Original Article

Immunoinformatics and Similarity Analysis of House Dust Mite Tropomyosin

Mohammad Mehdi Ranjbar[1], Sedigheh Nabian[2], Nayeb Ali Ahmadi[3*], Khodayar Ghorban[4], Alireza Sazmand[5], Maryam Dadmanesh[6], Seyed HosseinHekmatimoghaddam[7]

[1]Razi vaccine and Sera Institute, Karaj, Iran
[2] Department of Parasitology, Tick and Tick-Borne Diseases Center, Faculty of Veterinary Medicine, University of Tehran, Tehran, Iran
[3] Proteomics Research Center, Faculty of Paramedical Sciences, Shahid Beheshti University of Medical Sciences, Tehran, Iran
[4] Department of Immunology, AJA University of Medical Sciences, Tehran, Iran
[5] Department of Agriculture, Payame Noor University-Yazd Brancesh, Yazad, Iran
[6] Department of infectious diseases, AJA University of Medical Sciences, Tehran, Iran
[7] Department of Pathology, Shahid Sadoughi University of Medical Sciences, Yazd, Iran

Abstract

Background: *Dermatophagoides farinae* and *Dermatophagoides pteronyssinus* are house dust mites (HDM) that they cause severe asthma and allergic symptoms. Tropomyosin protein plays an important role in mentioned immune and allergic reactions to HDMs. Here, tropomyosin protein from *Dermatophagoides spp.* was comprehensively screened in silico for its allergenicity, antigenicity and similarity/conservation.

Materials and Methods: The amino acid sequences of *D. farinae* tropomyosin, *D. pteronyssinus* and other mites were retrieved. We included alignments and evaluated conserved/ variable regions along sequences, constructed their phylogenetic tree and estimated overall mean distances. Then, followed by with prediction of linear B-cell epitope based on different approaches, and besides *in-silico* evaluation of IgE epitopes allergenicity (by SVMc, IgE epitope, ARPs BLAST, MAST and hybrid method). Finally, comparative analysis of results by different approaches was made.

Results: Alignment results revealed near complete identity between *D. farina* and *D. pteronyssinus* members, and also there was close similarity among *Dermatophagoides* spp. Most of the variations among mites' tropomyosin were approximately located at amino acids 23 to 80, 108 to 120, 142 to 153 and 220 to 230. Topology of tree showed close relationships among mites in tropomyosin protein sequence, although their sequences in *D. farina*, *D. pteronyssinus* and *Psoroptes ovis* are more similar to each other and clustered. *Dermanyssus gallinae* (AC: Q2WBI0) has less relationship to other mites, being located in a separate branch. Hydrophilicity and flexibility plots revealed that many parts of this protein have potential to be hydrophilic and flexible. Surface accessibility represented 7 different epitopes. Beta-turns in this protein are with high probability in the middle part and its two terminals. Kolaskar and Tongaonkar method analysis represented 11 immunogenic epitopes between amino acids 7-16. From comparative analysis of predicted probable consensus epitope regions by machine learning approaches these epitopes were gained: AA_{23-48}, AA_{59-80}, AA_{91-110}, $AA_{114-143}$, $AA_{154-168}$, $AA_{182-200}$, $AA_{208-225}$, and $AA_{254-272}$. Prediction of allergenic proteins by AlgPred server showed 10 matches for IgE epitope, and prediction by hybrid approach showed that IgE epitope is undoubtedly the major allergen.

Conclusion: Immunoinformatic approaches in allergenic protein analysis are now reliable tools for explanation/interpretation of clinically observed complexities. Results of present study, would help in HDM immunotherapy against several species of parasites as a wide range epitopic desensitization or prevention (vaccine) regime.

Keywords: House dust mite, Tropomyosin, Allergenicity, Antigenicity, Similarity

***Corresponding Author:** Nayeb Ali Ahmadi, Proteomics Research Center, and Department of Medical Laboratory Technology, Faculty of Paramedical Sciences, Shahid Beheshti University of Medical Sciences, Tehran, Iran. Email: nayebalia@yahoo.com; nayebalia@sbmu.ac.ir

Please cite this article as: Ranjbar MM, Nabian S, Ahmadi NA, Ghorban Kh, Sazmand A, Dadmanesh M, Hekmatimoghaddam SH. Immunoinformatics and Similarity Analysis of House Dust Mite Tropomyosin. Novel Biomed. 2015;3(4):161-70.

Introduction

Allergy is a hypersensitivity reaction of the immune system to specific external substances called allergens, which are normally harmless substances in the environment. Environmental factors responsible for allergy include animal dander, foods, house dust mites (HDMs), pollen, insects, and chemical substances, to just name a few. Allergy to HDMs is popular and affects millions of people around the world. This allergy is not usually a direct life threatening danger, but could be a trigger for development of asthma, and may eventually lead to death due to respiratory complications[1-3].

HDMs belong to class Arachnida and include *Dermatophagoides farina* and *Dermatophagoides pteronyssinus*, which are defined as one of the major sources of aeroallergens that cause type I hypersensitivity disease (which sensitize and induce rhinitis, asthma, or atopic dermatitis) in a large portion of the world population[2,4]. The HDM extract used in allergy diagnosis and therapy is a complex mixture of allergens and non-allergen components from mite[5]. More than 30 different proteins from mite can bind human IgE[6].

Tropomyosin (TPM) protein (group 10 allergen from HDMs) has actin regulator activity and plays an important role in immune and allergic reactions. Despite its allergenicity; it is a candidate for vaccine production from some animal parasites[7-9]. This protein belongs to a family of highly conserved proteins with multiple isoforms found in both muscle and non-muscle cells of all species of vertebrates and invertebrates. Besides its role in the contractile activity of these cells, it also helps in regulation of cell morphology and motility. Its native structure consists of two parallel alpha-helical TPM molecules that are wound around each other forming a coiled-coil dimer[10,11].

B-cell epitope prediction plays an important role in vaccine design, allergy and immunodiagnostics research, and also in understanding immune system functions. Experimental methods for prediction of epitopes are costly and timely processes to find the antigen-antibody reaction sites. Computational methods accelerate reliable prediction of epitopes for different uses in biology. It is also a critical challenge in immunoinformatics and computational immunology. Such methods are cost effective and less time-consuming. B-cell epitopes can be linear or conformational by their structure. Linear epitopes consist of linear sequence of amino acids that can be recognized by either a particular antibody molecule or a particular B-cell receptor of the immune system. Parameters such as hydrophilicity, flexibility, accessibility, turns, exposed surface, and antigenic propensity of polypeptide chains are physico-chemical properties that have been correlated with the location of continuous epitopes[12]. Amino acid scale-based methods and machine learning methods are two main strategies that have various approaches and are used for prediction of linear B-cell epitopes[9]. Prediction and evaluation of allergenicity and antigenicity by hybrid approaches usually yields more accurate and reliable results[13], and we just used one such method.

Aims of this study were 1) Obtaining accurate details about allergenicity, antigenicity and epitope mapping of TPM in HDMs by various approaches; 2) Describe physico-chemical properties of this protein in deep insight; 3) Preparing a sound basis for integrated and uniform epitopic desensitization and prevention (vaccine) regimes, relying on cross-reactivity of TPM protein among mites; and 4) evaluation of similarity, conservation and evolutionary relationships in dust mites (not only HDMs) by phylogenetic tree.

Methods

Retrieving alignment and also conserved and variable regions of sequences

Complete protein sequences of *D. farinae* (AC: Q23939), *D. pteronyssinus* (AC: O18416) (two house dust mites), and other mites were retrieved from Uniprot KB database. The obtained sequences were aligned using Clustal X[14], analyzed and trimmed in CLC sequence viewer software version 6.6.2. Then, very short sequences and areas with ambiguous alignment or containing poly-N stretches were excluded from the analyses. The most highly conserved and variable regions were evaluated by CLC software, as depicted in figure 1.

Constructing the phylogenetic tree and overall mean distances

Selected, aligned and edited sequences directed according to phylogenetic tree were constructed by MEGA 5.3 software package[15], as seen in figure 2. We added homo sapiens tropomyosin alpha-1 chain sequence which was used as the out-group to mites tropomyosin dataset. All positions containing gaps and missing data were eliminated. Trees were constructed using the neighbor-joining (NJ) algorithm under the global gap removal option and Kimura's two-parameter substitution model[16]. Robustness of phylogenetic analysis was measured by bootstraps analysis with 10,000 replications. The percentage of replicate trees in which the associated taxa are clustered together in bootstraps test is shown next to the branches[17]. Then, the number of amino acid substitutions per site for overall means distances was calculated. Analyses were conducted using the Poisson correction model[18].

Prediction of linear B-cell epitope based on physico-chemical properties

These physico-chemical properties are shown in figure 3.

Antigenicity prediction based on hydrophobicity

Tropomyosin was scanned for hydrophobicity and Parker hydrophilicity index[19] with window size being seven. Hydrophobicity (or hydrophilicity) plots are designed to display the distribution of polar and non-polar residues along a protein sequence. Most commonly, this analysis has the goal of predicting membrane-spanning segments (highly hydrophobic) or regions that are likely exposed on the surface of proteins (hydrophilic domains) and therefore found useful for identifying potentially antigenic segments. IEDB analysis resource tools were applied for achieving this goal (http://tools.immuneepitope.org/tools/bcell/).

Antigenicity prediction based on assessment of solvent accessibility regions

The accessibility profile was developed using the formulae mentioned by Emini[20]. Each part of the sequences with surface probability greater than 1.0 indicates an increased probability for being found on the surface. These data and assessments may be useful for prediction of the peptides participating in antigenic activity, surface region peptides and useful domain(s) in the sequence.

Antigenicity prediction based on flexibility

We also concentrated on the flexibility data of protein to increase the prediction accuracy.

This was done with Karplus and Schulz flexibility prediction[21]. The calculation based on a flexibility scale is similar to classical calculations, except that the center is the fourth amino acid of the seven amino acid window.

Antigenicity prediction based on secondary structure (beta-turn prediction)

Secondary structures of HDM tropomyosin protein sequence were determined to predict the most probable regions and structures involved in antigenicity. The Chou and Fasman method was applied for prediction of the tropomyosin secondary structure[22]. Using these information parameters, the likelihood of a given residue assuming each of the four possible conformations alpha, beta, reverse turn, or coil was calculated, and the conformation with the most likelihood was assigned to the residue.

Antigenicity prediction based on Kolaskar and Tongaonkar method

Tropomyosin antigenic epitopes were determined using the method of Kolaskar and Tongaonkar[23], as is seen in figure 4. This prediction is based on a semi-empirical approach, depending on physico-chemical properties of amino acid residues (i.e. hydrophilicity, accessibility and flexibility). This approach has a good efficiency by detecting antigenic peptides with about 75% accuracy.

B-cell epitope prediction by machine learning approaches

Recently, several methods for B-cell epitope prediction using machine learning approaches have been published. We used four of them in our study as a hybrid method; they included hidden Markov model, feed forward and recurrent neural network, and subsequence Kernel based support vector machine (SVM), which were used in BepiPred[24], ABCPred[25], BCPred[26] and again ABCPred, respectively.

In-silico evaluation of allergenicity

World Health Organization (WHO) and Food and Agriculture Organization (FAO) proposed guidelines to assess the potential allergenicity (http://www.fao.org/es/ESN/food), and in this study we used AlgPred server[25] (http://www.imtech.res.in/raghava/algpred/). Algpred allows prediction of allergens based on similarity of known epitopes with any region of the target protein. Mapping of IgE epitope(s) feature of the server allows user to locate the position of epitope in their protein, and also allows predicting allergens based on SVM modules using amino acid or dipeptide composition. It facilitates BLAST search against 2890 allergen-representative peptides (ARPs) and assigning a protein allergen if it has a BLAST hit. Finally, the hybrid option of server allows predicting allergen using combined approach (SVMc+IgE epitope+ARPs BLAST+MAST).

Comparative analysis

Finally, we compared all of the analyses mentioned above, for interpretation of antigenicity and allergenicity regions of tropomyosin.

Results

Alignment

Alignment results revealed that there was near

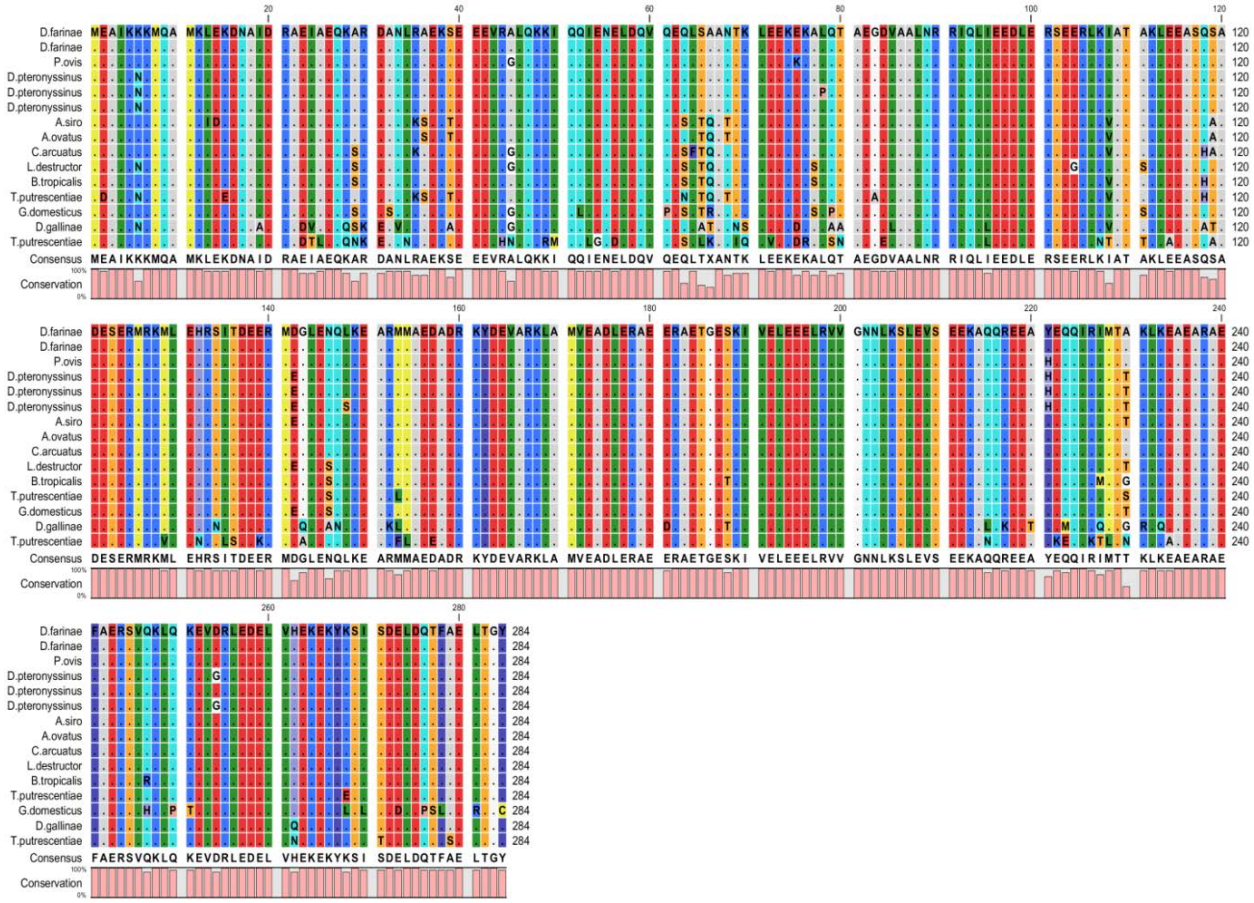

Figure 1. Comparison of HDM TPM against the TPM amino acid sequences of other mites by CLC sequence viewer 6.6.3. Gene bank accession numbers are not shown, and all sequences retrieved from databases have been used for alignment. Identical, conserved and semi-conserved regions along TPM sequence in parasites could be seen. All residues that are identical to the top sequence in an alignment are shown as a dot ('.'). Also, at button of the figure consensus and rate of conservation re-depicted.

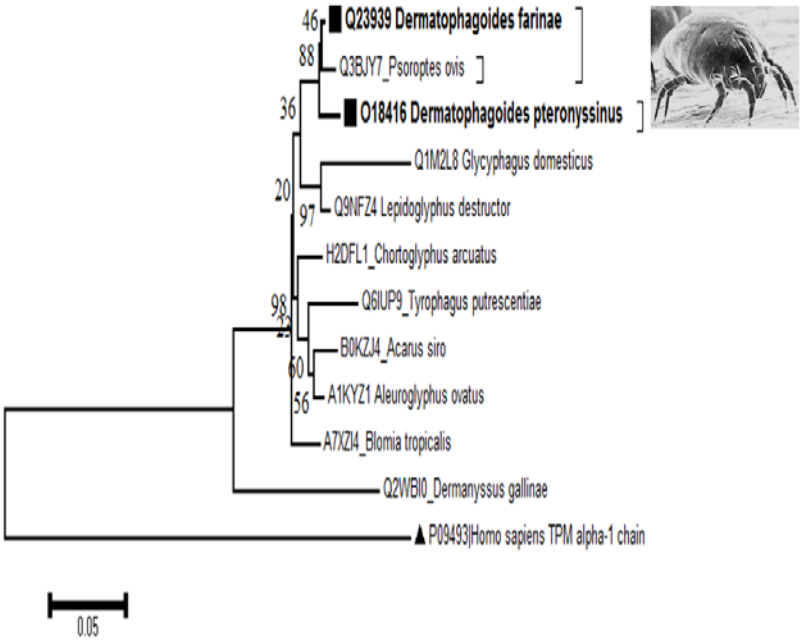

Figure 2. Phylogenetic tree constructed based on protein sequences with MEGA 5.3, illustrating the relationships in tropomyosin protein between *D. Farina* & *D. pteronyssinus* against other mites. Tropomyosin protein from *Dermatophagoides spp.* is marked by■ and also out-group (Homo sapiens tropomyosin alpha-1 chain) is shown by▲. Tree was constructed by the use of the neighbor-joining (NJ) algorithm based on differences in tropomyosin 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 number of the sequences and tick or mite name for each sequence.

complete identity between *D. farina* and *D. pteronyssinus* members, and also among *Dermatophagoides spp*. (Fig. 1). Most of the variations among mites tropomyosin were located at about amino acids 23 to 80, 108 to 120, 142 to 153 and 220 to 230.

Phylogenetic Tree

For the phylogenetic analysis, tropomyosin sequences of mites (12 sequences) were aligned, compared and edited using CLC sequence viewer 6.6.2. MEGA 5.3 software package was used for construction of phylogenetic tree and calculation of overall mean distances (Fig. 2). Topology of tree shows the close relationships among mites in tropomyosin sequence, although the sequences in *D. farina*, *D. pteronyssinus* and *Psoroptes ovis* are more similar to each other than to other mites and clustered. The nearest relatives to the mentioned cluster are *Glycyphagus domesticus* and *Lepidoglyphus destructor*. *Dermanyssus gallinae* (AC: Q2WBI0) has less relationship to other mites, being located in a separate branch. Our out-group, as is shown, was Homo sapiens tropomyosin alpha-1 chain.

Furthermore, estimates of average evolutionary divergence over sequence pairs in mites showed overall divergence over parasite sequence pairs (d) 0.087 with standard error (SE) estimate being 0.010. This reflects highly conservation of this protein among mites.

Pair wise alignments and similarity analysis showed that there are highly conserved areas and significant cross reactivity between *Dermatophagoides spp*. in tropomyosin. So, we used *D. farina* as representative of this genus for further analysis.

Antigenicity prediction based on different methods

Hydrophilic plot (Fig.3A) revealed that many parts of this protein have potential to be hydrophilic (locations above 3 are supposed to have hydrophilicity potential). Careful examination of the surface accessibility predicted peptides diagram using Emini algorithm (Fig. 3B) represents these epitopes: AA70-76 (KLEEKEK), AA98-104 (DLERSEE), AA119-126 (SADESERM), AA157-162 (DADRKY), AA177-184

Table 1: Predicted peptides by Kolaskar-Tongaonkar algorithm. Arrangement in table from ascending to descending is based on their position in sequence. Flexible length for peptides was allowed.

No.	Start position	End Position	Peptide	Peptide Length
1	43	51	VRALQKKIQ	9
2	57	68	LDQVQEQLSAAN	12
3	82	97	EGDVAALNRRIQLIEE	16
4	105	116	RLKIATAKLEEA	12
5	162	175	YDEVARKLAMVEAD	14
6	188	194	SKIVELE	7
7	196	203	ELRVVGNN	8
8	206	212	SLEVSEE	7
9	243	256	ERSVQKLQKEVDRL	14
10	258	264	DELVHEK	7
11	267	273	YKSISDE	7

(ERAEERAE), AA212-221 (EKAQQREEAY),

Table 2: Predicted B-cell epitopes by different servers that used machine learning approaches. Threshold or specificity was 85% in all servers (except Bepipred). Fixed length epitope prediction was 16(except Bepipred and FBCPred). Arrangement from ascending to descending is based on score (except Bepipred).

Server name	Epitopes	Positions	Features used	Machine Learning Technique
Bepipred	EIAEQKARDANLRAEKSEE EV, NELDQV, EQLSA, TKLEEKEKALQTAEGDVA, LEEASQSADESE, ITDEERMDG, AEDADRKYDE, LERAEERAETGE, VSEEKAQQREEAYE	23-43, 55-60, 62-66, 69-86, 113-125, 135-143, 155-164, 176-188, 209-222	Parker hydrophilicity scale and Levitt secondary structure	Hidden Markov model
ABCpred Server	HRSITDEERMDGLENQ, KEARMMAEDADRKYDE, SERMRKMLEHRSITDE, DELVHEKEKYKSISDE, IQLIEEDLERSEERLK, RAEIAEQKARDANLRA, SQSADESERMRKMLEH	132-148, 149-165, 123-139, 258-274, 92-108, 21-37, 117-133	Hydrophilicity, accessibility, flexibility, turns, antigenicity, polarity	Feed Forward and recurrent Neural network
BCPREDS Server 1.0	EVSEEKAQQREEAYEQ, TGESKIVELEEELRVV, QLIEEDLERSEERLKI, EKSEEEVRALQKKIQQ, EEASQSADESERMRKM, VDRLEDELVHEKEKYK, MEAIKKKMQAMKLEKD	208-224, 185-201, 93-109, 37-53, 114-130, 253-269, 1-17	Hydrophilicity, accessibility, flexibility, turns, antigenicity, Amino Acid Pair(AAP) antigenicity scale	Subsequence kernel based SVM
AAP (Cheng et al., 2007) (in BCPREDS Server 1.0)	EEKAQQREEAYEQQIR, QVQEQLSAANTKLEEK, EDADRKYDEVARKLAM, IEEDLERSEERLKIAT	211-227, 59-75, 156-172, 95-111	Hydrophilicity, accessibility, flexibility, turns, antigenicity, Amino Acid Pair (AAP) antigenicity scale	Support Vector Machine (SVM)

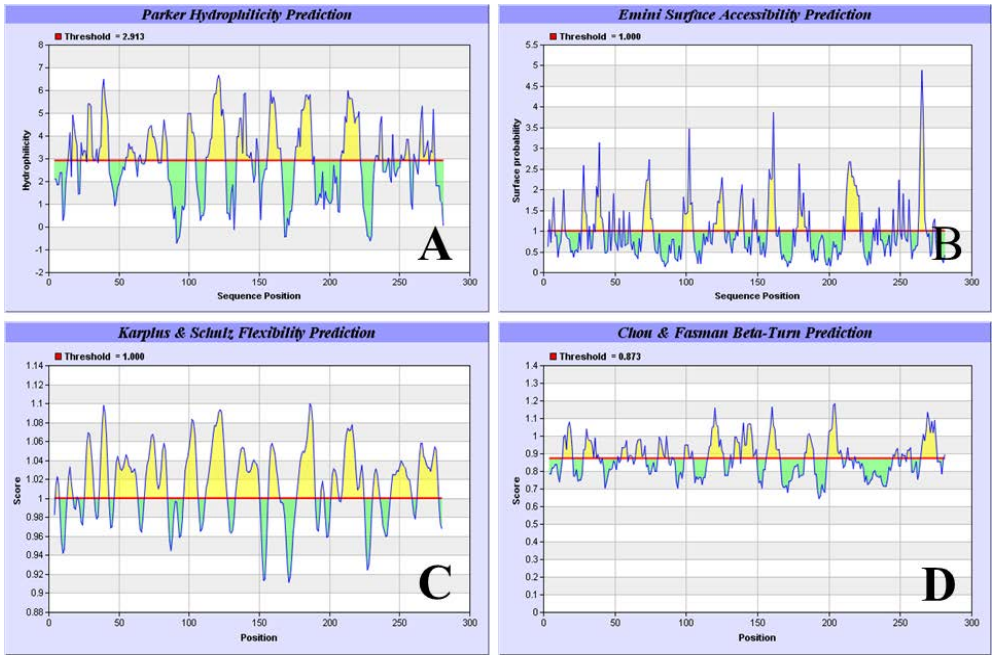

Figure 3. Physico-chemical plots. Hydrophilic plot (A), surface accessibility plot (B), flexibility plot (C) and beta-turn plot (D). A) Hydrophilicity profile diagram using Parker hydrophilicity prediction algorithms for *D. farina*. Size of window was 7 and tools calculated average (threshold), maximum and minimum hydrophilicity were 2.913, 6.671 and -0.714, respectively. In diagram A, green regions represent hydrophobic areas and are not likely antigenic regions, in which about 9 regions are credible. Diagram peak regions represent hydrophilicity and are antigenic regions, and about 16 regions are credible. Usually hydrophobic and hydrophilic regions are near each other. B) Emini surface accessibility prediction plot for HDM TPM (IEDB server). Average (threshold), maximum and minimum hydrophilicity were 1.000, 4.901 and 0.135, respectively. Size of window was 6, and center position 4. Also, details of predicted peptides by Emini algorithm are arranged in informative table.C) Flexibility prediction plot for HDM TPM (IEDB server). Threshold, maximum and minimum flexibility were 1.000, 1.100 and 0.912, respectively. A window of seven residues was used for analyzing epitope region. The corresponding value of the scale was introduced for each of the seven residues and the arithmetical mean of the seven residue value was assigned to the fourth (i+3) residue in the segment. D) Frequency plot of secondary structure prediction of α helixand coil (β turn) structures in TPM of HDM using Chou & Fasman algorithms. Average (threshold), maximum and minimum for beta-turn were 0.873, 1.186 and 0.646, respectively. Size of window was 7, and center position 4.

AA263-268 (EKEKYK). Karplus and Shulz flexibility prediction plot (Fig. 3C) showed that this protein is highly flexible in different parts of its sequence. Finally, beta-turn prediction (Fig. 3D) reflects that turns in this protein are in the middle parts and its two terminals with high probability.

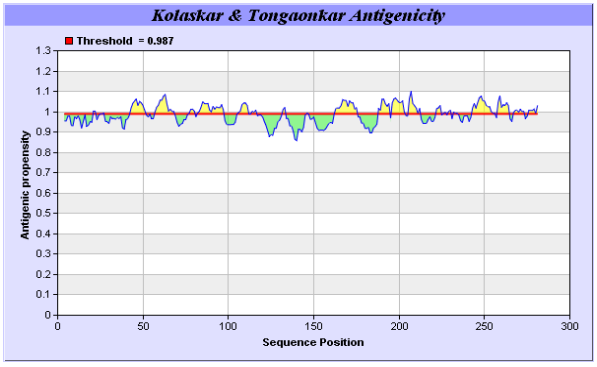

Figure 4. Antigenicity prediction plot of TPM protein using Kolaskar-Tongaonkar algorithm. Regions with antigenic propensity scale above 1 are antigenic regions. Average (threshold), maximum and minimum antigenicity are 0.987, 1.098 and 0.858, respectively. Window size and center position were 7 and 4, respectively

Considering the methods mentioned above, perhaps the simplest approach for prediction of antigenic determinants is Kolaskar and Tongaonkar[23], which is based on the occurrence of amino acid residues in experimentally determined epitopes. Antigenicity prediction plot of tropomyosin by use of the Kolaskar-Tongaonkar algorithm is represented in Figure 4. Also, details of predicted peptides by Kolaskar-Tongaonkar algorithm are arranged in informative Table 1. Kolaskar-Tongaonkar method evaluates hydrophilicity, accessibility and flexibility together to predict epitopes.

B-cell epitope prediction by machine learning approaches

Details of prediction by different servers are summarized in table 2.

Eventually, comparative analysis of probable

Table 3: The prediction of allergenicity of TPM protein by AlgPred server.

Name	length	Mapping of IgE Epitopes	MAST Results	Prediction by SVM method based on amino acid composition* (Threshold= -0.4)	Prediction based on SVM method based on dipeptide composition** (Threshold= -0.2)	Blast RESULT	Prediction by Hybrid Approach (By IgE method, ARPs, BLAST method)
D. farinae TPM	284	+	NON ALLERGEN (No Hits found)	[A]Potential ALLERGEN	[A]Potential ALLERGEN	Hits found with ARPs database: GESKIVELEEELR VVGNNLKSLEV	ALLERGEN

*[A]Score= 0.95877338, Positive Predictive Value= 85.64% Negative Predictive Value= 67.96%[B] Score= 1.0025221, Positive Predictive Value= 85.64% Negative Predictive Value= 67.96%

**[A]Score= 1.2319049 [Threshold= -0.2], Positive Predictive Value=100% Negative Predictive Value=59.74%[B] Score= 1.1705771, Positive Predictive Value=100% Negative Predictive Value=59.74%

consensus epitope regions by machine learning approaches, these epitopes were obtained: AA23-48, AA59-80, AA91-110, AA114-143, AA154-168, AA182-200, AA208-225, and AA254-272.

Allergenicity analysis

Chosen 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. Additionally, prediction by hybrid approach showed that this protein is undoubtedly allergen. Analysis results are shown in table 3 and 4.

Discussion

Various studies have reported that exposure to HDM allergens may be a primary cause or a risk factor for development of asthma[27,28]. About 50% of allergic patients and up to 80% of asthmatic children are sensitized to mite allergens[29]. Mite allergic airway diseases are conventionally defined as a type-I allergy, mediated by Th2 cells and IgE, that is characterized by chronic airway inflammation.

Immunoinformatic approaches in allergenic protein analysis are now reliable tools which help to provide explanation/interpretation for clinically observed diseases in terms of appearances, cross-reactivity, and co-variation of sensitization[30,31].

A number of methods have been developed to predict allergens, each having its own merits and demerits. In case of AlgPred, a systematic attempt has been made to integrate various approaches in order to predict allergenic proteins with high accuracy[25]. Results of allergenicity prediction by AlgPred are similar to other reports on allergenicity evaluation of tropomyosin in parasites and non-vertebrates[32-35].

Comparative analysis of predicted IgE epitopes with

Table 4: Full information on IgE epitopes of HDM TPM protein.

Protein name	IgE epitope	Sequence matched	Positions
***D .farinae* TPM**	AQLLAEEADRKYD, EKYKSITDELDQTFS, ELVNEKEKYKSITDE, ESKIVELEEELRVVG, MQQLENDLDQVQESLLK, QKLQKEVDRLEDELV, RIQLLEEDLERSEER, RSLSDEERMDALENQ, VAALNRRIQLLEEDL, VDRLEDELVNEKEKY	ARMMAEDADRKYD, EKYKSISDELDQTFA, ELVHEKEKYKSISDE, ESKIVELEEELRVVG, IQQIENELDQVQEQLSA, QKLQKEVDRLEDELV, RIQLIEEDLERSEER, RSITDEERMDGLENQ, VAALNRRIQLIEEDL, VDRLEDELVHEKEKY	151-164, 265-280, 259-274, 187-202, 50-67, 247-262, 91-106, 133-148, 85-100, 253-268

predicted consensus probable epitope regions by different machine learning approaches showed that all IgE epitopes are in the range of predicted consensus probable epitope regions by machine learning approaches, except AA23-48 and AA208-225. These two epitopes could be just immunogenic (not allergenic) or maybe allergen epitopes that are not yet known. Another comparative analysis of predicted IgE epitopes with predicted epitopes by Kolaskar and Tongaonkar method reveal that except allergenic epitopes of AA151-164 and AA133-148, other epitopes were predicted by Kolaskar and Tongaonkar. Also, epitope AA43-51 was only seen in Kalkaska and Tongaonkar method results. Sharique *et al.*, 2012 also evaluated tropomyosin similarity and allergenicity. A similar analysis was performed using reported B-cell IgE-binding epitopes from Met e1 (shrimp allergen) and Bla g7 (cockroach allergen) with other invertebrate tropomyosin. The percent identity in linear sequences was higher than 35% in mites, crustaceans, and cockroaches. The polar and hydrophobic regions in these groups were highly conserved. Also, results reflect that tropomyosin has different segments that are responsible for allergenicity and immunogenicity. Linear prediction servers (exactly machine learning approaches) were successful in prediction, and most of the epitopes predicted as B-cell epitope are allergenic.

Sequence analysis based on alignment, evaluation of conserved/variable regions and phylogenetic tree construction revealed that there is high conservation between mite species in this protein. It is in accordance with previous studies that identified tropomyosin as a conserved and cross-reactive allergen (sharing similar IgE binding epitopes) between mites and other invertebrates. Therefore, it may be a cause of clinically reported cross-reactivity[32,36,37].

In addition, tropomyosin is present in mites, ticks and insects. So, results of allergenicity and antigenicity analysis would help in HDM immunotherapy as a wide range epitopic desensitization or prevention regime against several species of parasites, although they might have other hosts[34,38,39]. Besides, a deep insight to sequence conservation/variability, similarity and phylogenetic relationships is crucial for study of sequence evolution and for identification of functional regions of the protein[40].

Our Study revealed some unique and valuable immunoinformatic aspects of tropomyosin in Dermatophagoides genus, but due to lack of the protein sequence data in mites, definition of a robust phylogeny and similarity remained unreached. Here, various approaches with different features for the protein linear B-cell epitope prediction were studied. By combining various features used in all of these approaches and doing a principal component analysis, several features which do not play a major role in epitope prediction were filtered out. There is a need for more studies about allergens of HDM and its role in allergies and asthma. Immunoinformatic approaches help toward deep understanding of the allergic response to HDMs. Comprehensive and comparative assessment of allergenicity, antigenicity, epitope mapping and phylogenetic relationships has never been easy, especially when we attempt to make statements from different aspects about a protein.

Acknowledgment

We appreciate the Proteomics Research Center head and all their colleagues who cooperated in this work.

References

1. Flaum M, Lung CL, Tinkelman D. Take control of high-cost asthma. J Asthma. 1997;34(1):5-14.

2. Platts-Mills TA, Vervloet D, Thomas WR, Aalberse RC& Chapman MD. Indoor allergens and asthma: report of the Third International Workshop. J Allergy ClinImmunol. 1997;100(6 Pt 1):S2–24.

3. Downs SH, Marks GB, Sporik R, Belosouva EG, Car NG, Peat JK. Continued increase in the prevalence of asthma and atopy. Arch Dis Child. 2001;84(1):20-3.

4. Weghofer M, Thomas WR, Pittner G, Horak F, Valenta R,Vrtala S. Comparison of purified Dermatophagoides pteronyssinus allergens and extract by two-dimensional immunoblotting and quantitative immunoglobulin E inhibitions. Clin Exp Allergy. 2005;35(10): 1384-91.

5. Batard T, Hrabina A, Bi XZ, et al. Production and proteomic characterization of pharmaceutical-grade Dermatophagoides pteronyssinus and Dermatophagoides farinae extracts for allergy vaccines.Int Arch Allergy Immunol. 2006;140(4):295-305.

6. Weghofer M, Thomas WR, Kronqvist M, et al. Variability of IgE reactivity profiles among European mite allergic patients. Eur J Clin

Invest. 2008;38(12):959-65.
7. Yi FC, Cheong N, Shek PC, Wang DY, Chua KY, Lee BW. Identification of shared and unique immunoglobulin E epitopes of the highly conserved tropomyosins in Blomiatropicalis and Dermatophagoide spteronyssinus. Clin Exp Allergy. 2002;32(8):1203-10.
8. Nisbet AJ, Huntley JF. Progress and opportunities in the development of vaccines against mites, fleas and myiasis-causing flies of veterinary importance. Parasit Immunol. 2006;28(4):165-72.
9. Kavitha KV, Saritha R, Vinod Chandra SS. Computational methods in linear B-cell epitope prediction. Int J Comp App. 2013;63(12):28-32.
10. Aki T, Kodama T, Fujikawa A, et al. Immunochemical characterization of recombinant and native tropomyosins as a new allergen from the house dust mite, Dermatophagoides farina. J Allergy ClinImmunol. 1995;96 (1):74–83.
11. Jenkins RE, Taylor MJ, Gilvary NJ, Bianco AE. Tropomyosin implicated in host protective responses to microfilariae in onchocerciasis. Proc Natl Acad Sci USA. 1998;95(13):7550–5.
12. Greenbaum JA, Andersen PH, Blythe M. Towards a consensus on datasets and evaluation metric for developing B-cell epitope prediction tools. J Mol Recognit. 2007;20(2):75–82.
13. Yang X& Yu X. An introduction to epitope prediction methods and software. Rev Med Virol 2009; 19(2):77–96. [Reviews in Medical Virology]
14. Thompson JD, 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 Res1994; 22(22): 4673–4680.
15. 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):2731-9.
16. Kimura M. A simple method for estimating evolutionary rate of base substitutions through comparative studies of nucleotide sequences. J Mol Evol. 1980;16(2):111–20.
17. Felsenstein J. Confidence limits on phylogeneies: an approach using the bootstrap. Evolution. 198539(4):783-91.
18. Zuckerkandl E, Pauling L. Evolutionary divergence and convergence in proteins. In: Bryson V, Vogel HJ. Evolving Genes and Proteins. New York: Academic Press. 1965;97-166.
19. Parker JM, Guo D, Hodges RS. New hydrophilicity scale derived from high-performance liquid chromatography peptide retention data: correlation of predicted surface residues with antigenicity and X-ray-derived accessible sites. Biochemistry. 1986;25(19):5425–32.
20. Emini EA, Hughes JV, Perlow DS, Boger J. Induction of hepatitis A virus neutralizing antibody by a virus-specific synthetic peptide. J Virol. 1985;53(3):836-9.
21. Karplus PA, Schulz GE. Prediction of chain flexibility in proteins: a tool for the selection of peptide antigen. Naturwissenschaften.1985;72(4): 212-3.
22. Chou PY, Fasman GD. Prediction of beta-turns. Biophys J. 1979;26(3):367-84.
23. Kolaskar AS, Tongaonkar PC. A semi-empirical method for prediction of antigenic determinants on protein antigens. FEBS Lett. 1990;276(1-2):172-4.
24. Larson JEP, Lund O, Neilsen M. Improved method for predicting linear B-cell epitopes. Immunome Res. 2006;2:2.
25. Saha S, Raghava GP. Prediction of continuous B-cell epitopes in an antigen using recurrent neural network. Proteins. 2006;65(1):40-48.
26. EL-Manzalawy Y, Dobbs D, Honavar V. Predicting linear B-cell epitopes using string kernels. J. Mol. Recognit. 2008;21(4):243-55.
27. Platts-Mills TAE, De Weck AL. Dust mite allergens and asthma – a worldwide problem. J Allergy ClinImmunol. 1989;83:416-27.
28. Squillace SP, Sporik RB, Rakes G, et.al. Sensitization to dust mites as a dominant risk factor for asthma among adolescents living in central Virginia.Multiple regression analysis of a population-based study. Am J Respir Crit Care Med. 1997;156(6):1760-4.
29. Boulet LP, Turcotte H, Laprise C, et al. Comparative degree and type of sensitization to common indoor and outdoor allergens in subjects with allergic rhinitis and/or asthma. Clin Exp Allergy. 1997;27(1):52-59.
30. Breiteneder H, Mills C. Structural bioinformatic approaches to understand cross-reactivity. Mol Nutr Food Res. 2006;50(7):628–32.
31. Schein CH, Ivanciuc O, Braun W. Bioinformatics approaches to classifying allergens and predicting cross-reactivity. Immunol Allergy Clin North Am. 2007;27(1):1–27.
32. Reese G, Ayuso R, Lehrer SB. Tropomyosin: An invertebrate pan-allergen. Int Arch Allergy Immunol. 1999;119(14):247-58.
33. 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(1):38–48.
34. Zhang R, Jise Q, Zheng W, et al. Characterization and evaluation of a Sarcoptesscabiei allergen as a candidate vaccine. Parasit Vectors. 2012;5:176.
35. Galán-Freyle N, Olivero-Verbel J, Díaz-López L. Modeling of allergen proteins found in sea food products. Ciênc Tecnol Aliment. 2012;32(2):393-400.
36. Acevedo N, Caraballo L. IgE cross-reactivity between Ascaris lumbricoides and mite allergens: possible influences on allergic sensitization and asthma. Parasit Immunol. 2011;33(6):309-21.
37. Shafique RH, Inam M, et al. Group 10 allergens (tropomyosins) from house-dust mites may cause co-variation of sensitization to allergens from other invertebrates. Allergy Rhinol (Providence). 2012;3(2):74-90.
38. Pajno GB, La Grutta S, Barberio G, Canonica GW, Passalacqua G. Harmful effect of immunotherapy in children with combined snail and mite allergy. J Allergy ClinImmunol. 2002;109(4):627-9.
39. Azofra J, Lombardero M. Limpet anaphylaxis: cross-reactivity between limpet and house-dust mite Dermatophagoides pteronyssinus. Allergy. 2003;58(2):146-9.
40. Xiaohui C, Haiyan H, Xiaoman L. A new measurement of sequence conservation. BMC Genomics. 2009;10:623.