

Potential allergenicity of *Medicago sativa* investigated by a combined IgE-binding inhibition, proteomics and *in silico* approach

Marwa Yakhlef,^{a,b} Ivana Giangrieco,^{b,c} Maria A Ciardiello,^b Immacolata Fiume,^b Adriano Mari,^{c,d} Lynda Souiki^e and Gabriella Pocsfalvi^{b*} 



Abstract

BACKGROUND: Alfalfa (*Medicago sativa* L) is one of the most planted crops worldwide primarily used to feed animals. The use of alfalfa in human diet as sprouts, infusions and nutritional supplements is rapidly gaining popularity. Despite this, allergenicity assessment of this novel plant food is largely lacking.

RESULTS: Here, leaf protein extract of alfalfa was studied using a combined proteomics, Immunoglobulin E (IgE)-binding inhibition assay and *in silico* approach to find potential allergens. We have identified and annotated 129 proteins using in-gel digestion proteomics and Blast2Go suit. A search against COMPARE database, using the identified proteins as query sequences, revealed high similarity with several allergenic proteins. The Single Point Highest Inhibition Achievable assay (SPHIAa) performed on the multiplex FABER[®] allergy testing system confirmed the *in silico* results and showed some additional potential allergens. This approach allowed the detection of proteins in alfalfa leaves cross-reacting with plant allergens from three different allergen families such as lipid transfer, thaumatin-like and Bet v 1-like protein families. In addition, the absence of structural determinants cross-reacting with seed storage allergenic proteins and with animal allergens was recorded.

CONCLUSION: This study reports for the first time potential allergenic proteins in alfalfa. The results suggest that this plant food can be safely introduced, as a protein-rich supplement, in the diet of patients allergic to animal food allergens. Allergic patients towards certain plant food allergens need to be careful about consuming alfalfa because they might have allergic symptoms.
© 2020 Society of Chemical Industry

Supporting information may be found in the online version of this article.

Keywords: alfalfa; *Medicago sativa*; allergens; proteins; mass spectrometry-based proteomics; IgE inhibition test; bioinformatics; food allergy

INTRODUCTION

According to the World Health Organization (WHO), allergic diseases constitute a significant cause of morbidity worldwide and a considerable burden on the health and medical systems of both developed and emerging economies.¹ Over the last two decades, the number of people with allergies has doubled and food allergy is a part of this increase. Treatment options are typically based on strict dietary elimination of foods to which the patients are sensitized to. The list of major food allergenic sources include milk, eggs, wheat, soy, peanut, tree nuts, fish, shellfish, etc.² However, several studies have raised concerns for poor growth and nutritional deficiencies related to avoidance diets.^{3,4} In this context, guiding food allergic patients in choosing appropriate alternatives that supply necessary nutrients becomes crucial to avoid deficiencies, and in pediatric patients also to avoid retardation in growth. Many attempts have been made to substitute the highly allergenic animal protein-based food stuffs with primarily from legumes, cereals, cassava, leaf and whole plant proteins in

the human diet. Nevertheless, plant-based diet can be allergenic too.

* Correspondence to: G Pocsfalvi, Institute of Biosciences and BioResources, National Research Council, 80131 Naples, Italy. E-mail: gabriella.pocsfalvi@ibbr.cnr.it

a Laboratoire de Biologie, Eau et Environnement, Department of Biology, Faculté des Sciences de la Nature et de la Vie et Sciences de la Terre et l'Univers, Université 8 Mai 1945 Guelma, Guelma, Algeria

b Institute of Biosciences and BioResources, National Research Council, Naples, Italy

c Allergy Data Laboratories (ADL), Latina, Italy

d Associated Centre for Molecular Allergology, Rome, Italy

e Department of Biology, Faculté des Sciences de la Nature et de la Vie et Sciences de la Terre et l'Univers, Université 8 Mai 1945 Guelma, Guelma, Algeria

Plant foods can cause local or systemic allergic reactions ranging from mild to severe, such as anaphylaxis. Proteins that are involved in physiology related basic mechanisms are generally expressed in all plant species. One example is the profilin that is known to be expressed in all plant tissues.⁵ Plant pathogenesis-related Bet v 1-like proteins (PR-10) are also ubiquitous and considered panallergens.⁶ Both profilin and Bet v 1 are classified as class 2 food allergens, that are heat-labile, prone to degrade by the gastrointestinal proteases⁷ and responsible for localized oral allergy symptoms.⁸ In contrast, class 1 food allergens are heat- and protease-stable that can cause severe and systemic allergic reactions.⁹ Lipid transfer proteins (LTPs), for example are class 1 food allergens that are widespread in plants where they function as intracellular and extracellular carrier proteins for hydrophobic molecules.¹⁰ Class 1 food allergens also includes (i) seed storage proteins, such as 2S albumins, 7S and 11S globulins, (ii) enzymes such as cysteine proteases, glucanases and chitinases and (iii) gibberellin regulated proteins (GRPs). In addition, there are some allergens not yet classified because they are poorly characterized or their clinical relevance is still unclear.

It is well known that homologs of allergenic proteins can produce cross-reactions which represent a risk for allergic people. Cross-reactions are caused by the Immunoglobulin E (IgE) epitopes shared at least partially between homologous proteins which are co-recognized by specific antibodies. This implies that higher is the structural similarity between two proteins, including primary, secondary and tertiary structure, the higher is the probability that the two molecules bear common epitopes (sequence and/or conformational epitopes) causing cross-reactions. The competition for IgE binding displayed by allergens with shared epitopes can be exploited for research purposes. For instance, it can be used to investigate immunological similarities between different molecules and also to detect possible allergens in a protein mixture, such as a food extract, by performing IgE inhibition test. In food allergen testing, multiplex systems that allow the analysis of a comprehensive panel of possible allergens in the proteome of a food extract are especially of value.^{11–13} Pasquariello *et al.*, for example, have used IgE inhibition tests performed with the Single Point Highest Inhibition Achievable assay (SPHIAa) in combination with a multiplex biochip-based immunoassay to analyse the allergen profiles of protein extracts from different apple cultivars¹¹ and to select potentially hypo- or hyper-allergenic apple cultivars. Similarly, the IgE inhibition with the SPHIAa method, combined with the allergen multiplex FABER test, was used to assess the cross-reaction between a protein extract and individual allergens, namely the cypress pollen extract and the peach and cypress GRPs.¹²

Alfalfa (*M. sativa*) meaning 'father of all foods' is a crop plant from the family of *Fabaceae*. It is the most popular and widespread, protein-rich crop which is grown in cool climate regions. Its protein content ranges from 170 to 220 g kg⁻¹ of shoot dry weight.¹⁴ This plant is grown for animal feed, soil improvement and medicinal uses. Alfalfa has a mild flavor and thus it could be well suited to human nutrition. In Europe, it is consumed in salads and sandwiches as raw sprouts or cooked and in the form of infusions.¹⁵ Alfalfa shows phytochemical activity even in small quantities¹⁶ and contains many vitamins and secondary metabolites. Alfalfa-based protein extract and powdered aerial part-based human nutritional supplements are emerging in the market. The increased use of this plant in human nutrition makes its more comprehensive analysis necessary. Despite the recent progress in genomic and proteomic characterization there are only

92 reviewed proteins of *M. sativa* L listed in the UniProtKB database.¹⁷ Watson *et al.* has identified cell wall proteins of *M. sativa* to generate proteome reference map for this important legume.¹⁸ In recent works, two-dimensional differential gel electrophoresis matrix-assisted laser desorption/ionization time-of-flight (2D-DIGE MALDI-TOF)^{19,20} mass spectrometry and isobaric tags for relative and absolute quantitation (iTRAQ)-based quantitative proteomics strategies have been applied to identify proteins and to detect changes upon developmental and environmental stimuli, like osmotic stress (drought),²¹ salt stress^{22,23} and atrazine stress.²⁴ These findings are important in the understanding of the molecular pathways and physiological mechanisms as well as the changes in protein expression in physiological and stress conditions in alfalfa.

Despite, the high economic and agricultural value, allergenicity assessment of the whole alfalfa plant and protein extract is rather limited and there are only some fragmentary data available today in the literature. A study based on skin prick test (SPT), shows that 36% of atopic patients living in Saud Arabia reacted to alfalfa.²⁵ Another study reported that 22.9% of a population suffering from asthma and living in the hot climate of the Arabian Desert were tested positive to alfalfa.²⁶ IgE co-reactivity between peanut and alfalfa was also observed using different immunological methods,²⁷ thus suggesting allergen sharing between these two species. However, a search in the WHO/International Union of Immunological Societies (IUIS) Allergen Nomenclature (www.allergen.org)²⁸ and in the Allergome database (www.allergome.org)²⁹ reveals no allergens identified in alfalfa so far.

Here, to investigate the allergenicity of *M. sativa* we have used a multidisciplinary approach. The strategy (Fig. 1) combines proteomics-based protein identification, immunological investigations, bioinformatics³⁰ and data mining to gain biological insights useful to evaluate the presence of potential allergens in protein extract. The possible allergenicity of the leaf protein extract of *M. sativa* L was investigated by IgE-binding inhibition test using the SPHIAa method^{11,12,31,32} associated to FABER test.¹² FABER is a nanotechnology-based multiplex *in vitro* serological test, which takes advantage of database and bioinformatics tools of the Allergome platform (<http://www.allergome.org/>) to analyze and store diagnostic and research data. To obtain the allergen profile of a protein extract, the SPHIAa method, combined with the FABER technology, represents a forefront tool that exploits a comprehensive panel of 244 allergens, including all of the most important allergy markers, in addition to exclusive allergens not available in other test systems.

MATERIAL AND METHODS

Sample preparation

The fresh leaves of *M. sativa* were manually separated from the stems, washed and dried on tissue paper in a stove at 38 °C for 48 h. The dried leaves were ground in a coffee grinder and passed through 0.3 mm sieve. The powder was stored in polyethylene plastic tubes in the dark until use.

Protein extract preparation

Protein extract was prepared from the dried powder of alfalfa using the protein extraction procedure developed for plant foods¹¹ with a few modifications. Briefly, 20 g of ground dried leaves were homogenized in a solution made with 80 mL of 0.5 mol L⁻¹ sodium chloride (NaCl), 0.8 g of poly(vinylpyrrolidone), 2 mL of 0.125 mol L⁻¹ ethylenediaminetetraacetic acid and 2.2 g of ascorbic acid dissolved

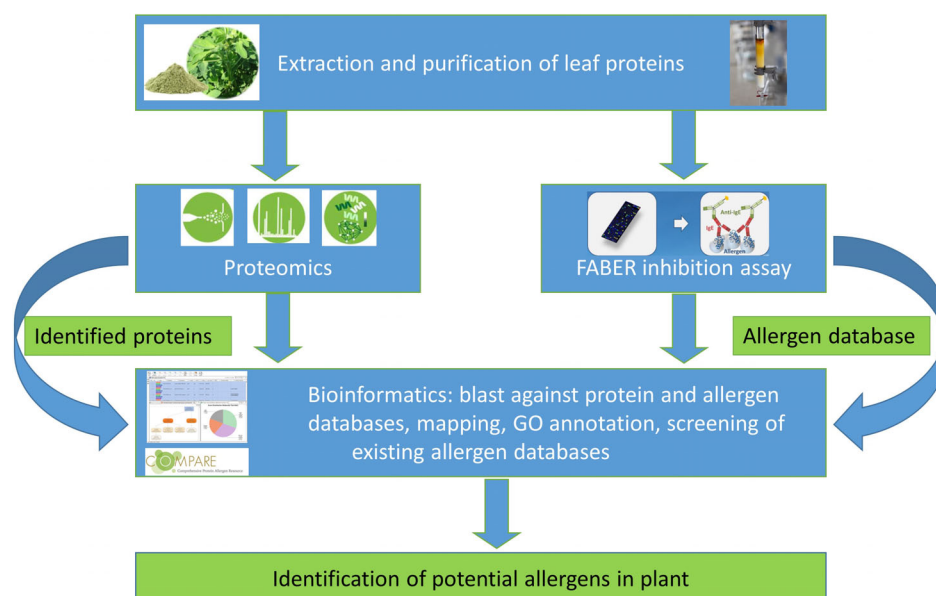


Figure 1. Scheme of the integrated multidisciplinary approach that combines proteomics, bioinformatics and SPHIAa assay for the identification of potential allergens in alfalfa protein extract.

in 25 mL of water and brought to pH 3.5 by addition of sodium hydroxide (NaOH). The sample was extracted for 2 h under stirring in an ice water bath. The sample was centrifuged at $10\,400 \times g$ for 1 h and about 100 mL of the supernatant were collected.

The proteins contained in 50 mL of the supernatant were salt precipitated using 90% (w/v) ammonium sulfate $(\text{NH}_4)_2\text{SO}_4$. The pellet obtained was resuspended in 8 mL of water and dialyzed against 50 mmol L^{-1} NaCl. The dialyzed sample (18 mL) was filtered with a $0.22 \mu\text{m}$ filter (Millex, Millipore, Bedford, MA, USA). The protein concentration of the extract determined by the Bradford method using the Bio-Rad Protein Assay (Bio-Rad, Milan, Italy) was 1.9 mg mL^{-1} . The sample was aliquoted and stored at -20°C until use.

IgE inhibition experiments with the SPHIAa method and FABER® testing system

The FABER version used to perform this study (FABER 244-122-122) contains 122 purified allergens and 122 multiple protein allergenic extracts, coupled to chemically activated nanoparticles. To detect possible allergens in alfalfa, the SPHIAa^{11,12} was performed using the FABER biochip containing 244 allergens. In the assay, pooled characterized sera (see later) containing the desired IgEs for the recognition of specific allergenic proteins were used. The assay was performed by incubating 0.12 mL of the sera pool with 0.12 mL of a solution containing 0.1 mg of the alfalfa leaf protein extract. The IgE-binding inhibition was evaluated by running the FABER test and recording the residual IgE binding on the allergens spotted on the biochip. Experiments were carried out in duplicates and the mean values are reported. Reference values for lack of IgE-binding inhibition were obtained by running control samples where the allergen solution was substituted with buffer only. The inhibition values were calculated in real time by the algorithm developed within the InterAll software package (version 5.0, Allergy Data Laboratories, Latina, Italy).³³ Only inhibition values equal or higher than 30% were considered in this study.

Patients' sera

Sera used in this study were selected among those stored in the serum bank of the Allergy Data Laboratories. These are residual sera deriving from venous blood sampling performed for the routine allergy diagnosis by FABER test.^{12,34,35} The features of each serum, in terms of content of IgE antibodies able to recognize and bind specific individual allergens (specific IgE) spotted on the FABER biochip, are registered in the InterAll databank (version 5.0, Allergy Data Laboratories).³²

Sera were selected based exclusively on the IgE content. In particular, for this study we have selected sera, which we were able to recognize and bind relevant purified allergenic proteins from plant and animal foods available on the FABER biochip. In addition, the chosen sera were free of IgE recognizing cross-reactive carbohydrate determinants (CCDs). In fact, they were tested negative against CCD-bearing proteins used as markers, namely bromelain from *Ananas comosus* and peroxidase from *Armoracia rusticana*. In the SPHIAa experiments, IgE is used as a probe to detect the presence of structural determinants, that is the epitopes of the proteins (purified or in mixture) under investigation. Therefore, the selection of sera was independent of the clinical history and/or symptoms of patients.

For the SPHIAa assay, a pool of four sera was prepared using 70 μL of each serum. Two sera were from patients sensitized to different plant foods. They contained IgE recognizing plant allergens, such as LTPs, profilin, Bet v 1-like, GRP, thaumatin-like protein and seed storage proteins. Two additional sera were from patients sensitized to allergens from animal foods, including egg, milk and fish. Thus, 0.12 mL of the pooled sera were mixed 1:1 with alfalfa protein extract resulting in a 1:8 final dilution of the individual serum.

All patients gave their informed consent to the use of their clinical data for research purposes in an anonymous form. In view of the purely comparative nature of this study, along with the fact that all venous blood samplings were part of routine clinical

practice and that a residual part of the routine sample was used for SPHIAa experiments, a formal approval by the Ethical Committee was not necessary.

Gel-based proteomic analysis

Leaf protein extract of *M. sativa* L (20 µg) was electrophoretically separated on a precast Novex 4–12% Bis-Tris NuPAGE gel using MOPS running buffer (Invitrogen, Carlsbad, CA, USA) according to the manufacturer's instructions and stained with colloidal Coomassie blue. The entire gel lane was cut into 5 mm bands, proteins were reduced, alkylated, *in-gel* digested using 20 µL of trypsin (Promega, Madison, WI, USA) at 6 ng µL⁻¹ concentration and extracted as described.³⁶ Extracted peptides were analyzed by nano-flow reversed-phase liquid chromatography electrospray ionization tandem mass spectrometry (LC-ESI-MS/MS; Perkin Elmer 200 series HPLC, PerkinElmer Life and Analytical Sciences, Shelton, CT, USA) coupled to the QStar Elite (Applied Biosystems/MDS Sciex, Ontario, Canada) mass spectrometer. Briefly, 10 µL sample was loaded, purified and concentrated on a reversed phase monolithic pre-column, 200 µm i.d. × 5 mm length (LCPackings, Sunnyvale, CA, USA) at 25 µL min⁻¹ flow rate. Peptides were separated at a flow rate of 300 nL min⁻¹ on a PepSwift Monolithic column, 100 µm i.d. × 5 cm length (LCPackings) using solvent A: 2% acetonitrile, 0.1% formic acid and solvent B: 98% acetonitrile, 0.1% formic acid with the following gradient: solvent B 5–50% in 20 min, solvent B 50% for 5 min and solvent B 50–98% in 6 s. Eluted peptides were analyzed by MS/MS using

information dependent acquisition (IDA) mode on a QSTAR Elite (Applied Biosystems, Foster City, CA, USA) equipped with a nano-flow electrospray ion source. Data acquisition, analysis and the extraction of the peak lists were performed by Analyst QS software.

Data analysis

Protein identification

Proteins were identified using the Byonic version 3.4.0 software against the UniProt database with *Medicago* taxonomy (93 591 entries). Note that today, there is only 97 reviewed protein sequences of *M. sativa* in the UniProt database that makes it necessary to use a larger database containing close relatives. The search criteria were as follows: enzyme: trypsin, 100 ppm precursor and 90 ppm fragment mass tolerance, peptide charges: +1, +2, +3. Two missed cleavages were allowed, and the following modifications were set: carbamidomethylation on cysteine as fixed modification, while methionine oxidation, asparagine and glutamine deamidation, ammonia loss, acetylation on protein as variable modifications. Protein identifications were validated by the Percolator algorithm³⁷ false discovery rate was < 1%. Only proteins with log prob > 3 values were included into the bioinformatics analysis.

Bioinformatics

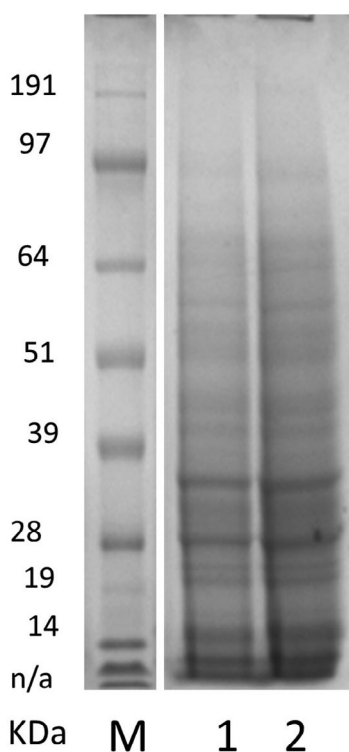
Blast2GO software package version 5.2.5 were used for *in silico* analysis.³⁰ Protein sequences were blasted using QBLAST service against two databases: (i) the UniProt public database using taxonomy filter green plants (*Viridiplantae*, 33 090 entries) and (ii) the COMPARE allergen database (2081 entries),³⁸ (<https://comparedatabase.org/>) using number of blast hits 20 and expectation value 1.0×10^{-3} .

The InterPro domain searches were performed using the public European Molecular Biology Lab-European Bioinformatics Institute to identify sequences against CDD, HAMAP, HMMpfam, HMMPIR, Fprintscan, BlastproDom Interpro's signatures. All sequences generated InterPro results. Annotated sequenced were mapped against exclusively created gene ontology (GO)-annotated proteins to obtain functional labels of GO-associated and Uniprot's ID mapping. Cellular localization was predicted using the plant subcellular location predictor (<http://bioinfo.usu.edu/PlantSubP/>).³⁹

RESULTS

Characterization of the leaf protein extract of *M. sativa*

We have performed *in-gel* digestion and LC-MS/MS-based proteomics on the leaf protein extract of *M. sativa*. The sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) image of the protein profile is shown in Fig. 2. Furthermore, 129 proteins were identified (Supporting Information Table S1) using the 'Medicago' taxonomy including 88 different species. Most of the proteins (114 sequences) were identified by homology with *M. truncatula* taxonomy and 14 with *M. sativa*. More than 98% of the identified proteins were computer-annotated. Table S2 shows the 20 top-ranking proteins identified in this sample. It is interesting to point out that most of them are enzymes with hydrolyze activity and not structural proteins. They are known for their essential function in plant defense mechanism against biotic and abiotic stresses, such as chitinases, thaumatin-like protein, enzymes involved in breaking down complex carbohydrates, like two glucan endo-1,3-beta-glucosidases, alpha-L-arabinofuranosidase, beta-1,3-glucanase 3, glycoside



M: molecular weight marker.

1: *M. medicago* leaf protein extract 20µg

2: *M. medicago* leaf protein extract 40µg

Figure 2. SDS-PAGE image of the leaf protein extract of *Medicago sativa*. Twenty bands were excised and analyzed by *in-gel* digestion proteomics.

hydrolase family 18 protein, beta-galactosidase and protease-like nepenthesin. Some of them are also involved in cell division and expansion, cell wall formation and organization like alpha-L- and glycerophosphoryl diester phosphodiesterase family protein.

Structural and functional annotation of identified proteins

To gain some insights into the cellular and biological roles as well as cellular localization of the identified proteins we performed functional annotation. Figure S1 (in Supporting Information) shows the node-directed acyclic graph (DAG) of the GO graphs for cellular component (CC), molecular function (MF) and biological processes (BPs). The major functional categories in CC were membrane (39 sequences, 46.4%), extracellular region (31 sequences, 37%), cytoplasm (25 sequences, 29.8%) and cell wall (15 sequences, 17.9%). For the MF, catalytic activity was dominant (95 sequences, 84.1%) including hydrolase activity: acting on ester bonds (14 sequences, 12.4%) and hydrolase activity: hydrolyzing O-glycosyl compounds (33 sequences, 29%) represented by chitinase activity (12 sequences, 10.7%). Finally, metabolic processes (100 sequences, 87.7%), cellular processes (61 sequences, 53.5%), and response to stimulus (20 sequences, 17.5%) were the most abundant groups in the BP.

In silico prediction of allergens in the leaf proteins of *M. sativa*

A BLAST search of the protein sequences identified in alfalfa (Table S1) was performed against the COMPARE database to find *Medicago* protein sequences with high similarity to known allergens (Table S3). Table 1 lists proteins showing more than 70% sequence similarity to allergens in the COMPARE database. The first four proteins (Table S3) have high similarities to Lactoylglutathione lyase, also called glyoxalase I, reported allergenic in rice seeds,⁴⁰ but it is not officially registered by WHO/IUIS. In addition, two proteins show high similarities to the apple thaumatin-like protein, Mal d 2.⁴¹ Thaumatin-like proteins belong to the family 5 of pathogenesis related proteins (PR-5). *In silico* analysis reveals four *Medicago* putative proteins belonging to the barwin family⁴² also known as family 4 of pathogenesis-related proteins (PR-4), showing high similarities to pro-hevein, that is a major latex allergen and a lectin-like glycoprotein with a chitin-binding domain

and a C-terminal with amyloid properties.⁴³ Another three proteins show similarities to plant LTPs and the highest identity with a lentil homolog (Len c 3). LTPs are low molecular mass panallergens ubiquitously expressed in plants belonging to the family PR-14. Significant similarity is also observed with a glucanase from latex, Hev b 2, with an endochitinase from chestnut, Cas s 5, and with a thioredoxin from maize, Zea m 25.

Identification of possible allergens in *M. sativa* by SPHIAa assay

The presence of potential allergenic proteins in *M. sativa* leaves was investigated by *in vitro* immunological tests. In particular, the SPHIAa method on the FABER system¹² was used. In the assay, proteins in the *M. sativa* extract competed with the allergens spotted on the FABER biochip for the binding to specific IgE contained in a pool of sera of allergic patients.

Figure 3 shows that, except for apricot and peach pulp, alfalfa extract inhibited the IgE binding to the tested fruits and vegetables, tubers and root. In contrast, out of 21 tested seed extracts, only four (maize, rice, pine nut and carob) were partially inhibited. No inhibition was observed on other seed extracts, although allergens isolated from some of these sources, such as the peanut Ara h 8 and Ara h 9 showed IgE binding competition (Fig. 4). This result suggests that the amount of allergens, such as Ara h 8 and Ara h 9 in the natural source is low.

Figure 4 shows that food allergenic proteins belonging to several protein families including LTPs, thaumatin-like protein, cysteine proteases, Bet v 1-like protein, GRP and chitinase (for additional details see Table 1 and Table S4) were inhibited by the alfalfa extract. In contrast, the food allergens classified as seed storage proteins, such as 2S albumin, 7S albumin, 11S globulin and lectin, were not inhibited, thus suggesting the absence of cross-reactive proteins in the protein extract of alfalfa.

IgE inhibitions were also observed on some of the extracts of latex and pollens spotted on the FABER biochip (Fig. 3) (see Table S4 for details). In addition, the presence of proteins cross-reacting with Bet v 1-like proteins, trypsin inhibitor and chitinase I was also confirmed by inhibition on their homologs from pollen sources. This observation suggests the possibility but does not prove the presence of alfalfa pollen in the starting material. In fact, inhibition could be due to structural similarity between

Table 1. Potential allergen families identified in the protein extract of leaf of *Medicago sativa* by Proteomics/BLAST search combined with IgE-binding inhibition experiments (Fig. 1)

Allergen family	Proteomics/bioinformatics			IgE-binding inhibition	
	UniProt accession number ^a	Similar allergens in COMPARE	Source	Homologous allergens on FABER	Source
Thaumatin-like protein (PR-5)	<i>I3T337</i> , <i>A0A396H709</i>	Mal d 2	Apple	Act d 2	Kiwifruit
Lipid transfer proteins (PR-14)	<i>Q45NK7</i> , <i>A0A396I3M7</i> , <i>G7JJJ6</i>	Len c 3	Lentil	Act d 10, Ara h 9, Cor a 8, Jug r 3, Pru p 3, Pun g 1, Zea m 14, Sola l 6	Kiwifruit, peanut, hazelnut, walnut, peach, pomegranate, corn
Bet v 1-like protein (PR-10)	D2Y175	Ara h 8	Peanut	Ara h 8, Mal d 1, Bet v 1	Peanut, apple, birch pollen

Proteins higher than 70% BLAST similarities in COMPARE database (Table S3) and homologous allergen/allergen family on FABER biochip with at least 30% inhibition (Table S4) were considered.

^a UniProt accession number indicating *M. trunculata* proteins are indicated in italics.

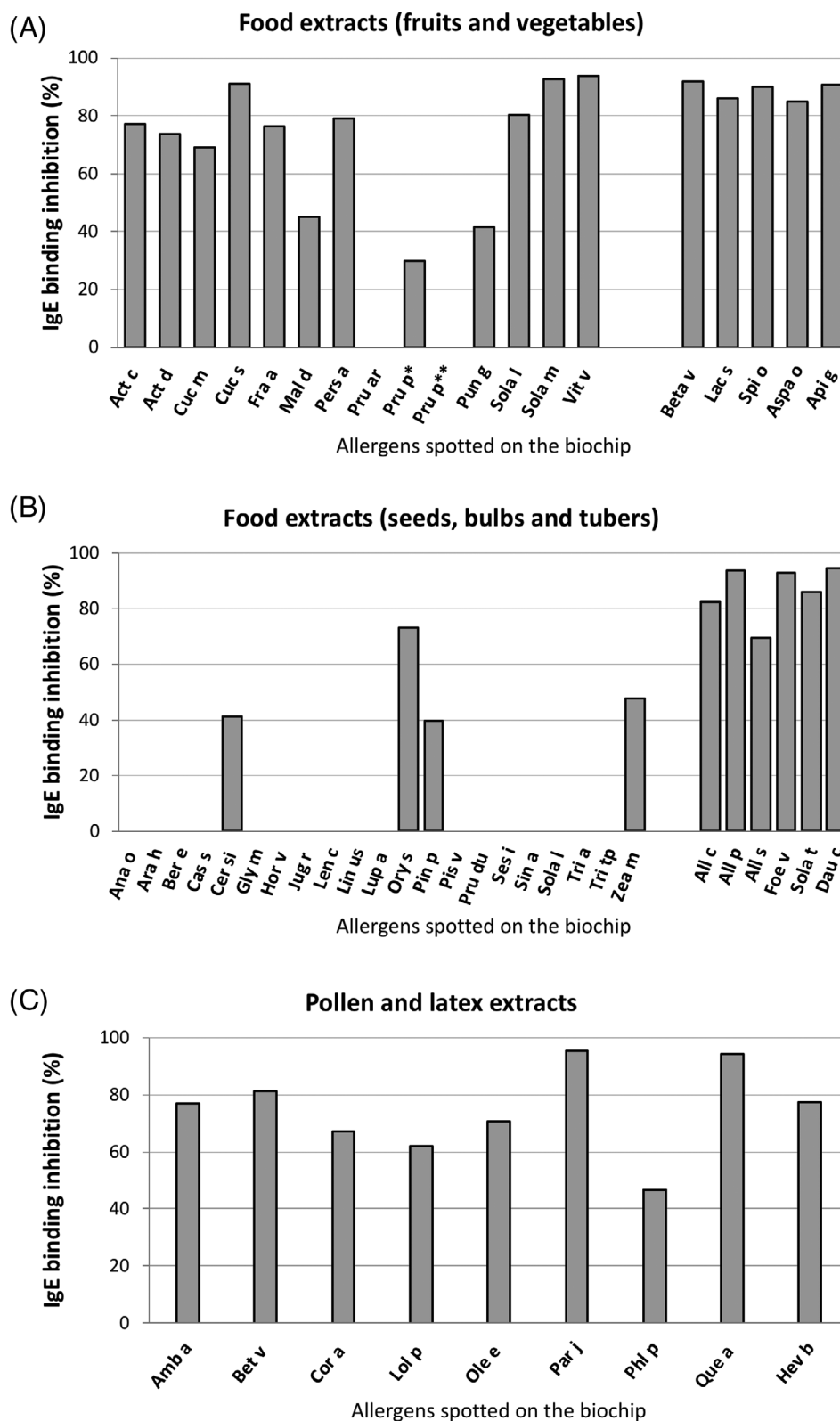


Figure 3. SPHIAa results obtained on extracts from plant food, pollen and latex. Allergenic sources of the same category are grouped. (A) Fruits and vegetables. (B) Seeds, bulbs and tubers; Pru p* and Pru p** represent the peach peel and pulp extracts, respectively. (C) Pollen and latex extracts.

homologous proteins, independently of the plant tissue source. The presence in alfalfa of other IgE-binding proteins, such as pectate lyase and expansin was also deduced on the basis of the IgE inhibitions on homologs from pollen sources.

Table 1 shows proteins belonging to three allergenic protein families, thaumatin-like proteins, Bet v 1-like proteins and LTP detected by the combined methods applied in this study (Fig. 1). Proteomics confirmed the identity of the proteins and

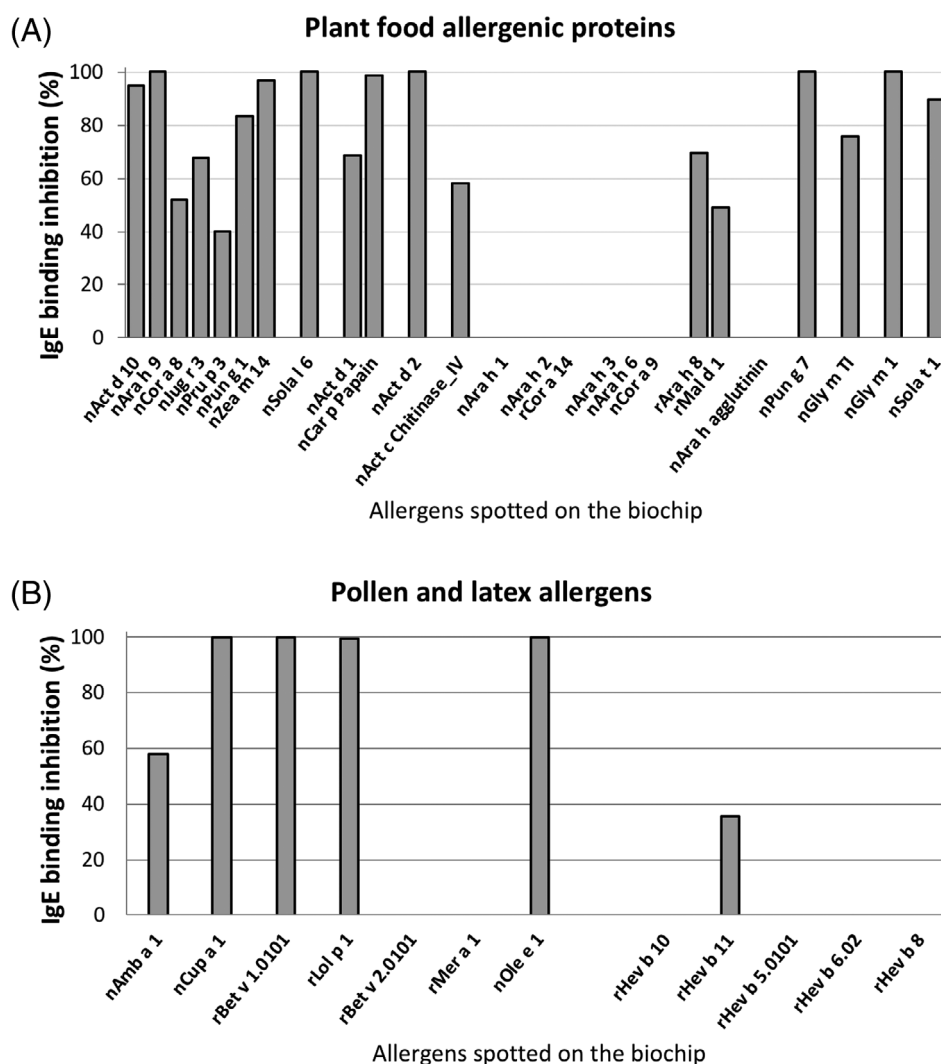


Figure 4. SPHIAa results obtained on allergenic molecules from plant foods (A) and on pollen and latex allergens (B). Allergens belonging to the same protein family are grouped. Allergen details are reported in Table S4.

their presence in the protein extract of alfalfa leaves. Bioinformatics highlighted their structural similarities to known allergens and the immunological results confirmed that they have functional IgE binding epitopes. These results highly support possible allergenicity of alfalfa proteins of these families.

It can be observed that components of the same protein family are inhibited at a different extent (Fig. 4), as observed for allergens belonging to the LTPs family. Results show that only one 9 k-LTP (Ara h 9) and 7 k-LTP (Sola l 6) are completely inhibited by the alfalfa extract. Slightly lower inhibition was observed on the 9 k-LTP from corn, kiwifruit and pomegranate, showing values of 97%, 95% and 83%, respectively. The lowest values were observed for LTPs from walnut, hazelnut and peach displaying 68%, 52% and 40% IgE binding inhibition, respectively.

Correlation between IgE binding inhibition and structural similarities in LTPs from different plant foods

The IgE binding inhibition observed on LTPs suggested the presence of at least one homologous protein in the alfalfa extract. An LTP (UniProt Accession number Q45NK7) was annotated in alfalfa based on experimental evidence at the transcript level.

The search performed in UniProt revealed that this is the only LTP isoform reported for *M. sativa* so far. The amino acid sequence of this LTP (82 amino acids length), although incomplete because of a lack of a short stretch of residues at the C-terminus, was used to analyze the structural similarity, in terms of sequence identity, with the allergenic homologs included in the FABER biochip. Figure 5 shows the alignment of alfalfa LTP with the analyzed homologous proteins. The isoform of *M. truncatula* LTP included in the alignment was selected because it gave the highest identity following the similarity search against the UniProt database, using the alfalfa LTP as query sequence. On the basis of this sequence alignment, the identity between alfalfa LTP and the homologous proteins was calculated (Table S5). In line with the taxonomic closeness between *M. truncatula* and *M. sativa*, the highest value of sequence identity (85%) is observed between the LTP of these two species. The comparison of alfalfa with allergenic 9 k-LTP reveals a correlation between sequence identity and IgE binding inhibitions. In fact, with the exception of the kiwifruit and peanut LTPs, gradually increasing values of sequence identity with the homologous molecule from peach 57%, hazelnut 58%, walnut 62%, pomegranate 63% and corn 67%, are observed and they

are in line with increasing values of IgE binding inhibition (Fig. 4). Kiwifruit and peanut LTPs are out of this correlation and their behavior might be explained by the presence of additional undetected isoforms of LTP in the alfalfa extract that could contribute to the IgE binding inhibition on those allergens.

DISCUSSION

The description of the allergen profile of a food is challenging and must be addressed with multidisciplinary approaches.⁴⁴ To classify a food protein as a cause of IgE-mediated allergy is necessary to demonstrate that (i) it is recognized by specific IgE and (ii) induces clinical allergy symptoms. IgE binding can be assessed using *in vitro* tests, but allergic reactions need to be evaluated on the basis of the clinical history of patients and/or by *in vivo* tests. These tests include prick-by-prick test (PPT), SPT and the provocation test (double-blind placebo-controlled food challenge, DBPCFC). DBPCFC is considered the 'gold standard' for the diagnosis of food allergy⁴⁵ and to confirm that a molecule is an allergen on the basis of clinical reactivity. However, this use of DBPCFC is limited especially by ethical issues due to important safety risks including anaphylaxis.⁴⁶ As a result many food allergens remain untested in DBPCFC and their characterization rely on alternative methodologies. The development of new combined strategies based on the integration of different established methodologies that produce orthogonal data and provide more comprehensive results could be helpful for the evaluation of potential allergenicity of a food source.⁴⁴

Here we set-up of a novel allergomics workflow (Fig. 1) that combines proteomics, *in silico* analysis and nanotechnology-based multiplex immunological tests. This approach takes advantages of blasting multiple sequences of the identified proteins in a single batch query against allergens present in curated databases and the SPHIAa assay performed on multiplex FABER biochip. A great advantage of the association of the SPHIAa assay with the

FABER test is the possibility to exploit a large panel of validated allergens available on a biochip and probed with characterized IgE antibodies, thus allowing a comprehensive analysis of the immunological features of the sample under investigation.

This approach allowed us to identify 129 proteins expressed in leaves of *M. sativa* (Table S1). Most of the proteins were identified by homology to known proteins of *M. truncatula* known to be a close relative to alfalfa and the most studied species in *Medicago* genus. Among them, several proteins resulted to be homologous to known allergens based on sequence similarities retrieved from the *in silico* analysis. A group of these proteins belongs to the barwin family, that includes barwin, prohevein, glucanases, class I and class II chitinases. Prohevein is composed of an N-terminal hevein-like domain and a C-terminal barwin domain. The hevein domain is a conserved chitin-binding structure that is widespread in various monocot and dicot plants where it is involved in innate immunity.⁴⁷ Several proteins share barwin-like and hevein-like structural modules that can be responsible for IgE co-recognition and cross-reactivity between different plant foods and between latex and plant foods. Experimental data, obtained using alfalfa protein extract as IgE binding inhibitor, show the absence of inhibition on the latex hevein (Hev b 6) and only a partial inhibition on the latex class I chitinase (Hev b 11) available on the FABER biochip. These results, together with the only partial IgE inhibition on the latex extract and the absence of inhibition on other latex allergens, suggest a low level of cross-reactions between alfalfa and latex. A partial inhibition is also observed on kiwifruit class IV chitinase, having a hevein module in the N-terminal region. Probably a low degree of cross-reactions between alfalfa proteins containing barwin-like and hevein-like domains and homologs present in plant foods also might occur. Further investigations are needed to assess these possibilities.

Class III chitinase was also identified by proteomics, but it did not show IgE binding inhibition on the chitinase III available on

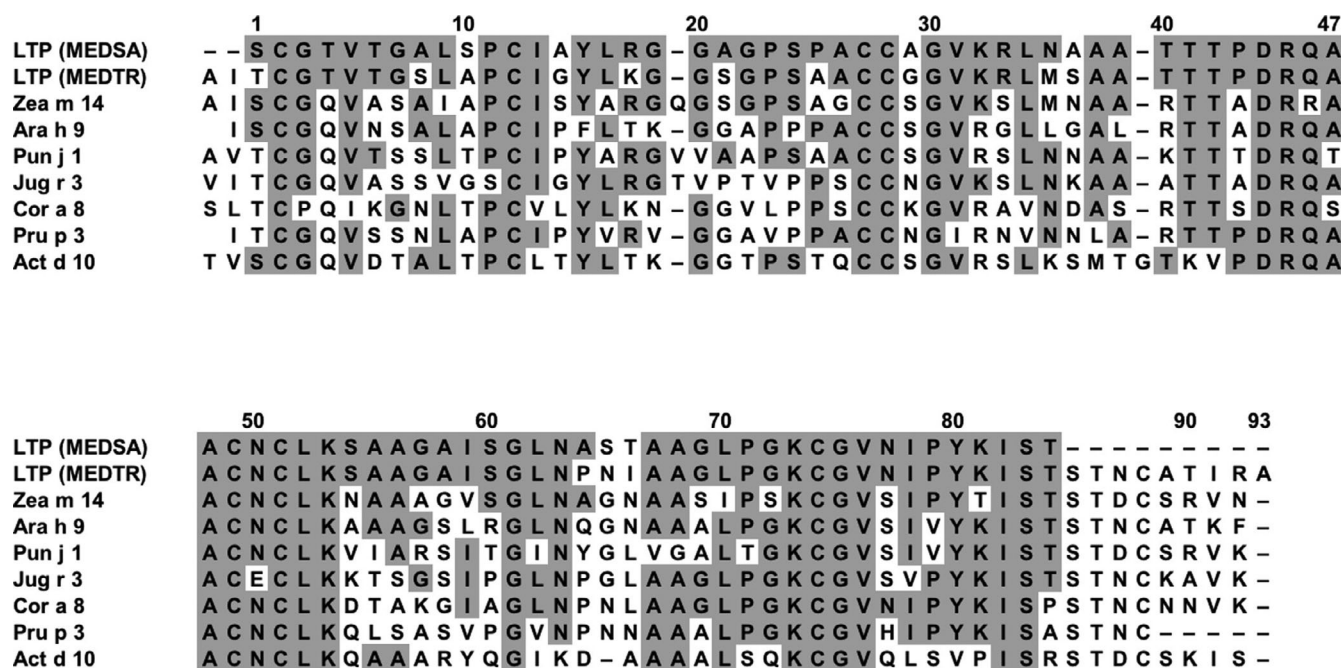


Figure 5. Multiple sequence alignment of *Medicago sativa* (MEDSA) LTP with a *Medicago truncatula* homolog (MEDTR) and with allergenic LTPs tested for IgE binding inhibition, namely the corn Zea m 14, peanut Ara h 9, pomegranate Pun j 1, walnut Jug r 3, hazelnut Cor a 8, peach Pru p 3 and green kiwifruit Act d 10. Their Uniprot accession numbers are Q45NK7, A0A396I974, Q2XX25, B6CEX8, A0A059SSZ0, C5H617, Q9ATH2, Q8H2B2 and P85206, respectively.

the FABER biochip, deriving from pomegranate fruit.³⁵ This result suggests that the IgE epitopes of this protein could be poorly conserved in homologous plant molecules.

LTPs are important class I plant food allergens.^{31,32} Several LTPs from different plant foods are available on the FABER biochip to allow an accurate analysis of patients sensitization to the sources of this allergy inducer. This panel of LTP was exploited to have indications about the correlation between allergen structural similarities and the experimental values of IgE binding inhibition. In fact, a high structural similarity suggests a possible high conservation of IgE binding epitopes responsible for IgE co-recognition and cross-reaction thus representing a risk for allergic people. A simplified method useful to obtain a preliminary estimation of the level of risk of an allergic reaction caused by a protein is based on the sequence similarity between the considered protein and that of known allergens. Although different criteria have been suggested by several authors, a value higher than 35% of amino acid identity with known allergens over a sliding window of 80 residues is the criterion used by the FAO/WHO/EFSA/Codex for the prediction of the allergenic risk of a new protein.⁴⁸ Although this method is not perfect and it does not take into consideration several contributing factors, such as the three-dimensional structure and the ligand binding that can affect the IgE epitope function, it can give a useful first indication. The sequence identity between the alfalfa LTP and the homologous molecules on the FABER biochip ranges from 51% and 67%. Therefore, they all are values predictive of allergenic risk based on the criteria indicated by the FAO/WHO/EFSA/Codex. The immunological tests indicated that alfalfa was able to compete for IgE binding with all the tested LTPs, thus indicating the sharing of antigenic epitopes with all these homologous proteins. In addition, with the exception of two LTPs, the results obtained show a good correlation between sequence identity and inhibition values, thus supporting the interpretation of the data obtained and the usefulness of the approach applied in this study. Clearly, we can expect that a high cross-reaction is correlated with a high structural similarity. For instance, the bioinformatics analysis showed the lentil LTP, Len c 3, as the most structurally similar to that from alfalfa. Therefore, we can expect an even higher cross-reaction with the lentil LTP, compared with the LTP available on the FABER biochip.

LTPs available on the FABER biochip are from kiwifruit, peanut, hazelnut, walnut, peach, pomegranate, maize and tomato. It can be observed that there is no correlation between IgE inhibition on the individual molecules and the extracts of the sources. For instance, purified peanut LTP, Ara h 9 is completely inhibited by the alfalfa extract, whereas the extract of the peanut seed (Ara h) shows no inhibition. This result is not surprising and it is in line with our knowledge on the relative abundance of different proteins in the natural source. In fact, the purified proteins are immobilized on the biochip at a constant and fixed amount. In contrast, the extracts contain protein components at different relative amounts like they are in the natural source. Therefore, the value of inhibition we can observe is essentially related to the IgE recognition towards the most abundant proteins of that natural source. An additional parameter that theoretically could have an effect on the IgE binding signals is the relative affinity towards IgE, but this is something that we cannot evaluate. In the case of peanut, the absence of inhibition on the extract spotted on the FABER biochip is reasonable due to the high abundance of seed storage proteins (2S albumin, 7S vicilin and 11S globulin) in this source. These proteins appear to be not inhibited by alfalfa extract. It is known that

the amount of LTP Ara h 9 is very low in peanut⁴⁹ and clearly that in the extract is not sufficient to give a detectable signal related to IgE binding.

Therefore, the inhibitions observed on the extracts have to be interpreted essentially taking into consideration the relative abundance of the protein components. On the basis of this concept, we can interpret the lack of inhibition on most of the tested seed extracts as the absence of cross-reaction between alfalfa proteins and the seed storage proteins that are the major components of these foods. Among seeds, the extracts from carob, rice and pine tree showed a partial inhibition that could probably be interpreted after the analysis of the major protein components of these foods. In contrast to seeds, plant foods represented by leaf, bulb, fruit, tuber and root generally showed high values of inhibition that suggest epitope sharing and cross-reactions between the alfalfa proteins and major protein components of these foods.

Profilins and Bet v 1-like proteins are panallergens found in all plant organisms. They both are found in pollens and in foods and can cause cross-reactions between these two allergy sources. Bet v 1-like protein was not detected by MS experiments, nevertheless it proved to be present in the alfalfa extract since it gave IgE inhibition on food and pollen homologous molecules available on the FABER biochip. Profilin was not detected by MS and did not show any IgE inhibition on the pollen and latex profilins available in the test. Since the presence of profilin in any plant cannot be questioned, it is conceivable that it was extracted from alfalfa leaves in amounts not sufficient to give detectable inhibition signals.

In addition, the immunological method detected some allergens that were not identified by proteomics in this work. Among these is a GRP, a class I allergen described in pollen¹² and food,⁵⁰ potato patatin, soybean trypsin inhibitor and hydrophobic peptide, thus suggesting that homologous proteins could be present in the alfalfa leaves.

CONCLUSION

The approach used in this study (Fig. 1) allowed the identification of still unknown components of the alfalfa proteome and provided us with a collection of data about the possible allergenicity of this plant source. As expected, alfalfa proteins did not show cross-reactions with allergens from animal foods, such as milk, egg, meat and fish, used as control markers. This suggests that alfalfa could be a valuable protein source for patients allergic to animal food proteins. In contrast, most of the analyzed fruits and vegetables showed cross-reaction with alfalfa proteins to different extents. Therefore, patients allergic to these plant foods have to be careful when consuming alfalfa.

Several potential allergens were identified in alfalfa by using a combined proteomic experiments, *in silico* analysis and immunological assay (Table 1). This strategy has proven useful to produce supporting information about the presence and/or absence of possible allergens in a food source. All together, the results obtained here strongly support the presence in alfalfa of at least three allergenic protein families. Likewise, they suggest the absence of proteins cross-reacting with animal food allergens. In addition, the results obtained suggest, for instance, that patients allergic to seed storage proteins could be allowed to introduce alfalfa in their diet. In contrast, patients sensitized to allergens, such as LTPs, GRP, thaumatin-like protein, cysteine protease, barwin-like, hevein-like and Bet v 1-like, have to be careful and

they need to make appropriate checks to verify whether alfalfa is a safe food for them. In fact, these patients could react to the alfalfa proteins homologous to those allergens when they have IgE recognizing shared antigenic determinants. This possibility needs to be assessed for each individual patient. In fact, as reported for LTPs,³¹ different patients recognizing the same allergen can have individual patterns of IgE recognizing the entire or partial panel of antigenic epitopes borne by the analyzed allergen.

Nowadays, the use of forefront technologies, combined with the precision medicine applied to allergology allows personalized diagnosis at the molecular level, thus providing us with information on the individual patterns of sensitizations revealed by detection of specific IgE. This approach is a very promising tool useful to increase our knowledge in the field of food allergy and can contribute to set a safe and nutritionally balanced diet for each allergic subject.

ACKNOWLEDGEMENTS

The authors are thankful for the traineeship 2018/2019 grant received from the Université 8 Mai 1945, Guelma, Algeria under the grant agreement 'Programme National exceptional 2019/2020' from the Ministry of Higher Education and Scientific Research of Algeria. This study was supported by Green and Circular Economy (GECE) project N.856, year 2019, funded by the Italian Ministry of University and Research (MUR).

CONFLICT OF INTEREST

AM as an ADL stakeholder, developed, manufactured and distributed the FABER test as a diagnostic service. IG as an employee of ADL, contributed to develop, manufacture and perform the FABER test. MAC received funding from ADL.

SUPPORTING INFORMATION

Supporting information may be found in the online version of this article.

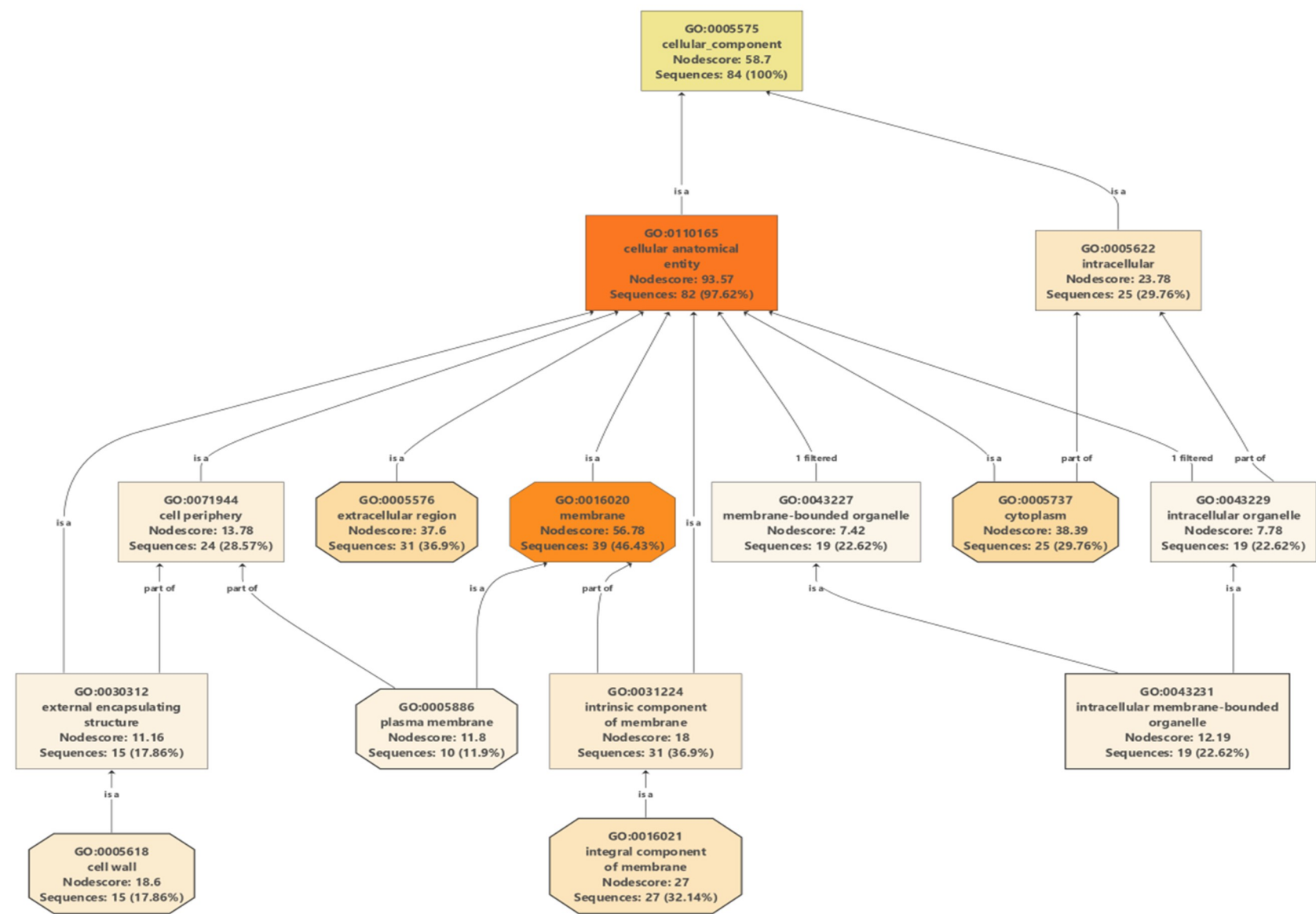
REFERENCES

- Chong SN and Chew FT, Epidemiology of allergic rhinitis and associated risk factors in Asia. *World Allergy Organ J* **11**:17 (2018).
- Oriel RC and Wang J, Diagnosis and management of food allergy. *Pediatr Clin North Am* **66**:941–954 (2019).
- Mehta H, Groetch M and Wang J, Growth and nutritional concerns in children with food allergy. *Curr Opin Allergy Clin Immunol* **13**: 275–279 (2013).
- Mehta H, Ramesh M, Feuille E, Groetch M, Rd MS and Wang J, Growth comparison in children with and without food allergies in 2 different demographic populations. *J Pediatr* **165**:842–848 (2014). <https://doi.org/10.1016/j.jpeds.2014.06.003>.
- Kapingidza AB, Pye SE, Hyde N, Dolamore C, Pote S, Schlachter CR *et al.*, Comparative structural and thermal stability studies of Cuc m 2.0101, art v 4.0101 and other allergenic profilins. *Mol Immunol* **114**:19–29 (2019).
- Ferreira F, Hirtenlehner K, Jilek A, Godnik-Cvar J, Breiteneder H, Grimm R *et al.*, Dissection of immunoglobulin E and T lymphocyte reactivity of isoforms of the major birch pollen allergen Bet v 1: potential use of hypoallergenic isoforms for immunotherapy. *J Exp Med* **183**:599–609 (1996).
- Breiteneder H and Ebner C, Molecular and biochemical classification of plant-derived food allergens. *J Allergy Clin Immunol* **106**:27–36 (2000).
- Alessandri C, Ferrara R, Bernardi ML, Zennaro D, Tuppo L, Giangrieco I *et al.*, Molecular approach to a patient's tailored diagnosis of the oral allergy syndrome. *Clin Transl Allergy* **10**:22 (2020).
- Sampson HA, Medical history. *J Allergy Clin Immunol* **103**:981–988 (1999).
- Scheurer S and Schülke S, Interaction of non-specific lipid-transfer proteins with plant-derived lipids and its impact on allergic sensitization. *Front Immunol* **9**:1389 (2018).
- Pasquariello MS, Palazzo P, Tuppo L, Liso M, Petriccione M, Rega P *et al.*, Analysis of the potential allergenicity of traditional apple cultivars by multiplex biochip-based immunoassay. *Food Chem* **135**:219–227 (2012).
- Tuppo L, Alessandri C, Giangrieco I, Ciancamerla M, Rafaiani C, Tamburrini M *et al.*, Isolation of cypress gibberellin-regulated protein: analysis of its structural features and IgE binding competition with homologous allergens. *Mol Immunol* **114**:189–195 (2019).
- Ciardiello MA, Palazzo P, Bernardi ML, Carratore V, Giangrieco I, Longo V *et al.*, Biochemical, immunological and clinical characterization of a cross-reactive nonspecific lipid transfer protein 1 from mulberry. *Allergy Eur J Allergy Clin Immunol* **65**:597–605 (2010).
- Apostol I, Iorga S, Mosoiu C, Racovita RC and Oana Mihaela Niculae GV, Alfalfa concentrate – a RICH source of nutrients for use in food products. *J Int Sci Publ* **5**:66–73 (2017).
- Mielmann A, The utilisation of lucerne: a review. *Br Food J* **115**:590–600 (2013).
- Gawel E, Grzelak M and Janyszek M, Lucerne (*Medicago sativa* L.) in the human diet—case reports and short reports. *J Herb Med* **10**:8–16 (2017).
- Hawkins C and Yu LX, Recent progress in alfalfa (*Medicago sativa* L.) genomics and genomic selection. *Crop J* **6**:565–575 (2018).
- Watson BS, Lei Z, Dixon RA and Sumner LW, Proteomics of *Medicago sativa* cell walls. *Phytochemistry* **65**:1709–1720 (2004).
- Gutsch A, Zouaghi S, Renaut J, Cuyper A, Hausman JF and Sergeant K, Changes in the proteome of *Medicago sativa* leaves in response to long-term cadmium exposure using a cell-wall targeted approach. *Int J Mol Sci* **19**(9):2498 (2018).
- Chen L, Chen Q, Zhu Y, Hou L and Mao P, Proteomic identification of differentially expressed proteins during alfalfa (*Medicago sativa* L.) flower development. *Front Plant Sci* **7**:1502 (2016).
- Zhang C and Shi S, Physiological and proteomic responses of contrasting alfalfa (*Medicago sativa* L.) varieties to peg-induced osmotic stress. *Front Plant Sci* **9**:242 (2018).
- Long R, Gao Y, Sun H, Zhang T, Li X, Li M *et al.*, Quantitative proteomic analysis using iTRAQ to identify salt-responsive proteins during the germination stage of two *Medicago* species. *Sci Rep* **8**:9553 (2018).
- Gao Y, Long R, Kang J, Wang Z, Zhang T, Sun H *et al.*, Comparative proteomic analysis reveals that antioxidant system and soluble sugar metabolism contribute to salt tolerance in alfalfa (*Medicago sativa* L.) leaves. *J Proteome Res* **18**:191–203 (2019).
- Sui X, Wu Q, Chang W, Fan X and Song F, Proteomic analysis of the response of *Funnellifor mismosseae/Medicago sativa* to atrazine stress. *BMC Plant Biol* **18**:289 (2018). <https://doi.org/10.1186/s12870-018-1492-1>.
- Suliaman FA, Holmes WF, Kwick S, Khouri F and Ratard R, Pattern of immediate type hypersensitivity reactions in the Eastern Province, Saudi Arabia. *Ann Allergy Asthma Immunol* **78**:415–418 (1997).
- Bener A, Safa W, Abdulhalik S and Lestrangier GG, An analysis of skin prick test reactions in asthmatics in a hot climate and desert environment. *Allergy Immunol* **34**:281–286 (2002).
- Jensen LB, Pedersen MH, Skov PS, Poulsen LK, Bindeslev-Jensen C, Andersen SB *et al.*, Peanut cross-reacting allergens in seeds and sprouts of a range of legumes. *Clin Exp Allergy* **38**:1969–1977 (2008).
- Pomés A, Davies JM, Gadermaier G, Hilger C, Holzhauser T, Lidholm J *et al.*, WHO/IUIS allergen nomenclature: providing a common language. *Mol Immunol* **100**:3–13 (2018).
- Mari A, Scala E, Palazzo P, Ridolfi S, Zennaro D and Carabella G, Bioinformatics applied to allergy: allergen databases, from collecting sequence information to data integration. The Allergome platform as a model. *Cell Immunol* **244**:97–100 (2006).
- Conesa A and Gotz S, Blast2GO: a comprehensive suite for functional analysis in plant genomics. *Int J Plant Genomics* **12**:1–12 (2008). <https://doi.org/10.1155/2008/619832>.
- Bernardi ML, Giangrieco I, Camardella L, Ferrara R, Palazzo P, Panico MR *et al.*, Allergenic lipid transfer proteins from plant-derived foods do not immunologically and clinically behave homogeneously: the kiwifruit LTP as a model. *PLoS One* **6**:e27856 (2011).

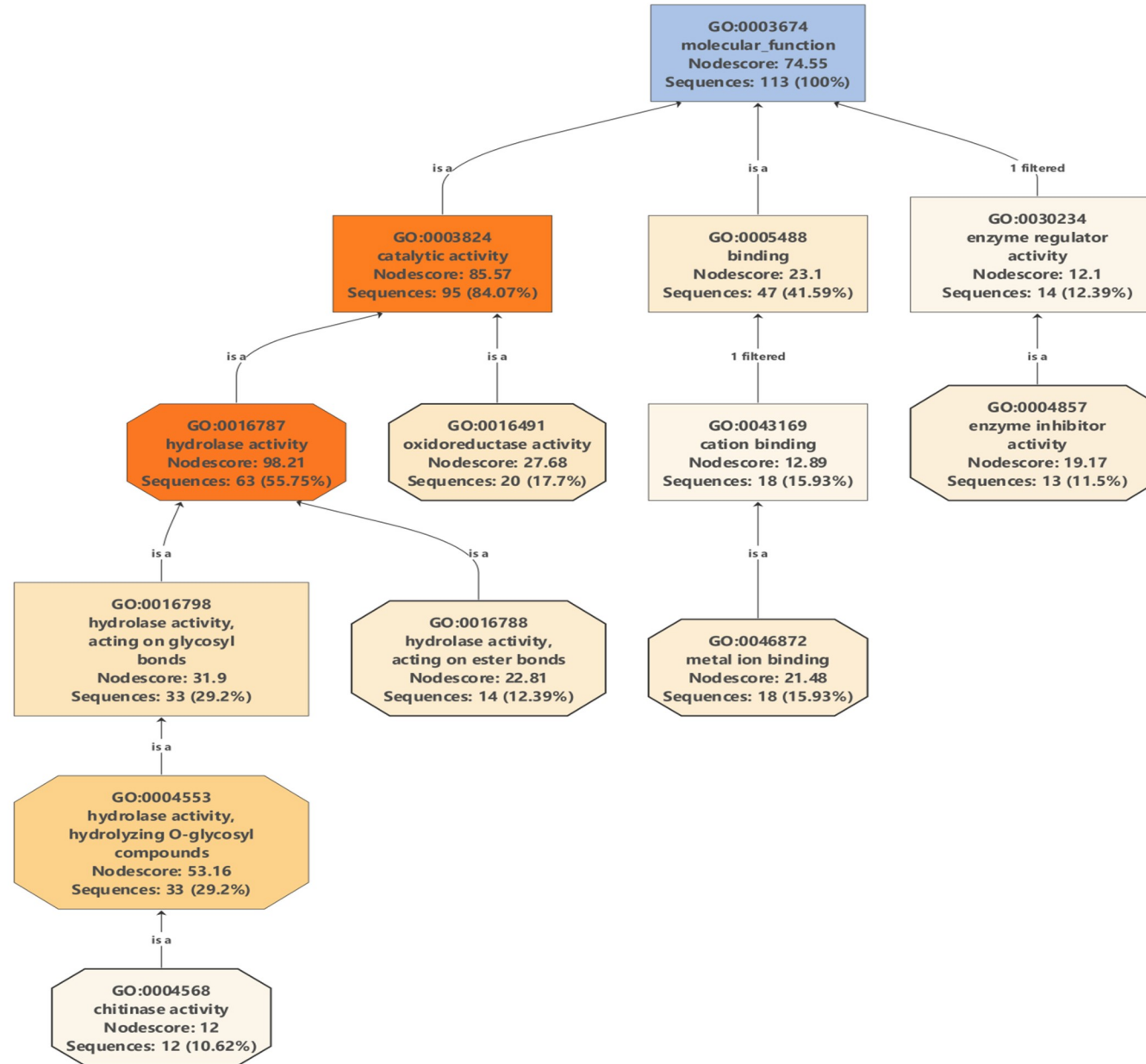
- 32 Giangrieco I, Alessandri C, Rafaiani C, Santoro M, Zuzzi S, Tuppo L *et al.*, Structural features, IgE binding and preliminary clinical findings of the 7 kDa lipid transfer protein from tomato seeds. *Mol Immunol* **66**:154–163 (2015).
- 33 Alessandri C, Ferrara R, Bernardi ML, Zennaro D, Tuppo L, Giangrieco I *et al.*, Diagnosing allergic sensitizations in the third millennium: why clinicians should know allergen molecule structures. *Clin Transl Allergy* **7**:21 (2017).
- 34 Giangrieco I, Ricciardi T, Alessandri C, Farina L, Crescenzo R, Tuppo L *et al.*, ENEA, a peach and apricot IgE-binding protein cross-reacting with the latex major allergen Hev b 5. *Mol Immunol* **112**:347–357 (2019).
- 35 Tuppo L, Giangrieco I, Alessandri C, Ricciardi T, Rafaiani C, Ciancamerla M *et al.*, Pomegranate chitinase III: identification of a new allergen and analysis of sensitization patterns to chitinases. *Mol Immunol* **103**:89–95 (2018).
- 36 Shevchenko A, Tomas H, Havliš J, Olsen JV and Mann M, In-gel digestion for mass spectrometric characterization of proteins and proteomes. *Nat Protoc* **1**:2856–2860 (2007).
- 37 Käll L, Canterbury JD, Weston J, Noble WS and Maccoss MJ, Semi-supervised learning for peptide identification from shotgun proteomics datasets. *Nat Methods* **4**:923–925 (2007).
- 38 Compare Database. Comp. Database 2019. <https://comparedatabase.org/> [16 December 2019].
- 39 Sahu SS, Loaiza CD and Kaundal R, Plant-mSubP: a computational framework for the prediction of single and multi-target protein sub-cellular localization using integrated machine-learning approaches. *AoB Plants* **12**:1–10 (2019).
- 40 Usui Y, Nakase M, Hotta H, Urisu A, Aoki N, Kitajima K *et al.*, A 33-kDa allergen from rice (*Oryza sativa* L. *japonica*): cDNA cloning, expression, and identification as a novel glyoxalase I. *J Biol Chem* **276**: 11376–11381 (2001).
- 41 Smole U, Bublin M, Radauer C, Ebner C and Breiteneder H, Mal d 2, the Thaumatin-like allergen from apple, is highly resistant to gastrointestinal digestion and thermal processing. *Int Arch Allergy Immunol* **147**:289–298 (2008).
- 42 Franco FP, Dias RO, Toyama D, Henrique-Silva F, Moura DS and Silva-Filho MC, Structural and functional characterization of PR-4 SUGAR-WINs from sugarcane and their role in plant defense. *Front Plant Sci* **9**:1916 (2019).
- 43 Berthelot K, Peruch F and Lecomte S, Highlights on *Hevea brasiliensis* (pro)hevein proteins. *Biochimie* **127**:258–270 (2016).
- 44 Ciardiello MA, Tamburrini M, Liso M, Crescenzo R, Rafaiani C and Mari A, Food allergen profiling: a big challenge. *Food Res Int* **54**:1033–1041 (2013).
- 45 Cerecedo I, Zamora J, Fox M, Voordouw J, Plana N, Rokicka E *et al.*, The impact of double-blind placebo-controlled food challenge (DBPCFC) on the socioeconomic cost of food allergy in Europe. *J Investig Allergol Clin Immunol* **24**:418–424 (2014).
- 46 Asero R, Fernandez-Rivas M, Knulst AC and Brujnzeel-Koomen CA, Double-blind, placebo-controlled food challenge in adults in everyday clinical practice: a reappraisal of their limitations and real indications. *Curr Opin Allergy Clin Immunol* **9**:379–385 (2009).
- 47 Slavokhotova AA, Shelenkov AA, Andreev YA and Odintsova TI, Hevein-like antimicrobial peptides of plants. *Biochemistry* **82**: 1659–1674 (2017).
- 48 Ladics GS, Assessment of the potential allergenicity of genetically-engineered food crops. *J Immunotoxicol* **16**:43–53 (2019).
- 49 Lauer I, Dueringer N, Pokoj S, Rehm S, Zoccatelli G, Reese G *et al.*, The non-specific lipid transfer protein, Ara h 9, is an important allergen in peanut. *Clin Exp Allergy* **39**:1427–1437 (2009).
- 50 Tuppo L, Spadaccini R, Alessandri C, Wienk H, Boelens R, Giangrieco I *et al.*, Structure, stability, and IgE binding of the peach allergen Pea-maclein (Pru p 7). *Biopolymers* **102**:416–425 (2014).

Figure S1. The node-directed acyclic graph (DAG) of the Gene Ontology (GO) graphs. **A:** represents the major functional categories of cellular component (CC), **B:** stands for the major functional categories of molecular function (MF) and **C:** is the graph showing the major functional categories of biological processes (BP).

A



B



C

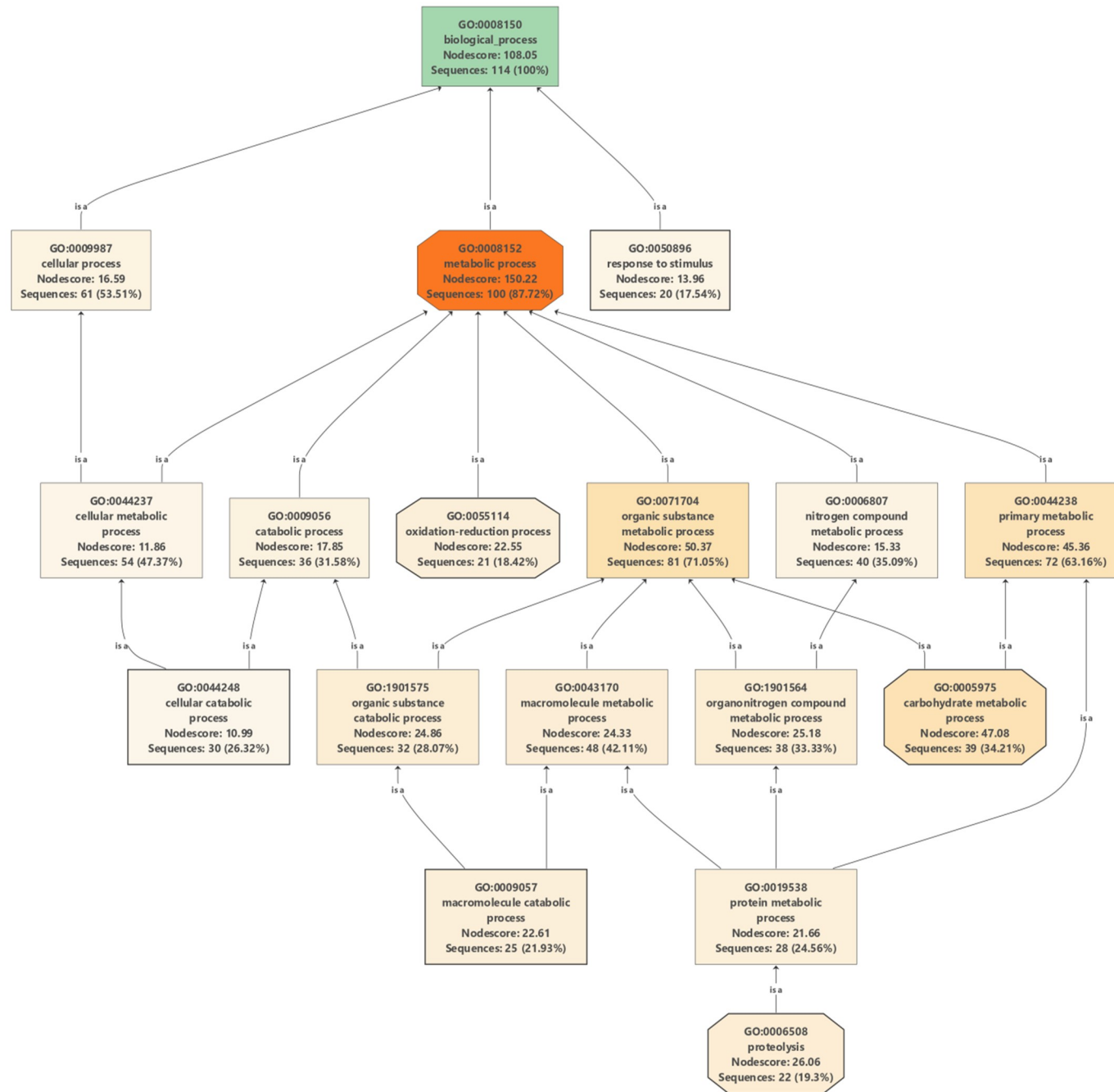


Table S2. List of top 20 ranking proteins identified in the leaf protein extract of *M. sativa*. Accession numbers according to UniProt database, “Taxa” shows the Medicago taxonomy associated to the identified protein in Byonic™ search engine using the UniProt database. MEDTR and MEDSA stand for *Medicago truncatula* and *Medicago sativa*, respectively. Protein names were assigned based on functional annotation limiting the blast search to green plant taxonomy. “Cellular localization” was predicted using the plant subcellular location predictor (<http://bioinfo.usu.edu/Plant-mSubP/>). “PM” indicates plasma membrane. Gene ontology terms “Biological process” and “Molecular functions” were determined by the InterPro functional analysis tool in Blast2GO.

UniProt Accession	Taxa	Protein Name	Cellular localization	Biological process	Molecular function
G7JQL4	MEDTR	Glucan endo-1,3-beta-glucosidase	PM	carbohydrate metabolic process	hydrolase activity, hydrolyzing O-glycosyl compounds
A0A1L2BU68	MEDSA	Class III-1 chitinase 10 (Fragment)	Plastid	carbohydrate metabolic process	hydrolase activity, hydrolyzing O-glycosyl compounds
A0A396GNA8	MEDTR	Endochitinase	Cell wall	Carbohydrate metabolic process. Chitin catabolic process. Cell wall macromolecule catabolic process.	chitinase activity chitin binding
G7IMV1	MEDTR	Alpha-L-arabinofuranosidase/beta-D-xylosidase	Cell wall	carbohydrate metabolic process	hydrolase activity, hydrolyzing O-glycosyl compounds
I3T3C6	MEDTR	LRRNT_2 domain-containing protein	Plastid		protein binding
C3VM17	MEDSA	Chitinase class III-1	Vacuole	carbohydrate metabolic process	hydrolase activity, hydrolyzing O-glycosyl compounds
A0A1L2BU75	MEDSV	Beta-1,3-glucanase 3 (Fragment)	Vacuoles	carbohydrate metabolic process	hydrolase activity, hydrolyzing O-glycosyl compounds
G7IMM5	MEDTR	FAD-binding berberine family protein	Cell wall	oxidation-reduction process	oxidoreductase activity FAD binding
I3SK49	MEDTR	Plant basic secretory protein (BSP) family protein	Plastid		
O48904	MEDSA	Malate dehydrogenase	Mitochondria	carbohydrate metabolic process tricarboxylic acid cycle	L-malate dehydrogenase activity

				malate metabolic process	
G7K4B9	MEDTR	Glycerophosphoryl diester phosphodiesterase family protein	PM	lipid metabolic process tubulin complex assembly post-chaperonin tubulin folding pathway	glycerophosphodiester phosphodiesterase activity beta-tubulin binding
I3STX0	MEDTR	Glycoside hydrolase family 18 protein	Extracellular	carbohydrate metabolic process chitin catabolic process	chitinase activity chitin binding
A0A072TUC3	MEDTR	Pathogenesis-related thaumatin family protein	Extracellular		
A0A396K0C6	MEDTR	Putative nepenthesin	PM	proteolysis	aspartic-type endopeptidase activity
Q5RLX9	MEDSA	Chitinase	Cell wall	polysaccharide catabolic process chitin catabolic process defense response cell wall macromolecule catabolic process	chitinase activity chitin binding
G7JQL1	MEDTR	Glucan endo-1,3-beta-glucosidase	PM	carbohydrate metabolic process	glucan exo-1,3-beta-glucosidase activity glucan endo-1,3-beta-D-glucosidase activity
A0A072VNF7	MEDTR	Putative tripeptidyl-peptidase II	PM	proteolysis	serine-type endopeptidase activity
A0A396GUY5	MEDTR	Putative thaumatin	Extracellular	defense response	
A0A072VD04	MEDTR	Beta-galactosidase	Extracellular	carbohydrate metabolic process	beta-galactosidase activity carbohydrate binding

A0A396JDT6	MEDTR	Putative nepenthesin	PM	proteolysis; protein catabolic process	aspartic-type endopeptidase activity
------------	-------	-------------------------	----	---	--

Table S4. Plant, animal food, pollen and latex extracts and allergens probed with the SPHIAa test using sera containing IgE recognizing them. Allergenic extracts were obtained from the natural source. Natural and recombinant allergenic proteins are labeled with n and r, respectively. Components of the same protein family are grouped.

Allergenic extracts from fruits and vegetables (IgE inhibition is shown in figure 3)			
Extracts	Allergen name	Source	Tissue
	Act c	Gold kiwifruit	Fruit
	Act d	Green kiwifruit	Fruit
	Cuc m	Melon canatloupe	Fruit
	Cuc s	Cucumber	Fruit
	Fra a	Strawberry	Fruit
	Mal d	Apple	Fruit
	Pers a	Avocado	Fruit
	Pru ar	Apricot	Fruit
	Pru p†	Peach peel	Fruit
	Pru p‡	Peach pulp	Fruit
	Pun g	Pomegranate	Fruit
	Sola l	Tomato	Fruit
	Sola m	Eggplant	Fruit
	Vit v	Grape	Fruit
	Beta v	Common beet	Leaf
	Lac s	Lettuce	Leaf
	Spi o	Spinach	Leaf
	Aspa o	Asparagus	Leaf
	Api g	Celery	Stalk
Extracts from seeds, bulbs and roots			
Extracts	Ana o	Cashew	Seed
	Ara h	Peanut	Seed
	Bar e	Brazil nut	Seed
	Cas s	Chestnut	Seed
	Cer si	Carob	Seed
	Gly m	Soy	Seed
	Hor v	Barley	Seed
	Jug r	Walnut	Seed
	Len c	Lentil	Seed
	Lin us	Flux	Seed
	Lup a	White lupine	Seed
	Ory s	Rice	Seed
	Pin p	Pine nut	Seed
	Pis v	Pistachio	Seed
	Pru du	Almond	Seed
	Ses i	Sesame	Seed
	Sin a	White mustard	Seed
	Sola l	Tomato	Seed
	Tri a	Wheat	Seed
	Tri tp	kamut	Seed
	Zea m	Corn	Seed
	All c	Onion	Bulb
	All p	Leek	Bulb

	All s	Garlik	Bulb
	Foe v	Fennel	Bulb
	Sola t	Potato	Tuber
	Dau c	Carrot	Root
Allergenic molecules from plant food (IgE inhibition is shown in figure 4)			
9K-LTP	Allergen name	Source	tissue
	nAct d 10	Green kiwifruit	Seed
	nAra h 9	Peanut	Seed
	nCor a 8	Hazelnut	Seed
	nJug r 3	Walnut	Seed
	nPru p 3	Peach	Fruit
	nPun g 1	Pomegranate	Fruit
	nZea m 14	Maize	Seed
7K-LTP	nSola l 6	Tomato	Seed
Actinidin (cystein protease)	nAct d 1	Green kiwifruit	Fruit
Papain (cystein protease)	nCar p Papain	papaya	Fruit
Thaumatococcus-like protein	nAct d 2	Green kiwifruit	Fruit
Classe IV chitinase	nAct c Chitinase_IV	Gold kiwifruit	Fruit
7S vicilin	nAra h 1	Peanut	Seed
2S albumin	nAra h 2	Peanut	Seed
	rCor a 14	Hazelnut	Seed
11S globulin	nAra h 3	Peanut	Seed
	nAra h 6	Peanut	Seed
	nCor a 9	Hazelnut	Seed
Bet v 1-like	rAra h 8	Peanut	Seed
	rMal d 1	Apple	Fruit
Lectin	nAra h agglutinine	Peanut	Seed
GRP	nPun g 7	Pomegranate	Fruit
Trypsin inhibitor	nGly m TI	Soybean	Seed
Hydrophobic peptide	nGly m 1	Soybean	Seed
Patatin	nSola t 1	potatoes	tuber
Extracts and allergens from animal food			
	Allergen name	Source	Tissue
Extracts	Bos d	Cow	Milk
	Bos d	Cow	Meat
	Gal d	Chicken	Egg white
	Gal d	Chicken	Egg yolk
	Sol so	Common sol	Muscle
	Sar m	Sardine	Muscle
Allergens	nBos d 4	Cow	Milk
	nBos d 5	Cow	Milk
	nBos d 6	Cow	milk, muscle, serum

	nBos d 8	Cow	Milk
	nBos d gelatin	Cow	Meat, bone
	nGal d 1	Chicken	Egg
	nGal d 2	Chicken	Egg
	nGal d 4	Chicken	Egg
	nGal d 5	Chicken	Egg, serum
	nMer mr 1	Hake	Muscle
Pollen and latex extracts (IgE inhibition is shown in figure 3)			
Extracts	Allergen name	Source	Tissue
	Amb a	mugwort	Pollen
	Bet v	Birch	Pollen
	Cor a	Hazelnut	Pollen
	Lol p	Rye grass	Pollen
	Ole e	Olive tree	Pollen
	Par j	Parietaria	Pollen
	Phel p	Timothy grass	Pollen
	Que a	White oak	Pollen
	Hev b	Rubber tree	Latex
Pollen and latex allergens (IgE inhibition is shown in figure 4)			
Protein name	Allergen name	Source	Tissue
Pectase lyase	nAmb a 1	ragweed	Pollen
Pectase lyase	nCup a 1	cypress	Pollen
Bet v 1	rBet v 1	Birch	Pollen
Expansin	rLol p 1	Rye grass	Pollen
Profilin	rBet v 2	Birch	Pollen
Profilin	rMer a 1	Annual mercury	Pollen
Trypsin inhibitor	nOle e 1	Olive tree	Pollen
Super oxide dismutase	rHev b 10	Rubber tree	Latex
Class I chitinase	rHev b 11	Rubber tree	Latex
Hev b 5	rHev b 5	Rubber tree	Latex
Hevein	rHev b 6	Rubber tree	Latex
Profilin	rHev b 8	Rubber tree	Latex

Table S5 Percentage of identity between LTPs from *M. sativa* (MEDSA), *M. truncatula* (MEDTR), corn Zea m 14, peanut Ara h 9, pomegranate Pun g 1, walnut Jug r 3, hazelnut Cor a 8, peach Pru p 3 and green kiwifruit Act d 10. Their UniProt accession numbers are Q45NK7, A0A396I974, Q2XX25, B6CEX8, A0A059SSZ0, C5H617, Q9ATH2, Q8H2B2 and P85206, respectively.

[illegible]