IDENTIFICATION OF GENE EXPRESSION ELEMENTS IN *HISTOMONAS MELEAGRIDIS* USING SPLINKERETTE PCR, A VARIATION OF LIGATED ADAPTOR PCR

by

ELIZABETH CAROLYN LYNN

(Under the Direction of Robert B. Beckstead)

ABSTRACT

Histomonas meleagridis is the causative agent of blackhead disease in gallinaceous birds, but little genetic information exists for this organism. The complete genome for this protozoan is unsequenced. The only available sequence information is for coding portions of genes. No information is available for expression elements. In this study, we demonstrate that splinkerette PCR procedure, a variation of ligated adaptor PCR, can be used to identify regions upstream and downstream of known coding sequences. Using this technique, we isolated the upstream sequence of 2 beta-tubulin genes. With sequence analysis of their upstream regions, we identified their upstream intergenic regions and 2 different open reading frames. The intergenic region contained putative polyadenylation and cleavage signals and initiator elements. Our research demonstrates that the use of splinkerette PCR is a valuable tool to identify regions of unknown DNA that are 5' or 3' to known sequences in parasites whose genomes remain unsequenced. The identification of the expression elements of *H. meleagridis* will provide tools for future studies on its gene expression.

INDEX WORDS: *Histomonas meleagridis*, molecular characterization, beta-tubulin, splinkerette PCR

IDENTIFICATION OF GENE EXPRESSION ELEMENTS IN *HISTOMONAS MELEAGRIDIS* USING SPLINKERETTE PCR, A VARIATION OF LIGATED ADAPTOR PCR

by

ELIZABETH CAROLYN LYNN

AS, Abraham Baldwin Agricultural College, 2007

BSA, University of Georgia, 2009

A Thesis Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment of the Requirements for the Degree

MASTER OF SCIENCE

ATHENS, GEORGIA

2011

IDENTIFICATION OF GENE EXPRESSION ELEMENTS IN *HISTOMONAS MELEAGRIDIS* USING SPLINKERETTE PCR, A VARIATION OF LIGATED ADAPTOR PCR

by

ELIZABETH CAROLYN LYNN

Major Professor: Robert B. Beckstead

Committee: Larry R. McDougald

Andrew R. Moorhead

Electronic Version Approved:

Maureen Grasso

Dean of the Graduate School

The University of Georgia

August 2011

DEDICATION

I would like to dedicate this thesis to my parents, Dr. Byron (Steve) and Mrs. Miriam Lynn, who have held my hands and supported my goals through my entire life. Thank you for instilling the drive to be an overachiever.

Secondly, this thesis will be dedicated to Dr. Robert B. Beckstead. Thank you for the opportunity to continue my education with you. I could not have done this without your faith in me.

Finally, I would like to dedicate this thesis to you, the reader. May you gain knowledge and truth from my research. Hopefully I will not lead you astray.

ACKNOWLEDGEMENTS

Special thanks to Dr. Robert B. Beckstead and Dr. Larry R. McDougald for the opportunity to explore this project and instilling me with a love for research. Thanks for putting up with me. I would like to acknowledge Thomas Ergen, Blake Duke, Nicholas Crowell, Zachary Williams, and Jacob Fluri for their contributions to this project and Dr. Lorraine Fuller for helping me get started with my project and her dedication to live animal studies. I would also like to acknowledge Dr. Richard Gerhold who showed me the ropes. Brian Jordan suggested splinkerette PCR; without it, this project would not have succeeded.

Table of Contents

Page

List of Tables

List of Figures

Page

CHAPTER 1

INTRODUCTION

Blackhead disease is caused by *Histomonas meleagridis* (Tyzzer and Fabyan, 1922). The disease infects gallinaceous birds with clinical signs including necrotic liver lesions, caseous secretions of the ceca, and sulfur-colored droppings. All treatment drugs have been removed from the consumer market due to residual health concerns with only one preventative feed additive Nitarsone (4-nitrophenylarsonic acid) still on the market in the USA (McDougald, 2005). Based on sequence information from ribosomes and proteins, *H. meleagridis* is classified in the Phylum Trichomonadae, Class Tritrichomonadea, Order Tritrichomonadida, and Family Dientamoebidae (Cepicka et al., 2010). Little genetic information exists for *H. meleagridis*, making study of the molecular genetics of this parasite imperative to finding a means of prevention or treatment for disease. Development of methods to genetically alter histomonads through generation of a gene product through expression constructs requires the identification and testing of expression elements.

Molecular data is limited to a few nucleotide and protein sequences (Felleisen, 1997; Gerbod et al., 2001; Mazet et al., 2008; Bilic et al., 2009; Hauck and Hafez, 2009b; Cepicka et al., 2010; Leberl et al., 2010). These sequences were obtained by homology PCR or the establishment of a partial cDNA library (Mazet, et al., 2008; Bilic, et al., 2009). These approaches do not yield DNA sequence information for the noncoding portions of the gene needed for their expression. Thus, no information is available in *H. meleagridis* regarding regulatory elements such as promoters, initiator elements, TATA boxes, and polyadenylation and cleavage signals.

Research in protozoan species has demonstrated differences in expression elements compared to mammalian genes. The presence of TATA boxes in protozoa is rare. Instead, these organisms rely on an initiator element (Inr), a ubiquitous expression element found within the promoter region of eukaryotes that allows for the binding of polymerase II (Quon et al., 1994). Metazoan Inr has a consensus sequence of $YYA^{+1}NWYY$ (Smale, 1997). Primitive eukaryotes loosely follow this consensus, but have a unique motif to each species. Since *H. meleagridis* has little genetic information available, other related parasites must be examined as models for expression elements in this organism. Ideally, it would be another member of the Dientamoebidae family (*Dientamoeba*, *Protrichomonas*, or *Parahistomonas*), but little genetic information is known about these species.

In this study, we analyze other amitochondriates, specifically *Trichomonas vaginalis* and *Giardia lamblia* that have yielded key information regarding expression elements associated with protozoans. Downstream expression elements in the 3'UTR include a polyadenylation and cleavage signal sequences (Espinosa et al., 2002). The 3' UTR itself is short in protozoan species (Vanacova et al., 2003). Metazoan polyadenylation signals have a consensus sequence of AAUAAA. However, the polyadenylation signal for primitive eukaryotes seems to be species specific (Zamorano et al., 2008). Research shows each protozoan has unique elements, therefore the elements specific to *H. meleagridis* must be identified, instead of extrapolating the sequence from a related organism.

The development of an expression vector for *H. meleagridis* containing a ubiquitous promoter element with polyadenylation and cleavage signals is crucial for further study of genetics for this organism. In this study, we aim to:

1) Test the use of splinkerette PCR as a technique to amplify unknown upstream region of beta-tubulin

2) Isolate expression elements for *H. meleagridis* within the intergenic region upstream of beta-tubulin

LITERATURE CITED

BILIC, I., M. LEBERL, and M. HESS. 2009. Identification and molecular characterization of numerous Histomonas meleagridis proteins using a cDNA library. *Parasitology, 136*(4), 379-391. doi: S0031182008005477 [pii]10.1017/S0031182008005477

CARLTON, J. M., R. P. HIRT, J. C. SILVA, A. L. DELCHER, M. SCHATZ, Q. ZHAO. 2007. Draft genome sequence of the sexually transmitted pathogen Trichomonas vaginalis. [Article]. *Science, 315*(5809), 207-212. doi: 10.1126/science.1138294

CEPICKA, I., V. HAMPL, and J. KULDA. 2010. Critical Taxonomic Revision of Parabasalids with Description of one New Genus and three New Species. [Article]. *Protist, 161*(3), 400-433. doi: 10.1016/j.protis.2009.11.005

DEVON, R. S., D.J. PORTEOUS, AND A.J. BROOKES. 1995. Splinkeretts--improved vectorettes for greater efficiency in PCR walking. *Nucleic Acids Research, 23*(9), 1644-1645.

ESPINOSA, N., R. HERNANDEZ, L. LOPEZ-GRIEGO, and I. LOPEZ-VILLASENOR. 2002. Separable putative polyadenylation and cleavage motifs in Trichomonas vaginalis mRNAs. [Article]. *Gene, 289*(1-2), 81-86.

FELLEISEN, R. S. J. 1997. Comparative sequence analysis of 5.8S rRNA genes and internal transcribed spacer (ITS) regions of trichomonadid protozoa. *Parasitology, 115*(2), 111-119.

GERBOD, D., V. P. EDGCOMB, C. NOEL, L. ZENNER, R. WINTJENS, P. DELGADO-VISCOGLIOSI. 2001. Phylogenetic position of the trichomonad parasite of turkeys, Histomonas meleagridis (Smith) Tyzzer, inferred from small subunit rRNA sequence. *J Eukaryot Microbiol, 48*(4), 498-504.

HAUCK, R., and H. M. HAFEZ. 2009a. Partial sequence of the beta-tubulin of Histomonas meleagridis and the activity of benzimidazoles against H-meleagridis in vitro. *Parasitology Research, 104*(5), 1183-1189. doi: DOI 10.1007/s00436-008-1309-5

_____. 2009b. Partial sequence of the beta-tubulin of Histomonas meleagridis and the activity of benzimidazoles against H. meleagridis in vitro. *Parasitol Res, 104*(5), 1183-1189. doi: 10.1007/s00436-008-1309-5

HENGEN, P. N. 1995. METHODS AND REAGENTS - VECTORETTE, SPLINKERETTE AND BOOMERANG DNA AMPLIFICATION. [Note]. *Trends in Biochemical Sciences, 20*(9), 372-373.

HORN C, H. J., SCHNUTGEN F, SEISENBERGER C, FLOSS T, IRGANG M, DE-ZOLT S, WURST W, VON MELCHNER H, NOPPINGER PR. 2007. Splinkerette PCR for more efficient characterization of gene trap events. *Nature Genetics, 39*(8), 933-934.

KATIYAR, S., and T. EDLIND. 1994. Beta-tubulin genes of *Trichomonas vaginalis*. *Molecular and Biochemical Parasitology, 64*, 33-42.

KIRK-MASON, K. E., M. J. TURNER, and P. R. CHAKRABORTY. 1989. Evidence for unusually short tubulin mRNA leaders and characterization of tubulin genes in *Giardia lamblia*. *Molecular and Biochemical Parasitology, 36*, 87-100.

LEBERL, M., M. HESS, and I. BILIC. 2010. Histomonas meleagridis possesses three alpha-actinins immunogenic to its hosts. *Mol Biochem Parasitol, 169*(2), 101-107. doi: 10.1016/j.molbiopara.2009.10.007

MAZET, M., M. DIOGON, J. F. ALDERETE, C. P. VIVARES, and F. DELBAC. 2008. First molecular characterisation of hydrogenosomes in the protozoan parasite Histomonas meleagridis. *Int J Parasitol, 38*(2), 177-190. doi: S0020-7519(07)00224-X [pii]10.1016/j.ijpara.2007.06.006

MCDOUGALD, L. R. 2005. Blackhead disease (histomoniasis) in poultry: a critical review. *Avian Dis, 49*(4), 462-476.

MORRISON, H., *ET AL.* 2007. Genomic Minimalism in the Early Diverging Intestinal Parasite *Giardia lamblia*. *Science, 317*, 1921-1926.

QUON, D. V. K., M. G. DELGADILLO, A. KHACHI, S. T. SMALE, and P. J. JOHNSON. 1994. SIMILARITY BETWEEN A UBIQUITOUS PROMOTER ELEMENT IN AN ANCIENT EUKARYOTE AND MAMMALIAN INITIATOR ELEMENTS. [Article]. *Proceedings of the National Academy of Sciences of the United States of America, 91*(10), 4579-4583.

SMALE, S. T. 1997. Transcription initiation from TATA-less promoters within eukaryotic protein-coding genes. [Review]. *Biochimica Et Biophysica Acta-Gene Structure and Expression, 1351*(1-2), 73-88.

TYZZER, E. E., and M. FABYAN. 1922. A Further Inquiry into the Source of the Virus in Blackhead of Turkeys, Together with Observations on the Administration of Ipecac and of Sulfur. *J Exp Med, 35*(6), 791-812.

VANACOVA, S., D. R. LISTON, J. TACHEZY, and P. J. JOHNSON. 2003. Molecular biology of the amitochondriate parasites, Giardia intestinalis, Entamoeba histolytica and Trichomonas vaginalis. [Review]. *International Journal for Parasitology, 33*(3), 235-255. doi: 10.1016/s0020-7519(02)00267-9

YEE, J., M. R. MOWATT, P. P. DENNIS, and T. E. NASH. 2000. Transcriptional analysis of the glutamate dehydrogenase gene in the primitive eukaryote, Giardia lamblia - Identification of a primordial gene promoter. [Article]. *Journal of Biological Chemistry, 275*(15), 11432-11439.

ZAMORANO, A., C. LOPEZ-CAMARILLO, E. OROZCO, C. WEBER, N. GUILLEN, and L. A. MARCHAT. 2008. In silico analysis of EST and genomic sequences allowed the prediction of cis-regulatory elements for Entamoeba histolytica mRNA polyadenylation. [Article]. *Computational Biology and Chemistry, 32*(4), 256-263. doi: 10.1016/j.compbiolchem.2008.03.019

CHAPTER 2

LITERATURE REVIEW

Histomonas meleagridis

History and Morphology

Blackhead disease was first described as an outbreak of enterohepatitis in 1893 on a turkey farm in Rhode Island (Cushman, 1893). However, it wasn't until 1920 that Tyzzer identified the causative agent as *Histomonas meleagridis,* a flagellated protozoan organism distinct from other Trichomonads, coccidia, and amoebas (Tyzzer, 1920). *H. meleagridis* is a member of the family Dientamoebidae, order Tritrichomonadida, and class Tritrichomonadea (Cepicka et al., 2010). The Dientamoebidae family contains four genera *Dientamoeba*, *Protrichomonas*, *Histomonas*, and *Parahistomonas*. These members lack an undulating membrane and costa. Also missing are suprakinetosomal and infrakinetosomal body. Organisms of Dientamoebidae are uninucleate to binucleate with zero to four flagella. Specifically, members of the genus *Histomonas* are uninucleate with one flagellum.

Histomonads are round or amoeboid in shape with a diameter of 8-15 microns (Smith, 1895). They utilize anaerobic metabolism. *Histomonas* has hydrogenosomes instead of mitochondria that generate ATP by converting pyruvate and malate into hydrogen and acetate (Muller, 1993). Food vacuoles can be observed microscopically and in culture can contain rice flour or bacteria. Recent work has proposed that *Histomonas* has a cyst stage although its significance is not clear (Munsch et al., 2009; Zaragatzki et al., 2010). Previous work had observed only the amoeboid and flagellated form (Lund and Chute, 1974).

The life cycle of *H. meleagridis* is complex and may utilize an intermediate host in order to survive in the environment outside the bird. Without an intermediate host, histomonads die quickly in the external environment (Graybill, 1920; Tyzzer and Collier, 1925; Niimi, 1937). The most commonly used carriers are *Heterakis gallinarum* and some species of earthworms, members of the class Oligochaeta. Histomonads infect the ova of *H. gallinarum* and are passed with the feces. The eggs may then be eaten by an earthworm where the larvae can hatch and live indefinitely, or be eaten directly by a bird and hatch during digestion (Graybill, 1920; Lund, Wehr et al., 1966). Histomonads are released when *Heterakis* larva molt, and migrate to the ceca, the primary site of infection (Fine, 1975). From there, histomonads may enter the bloodstream. The blood supply of the liver is connected to the blood of the intestines and ceca, providing a pathway for histomonads to migrate to the secondary site (See Fig.) (Clarkson, 1961). *Histomonas meleagridis* has also been found in the kidney, spleen, and lungs (Levine, 1947; Malewitz and Calhoun, 1957; Malewitz et al., 1958; Clarkson, 1961; Senties-Cue et al., 2009). Studies have indicated that *H. meleagridis* can survive for almost three years within viable *H. gallinarum* eggs (Farr, 1961). With such a long period of contamination and with the variety of inapparent hosts, it has long been recommended that turkeys be reared away from other gallinaceous birds to avoid cross-contamination.

Turkeys may also become infected through the uptake of histomonads from infected fecal droppings via reverse peristalsis of the cloaca. This route of transmission has only been shown in turkey flocks maintained at high stocking density. It is thought the behavior tendency of turkeys to huddle together when sick contributes to this mode of transmission where histomonads can pass from cloaca to cloaca or from fresh droppings to cloaca through the entire flock (Hu and McDougald, 2003). This mode of transmission does not seem to be an important factor in chickens (Hu et al., 2006).

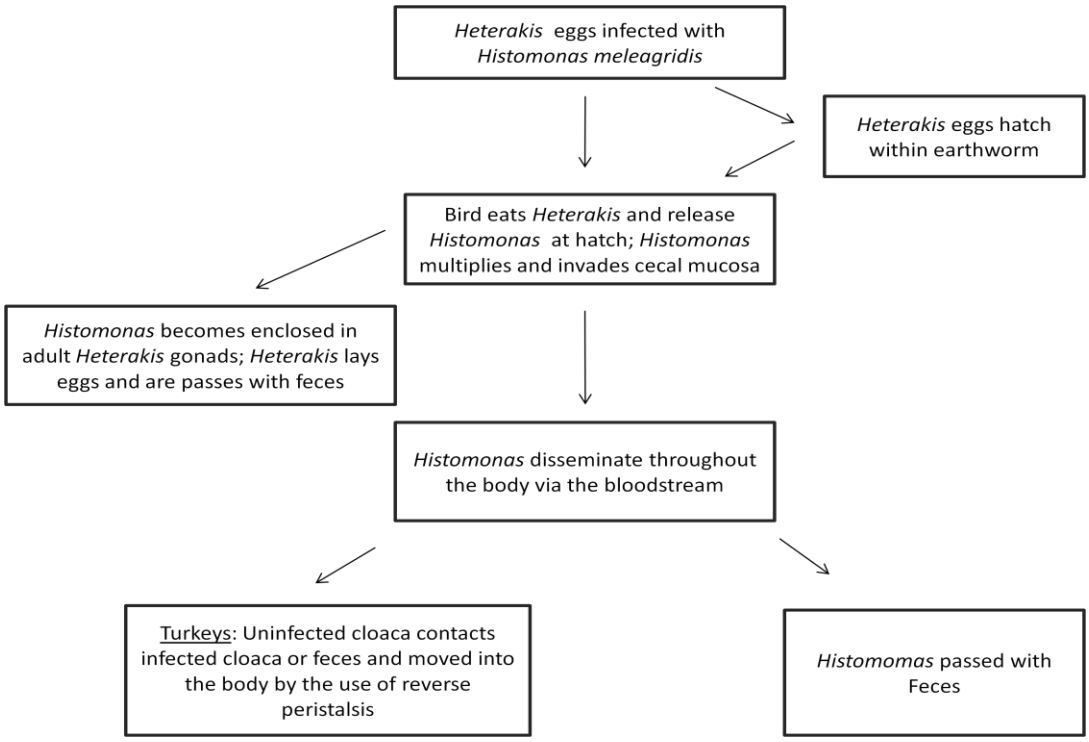


Figure LR1: Life cycle of *Histomonas meleagridis*

As reviewed by McDougald in 2003, clinical signs of histomoniasis include lethargy, anorexia, drooping of wings and head, malaise, fluffing of feathers, and sulfur colored droppings. The cecal walls may swell from an immune response against the histomonads and become thickened with caseous exudates. The histomonads infecting the ceca are flagellated. When histomonads move to the liver, lesions are necrotic foci of the liver that have raised borders with a dark, concave center. Variations of these lesions can be large white raised necrotic areas. Occasionally, the liver may take on a green tinge, presumably because of the clogging of the bile duct with histomonads or inflammation. The histomonads within these tissues take on an aflagellate, amoeboid form. Birds suffering from blackhead may eventually die of malnutrition due to the inability to absorb nutrients or of liver failure (McDougald, 2003).

Histomonas meleagridis can infect a variety of gallinaceous birds with varying virulence. The ring-necked pheasant is thought to be responsible for the introduction of this parasite to the Americas

(Lund and Chute, 1974). The pheasant, jungle fowl, and chicken are inapparent carriers sometimes showing no signs of infection. Moderate disease can be observed in guinea fowl, bobwhite quail, and Hungarian partridge. Turkeys in particular are subject to severe histomoniasis. In the industry, mortality can reach from 24% to 95% in flocks (AbdulRahman and Hafez, 2009; Senties-Cue, et al., 2009). The reason for this variation is unknown, but could be due to variations in genetic resistance to *H. meleagridis* or varying virulence of the *H. meleagridis* strain (AbdulRahman and Hafez, 2009; Lollis, submitted for publication).

In vitro Culture

Requirements for *H. meleagridis* have been established for successful culture. Histomonads require slightly acidic, anerobic environment with bacteria and a nutrient-rich medium and added carbohydrates (Hauck et al., 2010). An axenic culture has yet to be established suggesting that bacteria are somehow essential for survival, perhaps due to nutritional or environmental necessity (McDougald, 2005). Generally rice flour has been used for added carbohydrates, but other grain powders have also been used as additives (Hauck, et al., 2010). An anaerobic environment could be necessary due to the hydrogenosome's anaerobic metabolic nature (Muller, 1993).

Several media can support the growth of histomonads (Drbohlav, 1924; Laidlaw, 1928; Bishop, 1938; Devolt, 1943). Dwyer medium has been the primary medium employed by some investigators to support histomonad growth (Dwyer, 1970). It contains 85-95% of Medium 199, 5-10% horse serum, 5% chick embryo extract, and 1% rice powder. It is nutrient rich and provides for exponential growth of histomonads and bacteria alike. More recently, an improved Dwyer medium has been used which lacks the chick embryo extract and contains only 0.8% rice powder in an effort to provide better growth (van der Heijden and Landman, 2007). Histomonads grow logarithmically for 1-5 days in modified Dwyer medium. After the peak, numbers tend to decline due to death from depletion of nutrients. Cultures can

be preserved by cryogenic freezing using a cryoprotectant (8% DMSO) added to the culture (Chute and Chute, 1969). Maintenance of culture is accomplished by passing some of the histomonads to new media. However, repeated passage in culture can lead to loss of ability to establish infections in hosts (Lund, Augustin et al., 1966; Lund et al., 1967; Dwyer and Honigber, 1970). Virulence can be restored in cultures that have been passed up to 16 weeks by serial passages through susceptible hosts. After 16 weeks, virulence could not be restored (Dwyer and Honigber, 1970). The reason for loss of virulence is unknown.

Several hypotheses have been suggested concerning how virulence is lost upon passage in culture from bacterial changes and genetic selection (McDougald, 2005). Franker and Doll were unable to establish infection in gnotobiotic turkeys (Franker and Doll, 1964). Bacteria are normal to the turkey intestinal tract, and some thought to affect *H. meleagridis* virulence include *Escherichia coli, Clostridium perfringens, Bacillus subtilis, Klebsiella pneumoniae, Stenotrophomonas maltophilia, Pseudomonas aeruginosa*, and *Staphylococcus* sp., though their role is indefinite (Bradley and Reid, 1964; Hauck et al., 2010). Virulence of *H. meleagridis* could be influenced by genetic selection; however, there is no published data concerning the issue. Overall, virulence of *H. meleagridis* changes but the reason remains unclear.

Treatment

All preventative and treatments for *H. meleagridis* have been banned in Europe and the United States due to concerns of residues and carcinogens. Compounds formerly on the market for the effective treatment and/or prevention of histomoniasis were enheptin, nithiazide, nitroimidazoles (McDougald, 2005). Because of de-registration of the nitroimidazoles, only one treatment is on the market for the prevention of histomoniasis—Nitarsone (4-nitrophenyl-arsenic acid), a pentavalent arsenical feed additive. Even with this additive, infection occurs soon after withdrawal of medication (McGuire, 1952). New drugs for the treatment and prevention of histomoniasis are unlikely to be

developed because the small size of the turkey industry. The creation of a vaccine would be may be the only possibility for future preventative treatment.

Expression Elements in Related Species

Since little genetic information is available for *H. meleagridis*, other related parasites must be examined as models for expression elements in this organism. Ideally, another member of the Dientamoebidae family (*Dientamoeba*, *Protrichomonas*, or *Parahistomonas*) would be used as a model, but few genetic resources are available from these species. Instead, we must examine the genetic characteristics of members of Order Trichomonadida's *Trichomonas vaginalis,* Order Diplomonadida's *Giardia lamblia,* and Order Amoebida's *Entamoeba histolytica* on whom expression studies exist.

Trichomonas vaginalis

Of the primitive eukaryotes, *T. vaginalis* has the most defined expression elements. Transcription initiation utilizes an initiator element (Inr), a sequence that directs the binding of RNA polymerase II and various transcription factors at the start site for transcription; it also helps in accurate start site detection. The Inr for metazoan genes is YYA^{+1}NWYY, where Y is a pyrimadine, W is A or T, and $^{+1}$ signifies the A that is the start of transcription (Smale, 1997). The Inr in *T. vaginalis* has developed a consensus sequence of TCA^{+1}YTWYTCATTA. The ATG start codon is located downstream of the Inr at a maximum of 4 bp, suggesting that the 5'UTR is less than 4 bp (Quon et al., 1994). Analysis of the alpha-succinyl CoA sythase promoter elucidates other regulatory factors, as well. Deletion of the area -251 to -51 decimated promoter activity. Within this area, the sequence from -98 to -84 and -80 to -69 affect the activity most profoundly. The -98 to -84 sequence has also been found in other promoters in *T. vaginalis*, but the -80 to -69 seems to be gene specific (Liston et al., 1999). Conservation of these regions suggests the necessity for their sequence to bind transcription factors. No TATA boxes or TATA-like elements have been identified regulating transcription 5' of the coding region of genes.

Although this is a primitive eukaryote, transcription signaling exists. Presence of iron is known to alter the expression of some genes in *T. vaginalis* (Lehker, 1992). Housekeeping mRNAs seem unaffected by iron levels in the organisms' environment, while other mRNAs, particularly those coding for virulence factors, increase greatly with an increase in the concentration of environmental iron availability (Tsai, 2002). The host's iron contained in the hemoglobin in blood seems to activate a transcription factor that initiates the cascade of upregulation of virulence genes.

Genes also encode for the 3' end of an mRNA molecule that has two distinct motifs, a polyadenylation signal and a cleavage signal. In *T. vaginalis*, the polyadenylation signal is an extension of the stop codon TAA with an adenosine being the first base following it, although there are a few exceptions. Espinosa claims the polyadenylation signal becomes TAAA with the addition of an adenosine to the stop codon. The putative cleavage site is 4 bp downstream of the polyadenylation signal and has a conserved sequence of AATT (Espinosa et al., 2002). According to Minotto, the polyadenylation signal for *T. vaginalis* is AATAAA, which is characteristic of higher eukaryotic signals (Minotto et al., 2000). It is currently unclear which of these signals is the functioning polyadenylation signal.

Giardia lamblia

Giardia lamblia is another anaerobic protozoan organism with a minimalistic genome containing 12.3 Mb (Adam, 2001). The genome lacks introns and has relatively short 5' and 3' UTRs, both AT-rich. Evidence exists for bidirectional transcription (Teodorovic, 2007). The majority of 5' UTRs consists of 1-14 bp. A sequence of AATTA^{+1}AAAA has been suggested as the consensus initiation site. A second sequence of CAAAAAWYAGAKTCYGAA where K is T or G is conserved about -30 bp. A third consensus of CAATTT has been identified at -45 bp. The region necessary for promoter activity is -51 to -20 and was found by deletion and mutation mapping (Sun and Tai, 1999). Genes in *G. lamblia* have a short 3' UTR of only 10-30 bp past the stop codon. The sequence AGTRAA is found 6-9 bp downstream from the stop codon and has been proposed as the putative polyadenylation signal (Peattie, 1989). *Giardia* genes have

very short polyadenylated tails (Adam, 2001). A cleavage motif has not been identified. Overall, the regions required for expression of *G. lamblia* are compact with little unnecessary information.

Entamoeba histolytica

Entamoeba histolytica is the most genetically and phylogenetically dissimilar from *H. meleagridis* of the organisms mentioned here. A few genes have introns (Willhoeft, 2001) whereas genes from *T. vaginalis* and *G. lamblia* do not. Introns are short (46-115 bp) and follow the patterns of introns of higher eukaryotes, GU and AG are splice sites and the branchpoint sequence is conserved YNYYRAY, where R is a purine (Vanacova et al., 2003).

In *E. histolytica*, the 5'UTR is short. The sequence A^{+1}TTCA or A^{+1}TCA surrounds the transcription start-site with another TATA-like conserved motif of TATTTAAA, located approximately 30 bp upstream (Bruchhaus, 1993). A GAAC element has been found at approximately -10 to -20 (Purdy et al., 1996). A few more upstream promoter elements have been characterized, but these seem to have positive and negative regulatory controls and are species specific (Buss et al., 1995; Singh et al., 1997). The 3' UTR is short with 20 to 100 bp with few exceptions. The motif TAWTT is found about halfway within this region and functions as the putative transcription termination/polyadenylation signal. A run of pyrimadines has been noted upstream and adjacent to the polyadenylated tail (Bruchhaus, 1993).

From the data that has been published on *T. vaginalis, G. lamblia,* and *E. histolytica*, we can draw a few conclusions about the expression elements of *H. meleagridis*. Both the 5' and 3' UTRs will be short, probably containing less than 20 bp. The transcription start site is contained within the Inr. It is unlikely there are introns within the *H. meleagridis* genome (Mazet et al., 2008; Hauck and Hafez, 2010; Leberl et al., 2010). The polyadenylation signal will be mostly thymines and adenines and will be uniform through most transcripts. Similarities between the above mentioned species will allow putative expression elements to be identified and characterized in *H. meleagridis.*

Genes Identified in *Histomonas meleagridis*

A few molecular sequences are available in *H. meleagridis,* including ribosomal genes and protein coding regions. Ribosomal sequences known are the internal transcribed spacer (ITS)-1 region, 5.8S region, ITS-2, partial 18S ribosomal RNA (Felleisen, 1997), partial 28S ribosomal RNA, and the small subunit ribosomal RNA gene (Gerbod et al., 2001). Complete known coding regions include three hydrogenosomal proteins (Mazet, et al., 2008) and three alpha-actinins (Leberl, et al., 2010). Partial sequences are known for thirty-seven partial cDNA sequences for translation, ribosomal structure and biogenesis, structural proteins, energy conversion, cell envelope, protein kinases, oxidative stress proteins (Bilic et al., 2009), beta-tubulin (Hauck and Hafez, 2009), GADPH, alpha-tubulin, and enolase (Hauck and Hafez, 2010). These genes have been described either by PCR methods using sequence homology to other protozoa (Mazet, et al., 2008) or the sequencing of *H. meleagridis* cDNA libraries where proteins were identified by antibody response to antigens, and proteins eliciting responses were sequenced (Bilic, et al., 2009). However, these approaches do not provide sequence information for the noncoding portions of the gene needed for proper expression. Thus no information is available regarding the expression elements for *H. meleagridis* including promoters, initiator elements, TATA boxes, and polyadenylation and cleavage signals.

In this study, we chose to work with beta-tubulin. Beta-tubulin is an ideal gene to work with because it is a housekeeping gene ubiquitously expressed in all eukaryotic organisms. A partial sequence is already known for *H. meleagridis* which is necessary for our use of splinkerette PCR, a type of ligated adaptor PCR to be discussed later. Beta-tubulin is a globular protein that forms a dimer with alpha-tubulin making microtubules. Microtubules are an essential part of the mitotic spindle, helping cells to divide, and can also be part of the cell cytoskeleton, flagella, or cilia.

Benzimidazoles are antihelminthic drugs that act on beta-tubulin and inhibit growth and are parasiticidal at high concentrations (>.33ug/mL). Their activity has proved effective for the control and treatment of *T. vaginalis, G. lamblia,* and *E. histolytica.* However, their efficacy is dependent on the amino acid sequence of beta-tubulin. Certain mutant strains in which the beta-tubulin has been altered at specific positions are able to survive when treated with benzimidazoles (Katiyar et al., 1994). Previous studies have tested the effectiveness of benzimidazoles on *H. meleagridis*. Hegngi successfully used benzimidazole as an antihelminthic to stop the infection of birds with *Heterakis* eggs infected with *H. meleagridis*. However, this was due to the drug killing the worm before having the chance to molt and release the histomonads (Hegngi et al., 1999). Hauck and Hafez sequenced part of the beta-tubulin gene and compared his sequences to those of *T. vaginalis*. Amino acid sequences appeared benzimidazole susceptibility, but *in vitro* tests proved otherwise (Hauck and Hafez, 2009).

In *T. vaginalis*, beta-tubulin is commonly expressed gene as indicated by its high copy number (6-7, determined by Southern blot analysis) within the genome. Its size is around 1.7 kb. These genes are broken down into categories of *btub1, btub2*, and *btub3*. These categories are selected by amino acid sequence. Of the groups, *btub1* is most highly expressed with less expression of *btub2,* and *btub3* expression was hardly detectable. *Trichomonas vaginalis*'s beta-tubulin shares only and 84% identity with that of the beta-tubulin of *G. lamblia* (Katiyar and Edlind, 1994). The beta-tubulin coding region for *G. lamblia* is ~1.35 kb (Kirk-Mason et al., 1988). *Giardia* possesses three copies of the gene. Two of these beta-tubulins are oriented "head-to-head" within the genome (Kirk-Mason et al., 1989), possibly due to bidirectional transcription. For *E. histolytica*, only one copy of the beta-tubulin gene exists. It is highly divergent from other beta-tubulins including *T. vaginalis* and *G. lamblia* having a mere 54-58% identity (Katiyar and Edlind, 1996). These differences could be due in part to the lack of beta-tubulin within the cytoplasm and spindle structures within this organism (Orozco et al., 1988).

For the beta-tubulin of *H. meleagridis*, we expect multiple copies of the gene like *T. vaginalis* and *G. lamblia*. We also expect for the amino acid sequence to be 440-460 units long, because there are 447 amino acids for beta-tubulin of *T. vaginalis* (Katiyar and Edlind, 1994), 446 of *G. lamblia* (Kirkmason et al., 1988), and 459 of *E. histolytica* (Katiyar and Edlind, 1996). Hauck and Hafez already isolated ~438 bp of *H. meleagridis*'s beta-tubulin and determined it not to be benzimidazole susceptible. Their clones produced 21 different nucleotide sequences and 11 amino acid sequences. However, they were unable to conclusively say how many copies of the beta-tubulin gene are found in *H. meleagridis* (Hauck and Hafez, 2009).

Splinkerette PCR

In this research, we present a method for isolating upstream and downstream sequences of the beta-tubulin genes in *Histomonas meleagridis* using Splinkerette PCR, a type of ligated adaptor PCR (Devon, 1995). Splinkerette PCR is a variation on vectorette PCR. Vectorette PCR relies on a mismatch middle region of the adaptor. The vectorette specific primer is an exact match to the mismatched region of the adaptor, but it cannot anneal to or initiate elongation until its complementary sequence has been synthesized by the polymerase extension from the DNA specific primer. Non-specific priming is an issue with this method, in particular end-repair priming. The 5' overhang based on restriction sites of adaptors and digested DNA tend to self anneal to initiate priming. This leads to exponential amplification of undesired sequence, normally that of the vectorette (Devon, 1995).

Splinkerette is an improvement on the vectorette. Instead of mismatched middle region in an adaptor, a hairpin structure is incorporated into one of the strands of the adaptor (Devon, 1995). Once again, the primer for the adaptor has the same sequence as the adaptor and cannot participate in priming until its complement is generated. The elongation product that is initially generated from the DNA specific primer continues to be expanded until it reaches the adaptor region that has the hairpin

structure. The DNA can then flip back onto itself to form a stable double stranded structure. This inhibits end-repair priming. Through the adaptor design, only the strand of the adaptor without the hairpin can act as a non-specific primer (Devon, 1995).

To further decrease non-specific amplification, the splinkerette techniques uses nested primers and two rounds of PCR. The first set of primers is specific for the outer splinkerette adaptor and the known sequence. The second set of primers is specific for the inner splinkerette and inside of the original primer in the known sequence. This decreases extraneous sequence that might be specific to the first primer but not the second, giving the researcher the desired region of amplification (Horn et al., 2007). Splinkerette was designed as a tool for the mapping of gene trap events (Friedel and Soriano, 2010) and for generating comprehensive markers of sequence (Qureshi et al., 1994). Here we are going use splinkerette PCR to identify non-coding portion of the gene.

LITERATURE CITED

ABDULRAHMAN, L., and H. M. HAFEZ. 2009. Susceptibility of different turkey lines to Histomonas meleagridis after experimental infection. *Parasitol Res, 105*(1), 113-116. doi: 10.1007/s00436-009-1369-1

ADAM, R. 2001. Biology of *Giardia lamblia*. *Clinical Microbiology Reviews, 14*(3), 447-475.

BILIC, I., M. LEBERL, and M. HESS. 2009. Identification and molecular characterization of numerous Histomonas meleagridis proteins using a cDNA library. *Parasitology, 136*(4), 379-391. doi: S0031182008005477 [pii]10.1017/S0031182008005477

BISHOP, A. 1938. *Histomonas meleagridis* in domestic fowls (*Gallus gallus*) cultivation and experimental infection. *Parasitology, 30*, 181-194.

BRADLEY, R. E., and W. M. REID. 1966. *Histomonas meleagridis* and several bacteria as agents of infective enterohepatitis in gnotobiotic turkeys. *Experimental Parasitology*, *19*, 91-101.

BRUCHHAUS, I., M. LIEPPE, C. LIOUTAS, AND E. TANNICH. 1993. Unusual Gene Organization in the Protozoan Parasite *Entamoeba histolytica*. *DNA and Cell Biology, 12*(10), 925-933.

BUSS, H., C. LIOUTAS, S. DOBINSKY, R. NICKEL, and E. TANNICH. 1995. ANALYSIS OF THE 170-KDA LECTIN GENE PROMOTER OF ENTAMOEBA-HISTOLYTICA. [Article]. *Molecular and Biochemical Parasitology, 72*(1-2), 1-10.

CEPICKA, I., V. HAMPL, and J. KULDA. 2010. Critical Taxonomic Revision of Parabasalids with Description of one New Genus and three New Species. [Article]. *Protist, 161*(3), 400-433. doi: 10.1016/j.protis.2009.11.005

CHUTE, M. B., and A. M. CHUTE. 1969. Freeze preservation of histomonas meleagridis. *Poult Sci, 48*(6), 2189-2191.

CLARKSON, M. J. 1961. The blood supply of the liver of the turkey and the anatomy of the biliary tract with reference to infection with *Histomonas meleagridis*. *Research in Veterinary Science, 2*, 259-264.

CUSHMAN, S. 1893. The production of turkeys. *Agricultural Experiment Station, Bulletin 25*, 89-123.

DEVOLT, H. M. 1943. A new medium for the culture of *Histomonas meleagridis*. *Journal of Parasitology, 29*, 353-355.

DEVON, R. S., D.J. PORTEOUS, AND A.J. BROOKES. 1995. Splinkeretts--improved vectorettes for greater efficiency in PCR walking. *Nucleic Acids Research, 23*(9), 1644-1645.

DRBOHLAV, J. J. 1924. The cultivation of the protozoan of blackhead. *Journal of Medical Research, 44*, 677-678.

DWYER, D. M. 1970. An improved method for cultivating Histomonas meleagridis. *J Parasitol, 56*(1), 191-192.

_____, and B. M. HONIGBER. 1970. Effect of Certain Laboratory Procedures on Virulence of Histomonas-Meleagridis for Turkeys and Chickens. *Journal of Parasitology, 56*(4), 694-700.

ESPINOSA, N., R. HERNANDEZ, L. LOPEZ-GRIEGO, and I. LOPEZ-VILLASENOR. 2002. Separable putative polyadenylation and cleavage motifs in Trichomonas vaginalis mRNAs. [Article]. *Gene, 289*(1-2), 81-86.

FARR, M. M. 1961. Further observations on survival of the protozoan parasite, Histomonas meleagridis, and eggs of poultry nematodes in feces of infected birdes. *Cornell Vet, 51*, 3-13.

FELLEISEN, R. S. J. 1997. Comparative sequence analysis of 5.8S rRNA genes and internal transcribed spacer (ITS) regions of trichomonadid protozoa. *Parasitology, 115*(2), 111-119.

FINE, P. E. M. 1975. Quantitative sturdies on the transmission of *Parahistomonas wenrichi* by ova of *Heterakis gallinarum*. *Parasitology, 70*(3), 407-417.

FRANKER, C. K., and J. P. DOLL. 1964. Experimental histomoniasis in gnotobiotic turkeys. II. Effects of some cecal bacteria on pathogenesis. *Journal of Parasitology, 50*, 636-640.

FRIEDEL, R. H., and P. SORIANO. 2010. GENE TRAP MUTAGENESIS IN THE MOUSE *Methods in Enzymology, Vol 477: Guide to Techniques in Mouse Development, Part B: Mouse Molecular Genetics, Second Edition* (Vol. 477, pp. 243-269). San Diego: Elsevier Academic Press Inc.

GERBOD, D., V. P. EDGCOMB, C. NOEL, L. ZENNER, R. WINTJENS, P. DELGADO-VISCOGLIOSI. 2001. Phylogenetic position of the trichomonad parasite of turkeys, Histomonas meleagridis (Smith) Tyzzer, inferred from small subunit rRNA sequence. *J Eukaryot Microbiol, 48*(4), 498-504.

GRAYBILL, H. 1920. Production of fatal blackhead in turkeys by feeding embryonated eggs of *Heterakis papillosa*. *Journal of Experimental Medicine, 31*, 647-655.

HAUCK, R., P. L. ARMSTRONG, and L. R. MCDOUGALD. 2010. Histomonas meleagridis (Protozoa: Trichomonadidae): analysis of growth requirements in vitro. *J Parasitol, 96*(1), 1-7. doi: GE-1985 [pii]10.1645/GE-1985.1

_____, and H. M. HAFEZ. 2009. Partial sequence of the beta-tubulin of Histomonas meleagridis and the activity of benzimidazoles against H. meleagridis in vitro. *Parasitol Res, 104*(5), 1183-1189. doi: 10.1007/s00436-008-1309-5

_____, and _____. 2010. Systematic position of Histomonas meleagridis based on four protein genes. *J Parasitol, 96*(2), 396-400. doi: 10.1645/GE-2267.1 [pii]

HEGNGI, F. N., J. DOERR, T. S. CUMMINGS, R. D. SCHWARTZ, G. SAUNDERS, A. ZAJAC. 1999. The effectiveness of benzimidazole derivatives for the treatment and prevention of histomonosis (blackhead) in turkeys. *Vet Parasitol, 81*(1), 29-37. doi: S0304401798002337 [pii]

HORN, C., J. HANSEN, F. SCHNUTGEN, C. SEISENBERGER, T. FLOSS, M. IRGANG. 2007. Splinkerette PCR for more efficient characterization of gene trap events. *Nature Genetics, 39*(8), 933-934.

HU, J., L. FULLER, P. L. ARMSTRONG, and L. R. MCDOUGALD. 2006. Histomonas meleagridis in chickens: attempted transmission in the absence of vectors. *Avian Dis, 50*(2), 277-279.

_____, and L. R. MCDOUGALD. 2003. Direct lateral transmission of Histomonas meleagridis in turkeys. *Avian Dis, 47*(2), 489-492.

KATIYAR, S., and T. EDLIND. 1994. Beta-tubulin genes of *Trichomonas vaginalis*. *Molecular and Biochemical Parasitology, 64*, 33-42.

_____, and _____. 1996. *Entamoeba histolytica* Encodes a Highly Divergent Beta-Tubulin.

KATIYAR, S., V. GORDON, G. MCLAUGHLIN, and T. EDLIND. 1994. Antiprotozoal Activities of Benzimidazoles and Correlations with beta-Tubulin Sequence. *Antimicrobial Agents and Chemotherapy, 38*(9), 2086-2090.

KIRK-MASON, K. E., M. J. TURNER, and P. R. CHAKRABORTY. 1988. CLONING AND SEQUENCE OF BETA-TUBULIN CDNA FROM GIARDIA-LAMBLIA. [Note]. *Nucleic Acids Research, 16*(6), 2733-2733.

_____, _____, and _____. 1989. Evidence for unusually short tubulin mRNA leaders and characterization of tubulin genes in *Giardia lamblia*. *Molecular and Biochemical Parasitology, 36*, 87-100.

LAIDLAW, P. P., C. DOBELL, A. BISHOP. 1928. Further experiments of the action of emetine in cultures of *Entamoeba histolytica*. *Parasitology, 20*, 207-220.

LEBERL, M., M. HESS, and I. BILIC. 2010. Histomonas meleagridis possesses three alpha-actinins immunogenic to its hosts. *Mol Biochem Parasitol, 169*(2), 101-107. doi: 10.1016/j.molbiopara.2009.10.007

LEHKER, M. A. J. A. 1992. Iron regulates growth of *Trichomonas vaginalis* and the expression of immunogenic trichomonad proteins. *Molecular Microbiology, 6*(1), 123-132.

LEVINE, P. P. 1947. Histomoniasis in a kidney of a turkey. *The Cornell Veterinarian, 37*, 269-270.

LISTON, D. R., J. C. CARRERO, and P. J. JOHNSON. 1999. Upstream regulatory sequences required for expression of the Trichomonas vaginalis alpha-succinyl CoA synthetase gene. [Article]. *Molecular and Biochemical Parasitology, 104*(2), 323-329.

LUND, E. E., P. C. AUGUSTIN, and D. J. ELLIS. 1966. Immunizing Action of in Vitro-Attenuated Histomonas Meleagridis in Chickens and Turkeys. *Experimental Parasitology, 18*(3), 403-407.

_____, _____, and A. M. CHUTE. 1967. Histomonas Meleagridis after 1000 in Vitro Passages. *Journal of Protozoology, 14*(2), 349-351.

_____, and A. M. CHUTE. 1974. The reproductive potential of Heterakis gallinarum in various species of galliform birds: implications for survival of H. gallinarum and Histomonas meleagridis to recent times. *Int J Parasitol, 4*(5), 455-461. doi: 0020-7519(74)90061-7 [pii]

_____, E. E. WEHR, and D. J. ELLI. 1966. Earthworm transmission of Heterakis and Histomonas to turkeys and chickens. *J Parasitol, 52*(5), 899-902.

MALEWITZ, T. D., and L. CALHOUN. 1957. The normal hematological picture of turkey poults and blood alterations caused by enterohepatitis. *American Journal of Veterinary Research, 18*, 396-399.

_____, R. A. RUNNELS, and L. CALHOUN. 1958. The pathology of experimentally produced histomoniasis in turkeys. *American Journal of Veterinary Research, 19*, 181-185.

MAZET, M., M. DIOGON, J. F. ALDERETE, C. P. VIVARES, and F. DELBAC. 2008. First molecular characterisation of hydrogenosomes in the protozoan parasite Histomonas meleagridis. *Int J Parasitol, 38*(2), 177-190. doi: S0020-7519(07)00224-X [pii] 10.1016/j.ijpara.2007.06.006

MCDOUGALD, L. R. 2003. Other protozoan diseases of the intestinal tract--histomoniasis (blakchead). In Y. M. Saif & J. R. G. H. J. Barnes, A. M. Fadly, L. R. McDougald, and D. E. Swayne (Eds.), *Diseases of Poultry* (pp. 1001-1006). Ames, IA: Iowa State Univeristy Press.

_____. 2005. Blackhead disease (histomoniasis) in poultry: A critical review. *Avian Diseases, 49*(4), 462-476.

MCGUIRE, W. C. A. N. F. M. 1952. Chemotherapy Studies of Histomoniasis. *Poultry Science, 31*(4), 603-609.

MINOTTO, L., M. R. EDWARDS, and A. S. BAGNARA. 2000. Trichomonas vaginalis: Characterization, expression, and phylogenetic analysis of a carbamate kinase gene sequence. [Article]. *Experimental Parasitology, 95*(1), 54-62.

MULLER, M. 1993. THE HYDROGENOSOME. [Review]. *Journal of General Microbiology, 139*, 2879-2889.

MUNSCH, M., A. LOTFI, H. M. HAFEZ, S. AL-QURAISHY, and H. MEHLHORN. 2009. Light and transmission electron microscopic studies on trophozoites and cyst-like stages of Histomonas meleagridis from cultures. *Parasitol Res, 104*(3), 683-689. doi: 10.1007/s00436-008-1246-3

NIIMI, D. 1937. [Studies on blackhead II. Mode of Infection]. *Japanese Journal of Veterinary Science, 16*, 183-239.

OROZCO, E., F. SOLIS, J. DOMINGUEZ, B. CHAVEZ, and F. HERNANDEZ. 1988. *Entamoeba histolytica*: Cell cycle and nuclear division. *Experimental Parasitology, 67*(1), 85-95.

PEATTIE, D., ROGELIO ALONSO, ANN HEIN, JOHN CAULFIELD. 1989. Ultrastructural Localization of Giardins to the Edges of Disk Microribbons of Giardia lamblia and the Nucleotide and Deduced Protein Sequence of Alpha Giardin. *Journal of Cell Biology, 109*(5), 2323-2335.

PURDY, J. E., L. T. PHO, B. J. MANN, and W. A. PETRI. 1996. Upstream regulatory elements controlling expression of the Entamoeba histolytica lectin. [Article]. *Molecular and Biochemical Parasitology, 78*(1-2), 91-103.

QUON, D. V. K., M. G. DELGADILLO, A. KHACHI, S. T. SMALE, and P. J. JOHNSON. 1994. SIMILARITY BETWEEN A UBIQUITOUS PROMOTER ELEMENT IN AN ANCIENT EUKARYOTE AND MAMMALIAN INITIATOR ELEMENTS. [Article]. *Proceedings of the National Academy of Sciences of the United States of America, 91*(10), 4579-4583.

QURESHI, S. J., D. J. PORTEOUS, and A. J. BROOKES. 1994. Alu-Based Vectorettes and Splinkerettes More Efficient and Comprehensive PCR Amplification of Human DNA from Complex Sources. *Genetic Analysis, Techniques, and Applications, 11*(4), 95-101.

SENTIES-CUE, G., R. P. CHIN, and H. L. SHIVAPRASAD. 2009. Systemic Histomoniasis Associated with High Mortality and Unusual Lesions in the Bursa of Fabricius, Kidneys, and Lungs in Commercial Turkeys. [Article]. *Avian Diseases, 53*(2), 231-238.

SINGH, U., J. ROGERS, B. MANN, and J. WILLIAM PETRI. 1997. Transcription initiation is controlled by three core promoter elements in the *hgl5* gene fo the protozoan parasite *Entamoeba histolytica*. *Microbiology, 94*, 8812-8817.

SMALE, S. T. 1997. Transcription initiation from TATA-less promoters within eukaryotic protein-coding genes. [Review]. *Biochimica Et Biophysica Acta-Gene Structure and Expression, 1351*(1-2), 73-88.

SMITH, T. 1895. An infectious disease among turkeys cause by Protozoa (infectious entero-hepatitis). *USDA, Bureau of Animal Industry Bulletin, 8*, 3-27.

SUN, C. H., and J. H. TAI. 1999. Identification and characterization of a *ran* gene promoter in the protozoan pathogen Giardia lamblia. [Article]. *Journal of Biological Chemistry, 274*(28), 19699-19706.

TEODOROVIC, S., COLLEEN WALLS, AND HEIDI ELMENDORF. 2007. Bidirectional transcription is an ingerent feature of *Giardia lamblia* promoters and contributes to an abundance of sterile antisense transcripts throughout the genome. *Nucleic Acids Research, 35*(8), 2544-2553.

TSAI, C.-D., HSING-WEI LIU, AND JUNG-HSIANG TAI. 2002. Characterization of an Iron-responsive Promoter in the Protozoan Pathogen *Trichomonas vaginalis*. *The Journal of Biological Chemistry, 277*(7), 5153-5162.

TYZZER, E. E. 1920. The Flagellate Character and Reclasification of the Parasite Producing "Blackhead" in Turkeys: Histomonas (Gen. nov.) meleagridis (Smith). *The Journal of Parasitology, 6*(3), 124-131.

_____, and J. COLLIER. 1925. Induced an natural transmission of blackhead in the absence of *Heterakis*. *Journal of Infectious Disease, 37*, 265-276.

VAN DER HEIJDEN, H., and W. J. M. LANDMAN. 2007. Improved Culture of Histomonas meleagridis in a Modification of Dwyer Medium. [Article]. *Avian Diseases, 51*(4), 986-988.

VANACOVA, S., D. R. LISTON, J. TACHEZY, and P. J. JOHNSON. 2003. Molecular biology of the amitochondriate parasites, Giardia intestinalis, Entamoeba histolytica and Trichomonas vaginalis. [Review]. *International Journal for Parasitology, 33*(3), 235-255. doi: 10.1016/s0020-7519(02)00267-9

WILLHOEFT, U., EDUARDO CAMPOS-GONGORA, SASSIA TOUZNI, IRIS BRUCHHAUS, AND EGBERT TANNICH. 2001. Introns of *Entamoeba histolytica* and *Entamoeba dispar*. *Protist, 152*, 149-156.

ZARAGATZKI, E., H. MEHLHORN, F. ABDEL-GHAFFAR, K. A. S. AL-RASHEID, E. GRABENSTEINER, and M. HESS. 2010. Experiments to produce cysts in cultures of Histomonas meleagridis-the agent of histomonosis in poultry. *Parasitology Research, 106*(4), 1005-1007. doi: DOI 10.1007/s00436-010-1776-3

CHAPTER 3

IDENTIFICATION OF GENE EXPRESSION ELEMENTS IN *HISTOMONAS MELEAGRIDIS* USING SPLINKERETTE PCR, A VARIATION OF LIGATED ADAPTOR PCR[1]

[1] Lynn, E. C. and R. B. Beckstead, Submitted to *The Journal of Parasitology* 7/13/2011

Abstract

Histomonas meleagridis is the causative agent of blackhead disease in gallinaceous birds, but little genetic information exists for this organism. The complete genome for this protozoan is unsequenced. The only available sequence information is for coding portions of genes. No information is available for expression elements. In this study, we demonstrate that splinkerette PCR procedure, a variation of ligated adaptor PCR, can be used to identify regions upstream and downstream of known coding sequences. Using this technique, we isolated the upstream sequence of 2 beta-tubulin genes. With sequence analysis of their upstream regions, we identified their upstream intergenic regions and 2 different open reading frames. These intergenic regions contain putative polyadenylation and cleavage signals and putative initiator elements. Our research demonstrates that the use of splinkerette PCR is a valuable tool to identify regions of unknown DNA that are 5' or 3' to known sequences in parasites whose genomes remain unsequenced. The identification of the expression elements of *H. meleagridis* will provide tools for future studies on its gene expression.

INDEX WORDS: *Histomonas meleagridis*, molecular characterization, beta-tubulin, splinkerette PCR

Introduction

Blackhead disease is caused by the protozoan *Histomonas meleagridis* (Tyzzer and Fabyan, 1922). The disease manifests itself in gallinaceous birds with necrotic liver lesions, caseous secretions in the ceca, and sulfur-colored droppings. All treatment drugs have been banned due to residual health concerns with only one preventative feed additive Nitarsone (4-nitrophenylarsonic acid) still on the market in the USA (McDougald, 2005). Based on sequence information from the small subunit (SSU) RNA, the internal transcribed spacer (ITS)-1 region, and SSU rDNA genes, and glyceraldehyde-3-phosphate (GAPDH), enolase, and alpha- and beta-tubulins, *H. meleagridis* is classified in the Phylum Trichomonadae, Class Tritrichomonadea, Order Tritrichomonadida, and Family Dientamoebidae (Cepicka et al., 2010). Little other genetic information exists for *H. meleagridis*, making molecular study of this parasite essential to finding a means of prevention or treatment for disease.

Published molecular sequence data in *H. meleagridis* is limited to a small number of sequences including ribosomal and protein coding genes. Ribosomal genes include the internal transcribed spacer (ITS)-1 region, 5.8S region, ITS-2, partial 18S ribosomal RNA, partial 28S ribosomal RNA , and the small subunit ribosomal RNA gene (Felleisen, 1997; Gerbod, et al., 2001) . Complete protein coding sequences are known for three hydrogenosomal proteins and three alpha-actinins (Mazet, et al., 2008; Leberl, et al., 2010). Partial sequences include thirty-seven partial cDNA sequences that encode for translation, ribosomal structure and biogenesis, structural proteins, energy conversion, cell envelope, protein kinases, and oxidative stress proteins , beta-tubulin , GADPH, alpha-tubulin, and enolase (Bilic, et al., 2009; Hauck and Hafez, 2009a; Cepicka, et al., 2010). These genes have been identified either by PCR methods using sequence homology to genes identified in other protozoa or the sequencing of *H. meleagridis* cDNA library clones where proteins were identified by antibody response to antigens and plasmids eliciting responses were sequenced (Mazet, et al., 2008; Bilic, et al., 2009). Mazet, et al.

previously tried to look at the 3' region of hydrogenosome mRNA using oligodTs but were unable to discern anything from their short sequences (<34 bp) (Mazet, et al., 2008). These approaches do not yield DNA sequence information for the noncoding portions of the gene needed for proper expression. Thus no information is available on *H. meleagridis* regarding regulatory elements such as promoters, initiator elements, TATA boxes, and polyadenylation and cleavage signals.

In this paper, we present a method for isolating upstream and downstream sequences of the beta-tubulin genes in *H. meleagridis* using Splinkerette PCR (Devon, 1995). This technique uses DNA adaptors elements ligated to restricted genomic DNA. Splinkerettes are a variation on vectorettes, but they have a mismatch that forms hairpin structures to decrease end-repair priming and non-specific priming (Hengen, 1995). Nested primers to the adaptor and known genomic sequence allow PCR amplification of targeted regions of genomic DNA. Using this technique, partial sequences for two beta-tubulin genes, and 2 ORFs for a serine/threonine phosphotase and a ras-related protein, racG were identified.

Materials and Methods

Culturing of *Histomonas*

A *H. meleagridis* isolate from a backyard poultry flock located in Buford, GA was passed *in vitro* in modified Dwyer's medium containing 85% M199, 10% nonsterile deactivated horse serum, 0.1% sodium bicarbonate, and 1% rice powder. Cultures were grown at 40°C.

DNA isolation and Splinkerette PCR

DNA was isolated from *H. meleagridis* cultures using Qiagen DNeasy kit (Qiagen Inc., Valencia, CA) according to manufacturer's instruction. Splinkerette PCR method was adapted from Horn et al. (Horn C, 2007). 0.5 ug of *Histomonas* DNA was digested with either ApoI, BamHI, EcoRI, or XbaI

restriction endonucleases (Fermentas International Inc., Canada) followed by ligation to 10ng of an adaptor consisting of oligonucleotides (oligos) SpAa_ApoI and SpBb_ApoI for ApoI and EcoRI, adaptors consisting of SpBb_BamHI and SpAa_ApoI for BamHI, and adaptors consisting of SpBb_XbaI and SpAa_ApoI for XbaI using T4 DNA ligase (Fermentas, Glen Burnie, MD).

PCR amplification of ApoI and XbaI restricted DNA was performed under the following conditions. The first round of PCR was carried out in a 25uL reaction containing 12.5 uL of Dream Taq Green Master Mix (Fermentas, Glen Burnie, MD), 0.4ug of ligated DNA, Oligo Sp0F (5 pg), and Oligo H tubulin 0R (5 pg) (made from the GenBank accession no. FJ62493) for the ApoI reaction and oligo bTub0b (5 pg) for the XbaI reaction(Hauck and Hafez, 2009a). Cycling parameters for the amplification were 94°C for 2 minutes, 3 cycles of 94°C for 20 seconds, 64°C for 15 seconds, and 72°C for 5 minutes followed by 30 cycles of 94°C for 20 seconds, 58°C for 15 seconds, and 72°C for 3 minutes with a final extension period at 72°C for 10 minutes. The PCR reactions were diluted with 50uL of nuclease free water.

Second round of nested PCR was performed using a 25uL reaction containing 12.5uL Dream Taq Green Master Mix, 1uL of dilute first round PCR product, Oligo Sp1F (5 pg), and Oligo H tubulin 1R (5 pg) (made from the GenBank accession no. FJ62493) for the ApoI reaction and Oligo bTub1b (5 pg) for the XbaI reaction (Hauck and Hafez, 2009a). Cycling parameters for amplification of the second PCR product were 94°C for 2 minutes, 35 cycles of 94°C for 20 seconds, 60°C for 15 seconds, and 72°C for 3 minutes with a final extension period at 72°C for 5 minutes.

Upstream areas were amplified by PCR for EcoRI restricted DNA (Fermentas, Glen Burnie, MD) using 10uL of 5x LR buffer, 25ng of dNTPs, 5ng of oligo Sp0F, 5ng of oligo bTub0b, 1uL DMSO, .7uL of XL PCR enzyme mix, 0.25ug of the EcoRI DNA ligation, and nuclease free water up to 25uL. First round XL PCR was run with the parameters of one cycle of 92°C for 2 min, 10 cycles of 92°C for 10 sec, 63°C for 15

sec, and 68°C for 2 min, 25 cyles of 92°C for 10 sec, 63°C for 15 sec, and 68°C for 2 min + 20sec/cycle, final elongation of 68°C for 7 min, and hold at 4°C. XL PCR product was diluted in 50uL of nuclease free water and a second round of XL PCR was performed the same reagents as the first round except 1uL of dilute XL PCR product was added instead of the digested and adapted DNA oligo Sp1F and bTub1b. The PCR was run using the same parameters with the first round of XL PCR. The 3' region of beta-tubulin-2 was amplified using Splinkerette PCR method described above using oligos 1kb Splink F1 and 1kb Splink F2 and *H. meleagridis* DNA restricted with EcoRI and ligated to oligos SpAa_ApoI and SpBb_ApoI.

PCR amplicons were separated by gel electrophoresis using a 1% agarose gel containing ethidium bromide and visualized by UV light. Amplicons were gel-purified (Fermentas, Glen Burnie, MD) and cloned into Qiagen pDrive vector per manufacturer's instructions (Qiagen), and sequenced by The Georgia Genomics Facility (University of Georgia) using T7 and SP6 promoter oligos.

Sequence Analysis

Sequences were run through the Blastx or Blastn program with default settings (http://www.ncbi.nlm.nih.gov/blast/). Nucleic acid sequences were aligned using ApE, A Plasmid Editor (http://biologylabs.utah.edu/jorgensen/wayned/ape/). Amino acids for Figure 4 were aligned manually.

Results

Upstream sequences of the beta-tubulin gene Splinkerette PCR were identified using DNA digested with ApoI, XbaI, or EcoRI ligated to splinkerette adaptors (ApoI-SplA, XbaI-SplA, EcoRI-SplA) (Fig. 1). PCR with ApoI-SplA yielded an amplicon of 375 bp that included the ATG start site of the beta-tubulin-1 gene and 42 bp of upstream sequence (Fig. 2A). When comparing these results to known promoter lengths of genetically similar organisms, it was unlikely that the complete promoter was identified. Oligos and primers were redesigned to identify upstream sequence, and Splinkerette PCR

was performed on XbaI-SplA and EcoRI-SplA. An amplicon of 632 bp was obtained using XbaI-SplA and an amplicon of 1084 were obtained for EcoRI-SplA. Sequence analysis of the XbaI-SplA amplicon overlapped with the ApoI-SplA sequence and provided an additional 307 bp of upstream sequence. Blastx analysis of the sequence upstream of the beta-tubulin ATG identified an open reading frame of 129 bp at the start of the sequence, leaving 130 bp of intergenic sequence (Fig. 2A).

Sequence comparison of the EcoRI-SplA amplicon showed varying overlap with the entire ApoI-SplA and Xba-SplA beta-tubulin composite sequence (Fig. 3). Blastx of EcoRI-SplA sequence revealed an ORF of 375 bp at the beginning of the sequence and 334 bp that encodes for a beta-tubulin protein that includes the start ATG (Fig. 4A). As compared to the ApoI-SplA and Xba-SplA sequence, limited sequence similarity was seen upstream of the beta-tubulin ATG and 90% similarity was seen with the coding portion of the 299 bp of overlapping beta-tubulin sequences (Fig. 3). Based on these differences in nucleotide sequence and the homology of the second ORF to beta-tubulin is referred to the original gene as beta-tubulin-1 (bTub1) and the EcoRI-SplA sequence as beta-tubulin-2 (bTub2). As seen in Figure 4A and -B, this sequence includes an intergenic sequence of 434 bp between the btub2 ATG and the upstream ORF.

In order to determine downstream sequence variation between bTub1 and bTub2, PCR was performed on EcoRI-SplA using primers that amplify downstream sequence specific to btub2. An amplicon of 1145 bp was identified that overlapped with the original EcoRI-SplA sequence (Fig. 4A). Translation of the entire sequence yielded a beta-tubulin protein of 400 amino acids, which is 90% of the length of known beta-tubulin sequences. The EcoRI-SplA downstream underwent Blastn, it was found to have 99% homology with GenBank accession no. FJ624277 (Hauck and Hafez, 2009a).

Comparison of the amino acid sequences obtained for bTub1 and bTub2 to each other and representative member of the beta-tubulin family for *T. vaginalis* (Genbank accession no. AAA67546)

and *G. lamblia* (Genbank accession no. XP_001707375) shows the sequence homology (Fig 5) (Katiyar and Edlind, 1994; Morrison, 2007). BTub1 codes for 256 amino acids, and bTub2 codes for 400. For bTub1, an E-value of 3e-129 and an identity of 237/256 (93%) was observed concerning *T. vaginalis*, and an E-value of 6e-126 with an identity of 206/256 (80%) was observed with *G. lamblia*. For bTub2, an E-value of 0 and an identity of 380/400 (95%) was observed concerning *T. vaginalis*, and an E-value of 0 with an identity of 323/400 (80%) was observed with *G. lamblia*.

In order to identify the proteins encoded for by the ORFs found upstream of the beta-tubulins Blastx was performed. The ORF upstream of our bTub1showed limited homology to a 203 amino acid long Ras-related protein, racG, of *T. vaginalis* (GenBank accession no. XP_001305646) with an E-value of 0.77 and 46% identity over 46 amino acids (Carlton et al., 2007). Our ORF showed homology to the last 42 amino acids of the C-terminal portion of the protein. The ORF upstream of btub2 had high homology with Ser/Thr protein phosphatase of *Trichomonas vaginalis* (GenBank accession no. XP_001330801) with an E-value of 1e-47 and 73% identity over 124 amino acids (Carlton, et al., 2007) (Figure 5). This ORF showed homology with the last 124 amino acids of the C-terminal of the total 339 amino acid sequence.

Based on our analysis, each sequence contains the putative promoter elements for beta-tubulin and the stop codon and polyadenylation and cleavage signals for the upstream ORF. Using the consensus eukaryotic Inr sequence, we identified one possible Inr in the intergenic region of putative bTub1 with the sequence of TTATTTT 4 bp upstream of the start site. Using the same parameters for the intergenic region of bTub2, we identified 2 possible Inrs that overlap. One is 16 bp upstream of the start site with a sequence of TTAAACC. The other is 21 bp upstream of ATG with a sequence of TCAAATT. Intergenic regions varied in size with bTub1's sequence of 220 and bTub2's of 376. Sequence analysis found a consensus sequence of ~24 bp within the intergenic region (Figure 2C). The end of the ORFs upstream of the beta-tubulins contain stop codons similar to that of *T. vaginalis* of TAA. Both stop sites

are followed by an adenosine which could be a putative polyadenylation signal like that of *T. vaginalis* (Espinosa, et al., 2002). Downstream 6 and 12 bp from the TAAA, both sequences contain a TTTC.

Discussion

This research demonstrates the use of splinkerette PCR, a type of ligated adaptor PCR, to amplify unknown sequences of genomic DNA in unsequenced Trichomonad species. Using this technique, the 5' intergenic region of the bTub1 and bTub2 and the 3' region of bTub2 of *H. meleagridis* was identified. Additionally, we identified 2 ORFs upstream of bTub1 and bTub2 that show homology to *T. vaginalis* proteins racG and Ser/Thr phosphatase. We propose that within these intergenic regions, we have identified putative expression elements for the beta-tubulin genes and the polyadenylation and cleavage signals. Splinkerette technology will be useful in future studies for isolating unknown sequence flanking known sequence in organisms in which little genetic information is known.

The original study on beta-tubulin of *H. meleagridis* contained 21 different nucleic acid sequences and 11 translation sequences (Hauck and Hafez, 2009a). However, these were only partial sequences of the gene so it was unknown if the variations in sequence they observed were due to differences between sequences of the beta-tubulin gene or the presence of multiple beta-tubulin genes within the *H. meleagridis* genome. Our work expanded upon their sequence information by providing upstream sequence for the beta-tubulin gene. We identified two different beta-tubulins by isolating and sequencing flanking genomic sequence and observing divergent upstream sequence. Based on this data, *T. vaginalis* having 6-7 copies of the beta-tubulin gene, *G. lambia* having 3 copies of the gene, and the multiple different sequences Hauck and Hafez identified, we can speculate that *H. meleagridis* probably has more than 2 copies of the beta-tubulin gene within its genome (Kirk-Mason et al., 1989; Katiyar and Edlind, 1994; Carlton, et al., 2007; Hauck and Hafez, 2009a). Future experiment using splinkerette PCR could be to amplify region upstream or downstream of known regions of

differentiation and use sequence analysis to allow cDNA and partial sequences to be associated with specific copies of genes.

The identification and evaluation of expression elements is important for further study of virulence factors and other genetic information of *H. meleagridis*. Unlike metazoans that can have regulatory elements far away from the beginning or end of the coding regions, protozoans have all expression elements within their short intergenic regions (Vanacova, et al., 2003). This suggests that the sequences associated with the intergenic regions bTub1 and bTub2 contain all of the expression elements needed to express a gene. Similar to other protozoans, no TATA elements were observed. Instead, an initiator element (Inr) was found providing a place for the binding of RNA polymerase II. Inr are common among almost all metazoans, and have a consensus sequence of $YYA^{+1}NWYY$. Ideally, we compare these sequences to another member of the Dientamoebidae family (*Dientamoeba, Ptrotrichomonas* or *Parahistomonas*), but these, too, have received little attention from the scientific community. In *Trichomonas vaginalis*, its Inr sequence has further been characterized as $TCA^{+1}YTWYTCATTA$, and *Giardia lamblia*'s Inr is $ATTTTA^{+1}AAAT$ (Quon, et al., 1994; Yee et al., 2000). However, because we only have two sequences, we cannot conclusively identify such a pattern other than Inr for each gene was generally found 4-17 bp upstream of ATG. With the variation in size of the intergenic regions upstream of the beta-tubulins, it will be intriguing to note the difference in expression levels in later studies. Interestingly, we observed a large consensus sequence in both intergenic regions, which could be a *cis*-acting regulatory element.

Polyadenylation and cleavage signals are necessary for the proper expression of mRNAs. The identification of the 3' end of ORF provides us with information concerning these elements through sequence analysis of the site. Similar to *T. vaginalis*, a TAAA was observed in the ORF, leading us to believe this is a putative polyadenylation signal (Espinosa, et al., 2002). Although this sequence was

similar, we did not observe the consensus *T. vaginalis* cleavage signal. Instead, a conserved sequence of TTTC was observed. Analysis of more sequences 3' to a coding region will be needed to determine if this is the cleavage signal for *H. meleagridis*. The same holds true for the intergenic region of our sequences for the motif TAAA, our putative polyadenylation signal.

This paper provides the foundation for future experiments aimed at analyzing gene expression in histomonads using expression vectors. Development of these constructs is important for studying virulence genes, drugs that can act upon gene expression, and to determine other genetic information. Splinkerette PCR is used to discern non-expressed sequences through ligated adaptor PCR in organisms whose genome remains unsequenced. This could also provide the information necessary to study gene expression in other little-studied organisms.

Acknowledgements

We thank Dr. Larry McDougald for the *Histomonas* isolate, insights into blackhead disease and for critical reading of the manuscript. We thank Dr. Lorraine Fuller for technical assistance.

LITERATURE CITED

BILIC, I., M. LEBERL, and M. HESS. 2009. Identification and molecular characterization of numerous Histomonas meleagridis proteins using a cDNA library. *Parasitology, 136*(4), 379-391. doi: S0031182008005477 [pii] 10.1017/S0031182008005477

CARLTON, J. M., R. P. HIRT, J. C. SILVA, A. L. DELCHER, M. SCHATZ, Q. ZHAO. 2007. Draft genome sequence of the sexually transmitted pathogen Trichomonas vaginalis. [Article]. *Science, 315*(5809), 207-212. doi: 10.1126/science.1138294

CEPICKA, I., V. HAMPL, and J. KULDA. 2010. Critical Taxonomic Revision of Parabasalids with Description of one New Genus and three New Species. [Article]. *Protist, 161*(3), 400-433. doi: 10.1016/j.protis.2009.11.005

DEVON, R. S., D.J. PORTEOUS, AND A.J. BROOKES. 1995. Splinkeretts--improved vectorettes for greater efficiency in PCR walking. *Nucleic Acids Research, 23*(9), 1644-1645.

ESPINOSA, N., R. HERNANDEZ, L. LOPEZ-GRIEGO, and I. LOPEZ-VILLASENOR. 2002. Separable putative polyadenylation and cleavage motifs in Trichomonas vaginalis mRNAs. [Article]. *Gene, 289*(1-2), 81-86.

FELLEISEN, R. S. J. 1997. Comparative sequence analysis of 5.8S rRNA genes and internal transcribed spacer (ITS) regions of trichomonadid protozoa. *Parasitology, 115*(2), 111-119.

GERBOD, D., V. P. EDGCOMB, C. NOEL, L. ZENNER, R. WINTJENS, P. DELGADO-VISCOGLIOSI. 2001. Phylogenetic position of the trichomonad parasite of turkeys, Histomonas meleagridis (Smith) Tyzzer, inferred from small subunit rRNA sequence. *J Eukaryot Microbiol, 48*(4), 498-504.

HAUCK, R., and H. M. HAFEZ. 2009. Partial sequence of the beta-tubulin of Histomonas meleagridis and the activity of benzimidazoles against H-meleagridis in vitro. *Parasitology Research, 104*(5), 1183-1189. doi: DOI 10.1007/s00436-008-1309-5

HENGEN, P. N. 1995. METHODS AND REAGENTS - VECTORETTE, SPLINKERETTE AND BOOMERANG DNA AMPLIFICATION. [Note]. *Trends in Biochemical Sciences, 20*(9), 372-373.

HORN C, H. J., SCHNUTGEN F, SEISENBERGER C, FLOSS T, IRGANG M, DE-ZOLT S, WURST W, VON MELCHNER H, NOPPINGER PR. 2007. Splinkerette PCR for more efficient characterization of gene trap events. *Nature Genetics, 39*(8), 933-934.

KATIYAR, S., and T. EDLIND. 1994. Beta-tubulin genes of *Trichomonas vaginalis*. *Molecular and Biochemical Parasitology, 64*, 33-42.

KIRK-MASON, K. E., M. J. TURNER, and P. R. CHAKRABORTY. 1989. Evidence for unusually short tubulin mRNA leaders and characterization of tubulin genes in *Giardia lamblia*. *Molecular and Biochemical Parasitology, 36*, 87-100.

LEBERL, M., M. HESS, and I. BILIC. 2010. Histomonas meleagridis possesses three alpha-actinins immunogenic to its hosts. *Mol Biochem Parasitol, 169*(2), 101-107. doi: 10.1016/j.molbiopara.2009.10.007

MAZET, M., M. DIOGON, J. F. ALDERETE, C. P. VIVARES, and F. DELBAC. 2008. First molecular characterisation of hydrogenosomes in the protozoan parasite Histomonas meleagridis. *Int J Parasitol, 38*(2), 177-190. doi: S0020-7519(07)00224-X [pii] 10.1016/j.ijpara.2007.06.006

MCDOUGALD, L. R. 2005. Blackhead disease (histomoniasis) in poultry: a critical review. *Avian Dis, 49*(4), 462-476.

MINOTTO, L., M. R. EDWARDS, and A. S. BAGNARA. 2000. Trichomonas vaginalis: Characterization, expression, and phylogenetic analysis of a carbamate kinase gene sequence. [Article]. *Experimental Parasitology, 95*(1), 54-62.

MORRISON, H., *ET AL.* 2007. Genomic Minimalism in the Early Diverging Intestinal Parasite *Giardia lamblia*. *Science, 317*, 1921-1926.

PEATTIE, D., ROGELIO ALONSO, ANN HEIN, JOHN CAULFIELD. 1989. Ultrastructural Localization of Giardins to the Edges of Disk Microribbons of Giardia lamblia and the Nucleotide and Deduced Protein Sequence of Alpha Giardin. *Journal of Cell Biology, 109*(5), 2323-2335.

QUON, D. V. K., M. G. DELGADILLO, A. KHACHI, S. T. SMALE, and P. J. JOHNSON. 1994. SIMILARITY BETWEEN A UBIQUITOUS PROMOTER ELEMENT IN AN ANCIENT EUKARYOTE AND MAMMALIAN INITIATOR ELEMENTS. [Article]. *Proceedings of the National Academy of Sciences of the United States of America, 91*(10), 4579-4583.

SMALE, S. T. 1997. Transcription initiation from TATA-less promoters within eukaryotic protein-coding genes. [Review]. *Biochimica Et Biophysica Acta-Gene Structure and Expression, 1351*(1-2), 73-88.

TYZZER, E. E., and M. FABYAN. 1922. A Further Inquiry into the Source of the Virus in Blackhead of Turkeys, Together with Observations on the Administration of Ipecac and of Sulfur. *J Exp Med, 35*(6), 791-812.

VAN DER HEIJDEN, H., and W. J. M. LANDMAN. 2007. Improved Culture of Histomonas meleagridis in a Modification of Dwyer Medium. [Article]. *Avian Diseases, 51*(4), 986-988.

VANACOVA, S., D. R. LISTON, J. TACHEZY, and P. J. JOHNSON. 2003. Molecular biology of the amitochondriate parasites, Giardia intestinalis, Entamoeba histolytica and Trichomonas vaginalis. [Review]. *International Journal for Parasitology, 33*(3), 235-255. doi: 10.1016/s0020-7519(02)00267-9

YEE, J., M. R. MOWATT, P. P. DENNIS, and T. E. NASH. 2000. Transcriptional analysis of the glutamate dehydrogenase gene in the primitive eukaryote, Giardia lamblia - Identification of a primordial gene promoter. [Article]. *Journal of Biological Chemistry, 275*(15), 11432-11439.

ZAMORANO, A., C. LOPEZ-CAMARILLO, E. OROZCO, C. WEBER, N. GUILLEN, and L. A. MARCHAT. 2008. In silico analysis of EST and genomic sequences allowed the prediction of cis-regulatory elements for Entamoeba histolytica mRNA polyadenylation. [Article]. *Computational Biology and Chemistry, 32*(4), 256-263. doi: 10.1016/j.compbiolchem.2008.03.019

Table 1. Oligos and Primers Used in PCR.

Oligo	Sequence
SpAa_ApoI	CGAAGAGTAACCGTTGCTAGGAGAGACCGTGGCTGAATGAGACTGGTGTCGACA CTAGTGG
SpBb_ApoI	AATTCCACTAGTGTCGACACCAGTCTCTAATTTTTTTTTTCAAAAAAA
SpBb_XbaI	CTAGCCACTAGTGTCGACACCAGTCTCTAATTTTTTTTTTCAAAAAAA
SpBb_BamHI	GATCCCACTAGTGTCGACACCAGTCTCTAATTTTTTTTTTCAAAAAAA
Primer	**Sequence**
Sp0F	CGAAGAGTAACCATTACTAGGAGAGACC
Sp1F	GTGGCTGAATGAGACTGGTGTCGAC
H tubulin 1r	TGCCACCACCAAGAGAGTGAACAAG
H tubulin 0r	TGGGTGTGGTGAGCTTTAAGGTACG
bTUB0b	ACCATATGTGGGTGTGGTGAGCTT
bTUB1b	TTACCAGCACCAGATCGACCGAAA
1kbSplink F1	ATGGTCAGAGAGATTGTTCACCTCC
1kbSplink F2	CCAATGTGGTAACCAAATCGGCG

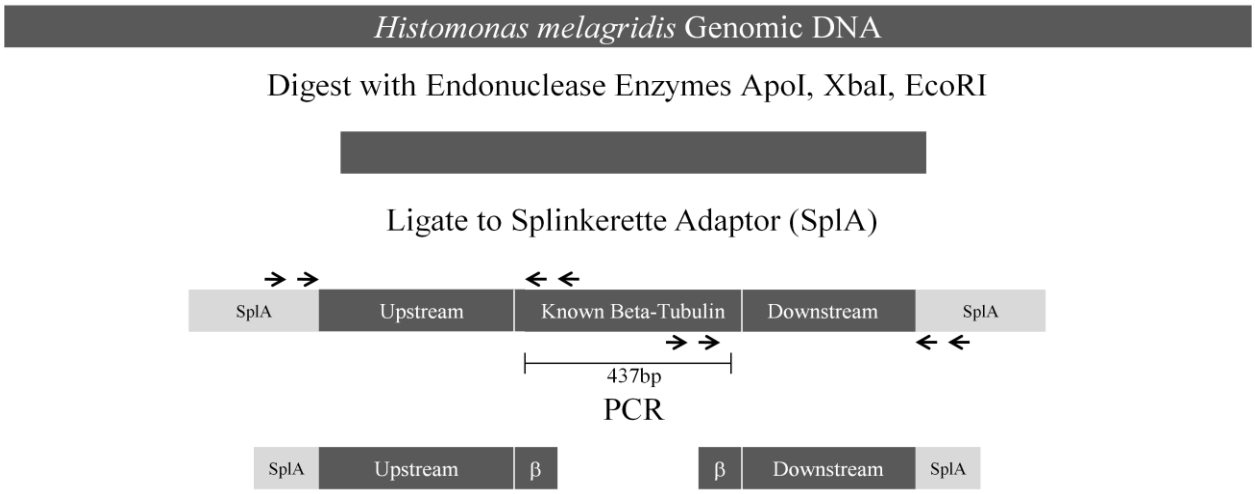
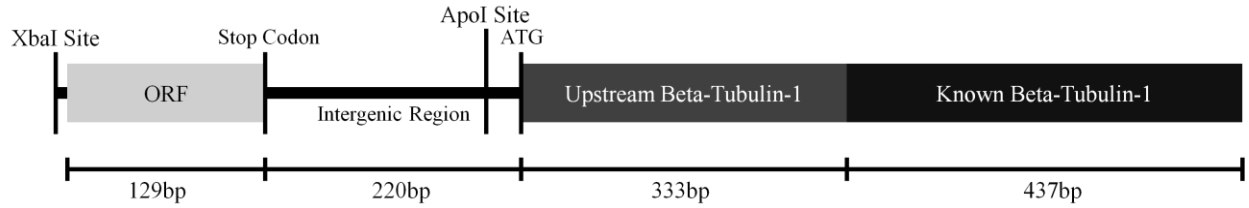


Figure 1: Principles of Splinkerette PCR. Ligation of splinkerette to digested DNA allows for amplification of nested sequences of DNA when used in conjunction with an oligo against known sequence. → Symbolize primers

A.



B.

TACTATTGAG TAAACGATTT TTCGAAACGA ATCATAAAGA TTTCTTTTAT

CTCTTAACCA ACCTTTAGAA GTTGAACCAG CTTCATTAAG CTTTCGGAAA

AGAACTAATT CATCAAATAA AGGAACTAAA AATGACCTAA AATTGTTAAG

CCTATCAGTG ATGAAGGTAC GTTTAACATT TTGAAGATAA ATTTATGAAA

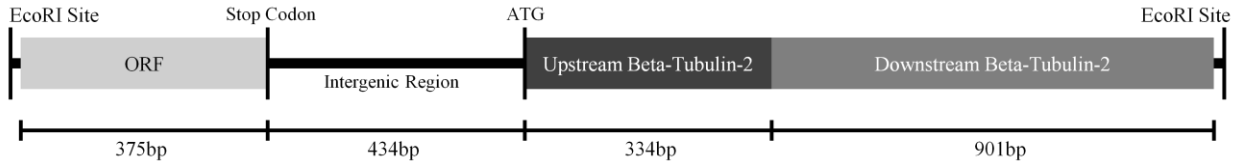
TTTCAGAGAA ATACGTTTTA ATTTATTTTT AGAATGGTGC GTGAAATTGT

Figure 2: Beta-Tubulin-1. A. Graphical representation of beta-tubulin-1 as discussed. B. Nucleotides of intergenic region beginning with the stop codon of the Ras-related ORF to the start codon of beta-tubulin-1. Potential Inr has a bar below the sequence. △ marks for putative sites for transcription initiation. Black rectangle marks stop site. Black oval marks start site. Boxed area is putative polyadenylation signal.

```
bTub1      GACCTAAAAT TGTTAAGCCT ATCAGTGATG AAGGTACGTT TAACATTTTG
bTub2      TTGCTCGATA AATGTTTATA AATAATTAAC GGTCGAGTGG TTTTGTGTTT

bTub1      AAGATAAATT TATGAAATTT CAGAGAAATA CGTTTTAATT TATTTTTAGA
bTub2      AACACTTTTT CAAGAGGAAA ATTCAAATTA AACCGTTTTT GACTTCACAA

bTub1      ATGGTGCGTG AAATTGTTCA CCTCCAAGCA GGACAATGTG GAAACCAAAT
bTub2      ATGGTCAGAG AGATTGTTCA CCTCCAAGCA GGCCAATGTG GTAACCAAAT

bTub1      CGGTGCCAAG TTCTGGGAAG TTACCTCAGA TGAACACGGC ATTGATCCAA
bTub2      CGGCGCTAAA TTCTGGGAAG TTATTTCAGA TGAACATGGT ATTGACCCAA

bTub1      CCGGTTCATT CCATGGTGAT AGTGATCTTC AATTAGAAAG AATTAACGTT
bTub2      CAGGTTCATT TCATGGCGAT AGTGATCTTC AATTAGAAAG AATCAACGTT

bTub1      TACTACAACG AAGCCACTGG TGGTAAGTAC GTTCCACGTG CTGTCCTTGT
bTub2      TATTATAACG AAGCCACTGG TGGTAAGTAT GTCCCACGTG CTGTCCTTGT

bTub1      CGATCTTGAA CCAGGTACTA TGGATGCAGT TCGTGCAGGA CAATATGGTC
bTub2      TGATCTTGAA CCAGGTACCA TGGATGCTGT TCGTGCTGGT CAATATGGCC

bTub1      AATTATTCCG CCCAGATAAC TTCGCTTTCG GTCGATCTGG TGCTGGTAA
bTub2      AATTATTCCG CCCAGATAAC TTCGTTTTCG GTCAATCTGG TGCTGGTAA
```

Figure 3: Alignment of 100 bp upstream and the coding region from start ATG to bTub1b primer of beta-tubulin-1 (bTub1) and beta-tubulin-2 (bTub2) of *Histomonas meleagridis*. The differences (shading) in the promoter regions strongly suggest 2 distinct beta-tubulin genes. Start site is underlined and primer region is boxed.

A.



B.

AAAACTACCT TAAACAGTTA ATACATTTCA AGTTTTTTTT ATTTGTGGAT

AACATTTATT TTATGAAGTT TCGTGTTATG ATTAAAAATG GTAATGGAAT

ATCGATCAAT TTCTTGATTT TTATTATTTT AATAATACTT AAAAATATTA

AGTAATGAGT GAAATTATTT TATACGGCAC TAATAATTTT AAATTTTGTA

AACAAAAAAT TTCTAAATTA AATCATTAGA AAATATAATT TCCTTAATCA

AATAATATAT CAAATAAAAA AGAGTATATT TTTATTAAAG GTTTCTTTTT

GTGAACCGAT GATTTAGCTA AACACTTTGC GATATGTGTA CTTAAAGTTG

CTCGATAAAT GTTTATAAAT AATTAACGGT CGAGTGGTTT TGTGTTTAAC

ACTTTTTCAA GAGGAAAATT CAAATTAAAC CGTTTTTGAC TTCACAAATG

GTCAGAGAGA

C.

```
bTub1   ATCAAATAA---AGGAACTAAAAA
bTub2   ATCAAATAATATATCAAATAAAAA
```

Figure 4: Beta-Tubulin-2. A. Graphical representation of beta-tubulin-2 as discussed. B. Nucleotides of intergenic region beginning 10 bp upstream of the stop codon of the ORF to 10 bp downstream of the start codon of beta-tubulin-2. Two potential Inrs: one has a bar below the sequence, the other has a bar above. △ marks for possible sites for transcription iniation. Black rectangle marks stop site. Black oval marks start site. Boxed area is putative polyadenylation signal. C. Region of consensus sequence in intergenic regions. Shading indicates mismatches.

```
Tv    MVREIVHIQAGQCGNQIGAKFWEVISDEHGIDPTGSYHRDSDLQLERINVYYNEATGAKY      60
Gl    M-REIVHIQAGQCGNQIGAKFWEVISDEHGVDPSGEYRGDSELQIERINVYFNEAAGGRY      59
Hm1   MVREIVHLQAGQCGNQIGAKFWEVTSDEHGIDPTGSFHGDSDLQLERINVYYNEATGGKY      60
Hm2   MVREIVHLQAGQCGNQIGAKFWEVISDEHGIDPTGSFHGDSDLQLERINVYYNEATGGKY      60

Tv    VPRAILVDLEPGTSESVRAGQFGQLFRPDNFVFGQSGAGNNWAKGYYTEGQELCESILDV      120
Gl    VPRAILVDLEPGTMDSVRAGPFGQIFRPDNFVFGQSGAGNNWAKGHYTEGAELVDAVLDV      119
Hm1   VPRAVLVDLEPGTMDAVRAGQYGQLFRPDNFAFGRSGAGNNWAKGYYTEGQELAESILDV      120
Hm2   VPRAVLVDLEPGTMDAVRAGQYGQLFRPDNFVFGQSGAGNNWAKGYYTEGQELAESILDV      120

Tv    IRKEAESCDCLQGFQLVHSLGGGTGAGLGTLLLNKLREEYPDRILSTYSIVPSPKVSDTV      180
Gl    VRKESEACDCLQGFQICHSLGGGTGAGMGTLLIAKIREEYPDRMMCTFSVVPSPKVSDTV      179
Hm1   VRKEAESCDALQGFQLVHSLGGGTGSGLGTLFLNKLREEYPDRILSTYSIVPSPKVSDTV      180
Hm2   VRKEAESCDALQGFQLVHSLGGGTGSGLGTLLLNKLREEYPDRILSTYSIVPSPKVSDTV      180

Tv    VEPYNCTLSVHQLVESADEVFCIDNEALYDICFRTLKLTTPTYGDLNHLVSMVMSGTTCA      240
Gl    VEPYNATLSVHQLVEHADEVFCIDNEALYDICFRTLKLTCPTYGDLNHLVSLVMSGCTSC      239
Hm1   VEPYNCTLSVHQLVESVDEVFCIDNEALYDICFRTLKLTTPTYGDLNHLVSMVMSGTTCS      240
Hm2   VEPYNCTLSVHQLVESADEVFCIDNEALYDICFRTLKLTTPTYGDLNHLVSMVMSGTTCS      240

Tv    LRFPGQLNSDLRKLAVNLVPFPRLHFFIVGFAPLTSRGSQQYRALTVPELTSQLFDNKNM      300
Gl    LRFPGQLNADLRKLAVNLIPFPRLHFFLVGFAPLTSRGSQIYRALTVPELVSQMFDNKNM      299
Hm1   LRFPGQLNSDLRKLA.........................................          256
Hm2   LRFPGQLNSDLRKLAVNLVPFPRLHFFIVGFAPLTSRGSPQYRALTVPELTSQLFDNKNM      300

Tv    MAACDPRRVSYLT-SAHFRGRMSSKEVDEQMLNIQARNTSYFVEWIPSNVKSAICDIPPR      359
Gl    MAASDPRHGRYLTAAAMFRGRMSTKEVDEQMLNIQNKNSSYFVEWIPNNMKVSVCDIPPR      359
Hm1   ..................................................................
Hm2   MAACDPRRGVYLTVSAHFRGRMSSKEVDEQMLNIQARNTSYFVEWIPNNVKSSICDIPPR      360

Tv    GLKMAATFIGNTTAFRELFTRVDSQFQKMYARRAFIHWYV                          399
Gl    GLKMAATFIGNSTCIQELFKRVGEQFTAMFRRKAFLHWYT                          399
Hm1   ....................................
Hm2   GLKMAATFIGNTTAFRELFTRVDSQFQKMYARRAFIHWYV                          400
```

Figure 5: Beta-tubulin alignments of amino acid coding sequences of the *Trichomonas vaginalis* (Tv), *Giardia lamblia* (Gl), *Histomonas meleagridis* beta-tubulin-1 (Hm1), and *H. meleagridis* beta-tubulin-2 (Hm2). Shading signifies mismatches in sequence. Dashes (-) mark misalignments. Periods (.) signify lack of sequence.

```
Tv    PIEGHHGVYFGEDVLENFFKLNGLTCMIRSKQLCMNGNTTSLNGKCITIWSAPNFNGWIQNA    62
Hm    PIENRHGYFFGEPKIDKFLKDNGLSCIIRSKQLCMNGHTNMFNGKCITVWPAPNFCGWIQNA    62

Tv    ASVVQMYAPTSQFGQKSDQKFPLMINTYKARPESERIERNRPPFQDVHSLDHIYIKHLPKYP    124
Hm    ASVVQMHAQTSQFGQKGGSAQTYTINTFKARPESERIEQNRPPFXDVHSLDHIYIKHLPKLP    124
```

Figure 6: Amino acid alignment of the last 124aa of Ser/Thr protein phosphatase of *Trichomonas vaginalis* (Tv) (GenBank accession no. XP_001330801) and 124aa ORF from *Histomonas meleagridis* (Hm). Shading signifies mismatching in sequence.

CHAPTER 4

CONCLUSIONS

This research demonstrates the use of splinkerette PCR, a ligated adaptor PCR, to amplify unknown sequences of genomic DNA in unsequenced Trichomonad species. Using this technique, the 5' intergenic region of the bTub1 and bTub2 and the 3' coding region of bTub2 in *H. meleagridis* was identified. In addition, we identified 2 ORFs upstream of bTub1 and bTub2 that show homology to *T. vaginalis* proteins racG and Ser/Thr phosphatase. We believe that within these intergenic regions, we have identified putative expression elements for the beta-tubulin genes and their polyadenylation and cleavage signals. This application of splinkerette technology will be useful in future studies for isolating unknown sequence flanking known sequence in organisms in which little genetic information is known.

The original study on beta-tubulin of *H. meleagridis* identified 21 different nucleic acid sequences and 11 amino acid sequences (Hauck and Hafez, 2009). However, these were only partial sequences of the gene so it was inconclusive on the number of genes contained within the *H. meleagridis* genome. Our work expanded on their information by providing upstream sequence for the beta-tubulin gene. We identified two different beta-tubulins by isolating and sequencing upstream divergent sequence. Based on this data, related organisms having 3 -7 copies of the gene, and the multiple different sequences Hauck and Hafez identified, we can speculate that *H. meleagridis* probably has more than 2 copies of the beta-tubulin gene within its genome (Kirk-Mason et al., 1989; Katiyar and Edlind, 1994; Carlton et al., 2007; Hauck and Hafez, 2009). Future experiment using splinkerette PCR

could be to amplify region upstream or downstream of known regions of differentiation and use sequence analysis to allow cDNA and partial sequences to be associated with specific copies of genes.

The identification and evaluation of expression elements is important for further study of virulence factors and other genetic information of *H. meleagridis*. Protozoans, unlike metazoans that can have regulatory elements distal to the beginning or end of the coding regions, have all expression elements within their short intergenic regions (Vanacova et al., 2003). This suggests that the intergenic regions upstream of bTub1 and bTub2 contain all of the expression elements needed to express a gene. Instead of a TATA element, an initiator element (Inr) was found providing a place for binding RNA polymerase II. Inr are common among almost all metazoans, and have a consensus sequence of $YYA^{+1}NWYY$. Because we only have two sequences, we cannot conclusively identify such a pattern other than the metazoan consensus Inr for each gene was generally found 4-17 bp upstream of ATG. With the variation in size of the intergenic regions upstream of the beta-tubulins, it will be intriguing to note the difference in expression levels in later studies. In addition to promoter expression elements, polyadenylation and cleavage signals are necessary for the proper expression of mRNAs. The identification of the 3' end of ORF provides us with information concerning these elements. Similar to *T. vaginalis*, a TAAA was observed in the ORF, leading us to believe this is a putative polyadenylation signal (Espinosa et al., 2002). Although this sequence was similar, we did not observe the consensus *T. vaginalis* cleavage signal.

This paper provides the foundation for future experiments aimed at analyzing gene expression in histomonads using expression vectors. Development of these constructs is important for studying virulent genes, drugs that can act upon gene expression, as well as determining other genetic information. Splinkerette PCR is used to discern non-expressed sequences through ligated adaptor PCR in organisms whose genome remains unsequenced. This could also provide the information necessary to study gene expression in other little-studied organisms.

LITERATURE CITED

CARLTON, J. M., R. P. HIRT, J. C. SILVA, A. L. DELCHER, M. SCHATZ, Q. ZHAO. 2007. Draft genome sequence of the sexually transmitted pathogen Trichomonas vaginalis. [Article]. *Science, 315*(5809), 207-212. doi: 10.1126/science.1138294

ESPINOSA, N., R. HERNANDEZ, L. LOPEZ-GRIEGO, and I. LOPEZ-VILLASENOR. 2002. Separable putative polyadenylation and cleavage motifs in Trichomonas vaginalis mRNAs. [Article]. *Gene, 289*(1-2), 81-86.

HAUCK, R., and H. M. HAFEZ. 2009. Partial sequence of the beta-tubulin of Histomonas meleagridis and the activity of benzimidazoles against H-meleagridis in vitro. *Parasitology Research, 104*(5), 1183-1189. doi: DOI 10.1007/s00436-008-1309-5

KATIYAR, S., and T. EDLIND. 1994. Beta-tubulin genes of *Trichomonas vaginalis*. *Molecular and Biochemical Parasitology, 64*, 33-42.

KIRK-MASON, K. E., M. J. TURNER, and P. R. CHAKRABORTY. 1989. Evidence for unusually short tubulin mRNA leaders and characterization of tubulin genes in *Giardia lamblia*. *Molecular and Biochemical Parasitology, 36*, 87-100.

VANACOVA, S., D. R. LISTON, J. TACHEZY, and P. J. JOHNSON. 2003. Molecular biology of the amitochondriate parasites, Giardia intestinalis, Entamoeba histolytica and Trichomonas vaginalis. [Review]. *International Journal for Parasitology, 33*(3), 235-255. doi: 10.1016/s0020-7519(02)00267-9