Comprehensive phylogenetic, similarity and allergenicity analysis of *Boophilus* genus tick Tropomyosin protein

Mohammad Mahdi Ranjbar[1], Sedigheh Nabian[2], Seyed Dawood Mousavi Nasab[3], Mohammad Taheri[4], Khodayar Ghorban[5], Maryam Dadmanesh[6], Mohammad Malekan[7], Nayeb Ali Ahmadi[*,8], Masoumeh Imanzad [9]

[1]Department of Immunology, University of Tehran, Tehran, Iran.
[2]Department of Parasitology and Tick and Tick Born Diseases Center, Faculty of Veterinary Medicine University of Tehran, Tehran, Iran.
[3]Department of Virology, Faculty of Medical Sciences, Tarbiat Modares University, Tehran, Iran.
[4]Rastegar Central Laboratory, Faculty of Veterinary Medicine, University of Tehran, Tehran, Iran.
[5]Department of Immunology, School of Medicine, AJA University of Medical Sciences, Tehran, Iran.
[6]Department of Clinical Science, School of Medicine, AJA University of Medical Sciences, Tehran, Iran.
[7]Department of Clinical Science, Faculty of Veterinary Medicine, University of Tehran, Tehran, Iran
[8]Proteomics Research Center, and Department of Medical Lab Technology, Faculty of Paramedical Sciences, Shahid Beheshti University of Medical Sciences, Tehran, Iran
[9]Faculty of Medicine, Tabriz University of Medical Sciences, Tabriz, Iran

*Corresponding Author: email address: **:** nayebalia@yahoo.com , nayebalia@sbmu.ac.ir (NA. Ahmadi)

ABSTRACT

Boophilus genus ticks are responsible for transferring some pathogens and reducing production factors in cattle. Tropomysin (TPM) protein has actin regulator activity and playing important role in immune and allergic reactions. The main goal is to determine different aspects of phylogenetic, similarity, homology, structure and allergenicity of TPM protein. In prior study, we identified TPM by using Mass-spectrometry in *Boophilus anulatus* larva proteins extraction. Analysis by NCBI and Mascot software showed complete similarity of this protein with *Boophilus microplus*. TPM Blasting, invertebrates TPM sequences retrieval, aligning and analyzing of conserved and variable regions along sequences were next steps. Also, construction the phylogenetic tree, overall mean distances estimation, homology protein secondary structure, allergencity analysis was achieved. The most similar sequences to *Boophilus* genus TPM are *Haemaphysalis* sp., *Scolopendra* sp. and etc., respectively. The multiple sequence alignment showed that conserved and variable regions stretched in different part of TPM. The close relationships in Phylogenetic tree between Ticks and Mites were seen, although the TPM sequences in ticks are more similar to each other than to mites and assume as the nearest relatives. Insects TPM like worms, located in two separated clades, and *Trichinella spiralis* in worm clades are more related taxa to members of ticks and mites groups. Furthermore, overall mean distances over sequence pairs reflects TPM conservation during speciation. TPM has high homology in different species and has two domain of α-helix that cannot form disulfide bonds. Finally, allergenicity analysis by separated and hybrid approach showed it undoubted is allergen and candidates some peptides as responsible for allergenicity of TPM. The comprehensive analysis of TPM has never been easy, especially when we attempt to make statements from different aspects about this protein. Our study revealed the some unique and valuable aspects of TPM protein of *Boophilus* genus, and will help to further studies on mentioned protein.

Keywords: Tropomyosin, *Boophilus* genus, Phylogenetic, Similarity, Allergenicity

INTRODUCTION

Boophilus genus ticks are responsible for transferring some pathogens (such as *Babesia bigemina*, *B.bovis*, *Anaplasma marginale* and some bacteria) and reducing production factors, thus are one of most important tick in cattle [1,2].

Tropomyosin (TPM) protein has actin regulator activity and playing important role in immune and allergic reactions, thus candidates for vaccine in some tick types. This molecule belongs to family of highly conserved proteins with multiple isoforms found in both muscle and non-muscle cells of all species of vertebrate and invertebrate. Its native structure consists of two parallel alpha-helical TPM molecules that are wound around each other forming a coiled-coil dime [3-5].

Comparative sequence analysis is an important tool for most of scientists that are occupy in different branches of biology. Phylogenetic tree showing the evolutionary relationships, similarities and differences among various species and inter species, based upon similarities and differences in their genetic characteristics. The amount of changes among the sequences reflects the evolutionary relatedness of the organisms. When sequences from two species are very similar, they are thought to be closely related and share a more recent common ancestor, and when sequences from two species are more dissimilar, the species are thought to be more distantly related [6-7]. Some area of phylogenetic method applications are detection of orthology and paralogy, estimating divergence times, reconstructing ancient proteins, finding the residues that are important to natural selection, detecting recombination points and determining the identity of new pathogens and vaccine design[7]. Also, knowledge of relationships is crucial to our understanding of the evolution of DNA and proteins and of development. It will also play an important role in the analysis of the sequence data that is being produced by worldwide genetic projects [8].

The prediction of allergenic proteins is becoming very important in present time due to use of modified proteins in foods (genetically modified foods), evaluation allergenicity protein of extrinsic protein, therapeutics, bio-pharmaceuticals etc. World Health Organization (WHO) and Food and Agriculture Organization (FAO) realize the importance of prediction and proposed guidelines to assess the potential allergenicity of proteins (http://www.fao.org/es/ESN/food).

Review on literature shows that there is not any comprehensive study on different aspects of phylogenetic, similarity, homology and allergenicity of TPM protein. Thus,

The aims of study are: 1) analysis of sequence and genetic characteristics of TPM protein in *Boophilus* tick; 2) to contribute to the understanding of genetic similarity and differentiation of *Boophilus* TPM protein in contrary with ticks, mites, worms and insects populations and vertebrates; 3) to further analyze the phylogenetic relationships, based on the TPM protein, between populations of ticks and mites and some other species; 4) to elucidate homology of this protein and prediction of its structure; 5) and at least, preparing suitable background for production wide range effective vaccines for parasites and allergencity researches.

MATERIALS AND METHODS

Identification of TPM by using Mass-spectrometry:

In 2012, our group by using of one and two dimensional electrophoresis and then Mass-spectrometry identified presents of immunogenic protein of TPM with 37 Kd weight in *B. anulatus* larva proteins extraction [9]. Analysis sequences of this protein by National Center for Biotechnology Information (NCBI) and Mascot software showed complete similarity (100%) of this protein with *B. microplus*. As there was not any sequences of *B. anulatus* in Databases, Thus sequence of *B. microplus* (nearest genus) retrieved and used as subjected sequence for next steps.

Sequence analysis:

Protein sequence statistics for TPM protein of *B. microplus* including length, the molecular weight (Mw), isoelectric point (pI) and amino acid distribution was calculated and arranged in Table 1 by using of Expasy tools (http://us.expasy.org/tools/pi_tool.html).

Retrieving, blasting, alignment and conserved and variable regions of sequences:

complete protein sequences of TPM of *B. microplus* and TPM sequences of other tick, mite, worms, Insects and some vertebrates were retrieved from GenBank (http://

ncbi.nlm.nih.gov/) and UniprotKB (http:// www.expasy.org/uniprot) databases. Also, *Boophilus* TPM directed to protein similarity blast search (http://blast.ncbi.nlm.nih.gov/Blast.cgi) for evaluation of similarity with sequences report in different species and we selected the closest hits and summarized in informative tables. Obtained sequences were aligned by using ClustalX and analyzed in Bioedit software version 7.7.9 [10].

At last step, very short sequences and areas with ambiguous alignment or containing poly-N stretches were excluded from the analyses. The most highly conserved and variable region were evaluated by T-coffee software and showed in Figure 1.

Construct the phylogenetic tree and overall mean distances:

Selected, aligned and edited sequences of TPM protein directed to phylogenetic tree by using MEGA5.3 software package [11]. The phylogenetic analysis was conducted based on the different sets of aligned sequences of the ticks, mites, worms, Insects and mammals and chicken. To each of datasets, added sequences which were used as the out-group. Trees were constructed using the neighbor-joining (NJ) algorithm under the global gap removal option and Kimura's two-parameter substitution model [12]. Robustness of phylogenetic analysis was measured by bootstrap analysis with 10,000 replications. The percentage of replicate trees in which the associated taxa clustered together in bootstrap test is shown next to the branches [13]. The number of amino acid substitutions per site from averaging over all sequence pairs is shown. Standard error estimate(s) are shown above the diagonal. Analyses were conducted using the Poisson correction model [14]. The analysis involved 29 (for parasites) and 13 amino acid sequences. All positions containing gaps and missing data were eliminated. There were a total of 256 (for parasites) and 238 (for vertebrates) positions in the final dataset. Evolutionary analyses were conducted in MEGA5 [11].

Homology, orthologous and paralogous of sequences:

TPM protein sequences of some mammals, chicken and parasites selected for homology study. TPM directed to phylogenetic tree by using MEGA5.3 software package. Also, parasites group assumed as the out-group. Trees were constructed using the neighbor-joining (NJ) algorithm and robustness of phylogenetic analysis was measured by bootstrap analysis with 10,000 replications.

Protein secondary structure prediction for TPM of *Boophilus genus*:

Prediction of secondary structure achieved by using of PSIPRED (http://bioinf.cs.ucl.ac.uk/psipred) and Scrach (http://scratch.proteomics.ics.uci.edu/) online servers. Then, results of these tools compare to each other to better and accurate understanding of structure TPM protein in *B.* genus.

***In-silico* evaluation of allergencity**

AlgPred (allergenicity prediction) server (http://www.imtech.res.in/raghava/algpred/) cause of its features in this study was used [15]. AlgPred allows prediction of allergens based on similarity of known epitope with any region of protein. The mapping of IgE epitope(s) feature of server allows user to locate the position of epitope in their protein and also, allows predicting allergens based on SVM (support vector machines) modules using amino acid or dipeptide composition. It facilitates BLAST search against 2890 allergen-representative peptides (ARPs) and assign a protein allergen if it have a BLAST hit. Finally, Hybrid option of server allows predicting allergen using combined approach (SVMc + IgE epitope + ARPs(allergen representative peptides) BLAST).

RESULTS

Protein sequence statistics for TPM protein of *B. microplus*

Some general sequence analysis of TPM protein of *B.microplus* are shown in Table 1.

Blasting and sequence alignment results

TPM of *B. microplus* (AC: O97162) used as entry data to BLAST search for finding sequences with similarity to our sequence. BLAST result revealed that most similar sequences to TPM protein are; *Haemaphysalis longicornis, Haemaphysalis qinghaiensis, Scolopendra sp*. and etc., respectively. Details of similarity are

summarizing in Table 2. The multiple sequence alignment (MSA) for ticks, mites, worms and insects produced by T-coffee [16]. Results showed that conserved and variable regions stretched in different part of protein (Fig.1). Column of sequences start with *Boophilus* sequences and other ticks, then mites, worms and insect from ascending to descending, respectively.

Table 1. Sequence information and amino acid distribution by using Expasy online tools servers. The molecular weight of protein was calculated using Compute pI/Mw (http://us.expasy.org/tools/pi_tool.html).

Organism (name & descriptions)	**Sequence type**	**Length**	**Weight (double-stranded)**	**Isoelectric point (PI) & Aliphatic index (AI)**	**Amino acid distribution**		
					aa name*	**aa count**	**Frequency**
Tick (TPM protein of *B. microplus*) AC: KC253896	Protein	284aa	33.001 kDa	PI= 4.9 AI= 76.056	A,C,D,E,F G,H,I,K,L M,N,P,Q,R S,T,V,W,Y	34,0,16,59,2,5 ,2,9,26,28,10, 7,0,23,23,13, 10,13,0,4	0.120,0, 0.056,0.208 0.007,0.018, 0.007,0.032, 0,92,0.099, 0.035,0.025, 0,0.081, 0.081,0.046, 0.035,0.046, 0.000,0.014

Table 2. NCBI protein BLAST results for TPM in *Boophilus* genus. Ranking of sequences is based on their Max score, Identity and E.value and arranged ascending to descending, respectively.

Number of protein sequence	**Name (**Description)	Max score	Max identity	E.value	Accesion number
1	*Haemaphysalis longicornis*	529	99%	0.0	Q8IT89.1
2	*Haemaphysalis qinghaiensis*	528	99%	0.0	ABQ96858.1
3	*Scolopendra sp.*	459	86%	1e-160	AAR87377.1
4	*Neoscona nautica*	459	86%	2e-160	AAR87381.1
5	*Dermanyssus gallinae*	455	89%	6e-159	CAJ44440.1
6	*Metaseiulus occidentalis*	454	89%	8e-159	XP_003745223.1
7	*Dermatophagoides farinae*	452	89%	2e-157	BAA04557.1
8	*Blomia tropicalis*	451	88%	3e-157	ABU97466.1
9	*Aleuroglyphus ovatus*	449	88%	2e-156	AAX37287.1
10	*Psoroptes ovis*	449	88%	2e-156	CAJ38272.1
11	*Dermatophagoides pteronyssinus*	448	88%	3e-156	AAB69424.1
12	*Chortoglyphus arcuatus*	447	88%	5e-156	AEX31649.1
13	*Sarcoptes scabiei*	446	88%	3e-155	AFH08744.1
14	*Tyrophagus putrescentiae*	438	86%	2e-152	AAT40866.1

Figure 1. Comparison of *Boophilus* genus TPM with the TPM amino acid sequences of other ticks, mites, insects and worms. This figure just showed part of the multiple sequence alignment result as produced by T-coffee and edited in Bioedit7.9 software (Genebank accession numbers not shown). Conserve, semi-conserved and variable regions along TPM sequence in parasites could be seen in this figure. All residues that are identical to the top sequence in an alignment as a dot ('.') used and ('-') character denotes gaps.

Construct the phylogenetic tree and overall mean distances

For the phylogenetic analysis of TPM proteins sequences of ticks (4 sequences) and mites (12 sequences), worms (8 sequences), Insects (4 sequences) and mammals and chicken (13 sequences) along with TPM protein sequence of *b. microplus* (AC: O97162), obtained from GenBank and UniprotKB. The sequences were aligned, compared and edited using Bioedit software version 7.7.9. MEGA 5.3 software packages used for construction of phylogenetic tree and calculation overall mean distances (Figure 2 and table 3). Phylogenetic tree in Figure 2 shows the close relationships between ticks and mites in TPM protein sequence, although the TPM sequences in ticks (like *Boophilus* genus, *Haemaphysalis* genus, *Ixodes scapularis* and *amblyomma maculatum*) are more similar to each other than to mites and assume as sister group (the nearest relatives).

Amblyomma maculatum (AC: AEO36033) has more distances relationship to other tick and neither mites, located in different branches near insects. *B.s* genus and *haemaphysalis* genus are in sister branches. In out-group branches, as it shown, TPM in mammals and chicken has three or four different chains (alpha 1, 2 (in human and mouse 4) and beta) and seems to alpha chains are more related to each other than to beta chains.

Insects TPM like worms, located in two separated clades, *Bombyx* mori and *Drosophila melanogaster* are in sister branches and are more related taxa to members of family of ticks and mites, respectively. Also, *Trichinella spiralis* in worm clades is more related taxa to ticks and mite groups. Furthermore, estimation of average evolutionary divergence over parasite Sequence pairs demonstrated in Table 3. Results show overall mean distances over sequence pairs in parasites and vertebrate (mammals and chicken) in not high and it reflects conservation of this protein

during speciation in these two groups. Also, overall mean distances over sequence pairs in vertebrates significantly, is lower than parasites.

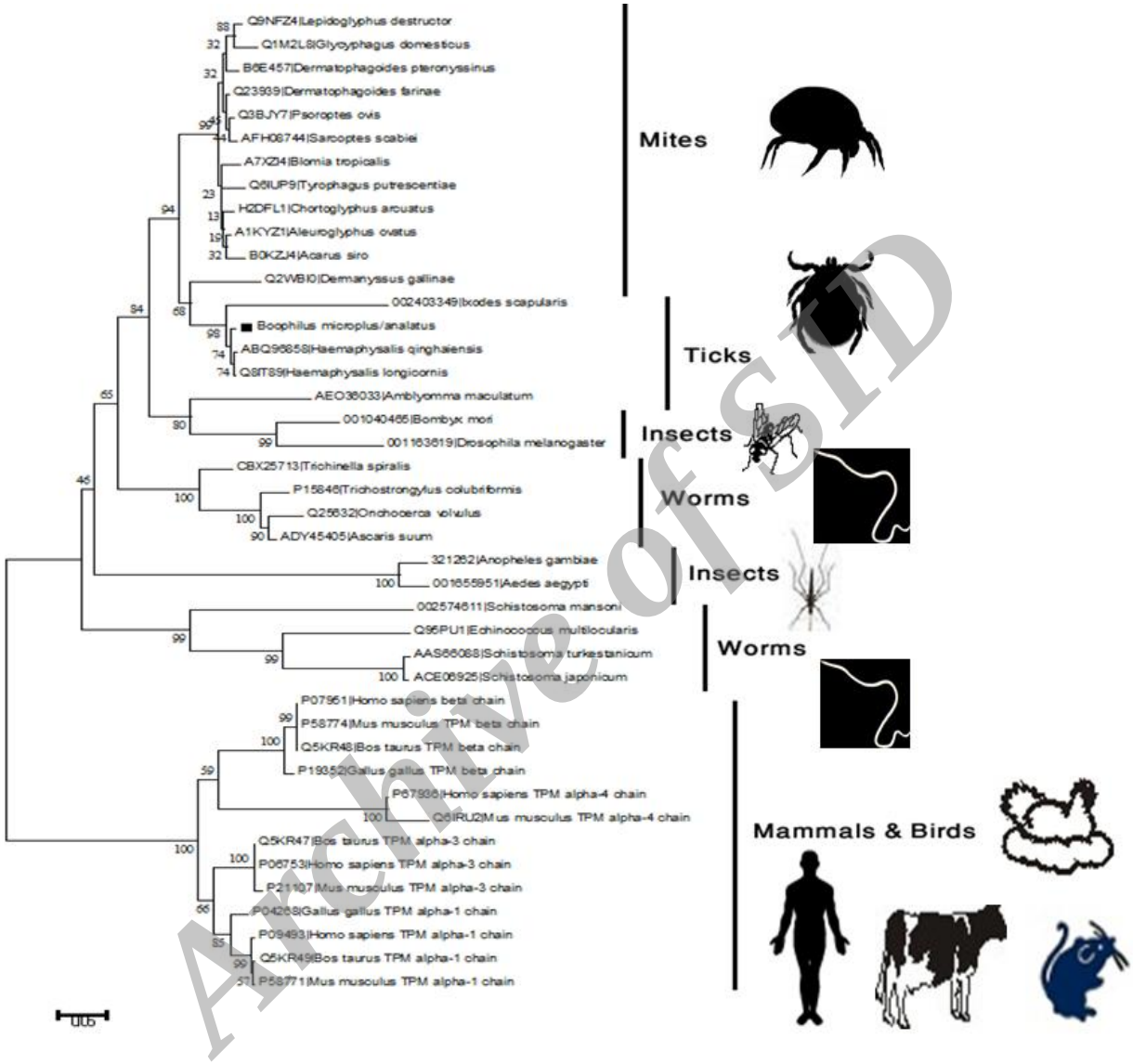

Figure 2. Phylogenetic tree constructed based on protein sequences with MEGA 5.3, illustrating the relationships in TPM protein from *B. microplus* among other parasite and mammals and chicken. TPM protein from *B. microplus* marked by ■ and also, mammal's TPM (alpha and beta chains) assumed as out group in tree. Tree was constructed by using the neighbor-joining (NJ) algorithm based on differences in TPM protein sequences of different species. Units at the bottom of the tree indicate the number of substitution events. The length of each pair of branches represents the distance between sequence pairs. The dataset was resampled 10,000 times using the bootstrap method. The sequence information at the tips of the branches includes an accession numbers of the sequences and tick or mite name for each sequence. Some very short and incomplete sequences or sequences with ambiguous alignment or containing poly-N stretches removed after alignment and tree not showing them.

Table 3. Comparison of estimation of average evolutionary divergence over sequence pairs between parasites and vertebrates.

Protein name	Organisms	Overall mean distances	Standard error estimates (S. E.)	Selected variance estimation method	Selected substitution mode for amino acid
TPM	Tick, mites , worms and insects	d = 0.341	0.023	Bootstrap method	Poisson model
¨	Some mammals and chicken	d = 0.151	0.016	Bootstrap method	Poisson model

Homology

Homologous sequences are the term that uses for orthologous and paralogous sequences. These sequences are result of horizontal transfer between 2 species, and not common ancestor. Homologous sequences as result of convergence. Two genes are orthologous if they diverged after a speciation event and two genes are paralogous if they diverged after a duplication event. It is likely that two orthologs have similar function, these functions not necessarily identical.

Paralogous usualy have different function [17, 18]. Figure 3, depicts homology (and also, orthology and paralogy) of TPM protein in mammals and chicken in comparison with some selected parasites and they classified in tree separated groups. A result notices that TPM protein has high homologoy in different species and include alpha and beta chains in higher animal species (especially vertebrates).

Alpha (1, 3 or 4(only in human and mouse) and beta chains are more related to each other and are in separated clades (groups), thus they are orthologous. Besides, alpha to beta chain have paralogous relationship but in lower animal species (non-vertebrates, such as *Boophilus* genus) TPM is just a single sequence and so, relationship between sequence is only orthologous.

Among alpha chains sequences, alpha 4 chains are nearest branches to beta chains and so located in same group. *Boophilus* genus (tick) and *S. scaibiei* (mite) are located in sister branches.

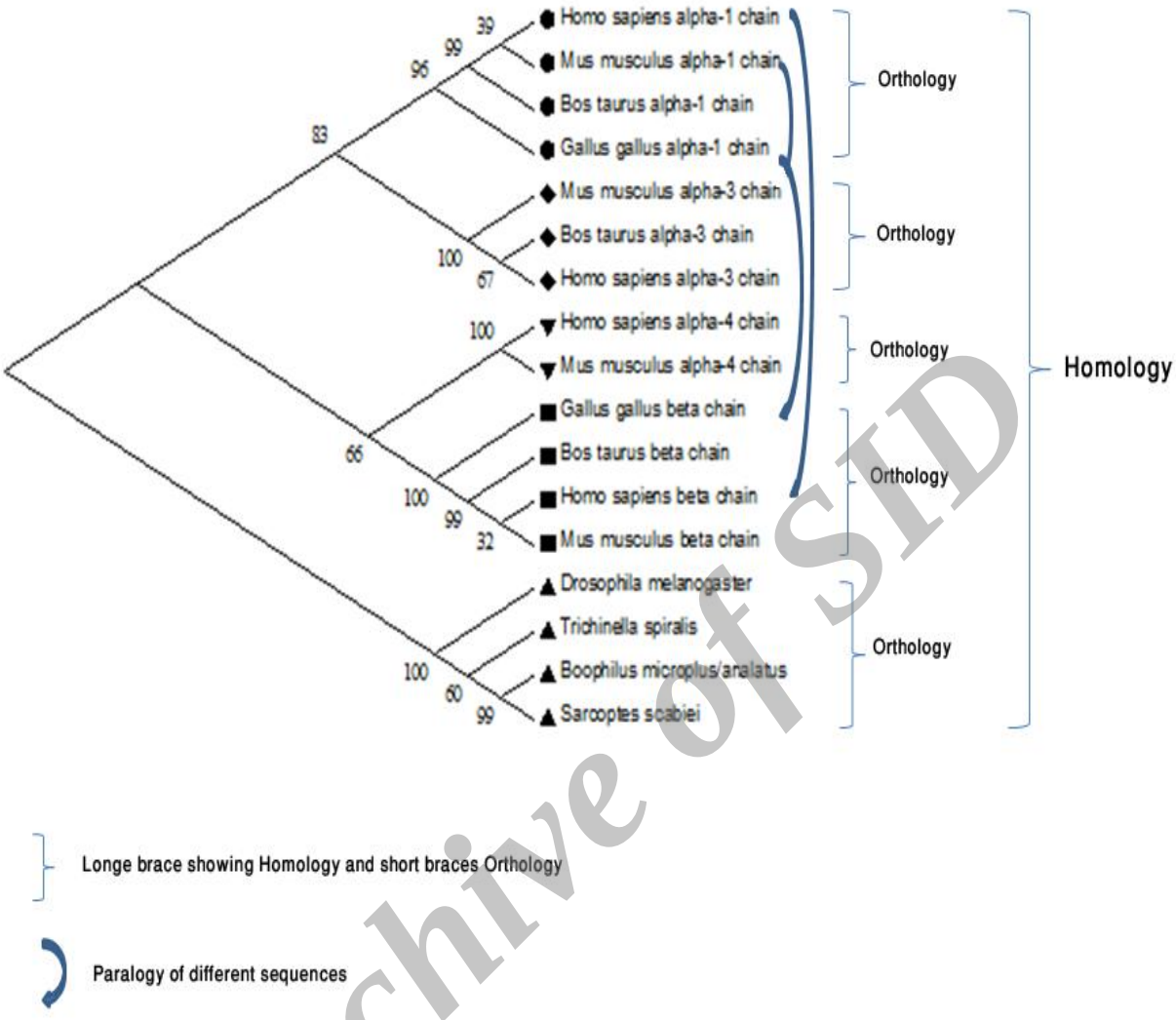

Figure 3. Homology relationships of TPM protein in some parasites and vertebrates. Mega5.3 software package used for construction tree for homology analysis of TPM protein. Tree was constructed by using the neighbor-joining (NJ) algorithm, and the length of each pair of branches represents the distance between sequence pairs. Straight branch style selected for showing homology-orthology and paralogy of TPM chains among taxa. The dataset was resampled 10,000 times using the bootstrap method. The sequence information at the tips of the branches includes an name for each sequence and symbols that; ● represents alpha 1 chains, ◆ alpha 3 chains, ▼ alpha 4 chains, ■ beta chains and ▲ parasites (also this group assumed as out-group).

Prediction frequency of secondary structure:

Results of Scratch server showed, this protein has two domain of α-helix that domain 1 stretch at amino acid 1-193 and second domain from amino acid 194- 284. Also, a TPM sequence has less than two cysteine and therefore cannot form disulfide bonds. Prediction secondary structure of protein revealed that this protein has two conformations helix and coil (β-Turn) in its structure. For right and confident interpretation, results of both PSIPRED V3.0 and Scratch were evaluated and compared with each other in table 4.

Table 4- Frequency results of secondary structure prediction by using of PSIPRED V3.0 and Scratch servers. Results of PSIPRED and Scratch servers show five and seven Coil regions in TPM protein, respectively.

Targeted Sequence	[1]MEAIKKKMQAMKLEKDNAVDRAETAEQQSREAALRAEKAEEEVRSLQKKIQQIENELDQV QEQLSQANSKLEEKDKALQAAEAEVAAHNRRIQLLEEDLERSEERLKIATQKLEEASQAA DESERMRKMLEHRSITDEERMDGLEGQLKEARTMAEDADRKYDEVARKLAMVEADLERAEERAETG ETKIVELEEELRVVGNNLKSLEVSEEKALQKEETYEMQIRQMTNRLQEAEARAE FAERSVQKLQKEVDRLEDELVQEKEKYKAISDELDQTFSELTGY[284]
Degree of Confidence*	92789999980003552207488889988999999998589997786667685512887899999985549999999999998787767 78899998642258889999999999976501578756785236788377764366618999755332156338876678884312233 4555543345667478889877543354565888852257778999999988887788777877777766767767777766667775 57899898889863089
Predicted secondary structure [a] **	CHHHHHHHHHHHHHHHHHCCHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHH HHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHH HHHHHHHCCCCCCHHHHHHHHHHHHHHHHHHHHHHCCCHHHHHHHHHHHHHHHHHHHHHHHHH HHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHH HHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHCCC
Predicted secondary structure [b]**	CHHHHHHHHHHHHHHCCHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHH HHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHH HHHHHHHHCCCCCHHHHCCCHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHCCC CEEEEHHHHHHHHCCCCCCCCCCHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHH HHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHCCC

[a] Results of PSIPRED server, [b] Results of Scratch server,* Confidence (0=low, 9=high), ** Predicted of secondary structure (H=helix, C=coil)

Table 5. Full information on IgE epitopes of TPM protein.

Number	IgE epitope	Sequence matched	position
1	AQLLAEEADRKYD	ARTMAEDADRKYD	151
2	EKYKSITDELDQTFS	EKYKAISDELDQTFS	265
3	ELVNEKEKYKSITDE	ELVQEKEKYKAISDE	259
4	ESKIVELEEELRVVG	ETKIVELEEELRVVG	187
5	MQQLENDLDQVQESLLK	IQQIENELDQVQEQLSQ	50
6	QKLQKEVDRLEDELV	QKLQKEVDRLEDELV	247
7	RIQLLEEDLERSEER	RIQLLEEDLERSEER	91
8	RSLSDEERMDALENQ	RSITDEERMDGLEGQ	133
9	VAALNRRIQLLEEDL	VAAHNRRIQLLEEDL	85
10	VDRLEDELVNEKEKY	VDRLEDELVQEKEKY	253

Allergenicity analysis:

Chose prediction approaches were mapping of IgE epitopes and PID, SVM module based on amino acid composition, SVM module based on dipeptide composition, BLAST search on allergen representative peptides (ARPs) and hybrid approach (SVMc+IgE epitope+ARPs BLAST+MAST), respectively.

Mapping of TPM IgE epitopes

TPM epitopes and their positions have been shown in Table 5.

Prediction by SVM method based on amino acid composition

Prediction by SVM method based on amino acid composition shown that this protein is potentially allergen and it its scores, positive predictive value, negative predictive value were 1.0345043 (Threshold= -0.4), 85.64% and 67.96%, respectively.

Prediction based on SVM method based on dipeptide composition

Also, prediction by SVM method based on amino acid composition shown that this protein is potentially allergen and its scores, positive predictive value, negative predictive value were 1.0003902 (Threshold= -0.2), 100% and 59.74%, respectively.

Prediction by Hybrid Approach

Finally, Prediction hybrid approach that include; SVMc, IgE epitope, ARPs methods revealed this protein undoubted an allergen.

DISCUSSION

Ticks have numerous compounds with diverse biological function[17-21], but certain tick proteins such as TPM of *Boophilus* genus, have immunogenicity and allergenicity nature and study them from different aspects of *in-vivo* (experimental infection), *in-vitro* (identification and characterization) and *in-silico* (similarity, phylogeny, homology, structure, allergenicity) could help to understanding diverse functions, preparing suitable anti-tick vaccines, and insight to cure allergenic diseases caused by parasites[22-27].

In this study, after experimental infection of cattle by *B. analatus* and identification and characterization of TPM protein in *B. analatus* larva proteins extraction by Mass-spectrometry, complete similarity (100%) of this protein with *B. microplus* by NCBI blast and Mascot software obtained [9]. So, *B. microplus* TPM sequence used as representative of *Boophilus* genus sequence.

Results of sequence analysis, Blasting, alignment, evaluation of conserved/variable regions of sequences, phylogenetic tree construction and homology analysis revealed that at first, there is high conservation between species in this protein and it is accordance with Previous study identified TPM as a conserved and cross-reactive allergen between mites and other invertebrates[28]. In addition, the TPM is present in ticks, mites and insects, so, it may serve as a wide range vaccine candidate antigen in several species of parasites along with myosin and paramyosin [29-30].

Blast search, phylogenetic and conservation/variablity analysis showed that if vaccine designs based on *Boophilus* genus (*B. analatus or microplus*) TPM it also, will coverage the *Haemaphysalis* genus, *Ixodes* genus and also cross-react with some other mites such as *Dermanyssus gallinae*, *Psoroptes ovis* and/or *S. scabiei* , although they might have another hosts[31-33]. Besides, understanding sequence conservation is important for the study of sequence evolution and for the identification of functional regions of the protein [34]. Also, according to literatures, a numerous proteins interspecies or between some species in parasite, mammals, bacteria and etc. identified that have high conservation like TPM. [29, 35, 36]. Our phylogenetic analysis and comparison of the TPM amino acid sequences among parasites are in agree with evaluation of this protein in *S. scabiei* and some parts analysis of barnacle TPM [33, 37].

For a variety of reasons, many of the proteins identified as TPM for different species and registered in NCBI and UniproKB were not included in our data set of TPM. In several cases, no complete sequence (very short sequences) and Areas with ambiguous alignment or containing poly-N stretches were excluded from the datasets.

Also, the results of phylogenetic TPM analysis in this study is in agree with current taxonomy of arthropod phylogeny [38-40].

Nearly all proteins have structural similarities with other proteins and, in some of these cases, share a common evolutionary origin. Knowledge of these relationships is crucial to our understanding of the evolution and development of proteins [8].

Structural analysis of TPM protein in *Boophilus* genus revealed that it has two domain of α-helix, no disulfide bonds and two conformations helix and coil (β-Turn) in its structure. Other study invertebrate reported that a microfilament protein with a α-helical Coiled-coil structure is found in all their cell types [41].

There is some proteins like TPM, that have alpha and beta chains and orthology and paralogy concept defines for them, such as albumin and hemoglobin [42-43]. Estimation of average evolutionary divergence over parasite sequence pairs is just window to understanding of evolutionary relationships and not fully representative of all the patterns of evolution, and calculation of this parameter in study showed conservation of TPM among parasite sequence pairs and vertebrate [11].

In past, number of approaches and methods has been developed to predict allergens. In AlgPred a systematic attempt has been made to integrate various approaches in order to predict allergenic proteins with high accuracy [15].

Results of allergenicity evaluation is similar to other reports on allergenicity evaluation of TPM in other parasites and non-vertebrates, and it

cause of its conservation that lead to cross-reactivity with other parasite such as house dust mite [26,28,33].

In modern biology, scientists in different field of biology agree that a comparison of relationships of species, their genes or proteins sequences in a phylogenetic context and evaluation of similarities, sequence characteristics and structures is important part of research, and provide deep insights into organism nature. In general, the output tree of a phylogenetic analysis is an estimate of the character′s phylogeny (i.e. a protein tree) and not the phylogeny of the taxa (i.e. species tree) from which these characters were sampled, though ideally, both should be very close they do not necessarily accurately represent the species evolutionary history.

The comprehensive analysis of similarity, phylogenetic relationships, homology and allergenicity has never been easy, especially when we attempt to make statements from different aspects about a protein. Our study revealed the some unique and valuable aspects of TPM protein of *Boophilus* genus, but for the lack of TPM sequences data in many species the definition of a robust phylogeny and similarity remained unreached.

REFERENCES

1.De la Fuente J, and Estrada-Peña A. Ticks and tick-borne pathogens on the rise. Ticks and tick-borne dis 2012; 3(3): 115-6.

2.Miller R, Estrada-Pena A, Almazan C, Yeater KM, Messenger M, Ellis D, et al. Exploring the use of an anti-tick vaccine as a tool for the integrated eradication of the cattle fever tick, Rhipicephalus (Boophilus) annulatus. Vac 2012; 30(38): 5682-5682.

3.Aki T, Kodama T, Fujikawa A, Miura K, Shigeta S, Wada T, et al. Immunochemical characterization of recombinant and native tropomyosins as a new allergen from the house dust mite, Dermatophagoides farina. J. Allergy Clin Immunol 1995; 96: 74–83.

4.Jenkins RE, Taylor MJ, Gilvary NJ, Bianco, AE. Tropomyosin implicated in host protective responses to microfilariae in onchocerciasis. Proc Natl Acad Sci 1998; 95: 7550–7555.

5.Ranjbar MM, Nabian S, Taheri M, Yousefi P, Ahmadzadeh M. IgE Allergenic Epitope Prediction of Boophilus Microplus Tropomysin. Iranian j of Immunol 2012; 10: 110.

6.Yang Z, and Bielawski JP. Statistical methods for detecting molecular adaptation. Trends Ecol Evol 2000; 15: 496–502.

7.Holder M, Paul O, Lewi P. Phylogeny estimation: Traditional and Bayesian approaches. Nature Gen rev 2003; 4: 275-284.

8.Murzin AG, Brenner SE, Hubbard T, Chothia C. SCOP: a structural classification of proteins database for the investigation of sequences and structures. J Mol Bio 1995; 247: 536-540.

9.Nabian S, Taheri M, Nikbakht GR, Mazaheri nejad fard R. Immunoproteomics survey of fetal extract of boophlius anolatus tick. J V Lab Res 2012; 4(1): 58.

10.Thompson J.D., Higgins DG, Gilbson TJ. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nuleic Acids Res 1994; 22: 46730–4680.

11.Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S. MEGA5: Molecular Evolutionary Genetics Analysis Using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony Methods. Mol Biol Evol 2011; 28(10): 2731–2739.

12.Kimura M. A simple method for estimating evolutionary rate of base substitutions through comparative studies of nucleotide sequences. J Mol Evol 1980; 16: 111–120.

13.Felsenstein J. Confidence limits on phylogenies: an approach using the bootstrap. Evolution 1985; 39: 783-791.

14.Zuckerkandl E, and Pauling L. Evolutionary divergence and convergence in proteins. Edited in Evolving Genes and Proteins by V. Bryson and H.J. Vogel. Academic Press 1965; 97-166.

15.Saha S, and Raghava GPS. AlgPred: prediction of allergenic proteins and mapping of IgE epitopes. Nuc. Acids Res 2006; 34: 202-209.

16.Di Tommaso P, Moretti S, Xenarios I, Orobitg M, Montanyola A, Chang JM, et al. Coffee: a web server for the multiple sequence alignment of protein and RNA sequences using structural information and homology extension. Nucleic Acids Res 2011; 39: 7-13.
17.Kondrashov FA, Rogozin IB, Wolf YI, Koonin EV. Selection in the evolution of gene duplications. Genome Bio 2002; 3: 1-0008.9.
18.Scotland RW. Deep homology: A view from systematics. BioEssays : news and reviews in mol., cellul and develop Bio 2010; 32(5): 438–449.
19.Valenzuela JG. Exploring tick saliva: from biochemistry to "sialomes' and functional genomics. Parasitol 2004; 129: 83–94.
20.Andreotti R, Pedroso MS, Caetano AR, Martins NF. Comparison of predicted binders in Rhipicephalus (Boophilus) microplus intestine protein variants Bm86 Campo Grande strain, Bm86 and Bm95. Rev Bras Parasitol Vet. 2008; 17(2): 93-98.
21.Krasky A, Rohwer A, Schroeder J, Selzer PM. A combined bioinformatics and chemoinformatics approach for the development of new antiparasitic drugs. Genomics 2007; 89(1): 36-43.
22.Xu Y, Bruno JF, Luft BJ. Identification of novel tick salivary gland proteins for vaccine development. Biochem Biophys Res Comm 2005; 326:901–904.
23.Das S, Banerjee G, DePonte K, Marcantonio N, Kantor FS, Fikrig E. Salp25D, an Ixodes scapularis antioxidant, is 1 of 14 immunodominant antigens in engorged tick salivary glands. J. Infect. Dis 2001; 184: 1056–1084.
24.Mulenga A, Sugimoto C, Onuma M. Issues in tick vaccine development: identification and characterization of potential candidate vaccine antigens. Microbes Infect 2000; 2: 1351–1361.
25.Schuijt TJ, Narasimhan S, Daffre S, DePonte K, Hoviius JW, et al. Identification and characterization of Ixodes scapularis antigens that elicit tick immunity using yeast surface display. PLoS ONE 2011; 6(1), e15926.
26.Ayuso R, Reese G, Leong-Kee S, Plante M, Lehrer SB. Molecular basis of arthropod cross-reactivity: IgE-binding cross-reactive epitopes of shrimp, house dust mite and cockroach tropomyosins. Int. Arch Allergy Immunol 2002; 129: 38–48.
27.Lehrer SB, Ayuso R, Reese G. Seafood allergy and allergens: a review. Mar. Biotechnol 2003; 5: 339–348.
28.Reese G, Ayuso R, Lehrer SB. Tropomyosin: An Invertebrate Pan-allergen. Int Arch Allergy Immunol 1999; 119(4): 247–258.
29.Chu KH, Wong SH, Leung PSC. Tropomyosin is the major mollusk allergen: reverse transcriptase polymerase chain reaction, expression and IgE reactivity. Mar Biotechnol 2000; 2(5): 499–509.
30.Nisbet A J, and Huntley JF. Progress and opportunities in the development of vaccines against mites, fleas and myiasis-causing flies of veterinary importance. Parasite Immunol 2006; 28(4):165 - 172.
31.Reese G, Jeoung BJ, Daul CB, Lehrer SB. Characterization of recombinant shrimp allergen Pen a 1 (tropomyosin). Int Arch Allergy Immunol 1997; 113: 240–242.
32.You MJ. Immunization effect of recombinant P27/30 protein expressed in Escherichia coli against the hard tick Haemaphysalis longicornis (Acari: Ixodidae) in rabbits. Korean J Parasitol 2004; 42(4): 195–200.
33.Zhang R, Jise Q, Zheng W, Ren Y, Nong X,
34.Wu X, et al. Characterization and evaluation of a Sarcoptesscabiei allergen as a candidate vaccine. Parasites and Vectors 2012; 5: 176.
35.Xiaohui C, Haiyan H, Xiaoman Li. A new measurement of sequence conservation. BMC Genomics 2009; 10: 623.
36.Bull JJ, Hillis DM, O'Steen S. Mammalian ZFY sequences exist in reptiles regardless of sex-determining mechanism. Science 1988; 28: 567-569.
37.Barondess JJ, Beckwith J.. Bor gene of phage lambda, involved in serum resistance, encodes a widely conserved outer membrane lipoprotein. J Bacteriol 1995; 177:1247–53
38.Suma Y, Ishizaki Sh, Nagashima Y, Lu Y, Ushio H, Shiomi K.. Comparative analysis of barnacle tropomyosin: Divergence from decapod tropomyosins and role as a potential allergen.

Comparative Biochem. and Physio 2007; 147:230–236.
39.Wheeler WC, Cartwright P, Hayashi CY. Arthropod phylogeny: A combined approach. Cladistics 1993; 9: 1-39.
40.Giribet G, and Ribera C. A review of arthropod phylogeny: New data based on ribosomal DNA sequences and direct character optimization. Cladistics 2000; 16: 204-231.
41.Regier JC, and Shultz JW. Molecular phylogeny of arthropods and the signi®cance of the Cambrian``explosion" for molecular systematics. Am. Zool 1998; 38: 918-928.
42.Perry SV. Vertebrate tropomyosin: distribution, properties and function. J Muscle Res Cell Motil 2001; 22(1): 5–49.
43.Studer RA, and Robinson-Rechavi M. How confident can we be that orthologs are similar, but paralogs differ?. Trends in Genetics 2009; 25(5): 210–216.
44.Nehrt NL, Clark WT, Radivojac P, Hahn MW. Testing the Ortholog Conjecture with Comparative Functional Genomic Data from Mammals, PLOS Comput Bio2011; 7(6): e1002073.