République Algérienne Démocratique et Populaire
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique
Université 8 Mai 1945 Guelma

Faculté Sciences de la nature et de la vie sciences de la terre et de l'univers
Département Biologie
Laboratoire de domiciliation Laboratoire de Biologie, Eau et Environnement (LBEE)

THÈSE

EN VUE DE L'OBTENTION DU DIPLOME DE DOCTORAT EN 3ème CYCLE

Domaine : Sciences de la nature et de la vie Filière : Sciences biologiques
Spécialité : Biochimie appliquée

Présentée par

YAKHLEF Marwa

Intitulée

Evaluation of the allergenic risk of proteins of food origin

Soutenue le : 13/01/2022 Devant le Jury composé de :

Nom et Prénom	Grade		
Mr. BENOUARETH Djamel Eddine	Professeur	Univ. de Guelma	Président
Mme. SOUIKI SOUMATI Lynda	Professeur	Univ. de Guelma	Encadreur
Mr. LADJAMA Ali	Professeur	Univ. d'Annaba	Examinateur
Mr. MESSARAH Mahfoud	Professeur	Univ. d'Annaba	Examinateur
Mme. GRARA Nedjoud	Professeur	Univ. de Guelma	Examinatrice

Année Universitaire : 2021/2022

Acknowledgements

First, thanks to Almighty Allah for giving me strength, ability and courage to achieve this work and complete this thesis.

I would like to thank all the jury members, Prof. BENOUARETH Djamel Eddine (President of the jury), Prof. LADJAMA Ali (examiner of the doctorate thesis), Prof. MESSARAH Mahfoud (examiner of the doctorate thesis), and Prof. GRARA Nedjoud (Examiner of the doctorate thesis) for their interest in my work.

I wish to express my sincere thanks to my supervisor Prof. SOUIKI SOUMATI Lynda for introducing me to such an interesting field of science, for providing me with the opportunity to carry out this study, and for her advice, patience and guidance throughout this work.

I would like to thank the Faculté des Sciences de la nature et de la vie sciences de la terre et de l'univers, Université 8 Mai 1945 Guelma and the ministry of the higher education and the scientific research for giving me the opportunity through Programme national exceptional (PNE) to achieve this work.

I wish to express my deepest gratitude to Dr. Gabriella POCSFALVI for accepting me in her laboratory and giving me the opportunity to work with her exceptional team. I had the chance to learn a lot of things from her on both professional and personal level.

My sincere thanks also go to Dr Maria Antonietta CIARDIELLO for accepting me in her laboratory to do the in vitro test and for all the help with my work.

Many thanks go to Miss Leila ABBAS technician of the laboratory Biologie, Eau et Environnement (LBEE) for all her help during this work, for her support advices and availability.

Many thanks go to Dr. Immacolata FIUME, for all her help and assistance during the manipulations, her explanations, scientific and personal support.

Many thanks to Dr. Ivana GIANGRIECO who helped me to conduct laboratory tests related to my research.

Abstract

Food allergy is a dangerous health issue worldwide especially in children. It is causing malnutrition problems and multiple deficiencies due to the eviction diet followed.

In our study, we worked on 3 distinct axes of food allergy. First, we started with a cross-sectional, random, questionnaire based survey in the schools of the city of Guelma (Algeria). Then, allergens were extracted from different foods: ovalbumin and ovomucoid from hen's egg, α- lactalbumin and β-lactoglobulin from cow's milk, protein M from sardines and gliadins from wheat. A biochemical characterization was done in order to find common characteristics of allergens. Finally, leaf protein extract of lucerne was studied using a combined proteomics, *in vitro* and *in silico* approach to suggest alfalfa as a protein substitute for allergic patients.

The results showed the prevalence of food allergy in schoolchildren to be 8.5% (95% CI 6.1–10.8). The top foods causing allergy were: Chocolate (1.7%), wheat (1.3%), milk (1.1%), eggs (1.1%), strawberry (1.1%), and fish (0.9%). The presence of family history influences the appearance of the allergy. On the other hand, the extracted allergens were characterized with an acidic pH. The multiple sequence alignment of the allergens didn't reveal any sequence similarity except for the allergens from the same food. The combination of *in silico* and *in vitro* allowed the detection of proteins in alfalfa leaves sharing similar sequences or cross-reacting with plant allergens from three different allergen families such as lipid transfer protein, thaumatin-like and Bet v 1-like protein families.

To conclude, the main risk factor leading to the appearance of food allergy was family history. The allergens can have some common characteristics like the acidic pH. However, no significant amino acid sequence was found to be similar between the studied allergens. The use of *in silico* and *in vitro* approach predicted the allergenicity of alfalfa. This plant can be safely introduced as a protein-rich supplement in the diet of patients allergic to animal food allergens.

Keywords: Food allergy, Allergens, Prevalence, Sequence alignment, *Medicago sativa.* IgE inhibition assay.

Résumé

L'allergie alimentaire est un problème de santé préoccupant dans le monde entier, en particulier chez les enfants. Elle engendre des problèmes de malnutrition et des carences multiples à cause du régime d'éviction suivi.

Dans notre projet, nous avons travaillé sur l'allergie alimentaire. Tout d'abord, une enquête transversale, aléatoire, par questionnaire dans les écoles de la ville de Guelma (Algérie) a été réalisée. Ensuite, des allergènes ont été extraits de différents aliments : l'ovalbumine et l'ovomucoïde de l'oeuf de poule, l'α-lactalbumine et la β-lactoglobuline du lait de vache, la protéine M de la sardine et les gliadines du blé. Une caractérisation biochimique a été faite afin de trouver des caractéristiques communes entre ces allergènes. Enfin, l'extrait protéique des feuilles de luzerne a été étudié en utilisant une approche protéomique, *in vitro* et *in silico* combinée pour suggérer la luzerne comme substitut protéique pour les patients allergiques.

Les résultats ont montré que la prévalence de l'allergie alimentaire chez les écoliers est de 8,5 % (IC à 95 % 6,1 à 10,8). Les principaux aliments en cause sont : le chocolat (1,7 %), le blé (1,3 %), le lait (1,1 %), les œufs (1,1 %), les fraises (1,1 %) et le poisson (0,9 %). La présence d'antécédents familiaux influence l'apparition de l'allergie. De l'autre part, les allergènes extraits sont caractérisés par un pH acide. L'alignement des séquences multiples des allergènes n'a révélé aucune similitude de séquence, à l'exception des allergènes du même aliment. La combinaison entre l'approche *in silico* et *in vitro* a permis la détection des séquences similaires ou ayant des réactions croisées entre les protéines de luzerne et des allergènes végétaux de quelques familles d'allergènes telles que les protéines de transfert de lipides, les familles de protéines de type thaumatine et Bet v 1.

En conclusion, le principal facteur de risque conduisant à l'apparition de l'allergie alimentaire était les antécédents familiaux. Les allergènes peuvent avoir des caractéristiques communes comme le pH acide. Cependant, aucune séquence commune d'acides aminés n'a été trouvée entre les allergènes étudiés. L'utilisation de l'approche *in silico* et *in vitro* a prédit l'allergénicité de la luzerne. Cette plante peut être introduite en toute sécurité comme complément riche en protéine dans l'alimentation des patients allergiques aux allergènes alimentaires d'origine animale.

Mots clés : Allergie alimentaire, Allergènes, Prévalence, Alignement de séquences,

Medicago sativa. Test d'inhibition des IgE.

الملخص

حساسية الطعام هي مشكلة صحية خطيرة في جميع أنحاء العالم خاصة عند الأطفال. تسبب مشاكل سوء التغذية بسبب نظام الإخلاء المتبع.

في دراستنا ، عملنا على 3 محاور متميزة للحساسية الغذائية. اولا ، بدأنا بمسح مقطعي عشوائي عن طريق الاستبيان في مدارس مدينة قالمة (الجزائر). ثانيا، تم استخلاص المواد المسببة للحساسية من أطعمة مختلفة: الزلال البيضاوي والمخاط البويضي من بيض الدجاج ، ألفا لاكتالبومين وبيتا لاكتوجلوبولين من حليب البقر ، بروتين M من السردين والجليدين من القمح. تم إجراء توصيف كيميائي حيوي من أجل إيجاد خصائص مشتركة بين هذه المواد المسببة للحساسية. أخيرًا ، تمت دراسة مستخلص البروتين من أوراق البرسيم باستخدام نهج بروتيني مركب ، في المختبر وفي السيليكو لاقتراح البرسيم كبديل للبروتين لمرضى الحساسية.

أظهرت النتائج أن انتشار حساسية الطعام لدى أطفال المدارس كان 8.5٪ (95٪ CI 6.1 إلى 10.8). كانت الأطعمة الرئيسية: الشوكولاته (1.7٪) ، القمح (1.3٪) ، الحليب (1.1٪) ، البيض (1.1٪) ، الفراولة (1.1٪) والأسماك (0.9٪). يؤثر وجود تاريخ عائلي على ظهور الحساسية. من ناحية أخرى ، تميزت المواد المسببة للحساسية المستخرجة بدرجة حموضة حمضية. لم تكشف محاذاة التسلسلات المتعددة لمسببات الحساسية عن أي تشابه في التسلسل باستثناء المواد من نفس الطعام. سمح الجمع بين نهج في السيليكو وفي المختبر باكتشاف متواليات متشابهة أو تفاعلية بين بروتينات البرسيم ومسببات الحساسية النباتية لبعض عائلات المواد المسببة للحساسية مثل بروتينات نقل الدهون وعائلات البرسيم الحجازي والبروتينات الشبيهة بالثوماتين و Bet v 1.

في الختام ، كان عامل الخطر الرئيسي الذي أدى إلى ظهور حساسية الطعام هو التاريخ العائلي. يمكن أن يكون لمسببات الحساسية خصائص مشتركة مثل درجة الحموضة الحمضية. ومع ذلك ، لم يتم العثور على تسلسل مشترك للأحماض الأمينية بين المواد المسببة للحساسية المدروسة. توقع استخدام النهج في السيليكو وفي المختبر حساسية البرسيم. يمكن إدخال هذا النبات بأمان كمكمل غذائي عالي البروتين في النظام الغذائي للمرضى الذين يعانون من الحساسية تجاه المواد الغذائية المسببة للحساسية من أصل حيواني.

الكلمات الرئيسية: حساسية الطعام ، مسببات الحساسية ، الانتشار ، محاذاة التسلسل ،البرسيم الحجازي، اختبار تثبيط IgE.

Summary

Summary

III. Chapter 3. Evaluation of potential allergenicity of protein extract from *Medicagosativa*

List of tables

List of figures

List of abbreviations

2D-DIGE MALDI-TOF: two-Dimensional Difference In Gel Electrophoresis Matrix-Assisted Laser Desorption/Ionization-Time Of Flight

ACN: Acetonitril

Act c: Golden kiwi fruit

Act d: Green kiwifruit

ALA: α-Lactalbumin

Amb a: Mugwort

AMBIC: Ammonium bicarbonate

Ana o: Cashew

Api g: Celery

Ara h: Peanut

Aspa o: Asparagus

Ber e: Brazil nut

Bet v: Birch

Beta v: Common beet

BLA: β-Lactoglobulin

Bos d:Bovine

BP: biological processes

CAN: Ceric ammonium nitrate

Cas s: Chestnut

CC: cellular component

Cer si: Carob

CI: confidence interval

CM: carboxymethyl

CMPA: Cow‘s Milk Proteins Allergy

Codex: Food Code

Cor a: Hazelnut

Cuc m: Melon

Cuc s: cucumber

DEAE-C: DiEthylAminoEthyl Cellulose

DTT: DiThioThreitol

EFSA: European Food Safety Authority

FA: Formic Acid

FABER: Friendly Allergens Nano-Bead Array

FAO: Food and Agriculture Organization

Fra a: Strawberry

Gal d: Chicken

GLIA: Gliadins

Gly m: Soy

GO: Gene Ontology

GRP: Gibberellin Regulated Proteins

Hev b: Rubber tree

Hor v: Barley

HPLC/MS/MS: High Performance Liquid Chromatography and Tandem Mass Spectrometry

IDA: Information-Dependent Acquisition

INFOSAN: International Food Safety Authorities Network

IOVO_CHICK: Ovomucoid

iTRAQ: Isobaric Tag for Relative and Absolute Quantitation

IUIS: International Union of Immunological Society

Jug r: Walnut

Lac s: Lettuce

LACB_BOVIN: β-Lactoglobulin

LALBA_BOVIN: α-Lactalbumin

LDS: low density solvent

Len c: Lentil

Lin us: Flux

Lol p: Rye grasse

LTP: Lipid Transfer Protein

Lup a: White lupine

M/F: Male/Female

Mal d: Apple

MEDSA: *Medicago sativa*

MEDTR: *Medicago truncatula*

MF: Molecular Function

MOPS: 3-(*N*-Morpholino) propanesulfonic acid

MW: Molecular weight

MWCO: Molecular weight cut off

OFC: Oral Food Challenge

Ole e: Olive tree

Ory s: Rice

OVA: Ovalbumin

OVAL_CHICK: Ovalbumin

OVO:Ovomucoid

Par j: Parietaria

Pers a: Avocado

Phl p: Timothy grass

pI: Isoelectric point

Pin p: Pine nut

Pis v: Pistachio

PPT: Prick-by-Prick Test

PR-10 family: Pathogenesis-Related

Pru ar: Apricot

Pru du: Almond

Pru p: Peach peel

Pun g: Pomegranate

PV: Parvalbumin

Que a: white oak

RAST: RadioAllergoSorbant Test

R_f: Ratio frontal

Sar m: Sardine

SD: Standard Deviation

SDS-PAGE: Sodium Dodecyl Sulphate–PolyAcrylamide Gel Electrophoresis

Ses i: Sesame

Sin a: White mustard

Sol so: Common sol

Sola l: Tomato

Sola m: Eggplant

SP: Sulfopropyl

SPHIAa: Single Point Higiest Inhibition Acheivement assay

Spi o: Spinach

SPT: Skin Prick Test

Tri a: Wheat

Tri tp: Kamut

Vit v: Grape

WHO: World Health Organization

Zea m: Maize

General introduction

General introduction

According to the world's health organization, 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 (Chong and Chew, 2018). During these last twenty years, the number of people having allergies has doubled in number and of course food allergy is a part of this increase.

On 2006, the report from the world health Organization defined the food allergy as an adverse reaction to food that is involving the immune system or not. For instance allergy to milk, eggs or fish is classified as IgE mediated food allergy. However in the case of celiac disease, it is classified as a non IgE mediated food allergy (World Health Organization (WHO), International Food Safety Authorities Network (INFOSAN) and Food and Agriculture Organization of the United Nations (FAO, 2006).

The mechanism of food allergy reaction covers all 4 types of the immune response of the classical division of Gell and Coombs. If the immune response is by type I reaction, the IgE antibodies are involved, and the clinical symptoms occur soon after ingestion of food or within a short period of time (usually about 2 h). The IgE antibodies play the main role in allergy, and in the healthy human they can be found in small quantities. Only in some people, especially with a hereditary risk of atopy, there is an excessive production of IgE antibodies which trigger allergic reactions. The presence of IgE antibodies in the pathogenesis of the disease is the main criterion for the definition of food allergy. IgE-dependent allergy in the primary form is especially frequent in children, in which foods are the main or the only cause of the disease or disorder in contrast to the secondary form, which is more typical of older children and adults, in which foods are one of the many factors causing the disorder. Adverse food-induced immune response can be a sign of other pathogenic mechanisms of allergic reaction – II, III or IV type when T cells, IgG, IgM, IgA, and other immunologically competent cells are involved. Due to the pathogenic mechanism of these reactions, they are described as IgE-independent (Zukiewicz-Sobczak WA *et al*., 2013).

Food allergy can manifest a wide range of symptoms. The appearance of itching of the lips or tongue, repeated vomiting, frequent diarrhea or urticaria may be defined as the most common symptoms, which reveal an allergic reaction to food. Fatigue caused by allergies can be felt the most in the morning and right after getting out of bed, or late in the afternoon, when any kind of rest or its length does not bring relief. Pain, stiffness and muscle tearing of

the shoulder, neck and back, which may occur with headache, may drag through the days and weeks. Psychiatric disorders (tension, nervousness, irritability, stubbornness, anxiety, confusion, nervousness combined with trembling, stammering, disorderly speech, lethargy, stupor, aphasia – loss of the ability to speak, and feeling dazed, depression, discouragement, melancholy) are common in allergic individuals (Zukiewicz-Sobczak WA *et al*., 2013).

Generally, there are two types of food allergy. The first type is characterized by immediate reaction, which is accompanied by symptoms occurring within a few minutes, or even seconds after consumption of the food, which is anaphylaxis (shock), urticaria, angioneurotic edema (skin swelling). Eggs, nuts, peanuts, fish and shellfish are often foods that often cause this type of allergy. The second type of food allergy is a late reaction, in which the symptoms (fatigue, irritability, depression, hyperactivity, insomnia, headache, poor concentration, paleness, itching limbs, involuntary bedwetting, asthma, colds, indigestion, colic, diarrhea, bloating and skin lesions) appear a few hours, and even a few days after food intake. Foods that cause this type of reaction are milk, chocolate, legumes, citrus and food additives. Because of this delay, it is difficult to determine what is the cause of food allergies (Zukiewicz-Sobczak WA *et al*., 2013)

Nevertheless of the reports available on the prevalence of food allergy all over the globe, it still not sufficient (Jorge *et al.*, 2017; Cabrera-Chávez *et al.*, 2018; Levin *et al.*, 2020). According to a survey of 83 World Allergy Organization member countries and six non-member countries, 57% of them had no data on food allergy prevalence, 25% had data based on patient/parent report, and only 10% had food allergy prevalence data based on oral food challenges (OFCs) (Prescott *et al.*, 2013).

To get an exact prevalence of food allergy, it is necessary to go through multiple clinical demonstrations like skin prick test, double-blind, placebo-controlled food challenges (DBPCFCs) or the detection of specific IgE. To do so, there must be adequate medical centers and specialized staff as well as the ethical aspect that should be respected (Messina and Venter, 2020). Due to the difficulties of this method, the main epidemiological studies on the prevalence of food allergy are usually based on self/ parent reported answers or telephone surveys. The determination of the prevalence of food allergy is extremely important. It helps to determine the most common food allergens in a certain population at a certain period of time. Consequently, the results will highlight on a medical issue that affects the quality of life of the allergic patients.

According to several studies, The list of major food allergenic sources worldwide includes: milk, eggs, wheat, soy, peanut, tree nuts, fish, shellfish with different extends and prevalences (Sampson *et al.*, 2018; Oriel and Wang, 2019). As reported in many works, the prevalence of food allergy depends on multiple factors like geographical location, nutrition habits and genetics (Poole *et al.*, 2006; Loh and Tang, 2018; Levin *et al.*, 2020).

Many studies are interested in the extraction and the physicochemical characterization of food allergens. This characterization is usually represented by the determination of the isoelectric point of the allergens, the molecular weight, the sequencing and the determination of the 3D structure of the allergen.

When it comes to the treatment of food allergy, the only option is the strict elimination of the food in question and the use of rescue medications like: Adrenaline if the reaction happens accidently (Ring, Klimek and Worm, 2018; Corica *et al.*, 2020). The eviction diet showed to cause multiple nutritional disorders especially in children (Venter *et al.*, 2017; Meyer *et al.*, 2019). The most common nutritional disorders include poor growth, micronutrient deficiencies and feeding difficulties (Meyer, 2018). In a cross sectional study in Singapore, it has been reported that children with food allergy are associated with poor growth compared to normal children and advised the early introduction of nutritional intervention (Chong *et al.*, 2018). In a review gathering the available evidence on children's growth having food allergy in the literature, it has been proved an impaired growth in children with food allergy. It is most likely caused by multifactorial origin (Pavić and Kolaček, 2017). Additionally, it was shown that, compared to healthy controls, children with food allergy were still smaller and lighter, even when no differences in energy and nutrient intakes were observed (Mehta, Groetch and Wang, 2013)

Consequently, it became crucial to search for new approaches either to reduce the risk related to allergens or to suggest new proteins for allergic patients. That is why; understanding the physicochemical properties of these proteins can help in the development of treatment techniques that can reduce the sensitivity to these proteins (Rahaman, Vasiljevic and Ramchandran, 2016). The choice of adequate processing methods needs a full understanding of the eventual changes that may occur to the treated proteins on both microscopic and macroscopic level as well as their gastrointestinal digestibility (Rahaman, Vasiljevic and Ramchandran, 2016)

This work was designed to investigate different aspects of food allergy in three different chapters. Each chapter treats a different aspect of food allergy.

Chapter 1: Contribution to the estimation of food allergy in the city of Guelma.

This chapter presents a cross sectional questionnaire based study having the following goals:

- Overview the current state of food allergies in schoolchildren in the city of Guelma by determining the prevalence of food allergy in this population.
- Highlight the most common foods causing allergy.
- Assess any correlation between this pathology and some of its risk factors (gender and family history).

Chapter 2: Extraction of some food allergens and determination of some of their biochemical characteristics.

The aims of this chapter are:

- The extraction of ovalbumin and ovomucoid from eggs, α-lactalbumin and β-lactoglobulin from cow's milk, protein M from sardine and gliadins from wheat.
- Determination of the isoelectric point and molecular weight of the allergens.
- Purification and Identification of the allergens through HPLC/MS/MS using Mascot search database.
- Multiple sequence alignments of the allergens to find a link between similar sequences and allergenicity.

Chapter 3: Evaluation of potential allergenicity of protein extract from *Medicago sativa*.

In this part we aim to propose the protein extract of *Medicago sativa* as a protein substitute for allergic patients. To achieve that, the objectives were:

- Make the extraction of proteins from the leaves of *Medicago sativa.*
- The use of proteomic techniques to identify proteins in the extract.
- Realization of *in silico* analysis in order to find potential allergens in the extract based on similar sequences with known allergens.
- The use of immunologic test based on the inhibition of IgE binding to find cross-reaction between the extract and antibodies generated against known allergens.

- Combine the results from both *in silico* and *in vitro* analysis to determine for which allergic patients this plant could be proposed.

Chapter 1

Contribution to the estimation of food allergy in the city of Guelma.

I.1. Introduction

Little information on the prevalence of food allergies is available despite the efforts underway to provide reliable data (Sha *et al.*, 2019). A rigorous estimation of the prevalence of food allergy would need to involve several critical features: starting from a study of the general population; clinical demonstration of adverse reactions to a food, preferably by Double-Blind, Placebo-Controlled Food Challenges (DBPCFCs) arriving to a clinical documentation of an IgE-mediated mechanism for the adverse reaction (Messina and Venter, 2020). Conducting such a large-scale study utilizing this approach may not be feasible. It takes a lot of time, money and effort which are not always available.

Cross sectional studies are the best choice to examine the prevalence of some outcome at a certain time. It can be used for both analytical and descriptive purposes (Thomas, 2020). This method is largely used for the estimation of food allergies' prevalence.

On 2019, a cross-sectional study of the prevalence of food allergies among children younger than ages 14 years was conducted in a Beijing urban region. The study showed a total prevalence of 3,2% with different prevalences between age, gender and foods showing at the same time that skin manifestations were the most common symptoms and fruits and seafood are the main allergens (Sha *et al.*, 2019). Another cross-sectional random questionnaire based survey on 2012 was conducted in Taiwan. It revealed that 6.95% was the prevalence of food allergies with predominance in children at the age between 4 to 18 years old (Wu *et al.*, 2012). In Toulouse schools, a cross-sectional, descriptive, questionnaire-based survey was conducted. The cumulative and point prevalences of food allergies were estimated at 6.7% and 4.7%, respectively. Cow's milk, eggs, and peanuts were the main food allergens (Rancé, Grandmottet and Grandjean, 2005). In Africa, the prevalence of reported adverse reactions assessed by the questionnaire in the Ghanaian schoolchildren was 11.0% and 5% showed a positive SPT reaction (Obeng *et al.*, 2011).

In Algeria, very few studies on the food allergy's prevalence are available in the literature. For instance; a cross-sectional descriptive study was made on 2008 in two private allergy's clinics in Skikda and Constatntine. The study gathered 103 patients came for suspicion in food allergy. 39 patients were diagnosed to have food allergies. According to prick test and specific IgE test, 13% were IgE related hypersensitivity and 39% were non-IgE related hypersensitivity. The main food allergen was egg with a percentage of 10% from the studied population (Latreche, 2009). On 2015, another cross-sectional study was conducted to

estimate the prevalence of cow's milk protein allergy in a pediatric population in Constantine. According to the survey results, the prevalence was 3.64% in this population. Boys were 2.11 times more affected by this pathology than girls. The main clinical manifestations were cutaneous (57.14%), digestive (46.42%), respiratory (25%) and anaphylactic (14.28%). The major risk factors were: positive history of atopy in the children's parents and early consumption of cow's milk (Boughellout, Benatallah and Zidoune, 2015). In the same context, a case-control study was done in the epidemiology service, hospital of Hussein Dey, Algiers during 6 years (2005-2010). 95 cases suffering from cow's milk protein allergy IgE-mediated (age of 113.095 days old ±13.94) were compared to 300 healthy subjects (119.92 days old ± 3.1). The risk factors related to this allergy were: cesarean ($p= 0.0001$), the use of artificial milk complements during the first week of birth before having breast milk ($p <0.0001$), atopy ($p<0.0001$), and breast feeding for a period superior than 3 months ($p= 0.006$) (Ibsaine *et al.*, 2010).

Facing the lack of information on the prevalence of food allergy and some fragmented studies available on this pathology in Algeria; we conducted for the first time a cross sectional survey study in the primary schools of the city of Guelma. The main goals of this research study are:

- Determine the prevalence of food allergy in this population.
- Find out the most common allergens among schoolchildren of the city.
- Make a correlation between this pathology and some of its risk factors.

II.2. Material and methods

1. Description of the study

Cross sectional study is one of the most common and well-known study designs. It is carried out at one time point or over a short period (Figure 1). The purpose of the study is to find the prevalence of the outcome of interest, for the population or subgroups within the population at a given time point (Levin, 2006).

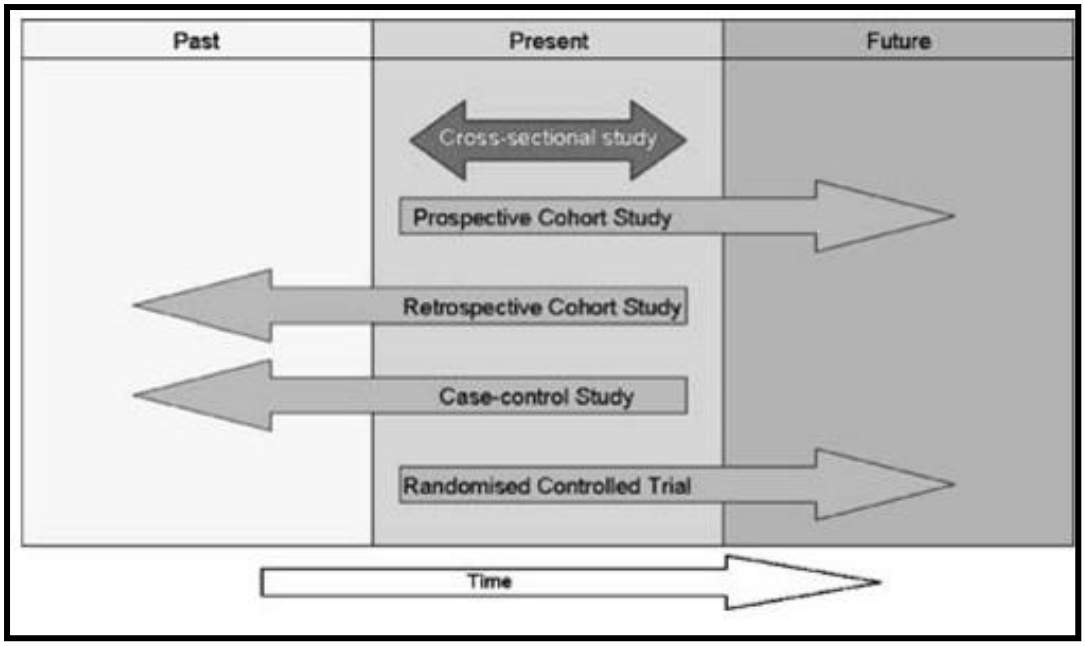


Figure1. Delimitation of cross sectional studies in time.(Levin, 2006)

In order to estimate the prevalence of food allergy in the schoolchildren in the town of Guelma, a cross-sectional, random, questionnaire based survey was performed during the last 2 weeks of April 2018.

2. Description of the population and sampling

In total, the city of Guelma has 40 primary schools with over than 11000 students according to the Directorate of Education of the province of Guelma on 2018.

8 schools (20%) accepted to participate in the study. The sample size was calculated according to the method shown in the work of Charan and Biswas (Charan and Biswas, 2013).

$$\textbf{Sample size} = Z_{1-\alpha/2}^2 p(1-p)/d^2$$

Here:

$\mathbf{Z_{1-\alpha/2}}$= is standard normal variate at 5% type 1 error (P<0.05) it is 1.96 and at 1% type error (P<0.01) it is 2.58).

As in majority of studies P values are considered significant below 0.05 hence 1.96 is used in formula.

p=Expected proportion in population based on previous studies or pilot studies.

d=Absolute error or precision-Has to be decided by the researcher.

With a 95% confidence interval (95% CI), assuming the prevalence was 5% or less and 2% absolute error. The sample size calculated was 457 children.

Anticipating a response less than 80%, 600 surveys were distributed. Children were selected by a multistage stratified random sampling strategy from all grades (preparatory to 5^{th} grade) in order to have all age categories.

3. Questionnaire

The survey was approved by the evaluation office in the education department of Guelma. The survey was written in Arabic (Appendix 1) and was addressed to the parents considering the fact that the children in the study are too young to answer by themselves.

The questionnaire was reviewed in several stages of development to reduce ambiguity, improve the flow of questions and to ensure that questions were specific. The interviewers checked that the questions were clear, easy to understand and practical for the respondents to answer.

A standard, anonymous questionnaire was distributed to gather personal data and the answer to the following question: ‘does your child have or ever had an allergic reaction to food?’ If the response was ‘Yes’, the parents were asked to keep on answering to other questions. There was no pre-determined list and the parents were asked to write down each

food involved in an adverse reaction. In the case of a positive response to the first question, the parents were asked to specify the food that caused the adverse reaction. For each food reported as causing an adverse reaction, the parents were asked to determine the age at the first reaction, the clinical signs, whether the allergy had disappeared or had persisted and whether there was a personal or family history of allergy.

4. Statistical analysis

The quantitative results are presented as means± SD. The Chi2 test was used for comparisons between proportions and kruskal-wallis test was performed for comparison between means. A 95% confidence interval (CI) was used and a p-value <0.05 was considered statistically significant. The statistical analyses were performed using XLSTAT 2016 (version 1802.01.28451).

I.3. Results and discussion

I.3.1. Results

1. General aspect of the results

In this cross-sectional questionnaire based study, 600 questionnaires were distributed randomly in 100 classes in eight different schools in the city of Guelma. The return rate was 88% (528).

There were 232 boys and 296 girls (sex ratio M/F 0.73). The overall average age was 10.3 ±3.6 years (5–13 years of age). The total of 528 children was classified into three age groups. 203 (38.4%) children had the age between 5-7 years, 290 (54.9%) children aged (8-10 years) and 35 (6.7%) children 11-13 years (table 1).

From the 53 subjects reporting food allergy, 8 were excluded because of the absurd answers on the food section (food coloring and conservatives (4), street food (2), mortadella (1) and soda (1)).

The prevalence of food allergies estimated according to the questionnaires was 8.5% (45 answered yes) (95% CI 6.1–10.8).

There were 12 (5.9%) children with food allergies in the first age group (5-7 years), 29 children (10%) allergic aged 8-10 years and 4 allergic children (11.4%) aged 11-13 years (Table 1).

Table1. Classification of children by age group gender and allergy.

Age group (years)	**Total number N**	**Gender**		**Allergic**	
		Girls	**Boys**	**Number n**	**Percentage (n/N)**
[5-7]	203	110	93	12	5.9
[8-10]	290	170	120	29	10
[11-13]	35	16	19	4	11.4
Total	528	296	232	45	

From subjects answering 'yes' to the allergy question: 51% (23) were girls and 49% (22) were boys. Figure 2 shows that the prevalence of food allergy in girls was 7.7% (95% CI 4.6–10.7) and in boys 9.5% (95% CI 5.7–13.2). According to the statistical analysis there is no significant influence of the gender on the appearance of allergy (p=0.429).

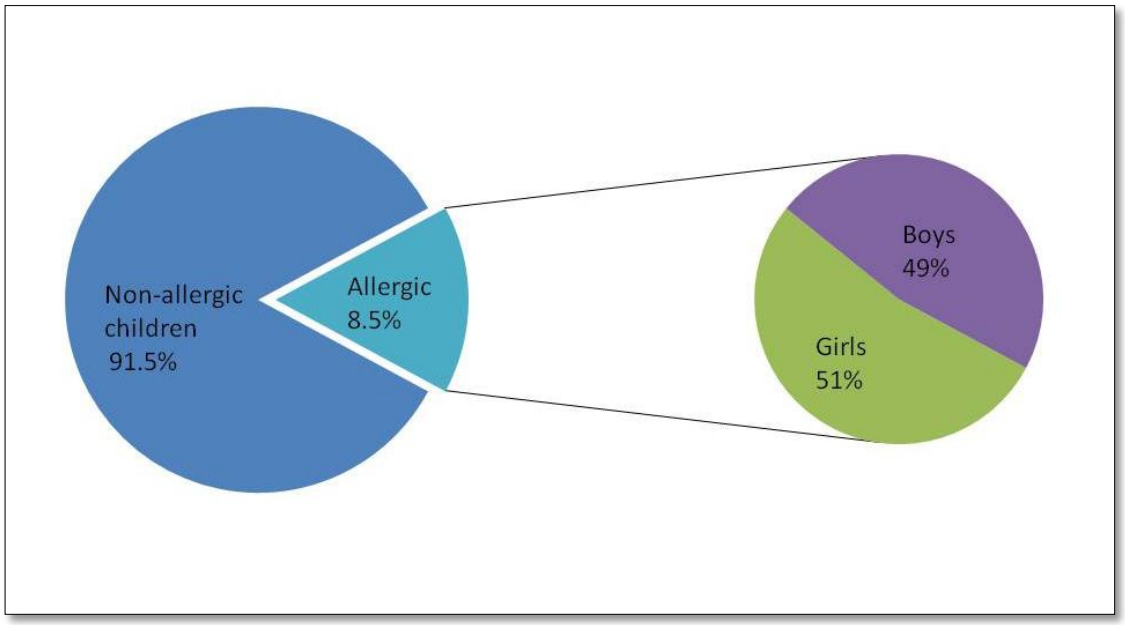


Figure 2. Detailed information of the percentage of allergic children by gender.

2. Prevalence of allergy by foods

In total 20 foods were cited 62 times (Figure 3). The food more cited was chocolate that was mentioned 9 times. After, comes wheat mentioned 7 times, followed by milk, eggs and strawberry cited 6 times each. Fish was represented with sardine was mentioned 5 times. faba beans was cited 4 times . Sunflower seeds and Peanut were cited 3 times each. Potatoes, green peas and apple were cited twice each. Finally, red meat (beef), beetroot, peach, tomatoes, melon, olive, raspberry and green beans were mentioned only once each.

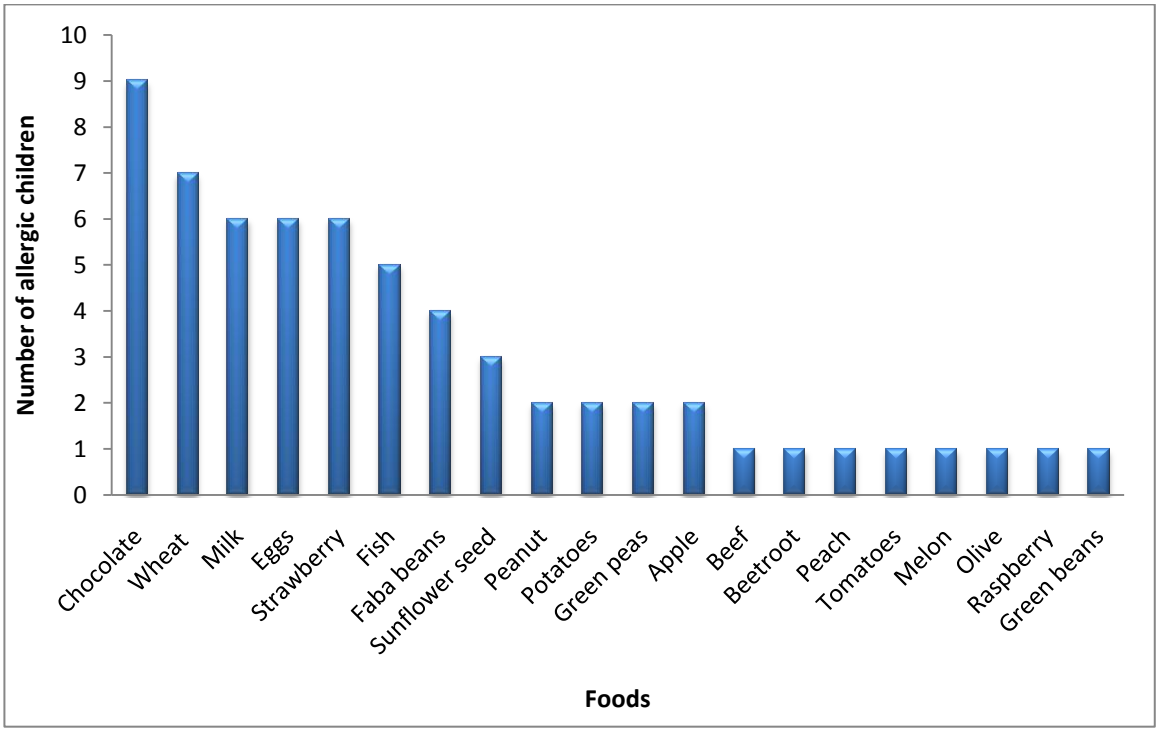


Figure 3. Foods reported by the parents causing food allergy to their children.

Table 2 gathers the details about the prevalence (%) of each food with the corresponding 95% confidence interval (CI).

The prevalence of food allergy to chocolate was estimated at (1.7%), Wheat 7 times (1.3%), milk, eggs and strawberries 6 times each (1.1%), fish and beans 5 times each (0.9%).

Table 2: Prevalence of foods causing food allergies and their confidence interval (IC)

Foods	Prevalence	95% CI
Chocolate	1.7	0.6- 2.8
Wheat	1.3	0.3-2.3
Milk	1.1	0.1-1.8
Egg	1.1	0.1-1.8
Strawberry	1.1	0.1-1.8
Fish	0.9	0.09-1.7
Faba beans	0.7	0.1-1.4
Sunflower seeds	0.5	0.1-1.4
Peanut	0,3	0.1-0.7

Potatoes	0,3	0.1-0.7
Green peas	0,3	0.1-0.7
Apple	0,3	0.1-0.7
Beef	0,1	0.03-0.3
Beetroot	0,1	0.03-0.3
Peach	0,1	0.03-0.3
Tomatoes	0,1	0.03-0.3
Melon	0,1	0.03-0.3
Olive	0,1	0.03-0.3
Raspberry	0,1	0.03-0.3
Green beans	0.1	0.03-0.3

From the 45 children reported with food allergy: 34 had only one food allergy, 9 had two food allergies, 1 had three food allergies and 1child had six food allergies.

Figure 4 points out the number of children with food allergies and the number of recovery in every age group. The recovery was higher in children with the age of 8-10 years (11 children, 42.8%), followed by 30.7% of recovery in the group of age 5-7 years and at last 20% of recovery for children with the age 11-13 years. There is no significant difference between the percentage of recovery between the age groups (p=0.178).

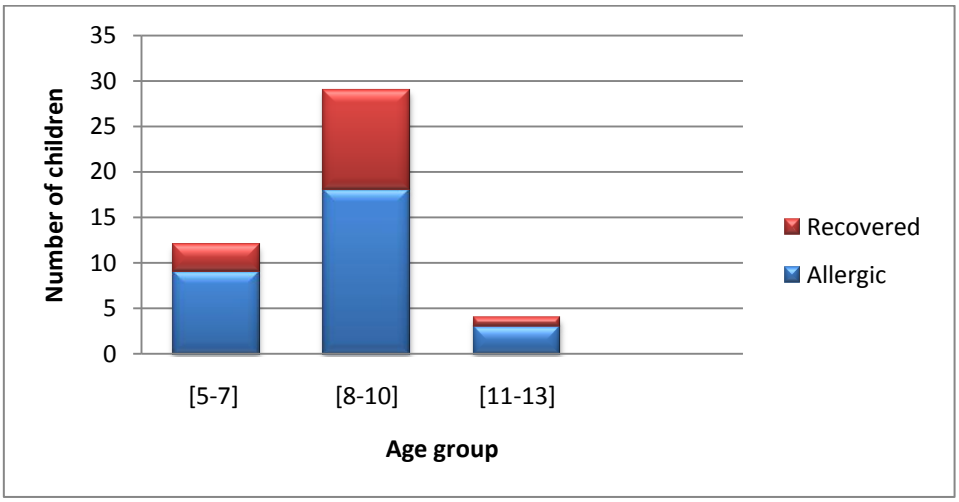


Figure 4. Representation of the allergic children and the recovered ones in each age group.

3. Main clinical characteristics

The average age when the food allergies were detected was 4.5 ±2.5 years (with a range of 0-10). Twenty children (38.5%) outgrew their food allergy. The average age when the allergy disappeared was 6.5±3.0 years.

Table 3 shows the age of appearance and recovery of food allergy in the most common foods. Milk had the youngest age of appearance and recovery (1.6 and 4 years old respectively) among the other foods. It had the highest percentage of recovery (66.6%) compared to all other foods.

Table 3. Appearance and recovery of the main food allergy

Food	Average age of appearance (years±SD)	Percentage of recovery	Average age of recovery (years±SD)
Chocolate	3.4±2.3†	33.3	6.3±3.5
Wheat	5.5±1.3‡	42.8	6.3±0.5
Milk	1.6±1.8¥	66.6	4.0±4.1
Egg	2.5±2.1†	60.0	7.6±4.0
Strawberry	5.0±4.0†	28.5	7.5±3.5
Fish	5.7±2.1§	60.0	10.3±0.5
Faba beans	4.0±2.7§	20.0	9.0±0.0

† Four subjects didn't mention the age of first reaction

‡ Three subjects didn't mention the age of first reaction

¥Two subjects didn't mention the age of first reaction

§One subject didn't mention the age of first reaction

The three clinical signs of allergy were cited 51 times. Cutaneous signs were mentioned 35 times (69%). It was the most frequent sign among the top 7 common foods except for wheat that caused digestive symptoms in most cases. Digestive signs were reported 10 times (19%). The respiratory symptoms were present in 6 cases (12%) (Figure 5).

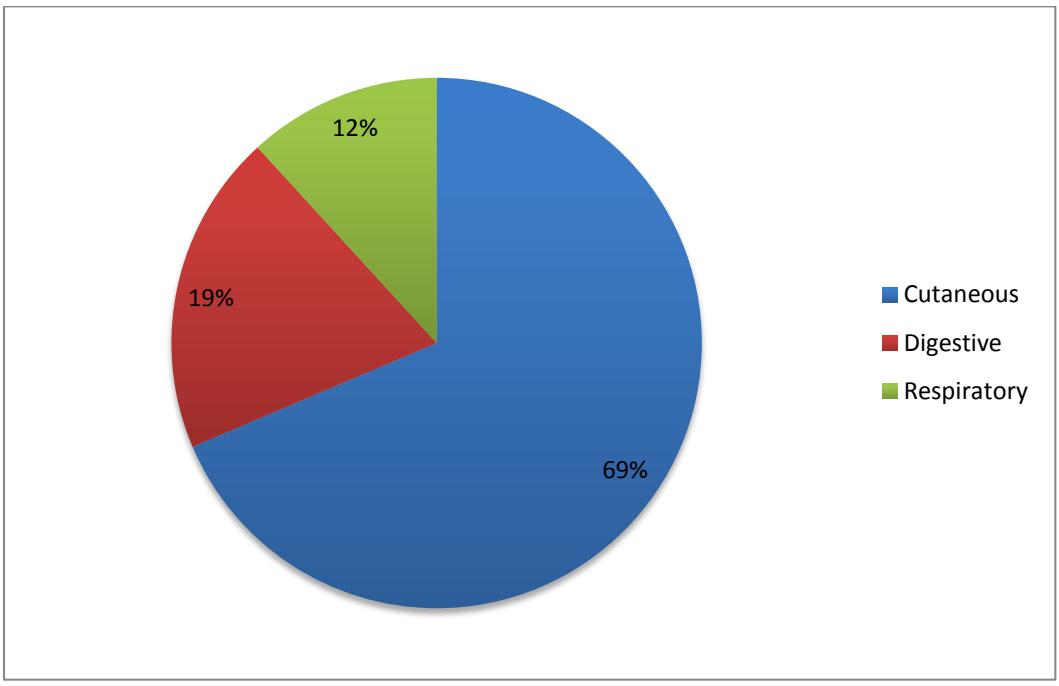


Figure 5. Percentage of the three clinical symptoms appearing on children with food allergy.

In some cases, children suffered from combined symptoms. Three children had cutaneous and respiratory issues at once. Two children had simultaneously cutaneous and digestive symptoms and one child had respiratory and digestive symptoms at the same time.

66.6% of children with family history of food allergy turned to be allergic while only 5.7% of children without a family history had food allergy. The statistical study shows a positive relation between the family history and the appearance of food allergy in children (p= 0.0001).

I.3.2. Discussion

In the present work, we conducted for the first time a cross-sectional study of the prevalence of food allergy in schoolchildren in the city of Guelma (Algeria). This type of studies has certain limitations and bias that we are aware of; especially the subjectivity of the answers. We examined food allergies without focusing on a single food in children aged from 5-13 years. The prevalence of parent-reported food allergy was estimated at 8.5% (95% CI 6.1–10.8).

In Africa, children aged between 5 and 16 years old from 9 Ghanaian schools were recruited after getting parental consent to participate in the study. Adverse reactions and food consumption were determined by a questionnaire and atopy by skin prick testing (SPT) to peanut and 6 fruits. The prevalence of reported adverse reactions assessed by the questionnaire in the Ghanaian schoolchildren was 11.0% and 5% showed a positive SPT reaction (Obeng *et al.*, 2011). In France, a cross-sectional, descriptive, questionnaire-based survey conducted in Toulouse (France) schools showed that the cumulative prevalence of food allergy in schoolchildren (2-14 years old) was 6.7% (95% (CI) 5.8–7.6) (Rancé, Grandmottet and Grandjean, 2005). In European, a study aimed to determine the prevalence of self-reported food allergy, food sentitization, probable food allergy (symptoms plus IgE sensitization), and challenge-confirmed food allergy in European school-age children. The prevalence of self-reported food allergy ranged from 6.5% in Athens to 24.6% in Lodz showing large geographical differences in the prevalence of food allergy in school-age children across Europe (Lyons *et al.*, 2020).

The previous prevalences of food allergy from an African country (Ghana) and European countries are showing some similarities to the prevalence found in our study.

In the question about the allergy, we aimed to know current and outgrown allergies. After dividing the cases into 3 age categories, we expected that the prevalence of food allergy decreases with the age. However the prevalence of recovery of the last category [11-13years] dropped. We can suppose that the recovery of food allergy reaches the maximum before 11 years old. Similar result was shown in the investigation of Rancé and his team, where they explained this result by a possible generational effect conducting to a real rise in food allergies in the younger children.(Rancé, Grandmottet and Grandjean, 2005)

According to our results the prevalence between girls and boys was not significant. In contrast, a review about the gender aspects in food allergy mentioned that gender differences likely plays a role in food allergy development. Boys are more prone to have allergies before puberty whereas in adulthood women are more likely to have allergies (Pali-Schöll and Jensen-Jarolim, 2019).

In the other hand, the family history appeared to be the main risk factor influencing the appearance of food allergy in children (66.6% of the allergic children have a family background with the pathology). Similar results were found in other studies showing that family history is known to be one of the main factors in the appearance of food allergy (Ibsaine *et al.*, 2010; Boughellout, Benatallah and Zidoune, 2015; Topçu *et al.*, 2018).

While most studies focus on one food, our study gathered information about food allergy in the population of schoolchildren despite the food. Many reports on food allergies confirmed the relation between geographical region and the foods responsible for the allergy (Rona *et al.*, 2007; Ontiveros *et al.*, 2014; Lyons *et al.*, 2020). Some studies reported that atopy prevalence is higher in urban children than rural children (Nafti *et al.*, 2009; Chu, 2014; Sha *et al.*, 2019). Knowing that the population studied in our survey was urban; this factor can have some impact in our results.

In addition to regional factors, The hypersensitivity reaction is also influenced by genetics, customs and traditions, and is also related to dietetic exposure to new allergenic food products over the world (Ansari and Mu, 2018).

Chocolate was found to be the most common food causing allergy to children in our study (1.7% 95% (CI) 0.6- 2.8). Here comes again the possibility of bias in the questionnaire and the subjectivity in answering the questions. However, nevertheless of the absence of evidence of food allergy to chocolate, according to the study in Salvadoran schoolchildren, the prevalence of self-reported chocolate/cocoa allergy is 0.5% to 0.7% in select populations (Cabrera-Chávez *et al.*, 2018). Some studies are suggesting that chocolate or cocoa allergy is a result of cross-contamination from some allergens such as: milk (Planque *et al.*, 2017), tree nuts (Scheibe *et al.*, 2001) or peanuts (Vadas and Perelman, 2003). In addition to that, chocolate is considered to be a food that contains histamine releasers, when consumed, it can cause pseudo-allergic reactions (Maintz and Novak, 2007).

Milk, eggs, wheat and fish are known to be part of the most common causes of food allergies worldwide in children (Ansari and Mu, 2018; Ochfeld and Pongracic, 2019). This

suggestion was confirmed in our results where these foods were in the top eight allergy reasons with different prevalences.

The allergy to wheat affected 1.3% (95% (CI) 0.3-2.3) of the studied population. Knowing that wheat is associated to multiple disorders like: IgE mediated wheat allergy, non IgE mediated wheat allergy and celiac disease; it could explain the high prevalence of this allergy.

In a study on the risks of wheat allergy and early exposure to cereal grains, it has been shown that allergy to wheat is more frequent when wheat is introduced after 6 months of age into the children's diet (Poole *et al.*, 2006). This statement could explain the result in our study assuming that the Algerian diet involves eating wheat in different forms and dishes especially for the young aged kids. More than 62% suffered from digestive symptoms, then comes cutaneous and respiratory signs (25% and 12.5% respectively). According to the literature, symptoms of wheat allergy depend on the age, at a young age. It starts with gastroenterological symptoms. They recede with age; therefore, older children suffer mostly from dermatitis and respiratory disorders (wheeze, stridor, persistent cough, hoarse voice, respiratory distress, nasal congestion) (Czaja-Bulsa and Bulsa, 2017).

According to our findings, the prevalence of milk allergy was 1.1% (95% (CI) 0.1-1.8). The same result was found in cases follow up study. It showed that the prevalence of IgE-mediated cow's milk proteins allergy (CMPA) in Algeria was 1.1% (Ibsaine *et al.*, 2010). In the same context, another study on 2015 revealed the prevalence of CMPA to be 3.64% in a pediatric population in Constantine, Algeria (Boughellout, Benatallah and Zidoune, 2015). The average age of appearance of allergy to milk was 2.62± 1.79 years. The allergy to milk is known to affect children at an early age this was confirmed by the mean age of the appearance of this allergy in our study (1.6±1.8 years). This fluctuation could be explained by a late weaning in the Algerian society. Due to religious reasons (Al-Jassir *et al.*, 2006), Algerian mothers tend to wean children around the age of 2 years old.

The prevalence of allergy to eggs was estimated at 1.1% (95% (CI) 0.1-1.8). This food is also one of the most common foods causing allergy worldwide in children. The prevalence of eggs' allergy among kids varies from 0.3% to 19% according to several studies (Nwaru *et al.*, 2014; Irani and Maalouly, 2015; Peters *et al.*, 2017; Loh and Tang, 2018; Sha *et al.*, 2019; Lyons *et al.*, 2020). Those findings are comparable to ours, showing that the studied population is in accordance with the international data.

Nevertheless of the high number of studies focusing on the identification and characterization of the allergens in strawberry (Marzban *et al.*, 2008; Casañal *et al.*, 2013; Franz-oberdorf *et al.*, 2016; Ishibashi *et al.*, 2017; Orozco-Navarrete *et al.*, 2019), studies on the prevalence of this fruit are not enough. Among schoolchildren (3-11 years old) from Central Portugal, a study reported that the most frequent food group was fresh fruits including strawberry as the top fruit in the list(Jorge *et al.*, 2017). In another work, the parent-reported prevalence of food allergy to strawberry was 0.6% in schoolchildren in a Mexican population (Ontiveros *et al.*, 2016). According to our work, strawberry appeared to be one of the most common foods causing allergy in schoolchildren with a prevalence of 1.1% (95% (CI) 0.1-1.8). In the literature, allergy to this fruit is often reported but rarely confirmed. The work of Collins et al., reported one case study of 9 year old boy with a history of strawberry and raspberry anaphylaxis.(Collins *et al.*, 2020). The presence of histamine releasers in strawberry can also explain the prevalence of the allergy of this fruit. In certain subjects the consumption of this fruit can cause non-specific histamine-releasing resulting allergic responses (Malone and Metcalfe, 1986; Lorenz, Scheurer and Vieths, 2015).

The study made by Lyons et al., reported the prevalence (95% CI) of food sensitization to fish in some European cities. The prevalence was 0.91% (0.40%-1.42%) in Madrid and 0.76% (0.44%-1.08%) in Vilnius (Lithuania) (Lyons *et al.*, 2020). These results are comparable to ours, where the prevalence of parent-reported food allergy to fish was estimated at 0.9% (0.09%-1.7%).

Food allergy to Faba beans had a prevalence of 0.7% (0.1%-1.4%). Surprisingly, the allergy to these beans is rare. It has been reported for the first time on 2007 when a 25 year old Spanish woman showed itching in the tongue and pharynx 5 min after eating a sandwich containing broad bean flour (Mur Gimeno *et al.*, 2007). Another case was a 49-year-old Italian woman, farmer, showed adverse reactions to raw and boiled faba beans after ingestion or handling fresh faba bean vegetable. In an investigation in a Moroccan population, it has been shown that 79.3% of children and 80.4% of 94 adults showed high levels of IgE against faba beans (Bousfiha and Aarab, 2014). In an Egyptian investigation, it has been reported that 12.5% of children and 50% of adults had positive IgE against faba beans. It has been shown that this allergy is less common in children than adults (Hamad *et al.*, 2020).

The prevalence of allergy to sunflower seeds was 0.5%. Although uncommon, sunflower seed allergy has been reported as a cause of IgE-mediated food allergic reaction

and anaphylaxis (Axelsson, Ihre and Zetterström, 1994; Iwaya *et al.*, 1994; Palma-Carlos, Palma-Carlos and Tengarrinha, 2005; Lavine and Ben-Shoshan, 2015). In some cases allergy to sunflower seeds may be associated with inhalant allergy to pollen from plants of the family Compositae, especially Artemisia (mugwort)(Yagami, 2010; Ukleja-Sokołowska *et al.*, 2016). In the other hand some cases linked this allergy to ownership of household pet birds where sensitization happens during the handling of birdseed containing sunflower seeds (Palma-Carlos, Palma-Carlos and Tengarrinha, 2005; Lavine and Ben-Shoshan, 2015).

Allergy to peanut is known to be one of the most common allergies in children (Abrams, Chan and Sicherer, 2020). IgE-mediated peanut allergy has an estimated prevalence of between 0.2% and 4.5% (Greenhawt *et al.*, 2020). Prevalence estimates are difficult to ascertain because the gold standard for determining diagnosis is a double-blind, placebo-controlled oral food challenge (DBPCFC), but many epidemiologic studies estimate prevalence based on self-reported food allergy or sensitization only. That was the case in our study (Oriel and Wang, 2019). According to our results, the prevalence to parent reported allergy to peanut was estimated at 0.3% (0.1% -0.7%).

Potatoes were guilty of causing food allergy in 0.3% (0.1% -0.7%) of the studied population. Hypersensitivity to raw potato is probably owing to patatin which is the main storage protein of potato tubers (Ansari and Mu, 2018). In a cross sectional study in France during two years with all ages included (more than 2 years old); Sensitization to potato was found in one-tenth of the study population, which is not a negligible rate in Mediterranean population. They presumed that those data prove a good profile of tolerance to potato, in a large population (Chiriac *et al.*, 2017).

According to our results, allergy to green peas was estimated at 0.3% (0.1% -0.7%). It has been estimated that 0.8% of Europeans suffer from allergies to green pea proteins (Chudzik-Kozłowska, Wasilewska and Złotkowska, 2020). In a review published in 2021, it has been reported that 2.3% of children with IgE-mediated clinically-relevant allergy to pea was found. Furthermore cases of anaphylaxis were reported to different legumes and peas were one of them (Vergeer *et al.*, 2019). This high prevalence was explained by the recent shift towards plant-based diets in the Western world. Moreover, legume protein, specifically pea, has recently been appearing in more manufactured pre-packaged food products due to economic decisions by food manufacturers to increase protein content in foods using cheaper sources of protein (Hildebrand *et al.*, 2021).

Allergy to apple and peach had a prevalence of 0.3% (0.1% -0.7%) and 0.1% (0.03%-0.3%) respectively according to our study. Sensitization to plant-source foods such as apple and peach in the Mediterranean is more likely due to primary sensitization, and partly through lipid transfer protein (Lyons *et al.*, 2020).

Both prevalences in our study are lower than results in western countries (Gomez *et al.*, 2014; Lyons *et al.*, 2020). This difference could be explained by the fact that allergy to apple and peach in western countries is related to birch pollen-food syndrome (Shirasaki *et al.*, 2017; Čelakovská *et al.*, 2021). The sensitization to the major birch pollen allergen Bet v 1 generates IgEs against Bet v 1 in birch pollen allergic patients. These ones cross-react with structural homologue members of the PR-10 family that are present in fruits; the major allergen in apple is Mal d 1 and rPru p 1 in peach which are also PR-10 proteins (Shirasaki *et al.*, 2017; Orozco-Navarrete *et al.*, 2019; van der Valk *et al.*, 2020). Due to the absence of birch trees in the studied region this could explain the low prevalence of allergy to both apple and peach.

Mammalian meat allergy is a recently described disease with a characteristic clinical presentation resulting from an allergic reaction mediated by IgE antibodies directed against the mammalian oligosaccharide epitope galactose-α-1,3-galactose (Wong and Sebaratnam, 2018). The overall incidence and prevalence of beef allergy in childhood is unknown (Topçu *et al.*, 2018). According to our study the prevalence of beef allergy was 0.1% (0.03%-0.3%). In a cross sectional study among urban children in turkey, the prevalence of IgE-mediated beef allergy confirmed by double-blind, placebo controlled food challenge (DBPCFC) was 0.30% (95% confidence interval, 0.18–0.5%) (Topçu *et al.*, 2018).

Allergy to beetroot is rare. Usually reports of allergic reactions to beet and closely related plants mostly involve rhinitis or asthma symptoms, rather than food allergy (El-Hosni, Montejo and Schuler, 2020). Only two cases of anaphylaxis of this vegetable are available. The first one, a 13-years old Brazilian girl with food anaphylaxis was attributed to beetroot. She has complained of urticaria and asthma about 40 minutes after ingesting boiled beetroot on a meal confirmed with a positive oral challenge to boiled beetroot (Lopes de Oliveira *et al.*, 2011). The second case of anaphylaxis belongs to a 22-month-old girl who was referred for evaluation of anaphylaxis after ingestion of boiled beetroot supported by positive skin allergy testing (El-Hosni, Montejo and Schuler, 2020). The prevalence of allergy to beetroot is not available in the literature. Our results showed a prevalence of 0.1% (0.03%-0.3%).

The prevalence of tomato allergy ranges from 1.5% in northern Europe up to 16% in Italy among the food allergic population (Dölle *et al.*, 2011). An Italian epidemiological study showed that among adults allergic to plant-derived foods (tomato included), only 0.5% had allergy to genuine vegetable food allergy (Asero *et al.*, 2009) .As tomato allergy results partly from pollen cross-reactivity, a geographical difference in sensitization pattern between patient populations exists, mainly due to the specific geographical distribution of pollen and to local dietary habits (Dölle *et al.*, 2011). This could explain the prevalence of tomato allergy in our work estimated at 0.1% (0.03%-0.3%).

Our study revealed a prevalence of melon allergy at 0.1% (0.03%-0.3%) among schoolchildren. Allergic reactions that develop after eating melon are frequent. Oral allergy syndrome and systemic reactions have been reported after ingestion of melon pulp (Gandolfo-Cano *et al.*, 2014). In patients with oral allergy syndrome (10years and above), who visited the Department of Otorhinolaryngology at the University of Fukui; melon was the causative food in 42% of the cases (Osawa *et al.*, 2020).

Our study revealed a prevalence of olive allergy at 0.1% (0.03%-0.3%). Nevertheless of the high consumption of olive in the region, its allergy was very low. Similar observation was reported in a case report of a Tunisian woman having palatal itching and generalized urticaria following ingestion of olive fruit. The oral provocation test was positive for olives and negative for olive oil (Racil *et al.*, 2015). Allergy to olive is mainly due to the pollen produced in these flowers produced by the tree, being allergic to olive fruit and olive oil less common (Esteve *et al.*, 2012).

Allergy to raspberry had a prevalence of 0.1% (0.03%-0.3%) according to our investigation. Reports on the allergenicity of berries such as raspberries, blackberries and blueberries are rare, but they do exist (Dosanjh, 2019; Hallmann *et al.*, 2020). It has been shown in a study of the prevalence of food allergy among children that berries are not on the list of the most common food allergies (Gupta *et al.*, 2018).

Allergy to Green beans had a prevalence of 0.1% (0.03%-0.3%) in our study. Few cases in the literature have described allergic reactions upon the exposure to green bean boiling steam or ingestion. In 2010, five patients with IgE positivity toward an allergen in green beans were reported (Pastorello *et al.*, 2010).

Foods with animal origins like: milk, eggs and fish have shown to have a high percentage of recovery (66.6%, 60%, and 60% respectively). In the other hand, foods with plant origins have less percentage of recovery such as chocolate and strawberry (33.3% and 28% respectively). This could be explained by the involvement of the pollen-food allergy syndrome. The sensitization to some aeroallergens can cause cross reaction with food allergens when ingested; for instance, an aeroallergen in date palms from the profilin family (Pho d 2) can cause cross reaction with food allergens belonging to the same family like: Cor a 2 in hazelnuts, Ara h 5 in peanuts and Fra a 4 in strawberry (Carlson and Coop, 2019). Among the clinical manifestations, the dermatologic signs were popular with 68.6%, followed by digestive signs (19.6%) then respiratory symptoms (11.8%). Many studies confirm these findings; for instance: the main clinical manifestation in the IgE adverse food reaction in Portuguese children was mucocutaneous. (Jorge *et al.*, 2017) In the same context, another investigation reported that food allergy related adverse reaction in children having immediate food allergy was mainly skin symptoms (62%) (Ontiveros *et al.*, 2016)

I.4. Conclusion

Food allergy is an important health issue affecting children and adults worldwide. This pathology appears to be increasing and has become an important health concern in developing and developed countries. Studies on the prevalence of food allergies became crucial. It allows the determination of the most frequent foods causing this reaction especially that food allergy depends on the most consumed foods in the region. In the other hand it is important to study the total population and determine the main risk factors that interfere in the appearance of this pathology.

Only few fragmented and rare studies reporting food allergies in Algeria and none in the city of Guelma are available. In this work we aimed to get an estimation of food allergies in schoolchildren of the city of Guelma, highlight the main foods causing the allergy and the risk factors related to this health issue.

The results showed for the first time the prevalence of parent-reported food allergy among schoolchildren in 20% of primary schools of the city of Guelma (Algeria). The prevalence of reported food allergy was estimated at 8.5% (95% CI 6.1-10.8).

Surprisingly, the most common food allergy was allergy to chocolate with a prevalence of 1.7%. This high prevalence was associated with different suggestions: first, the fact that chocolate can contain allergens from other sources like: milk, peanut or hazelnuts. Second, chocolate is known to have histamine releaser that can induce similar symptoms to the ones caused by real food allergy.

Allergy to wheat had also a high prevalence (1.3%). Knowing that wheat is associated to multiple disorders like: IgE mediated wheat allergy, non IgE mediated wheat allergy and celiac disease; it could explain the high prevalence of this allergy.

In accordance to the results available in the literature, milk, eggs and fish were from the top foods causing allergy in kids.

For risk factors related to food allergy, family history was found to be the main risk factor influencing the appearance of food allergy in children according to our results. The prevalence of food allergy in girls was 7.7% (95% CI 4.6–10.7) and in boys 9.5% (95% CI 5.7–13.2). Statistically this difference was not significant, suggesting that according to our results; gender was not a risk factor of food allergy.

According to our results, cutaneous reactions were the most common symptom (68.6%), followed by digestive (19.6%) and respiratory reactions (11.8%).

Overall, the results provided by the survey were comparable. Additional investigation should be done in order to provide more precise data on food allergy by making follow up of patients of food allergy and diagnosing using clinical demonstrations such as skin prick test, search of specific IgE or Placebo-Controlled Food Challenges (DBPCFCs).

Scientific paper 1

Revue française d'allergologie xxx (xxxx) xxx–xxx

Disponible en ligne sur
ScienceDirect
www.sciencedirect.com

Elsevier Masson France
EM|consulte
www.em-consulte.com

Original article

Contribution to the estimation of the prevalence of food allergy in schoolchildren in the city of Guelma (Algeria)

Contribution à l'estimation de la prévalence de l'allergie alimentaire chez les écoliers de la ville de Guelma (Algérie)

M. Yakhlef[a,b,*], L. Souiki[b,c]

[a] *Laboratoire Biologie, Eau et Environnement (LBEE), Université 8 Mai 1945 Guelma. BP. 401 Guelma 24000, Algeria*
[b] *Département de biologie, faculté des sciences de la nature et de la vie et sciences de la terre et l'univers, université 8 mai 1945 Guelma, BP 401, 24000 Guelma, Algeria*
[c] *Laboratoire de biochimie et de toxicologie environnementale, université Badji-Mokhtar, BP 12, 23000 Annaba, Algeria*

ARTICLE INFO

Article history:
Received 6 March 2021
Accepted 22 June 2021
Available online xxx

Keywords:
Food allergy
Prevalence
Schoolchildren
Cross sectional study
Epidemiology
Algeria

ABSTRACT

Goal. – The prevalence of food allergy has never been estimated in the city of Guelma (east Algeria). The goal in this work was to determine parent-reported prevalence of food allergy in a population of schoolchildren (5 to 13 years old).
Methods. – Cross-sectional, random, questionnaire based survey was performed from 15th to 29th April 2018 among primary schools of Guelma. Six hundreds questionnaires were distributed in 8 schools of the city.
Results. – The return rate was 88% (528). Forty-five subjects were retained as having probable food allergy at a prevalence of 8.5% (95% CI 6.1–10.8). In total, 49% were boys and 51% girls with no significant difference ($P = 0.429$). The average age of appearance of the potential allergy was 4.41 ± 2.51 years. Among allergic children, 26.7% were aged from 5–7 years old, 64.4% had 8–10 years old and 8.9% had 11–13 years old. The main foods reported by the parents were: chocolate (1.7%), wheat (1.3%), milk (1.1%), eggs (1.1%), strawberry (1.1%), fish and beans (0.9%). The prevalence of the three symptoms according to the observations of the parents were cutaneous (68.6%), digestive (19.6%) and respiratory (11.8%). Among the allergic subjects, 66.6% had at least one of their parents atopic ($P = 0.0001$).
Conclusion. – The prevalence of parent-reported food allergy in schoolchildren was assessed for the first time in the city of Guelma. It showed a prevalence of 8.5% with no significant difference between the boys and the girls. The main risk factor in the appearance of food allergy was the presence of family history with this pathology.

Mots clés :
Allergie alimentaire
Prévalence
Écoliers
Enquête transversale
Épidémiologie
Algérie

RÉSUMÉ

But. – La prévalence de l'allergie alimentaire n'a jamais été estimée dans la ville de Guelma (est de l'Algérie). Le but de ce travail était de déterminer la prévalence des allergies alimentaires déclarée par les parents dans une population d'écoliers (de 5 à 13 ans).
Méthode. – Une enquête transversale, aléatoire, basée sur un questionnaire a été réalisée du 15 au 29 avril 2018 dans les écoles primaires de Guelma. Au total, 600 questionnaires ont été distribués dans 8 écoles de la ville.

* Corresponding author.
E-mail address: yakhlef.marwa@univ-guelma.dz (M. Yakhlef).

https://doi.org/10.1016/j.reval.2021.06.005
1877-0320/

Please cite this article as: Yakhlef M, Souiki L, Contribution to the estimation of the prevalence of food allergy in schoolchildren in the city of Guelma (Algeria), Rev Fr Allergol, https://doi.org/10.1016/j.reval.2021.06.005

Résultats. – Le taux de retour était de 88 % (528). Quarante-cinq sujets ont été retenus comme ayant une allergie alimentaire probable, soit une prévalence de 8,5 % (IC à 95 % 6,1–10,8). Quarante-neuf pour cent étaient des garçons et 51 % des filles sans différence significative ($p = 0{,}429$). L'âge moyen d'apparition de l'allergie potentielle était de 4,41 ± 2,51 ans. Parmi les enfants allergiques, 26,7 % étaient âgés de 5 à 7 ans, 64,4 % de 8 à 10 ans et 8,9 % de 11 à 13 ans. Les principaux aliments déclarés par les parents étaient : le chocolat (1,7 %), le blé (1,3 %), le lait (1,1 %), les œufs (1,1 %), la fraise (1,1 %), le poisson et les haricots (0,9 %). La prévalence des trois symptômes selon les observations des parents était cutanée (68,6 %), digestive (19,6 %) et respiratoire (11,8 %). Parmi les sujets allergiques, 66,6 % avaient au moins un de leurs parents atopique ($p = 0{,}0001$).
Conclusion. – La prévalence des allergies alimentaires déclarées par les parents d'enfants scolarisés a été estimée à 8,5 %, sans différence significative entre les garçons et les filles. Le principal facteur de risque d'apparition de l'allergie alimentaire était la présence d'antécédents familiaux avec cette pathologie.

1. Introduction

Little information on the prevalence of food allergies is available despite the efforts underway to provide reliable data [1]. A rigourous estimation of the prevalence of FA in a general population is starting by studying the history and the clinical symptoms and assessing allergic sensitization (skin prick tests, specific IgE dosage), and if any doubt, DBPCFC [2]. In order to make such a large-scale study using this approach may not be feasible. It takes a lot of time, money and effort which are not always available. Cross sectional studies are the best choice to examine the prevalence of some outcome at a certain moment in time. It can be used for both analytical and descriptive purposes [3]. This method is largely used for the estimation of food allergies' prevalence.

Facing the lack of information on the prevalence of food allergy and some fragmented studies available on this pathology in Algeria [4–6]; we conducted a cross sectional questionnaire based study in the primary schools of the city of Guelma, targeting children aged between 5 and 13 years old. The main goals of this research study are to overview the current state of food allergies in schoolchildren in the city of Guelma; to determine the prevalence of food allergy in this population, to highlight the most common foods causing allergy and finally to assess any correlation between this pathology and some of its risk factors.

2. Material and methods

In order to estimate the prevalence of food allergy in the schoolchildren in the town of Guelma, a cross sectional questionnaire based study was conducted in 8 different primary schools. The study was made from 15th to 29th April 2018.

2.1. Study design and population

Cross-sectional, random, questionnaire based survey was performed during the last 2 weeks of April 2018. Among 40 primary schools of the city, 8 schools (20%) accepted to participate in the study.

The sample size was calculated according to the method shown in the work of Charan and Biswas [7]. With a 95% confidence interval (95% CI), assuming the prevalence was 5% or less and 2% absolute error, the sample size calculated was 457 children. Anticipating a response less than 80%, 600 surveys were distributed. Children were randomly recruited from all grades (1st to 5th grade) in order to have all age categories. There was no direct contact between the recruited children and the researchers.

2.2. Questionnaire

The survey was approved by the evaluation office in the education department of Guelma. The survey was written in Arabic and was addressed to parents assuming that the targeted children are too young to answer by themselves.

A standard, anonymous questionnaire was distributed to gather personal data (gender and age). The first question to the parents was: "does your child have or ever had an allergic reaction to food?" If the answer was "Yes", the parents were asked to determine the food/foods causing the allergy because there was no pre-determined list of foods. For each food reported as causing an adverse reaction, the parents were asked to determine the age of the first reaction, the clinical signs, whether the allergy had persisted or had disappeared with the age of recovery and whether there was family history of allergy. The survey is provided as supplementary figure (Online material Supplementary Fig. 1).

2.3. Statistical analysis

The quantitative results are presented as median with interquartile range (IQR). The Chi2 test was used for comparisons between proportions and kruskal-wallis test was performed for comparison between means. The significance level was set at $P < 0.05$. The statistical analyses were performed using XLSTAT 2016 (version 1802.01.28451).

3. Results

3.1. General aspect of the results

In this cross-sectional questionnaire based study, 600 questionnaires were distributed randomly in 100 classes in eight different schools in the city of Guelma. The return rate was 88% (528).

The overall average age was 10.3 ± 8.6 years (5–13 years of age). There were 232 boys and 296 girls (sex ratio M/F 0.73). In total, 203 (38.4%) children had the age between 5–7 years, 290 (54.9%) children aged (8–10 years) and 35 (6.7%) children 11–13 years (Table 1).

From the 53 subjects reporting food allergy, 8 were excluded because of the absurd answers on the food section (food coloring and conservatives (4), street food (2), mortadella (1) and soda (1)).

3.2. Prevalence of food allergy

The prevalence of food allergies estimated according to the questionnaires was 8.5%, 45 answered yes (95% CI 6.1–10.8).

Table 1
Classification of children by age group, allergic subjects and recovery.

Age group (years)	Total number *n*	Gender		Allergic		Recovery	
		Girls	Boys	Number *n*	Percentage (*n*/N)	Number *n*1	Percentage (*n*1/*n*)
[5–7]	203	110	93	12	5.9	4	33.3
[8–10]	290	170	120	29	10	15	51.7
[11–13]	35	16	19	4	11.4	1	25.0
Total	528	296	232	45		20	

Table 2
Prevalence of foods causing food allergies.

Foods	Number	Percentage
Chocolate	9	1.7
Wheat	7	1.3
Milk	6	1.1
Egg	6	1.1
Strawberry	6	1.1
Fish[a]	5	0.9
Beans[b]	5	0,9
Sunflower seeds	3	0,5
Peanut	2	0,3
Potatoes	2	0,3
Green peas	2	0,3
Apple	2	0,3
Meat[c]	1	0,1
Beetroot	1	0,1
Peach	1	0,1
Tomatoes	1	0,1
Melon	1	0,1
Olive	1	0,1
Raspberry	1	0,1

[a] Fish: the only type mentioned in the 5 cases was sardine.
[b] Beans regroup 4 children allergic to faba beans and one child allergic to green beans.
[c] Meat refers to beef.

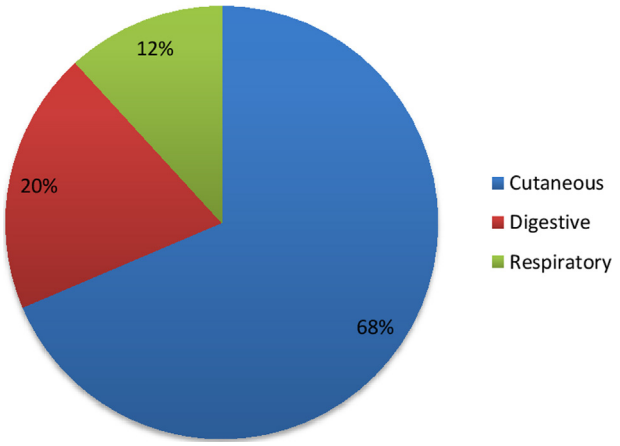


Fig. 1. Percentage of the three clinical symptoms appearing on children with food allergy.

3.2.1. Prevalence of food allergy by age group and gender

There were 12 (5.9%) children with food allergies in the first age group (5–7 years), 29 children (10%) allergic aged 8–10 years and 4 allergic children (11.4%) aged 11–13 years (Table 1).

From subjects answering "yes" to the allergy question: 51% (23) were girls and 49% (22) were boys. The prevalence of food allergy in girls was 7.7% (95% CI 4.6–10.7) and in boys 9.5% (95% CI 5.7–13.2). According to our study there is no significant influence of the gender on the appearance of allergy ($P = 0.429$).

3.2.2. Prevalence of allergy by foods

In total, 20 foods were cited 62 times. Chocolate was mentioned 9 times (1.7%), wheat 7 times (1.3%), milk, eggs and strawberries 6 times each (1.1%), fish and beans (faba beans and green beans) 5 times each (0.9%). Greater detail is provided in Table 2.

From the 45 children reported with food allergy: 34 had only one food allergy, 9 had two food allergies, 1 had three food allergies and 1child had six food allergies.

Table 1 points out the number of children with food allergies and the number of recovery in every age group. The recovery was higher in children with the age of 8–10 years (15 children, 51.7%), followed by 33.3% of recovery in the group of age 5–7 years and at last 25% of recovery for children with the age 11–13 years. There is no significant difference between the percentage of recovery between the age groups ($P = 0.178$).

3.3. Main clinical characteristics

The median age when the food allergies were detected was 4 years (with a range of 0–10 years). Twenty children (44.4%) mentioned that they outgrew their food allergy. The median age when the allergy disappeared was 6.5 (with a range of 1–11 years).

Table 3 shows the median age (range) of appearance and recovery of food allergy in the most common foods. Even though milk had the youngest age of appearance and recovery (1.2 and 2 years old respectively) among the other foods, the differences for both appearance and recovery were not statistically significant ($P = 0.166$ and $P = 0.325$ respectively). In the other hand, milk had the highest percentage of recovery (66.6%) compared to all other foods but no statistical significance ($P = 0.714$).

The three clinical signs of allergy were cited 51 times. Cutaneous signs were mentioned 35 times (69%). It was the most frequent sign among the top 7 common foods except for wheat that caused digestive symptoms in most cases. Digestive signs were reported 10 times (19%). The respiratory symptoms were present in 6 cases (12%) (Fig. 1).

In some cases, children suffered from combined symptoms. Three children had cutaneous and respiratory issues at once. Two children had simultaneously cutaneous and digestive symptoms and one child had respiratory and digestive symptoms at the same time.

3.3.1. Clinical signs by food

Table 4 shows the clinical signs for the top 7 foods (chocolate, wheat, milk, egg, strawberry, fish, and beans). Cutaneous symptoms were dominant in all those foods except for wheat that had a higher percentage in digestive symptoms. Both strawberry and beans didn't cause any digestive problems while the allergy to eggs was exclusively cutaneous.

3.4. History of atopy in allergic children

In total, 66.6% of children with family history of food allergy turned to be allergic while only 5.7% of children without a family history had food allergy. The statistical study shows a positive

Table 3
Appearance and recovery of food allergypresented by median (rage) and interquartile range IQR (7 most common foods).

Food	Number	Age of appearance median (range)	IQR	Percentage of recovery	IQR	Age of recovery median (range)
Chocolate	9	3.0 (1–7)[a]	2	33.3	3.5	6 (3–10)
Wheat	7	5.5 (4–7)[b]	1.5	42.8	0.5	6 (6–7)
Milk	6	1.2 (0–4)[c]	2.12	66.6	4.5	2 (1–10)
Egg	6	2.5 (1–4)[a]	1.5	60.0	3.5	10 (3–10)
Strawberry	6	5.0 (1–9)[a]	4.0	28.5	2.5	7.5 (5–10)
Fish	5	6.0 (3–8)[d]	1.25	60.0	0.5	10 (10–1)
Beans	5	3.0 (2–8)[d]	1.5	20.0	/	9.0 (/)[e]

[a] Four subjects didn't mention the age of first reaction.
[b] Three subjects didn't mention the age of first reaction.
[c] Two subjects didn't mention the age of first reaction.
[d] One subject didn't mention the age of first reaction.
[e] Only one subject recovered.

Table 4
Clinical signs of the top 7 most common foods.

Food	Cutaneous (%)	Digestive (%)	Respiratory (%)
Chocolate	77.8	11.1	11.1
Wheat	25	62.5	12.5
Berries	87.5	0	12.5
Milk	83.3	16.7	0
Egg	100	0	0
Fish	86.6	0	16.7
Beans	60	20	20

relation between the family history and the appearance of food allergy in children ($P = 0.0001$).

4. Discussion

In the present work, we conducted for the first time a cross-sectional study of the prevalence of food allergy in schoolchildren in the city of Guelma (Algeria). In our study, we examined food allergies without focusing on a single food in children aged from 5–13 years. The prevalence of parent-reported food allergy was estimated at 8.5% (95% CI 6.1–10.8).

4.1. Similarities of prevalence with previous studies

In Africa, children aged between 5 and 16 years old from 9 Ghanaian schools were recruited after getting parental consent to participate in the study. Adverse reactions and food consumption were determined by a questionnaire and atopy by skin prick testing (SPT) to peanut and 6 fruits. The prevalence of reported adverse reactions assessed by the questionnaire in the Ghanaian schoolchildren was 11.0% and 5% showed a positive SPT reaction [8]. In France, a cross-sectional, descriptive, questionnaire-based survey conducted in Toulouse (France) schools showed that the cumulative prevalence of food allergy in schoolchildren (2–14 years old) was 6.7% (95% (CI) 5.8–7.6) [9]. In European, a study aimed to determine the prevalence of self-reported food allergy, food sentitization, probable food allergy (symptoms plus IgE sensitization), and challenge-confirmed food allergy in European school-age children. The prevalence of self-reported food allergy ranged from 6.5% in Athens to 24.6% in Lodz showing large geographical differences in the prevalence of food allergy in school-age children across Europe [10].

The previous prevalences of food allergy from an African country (Ghana) and European countries are showing some similarities to the prevalence found in our study.

In the question about the allergy, we aimed to know current and outgrown allergies. After dividing the cases into 3 age categories, we expected that the prevalence of food allergy decreases with the age coupled with an increase in the prevalence of recovery, however the prevalence of recovery of the last category [11–13 years] dropped. We suppose that the recovery of food allergy reaches the maximum before 11 years old. We might suggest that, compared with younger children, the adolescents have persistent, possibly long life food allergies.

4.2. Risk factors influencing food allergy

According to our results the prevalence between girls and boys was not significantly different. In contrast, a review about the gender aspects in food allergy mentioned that gender differences likely play a role in food allergy development. Boys are more prone to have allergies before puberty whereas in adulthood women are more likely to have allergies [11]. In the other hand, the family history appeared to be the main risk factor influencing the appearance of food allergy in children (66.6% of the allergic children have a family background with the pathology). Similar results were found in other studies showing that family history is known to be one of the main factors in the appearance of food allergy [4,5].

While most studies focus on one food, our study gathered information about food allergy in the population of schoolchildren despite the food. Many reports on food allergies confirmed the relation between geographical region and the foods responsible for the allergy [10,12,13]. Some studies reported that atopy prevalence is higher in urban children than rural children [14,15]. Knowing that the population studied in our survey was urban, this factor can have an impact on our results.

4.3. Foods declared in food allergic children

Chocolate was found to be the most common food causing allergy to children in our study (1.7%). Nevertheless of the absence of evidence of food allergy to chocolate, according to the study in Salvadoran Schoolchildren, the prevalence of self-reported chocolate/cocoa allergy is 0.5% to 0.7% in selected populations [16]. Some

studies are suggesting that chocolate or cocoa allergy is a result of cross-contamination from some allergens such as: milk, tree nuts or peanuts [17]. The fact that chocolate may contain different allergens at once could explain its high prevalence. In addition to that, chocolate is considered to be a food that contains histamine releasers, when consumed it can cause pseudo-allergic reactions [18].

Milk, eggs, wheat and fish are known to be part of the most common causes of food allergies worldwide in children [19]. This suggestion was confirmed in our results where these foods were in the top eight allergy reasons with different prevalences.

The allergy to wheat affected 1.3% of the studied population. Knowing that wheat is associated to multiple disorders like: IgE mediated wheat allergy, non IgE mediated wheat allergy and celiac disease; it could explain the high prevalence of this allergy.

In a study on the risks of wheat allergy and early exposure to cereal grains, it has been shown that allergy to wheat is more frequent when wheat is introduced after 6 months of age into the children's diet [20]. This statement could explain the result in our study assuming that the Algerian diet involves eating wheat in different forms and dishes especially for the young aged kids. More than 62% suffered from digestive symptoms, then comes cutaneous and respiratory signs (25% and 12.5%, respectively). According to the literature, symptoms of IgE mediated wheat allergy may change with age [21].

According to our findings, the prevalence of milk allergy was 1.1%. The same result was found in cases follow up study. It showed that the prevalence of IgE-mediated cow's milk proteins allergy (CMPA) in Algeria was 1.1% [5]. In the same context, another study on 2015 revealed the prevalence of CMPA to be 3.64% in a pediatric population in Constantine, Algeria [4]. The allergy to milk is known to affect children at an early age [9], the median age of appearance of this allergy in our work was 1.2 years. This fluctuation could be explained by a late weaning in the Algerian society. Du to religious reasons [22], Algerian mothers tend to wean children around the age of 2 years old.

The prevalence of allergy to eggs was estimated at 1.1%. This food is also one of the most common foods causing allergy worldwide in children. The prevalence of eggs' allergy among kids varies from 0.3% to 9.9% according to several studies [10,15]. According to an inversigation in a population-based cohort study, the prevalence of challenge-confirmed food allergy to eggs at age 1 and 4 years was 1.2% (95% CI, 0.9% to 1.6%) [23]. Those findings are comparable to ours, showing that the studied population is in accordance with the international data.

According to our work, strawberry appeared to be one of the most common foods causing allergy in schoolchildren with a prevalence of 1.1%. In the literature, allergy to this fruit is often reported but rarely confirmed. The work of Collins et al. reported one case study of 9-year-old boy with a history of strawberry and raspberry anaphylaxis in France [24]. Among schoolchildren (3–11 years old) from Central Portugal, a study reported that the most frequent food group causing food allergy was fresh fruits including strawberry as the top fruit in the list [25]. The presence of histamine releasers in strawberry can also explain the prevalence of the allergy of this fruit. In certain subjects the consumption of this fruit can cause non-specific histamine-releasing resulting allergic responses [26,27].

The study made by Lyons et al. reported the prevalence (95% CI) of food sensitization to fish in some European cities. The prevalence was 0.91% (0.40–1.42%) in Madrid (Spain) and 0.76% (0.44–1.08%) in Vilnius (Lithuania) [10]. These results are comparable to ours, where the prevalence of parent-reported food allergy to fish was estimated at 0.9%.

4.4. Food allergy symptoms

Among the clinical manifestations, the dermatologic signs affected 68.6% of the children, followed by digestive signs (19.6%) then respiratory symptoms (11.8%). Many studies confirm these findings. For instance, the main clinical manifestation in the IgE mediated adverse food reaction in Portuguese children was mucocutaneous [25]. In the same context, another investigation reported that food allergy related adverse reaction in children having immediate food allergy was mainly skin symptoms (62%) [28].

Anaphylaxis is the most severe presentation of food allergy; it can ultimately lead to death if not treated. As a systemic reaction, anaphylaxis may present with symptoms targeting different organs [29]. According to our results, 13% of the children had cutaneous and respiratory issues, cutaneous and digestive symptoms or respiratory and digestive symptoms at once. Even though, no anaphylactic reaction was reported by the parents, but the presence of combined symptoms in those children could be a characteristic of anaphylaxis.

4.5. Limitations of the study

This type of studies has certain limitations and bias that we are aware of; especially the subjectivity of the answers. One of the main biases that could interfere in the prevalence of food allergy is the absence of a question about the diagnostic of the allergy in the survey. For instance, the presence of pseudo-allergic reaction like the ones caused by foods containing histamine releasers (chocolate, strawberries) can cause an overestimation of some allergies. Furthermore, some "yes" answers could be confusing food allergy with other reactions. An additional bias in such studies could be due to avoidance of eating. Some people simply refuse the consumption of some foods because they say they don't like them. Actually, sometimes they are allergic and instinctively avoid the consumption even if they are not aware of allergy. It happens, for instance, in people with digestive problems due to allergy.

5. Conclusion

The results showed for the first time the prevalence of parent-reported food allergy among schoolchildren in the city of Guelma (Algeria). The prevalence of reported food allergy was estimated at 8.5% (95% CI 6.1–10.8). On contrary to gender, family history was found to be the main risk factor influencing the appearance of food allergy in children. Overall, the results provided by the survey were comparable with the data available in the literature.

In contrast, an outstanding prevalence of food allergy to chocolate (1.7%) was found. Additional investigations should be done in order to provide more precise data on this food allergy. Clinical demonstrations such as skin prick test, search of specific IgE or Placebo-Controlled Food Challenges (DBPCFCs) should be used to confirm this allergy.

Disclosure of interest

The authors declare that they have no competing interest.

Acknowledgement

We sincerely thank all the primary schools' directors and teachers for facilitating the run of the study in the schools in a short time and during the period of exams.

We are also grateful to the parents for taking time to answer the survey's questions and giving it back in time.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at https://doi.org/10.1016/j.reval.2021.06.005.

References

[1] Hadley C. Food allergies on the rise? Determining the prevalence of food allergies, and how quickly it is increasing, is the first step in tackling the problem. EMBO Rep 2006;7:1080–3.
[2] Messina M, Venter C. Recent surveys on food allergy prevalence. Nutr Today 2020;55:22–9.
[3] Levin KA. Study design III: cross-sectional studies. Evid Based Dent 2006;7:24–5.
[4] Boughellout H, Benatallah L, Zidoune MN. Prévalence de l'allergie aux protéines du lait de vache chez des enfants âgés de moins de 3 ans de la ville de Constantine (Algérie). Rev Fr Allergol 2015;55:288–92.
[5] Ibsaine O, Djenouhat K, Lemdjadani N, Arrada Z, Berrah H. Etude cas témoins des facteurs de risque de lAPLV IgE-médiée. Arch Pediatr 2010;21:847.
[6] Latreche A. L'approche diagnostic de l'allergie alimentaire dans l'Est algérien [thèse]. Constantine: université Mentouri; 2009. p. 94.
[7] Charan J, Biswas T. How to calculate sample size for different study designs in medical research? Indian J Psychol Med 2013;35:121–6.
[8] Obeng BB, Amoah AS, Larbi IA, Yazdanbakhsh M, Van Ree R, Boakye DA, et al. Food allergy in Ghanaian schoolchildren: data on sensitization and reported food allergy. Int Arch Allergy Immunol 2011;155:63–73.
[9] Rancé F, Grandmottet X, Grandjean H. Prevalence and main characteristics of schoolchildren diagnosed with food allergies in France. Clin Exp Allergy 2005;35:167–72.
[10] Lyons SA, Clausen M, Knulst AC, Ballmer-Weber BK, Fernandez-Rivas M, Barreales L, et al. Prevalence of food sensitization and food allergy in children across Europe. J Allergy Clin Immunol Pract 2020;8:2736–46.
[11] Pali-Schöll I, Jensen-Jarolim E. Gender aspects in food allergy. Curr Opin Allergy Clin Immunol 2019;19:249–55.
[12] Rona RJ, Keil T, Summers C, Gislason D, Zuidmeer L, Sodergren E, et al. The prevalence of food allergy: a meta-analysis. J Allergy Clin Immunol 2007;120:638–46.
[13] Ontiveros N, Flores-Mendoza L, Canizalez-Roman V, Cabrera-Chavez F. Food allergy: prevalence and food technology approaches for the control of IgE-mediated food allergy. Austin J Nutr Food Sci 2014;2:1029.
[14] Chu ML. Prevalence and determinants of atopy among school-age children in rural Saskatchewan, Canada [thèse]. Saskatoon: University of Saskatchewan; 2014. p. 142.
[15] Sha L, Shao M, Liu C, Wu Y, Chen Y. A cross-sectional study of the prevalence of food allergies among children younger than ages 14 years in a Beijing urban region. Allergy Asthma Proc 2019;40.
[16] Cabrera-Chávez F, Rodríguez-Bellegarrigue CI, Figueroa-Salcido OG, Lopez-Gallardo JA, Arámburo-Gálvez JG, Vergara-Jiménez MJ, et al. Food allergy prevalence in salvadoran schoolchildren estimated by parent-report. Int J Environ Res Public Health 2018;15.
[17] Lopes JP, Kattan J, Doppelt A, Nowak-Węgrzyn A, Bunyavanich S. Not so sweet: true chocolate and cocoa allergy. J Allergy Clin Immunol Pract 2019;7:2868–71.
[18] Maintz L, Novak N. Histamine and histamine intolerance. Am J Clin Nutr 2007;85:1185–96.
[19] Ochfeld EN, Pongracic JA. Food allergy: diagnosis and treatment. Allergy Asthma Proc 2019;40:446–9.
[20] Poole JA, Barriga K, Leung DYM, Hoffman M, Eisenbarth GS, Rewers M, et al. Timing of initial exposure to cereal grains and the risk of wheat allergy. Pediatrics 2006;117:2175–82.
[21] Czaja-Bulsa G, Bulsa M. What do we know now about IgE-mediated wheat allergy in children? Nutrients 2017;9.
[22] Al-Jassir MS, El-Bashir BM, Moizuddin SK, Abu-Nayan AAR. Infant feeding in Saudi Arabia: mothers' attitudes and practices. La Revue de Santé de la Méditerranée orientale 2006;12.
[23] Peters RL, Koplin JJ, Gurrin LC, Dharmage SC, Wake M, Ponsonby AL, et al. The prevalence of food allergy and other allergic diseases in early childhood in a population-based study: HealthNuts age 4-year follow-up. J Allergy Clin Immunol 2017;140:145–53.
[24] Collins A, Derkenne B, Giebels K, Carvelli T. Allergie immédiate à la fraise et à la framboise. Rev Med Liege 2020;75:494–6.
[25] Jorge A, Soares E, Sarinho E, Lorente F, Gama J, Taborda-Barata L. Prevalence and clinical features of adverse food reactions in Portuguese children. Allergy Asthma Clin Immunol 2017;13.
[26] Lorenz AR, Scheurer S, Vieths S. Food allergens: molecular and immunological aspects, allergen databases and cross-reactivity. Chem Immunol Allergy 2015;101:18–29.
[27] Malone MH, Metcalfe DD. Histamine in foods: its possible role in non-allergic adverse reactions to ingestants. N Engl Reg Allergy Proc 1986;7:241–5.
[28] Ontiveros N, Valdez-Meza EE, Vergara-Jiménez MJ, Canizalez-Román A, Borzutzky A, Cabrera-Chávez F. Parent-reported prevalence of food allergy in Mexican schoolchildren: a population-based study. Allergol Immunopathol (Madr) 2016;44:563–70.
[29] Sharma HP, Bansil S, Uygungil B. Signs and symptoms of food allergy and food-induced anaphylaxis. Pediatr Clin North Am 2015;62:1377–92.

Chapter 2

Extraction of some food allergens and determination of some of their biochemical characteristics.

II.1. Introduction:

One of the main and important questions in the field of allergy is what makes some proteins become allergens? Some proteins have the ability to provoke an allergic reaction in certain individuals. Symptoms may differ from a simple skin rash to severe or even deadly anaphylaxis (Ito *et al.*. 2016).

Many scientific studies focused on allergens' characteristics like: stability, solubility, molecular properties and molecular size in order to have a link between those characteristics and allergenicity (Pekar. Ret and Untersmayr. 2018).

The comprehension of the characteristics might help predicting the allergenicity of some proteins. However, due to the complexity of the mechanisms involved in food allergy between the allergens and the immune system, those reactions are not easy to predict (Huby, Dearman and Kimber, 2000).

Many attempts for extracting allergens from different foods are available. Some researchers are focusing on finding the best extracting protocols. For instance, Abeyrathne and his collaborators reported two different methods for the separation of ovotransferrin and ovomucoid from chicken eggs: by high level ethanol method and acidic salt precipitation method (Abeyrathne, Lee and Ahn, 2014a). Another study of the same group aimed to present a simple and easy method to separate ovotransferrin without using organic solvents (Abeyrathne, H. Lee, *et al.*, 2013).

Others researchers are more interested in the biochemical and physicochemical characterization of the allergens. An investigation on parvalbumins from different fish species focused on the determination of their isoelectric points, molecular weight by SDS-PAGE and studied their thermostable capacity (Hasan Arif, Jabeen and Hasnain, 2007).

In the present work, we chose to focus on the main allergens of the most consumed allergenic foods among children in the city of Guelma. According to the results in chapter 1; Eggs, Cow's milk, sardine and wheat represent the top foods causing allergy in children being essential daily foods. The goals set in this study are:

- Extraction of ovalbumin and ovomucoid from eggs. α-lactalbumin. β-lactoglobulin from cow's milk. parvalbumin from sardine and gliadins from wheat.

- Biochemical characterization of the six allergens by the estimation of their ispoelectric points as well as the determination of their molecular weight using the electrophoresis profile of the SDS-PAGE.
- Identification of the proteins by HPLC/MS/MS and data analysis with Mascot software.
- Make the multiple sequence alignment of allergens to study the sequence similarity between them and check the presence of repeating sequences causing allergy

The extraction and the lyophilisation of the allergens was done in the university 8 Mai 1945 Guelma and the institute CRBT in constantine. The SDS-PAGE and the proteomics were done in the CNR (ibbr institute) in Italy.

II.2. Material and methods

1. Materials

Raw materials were: cow's milk provided from a local farm in El-fedjouj, chicken eggs purchased from a local food store, sardine bought from the fish market of Guelma and wheat a local variety of the region of Guelma known as "Hadba".

2. Extraction of food allergens

2.1. Cow's milk

1. Extraction of α-lactalbumin and β-lactoglobuline

250 ml of fresh cow's milk were heated to 40°C then 50 g of sodium sulfate (Na_2SO_4) were added. After dissolving the salt and when the temperature drops to 25°C, the solution is filtered. To 150ml of the filtrate, 1.5ml of concentrated hydrochloric acid was added while stirring vigorously. At pH close to 2, α lactalbumin forms a precipitate containing other proteins. The β lactoglobulin remains in solution. The precipitate and the β lactoglobulin are separated by centrifugation at 2000 rpm/min for 15 minutes. Little β lactoglobulin is dissolved in 15 ml of diluted ammonia. The final volume must be equal to 1/10 of the filtrate. The pH of the solution is brought to 3.5 with 0.1N hydrochloric acid. The β lactoglobulin remaining in solution is recovered by centrifugation at 2000 rpm / min for 15 minutes. The precipitate is dissolved in a volume of diluted ammonia equal to a quarter of the previous volume. The pH of the solution is brought to 4 with 0.1N hydrochloric acid. The α-lactalbumin is placed alone in the refrigerator overnight to sediment. It was recovered by centrifugation at 2000 rpm / min for 20 minutes (Souiki. 2000).

2.2. Chiken's egg

1. Extraction of ovalbumin from egg white

Approximately 250 ml of egg white (around 8 eggs) were diluted in the same volume of distilled water, stirred for 10 minutes then filtered. 100 g of sodium sulfate (Na_2SO_4) (Sigma Aldrich) were added to the filtrate. The solution was stirred for 75 minutes then centrifuged at 2000 rpm for 15 minutes. The supernatant contains ovalbumin which was recovered by heating at 70 ° C. A new centrifugation was carried out at 2000 rpm for 15 minutes. Ovalbumin remains soluble in the supernatant (Abeyrathne *et al.*, 2014).

2. Extraction of ovomucoid from egg white

40 ml of the egg white (around 2 eggs) were added to 240 ml of boiling distilled water slightly acidified with acetic acid prepared at 0.1% at pH around 4. After filtration, the filtrate was centrifuged at 2000 rpm for 15 minutes to recover the ovomucoid in the precipitate (Abeyrathne. Lee and Ahn. 2014a).

2.3. Sardine

1. Extraction of protein M

The muscle sample was minced and mixed with a mincer. 10 g of the mince was homogenized on ice with 9 volumes of 150 mM NaCl-10 mM phosphate buffer (pH 7.0. PBS). To prepare a heated extract, the homogenate was heated at 100 °C for 10 min in a heating block. Then, it was cooled on ice. The homogenate was centrifuged at 8.000g for 10 min. The supernatant was used as a heated extract containing the protein M. (Kobayashi *et al.*. 2016)

2.4. Wheat

1. Extraction of gliadins

50 g of milled wheat were added to 500 ml of 0.5 N NaCl. The solution was stirred for 2 hours and centrifuged at 3000 rpm for half an hour. 250 ml of 0.5N NaCl were added to the pellet collected. The solution was stirred for 1 hour and centrifuged at 3000 rpm for 30 minutes. To the residue containing the gliadins were added 150 ml of 70% ethanol and then the solution was stirred for 2 hours and centrifuged at 3000 rpm for 15 minutes. The supernatant contains soluble gliadins. Dialysis against cold distilled water for 24 hours allowed the recovery of gliadins (Mahroug. 2010).

All the allergens were lyophilized in the institute CRBT Constantine.

3. Protein estimation by BicinChoninic Acid assay (BCA assay)

The protein concentration of the samples was determined by micro-BCA (Bicinchoninic Acid Assay) according to the manufacturer's instructions. Micro-BCA Protein Assay Kit (Thermo Scientific), Bovine serum albumin (BSA) and MilliQ water were used for the assay (Appendix 2) (He. 2011).

3.1. Samples preparation

All the lyophilized allergens (α-lactalbumin. β-lactoglobulin. ovalbumin. ovomucoid. Protein M and gliadins) were solubilized in Milli-Q water to make stock solutions at a concentration of 1µg/µl.

1 µL of each sample was added to 9 µL of H2O and 10 µL of working reagent and incubated at 60° for 1 hour. The absorbance was determined with the spectrophotometer Nanodrop 1000 at wavelength 562 nm and the concentration was determined from the calibration curve.

4. Determination of the isoelectric point (pI) of the allergens

To each protein solution, either acetic acid (CH_3COOH) at 0.1M or sodium hydroxid (NaOH) at 0.1M was added slowly until the protein precipitates. The values were indicated in the pH meter that was previously equilibrated.

5. Sodium Dodecyl Sulfate Dodecyl polyacrilamide gel electrophoresis (SDS-PAGE)

SDS-PAGE is a very useful tool to separate proteins based on their molecular weight. SDS is a detergent that denatures secondary and nondisulfide-linked tertiary structures and coats them with a negative charge that correlates with their length allowing molecular weights to be estimated. SDS-PAGE system is a discontinuous gel with an upper stacking gel and lower resolving gel that have different pH values and polyacrylamide concentrations. The upper stacking gel has a lower percentage of polyacrylamide allowing proteins to move through quickly and ‘stack’ into a tight band before entering into the higher percentage polyacrylamide resolving gel for separation. (Brunelle and Green. 2014)

5.1. Procedure

Protein samples (20 µg) were loaded on a precast Novex 4-12% Bis-Tris NuPAGE gel (Invitrogen. Carlsbad. CA. USA) using MOPS running buffer according to the manufacturer’s instructions and stained with colloidal coomassie blue.

5.2. Preparartion of the tank and the gel

400 ml of MOPS SDS running buffer (X1) was prepared from the stock solution (X20) by diluting with milliQ water. The buffer was poured in the tank slowly to avoid the

formation of bubbles. The pre-casted gel was provided by Novex Native PAGE™ 4-16%Bis-Tris gel with 1.0mmX 10 wells. After the pre-casted gel was fixed in the tank, the comb was removed and the wells were one by one washed with the running buffer multiple times.

5.3. Samples preparation

The samples were prepared according to the manufactures‘ instructions. All the samples were dried in the savant. To every sample; 2.5 µl LDS, 1µl reducing agent and 6.5µl milliQ water were added and after a quick spin the tubes were incubated at 70°C for 10 minutes. Afterwards, the samples were centrifuged to cool down.

For the preparation of the standard, 7.5µl of the standard (thermo fisher) was added to 2.5 µl LDS. The volume loaded was 10µl for all the samples and the standard. The electrophoresis equipment was connected to the current (165A) for 35 to 40 minutes.

6. Determination of the molecular weight of the allergens

After separation on SDS-PAGE, the molecular weight of a protein can be estimated using its migration distance on the gel. Established around 1969, this method is tilled now utilized due to its simplicity. Nevertheless of some reports on the deviation of the molecular weight due to post-translational modifications of a protein. this method is reliable and used in biochemical works (Matsumoto. Haniu and Komori. 2019).

All the six allergens were run on SDS-PAGE with a standard with known molecular weight (MW). A graph of log MW versus the relative migration distance (R_f) was made based on the values measured from the migration of the bands of the standard (Appendix 3).

7. In-gel based proteomic analysis

The bands corresponding to the allergens were cut from the gel. The gel pieces were washed using 100 mM AMBIC in 50% ACN thrice for 10 min; once with 100% ACN for 10 min and then dried using a SpeedVac. Reduction was done using 10 mM DTT in 100 mM AMBIC in 5% ACN at 55°C for 1h and alkylated with 55 mM IAA in 100 mM AMBIC at room temperature for 30 min in dark. The gel pieces were washed with 100 mM AMBIC and dehydrated with 100% CAN and dried using the SpeedVac. The samples were rehydrated in a solution containing 6ng/µL trypsin (Promega. Madison. WI) in 50 mM AMBIC, 10% ACN for 30 min on ice and then 16h at 37°C. Peptides were extracted sequentially using 30% CAN/3.5% FA, 50%ACN/0.5% FA and 100% ACN.

The peptides were analyzed by nano-flow reversed-phase LC-ESI-MS/MS. 10 µL samples were loaded, purified and concentrated on a reversed phase monolithic pre-column. 200 µm ID x 5mm length (LC Packings. Sunnyvale. CA USA) at 25 µL/min flow rate. Peptides were separated at a flow rate of 300 nL/min on a PepSwift Monolithic column. 100µm ID x 5 cm (LC Packings. Sunnyvale. CA USA) using the following gradient: (solvent A: 2% CAN, 0.1% FA. Solvent B: 98% CAN. 0.1% FA) 5-50% B in 60 min. 5-98% B in 6 sec for 10 min. Eluted peptides were analyzed in IDA mode using the QSTAR Elite (Applied Biosystems Foster City. CA. USA) equipped with a nanoflow electrospray ion source.

8. Data Analysis

Analyst QS 2.0 software was used with default parameters to generate and to analyze peak lists extracted from IDA mass spectra. Mascot 2.2 was used to search data against SwissProt 2013-2014 database using trypsin with one possible missed cleavage. Proteins identified by in-gel digestion proteomics. carbamidomethylation of cysteine and oxidation of methionine were considered as fixed and variable modifications respectively. An analysis false-positive rate of the protein identifications was performed by searching all tandem mass spectra from the nano-HPLC-ESI-MS/MS analyses against an in-house curated decoy SwissProt human protein database containing forward and reversed sequences. In addition, contaminants such as human keratins and porcine trypsin were included in this database.

9. Multiple sequence alignment tools

Multiple sequence alignment is an important approach in comparative analyses of biological sequences. It also plays an important role in protein structural and functional analysis. Online align tool in UniProt was used to determine similar and identical sequences between the allergens. All the sequences were provides from the UniProt database.

Appendix 4 represents the single letter abbreviation used for the 20 amino acids found in proteins. In addition, pyrrolysine, used in the biosynthesis of proteins in some archaea and bacteria but not present in humans, and selenocysteine, a cysteine analogue only found in some lineages, are included. Finally, codes used for amino acid residues with more than one potential identity are shown to complete the alphabet of single letter abbreviations (Nigoskar, 2007).

10. Statistical analysis

The quantitative results are presented as means± SD. Every measurement was done in triplicate (n=3). Statistical significance ($p < 0.05$) was determined by Student's t test and ANOVA test. Student's t test was used to compare two independent means and the ANOVA test was used to compare more than two independent means. Tukey's HSD (honestly significant difference) test was used as post-hock to check the difference between every 2 proteins. The statistical analysis was done using XLSTAT 2016 (version 1802.01.28451).

II.3. Results and discussion

II.3.1. Results

1. Protein estimation of the extracted allergens

Table 4 shows the protein concentration of each allergen (mean±SD). Each sample was measured in triplicate

The concentrations of ovalbumin and ovomucoid were 1.63 mg/ml and 1.07mg/ml respectively. For the allergens from cow's milk; the concentration of α-lactalbumin was estimated at 0.5mg/ml and 0.62mg/ml for β-lactoglobuline. In the other hand, protein M extracted from sardine had a concentration of 0.009mg/ml. Finally, gliadins had a concentration of 0.28mg/ml.

Table 4. Values of the concentrations of the extracted allergens.

Allergens	Ovalbumin	Ovomucoid	α-Lactalbumin	β-Lactoglobuline	Protein M	Gliadins
Concentration (mg/ml)	1.63±0.15	1.07±0.02	0.5±0.094	0.62±0.24	0.009±0.001	0.28±0.003

2. Isoelectric point (pI) of the extracted allergens

The table 5 below shows the pIs values (mean±SD) of the six extracted allergens and the corresponding pI in the literature and the p-value given in the statistical test.

All the values were acidic in a range between 4.20±0.094 to 6.20±0.133. Ovalbumin had a pI of 4.56±0.260 and ovomucoid was 4.34±0.210. For α-lactalbumin and β-lactoglobuline their pIs were 4.20±0.094 and 4.95±0.062 respectively. Protein M had a pI of 4.76±0.057. When it comes to gliadins from wheat, the pI was estimated at 6.20±0.133

The comparison between every measured pI and its corresponding in the literature showed no significant difference ($p>0.05$).

Table 5. The measured values of isoelectric point (pI) for the extracted allergens.

Allergens	Ovalbumin	Ovomucoid	α-Lactalbumin	β-Lactoglobuline	Protein M	Gliadins
pI	4.56±0.260	4.34±0.210	4.20±0.094	4.95±0.062	4.76±0.057	6.20±0.133
pI in Theory	4.50	4.82	4.40	4.85	4.82	6.00
p-value	0.204	0.973	0.816	0.500	0.387	0.113

The statistical analysis made for the comparison between the pIs of the six allergens showed a significant difference between the pIs of the allergens $p < 0.0001$.

Tukey test was used as a post-hock to find means that are significantly different from each other (Appendix 5). It has been shown that the isoelectric point of gliadins has a highly significant difference from the other values ($p < 0.0001$). In the other hand no other significant differences were reported between other proteins.

3. Protein profile and molecular weight of the allergens by SDS-PAGE

The protein profile on SDS-PAGE for the six allergens is shown in appendix 6. This gel served to estimate the molecular weight of the extracted allergens. The standard (S) used has a range of molecular weight between 14KDa to 191KDa.

The results calculated of the molecular weight of the bands in every lane are collected bellow in the table 6. For ovalbumin and ovomucoid the molecular weight measured was 45.31 KDa and 34.38 KDa respectively. α-lactalbumin, β-lactoglobuline and protein M gave one band each with a molecular weight of 14.10; 17.88 and 13.20 KDa respectively. Finally for gliadins, there were multiple bands with a molecular weight ranging from 34.9 to 109.2KDa.

Table 6. The molecular weight estimated from the SDS-PAGE.

	OVA	OVO	ALA	BLA	Protein M	GLIA
R_F	0.637	0.677	0.950	0.883	0.970	0.672 0.646 0.560 0.512 0.429 0.390 0.312
MW (KDa)	45.31	34.38	14.10	17.88	13.20	34.9 37.9 49.8 58 75.4 85.3 109.2

4. Allergens identification

After running the generated spectrum from HPLC/MS/MS in Mascot software online, the results are gathered in Table 7.

Each allergen was presented with its identification (protein name, source), molecular weight (MW) and isoelectric point (pI). Both extracts from eggs and cow's milk were identified. Ovalbumine had a molecular weight of 43.196KDa and isoelectric point of 5.19. Ovomucoid was 20.223KDa of weight and an isoelectric point of 4.85. α-Lactalbumin and β-Lactoglobulin were identified with 16,247KDa and 19,883KDa for their molecular weight and 4.45and 4.80 for their isoelectric point respectively. Protein M from sardine was identified as Parvalbumin alpha from *Cyprinus carpio* with a molecular weight of 11.501 KDa and an isoelectric point of 4.43. For the extract from wheat, three proteins were identified with their molecular weight and isoelectric point: α/β-Gliadin, γ-Gliadin and ω-Gliadin (32,963KDa; 34,300KDa and 38,457KDa respectively).

Table 7. Mascot Results for the identified proteins.

Extract	Identification	MW (KDa)	pI
OVA	Ovalbumin *Gallus gallus* (Chicken)	43.196	5.19
OVO	Ovomucoid *Gallus gallus* (Chicken)	20.223	4.85
ALA	Alpha-lactalbumine *Bos Taurus* (Bovine)	16,247	4.45
BLA	Beta-lactoglobulin *Bos Taurus* (Bovine)	19,883	4.80
PV	Parvalbumin alpha *Cyprinus carpio*	11.501	4.43
GLIA	Alpha/beta-gliadin *Triticum aestivum* (Wheat)	32,963	6.75
	Gamma-gliadin *Triticum aestivum* (Wheat)	34,300	6.93
	Omega gliadin *Triticum aestivum* (Wheat)	38,457	6.84

5. Identification by protein sequences match

After running the generated results from HPLC/MS/MS in Mascot software online, it gave the amino acid sequences permitting the identification of the allergens. Figures 6, 7, 8, 9, 10, 11, 12 and 14 show the sequences for the identification of the allergens: Ovalbumin,

Ovomucoid, α-Lactalbumin, β-Lactoglobulin, Protein M, α/β Gliadin, γ-Gliadin and ω-Gliadin.

Figure 6 bellow shows the sequence given by mascot search engine. The Matched peptides are shown in red giving coverage of 36% with the sequence of Ovalbumin in SwissProt database.

```
  1 MGSIGAASME FCFDVFKELK VHHANENIFY CPIAIMSALA MVYLGAKDST
 51 RTQINKVVRF DKLPGFGDSI EAQCGTSVNV HSSLRDILNQ ITKPNDVYSF
101 SLASRLYAEE RYPILPEYLQ CVKELYRGGL EPINFQTAAD QARELINSWV
151 ESQTNGIIRN VLQPSSVDSQ TAMVLVNAIV FKGLWEKAFK DEDTQAMPFR
201 VTEQESKPVQ MMYQIGLFRV ASMASEKMKI LELPFASGTM SMLVLLPDEV
251 SGLEQLESII NFEKLTEWTS SNVMEERKIK VYLPRMKMEE KYNLTSVLMA
301 MGITDVFSSS ANLSGISSAE SLKISQAVHA AHAEINEAGR EVVGSAEAGV
351 DAASVSEEFR ADHPFLFCIK HIATNAVLFF GRCVSP
```

Figure 6. Protein sequence coverage for the extracted ovalbumin.

From over 200 sequences identified in the extract of ovalbumin, 50 sequences were identical shown in red color (figure 7) to the sequence on the database. The protein was identified as ovomucoid based on 24.8% sequence match.

```
1   MTLTLSHFGK AAFGAEVDCS RFPNATDKEG KDVLVCNKDL SFVLCGFLPD
51  RPICGTDGVT YTNDCLLCAY SIEFGTNISK PMNCSSYANT EHDGECKETV
101 TSEDGKVMVL CNRAFNPVCG TDGVTYDNEC LLCAHKVEQG ASVDKRHDGG
151 CRKELAAVSV DCSEYPKPDC TAEDRPLCGS FCNAVVESNG DNKTYGNKCN
201 CAMAGVFVLF
```

Figure 7. Protein sequence coverage for the extracted ovomucoid.

The similar suequence are highlits in red (figure 8) permitting the identification of α-Lactalbumin. 53 sequences from 142 (36.3%) were similar to the sequence available in the SwissProt database.

```
1   MMSFVSLLLV GILFHATQAE QLTKCEVFRE LKDLKGYGGV SLPEWVCTTF

51  HTSGYDTQAI VQNNDSTEYG LFQINNKIWC KDDQNPHSSN ICNISCDKFL

101 DDDLTDDIMC VKKILDKVGI NYWLAHKALC SEKLDQWLCE KL
```

Figure 8. Protein sequence coverage for the extracted α-Lactalbumin.

β-Lactoglobulin was identified based on the sequence similarity with a percentage of 17.41% match.

```
1   MKCLLLALAL TCGAQALIVT QTMKGLDIQK VAGTWYSLAM AASDISLLDA

51  QSAPLRVYVE ELKPTPEGDL EILLQKWENG ECAQKKIIAE KTKIPAVFKI

101 DALNENKVLV LDTDYKKYLL FCMENSAEPE QSLACQCLVR TPEVDDEALE

151 KFDKALKALP MHIRLSFNPT QLEEQCHI
```

Figure 9. Protein sequence coverage for the extracted β-Lactoglobulin.

For the allergen extracted from sardine. The protein sequence is shown in figure 10. Matched peptides are in red. The coverage was estimated at 11%. The result identified the allergen based on sequence similarity to Parvalbumin alpha from *Cyprinus carpio*.

```
  1 MAYGGILNDA  DITAALEACX  AXDSFNAKSF  FAKVGLSAKT  PDDIKKAFAV
 51 IDQDKSGFIE  EDELKLFLQN  FSAGARALTD  AETKAFLKAG  DSDGDGKIGV
101 DEFAALVKA
```

Figure 10. Protein sequence coverage for the extracted Parvalbumin.

The figures 11, 12, 13 represent the proteins identified from the extract of wheat. It has given three types of gliadins (α/β gliadins, γ-gliadin, ω-gliadin).

The sequence comparison showed 20.19% similarity to Alpha/beta-gliadin from wheat highlighted in red color. From a total of 307 sequences, 62 sequences were matching to alpha/beta gliadin in SwissProt database.

```
1   MKTFLILALL VRVPVPQLQP QVPLVQQQQF QNPSQQQPQE AIVATTARIA

51  AIVATTARIA SQQPYLQLQP QPYPQPQPFP FPQPQLPYPQ PQLPYPQPQL

101 AIVATTARIA QQPYPQSQPQ YSQPQQPISQ QQQQQQQQQQ QKQQQQQQQQ

151 AIVATTARIA ILQQILQQQL YSIAYGSSQL LQQSTYQLVQ QLCCQQLWQI

201 PEQSRCQAIH NVVHAIILHQ QQQQQQQQQQ PQQQYPSGQG PQFEEIRNLA

251 CRQPSQQNPQ VQPQQAQGSL QQQQQQQQQQ LETLPAMCNV QPLSQVSFQQ

301 SFIFGTN
```

Figure 11. Protein sequence coverage for the extracted α/β-Gliadin.

Furthermore, figure 12 shows the identification of γ-Gliadin. The sequence similarity has been estimated at 18.21%. The matching sequences are written in red color.

```
1   MKTLLILTIL MQVDPSGQVQ AMATTIATAN QQPFCQQPQR WPQQQPFPQP

51  KTLPTMCNVY HQPQQTFPQP KTLPTMCNVY QFPQTQQPQQ PFPQPQQTFP

101 PQQPFPQPQQ PQQPFPQSQQ QQPQLPFPQQ PQQPFPQPQQ PFPQPQQTFP

151 MQQQCCQQLA QSFLQQQMNP PFPQPQQTFP PFPQPQQTFP IILPRSDCQV

201 SFPQQQQPAI QIPQQLQCAA IHSVAHSIIM QQEQQQGVPI LRPLFQLAQG

251 LGIIQPQQPA QLEGIRSLVL QQTYPHQPQQ VPPDCSTINV IHSVAHSIIM

301 QG
```

Figure 12. Protein sequence coverage for the extracted γ-Gliadin.

Finally, the identification of ω-Gliadin was made based on 63 sequences similar from a total of 354 sequences. The sequence match was estimated at 17.8%.

```
  1  MKPHHDGYKY TCSIIVTFHY PNFKHQDQKH QFQESIKHKS LSMPMSIVIA

 51  KMKTFIIFVL ARHLNPSDQE LQSPQQQTIS KTIISAAFLE TSTIFTTTTI

101  ATTTTISLAP TNPPTTTMTI TTTTISPTPT FSPAPTTISL PLATPTTTTT

151  SHTPTIFPPS TTNSPITTTT YASTATTISL TIPPATRTNN IPAATPETTT

201  TTIPPATRTN PTTILSATTT PATRTNNSLA TISPAPTIIS LTATTTPPAT

251  AQQPFPQQPG NSPATATTIP PAPQQRFPHT RQKFPRNPNN HSLCSTHHFP

301  TPTTIPPATA QIIPQQPQQP LPLQPQQPFP WQPEQRSSQQ PQQPFSLQPQ

351  FSQP
```

Figure 13. Protein sequence coverage for the extracted ω-Gliadin.

6. Multiple sequence alignment

Appendix 7 represents the multiple sequence alignment for all the allergens together. Comparing all the sequences by the multiple sequence alignment algorithm tools in UniProt didn't reveal significant similarities (0 similar positions, 0% identity).

In order to see the similarities in more close proteins, like allergens from the same source, we made the sequence alignments for ovalbumin and ovomucoid then between α-Lactalbumin and β-Lactoglobulin. Then, a comparison between the allergens from animal souce (Ovalbumin, Ovomucoid, between α-Lactalbumin, β-Lactoglobulin and Parvalbumin) was also made. Figures 14, 15 and 16 represent the multiple sequence alignments resulted from Uniprot align tool.

The sequence alignment between Ovalbumin and Ovomucoid gave a percentage of 7.44% identity with 34 identical positions (*). There were 49 similar positions represented by 27 conservative replacements (**:**) and 22 semi conservative mutations (**.**).

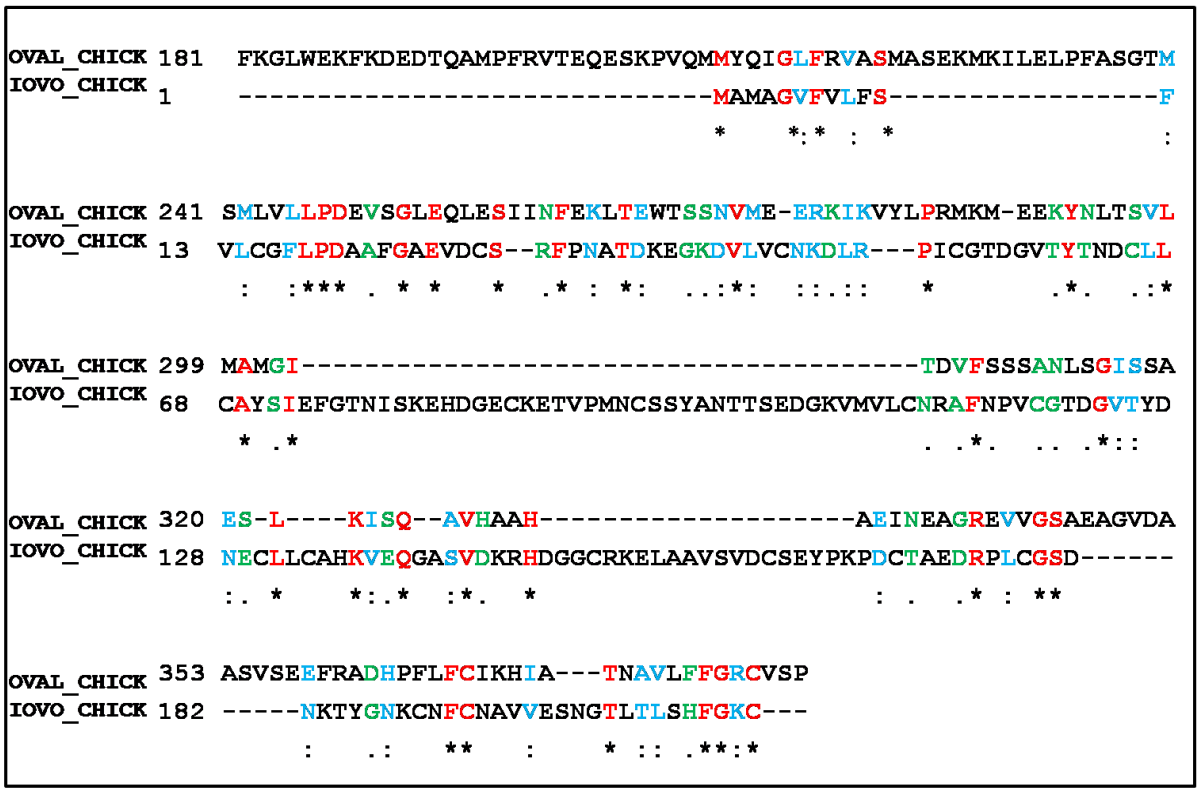


Figure 14. Sequence alignment of ovalbumin (OVAL_CHICK) and ovomucoid (IOVO_CHICK).

In the other hand, the sequence alignment between the two allergens from cow's milk (α-Lactalbumin and β-Lactoglobulin) showed an identity percentage of 13.706%. It had 27 identical positions shown in the figure 15 with a (*). A total of 49 similar positions were noted. It is presented by 30 conservative replacements shown with two points (**:**) and 19 semi conservative mutations shown with a point (**.**).

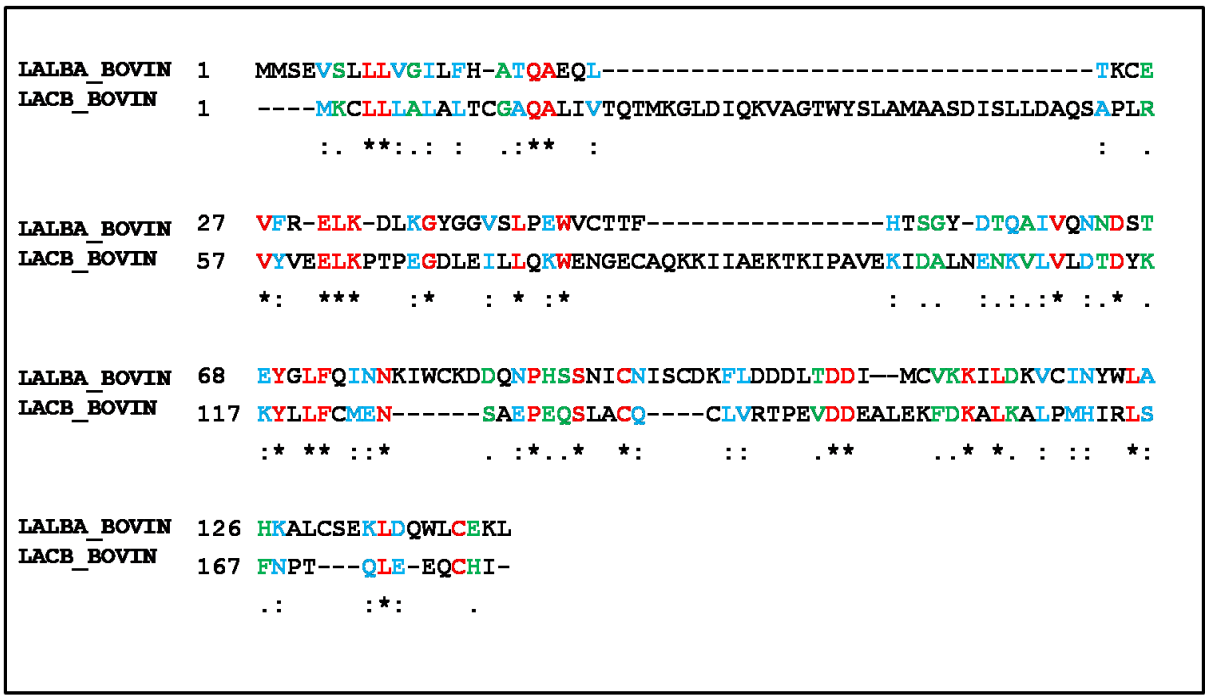


Figure 15. Sequence alignment of α-Lactalbumin (LALBA_BOVIN) and β-Lactoglobulin (LACB_BOVIN).

The sequence alignment of the allergens from animal source (Ovalbumin, Ovomucoid, α-Lactalbumin and β-Lactoglobulin and Parvalbuin) gave 0% identity with only 4 similar positions (one conservative replacements (**:**) and three semi conservative mutations (**.**)).

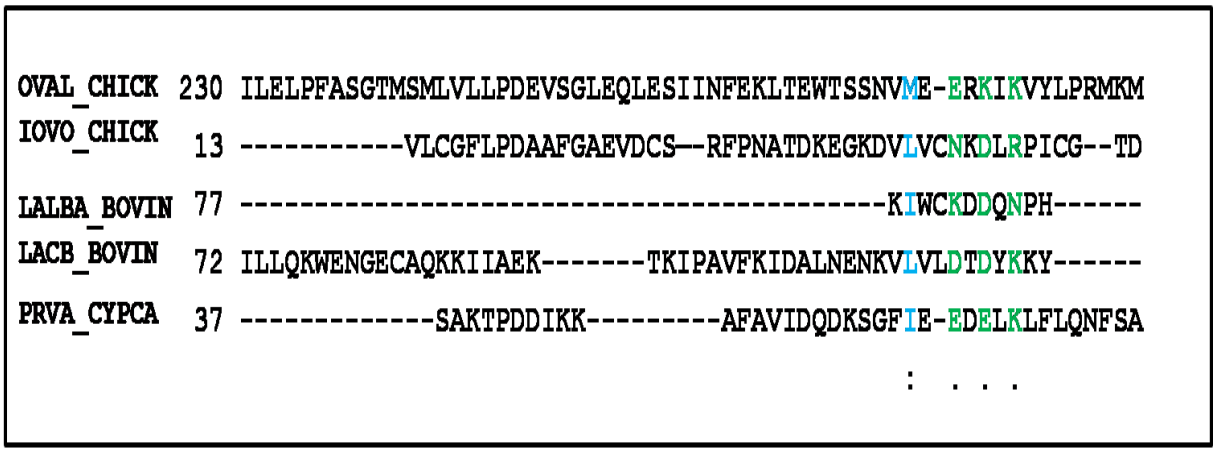


Figure 16. Multiple sequence alignment of allergens from animal source.

II.3.2. Discussion

During this study, we extracted 6 different allergens from both animal and plant sources: ovalbumin and ovomucoid from egg white. α-lactalbumin and β-lactoglobulin from cow's milk, protein M from sardine and gliadins from wheat.

Egg white contains over than 40 proteins from which: ovomucoid. ovalbumin. ovotransferrin. and lysozyme have been recognized as the four major allergens (Ma *et al.*. 2020). Ovalbumin (Gal d 2) represents 54% of the total protein content. It has 4 sulphydryl groups with a single disulphide bridge and it is found to be sensitive to heat denaturation (Verhoeckx *et al.*. 2015). Ovomucoid (OVO; Gal d) representing 11% of the total protein content consists of 3 subdomains, each one is internally linked by disulphide bonds which explains its high resistance to heat denaturation and proteolytic digestion. Domain 3 is very stable. It has most dominant IgE and IgG-binding epitopes and is considered the major determinant of the strong allergenicity of the protein (Verhoeckx *et al.*. 2015).

When it comes to the isoelectric point of ovalbumine, it was determined at 4.56±0.260 according to our results. This value is supported by the ones in the literature (pI=4.5) (Abeyrathne *et al.*. 2014). Likewise for ovomucoid, the measured pI (4.34±0.210) was in accordance with the values reported by other works (Abeyrathne *et al.*. 2013). Both results showed no significant difference to the values in the literature (p=0.204 and p=0.973 respectively.

For the molecular weight deduced from the SDS-PAGE, the measured molecular weight of ovalbumin was 45.31KDa. This finding is in perfect accordance with the results of many reports (Mine, 1995; Liu *et al.*, 2020); showing that the molecular weight of this protein is around 45 KDa.

On the other hand, the molecular weight measured for ovomucoid was 34.38 KDa. This protein was shown to have a real molecular weight at a range of 20KDa to 30KDa. Although, it has been reported that this protein appears in the SDS-PAGE at a range between 30-40 KDa (Abeyrathne. Lee and Ahn. 2014b). This statement has been confirmed in our results.

For proteins from cow's milk: α-lactalbumin (a-Lac. Bos d 4) and β-lactoglobulin (b-Lg. Bos d 5) are considered to be the most important allergens in milk. It has been reported

that more than 50% of the individuals with cow's milk allergy are sensitized to those proteins (Tsabouri. Douros and Priftis. 2014). α-lactalbumin is a small. monomeric Ca2+-binding protein. In addition to calcium binding, it has four disulfide bridges stabilizing the structure of the molecule (Linhart *et al.*. 2019). β-lactoglobulin is present in the milk of many mammals, except for humans. This allergen belongs to the family of lipocalins, made of nine antiparallel β-strands and one α-helix (Varlamova and Zaripov, 2020).

When it comes to our results for the allergens from cow's milk, both proteins showed comparable values. α-lactalbumin had a pI of 4.20 and β-lactoglobulin had a pI of 4.95. Those results are in agreement with the results in the literature where α-lactalbumin has a pI between 4.2–4.6 and β-lactoglobulin has pI of 4.85 (Wal. 2001; Uniacke-Lowe and Fox. 2011).

For the measurements of the molecular weight of those proteins, the results were 14.10KDa and 17.88 KDa for α-lactalbumin and β-lactoglobulin respectively. Compared to the theoretical values, our results are in accordance proving a high match-up between the measured and theoretical molecular weight being 14KDa and 18KDa for α-lactalbumin and β-lactoglobulin respectively (Wal. 2001; Uniacke-Lowe and Fox. 2011).

In fish, the dominant allergen is the homologues of Gad c 1 from cod, formerly known as protein M. A close cross-reactivity exists within different species of fish between this protein, denominated the parvalbumins (Poulsen *et al.*, 2001).

Parvalbumin is a small, water-soluble, calcium-binding muscle protein involved in the muscle relaxation process belong to the EF-hand protein superfamily. α- and β-parvalbumin were identified as two separate phylogenetic lineages of parvalbumin (Lee *et al.*, 2011). According to a mini review on 2014, the official database of allergens (the International Union of Immunological Societies Allergen Nomenclature Subcommittee database) contains 21 parvalbumins from 12 fish species including: sardine (*Sardinops sagax*) and common carp (*Cyprinus carpio*) (Kuehn *et al.*, 2014)

The parvalbumin extracted from sardines had a pI of 4.75. This value exhibited no significant difference (p=0.387) with the pI reported by previous studies (pI 4.55–5.10) (Hasan Arif. Jabeen and Hasnain. 2007). The molecular weight measured from the SDS-PAGE was 13.20KDa. Similar results were reported in other studies qualifying parvalbumin as the major fish allergens (PVs) with a molecular weight between 11–13 kDa (Van Do *et al.*, 2005; Kalic *et al.*, 2019).

Gliadins are part of the wheat gluten. They account for 40–50% of the total storage proteins of wheat and they are classified into four subcategories, α-, β-, γ -, and ω-gliadin. Gliadins are usually regarded to have globular protein structure, however recent studies report that the α/β-gliadins have compact globular structures and γ - and ω-gliadins have extended rod-like structures (Mahroug, 2010; Schalk *et al.*, 2017).

When it comes to the characterization of gliadins extracted from wheat, the pI was 6.20. The value showed no significant difference (p=0.113) to the values in the literature (pI 6-8.24) (Dziuba *et al.*, 2014).

For the MW of gliadins, multiple bands were visible on the SDS-PAGE. It had a range from 30 to 40 kDa. According to other studies, that of ω-GLIAdins. range from 40 to 50 kDa although their apparent MW in SDS-PAGE is much higher (55±75 kDa) (Battais *et al.*. 2003).

All the values of pI of the studied allergens were acidic with a range between 4.20 to 6.20. These proteins have the same form of ionization. Remarkably, the pI of gliadins was significantly higher than the other allergens ($p < 0.0001$).

The acidic pI is usually a characteristic of allergens. An *in silico* study tried to predict the allergenicity of proteins based on their physical properties. The results showed that the calculated net charge for the entire protein in a neutral aqueous solution was more negative for allergens compared to non-allergens with a significant acidic pI for all the tested allergens. It has been reported that 88% of the tested food allergens had an acidic pI (5.6; $P<0.0001$) (Singh *et al.*. 2009).

The physical properties of food allergens like the isoelectric point, molecular weight and structure give them their specific allergenic characteristics like being resistant to heat, to proteolytic, to acidic conditions and to digestive enzymes (Jeong. 2016).

Food allergens are known to be stable to the proteolytic and acidic conditions of the digestive tract. This increases their probability of reaching the intestinal mucosa, where absorption occurs. Even though stability of allergens has been demonstrated only little information is known about why these proteins have the ability to resist degradation (Pekar, Ret and Untersmayr, 2018).

Due to these characteristics, those allergens are able to get to the gut where the absorption happens. Trophallergens are even able to attack the immune system by elements other than antigens. The simple example is β-lactoglobulin in milk. This one is able to bind to phosphatidylcholine when it is not fully digested causing the persistence of its antigenicity (Jeong. 2016).

On the other hand, the results generated from the HPLC/MS/MS and Mascot search engine confirmed the efficiency of the protocols of extraction used in our work to extract the targeted proteins with sufficient amounts.

All the allergens were confirmed by isoelectric points, molecular weight and amino acid sequence match except for parvalbumin from sardine. Based on amino acid sequence, it was identified as parvalbumin from *Cyprinus carpio.* It could be explained by the contamination of the sardines bought from the market with different types of fish; in this case: *Cyprinus carpio.* In addition, a search for parvalbumin from sardine in the database of SwissProt, allergome and UniProt revealed that the total sequence was not available. This could be another reason why it was not identified in our result.

In most cases, food allergens have been characterized to be water-soluble glycoproteins with a molecular weight of 10 to 70 kDa; However. many food allergens are found not to share such characteristics (Bøgh and Madsen. 2016).

According to a review dealing with effects of chemical, physical and technological processes on the nature of food allergens, the molecular weight of allergens was given to be between 5-70KDa in monomers and can reach more than 200KDa in case of oligomers (Poms and Anklam. 2004).

Multiple sequence alignments are one of the most helpful tools in bioinformatics. They are used when sets of homologous sequences are compared and are important base to many additional analyses (Wallace, Blackshields and Higgins, 2005).

When it comes to our results of the multiple sequence alignments, no similarities were revealed between the six studied allergens.

Based on the absence of similarities in the primary structure of these proteins, we can suggest the absence of a repeating sequence causing the allergy. At the same time, similar sequences between an allergen and a proteins doesn’t mean necessary that they are both

allergens or because they are all allergens they will necessary share similar sequences. The best example is α-lactalbumin from bovine and from human. They show a high amino acid sequence similarity that exceeds 70% but only the one from bovine is an allergen (Linhart *et al.*. 2019).

In earlier studies, major allergens from peanut were isolated and their sequencing was made and the peptides that react with IgE from patients were identified. It has been found that similar proteins to those allergens were present in other foods (tree nuts and soy) known to cause cross reactions in peanut allergic patients (Schein, Ivanciuc and Braun, 2007).

Protein stability, cleavage sites, post-translational modifications and physico-chemical properties are known factors contributing in protein allergenicity. However, allergens must be recognized by the cells of the immune system. This specific immune recognition is controlled by the amino acid sequence and the 3D structure of the proteins (Maurer-Stroh *et al.*, 2019).

II.4. Conclusion

The different allergens from hen's egg (ovalbumin. ovomucoid), cow's milk (α-lactalbumin, β-lactoglobulin), sardine (parvalbumin) and wheat (gliadins) were successfully extracted. The extraction protocols' used were basic ones through solvents precipitation or salt precipitation with ammonium sulfate.

The determination of the isoelectric point of the extracted allergens revealed acidic values for all of them. They had values from 4.20 for α-lactalbumin from milk to 6.20 for gliadins from wheat. These results are supported by the fact that allergens are characterized with an acidic isoelectric point. It shows also that those proteins have the same forme of ionization.

Except for ovomucoid, the estimation of the molecular weight of the different allergens based on the appearing bands on SDS-PAGE gave very close results to the real proteins. For the case of ovomucoid, this protein is known to figure with a different molecular weight in SDS-PAGE around 30-40KDa, however its real molecular weight is around 20KDa. The results of the molecular weight given by HPLC/MS/MS were also in accordance with the known values of those proteins.

The use of Mascot search database linked to the results from HPLC/MS/MS confirmed the identity of the extracted allergens with the exception of parvalbumin from sardine. This one was identified as parvalbumin from common carp. This result could be explained by the contamination of the sardine bought with common carp. Another suggestion is the absence of the amino acid sequence of parvalbumin from sardine in this case our extract was identified based on similarity to available sequences.

The multiple sequence alignment between all the allergens (ovalbumin and ovomucoid, α-lactalbumin and β-lactoglobulin, parvalbumin and gliadins) showed no similarity. However, it was found between proteins from the same source like ovalbumin and ovomucoid with a similarity of 7.44% and α-lactalbumin and β-lactoglobulin with a similarity of 13.706%.

These results represent a solid start for future investigations. It is interesting to make a total characterization of the parvalbumin from sardine and determine its amino acid sequence.

In addition to that, the determination of the 3D structure trough X-ray crystallography could give promising results.

Chapter 3

Evaluation of potential allergenicity of protein extract from *Medicago sativa*.

III.1. Introduction

Till now, the avoidance diet is the only treatment for food allergies (Ochfeld and Pongracic, 2019). Many studies have reported observations on poor growth and nutritional deficiencies caused by the avoidance diets in children (Mehta, Groetch and Wang, 2013; Giovannini *et al.*, 2014; Mehta *et al.*, 2014; Sharma, Bansil and Uygungil, 2015). Monitoring growth and guiding food allergic patients in choosing appropriate alternatives to supply necessary nutrients becomes crucial to avoid deficiencies and retardation in growth. A long record of attempts to substitute proteins originating primarily from legumes, cereals, cassava, leaf proteins and whole plant lucerne proteins for animal proteins in the human diet exists (Gaweł, Grzelak and Janyszek, 2017).

Alfalfa (*Medicago sativa*) meaning "father of all foods" also known as lucerne 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 grams per kilogram of shoot dry weight (Livia Apostol, Sorin Iorga, Claudia Mosoiu, Radu Ciprian Racovita, Oana Mihaela Niculae, 2017). Alfalfa contains many vitamins, secondary metabolites and reported as having phytobiotic activity even in small quantities It has been used in different fields in particulary: animal feed, soil improvement and medicinal uses (Gaweł, Grzelak and Janyszek, 2017). It is available in the market in multiple forms like: protein extract and powdered aerial part-based human nutritional supplements (capsules, tablets, and powder) in sandwiches and salads as raw sprouts or even cooked and in the form of infusions from leaves and seeds (Mielmann, 2013).

Besides of the economic and agricultural importance of alfalfa, allergenicity assessment of the whole plant and its protein extract is very limited where only few fragmentary data are available today in the literature. A study reported that 22.9% of a population suffering from asthma and living in hot climate of Arabian Desert were tested positive to alfalfa (Bener *et al.*, 2002). Another study based on skin prick test (SPT), shows that 36% of atopic patients living in Saud Arabia reacted to alfalfa (Suliaman *et al.*, 1997). IgE co-reactivity between peanut and alfalfa was also observed using different immunological methods,(Jensen *et al.*, 2008) thus suggesting allergen sharing between these two species. However, a search in the WHO/IUIS Allergen Nomenclature (Pomés *et al.*, 2018) and in the Allergome database (Mari *et al.*, 2006) reveals no allergens identified in alfalfa so far.

In order to describe the allergenic profile of a any food, it has to be conducted with multidisciplinary approaches (Ciardiello *et al.*, 2013). Two criteria have to be found to qualify a food protein as a cause of IgE-mediated allergy: in one hand, the recognition by specific IgE. In the other hand, the induction of 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), skin prik test (SPT) and the provocation test (double-blind placebo-controlled food challenge, DBPCFC). DBPCFC is considered as the gold standard for the diagnosis of allergy (Cerecedo *et al.*, 2014) and to confirm that a molecule is an allergen on the basis of clinical reactivity. However, this procedure is limited especially by ethical issues due to important safety risks as the procedure can trigger anaphylaxis (Oriel and Wang, 2019). As a result many food allergens remain untested in DBPCFC and their characterization rely on alternative methodologies. Therefore, the adoption of new strategies based on a dynamic combination of different methodologies, spanning from the classic biochemistry-based ones to the innovative microtechnologies and bioinformatics, could provide the best results giving a collection of data useful to obtain as reliable as possible estimations of the potential allergenicity of a food source (Ciardiello *et al.*, 2013).

The objective of this study is to investigate the allergenicity of *M. sativa*, we have used a multidisciplinary approach. The study analysed the following points:

- Identification of the most abondant proteins in the leaf protein extract of *M. sativa*
- *In silico* analysis using proteomics-based protein identification, immunological and Blast2Go bioinformatics to search for hologous sequences
- *In vitro* test using IgE-binding inhibition test using the SPHIAa method associated to FABER test to check the presence of potentiel allergens in the leaf extract of . *M. sativa*
- Comparaison between the results given by the *in vivo* and *in silico* tests.

This part of the work was done in the CNR (ibbr institute) in italy during PNE grant.

III.2. Material and methods

1. Sample preparation

Alfalfa (*Medicago Sativa*) was provided from a local farm in the region of Guelma (East Algeria) The fresh leaves of *Medicago sativa* were manually separated from the stems, washed and dried on tissue paper in a stove at 38°C for 48 hours. 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 dark until use.

2. Total protein extract

Protein extract was prepared from the dried powder of alfalfa. 20g of ground dried leaves were homogenized in a solution made with 80 ml of 0.5 M NaCl, 0.8 g of poly (vinylpolypyrrolidone) (Sigma, Steinheim, Germany), 2 mL of 0.125 M ethylene diamine tetraacetic acid (EDTA) (Sigma, St. Louis, USA), and 2.2 g of ascorbic acid then brought to pH3.5 using NaOH.

The powder was extracted for 2 hours under stirring in an ice water bath. The sample was centrifuged at 10400 x *g* for 1 hour and about 100 mL of the supernatant were collected. The supernatant collected represents the protein extract.

In order to improve the color of the extract and have a clear colored one, the solution was divided into two parts. Part 1 was precipitated by 90% (w/v) ammonium sulfate (NH4)2SO4. Part 2 was purified using 3 different ionic exchange chromatography.

A. Ammonium sulfate precipitation

In brief, 30.15g of (NH4)2SO4 was added to 50ml of the extract, agitated for 3 hours then centrifuged at 12500rpm/ 1 hour at 4°C. The pellet obtained was resuspended in 8 mL of water and dialyzed against 50 mM NaCl for 2 hours. The dialyzed sample (18 mL) was filtered with a 0.22 µm filter (Millex, Millipore, Bedford, MA, USA).

The filtrate was purified in Diethylaminoethyl cellulose (DEAE-C) chromatography. The elution was done with 0.5M NaCl and gave 3 fractions of 5ml each.

B. Ionic exchange chromatography

Part 2 of the extract was dialyzed against 10mM Tris-Cl for 4 hours with change of the buffer every 30 minutes. Then it was purified in 3 steps using 3 different chromatographies.

A. Diethylaminoethyl cellulose (DEAE-C) eluted with 0.5M NaCl. The volume recovered was 20ml representing the bound proteins.
B. The flow through was purified in carboxymethyl CM chromatography. The buffer used was 10mM NaAc to equilibrate the resine. The elution was done by 0.5M NaCl, recovered 12ml as bound proteins.
C. The flow through was purified in sulfopropyl (SP) Sepharose colomn. The resine was washed with 50mM NaAc and 500mM NaCl. The elution was done by 0.5M NaCl. The volume recovered from this column was 10ml considered as the bound proteins.

The elutions from the 3 columns were poured together and dialyzed against 50mM NaCl to recover a final volume of 27.5ml. Afterwards, it was concentrated in amicon tube (3000 MWCO) by centrifuging at 3500rpm until getting a final volume of 5ml.

In every step, the protein concentration was determined by the Bio-Rad protein assay (Biorad, Milan, Italy) using a calibration curve made with Bovin Serum Albumin (BSA). The sample was aliquoted and stored at -20°C until use.

3. Protein dosage

The protein concentration of the extract was estimated with Bradford method (Bio-Rad protein assay) using bovine serum albumin (BSA) as standard (Appendix 2).

The principle of this method is based on the absorbance of Coomassie Brilliant Blue G-250 dye from 465 to 595 nm (maximum absorbance is at 595 nm) (Kielkopf, Bauer and Urbatsch, 2020). The Coomassie Blue G-250 dye originally reddish/brown changes its color to bleu due to the acidic conditions caused by the reaction with arginine, lysine, histidine, tyrosine, tryptophan, and phenylalanine residues in proteins. Hydrophobic interactions are also taking place (Goldring, 2019). It is an easy, fast and sensitive method allowing the determination of the concentrations of proteins.

3.1. Sample preparation

20µl of the extract was added to 980µl of milli-Q water and 250µl of Bio-Rad Dye. For the preparation of the blank; 1ml of milli-Q water was added to 250µl of Bio-Rad dye. After 10 minutes of incubation the absorbance was measured at 595nm.

4. IgE inhibition experiments

4.1. SPHIAa method

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. Pooled characterized sera containing the desired IgEs for the recognition of specific allergenic proteins were used in the assay.

4.2. Informed consent of patients

All the patients presented their informed consent for the use of their clinical data for research purposes anonymously. 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.

4.3. Patients' sera

Sera used in this study were selected among those stored in the serum bank of Allergy Data Laboratories s.r.l. These are residual sera deriving from venous blood sampling performed for the routine allergy diagnosis by FABER test. 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 s.r.l).

The choice of sera was selected exclusively on the IgE content. In particularly, for this study we have selected sera, which were able to recognize and bind relevant purified allergenic proteins from plant and animal foods available on the FABER biochip. In the SPHIAa experiments, IgE is used as a probe to detect the presence of structural determinants, i.e. 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. 0.12 mL of the pooled sera was mixed 1:1 with alfalfa protein extract resulting in a 1:8 final dilution of the individual serum.

4.4. Procedure

In the assay, pooled characterized sera 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. 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 s.r.l.) Only inhibition values equal or higher than 30% were considered in this study.

5. Gel-based proteomic analysis

5.1. SDS-PAGE

Leaf protein extract of *Medicago 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.

5.2. In gel digestion

In gel digestion was done according to the protocol described by Shevchenko.S *et al* on 2007. The entire gel lane was cut into 5 mm bands, proteins were reduced, alkylated, in-gel digested using 20 µL of trypsine (Promega, Madison, WI, USA) at 6 ng/µL concentration and extracted as follows:

1. Destain gel pieces excised from Coomassie-stained gels:

Each band was put in a tube with 300µl of Ammonium Bicarbonate (100mM) then in sonic bath for 15 minutes. After removing the Ammonium Bicarbonate (AMBic), 300µl of Ammonium Bicarbonate (100mM), 50% Acitonitril then left in sonic bath for another 15 minutes. After removing the Ammonium Bicarbonate (100mM), 50% Acitonitril, 300%l of acetonitril 100% was added, left under the hood until the bands shrank and became white. The acetonitril was removed and the tubes were left under the hood to remove the acetonitril. 30µl of DTT (10mM) in Ammonium Bicarbonate (100mM, 5% Acetonitril) were added and incubated for 30 minutes at 56°C. The DTT was removed and 300µl of Acetonitril 100% was added and left under the hood until the bands became white and shrank. After removing the Acetonitril, 30µl of IAA (55mM) was added and incubated for 20 minutes in the dark. When the IAA was removed, 300µl of ACN (100%) was added until the bands became white then the ACN was removed.

2. Digestion

The last step was the digestion with trypsine , 20µl of trypsine (6ng/µl) was added for each tube and left in ice for 20 minutes. After that, the bands were covered with AMBic (50mM) and incubated at 37°C overnight.

3. HPLC/MS/MS analysis

Extracted peptides were analyzed by nano-flow reversed-phase 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.

10 µl sample was loaded, purified and concentrated on a reversed phase monolithic pre-column, 200 µm ID x 5 mm length (LCPackings, Sunnyvale, CA USA) at 25 µl/min flow rate. Peptides were separated at a flow rate of 300 nl/min on a PepSwift Monolithic column, 100 µm ID x 5 cm length (LCPackings, Sunnyvale, CA USA) 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 sec. 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 nanoflow electrospray ion source. Data acquisition, analysis and the extraction of the peak lists were performed by Analyst QS software.

6. Data analysis

6.1 Protein identification

Proteins were identified using the Byonic version 3.4.0 software against the UniProt database with *Medicago* taxonomy (93591 entries). Because there were only 97 reviewed protein sequences of *M sativa* in the UniProt database it was necessary to use a bigger database containing close relatives. The search criteria were set 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 glutamin deamidation, ammonia loss, acetylation on protein as variable modifications. Protein identifications were validated by the Percolator algorithm false discovery rate was <1%. Only proteins with log prob>3 values were included into the bioinformatics analysis.

6.2. Bioinformatics

Blast2GO software package version 5.2.5 were used for *in silico* analysis. Protein sequences were blasted using QBlast service against two databases:

1. The UniProt public database, using taxonomy filter green plants (*Viridiplantae*, 33090 entries)
2. The COMPARE allergen database, (2081 entries), (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 Bioiformatics 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 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/Plant-mSubP/).(Sahu, Loaiza and Kaundal, 2019)

I. Results and discussion

III.1. Results

1 Protein concentration

Table 8 presents the protein's concentrations of the leaf extract during all the steps of the extraction. The protein estimation was made in every step. It permitted to get the elution and recycle the flow through. After putting all the elutions together from the three chromatographies and the ammonium sulfate precipitate, the final extract has a concentration of 1mg/ml (6mg in total).

Table 8. Protein concentration measured during the extraction of total protein extract of *M.sativa.*

Solution	Concentration (mg/ml)	Protein (mg)
Ammonium sulfate precipitate	0.16	0.8
DEAE elution	0.45	4.74
CM elution	0.028	0.34
Sepharose elution	0.012	0.12
Final extract	1.0	6

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

Appendix 8 represents the SDS-PAGE image of the protein profile of the leaf protein extract of *M. sativa.* The extract showed bands with a low molecular weight starting from less than 14KDa extending until a high molecular weight around 70KDa and some fading bands with a molecular weight superior than 70KDa. In order to not miss any band from the extract, the totality of the lane was used for the in-gel digestion and LC-MS/MS-based proteomics. Twenty bands were cut and analyzed by in-gel digestion proteomics.

3. Identification of proteins in the leaves of *M. sativa*

From the 129 proteins identified in the sample (appendix 9), table 9 shows the 20 top-ranking proteins in the extract. Both "Biological process" and "Molecular functions" were determined by the InterPro functional analysis tool in Blast2GO. The identified proteins represent the most abundant alfalfa proteins in the sample enabled for the proteomics characterization.

The main proteins identified were: three Glucan endo-1,3-beta-glucosidase, three chitinases and one endochitinase, Alpha-L-arabinofuranosidase/beta-D-xylosidas, LRRNT_2 domain-containing protein, FAD-binding berberine family protein and Plant basic secretory protein (BSP) family protein. It is interesting to point out that most of the identified proteins were enzymes with hydrolyze activity and not structural proteins.

Table 9. List of top 20 ranking proteins identified in the leaf protein extract of *M. sativa*.

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

G7K4B9	MEDTR	Glycerophosphoryl diester phosphodiesterase family protein	lipid metabolic process tubulin complex assembly post-chaperonin tubulin folding pathway	glycerophosphodiester phosphodiesterase activity beta-tubulin binding
I3STX0	MEDTR	Glycoside hydrolase family 18 protein	carbohydrate metabolic process chitin catabolic process	chitinase activity chitin binding
A0A072TUC3	MEDTR	Pathogenesis-related thaumatin family protein	/	/
A0A396K0C6	MEDTR	Putative nepenthesin	Proteolysis	aspartic-type endopeptidase activity
Q5RLX9	MEDSA	Chitinase	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	carbohydrate metabolic process	glucan exo-1,3-beta-glucosidase activity glucan endo-1,3-beta-D-glucosidase activity
A0A072VNF7	MEDTR	Putative tripeptidyl-peptidase II	Proteolysis	serine-type endopeptidase activity
A0A396GUY5	MEDTR	Putative thaumatin	defense response	
A0A072VD04	MEDTR	Beta-galactosidase	carbohydrate metabolic process	beta-galactosidase activity carbohydrate binding
A0A396JDT6	MEDTR	Putative nepenthesin	proteolysis; protein catabolic process	aspartic-type endopeptidase activity

4. Structural and functional annotation of identified proteins

Functional annotation performed gave some insights into the cellular and biological roles as well as cellular localization of the identified proteins. The Gene Ontology (GO) terms represented by: cellular component (CC), molecular function (MF) and biological processes (BP) were given for most of the identified proteins.

Figure 17 gathers the cellular components related to proteins of the extract. The major three functional categories in CC were extracellular region (36%), integral component of membrane (26%) and cell wall (18%).

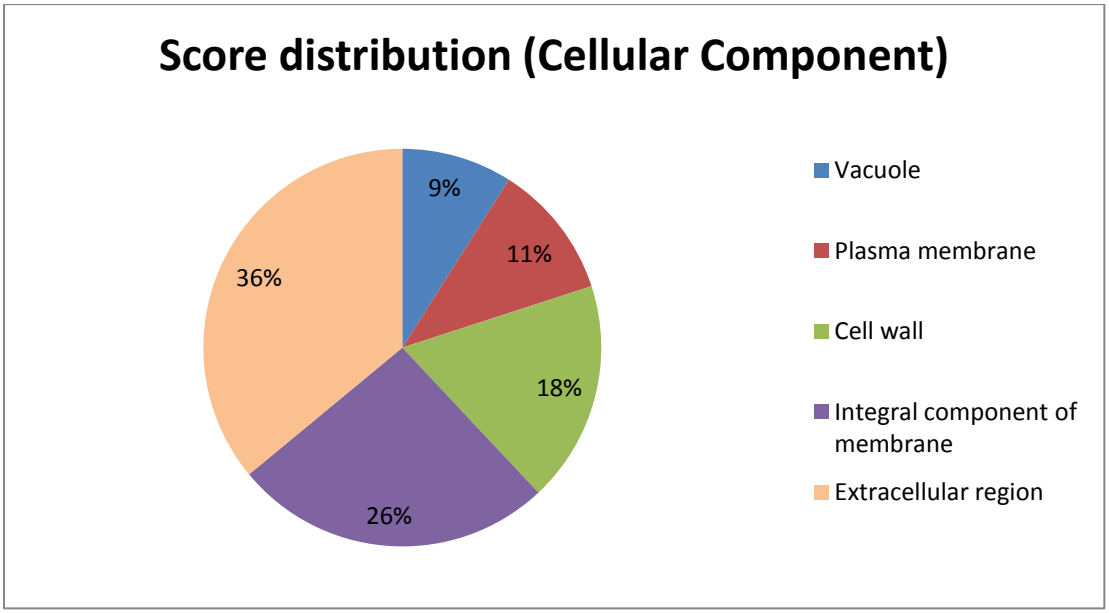


Figure 17. Cellular components given by the functional annotation of the protein extract by Blast2GO.

Figure 18 shows the score distribution of the molecular function given for the identified proteins in the extract. For the MF, the most abundant were: Oxidoreductas activity (24), hydrolase activity: acting on ester bonds (20%), metal ion binding (19%), enzyme inhibitor activity (17%) and chitinase activity (11%).

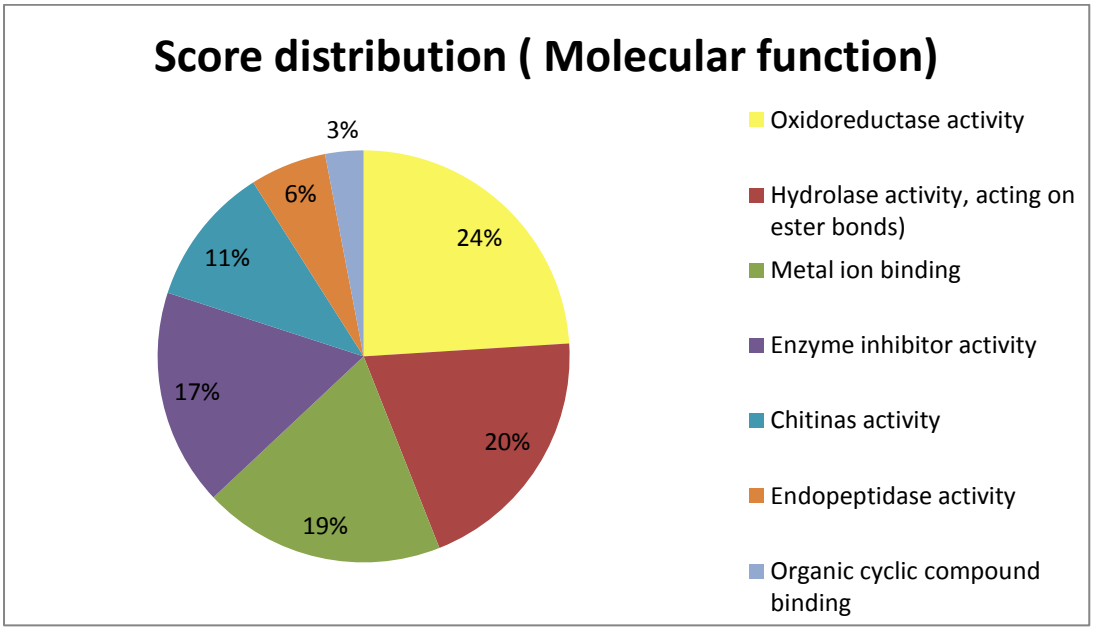


Figure 18. Molecular Function given by the functional annotation of the protein extract by Blast2GO.

Finally, figure 19 bellow represents the third GO term. The main Biological Processes were metabolic process (47%), cellular process (28%) and biological regulation (10%).

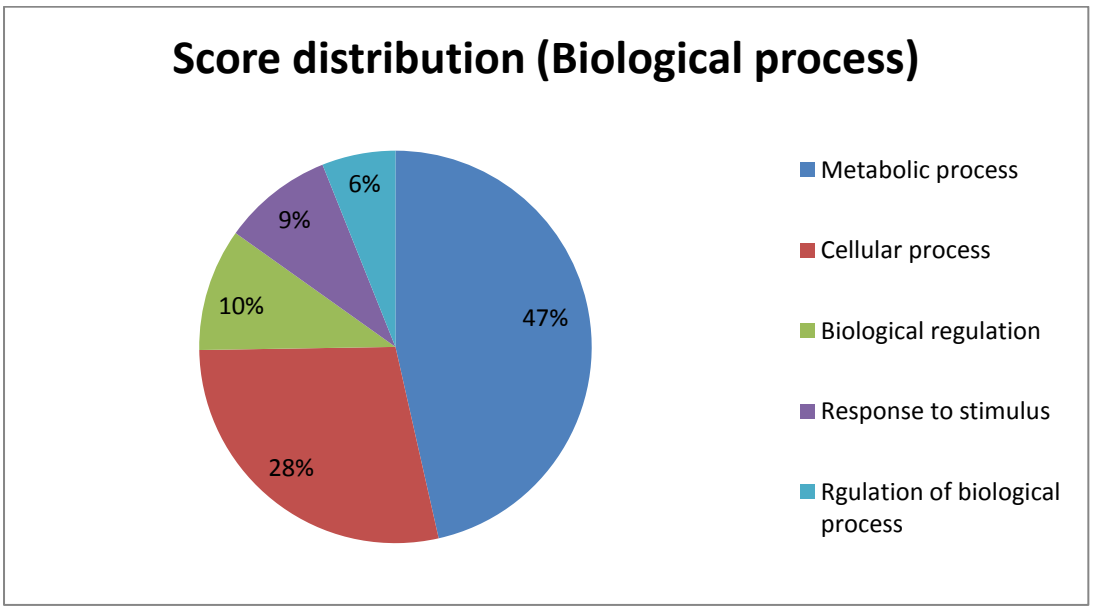


Figure 19. Biological process given by the functional annotation of the protein extract by Blast2GO.

5. *In silico* prediction of allergens in the leaf protein extract of *M sativa*

Table 10 lists proteins that showed high similarity means (≥70%) to allergens or putative allergens from COMPARE database. The first four high ranked proteins from the extract exhibited resemblance ranging from 86% to 84% to Putative Lactoylglutathione lyase from Asian rice. In addition, two different proteins from *M. sativa* revealed sameness to Putative thaumatin-like protein precursor from apple, six proteins were similar to three different allergens from *Hevea brasiliensis* and three proteins expressed ressemblence to non-specific lipid transfer protein from Lens culinaris. Also, proteins from the extract showed similarity to allergens from chestnuts, peanuts and corn.

Table 10. Potential allergens in *M.sativa* based on blast against COMPARE database.

Accession Number.	Name of the identified protein	Similarity %	Name of the allergens in COMPARE	Taxonomy of the allergen	Gene identifier
I3S2W4	Lactoylglutathione lyase	86.1	Putative Lactoylglutathione lyase	*Oryza sativa*	*Q948T6.2*
A0A396IHC0	Putative lactoylglutathione lyase	84.6			
G7L865	Lactoylglutathione lyase	84.5			
I3S4Q3	lactoylglutathione lyase GLX1-like	84.3			
I3T337	thaumatin-like protein 1b	78.3	Putative thaumatin-like protein precursor	*Malus domestica*	*Q9FSG7.1*
I3SBN3	Pro-hevein	77.2	Putative Chitin-binding allergen Bra r 2	*Hevea brasiliensis*	*ABW34946.1*
Q45NK7	Non-specific lipid-transfer protein (Fragment)	76.9	Allergen lipid transfer protein precursor	*Lens culinaris*	*AAX35807.1*
A0A1L2BU75	Beta-1,3-glucanase 3 (Fragment)	76.8	Allergen beta-1,3-glucanase	*Hevea brasiliensis*	*AEV41413.1* *CAA05978.1* *AEV41413.1*
A0A072VN05	Chitinase / Hevein / PR-4 / Wheatwin2	76.5	prohevein		
A0A072V710	Glucan endo-1,3-beta-glucosidase	75.5	Allergenbeta-1,3-glucanase		
A0A396I3	Non-specific	75.5	Allergen lipid	*Lens*	*AAX35807*

M7	lipid-transfer protein		transfer protein precursor	*culinaris*	*.1*
A0A396H709	Putative thaumatin	75.4	Putative thaumatin-like protein precursor	*Malus xdomestica*	*Q9FSG7.1*
A0A072VNL8	Chitinase / Hevein / PR-4 / Wheatwin2	75.0	prohevein	*Hevea brasiliensis*	*CAA05978.1*
I3SZI9	Barwin domain-containing protein	73.9	prohevein		
A0A396GNA8	Endochitinase	72.1	Allergen class I chitinase isoform 2	*Castanea sativa*	*ADN39439.1*
D2Y175	Harvest-induced protein	71.4	AllergenFag s 1 pollen allergen	*Arachis hypogaea*	*AAQ91847.1*
G7JJJ6	Non-specific lipid-transfer protein	71.3	Allergen non-specific lipid transfer protein	*Lens culinaris*	*AAX35807.1*
Q45NL7	Thioredoxin	70.6	Putative thioredoxin h1 protein	*Zea mays*	*CAI64400.1*

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

The results were classified into four categories: plant food extracts, plant allergic proteins, pollen and latex extracts and animal food extracts.

6.1. Plant food extracts

Figure 20 shows the IgE binding inhibition of the extract with food extracts (fruits and vegetables) Except for apricot (Pru ar) and peach pulp (Pru p‡), alfalfa extract inhibited the IgE binding to the tested fruits and vegetables. Extracts from; Cucumber (Cuc s), Eggplant (Sola m), grape (Vit v), common beet (Beta v), spinach (Spi o) and celery (Api g) had IgE inhibition equal or higher than 90%.

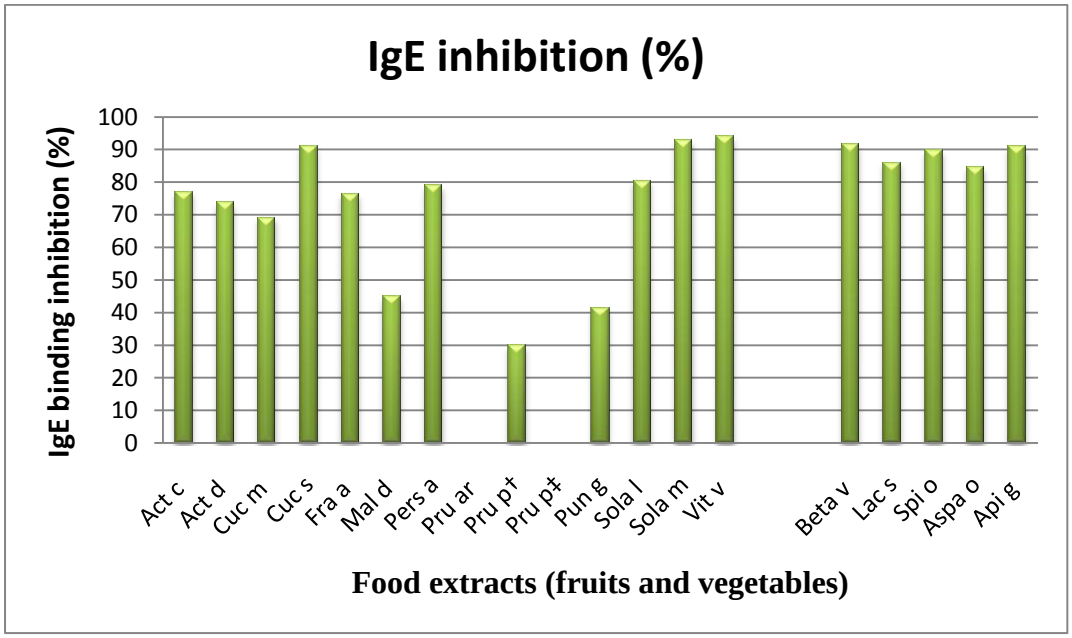


Figure 20. IgE inhibition results obtained on extracts from fruits and vegetables. (Golden kiwi fruit (Act c), Green kiwifruit (Act d), Melon (Cuc m), cucumber (Cuc s), Strawberry (Fra a), Apple (Mal d), Avocado (Pers a), Apricot (Pru ar), Peach peel (Pru p†), Peach pulp (Pru p‡), Pomegranate (Pun g), Tomato (Sola l), Eggplant (Sola m), Grape (Vit v) Common beet (Beta v), Lettuce (Lac s), Spinach (Spi o), Asparagus (Aspa o), Celery (Api g)).

On the other hand, all the bulbs and tubers used in the extract were inhibited onion (All c), Leek (All p), Garlic (All s), Fennel (Foe v), Potato (Sola t), Carrot (Dau c) with a percentage higher than 70% (Figure 21).

In contrast, out of 21 tested seed extracts, only four (carob (Cer si), rice (Ory s), pine nut (Pin p) and maize (Zea m)) were partially inhibited. No inhibition was observed on other seed extracts,

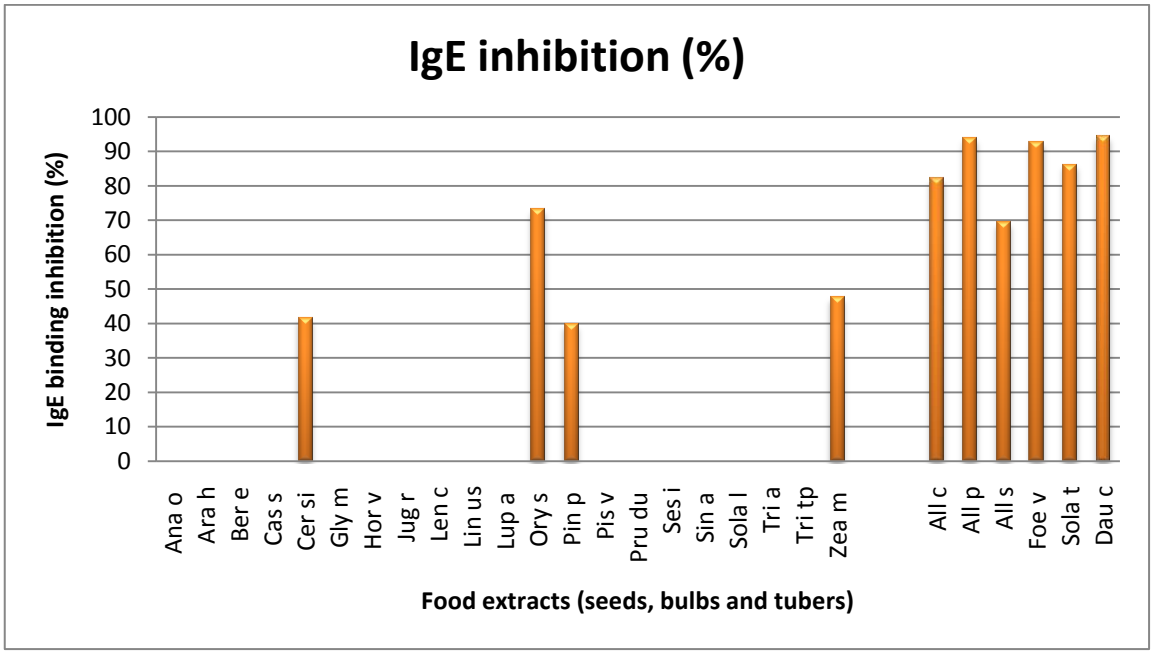


Figure 21. IgE inhibition results obtained on extracts from seeds, bulbs and tubers. Cashew (Ana o), Peanut (Ara h), Brazil nut (Ber e), Chestnut (Cas s), Carob (Cer si), Soy (Gly m), Barley (Hor v), Walnut (Jug r), Lentil (Len c), Flux (Lin us), White lupine (Lup a), Rice (Ory s), Pine nut (Pin p), Pistachio (Pis v), Almond (Pru du), Sesame (Ses i), White mustard (Sin a), Tomato (Sola l), Wheat (Tri a), Kamut (Tri tp), Maize (Zea m), carob (Cer si), rice (Ory s), pine nut (Pin p) and maize (Zea m).

6.2.Plant allergenic proteins

Figure 22 shows that food allergenic proteins belonging to several protein families including: 9k-LTP are dark green, 7k-LTP is light green, cysteine proteases are yellow, thaumatin is pink, chitinase is grey, Bet v 1-like are in orange, GRP is red, trypsin inhibitor is blue, hydrophobic peptide is black, and patatin is brown

LTPs, thaumatin-like protein, cysteine proteases, Bet v 1-like protein, GRP and chitinase 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. It is observed that components of the same protein family are inhibited at a different extent, as observed for allergens belonging to the LTPs family. Results show that only one 9k-LTP (Ara h 9) and 7k-LTP (Sola l 6) are completely inhibited by the alfalfa extract. Slightly lower inhibition was observed on the 9k-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.

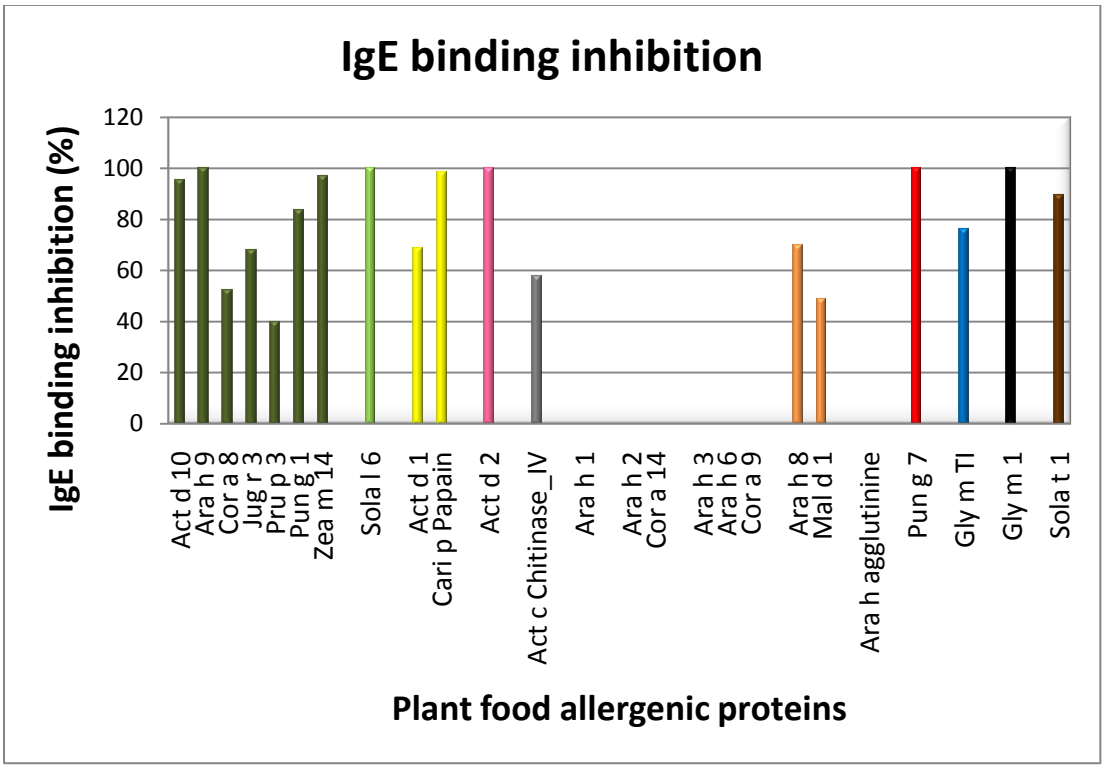


Figure 22. IgE inhibition results obtained on allergenic molecules from plant foods allergens belonging to the same protein family are grouped.

6.3.Latex and pollen extracts

Figure 23 shows that IgE inhibitions were also observed on some of the extracts of latex and pollens spotted on the FABER biochip (Amb a, Bet v, Cor a, Lol p, Ole e, Par j, Phl p, Que a and Hev b). The inhibition in all extracts was higher than 50% except for Phl p (47%). The higher values of IgE inhibition belong to Par j (95.5%) followed by Que a (94%), Bet v, Hev b, Amb a, Ole e, Cor a and Lol p.

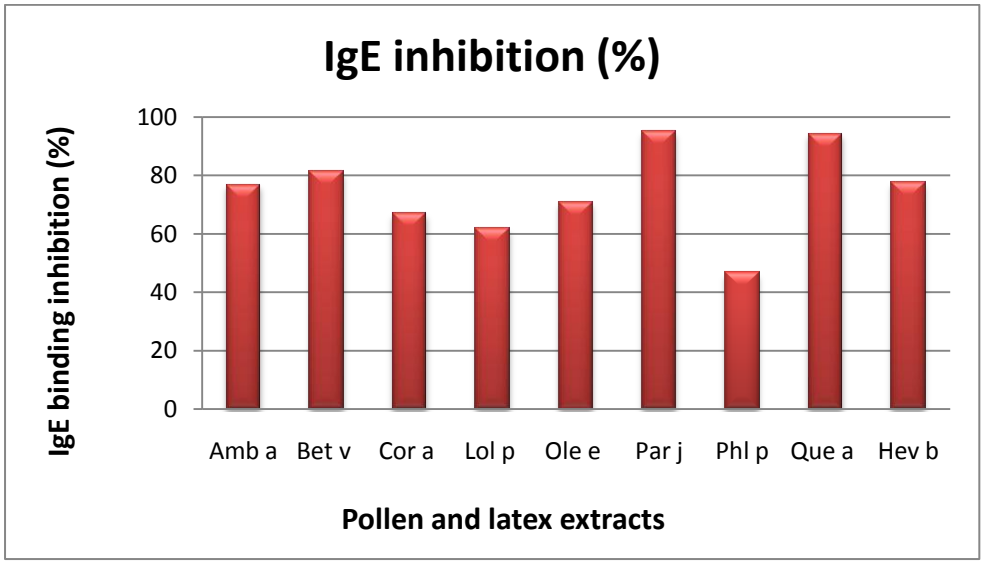


Figure 23. IgE binding inhibition on pollen and latex extracts given in the Faber test. Mugwort (Amb a), Birch (Bet v), Hazelnut (Cor a), Rye grasse (Lol p), Olive tree (Ole e), Parietaria (Par j), Timothy grass (Phl p), white oak (Que a), Rubber tree (Hev b).

6.4.Animal food extracts

There were no IgE inhibitions with any of the animal food extracts used in Faber test. Table 11 shows the extracts used in the test. Different sources and tissues were used such as: cow's milk, cow's meat, chicken's egg white and egg yolk, both common sol and sardines' meat.

Table 11. Animal food extracts and allergens used in Faber biochip.

	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	Bos d 4	Cow	Milk
	Bos d 5	Cow	Milk
	Bos d 6	Cow	milk, muscle, serum
	Bos d 8	Cow	Milk
	Bos d gelatin	Cow	Meat, bone

	Gald 1	Chicken	Egg
	Gald 2	Chicken	Egg

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

Figure 24 shows the alignment of alfalfa LTP with the analyzed homologous proteins: Zea m 14 from corn, Ara h 9 from peanut, Pun j 1 from pomegranate, Jug r 3 from walnut, Cor a 8 from chestnut, Pru p 3 from peach and Act d 10 from green kiwifruit.

The isoform of *M. truncatula* (MEDTR) 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.

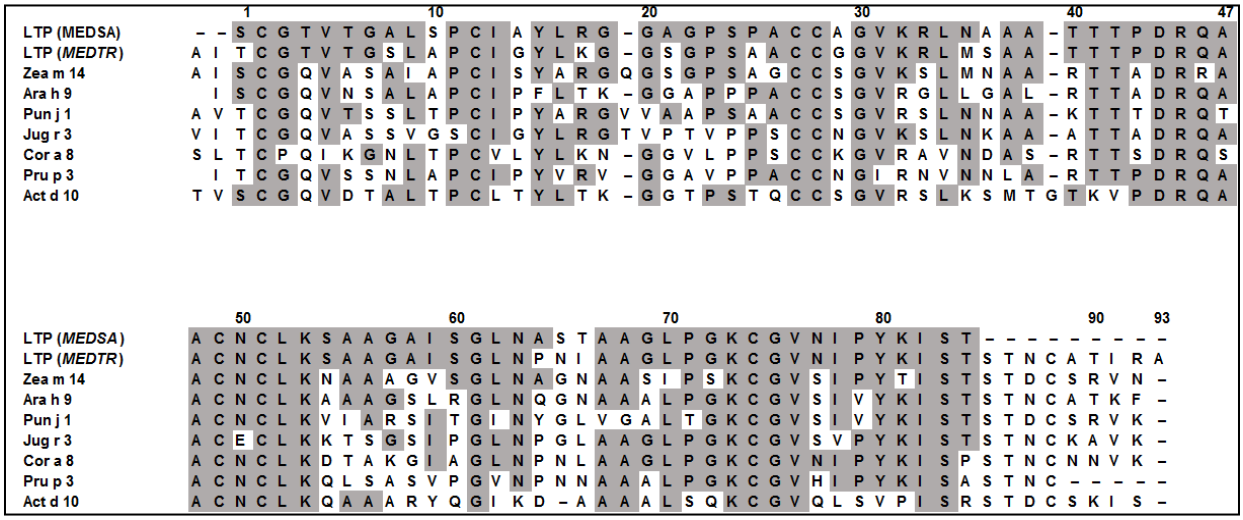


Figure 24. Multiple sequence alignement of LTP sequences from allergens that gave cross reactivity with alfalfa leaf extract compared with LTP sequence from *Medicago sativa* (MEDSA).

On the basis of this sequence alignment, the identity between alfalfa LTP and the homologous proteins was calculated (Table 12). 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. 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.

Table 12. Percentage of identity between LTP from *M. sativa*.

LTP	M. sativa	*M. Truncatula*	Zea m 14	Ara h 9	Pun g 1	Jug r 3	Cor a 8	Pru p 3	Act d 10
M. sativa	100	85	67	65	63	62	58	57	51
M. truncatula		100	63	64	61	62	60	61	49
Zea m 14			100	61	63	58	49	60	49
Ara h 9				100	60	60	55	67	52
Pun g 1					100	59	53	62	50
Jug r 3						100	58	62	39
Cor a 8							100	61	41
Pru p 3								100	48
Act d 10									100

8. Comparison between *in vivo* and *in silico* results

Table 13 represents the results found by the *in silico* study and confirmed by *in vivo* study. The first four proteins have high similarities to Lactoylglutathione lyase, also called glyoxalase I. In addition, two proteins show high similarities to the apple thaumatin-like protein, Mal d 2. *In silico* analysis reveals four *Medicago* putative proteins belonging to the barwin family also known as family 4 of pathogenesis-related proteins (PR-4), showing high similarities to pro-hevein, a major latex allergen. Other three proteins show similarities to plant LTPs and the highest identity with a lentil homolog (Len c 3).

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.

Table 13. Potential allergen families identified in the protein extract of leaf of *M. sativa* based on Proteomics/BLAST search combined with IgE-binding inhibition experiments

	Proteomics/Bioinformatics			**IgE-binding inhibition**	
Allergen family	**UniProt Accession number[a]**	**Similar allergens in COMPARE**	**Source**	**Homologous allergens on FABER**	**Source**
Thaumatin–like protein (PR-5)	*I3T337*, *A0A396H709*	Mal d 2	Apple	Act d 2	kiwifruit
Lipid transfer proteins (PR-14)	Q45NK7, *A0A396I3M7*, *G7JJJ6*	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
Pro-hevein (PR-4)	*I3SBN3*, *A0A072VN05*, *A0A072VNL8*, I3SZI9	Hev b 6	Latex	NI (Hev b 6)	latex
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

III.3.2. Discussion

This approach allowed us to identify 129 proteins expressed in *M. sativa* leaves. Most of the proteins were putatively identified by homology to known proteins of *M. truncatula* (144/129) known to be a close relative to alfalfa and the most studied species in *Medicago* genus (Confalonieri and Sparvoli, 2019).

In the literature, the genomic and proteomic characterization of this plant still needs investigation. There are only 92 reviewed proteins of *M. sativa L* listed in the UniProtKB database (Hawkins and Yu, 2018). A previous study has identified cell wall proteins of *M. sativa* providing proteome reference map for this important legume (Watson *et al.*, 2004). In recent works, 2D-DIGE MALDI-TOF (Gutsch *et al.*, 2018) and iTRAQ-based quantitative proteomics strategies (Chen *et al.*, 2016) have been applied to identify proteins and to detect changes after developmental and environmental stimuli, like osmotic stress (Zhang and Shi, 2018), salt stress(Long *et al.*, 2018; Gao *et al.*, 2019) and atrazine stress(Sui *et al.*, 2018). Those results are very important to understand the molecular pathways and physiological mechanisms as well protein expression of alfalfa in physiological and stress conditions in addition those results may serve for the enrichment of the database of the identified proteins in *M. sativa*.

The main proteins identified are known for their essential function in plant defense mechanism against biotic and abiotic stresses such as: chitinases, pathogenesis-related thaumatin family protein (Boccardo *et al.*, 2019), Glucan endo-1,3-beta-glucosidase, Beta-1,3-glucanase 3 and Beta-galactosidase (Faghani *et al.*, 2015). Some of them are also involved in cell division and expansion, cell wall formation and organization for instance: Alpha-L-arabinofuranosidase (Chávez Montes *et al.*, 2008) and Glycerophosphoryl diester phosphodiesterase family protein (Hayashi *et al.*, 2008).

According to the results of the IgE inhibition test, alfalfa proteins did not show any cross-reactions with allergens from animal foods, such as milk, egg, meat and fish, used as control markers. This cross reactivity between a plant and animal food was checked in the light of the presence of some repots in the literature suggesting a possible cross reaction between unrelated allergens like cross reaction between mammalian proteins and seed storage proteins like: soy and milk (Bublin and Breiteneder, 2019) .Our results suggest that alfalfa could be a valuable protein source for patients allergic to animal food proteins.

Among the identified proteins, several ones were recognized to be homologous to known allergens according to the sequence similarities retrieved from the *in silico* analysis. Some of those proteins (glucanases, class I chitinases and proheveins) belongs to the barwin family, that includes barwin, prohevein, glucanases, class I and class II chitinases. These proteins have in common a protein module, barwin that is involved in plant defense mechanisms. In fact, they are known as family 4 of pathogenesis-related proteins (PR-4). Furthermore, prohevein is made of an N-teminal hevein-like domain and a C-terminal barwin domain. In latex, the protein is cleaved into its constituent domains during wound-induced coagulation (Wahler *et al.*, 2009). 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 (Slavokhotova *et al.*, 2017). This protein module can be found at the N-terminus of proteins such as class I and IV chitinases, whereas class V chitinases holds two of them. Therefore, barwin-like and hevein-like modules are ubiquitous and evolutionary conserved protein structures. Several proteins share these 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, Hev b 10, Hev b 5 and Hev b 8, suggest a low level of cross-reactions between alfalfa and latex.

Medicago proteins showed also competition for IgE binding with the tested cysteine proteases, namely the kiwifruit Act d 1 and the papaya Cari p papain.

In the same context, similar results were demonstrated in a clinical report suggesting the existence of a "latex-fruit syndrome (Blanco *et al.*, 1994). 52% of patients with latex allergy cross reacted to some fruits like: avocado, chestnut, banana, kiwi and papaya. According to their findings and because of the results with the alfalfa leaf extract, it could also explain some of the cross reactivity of the leaf extract with: avocado (Pers a), kiwi (Act d 2, Act d 10, Act d [Fruit], Act d 1) and papaya (Cari p Papain).

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 this possibility.

Cross-reactivity between alfalfa leaf extract and extracts from rice, lentils, peanuts and corn was in accordance with some results in the literature. In a work using Radioallergosorbant test (RAST), cross-reactivity was shown between corn and soybean, corn and rice, rice and soybean, peanut and soybean. However no significant correlation was found between corn and peanut or peanut and rice (Lehrer *et al.*, 1999). Many works were showing results of cross reactivity between legumes (Nicolaou and Custovic, 2011). A study combining *in vivo* and *in vitro* approaches investigating the cross-reactivity in 12 peanut-allergic children from Italy with other legumes (lupin, lentils, kidney beens and soybeans) unveiled a large IgE-binding response within and between patients (Ballabio *et al.*, 2010). In another study, 39 peanut-sensitized patients were evaluated to lupine, pea, and soy by skin prick tests (SPT) and ImmunoCAP by skin prick tests and ImmunoCAP. It found out that 82% of patients were sensitized to lupine, 55% sensitized to pea and 87% sensitized to soy (Peeters *et al.*, 2009).

LTP is a relevant plant food allergen, classified as class I allergen because it is resistant to heat and gastrointestinal digestion and can cause severe and systemic allergic reactions, including the life-threatening anaphylaxis (Alessandri *et al.*, 2020). Due to its relevance, several LTP from different plant foods were included in the FABER biochip to allow a careful analysis of patients' sensitization to the sources of this allergy inducer. This panel of LTP resulted useful to investigate the cross-reaction between alfalfa LTP and the homologous molecules from several other foods. In addition, it 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 in secondary and tertiary structure suggests a possible high conservation of IgE binding epitopes responsible of IgE co-recognition and cross-reaction thus representing a risk for allergic people. It is usually supposed that cross reaction will occur when two allergens have more than 70% sequence identity and less than 50% sequence similarity is rarely causing cross reaction (Bublin and Breiteneder, 2019).

A simplified method useful to obtain a preliminary estimation of the level of risk of an allergic reaction being caused by a protein is based on the sequence identity 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 (Ladics, 2019).

Altough this method is not perfect and it does not take into consideration several contributing factors, such as the 3D-structure and the ligand binding that can affect the IgE epitope function, it can give a useful first indication.

The sequence identity between the LTP from alfalfa and the homologous molecules present on the FABER biochip varies from 51% to 67%. On that account, the values are giving a prediction of possible cross reactivity that could lead to an allergenic risk according to 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 LTP, thus indicating the sharing of antigenic epitopes with all these homologous proteins. In addition, with the exception of two LTP from kiwifruit and peanut, the results obtained show a good correlation between sequence identity and inhibition values. Supporting the interpretation of the data obtained and the help of the approach applied in this study. The two exceptions should be evaluated taking into consideration factors that can affect these results, such as the presence of undetected LTP isoforms in the alfalfa extract. Clearly, we can expect that a high cross-reaction between alfalfa LTP and homologous proteins having a higher structural similarity, compared to the tested eight ones, may occur. For example, the bioinformatics analysis revealed 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.

LTP molecules present on the FABER biochip are from kiwifruit, peanut, hazelnut, walnut, peach, pomegranate, maize and tomato. It has been noted the absence of correlation between IgE inhibition on the individual molecules and the extracts of the sources. For explanation, LTP purified from peanut, Ara h 9 is completely inhibited by the alfalfa extract, however no inhibition was shown with the extract of the peanut seed (Ara h). these results could be predicted and in accordance with the principle of the relative abundance of different

proteins in the natural source. Indeed, the purified proteins are fixed on the biochip at a constant amount. Likewise, the extracts contain protein components at different relative amounts like they are in the natural source. Therefore, the observed value of inhibition is essentially related to the IgE recognition towards the most abundant proteins of that natural source. Theoretically, the relative affinity towards IgE is another parameter that could interfere on the IgE binding signals; unfortunately this is a factor that we cannot evaluate.

In the case of peanut, there was no inhibition on the extract spotted on the FABER biochip; it is explained by the high amount of seed storage proteins (2S albumin, 7S vicilin and 11S globulin) in this source. These proteins are not inhibited by alfalfa extract. It is known that the amount of LTP Ara h 9 is very low in peanut and apparently its amount 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.

Thaumatin-like protein is a protein found in quite high amount in some plant foods, such as kiwifruit, tomato and apple. However, the clinical importance of this allergen is still debated. It belongs to the family 5 of pathogenesis-related proteins (PR-5). It is an additional protein contained in the *Medicago* proteome and found to have a significant similarity with homologous allergens listed in the COMPARE database, in particular with the apple Mal d 2. This apple allergen is not present on the FABER biochip, but the homologous kiwifruit protein is available and proved to be completely inhibited by the Medicago homologous protein, thus showing IgE cross-reaction.

Profilins and Bet v 1-like proteins are panallergens found in all plant organisms (Ruiz-Hornillos *et al.*, 2020; Morris *et al.*, 2021). They are classified as class 2 food allergens, responsible for OAS but not for systemic reactions, since they are readily denatured and

degraded in the gastrointestinal tract (Fernández-Rivas, 2015). 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 mass spectrometry experiments, nevertheless it proved its presence in the alfalfa extract because it gave IgE inhibtion on food and pollen homologous molecules available on the FABER biochip. In fact, it competed for IgE binding with the peanut seed Ara h 8, the apple fruit Mal d 1 and the birch pollen Bet v 1. Profilin was not detected by mass spectrometry and did not show any IgE inhibition on the pollen and latex profilins available in the test. Since the presence of profilin in alfalfa cannot be questioned, it is conceivable that it was extracted from alfalfa leaves in amounts not sufficient to give detectable inhibition signals. On the other hand, it is well known that profilin requires a specific protocol to be efficiently extracted (Cases, Pastor-Vargas and Perez-Gordo, 2017).

In addition, the immunological method detected some allergens that were not identified by mass spectrometry based proteomics in this work. Among these is a gibberellin regulated proteins (GRP) that is an allergen described in pollen and food and included in the class I because it is heat stable, protease resistant and can cause severe allergic reactions (Inomata, 2019). Inhibition was also observed on other allergens, such as the potato patatin and the soybean trypsin inhibitor and hydrophobic peptide, thus suggesting that homologous proteins are present in the alfalfa leaves although they were not detected by MS based proteomics in the present work.

Gathering all the results above, it is important to call attention to allergic patients to seed storage proteins who could be allowed to introduce alfalfa in their diet. In the other hand, patients sensitized to allergens belonging to LTPs, GRP, thaumatin-like protein, cysteine protease, barwin-like, hevein-like and Bet v 1-like, have to be vigilant and aware of the possible cross reactions that could occur. In fact, these patients could react with alfalfa proteins homologous to those allergens in case of presence of IgE recognizing shared antigenic determinants. This risk changes from a person to another depending on their genetic patrimony. In fact, as reported for LTPs 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 (Bernardi *et al.*, 2011).

III.4. Conclusion

Here we set-up of a novel workflow that combines different methodologies involving proteomics, *in silico* analysis and *in vitro* test based on nanotechnology multiplex immunological tests. This method takes advantages of the multiple query sequences in a single BLAST search of identified proteins against allergens present in curated databases and the SPHIAa assay performed on multiplex biochips which were *in vitro* probed with sera of allergic patients. A great advantage of the association of those methodologies rely on the exploitation of a large panel of validated allergens that were available on a biochip and probed with characterized IgE antibodies, thus allowing a comprehensive analysis of the immunological features of the sample under investigation.

The method followed in our study permitted the identification of 129 proteins of the alfalfa proteome from which some are still unknown components and provided a collection of data about the potential allergenicity of this plant. The annotation of the identified proteins revealed that most of them were enzymes. They are involved in multiples mechanisms like: plant defence and cell wall expension.

Remarkably, alfalfa proteins did not reveal cross-reactions with allergens from animal foods, such as milk, egg, meat and fish, used as control markers suggesting a save consumption of this plant for people allergic to animal food.

In contrast, most of the analyzed fruits and vegetables showed cross-reaction with alfalfa proteins at different extents like: Kiwifruit, strawberry, melon, tomato, grape, common beet, lettuce and many others. This result is explained by the fact that plants share some protein families.Therefore, patients allergic to these plant foods have to be careful when consuming alfalfa.

In addition, the IgE inhibition test revealed the presence of cross reactivity between alfalfa extract and some pollen and latex extracts like: Parietaria (Par j), Timothy grass (Phl p), white oak (Que a) and rubber tree (Bet v).

Gathering the results obtained from *in silico* and *in vitro* methods: four allergen families have been shown to give cross reaction with alfalfa extracts in both techniques: Thaumatin-like protein, Lipid transfer protein, Pro-hevein-like and Bet v 1-like protein. Consequently, people allergic to these protein families must cautious when consuming them. Further investigations should be done to make appropriate checks individually in order to check if consuming alfalfa for them or not.

Scientific paper 2

Research Article

Received: 23 April 2020 Revised: 23 July 2020 Accepted article published: 13 August 2020 Published online in Wiley Online Library:

(wileyonlinelibrary.com) DOI 10.1002/jsfa.10730

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.

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 staffs 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 proteases[7] 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 cultivars[11] 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 phytobiotic activity even in small quantities[16] 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 ofthe 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, bioinformatics[30] 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 foods[11] 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(vinylpolypyrrolidone), 2 mL of 0.125 mol L^{-1} ethylenediaminetetraacetic acid and 2.2 g of ascorbic acid dissolved

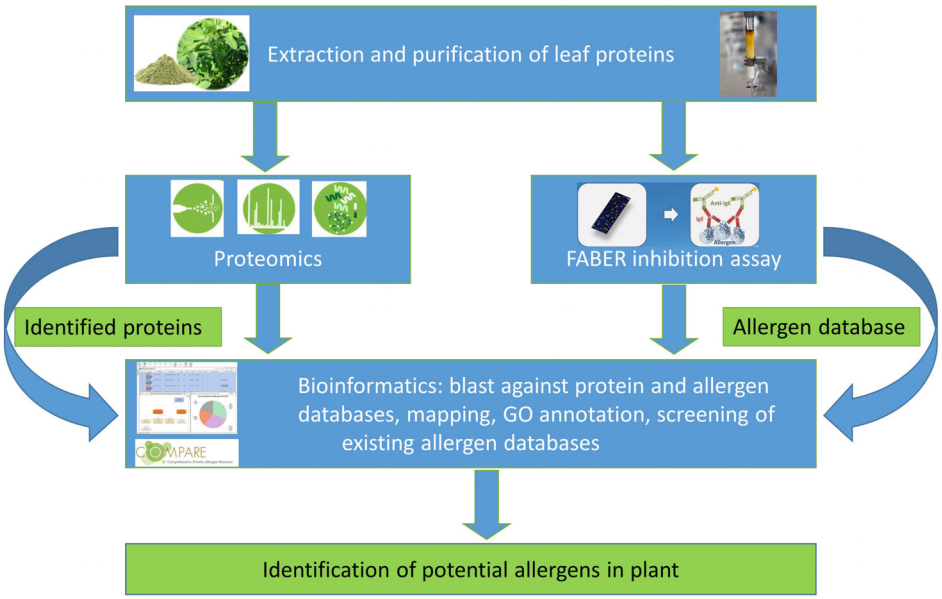


Figure 1. Scheme of the integrated multidisciplinary approach that combines proteomics, bioinformatics and SPHIAa assay for the identification of potential allergens in alfalfa protein extract.

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 × *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 ($(NH_4)_2SO_4$). 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 µm 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 particularly, 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.

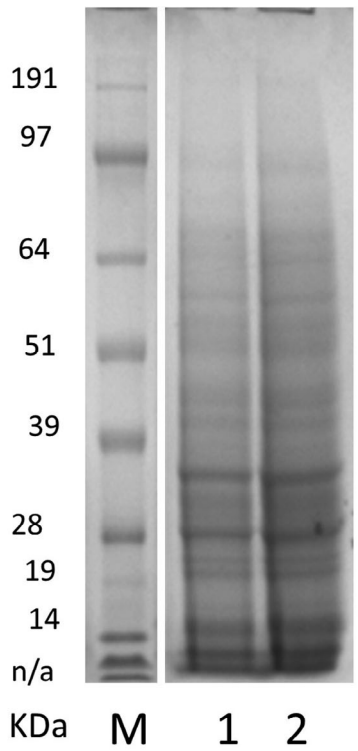
For the SPHIAa assay, a pool of four sera was prepared using 70 µL of each serum. Two sera were from patients sensitized to different plant foods. They contained IgE recognizing plant allergens, such as LTPs, profilin, Bet v 1-like, GRP, thaumatin-like protein and seed storage proteins. Two additional 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 trypsine (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 $\mu L\ min^{-1}$ 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.



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.

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 glutamin deamidation, ammonia loss, acetylation on protein as variable modifications. Protein identifications were validated by the Percolator algorithm[37] 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 Bioiformatics 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/PlantmSubP/).[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

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 proteina 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 family[42] 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 system[12] 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)

	Proteomics/bioinformatics			IgE-binding inhibition	
Allergen family	UniProt accession number[a]	Similar allergens in COMPARE	Source	Homologous allergens on FABER	Source
Thaumatin–like protein (PR-5)	*I3T337*, *A0A396H709*	Mal d 2	Apple	Act d 2	Kiwifruit
Lipid transfer proteins (PR-14)	Q45NK7, *A0A396I3M7*, *G7JJJ6*	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.

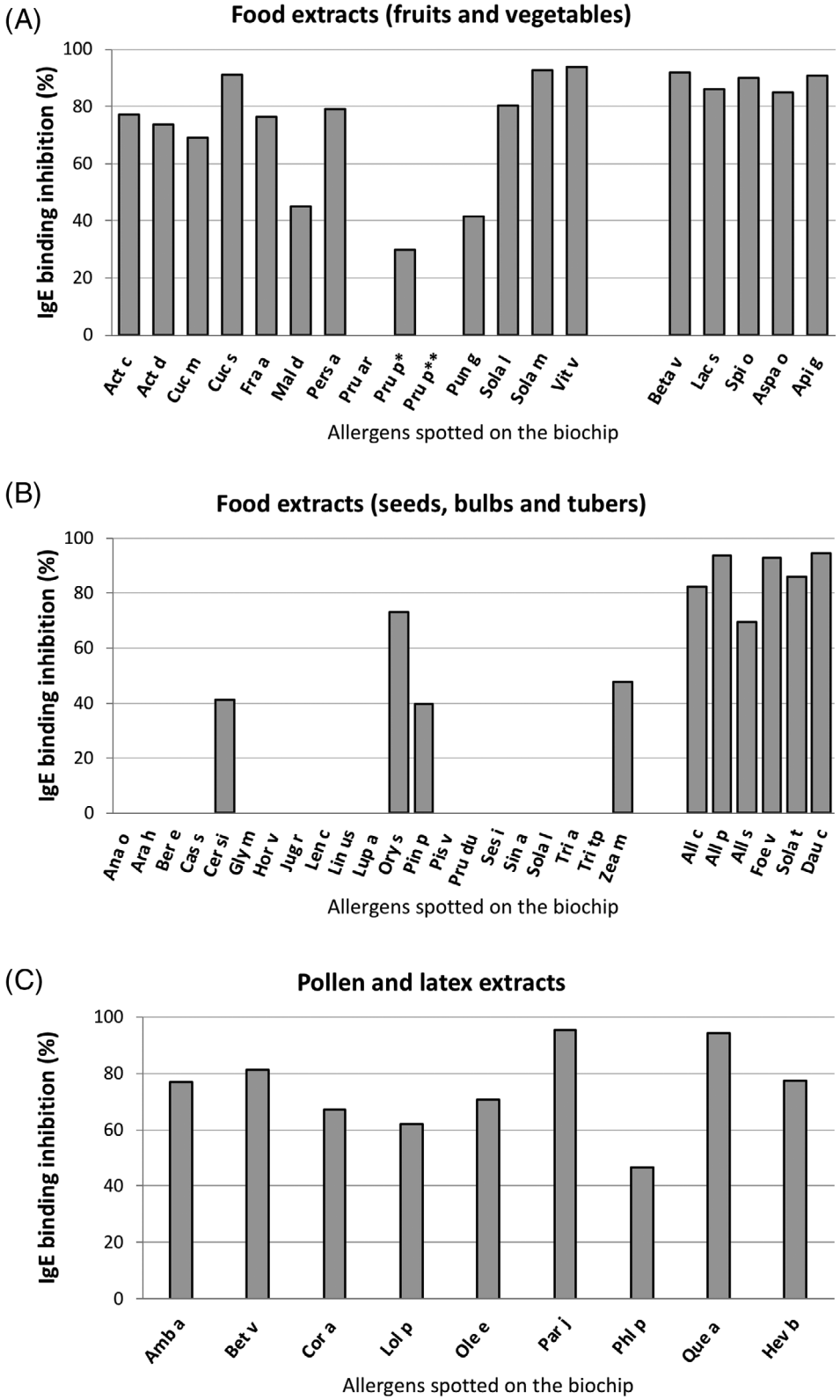


Figure 3. SPHIAa results obtained on extracts from plant food, pollen and latex. Allergenic sources of the same category are grouped. (A) Fruits and vegetables. (B) Seeds, bulbs and tubers; Pru p* and Pru p** represent the peach peel and pulp extracts, respectively. (C) Pollen and latex extracts.

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

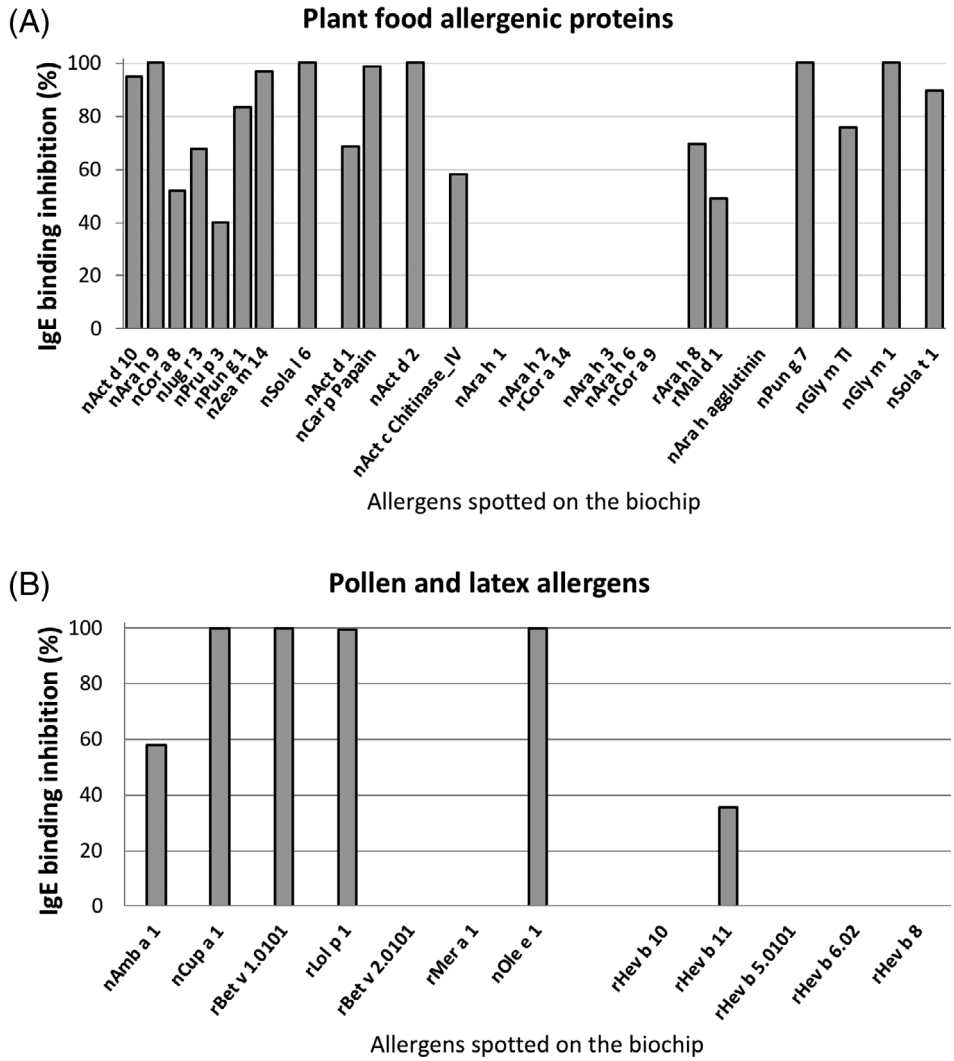


Figure 4. SPHIAa results obtained on allergenic molecules from plant foods (A) and on pollen and latex allergens (B). Allergens belonging to the same protein family are grouped. Allergen details are reported in Table S4.

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 UnipProt 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 challanging 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 allergy[45] 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-teminal 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

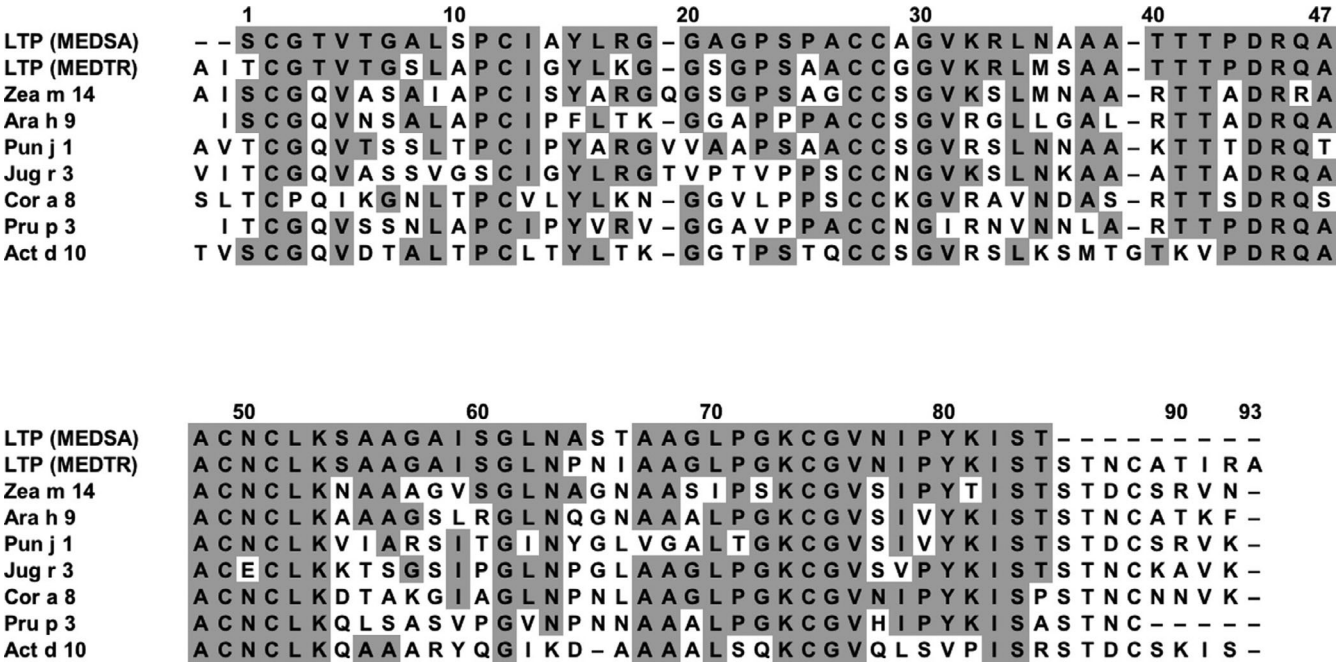


Figure 5. Multiple sequence alignment of *Medicago sativa* (MEDSA) LTP with a *Medicago truncatula* homolog (MEDTR) and with allergenic LTPs tested for IgE binding inhibition, namely the corn Zea m 14, peanut Ara h 9, pomegranate Pun g 1, walnut Jug r 3, hazelnut Cor a 8, peach Pru p 3 and green kiwifruit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

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 peanut[49] 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 inhibtion 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 pollen[12] 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

1 Chong SN and Chew FT, Epidemiology of allergic rhinitis and associated risk factors in Asia. *World Allergy Organ J* **11**:17 (2018).
2 Oriel RC and Wang J, Diagnosis and management of food allergy. *Pediatr Clin North Am* **66**:941–954 (2019).
3 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).
4 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.
5 Kapingidza AB, Pye SE, Hyduke 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).
6 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).
7 Breiteneder H and Ebner C, Molecular and biochemical classification of plant-derived food allergens. *J Allergy Clin Immunol* **106**:27–36 (2000).
8 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).
9 Sampson HA, Medical history. *J Allergy Clin Immunol* **103**:981–988 (1999).
10 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).
11 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).
12 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).
13 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).
14 Apostol L, 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).
15 Mielmann A, The utilisation of lucerne: a review. *Br Food J* **115**:590–600 (2013).
16 Gaweł 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).
17 Hawkins C and Yu LX, Recent progress in alfalfa (*Medicago sativa* L.) genomics and genomic selection. *Crop J* **6**:565–575 (2018).
18 Watson BS, Lei Z, Dixon RA and Sumner LW, Proteomics of *Medicago sativa* cell walls. *Phytochemistry* **65**:1709–1720 (2004).
19 Gutsch A, Zouaghi S, Renaut J, Cuypers 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).
20 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).
21 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).
22 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).
23 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).
24 Sui X, Wu Q, Chang W, Fan X and Song F, Proteomic analysis of the response of Funnelifor mismosseae/*Medicago sativa* to atrazine stress. *BMC Plant Biol* **18**:289 (2018). https://doi.org/10.1186/s12870-018-1492-1.
25 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).
26 Bener A, Safa W, Abdulhalik S and Lestringant GG, An analysis of skin prick test reactions in asthmatics in a hot climate and desert environment. *Allerg Immunol* **34**:281–286 (2002).
27 Jensen LB, Pedersen MH, Skov PS, Poulsen LK, Bindslev-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).
28 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).
29 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).
30 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.
31 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 subcellular 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 Bruijnzeel-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).

General conclusion

General conclusion and prospects

Food allergy is an important health issue affecting children and adults worldwide. This pathology appears to be increasing and has become an important health concern in developing and developed countries. Studies on the prevalence of food allergies became crucial. It allows the determination of the most frequent foods causing this reaction especially that food allergy depends on the most consumed foods in the region. On the other hand, it is important to study the total population and determine the main risk factors that interfere in the appearance of this pathology.

In Algeria, the studies about the prevalence of food allergy are rare. In our work, we are the first to determine time the prevalence of food allergy among schoolchildren of the city of Guelma during April 2018. The prevalence of reported food allergy was estimated at 8.5% (95% CI 6.1-10.8). The average age of appearance of the potential allergy was 4.41±2.51 years. Among allergic children, 26.7% were aged from 5-7 years old, 64.4% had 8-10 years old and 8.9% had 11-13 years old. According to our findings, the gender of children was not significantly influencing the appearance of this pathology. On the other hand, the main factor of risk was found to be the presence of family history where 66.6% had at least one of their parents atopic (p=0.0001). For the symptoms reported, 68.6% were Cutaneous, 19.6% were digestive and 11.8% were respiratory.

The main foods reported by the parents were: chocolate (1.7%), wheat (1.3%), milk (1.1%), eggs (1.1%), strawberry (1.1%) and fish (0.9%). Every prevalence was related to different suggestions and interfering factors to explain the given results. For instance chocolate appeared to be the most common food. This high prevalent result was explained by the fact that chocolate can contain different allergens at once such as: milk and different types of nuts (peanut, hazelnut, almond…….etc). In addition to that, chocolate is known to contain histamine releasers that could cause reactions that are similar to food allergy. For wheat, its high prevalence was related to the eating habits of the studied population where wheat is consumed at early age even for kids in Guelma or Algeria. It is conventional that the earlier the food is introduced the more it has a possibility to cause food allergy. Strawberry was also one of the main foods causing allergy. This fruit is also known to contain histamine releasers that induce similar symptoms to food allergy. Another suggestion is that strawberry is related to pollen-food syndrome. In this type of allergy, people get sensitized by pollen from different

types of trees. The antibodies generated against this pollen cross react with allergens from strawberry.

Those results opened new prospects to deepen our knowledge about our findings on children of the city of Guelma. We are interested in the high prevalent allergies like chocolate and strawberries. An investigation on those two foods is necessary by doing *in vivo* tests: skin prick test, food oral challenge and search of specific IgE.

In the second part, we did the extraction of different allergens from different sources: ovalbumin and ovomucoid from eggs, α lactalbumin and β Lactoglobulin from cow's milk, parvalbumin from sardines and gliadins from wheat. It has been shown that all the allergens had an acidic isoelectric point with no significant difference except for gliadins (pI=6.20). The acidic isoelectric point is usually a characteristic of allergens conferring them resistance in acidic environments. The purification and the identification of the allergens by HPLC/MS/MS and data analysis confirmed the efficiency of the protocols of the extraction.The identification of allergens based on the amino acid sequence was successful except for parvalbumin from sardine that was identified as parvalbumin from common carp.

This was explained by two possible explanations. Either the contamination of the sardine bought from the market with the other fish or the fact that the total sequence of parvalbumin from sardine is not available so the identification was made according to resemblance to the available sequences from other fish.

The multiple sequence alignment between all the allergens didn't show a significant similarity. However, the similarity was present in proteins from the same food, like ovalbumin and ovomucoid where both are from eggs. They revealed 7.44 % identity with 49 similar positions represented by 27 conservative replacements and 22 semi conservative mutations. For α-lactalbumin and β-lactoglobulin from cow's milk, the identity was 13.706% with 49 similar positions presented by 30 conservative replacements and 19 semi conservative mutations. Surprisingly, when comparing the allergens from animal source (ovalbumin, ovomucoid, α-lactalbumin, β-lactoglobulin and parvalbumin) the identity 0% with only 4 similar positions.

Till now, now relation between the similarity of the primary sequence alignment and the allergenicity of a protein. In some cases, two proteins have more than 70% similar sequence but one of them is an allergen but the other is not. On contrary, some proteins don't share similar sequences but are both allergens like the case in our study. This result is

explained by the presence of other factors like the 3D structure of the allergen that causes the allergic reaction. Also, the type of the epitope conformational epitopes or sequential epitopes could be another reason. In addition, the glycolisation of some proteins might also interfere in this result.

Our findings are interesting, from the biochemical and immunological point of vu. We would like to focus on the continuity of this work by:

- The determination of the amino acid sequence of parvalbumin from sardine.
- The study of the 3D structure of the allergens and their epitops.
- The study of possible chemical and biological treatments for allergens in order to reduce or eliminate their allergenicity.

In the last part of our work, we aimed to focus on the applied field of biochemistry. We studied the leaf protein extract from *Medicago sativa*. We investigated the potential allergenicity of this plant in order to see its safety to be used as a protein substitute for allergic patients.

Starting with *in silico* analysis, it revealed putative allergens in the leaf extract. Over than 10 proteins were similar (>70%) to known allergens like Putative Lactoylglutathione lyase, Putative thaumatin-like protein precursor from apple and Putative Chitin-binding allergen Bra r 2 from rubber tree.

In the other hand, *in vitro* analysis by IgE inhibition revealed the presence of cross reaction between known allergens in plant kingdom and the proteins of *Medicago sativa*. LTP was the main family giving cross reaction with our extract. Most of the analyzed fruits and vegetables showed cross-reaction with alfalfa proteins at different extent like: Kiwifruit, strawberry, melon, tomato, grape, common beet, lettuce and many others. Therefore, patients allergic to these plant foods have to be careful when consuming alfalfa.

However, no cross reaction was observed with allergens from animal foods such as milk, egg, meat and fish, used as control markers. This result permitted us to suggest the use of alfalfa proteins safely by allergic people to animal foods allergens like cow's milk, eggs, fish etc as a protein substitute.

According to our findings, four allergen families had high amino acid sequences similarities *in silico* analysis with our extract and supported by the *in vitro* test. Those four

families are: Thaumatin-like protein, Lipid transfer protein, Pro-hevein-like and Bet v 1-like protein. Consequently, people allergic to these protein families must be cautious when consuming this plant.

The comparison between *in vitro* and *in silico* results permitted to conclude that the presence of a high similarity in amino acid sequence doesn't lead necessary to a cross reaction and vice versa.

This study opens to new perspectives for the *in vivo* testing:

- It is important to study the reaction of ingestion of the plant or its proteins in animal modal with food allergy to different foods.
- Further investigations should be done to make appropriate checks individually in order to check the safety of the consumption of alfalfa.

References

References

Abeyrathne EDNS., Lee JS., Ham H and Ahn DU, 2013. Separation of Ovotransferrin from Chicken Egg White without Using Organic Solvents. *Poultry Science*, 92,1091–1097.

Abeyrathne EDNS., Lee JS and Ahn DU, 2014a. Separation of Ovotransferrin and Ovomucoid from Chicken Egg White. *Poultry Science*, 93,1010–1017.

Abeyrathne EDNS., Lee JS., Ham H and Ahn DU, 2014b. Sequential Separation of Lysozyme, Ovomucin, Ovotransferrin, and Ovalbumin from Egg White. *Poultry Science*, 93,1001–1009.

Abeyrathne EDNS., Lee CJo,. Nam KC and Ahn DU. 2014. Enzymatic Hydrolysis of Ovalbumin and the Functional Properties of the Hydrolysates SDS-PAGE. *Poultry Science*, 93,2678–2686.

Abeyrathne EDNS., Lee CJo., Suh JW and Ahn DU, 2013. Enzymatic Hydrolysis of Ovomucoid and the Functional Properties of Its Hydrolysates. *Poultry Science* 94, 2280–2287.

Abrams Elissa M., Chan Edmond S and Sicherer S, 2020. Peanut Allergy: New Advances and Ongoing Controversies. *Pediatrics*, 145, 2019-2022.

Al-Jassir MS., El-Bashir BM., Moizuddin S.K and Abu-Nayan AAR, 2006. *Infant Feeding in Saudi Arabia: Mothers' Attitudes and Practices*, 12,6-13.

Alessandri C., Ferrara R., Bernardi ML., Zennaro D., Tuppo L., Giangrieco I., Ricciardi T., Tamburrini M., Ciardiello MA and Mari A, 2020. Molecular Approach to a Patient's Tailored Diagnosis of the Oral Allergy Syndrome. *Clinical and Translational Allergy*, 10,22.

Ansari IT and Mu T, 2018. A Murine Model of Wheat versus Potato Allergy: Patatin and 53kDa Protein Are the Potential Allergen from Potato. *Molecular Immunology*, 101,284–293.

Asero R., Antonicelli A., Arena L., Bommarito B., Caruso M., Crivellaro M., De Carli E., Della Torre F., Della Torre E., Heffler F., Lodi Rizzini R., Longo G., Manzotti M., Marcotulli A., Melchiorre P., Minale P., Morandi B., Moreni A., Moschella F.

,Murzilli F., Nebiolo M., Poppa S., Randazzo G., Rossi and GE, 2009. Epidem AAITO: Features of Food Allergy in Italian Adults Attending Allergy Clinics: A Multi-Centre Study.*Clinical and Experimental Allergy*, 39(4),547–555.

Axelsson IGK., Ihre E and Zetterström O, 1994. Anaphylactic Reactions to Sunflower Seed. *Allergy*, 49(7):517–520.

Ballabio C., Magni C., Restani P., Mottini M., Fiocchi A., Tedeschi G and Duranti M, 2010. IgE-Mediated Cross-Reactivity among Leguminous Seed Proteins in Peanut Allergic Children. *Plant Foods for Human Nutrition*, 396–402.

Battais F., Pineau F., Popineau Y., Aparicio C., Kanny G and Guerin L, 2003. Food Allergy to Wheat : Identification of Immunogloglin E and Immunoglobulin G-Binding Proteins with Sequential Extracts and Purifed Proteins from Wheat Four. *Clinical and Experimental Allergy*, 33:962–970.

Bener A., Safa W., Abdulhalik S and Lestringant GG, 2002. An Analysis of Skin Prick Test Reactions in Asthmatics in a Hot Climate and Desert Environment. *Allergie et Immunologie*, 34(8),281–286.

Bernardi ML., Giangrieco I., Camardella L., Ferrara R., Palazzo P., Panico MR., Crescenzo R., Carratore V., Zennaro D., Liso M., Santoro M., Zuzzi S., Tamburrini M., Ciardiello MA and Mari A, 2011. Allergenic Lipid Transfer Proteins from Plant-Derived Foods Do Not Immunologically and Clinically Behave Homogeneously: The Kiwifruit LTP as a Model. *PLoS ONE*, 6(11),e27856.

Blanco C., Carrillo T., Castillo R., Quiralte J and Cuevas M, 1994. Latex Allergy: Clinical Features and Cross-Reactivity with Fruits. *Annals of Allergy*, 73(4),309–314.

Boccardo N A., Segretin ME., Hernandez I., Mirkin FG., Chacón O., Lopez Y., Borrás-Hidalgo O and Bravo-Almonacid FF? 2019. Expression of Pathogenesis-Related Proteins in Transplastomic Tobacco Plants Confers Resistance to Filamentous Pathogens under Field Trials. *Scientific Reports* 9(1),1-13.

Bøgh KL and Madsen CB, 2016. Food Allergens: Is There a Correlation between Stability to Digestion and Allergenicity? *Critical Reviews in Food Science and Nutrition*, 56(9),1545–1567.

Boughellout, H., Benatallah L., and Zidoune MN, 2015. Prévalence de l'allergie Aux Protéines Du Lait de Vache Chez Des Enfants Âgés de Moins de 3 Ans de La Ville de Constantine (Algérie). *Revue Francaise d'Allergologie,* 55(4), 288–292.

Bousfiha A and Aarab L, 2014. Modulation of IgE Immunoreactivity to Broad Bean Proteins after Food Processing in a Moroccan Population. *Allergologia et Immunopathologia,* 42(1),29–34.

Brunelle JL and Green R, 2014. One-Dimensional SDS-Polyacrylamide Gel Electrophoresis (1D SDS-PAGE). in *Methods in Enzymology*,541,151–59

Bublin M and Breiteneder H, 2019. Cross reactivities of Non homologous Allergens. *Allergy*, 75:1019–1022.

Cabrera-Chávez F., Rodríguez-Bellegarrigue CI., Figueroa-Salcido OG., Lopez-Gallardo JA., Arámburo-Gálvez JG., Vergara-Jiménez MJ., Castro-Acosta ML., Sotelo-Cruz N., Gracia-Valenzuela MH and Ontiveros N, 2018. Food Allergy Prevalence in Salvadoran Schoolchildren Estimated by Parent-Report. *International Journal of Environmental Research and Public Health*, 15(11).

Carlson G., Coop C, 2019. Pollen food allergy syndrome (PFAS): A review of current available literature. *Ann Allergy, Asthma Immunol*, 123, 359–365

Casan A and Zander U, 2013. Crystallization Communications Purification , Crystallization and Preliminary X-Ray Analysis of the Strawberry Allergens Fra a 1E and Fra a 3 in the Presence of Catechin. *Crystallization Communications*,69, 510–14.

Cases B., Pastor-Vargas C and Perez-Gordo M, 2017. Allergen Extraction and Purification from Natural Products: Main Chromatographic Techniques. *Methods in Molecular Biology*,1592, 13–22.

Čelakovská J., Bukač J., Vaňková R., Krejsek J and Andrýs C, 2021. Food Allergy to Apple, Peach and Celery in Atopic Dermatitis Patients, Analysis of Sensitisation to Molecular Components. *Food and Agricultural Immunology*,32, 184-203.

Cerecedo I., Zamora J., Fox M., Voordouw J., Plana N., Rokicka E., Fernandez-Rivas M., Vázquez Cortés S., Reche M., Fiandor A., Kowalski M., Antonides G., Mugford M., Frewer LJand Hoz B, 2014. The Impact of Double-Blind Placebo- Controlled Food

Challenge (DBPCFC) on the Socioeconomic Cost of Food Allergy in Europe.*Journal of Investigational Allergology and Clinical Immunology*, 24(6),418–424.

Charan J and Biswas T, 2013. How to Calculate Sample Size for Different Study Designs in Medical Research? *Indian Journal of Psychological Medicine*, 35(2),121–126.

Chávez M., Ricardo A., Ranocha P., Martinez Y., Minic Z., Jouanin L., Marquis M., Saulnier L., Fulton LM., Cobbett CS., Bitton F., Renou J P., Jauneau A and Goffner D, 2008. Cell Wall Modifications in Arabidopsis Plants with Altered α-l-Arabinofuranosidase Activity. *Plant Physiology*, 147(1),63–67.

Chen L., Chen Q., Zhu Y., Hou L and Mao P, 2016. Proteomic Identification of Differentially Expressed Proteins during Alfalfa (Medicago Sativa l.) Flower Development. *Frontiers in Plant Science*, 7,1502.

Chiriac AM., Bourrain JL., Lepicard E., Molinari N and Demoly P, 2017. Prevalence of Sensitization and Allergy to Potato in a Large Population. *Journal of Allergy and Clinical Immunology: In Practice*, 5(2),507–509.

Chong KW., Wright K., Goh A., Meyer R and Rao R, 2018. Growth of Children with Food Allergies in Singapore. *Asia Pacific Allergy*, 8(4),e34.

Chong SN and Chew FT, 2018. Epidemiology of Allergic Rhinitis and Associated Risk Factors in Asia. *World Allergy Organization Journal*,*11*.

Chu ML, 2014. PREVALENCE AND DETERMINANTS OF ATOPY AMONG SCHOOL-AGE CHILDREN IN RURAL SASKATCHEWAN, CANADA. *Applied Microbiology and Biotechnology*, 85(1):2071–2079.

Chudzik-Kozłowska J., Wasilewska E and Złotkowska D. 2020. Evaluation of Immunoreactivity of Pea (Pisum Sativum) Albumins in BALB/c and C57BL/6 Mice. *Journal of Agricultural and Food Chemistry*, 68(13),3891–3902.

Ciardiello MA, Tamburrini M., Liso M., Crescenzo R., Rafaiani C and Mari A. 2013. Food Allergen Profiling: A Big Challenge. *Food Research International*, 54(1),1033–1041.

Collins A., Derkenne B., Giebels K and Carvelli T, 2020. Allergie Immédiate à La Fraise et à La Framboise. *Revue Medicale de Liege*, 75,494–496.

Confalonieri M and Sparvoli F, 2019. Recent Advances in Medicago Spp. Genetic Engineering Strategies in *The Model Legume Medicago truncatula*, *14*, 1149–1161 .

Corica D, Aversa T., Caminiti L., Lombardo F., Wasniewska M and Pajno GB, 2020. Nutrition and Avoidance Diets in Children With Food Allergy. *Frontiers in Pediatrics*, 8,518.

Czaja-Bulsa G and Bulsa M, 2017. What Do We Know Now about IgE-Mediated Wheat Allergy in Children? *Nutrients*, 9(1).

Dölle S., Schwarz D., Lehmann K., Weckwerth W., George E., Worm M and Franken P, 2011. Tomato Allergy: Impact of Genotype and Environmental Factors on the Biological Response. *Journal of the Science of Food and Agriculture*, 91(12),2234–2240.

Dosanjh A, 2019. Raspberry Allergen and Rosaceae Family Member Allergens: Clinical and Cellular Responses. *Journal of Interferon and Cytokine Research*, 39(5),273.

Dziuba M., Nałęcz D., Szerszunowicz I and Waga J, 2014. *Proteomic Analysis of Wheat α/A-and β-Gliadins*, 32,437-442.

El-Hosni M., Montejo JM and Schuler CF, 2020. Food Anaphylaxis Due to Boiled Beetroot. *Journal of Allergy and Clinical Immunology: In Practice*, 8(9),3207–3208.

Esteve C., Montealegre C., Marina ML and García MC, 2012. Analysis of Olive Allergens. *Talanta*, 92,1–14.

Faghani E., Gharechahi J., Komatsu S., Mirzaei M., Khavarinejad RA., Najafi F., Farsad LK and Salekdeh GH, 2015. Comparative Physiology and Proteomic Analysis of Two Wheat Genotypes Contrasting in Drought Tolerance. *Journal of Proteomics*, 114,1–15.

Fernández-Rivas M, 2015. Fruit and Vegetable Allergy. *Chemical Immunology and Allergy*, 101,162–170.

Franz-oberdorf K., Eberlein B., Edelmann K., Hu S., Besbes F., Darsow U., Ring J and Schwab W, 2016. Fra a 1.02 Is the Most Potent Isoform of the Bet v 1-like Allergen in Strawberry Fruit. *J. Agric. Food Chem*, 64(18), 3688–3696.

Gandolfo-Cano M., Bartra J., González-Mancebo E., Feo-Brito F., Gõmez E., Bartolomé

B., Muñoz-García E., Sanz Maroto A., Vivanco F., Cuesta-Herranz Jand Pastor-Vargas C, 2014. Molecular Characterization of Contact Urticaria in Patients with Melon Allergy. *British Journal of Dermatology*, 170(3),651–656.

Gao Y., Long R., Kang J., Wang Z., Zhang T., Sun H., Li X., and Yang Q, 2019. Comparative Proteomic Analysis Reveals That Antioxidant System and Soluble Sugar Metabolism Contribute to Salt Tolerance in Alfalfa (Medicago Sativa L.) Leaves. *Journal of Proteome Research*, 18(1),191–203.

Gaweł E., Grzelak Mand Janyszek M, 2017. Lucerne (Medicago Sativa L.) in the Human Diet—Case Reports and Short Reports. *Journal of Herbal Medicine*, 10,8–16.

Giovannini M., D'Auria E., Caffarelli C., Verduci E., Barberi S., Indinnimeo L., Iacono ID., Martelli A., Riva E., and Bernardini R, 2014. Nutritional Management and Follow up of Infants and Children with Food Allergy: Italian Society of Pediatric Nutrition/Italian Society of Pediatric Allergy and Immunology Task Force Position Statement. *Italian Journal of Pediatrics*, 40(1),1–9.

Goldring JPD, 2019. Measuring Protein Concentration with Absorbance, Lowry, Bradford Coomassie Blue, or the Smith Bicinchoninic Acid Assay before Electrophoresis. *Methods in Molecular Biology*, 1855,31–39.

Gomez F., Aranda A., Campo P., Diaz-Perales A., Blanca-Lopez N., Perkins J., Garrido M., Blanca M., Mayorga C and Torres MJ, 2014. High Prevalence of Lipid Transfer Protein Sensitization in Apple Allergic Patients with Systemic Symptoms. *PLoS ONE*, 9(9).

Greenhawt M., Shaker M., Wang J., Oppenheimer JJ., Sicherer S., Keet C., Swaggart K., Rank M., Portnoy J M., Bernstein J., Chu DK., Dinakar C., Golden D., Horner C., Lang DM., Lang ES., Khan DA., Lieberman J., Stukus D and Wallace D, 2020. Peanut Allergy Diagnosis: A 2020 Practice Parameter Update, Systematic Review, and GRADE Analysis. *Journal of Allergy and Clinical Immunology*, 146(6),1302–1334.

Gupta RS., Warren CM., Smith BM., Blumenstock JA., Jiang J., Davis MM and Nadeau KC, 2018. The Public Health Impact of Parent-Reported Childhood Food Allergies in the United States. *Pediatrics*, 142(6).

Gutsch A., Zouaghi S., Renaut J., Cuypers A., Hausman JF and Sergeant K, 2018.

Changes in the Proteome of Medicago Sativa Leaves in Response to Long-Term Cadmium Exposure Using a Cell-Wall Targeted Approach. *International Journal of Molecular Sciences*, 19(9).

Hallmann E., Ponder A., Aninowski M., Narangerel T and Leszczyńska J, 2020. The Interaction between Antioxidants Content and Allergenic Potency of Different Raspberry Cultivars. *Antioxidants* 9(3).

Hamad NA., Al-Falahi N., Lafi SA and Hassan AS, 2020. Role of Alpha Trypticase Enzyme, Total IgE, Specific IgE & IgG4 In Patients with Food Allergy.*Egyptian Academic Journal of Biological Sciences G. Microbiology*, 12(2),49–58.

Hasan AS., Jabeen M and Ul Hasnain A, 2007. *BIOCHEMICAL CHARACTERIZATION AND THERMOSTABLE CAPACITY OF PARVALBUMINS: THE MAJOR FISH-FOOD ALLERGENS*. *Journal of Food Biochemistry,31(1),121-137*.

Hawkins C and Yu XL, 2018. Recent Progress in Alfalfa (Medicago Sativa L.) Genomics and Genomic Selection. *Crop Journal*, 6(6),565–575.

Hayashi S., Ishii T., Matsunaga T., Tominaga R., Kuromori T., Wada T., Shinozaki K and Hirayama T, 2008. The Glycerophosphoryl Diester Phosphodiesterase-like Proteins SHV3 and Its Homologs Play Important Roles in Cell Wall Organization."*Plant and Cell Physiology*, 49(10),1522–1535.

He F, 2011. BCA (Bicinchoninic Acid) Protein Assay.*Bio-protocol*, 1(5).

Hildebrand HV., Arias A., Simons E., Gerdts J., Povolo B., Rothney J and Protudjer JLP, 2021. Adult and Pediatric Food Allergy to Chickpea, Pea, Lentil, and Lupine: A Scoping Review. *Journal of Allergy and Clinical Immunology: In Practice*, 9(1):290-301.

Huby RDJ., Dearman RJ and Kimber I, 2000. *Why Are Some Proteins Allergens? Toxicological sciences*, 55, 235–246.

Ibsaine O., Djenouhat K., Lemdjadani N., Arrada Z and Berrah H, 2010. Etude Cas Témoins Des Facteurs de Risque de LAPLV IgE-Médiée. *Archives de Pédiatrie*, 21(5),847.

Inomata N, 2019. Gibberellin-Regulated Protein Allergy: Clinical Features and Cross-

Reactivity. *Allergology International*, 69:11–18.

Irani C and Maalouly G, 2015. Prevalence of Self-Reported Food Allergy in Lebanon: A Middle-Eastern Taste. *International Scholarly Research Notices*, 2015,1–5.

Ishibashi M., Nabe T., Nitta Y., Tsuruta H., Iduhara M and Uno Y, 2017. Analysis of Major Paralogs Encoding the Fra a 1 Allergen Based on Their Organ-Specificity in Fragaria × Ananassa. *Plant Cell Reports* 37, 411–424.

Ito K., Yamamoto T., Oyama Y., Tsuruma R and Saito E, 2016. Food Allergen Analysis for Processed Food Using a Novel Extraction Method to Eliminate Harmful Reagents for Both ELISA and Lateral-Flow Tests. *Analytical and Bioanalytical Chemistry*, 480, 5973–5984.

Iwaya M., Murakami G., Kurose K., Matsuno M., Onoue Y., Takayanagi M., Kayahara M., Adachi Y., Adachi Y and Okada T, 1994. A Case of Anaphylaxis Caused by Sunflower Seed. *Arerugī = [Allergy]*, 43(6),696–700.

Jensen LB., Pedersen MH., Skov PS., Poulsen LK., Bindslev-Jensen C., Andersen SB and Torp A M, 2008. Peanut Cross-Reacting Allergens in Seeds and Sprouts of a Range of Legumes. *Clinical and Experimental Allergy*, 38(12),1969–1977.

Jeong KY, 2016. Physical and Biochemical Characteristics of Allergens. *Allergy, Asthma & Respiratory Disease*, 4(3),157.

Jorge A., Soares E., Sarinho E., Lorente F., Gama J and Taborda-Barata L, 2017. Prevalence and Clinical Features of Adverse Food Reactions in Portuguese Children. *Allergy, Asthma and Clinical Immunology*, 13(40).

Kalic T., Ellinger I., Kamath SD., Palladino C., Mayr V., Tscheppe A., Ruethers T., Waltl EE., Niederberger V., Lengger N., Radauer C., Hafner C., Lopata AL., Bublin M and Breiteneder H, 2019. Fish-Derived Low Molecular Weight Components Modify Bronchial Epithelial Barrier Properties and Release of pro-Inflammatory Cytokines. *Molecular Immunology* 112,140–50.

Kielkopf CL., Bauer W and Urbatsch IL, 2020. Bradford Assay for Determining Protein Concentration.*Cold Spring Harbor Protocols*, 2020(4):136–38.

Kobayashi Y., Yang T., Yu CT., Ume C., Kubota H., Shimakura K., Shiomi K and

Hamada-Sato N, 2016. Quantification of Major Allergen Parvalbumin in 22 Species of Fish by SDS-PAGE. *Food Chemistry*, 194,345–353.

Kuehn A., Swoboda I., Arumugam K., Hilger C and Hentges F, 2014. Fish Allergens at a Glance: Variable Allergenicity of Parvalbumins, the Major Fish Allergens. *Frontiers in Immunology* 5(179).

Ladics GS, 2019. Assessment of the Potential Allergenicity of Genetically-Engineered Food Crops. *Journal of Immunotoxicology*, 16(1):43–53.

Latreche A, 2009. L ' Approche Diagnostic de l ' Allergie Alimentaire Dans l ' Est Algérien. Thesis. Université Mentouri Constantine, Algeria.

LavineE and Ben-Shoshan M, 2015. Allergy to Sunflower Seed and Sunflower Butter as Proposed Vehicle for Sensitization. *Allergy, Asthma and Clinical Immunology*, 11(1).

Lee PW., Nordlee JA., Koppelman SJ., Baumert JL and Taylor SL, 2011. Evaluation and Comparison of the Species-Specificity of 3 Antiparvalbumin IgG Antibodies. *Journal of Agricultural and Food Chemistry*, 59(23),12309–12316.

Lehrer SB., Reese G., Malo JL., Lahoud C., Leong-Kee S., Goldberg B., Carle T and Ebisawa M, 1999. Corn Allergens: IgE Antibody Reactivity and Cross-Reactivity with Rice, Soy, and Peanut. *International Archives of Allergy and Immunology*, 118, 298–299.

Levin KA, 2006. Study Design III: Cross-Sectional Studies. *Evidence-Based Dentistry*, 7(1):24–25.

Levin ME., Botha M., Basera W., Facey-Thomas HE., Gaunt B., Gray CL., Kiragu W., Ramjith J., WatkinsA and Genuneit J, 2020. Environmental Factors Associated with Allergy in Urban and Rural Children from the South African Food Allergy (SAFFA) Cohort. *Journal of Allergy and Clinical Immunology*, 145(1),415–426.

Linhart B., Freidl R., Elisyutina O., Khaitov M., Karaulov A and Valenta R, 2019. Molecular Approaches for Diagnosis, Therapy and Prevention of Cow´s Milk Allergy. *Nutrients*, 11(7).

Liu L., Dai X., Kang H., Xu Y and Hao W, 2020. Structural and Functional Properties of Hydrolyzed/Glycosylated Ovalbumin under Spray Drying and Microwave Freeze

Drying. *Food Science and Human Wellness*, 9(1):80–87.

Livia A., Iorga S., Mosoiu C., Racovita RC., Niculae OM and Vlasceanu G, 2017. ALFALFA CONCENTRATE – A RICH SOURCE OF NUTRIENTS FOR USE IN FOOD PRODUCTS. *Journal of International Scientific Publications* 5, 66–73.

Loh W and Tang MLK, 2018. The Epidemiology of Food Allergy in the Global Context. *International Journal of Environmental Research and Public Health*, 15(9).

Long R., Gao Y., Sun H., Zhang T., Li X., Li M., Sun Y., Kang J., Wang Z., Ding W and Yang Q, 2018. Quantitative Proteomic Analysis Using ITRAQ to Identify Salt-Responsive Proteins during the Germination Stage of Two Medicago Species. *Scientific Reports*, 8(1).

Lopes O., Camargo L., Genov IR., Cabral EC., Mello YAMF., Mallozi MC and Solé D, 2011. Anaphylaxis to Beetroot (Beta Vulgaris): A Case Report. *Clinical and Translational Allergy*, 1,1.

Lorenz AR., ScheurerS and Vieths S, 2015. Food Allergens: Molecular and Immunological Aspects, Allergen Databases and Cross-Reactivity. *Chemical Immunology and Allergy*, 101,18–29.

Lyons, S A., Clausen M., Knulst AC., Ballmer-Weber BK., Fernandez-Rivas M., Barreales L., Bieli C., Dubakiene R., Fernandez-Perez C., Jedrzejczak-Czechowicz M., Kowalski ML., Kralimarkova T., Kummeling I., Mustakov TB., Papadopoulos NG., Popov TA., Xepapadaki P., Welsing PMJ., Potts J., Mills ENC., Van-Ree R., Burney PGJ and Le TM, 2020. Prevalence of Food Sensitization and Food Allergy in Children Across Europe. *Journal of Allergy and Clinical Immunology: In Practice*, 8(8),2736-2746.

Ma X., Liang R., Yang X., Gou J., Li Y and Lozano-Ojalvo D, 2020. Simultaneous Separation of the Four Major Allergens of Hen Egg White. *Journal of Chromatography B: Analytical Technologies in the Biomedical and Life Sciences*, 1152.

Mahroug H, 2010. Contribution à l'etude de Certaines Proteines Allergènes Alimentaires d'origine Végétale et Dertermination de Relations Entres Différents Parametres Physicochimiques.Thesis. Université Mentouri Constantine,Algerie.

Maintz L and Novak N, 2007. Histamine and Histamine Intolerance. *American Journal of Clinical Nutrition,* 85(5),1185–1196.

Malone MH and Metcalfe DD, 1986. Histamine in Foods: Its Possible Role in Non-Allergic Adverse Reactions to Ingestants. *New England and Regional Allergy Proceedings,* 7(3),241–245.

Mari A., Scala E., Palazzo P., Ridolfi S., Zennaro D and Carabella G, 2006. Bioinformatics Applied to Allergy: Allergen Databases, from Collecting Sequence Information to Data Integration. The Allergome Platform as a Model. *Cellular Immunology,* 244(2),97–100.

Marzban G., Maghuly F., Herndl A and Katinger LH,.2008. Screening and Identification of Putative Allergens in Berry Fruits of the Rosaceae Family : Technical Challenges. *IOS Press*, 34, 37–46.

Matsumoto H., Haniu H and Komori N, 2019. Determination of Protein Molecular Weights on SDS-PAGE. *Methods in Molecular Biology*,1855,101–105.

Maurer-Stroh S., Krutz NL., Kern PS., Gunalan V., Nguyen MN., Limviphuvadh V., Eisenhaber F and Gerberick GF, 2019. AllerCatPro-Prediction of Protein Allergenicity Potential from the Protein Sequence. *Bioinformatics,* 35(17),3020–3027.

Mehta H., Groetch M and Wang J, 2013. Growth and Nutritional Concerns in Children with Food Allergy. *Curr Opin Allergy Clin Immunol*, 13(3), 275–79.

Mehta H., Ramesh M., Feuille E., Groetch M and Wang J, 2014. Growth Comparison in Children with and without Food Allergies in 2 Different Demographic Populations. *The Journal of Pediatrics, 165(4), 842-848.*

Messina M and Venter C, 2020. Recent Surveys on Food Allergy Prevalence. *Nutrition Today,* 55(1),22–29.

Meyer R., Wright K., Vieira MC., Chong KW., Chatchatee P., Vlieg-Boerstra BJ., Groetch M., Dominguez-Ortega G., Heath S., Lang A., Archibald-Durham L., Rao R., De Boer R., Assa'ad A., Trewella E and Venter C, 2019. International Survey on Growth Indices and Impacting Factors in Children with Food Allergies. *Journal of Human Nutrition and Dietetics,* 32(2),175–184.

Meyer R, 2018. Nutritional Disorders Resulting from Food Allergy in Children. *Pediatric Allergy and Immunology*, 29(7),689–704.

Mielmann A, 2013. The Utilisation of Lucerne: A Review. *British Food Journal*, 115(4),590–600.

Mine Y,1995. Recent Advances in the Understanding of Egg White Protein Functionality. *Trends in Food Science and Technology*, 6(7),225–232.

Morris JS., Caldo KMP., Liang S and Facchini PJ, 2021. PR10/Bet v1-like Proteins as Novel Contributors to Plant Biochemical Diversity. *ChemBioChem*, 22(2),264–287.

Mur GP., Brito FF., Iglesias MA., Vega LM., and Martínez PB, 2007. Allergic Reaction Caused by a New Hidden Food, Broad Bean Flour. *Allergy: European Journal of Allergy and Clinical Immunology*, 62(11),1340–1341.

Nafti S., Taright S., El Ftouh M., Yassine N., Benkheder A., Bouacha H., Fakhfakh H., Ali-Khoudja M., Texier N and El Hasnaoui A, 2009. Prevalence of Asthma in North Africa: The Asthma Insights and Reality in the Maghreb (AIRMAG) Study. *Respiratory Medicine* 103,S2–S11.

Nicolaou N and Custovic A, 2011. Molecular Diagnosis of Peanut and Legume Allergy. *Current Opinion in Allergy and Clinical Immunology*, 11,222–228.

Nigoskar S, 2007. Amino Acids (The Building Blocks of Proteins). *Biochemistry for Dental Students (Theory and Practical)*,39.

Nwaru BI., Hickstein L., Panesar SS., Roberts G., Muraro A and Sheikh A, 2014. Prevalence of Common Food Allergies in Europe: A Systematic Review and Meta-Analysis. *Allergy: European Journal of Allergy and Clinical Immunology*, 69(8),992–1007.

Obeng BB., Abena AS., Irene A., YazdanbakhshLM., Van Ree R., Boakye DA and Hartgers FC, 2011. Food Allergy in Ghanaian Schoolchildren: Data on Sensitization and Reported Food Allergy. *International Archives of Allergy and Immunology*, 155(1),63–73.

Ochfeld ENand Pongracic JA, 2019. Food Allergy: Diagnosis and Treatment. *Allergy and Asthma Proceedings*, 40(6),446–449.

Ontiveros N., Flores-Mendoza LK., Canizalez-Roman VA and Cabrera-Chavez F, 2014. Food Allergy : Prevalence and Food Technology Approaches for the Control of IgE-Mediated Food Allergy. *Austin Journal of Nutrition and Food Sciences*, 2(5),1029.

Ontiveros N., Valdez-Meza EE., Vergara-Jiménez MJ., Canizalez-Román A., Borzutzky A and Cabrera-Chávez F, 2016. Parent-Reported Prevalence of Food Allergy in Mexican Schoolchildren: A Population-Based Study. *Allergologia et Immunopathologia*, 44(6),563–70.

Oriel RC and Wang J, 2019. Diagnosis and Management of Food Allergy. *Pediatric Clinics of North America*, 66(5),941–954.

Orozco-Navarrete B., Kaczmarska Z., Dupeux F., Garrido-Arandia M., Pott D., Perales AD., Casan A., Ma A., Valpuesta Vand Merchante C, 2019. Structural Bases for the Allergenicity of Fra a 1.02 in Strawberry Fruits. *Journal of Agricultural and Food Chemistry*, 68,10951–10961.

Osawa Y., Ito Y., Takahashi N., Sugimoto C., Kohno Y., Mori S., Morikawa T., Kato Y., Okamoto M., Kanno M., Takabayashi T and Fujieda S, 2020. Epidemiological Study of Oral Allergy Syndrome in Birch Pollen Dispersal-Free Regions. *Allergology International*, 69(2),246–252.

Pali-Schöll I and Jensen-Jarolim E, 2019. Gender Aspects in Food Allergy. *Current Opinion in Allergy and Clinical Immunology*, 19(3),249–255.

Palma-Carlos AG., Palma-Carlos ML and Tengarrinha F, 2005. Allergy to Sunflower Seeds. *European Annals of Allergy and Clinical Immunology*, 37(5),183–186.

Pastorello E A., Pravettoni V., Farioli L., Primavesi L., Scibilia J., Piantanida M., Mascheri A and Conti A, 2010. Green Bean (Phaseolus Vulgaris): A New Source of IgE-Binding Lipid Transfer Protein. *Journal of Agricultural and Food Chemistry*, 58(7),4513–4516.

Pavić I and Kolaček S, 2017. Growth of Children with Food Allergy. *Hormone Research in Paediatrics*, 88(1),91–100.

Peeters KABM., Koppelman SJ., Penninks AH., Lebens A., Hefle SL., Taylor SL., Van Hoffen E and Knulst AC, 2009. Clinical Relevance of Sensitization to Lupine in

Peanut-Sensitized Adults. *Allergy*, 64(4),549–555.

Pekar J., Ret D and Untersmayr E, 2018. Stability of Allergens.*Molecular Immunology Journal*, 100,14–20.

Peters RL., Koplin JJ., Gurrin LC., Dharmage SC., Wake M., Ponsonby AL., Tang MLK., Lowe AJ., Matheson M., Dwyer T and Allen KJ, 2017. The Prevalence of Food Allergy and Other Allergic Diseases in Early Childhood in a Population-Based Study: HealthNuts Age 4-Year Follow-Up. *Journal of Allergy and Clinical Immunology*, 140(1),145-153.

Planque M., Arnould T., Dieu M., Delahaut P., Renard P and Gillard N, 2017. Liquid Chromatography Coupled to Tandem Mass Spectrometry for Detecting Ten Allergens in Complex and Incurred Foodstuffs. *Journal of Chromatography A*, 1530,138–151.

Pomés A., Davies JM., Gadermaier G., Hilger C., Holzhauser T., Lidholm J., Lopata AL., Mueller GA., Nandy A., Radauer C., Chan SK., Jappe U., Kleine-Tebbe J., Thomas WR., Chapman MD., Hage M., Van-Ree R., Vieths S., Raulf M and Goodman RE, 2018. WHO/IUIS Allergen Nomenclature: Providing a Common Language. *Molecular Immunology*, 100, 3–13.

Poms RE and Anklam E, 2004. Effects of Chemical, Physical, and Technological Processes on the Nature of Food Allergens. *Journal of AOAC International*, 87(6),1466–1474.

Poole JA., Barriga K., Leung DYM., Hoffman M., Eisenbarth GS., Rewers M and Norris JM, 2006. Timing of Initial Exposure to Cereal Grains and the Risk of Wheat Allergy. *Pediatrics*, 117(6),2175–2182.

Poulsen LK., Hansen TK., Nørgaard A.,. Vestergaard H., Stahl Skov P and Bindslev-Jensen C, 2001. Allergens from Fish and Egg. *Allergy*, 56 (6), 39–42.

Prescott SL., Pawankar R., Allen KJ., Campbell DE., Sinn JKH., Fiocchi A., Ebisawa M., Sampson HA., Beyer K and Lee BW, 2013. A Global Survey of Changing Patterns of Food Allergy Burden in Children. *World Allergy Organization Journal*, 6(1),1–12.

Racil H., Loukil M., Ferah Y., Rouhou C., Chaouch N., Yaalaoui S and Chabbou . 2015. Le 'Pollen-Food Syndrome' à l'olive. *Tunisie Medicale*, 93(5), 326–27.

Rahaman T., Vasiljevic T and Ramchandran L, 2016. Effect of Processing on

Conformational Changes of Food Proteins Related to Allergenicity. *Trends in Food Science and Technology*, 49, 24–34.

Rancé F., Grandmottet X and Grandjean H, 2005. Prevalence and Main Characteristics of Schoolchildren Diagnosed with Food Allergies in France. *Clinical and Experimental Allergy*, 35(2),167–172.

Ring J., Klimek L and Worm M, 2018. Adrenaline in the Acute Treatment of Anaphylaxis. *Deutsches Arzteblatt International*, 115, 528–534.

Rona RJ., Keil T., Summers C., Gislason D., Zuidmeer L., Sodergren E., Sigurdardottir ST., Lindner T., Goldhahn K., Dahlstrom J., McBride D and Madsen C, 2007. The Prevalence of Food Allergy: A Meta-Analysis. *Journal of Allergy and Clinical Immunology*, 120(3), 638–646.

Ruiz-Hornillos JM., López-Matas A., Jimeno PB, Henríquez A., Blanco S., Seoane-Rodríguez M., Mahíllo I and Carnés J, 2020. Profilin Is a Marker of Severity in Allergic Respiratory Diseases. *Allergy: European Journal of Allergy and Clinical Immunology*, 75(4),853–861.

Sahu SS., Loaiza CD and Kaundal R, 2019. Plant-MSubP: A Computational Framework for the Prediction of Single and Multi-Target Protein Subcellular Localization Using Integrated Machine-Learning Approaches. *AoB PLANTS*, 12(2),1–10.

Sampson HA., O'Mahony L., Burks AW., Plaut M., Lack G and Akdis CA, 2018. Mechanisms of Food Allergy. *Journal of Allergy and Clinical Immunology*, 141(1), 11–19.

Schalk K., Lexhaller B., Koehler P and Scherf KA, 2017. Isolation and Characterization of Gluten Protein Types from Wheat , Rye , Barley and Oats for Use as Reference Materials. *PloS one*, 12(2).

Scheibe B., Weiss W., Ruëff F., Przybilla B and Görg A, 2001. Detection of Trace Amounts of Hidden Allergens: Hazelnut and Almond Proteins in Chocolate. *Journal of Chromatography B: Biomedical Sciences and Applications*, 756, 229–37.

Schein CH., Ivanciuc O and Braun W, 2007. Bioinformatics Approaches to Classifying Allergens and Predicting Cross-Reactivity. *Immunology and Allergy Clinics of North*

America, 27(1), 1–27.

Sha L., Shao M., Liu C., Wu Y and Chen Y, 2019. A Cross-Sectional Study of the Prevalence of Food Allergies among Children Younger than Ages 14 Years in a Beijing Urban Region. *Allergy and Asthma Proceedings*, 40(1), E1–E7.

Sharma HP., Bansil S and Uygungil B, 2015. Signs and Symptoms of Food Allergy and Food-Induced Anaphylaxis. *Pediatric Clinics of North America*, 62(6), 1377–1392.

Shirasaki H., Yamamoto T., Abe S., Kanaizumi E., Kikuchi M and Himi T, 2017. Clinical Benefit of Component-Resolved Diagnosis in Japanese Birch-Allergic Patients with a Convincing History of Apple or Peach Allergy. *Auris Nasus Larynx*, 44(4), 442–446.

Singh S., Taneja B., Salvi SS and Agrawal A, 2009. Physical Properties of Intact Proteins May Predict Allergenicity or Lack Thereof. *PLoS ONE*, 4(7).

Slavokhotova AA., Shelenkov AA., Andreev YA and Odintsova TI, 2017. Hevein-like Antimicrobial Peptides of Plants. *Biochemistry (Moscow)*, 82(13), 1659–1674.

Souiki L, 2000. Contribution à l'éxtraction et à la caractérisation de certaines protéines allergènes d'origine alimentaire, thèse magister, Annaba, Université Badji Mokhtar.

Sui X., Wu Q., Chang W., Fan X and Song F, 2018. Proteomic Analysis of the Response of Funnelifor Mismosseae/Medicago Sativa to Atrazine Stress. *BMC Plant Biology*, 18(1).

Suliaman FA., Holmes WF., Kwick S., Khouri F and Ratard R, 1997. Pattern of Immediate Type Hypersensitivity Reactions in the Eastern Province, Saudi Arabia. *Annals of Allergy, Asthma & Immunology : Official Publication of the American College of Allergy, Asthma, & Immunology*, 78(4), 415–418.

Thomas L, 2020. Cross-Sectional Study | Definitions, Uses & Examples. [on ligne], Retrieved: August 14, 2020. Available on: https://www.scribbr.com/methodology/cross-sectional-study/.

Topçu ZI K., Kaklikkaya N., Baki A and Orhan F, 2018. Characteristics of Beef Allergy in Schoolchildren in Turkey. *Allergy and Asthma Proceedings*, 39(1), 59–65.

Tsabouri S., Douros K and Priftis K, 2014. Cow's Milk Allergenicity. *Endocrine,*

Metabolic & Immune Disorders-Drug Targets, 14(1),16–26.

Ukleja-Sokołowska N., Gawrońska-Ukleja E., Zbikowska-Gotz M., Bartuzi Z and Sokołowski Ł, 2016. Sunflower Seed Allergy. *International Journal of Immunopathology and Pharmacology*, 29(3), 498–503.

Uniacke-Lowe T and Fox P. F, 2011. Milk: Equid Milk. in *Encyclopedia of Dairy Sciences: Second Edition*. Elsevier Inc, 518–529.

Vadas P and P Boris, 2003. Presence of Undeclared Peanut Protein in Chocolate Bars Imported from Europe. *Journal of food protection*, 66(10), 1932-1934.

Van der Valk J., Nagl B., Van Wljk RG., Bohle B and De Jong N, 2020. The Effect of Birch Pollen Immunotherapy on Apple and Rmal d 1 Challenges in Adults with Apple Allergy. *Nutrients*, 12(2).

Van Do T., Hordvik I., Endresen C and Elsayed S, 2005. Characterization of Parvalbumin, the Major Allergen in Alaska Pollack, and Comparison with Codfish Allergen M. *Molecular Immunology*, 42(3),345–353.

Varlamova EG and Zaripov OG, 2020. Beta–Lactoglobulin–Nutrition Allergen and Nanotransporter of Different Nature Ligands Therapy with Therapeutic Action. *Research in Veterinary Science*, 133,17–25.

Venter C., Mazzocchi A., Maslin K and Agostoni C, 2017. Impact of Elimination Diets on Nutrition and Growth in Children with Multiple Food Allergies. *Current Opinion in Allergy and Clinical Immunology*, 17(3), 220–226.

Vergeer L., Vanderlee L., White CM., Rynard VL and Hammond D, 2019. Vegetarianism and Other Eating Practices among Youth and Young Adults in Major Canadian Cities. *Public health nutrition*, 23(4), 609-619.

Verhoeckx KCM., Vissers YM., Baumert JL., Faludi R., Feys M., Flanagan S., Herouet-Guicheney C., Holzhauser T., Shimojo R., Van der Bolt N., Wichers H and Kimber I, 2015. Food Processing and Allergenicity. *Food and Chemical Toxicology*, 80,223–240.

Wahler D., Gronover CS., Richter C., Foucu F., Twyman RM., Moerschbacher BM., Fischer R., Muth J and Dirk P, 2009. Polyphenoloxidase Silencing Affects Latex

Coagulation in Taraxacum Species. *Plant Physiology*, 151(1), 334–346.

Wal JM, 2001. Structure and Function of Milk Allergens. *Allergy,* 56(67),35–38.

Wallace IM., Blackshields G and Higgins DG, 2005. Multiple Sequence Alignments. *Current Opinion in Structural Biology,* 15(3), 261–66.

Watson BS., Lei Z., Dixon RA and Sumner LW, 2004. Proteomics of Medicago Sativa Cell Walls. *Phytochemistry*, 65(12),1709–1720.

Wong XL and Sebaratnam DF, 2018. Mammalian Meat Allergy. *International Journal of Dermatology*, 57(12),1433–1436.

World Health Organization (WHO), International Food Safety Authorities Network (INFOSAN), and Food and Agriculture Organization of the United Nations (FAO). 2006. *International Food Safety Authorities Network (INFOSAN) Food Allergies*.

Wu TC., Tsai TC., Huang CF., Chang FY., Lin CC., Huang IF., Chu CH., Lau BH., Wu L., Peng HJ and Tang RB, 2012. Prevalence of Food Allergy in Taiwan: A Questionnaire-Based Survey. *Internal Medicine Journal*, 42(12),1310–1315.

Yagami A, 2010. Anaphylaxis to Lipid Transfer Protein from Sunflower Seeds. *Allergy: European Journal of Allergy and Clinical Immunology*, 65(10),1340–1341.

Zhang C and Shangli S, 2018. Physiological and Proteomic Responses of Contrasting Alfalfa (Medicago Sativa L.) Varieties to Peg-Induced Osmotic Stress. *Frontiers in Plant Science*, 9.

Zukiewicz-Sobczak WA., Wróblewska P., Adamczuk P and Kopczyński, P, 2013. Causes, symptoms and prevention of food allergy. *Postepy dermatologii i alergologii*, *30*(2), 113–116.

Appendixes

Appendix 1. Example of the survey distributed to the parents

استطلاع حول الحساسية الغذائية عند الأطفال

Les allergies alimentaires chez les enfants

جنس الطفل **أنثى** ☐ **ذكر** ☐

العمر : .....

هل عانى/يعاني ابنك من حساسية غذائية ؟

نعم ☐ **لا** ☐

في حالة الإجابة بنعم ماهو الغذاء أو الأغذية المسببة لهذه الحساسية مع ذكر العمر عند ظهور أول تحسس ؟

.....

.....

كيف ظهر التحسس على ابنك؟

طفح جلدي ☐ **مشاكل تنفسية** ☐ **مشاكل هضمية** ☐

هل تخلص ابنك من هذه الحساسية؟ ☐ ☐

في حالة الإجابة بنعم في اي سن تخلص منها؟

.....

هل يوجد افراد اخرون في العائلة يعانون من الحساسية الغذائية ؟

☐ **نعم** ☐ **لا**

في حالة الاجابة بنعم اذكر الاغذية التي يتحسسون منها ؟

الاب.....

الام .....

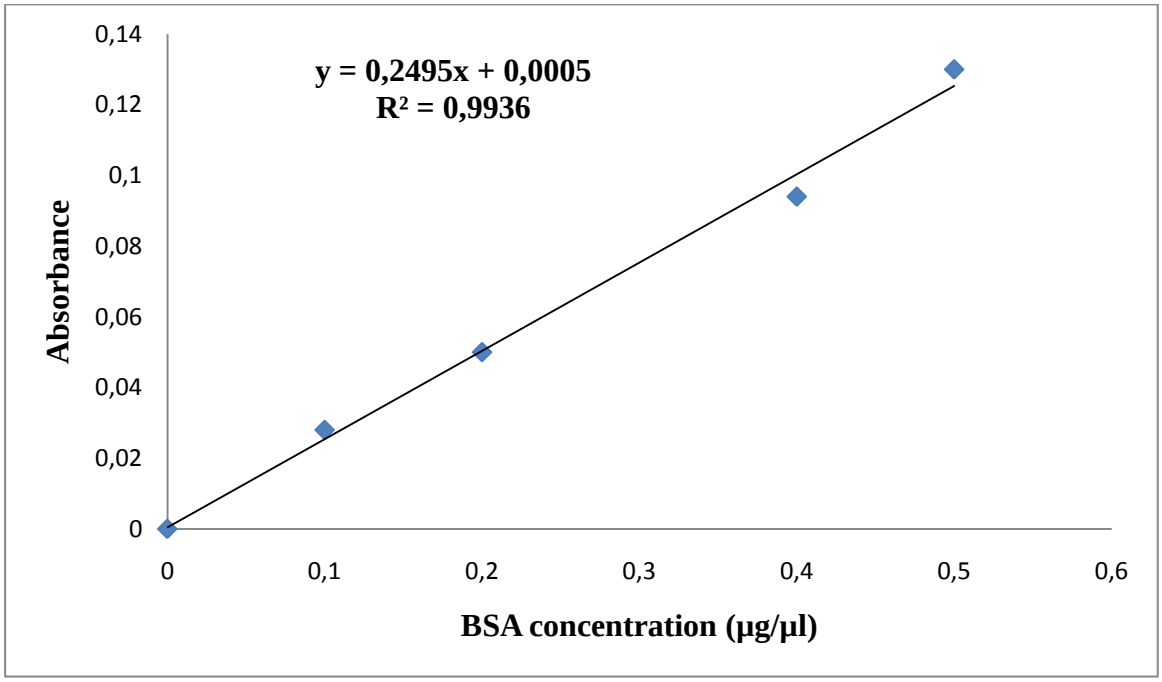
الإخوة.....

Appendix 2. Protein estimation by BicinChoninic Acid assay (BCA assay)

The measurements of the concentration of the BSA and the standard curve are given in table bellow and figure represents the standard curve.

Concentrations of BSA solutions used for the standard curve.

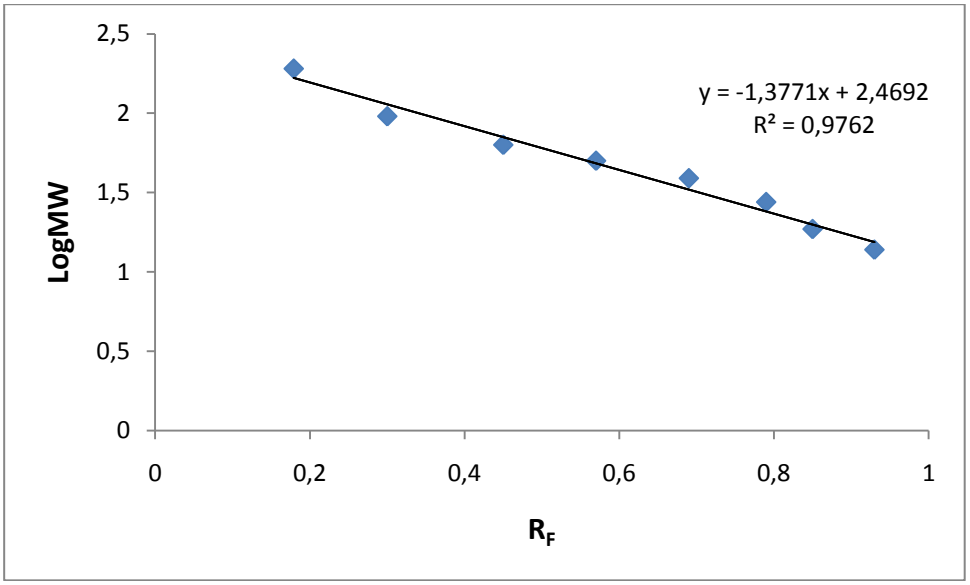
BSA concentration (µg)	0	0.1	0.2	0.4	0.5	0.6	0.8	1
Absorbance (562 nm)	0	0.028	0.05	0.094	0.13	0.144	0.195	0.243



Standard curve for the determination of the concentration

Appendix 3. Determination of the molecular weight through SDS-PAGE

R_f	0.93	0.85	0.79	0.69	0.57	0.45	0.30	0.17
LogMW	1.14	1.27	1.44	1.59	1.7	1.8	1.98	2.28



Standard curve for the determination of the molecular weight

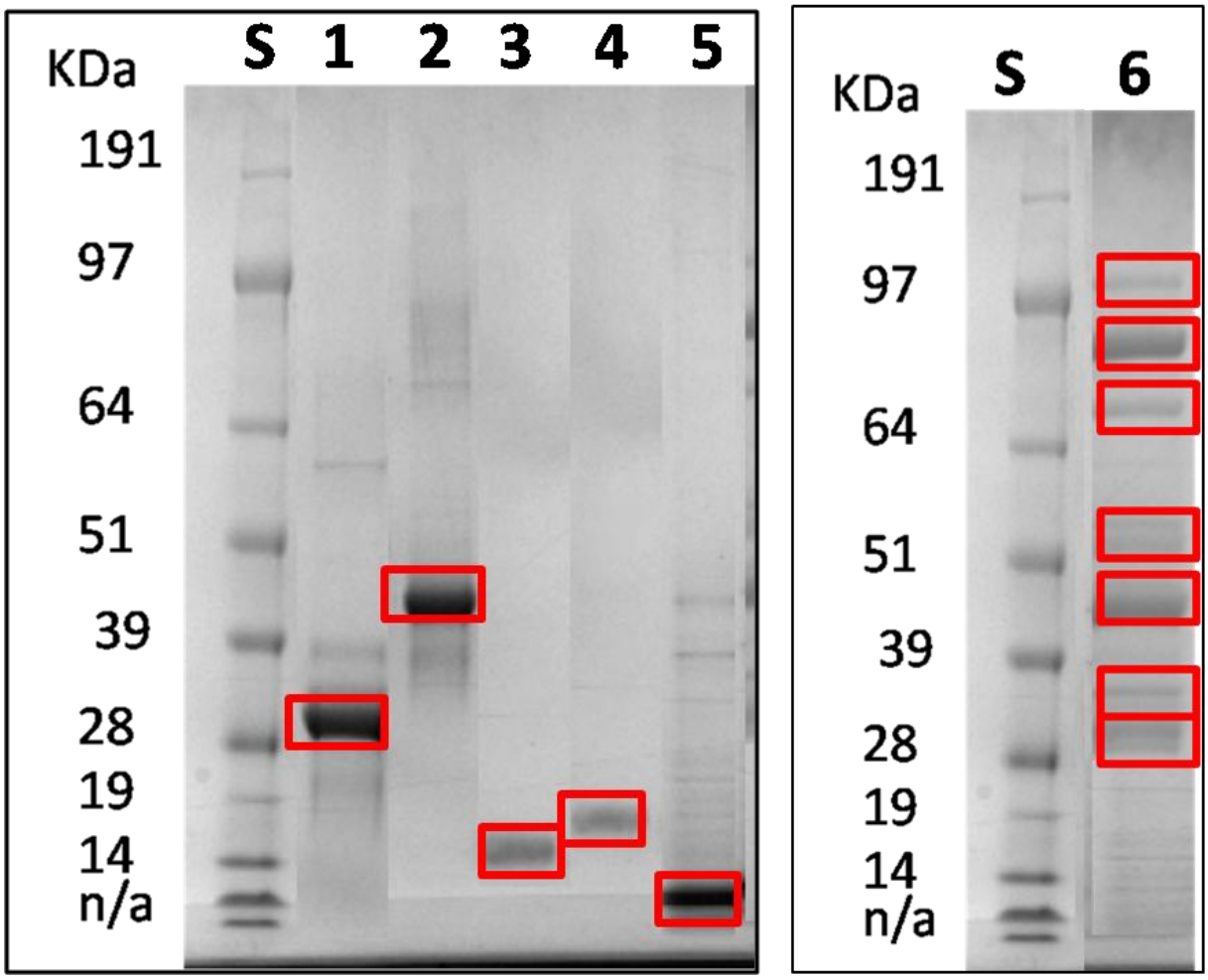
Appendix 4. Amino acid codes

Amino acid	Abreviation	Amino acid	Abreviation
Alanine	A	**leucine**	L
Arginine	R	**lysine**	K
Asparagine	R	**Methionine**	M
Aspartic acid	D	**Phenylalanine**	F
Cysteine	C	**Proline**	P
Glutamine	Q	**Serine**	S
Glutamic acid	E	**Threonine**	T
Glycine	G	**Tryptophan**	W
histidine	H	**Tyrosine**	Y
Isoleucine	I	**Valine**	V
Pyrrolysine	O	**Selenocysteine**	U
Aspartic acid or Asparagine	B	**Any amino acid**	X
Glutamic acid or Glutamine	Z	**Leucine or Isoleucine**	J

Appendix 5. Analysis of differences between modalities with a 95% confidence interval byTukey test.

Pairs	Standardized difference	Critical values	p value	Significance
GLIA vs ALA	14.035	3.359	< **0.0001**	Yes
GLIA vs OVO	13.001	3.359	< **0.0001**	Yes
GLIA vs OVA	11.376	3.359	< **0.0001**	Yes
GLIA vs PV	9.357	3.359	< **0.0001**	Yes
GLIA vs BLA	9.234	3.359	< **0.0001**	Yes
BLA vs ALA	4.802	3.359	0.084	No
BLA vs OVO	3.767	3.359	0.250	No
BLA vs OVA	2.142	3.359	0.329	No
BLA vs PV	0.123	3.359	1.000	No
PV vs ALA	4.678	3.359	0.455	No
PV vs OVO	3.644	3.359	0.311	No
PV vs OVA	2.019	3.359	0.386	No
OVA vs ALA	2.659	3.359	0.156	No
OVA vs OVO	1.625	3.359	0.599	No
OVO vs ALA	1.034	3.359	0.897	No

Appendix 6. SDS-PAGE of the extracted allergens



A: allergens from aminal source (S: Standars 1: Ovomucoid 2: Ovalbumin 3: α-lactalbumin 4: β-lactoglobulin5: parvalbumin)

B: Allergen from vegetal source (wheat). S: Standard. 6: Gliadins.

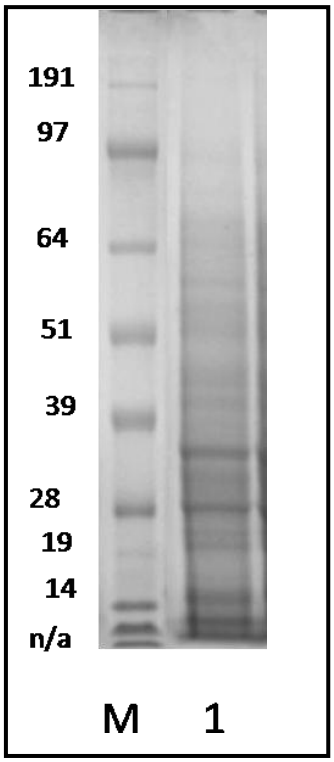
Appendix 7. Multiple sequence alignment of the six extracted allergens

```
P01012 OVAL_CHICK      1   ------------------------------------------------------------   0
P01005 IOVO_CHICK      1   ------------------------------------------------------------   0
P00711 LALBA_BOVIN     1   ------------------------------------------------------------   0
P02754 LACB_BOVIN      1   ------------------------------------------------------------   0
P09227 PRVA_CYPCA      1   ------------------------------------------------------------   0
P02863 GDA0_WHEAT      1   ------------------------------------------------------------   0
P08079 GDB0_WHEAT      1   ------------------------------------------------------------   0
Q6PNA3 Q6PNA3_WHEAT    1   MKPHHDGYKYTCSIIVTFHYPNFKHQDQKHQFQESIKHKSKMKTFIIFVLLSMPMSIVIA   60

P01012 OVAL_CHICK      1   ------------------------------------------------------------   0
P01005 IOVO_CHICK      1   ------------------------------------------------------------   0
P00711 LALBA_BOVIN     1   ------------------------------------------------------------   0
P02754 LACB_BOVIN      1   ------------------------------------------------------------   0
P09227 PRVA_CYPCA      1   ------------------------------------------------------------   0
P02863 GDA0_WHEAT      1   ------------------------------------------------------------   0
P08079 GDB0_WHEAT      1   ------------------------------------------------------------   0
Q6PNA3 Q6PNA3_WHEAT   61   ARHLNPSDQELQSPQQQFLEKTIISAATISTSTIFTTTTISHTPTIFPPSTTTTISPTPT   120

P01012 OVAL_CHICK      1   ---------------------------------------------MGSIGAASMEFCF   13
P01005 IOVO_CHICK      1   ------------------------------------------------------------   0
P00711 LALBA_BOVIN     1   ------------------------------------------------------------   0
P02754 LACB_BOVIN      1   ------------------------------------------------------------   0
P09227 PRVA_CYPCA      1   ------------------------------------------------------------   0
P02863 GDA0_WHEAT      1   ------------------------------------------------------------   0
P08079 GDB0_WHEAT      1   ------------------------------------------------------------   0
Q6PNA3 Q6PNA3_WHEAT  121   TNPPTTTMTIPLATPTTTTTFSPAPTTISLATTTTISLAPTTNSPITTTTIPAATPETTT   180
P01012 OVAL_CHICK     14   ------------DVFKELKVHHANENIFYCPIAIMSALAMVYLGA----KDSTR------    51
P01005 IOVO_CHICK      1   ------------------------------------------------------------    0
P00711 LALBA_BOVIN     1   ------------------------------------------------------------    0
P02754 LACB_BOVIN      1   ------------------------------------------------------------    0
P09227 PRVA_CYPCA      1   ----------------------------------------MAYGGI---------------    6
P02863 GDA0_WHEAT      1   --------------MKTFLILV-----------LL-AIVATTATTAVR------------   22
P08079 GDB0_WHEAT      1   --------------MKTLLILT-----------IL-AMAITIGTANMQVDPSS-------   27
Q6PNA3 Q6PNA3_WHEAT  181   TIPPATRTNNYASTATTISLLTATTTPPATPTTILSATTTTISPAPTIISPATRTNNSLA   240

P01012 OVAL_CHICK     52   ----------TQINKVVRFDKLPGFGDSIEAQCGTS-----------VNVHSSLRDIL--    88
P01005 IOVO_CHICK      1   ------------------------------------------------------------    0
P00711 LALBA_BOVIN     1   ------------------------------------------------------------    0
P02754 LACB_BOVIN      1   ------------------------------------------------------------    0
P09227 PRVA_CYPCA      7   ------------------------------------------------------------    6
P02863 GDA0_WHEAT     23   ------------FPVPQ-----LQPQNPSQQQPQEQVPL----------------VQQQ    48
P08079 GDB0_WHEAT     28   ---------QVQWPQQQ---PVPQPHQPFSQQPQQTFPQPQQTFPHQPQQQFPQPQQPQQ    75
Q6PNA3 Q6PNA3_WHEAT  241   TPTTIPPATATTIPPATRTNNSPATATTIPPAPQQRFPHTRQKFPRNPNNHS----LCST   296

P01012 OVAL_CHICK     89   ----NQ---ITKPNDVYSFS----LASRLYAEERYPILPEYLQCV--------------   122
P01005 IOVO_CHICK      1   ------------------------------------------------------------     0
P00711 LALBA_BOVIN     1   ------------------------------------------------------------     0
P02754 LACB_BOVIN      1   ------------------------------------------------------------     0
P09227 PRVA_CYPCA      7   ------------------------------------------------------------     6
P02863 GDA0_WHEAT     49   QFLGQQQPFPPQ--QPYPQPQPFPSQLPYLQLQPFPQPQLPY-----SQPQPFRPQQPYP   101
P08079 GDB0_WHEAT     76   QFLQPQQPFPQQPQQPYPQQ----------PQQPFPQTQQPQ----QLFPQSQQPQQQFS   121
Q6PNA3 Q6PNA3_WHEAT  297   HHFPAQQPFPQQPGQIIPQQPQQ--PLPLQPQQPFPWQPEQRSSQQPQQPFSLQPQQPFS   354
```

Appendixes

```
P01012 OVAL_CHICK      303  ------------------------------ITDVFSSSANLSGISSAESLKISQAVHA  330
P01005 IOVO_CHICK       89  TVPMNCSSYANTTSEDGKV----------MVLCNRAFNPVCGTDGVTYDNECLL-----C  133
P00711 LALBA_BOVIN      89  SN--ICNISCDKFLDDD-----------------------------------LTDDIMC  110
P02754 LACB_BOVIN      132  SLACQCLVRTPEV-DDEALEKFDKALKALPMHIRLSFNPT------QLEEQCHI------  178
P09227 PRVA_CYPCA       82  -------------------------------ETKAFLKA-------------------   89
P02863 GDA0_WHEAT      263  -LPAMCNVYIPPYCTIAPFGIFGTN---------------------------------  286
P08079 GDB0_WHEAT      252  ------------------------------------------------------------  251
Q6PNA3 Q6PNA3_WHEAT    355  ------------------------------------------------------------  354

P01012 OVAL_CHICK      331  AHAEINEAGREVV----GSAEAGVDAASV------------------SEEFRADHPFLF  367
P01005 IOVO_CHICK      134  AHK--VEQGASVDKRHDGGCRKELAAVSVDCSEYPKPDCTAEDRPLCGSDNKTYGNKCNF  191
P00711 LALBA_BOVIN     111  VKKILDKVGINYWLAHKALCSEKLDQWL-------------------------------  138
P02754 LACB_BOVIN      179  ------------------------------------------------------------  178
P09227 PRVA_CYPCA       90  --------G-DSD----GDGKIGVDEFA-----------------------A------  105
P02863 GDA0_WHEAT      287  ------------------------------------------------------------  286
P08079 GDB0_WHEAT      252  ------------------------------------------------------------  251
Q6PNA3 Q6PNA3_WHEAT    355  ------------------------------------------------------------  354

P01012 OVAL_CHICK      368  CIKHIA---TNAVLFFGRCVSP                                      386
P01005 IOVO_CHICK      192  CNAVVESNGTLTLSHFGKC---                                      210
P00711 LALBA_BOVIN     139  CEKL-----------------                                       142
P02754 LACB_BOVIN      179  ---------------------                                       178
P09227 PRVA_CYPCA      106  LVKA-----------------                                       109
P02863 GDA0_WHEAT      287  ---------------------                                       286
P08079 GDB0_WHEAT      252  ---------------------                                       251
Q6PNA3 Q6PNA3_WHEAT    355  ---------------------                                       354
P01012 OVAL_CHICK      303  ------------------------------ITDVFSSSANLSGISSAESLKISQAVHA  330
P01005 IOVO_CHICK       89  TVPMNCSSYANTTSEDGKV----------MVLCNRAFNPVCGTDGVTYDNECLL-----C  133
P00711 LALBA_BOVIN      89  SN--ICNISCDKFLDDD-----------------------------------LTDDIMC  110
P02754 LACB_BOVIN      132  SLACQCLVRTPEV-DDEALEKFDKALKALPMHIRLSFNPT------QLEEQCHI------  178
P09227 PRVA_CYPCA       82  -------------------------------ETKAFLKA-------------------   89
P02863 GDA0_WHEAT      263  -LPAMCNVYIPPYCTIAPFGIFGTN---------------------------------  286
P08079 GDB0_WHEAT      252  ------------------------------------------------------------  251
Q6PNA3 Q6PNA3_WHEAT    355  ------------------------------------------------------------  354

P01012 OVAL_CHICK      331  AHAEINEAGREVV----GSAEAGVDAASV------------------SEEFRADHPFLF  367
P01005 IOVO_CHICK      134  AHK--VEQGASVDKRHDGGCRKELAAVSVDCSEYPKPDCTAEDRPLCGSDNKTYGNKCNF  191
P00711 LALBA_BOVIN     111  VKKILDKVGINYWLAHKALCSEKLDQWL-------------------------------  138
P02754 LACB_BOVIN      179  ------------------------------------------------------------  178
P09227 PRVA_CYPCA       90  --------G-DSD----GDGKIGVDEFA-----------------------A------  105
P02863 GDA0_WHEAT      287  ------------------------------------------------------------  286
P08079 GDB0_WHEAT      252  ------------------------------------------------------------  251
Q6PNA3 Q6PNA3_WHEAT    355  ------------------------------------------------------------  354

P01012 OVAL_CHICK      368  CIKHIA---TNAVLFFGRCVSP                                      386
P01005 IOVO_CHICK      192  CNAVVESNGTLTLSHFGKC---                                      210
P00711 LALBA_BOVIN     139  CEKL-----------------                                       142
P02754 LACB_BOVIN      179  ---------------------                                       178
P09227 PRVA_CYPCA      106  LVKA-----------------                                       109
P02863 GDA0_WHEAT      287  ---------------------                                       286
P08079 GDB0_WHEAT      252  ---------------------                                       251
Q6PNA3 Q6PNA3_WHEAT    355  ---------------------                                       354
```

Appendix 8. SDS-PAGE of the extract of the leaf extract from *Medicago sativa*



M: marker. 1: plant leaf protein extract.

Appendixes

Appendix 9. Total 129 proteins identified in the extract

Rank	Description	|Log Prob|
1	>tr|G7JQL4|G7JQL4_MEDTR Glucan endo-1,3-beta-glucosidase OS=Medicago truncatula OX=3880 GN=11440651 PE=3 SV=1	83,00
2	>tr|A0A1L2BU68|A0A1L2BU68_MEDSV Class III-1 chitinase 10 (Fragment) OS=Medicago sativa subsp. varia OX=36902 GN=chit3-1-10 PE=3 SV=1	70,21
3	>tr|A0A396GNA8|A0A396GNA8_MEDTR Endochitinase OS=Medicago truncatula OX=3880 GN=MtrunA17_Chr8g0370821 PE=4 SV=1	69,54
4	>tr|G7IMV1|G7IMV1_MEDTR Alpha-L-arabinofuranosidase/beta-D-xylosidase OS=Medicago truncatula OX=3880 GN=11412498 PE=4 SV=1	62,94
5	>tr|I3T3C6|I3T3C6_MEDTR LRRNT_2 domain-containing protein OS=Medicago truncatula OX=3880 PE=2 SV=1	52,53
6	>tr|C3VM17|C3VM17_MEDSA Chitinase class III-1 OS=Medicago sativa OX=3879 PE=2 SV=1	50,16
7	>tr|A0A1L2BU75|A0A1L2BU75_MEDSV Beta-1,3-glucanase 3 (Fragment) OS=Medicago sativa subsp. varia OX=36902 PE=3 SV=1	49,15
8	>tr|G7IMM5|G7IMM5_MEDTR FAD-binding berberine family protein OS=Medicago truncatula OX=3880 GN=11420001 PE=3 SV=1	46,29
9	>tr|I3SK49|I3SK49_MEDTR Plant basic secretory protein (BSP) family protein OS=Medicago truncatula OX=3880 GN=25485381 PE=2 SV=1	46,12
10	>tr|O48904|O48904_MEDSA Malate dehydrogenase OS=Medicago sativa OX=3879 GN=mmdh PE=2 SV=1	44,59
11	>tr|G7K4B9|G7K4B9_MEDTR Glycerophosphoryl diester phosphodiesterase family protein OS=Medicago truncatula OX=3880 GN=11436650 PE=4 SV=1	44,02
12	>tr|I3STX0|I3STX0_MEDTR Glycoside hydrolase family 18 protein OS=Medicago truncatula OX=3880 GN=25490142 PE=2 SV=1	44,01
13	>tr|A0A072TUC3|A0A072TUC3_MEDTR Pathogenesis-related thaumatin family protein OS=Medicago truncatula OX=3880 GN=25502021 PE=4 SV=1	39,71
14	>tr|A0A396K0C6|A0A396K0C6_MEDTR Putative nepenthesin OS=Medicago truncatula OX=3880 GN=MtrunA17_Chr1g0185661 PE=3 SV=1	38,77
15	>tr|Q5RLX9|Q5RLX9_MEDSA Chitinase OS=Medicago sativa OX=3879 GN=Chi PE=2 SV=1	38,07
16	>tr|G7JQL1|G7JQL1_MEDTR Glucan endo-1,3-beta-glucosidase OS=Medicago truncatula OX=3880 GN=11444942 PE=3 SV=1	35,23
17	>tr|A0A072VNF7|A0A072VNF7_MEDTR Putative tripeptidyl-peptidase II OS=Medicago truncatula OX=3880 GN=25484584 PE=3 SV=1	34,58
18	>tr|A0A396GUY5|A0A396GUY5_MEDTR Putative thaumatin OS=Medicago truncatula OX=3880 GN=MtrunA17_Chr8g0386351 PE=4 SV=1	34,48
19	>tr|A0A072VD04|A0A072VD04_MEDTR Beta-galactosidase OS=Medicago truncatula OX=3880 GN=11412797 PE=3 SV=1	32,28
20	>tr|A0A396JDT6|A0A396JDT6_MEDTR Putative nepenthesin OS=Medicago truncatula OX=3880 GN=MtrunA17_Chr2g0310821 PE=3 SV=1	31,96
21	>tr|G7J6L9|G7J6L9_MEDTR Glycoside hydrolase family 3 protein OS=Medicago truncatula OX=3880 GN=11435958 PE=4 SV=1	30,98

22	>tr|A0A072U3V6|A0A072U3V6_MEDTR Pectinesterase OS=Medicago truncatula OX=3880 GN=25499611 PE=3 SV=1	29,86
23	>sp|O22585|AMYB_MEDSA Beta-amylase OS=Medicago sativa OX=3879 GN=BMY1 PE=2 SV=1	29,30
24	>tr|A0A072UTK4|A0A072UTK4_MEDTR Glycoside hydrolase family 1 protein OS=Medicago truncatula OX=3880 GN=25488733 PE=3 SV=1	29,05
25	>tr|G7L6U5|G7L6U5_MEDTR Germin-like protein OS=Medicago truncatula OX=3880 GN=11429465 PE=2 SV=1	28,63
26	>tr|A0A072VN05|A0A072VN05_MEDTR Chitinase / Hevein / PR-4 / Wheatwin2 OS=Medicago truncatula OX=3880 GN=25484634 PE=4 SV=1	27,89
27	>tr|G7ICF3|G7ICF3_MEDTR Subtilisin-like serine protease OS=Medicago truncatula OX=3880 GN=11418150 PE=3 SV=2	27,31
28	>tr|A0A072VDP0|A0A072VDP0_MEDTR Multi-copper oxidase-like protein OS=Medicago truncatula OX=3880 GN=25490285 PE=4 SV=1	26,96
29	>tr|A0A072VFN2|A0A072VFN2_MEDTR Multi-copper oxidase-like protein OS=Medicago truncatula OX=3880 GN=25482222 PE=4 SV=1	26,20
30	>tr|A0A072TS46|A0A072TS46_MEDTR Putative 2-alkenal reductase (NAD(P)(+)) OS=Medicago truncatula OX=3880 GN=25501340 PE=4 SV=1	26,10
31	>tr|G7I467|G7I467_MEDTR Putative nepenthesin OS=Medicago truncatula OX=3880 GN=11408499 PE=3 SV=1	26,04
32	>tr|I3SGS4|I3SGS4_MEDTR Germin-like protein OS=Medicago truncatula OX=3880 PE=2 SV=1	24,62
33	>tr|I3SEU9|I3SEU9_MEDTR Thioredoxin domain-containing protein OS=Medicago truncatula OX=3880 PE=2 SV=1	24,37
34	>tr|G7JQL0|G7JQL0_MEDTR Glucan endo-1,3-beta-glucosidase OS=Medicago truncatula OX=3880 GN=11445454 PE=3 SV=1	23,37
35	>tr|A0A072V908|A0A072V908_MEDTR Kunitz type trypsin inhibitor OS=Medicago truncatula OX=3880 GN=25487174 PE=4 SV=1	22,38
36	>tr|G7JK03|G7JK03_MEDTR Serine/Threonine kinase, plant-type protein, putative OS=Medicago truncatula OX=3880 GN=11446209 PE=4 SV=2	21,03
37	>tr|G7L865|G7L865_MEDTR Lactoylglutathione lyase OS=Medicago truncatula OX=3880 GN=11416731 PE=3 SV=1	20,81
38	>tr|A0A072UTA7|A0A072UTA7_MEDTR Pectinesterase OS=Medicago truncatula OX=3880 GN=25491636 PE=3 SV=1	20,17
39	>tr|A0A396H709|A0A396H709_MEDTR Putative thaumatin OS=Medicago truncatula OX=3880 GN=MtrunA17_Chr7g0265011 PE=4 SV=1	19,94
40	>tr|G7L9V7|G7L9V7_MEDTR Pectinesterase OS=Medicago truncatula OX=3880 GN=11445926 PE=3 SV=1	19,10
41	>tr|A0A396I7P7|A0A396I7P7_MEDTR Putative glucan endo-1,3-beta-D-glucosidase OS=Medicago truncatula OX=3880 GN=MtrunA17_Chr4g0038941 PE=3 SV=1	19,10
42	>tr|A0A072UWF5|A0A072UWF5_MEDTR Eukaryotic aspartyl protease family protein OS=Medicago truncatula OX=3880 GN=25489024 PE=3 SV=1	18,68
43	>tr|G7ILA8|G7ILA8_MEDTR Beta-xylosidase/alpha-L-arabinofuranosidase-like protein OS=Medicago truncatula OX=3880 GN=11424908 PE=4 SV=1	18,40
44	>tr|I3S355|I3S355_MEDTR AAI domain-containing protein OS=Medicago truncatula OX=3880 PE=2 SV=1	18,12
45	>tr|D7RIC7|D7RIC7_MEDTR Alpha-L-arabinofuranosidase OS=Medicago truncatula OX=3880 GN=Araf1 PE=2 SV=1	17,72

46	>tr|H8Y256|H8Y256_MEDSV Polygalacturonase-inhibiting protein 1 (Fragment) OS=Medicago sativa subsp. varia OX=36902 GN=PGIP1 PE=4 SV=1	17,43
47	>tr|A0A072TQN9|A0A072TQN9_MEDTR Plastocyanin-like domain protein OS=Medicago truncatula OX=3880 GN=25502506 PE=4 SV=1	17,35
48	>tr|I3SZI9|I3SZI9_MEDTR Barwin domain-containing protein OS=Medicago truncatula OX=3880 PE=2 SV=1	16,85

49	>tr|A0A072V710|A0A072V710_MEDTR Glucan endo-1,3-beta-glucosidase OS=Medicago truncatula OX=3880 GN=11416839 PE=3 SV=1	16,40
50	>tr|Q45NK7|Q45NK7_MEDSA Non-specific lipid-transfer protein (Fragment) OS=Medicago sativa OX=3879 PE=2 SV=1	15,72
51	>tr|G7KWW4|G7KWW4_MEDTR Pectinesterase OS=Medicago truncatula OX=3880 GN=11437869 PE=3 SV=1	15,57
52	>tr|G7ID31|G7ID31_MEDTR Acidic endochitinase OS=Medicago truncatula OX=3880 GN=11428131 PE=3 SV=1	15,37
53	>tr|G7KVR2|G7KVR2_MEDTR Polygalacturonase inhibitor protein OS=Medicago truncatula OX=3880 GN=11437801 PE=4 SV=1	15,37
54	>tr|A0A396J9I7|A0A396J9I7_MEDTR Putative tetrahydroberberine oxidase OS=Medicago truncatula OX=3880 GN=MtrunA17_Chr2g0293191 PE=3 SV=1	14,48
55	>tr|A0A396J8U9|A0A396J8U9_MEDTR Putative alpha-glucosidase OS=Medicago truncatula OX=3880 GN=MtrunA17_Chr2g0293101 PE=3 SV=1	14,46
56	>tr|Q2HU31|Q2HU31_MEDTR Glycoside hydrolase family 18 protein OS=Medicago truncatula OX=3880 GN=11422665 PE=3 SV=1	13,75
57	>tr|I3SK73|I3SK73_MEDTR Perchloric acid soluble translation inhibitor-like protein OS=Medicago truncatula OX=3880 GN=25488952 PE=2 SV=1	13,74
58	>tr|I3SBN3|I3SBN3_MEDTR Uncharacterized protein OS=Medicago truncatula OX=3880 PE=2 SV=1	13,59
59	>tr|I3S4B7|I3S4B7_MEDTR Pectinesterase OS=Medicago truncatula OX=3880 PE=2 SV=1	13,33
60	>tr|G7JJJ6|G7JJJ6_MEDTR Non-specific lipid-transfer protein OS=Medicago truncatula OX=3880 GN=11434678 PE=3 SV=1	13,23
61	>tr|I3S2W4|I3S2W4_MEDTR Lactoylglutathione lyase OS=Medicago truncatula OX=3880 PE=2 SV=1	12,72
62	>tr|A0A072TYE4|A0A072TYE4_MEDTR Putative tripeptidyl-peptidase II OS=Medicago truncatula OX=3880 GN=25500620 PE=3 SV=1	11,61
63	>tr|G7KA59|G7KA59_MEDTR Glycoside hydrolase family 3 amino-terminal domain protein OS=Medicago truncatula OX=3880 GN=11432485 PE=4 SV=1	11,47
64	>tr|A0A396HHV5|A0A396HHV5_MEDTR Kunitz-type trypsin inhibitor-like 1 protein OS=Medicago truncatula OX=3880 GN=MtrunA17_Chr6g0476831 PE=4 SV=1	11,34
65	>tr|A0A396J9B4|A0A396J9B4_MEDTR Putative reverse transcriptase, RNA-dependent DNA polymerase OS=Medicago truncatula OX=3880 GN=MtrunA17_Chr2g0280291 PE=3 SV=1	11,10
66	>tr|A0A072UR60|A0A072UR60_MEDTR Glycoside hydrolase family 18 protein OS=Medicago truncatula OX=3880 GN=MTR_4g116920 PE=3 SV=1	10,93
67	>tr|Q45NL7|Q45NL7_MEDSA Thioredoxin OS=Medicago sativa OX=3879 GN=Trxh-1 PE=2 SV=1	10,82

68	>tr|Q1KK73|Q1KK73_MEDSA Cysteine protease OS=Medicago sativa OX=3879 GN=CP1 PE=2 SV=1	10,73
69	>tr|A0A396HJV1|A0A396HJV1_MEDTR Putative leucine-rich repeat domain, L domain-containing protein OS=Medicago truncatula OX=3880 GN=MtrunA17_Chr6g0485141 PE=4 SV=1	10,69
70	>tr|B7FHT3|B7FHT3_MEDTR Uncharacterized protein OS=Medicago truncatula OX=3880 PE=2 SV=1	10,21
71	>tr|Q43791|Q43791_MEDSA Peroxidase OS=Medicago sativa OX=3879 GN=prx1C PE=2 SV=1	10,16
72	>tr|B7FII6|B7FII6_MEDTR Uncharacterized protein (Fragment) OS=Medicago truncatula OX=3880 PE=2 SV=1	10,15

73	>tr|A0A072VNL8|A0A072VNL8_MEDTR Chitinase / Hevein / PR-4 / Wheatwin2 OS=Medicago truncatula OX=3880 GN=25484636 PE=4 SV=1	10,07
74	>tr|A0A072UXA0|A0A072UXA0_MEDTR Peroxidase OS=Medicago truncatula OX=3880 GN=25491018 PE=3 SV=1	10,00
75	>tr|B7FGU7|B7FGU7_MEDTR Cytochrome b6-f complex iron-sulfur subunit OS=Medicago truncatula OX=3880 GN=11427739 PE=2 SV=1	9,73
76	>tr|I3SSE3|I3SSE3_MEDTR ML domain-containing protein OS=Medicago truncatula OX=3880 PE=2 SV=1	9,62
77	>tr|I3SZS6|I3SZS6_MEDTR Uncharacterized protein OS=Medicago truncatula OX=3880 PE=2 SV=1	9,59
78	>tr|D2Y175|D2Y175_MEDSA Harvest-induced protein OS=Medicago sativa OX=3879 PE=2 SV=1	9,34
79	>tr|A0A072TNE0|A0A072TNE0_MEDTR GDSL-like lipase/acylhydrolase OS=Medicago truncatula OX=3880 GN=25500921 PE=3 SV=1	9,33
80	>tr|I3T337|I3T337_MEDTR Uncharacterized protein OS=Medicago truncatula OX=3880 PE=2 SV=1	9,29
81	>tr|A0A072V4S0|A0A072V4S0_MEDTR GDSL-like lipase/acylhydrolase OS=Medicago truncatula OX=3880 GN=25486026 PE=3 SV=1	8,99
82	>tr|G7LGQ4|G7LGQ4_MEDTR Putative Acid phosphatase OS=Medicago truncatula OX=3880 GN=11431124 PE=2 SV=1	8,97
83	>tr|A0A396GKD7|A0A396GKD7_MEDTR Putative chitinase OS=Medicago truncatula OX=3880 GN=MtrunA17_Chr8g0350451 PE=4 SV=1	8,86
84	>tr|G7KKP1|G7KKP1_MEDTR Kunitz type trypsin inhibitor / miraculin OS=Medicago truncatula OX=3880 GN=11420195 PE=2 SV=1	8,80
85	>tr|A0A396HWQ4|A0A396HWQ4_MEDTR Fructose-bisphosphate aldolase, cytoplasmic isozyme 2 OS=Medicago truncatula OX=3880 GN=MtrunA17_Chr5g0429021 PE=4 SV=1	8,74
86	>tr|A0A167SVS6|A0A167SVS6_9FABA Putative ubiquitin-60S ribosomal L40 (Fragment) OS=Medicago ruthenica OX=70973 PE=2 SV=1	8,66
87	>tr|A0A072U1H6|A0A072U1H6_MEDTR Alpha-galactosidase OS=Medicago truncatula OX=3880 GN=25498854 PE=3 SV=1	8,32
88	>tr|G7ILY5|G7ILY5_MEDTR 1-O-acylglucose:anthocyanin acyltransferase OS=Medicago truncatula OX=3880 GN=11438858 PE=3 SV=1	7,89
89	>tr|A0A072UT49|A0A072UT49_MEDTR NAD(P)-binding rossmann-fold protein OS=Medicago truncatula OX=3880 GN=25491591 PE=4 SV=1	7,66

Appendixes

90	>tr|I3T683|I3T683_MEDTR Thioredoxin domain-containing protein OS=Medicago truncatula OX=3880 PE=2 SV=1	7,63
91	>tr|A0A396HM61|A0A396HM61_MEDTR Putative cucumisin OS=Medicago truncatula OX=3880 GN=MtrunA17_Chr5g0398331 PE=3 SV=1	7,57
92	>tr|G7LEC0|G7LEC0_MEDTR Lectin kinase family protein OS=Medicago truncatula OX=3880 GN=11430405 PE=4 SV=2	7,54
93	>tr|G7L7L3|G7L7L3_MEDTR Putative tripeptidyl-peptidase II OS=Medicago truncatula OX=3880 GN=11434053 PE=3 SV=1	7,53
94	>tr|A0A072UMJ2|A0A072UMJ2_MEDTR Multi-copper oxidase-like protein OS=Medicago truncatula OX=3880 GN=25492860 PE=4 SV=1	7,41
95	>tr|I3S0L8|I3S0L8_MEDTR PKS_ER domain-containing protein OS=Medicago truncatula OX=3880 PE=2 SV=1	7,01
96	>tr|I3S399|I3S399_MEDTR Putative SOUL heme-binding protein OS=Medicago truncatula OX=3880 GN=25502484 PE=2 SV=1	5,78
97	>tr|I3SKM9|I3SKM9_MEDTR Germin-like protein OS=Medicago truncatula OX=3880	5,77

	PE=2 SV=1	
98	>tr|I3T8N2|I3T8N2_MEDTR Thioredoxin domain-containing protein OS=Medicago truncatula OX=3880 PE=2 SV=1	5,73
99	>tr|G7JAV5|G7JAV5_MEDTR Alginate lyase OS=Medicago truncatula OX=3880 GN=11419455 PE=2 SV=1	5,67
100	>tr|A0A072UDD5|A0A072UDD5_MEDTR Pathogenesis-related thaumatin family protein OS=Medicago truncatula OX=3880 GN=25494521 PE=4 SV=1	5,54
101	>sp|P16346|IBBWT_MEDSA Bowman-Birk type wound-induced trypsin inhibitor OS=Medicago sativa OX=3879 PE=1 SV=1	5,53
102	>tr|A0A396IHC0|A0A396IHC0_MEDTR Putative lactoylglutathione lyase OS=Medicago truncatula OX=3880 GN=MtrunA17_Chr4g0075021 PE=4 SV=1	5,45
103	>tr|Q40366|Q40366_MEDSA Peroxidase OS=Medicago sativa OX=3879 GN=pxdC PE=2 SV=1	5,14
104	>tr|I3SW13|I3SW13_MEDTR Peptidylprolyl isomerase OS=Medicago truncatula OX=3880 PE=2 SV=1	5,13
105	>tr|A0A072URQ0|A0A072URQ0_MEDTR Glycerophosphoryl diester phosphodiesterase family protein OS=Medicago truncatula OX=3880 GN=25493892 PE=4 SV=1	4,53
106	>tr|A0A396I3M7|A0A396I3M7_MEDTR Non-specific lipid-transfer protein OS=Medicago truncatula OX=3880 GN=MtrunA17_Chr4g0012991 PE=3 SV=1	4,51
107	>tr|A0A396IET8|A0A396IET8_MEDTR Putative glycoside hydrolase superfamily OS=Medicago truncatula OX=3880 GN=MtrunA17_Chr4g0065461 PE=4 SV=1	4,42
108	>tr|G7J3A3|G7J3A3_MEDTR Alpha-mannosidase OS=Medicago truncatula OX=3880 GN=11421286 PE=3 SV=1	4,41
109	>tr|I3SQJ7|I3SQJ7_MEDTR Uncharacterized protein OS=Medicago truncatula OX=3880 PE=2 SV=1	4,39
110	>tr|A0A072V9U1|A0A072V9U1_MEDTR Receptor-like kinase OS=Medicago truncatula OX=3880 GN=11411817 PE=4 SV=1	4,37
111	>tr|A0A396IUP0|A0A396IUP0_MEDTR Putative gamma-glutamyltransferase, Glutathione hydrolase OS=Medicago truncatula OX=3880 GN=MtrunA17_Chr3g0124611 PE=4 SV=1	4,36

Appendixes

112	>tr|A0A072UGX9|A0A072UGX9_MEDTR Phosphatidylethanolamine-binding protein OS=Medicago truncatula OX=3880 GN=25491475 PE=4 SV=1	4,33
113	>tr|I3T4G2|I3T4G2_MEDTR Biotin carboxyl carrier protein of acetyl-CoA carboxylase OS=Medicago truncatula OX=3880 PE=2 SV=1	4,32
114	>tr|A0A072UTV7|A0A072UTV7_MEDTR Carboxypeptidase OS=Medicago truncatula OX=3880 GN=25490374 PE=3 SV=1	3,99
115	>tr|G7K6M0|G7K6M0_MEDTR MAP3K-like kinase OS=Medicago truncatula OX=3880 GN=11405797 PE=4 SV=1	3,99
116	>tr|I3T7C9|I3T7C9_MEDTR SERPIN domain-containing protein OS=Medicago truncatula OX=3880 PE=2 SV=1	3,94
117	>tr|G7J0U5|G7J0U5_MEDTR Alpha-N-acetylglucosaminidase family protein OS=Medicago truncatula OX=3880 GN=11410414 PE=4 SV=2	3,89
118	>tr|I3S4Q3|I3S4Q3_MEDTR Uncharacterized protein OS=Medicago truncatula OX=3880 PE=2 SV=1	3,89
119	>tr|G7KVQ4|G7KVQ4_MEDTR Polygalacturonase-inhibiting protein, putative OS=Medicago truncatula OX=3880 GN=11443840 PE=4 SV=1	3,83
120	>tr|A0A396JMH1|A0A396JMH1_MEDTR Putative chitinase OS=Medicago truncatula OX=3880 GN=MtrunA17_Chr2g0329761 PE=4 SV=1	3,78
121	>tr|B7FM03|B7FM03_MEDTR Uncharacterized protein OS=Medicago truncatula	3,72
	OX=3880 PE=2 SV=1	
122	>tr|A2Q3E5|A2Q3E5_MEDTR Virulence factor, pectin lyase fold OS=Medicago truncatula OX=3880 GN=MtrDRAFT_AC155881g2v1 PE=3 SV=1	3,67
123	>tr|A0A072UQ15|A0A072UQ15_MEDTR Peptidylprolyl isomerase OS=Medicago truncatula OX=3880 GN=25493451 PE=4 SV=1	3,66
124	>tr|G7IXH3|G7IXH3_MEDTR Legume lectin beta domain protein OS=Medicago truncatula OX=3880 GN=11426365 PE=3 SV=2	3,65
125	>tr|G7JCT4|G7JCT4_MEDTR Putative tripeptidyl-peptidase II OS=Medicago truncatula OX=3880 GN=11410429 PE=3 SV=1	3,46
126	>tr|A0A396IB37|A0A396IB37_MEDTR Putative tetrahydroberberine oxidase OS=Medicago truncatula OX=3880 GN=MtrunA17_Chr4g0051531 PE=3 SV=1	3,44
127	>tr|A0A072TU31|A0A072TU31_MEDTR Pathogenesis-related thaumatin family protein OS=Medicago truncatula OX=3880 GN=25502023 PE=4 SV=1	3,27
128	>tr|A0A1L2BU61|A0A1L2BU61_MEDSV Class III-1 chitinase 11 (Fragment) OS=Medicago sativa subsp. varia OX=36902 GN=chit3-1-11 PE=3 SV=1	3,24
129	>tr|G7J6V7|G7J6V7_MEDTR Pectinesterase OS=Medicago truncatula OX=3880 GN=11441469 PE=3 SV=1	3,22