2015

Determination of the Dermacentor variabilis Ferritin Gene Exon and Intron Structure

Dana Sylvestre
Georgia Southern University

Follow this and additional works at: https://digitalcommons.georgiasouthern.edu/honors-theses

Part of the Molecular Genetics Commons

Recommended Citation
https://digitalcommons.georgiasouthern.edu/honors-theses/138
Determination of the Dermacentor variabilis Ferritin Gene Exon and Intron Structure:

An Honors Thesis submitted in partial fulfilment of the requirements for Honors in Biology

By Dana Sylvestre

Under the mentorship of Dr. Quentin Q. Fang

The American dog tick *Dermacentor variabilis* is a hard body, 3-host tick that relies on the blood of small and large mammals to satisfy its metabolic demands. The excess iron consumed in its blood diet can cause cellular damage through oxidation. Ticks rely on the ferritin protein to metabolize iron. For this reason, it is important to further characterize the ferritin gene. After extracting DNA from *D. variabilis* ticks, the suspected ferritin gene was amplified through polymerase chain reaction (PCR) and purified. The putative ferritin DNA fragment was inserted into plasmids and isolated in a long PCR length gene cloning. Overlapping DNA fragments were sequenced from the genomic DNA of the entire ferritin coding region and aligned with cDNA of the *D. variabilis* ferritin heavy chain homologue (HCH) gene to determine the intron/exon structure. The results revealed that the *D. variabilis* ferritin gene coding region has three exons and two introns. The entire *D. variabilis* ferritin (HCH) gene coding region is 3,805 base pairs long. The three exons are 106 bp, 277 bp, and 135 bp in length, respectively, with a G+C content around 53%. The two introns are 2268 and 1019 bp in length, with an average A+T content of 56%. It is hypothesized that the long first intron of the *D. variabilis* ferritin gene might play a role in alternative splicing and possibly also ferritin gene expression.

Thesis Mentor: ____________________________

Dr. Quentin Q. Fang

Honors Director: ____________________________

Dr. Steven Engel

April 2015
Acknowledgements:

I would like to thank Dr. Quentin Q. Fang for the investment of his time and resources, as well as his expertise in primer construction, sequence alignment, and gene cloning. I also want to thank the students, Weibo Fu, Kayla Perry, Nick Sasser, and Heather Lambert, who worked in Dr. Fang’s lab, for their contributions to this project.
Introduction:

Tick

Ticks are small arachnids in the order Acari, along with mites, spiders, scorpions, and kin, they constitute the class Arachnida. Ticks are a part of the suborder Ixodida, and they belong to the Superfamily Ixodoidea. Three families are identified in ticks, which are the Ixodidae (hard ticks), Argasidae (soft ticks), and Nuttalliellidae (one species) (Hoogstraal and Aeschlimann 1982; Horak et al. 2002, Nava et al. 2009). There are over 900 tick species within these three tick families. Ticks are solely blood-feeder ectoparasites (external parasites), and they rely on the blood of mammals, birds, reptiles, and amphibians for their nutrition and development (Oliver 1989; Sonenshine 1991, Sonenshine and Roe 1993, Klompen et al. 2000). The ticks within the family Ixodidae are hard-bodied ticks, which consist of two major groups, the Prostriata (subfamily Ixodinae) and Metastriata (subfamilies Amblyomminae, Haemaphysalinae, Hyalomminae, and Rhipicephalinae) (Nava et al. 2009).

Most of the ixodid ticks are three-host ticks, and the others are one-host ticks. The three-host tick lifecycle consists of one separate host during the larval stage, nymphal stage, and the adult stage, as well as dropping off of its host once it is engorged for molting between stages on the ground (Hoogstraal and Aeschlimann 1982). This lifecycle takes roughly two to three years to complete (Sonenshine and Roe 1993). The one-host ticks will lay their eggs in the surrounding environment, and the larvae will attach to their
host after hatching. They will spend all of their life stages on this single animal. Ticks are unique in comparison to other hematophagous arachnids, in that they are known to transmit a larger variety of pathogens, have a longer lifespan that takes place over years, and will engorge to much larger sizes (4-5 mL vs. < 1 mL) after a feeding period (Sonenshine 1991).

Certain tick species commonly feed on mammals, but will also feed on humans (Parola and Raoult 2001). The long feeding period that ixodid ticks experience contributes to their function as vectors and reservoirs for many pathogens (Parola and Roult 2011). Proteins in the tick’s saliva can attach to cytokines and inhibit the cytokines and chemokines that are involved in the immune response, which can delay detection by the host (Francischetti et al. 2009). Also, a variety of other chemicals in the tick’s saliva, such as anticoagulants, prostaglandins, and antihistamines inhibit the host’s vasoconstriction and platelet aggregation (Sonenshine 1991). By inhibiting these and other vital functions, the tick is able to weaken the host’s inflammatory response and immune cellular response to the cell injury at the lesion caused by tick feeding (Sonenshine 1991). Also, in its immature, unfed form, the ixodid tick is rather small, but its size will rapidly increase upon blood ingestion. For example, the *Ixodes ricinus* tick can increase in weight from 2 mg to roughly 250-450 mg after feeding (Arthur 1962). The tick’s small size at the beginning of feeding can help it evade detection (Parola and Raoult 2001).

Some of the potentially lethal diseases caused by ixodid ticks include tularemia, Rocky Mountain Spotted Fever, Q-fever, and Lyme borreliosis. Tick paralysis can also possibly result from the neurotoxins secreted in the tick’s saliva upon feeding, as opposed
to tick-vector diseases caused by a microbial agent (Sonenshine and Roe 1993). These properties contribute to ticks being the most important vector of human pathogens in North American countries (Parola and Raoult 2001). Additionally, the transmission of pathogens to cattle can cause an economic strain in developing countries. Conditions, such as babesiosis, tick paralysis, and anaplasmosis can severely diminish the health of a susceptible cattle population (Jongejan and Uilenberg 2004).

*Dermacentor variabilis*

The American Dog tick, *Dermacentor variabilis* (Say, 1821), is an ixodid tick species that is commonly found east of the Rocky Mountains in North America (Arthur 1962). The *D. variabilis* ticks are rather small, with males and females roughly 3.6 mm x 2.4 mm and 3.8 mm x 2.5 mm, respectively (Arthur 1962). However, engorged female ticks experience a relatively large increase in size and can potentially reach 15 x 10 mm (Arthur 1962). Immature *D. variabilis* ticks, such as larvae and nymphs, commonly parasitize a variety of small mammals, such as mice, rats, rabbits, squirrels and cats (Arthur 1962). However, *D. variabilis* adults prefer larger mammalian hosts, such as dogs, humans, horses, sheep, and cattle (Kollars et al. 2000, Burg 2001). It is common for female *D. variabilis* ticks to survive from 2-3 years in ideal conditions without a blood meal. *D. variabilis* ticks are most commonly found during the summer months, and it is the adult stage ticks that are most commonly associated with tick vector-borne diseases in humans (Arthur 1962).
Although *D. variabilis* larvae and nymphs parasitize smaller mammals, adult *Dermacentor variabilis* prefer larger mammals as hosts, including humans (Arthur 1962). This tick species can transmit bacteria, such as *Francisella tularemi*a, through its saliva and feces that can infect humans (Parola and Raoult 2001). *D. variabilis* is also an important vector of the disease Rocky Mountain Spotted Fever (RMSF) in America (Sonenshine and Roe 1993). Some of the signs and symptoms of RMSF include vomiting, headache, rash formation, and nausea roughly a week after the *Rickettsia rickettsii* bacteria is transmitted to its human host (Sonenshine and Roe 1993). In severe cases, the signs and symptoms could be gangrene, renal failure, and cerebral edema (Sonenshine and Roe 1993). Thus, RMSF can be deadly if the patient is not treated with antibiotics (Sonenshine and Roe 1993). *D. variabilis* has also been known to cause tick paralysis in the southeastern region of the United States (Arthur 1962). Tick paralysis is more likely to affect dogs and male humans, particularly children in the event that an attached tick is not detected (Arthur 1962).

**Ferritin Gene and Protein**

The ferritin protein is an intracellular inorganic protein that functions in the metabolism of excess iron (Torti and Torti 2002, Harrison et al. 1998, Chasteen and Harrison 1999). Because ticks are hematophagous parasites, it is crucial that they are able to sequester the excess iron present in the host’s blood (Sonenshine 1991). In some invertebrate species, such as *Drosophila melanogaster*, having a complete deficiency of either the ferritin light chain or heavy chain homologue is lethal to the developing embryo (Missirlis et al. 2007). At optimal levels, iron plays an indispensable role as a cofactor for many essential processes, such as electron transport and the cell replication
cycle (Torti and Torti 2002). Iron metabolism likely also plays an important role in tick reproduction, as the silencing of both types of ferritin have been shown to decrease both the hatching rate and oviposition of female ticks (Hadjuseka et al. 2009). At high concentrations, the excess iron can contribute to the formation of radicals which can ultimately damage the organelles of cells and cause harmful oxidative damage (Wang and Pantopoulos 2011).

The ferritin protein consists of 24 H-chain homologue or L-chain homologue subunits, and it can sequester around 4,500 iron atoms per molecule (Torti and Torti 2002). When excessive levels of iron are present, the ferritin protein can take in the additional Fe\(^{2+}\) ions and oxidize them to Fe\(^{3+}\), storing them in an iron core at the center of the protein (Harrison and Arosio 1996). The ferritin heavy chain protein provides the ferroxidase function of the ferritin structure that oxidizes iron, whereas the ferritin light chain uses its structure to increase the rate of this oxidation and the subsequent mineralization of iron (Chasteen and Harrison 1999). The ferritin protein in insects contains homologues of these proteins called the heavy chain homologue and the light chain homologue (Dunkov and Georgieva 1999). It is likely that ferritins play an important role in managing oxidative stress caused by excessive iron in the ixodid tick, as it has been found that ticks with a silenced ferritin gene present with higher levels of oxidative stress biomarkers than those with the ferritin gene intact (Galay et al. 2014).

**Exons and Introns in Eukaryotic Genes:**

All eukaryotic genes contain exons and introns. To express proteins in the cell, RNA polymerase transcribes messenger RNA (mRNA) from the template DNA strand.
and transfers this mRNA strand to the cytoplasm of the cell (Clancy 2008). Once it is in the cytoplasm and attached to the ribosome, the mRNA will be translated into an amino acid sequence, which will form the specific protein encoded for by the DNA. However, before the mRNA exits the nucleus, it is chemically modified to become a mature mRNA sequence. This happens through the addition of a poly-adenine tail at the 3’ end, a guanine cap synthesized at the 5’ end, and the splicing of introns out of the original mRNA sequence by small nuclear ribonucleoproteins in the spliceosome. When the introns are removed from the original mRNA sequence, the exon sequences are joined together to form the coding sequence. Introns are, thus, sections of the mRNA sequence that are spliced out of the immature mRNA sequence, and they do not contribute to the final protein product’s amino acid sequence (Clancy 2008).

Although introns were once considered “junk” DNA with no useful purpose, they are now known to have multiple necessary functions. Introns are involved in the formation of cis-regulatory elements, and they also function in transcription initiation, transcription termination, and genome organization (Chorev and Carmel 2012). The 5’ UTR intron is hypothesized to be heavily involved in transcription initiation, and it also possesses cis-regulatory sequences that support this possible function. The first intron frequently contains regulatory motifs, and the introns that are closer to the 5’ end of the transcribed sequence are more often conserved (Park et al. 2014). It is believed that long intron sequences play an instrumental role in the regulation of transcription, although a strong correlation between human gene expression and intron length has not yet been confirmed (Chorev and Carmel 2012). It is also hypothesized that the second intron is involved in processing of the mRNA sequence 3’ end (Chorev and Carmel 2012). Introns
likely also have a role in nucleosome formation, which has been supported in various experiments that remove intron sequences. For these reasons, knowing the intron sequences of the tick ferritin gene has the potential to provide greater knowledge on how the ferritin gene is regulated.

Objective

Although the coding sequence (cDNA) of a gene is more instrumental in revealing the specific amino acid sequence of a protein, knowing the intron sequence can provide additional valuable knowledge of the gene (Chorev and Carmel 2012). The primary goal of this project is to determine the intron/exon structure of the *D. variabilis* ferritin heavy-chain gene. After determining the intron and exon structure of the ferritin heavy-chain gene in *D. variabilis*, future projects could involve determining the intron and exon structure for other ixodid tick species. These sequences could then be compared to determine the degree of homology between species. Because the ferritin gene contains a 5’ untranslated region that has been shown to be highly conserved between tick species, it has the potential to be used in phylogenetic reconstruction (Xu et al. 2004). It has also been hypothesized and supported that there was a common ancestor for the ferritin gene of eukaryotes, further supporting its potential in phylogenetic reconstruction. Therefore, determining the intron structure of the ferritin gene in multiple tick species could provide a better understanding of the multiple functions of the ferritin gene and of how it evolved over time.

With this being stated, the first objective of this project was to successfully isolate the *D. variabilis* ferritin heavy-chain gene using PCR amplification. The putative *D.*
Dermacentor variabilis ferritin heavy-chain gene that is isolated can then be aligned with the ferritin gene coding sequence, which was established by work previously done in the same lab (Xu et al. 2004), to identify the intron/exon structure.

**Materials and Methods:**

*Tick Samples and Genomic DNA Extraction*

Alive adult male *Dermacentor variabilis* ticks were obtained from the Tick Rearing Facility at Oklahoma State University (OSU). The ticks were kept alive until use and stored at -75°C. Total genomic DNA was extracted from a single tick using the Epicentre MasterPure Complete DNA and RNA Purification kit (Madison, WI). The tick was first cut into small pieces and incubated at 55°C in 300µl of 2X Tissue and Cell buffer with 1.0µl of Proteinase K (50µg/µl) for one hour to inhibit DNAse activity. Then, the tick was ground, either with a sterile blue pestle or a homogenizer, and incubated at 85°C for another 30 minutes. The tick sample was then placed on ice for 10 minutes. Next, 220µl of the MPC Protein Precipitation Reagent was added to the sample. After vortexing for 10 seconds, the sample was centrifuged at 15.8kg for 10 minutes at 4°C. The supernatant was transferred to a new clean tube, and 700µl of 100% isopropanol was added to precipitate the DNA. The DNA pellet was then washed with 70% and 100% ethanol. Finally, the DNA pellet was resuspended with 75 to 100µl of ddH₂O and visualized on a 1.0% agarose gel electrophoresis to determine the yield. The extracted DNA was stored at -75°C until it was used.
PCR Amplification and DNA Purification

The *D. variabilis* tick ferritin H-chain gene was amplified from the genomic DNA through polymerase chain reaction (PCR) using the Invitrogen Long PCR kit (Grand Island, NY) on a DNA Engine Peltier thermal cycler 200 (MJ Research Inc., Waltham MA), following the manufacturer’s protocol. The primers used were the Ferritin universal primer pairs (DVAFERUniv 1F 5’-TMTGCTTCAACAGTGATTGAA and DVAFERUnivRC 5’-CTA GTC TGA CAG GGT CTC CTT G), which flank the genomic region between the Ferritin 5’UTR and the end of stop codon. The PCR products were gel purified using the QIAQuick Gel Extraction Kit and sequenced using the Sanger’s method at Clemson University Genomics Institute using PCR primers. The DNA sequences were blasted to GenBank. After blast conformation and alignment with *D. variabilis* cDNA, the PCR products were reamplified with the primers DVAFERUnivF1 (5’-ATG GCC GCT ACT CAG C) and DVAFERUnivRC (5’-CTA GTC TGA CAG GGT CTC CTT G) which flank the genomic region between start codon and stop codon, using Thermo Scientific (Waltham, MA) Phire Hot Start II DNA polymerase. The thermocycler had an initial heat step at 98°C for 30 s to activate the Phire Hot Start II DNA Polymerase and amplifications through 34 cycles at 95°C for 10 s, 58°C for 20 s, and 72°C for 90 s. The final extension reaction proceeded at 72°C for 5 min. The DNA was gel-purified using the QIAQuick Gel Extraction Kit and visualized using a 1% agarose gel electrophoresis.

DNA Cloning
The purified DNA fragments were cloned using the TOPO® XL PCR Cloning Kit (Invitrogen, Grand Island, NY). The gel-purified re-amplified ferritin gene product (4 µL of 2-40 ng/µL) was combined with the pCR-XL-TOPO vector (1µL) and incubated at 25°C for five minutes. Then, 1.0 µL of 6X TOPO Cloning Stop Solution was added to stop the reaction. Chemical transformation was used to induce uptake of the plasmid by E. coli cells. Following incubation of the cells and cloning reaction on ice for 30 minutes, the cells experienced heat shock at 42°C for 30 seconds and were returned to ice for two minutes. Next, 250 µL of 25°C S.O.C. medium (Super Optimal Broth for E. coli) was added to the tubes and incubated for two minutes on ice. The samples were shaken horizontally at 225 rpm for 60 minutes at 37°C. A small sample (120 µL) of the E. coli bacteria solution was spread across lysogenic broth plates (1.0% tryptone, 0.5% yeast extract, 1.0% NaCl, pH 7.0) containing kanamycin, and the plates were incubated at 37°C overnight. The control plate was prepared under the same conditions using untransformed cells to ensure that the antibiotics were inhibiting bacterial growth when the plasmids did not contain the resistance genes for kanamycin and ampicillin.

Following incubation at 37°C for 24 hours, the colonies that grew on the LB plates with kanamycin were used as template DNA in a PCR amplification using Phire Hot Start II Polymerase and FerM13RF/ FerM13FR primers to check for the expected size of insertion. A 1.0% agarose gel electrophoresis was used to check the PCR products for the presence of positive clones. Clones with the expected size of insertions were then re-grown in 15mL sterile tubes containing kanamycin.
Plasmid Isolation

Following the detection of positive clones through PCR amplification, the remaining colony residue was transferred to a tube containing 250 µL of lysogenic broth to create a bacterial culture. This solution was shaken overnight at 37°C to stimulate *E. coli* growth. Plasmid DNA was isolated from the *E. coli* broth using the Eppendorf Perfectprep Plasmid Mini kit (Hamburg, Germany) or the Thermo Scientific GeneJET plasmid MiniPrep kit (Waltham, MA).

DNA Sequencing

The plasmid DNA was sequenced with PCR primers at Clemson University Genomics Institute using the ABI automatic sequencer. After the sequence was confirmed by searching online using Blastn at GenBank and aligned with the previously obtained *D. variabilis* ferritin coding sequence, the positive colonies were further grown in 7.5ml volume. Internal sequencing primers were designed near the 3’ end of the sequence and the plasmid DNA were sequenced again using the newly designed internal primers. This sequencing strategy was repeated several times until the primers walking over the entire plasmid insertion region and all DNA sequences were obtained from the insertion in positive plasmids.
Sequence Alignment

After receiving the sequences from the Clemson University Genomics Institute, the sequenced contigs were assembled using the Staden software package (Staden 1986). Base calling errors were checked back to ABI sequencing chromatograms on Staden. A single consensus sequence from all of the contigs sequenced from the entire cloned *D. variabilis* genomic sequences was constructed. This consensus of ferritin genomic sequence was aligned with *D. variabilis* cDNA that was previously obtained in Dr. Fang’s lab (Xu et al. 2004) using BioEdit, a sequence alignment editor program (Hall 1999). The location and size of exons and introns were determined from the DNA sequencing alignment.

Results:

High-molecular weight genomic DNA was successfully isolated from a single tick (Fig. 1). This genomic DNA was used for successive PCR amplifications and cloning. Using the DVAFERUniv1F/RC primers, the PCR successfully amplified multiple DNA fragments from the genomic DNA. The largest DNA band was over 4.2kb. The reamplification of PCR products using the DVAFERUnivF1/RC primers yielded two DNA bands with sizes of ~4.0kb (top) and ~2.4kb (bottom), respectively (Fig. 2). Both DNA bands were excised off of the agarose gel and purified using the QiaGen gel purification kit (Fig.3). Both the top and the bottom DNA bands were sent out for sequencing. The sequencing results confirmed that the top band was the heavy-chain ferritin sequence, not the bottom band of DNA. The top band was then ligated into the plasmid pCR-XL-TOPO® for TOPO® XL PCR cloning. The positive clones (Fig. 4),
PCR amplification of the positive clones (Fig. 5), and gel purification of the positive clones (Fig. 6) indicated that the TOPO® XL PCR cloning had been successful in isolating the ~4.0kb DNA fragment containing the putative ferritin heavy-chain gene.

Sequencing of the positive clones using the PCR primers showed that the inserted DNA was, indeed, the heavy-chain ferritin gene. Over 80 plasmid DNA templates were sequenced with six internal sequencing primers, which walked through the entire insertion of the plasmid. Over 60 of the sequences returned from Clemson were of good quality. Assembly of these ~60 contigs yielded a consensus sequence of 4061 bp between the start codon and stop codon. Aligning this ferritin genomic consensus sequence with the D. variabilis cDNA showed that the D. variabilis heavy-chain gene coding region contains three exons and two introns in addition to the 5’ and 3’ transcribed but not translated regions (5’ UTR and 3’ UTR) (Fig. 7). The first exon begins with the start codon (UTG), which codes for methionine, is only 106 base pairs in length, and codes for 35.3 amino acids. The second exon is 277 bp in length and codes for 92.3 amino acids. The third exon is 135 bp in length, ending at the stop codon (UAG), and it codes for 45 amino acids. Because the genomic DNA was amplified between the start and stop codons, no 5’ and 3’ UTR regions were obtained in this study. The 5’ and 3’ UTR regions were obtained in a previous study in the same lab (Xu et al. 2004).

Introns in the D. variabilis H-chain ferritin gene are very large in length. The length of the first intron is 2268 bp, whereas the second intron is 1019 bp (Fig. 7). The first intron starts between the codon positions 1 and 2 of alanine. The second intron begins at the end of codon cysteine (after UGU). The 5’ and 3’ intron splice junctions are different in the two introns. The 5’ splice donors are AATGGtaatac and TGTGTcctgt,
whereas the 3’ splice donors are gcaggCCTAC and tgtgtGACTT (Fig. 7, 8). The first intron is classified as a phase 1 intron, and the second intron is classified as phase 0 intron (Fig. 7).

The nucleotide composition differs between the intron and exon sequences. Both introns are A+T rich nucleotide sequences (Table 1), whereas the exons are G+C rich (Table 2). There is around 56% A+T content in the two introns and approximately 53% G+C content in the three exons.

**Discussion:**

The ferritin gene has previously been studied in a number of ticks (Kopacek et al. 2003, Xu et al. 2004, Galay et al. 2013, 2014, 2015). However, no tick ferritin exon/intron structure has been published. This is the first time that the tick ferritin exon/intron structure has been reported. From the results achieved in this project, the ferritin gene exon/intron structure in the *Dermacentor variabilis* tick can now be further characterized. The *D. variabilis* ferritin heavy-chain gene cDNA is not very long, as it only consists of 522 base pairs and encodes for 172 amino acids plus a stop codon (UAG). However, in the *D. variabilis* genome, the ferritin heavy-chain coding region sequence of the pre-mature mRNA is 4061 base pairs long.

It is unknown why the *D. variabilis* ferritin H-chain has such large introns.

Because this project was the first to determine the ferritin gene structure in ticks, it is not yet possible to compare the ferritin exon/intron structure of *D. variabilis* ticks to that of other *Dermacentor* species or ixodid ticks. Once the ferritin exon/intron structure is discovered in other *Dermacentor* and ixodid ticks, it will be possible to determine if
there is a high degree of variation between the intron composition and sizes and if the intron/exon junctions are conserved. This future experiment has the potential to provide useful information, as it has previously been discovered that the ferritin gene intron positions are conserved to an extent within kingdoms (Proudhon et al. 1996). This same observation was made in comparing the ferritin genes of the higher eukaryotes mice and humans, where the intron/exon arrangement was conserved, despite a lack of conservation in the number and length of introns (Leibold and Munroe 1987). Also, it has previously been discovered that the ferritin intron splice junction is related to the ferritin structure, as most are found on the surface positions of the ferritin molecule (Harrison et al. 1991). It has also been found that the amino acid compositions around the intron splice junction are highly conserved at certain junctions more than others (Harrison et al. 1991).

The gene structure has been elucidated in other organisms both distantly and more closely related to the tick, such as the mouse *Mus musculus* (Beaumont et al. 1994) and the fly *Drosophila melanogaster* (Dunkov and Georgieva 1999), respectively. As suspected, the number and size of the *D. variabilis* ferritin introns differ from those found in distantly related species. The ferritin heavy chain homologue gene in the Sprague-Dawley rat consists of three introns with lengths of ~1500 base pairs, ~200 base pairs, and 92 base pairs, respectively (Murray et al. 1987). Although the rat is a vertebrate animal and a higher eukaryote in comparison to the distantly related invertebrate arachnid tick, it is interesting to note that the length of the first intron in both the ferritin heavy chain gene in the rat and the ferritin heavy-chain homologue in the *D. variabilis* tick are significantly longer than the subsequent intron(s).
The number of introns is more similar to the insect *Drosophila melanogaster* heavy-chain homologue ferritin gene, with the *D. melanogaster* ferritin heavy chain homologue also having two introns and three exons between the start and stop codons, and the light chain homologue having two exons and one intron between the start and stop codons (Dunkov and Georgieva 1999). The introns after the start codon of the *D. melanogaster* ferritin heavy-chain homologue are 877 and 62 bp in length (Dunkov and Georgieva 1999). There is also an intron within the 5’UTR that is 224 bp in length. The first intron after the start codon in the *D. melanogaster* ferritin heavy chain homologue is also much longer than the successive intron (877 vs. 62 base pairs) (Dunkov and Georgieva 1999). This consistent observation of increased first intron length could support the hypothesis that the first intron might contain cis-regulatory elements that are essential for the ferritin gene.

It has previously been discovered that the iron responsive element (IRE) is located within the 5’ UTR region of the *D. variabilis* ferritin heavy-chain homologue (Xu et al. 2004), and it has also been found in the same region in the heavy-chain homologue gene of *D. melanogaster* and *A. aegypti* (Dunkov et al. 1995), which indicates possible iron responsive protein (IRP) control of ferritin gene translation in insects (Dunkov and Georgieva 1999). In the *D. melanogaster* ferritin heavy-chain homologue, alternative splicing has been demonstrated in the intron within the 5’ UTR, which can produce three unique mRNA sequences without the IRE and one mRNA that retains the IRE (Lind et al. 1998). The different mRNA sequences are found in varying amounts at different life stages (Lind et al. 1998). It is unknown whether or not the “flexible” IRE/IRP control mechanism also exists in the *D. variabilis* heavy-chain homologue ferritin gene or if it is
consistently present in the 5’ UTR, as in mammalian ferritin genes. It has been predicted by other researchers that the D. melanogaster heavy-chain homologue IRE within the 5’ UTR can be alternatively spliced due to its diet, which can vary in iron (Dunkov and Georgieva 1999). Following this logic, the “flexible IRE/IRP” mechanism would probably not be present in D. variabilis ticks, as they have a completely hematophagous, iron-rich diet. This is also supported by the observation that the D. variabilis ferritin H-chain gene 5’ UTR is more similar to the vertebrate 5’ UTR in comparison to the 5’ UTR in insects (Xu et al. 2004). This idea is also supported by the previous finding that the tick species ferritins have coding sequences that are more similar to vertebrate ferritins than insect ferritins (Kopacek et al. 2003).

As mentioned above, the first intron of the D. variabilis ferritin gene is more than twice as long as the second ferritin gene intron (2268 vs. 1019 base pairs). It is unknown why the first intron is significantly larger than the second intron. We did not analyze the intron for repeat sequences, so it is possible that the length is due to a large number of repeat sequences. It has been hypothesized that longer introns are conducive for alternative splicing and that their longer size can also impact gene expression (Kandul and Noor 2009). This hypothesis is further supported by the observation that alternative splicing takes place in the ferritin gene of the Bactrocera dorsalis insect (Jiang et al. 2014). Insect ferritin heavy chain homologue and light chain homologue genes are highly conserved (Dunkov and Georgieva 2006). Due to this high degree of conservation in insect ferritins, it is plausible that alternative splicing might also take place in the ferritin gene of other insects (Jiang et al. 2014). Thus, alternative splicing could occur in the D. variabilis tick. The long intron might also possibly serve as a source of genetic diversity.
in the gene, as there are a greater number of locations for possible errors in replication or mutations that can take place in longer introns (Kandul and Noor 2009). It is possible that having a long first intron in the *D. variabilis* ferritin gene is vital for the gene. This prediction would be supported if the first intron in other *Dermacentor* or ixodid tick species is also much longer than the second intron. However, it is more likely that the exon/intron junction would be conserved between species as opposed to the intron composition or length, such as in the hard tick calreticulin intron/exon structure (Xu et al. 2005).

There were a few limitations to this study. To strengthen the data collected, it would have been ideal to have isolated and sequenced the *D. variabilis* ferritin heavy-chain homologue gene from several different *D. variabilis* ticks to verify the identity of the intron/exon structure. It would have also been informative to further characterize the intron/exon structure by using computer analysis to identify any possible metal regulatory element (MRE) sites in the introns, as it was discovered that four MRE sites exist in the first intron after the start codon of the *D. melanogaster* ferritin heavy-chain homologue gene (Dunkov and Georgieva 1999).

In conclusion, the objective of elucidating and characterizing the intron and exon structure of the *D. variabilis* ferritin heavy-chain gene was accomplished. This project was the first to reveal the intron/exon structure of the ferritin gene in ticks, and it has resulted in more questions that will, hopefully, be answered by continued research in the future.
Table 1. The A+T content (%) and corresponding G+C content (%) of the *D. variabilis* ferritin gene intron sequences. The averages and standard deviations are also displayed.

<table>
<thead>
<tr>
<th>Intron</th>
<th>A+T Content (%)</th>
<th>G+C Content (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>54.36</td>
<td>45.64</td>
</tr>
<tr>
<td>2</td>
<td>58.29</td>
<td>41.71</td>
</tr>
<tr>
<td>Mean</td>
<td>56.33 ± 2.78</td>
<td>43.68 ± 2.78</td>
</tr>
</tbody>
</table>

Table 2. The A+T content (%) and corresponding G+C content (%) of the *D. variabilis* ferritin gene exon sequences. The averages and standard deviations are also displayed.

<table>
<thead>
<tr>
<th>Exon</th>
<th>A+T Content (%)</th>
<th>G+C Content (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>46.23</td>
<td>53.78</td>
</tr>
<tr>
<td>2</td>
<td>45.68</td>
<td>54.32</td>
</tr>
<tr>
<td>3</td>
<td>47.41</td>
<td>52.59</td>
</tr>
<tr>
<td>Mean</td>
<td>46.44 ± 0.88</td>
<td>53.56 ± 0.88</td>
</tr>
</tbody>
</table>
Figure 1. Genomic DNA extracted from a single *D. variabilis* tick visualized on a 1.0% agarose gel with 1X TAE buffer. Lane 1: 1 kb ladder, Lane 2 and 3: *D. variabilis* genomic DNA.
Figure 2. PCR amplification of the putative *D. variabilis* ferritin gene visualized on a 1.0% agarose gel with 1X TAE buffer. Lanes 1: 1 kb ladder, Lane 2: PCR negative control, Lanes 3-6: PCR amplified *D. variabilis* genomic DNA with bright gel bands at ~4.0 kb and ~2.4 kb.
Figure 3. Gel-purification of the putative *D. variabilis* ferritin gene visualized using 1.0% agarose gel electrophoresis with 1X TAE buffer. Lane 1: 1 kb ladder, Lanes 2-5: purified 4.0 kb DNA fragment from the PCR, Lanes 6-9: purified 2.4 kb DNA fragment from the PCR.
Figure 4. Positive gene cloning products isolated from the *E. coli* plasmid visualized on a 1.0% agarose gel electrophoresis using 1X TAE buffer. The Invitrogen TOPO® XL PCR Cloning Kit was used to isolate the DNA segment inserted in gene cloning. Lane 1: 1 kb ladder, Lanes 2-12: positive gene cloning products with bright gel band at ~4.0 kb.
Figure 5. PCR amplification products of the inserts from the positive gene cloning plasmids visualized on a 1.0% agarose gel electrophoresis using 1X TAE buffer. Lane 1: 1 kb ladder, Lane 2: PCR negative control, Lanes 3-8: amplified gene cloning plasmid inserts.
Figure 6. Gel-purification of the positive gene cloning plasmids visualized on a 1.0% agarose gel electrophoresis with 1X TAE buffer. Lane 1 and 2: 1 kb ladder, Lane 3-6 and Lanes 10-12: gel-purified ~4.0kb fragments, Lane 7-9 and Lanes 13-15: gel purified ~2.4kb fragments.
Figure 7. Complete genomic nucleotide sequence of the *D. variabilis* ferritin H-chain gene, aligned with *D. variabilis* cDNA and the deduced amino acid sequences. The comprehensive *D. variabilis* putative ferritin heavy-chain gene discovered from this project was aligned with the published *D. variabilis* ferritin heavy-chain cDNA sequence (DVA_cDNA_nt), and it is displayed below the consensus sequence. The third line (DVA_cDNA_aa) displays the deduced amino acid sequence.
Figure 8. Map of the *D. variabilis* ferritin heavy-chain homologue gene. The black lines represent introns 1 and 2 of the gene. The three exons are represented by the gray boxes. The start codon, stop codon, 5’ intron splice junction sequences, and 3’ intron splice junction sequences are included in the visual. The 5’ UTR and 3’ UTR are not included in this visual, as they were not elucidated in this study. The intron sizes (bp) are included below the introns, whereas the exon sizes are displayed above the exons. The map is drawn to scale.
References Cited:


revealed by genetic manipulations and live imaging of *Drosophila* ferritin.

Genetics. 177(1): 89-100.


