2023

Methods for Chromatin Immunoprecipitation of Biomphalaria glabrata NF-κB

Marissa A. Zintel

Methods for Chromatin Immunoprecipitation of *Biomphalaria glabrata* NF-κB

Marissa Zintel
B.A. Biology, Biochemistry
Lawrence University, ‘23
May 1, 2023
Advisor: Judith Humphries, Ph.D.

A Thesis Submitted in Candidacy for Honors at Graduation

I hereby reaffirm the Lawrence University Honor Code.

Table of Contents

Abstract

Biomphalaria glabrata is a species of freshwater snail that is an intermediate host to the parasite *Schistosoma mansoni*, a causative agent of human schistosomiasis. Substantial research has focused on the *B. glabrata* immune system, but progress is restricted by the limited molecular tools available for molluscan organisms. While many studies have described the anti-parasitic defenses of *B. glabrata*, the initiation and regulation of these defenses remain unclear. One important immune regulator, well known in both human and *Drosophila* immunology, is nuclear factor kappa B (NF-κB). Research has shown that *B. glabrata* NF-κB transcripts are upregulated in response to *S. mansoni* infection, but few functional assays have been performed to describe the activity of *B. glabrata* NF-κBs. The present study addresses this knowledge gap by developing a Chromatin Immunoprecipitation (ChIP) assay to determine the binding sequences, gene targets, and relative activity of *B. glabrata* NF-κBs. The results of this assay will further our understanding of *B. glabrata* immunity, which may then be exploited to disrupt the *S. mansoni* life cycle and may provide insight into the evolution of NF-κB binding sequences. Additionally, this ChIP procedure may be easily adapted to suit other *B. glabrata* transcription factors, providing a valuable tool for future *B. glabrata* studies.

Introduction

Biomphalaria glabrata

Biomphalaria glabrata is a species of freshwater snail known for its role as an intermediate host to the parasite *Schistosoma mansoni*. Because *S. mansoni* is a causative agent of human schistosomiasis, *B. glabrata* has been the subject of significant research aiming to understand the parasite-host dynamics within the snail. Additionally, *B. glabrata* is a member of the under-studied superphylum Lophotrochozoa, and thus all *B. glabrata* studies contribute towards a more complete understanding of this significant invertebrate superphylum.

Schistosomiasis

Human schistosomiasis is a neglected tropical disease caused primarily by six *Schistosoma* species that infect over 200 million people in tropical and subtropical regions of sub-Saharan Africa, Latin America, the Middle East, and Asia ("CDC - Schistosomiasis," 2021). Nearly 90% of cases occur in Africa, where *S. mansoni* is one of two dominant *Schistosoma* species (Montgomery, 2019; "Schistosomiasis," 2023). *S. mansoni* is estimated to account for 54 million intestinal schistosomiasis infections (Aula et al., 2021), primarily in Africa but also in Latin American, Caribbean, and Middle Eastern countries ("Schistosomiasis," 2023).

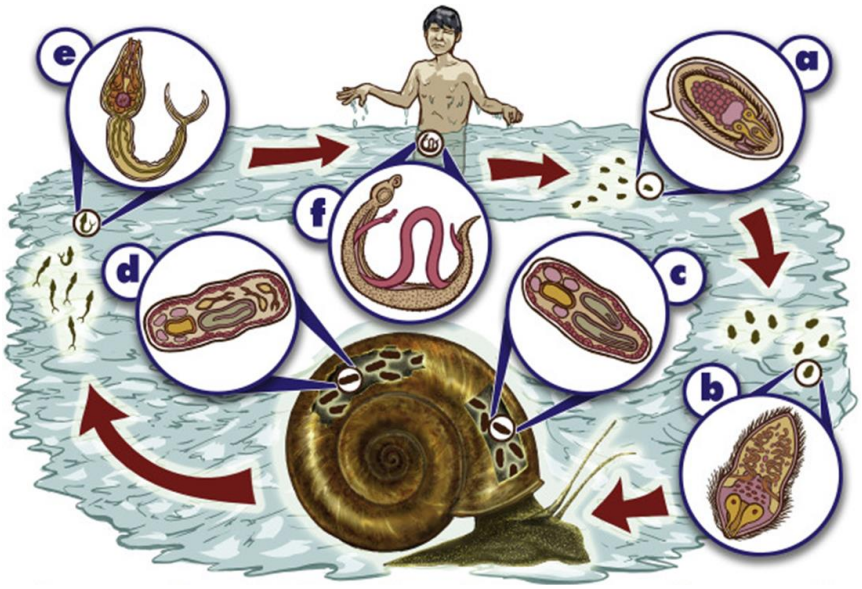


Figure 1. The life cycle of *S. mansoni*, causative agent of schistosomiasis. (a) Egg, (b) miracidium, (c) sporocyst, (d) daughter sporocyst, (e) cercariae, (f) adult schistosomes. Adapted from Castillo et al., 2020.

The life cycle of *S. mansoni* requires both humans and snails, with transmission between both hosts occurring through water sources. When *S. mansoni* eggs (Fig. 1a) are deposited into fresh water sources, miracidia (Fig. 1b), the parasitic form infective to snails, hatch from these eggs. Snail hosts are penetrated by the miracidia, and within the snail the miracidia transform into sporocysts (Fig. 1c), which undergo asexual reproduction to sequentially produce daughter sporocysts (Fig. 1d) and cercariae (Fig. 1e). An infected snail can shed 250-600 cercariae per day, and these cercariae are infectious to humans. Cercariae swim through water and can infect humans through exposed skin. Inside the human body, cercariae become schistosomula, which move through the bloodstream to the liver, where they mature into adult worms (Fig. 1f). Adult worms generally reside in the mesenteric veins near either the small or large intestine ("CDC - Schistosomiasis," 2021) where they undergo sexual reproduction to produce eggs. The eggs pass through the intestines and are eliminated with human feces. When infected fecal matter contaminates a freshwater source, the eggs hatch into miracidia to begin the cycle again in a snail host (Nelwan, 2019).

The symptoms of schistosomiasis are primarily due to the body's immune response to *S. mansoni* eggs. In experimental settings, only 20-55% of eggs were successfully passed by the host, with the remainder being left to circulate through the blood until becoming lodged in body tissues as distant as the eyes or central nervous system (Costain et al., 2018), although eggs most commonly accumulate in the liver. The immune system attacks these foreign eggs by encapsulating them in granulomas (Fig. 2), regions of intense immune activity which often leave the surrounding tissue scarred (Costain et al., 2018). Acute schistosomiasis is mainly caused by this intense immune activity and can result in fever, chills, cough, and muscle aches ("CDC - Schistosomiasis," 2021). Chronic schistosomiasis results from prolonged infection, and can lead to abdominal pain, enlarged liver, blood in the stool, and liver scarring ("CDC - Schistosomiasis," 2021). In

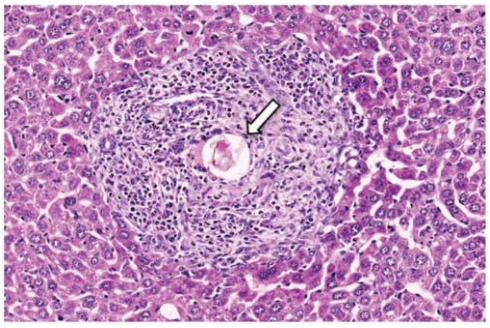


Figure 2. Granuloma surrounding *S. mansoni* egg.
Stained via H&E. From Hams et al., 2013.

rare cases, eggs deposited in the spinal cord can lead to seizures or paralysis ("CDC - Schistosomiasis," 2021).

Left untreated, the egg-producing parasites can remain in the body for years, producing hundreds to thousands of eggs per day (Colley et al., 2014). Currently, the dominant treatment for schistosomiasis is praziquantel, a drug active against a variety of parasitic flatworms. The World Health Organization-recommended schistosomiasis control strategy is to mass administer praziquantel to school-age children every 1-3 years depending on the area's infection prevalence (Deol et al., 2019). However, there is concern that *Schistosoma* species might develop resistance to praziquantel (Vale et al., 2017), and thus there is a need to develop alternative strategies that disrupt the parasite's life cycle. The intermediate snail host may be the key to such a disruption, as one study found that anti-schistosomiasis strategies utilizing snail control reduced schistosomiasis prevalence by 92%, compared to only a 37% reduction without snail control (Sokolow et al., 2016). Snail control involves reducing the population of intermediate snail hosts, which can be achieved using either molluscicides or by increasing the population of parasite-resistant snails (Sokolow et al., 2016). The use of molluscicides is undesirable due to their potential to disrupt the food chain and their toxicity against other organisms, such as fish (Pila, 2018). Reducing the susceptibility of intermediate snail hosts to *S. mansoni* is a more attractive target, as this method preserves the ecosystem as much as possible. Informed by immunological studies, this may be achieved by modifying *Biomphalaria* populations to render snails resistant to infection, thus interrupting the *S. mansoni* life cycle (Pila, 2018). Genetic manipulation strategies leveraging immunological data are being pursued in other parasitic disease vectors, such as mosquitoes transmitting *Plasmodium* parasites (Pila, 2018).

Model Organism for Lophotrochozoa

In addition to the schistosomiasis-related benefits of exploring the *B. glabrata* immune system, studies of *B. glabrata* represent opportunities to fill the gap in scientific understanding of the superphylum Lophotrochozoa. Lophotrochozoa are one of three superphyla of organisms displaying bilateral symmetry, along with Ecdysozoa and Deuterostomia (Fig. 3). Ecdysozoa include well-known model organisms such as *Drosophila melanogaster* and *Caenorhabditis elegans*, while Deuterostomia includes mice and humans, arguably the most well-studied organisms. Well-established model organisms are comparatively lacking within Lophotrochozoa, and thus all new knowledge from studies of *B. glabrata* can contribute towards this model organism gap. Increased knowledge of Lophotrochozoa may also facilitate a deeper evolutionary understanding of the divergence of bilateral organisms through comparisons of novel Lophotrochozoa findings to published Ecdysozoa and Deuterostomia studies (Tessmar-Raible, 2003).

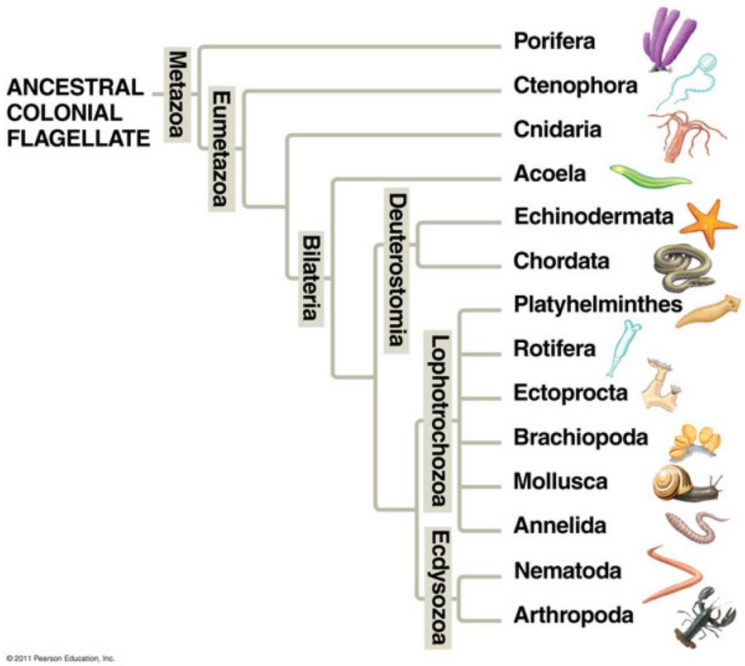


Figure 3. Simplified phylogeny of animals. From Pearson Education, 2009.

Significant research focusing on *B. glabrata* has already been completed to advance these goals (albeit primarily within the immune system), and many research tools are now available for *B. glabrata* that are rare among Lophotrochozoa organisms. For instance, the genome of *B. glabrata* is fully sequenced, providing an accessible foundation for bioinformatic analyses. Many bioinformatics projects have already identified gene homologs for further study (Castillo et al., 2020). Additionally, *B. glabrata* are the only members of Lophotrochozoa for which there is an immortal cell line (Wheeler et al., 2018). The Bge cell line was developed in 1976 from *B. glabrata* embryonic cells (Maramorosch, 2012), and has since facilitated many valuable *in vitro* studies. Bge cells display anti-parasitic behavior similar to *B.*

glabrata immune cells and have therefore been used to study immune-related cellular interactions between *S. mansoni* and *B. glabrata*. Bge cells can also be used to cultivate *S. mansoni in vitro*, allowing the parasite to be studied during its transformation from miracidia to sporocyst to cercariae. Furthermore, immortal cell lines allow researchers to perform assays that might otherwise be limited by source material, such as gene editing experiments (Coelho et al., 2020; Rinaldi et al., 2015) or the Chromatin immunoprecipitation (ChIP) experiments of this study.

B. glabrata Immune Regulation

Immune System of *B. glabrata*

Although less is known about *B. glabrata* immunity compared to human immunity, decades of significant research have revealed many functional elements of the *B. glabrata* immune system. *B. glabrata* is considered to have an innate (not adaptive) immune system, which is traditionally characterized by general defense systems against many types of pathogens. Notable innate immune molecules identified in *B. glabrata* include pattern recognition receptors (PRRs) to recognize pathogen-associated molecular patterns (PAMPs) and humoral defenses such as antimicrobial peptides, lectins, and complement-related proteins (Castillo et al., 2020). A notable class of PRRs includes fibrinogen-related proteins (FREPs), which are analogous to antibodies in their ability to undergo somatic mutation to develop reactivity against a broad range of non-self motifs. Additionally, the snail possesses immune cells called hemocytes, which perform critical functions such as phagocytosis (Fryer and Bayne, 1990), encapsulation of pathogens such as *S. mansoni* sporocysts, and the production of toxic compounds (Adema, 2015).

These innate immune features are often contrasted with the adaptive immune system, which is possessed solely by vertebrates and can develop immunological memory of a specific pathogen. Vertebrate adaptive immunity involves T cells, B cells, and antibodies that can facilitate an enhanced immune response during a secondary infection. Although *B. glabrata*, along with all invertebrates, lacks

this adaptive immune system, the snail is nevertheless capable of mounting a more successful immune response during a repeat infection with *S. mansoni* (Pinaud et al., 2016). This phenomenon is known as innate immune memory and has been demonstrated in both vertebrates and invertebrates (Gourbal et al., 2018; Netea et al., 2011).

Defense Against *S. mansoni*

Through an unknown mechanism, *B. glabrata* displays innate immune memory in response to multiple *S. mansoni* infections by deploying a different, more effective immune response during a secondary infection. The *B. glabrata* immune system responds to a primary infection with *S. mansoni* with a cell-mediated response, encapsulating *S. mansoni* sporocysts with hemocytes and releasing toxic compounds such as reactive oxygen species (ROS). However, a secondary *S. mansoni* infection results in no encapsulation, with *B. glabrata* instead employing a humoral response, principally involving the production of biomphalysin, a pore-forming protein extremely toxic to *S. mansoni* (Pinaud et al., 2016). Pinaud et al. found these secondary responses to be 100% effective in snails susceptible to infection, while primary responses were only 60-70% effective (Pinaud et al., 2016). Yet the mechanisms of this memory remain unknown. Research in vertebrates has shown that innate immune memory can be encoded in the epigenome by immune transcription factors, but this mechanism has not been validated for mollusks (Gourbal et al., 2018). Indeed, while researchers have compiled significant data cataloging *B. glabrata* immune responses, including specific gene expression changes during different stages of *S. mansoni* infection (Dinguirard et al., 2018), the signaling pathways and transcription factors responsible for initiating and regulating these immune responses are significantly under-studied.

Immune Signaling Pathways

Signaling pathways and the transcription factors (TFs) they activate are crucial components of successful immune responses, allowing cells to respond to a single pathogenic signal with a complex defense involving hundreds of genes. TFs are proteins which bind to a specific regulatory DNA sequence

near a gene, thus impacting the expression of this gene. TFs can increase expression by recruiting complexes for RNA polymerase activity, decrease expression by blocking such factors, or interact with chromatin remodeling proteins to either increase or decrease gene accessibility (and transcription) (Lambert et al., 2018). To produce an immune response, TFs are required to activate various defensive genes and temporarily deactivate nonessential cellular functions. Immune TFs are typically activated by a signaling pathway which begins with the recognition of a pathogenic molecule and ends with nuclear translocation of the TF to perform its function (Medzhitov and Janeway, 2000).

Many immune transcription factors well-known in vertebrates have been identified in *B. glabrata*, and transcript expression studies have suggested immune roles for STAT1, STAT2, NF-κB p65, and NF-κB p50 homologs (Zhang and Coultas, 2011). However, none of these have yet been demonstrated to play a functional role in *B. glabrata* immune responses. Toll-like receptors (TLRs) are evolutionarily conserved initiators of many immune pathways, and TLRs have been shown to play a role in the *B. glabrata* immune system (Pila et al., 2016). Additionally, some signaling pathways that can be activated downstream of TLRs have been demonstrated to play a role in *B. glabrata* immunity, such as the MAPK, ERK, and PI3K pathways (Humphries et al., 2001; Zelck et al., 2007; Humphries and Yoshino, 2008). Notably, NF-κB proteins are also often activated downstream of TLR signaling (Hayden et al., 2006), but the function of *B. glabrata* NF-κB proteins remains unclear. Further studies of the regulatory pathways in the *B. glabrata* immune system can help both to provide an evolutionary perspective to the development of Bilateria immune systems, and to identify immune factors that might be of use in combating *S. mansoni* infections of *B. glabrata*.

NF-κB

Nuclear factor kappa B (NF-κB) is a family of transcription factor proteins that affects the expression of hundreds of genes, thereby impacting processes such as the inflammatory response, apoptosis, and development, among others. NF-κB was named for its originally observed role: binding to

the enhancer of the kappa light chain gene in vertebrate B cells (Sen and Baltimore, 1986). The kappa light chain is a component of antibodies, and thus this discovery immediately indicated the important role NF-κB proteins play in the immune system. NF-κB has since become most well-known for its role in regulating innate immunity.

Structure

All NF-κB proteins are identifiable by a characteristic amino acid sequence known as the Rel Homology Domain (RHD), an evolutionarily conserved motif shared by all NF-κB proteins. The RHD is found at the N-terminal end of NF-κB proteins and contains motifs essential for the function of NF-κBs as transcription factors: a nuclear localization sequence and DNA-binding domain that recognizes κB sequences. The RHD also contains a dimerization domain, through which all NF-κB proteins function as either homodimers or heterodimers (Williams and Gilmore, 2020).

The NF-κB family can be categorized into two subfamilies ("NF-κB" or "Rel") on the basis of each protein's C-terminal motifs (Fig. 4). NF-κB proteins contain ankyrin (ANK) repeat domains and a death domain (DD), while Rel proteins contain a transactivation domain (TAD) (Oeckinghaus and Ghosh, 2009). The C-terminal end of NF-κB proteins acts as an inhibitor, with ANK repeats binding to NF-κB dimers to inhibit them (Hatada et al., 1992). Less is known about NF-κB death domains, but they have been shown to act as docking sites to facilitate degradation of the inhibitory protein (Beinke et al., 2002). Meanwhile, Rel transactivation domains function to recruit transcriptional proteins for increased expression of Rel-targeted genes.

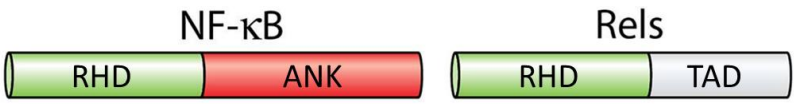


Figure 4. Schematic of Rel and NF-κB protein structures. Adapted from Williams & Gilmore, 2020.

Evolutionary History

While NF-κB proteins have been most extensively studied in humans, homologs have been identified in organisms as evolutionarily distant as unicellular choanoflagellates and the protist

Capsaspora owczarzaki, whose most recent common ancestor (MRCA) shared with humans existed approximately 1,000-600 million years ago. Recent research suggests that these first NF-κB proteins originated due to the genetic fusion of the RHD and ANK domains, formerly two separate proteins. Such RHD-ANK fusion proteins (NF-κB subfamily) later experienced gene duplication events, so that by the divergence of Lophotrochozoa, Ecdysozoa, and Deuterostomia about 550 million years ago, the RHD-TAD fusions (Rel subfamily) had appeared (Williams and Gilmore, 2020). While NF-κB genes have been identified and compared across each of these superphyla, fewer studies have specifically described NF-κB binding sites. Notably, binding site data is available for both Deuterostomia (from humans) and Ecdysozoa (from *Drosophila*) (Siggers et al., 2015), but very few binding sites are known for Lophotrochozoa organisms (Humphries and Harter, 2015; Humphries and Deneckere, 2018). Identifying the binding sites of *B. glabrata* NF-κB proteins will allow comparisons among these superphyla that may illuminate the evolution of NF-κB proteins' binding specificity.

NF-κB in Mammals

The NF-κB family in mammals includes five transcription factors: the NF-κB family of p105/p50 and p100/p52, and the Rel family of RelA (p65), RelB, and c-Rel. Like all known NF-κB proteins, under normal conditions all are found as dimers in the cytoplasm of a cell, bound by some type of inhibitor. The canonical pathway of NF-κB activation in humans proceeds

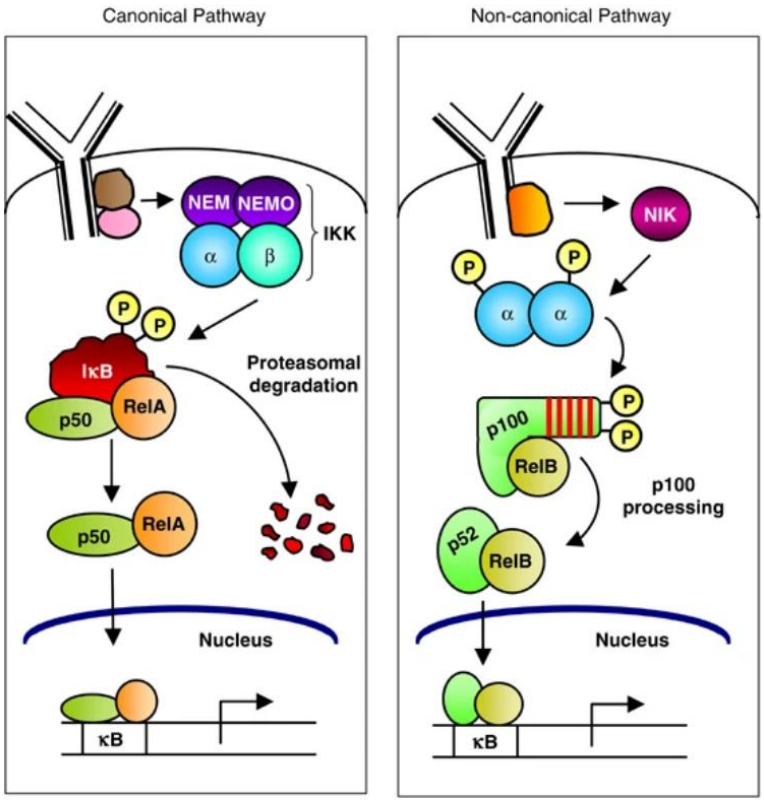


Figure 5. Canonical and Non-canonical NF-κB pathways in mammals. Adapted from Gilmore, 2006.

when IκB (inhibitor of kappa B) is degraded, freeing an NF-κB dimer to translocate to the nucleus. The non-canonical pathway involves the degradation of the ANK domain of p100, resulting in its activated p52 form, which translocates to the nucleus (Fig. 5). Although these are the most well-characterized pathways, there are additional variations (such as the degradation of p105 to p50), which are activated by different cellular signals (Gilmore, 2006). All human NF-κB proteins have been demonstrated to form both homodimers and heterodimers with all other human NF-κB proteins (although not always *in vivo* (Gilmore, 2006)), with each distinct dimer displaying slightly different binding preferences, and likely being activated by different stimuli to perform different functions (Ghosh et al., 2012).

Indeed, NF-κB has been shown to function in numerous contexts in humans, including both the innate and adaptive immune systems, the nervous system, and in some cancers. NF-κB was first named for its role in the adaptive immune system, where it binds a regulatory sequence necessary to transcribe a component of antibodies (Sen and Baltimore, 1986). Further studies on NF-κB in adaptive immunity have distinguished NF-κB signaling as important for survival and specialization of T and B cells (Hayden et al., 2006), with NF-κB signaling being critical for naïve T and B cells' development after their first contact with their antigen. Innate immune functions of NF-κB are generally initiated when PAMPs from bacteria, fungi, or viruses bind host cell PRRs. Such NF-κB activation leads to immediate anti-pathogen responses such as the release of defensins (to permeabilize bacterial membranes) and reactive oxygen and nitrogen species (toxic to many microbes) (Hayden et al., 2006). NF-κB is also critical to the inflammatory response by driving cytokine and chemokine production to recruit effector cells (such as neutrophils) to an infection site. Beyond immunity, NF-κB is highly expressed in the nervous system, and may play important roles in neurogenesis, brain injuries, neurodegenerative models, and neural signaling (O'Neill and Kaltschmidt, 1997; Shih et al., 2015). Additionally, over-active NF-κB has been identified in more than 40 cancer types, where it contributes towards cancer cell survival and proliferation, inhibition of immune cells, and metastasis, among other roles (Gilmore, 2021).

NF-κB in *Drosophila*

The most well-studied group of NF-κB proteins in invertebrates are those found in *Drosophila*, which include three NF-κB family members: Dorsal and DIF (Rel family or "Class I") and Relish (NF-κB family or "Class II"). *Drosophila* also expresses an IκB protein known as Cactus. Two NF-κB immune pathways have been characterized in *Drosophila*, which are very similar to their mammalian counterparts. The Toll pathway is analogous to the mammalian canonical pathway, in which Dorsal and/or DIF dimers are activated upon the degradation of Cactus. The immune deficiency (IMD) pathway is analogous to the mammalian non-canonical pathway, with the cleavage of the ANK domain of Relish resulting in the active Rel-N form that translocates to the nucleus (Fig. 6). Both the Toll and IMD pathways are activated by PAMPs from bacteria or fungi binding *Drosophila* PRRs to signal an immune threat (Minakhina and Steward, 2006).

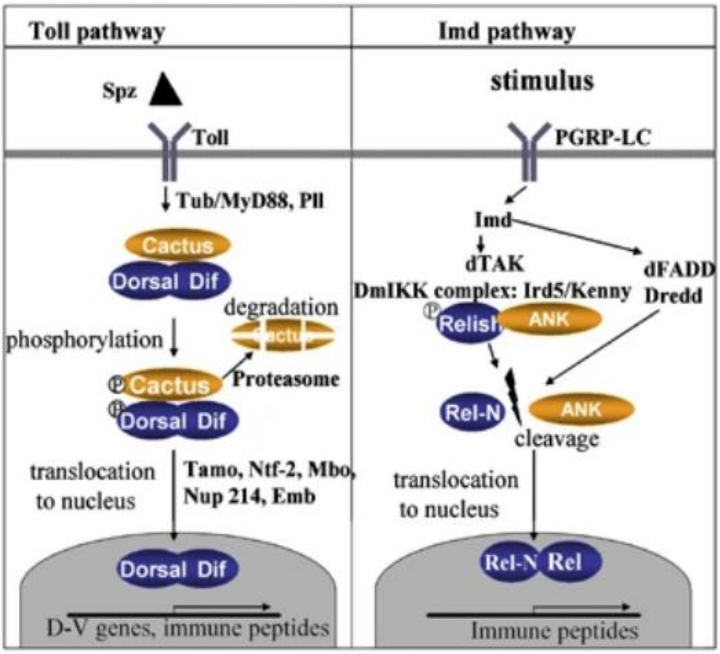


Figure 6. Toll and IMD pathways in *Drosophila*. Adapted from Minakhina & Steward, 2006.

Interestingly, the Dorsal *Drosophila* protein was first discovered and named due to its function in creating the dorsal-ventral polarity of developing embryos. Embryos rely on a gradient of the Dorsal protein in the nuclei of outer embryonic cells, so that cells with high levels of Dorsal become ventral cells, and cells without Dorsal become dorsal cells (Minakhina and Steward, 2006). Additional non-immune roles in *Drosophila* have been observed, including the regulation of neurogenesis in sensory organ precursors (Ayyar et al., 2007). The prevalence of non-immune roles of NF-κB suggests that additional surprising functions might be found upon investigation of *B. glabrata* NF-κB proteins.

NF-κB in *B. glabrata*

Two NF-κB proteins have been identified in *B. glabrata* (Zhang and Coultas, 2011), termed BgRel (Rel family, class I) and BgRelish (NF-κB family, class II). Both contain an RHD with a DNA-binding motif and a nuclear localization signal. BgRelish also contains six ANK repeats and a death domain (Fig. 7).

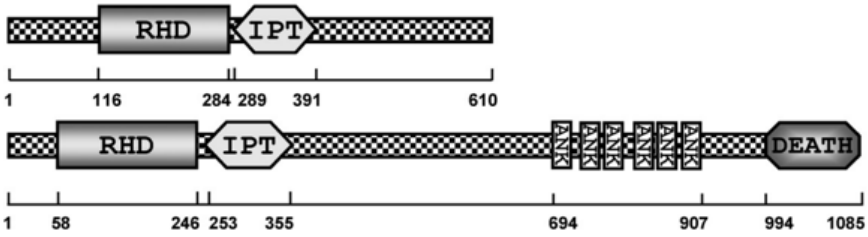


Figure 7. Schematic of NF-κB structure in B. glabrata. Adapted from Zhang & Coultas, 2011.

Studies of BgRel and BgRelish have thus far been limited, but several notable findings have emerged. Zhang and Coultas found increased expression of BgRel and BgRelish in *S. mansoni*-resistant whole snail tissues six hours after exposure to *S. mansoni*. Additionally, expression of BgRel decreased after 24 hours (BgRelish expression also decreased after 24 hours, but the decrease did not reach statistical significance), and both BgRel and BgRelish expression returned to baseline by 72 hours post-exposure (Zhang and Coultas, 2011). This finding suggests *B. glabrata* NF-κB proteins are present during an immune response but does not confirm any active role for NF-κB proteins. Humphries and Harter provided evidence that BgRel has functional potential by demonstrating that the BgRel RHD can bind to κB sequences found in the snail genome (Humphries and Harter, 2015). The κB sites tested lie upstream of the *B. glabrata* IκB and MAPK p38 genes, both of which are targets of NF-κB proteins in humans during an immune response. Li et al. found expression of NF-κB proteins in *B. glabrata* hemocytes and identified more expression in resistant than susceptible strains (Li et al., 2022). Additionally, Marsh identified the presence of BgRelish during embryogenesis, and in the eyespot, muscle-like tissue, and putative central nervous system tissue in *B. glabrata* embryos (Marsh, 2021). BgRelish has also been found localizing to the ganglia in adult *B. glabrata* (personal communication, Ben Glazer, 2022). Yet, while growing evidence suggests putative physiological roles of NF-κB proteins in *B. glabrata*, functional characterizations have remained elusive.

Chromatin Immunoprecipitation (ChIP)

The ChIP assay is a powerful tool for understanding epigenetics and can be used to determine the activities of transcription factors. In brief, ChIP allows researchers to isolate short pieces of DNA to which a TF or other DNA-binding protein-of-interest is bound. One notable downstream analysis is ChIP-seq, which adds next-generation sequencing (NGS) to the end of this procedure, allowing researchers to identify the DNA sequences being bound by their protein-of-interest.

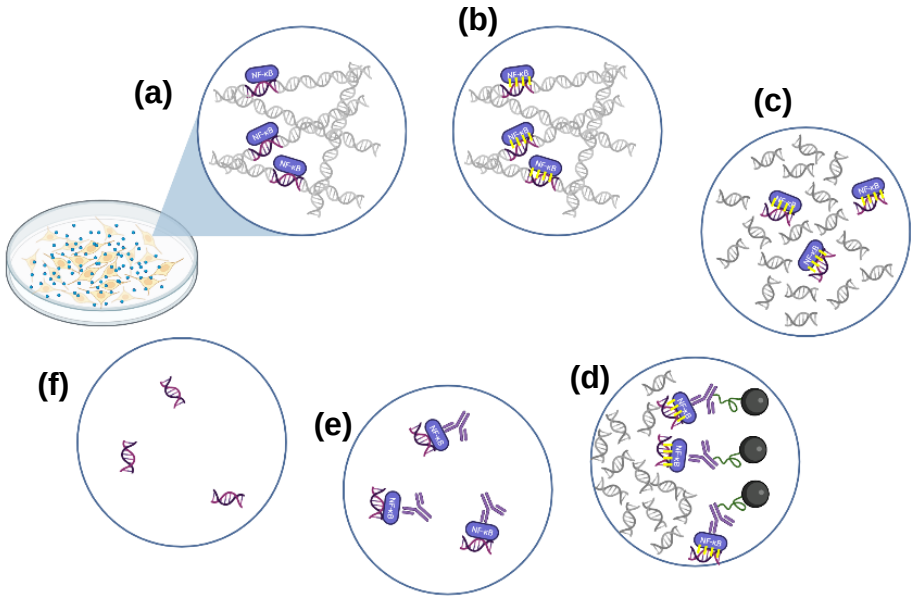


Figure 8. ChIP Methodology. (a) Sample treatment, (b) crosslinking, (c) sonication, (d) immunoprecipitation, (e) elution and reverse crosslinking, (f) purified DNA. Created using BioRender.

Methodology

ChIP experiments are most easily performed using cultured cells as the biological sample, due to the necessity of sufficient quantities of starting material. Cells are first treated with any stimuli necessary for the experiment (Fig. 8a), if applicable (for example: treatment with *S. mansoni* immunogenic molecules). Then, cells are chemically crosslinked (Fig. 8b) to preserve all transient molecular interactions, including the binding of regulatory proteins to DNA. Crosslinked cells are lysed to release nuclei, and these nuclei are then isolated and lysed to release chromatin. Chromatin is the tightly packed complex of DNA wound around histone proteins, which is the usual configuration of DNA within cells. The nuclear lysate containing chromatin is then subjected to sonication (Fig. 8c), during which high frequency sound waves shear the DNA, breaking it into small fragments. Some of these DNA fragments are bound by DNA-binding proteins (for example: NF-κB proteins). Using antibodies that specifically

recognize the DNA-binding protein-of-interest, the DNA-binding protein and its bound DNA fragment are immunoprecipitated (Fig. 8d) using commercially available beads coated with a bacterial protein which binds strongly to antibodies (in this study: protein A). Finally, the immunoprecipitated sample can be reverse-crosslinked, the protein contents (antibodies and protein-of-interest) can be digested (Fig. 8e), and the bound DNA fragments can be purified for downstream applications (Fig. 8f) (Kidder et al., 2011).

NGS is a particularly useful downstream application of ChIPed DNA, particularly for transcription factors, as it allows researchers to identify the binding sequences for their protein-of-interest. Short fragments of DNA can be sequenced, then algorithms can be used to align sequences to the organism's genome, mapping the TF binding sites. By analyzing the genes near these TF binding sites, ChIP-seq can suggest which specific genes might be targets of the TF's activity (Park, 2009).

Another downstream application of ChIPed DNA is quantitative polymerase chain reaction (qPCR,) which is often used as a quality control measure for ChIP-seq experiments. PCR is an assay which copies a template DNA sequence millions to billions of times. qPCR allows for quantification of the original amount of template DNA by using a fluorescence reporter which fluoresces brighter as more DNA is made. Quality control qPCR requires researchers to know at least one binding site for their protein-of-interest, and an immunoprecipitated sample should demonstrate significant enrichment of that known binding site (measured by qPCR) when compared to a pre-immunoprecipitation sample (Landt et al., 2012). qPCR can also report relative DNA-binding activity of the protein-of-interest for a specific condition. This application necessitates that all chromatin samples to be compared are normalized before immunoprecipitation and requires multiple conditions to be compared. Normalization can be achieved after crosslinking by counting cells and proceeding to the lysis step with the same number of cells for each treatment condition. Provided the proper procedures are followed, ChIP-qPCR can suggest a functional role for a TF by reporting differences in relative TF-binding activity under control and treatment conditions.

ChIP Optimization

ChIP experiments are notoriously complex, requiring optimization at multiple levels to ensure reproducible data. While the primary source of a failed ChIP experiment is an incompatible antibody, researchers must also optimize variables during the crosslinking and sonication steps.

Antibody validation is required to demonstrate that an antibody is specific enough for use in ChIP. In an attempt to improve the quality of publicly available ChIP-seq data, the ENCODE (Encyclopedia of DNA Elements) consortium requires ChIP antibodies to pass both a primary and secondary validation assay in order for a ChIP-seq dataset to be submitted to the ENCODE database. For transcription factor ChIP experiments, the primary validation of an antibody is assessed by an immunoblot or immunofluorescence. Immunoblot validation requires that when the sample is probed with the intended antibody in an immunoblot, the result is a band at the expected size of the protein-of-interest. This correct band must also be the dominant band, and if other bands are present the correct band must make up at least 50% of the total signal within the lane. Immunofluorescence validation requires that a signal is observed localizing as expected; for example, NF-κB is expected to localize primarily in the cytoplasm under basal conditions, and in the nucleus during an immune response. The expected localization of immunofluorescence should be combined with some method to reduce the amount of the protein-of-interest, which must correspond to a decrease in signal (Landt et al., 2012; Wardle and Tan, 2015).

Secondary validation can be satisfied by a variety of methods, including knockdown/knockout of the protein-of-interest, mass spectroscopy of all immunoprecipitated proteins, use of multiple antibodies recognizing different parts of the protein-of-interest, immunoprecipitation with an epitope-tagged version of the protein-of-interest, or ChIP-seq data reporting expected binding sites. In this study, neither mass spectroscopy nor the acquisition of additional antibodies is possible due to prohibitive costs, and so these are not viable secondary validation methods. Due to limitations of time, the creation of knockdown/knockout or epitope-tagged Bge strains is similarly impractical. Thus, this experiment will

rely on secondary validation via expected motif enrichment generated by ChIP-seq data (Landt et al., 2012; Wardle and Tan, 2015).

Designing an effective ChIP experiment also involves the optimization of crosslinking and sonication conditions to ensure reproducibility and good sample quality for downstream applications. Insufficient crosslinking can lead to dissociation of the protein-of-interest from the DNA, resulting in low final DNA recovery. Over-crosslinking may obscure the protein epitopes recognized by the antibody, leading to low immunoprecipitation yields (and thus also low DNA yields). Over-crosslinking can also cause DNA to become resistant to sonication (Hoffman et al., 2015). Incomplete sonication results in long DNA fragments that are unsuitable for NGS, while over-sonication yields fragments that are too short and likewise unsuitable for NGS (Keller et al., 2021).

ChIP for *B. glabrata* NF-κB

As a likely immune transcription factor, *B. glabrata* NF-κB is an attractive target for ChIP experiments. Some binding sequences for *B. glabrata* NF-κB proteins have been predicted (Humphries and Deneckere, 2018) and confirmed using partial NF-κB proteins (Humphries and Deneckere, 2018, Humphries and Harter, 2015), but it is likely that most binding sites remain unknown. By identifying these as-yet unknown binding sequences and comparing between different evolutionary lineages, we may gain insights into the evolution of NF-κB binding sequences. NF-κB binding sequence data is already available for humans, *D. melanogaster*, and *Aiptasia* using an alternate method of binding site identification: protein-binding microarrays (PBMs). PBMs measure binding affinities to a random collection of thousands of short DNA sequences, allowing for high-throughput determination of binding sequences (Annala et al., 2011). No PBMs have been performed for Lophotrochozoa NF-κB proteins, but data exists for Deuterostomia (humans), Ecdysozoa (*Drosophila*), and *Aiptasia*, a Cnidarian whose MRCA with mollusks is shared by all Bilateria (Deuterostomia, Ecdysozoa, and Lophotrochozoa, Fig. 3) (Siggers et al., 2015). Although PBM and ChIP-seq data are not directly comparable, the identification of

NF-κB binding sequences for *B. glabrata*, a Lophotochozoan, by ChIP-seq may provide insight into the evolution of NF-κB binding sequences as Bilateria superphyla diverged.

Additionally, the identification of NF-κB binding sites in immune-challenged samples will suggest potential immune gene targets of NF-κB, addressing our gap in understanding of *B. glabrata* immune regulation. DNA regulatory regions are often proximal to the DNA sequences they control, so by searching the genes surrounding the binding sites we identify, we can hypothesize the gene targets of NF-κB. Furthermore, ChIP-qPCR studies may suggest the functionality of *B. glabrata* NF-κB in the immune system. For example, ChIPed DNA from control-treated and *S. mansoni*-treated cells can be compared, and qPCR can be used to measure the relative difference in κB DNA sequences. If the *S. mansoni* sample contains significantly more κB DNA sequences than the control, this indicates that the *S. mansoni* immune challenge increased the number of NF-κB binding events, thus suggesting functional activity of *B. glabrata* NF-κB.

Finally, to the best of my knowledge, no ChIP experiment has yet been performed in any molluscan organism. Given the availability of the Bge cell line for *B. glabrata*, a ChIP protocol for Bge cells has the potential to facilitate future useful epigenetic studies in *B. glabrata*. Logistically, the availability of custom antibodies specific to *B. glabrata* NF-κB p65 (BgRel) and p50 (BgRelish) in the Humphries lab represents a valuable opportunity to design a Bge ChIP protocol while collecting data to fill a critical gap in our knowledge of *B. glabrata* immunity.

Methods

Bge Cell Culture

Bge cells were obtained from either Biomedical Research Institute (BRI, Rockville, MD, U.S.A.) or the laboratory of M. G. Castillo (New Mexico State University, Las Cruces, NM, USA). Cells were cultured in complete Bge medium (cBge, 22% Schneider's Insect Media (Sigma-Aldrich; MilliporeSigma, St. Louis, MO, U.S.A.), 0.45% lactalbumin hydrolysate (MP Biomedicals, Santa Ana, CA, U.S.A.), 0.13% galactose (Sigma-Aldrich), 10% FBS (Sigma-Aldrich), 100units/mL penicillin and 100ug/mL streptomycin (Sigma-Aldrich), pH 7.2), and grown at 26°C. Flasks were passaged as needed (approximately every 1-3 weeks) and re-seeded at densities to maintain >70% confluence.

Antibody Preparation

Custom antibodies targeting BgRel and BgRelish were generated in New Zealand rabbits by Pacific Immunology (San Diego, California, U.S.A). Anti-BgRel was designed against amino acids 590-604 (IHILDHDPETEAALR) on the C-terminus of BgRel, while anti-BgRelish was designed against amino acids 2-22 (SSYGSSSNDSDTLNENNLPVD) on the N-terminus of BgRelish. All antibody procedures in this study were performed using the same batch of each antibody, affinity purified as follows. Terminal-bleed serum was filtered through a 0.2μM filter into an equal volume of Phosphate Buffered Saline (PBS) pH 7.2 (135 mM NaCl, 2.7 mM KCl,11 mM $NaPO_4$). The serum in PBS was loaded onto affinity columns from Pacific Immunology that contained the epitope (the amino acid sequence against which the antibody was raised) bound to resin. Each column was washed with PBS until the absorbance at 280nm of the flow-through matched a PBS blank. Antibodies were eluted using 0.2M glycine (Research Products International, Mt. Prospect, IL, U.S.A.) pH 1.85, and the eluate was collected directly into Tris-HCl pH 8.5 to neutralize the pH. Antibodies were concentrated using 10kDa molecular weight cutoff Spin-X® UF columns (Corning, Corning, NY, U.S.A.) at 4,000 rcf. The spin column was

also used to exchange the buffer to PBS with 0.02% NaN_3, pH 7.2. Antibodies were concentrated to ~1μg/mL, aliquoted, and stored at -20°C.

Western Blotting

In preparation for SDS-PAGE, all protein samples were boiled at 95-100°C for 5 minutes, then centrifuged for 2 minutes at 13,100 rcf. SDS-PAGE was performed using 10% acrylamide hand-cast gels. After SDS-PAGE, proteins were transferred to a nitrocellulose membrane either overnight at 25 volts (4°C) or at 100 volts (room temperature) for one hour. Membranes were incubated in blocker (Tris Buffered Saline (TBS; 25 mM Tris, 140 mM NaCl, 3 mM KCl), 0.05% Tween-20 v/v, 5% nonfat dried milk w/v) overnight at 4°C or for at least one hour at room temperature. Following, blots were incubated with 0.5μg/mL primary antibody diluted in blocker overnight at 4°C or for one hour at room temperature, then briefly rinsed once in TBST (TBS, 0.05% Tween-20 v/v) and washed three times for 15 minutes each in TBST. Blots were incubated with 1:2000 dilution anti-rabbit alkaline phosphatase (AP)-linked antibody (Cell Signaling, Danvers, MA, U.S.A.) in blocker for one hour at room temperature, then subjected to the same wash procedure. Blots were washed at least 5 minutes in AP buffer (100mM Tris-HCl, 100mM NaCl, 5mM MgCl2, 0.05% Tween 20 v/v, pH 9.5), incubated with NBT and BCIP according to manufacturer's instructions (Promega, Madison, WI, U.S.A.), and imaged after 30 minutes using a Pixel 5a phone camera.

qPCR

qPCR experiments were developed for downstream confirmation that the ChIP procedure can specifically isolate DNA sequences bound to NF-κB proteins. Primers were designed to amplify a region of DNA containing a hypothesized κB sequence (positive control) and a region of DNA without a κB sequence (negative control). Successful ChIP experiments should ideally exhibit at least five-fold enrichment of the κB region over the negative control region, which can be measured by observing

differences in the threshold cycle (C_T) of reactions. The qPCR reactions used here employ SYBR Green as a reporter, which fluoresces when it binds to double stranded DNA. As the qPCR reaction exponentially copies template DNA, the reaction fluoresces brighter. The C_T value is the cycle number at which this fluorescence signal reaches a set threshold. Lower C_T values mean this threshold was reached sooner, indicating more of the DNA template was present in the sample.

Previous research in the Humphries lab has confirmed *in vitro* binding of the BgRel RHD to putative κB sites upstream of the *B. glabrata* IκB gene (Humphries and Harter, 2015). Thus, a BgRel positive control qPCR primer set was designed to amplify within the 350bp region flanking this κB site upstream of IκB (see Appendix II). To develop a negative control primer set, the 2,000bp upstream of various *B. glabrata* genes was screened for predicted κB sites. The upstream region of the integrin beta-PS-like gene (see Appendix II) was chosen as a negative control because no significant κB sites were predicted.

Primers were designed to amplify a 75-200bp region within the positive or negative control regions. Primers for IκB and integrin were designed using Primer3 (Koressaar and Remm, 2007, Untergasser et al., 2012, Koressaar et al., 2018), then BLASTed (Altschul et al., 1990) to check their specificity within the *B. glabrata* genome. IDT OligoAnalyzer was used to assess the tendency of primers to form self-dimers, hairpins, or heterodimers by retaining only primers with $\Delta G > -7$, or $\Delta G > -5$ for homology at the 3' ends. Primers were obtained from Invitrogen. Sequences can be found in Appendix I, Table 2.

qPCR experiments were performed and analyzed using Applied Biosystems 7500 Real Time PCR System (Applied Biosystems; Thermo Fisher Scientific, Waltham, MA, U.S.A.) and associated software. Initial experiments were performed using G2 qPCR Master Mix (Promega), and further troubleshooting and final standard curves were performed using SsoAdvanced Universal SYBR Green Supermix (Bio-Rad, Hercules, CA, U.S.A.). Each primer set was validated by performing a standard curve using 100, 20, 4, 0.8, and 0.16ng of BB02 genomic DNA (BRI) in triplicate. No template controls (NTCs) for each primer set were also performed in triplicate. Final amplification curves and melt curves for each primer

set can be found in Appendix II, Supplemental Fig.4. Cycling conditions for G2 qPCR Master Mix were as follows: 120 sec at 95°C; [15 sec at 95°C; 60 sec at 55-67.5°C] x 40. Cycling conditions for SsoAdvanced Universal SYBR Green Supermix were as follows: 180 sec at 98°C; [15 sec at 98°C; 60 sec at 65-72°C] x 40.

Chromatin Immunoprecipitation

Crosslinking and Cell Lysis

Bge cells were prepared for chromatin immunoprecipitation using a modified protocol from Schmidt et al., 2009. Cells were grown in cBge medium in culture dishes to densities of approximately 10^6 cells per 10cm dish. cBge media was then aspirated and replaced with room temperature CBSS, taking care not to disturb the adherent cells. After incubating one hour at room temperature, 16% methanol-free formaldehyde (Pierce; Thermo Fisher Scientific) was added to a final concentration of 1%. Cells were left to crosslink while rocking for 2.5 to 15 minutes, then the reaction was quenched with 2M glycine to a final concentration of 0.125M glycine. After rocking for 10 minutes, cells were rinsed twice with ice-cold snail PBS (sPBS; 8.41mM Na_2HPO_4, 1.65mM NaH_2PO_4, 45.34mM NaCl, pH 7.2), then collected in ice-cold sPBS with a cell scraper. Cells were kept on ice and resuspended in only ice-cold solutions for the remainder of the protocol until the DNA purification steps. Cell suspensions were centrifuged at 300 rcf for 5 minutes at 4°C, resuspended in sPBS, counted via hemocytometer, and centrifuged again. Pellets were then either flash frozen and stored at -80°C before continuing or immediately resuspended in LB1 (50 mM Hepes–KOH, pH 7.5, 140 mM NaCl, 1 mM EDTA, 10% glycerol, Igepal, 0.25% Triton X-100, 1:500 dilution CalBiochem protease inhibitor cocktail IV (Millipore, MilliporeSigma)). Cells were incubated in LB1 with rocking for 10 minutes on ice, then centrifuged at 2,000 rcf for 4 minutes at 4°C. Pellets were resuspended in LB2 (10 mM Tris–HCL, pH8.0, 200 mM NaCl, 1 mM EDTA, 0.5 mM EGTA, 1:500 dilution CalBiochem protease inhibitor cocktail IV) and incubated on a rocker for 5 minutes on ice before being pelleted at 2,000 rcf for 5 minutes at 4°C to

yield nuclear pellets. Nuclear pellets were resuspended in LB3 (10 mM Tris–HCl, pH 8, 100 mM NaCl, 1 mM EDTA, 0.5 mM EGTA, 0.1% Na–deoxycholate, 0.5% N-lauroylsarcosine, 1:500 dilution CalBiochem protease inhibitor cocktail IV) with 0-0.5% SDS at volumes to yield densities of approximately 3,000 or 10,000 nuclei/μL. Nuclear suspensions were aliquoted into polypropylene tubes or TPX tubes (Diagenode, Denville, NJ, U.S.A.) at volumes of 100-300μL, making sure all samples were at the same volume and concentration.

Sonication

Sonication was performed using a Bioruptor® Standard (Diagenode) water bath sonicator. The sonicator was always loaded with 6 microtubes of the same volume to ensure consistent sonication among all samples. The water bath was pre-cooled to 4°C, and samples were vortexed for 2-5 seconds then centrifuged 2,000 rcf for 3-5 seconds before being loaded into the sonicator. Sonication was performed at medium or high settings for 10-60 cycles of 30 seconds on, 60 seconds off. After sonication, 10% TritonX-100 was added to a final concentration of 1% and samples were centrifuged at 16,000 rcf for 10 minutes. An aliquot for each condition was saved as an input control (no immunoprecipitation).

Immunoprecipitation

To prepare for immunoprecipitation, Dynabeads™ Protein A (Invitrogen; Thermo Fisher Scientific) were washed twice with 0.1M $NaPO_4$, pH 8.0. For each immunoprecipitation reaction, 3μg of anti-BgRel was added per 20μL Dynabeads™ in a 0.1M $NaPO_4$ solution. To allow protein A to bind to the rabbit antibody, anti-BgRel was incubated with Dynabeads™ for at least two hours at room temperature while rotating. When being used for binding, Dynabeads™ were never resuspended in volumes exceeding five times the original volume of Dynabeads™ (e.g., 20μL Dynabeads™ can be resuspended in a maximum of 250μL for binding). The antibody-bead complexes were washed gently three times with 0.1M $NaPO_4$ pH 8.0 with 0.1% Tween-20, then sonicated chromatin samples were added. The mixtures were incubated while rotating at least 2 hours, then bead complexes were washed

gently three times with wash solution (0.1M $NaPO_4$, 0.1% Tween-20, pH 8.0) and eluted with 50mM glycine pH 2.8 for two minutes while rotating. Each eluate was transferred to a new tube with one quarter its volume of 1M Tris pH 8 to neutralize the pH. Beads were resuspended in 0.1M $NaPO_4$ pH 8.0 with 0.02% NaN_3 and stored at 4°C.

Dot Blot

To confirm the successful capture and elution of the anti-BgRel antibody in ChIP, 2µL of eluate from the immunoprecipitation was blotted to a nitrocellulose membrane, blocked with 5% skim milk in TBS with 0.1% Tween-20 and probed with a 1:2000 dilution of HRP-linked anti-rabbit antibody (Cell Signaling). Blots were washed 3 times in TBS with 0.1% Tween-20, then once for at least 5 minutes in TBS before adding SuperSignal™ West Femto Maximum Sensitivity Substrate (Thermo Fisher Scientific). Blots were imaged within 15 minutes of adding the luminol-peroxide substrate using UVP ChemiDoc-It®2 Imager (Analytik Jena GmbH, Jena, Germany).

DNA Purification

To prepare DNA for purification, input DNA and immunoprecipitation eluate samples were diluted with an equal volume of TE buffer (10mM Tris, 0.5mM EDTA), then RNase A was added to a final concentration of 50µg/mL. This mixture was incubated 2 hours at 37°C, then proteinase K was added to a final concentration of 50µg/mL and incubated 2 hours at 37°C. To reverse crosslinks, the solution was incubated at 65°C overnight. DNA purification was performed using either a PCR Purification Kit (Bio Basic, Amherst, NY, U.S.A.) or ChIP DNA Clean and Concentrator (Zymo Research, Irvine, CA, U.S.A.).

Results

qPCR Assay Development

qPCR assays are a necessary quality control step for ChIP, validating that the appropriate sequences are enriched in ChIPed samples over negative controls. Thus, qPCR reactions for two primer sets (one positive and one negative control for κB sequences) were validated for use in our optimizations of ChIP methods.

First, a primer matrix was performed with primer concentrations of 200, 400, 600, and 800nM for both forward and reverse primers. Reactions used G2 qPCR Master Mix, the Promega-suggested annealing temperature of 60°C, and 100ng of BB02 gDNA as the template. All integrin reactions performed similarly well (Appendix II), and so the minimum primer concentration (200nM) was used for both integrin primers. However, some IκB reactions exhibited primer dimer formation, as evidenced by the melt curves (Appendix II), especially at high concentrations of the forward primer. Thus, I decided to use a forward primer concentration of 200nM and reverse primer concentration of 800nM for the IκB primer set, as this condition clearly resulted in no primer dimer formation, as evidenced by the melt curve (Supplemental Fig.1).

Next, the reaction annealing temperature was optimized with these primer concentrations. Using the G2 qPCR Master Mix, both integrin and IκB primer sets were run with 4ng of BB02 gDNA at annealing temperatures of 55, 57.5, 60, and 62.5°C. The integrin primer set produced amplification curves with an odd bent shape that appeared less prominent at high annealing temperatures (Appendix II, Supplemental Fig.2), so integrin qPCR was also performed at 65 and 67.5°C. On average, the C_T for both primer sets increased as annealing temperatures increased, indicating a less efficient reaction at higher temperatures. However, for integrin, reactions run at 65°C amplified only about one cycle later than reactions run at 55°C (Fig.9a). Because the amplification curve shape was improved by the higher temperature, the optimal annealing temperature for integrin was determined to be 65°C. Integrin reactions run at 62.5°C and 65°C also exhibited a lower standard deviation between replicates (Fig.9a) than all

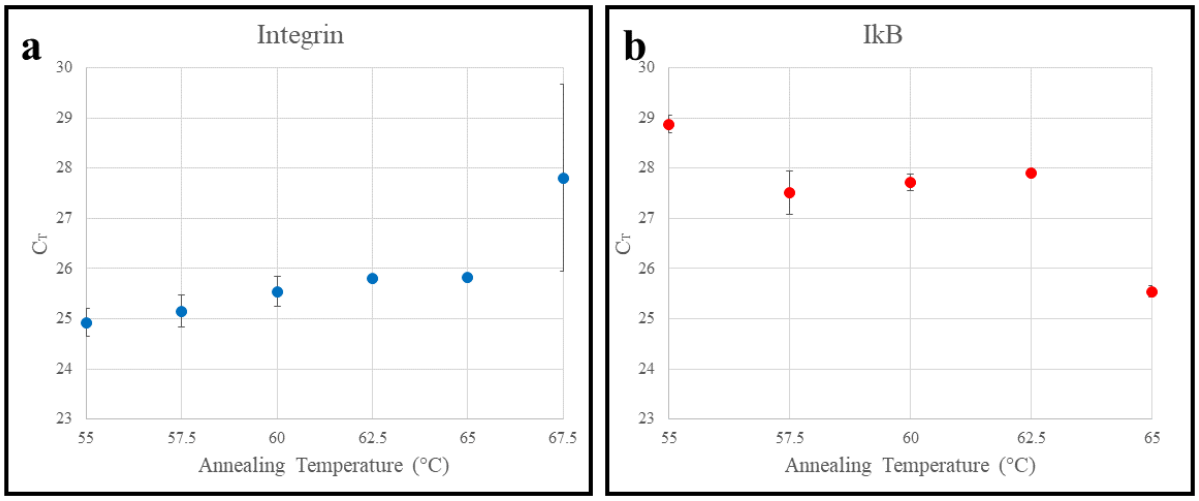


Figure 9. Annealing temperature optimization for integrin and IκB primer sets using G2 qPCR master mix. (a) Integrin, (b) IκB. 4ng of BB02 gDNA was added as the template to each reaction. Each annealing temperature condition was performed in triplicate, and error bars denote standard deviation.

other temperatures, indicating more reproducible assay conditions. The IκB primer set also exhibited a lower standard deviation between replicates at 62.5°C while maintaining a C_T within one cycle of the 57.5°C reaction condition (Fig.9b). At 55°C, primer dimers formed during IκB reactions (Supplemental Fig.3a), and to maximize primer specificity, it was determined to run IκB reactions at high annealing temperatures. For ease, the final standard curve reaction was planned to be run with integrin at 65°C.

Because of limited reagent availability, all following reactions were performed using SsoAdvanced Universal SYBR Green Supermix, rather than G2 qPCR Master Mix. Using the new master mix, standard curves for both integrin and IκB primer sets were performed using an annealing temperature of 65°C. The integrin reaction performed well with the SsoAdvanced Universal SYBR Green Supermix, generating a standard curve (Fig.10) with efficiency of 99.35%, as well as an improved amplification curve shape (Appendix II, Supplemental Fig.4c). However, the IκB primer set amplified primer dimers (Fig.11a) after switching to the new master mix, and therefore necessitated an additional round of optimization for its use with the SsoAdvanced Universal SYBR Green Supermix.

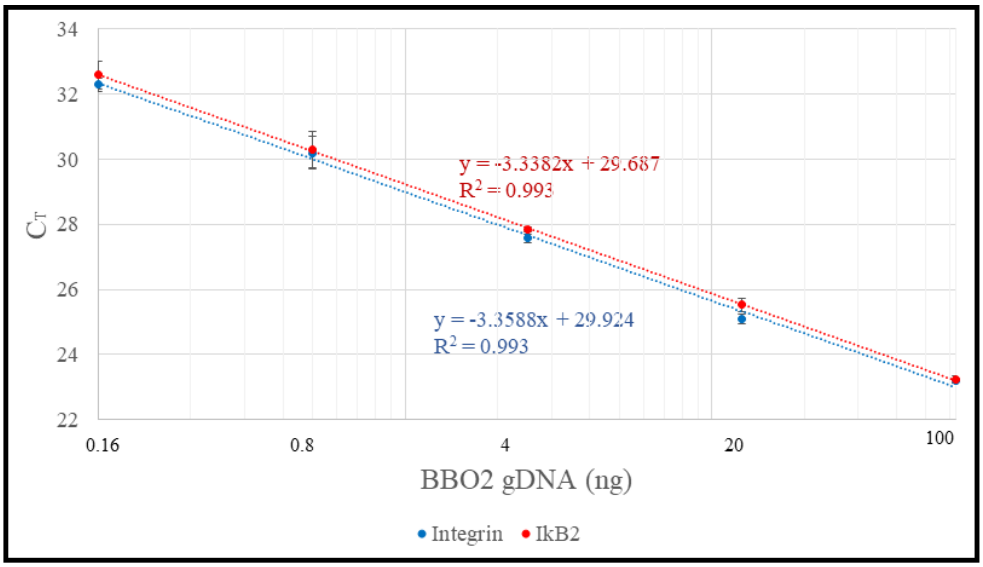


Figure 10. Standard curves for integrin and IκB primer sets. Reactions were performed using SsoAdvanced Universal SYBR Green Supermix and five-fold dilutions from 100ng of BB02 gDNA. Integrin standard curve (blue) was performed at an annealing temperature of 65°C and forward and reverse primer concentrations of 200nM. IκB standard curve (red) was performed at an annealing temperature of 67.5°C and forward primer concentration of 100nM and reverse primer concentration of 400nM. Each gDNA concentration was performed in triplicate; error bars denote standard deviation. Equations and R^2 were calculated using Applied Biosystems 7500 software.

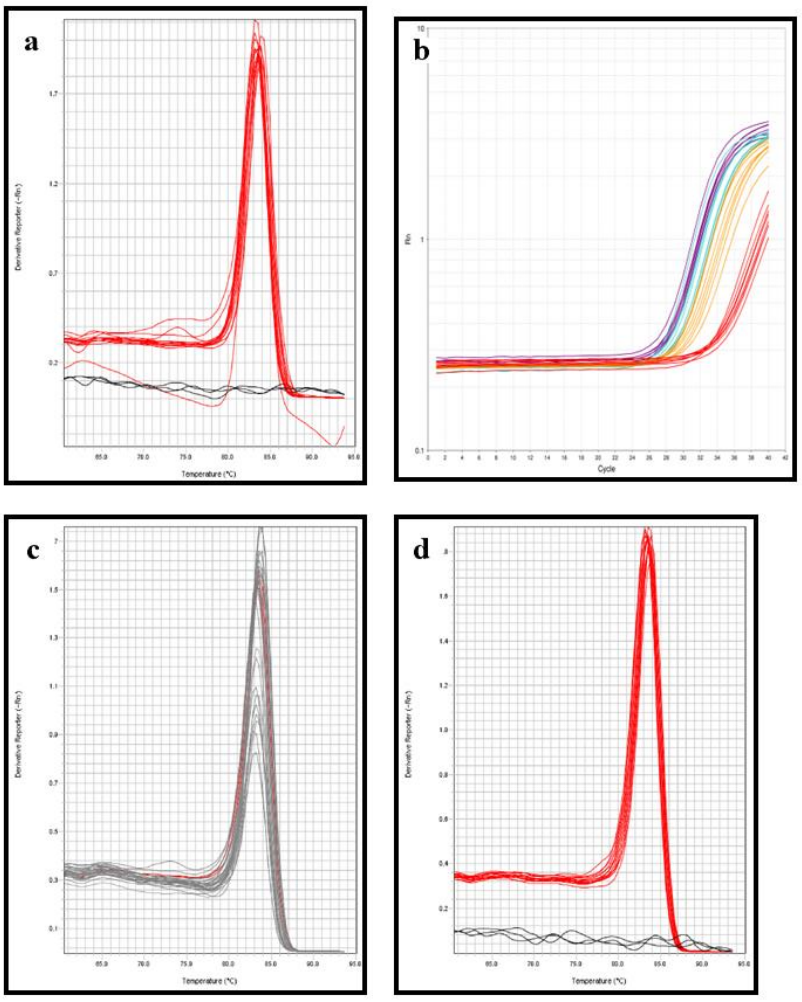


Figure 11. Melt and amplification curves for IκB SsoAdvanced master mix optimization. (a) Melt curve for standard curve run at annealing temperature = 65°C and primer concentrations of forward = 200nM and reverse = 800nM. Standard curve reactions are in red, NTC reactions are in black. (b) Plot of amplification curves for primer matrix including 100, 200, 300, and 400nM concentrations of each primer, two replicates for each condition. Color denotes reverse primer concentration; red = 100nM, orange = 200nM, blue = 300nM, purple = 400nM. (c) Melt curve for primer matrix. Chosen primer concentration (forward = 100nM, reverse = 400nM) in red, all other primer concentrations in grey. (d) Melt curve for standard curve run at annealing temperature = 67.5°C and primer concentrations of forward = 100nM and reverse = 400nM. Standard curve reactions are in red, NTC reactions are in black. Plots generated by Applied Biosystems 7500 software.

To reduce the likelihood of primer dimer formation, I decided to first reduce the primer concentration in the IκB reaction. Another primer concentration matrix was performed, this time using concentrations of 100, 200, 300, and 400nM of each primer. This revealed that low reverse primer concentrations had the greatest negative impact on C_T values (Fig.11b), but some of these conditions still led to primer dimer formation (Fig.11c). To minimize both C_T and overall primer concentration, I decided to use forward primer concentrations of 100nM, and reverse primer concentrations of 400nM, a condition which also demonstrated a single melt curve peak during testing (Fig.11c).

To finalize IκB qPCR reaction conditions, standard curves were run with the new, lower primer concentrations at annealing temperatures of both 65°C and 67.5°C. Amplification specificity was slightly improved at 67.5°C in that no primer dimer formation was visible in the melt curve (Fig.11c). The 65°C condition resulted in lower C_T values, but also saw primer dimer amplification in the NTC at a C_T of 37 (Appendix II, Supplemental Fig.3b). Thus, the optimal qPCR reaction conditions for the IκB primer set were determined to be: forward primer concentration = 100nM, reverse primer concentration = 400nM, and annealing temperature = 67.5°C. These conditions resulted in an efficiency of 98.46%, as calculated from the standard curve (Fig.10).

Bge Cell ChIP Optimization

Sonication

Optimal ChIP sonication conditions will yield DNA fragments between 100 and 500bp (Wardle and Tan, 2015, Kidder et al., 2011). This length results in short enough fragments to promote accurate identification of NF-κB target genes and κB sequences, while being large enough to enable high quality next generation sequencing. To optimize the

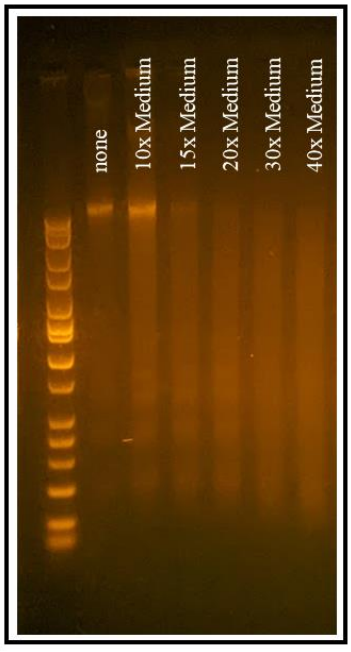


Figure 12. Bge chromatin sonication in microtubes. Cells were previously crosslinked for 10 minutes, and after sonication the DNA was purified using a PCR Purification kit. Samples were run with Hilo (Bionexus; Kansas City, MO, U.S.A.) DNA marker on 1% agarose gel with GelRed® (Biotium) in TAE (40mM Tris, 20mM acetic acid, 1mM EDTA). Gel was imaged using Fotodyne™ FOTO/Phoresis™ UV Transilluminator (Hartland, WI, U.S.A.) and Pixel 5a phone camera.

shearing of Bge cell chromatin, I first attempted Bioruptor® sonication of 10 minute-crosslinked cells in 1.5mL microtubes for 10, 15, 20, 30, and 40 cycles of 30 seconds ON, 60 seconds OFF at medium power. Very poor shearing was observed (Fig.12) for all conditions; although the large genomic fragment was clearly degraded after 20 cycles, no consistent enrichment in the 100-500bp range was observed. To expedite preliminary sonication optimization, I simplified the protocol by testing sonication conditions using commercially available DNA sources.

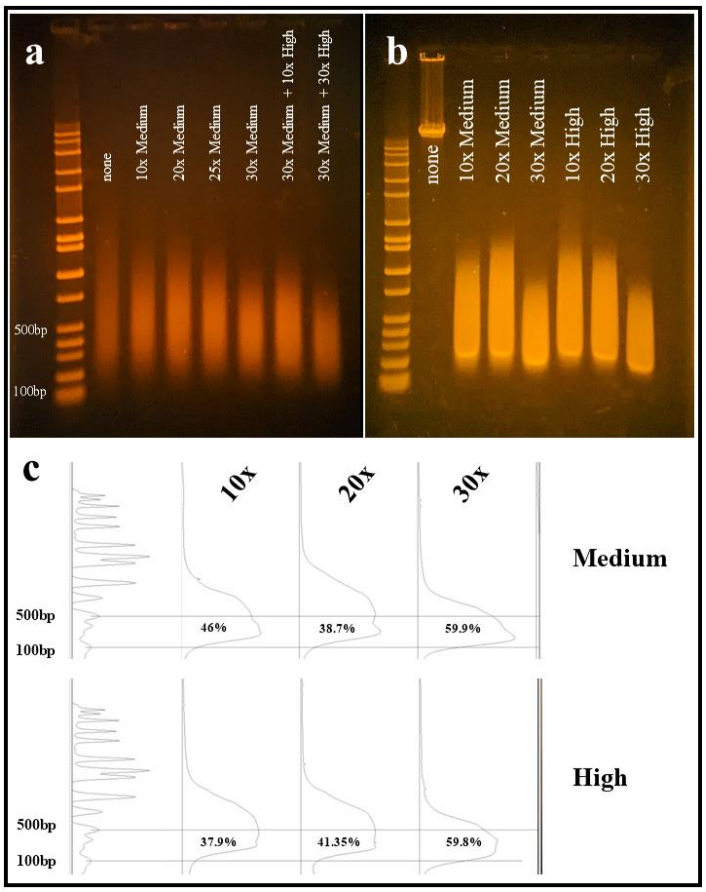


Figure 13. Sonication optimization using purified, commercially available DNA. (a) Herring sperm DNA sonicated at a concentration of 750µg /mL. (b) Lambda DNA sonicated at a concentration of 50ng/µL. (c) Lambda DNA quantification using ImageJ. Lines were added at 100 and 500bp (using HiLo as a reference) to identify the ideal range of fragment sizes, and the percentage of gel signal within this ideal range is listed for each lane. Samples were run with HiLo DNA marker on 1.25% agarose gels, then post-stained for 30 minutes in an 8% GelRed solution. Gels were imaged using Fotodyne™ FOTO/Phoresis™ UV Transilluminator and Pixel 5a phone camera.

First, Herring Sperm DNA (Promega) was sonicated in TPX tubes for 10, 15, 20, and 30 cycles of 30 seconds ON, 60 seconds OFF at medium power. Additional samples were sonicated 30 cycles at medium power followed by 10 or 30 cycles at high power. Sonication efficiency was markedly improved (Fig.13a), with sonication at 30 cycles in TPX tubes generating 100-500bp sheared DNA. Due to this improvement, TPX tubes were used for sonication in all subsequent experiments. Next, lambda phage DNA (Sigma-Aldrich) was sonicated for 10, 20, or 30 cycles at medium or high power. The high molecular weight band for lambda DNA was clearly sheared in all conditions, with 30 cycles at either medium or high power resulting in the highest percentage of chromatin fragments within the 100-500bp range (Fig.13b-c). The sonication performance was very similar between medium and high power

conditions (Fig.13c), but I ultimately elected to use high power in all subsequent experiments, anticipating that Bge chromatin would be more resistant to sonication than herring sperm or lambda DNA. Bge chromatin for ChIP will be crosslinked and will contain nuclear proteins and other debris that might inhibit sonication efficiency.

With general sonication conditions established, I next attempted sonication of Bge chromatin that had been cross-linked for 10 minutes. Nuclei obtained from concentrations of approximately 3,000 cells/µL were sonicated for 10, 20, 25, and 30 cycles. Sonication at 25 and 30 cycles showed clear improvement in sonication efficiency (Fig.14a) over the previous attempt (Fig.12), yet a large portion of chromatin remained at lengths >500bp. SDS is often used in ChIP experiments (Kidder et al., 2011) to improve sonication efficiency by disrupting the tight packing of chromatin, thus better exposing the DNA to shearing sound waves. Accordingly, I next supplemented the sonication buffer (LB3) with 0, 0.2%, and 0.5% SDS before sonication. I tested these three SDS concentrations using Bge cell nuclei from both 10,000 cells/µL and 3,000 cell/µL. In both cases, the optimum conditions for obtaining 100-500bp chromatin fragments were achieved when the sonication buffer was supplemented with 0.2% SDS (Fig.14b).

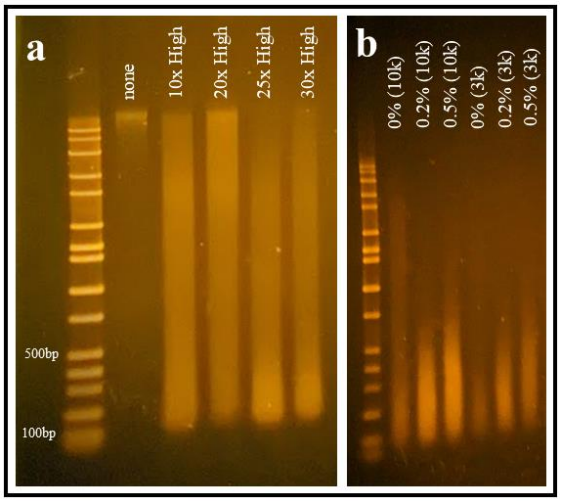


Figure 14. Sonication optimization of crosslinked Bge chromatin with/without SDS supplementation. (a) Bge chromatin sonication without SDS. (b) Bge chromatin sonication at 30 cycles on high, with 0%, 0.2%, or 0.5% SDS. Number of cells used is indicated with "10k" (10,000 cells/uL) or "3k" (3,000 cells/uL) DNA was purified using Zymo ChIP DNA Clean and Concentrator kit. Samples were run with HiLo DNA marker on 1.25% agarose gels in TAE, then post-stained for 30 minutes in an 8% GelRed solution. Gels were imaged using Fotodyne™ FOTO/Phoresis™ UV Transilluminator and Pixel 5a phone camera.

Thus, the optimum sonication conditions for 10-minute crosslinked Bge chromatin samples are as follows: Bioruptor® sonication in TPX tubes for 30 cycles of 30 seconds ON, 60 seconds OFF in sonication buffer supplemented with 0.2% SDS.

Crosslinking

Sufficient crosslinking times are necessary to ensure NF-κB proteins remain bound to DNA throughout the ChIP procedure, but over-crosslinking can render chromatin resistant to sonication and might obscure NF-κB epitopes, preventing successful antibody capture of the NF-κB-DNA complex. I approached crosslinking optimization with the goal of finding a crosslinking time best suited to deliver high sonication efficiency and high immunoprecipitation yield. To ensure the optimal balance could be found, I tested a broad range of crosslinking durations: 2.5, 5, 7.5, 10, 12.5, and 15 minutes.

First, each of these crosslinking conditions was subjected to the optimized sonication protocol and run on a gel to determine sonication efficiency. All crosslinking conditions yielded high sonication efficiency, although crosslinking times of 12.5 and 15 minutes appeared to render chromatin slightly more resistant to shearing (Fig.15). Since all crosslinking conditions appeared suitable for Bge sonication conditions, I next aimed to test their effect on immunoprecipitation yield. I anticipated that I could identify the optimal crosslinking condition as the condition that yields the most DNA after immunoprecipitation. Then, qPCR could be used to verify that the immunoprecipitated DNA was enriched in κB sequences as compared to an input (non-immunoprecipitated) control.

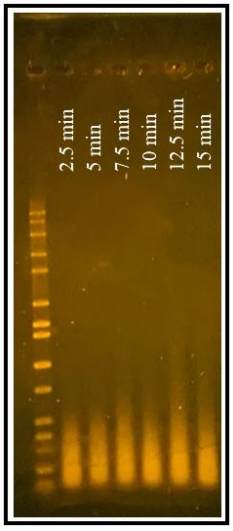


Figure 15. Sonication of various crosslinking conditions. Optimized sonication protocol was used for all crosslinking times. DNA was purified using Zymo ChIP Clean and Concentrator kit. Samples were run with HiLo DNA maker on a 1.25% agarose gel in TAE, then post-stained with 8% GelRed solution for 30 minutes. Gel was imaged using Fotodyne™ FOTO/Phoresis™ UV Transilluminator and Pixel 5a phone camera.

I immunoprecipitated samples from four crosslinking conditions: 2.5, 5, 10, and 12.5 minutes. These conditions were selected to encompass a broad range of crosslinking, while also including the samples with the largest chromatin yields. Unfortunately, after immunoprecipitation with anti-BgRel and Dynabeads™, the yield of purified DNA in all conditions was too low to quantify using a Nanodrop (Appendix III, Supplemental Table 3.). However, qPCR was still performed on all samples as a more sensitive DNA-detection method, enabling comparisons in κB site enrichment across the four crosslinking conditions.

qPCR assays were carried out using two primer sets; the integrin primer set was used as a negative control (amplifies a region with no predicted κB site), and the IκB primer set was used as a positive control (amplifies a site with a hypothesized κB region (Humphries and Harter, 2015)). Additionally, both input and ChIPed DNA samples were included for each crosslinking condition. The input samples served as a positive control to ensure the primer sites were present in the Bge genome, since the qPCR assays were designed and optimized using BB02 snail gDNA. A known concentration of 2.5ng of input DNA was added to each input qPCR, corresponding to a predicted C_T of 28.35 and 28.59 for integrin and IκB reactions, respectively (Fig.16a-b, dashed lines).

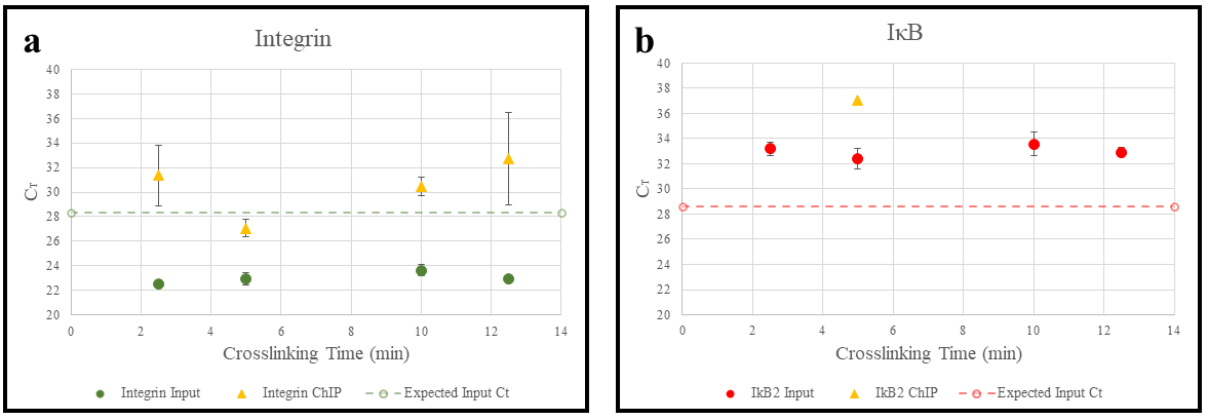


Figure 16. qPCR C_T values for input and ChIPed DNA across four crosslinking conditions. (a) Integrin enrichment. Dark green circles denote input samples, yellow triangles denote ChIPed samples. (b) IκB enrichment. Red circles denote input samples, yellow triangles denote ChIPed samples. 2.5ng of Bge gDNA was added as the input sample for both primer sets, and an unknown amount of DNA was added as the ChIPed sample. Dashed lines represent the expected C_T for 2.5ng of input sample based on BB02 gDNA standard curves (Fig.10). All reactions were performed in triplicate; error bars denote standard deviation. Only one replicate was amplified for the IκB 5-minute-crosslinked ChIP sample, so no error bars are displayed.

Unexpectedly, input C_T values for 2.5ng of Bge DNA averaged 23.00 for integrin (Fig.16b), much lower than expected, implying enrichment of the integrin amplicon in Bge gDNA as compared to BB02 gDNA. Oppositely, input C_T values for 2.5ng of Bge DNA averaged 33.03 for IκB (Fig.16a), implying a lower prevalence of the IκB amplicon in Bge gDNA as compared to BB02 gDNA. Furthermore, the IκB amplicon was amplified in only one sample of twelve total ChIPed replicate qPCR reactions, and this single amplification reaction had a C_T of 37.08 (Fig.16a). Meanwhile, amplification was observed in eleven of twelve total ChIP replicates for integrin-qPCR, with an average C_T of 30.39 (Fig.16b).

To verify that the immunoprecipitation successfully isolated anti-BgRel antibodies, a small portion of the immunoprecipitation eluate was assessed via dot blot. As expected, a strong signal was detected in the eluates from all crosslinking conditions, but not in the elution buffer-only control (Fig.17).

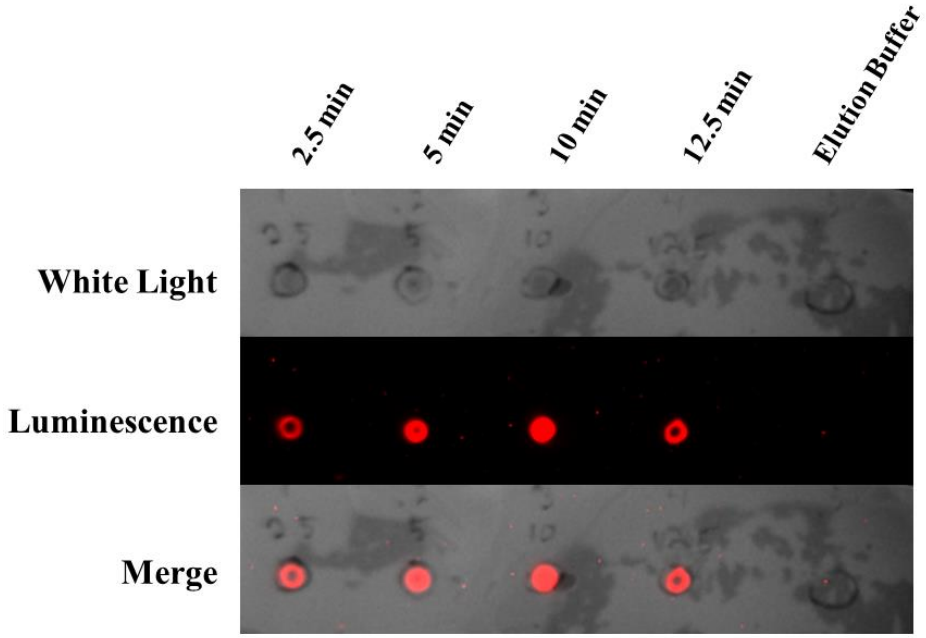


Figure 17. Dot blot of immunoprecipitation eluate. Eluate was probed using HRP-conjugated anti-rabbit IgG. Membrane was imaged using UVP ChemiDoc-It®2 Imager. Images overlaid and falsely colored using ImageJ.

Antibody Validation

Concurrent with the ChIP optimization process, primary validation of anti-BgRel and anti-BgRelish was attempted by performing western blots against Bge cell samples. Bge cell lysate probed with anti-BgRelish displayed bands at 50kDa (the expected size of active BgRelish) which made up >50% of the total signal (74.4%, Fig.18b). However, Bge cell lysate probed with anti-BgRel displayed two distinct bands at 50 and 75kDa. The expected size of BgRel is 75kDa, but the band at 50kDa makes up >50% of the total signal (81.6%, Fig.18a).

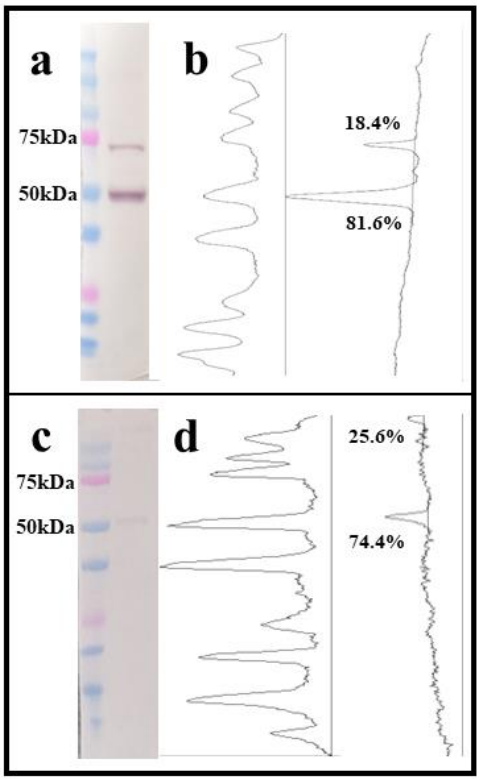


Figure 18. Quantified western blots for B. glabrata NF-kB proteins. (a) Anti-BgRel blot and (b) quantification. Bge cells were lysed by resuspending a cell pellet in 1x Laemmli buffer. (c) Anti-BgRelish blot and (d) quantification. Bge cells were lysed using REAP procedure (see Appendix I), and whole cell extract was analyzed. Blot signal was quantified using ImageJ by measuring the area under the curve of each peak (marked off with a line) and dividing each peak's measurement by the sum of all peaks to obtain each peak's percentage of total blot signal.

Discussion

Optimal Bge ChIP Conditions

Here I present the first chromatin immunoprecipitation procedure for Bge cells, and (to the best of my knowledge) the first ChIP procedure in any mollusk. While this procedure requires further optimization, I have successfully developed a set of sonication conditions to consistently shear Bge gDNA to the 100-500bp fragments suitable for ChIP. I have also made progress towards identifying the optimal crosslinking conditions and have partially validated an anti-BgRelish antibody for use in ChIP. Additionally, I have developed qPCR assays which can be used to validate the ChIP procedure by identifying relative enrichment of ChIPed DNA in DNA sequences with or without a predicted *B. glabrata* κB sequence.

Sonication and Crosslinking

Optimal shearing of Bge gDNA was obtained via Bioruptor® Standard sonication in TPX tubes for 30 cycles of 30 seconds ON, 60 seconds OFF in sonication buffer supplemented with 0.2% SDS. In particular, the use of TPX tubes and supplementation of the sonication buffer with 0.2% SDS was found to be critical for achieving efficient sonication (Figs.12-14). While standard microtubes are composed of polypropylene, TPX tubes are composed of polymethylpentene, a more rigid plastic. The more flexible polypropylene tubes were partially absorbing the ultrasonic frequencies emitted by the Bioruptor® (Sen et al., 2021), which prevented initial Bge chromatin trials from experiencing the full effect of sonication (Fig.12).

Now using the appropriate sonication tubes, I next discovered there is not much difference in sonication efficiency between medium and high power (Fig.13a). At 30 cycles (necessary to generate smaller fragments), ImageJ analysis indicated that medium and high power resulted in an almost identical percentage of fragments within the 100-500bp region (Fig.13b). Since these power setting experiments were conducted using purified DNA, I chose to use the high power setting for later Bge chromatin

samples, since Bge chromatin contains additional proteins and tightly packed DNA. These extra factors are likely to absorb some of the sound waves and decrease sonication efficiency when working with Bge nuclear lysates.

Indeed, sonication of Bge chromatin in TPX tubes at high power did not result in as efficient sonication as was observed during the previous experiments with purified DNA (Fig.14a). However, adding SDS (an anionic detergent) to the sonication buffer greatly increased the sonication efficiency (Fig.14b). SDS disrupts molecular interactions, and so can "loosen" chromatin to make it more susceptible to sonication (Kidder et al., 2011). However, because SDS interferes with protein interactions, excess SDS can complicate downstream immunoprecipitation and DNA purification procedures. Thus, it is desirable to use a low concentration of SDS. Our results indicate 0.2% SDS is sufficient to greatly improve sonication efficiency while, surprisingly, 0.5% SDS appears to decrease sonication efficiency (Fig.14b). This lowered efficacy for 0.5% SDS might be because SDS in both 0.2% and 0.5% conditions precipitated in the ice-cold sonication buffer. It is possible that these SDS aggregates absorbed some of the ultrasonic frequencies, interfering with complete sonication of the sample. Regardless, I proceeded with 0.2% SDS supplementation both to maximize sonication efficiency and to minimize SDS concentration for downstream procedures.

After determining the optimal sonication conditions, I next sought to optimize crosslinking conditions by testing crosslinking times from 2.5 – 15 minutes. Sufficient crosslinking is necessary to ensure that NF-κB remains bound to its DNA binding sequence during the entire ChIP procedure, but excessive crosslinking can make chromatin resistant to sonication and might mask NF-κB epitopes, preventing its immunoprecipitation. As expected, with increasing crosslinking durations, sonication efficiency decreased, yet all tested crosslinking conditions led to satisfactory sonication (Fig.15). Thus, I proceeded to test the immunoprecipitation efficiency of these crosslinking conditions.

Immunoprecipitation was performed on 4.5μg of chromatin from each crosslinking condition to determine which condition would result in the highest final DNA yield. Unfortunately, very little DNA was isolated from any crosslinking condition (Appendix III, Supplemental Table 3), and none of the

ChIPed DNA appeared significantly enriched in the κB region upstream of IκB (Fig.16). Of twelve total reactions (four crosslinking conditions performed in triplicate), only one qPCR reaction amplified the positive control IκB κB site. The single amplification reaction had a C_T of 37 (Fig.16), indicating very little of this sequence was present in the reaction. Because almost no κB sequence DNA was isolated, the efficacy of these crosslinking conditions could not be compared. Further optimization of the immunoprecipitation and/or sample preparation procedure is required to increase the yield of immunoprecipitated DNA and ensure the differences in crosslinking durations on immunoprecipitation efficacy can be detected. Possible explanations for the low observed yield are discussed below.

Troubleshooting Immunoprecipitation

To ensure the Dynabeads™ were effectively capturing the anti-BgRel antibody, a dot blot was performed on the ChIP eluate to probe for the anti-BgRel antibody. While no quantification was performed, adding just 5% of the total eluate resulted in a very strong signal in all immunoprecipitations (Fig.17). This indicates that the Dynabeads™ Protein A can successfully bind and retain the anti-BgRel antibody during all immunoprecipitation wash steps, and anti-BgRel can be successfully eluted from the Dynabeads™. Thus, the low ChIPed DNA yield I observed is due to a problem in complex formation between anti-BgRel, BgRel, and/or κB DNA sequences.

One possibility is that anti-BgRel is not a suitable antibody for ChIP. The epitope recognized by anti-BgRel might be hidden within the 3-D structure of BgRel or by the BgRel-DNA complex, preventing anti-BgRel from isolating BgRel during immunoprecipitation. Anti-BgRel was generated against a very short peptide sequence, and if that sequence is hidden in the chromatin environment, anti-BgRel will be unable to bind. If this is the problem, optimizing Bge cell ChIP will require the use of an alternate antibody. Although anti-BgRelish was also generated against a short peptide sequence, if the epitope it recognizes is more exposed *in vivo*, then anti-BgRelish might be suitable for ChIP. If neither antibody appears compatible with ChIP, it will be necessary to find a new antibody. Generating a polyclonal antibody (raised against the complete, folded proteins BgRel or BgRelish) could result in a better ChIP

signal, as the polyclonal antibodies could bind at multiple locations on the NF-κB proteins, reducing the likelihood that the antibodies fail due to one masked epitope.

Although not a complete explanation for the low DNA yield, the sonication conditions could account for the low IκB qPCR signal. It may be that the IκB κB site is preferentially sheared during sonication, and/or the sonication process is degrading NF-κB proteins. Regions of chromatin can be categorized as heterochromatin or euchromatin based on the availability of the DNA regions to transcriptional proteins. Heterochromatin comprises genomic regions that are tightly packed and whose genes are mostly inaccessible to proteins. Euchromatin comprises the regions of DNA that are open and available for transcriptional proteins to bind. Regions of euchromatin are more likely to be sheared during sonication because the DNA in these regions is more accessible. Thus, if the IκB κB site is in an area of euchromatin, it will be more susceptible to shearing during the sonication procedure. Over-sonication is also known to degrade proteins, especially high molecular weight proteins (Pchelintsev et al., 2016). The elevated temperatures and stress of sonication might be disrupting BgRel's structure and leading to its degradation. Potential sonication issues may be ameliorated by reducing sonication power (to medium or low) and/or sonication cycles.

Another likely explanation for the low ChIPed DNA yield (that would not be indicative of any technical failures in the ChIP procedure) is that these samples are from Bge cells under normal conditions – cells were not exposed to any immune challenges. Although I hypothesize that some BgRel nuclear shuttling occurs under basal conditions, the concentration and activity of BgRel in the nucleus is expected to be very low. Thus, perhaps there was not enough BgRel to detect with ChIP performed on only 4.5µg of chromatin. If the amount of starting chromatin is increased (by increasing the number of cells used), the (potentially) low levels of BgRel binding to κB sequences in the nucleus might become detectable. Alternatively, if Bge cells are exposed to an immune stimulus, this might activate BgRel to translocate to the nucleus and bind its κB sequences, and therefore enable the detection of BgRel-DNA complexes with the amount of chromatin that was prepared in this study.

Finally, an even simpler explanation is that BgRel does not bind DNA sequences *in vivo*, or at least does not bind the κB sequences upstream of IκB. To address this, I am optimizing an additional primer set as an alternate positive control (MAPK p38, Appendix I) for BgRel, and I also plan to perform this assay with BgRelish using an additional positive control primer set (C3-1, Appendix I). We do not know the functional activity of any *B. glabrata* NF-κB proteins, but by testing different possible binding sites I hope to identify at least one *in vivo* target that can be used as a readout of immunoprecipitation quality. However, if no binding sequences can be identified for BgRel or BgRelish, the optimization of this ChIP assay can be completed using other *B. glabrata* transcription factors.

qPCR and the Bge Genome

As previously discussed, the ChIP-qPCR results for the IκB primer set suggest a failure of the immunoprecipitation procedure. However, the input-qPCR results for both IκB and integrin primer sets suggest an important genetic difference between the BB02 and Bge genomes.

Each input sample for qPCR contained 2.5ng of Bge gDNA, yet none of the observed input C_T values matched the predicted C_T values for 2.5ng of DNA based on the BB02 gDNA standard curves. Integrin input samples had an average C_T of 23.00, lower than the predicted C_T of 28.35. IκB input samples had an average C_T of 33.03, higher than the predicted C_T of 28.59. Based on the observed C_T values, the BB02 standard curve would have predicted the integrin input samples to contain an average of 103.5ng of DNA per reaction, and the IκB input samples to contain an average of 0.0064ng of DNA per reaction. Clearly, neither of these predictions is correct, as the same known DNA quantity (2.5ng) was added to each reaction. This mismatch is very likely a result of the significant differences between the Bge genome and the BB02 genome.

The BB02 genome is 0.915 gigabases (Gb) in length, while the Bge genome is approximately 38.7Gb. If 2.5ng of BB02 gDNA were added to each reaction, this would correspond to 2,663 total genomes (NEBioCalculator, NEB, Ipswich, MA, U.S.A.), and 2,663 total copies of both the integrin and IκB amplicons (since these primer sets were designed to be specific to a single BB02 genomic locus).

However, 2.5ng of Bge DNA would correspond to only 62.9 total genomes, so only 62.9 total copies of both integrin and IκB (if both are present only once in the Bge genome). For the IκB primer set, this difference in genome size may explain the high C_T values which were observed: the average input C_T for IκB reactions was 33.03, corresponding to approximately 6.8 copies per reaction, on average. While this is about a ten-fold difference from the actual 62.9 copies that were added, quantification is less reliable at these very low copy numbers, and the sonication process may have sometimes sheared the DNA at the IκB amplicon, resulting in a lower observed copy number than would have been present in an intact Bge genome. This genome size disconnect can be easily solved by creating a new standard curve using Bge gDNA instead of BB02 gDNA and considering that more Bge chromatin must be collected for future qPCR reactions to increase template copy numbers.

While the high C_T values for IκB input reactions may be explained by the Bge genome's large size relative to the BB02 genome, this difference does not explain the integrin input C_T results. Opposite the IκB trend, the average observed C_T for integrin input samples was lower than expected: 23.00. This corresponds to 110,260 copies of the integrin amplicon. Since only 62.9 copies of the Bge genome were added (2.5ng), this suggests that each Bge genome has approximately 1,700 copies of the integrin amplicon. One possible explanation for this result is that the integrin amplicon contains a *B. glabrata* κB sequence, and this sequence was successfully immunoprecipitated with BgRel. Integrin is known to function in the immune system of vertebrates, notably by facilitating signaling which can activate immune pathways (Kinashi, 2011). In *B. glabrata* Bge cells, Humphries et al. have even suggested a possible immune signaling pathway downstream of integrin interactions (Humphries et al., 2001). Therefore, integrin is a plausible NF- κB target, and because LASAGNA 2.0 is designed for human κB sequences, we may have missed a *B. glabrata* κB site when designing integrin primers, resulting in this amplicon not being a true negative control region. However, given that we saw no enrichment at the IκB κB site, which has demonstrated *in vitro* binding to the BgRel RHD (Humphries and Harter, 2015), it seems unlikely that we would observe such a dramatic enrichment in an integrin upstream region with no similar enrichment in the IκB upstream region.

Regardless of whether the integrin amplicon contains a functional κB sequence, the cause of these huge copy numbers might be the eccentric Bge genome. Bge cells have been continuously cultured for the past nearly 50 years, during which time the Bge genome has changed dramatically (Odoemelam et al., 2008, Wheeler et al., 2018). Because integrin aids in cell-cell and cell-surface connections (Yamada et al., 2002), it seems particularly likely that any integrin gene duplication events would be selected for in cell culture conditions. When cell lines are maintained in a laboratory setting, cells which can adhere to tissue culture plates are under positive selection, possibly providing an advantage to cells with multiple gene copies of integrin. Alternatively, it is also possible that the integrin amplicon used in this qPCR is in a region of the Bge genome that has experienced replication slippage, leading to massive repeats of this sequence. To interrogate these possibilities, *in silico* exploration of the Bge genome is warranted.

It is also important to note that the Bge genome has been described as a mixed, non-clonal population of highly aneuploid cells (Wheeler et al., 2018). This means that individual cells in a Bge culture plate can have different genomes. The original snail embryos from which Bge was isolated had a chromosome count of 2n = 36, as a diploid organism with 18 distinct chromosome pairs. Yet a 2018 study of the Bge genome found that chromosome counts now range from 57 to 67 chromosomes in each cell, and some (possibly tetraploid) cells have even higher counts, around 120 chromosomes. Because none of these chromosome counts are exact multiples of 18, these cells are considered aneuploid, meaning they do not have the exact same quantity of all their chromosomes (e.g., Fig.19). Practically, this means that some Bge cells might have more copies of a target gene than others, because some cells will have more of that gene's chromosome than others. It also means the sequenced Bge genome is a combination of many significantly different genomes, so it is most useful when considered to be an "average" of many Bge cells. Furthermore, stocks of the Bge cell line are no longer available in the

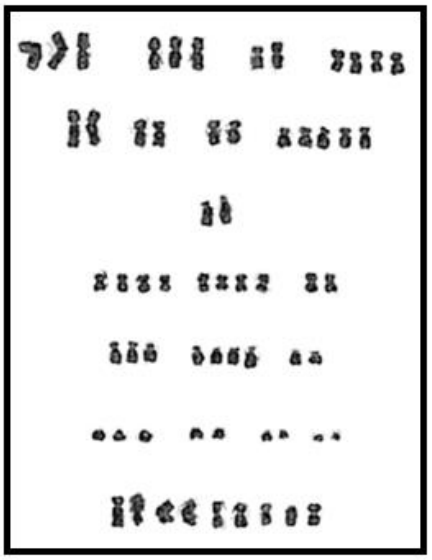


Figure 19. Representative karyotype of a Bge cell. Similar chromosomes are grouped together. From Wheeler, et al., 2018.

ATCC collection, and common practice for Bge cell culture is to keep a continuous culture, as the cells are difficult to revive from frozen stocks. This has resulted in the development of different cell populations in different labs employing slightly different culture conditions (Odoemelam et al., 2008, Wheeler et al., 2018). The Bge cell line used in this study was generously donated from the lab of M.G. Castillo, which sequenced their Bge population in 2018 and termed it "Bge3" (Wheeler et al., 2018). Thus, the cell line used in this study is as close as possible to the available Bge genome, but 5 years of continuous culture and several months of culture in the Humphries lab may have introduced new genomic differences.

Clearly, the Bge genome has some undesirable properties for its use as a model for *B. glabrata* snails. While these genomic abnormalities complicate the process of ChIP assay design and the translation of knowledge from Bge to snail systems, the Bge cell line remains the only Lophotrochozoan cell line, and this continues to make it an invaluable tool. Cell lines are ideally suited for optimizing new procedures, such as the ChIP assay described here, because large quantities of cellular material can be easily obtained in a short period of time. Having more cells available allows for more experimental freedom: more variables can be tested, greater samples sizes can be used, and the cellular material is easier to work with. After thoroughly optimizing a procedure with Bge cells, an improved protocol might then be run using snail tissue samples, which are a more limited resource. Additionally, the data obtained from Bge cells while optimizing can be a useful initial screen to be later verified in snail systems.

Antibody Validation

Finally, for a ChIP study to be successful, it is imperative that the antibody used is highly specific to the target protein and can readily bind to the target's *in vivo*, folded conformation. Anti-BgRel and anti-BgRelish are custom antibodies, and so have not been commercially validated for use in ChIP studies. To confirm their suitability for ChIP, we have subjected these antibodies to primary validation via western blots, as per the ENCODE guidelines (Landt et al., 2012). Secondary validation will be performed by submitting ChIP samples for NGS (after full optimization of the protocol) and verifying

that the sequencing results include κB sequences similar to those found in other organisms, and/or the hypothesized κB sites upstream of IκB and MAPK p38.

The present study has described primary validation experiments for both anti-BgRelish and anti-BgRel. Primary validation for anti-BgRelish was successful, as Bge cell lysates probed with anti-BgRelish produced a 50kDa band (expected size of BgRelish) which made up 74.4% of the total signal. However, anti-BgRel did not pass primary validation, as Bge cell lysates probed with anti-BgRel produced two strong bands at 75kDa (18.4%) and 50kDa (81.6%). The expected size of BgRel is 75kDa, so the presence of the 50kDa band is concerning. To ensure the 50kDa band is not an artifact created by non-specific binding of the secondary antibody, control experiments were performed using the secondary antibody without adding primary antibody. However, no bands were observed in the secondary-only negative control (Appendix II, Supplemental Fig.5), indicating the 50kDa band is being recognized by anti-BgRel.

Interestingly, at the time primary validation westerns were performed, I began using a new batch of cell culture media which was found to kill the Bge cells. It is possible that these two bands for the BgRel protein are a result of BgRel cleavage during Bge cell apoptosis. The human RelA (p65) NF-κB protein has been shown to undergo caspase-3-mediated cleavage of 97 amino acids at its N terminus (Wier et al., 2016) during apoptosis. Anti-BgRel recognizes the C terminus of BgRel, so a similar cleavage could result in the shift from 75 to 50kDa. Future experiments with healthy Bge cells compared to Bge cells incubated with caspase 3 or an apoptosis-inducing factor can test this hypothesis and may yet lead to successful primary validation of anti-BgRel.

Conclusions and Future Directions

This study has determined the optimal sonication conditions for Bge ChIP to be: Bioruptor® Standard sonication for 30 cycles of 30 seconds ON, 60 seconds OFF at high power, with sonication buffer supplemented with 0.2% SDS. I have also determined that crosslinking times from 2.5 – 15

minutes with 1% formaldehyde do not interfere with these sonication conditions, although I do not yet know which of these crosslinking durations is optimal for immunoprecipitation. Additionally, I developed a qPCR assay to test enrichment at the κB site upstream of *B. glabrata* IκB and the upstream region of *B. glabrata* integrin, although this assay may be limited to use with snail tissue samples.

Next steps towards optimizing the Bge cell ChIP procedure will focus on improving yield from the immunoprecipitation, which might involve increasing the number of cells used per experiment, optimizing with immune-challenged cells, modifying the antibody used, and adjusting the immunoprecipitation incubations. Once a reliable immunoprecipitation protocol is established, I will again be able to test a range of crosslinking conditions to determine the optimal crosslinking duration. Finally, the described qPCR assays will need to be adjusted for use with the Bge genome. New standard curves will be needed, and primer sets should be BLASTed against the Bge3 genome to verify their specificity within the Bge cell genome.

Once an optimized Bge cell ChIP procedure is determined, this assay can be used to probe the *in vivo* activity of *B. glabrata* NF-κB proteins. ChIP-qPCR can be used to determine relative binding activity of BgRel or BgRelish under different conditions, such as basal conditions, stimulation with *S. mansoni* larval transformation product, peptidoglycan, LPS, or other immune/stress signals. If the same quantity of chromatin is loaded into the immunoprecipitation for each condition, differences in κB site yield in the eluted DNA can be compared to suggest relative NF-κB binding activity. ChIP samples can also be used to investigate *in vivo* heterodimer formation between BgRel and BgRelish. Heterodimers of Rel-Relish (or p65-p50) are the most common forms of NF-κB in human cells (Gilmore, 2006), but we do not know if BgRel and BgRelish also follow this pattern. To investigate, we can perform ChIP to immunoprecipitate BgRel, then perform a western blot on the ChIP eluate to probe for BgRelish. If the western blot identifies BgRelish, this indicates that a heterodimer forms between BgRel and BgRelish. The assay can also be performed with the opposite antibodies, using anti-BgRelish for ChIP and anti-BgRel for the western.

Additionally, ChIP-seq can be performed to determine the binding sequences of BgRel and BgRelish under various conditions. Determining these binding sequences can indicate which genes BgRel and BgRelish might be regulating, and if those genes change in response to different immune stimuli. Knowing *B. glabrata* NF-κB binding sites also allows for comparisons between *B. glabrata*, *Drosophila*, human, and *Aiptasia* binding sequences, which may provide insights into the evolution of NF-κB in Bilateria organisms. At minimum, ChIP-seq will provide a wealth of information which will spawn future studies to better understand *B. glabrata* gene regulation and can provide insight into the activity of *B. glabrata* NF-κB proteins in Bge cells.

Finally, a complete and optimized Bge cell ChIP assay for NF-κB proteins can be easily modified to suit other *B. glabrata* proteins, requiring only that new antibodies are validated, and new control qPCR primer sets are designed. ChIP is a powerful tool for investigating protein-DNA interactions, and our optimization of this assay will soon make this tool readily available for all *B. glabrata* researchers. Future ChIP experiments in *B. glabrata* may focus on any transcription factors or histone proteins, which promises to further our understanding of gene regulation in *B. glabrata* and similar molluscan systems.

Acknowledgments

I would not be the scientist I am today without the generous support and mentorship from the Lawrence community. First and foremost, thank you to Judith Humphries for introducing me to *B. glabrata* research and giving me the freedom to explore in the lab. I have learned so much from my time in your lab, and I am a stronger and more confident researcher because of those experiences and your mentorship. Also, thank you for all the coffee and chocolates (and proofreading papers right up to the deadline), and always being an advocate for my well-being. You have been an incredibly positive presence at Lawrence, and I'm so grateful I was able to spend so much time with you in your lab.

Thank you to Eric Lewellyn for giving me my first science gig at Lawrence and sharing excellent research advice with me ever since. You always have the answers to my questions, and you have great experimental design ideas for every scenario, many of which I've used in this project. I've also relied on my skills from the biology stockroom to work effectively on this project, so thank you for always having the time to teach. Also, thank you to Kim Dickson for helping me understand ChIP and sharing your protocols, experience, and equipment – I could not have done ChIP without you!

Thank you to Samantha George, Beth De Stasio, and Antoinette Powell for reading this thesis and overseeing my honors committee evolution. Special thanks to Professor George for taking on this science project; I'm very happy you could join this committee for one of my final Lawrence projects, because you have also been a hugely positive force in my Lawrentian life these past four years.

Thanks to all the members of the Humphries lab, Meghan Raebel, and the biology stockroom team for sharing/supplying reagents and helping me with troubleshooting. A special thank you to Ben Glazer for listening to all my experimental successes and failures, and always being willing to help me brainstorm. I've learned a lot from you and had fun doing it!

Finally, I am extremely grateful for the George '51 and Marjorie '44 Chandler Senior Experience grant, which funded this research project. Their generous funding enabled me to take on this exciting and ambitious project, which has been invaluable in my growth as an independent researcher and has taught me so much about ChIP!

References

Adema, C.M., 2015. Fibrinogen-Related Proteins (FREPs) in Mollusks, in: Hsu, E., Du Pasquier, L. (Eds.), Pathogen-Host Interactions: Antigenic Variation v. Somatic Adaptations, Results and Problems in Cell Differentiation. Springer International Publishing, Cham, pp. 111–129. https://doi.org/10.1007/978-3-319-20819-0_5

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J., 1990. Basic local alignment search tool. J Mol Biol 215, 403–410. https://doi.org/10.1016/S0022-2836(05)80360-2

Annala, M., Laurila, K., Lähdesmäki, H., Nykter, M., 2011. A Linear Model for Transcription Factor Binding Affinity Prediction in Protein Binding Microarrays. PLOS ONE 6, e20059. https://doi.org/10.1371/journal.pone.0020059

Aula, O.P., McManus, D.P., Jones, M.K., Gordon, C.A., 2021. Schistosomiasis with a Focus on Africa. Trop. Med. Infect. Dis. 6, 109. https://doi.org/10.3390/tropicalmed6030109

Ayyar, S., Pistillo, D., Calleja, M., Brookfield, A., Gittins, K., Goldstone, C., Simpson, P., 2007. NF-κB/Rel-Mediated Regulation of the Neural Fate in *Drosophila*. PLOS ONE 2, e1178. https://doi.org/10.1371/journal.pone.0001178

Beinke, S., Belich, M.P., Ley, S.C., 2002. The Death Domain of NF-κB1 p105 Is Essential for Signal-induced p105 Proteolysis. J. Biol. Chem. 277, 24162–24168. https://doi.org/10.1074/jbc.M201576200

Castillo, M.G., Humphries, J.E., Mourão, M.M., Marquez, J., Gonzalez, A., Montelongo, C.E., 2020. *Biomphalaria glabrata* immunity: Post-genome advances. Dev. Comp. Immunol. 104, 103557. https://doi.org/10.1016/j.dci.2019.103557

CDC - Schistosomiasis [WWW Document], 2021. Cent. Dis. Control Prev. URL https://www.cdc.gov/parasites/schistosomiasis/index.html (accessed 3.6.23).

Coelho, F.S., Rodpai, R., Miller, A., Karinshak, S.E., Mann, V.H., dos Santos Carvalho, O., Caldeira, R.L., de Moraes Mourão, M., Brindley, P.J., Ittiprasert, W., 2020. Diminished adherence of *Biomphalaria glabrata* embryonic cell line to sporocysts of *Schistosoma mansoni* following

programmed knockout of the allograft inflammatory factor. Parasit. Vectors 13, 511. https://doi.org/10.1186/s13071-020-04384-9

Colley, D.G., Bustinduy, A.L., Secor, W.E., King, C.H., 2014. Human schistosomiasis. Lancet 383, 2253–2264. https://doi.org/10.1016/S0140-6736(13)61949-2

Costain, A.H., MacDonald, A.S., Smits, H.H., 2018. Schistosome Egg Migration: Mechanisms, Pathogenesis and Host Immune Responses. Front. Immunol. 9.

Deol, A.K., Fleming, F.M., Calvo-Urbano, B., Walker, M., Bucumi, V., Gnandou, I., Tukahebwa, E.M., Jemu, S., Mwingira, U.J., Alkohlani, A., Traoré, M., Ruberanziza, E., Touré, S., Basáñez, M.-G., French, M.D., Webster, J.P., 2019. Schistosomiasis — Assessing Progress toward the 2020 and 2025 Global Goals. N. Engl. J. Med. 381, 2519–2528. https://doi.org/10.1056/NEJMoa1812165

Dinguirard, N., Cavalcanti, M.G.S., Wu, X.-J., Bickham-Wright, U., Sabat, G., Yoshino, T.P., 2018. Proteomic Analysis of *Biomphalaria glabrata* Hemocytes During in vitro Encapsulation of *Schistosoma mansoni* Sporocysts. Front. Immunol. 9.

Fryer, S.E., Bayne, C.J., 1990. *Schistosoma mansoni* Modulation of Phagocytosis in *Biomphalaria glabrata*. J. Parasitol. 76, 45–52. https://doi.org/10.2307/3282626

Ghosh, G., Ya-Fan Wang, V., Huang, D.-B., Fusco, A., 2012. NF-κB Regulation: Lessons from Structures. Immunol. Rev. 246, 36–58. https://doi.org/10.1111/j.1600-065X.2012.01097.x

Gilmore, T.D., 2021. NF-κB and Human Cancer: What Have We Learned over the Past 35 Years? Biomedicines 9, 889. https://doi.org/10.3390/biomedicines9080889

Gilmore, T.D., 2006. Introduction to NF-κB: players, pathways, perspectives. Oncogene 25, 6680–6684. https://doi.org/10.1038/sj.onc.1209954

Gourbal, B., Pinaud, S., Beckers, G.J.M., Van Der Meer, J.W.M., Conrath, U., Netea, M.G., 2018. Innate immune memory: An evolutionary perspective. Immunol. Rev. 283, 21–40. https://doi.org/10.1111/imr.12647

Hatada, E.N., Nieters, A., Wulczyn, F.G., Naumann, M., Meyer, R., Nucifora, G., McKeithan, T.W., Scheidereit, C., 1992. The ankyrin repeat domains of the NF-kappa B precursor p105 and the

protooncogene bcl-3 act as specific inhibitors of NF-kappa B DNA binding. Proc. Natl. Acad. Sci. 89, 2489–2493. https://doi.org/10.1073/pnas.89.6.2489

Hayden, M., West, A., Ghosh, S., 2006. NF-κB and the immune response. Oncogene 25, 6758–80. https://doi.org/10.1038/sj.onc.1209943

Hoffman, E.A., Frey, B.L., Smith, L.M., Auble, D.T., 2015. Formaldehyde Crosslinking: A Tool for the Study of Chromatin Complexes. J. Biol. Chem. 290, 26404–26411. https://doi.org/10.1074/jbc.R115.651679

Humphries, J., Harter, B., 2015. Identification of nuclear factor kappaB (NF-κB) binding motifs in *Biomphalaria glabrata*. Dev. Comp. Immunol. 53, 366–370. https://doi.org/10.1016/j.dci.2015.08.004

Humphries, J.E., Deneckere, L.E., 2018. Characterization of a Toll-like receptor (TLR) signaling pathway in *Biomphalaria glabrata* and its potential regulation by NF-kappaB. Dev. Comp. Immunol. 86, 118–129. https://doi.org/10.1016/j.dci.2018.05.003

Humphries, J.E., Elizondo, L., Yoshino, T.P., 2001. Protein kinase C regulation of cell spreading in the molluscan *Biomphalaria glabrata* embryonic (*Bge*) cell line. Biochim. Biophys. Acta BBA - Mol. Cell Res. 1540, 243–252. https://doi.org/10.1016/S0167-4889(01)00136-7

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, 554–562. https://doi.org/10.1016/j.dci.2007.09.001

Keller, C.A., Wixom, A.Q., Heuston, E.F., Giardine, B., Hsiung, C.C.-S., Long, M.R., Miller, A., Anderson, S.M., Cockburn, A., Blobel, G.A., Bodine, D.M., Hardison, R.C., 2021. Effects of sheared chromatin length on ChIP-seq quality and sensitivity. G3 GenesGenomesGenetics 11, jkab101. https://doi.org/10.1093/g3journal/jkab101

Kidder, B.L., Hu, G., Zhao, K., 2011. ChIP-Seq: technical considerations for obtaining high-quality data. Nat. Immunol. 12, 918–922. https://doi.org/10.1038/ni.2117

Kinashi, T., 2011. Overview of Integrin Signaling in the Immune System, in: Shimaoka, M. (Ed.),

Integrin and Cell Adhesion Molecules, Methods in Molecular Biology. Humana Press, Totowa, NJ, pp. 261–278. https://doi.org/10.1007/978-1-61779-166-6_17

Kõressaar, T., Lepamets, M., Kaplinski, L., Raime, K., Andreson, R., Remm, M., 2018. Primer3_masker: integrating masking of template sequence with primer design software. Bioinformatics 34, 1937–1938. https://doi.org/10.1093/bioinformatics/bty036

Koressaar, T., Remm, M., 2007. Enhancements and modifications of primer design program Primer3. Bioinformatics 23, 1289–1291. https://doi.org/10.1093/bioinformatics/btm091

Lambert, S.A., Jolma, A., Campitelli, L.F., Das, P.K., Yin, Y., Albu, M., Chen, X., Taipale, J., Hughes, T.R., Weirauch, M.T., 2018. The Human Transcription Factors. Cell 172, 650–665. https://doi.org/10.1016/j.cell.2018.01.029

Landt, S.G., Marinov, G.K., Kundaje, A., Kheradpour, P., Pauli, F., Batzoglou, S., Bernstein, B.E., Bickel, P., Brown, J.B., Cayting, P., Chen, Y., DeSalvo, G., Epstein, C., Fisher-Aylor, K.I., Euskirchen, G., Gerstein, M., Gertz, J., Hartemink, A.J., Hoffman, M.M., Iyer, V.R., Jung, Y.L., Karmakar, S., Kellis, M., Kharchenko, P.V., Li, Q., Liu, T., Liu, X.S., Ma, L., Milosavljevic, A., Myers, R.M., Park, P.J., Pazin, M.J., Perry, M.D., Raha, D., Reddy, T.E., Rozowsky, J., Shoresh, N., Sidow, A., Slattery, M., Stamatoyannopoulos, J.A., Tolstorukov, M.Y., White, K.P., Xi, S., Farnham, P.J., Lieb, J.D., Wold, B.J., Snyder, M., 2012. ChIP-seq guidelines and practices of the ENCODE and modENCODE consortia. Genome Res. 22, 1813–1831. https://doi.org/10.1101/gr.136184.111

Lee, C., Huang, C.-H., 2013. LASAGNA-Search: an integrated web tool for transcription factor binding site search and visualization. BioTechniques 54, 141–153. https://doi.org/10.2144/000113999

Li, H., Gharamah, A.A., Hambrook, J.R., Wu, X., Hanington, P.C., 2022. Single-cell RNA-seq profiling of individual *Biomphalaria glabrata* immune cells with a focus on immunologically relevant transcripts. Immunogenetics 74, 77–98. https://doi.org/10.1007/s00251-021-01236-3

Maramorosch, K., 2012. Invertebrate Tissue Culture: Research Applications. Elsevier.

Marsh, A., 2021. Identification and Localization of NF-κB During the Embryonic Development of the

Schistosome-transmitting Snail *Biomphalaria glabrata*. Lawrence Univ. Honors Proj.

Medzhitov, R., Janeway, C.J., 2000. Innate immune recognition: mechanisms and pathways: Innate immune recognition. Immunol. Rev. 173, 89–97. https://doi.org/10.1034/j.1600-065X.2000.917309.x

Minakhina, S., Steward, R., 2006. Nuclear factor-kappa B pathways in *Drosophila*. Oncogene 25, 6749–6757. https://doi.org/10.1038/sj.onc.1209940

Montgomery, S., 2019. Schistosomiasis - Chapter 4 - 2020 Yellow Book | Travelers' Health | CDC [WWW Document]. Cent. Dis. Control Prev. URL https://wwwnc.cdc.gov/travel/yellowbook/2020/travel-related-infectious-diseases/schistosomiasis (accessed 3.6.23).

Nelwan, M.L., 2019. Schistosomiasis: Life Cycle, Diagnosis, and Control. Curr. Ther. Res. 91, 5–9. https://doi.org/10.1016/j.curtheres.2019.06.001

Netea, M.G., Quintin, J., van der Meer, J.W.M., 2011. Trained Immunity: A Memory for Innate Host Defense. Cell Host Microbe 9, 355–361. https://doi.org/10.1016/j.chom.2011.04.006

Odoemelam, E., Raghavan, N., Miller, A., Bridger, J.M., Knight, M., 2009. Revised karyotyping and gene mapping of the *Biomphalaria glabrata* embryonic (Bge) cell line. International Journal for Parasitology 39, 675–681. https://doi.org/10.1016/j.ijpara.2008.11.011

Oeckinghaus, A., Ghosh, S., 2009. The NF-κB Family of Transcription Factors and Its Regulation. Cold Spring Harb. Perspect. Biol. 1, a000034. https://doi.org/10.1101/cshperspect.a000034

O'Neill, L.A.J., Kaltschmidt, C., 1997. NF-κB: a crucial transcription factor for glial and neuronal cell function. Trends Neurosci. 20, 252–258. https://doi.org/10.1016/S0166-2236(96)01035-1

Park, P.J., 2009. ChIP–seq: advantages and challenges of a maturing technology. Nat. Rev. Genet. 10, 669–680. https://doi.org/10.1038/nrg2641

Pchelintsev, N.A., Adams, P.D., Nelson, D.M., 2016. Critical Parameters for Efficient Sonication and Improved Chromatin Immunoprecipitation of High Molecular Weight Proteins. PLOS ONE 11, e0148023. https://doi.org/10.1371/journal.pone.0148023

Pila, E.A., 2018. The Role of Haemocytes in the Immune Response of *Biomphalaria glabrata* against *Schistosoma mansoni*: Investigations into Haematopoiesis and Pattern Recognition Receptors.

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

Pinaud, S., Portela, J., Duval, D., Nowacki, F.C., Olive, M.-A., Allienne, J.-F., Galinier, R., Dheilly, N.M., Kieffer-Jaquinod, S., Mitta, G., Théron, A., Gourbal, B., 2016. A Shift from Cellular to Humoral Responses Contributes to Innate Immune Memory in the Vector Snail *Biomphalaria glabrata*. PLOS Pathog. 12, e1005361. https://doi.org/10.1371/journal.ppat.1005361

Rinaldi, G., Yan, H., Nacif-Pimenta, R., Matchimakul, P., Bridger, J., Mann, V.H., Smout, M.J., Brindley, P.J., Knight, M., 2015. Cytometric analysis, genetic manipulation and antibiotic selection of the snail embryonic cell line *Bge* from *Biomphalaria glabrata*, the intermediate host of *Schistosoma mansoni*. Int. J. Parasitol. 45, 527–535. https://doi.org/10.1016/j.ijpara.2015.02.012

Schistosomiasis [WWW Document], 2023. World Health Organ. URL https://www.who.int/news-room/fact-sheets/detail/schistosomiasis (accessed 3.6.23).

Schmidt, D., Wilson, M.D., Spyrou, C., Brown, G.D., Hadfield, J., Odom, D.T., 2009. ChIP-seq: Using high-throughput sequencing to discover protein–DNA interactions. Methods 48, 240–248. https://doi.org/10.1016/j.ymeth.2009.03.001

Sen, I., Kavšek, A., Riedel, C.G., 2021. Chromatin Immunoprecipitation and Sequencing (ChIP-seq) Optimized for Application in *Caenorhabditis elegans*. Current Protocols 1, e187. https://doi.org/10.1002/cpz1.187

Sen, R., Baltimore, D., 1986. Inducibility of κ immunoglobulin enhancer-binding protein NF-κB by a posttranslational mechanism. Cell 47, 921–928. https://doi.org/10.1016/0092-8674(86)90807-X

Shih, R.-H., Wang, C.-Y., Yang, C.-M., 2015. NF-kappaB Signaling Pathways in Neurological

Inflammation: A Mini Review. Front. Mol. Neurosci. 8.

Siggers, T., Gilmore, T.D., Barron, B., Penvose, A., 2015. Characterizing the DNA Binding Site Specificity of NF-κB with Protein-Binding Microarrays (PBMs), in: May, M.J. (Ed.), NF-Kappa B, Methods in Molecular Biology. Springer New York, New York, NY, pp. 609–630. https://doi.org/10.1007/978-1-4939-2422-6_36

Sokolow, S.H., Wood, C.L., Jones, I.J., Swartz, S.J., Lopez, M., Hsieh, M.H., Lafferty, K.D., Kuris, A.M., Rickards, C., Leo, G.A.D., 2016. Global Assessment of Schistosomiasis Control Over the Past Century Shows Targeting the Snail Intermediate Host Works Best. PLoS Negl. Trop. Dis. 10, e0004794. https://doi.org/10.1371/journal.pntd.0004794

Suzuki, K., Bose, P., Leong-Quong, R.Y., Fujita, D.J., Riabowol, K., 2010. REAP: A two minute cell fractionation method. BMC Research Notes 3, 294. https://doi.org/10.1186/1756-0500-3-294

Tessmar-Raible, K., 2003. Emerging systems: between vertebrates and arthropods, the Lophotrochozoa. Curr. Opin. Genet. Dev. 13, 331–340. https://doi.org/10.1016/S0959-437X(03)00086-8

Untergasser, A., Cutcutache, I., Koressaar, T., Ye, J., Faircloth, B.C., Remm, M., Rozen, S.G., 2012. Primer3—new capabilities and interfaces. Nucleic Acids Research 40, e115. https://doi.org/10.1093/nar/gks596

Vale, N., Gouveia, M.J., Rinaldi, G., Brindley, P.J., Gärtner, F., Correia da Costa, J.M., 2017. Praziquantel for Schistosomiasis: Single-Drug Metabolism Revisited, Mode of Action, and Resistance. Antimicrob. Agents Chemother. 61, e02582-16. https://doi.org/10.1128/AAC.02582-16

Wardle, F.C., Tan, H., 2015. A ChIP on the shoulder? Chromatin immunoprecipitation and validation strategies for ChIP antibodies. F1000Research 4, 235. https://doi.org/10.12688/f1000research.6719.1

Wheeler, N.J., Dinguirard, N., Marquez, J., Gonzalez, A., Zamanian, M., Yoshino, T.P., Castillo, M.G., 2018. Sequence and structural variation in the genome of the *Biomphalaria glabrata* embryonic (*Bge*) cell line. Parasit. Vectors 11, 496. https://doi.org/10.1186/s13071-018-3059-2

Wier, E.M., Fu, K., Hodgson, A., Sun, X., Wan, F., 2015. Caspase-3 cleaved p65 fragment dampens NF-κB-mediated anti-apoptotic transcription by interfering with the p65/RPS3 interaction. FEBS Letters 589, 3581–3587. https://doi.org/10.1016/j.febslet.2015.10.019

Williams, L.M., Gilmore, T.D., 2020. Looking Down on NF-κB. Mol. Cell. Biol. 40, e00104-20. https://doi.org/10.1128/MCB.00104-20

Yamada, K.M., Pankov, R., Cukierman, E., 2003. Dimensions and dynamics in integrin function. Braz J Med Biol Res 36, 959–966. https://doi.org/10.1590/S0100-879X2003000800001

Ye, J., Coulouris, G., Zaretskaya, I., Cutcutache, I., Rozen, S., Madden, T.L., 2012. Primer-BLAST: a tool to design target-specific primers for polymerase chain reaction. BMC Bioinformatics 13, 134. https://doi.org/10.1186/1471-2105-13-134

Zelck, U.E., Gege, B.E., Schmid, S., 2007. Specific inhibitors of mitogen-activated protein kinase and PI3-K pathways impair immune responses by hemocytes of trematode intermediate host snails. Dev. Comp. Immunol. 31, 321–331. https://doi.org/10.1016/j.dci.2006.06.006

Zhang, S.-M., 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, 1868–1881. https://doi.org/10.1016/j.molimm.2011.05.017

Appendices

Appendix I: Supplemental Methods

Bge Cell Passaging

To passage, cBge medium was removed and cells were rinsed with room temperature Chernin's Balanced Salt Solution (CBSS, 48mM NaCl, 2mM KCl, 0.5mM Na_2HPO_4, 2mM $CaCl_2$, 0.64mM $NaHCO_3$, 6mM glucose, 3mM trehalose, pH 7.2) as needed to remove clumped or dead cells. Cold CBSS was then added to detach cells, and after 10-15 minutes cells were resuspended in the CBSS, then pelleted at 300 rcf for 5 minutes. The cell pellet was resuspended in cBge medium, then aliquoted into new T25 or T75 Falcon tissue culture flasks (Corning).

Nuclear Fractionation

Nuclear- and cytoplasmic-enrichment of Bge cells was achieved using a modification of the Rapid Efficient And Practical (REAP) procedure (Suzuki et al., 2010). Briefly: cells grown in culture dishes were rinsed with sPBS, collected in sPBS with a cell scraper, and pelleted at 2,000 rcf for 10 seconds at room temperature. The cell pellet was resuspended in REAP lysis buffer (sPBS with 0.1% Nonidet P-40 alternative (VWR International) and 1:500 dilution of CalBiochem protease inhibitor cocktail IV). An aliquot was taken as the whole cell extract (WCE) directly into an equal volume of 2x Laemmli sample buffer (Bio-Rad.), while the remaining solution was centrifuged at 2,000 rcf for 30 seconds at room temperature. An aliquot of the supernatant was saved as the cytoplasmic fraction directly into an equal volume of 2x Laemmli buffer, and the pellet was resuspended in REAP lysis buffer. This solution was centrifuged for 20 seconds, and the resulting pellet was resuspended in 1x Laemmli sample buffer for western blotting. WCE and nuclear fractions were sonicated on ice at 10% power with a Branson 450 Digital Sonifier (Branson Ultrasonics, Brookfield, CT, U.S.A.) for 3-5 second bursts totaling 20-30 seconds.

Extended qPCR Design

Previous research in the Humphries lab using EMSAs has confirmed *in vitro* binding of the BgRel RHD to putative κB sites upstream of the *B. glabrata* IκB and mitogen-activated protein kinase (MAPK) p38 genes (Humphries and Harter, 2015). In the same study, BgRel was found not to bind to a predicted κB site upstream of the superoxide dismutase [Cu-Zn]-like (SOD) gene. Thus, qPCR primers were designed to amplify within the 350bp region surrounding these three κB sites, with IκB and MAPK p38 serving as known targets and SOD serving as a negative control for BgRel binding. As of yet, no putative BgRelish binding sites have been examined using EMSAs, but the upstream region of IκB is predicted to contain two BgRelish binding sites, in addition to (and within 10bp of) the BgRel binding site. To increase the likelihood of targeting a true BgRelish binding sequence and a true sequence with no κB sites, I designed additional primer sets targeting a predicted BgRelish binding sequence and an upstream region with no predicted κB sites. These primer sets targeted regions upstream of complement protein C3-1 (C3-1) and integrin, respectively.

κB sites were predicted with LASAGNA 2.0 (Lee and Huang, 2013), using default settings and Matrix-Derived JASPAR CORE Models REL and RELA to predict BgRel binding, and Matrix-Derived JASPAR CORE Models for NFKB1 and LASAGNA-ChIP-Aligned PAZAR Models for NFKB1 to predict BgRelish binding. In addition, LASAGNA-identified binding sites were only considered if they contained a G in position 2 and a C at position 10 (Humphries and Deneckere, 2018).

Primers were designed to amplify a 75-200bp region within the 350bp surrounding known or predicted NF-κB binding sites. The binding sites targeted were upstream of IκB (Rel and Relish binding sites), MAPK p38 (Rel binding site), and C3-1 genes (Relish binding site). Negative control primers were also designed against the upstream regions of integrin and SOD genes.

Primers for MAPK p38, C3-1, and SOD were designed using the NCBI Primer tool (Ye et al., 2012). All primers were BLASTed to check their specificity within the *B. glabrata* genome, and self-dimer, hairpin, and heterodimer formation was checked using IDT OligoAnalyzer.

Supplemental Table 1. Genome reference information for target κB sequences and negative control regions for qPCR amplification. NCBI Ref Seq or GenBank IDs are provided, as these sequences were used to find each protein's VectorBase locus. Amplicon location provides the location within the VectorBase scaffold, and +/- indicates positive strand / reverse complement. Although not all amplicons include the predicted κB site, all are located within 100bp of the predicted κB site.

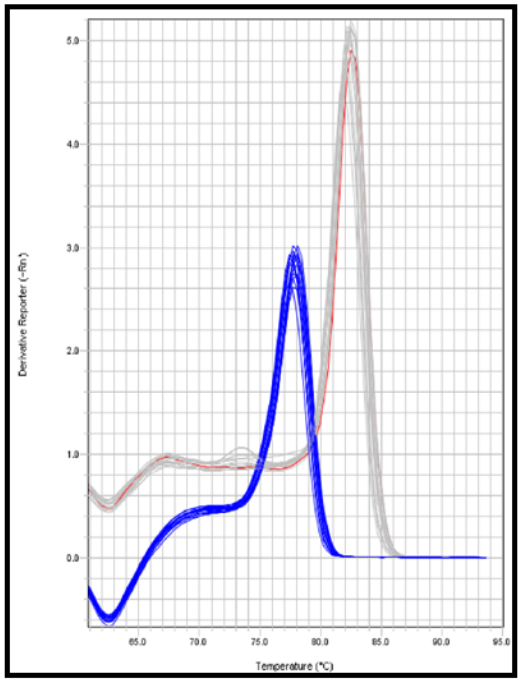
	Protein Ref Seq / GenBank ID	**Scaffold Region (VectorBase Identifier)**	**Amplicon Location**	**Predicted κB Site**
IκB	NP_001298198.1	KE719853	+391363:391515	AGGGCCTTTCC
MAPK p38	AAY89301.1	KE714862	-32157:32240	CGGATTTTCC
C3-1	QEQ12614.1	KE720537	-282284:82366	GGAATTTCTC
SOD	NP_001298216.1	KE706421	+612433:612507	-
Integrin	NP_001298213.1	KE711506	+24456:24633	-

Supplemental Table 2. Primer sequences with amplicon length, T_m, and GC content. *Self-dimer formation ΔG= -7.05.

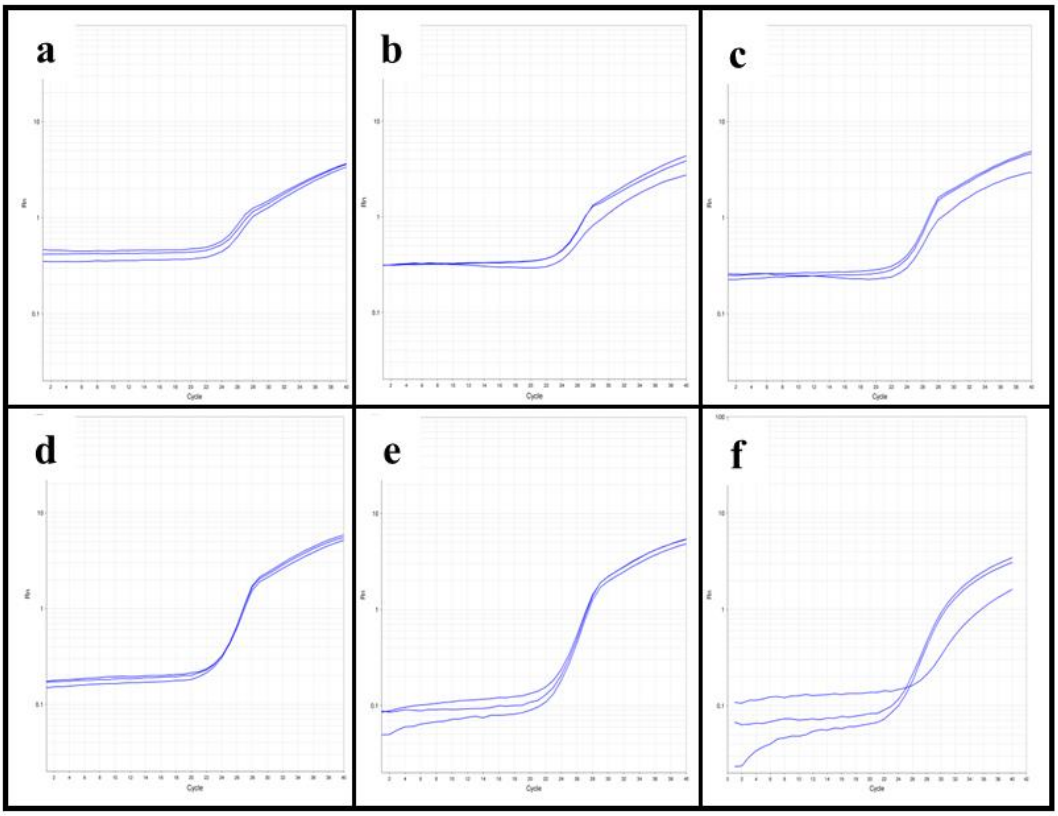
	Sequence (5' - 3')	Amplicon Length (bp)	T_m (°C)	%GC
IκB fwd	**CCCTTCAACTGCGATTCAAGTG**	153	56.5	50
IκB rev	**GGTTTTCCCATACGAATAGACAAA**		53.2	37.5
MAPK p38 fwd	***TATTGCCCTTGCACAATGCG**	84	56.7	50
MAPK p38 rev	**TTGCTAAGGTTTTCCTGTCCT**		56.7	50
C3-1 fwd	**TACTTCTGCCGAGCGTTCTGC**	83	59.7	57.1
C3-1 rev	**GAATGTAACGTCGCCCTCCG**		58.2	60
SOD fwd	**AATTCCTGGCAAGTCTGTCGT**	75	56.7	47.6

SOD rev	**TGCTGGGTCTTATGGACAAC**		54.8	50
Integrin fwd	**TGTGGTGCCAATACTTCCGT**	178	56.8	50
Integrin rev	**GCCAATATGACACTTGACAGGG**		55.9	50

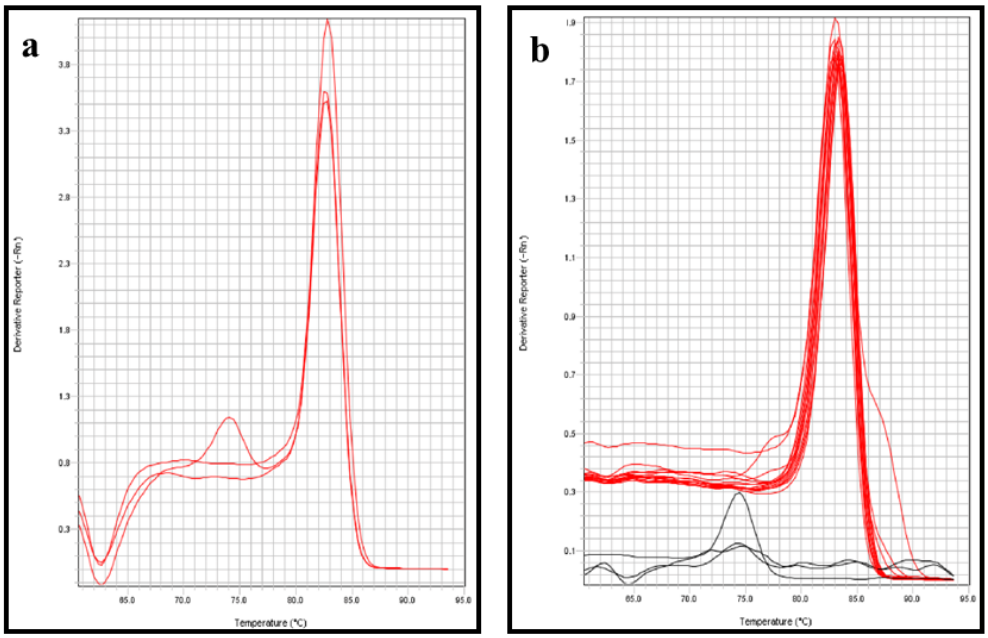
Appendix II: Supplemental Figures



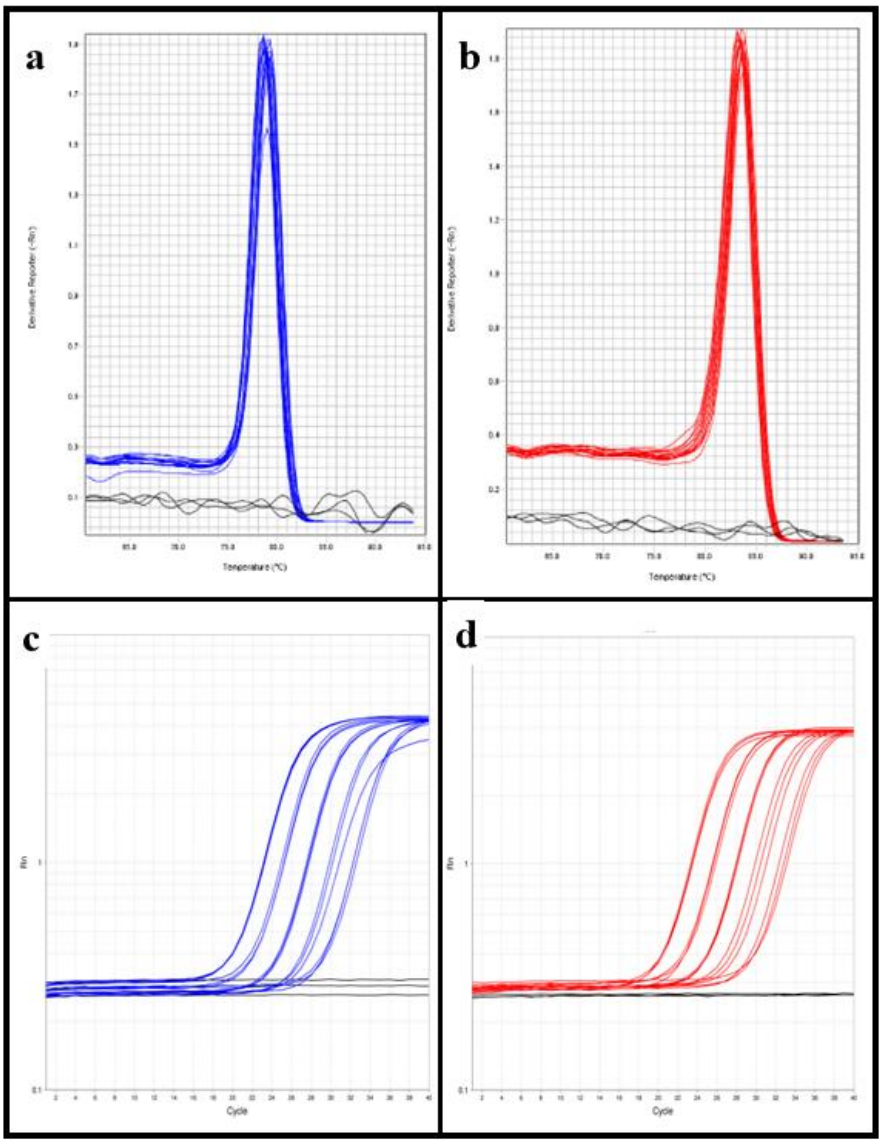
Supplemental Figure 1. Integrin and IκB melt curve for initial primer matrix. Reactions were performed with 100ng BB02 gDNA and G2 qPCR Master Mix. No replicates were performed for each primer concentration combination. Integrin melt curve is in blue, and IκB melt curve is in grey, except for the chosen primer concentration (forward = 200nM, reverse = 800nM) which is in red. Figures were generated using Applied Biosystems 7500 software.



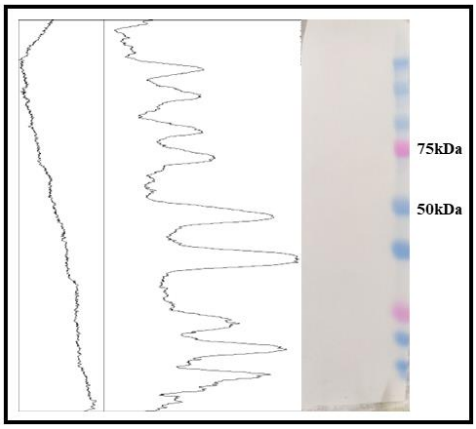
Supplemental Figure 2. Integrin annealing optimization with G2 qPCR Master Mix. (a) 55°C, (b) 57.5°C, (c) 60°C, (d) 62.5°C, (e) 65°C, (f) 67.5°C. Each temperature was performed in triplicate with 4ng BB02 gDNA. Images generated using Applied Biosystems 7500 software.



Supplemental Figure 3. Melt curves for IκB optimization. (a) G2 qPCR Master Mix, annealing temperature = 55°C. (b) SsoAdvanced Universal SYBR Green Supermix, annealing temperature = 65°C. Standard curve reactions in red, NTC reactions in black. Images generated using Applied Biosystems 7500 software.



Supplemental Figure 4. Integrin and IκB melt and amplification curves for final assay conditions. (a) Integrin melt curve, (b) IκB melt curve. (c) Integrin amplification plot, (d) IκB amplification plot. Images generated using Applied Biosystems 7500 software.



Supplemental Figure 5. Secondary-only control for anti-BgRel primary validation. Bge cells were lysed by resuspending a cell pellet in 1x Laemmli buffer. Blot signal analyzed using ImageJ.

Appendix III: Input and ChIPed DNA Quality

Supplemental Table 3. Nanodrop quantification for input and ChIPed DNA from crosslinking optimization. *No DNA curve observed on nanodrop; quantification and ratios are unreliable.

	Concentration (ng/uL)	**260/280 Ratio**	**260/230 Ratio**
2.5 min Crosslinking (Input)	53.6ng/uL	1.85	2.09
5 min Crosslinking (Input)	50.8ng/uL	1.82	2.10
7.5 min Crosslinking (Input)	58.4ng/uL	1.84	2.16
10 min Crosslinking (Input)	34.6ng/uL	1.84	2.10
12.5 min Crosslinking (Input)	42.3ng/uL	1.83	1.77
15 min Crosslinking (Input)	33.2ng/uL	1.83	2.02
2.5 min Crosslinking (ChIP)	*3.2ng/uL	*1.40	*0.52
5 min Crosslinking (ChIP)	*2.9ng/uL	*1.62	*0.50
10 min Crosslinking (ChIP)	*3.2ng/uL	*1.59	*0.44
12.5 min Crosslinking (ChIP)	*1.2ng/uL	*1.98	*0.37