Přírodovědecká fakulta
Faculty of Science

Jihočeská univerzita v Českých Budějovicích
University of South Bohemia in České Budějovice

Bachelor's Thesis

Isolation and molecular characterization of the histone methyltransferase EHMT from the soft tick *Ornithodoros moubata*

Laboratory of Host-Pathogen Interaction Evolution

Department of Parasitology

Author

Sewmi M A Delana Mudiyanselage

Supervisor

Dr. Alejandro Cabezas-Cruz

Co-supervisor

Dr. Ryan O. M. Rego

Bachelor Thesis: Sewmi M A Delana Mudiyanselage, 2017: Isolation and molecular characterization of the histone methyltransferase EHMT from the soft tick *Ornithodoros moubata*, Bachelor Thesis, in English – 39 pages, Department of Parasitology, Faculty of Science, University of South Bohemia in České Budějovice.

Annotation: The aim of this thesis was to isolate and characterize the histone methyltransferase EHMT from *Ornithodoros moubata*.

Affirmation: I hereby declare that I have worked on the submitted bachelor thesis independently. All additional sources are listed in the bibliography section. I hereby declare that, in accordance with Article 47b of Act No. 111/1998 in the valid wording, I agree with the publication of my bachelor thesis, in full form to be kept in the Faculty of Science archive, in electronic form in publicly accessible part of the STAG database operated by the University of South Bohemia in České Budějovice accessible through its web pages. Further, I agree to the electronic publication of the comments of my super visor and thesis opponents and the record of the proceedings and results of the thesis defense in accordance with aforementioned Act No. 111/1998. I also agree to the comparison of the text of my thesis with the Theses.cz thesis database operated by the National Registry of University Theses and plagiarism detection system

Date **Sewmi M A Delana Mudiyanselage**

Table of Contents

1 ACKNOWLEDGEMENT

First of all I would like to express my sincere thanks to Dr. Alejandro Cabezas-Cruz for giving me an opportunity to carry out my bachelor thesis and providing me with all the necessary facilities for the bachelor thesis research and also extremely thankful to him sharing his expertise and valuable guidance with me.

I am also grateful to Dr. Ryan Rego for his valuable advice and help during my bachelor thesis. Also, I like to give thanks to all the members of Dr. Rego's laboratory and all the members of Institute of parasitology.

I would like to give thanks to Dr. James Valdes for his support during my research project as well.

I take this opportunity to express my gratitude to Prof. Libor Grubhoffer who supported me in an incredible way during my whole studies at the University of South Bohemia.

I really appreciate and also grateful to my colleagues specially Julia Gobl and Hellen Chemtai for their support during my research project.

Finally I would like to thank my parents for providing me with unfailing support and continued encouragement of my years of studies. Also I would like to give my thanks to all of my colleges.

2 ABSTRACT

Epigenetics has bloomed in the last 20 years as a promising area of research. However, little is known about the epigenetic machinery of ticks as well as the epigenetic mechanisms that tick-borne pathogens may use to manipulate tick physiology.Initial results showed that the *Ixodes scapularis* genome has major histone modifying enzymes. However, this tick is a difficult model to study tick epigenetics due to a complex life cycle that results in long generation intervals that can take up to two years. Soft ticks offer an alternative to solve this problem because they have shorter generation intervals. The initial challenge in this direction is to characterize histone modifying enzymes in soft tick models. The objective of this thesis was to isolate the euchromatic histone-lysine N- methyltransferase (EHMT) from the soft tick *Ornithodoros moubata*. Degenerate primers were designed using available EHMT from chelicerates. Soft tick EHMTs were made available on the course of this project and more specific primers were designed. Total RNA was extracted from adult *O. moubata* and cDNA synthesized. Using this cDNA as template, a fragment of the expected size was amplified by PCR. This DNA fragment was then inserted by ligation into the TOPO vector system. *Escherichia coli* competent cells were then transformed using the plasmid. Recombinant colonies were selected using blue-white screening followed by plasmid propagation, purification and sequencing. BLAST and sequence alignment confirmed the isolation of the C-terminus of *O. moubata* EHMT which contains the typical SET domain found in members of histone methyltransferase family. The 3D structure of this SET domain was conserved with the SET domain of mammalian EHMT. Further phylogenetic analysis showed that *O. moubata* EHMT was closely related to EHMT from chelicerates. This is the first characterization of a histone methyltransferase from the soft tick *O. moubata*. Further studies should evaluate the potential role of this enzyme in tick physiology regulation and tick-pathogen interactions.

3 INTRODUCTION

3.1 Generalities and taxonomy of ticks

Ticks are hematophagous ectoparasites that feed from different species of vertebrate hosts (Brites-Neto et al., 2015). According to their structure, ticks are divided in three families, called Ixodidae, Argasidae and Nuttalliellidae. Ixodidae and Argasidae are known as hard ticks and soft ticks, respectively (Figure 1)[1].

Soft and hard ticks can be differentiated based on life cycle and morphology. The Ixodidae, which include over 700 tick species, have hard shield and their capitulum is visible. The Argasidae family includes 200 tick species and they do not have hard shield and the capitulum is concealed beneath the body[1].The family Nuttalliellidae has a single species, *Nuttalliella namaqua* that can be found in South Africa (Guglielmone et al., 2010). Argasids are endophilic and nidicolous parasites, thus they colonize the nest and burrows of their hosts and feed when the host arrives (Manzano-Roman et al., 2012). Ixodids are exophilic ticks that seek actively for a host. Most soft ticks are fast feeders, ingesting relatively small amount of blood per meal and the adult female can lay eggs repeatedly. They are very resistant to environmental stress. Also they can survive for several years without feeding (Manzano-Roman et al., 2012).

Ticks are only second to mosquitoes as vectors of human diseases. However, ticks are the most important vectors affecting simultaneously human and animal health (Manzano-Roman et al., 2012). Importantly, ticks can also affect humans by causing allergic diseases (Cabezas-Cruz et al., 2014). Interactions between ticks, pathogens and hosts build an ecology complexity of parasitic associations in the environment (Parola and Raoult. 2001).

Ixodes scapularis, also known as the deer tick, is a major vector of human and animal pathogens in North America. For example, this tick transmits *Borrelia burgdorferi* (causing Lyme disease), *Anaplasma phagocytophilum* (causing human granulocytic anaplasmosis, HGA), *Babesia microti*, *Babesia odocoilei* (causing babesiosis) and Powassan (POW) virus (causing POW virus disease)(Telford and Goether. 2004).The life cycle of *I. scapularis*

[1] Figure 1 taken from http://www.microbiologybook.org/parasitology/ticks.htm accessed 01-07-2017

involves, excluding eggs, three stages (larva, nymph, and adult) and each one of them feed on a different host. Thus, *I. scapularis* is a three-host tick (Keirans et al., 1996). In Europe, *Ixodes ricinus* transmits *A.phagocytophilum* that in addition to HGA causes tick-borne fever in ruminants (Telford and Goether. 2004).

Ornithodoros moubata is a soft tick that parasitizes domestic animals, especially pigs. They transmit important pathogens such as by *Borrelia duttoni*, African swine fever virus and West Nile virus (Manzano-Roman et al., 2012).

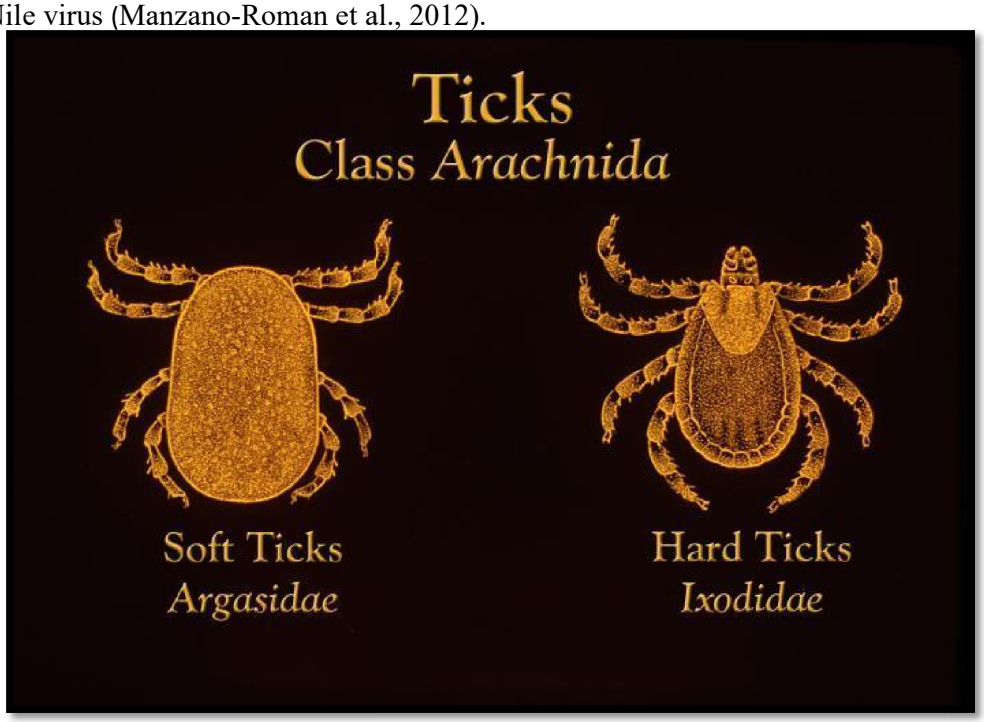

Figure 1. There are two types of ticks living within the environment, such as soft ticks and hard ticks. Hard ticks have hard shield and their capitulum is visible. Soft ticks do not have hard shield and their capitulum is concealed beneath the body (Brites-Neto et al., 2015)

3.2 Life cycle of ticks

As mentioned above, ticks have four stages in their life cycle including egg, larvae, nymph and adult (Figure 2)[2].

[2] Figure 2 taken from http://www.scalibor.com.au/ticks-on-dogs/. accessed 01-07-2017

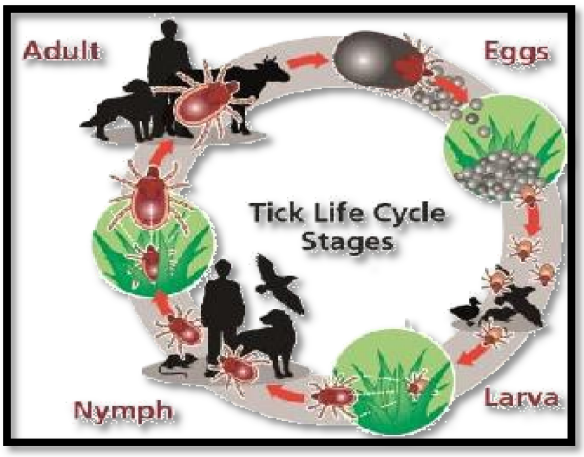

Figure 2.Typical life cycle of hard ticks. The duration of the life cycle swill depend on the tick species and, within the same species, on the geographic regions.

The life cycle of Ixodid ticks can take up to two years (*I. scapularis*), although some species can complete their life cycle within 21 days (i.e. *Rhipicephalus microplus*) (Allen et al., 2001).Soft ticks will take two or more nymphal stages to complete their life cycle (Aeschlimann et al., 1995). The larvae and nymphs require small rodents, birds and lizards to feed. Adult ticks can feed on large animals such as cattle, deer and humans (Brites-Neto et al., 2015).

Ticks live mostly in warm and humid climates. They need certain amount of moisture in the air to undergo metamorphosis and low temperatures might disrupt their life cycle (Nuttall, George. 1905).

3.3 Ticks and their environment

Ticks are not insects and they are incapable of flying and jumping like mosquitoes or fleas, respectively. Therefore, they can only move long distances on their host. If the humidity is high enough, ticks can remain long time on the vegetation until they find a host. Otherwise, they must go down to the ground for rehydration. Once on the host, ticks will find a soft-skin spot, insert its mouthparts and will start the feeding (Figure 3, Cabezas-Cruz et al., 2016b).In addition to humidity, temperature is also an important factor in the life cycle of tick. For example, while 20˚C of temperature is an ideal temperature for egg development, temperatures higher than 20˚C can affect significantly the embryonic development (Brites-Neto et al., 2015).

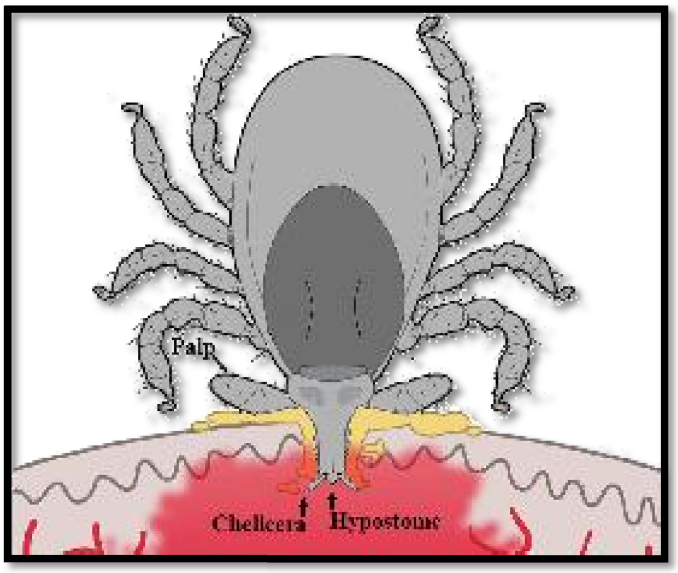

Figure 3.While feeding, ticks completely introduce the chelicerae and hypostome in the skin

3.4 Epigenetics

Epigenetics refers to stimulus-trigged changes in gene expression that does not involve changes to the underlying DNA sequences. Gene expression can be regulated by three main epigenetic mechanisms: DNA methylation, histone modification and non-coding RNA (Figure 4, Gomez-Dıaz et al., 2012). DNA methylation has a role in silencing gene expression and heterochromatin remodeling among other functions. The DNA of most species is methylated and this modification take place at 5-methylcytosine in eukaryotes and N6-methyladenine act as modified base for prokaryotes (Wion and Casadesus. 2006). Histone modifications are a set of posttranslational modifications on histones including methylation, phosphorylation, acetylation, ubiquitylation and sumoylation (Gomez-Dıaz et al., 2012). Histone modifying enzymes catalyze the addition or removal of an array of covalent modifications in histone and non-histone proteins (Marmorstein and Trievel. 2009).

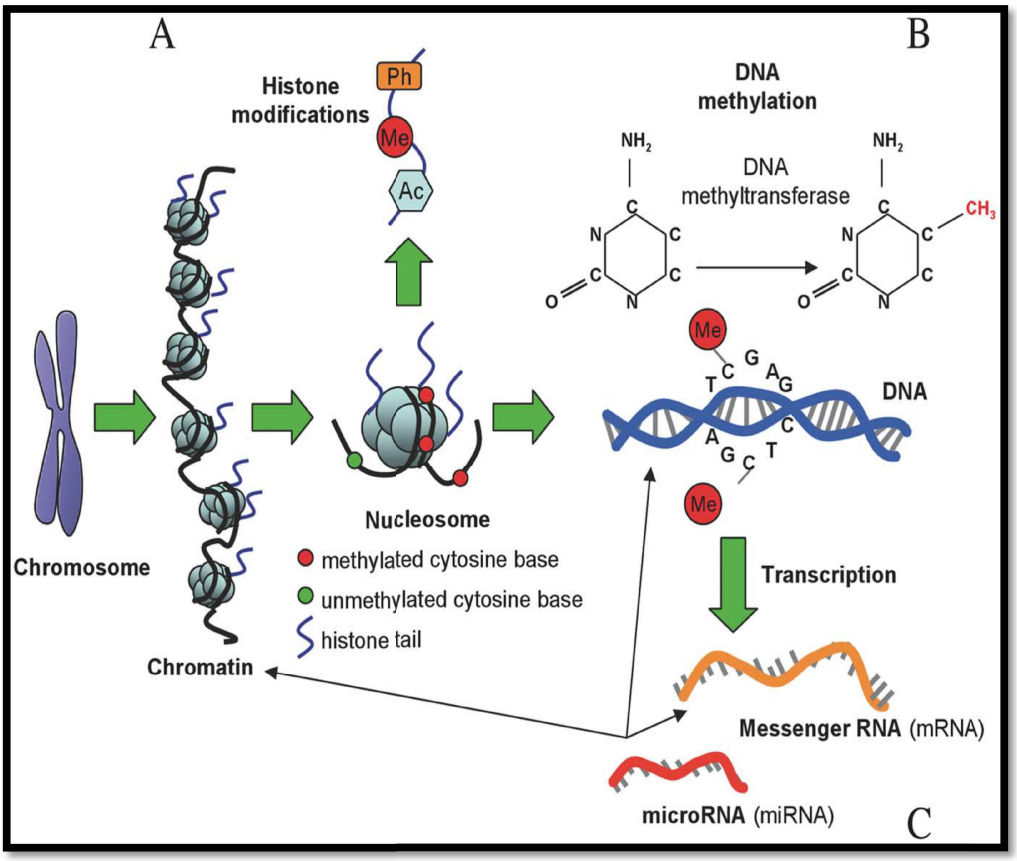

Figure 4.Types of epigenetic modifications. (A) Histones can undergo phosphorylation (Ph), methylation (Me), and acetylation (Ac), among other chemical modifications. These modifications are involved in chromatin remodeling and transcriptional regulation. (B) DNA molecules are methylated by the addition of a methyl group to carbon position 5 on cytosine bases, a reaction catalyzed by DNA methyltransferase enzymes, which maintains repressed gene activity. (C) mRNA is translated into a protein product, but this process can be repressed by binding of micro RNAs (miRNA), a class of non coding RNA (ncRNA)

There are two definitions that are frequently used to describe epigenetics: epigenetic plasticity and epigenetic inheritance. Epigenetic plasticity refers as all transient changes in gene expression that occur at individual levels and also these are propagated during the mitosis in multi-cellular organisms. Epigenetic inheritance refers as non-genetic variations that were obtained during an organism's life can be possibly being passed on to that organism's offspring (Gomez-Dıaz et al., 2012).

3.5 Tick epigenetics

Several classes of enzymes can modify histones at different lysine residues. A previous study showed that *I. scapularis* have five putative histones in their genome and transcriptome. They were classified as IsH1, IsH2A, IsH3 and IsH4 (Cabezas-Cruz et al., 2016a). Histones are major components of the chromatin in the eukaryotic nucleus. They can undergo different posttranslational modifications such as methylation, acetylation and phosphorylation (Zhang and Reinberg. 2001). These posttranslational modifications participate in the remodeling of chromatin and can regulate transcription, replication, DNA repairs and apoptosis (Nowak and Corces. 2004). There are four main histone modifying enzymes that may contribute to tick epigenetics. They are histone methylases, histone demethylases, histone acetylases and histone deacetylases. In general, those enzymes that add, remove and recognize histones modifications are regarded as 'writers', 'erasers' and 'readers', respectively.

The histone acetyltransferases (HATs) protein family contains enzymes which are capable of adding acetyl group to histone. Therefore, they have been called "writer" enzymes (Yang. 2015), while HDACs and sirtuins are "eraser" enzymes that remove acetyl groups on or from histones respectively (Minucci et al., 2006; Greiss et al., 2009). Histone acetylation is regulated by HATs and different levels of this modification regulate chromatin structure and the binding of repressors and activators of gene transcription. It was previously reported that *I. scapularis* genome has histone acetyltransferase1 (IsHAT1), E1A-binding protein p300/CREB-binding protein (Isp300/CBP) and several members of the MYST family (Cabezas-Cruz et al., 2016a).

The histone deacetylases are a class of enzyme capable to remove acetyl group from ε-N-acetyl lysine amino acid on a histone. As a consequence, deacetylation allows the histones to wrap the DNA more tightly due to removing of acetyl group makes the histones more positively charged, hence negatively charged DNA wrap more tightly around the histone proteins (Choudhary et al., 2009). Histone deacetylases are a protein family evolutionarily conserved in prokaryotes and eukaryotes, including mammals. Human histone deacetylases are separated in four classes: class I, class II, class III and class IV (Minucci and Pelicci. 2006) In *I. scapularis* three class I HDACs (IsHDAC1, 3 and 8) and one class II HDAC (IsHDAC4) enzymes were identified (Cabezas-Cruz et al., 2016a).

The histone methyltransferase (HMT) regulates chromatin organization and depending on the lysine residues it will target to induce either activation or repression of gene expression (Mozzetta et al., 2015). HMTs are divided in seven families classified according to their SET domain. These families include SUV39, EZH, SET1, SET2, PRDM and SMYD and other HMTs containing SET domain. Only one HMT (DOT1-like, DOT1L) does not contain a SET domain (Mozzetta et al., 2015; Volkel et al., 2007).The *I. scapularis* genome contains SET2 (IsSETD2, IsSEDT4and IsSEDT7), SUV39 (SETDB1, IsSETDB1-A, IsSETDB1-B and IsEHMT) and SMYD (IsSMYD3-5) family members. The tick homolog of human DOT1L (IsDOT1L) was also identified (Cabezas-Cruz et al., 2016a).

The histone demethylase protein family is separated into the JumonjiC (JMJC) demethylases and the lysine-specific histone demethylase (LSD) (Kooistra and Helin. 2012). The LSD family is composed by LSD1 and LSD2. LSD and JMJC protein families use different mechanisms to catalyze the demethylation reaction (Kooistra and Helin. 2012). In *I. scapularis*, there two copies of LSD1 (IsLSD1-A and IsLSD1-B), but no homolog of LSD2 was found. IsLSD1-A and IsLSD1-B have major differences between them. As an example N-terminus of IsLSD1-B is larger than IsLSD1-B, IsLSD1-B does not possess a "spacer region" like IsLSD1-A and IsLSD1-A does not possess a NAD-binding domain present in human LSD1 (Cabezas-Cruz et al., 2016a). There were six members of the JMJC demethylases identified in *I. scapularis* including IsUTY, IsJARID2-A, IsJMJD1B, IsJARIDIA, IsJMJD2B and IsJARID2-B (Cabezas-Cruz et al., 2016a). According to the studies all IsJMJCs are structurally and functionally related, but IsLSDI-A and IsLSD1-B were suggested to have different functions because they are structurally different from each other (Cabezas-Cruz et al., 2016a).

3.6 Histone methyltransferases

Histone methyltransferases can catalyze the transference of one, two or three methyl groups into lysine and arginine residues of histone proteins. Histone methylation will occur mainly at specific lysine or arginine residues on histones H3 and H4 (Wood and Shilatifard. 2004). These modifications are called as lysine specific and arginine specific methylations. For both types of methylation processes, S-Adenosyl methionine (SAM) serves as cofactor and methyl group donor (Wood and Shilatifard. 2004).

The lysine specific histone methyl transferases include euchromatic histone-lysine N-methyltransferase (EHMT1 and 2) and DOT1L (Mozzettaet al., 2015; Volkel et al., 2007). EHMT1 encodes a protein that methylates the Lys-9 of histone H3 and contributes to repress gene transcription. Trimethylation of H3 at Lys-27 and trimethylation of H4 at Lys-20 contribute to repress gene transcription same as EHMT1. In contrast trimethylation of H3 Lys-4 and Lys-36 contribute to activate gene transcription (Chen et al., 2011).

Apart from the SET domain, the EHMT have two other domains known as pre-SET and the post-SET domain. The pre-SET domain contains cystein residues that form triangular zinc clusters and its help to stabilize the structure of EHMT. The SET domain is composed by approximately 130 amino acids and contains a catalytic core rich in β-strands. These β-strands will give slightly variations to the structure and it's altering the specificity for the methylation (Wood and Shilatifard. 2004). In order for the reaction to proceed, SAM and lysine residue of the substrate histone tail must be properly bound and oriented in the catalytic pocket in the SET domain. Then nearby tyrosine residue deprotonates the ε-amino group of the lysine residues. Finally, the lysine chain will produce a nucleophilic attack on the methyl group on the sulfur atom of the SAM molecule, then transferring the methyl group to the lysine chain (Trievel et al., 2002).

DOT1L is the only enzyme that can methylate a lysine residue in the globular core of the histone (Wood and Shilatifard. 2004). DOT1L contains its active sites in the N-terminal. Binding sites of SAM are link N-terminal and C-terminal domains of the DOT1L catalytic domain. C-terminal of DOT1L is important for substrate specificity because its carries positively charge region, these regions make attraction with negatively charge DNA backbone and DOT1L only methylate histone H3 (Min et al., 2003).

Arginine specific histone methyltransferases have two types of protein arginine methyltransferases. First type of protein arginine methyltransferase produces monomethylarginine and asymmetric dimethylarginine. Second types of protein arginine methyltransferase produce monomethyl (Branscombe et al., 2001). Protein arginine methyltransferases consists with SAM binding domain and substrate binding domain. Every protein arginine methyltransferases have unique N-terminal region and catalytic core. Catalytic mechanism governs by interactions between glutamate and nitrogen atoms on the arginine residue. These interactions are help to redistribute the positive charges and help to

deprotonation of one nitrogen group, then make a nucleophilic attack on the methyl group of SAM (Zhang et al., 2000).

The main aim of this research was to isolate the homolog of EHMT from the soft tick *O. moubata.*

4 HYPOTHESIS

Ornithodoros moubata has an EHMT with molecular signatures conserved in other EHMT homologs.

5 OBJECTIVES

1- Amplification and sequencing of *Ornithodoros moubata* EHMT.

2- Molecular and phylogenetic characterization of *Ornithodoros moubata* EHMT

6 MATERIALS AND METHODS

To accomplish our objectives various molecular biology and bioinformatics techniques were used in following order:

- Isolation of total RNA from *O. moubata*
- cDNA synthesis
- Gene amplification using polymerase chain reaction (PCR)
- Cloning and transformation in *Escherichia coli*
- Sequencing and sequence analysis

6.1 Isolation of total RNA from *O. moubata*

For isolation of total RNA from *O. moubata* the RNeasy®minikit (QIAGEN GmbH, Hilden, Germany, Cat. No.74104) was used.

Firstly, the ticks were washed by adding 1 ml of ethanol to 1.5ml centrifuge tube where the two adult ticks were placed. The ticks were placed on the filter paper to dry. Buffer RLT 700 µl was added to a clean 1.5ml centrifuge tube and the ticks were ground within the buffer.

The resulting lysate was then centrifuged at 12000 revolutions per minute (rpm) for three minutes. The supernatant was transferred to a new 1.5ml centrifuge tube. 70% ethanol 700µl was added to the lysate and mixed by pipetting. 700µl of the suspension was poured onto the RNeasy spin column and placed in a 2ml collection tube. RNeasy spin column was centrifuged at 9000 rpm for fifteen seconds.

After centrifugation, 700µl of buffer RW1 were added to RNeasy column and centrifuged at 9000 rpm for fifteen seconds. The flow-through was discarded and 500µl of buffer RPE was added to the RNeasy spin column. Further centrifugation took place at 9000 rpm for two minutes and the flow-through discarded.

The RNeasy spin column was placed in 1.5ml collection tube. 50µl of RNase free water was added directly onto spin column membrane and centrifuge for one minute at 13000 rpm to elute the RNA.

6.2 Primers Design

For primers design, chelicerate EHMT sequences were collected from GenBank using Basic Local Alignment Searching Tool (BLAST) (Altschul et al., 1990; Madden et al., 1996). Sequences were aligned using Multiple Alignment using Fast Fourier Transform (MAFFT) (Katoh et al., 2005; Katoh et al., 2013) tool and conserved regions were identified. These conserved regions were then used as input setting for the Primer Designing Tool (Ye et al., 2012) implemented in BLAST. Firstly, we aimed to amplify the C-terminus of *O. moubata* EHMT. With the C-terminus sequence information we then attempted to amplify the N-terminus of *O. moubata* EHMT. Both strategies are explained in more details below.

Obtaining the C-terminus of *O. moubata* EHMT

Initially, *Ornithodoros* spp. sequences were not available and degenerated primers were designed based on chelicerate EHMT protein sequences. For this, *I. scapularis* EHMT protein sequence was blasted against 'chelicerate database' to retrieve chelicerate homologs with the highest identity. Selected sequences were aligned in MAFFT and conserved regions were identified. Proximal and distal conserved regions identified in the C-terminus of selected EHMT were used to design forward and reverse degenerate primers. Initial attempts of amplifying the targeted fragment with the degenerated primer were unsuccessful. A further refined search strategy, including Transcriptome Shotgun Assembly (TSA) database, retrieved a transcript corresponding to the *Ornithodoros rostratus* EHMT (Genbank accession number GCJJ01002040). New primers were then designed based in conserved regions identified after aligning *I. scapularis* and *O. rostratus* EHMT nucleotide sequences.

Obtaining the N-terminus of *O. moubata* EHMT

After C-terminus of *O. moubata* EHMT was obtained, we proceed to obtain the N-terminus of *O. moubata* EHMT. The forward primer was selected from a conserved region between *I. scapularis* and *O. rostratus* sequences. For the reverse primer, the nucleotide sequences of the C-terminus of *O. moubata* (amplified as mentioned above) and *O. rostratus* were aligned. The reverse primer was selected from a conserved region between *O. moubata* and *O. rostratus*.

6.3 Reverse Transcription Polymerase Chain Reaction

The polymerase chain reaction takes place in two different stages. The first stage is the synthesis of the cDNA strand and the second step is the preparation of the PCR reaction.

First-strand cDNA synthesis

Invitrogen Superscript™ III Reverse transcriptase (Thermo Fisher Scientific, Carlsbad, USA, Cat. No. 18080044) kit was used for first cDNA strand synthesis. Briefly, 1µl of Oligo(dT) 20-mer, 1µl of total RNA (500ng/µl), 1µl of 10mM dNTP and 10µl of distilled water were added to nuclease-free microcentrifuge tube. The mixture was heated to 65˚C for five minutes and quickly chilled on ice for one minute. The content of the tube was spin-down by brief centrifugation and 4µl of 5X first strand buffer, 1µl of 0.1M DTT, 1µl of reverse transcriptase SuperScript™ II and 1µl RNase OUT were added to the above first-strand cDNA synthesis. The content of the tube was mixed gently and incubated at 50˚C for thirty to sixty minutes. Last step of the process was inactivated the reaction mixture by heating at 70˚C for fifteen minutes.

Preparation of PCR mixture

PCR reaction mixture was prepared in a final volume of 50µl. 25µl of 2X Master mix (containing reaction buffer, Taq polymerase and dNTPs, Thermo Fisher Scientific, Waltham, MA, USA, Cat. No. K0171), 1µl of forward primer (10µM), 1µl of reverse primer (10µM), 2µl of cDNA and 21µl of DNase-free water were added to 0.5ml microcentrifuge tube. The microcentrifuge tube was placed in thermo cycler and proceeds with the PCR synthesis.

Thermo cycler steps for PCR synthesis

- 95 ˚C for two minutes
- 95 ˚C for thirty seconds
- Annealing temperature (depend on the Tm values on the primers) for thirty seconds
- 72 ˚C for one minute
- Go to step 2, 40 cycles
- 72 ˚C for fifteen minutes
- 4 ˚C for forever or thirty minutes
- End

6.4 Gel electrophoresis

1.2% agarose gel was prepared using 1X TAE buffer.

The samples (15µl of each) and the 1Kb plusDNA ladder (5µl) (Thermo Fisher Scientific, Carlsbad, USA, Cat. No. 10787018) were mixed with 3µl and 1µl of 6X DNA Loading Dye

(Thermo Fisher Scientific, Carlsbad, USA, Cat. No. R0611), respectively.Samples and 1Kb plus DNA ladder were loaded into the wells of the agarose gel. The electrophoresis was run at 100 volts. The SYBR dye (Thermo Fisher Scientific, Prague, Czech Republic, Cat. No. S33102) was used for DNA staining and the DNA was visualized using a UV transilluminator.

6.5 Cloning

After successful PCR, the products were cloned into a plasmid DNA vector and then transformed into *E.coli*. Invitrogen TOPO TA Cloning® Kit (Thermo Fisher Scientific, Carlsbad, USA, Cat. No. K450002) was used.

Preparation of ligation mixture

4µl of PCR product, 1µl of salt solution, 1µl of TOPO vector were added to 0.5ml microcentrifuge tube. The reaction was mixed gently and incubated for five minutes at room temperature. The reaction was placed on ice and we proceed to the protocol of transformation of *E. coli* cells.

Transformation of *E. coli* cells

The 6µl of the ligation reaction were directly transferred to the *E. coli* competent cell tube and mixed gently using the micropipette. Then, the mixture was incubated on ice at thirty minutes; followed by a heat shock of 42 ˚C for thirty seconds. The cells were kept on ice for another two minutes. Afterwards, 250µl of Super Optimal broth with Catabolite repression (S.O.C) medium were added to the cell suspension. The cells were then incubated at 37˚ C in a shaking incubator for one hour. Different volumes of the cell suspension (100µl and 150µl) were plated on prewarmed agar plates containing the carbenicillin antibiotic (100µg/ml) (Sigma- Aldrich, Prague, Czech Republic, Cat. No. C1389-1G) and a layer of X-Gal (Thermo Fisher Scientific, Prague, Czech Republic, Cat. No. R0941).

E. coli plating

40µl of X-Gal were spread over the surface of the agar plate using a sterilized loop. Plates were incubated at 37˚ C until used. As mentioned above, different volumes of cell suspension were plated on the agar plates. The plates were incubated overnight at 37˚C.

Blue-white screening

Blue-white colonies appeared on the agar plates because of the function of the LacZ gene. The lacZ gene encodes β-galactosidase, which hydrolyzes X-Gal and formed the blue product. A successful cloning event disrupts the lacZ gene. Thus, no functional β-galactosidase was produced, X-Gal remained colorless.

6.6 Inoculation of bacterial cultures

6ml of Lysogeny Broth (LB) and 3µl of ampicillin antibiotic (100µg/ml) (Sigma-Aldrich, Prague, Czech Republic, Cat. No. A9393-5G) were added to 15ml tubes. The white *E. coli* colonies were picked from blue/white colonies that grew on the agar plates. The tubes were kept in shaking incubator at 37˚C for eight to ten hours.

6.7 Preparation of LB broth

5g of yeast extract, 10g of sodium chloride and 10g of tripton were placed in 1000ml Erlenmeyer flask. 1L of distilled water was poured to Erlenmeyer flask and autoclaved for two hours (Figure 5).

6.8 Preparation of agar plates

5g of yeast extracts, 10g of sodium chloride, 10g of tripton and 15g of agar were placed in 1000 ml Erlenmeyer flask. 1L of distilled water was poured to Erlenmeyer flask and autoclaved for one hour. Afterwards, carbenicillin antibiotic (100µg/ml) was added to Erlenmeyer flask before content was solidified (Figure 5).

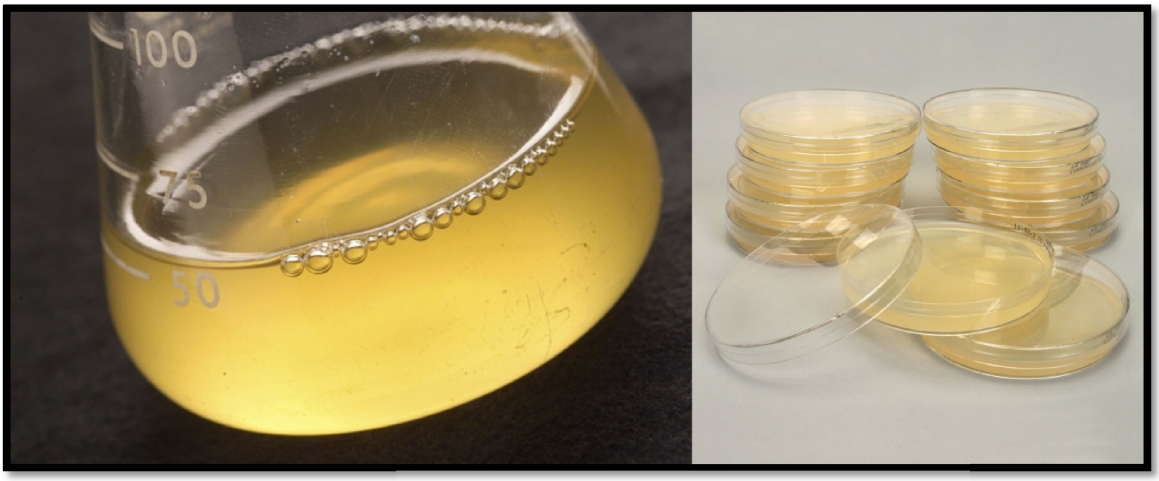

Figure 5.LB broth (left side), Agar plates (right side)

6.9 Isolation of high copy number plasmid DNA from *E. coli*

After inoculation, the falcon tubes were centrifuge for thirty seconds at 10000 rpm. The supernatant was pipetted out and the cell pellet was used for plasmid purification using the NucleoSpin plasmid Mini prepkit (Macherey-Nagel, Düren, Germany, Cat. No. 740588.250)

The cell pellet was resuspended in 250µl of Buffer A1. 250µl of Buffer A2 (Sodium dodecyl sulfate/alkaline lysis) was added and mixed gently by inverting the tube 6-8 times and incubated at room temperature for five minutes. Afterwards, 300µl of Buffer A3 (guanidine hydrochloride) was added and mixed by inverting the tube 6-8 times. The tubes were then centrifuged at 10000 rpm of five minutes. The columns were placed in the new collection tubes. Afterwards, 750µl of supernatant was pipetted out onto the nucleospin columns and centrifuged for one minute at 10000 rpm. Then, 600µl of Buffer A4 (ethanol) was added to columns and centrifuged for one minute at 10000 rpm. The flow-through was discarded and columns were placed in empty collection tubes. The columns were centrifuged for two minutes at 10000 rpm and discarded the collection tubes.

The Nucleospin columns were placed in new1.5ml microcentrifuge tube and 50µl of elution Buffer AE (5mMTris/ HCl) were added to the center of Nucleospin columns. The columns were incubated one minute at room temperature and centrifuged for one minute at 10000 rpm.

6.10 Restriction enzyme digestion

Restriction digest experiment was performed to detect plasmids containing inserts. Restriction endonucleases recognize specific sequences in the DNA. To check for inserts present in the isolated plasmid DNA, EcoRI (recognition sequence 5′-G/AATTC-3′) restriction enzyme was used. In the TOPO cloning vector, EcoRI sites flank the insertion site. Therefore, after EcoRI digestion, recombinant clones will have a band of the expected size (500bp) and another band of the size of the plasmid (3.9 kb). No recombinant clones will have only one band of the size of the plasmid. For the restriction digestion, firstly14µl of DNase water, 2µl of 10x NE buffer and 0.5µl (10000 units/µl) of restriction enzyme were added to a 0.5ml microcentrifuge tube. 3µl of the purified plasmid were added to the reaction tube. Final volume was 20µl.Then the samples were incubated for one hour at 37˚C and the result of the digestion reaction was visualized by gel electrophoresis.

6.11 Sample preparation for sequencing

For sequencing, 5µl of plasmid DNA sample and 5µl (5pmol/µl) of M13 forward primer (available in the TOPO TA Cloning® Kit) were added into 1.5ml centrifuge tube. The samples were sent to the GATC Biotech Lightrun sequencing service.

6.12 Sequence analysis

To confirm that the nucleotide sequence of interest was successfully amplified and sequenced, EcoRI sites were identified. Afterwards, the nucleotide sequence was blasted using the Blastn tool of BLAST software. The nucleotide sequence was translated to protein using ExPASy translate tool (Gasteiger et al., 2003).

Phylogenetic tree

To identify EHMT homologs, we searched arthropod and mammal databases using the Blastp tool from BLAST (Altschul et al., 1990; Madden et al., 1996). Selected sequences were belonged to *Apis mellifera, Bombus terrestris, Calliotropi scalcarata, Pogonomyrmex barbatus, Acryomyrmex echinatior, Solenopsis invicta, Orussus abietinus, Limulus polyphemus, Amblyomma aureolatum, I. scapularis, O. moubata, O. rostratus*, *Pteropus alecto, Pteropu vampyreus, Homo sapiens* and *Mus musculus*. Amino acid sequences were aligned with MAFFT configured for the highest accuracy using the scoring matrix 200PAM/Kd2, alignment strategy MAFFT-FFT-NS-I, gap opening penalty 1.53 and offset value 0.123 (Katoh et al., 2005; Katoh et al., 2013). Non-aligned regions were removed with Molecular Evolutionary Genetics Analysis (MEGA, version 6) software (Tamura K et al., 2013).The best-fit model of sequence evolution was selected based on corrected Akaike Information Criterion (cAIC) and Bayesian Information Criterion (BIC) implemented in MEGA. The JTT model, which had the lowest value of cAIC and BIC, was chosen for tree reconstruction (Jones et al., 1992). The neighbor joining (NJ) and maximum likelihood (ML) methods, implemented in MEGA, were used to obtain the best tree topologies. Reliability of internal branches was assessed using the bootstrapping method (100 bootstrap replicates).

Secondary and tertiary structure prediction of *O. moubata* EHMT.

The secondary structure was predicted by PSIPRED (Jones et al., 1999) software accessed through Phyre2 (Kelley et al., 2015) software. Briefly, an'evolutionary profile' was created to gather a large number of true EHMT homologs to capture the residue preferences at each position along the protein sequence length. To create this 'evolutionary profile', PSIPRED use the sequence of *O. moubata* EHMT as a query to search EHMT homologs using HHblits (Remmert et al., 2012), a sequence-search tool more accurate than PSI-BLAST. The resulting multiple sequence alignment is then used by PSIPRED to predict the secondary structure of the query.

For 3D modeling a PSI-BLAST was performed with five iterations using a non-redundant database to obtain the position-specific scoring matrix (PSSM). The PSI-BLAST was performed again using the generated PSSM against the Protein Data Bank (PDB) to obtain the structural template homologs for modeling. MAFFT was used to perform a multiple sequence alignment; the alignment was then transformed to the CLUSTAL format, saved as ALN file and submitted to MODELLER, a tertiary structure prediction tool (Webb et al., 2014).The 3D structure was visualized using UCSF Chimera (Pettersen et al., 2004).The quality of the models was evaluated using the protein model-qualifying server RESPROX (Berjanskii et al., 2012).

7 RESULTS

7.1 Primers for N-terminus and C-terminus fragments.

In order to amplify the EHMT gene of *O. moubata,* two set of primers were designed (Table 1). Firstly, we attempted to amplify the C-terminus of this gene using degenerated primers. However, the amplification using degenerated primers was not successful. A new set of primers (EHMTF1 and EHMTR1, Table 1) was then designed based on conserved regions identified after aligning *I. scapularis* and *O. rostratus* EHMT nucleotide sequences. Then the primers FullSF and FullSR (Table 1) were designed for amplification of the N-terminus (see methods for details).

Table 1.Primers used for amplification of N-terminus and C-terminus fragments of *O. moubata* EHMT.

Fragment	Names of the primers	Sequence 5'to 3'
N-terminus	FullSF	GTACTTGGACAGGGGCTGTG
	FullSR	GAAGGGCTGCAAGTGACACG
C-terminus	EHMTF1	GATCCACCRATGCTGTTTGAG
	EHMTR1	GYCACACTCACASGTGAAC

7.2 Amplification of the C-terminus of *O. moubata* EHMT

The PCR products were visualized in a DNA electrophoresis using a 1.2% agarose gel. The SYBR dye was used for DNA staining and the DNA was visualized using a UV transilluminator. The 1Kb plus DNA ladder was used to estimate the size of the PCR product. The size of the amplified fragment of C-terminus was between 500bp and 600bp (Figure 6).This DNA fragment corresponded to the expected size of the C-terminus of *O. moubata* EHMT. Attempts to amplify the N-terminus were not successful.

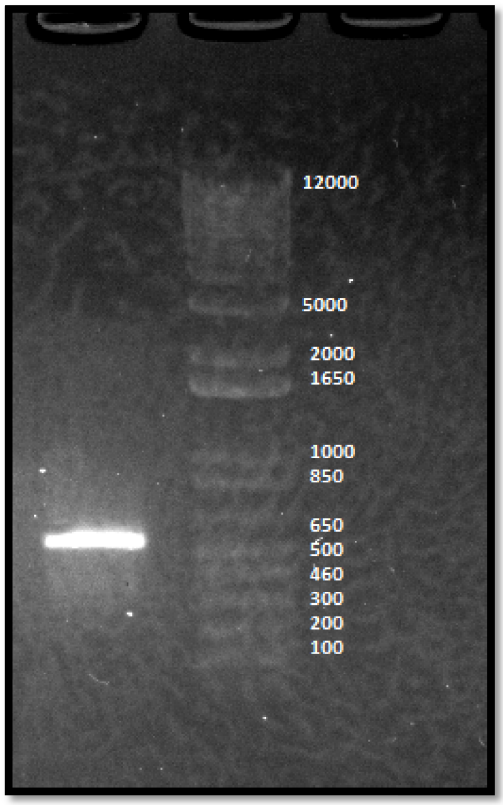

Figure 6.Agarose gel showing the C-terminus of O. moubata EHMT (right).1Kb plus DNA ladder (left) was used to estimate the size of the amplified fragment.

7.3 Transformation of *E. coli* competent cells and plasmid propagation

The PCR-amplified DNA fragment was cloned into the TOPO vector system and the *E. coli* competent cells were transformed with the circularized plasmid. After incubating agar plates at 37˚C for eight to ten hours, blue-white colonies grew in the agar plates. During inoculation, *E. coli* white colonies were picked and placed in 15ml tubes containing LB media. Cells were incubated for eight to ten hours at 37˚C in a shaking incubator.Cells were centrifuged and the pellet was conserved until further use (Figure 7).

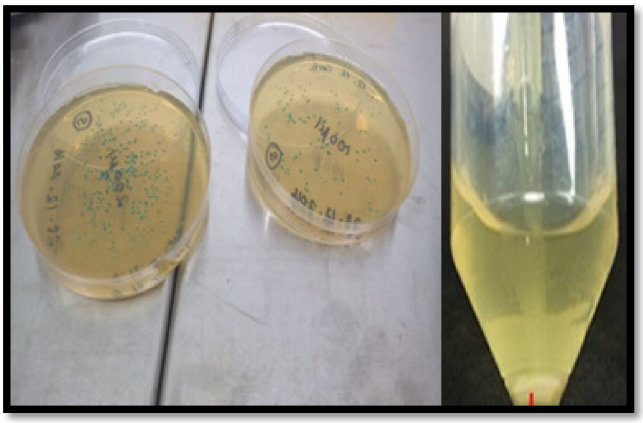

Figure 7.Blue-white colonies in plates prepared with the antibiotic carbenicillin (left side). Pellet of E. coli cells after the incubation (eight to ten hours) step (right side).

7.4 Plasmid purification and digestion

The cell pellet was used for plasmid purification and recombination events were detected using EcoRI restriction enzyme digestion. The presence of two bands (one of approximately 500 bp and the other of 3.9 kb) after the EcoRI digestion indicated the presence of recombinant clones that were sent for sequencing (Figure 8).

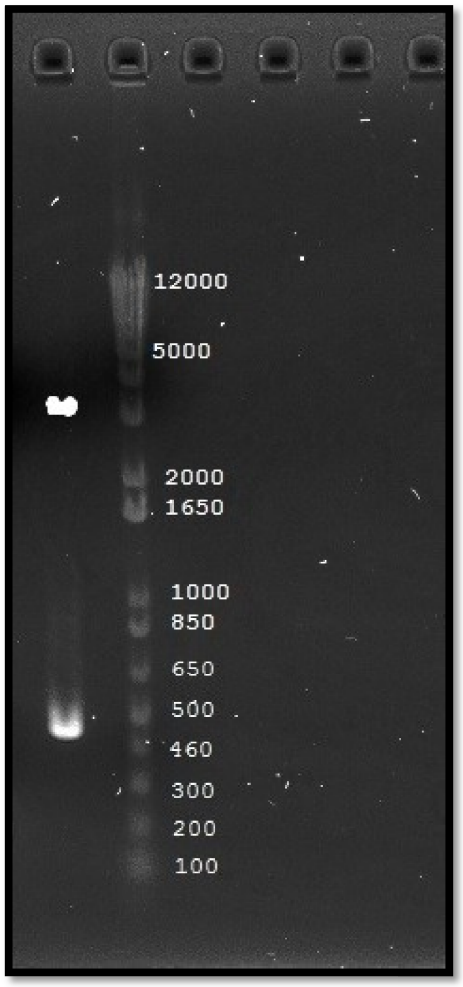

Figure 8.Agarose gel showing an example of the result of the restriction enzyme digestion (right). 1Kb plus DNA ladder (left) was used to estimate the size of the insert and the vector.

7.5 Sequence analysis

The sequence was submitted to Blastn and it showed high identity with EHMT of other species. Amino acid sequence analysis of the translated sequence confirmed the isolation of the C-terminus of *O. moubata* EHMT. The *O. moubata* EHMT sequence was submitted in Genbank and an accession number was provided (MF118910). Nucleotide sequence of *O. moubata* EHMT consist with 498 bp long, encoding a166 amino acid sequence.

7.6 Bioinformatics analysis of C-terminus of *O. moubata* EHMT

Amino acid sequence of *O. moubata* EHMT showed not only changes in amino acids but also the properties of these amino acids varied, compared to amino acid sequences of *I. scapularis, H. sapiens* and *L. polyphemus* EHMT. Such as, sequence alignment has shown *O. moubata* and *O. rostratus* were conserved with histidine amino residue while others were conserved with asparagine amino residues. Also *O. moubata* has alanine amino residue and others were conserved with methionine amino residues. However, several amino acid residues were found to be conserved between *O. moubata* and *O. rostratus*, *I. scapularis*, *H. sapiens* and *L. polyphemus* EHMTs (Figure 9). As expected, the sequence of *O. moubata* EHMT was highly similar to that of *O. rostratus.*

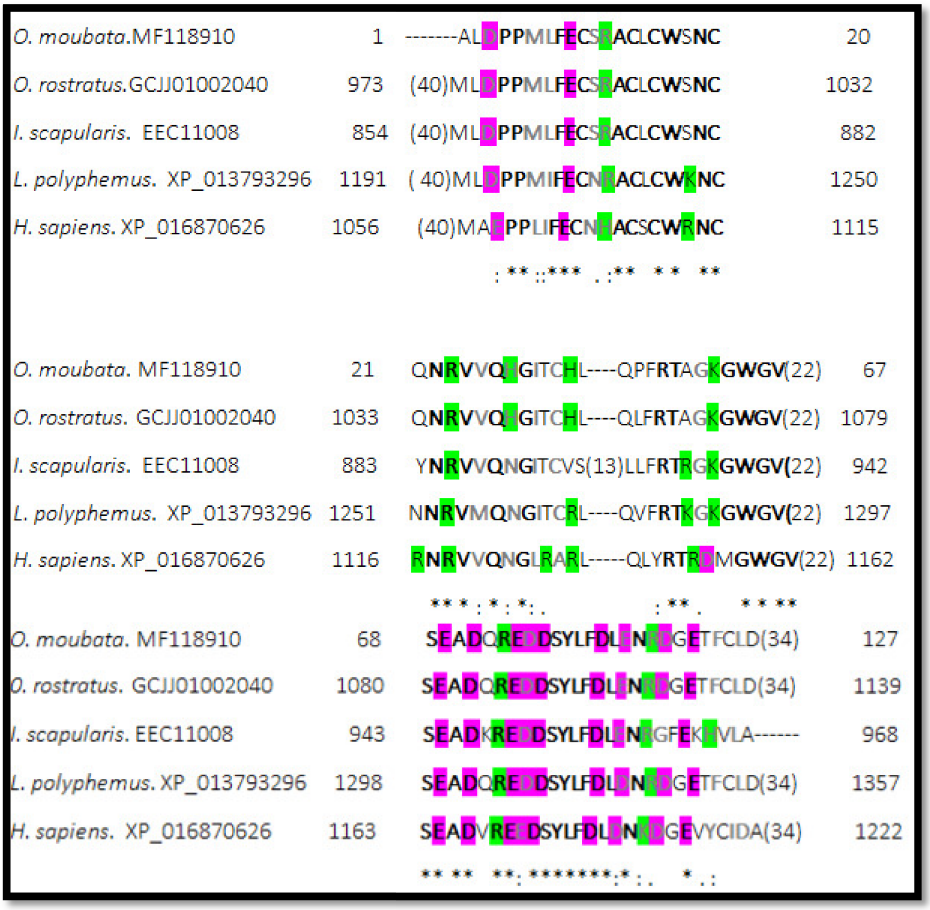

Figure 9.Multiple sequence alignment of O. moubata EHMT and other homologs of EHMTs. Bold color words represent conserved amino residues, green color represent positive charged amino residues and pink color represent negative charged amino residues. Multiple sequence alignment obtained using Clustal Omega (Sievers et al., 2011)

Conserved domain of *O. moubata* EHMT (Figure 10) was obtained using NCBI my domain search tool (Marchler-Bauer et al., 2015). The result confirmed that the C-terminus of *O. moubata* EHMT contains the typical SET domain found in members of histone methyltransferase family. N-terminus of *O. moubata* was not successfully amplified. Thus, SET domain was the only domain identified within C-terminus *O. moubata* EHMT. SET domain of *O. moubata* EHMT shared 89.36% and 74.36% identity with the SET domain of *I. scapularis* EHMT and *H. sapiens* EHMT2 respectively.

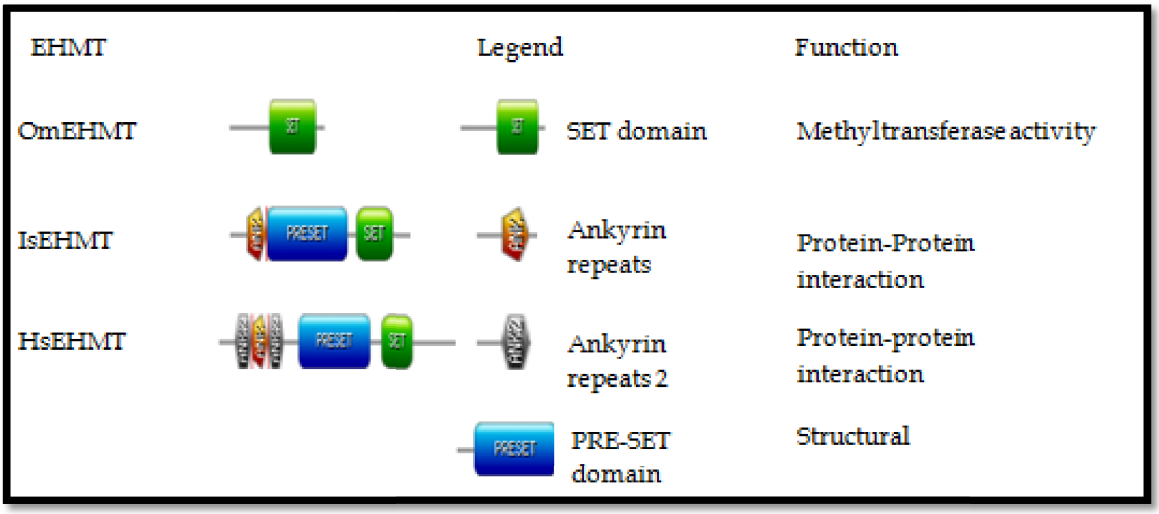

Figure 10.Major protein domains found in O. moubata (OmEHMT), I. scapularis (IsEHMT), Homo sapiens (HsEHMT). Graphic representations of the O. moubata, I. scapularis and H. sapiens EHMTs were obtained using ExPASy prosite my domain tool (Huloet al. 2008).

Phylogenetic tree

The phylogenetic tree was constructed based on EHMT homologs of arthropods and mammals. The phylogenetic tree analysis showed that the amino acid sequence of *O. moubata* EHMT was closely related to *O. rostratus* and other EHMT sequences from other chelicerates. EHMT from mammals clustered separated from insects and chelicerate sequences (Figure 11).

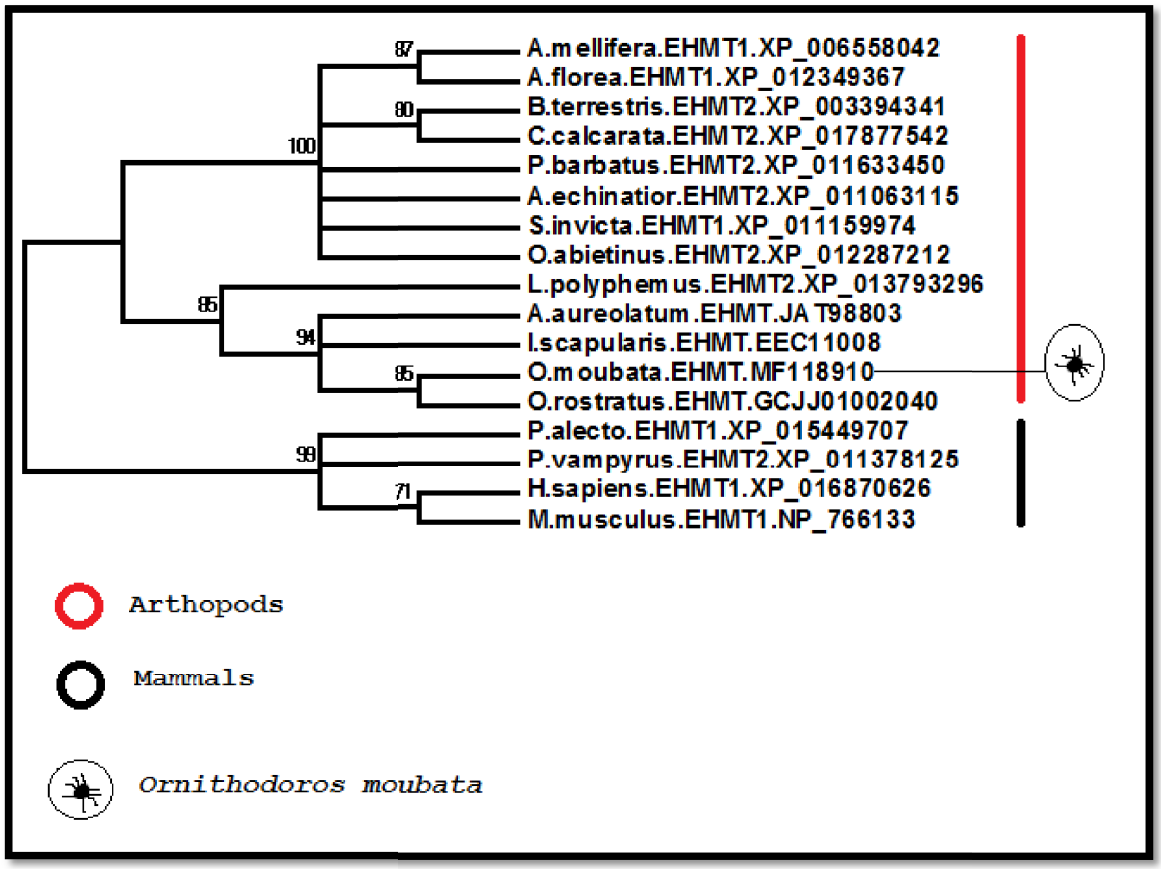

Figure 11.Phylogenetic analysis of EHMT amino acid sequences. A maximum likelyhood phylogenetic tree was built using the amino acid sequences of the EHMT identified in O. moubata and EHMT homologs.

7.7 Secondary and tertiary protein structure of C-terminus of *O. moubata* EHMT

In figure 12 below is a graphic representation of predicted secondary structure of *O. moubata* EHMT and figure 13 represents the tertiary structure of the SET domain of *O. moubata* EHMT. Secondary structure of *O. moubata* EHMT contained ten beta strands and one alpha helix.

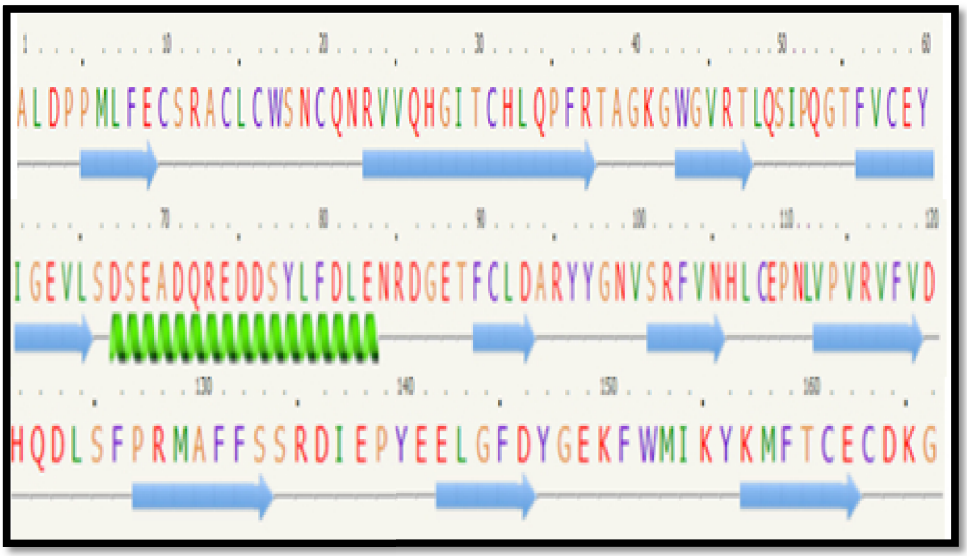

Figure 12.Secondary structure prediction of C-terminus of O. moubata EHMT by Phyre 2. Green color spiral represent alpha helix and blue color arrow represent beta strands. Colors represent difference in polarity of the amino acids; orange: small non polar amino acids (AA), green: hydrophobic AA, red: polar negatively charged AA, blue: polar charged AA.

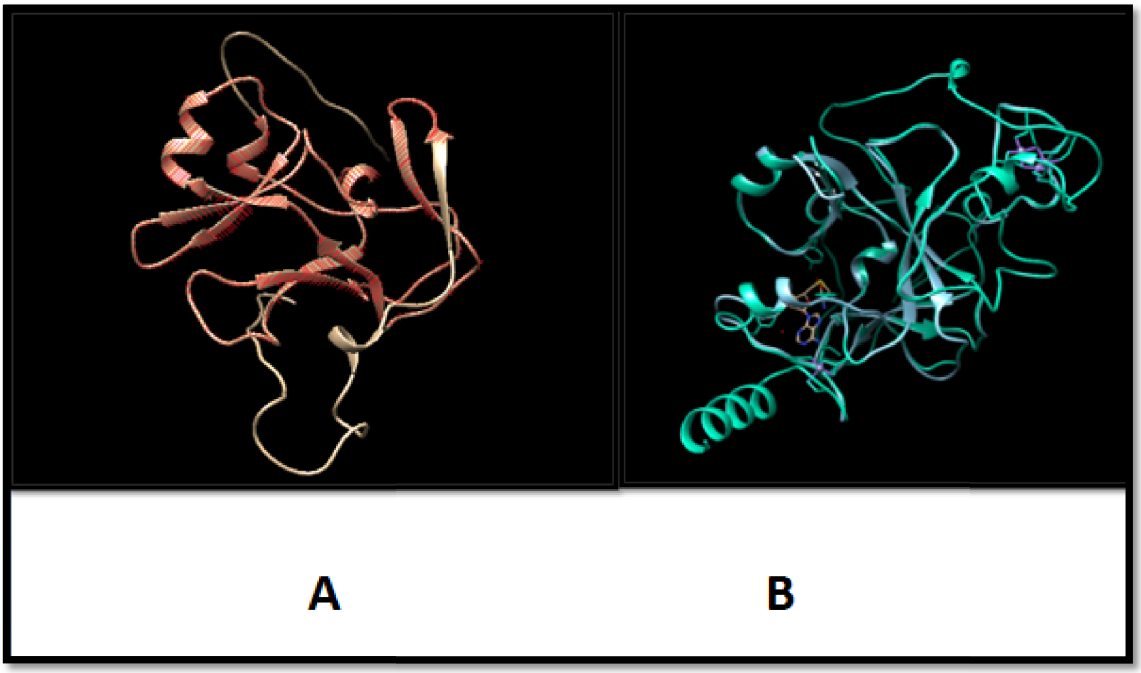

Figure 13.Predicted 3D protein structure of C-terminus of O. moubata EHMT (A).Red color represent SET domain of the predicted 3D protein structure. Super imposed image of predicted 3D protein structure of O. moubata EHMT and H. sapiens EHMT2 (B).

8 DISCUSSION

The main aim of this research was to isolate the euchromatic histone-lysine N-methyltransferase (EHMT) from the soft tick *O. moubata* (OmEHMT).This is the first characterization of a histone methyltransferase from the soft tick *O. moubata*.

Protein domains are distinct functional and structural units of proteins. They are responsible for particular functions or interactions with other proteins. The SET-domain was first recognized in conserved sequences in three *Drosopila melanogaster* proteins (Tschiersch et al., 1994; Jones and Gelbart. 1993; Stassen et al., 1995). These proteins were the suppressor of variegation 3-9 (Su(var)3-9) (Tschiersch et al., 1994), the Polycomb-group chromatin regulator Enhancer of zeste (E(z)) (Jones and Gelbart. 1993), and the trithorax-group chromatin regulator trithorax (Trx) (Stassen et al., 1995). Seven main families of SET-domain proteins were found; the SUV39, SET1, SET2, EZ, RIZ, SMYD,SUV4-20 and few orphan members such as SET7/9 and SET8 (Dillon et al., 2005). Each protein families have similar sequence motifs surrounding the SET domain and shared higher similarity in the SET domain.Previous study hasshown that the *I. scapularis* EHMT (IsEHMT) contains Ankyrin repeats (ANK), PRE-SET and SET-domain (Cabezas-Cruz et al., 2016a).In this study we amplified the C-terminus of *O. moubata* EHMT which contains only the SET domain. Multiple sequence alignment results showed that the SET domain of *O. moubata* EHMT shared 89.36% identity with SET-domain of IsEHMT, and it also shares 74.36% identity with SET domain of *H.sapiens* homolog (HsEHMT). These results showed that the SET domain of EHMT is highly conserved among soft and hard ticks and also other organisms. These suggested a conserved function of this domain in distantly-related taxa.

Previous studies have shown that the SET domain is important for the methyl transferase activity, ANK domain assists with the protein-protein interactions and PRE- SET domain is structurally important to protein structure (Cabezas-Cruz et al., 2016a). However, the function of SET domain-containing enzymes is not restricted to histone methylation. SET7/9 has been reported to methylate Lys189 in the general transcription factor TAF10 (Kouskouti et al., 2004), also been reported to methylate transcription factor p53 (Chuikov et al., 2004). These methylations resulted to increased affinity for RNA polymerase II and transcriptional activation of certain TAF10-dependent genes and increasing the stability of this short lived tumor-suppressor protein respectively. These observations suggested that SET domain-

containing proteins should not narrowly defined as histone lysine methyltransferases but instead call them protein lysine methyltransferases.

To identify the evolutionary relationship between *O. moubata* EHMT and other EHMT homologs, multiple sequence alignment and phylogenetic analysis were carried out. Sequence alignment results showed that the C-terminus fragment of *O. moubata* EHMT shared 94.58% identity with *O. rostratus* EHMT. The phylogenetic tree analysis also confirmed that the amino acid sequence of *O. moubata* EHMT was closely related to *O. rostratus* within chelicerates. Sequence alignment results showed that the C-terminus fragment of *O. moubata* shared 68.67% identity with *H. sapiens* EHMT2 and shared 67.47% identity with *H. sapiens* EHMT1. However, due to the high degree of identity among SET domains it is not possible to conclude whether the amplified fragment in *O. moubata* belong to EHMT1 or EHMT2. Further studies should attempt the amplification of the full length of the gene. A full length protein sequence may assist a more accurate classification.

The EHMTs are an evolutionarily conserved family of proteins that regulate H3K9 methylation of euchromatic DNA. Recent study has shown EHMT can act as a regulator of peripheral dendrite development, larval locomotor behavior, non-associative learning and courtship memory (Kramer et al., 2011). A recent study of *Drosophila* EHMT/G9a showed EHMT as key regulator of cognition that orchestrates an epigenetic program featuring classic learning and memory genes (Kramer et al., 2011). This observation was suggested EHMT not only has function of repressing the gene transcription. Further functional studies (e.g. RNA interference) should explore the functional role of EHMT in soft and/or hard ticks.

9 CONCLUSIONS

1. The C-terminus of a putative *O. moubata* EHMT was successfully amplified using the designed primers but attempts to amplify the N-terminus were not successful.
2. According to the sequencing, molecular features and phylogenetic analysis, we concluded that the fragment amplified corresponded to the SET domain of the homolog of EHMT in *O. moubata*.
3. Due to the high identity and conserved sequence features of SET domains, we do not have enough evidence to classify the amplified *O. moubata* EHMT fragment as EHMT1 or EHMT2.

10 REFERENCES

Aeschlimann., et al. Biology and distribution of ticks of medical importance. Crc Press. 1995; 177-189.

Altschul.S.F.,et al. Basic local alignment search tool. J Mol Biol. 1990; 215:403-10.

Allen.S.A. Ticks (Class Arachnida: order Acarina).Parasitic diseases of wild mammals. Iowa State Press.2001; 72-106.

Berjanskii. M., et al. Resolution-by-proxy: a simple measure for assessing and comparing the overall quality of NMR protein structures. J Bio mol NMR.2012; 53:167-180.

Branscombe.T.L., et al. PRMT5 (Janus kinase-binding protein 1) catalyzes the formation of symmetric dimethylarginine residues in proteins. J Biol Chem. 2001; 276(35): 32971-32976.

Brites-Neto J, Duarte KM, Martins TF. Tick-borne infections in human and animal population worldwide. Vet World. 2015; 8(3): 301-15.

Cabezas-Cruz.A., et al. *Anaplasma phagocytophilum* increases the levels of histone modifying enzymes to inhibit cell apoptosis and facilitate pathogen infection in the tick vector Ixodes scapularis. Epigenetic. 2016a; 11(4): 303-319.

Cabezas-Cruz.A., et al. Be aware of ticks when strolling through the park, 2016b; 4: 1-6.

Cabezas-Cruz.A., et al. Cancer research meets tick vectors for infectious diseases. Lancet Infect Dis. 2014; 14(10): 916-917.

Chen.F., et al. Methylation of lysine 9 of histone H3: role of heterochromatin modulation and tumorigenesis. Elsevier Inc. 2011; 149-157.

Choudhary.C., et al. Lysine acetylation targets protein complexes and co-regulates major cellular functions. Science. 2009; 325(5942): 834-840.

Chuikov.S., et al. Regulation of p53 activity through lysine methylation. Nature.2004; 432: 353.

Dillon. S., et al.The SET-domain protein super family: protein lysine methyltransferases. Genome Biol. 2005; 6(8): 227.

Gasteiger E., et al. ExPASy: the proteomics server for in-depth protein knowledge and analysis. Nucleic Acids Res. 2003; 31: 3784-3788.

Gomez-Dıaz.E., et al. Epigenetics of host–pathogen interactions: the road ahead and the road behind. Plos Pathog. 2012; 8(11): 1-5.

Gouy.M., et al. SeaView version 4: a multiplatform graphical user interface for sequence alignment and phylogenetic tree building. Mol Biol Evol. 2010; 27(2):221-224.

Greiss.S, Gartne.A.Sirtuin/Sir2 phylogeny, evolutionary considerations and structural conservation. Mol Cells.2009; 28: 407-415.

Guglielmone.A., et al. The Argasidae, Ixodidae and Nuttalliellidae (Acari: Ixodida) of the world: a list of valid species names. Zootaxa. 2010; 2528: 1-28.

Hulo.N., et al. The 20 years of PROSITE. Nucleic Acids Res. 2008; 245-249.

Jones.D.T. Protein secondary structure prediction based on position-specific scoring matrices. J Mol Biol. 1999; 292: 195-202.

Jones.D.T., et al. The rapid generation of mutation data matrices from protein sequences. Com App Bio. 1992; 8: 275-282.

Jones.R.S,Gelbart.W.M.The *Drosophila Polycomb*-group gene Enhancer *of* zeste contains a region with sequence similarity to trithorax. Mol Cell Biol. 1993; 13: 6357–6366.

Katoh.K., et al. MAFFT version 5: improvement in accuracy of multiple sequence alignment. Nucleic Acids Res. 2005; 33: 511-518.

Katoh.K., et al. MAFFT multiple sequence alignment software version7: improvements in performance and usability. Mol Biol Evol. 2013; 30: 772-780.

Keirans.J.E., et al. *Ixodes* (*Ixodes*) *scapularis* (Acari: Ixodidae): redescription of all active stages, distribution, hosts, geographical variation, and medical and veterinary importance. J Med Entomol. 1996; 33(3): 297-318.

Kelley.L.A., et al. The Phyre2 web portal for protein modeling, prediction and analysis. Nature Protocols 10. 2015; 845-858.

Kooistra.S.M, Helin.K.Molecular mechanisms and potential functions of histone demethylases. Nat Rev Mol Cell Biol. 2012; 13: 297-311.

Kouskouti.A., et al. Gene-specific modulation of TAF10 function by SET9-mediated methylation. Mol Cell. 2004; 14: 175–182.

Kramer. J.M., et al. Epigenetic regulation of learning and memory by *Drosophila* EHMT/G9a. Plos Biol. 2011; 9(1): e1000569

Madden.T.L., et al. Applications of network BLAST server. Methods Enzymol. 1996; 266:131-141

Manzano-Roman.R., et al. Soft ticks as pathogen vectors: distribution, surveillance and control. In Tech Open. 2012; 126-140.

Marchler-Bauer.A., et al. CDD: NCBI's conserved domain database. Nucleic Acids Res. 2015; 222-226.

Marmorstein.R, Trievel.R.C**.** Histone modifying enzymes: structures, mechanisms, and specificities. Biochim Biophys Acta. 1789; (1): 58-68.

Minucci.S, Pelicci.P.G. Histone deacetylase inhibitors and the promise of epigenetic (and more) treatments for cancer.Nat Rev Cancer. 2006; 6: 38-51.

Min. J., et al. Structure of the catalytic domain of human DOT1L, a non-SET domain nucleosomal histone methyltransferase. Cell. 2003; 112(5): 711-723.

Mozzetta.C., et al. Sound of silence: the properties and functions of repressive Lys methyltransferases. Nat Rev Mol Cell Biol. 2015; 16(8): 499-513.

Nowak.S.J, Corces.V.G. Phosphorylation of histone H3: a balancing act between chromosome condensation and transcriptional activation. Trends Genet. 2004; 20(4): 214-220

Nuttall, George H. F. Ticks and tick-transmitted diseases.Transactions of the Epidemiological Society of London. 1905; 24: 12-26.

Parola.P, Raoult.D. Tick-borne bacterial diseases emerging in Europe.ClinMicrobiol Infect. 2001: 7(2): 80-83.

Pettersen.E.F., et al. UCSF Chimera--a visualization system for exploratory research and analysis.J Comput Chem. 2004; 25(13): 1605-1612.

Remmert.M., et al. HHblits: lightning-fast iterative protein sequence searching by HMM-HMM alignment. Nat Methods **9**. 2012; 173–175.

Sievers.F., et al. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Mol Sys Bio. 2011; 539.

Stassen.M.J., et al. The *Drosophila* trithorax proteins contain a novel variant of the nuclear receptor type DNA binding domain and an ancient conserved motif found in other chromosomal proteins. Mech Dev. 1995; 52: 209–223.

Tamu.ra.K., et al. MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. Mol Biol Evol. 2013; 30: 2725-2729.

Telford.S.R, Goether. H.K. Emerging and emergent tick-borne infections: discovered and better characterized or truly "New". Parasitology. 2004; 129(1): 301-327.

Trievel.R.C., et al. Structure and catalytic mechanism of a SET domain protein methyltransferase. Cell. 2002; 111(1): 91-103.

Tschiersch.B., et al. The protein encoded by the *Drosophila* position-effect variegation suppressor gene Su(var)3-9 combines domains of antagonistic regulators of homeotic gene complexes. EMBO J. 1994; 13: 3822–3831.

Volkel.P., et al. The control of histone lysine methylation in epigenetic regulation. Biochimie. 2007; 89:1-20.

Wion.D, Casadesus.J. N6-methyl-adenine: an epigenetic signal for DNA protein interactions. Nat Rev Microbiol. 2006; 4(3): 183-192

Wood.A, Shilatifard. A. Posttranslational modifications of histones by methylation. Adv Protein Chem. 2004; 67: 201-222.

Webb.B., et al. Comparative protein structure modeling using modeller. CurrProtoc Bioinformatics. 2014; 47:1-32.

Yang.X.J. MOZ and MORF acetyltransferases: molecular interaction, animal development and human disease. Biochim Biophys Acta. 2015; 1853(8): 1818-1826.

Ye J., et al. Primer-BLAST: a tool to design target-specific primers for polymerase chain reaction**.** BMC Bioinformatics. 2012; 13: 134.

Zhang.X., et al. Crystal structure of the conserved core of protein arginine methyltransferase PRMT3. Embo J. 2000; 19(14): 3509-3519.

Zhang.Y., et al. Transcription regulation by histone methylation: interplay between different covalent modifications of the core histone tails. Genes Dev. 2001; 15(18): 2343-2360.