**In silico** analysis of cross reactivity among phospholipases from Hymenoptera species [version 1; peer review: awaiting peer review]

Yuliana Emiliani1, Andrés Sánchez1,2, Marlon Munera1, Jorge Sánchez1,2, Dilia Aparicio1

1Health Faculty - GINUMED, Corporation University Rafael Nuñez, Cartagena, Colombia
2Group of Clinical and Experimental Allergy, University of Antioquia, Medellin, Colombia

**Abstract**

**Background:** Phospholipases are enzymes with the capacity to hydrolyze membrane lipids and have been characterized in several allergenic sources, such as hymenoptera species. However, cross-reactivity among phospholipases allergens are little understood. The objective of this study was to determine potential antigenic regions involved in cross-reactivity among allergens of phospholipases using an in silico approach.

**Methods:** In total, 18 amino acids sequences belonging to phospholipase family derived from species of the order hymenoptera were retrieved from the UniProt database to perform phylogenetic analysis to determine the closest molecular relationship. Multialignment was done to identify conserved regions and matched with antigenic regions predicted by ElliPro server. 3D models were obtained from modeling by homology and were used to locate cross-reactive antigenic regions.

**Results:** Phylogenetic analysis showed that the 18 phospholipases split into four monophyletic clades (named here as A, B, C and D). Phospholipases from A clade shared an amino acid sequences’ identity of 79%. Antigenic patches predicted by ElliPro were located in highly conserved regions, suggesting that they could be involved in cross-reactivity in this group (Ves v 1, Ves a 1 and Ves m 1).

**Conclusions:** At this point, we advanced to the characterization of potential antigenic sites involved in cross-reactivity among phospholipases. Inhibition assays are needed to confirm our finding.

**Keywords**

phospholipase, Hymenoptera, allergen, in silico, epitope, cross-reactivity.
Introduction

Allergic diseases have become a public health problem; the genetic background of patients (atopy) and the environmental conditions are considered the cause of the increased risk to develop allergic diseases. Exposure to allergens (typically harmless antigens in the environment) also promotes an immune response mediated by IgE. Over the last few years, species belonging to the order Hymenoptera have been characterized as potential allergenic sources. They represent a common source of sensitization, with more than 200,000 species including bees, wasps, and ants. Most of the member of this order are cosmopolitan species, but some of them have an endemic distribution with a capacity of sensitization, like the Bombus sp. located more frequently in central and northern Europe, whereas the yellowjacket (YJ) (Vespa spp.) and honeybee (HB) (Apis mellifera) are allergenic sources in North America. Other wasps such as Polistinae are found in southern Europe and America.

Allergic immune response from hymenopteran allergens has been studied in detail due to a high incidence of sting reactions to these insects. Approximately, 9.2% to 28.7% of the adult population is sensitized to the venom of hymenoptera. Allergic response to Hymenoptera venom is one of the leading causes of anaphylaxis worldwide with a frequency of 27%, as compared to medications (41%) and foods (20%)\(^1\). Molecular, structural, and immunological characterization of hymenopteran venom allergens is advanced, in total 75 allergens from 31 different species have been explored, and since phospholipases are a family of allergens with clinical and biological relevance, some proteins belonging to this order such as hyaluronidase and antigen V are also considered relevant to sensitization to this allergenic\(^2\). Exposure to hymenoptera allergens is associated with bites and stings; it is considered that 56.6–94.5% of the general population have been bitten at least once in their life\(^9\).

Phospholipases (PLA) are a major component of the venom of these species, representing 75% of the total mass of the poison and has been characterized as one of the main allergens in Hymenoptera\(^3\). They can be found in venoms from other arthropods such as chelicerates, in the venom of ophidians, as well as in different tissues of mammals such as pancreatic juice, synovial arthritic fluid. The superfamily includes 42 groups distributed in four types: A, B, C, and D\(^4\).

Phospholipases belonging to class A split into two groups: class A1 hydrolyzes the phospholipid ester bond between the first acyl and glycerol (1 acyl-SN-glycerol phosphate), while class A2 hydrolyzes the bond between the second acyl and glycerol (2 Acyl-SN-glycerol phosphate). They are a family of enzymes with different molecular weights, PLA1 has a molecular weight of 28 kDa, while PLA2 are classified as high molecular weight cytotoxic PLA2 (40–85 kDa) and low molecular weight secretory PLA2 (14–18 kDa) with the capacity to hydrolyze fatty acids that are present on the cell membrane and other types of lipophilic substances or participate in the mechanism of regulation of gene expression through the production of free fatty acids, from which cyclooxygenases synthesize prostaglandins\(^12\–14\).

The structure, function, mechanisms, and cell signaling of PLA have been extensively studied; one important aspect of PLA is their capacity to induce allergic responses. Several epitopes involved in the co-sensitization of some PLA that share structural homology and identity have been studied; this suggests a potential role in cross-reactivity. However, this is little understood and studies are needed to complement what has been reported. The aim of this work was to explore cross-reactivity and antigenicity of allergenic PLA using an in silico approach, using bioinformatics tools, where we identified several antigenic regions that may be involved in cross-reactivity among phospholipases.

Today it is evident how the use of bioinformatics tools for science has grown; it is considered the first step to carry out experimental studies because they create a functional prediction. Understanding and predicting an individual clinical cross-reactivity to allergens is key to better management, treatment, and progression of new therapies for allergy to Hymenoptera; prediction can be performed by methods for the identification and computational mapping of specific IgE epitopes or epitopes reported in the Immune Epitope Database and Analysis Resource, which can help identify the areas that may be affecting patients’ health. Various studies have carried out on this methodology for predicting food allergen epitopes\(^16\–17\).

The in silico methodology has been used in other work to report possible cross-reactivity based on proteins in studies of structural or functional homology, through bioinformatics tools\(^18\).

Methodology

Selection of phospholipases and alignment

The amino acid sequences of phospholipases type A (A1 and A2) from 18 Hymenoptera species were selected according to the allergenic capacity reported. The sequences were obtained from the UniProt database (see Table 1 for a list of accession numbers). All Allergens that were reported in the WHO/IUIS Allergen Nomenclature Sub-Committee with a complete sequence were used. We did not include incomplete sequences for analysis. Three sequences are not reported as allergenic but were chosen to observe the differences in identity and the structures of several phospholipases. The identity degree among phospholipases was determined using the PRALINE web server. The parameters to perform the alignment were configured to use BLOSUM62 as the exchange matrix. The interactions used were 3 with an E value of 0.001.

Phylogenetic analysis

The Molecular Evolutionary Genetic Analysis (MEGA) program, version X was used to obtain phylogenetic trees, using the method of maximum parsimony of the taxa with the support of Bootstrap with 1000 repetitions as a measure of reliability and robustness under the assumption of a minimum
evolution. In the topology, this model uses a comparative matrix to find the similarity between the amino acids of 18 sequences to establish the evolutionary proximity between the species. The matrix was constructed with all the amino acid sequences of the phospholipases recovered from the UniProt database and reported to the WHO/IUIS. Therefore, the more positive identity values found between the sequences, the greater their relationship will be, and the closer they will be located in the tree. All empty spaces were eliminated (complete deletions). From the global comparison and the homologies, the sum of the length of the branches (SBL) will be presented, which will determine the number of nodes and their position, including the “groups” of the evolutionarily closest sequences. Phylogenetic sub-analyses were carried out in order to identify the degree of identity of the groups formed. The alignment for phylogenetic analysis was carried out using CLUSTAL W, which performs alignments. The parameters to perform the multiple alignment were configured to use gap opening penalty of 10.00 and gap extension penalty of 0.20, and the divergent cutoff delay was 30%.

**Results**

**Phospholipases found and phylogenetic results**
We selected 18 sequences of allergenic phospholipases and three not allergenic to include in the analysis with 361 positions in the final dataset. The sequences were derived from several biological sources: five from bees, six wasps, three ants, and three sources not described as an allergen, mosquito, spider, and scorpion. The allergens of bees and wasps belong to group 1 and the ants to groups 1 and 2 (Table 1).

The phylogenetic tree had a consistency index of 0.857256 with a retention index of 0.779682 and a composite index of

---

**Table 1. Phospholipases Allergens used to compare sequences.** The name of the allergen, source, and type of phospholipases and Uniprot code are detailed.

| Allergens | Allergen sources                  | Phospholipase | Uniprot  |
|-----------|-----------------------------------|---------------|----------|
| Api c 1   | Apis cerena (honeybee)            | Phospholipase A2 | V9IM80  |
| Api m 1   | Apis mellifera (honeybee)         | Phospholipase A2 | P00630  |
| Bom p 1   | Bombus pennsylvanicus (american bumblebee) | Phospholipase A2 | Q7M416  |
| Bom t 1   | Bombus terrestris (buff-tailed bumblebee) | Phospholipase A2 | P82971  |
| Pol a 1   | Polistes annulareis (paper wasp)  | Phospholipase A1 | Q9U6W0  |
| Pol d 1   | Polistes dominula (cardboard wasp) | Phospholipase A1 | Q6Q252  |
| Poly p 1  | Polybia paulista (South America wasp) | Phospholipase A1 | A2BC4   |
| Vesp c 1  | Vespa crabro (european hornet)    | Phospholipase A1 | P0CH87  |
| Dol m 1   | Dolichovespula maculate (baldfaced hornet) | Phospholipase A1 | P53357  |
| Ves v 1   | Vespuca vulgaris (common wasp)    | Phospholipase A1 | P49369  |
| Ves m 1   | Vespuca musculifrons (wasp of east yellow jacket) | Phospholipase A1 | P51528  |
| Ves s 1   | Vespuca squamosa (wasp of south yellow jacket) | Phospholipase A1 | P0CH86  |
| Sol i 1   | Solenopsis invicta (red imported fire ant) | Phospholipase A1 | Q68KK0  |
| Sol i 2   | Solenopsis invicta (red imported fire ant) | Phospholipase A1 | P35775  |
| Sol s 2   | Solenopsis saevissima (red ant)   | Phospholipase like A1 | A5X2H7  |
| Not allergen | Culex quinquefasciatus (house mosquito) | Phospholipase A2 | B0WT10  |
| Not allergen | Centruroides hentzi (scorpion)     | Phospholipase A1 | A0WD9LPH1 |
| Not allergen | Parasteatoda tepidarium (house spider) | Phospholipase A2 | A02L2Y6H2 |
0.683688 (0.668387) for all sites and parsimony-informative sites. A closed relationship among phospholipase allergens as shown, formed four nodes with a high phylogenetic relationship among them (Figure 1). According to the tree, group A grouped three phospholipase A1 all belonging to the Vespuca genus (Ves v 1, Ves m 1, and Ves s 1). This group presents the greatest relationship among the groups with the closest distance between branches. Meanwhile, group B contains the highest number of phospholipases A2 phylogenetically related, including allergens of the Bombus and Apis genera (Bom p 1, Bom t 1, Api m 1, Api c 1) and two non-allergic phospholipases from Parasteatoda tepidariorum (Common house spider) and Centruroides hentzi (Hentz striped scorpion). Group C included four proteins, three of them from ants belonging to Solenopsis genus (Sol i 1, Sol i 2, Sol s 2) and one belonging to the mosquito C. quinquefasciatus. In group D we found all the wasp allergens that belong to the genus Polistes (Pol a 1, Pol d 1), P. Paulista (Poly p 1), and D. maculate (Dol m 1) and V. crabro (Vesp c 1).

Identification of potential cross-reactive antigenic sites
Multiple alignments of the phospholipases of the different groups obtained from the phylogenetic analyzes were made. We built four 3D models of the 18 phospholipases Ves s 1, Sol i 1, Culex quinquefasciatus and Centruroides hentzi. The remaining proteins were reported on the UniProt database. We considered structures for better visibility of antigenic patches, the parameters for structural quality control for homology models are found in Table 3. To compare the ElliPro results, we chose the main antigen patches with a score higher than 0.7 and more than three residues, taking as reference the epitope of one phospholipase of each group; group A: Ves m 1; group B: Bom p 1; group C: Sol i 1; Group D: Pol d 1 (Table 2). The constitutional antigenic patches are shown in Figure 2.

Phospholipases from group A had a shared identity of 79% between their amino acid sequences (Figure 3). A total of 704 residues were identified and conserved among the phospholipases analyzed, and for these group, we used Ves m 1 to identify the possible epitopes. We found three common linear antigenic patches and two constitutive antigenic patches with a score greater than 0.7.

Group B shares an identity of 35% between their amino acid sequences but when we exclude Api c 1, the identity increases to 64%. In total, 259 identical residues among the

Figure 1. Phylogenetic tree based on the amino acid sequences of the phospholipases studied. The formation of fourth clades (A–D) with the highest degree of identity is observed (79% for clade A). The evolutionary history was inferred using the Maximum Parsimony method. The most parsimonious tree with length equal to 1462 is shown.
sequences were found. We found and included three linear epitopes and two discontinuous antigenic patches in Bom p 1 with a score >0.7. (Figure 4).

Group C, which includes allergens from ants, showed the lowest identity, with only 23% and the highest number of gaps (600 residues missing). Sol i 1 was the protein furthest away from any of the Hymenoptera allergens and appears to be closely related with wasps’ allergens. No common antigenic patches were detected; however, Sol l 1 presents an interesting antigenic patch with 46 residues and a score of 0.711.

For group D, 1916 residues exhibit an identity among the five sequences of allergens. This group exhibit a high identity

| Groups of Phospholipases | Residues preserved and antigenic capacity |
|--------------------------|------------------------------------------|
| A (Ves m 1)              | T17, N20, N22, D24, L25, Y26, T27, L28, Q29, T30, L31, Q32, N33, H34, P35, E36, F37, K38, K39, K40, T41, I42, T43, R44, P45, D73, N74, Y75, Q118, K119, V121, K122, D123, Y124, K125, I126, S127, M128, A129, N130, R149, Q151, E152, L153, K154, K155, I156, P157, A159, L183, I184, N185, K186, T187, P189 |
| B (Bom p 1)              | L98, Y100, P101, I102, V103, K104, C105, K106, V107, K108, S109, T110, I111, L112, C115, K116, E117, Y118, E119, F120, D121, T122, N123, A124, P125, Q126, A127 |
| C (Sol i 1)              | L52, Y53, N54, S55, F57, Q58, G59, K60, N61, L62, G63, N64, Q65, Q66, S67, C68, E69, D70, I71, N72, A73, S74, L75, A87, F100, V101, Q102, K103, G104, H105, V148, D149, M151, N152, K153, C154, K155, I156, P157, A159, L183, I184, N185, K186, T187, P189 |
| D (Pol d 1)              | W285, K286, S287, Y288, F289 |

| PARAMETERS | PROTEINS | Ramachandran favoured | GMQE | QMEAN4 | CB All Atom | Solvation | Torsion |
|------------|----------|-----------------------|------|--------|-------------|-----------|---------|
| Api c 1    | 97.93%   | 0.67                  | -3.02| -2.28  | -0.40       | 0.54      | -2.81   |
| Api m 1    | 96.21%   | 0.16                  | -1.58| -0.48  | 0.53        | -1.89     | -0.91   |
| Bom p 1    | 96.27%   | 0.77                  | -2.85| -0.97  | -1.52       | -1.18     | -2.40   |
| Bom t 1    | 97.39%   | 0.78                  | -2.85| -0.25  | -1.80       | -0.74     | -2.69   |
| Pol a 1    | 83.96%   | 0.33                  | -6.49| -3.17  | -3.20       | -2.02     | -5.93   |
| Pol d 1    | 90.41%   | 0.64                  | -1.97| 0.32   | -0.96       | 0.96      | -2.47   |
| Poly p 1   | 91.19%   | 0.31                  | -2.05| 0.46   | -1.41       | 0.94      | -2.54   |
| Vesp c 1   | 92.91%   | 0.96                  | -1.23| 0.18   | -1.08       | 0.53      | -1.49   |
| Dol m 1    | 90.17%   | 0.81                  | -1.98| 1.16   | -0.87       | 0.86      | -2.61   |
| Ves v 1    | 90.82%   | 0.78                  | -1.42| 0.98   | -0.78       | 0.82      | -1.97   |
| Ves m 1    | 91.16%   | 0.88                  | -1.20| 0.87   | -0.78       | 0.95      | -1.75   |
| Sol i 1    | 87.06%   | 0.56                  | -4.91| -1.91  | -2.20       | -0.26     | -4.79   |
| Sol i 2    | 99.12%   | 0.88                  | 0.23 | 0.85   | 1.20        | 0.80      | -0.35   |
| Sol s 2    | 99.12%   | 0.76                  | 0.06 | 0.99   | 0.85        | 1.33      | -0.76   |
| Culex quinquefasciatus | 91.07% | 0.43                  | -2.85| -0.53  | -1.58       | -1.08     | -2.36   |
| Centruroides hentzi | 89.74% | 0.64                  | -4.75| -2.76  | -2.82       | -0.28     | -1.88   |
| Parasteatoda tepidariorum | 91.60% | 0.54                  | -2.70| -2.24  | -1.85       | -4.26     | -1.69   |
Figure 2. Constitutional antigenic patches with the higher score among groups. The Sol I 1 structure was obtained by homology, using the Vespid basal is sequence as a template (34.32% de identical; QMEAN -5, 45).

Figure 3. Analysis of group A phospholipases. Unconserved sequence are shown with blue color and high conserved sequence with red color. Moderately conserve sequence are showed with green and orange color. The alignment score was 14,674 with a total of 704 identities residues. The percent sequence identity was 79%.
of 64%, the second highest after clade A. For the identification of antigen patches in this group, we used Pol d 1 finding four linear epitopes but only one continuous epitope with a valid score (Figure 5).

Discussion

Phospholipases A1 and A2 are allergens of insects, which provide a diagnostic benefit for the differentiation of genuine cross-reactivity sensitization. However, the cross-reactivity of this group of allergens has scarcely been holistically explored. In this study, we were able to predict those possible antigenic regions that could explain the cross-reactivity of phospholipases in Hymenoptera through in silico analyses.

The 18 amino acid sequences of the allergens were aligned, and a phylogenetic analysis was carried out which yielded four monophylogenetic groups (A, B, C, D). Group A yielded the highest degree of identity among their amino acid sequences (79%). All the allergens of this group belong to the Vespuia genus, one of the most studied sources of wasp allergens. In group B (Bom p 1, Bom t 1, Api m 1, Api c 1) two analyses were conducted, the first with the presence of the Api c 1 allergen where a degree of identity of (35%) was found and the second without the allergen, where we found a higher degree of identity at (64%). This showed that the alignment of these three species could explain a possible cross-reactivity. Group C (Pol a 1, Pol d 1, Poly p 1, Vesp c 1, Dol m 1.02) showed a level of identity of (64%). However, analysis of conserved and affected residues showed that Group A shares three antigenic regions that could contribute to their cross-reactivity.

IgE against cross-reactive carbohydrate determinants (CCD) is one of the main causes of double positivity and is present in most hymenopteran venom allergens with more frequency in venom from HB and YJ in patients that are allergic to insect bites. The prevalence of this allergen has been described in more than 20% of patients allergic to honeybee venom; approximately one of four HB poisons and one of 10 YJ venom allergens have been found to be CCDsIgE-positive. The PLA2 structure contains the insect CCDs that are specified by the presence of a 3-core α-1 fucose. Insect CCD causes 69% to 75% double positive test results for HBV and YJV during allergy diagnosis. Hemmer et al. propose that the Radio Allergo Sorbent Test (RAST) results to OSR pollen appear to be a simple and practicable measure to detect sugar specific IgE in individual sera. This could be useful to discriminate between patients who cross react through CCD and doubly sensitized patients who may require immunotherapy with two poisons.

Currently, CCD-free allergens have been known to allow cross-reactivity between proteins to be found without having a double positivity. Ves v 1, Api m 1, Dol m 1, Pol d 1 are allergens that lack cross-reactivity based on CCD and allow diagnoses without interference. However, it should be clarified that these are mostly of recombinant origin because in its purified natural form possess CCD; for example, Api m 1
of natural origin has CCD and makes diagnosis difficult. On the other hand, Sol i 1 is the only PLA1 hymenopteran venom known to have CCD, which could make the specific diagnosis of fire ant allergy difficult.

Research on the allergenic capacity of Hymenoptera allergens has been characterized by individualized studies, with Api m 1, Sol i 1, Pol d 1, Ves m 1 among those most studied so far, but the possible cross-reactivity between phospholipase allergens A1 and A2 has not been holistically evaluated.

No cross-reactivity between A. mellifera, S. invicta and V. vulgaris was detected, which supports our results, since there was no relationship between these allergens. However, when analyzed along with other allergens, it was observed that a certain degree of identity is maintained between these two proteins, suggesting a possible cross reactivity without CCD.

Figure 5. Analysis of group D phospholipases. Five sequences were studied with a total of 1564 residues. Unconserved sequences are shown with blue color and highly conserved sequence with red color. Middle conserve sequence are shown with green and orange colors. The percent sequence identity was 64% with a total of 1916 Aidentities residues.
Group A (Ves m 1, Ves s 1 and Ves v 1) being the most representative, the cross reactivity between Vespula spp. is strong due to the similarities in the composition of the poison and the structure of the individual allergens\textsuperscript{21}. Different studies evaluate the identity of the yellow jackets; for example, a 1996 study reported that Ves v 1 had 95% identity with Ves m 1 and both yellow jacket phospholipases have about 67% sequence identity with the hornet protein Dol m 1\textsuperscript{7}. Other authors demonstrated that Ves v 1 also shows an identity of 54% with Poly p 1, it being the lowest among the allergens studied and a study carried out in Spain with 59 previously diagnosed allergic patients with an allergy to vespids found that there could be a double sensitization between Ves v 1 and Pol d 1 because in 31% of patients they could not be clearly defined as sensitized only to Vespula or Polistes\textsuperscript{22,29}. Consequently, the different Vespula poisons react strongly in a crossed manner, which would explain the high degree of identity found in the study (Group A (79%)). Of the three proteins, only Ves v 1 has been described as a CCD allergen, showing that this interaction between the Vespula phospholipases could be CCD-independent and related only by protein structure\textsuperscript{19}.

The quaternary structure of the three Vespula phospholipases is also very similar, suggesting the possibility of present both linear and conformational epitopes (Figure 6A). Therefore, we suggested that fragment inhibition studies be carried out to identify the possible antigenic peptide described in this study.

Group B showed a degree of identity of 35%, however, in the analysis, we found that if we performed the alignment without the Api c 1 allergen, the degree of conservation between Api m 1, Bom p 1 and Bom t 1 increased to 64%. So far, we have found no more information about the possible cross reactivity in these allergens. In this group, Api m 1 is the most characterized allergen; It contains the cross-reactive carbohydrate (CCD) determinants of insects that are defined by the presence of a 3-core α-1 fucose\textsuperscript{30}.

For years, the detection of Api m 1 CCD challenges the differentiation of HB and YJ allergy. However, \textit{in vitro} detection of immunoreactive sIgE from these insects showed double positivity in up to 59% of the patients\textsuperscript{34}. PLA2s possess important venom allergens in other members of the genus

\textbf{Figure 6. Overlapping of phospholipases.} (A) Consisting of Ves v 1 in green, Ves m 1 in orange and Ves s 1 in blue. (B) Consisting of Api c 1, Api m 1, Bom p 1, Bom t 1, Centruroides hentzi and Parasteatoda tepidariorum, but we only found structural homology in Api m 1 of orange color, Bom p 1 of yellow color and Bom t 1 of red color. (C) Consisting of Sol i 1, Sol i 2, Sol s 1, and Culex quinquefasciatus but we only found structural homology in Sol i 2 of a magenta color, Sol s 1 of blue color. (D) Consisting of Pol a 1 of green color, Pol d 1 of yellow color, Poly p 1 of blue color, Dol m 1 of cyan color and Vesp c 1 of red color.
Apis and Bombus that have been shown to have homology. A. cerena (Api c 1) have been little explored but have been described as having high identity levels with other phospholipases, like A. mellifera (95%)\(^9\). In our study, we observed that when comparing the sequences of these phospholipases with those of the genus Bombus, that identity was not preserved since the identity we found was very low and when excluding it from the alignment, the sequences were more conserved\(^1\). Studies conducted on the genus Bombus found that the primary sequences of Bom t 1 and Api m 1 have an identity of 53% and their three-dimensional structures show conserved low protein surfaces\(^32\). However, the allergens selected from group B in our study showed a high conservation and structural homology leading to possible cross-reactivity (Figure 6B).

As for Group C, we highlight that it was the only group that included phospholipases A1 and A2 in the clade, so a low identity was expected. We found that the ant phospholipases Sol i 1, Sol i 2 and Sol s 2 showed a degree of alignment identity with the other phospholipases in the primary sequences of 23%. This low identity is not enough to explain cross-reactivity in silico, even though allergen Sol i 1 has been extensively analyzed and other studies suggest that it may have a possible reactivity with the Centruroides species\(^33,34\).

The phylogenetic analyzes reported in this study revealed that Sol i 1 is the most divergent member among the currently identified hymenopteran venom group PLA1. As noted, Sol i 1 is in a group (group C) completely isolated from the clade consisting of wasp allergenic PLA1 (group A) and showed no structural homology (Figure 6C). Furthermore, in multiple alignments, the fire ant exhibits the lowest level of sequence identity. However, studies have shown cross-reactivity between Sol i 1 and its wasp counterparts with amino acid sequence identity levels of 38% with Ves m 1, 36% with Ves v 1, 40% with Dol m, 1.35% with Pol d, 1.36% with Poly p 1\(^1\). However, a recent study suggests that peptide-based cross-reactivity between Sol i 1 and PLA1 of Polistinae wasps does not occur because the alignments and the phylogenetic and structural analyzes showed that it is an allergen further from its counterparts, in addition to possessing the lowest level of identity among the sequences studied, with 36%, and the highest RMSD value with 0.172\(^2\).

Several works have attempted to demonstrate cross-reactivity between A1 phospholipases\(^35,36\). The cross-reactivity based on PLA1 of the venoms of eight hymenoptera was analyzed and it was described that the identity of the primary sequence of Poly p 1 was conserved in 36% with Sol i 1, 74% with Pol d 1 and 71% with Pol a 1. In our study no relationship was found between Poly p 1 with Sol i 1. However, group D, where we found the different species of Polistes (Pol a 1 and Pol s 1), Poly p 1, Dol m 1 and Vesp c 1, showed a high degree of identity of 64% and structure homology (Figure 6D), enough to explain cross-reactivity\(^2\). An attempt was made to look for cross reactivity between Dol m 1, Ves v 1 and pol a 1 with mice; partial cross-reactivities in the T-cell epitopes of homologous vespid allergens was found, which supports our findings\(^2,39,36\).

Of the species chosen, three non-allergenic phospholipases (Centruroides hentzi, Parasteatoda tepidariorum, Culex quinquefasciatus) were taken to adjust the phylogenetic analysis, so as the results were produced, we observed that these phospholipases separated into two clades showing some affinity for some phospholipases allergens.

A study identified allergens in the venom of common striped scorpions. Eleven patients with scorpion venom allergy were assessed, where four patients had a history of anaphylaxis (with positive skin test responses) to imported fire ant venom (IFA) and at least two other had a history of large local reactions, suggesting that there could be a cross reactivity between proteins of these insects; this association would be clinically relevant\(^2\). This shows that despite not being described as allergens, it is necessary to carry out studies to verify their capacity to trigger sensitization.

Bioinformatic studies are high impact tools of great importance. Currently they are recognized as the first step to conducting an investigation, since they are in silico analyzes that facilitate a possible approximation to expected results, allow predictions or models, and serve as the basis for the emergence of large projects. In our study, we show possible antigenic regions involved in cross-reactivity between phospholipases A1 and A2, based on what was found with the use of in silico analysis we can say that they are proteins with a high degree of identity and that three antigenic regions were found, which would explain possible co-sensitization.

**Conclusion**

Potential antigenic sites were identified for the generation of cross-reactivity between the phospholipases analyzed in this study. The identity between these proteins of different species is relatively high, which shows that cross-reactivity between them is possible and their frequency in most cases can be high. These studies support diagnostic testing by component studies for venom allergy and the need to carry targeted mutagenesis tests is important to confirm their relevance in the allergic capacity of phospholipases.

**Data availability**

All data underlying the results are available as part of the article and no additional source data are required.
The benefits of publishing with F1000Research:

- Your article is published within days, with no editorial bias
- You can publish traditional articles, null/negative results, case reports, data notes and more
- The peer review process is transparent and collaborative
- Your article is indexed in PubMed after passing peer review
- Dedicated customer support at every stage

For pre-submission enquiries, contact research@f1000.com