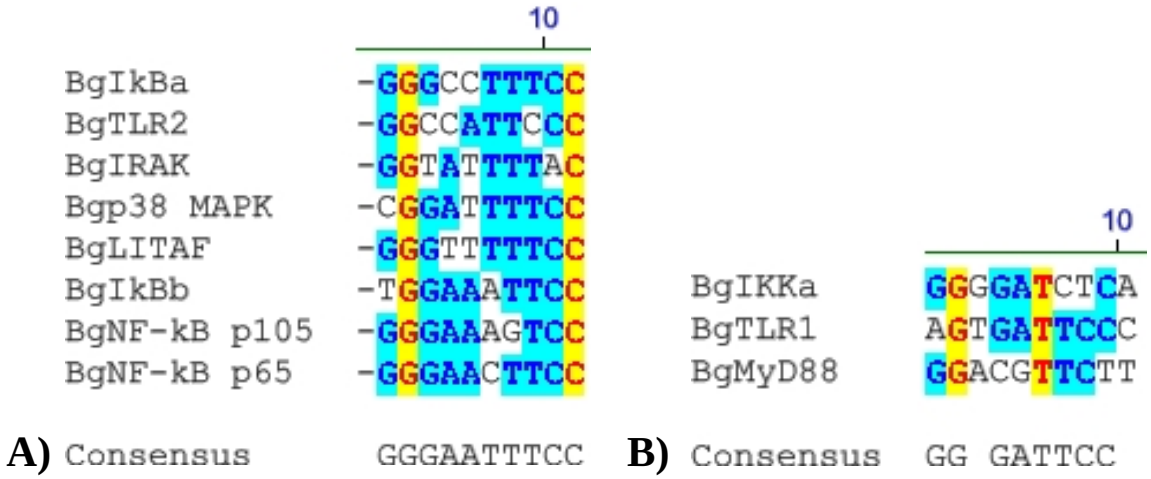

Figure 16. Sequence alignment of *B. glabrata* kappa-binding sites predicted upstream of protein coding regions for members of the Toll-like receptor pathway. A) Alignment with sequences showing specific binding via EMSAs. Note the 5' and 3' conserved residues. **B)** Alignment with sequences that did not show binding to the *Bg*RHD via EMSAs. Sequences (10 bp) were aligned using Vector NTI®. Note that *Bg*IκBα and *Bg*p38 MAPK sequences were obtained from Humphries and Harter's (2015) study. Yellow regions represent nucleotides that are identical across all aligned sequences, whereas blue regions show highly conserved with an identity level ranging between 50% and 99.9%.

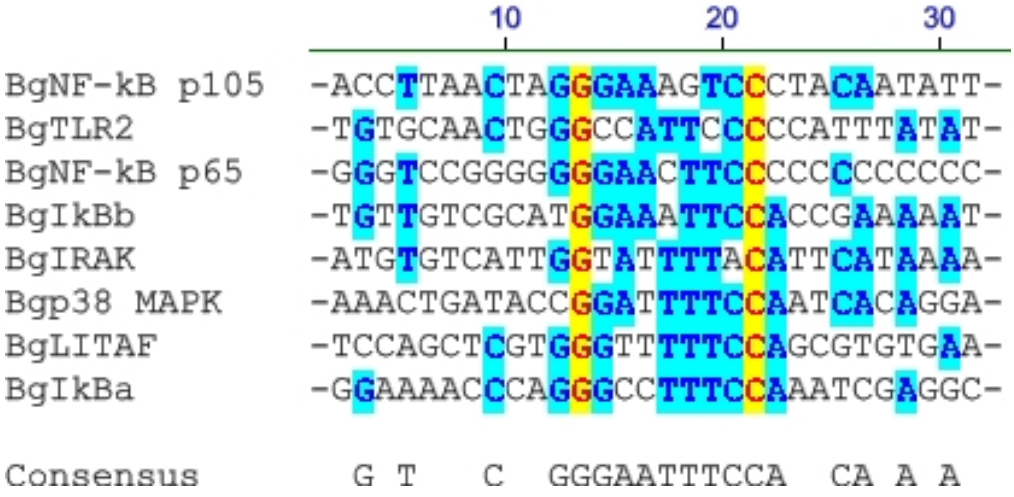

Figure 17. Multiple sequence alignment of *B. glabrata* oligonucleotide probe sequences predicted upstream of protein coding regions for members of the Toll-like receptor pathway. Sequences (30 bp) were aligned using Vector NTI®. Note that *Bg*IκBα and *Bg*p38 MAPK sequences were taken from Humphries and Harter's (2015) study. Yellow regions represent nucleotides that are identical across all aligned sequences. Blue regions show highly conserved with an identity level ranging between 50% and 99.9%.

Figure 18. *B. glabrata* kappa-binding site consensus sequence. All κB sites that bound successfully to the *Bg*RHD were aligned using MEME-ChIP (http://meme-suite.org/tools/meme-chip). Note the 5' and 3' conserved residues.

DISCUSSION

4.1 Characterization of κB sites upstream of genes involved in the TLR–NF-κB pathway

NF-κB related homologues have been identified in several molluscan species, although their functions and regulatory mechanisms are not well understood. One principal strategy for characterizing transcription factors is to identify and characterize DNA-binding sites that are compatible with the transcription factor of interest. The present study suggests that *Bg*NF-κB can potentially regulate the expression of protein coding regions for a variety of members of the TLR pathway. Bioinformatics analyses revealed the presence of several putative NF-κB binding sites upstream of 16 genes of interest (Table 2). Additionally, several putative κB sites were predicted within introns of 12 genes of interest (Table 3). Out of the nine oligonucleotide probes that were subjected to EMSA analyses, six (*Bg*TLR2, *Bg*IRAK, *Bg*IκBβ, *Bg*NF-κB p105, *Bg*NF-κB p65 and *Bg*LITAF) showed sequence specific binding interactions. In each of these cases, *Bg*RHD recognized and bound to the labeled κB site of interest. The interactions between the NF-κB protein and respective κB sites were sequence specific, as very little to no binding was detected between the *Bg*RHD and the probe that contained the altered 5' half-site. Furthermore, competition assays demonstrated that the *Bg*RHD and labeled wild-type probe still bound in the presence of a 500-fold excess amount of unlabeled mutant probe, further supporting the specificity of this interaction. Our successful EMSA analyses suggest that these genes may all be regulated in part by the *B. glabrata* NF-κB p65 transcription factor. These results expand upon those of Humphries and Harter (2015), in that six new κB sites have now been shown to bind to the same *B. glabrata* RHD. In response to our initial research question, our findings suggest that *Bg*NF-κB may regulate selective components of the TLR–NF-κB immune response.

Our findings add to our current understanding of overall NF-κB function. Until now, no other reports have shown the potential for NF-κB to play a role in regulating TLR expression in mollusks, so it is of particular interest that *Bg*RHD bound specifically to the κB site that was discovered upstream of *Bg*TLR2. Our findings can be compared to studies done in vertebrates (Musikacharoen *et al.* 2001; Tapping *et al.* 2007). For example, Musikacharoen *et al.* (2001) used EMSAs to show specific protein binding to a κB site (5'-GGGAACTTCC-3') that was upstream of a mouse TLR. Furthermore, when comparing the mouse κB site with our snail κB site reported upstream of *Bg*TLR2, it was seen that they shared an 80% identity in sequence. The fact that studies have shown the potential for NF-κB regulation of TLR expression in vertebrates supports our hypothesis that NF-κB may regulate TLR expression in *B. glabrata.*

In addition to surveying upstream regions of TLRs for potential binding sites, we also used bioinformatics to search for putative κB sites upstream of several kinases that are activated downstream of TLRs. Three genes (*Bg*IRAK, *Bg*IKKα and *Bg*IKKγ) showed potential κB sites upstream of the start codon. In the present study, the presence of specific binding between the *Bg*RHD and the κB upstream of *Bg*IRAK implies that *Bg*NF-κB might be able to regulate *Bg*IRAK expression. NF-κB regulation of IRAK has not been shown previously in an invertebrate species, however Cui *et al.* (2010) reported a putative NF-κB binding site (5'-GGGACGCTGG-3') -111 bp upstream of the transcriptional start site of the human *IRAK-2* gene of stressed human astroglial cells. Although functional assays were not directly performed on the predicted κB site, other experiments showed that NF-κB mediated transcriptional control occurred between -119 and +12 bp of the IRAK-2 promoter (Cui *et al.* 2010). This region encompasses the predicted κB site, which suggests that the human *IRAK-2* κB site might be

under regulatory control by NF-κB. Our EMSA data are the first to suggest that IRAK expression may be regulated by NF-κB in mollusks.

Highly specific binding observed between *Bg*RHD and the putative κB site reported upstream of IκBβ is similar to the interaction reported in Humphries and Harter's (2015) study, where the same *B. glabrata* RHD also bound to a κB site discovered upstream of IκBα. Our EMSA data demonstrate that the β-subunit can function in a similar manner to that of IκBα. More broadly, these results imply that *Bg*NF-κB might be able to self-regulate by promoting the transcription of its inhibitory protein, *Bg*IκBβ, in a negative feedback loop. Additionally, NF-κB regulation of IκBs has also been seen in other studies, which may suggest that this negative feedback loop can be conserved among other species (De Martin *et al.* 1992; Brown *et al.* 1993; Sun *et al.* 1993; Thompson *et al.* 1995).

Until now, mostly RT-PCR studies have been utilized to examine NF-κB proteins in mollusks (Adema *et al.* 2009; Zhang & Coultas, 2011). In vertebrates, p105 is a precursor protein that can undergo co-translational processing by the 26S proteasome to generate p50 (Moorthy *et al.* 2006). The p105 protein has been shown to function as both an NF-κB precursor as well as an IκB protein (Wan *et al.* 2009). It is of great interest that the κB site reported upstream of *Bg*NF-κB p105 showed specific binding with the *Bg*RHD, as this implies that *Bg*NF-κB may self-regulate by promoting the transcription of p105. Alternatively, perhaps p105 could function as a heterodimer with p65 if p105 can be cleaved to p50, which has been observed in other species (Moorthy *et al.* 2006). Additionally, our EMSA data suggest specific binding between the *Bg*RHD and the κB site upstream of *Bg*NF-κB p65. Interestingly, the 5' (5'-GGGTCCGGGG-3') and 3' (5'-CCCCCCCCCC 3') regions of the oligonucleotide probe were very G rich and C rich, respectively. These regions may be contributing to the smearing,

which is seen in Fig. 14. Although we do not know the exact mechanism by which this could be occurring, perhaps the flanking regions may be folding over one another to create hairpins, which would interfere with the proper annealing of the forward and reverse strands. Ultimately, our findings suggest that *Bg*NF-κB can regulate its own expression through an auto-regulatory feedback loop (Scott *et al.* 1993).

In general, the characterization of NF-κB binding sites suggests that NF-κB may be regulating the expression of members of the TLR pathway. However, we do not expect NF-κB to be the only transcription factor regulating immune genes in *B. glabrata*. Homologues of additional transcription factors, e.g. STAT, STAT2, and CREB, have been identified in *B. glabrata* (Zhang & Coultas, 2011). In other species, these transcription factors have also been shown to regulate immune responses (Wang *et al.* 2003; Wen *et al.* 2010).

In the present study, we searched for putative κB sites within the upstream region of three different transcription factors (*Bg*AP-1, *Bg*IRF, and *Bg*LITAF) that regulate gene expression for various signaling proteins. We also expanded our search to look for putative κB sites within introns of the coding regions for these genes. Based on our bioinformatics results, it is not expected that the predicted NF-κB binding site for *Bg*AP-1 might interact specifically with *Bg*RHD because the low associated score (6.88) (Table 4). In comparison, no predicted binding sites were reported upstream of *Bg*IRF. This contrasts with findings done by Lu *et al.* (2002), who characterized a functioning κB site next to the TATA box for the human *IRF7* gene. However, EMSA analyses show that *Bg*RHD bound to a κB site predicted upstream of *Bg*LITAF, which suggests that NF-κB may be able to regulate LITAF expression in *B. glabrata*. Likewise, putative NF-κB binding sites have been reported upstream of LITAF homologues in *C. gigas* as well as the Whiteleg shrimp (*Litopenaeus vannamei*) (Park *et al.* 2008; Wang *et al.*

2012). Interestingly, our functional κB site shares 60% identity with the predicted κB site (5'-GGGCATCGCA-3') reported within the promoter region of the LITAF gene in *L. vannamei* (Wang *et al.* 2012).

Despite these overall findings, no sequence specific binding interactions were reported between *Bg*RHD and three other putative κB sites reported upstream of *Bg*TLR1, *Bg*MyD88, and *Bg*IKKα. Interestingly, the κB site upstream of *Bg*TLR1 displayed a higher score and lower associated P and E values when compared to other upstream κB sites (*Bg*TLR2 and *Bg*IRAK) that were shown to bind successfully to *Bg*RHD (Table 4). However, sequence similarity with other related transcription factor binding sites does not necessarily imply that the predicted site will be functional. In fact, the κB site for *Bg*TLR1 shows 50% similarity with a mouse TLR2 κB site (5'GGGAACTTCC-3') that bound to a nuclear NF-κB protein after exposure to LPS and TNF-α (Musikacharoen *et al.* 2001). In comparison, our EMSA data for predicted κB sites reported upstream of *Bg*MyD88 and *Bg*IKKα did not show any binding with the *Bg*RHD. Bioinformatics analysis revealed low scores for both κB sites upstream of *Bg*MyD88 and *Bg*IKKα (Table 4). Although we were unable to confirm successful binding for these three predicted κB sites, other κB sites have been predicted within introns. It is possible that NF-κB may regulate the expression of these immune genes, but further EMSAs would need to be done.

4.2 Sequence alignment of kappa-binding sites in *B. glabrata*

All of the *B. glabrata* κB sites that have successfully bound to the *Bg*RHD were aligned using Vector NTI® (Fig. 16-17). The alignment includes sequences from two upstream binding sites (*Bg*p38 MAPK and *Bg*IκBα) that bound successfully to the same *B. glabrata* RHD used in the present study (Humphries & Harter, 2015). A novel, *B. glabrata* consensus sequence

(5'-GGGAATTTTCC-3') was generated by aligning all κB sites that bound successfully to the *Bg*RHD using MEME-ChIP (http://meme-suite.org/tools/meme-chip) (Fig. 18; Machanick & Bailey, 2011). This is the first demonstration of a molluscan consensus sequence for an NF-κB binding site. In all eight sequences, the consensus sequence contained a G and C residue within the second and tenth positions, respectively (Fig. 18). The 5' terminal G nucleotides and 3' terminal C residues displayed within the consensus sequences have been shown to be essential for facilitating protein-DNA interactions in other invertebrate species (Mrinal *et al.* 2011). Additionally, highly conserved T nucleotides are reported within the second half of our consensus sequence, which are characteristic of activator sequences in *Drosophila melanogaster* (Mrinal *et al.* 2011). Perhaps these κB sites are involved in initiating gene regulation. Furthermore, when comparing sequences that did not bind specifically to *Bg*RHD to those that did bind, we see two major differences (Fig. 16). We observe a G in the fourth position in two of the three sequences. Additionally, one sequence shows a T residue at both the ninth and tenth positions. Both are characteristics that were not seen in any of the probes that showed positive binding. This implies that these particular positions are important in determining binding to the *Bg*RHD.

Our consensus sequences can be compared to that of the vertebrate consensus sequence, which reads as 5'-GGGRNYYYCC-3', where R is a purine (A or G), Y is a pyrimidine (T or C), and N is any nucleotide (Chen *et al.* 1998). It is not surprising that our *B. glabrata* consensus sequences exhibits a high degree of similarity with the vertebrate consensus sequence, given that other studies have shown that the *Bg*RHD can exhibit cross reactivity with vertebrate κB sites (Jiang & Wu, 2007; Humphries & Harter, 2015).

Perhaps this new *B. glabrata* consensus sequence can be used to create a new transcription factor binding site search algorithm that is better at predicting κB sites within immune genes of the snail. This would allow for the unique opportunity to search large genomic sequences with a much higher degree of specificity. Additionally, the use of this consensus sequence could help to search for binding sites within other invertebrate species.

4.3 Challenges with the available transcription factor binding site search algorithms

In this study, the identification of putative binding sites is based on sequence similarity but the matrices used to predict these conserved sites come from vertebrate studies. In order to improve the available search engines, additional functional studies need to be done within invertebrates to characterize binding sites. Subsequently confirmed binding sites would be useful for creating a new search engine that is trained specifically within invertebrate species, such as *B. glabrata, Caenorhabditis elegans,* and *D. melanogaster.*

4.4 Further research

This research project could be expanded in several ways in order to build upon our understanding of the regulatory mechanisms for molluscan NF-κB activity. Future studies might include a focus on the prediction of κB sites downstream of the stop codon, as downstream κB sites have been characterized in vertebrate models (Xie *et al.* 1994; Rahman *et al.* 1999). I have begun preliminary bioinformatics analysis by searching 2000 bp downstream of the stop codon for several members of the NF-κB pathway, including *Bg*IκBβ, *Bg*NF-κB p105, and *Bg*NF-κB p65. Two putative κB sites for IκBβ (5'-GGGAAATACT-3' and 5'-GGGGACCCCCC-3') were reported downstream of the stop codon, but no potential binding sites were predicted for either *Bg*NF-κB p105 or *Bg*NF-κB p65.

Additionally, information from the generated *B. glabrata* consensus sequence gives us an initial insight into the determination of which nucleotides might play a role in facilitating successful binding interactions with the *Bg*RHD. In order to expand upon our knowledge of these sequence specific interactions, a series of single nucleotide substitutions could be performed to generate new oligonucleotide probes that resemble the *B. glabrata* consensus sequence. After modifying specific residues, e.g., $G_2 \rightarrow C_2$, we could use EMSAs to determine if the mutation could prevent or inhibit protein-DNA binding. Functional studies done by Mrinal *et al.* (2011) have generated a $T_6 \rightarrow A_6$ mutation in the *D. melanogaster* binding site shown upstream of IκB and found that the activator function was switched to that of a repressor. An additional study done by Muroi *et al.* (1993) indicated that a nucleotide switch at the same position showed a similar transformation from an enhancer to a repressor.

We could also use EMSAs to determine if NF-κB translocates to the nucleus following exposure to a PAMP, e.g. PGN or *S. mansoni* larval products. Preliminary work has been done via immunoblot assays and immunocytochemistry to investigate NF-κB nuclear translocation using nuclear extracts from the *B. glabrata* embryonic (Bge) cell line. If NF-κB can be induced to locate to the nucleus, nuclear fractions containing the native *Bg*NF-κB, rather than the recombinant *Bg*RHD, could be used in EMSAs.

In conclusion, the characterization of six novel κB sites upstream of *Bg*TLR2, *Bg*IRAK, *Bg*IκBβ, *Bg*NF-κB p105, *Bg*NF-κB p65, and *Bg*LITAF has helped to elucidate whether NF-κB may be able to regulate members of the TLR pathway in *B. glabrata*. Additionally, we report the discovery of the first molluscan consensus sequence for an NF-κB binding site. Furthermore, the binding of the κB sites by the *Bg*RHD in EMSAs supports that the *B. glabrata* NF-κB p65 homologue may function as a transcription factor. Based on our overall findings, we hypothesize

that NF-κB may regulate selective components of the TLR pathway. However, further work needs to be done to understand these regulatory processes in the snail, especially due to its crucial interaction with the schistosome parasite.

WORKS CITED

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.* 94(16): 8691-8696.

Adema, C.M., Hanington, P.C., Lun, C., Rosenberg, G.H., Aragon, A.D., Stout, B.A., Lennard-Richard, M.L., Gross, P.S., & Loker, E.S. (2009). Differential transcriptomic responses of *Biomphalaria glabrata* (Gastropoda, Mollusca) and metazoan parasites, *Schistosoma mansoni* and *Echinostoma paraensei* (Digenea, Platyhelminthes). *Mol Immunol.* 47(4): 849-860.

Adema, C.M., Anderson, K.V., Bayne, C.J., Belvin, M.P., Buckley, P.M., *et al.* (2016). Whole genome analysis of *Biomphalaria glabrata* (Lophotrochozoa), a snail intermediate host for schistosomiasis. *Science.* Manuscript in review.

Altschul, S.F., Gish, W., Miller, W., Meyers, E.W., & Lipman, D.J. (1990). Basic local alignment search tool. *J Mol Biol.* 215(3): 403-410.

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

Bulet, P., Stocklin, R., & Menin, L. (2004). Anti-microbial peptides: From invertebrates to vertebrates. *Immunol Rev.* 198: 169-184.

Bunje, P. (2003). "The Mollusca: Sea slugs, squid, snails, and scallops." Berkeley http://www.ucmp.berkeley.edu/taxa/inverts/mollusca/mollusca.php

Boros, D.L. (1989). Immunopathology of *Schistosoma mansoni* infection. *Clin Microbiol Rev.* 2(3): 250-269.

Brown, K., Park, S., Kanno, T., Franzoso, G., & Siebenlist, U. (1993). Mutual regulation of the transcription activator NF-kappa B and its inhibitor, I kappa B-alpha. *Proc Natl Acad Sci.* 90(6): 2532-2536.

Centers for Disease Control and Prevention. (2012). Parasites – Schistosomiasis. Retrieved from http://www.cdc.gov/parasites/schistosomiasis/

Chen, F.E., Huang, D.B., Chen, Y.G., & Ghosh, G. (1998). Crystal structure of p50/p65 heterodimer of transcription factor NF-κB bound to DNA. *Nature.* 391(6665): 410-413.

Chen, F.E. & Ghosh, G. (1999). Regulation of DNA binding by Rel/NF-κB transcription factors: Structural views. *Oncogene.* 18(49): 6845-6852.

Colley, D.C., Bustinduy, A.L., Secor, W.E., & King, C.H. (2014). Human schistosomiasis. *Lancet.* 383(9936): 2253-2264.

Cornet, V., Henry, J., Corre, E., Le Corguille, G., & Zatylny-Gaudin, C. (2015). The Toll/NF-κB pathway in cuttlefish symbiotic accessory nidamental gland. *Dev Comp Immunol.* 53(1): 42-46.

Cui, J.G., Li, Y., Zhao, Y., Bhattacharjee, S., & Lukiw, W.J. (2010). Differential regulation of interleukin-1 receptor-associated kinase-1 (IRAK-1) and IRAK-2 by microRNA-146a and NF-κB in stressed human astroglial cells and in Alzheimer disease. *J Biol Chem.* 285(50): 38951-38960.

De Martin, R., Vanhove, B., Cheng, Q., Hofer, E., Csizmadia, V., Winkler, H., & Bach, F.H. (1993). Cytokine-inducible expression in endothelial cells of an I kappa B alpha-like gene is regulated by NF kappa B. *EMBO J.* 12(7): 2773-2779.

Franzenburg, S., Fraune, S., Kunzel, S., Baines, J.F., Domazet-Loso, T., & Bosch, T.C. (2012). MyD88-deficient *Hydra* reveal ancient function of TLR signaling in sensing bacterial colonizers. *PNAS.* 109(47): 19374-19379.

Gauthier, M. & Degnan, B.M. (2008). The transcription factor NF-kappaB in the demosponge *Amphimedeon queenslandsica:* Insights on the evolutionary origin of the Rel homology domain. *Dev Genes Evol.* 218(1): 23-32.

Geertz, M. & Maerkl, S.J. (2010). Experimental strategies for studying transcription factor-DNA binding specificities. *Brief Funct Genomics.* 9(506): 362-373.

Gilmore, T.D. & Wolenski, F.S. (2012). NF-κB: Where did it come from and why? *Immunol Rev*. 246(1): 14-35.

Glinski, Z. & Jarosz, J. (1997). Molluscan immune defenses. *Arch Immunol Ther Exp (Warsz).* 45(2-3): 149-155.

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.

Goodson, M.S., Kojadinovic, M., Troll, J.V., Schetz, T.E., Casavant, T.L., Soares, M.B., & McFall-Ngai, M.J. (2005). Identifying components of the NF-κB pathway in the beneficial *Euprymna scolopes-Vibrio fischeri* organ symbiosis. *Appl Environ Microbiol.* 71(11): 6934-6946.

Gray, D.J. (2011). Diagnosis and management of schistosomiasis. *BMJ.* 342: d2651-d2662.

Gryseels, B., Polman, K., Clerinx, J., & Kestens, L. (2006). Human schistosomiasis. *Lancet.* 368(9541): 23-29.

Guillou, F., Mitta, G., Galinier, R., & Coustau, C. (2007). Identification and expression of gene transcripts generated during an anti-parasitic response in *Biomphalaria glabrata. Dev Comp Immunol.* 31(7): 657-671.

Hahn, U.K., Bender, R.C., & Bayne, C.J. (2001). Killing of *Schistosoma mansoni* sporocyts by hemocytes from resistant *Biomphalaria glabrata:* Role of reactive oxygen species. *J Parasitol.* 87(2): 292-299.

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. *PNAS.* 107(49): 21087-21092.

He, C.B., Wang, Y., Liu, W.D., Gao, X.G., Chen, P.H., Li, Y.F., & Bao, X.B. (2013). Cloning, promoter analysis and expression of the tumor necrosis-associated factor 6 (TRAF6) in Japanese scallop *(Mizuhopecten yessoensis*). *Mol Biol Res.* 40(8): 4769-4779.

Hellman, L.M. & Fried, G. (2007). Electrophoretic mobility shift assay (EMSA) for protein-nucleic acid interactions. *Nat Protoc.* 2(8): 1849-1861.

Hinkle, J.L. & Bowman, L. (2003). Neuroprotection for ischemic stroke. *J Neurosci Nurs.* 35(2): 114-118.

Humphries, J.E. & Yoshino, T.P. (2008). Regulation of hydrogen peroxide release in circulating hemocytes of the planorbid snail *Biomphalaria glabrata. Dev Comp Immunol.* 32(5): 554-562.

Humphries, J. & Harter, B. (2015). Identification of nuclear factor-kappaB (NF-κB) binding motifs in *Biomphalaria glabrata. Dev Comp Immunol.* 53(2): 366-370.

Iwanaga, S. & Lee, B.L. (2005). Recent advances in innate immunity of invertebrate animals. *J Biochem Mol Biol.* 38(2): 128-150.

Jiang, Y.S. & Wu, X.S. (2007). Characterization of a Rel\NF-κB homologue in a gastropod abalone, *Haliotis diversicolor* supertexta. *Dev Comp Immunol.* 31(2): 121-131.

Johnson, C.M. & Tapping, R.I. (2007). Microbial products stimulate human Toll-like receptor 2 expression through histone modification surrounding a proximal NF-kappaB-binding site. *J Biol Chem.* 282(43): 31197-31205.

Kawai, T. & Akira, S. (2006). TLR signaling. *Cell Death Differ.* 13: 816-825.

Kurtz, J. & Franz, K. (2003). Evidence for memory in invertebrate immunity. *Nature.* 425(6953): 37-38.

Lee, C. & Huang, C.H. (2014). LASAGNA-Search 2.0: Integrated transcription factor binding site search and visualization in a browser. *Bioinformatics.* 30(13): 1923-1925.

Lemaitre, B., Nicolas, E., Michaut, L., Reichhart, J.C., & Hoffman, J.A. (1996). The dorsoventral regulatory gene cassette *spätzle/Toll/cactus* controls the potent antifungal response in *Drosophila* adults. *Cell.* 86(6): 973-983.

Liang, Y., Zhong, Y., & Shen, P. (2004). NF-κB and its regulation on the immune system. *Cell Mol Immunol.* 1(5): 343-350.

Little, T.J., Read, A.F., & Hultmark, D. (2005). Invertebrate immunity and the limits of mechanistic immunology. *Nat Immunol.* 6(7): 651-654.

Lockyer, A.E., Emery, A.M., Kane, R.A., Walter, A.J., Mayer, C.D., Mitta, G., Coustau, C., Adema, C.M., Hanelt, B., Rollinson, D., Noble, L.R., & Jones, C.S. (2002). Early differential gene expression in haemocytes from resistant and susceptible *Biomphalaria glabrata* strains in response to *Schistosoma mansoni. PLoS ONE.* 7(12): e51102.

Lodish, H., Berk, A., Zipursky, S.L, Matsudaira, P., Baltimore, D., & Darnell, J. (2000). *Molecular Cell Biology* (4th ed.). New York, NY: W.H. Freeman and Company.

Loker, E.S., Adema, C.M., Zhang, S., & Kepler, T.B. (2004). Invertebrate immune systems – not homogenous, not simple, not well understood. *Immunol Rev.* 198: 10-24.

Lu, R., Moore, P.A., & Pitha, P.M. (2002). Stimulation of IRF-7 gene expression by tumor necrosis factor alpha: Requirement for NF-kappaB transcription factor and gene accessibility. *J Biol Chem.* 277(19): 16592-16598.

Machanick, P. and Bailey, T.L. (2011). MEME-ChIP: Motif analysis of large DNA datasets. *Bioinformatics*, 27**(**12): 1696-1697.

Matricon-Gondran, M. & Letocart, M. (1999). Internal defenses of the snail *Biomphalaria glabrata. J Invertebr Pathol.* 74(3): 235-247.

Mitta, G., Galinier, R., Tisseyre, P., Allienne, J.F., Girerd-Chambraz, Y., Guillou, F., Bouchut, A., & Coustau, C. (2005). Gene discovery and expression analysis of immune-relevant genes from *Biomphalaria glabrata* hemocytes. *Dev Comp Immunol.* 29(5): 393-407.

Montagnani, C., Kappler, C., Reichhart, J.M., & Escoubas, J.M. (2004). Cg-Rel, the first Rel/NF-kappaB homolog characterized in a mollusk, the Pacific oyster *Crassostrea gigas*. *FEBS Letters*. 561:75-82.

Montgomery, S.B., Griffith, O.L., Sleumer, M.C., Bergman, C.M., Bilenky, M., Pleasance, E.D., Prychyna, Y., Zhang, X., & Jones, S.J.M. (2005). ORegAnno: An open access database and curation system for literature-derived promoters, transcription factor binding sites, and regulatory variation. *Bioinformatics.* 22(5): 637-640.

Moorthy, A.K., Savinova, O.V., Ho, J.Q., Wang, V.Y., Vu, D., & Ghosh, G. (2006). The 20S proteasome processes NF-κB1 p105 into p50 in a translational-independent manner. *EMBO J.* 25(9): 1945-1956.

Moynagh, P.N. (2005). The NF-κB pathway. *J Cell Sci.* 118: 4389-4392.

Mrinal, N., Tomar, A., & Nagaruju, J. (2011). Role of sequence encoded κB DNA geometry in gene regulation by Dorsal. *Nuc Acids Res.* 39(22): 9574-9591.

Muroi, M., Muroi, Y., Yamamoto, K., & Suzuki, T. (1993). Influence of 3' half-site sequence of NF-κB motifs on the binding of lipopolysaccharide-activatable macrophage NF-κB proteins. *J Biol Chem.* 268(26): 19534-19539.

Musikacharoen, T., Matsuguchi, T., Kikuchi, T., & Yoshikai, Y. (2001). NF-κB and STAT5 play important roles in the regulation of mouse Toll-like receptor 2 gene expression. *J Immunol.* 166(7): 4516-4524.

Nowak, T.S., Woodards, A.C., Jung, Y., Adema, C.M., & Loker, E.S. (2004). Identification of transcripts generated during the response of resistant *Biomphalaria glabrata* to *Schistosoma mansoni* infection using suppression subtractive hybridization. *J Parasitol.* 90(5): 1034-1040.

Nyholm, S.V. & Graf, J. (2012). Knowing your friends: Invertebrate innate immunity fosters beneficial bacterial symbioses. *Nat Rev Microbiol.* 10(12): 815-827.

Owen, J., Pundt, J. & Stranford, S. (2013). *Kuby Immunology* (7th ed.). New York, NY: W.H. Freeman.

Park, E., Kim, Y., Nam, B., Kong, H., Kim, W., Lee, S., Kong, I., & Choi, T. (2008). Cloning, characterization and expression analysis of the gene for a putative lipopolysaccharide-induced TNF-α factor of the Pacific oyster, *Crassostrea gigas. Fish Shellfish Immunol.* 24(1): 11-17.

Putnam, N.H., Srivastava, M., Hellsten, U., Dirks, B., Chapman, J., Salamov, A., Terry, A., Shapiro, H., Lindquist, E., Kapitonov, V.V., Jurka, J., Gentikhovich, G., Griqoriev, I.V., Lucas, S.M., Steele, R.E., Finnerty, J.R., Technau, U., Martindale, M.Q., & Rokhsar, D.S. (2007). Sea anemone genome reveals ancestral eumetazoan gene repertoire and genomic organization. *Science.* 317(5834): 86-94.

Rahman, A., Khandaker, N., Anwar, N., True, A.L., & Malik, A.B. (1999). Thrombin-induced p65 homodimer binding to downstream NF-κB site of the promoter mediates endothelial ICAM-1 expression and neutrophil adhesion. *J Immunol.* 162(9): 5466-5476.

Rast, J.P. & Messier-Solek, C. (2008). Marine invertebrate genome sequences and our evolving understanding of animal immunity. *Biol Bull.* 214(3): 274-283.

Roberts, L., Janovy, J., & Nadler, S. (2012). *Foundations of Parasitology* (9th ed.). New York, NY: McGraw-Hill Education.

Rowe, S.E. & O'Gara, J.P. (2016). Electrophoretic mobility shift assays. *Methods Mol Biol.* 1373: 155-167.

Rowley, A.F. & Powell, A. (2007). Invertebrate immune systems-specific, quasi-specific, or nonspecific? *J Immunol*. 179(11): 7209-7214.

Sadd, B.M. & Schmid-Hempel, P. (2006). Insect immunity shows specificity in protection upon secondary pathogen exposure. 16(12): 1206-1210.

Scott, M.L., Fujita, T., Liou, H., Nolan, G.P., & Baltimore, D. (1993). The p65 subunit of NF-κB regulates IκB by two distinct mechanisms. *Genes Dev.* 7: 1266-1276.

Song, L.S., Wang, L.L., Qiu, L.M., & Zhang, H. (2010). Bivalve Immunity. In K. Söderhäll (Ed.), *Invertebrate Immununity* (pp. 44-65). New York, NY: Springer.

Sun, S.C., Ganchi, P.A., Ballard, D.W., & Greene, W.C. (1993). NF-kappa B controls expression of inhibitor I kappa B alpha: Evidence for an inducible autoregulatory pathway. *Science.* 259(5103): 1912-1915.

Takeshita, F., Suzuki, K., Sasaki, S., Ishii, N., Klinman, D.M., & Ishii, K.J. (2004). Transcriptional regulation of the human TLR9 gene. *J Immunol.* 173(4): 2552-2561.

Thompson, J.E., Phillips, R.J., Erdjument-Bromage, H., Tempst, P., & Ghosh, S. (1995). IκB-β regulates the persistent response in a biphasic activation of NF-κB. *Cell.* 80(4): 573-582.

Toubiana, M., Gerdol, M., Rosani, U., Pallavicini, A., Venier, P., & Roch, P. (2013). Toll-like receptors and MyD88 adaptors in *Mytilus:* Complete cds and gene expression levels. *Dev Comp Immunol.* 40(2): 158-166.

Valenzuela-Munoz, V., Bueno-Ibarra, M.A., & Gallardo-Escarate, C. (2014). Characterization of the transcriptomes of *Haliotis rufescens* reproductive tissues. *Aquaculture Res.* 45(6): 1026-1040.

Wan, F., Anderson, D.E., Barnitz, R.A., Snow, A., Bidere, N., Zheng, L., Hedge, V., Lam, L.T., Staudt, L.M., Levens, D., Deutsch, W.A., & Lenardo, M.J. (2007). Ribosomal protein S3: A KH domain subunit in NF-kappaB complexes that mediates selective gene regulation. *Cell.* 131(5): 927-939.

Wang, T., Niu, G., Kortylewski, M., Burdelya, L., Shain, K., Zhang, S., Bhattacharya, R., Gabrilovich, D., Heller, R., Coppola, D., Dalton, W., Jove, R., Pardoll, D., & Yu, H. (2004). Regulation of the innate and adaptive immune responses by Stat-3 signaling in tumor cells. *Nat Med.* 10(1): 48-54.

Wang, P.H., Wan, D.H., Pang, L.R., Gu, Z.H., Qiu, W., Weng, S.P., Yu, X.Q., & He, J.G. (2012). Molecular cloning, characterization and expression analysis of the tumor necrosis factor (TNF) superfamily gene, TNF receptor superfamily gene and lipopolysaccharide-induced TNF-α factor (LITAF) gene from *Litopenaeus vannamei*. *Dev Comp Immunol.* 36(1): 39-50.

Wen, A.Y., Sakamoto, K.M., & Miller, L.S. (2010). The role of the transcription factor CREB in immune function. *J Immunol.* 185(11): 6413-6419.

Wingender, E., Chen, X., Hehl, R., Karas, H., Liebich, I., Matys, V., Meinhardt, T., Prub, M., Reuter, I., & Schacherer, F. (2000). TRANSFAC: An integrated system for gene expression regulation. *Nuc Acids Res.* 28(1): 316-319.

World Health Organization. (1993). Report of the WHO Expert Committee: The control of Schistosomiasis. Geneva: WHO. Ser 830. p. 86.

World Health Organization. (2015). Schistosomiasis, Fact Sheet No 115. Retrieved from http://www.who.int/mediacentre/factsheets/fs115/en/

Xie, Q., Kashiwabara, Y., & Nathan, C. (1994). Role of transcription factor NF-κB/Rel in induction of nitric oxide synthase. *J Biol Chem.* 269(7): 4705-4708.

Xiong, K., Feng, Q., Chen, L., Xie, L., & Zhang, R. (2008). Cloning and characterization of an IKK homologue from the pearl oyster, *Pinctada fucata. Dev Comp Immunol.* 32(1): 15-25.

Yoshino, T.P. & Cousteau, C. (2011). Immunobiology of *Biomphalaria* – Trematode Interactions. In R. Toledo and B. Fried (Eds.). *Biomphalaria Snails and Larval Trematodes* (pp. 159-190). New York, NY: Springer.

Zhang, G. & Ghosh, S. (2001). Toll-like receptor-mediated NF-kappaB activation: A phylogenetically conserved paradigm in innate immunity. *J Clin Invest.* 107(1): 13-19.

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 Immunol.* 14(3): 175-189.

Zhang, S. & Coultas, K.A. (2011). Identification and characterization of five transcription factors that are associated with evolutionarily conserved immune signaling pathways in the schistosome-transmitting snail *Biomphalaria glabrata. Mol Immunol.* 48(15-16). 1868-1881.

APPENDIX

Appendix I. Bioinformatics analyses for *Bg*p38 MAPK and *Bg*IκBα.

Table 6. Identification of putative κB sites within 2000 bp upstream of protein coding regions for *Bg*IκBα and *Bg*p38 MAPK. Length-Alignment Site Guided by Nucleotide Association 2.0 (LASAGNA 2.0) data highlighted in blue represents genes and their respective sequences that were selected for further analyses in Humphries and Harter's (2015) study. Sequences in black represent the additional putative κB sites identified in the current study.

Control Region	Sequence	Number of κB sites
***Bg*p38 MAPK**	cggattttcc tggaatttct gggaggttcc tcggatttcc	4
***Bg*IκBα**	cggaaatccc gggcctttcc ggaactttct gaaatcccct	4

Appendix II. Troubleshooting guide

In principle, an EMSA is a very simple and rapid tool used to test transcription factor functionality. However, this assay may require much optimization and a variety of parameters typically need to be considered in order to obtain clear gel shifts. More common experimental factors can include buffer pH, buffer composition, salinity, and percentage acrylamide (Hellman & Fried, 2007). In addition, DNA-protein binding can be disrupted by improper preparation of binding reactions. Poly I:C (nonspecific blocker), KCl, 50% glycerol, magnesium chloride ($MgCl_2$), and NP-40 (detergent) can all be added in order to optimize the binding reaction. In addition, the amount of added protein has the potential to disrupt the binding interactions. An excess amount of protein can cause aggregate formation, causing the nucleic acid to become stuck in the gel. A troubleshooting guide has been assembled based on my experiences in order to help improve optimization and reduce issues that were mentioned above.

Table 7. Electrophoretic mobility shift assay troubleshooting guide. This information was compiled based on my own experiences troubleshooting the EMSA protocol.

Problem	**Possible cause**	**Possible solution**
Bands are smeared or streaked	Uneven gel polymerization	Use fresh gel components.
	Wrong buffer	Switch to TAE buffer and lower concentration (0.25 X to 0.5 X)
	Buffer went off	Prepare fresh buffer.
	Overheating occurred	Reduce voltage during transfer or add ice pack.
	Protein degraded	Use fresh protein
No shift appears	Poor transfer	Increase transfer time
	Binding buffer went off	Prepare fresh binding buffer
	Protein degraded	Use fresh protein.
	pH is not optimal	Increase buffer pH to promote protein-DNA binding
	Binding interaction was disrupted	Increase incubation time Incubate on a rocker
No visible bands	Streptavidin degraded	Perform "dark room test" and add 1 µL SA-HRP to solution containing equal parts of ECL reagents.
	Inefficient transfer	Increase transfer time.
	Samples cannot move from gel to membrane due to small pore size	Decrease amount of acrylamide in gel.
	Biotin degraded	Increase volume of DNA. Re-label probes
Bubbles or wrinkles in gel	Excessive heating.	Decrease voltage during transfer or add an ice pack during transfer.
Dark bands on membrane.	Too much HRP in the system.	Decrease amount of SA-HRP added to blocking buffer during detection.