University of Trás-os-Montes and Alto Douro

Nutriproteogenomic approach for detection and characterization of almond allergens

Master in Comparative and Technological Molecular Genetics

Mário Joaquim Oliveira Bezerra

Professor Dr. Gilberto Igrejas

Dr. Miguel Ribeiro

Vila Real, 2021

University of Trás-os-Montes and Alto Douro

Nutriproteogenomic approach for detection and characterization of almond allergens

Master in Cycle in Comparative and Technological Molecular Genetics

Mário Joaquim Oliveira Bezerra

Professor Dr. Gilberto Igrejas

Dr. Miguel Ribeiro

Composition of the jury

_____

_____

_____

_____

Vila Real, 2021

"If you want to be successful you need total dedication, seek your last limit and give the best of yourself"

Ayrton Senna

Assinado por: **GILBERTO PAULO PEIXOTO IGREJAS**
Num. de Identificação: 09594163
Data: 2021.12.26 10:57:49+00'00'

Scientific Supervision

Professor Dr. Gilberto Paulo Peixoto Igrejas
Associate Professor with Affiliation (DGB/UTAD)

Assinado por : **JOSÉ MIGUEL SILVA FERREIRA RIBEIRO**
Num. de Identificação: 13018371
Data: 2021.12.22 14:36:56+00'00'

Scientific Co-Supervision

Dr. Miguel Ribeiro
Researcher (DGB/UTAD)

Acknowledgments

Com a finalização da minha dissertação de mestrado deixo umas palavras de gratidão a todas as pessoas que me auxiliaram neste percurso quer a nível pessoal quer profissional.

À direção do Mestrado de Genética Molecular Comparativa e Tecnológica, e a todos os Professores que lecionaram este mestrado e contribuíram para o meu desenvolvimento profissional.

Ao meu orientador Professor Doutor Gilberto Paulo Peixoto Igrejas, por me acolher tão bem no seu grupo de investigação e pelos ensinamentos, pareceres e empenho em proporcionar-me as melhores condições de trabalho e de desenvolvimento pessoal e profissional.

Ao Doutor Miguel Ribeiro, pela coorientação, mas acima de tudo, um muito obrigado por tão bem ter desempenhado esse papel sendo uma ajuda e um amigo em todo o processo de realização deste trabalho e ter sempre em mente o maior sucesso não só do trabalho como também pessoal.

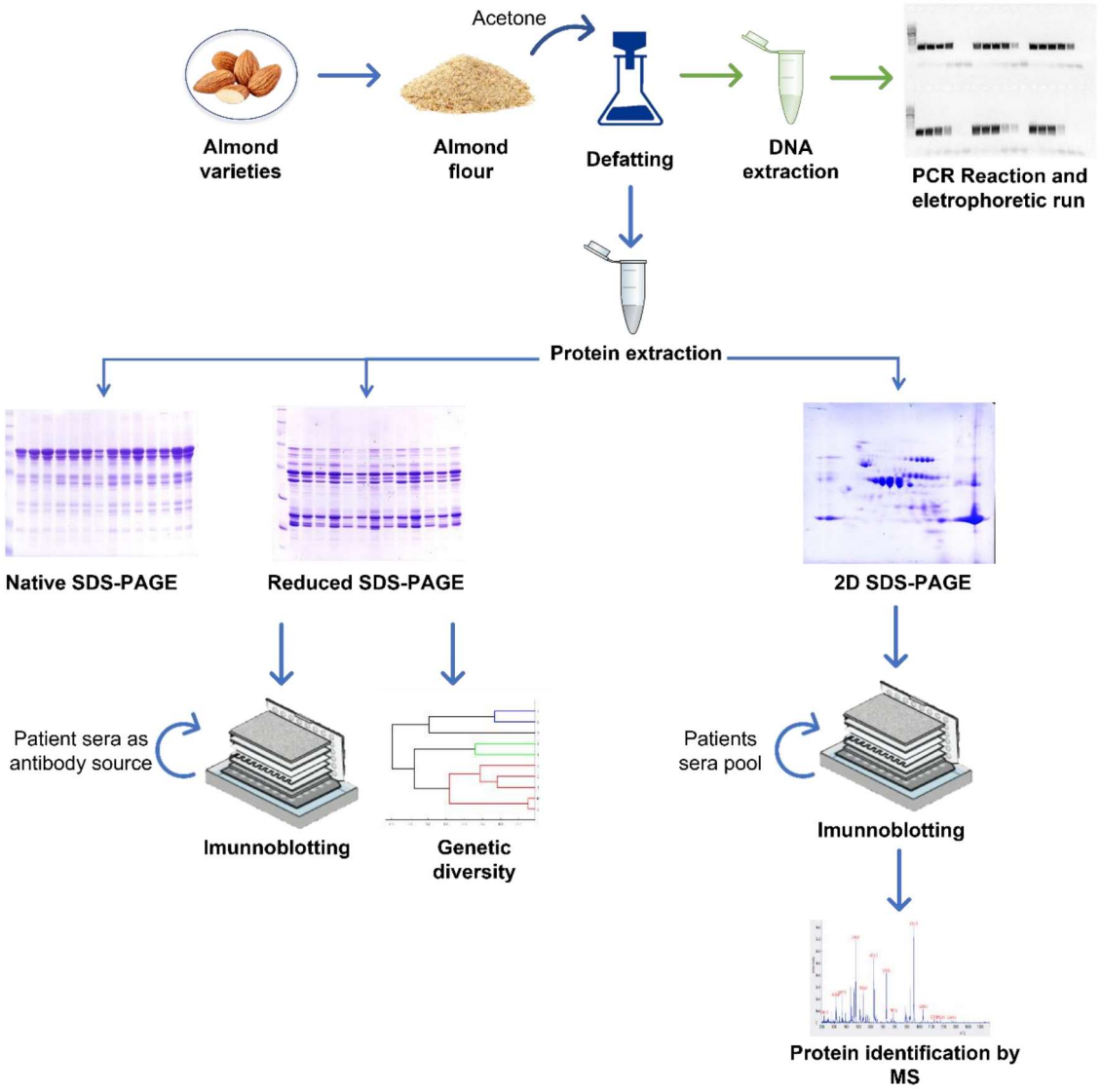
Ao Laboratório de Genómica e Proteómica Funcional e grupos associados, sem exceção, muito obrigado pela forma que me receberam fazendo-me sentir parte integrante dos mesmos, tendo encontrado não só colegas de trabalho, mas também amigos.

Em especial, um obrigado ao REQUIMTE, Laboratório de Bromatologia e Hidrologia, Faculdade de Farmácia da Universidade do Porto, nomeadamente à Professora Doutora Isabel Mafra e à Doutora Joana Costa, pela forma como me acolheram e pela sua contribuição na realização deste trabalho.

A todos os meus amigos, feitos durante estes últimos cinco anos ou já de longa data, um muito obrigado por sempre terem acreditado em mim e estarem presentes sempre algo não corria tão bem.

À Telma, por não só ter sempre acreditado em mim, mas também por me ter feito acreditar em mim próprio e por sempre ter estado ao meu lado em todas adversidades bem como momentos de felicidade. Um enorme obrigado!

Graphical Abstract



Resumo

Os frutos de casca rija são considerados um alimento importante numa dieta saudável. No entanto são uma das fontes mais comuns de alergénios alimentares responsáveis por reações alérgicas agudas que, em alguns casos, podem tornar-se mortais. Este grupo alimentar, pertence ao chamado "*Big Eight Foods"* que são responsáveis por mais de 90% dos casos de alergia alimentar nos Estados Unidos e neste grupo, as alergias à amêndoa e ao amendoim são persistentes e normalmente problemáticas. A amêndoa é geralmente consumida crua, tostada ou como parte de outros alimentos. O seu consumo está geralmente associado à redução do aparecimento de doenças cardiovasculares. Várias proteínas da amêndoa foram reconhecidas como alergénios. Seis deles, Pru du 3, Pru du 4, Pru du 5, Pru du 6, Pru du 8 e Pru du 10, foram incluídos na lista de alergénios da WHO-IUIS. Por outro lado, muitas proteínas de amêndoa já foram descritas como potenciais alergénios, embora apenas uma parte delas tenha sido reconhecida como alergénica e a autenticidade de algumas designações tenha sido questionada principalmente devido identificações erradas sendo, por isso necessário o desenvolvimento de linhas de investigação sobre a caracterização dos mesmos. Neste trabalho, o perfil proteico de 27 variedades de amêndoa, de quatro regiões diferentes da região de Trás-os-Montes, foi obtido através de eletroforese monodimensional. A ausência/presença de proteínas específicas foi utilizada para estabelecer relações de amostra através de dendrogramas. Além disso, o potencial alergénico destas variedades foi avaliado utilizando soro de 10 pacientes alérgicos como fonte de anticorpos IgE. Os resultados mostraram que diferentes variedades têm pequenas diferenças em relação ao seu perfil eletroforético, mas diferentes propriedades de ligação igE. Em particular, o soro de 5 pacientes apresentou um potencial alergénico com diferenças inter-varietais. A partir desses perfis surgiram resultados interessantes, como a identificação de possíveis reações aos putativos Pru du 8 e Pru du 10, apenas recentemente reconhecidos como alergénios da amêndoa e reação à putativa Pru du 1, uma proteína ainda não reconhecida como alergénio. A eletroforese bidimensional e o *imunoblot* adicional usando um conjunto dos soros de todos os indivíduos alérgicos usados neste estudo mostraram o complexo alergoma da amêndoa, com diferentes subunidades proteicas que possuem propriedades alergénicas. Finalmente, foi realizada uma breve caracterização genómica e Pru du 3 e Pru du 6 foram identificados em todas as variedades.

Palavras-chave: alergia alimentar, amêndoa, alergénios, nutrição.

Abstract

Tree nuts are widely considered an important food in a healthy diet. However, for a part of the world population, they are one of the most common sources of food allergens causing acute allergic reactions that, in some cases, can become life-threatening. They are part of the Big Eighth food groups which are responsible for more than 90% of food allergy cases in the United States and within this group, almond and peanut allergies are persistent and normally severe and life-threatening. Almond is generally consumed raw, toasted or as an integral part of other foods. Its dietary consumption is generally associated with a reduced risk of cardiovascular diseases. Several almond proteins have been recognized as allergens. Six of them, namely Pru du 3, Pru du 4, Pru du 5, Pru du 6, Pru du 8 and Pru du 10, have been included in WHO-IUIS list of allergens. On the other hand, a lot of almond proteins have been already described as potential allergens although only a part of them have been recognized as allergenic and the authenticity of some designations have been questioned mainly due to misidentification problems. Therefore, research about the accurate characterization of these allergens/proteins is needed. In this work, the protein profile of 27 almond varieties, from four different regions of the Trás-os-Montes's region, was obtained using 1D-eletrophoresis. The absence/presence of specific proteins was used to establish sample relationships trough dendrograms. Furthermore, the allergenic potential of these varieties was assessed using sera from 10 allergic patients as a source of IgE-antibodies. Results showed that different varieties have small differences regarding their electrophoretic profile, but different IgE-binding properties. In particular, the sera of 5 patients showed significant inter-varietal allergenic potential differences. Also, it was verified the IgE binding to the putatives Pru du 8 and Pru du 10, which were only recently recognized as almond allergens and most important to the putative Pru du 1, a protein not yet recognized as an almond allergen. Two-dimensional electrophoresis and further immunoblotting using a pooled serum from all allergic individuals enrolled in the present study showed a complex almond allergome, with different protein subunits possessing allergenic properties. Finally, a brief genomic characterization was conducted and Pru du 3 and Pru du 6 were identified across all varieties.

Keywords: food allergy, almond, almond allergens, nutrition

Scientific production related to this master's thesis

Article

Bezerra, M.; Ribeiro, M.; Igrejas, G. **An Updated Overview of Almond Allergens.** Nutrients 2021, 13, 2578. https://doi.org/10.3390/nu13082578

MDPI Encyclopedia entry:

Bezerra, M.; Ribeiro, M.; Igrejas, G. **Almond Allergens,** Entry ID: 15870**,** https://encyclopedia.pub/15870

Poster

Bezerra, M.; Ribeiro, M.; Igrejas, G. Protein profiling and allergenic potential of different almond varieties. In Book of Abstracts of XIII Conference on Genetics and Biotechnology and III Iberian Conference on Genetics and Biotechnology- University of Trás-os-Montes and Alto Douro, March 14-16, 2021. Awarded with the **Prize for Best Poster in Human Genetics and Biotechnology** category.

Abbreviations

ANOVA, Analysis of variance

BAT, Basophil activation test

BSA, Bovine serum albumin

BP, base pairs

DNA, desoxyribonucleic acid

DTT, Dithiothreitol

ELISA, Enzyme-linked immunosorbent assay

ESI, Electrospray ionization

IEF, Isoelectric focusing

IGE, Immunoglobulin E

IPG, Immobilized pH gradient

kDa, kilo Dalton

LC, Liquid chromatography

MM, Molecular marker

MS, Mass spectrometry

MS/MS, tandem mass spectrometry

MW, Molecular weight

PCR, Polymerase chain reaction

SD, Standard deviation

SDS, Sodium dodecyl sulphate

SDS-PAGE, Sodium dodecyl sulfate-polyacrylamide gel electrophoresis

TBS, triphosphate buffered saline

UPGMA, unweighted pair group method with arithmetic average

UV, Ultraviolet

1-DE, One-dimensional electrophoresis

2-DE, Two-dimensional electrophoresis

Index

Figure index

Table index

Chapter 1: Introduction

Food allergies are a concerning health issue affecting the worldwide population and their prevalence have been increasing for the last couple of decades (Gupta et al. 2011; Gupta et al. 2019; Kumfer and Commins 2019). For example, in the United States, around twenty-six million adults (Sicherer and Sampson 2018) and six million children (Gupta et al. 2011) suffer from this condition. Although there is no cure to food allergies and food avoidance is considered the best strategy, vast research has been made in this area and potential therapies can be generally divided into two categories: allergen non-specific like the use of monoclonal antibodies and allergen specific where the treatment is performed using recombined or native food antigens (Moore et al. 2017). Among those, immunotherapy, a food allergen-specific therapy, which refers to the administration of gradual and increasing doses of an antigen over a certain time (Kulis et al. 2011; Albin and Nowak-Węgrzyn 2015) is considered as a solid option since that, in the majority of cases, the side effects are mild, like itching and if successful, immunotherapy can induce desensitization and less commonly sustained unresponsiveness, also known as tolerance (Moore et al. 2017). However, less commonly, adverse side effects can range from mild to anaphylaxis or eosinophilic esophagitis (Sánchez-García et al. 2012) and due to their unpredictable character (Varshney et al. 2009), new and innovating therapies must be pursued. Moore et al. (2017) believes that the tolerance induced by immunotherapy with or without the administration of monoclonal antibodies could significantly shift the allergic diseases field.

To scientific research go further, food allergy, allergic diseases and allergens must be firstly identified and characterized. For allergens, when new ones from specific species are identified, a distinctive name is given by the WHO/IUIS Allergen Nomenclature Sub-Committee alongside the additional information about it. A vast number of allergens from more than one hundred and sixty species have been identified and most of them belong to a restricted number of protein families. In almond, several proteins of these protein families have been already identified as allergens namely, Pru du 6 (11S globulin legumin-like protein), Pru du 4 (profilin) and Pru du 3 (nonspecific LTP) and are several other proteins, belonging to other protein families and/or that do not have a name attributed by the Allergen Nomenclature Sub-Committee.

1.1. Food allergy

By definition, food allergy is “an adverse food health effect arising from a specific immune response that occurs reproducibly on exposure to a given food” (Boyce et al. 2011). It is also important to clarify that the immune reaction is key, otherwise food allergy could probably be described as food intolerance which is a non-immune response but may reproduce food allergy clinical symptoms (Green et al. 2015).

Nowadays, the worldwide prevalence of atopic diseases, disorders characterized for a pathological immune response to normally harmless antigens like asthma, allergic rhinitis and food allergy (Yu et al. 2016) are increasing and seen as a public health problem (Renz et al. 2018). Particularly, food allergy is one of the most common causes of anaphylaxis that can lead to fatality cases (Moore et al. 2017) and has been given a lot of interest by the scientific community to develop new strategies and therapies to prevent it since the usual approach, avoidance, is hard to achieve in most cases (Nowak-Węgrzyn and Sampson 2011).

Evidences that show global variation of food allergy as well as changes in its prevalence associated with migration (Prescott and Allen 2011) are increasing the interest on the epidemiological stand of food allergy and may promote hypothesis for why is food allergy a rising issue in some parts of the world and not on others (Panjari et al. 2016). Some authors proposed various hypotheses to the increasing prevalence of food allergy in association with geographical sites, being the most accepted ones hygiene increase which lead to less pathogen exposure, changes in the human microbiome, avoidance of certain allergens in the early stages of life causing allergen exposure reduction, obesity, diets lacking antioxidants and vitamin D deficiency (du Toit et al. 2016; Sicherer et al. 2017).

Tree nuts are one of the Big Eight food groups among peanut, milk, shellfish, soy, wheat, egg and fish which are responsible for more than 90% of food allergy cases in the United States (Zhang and Jin 2020) and, in particular, the number of people sensitized to tree nuts and peanuts have been concerningly growing in Europe and United States (Alasalvar and Shahidi 2008). In this group of foods, almond and peanut allergies are persistent and normally severe and life-threatening in opposition to allergies caused by milk or eggs which are normally mild and transient (Yunginger et al. 1988; Sampson et al. 1992; Panel 2010).

Tree nuts allergy prevalence data is very limited and even more for a specific nut species like almond (Mandalari and Mackie 2018). Although being known that tree nut allergy rates vary according to geographical regions, ethnic differences and dietary habits (Luyt et al. 2016), the real prevalence is not actually known (Buchanan et al. 2008) and is very difficult to present coherent data and perform geographical comparisons since country specific studies are not based on the same parameters like age or diagnosis method (Nollet and Van Hengel 2016).

Even so, some studies were performed and are useful to enlighten the general distribution of tree nut allergy around the world (Figure 1). For example, in the US general population is estimated to be 0.35% with that prevalence being higher in adults (0.5%) than in children (0.2%) (McWilliam et al. 2015); Canada's self-reported prevalence in the general population corresponds to 1.2% (Ben-Shoshan et al. 2012); 1.7% in the UK (Young et al. 1994); 0.1% in children from Turkey (Orhan et al. 2009); 1.3% from the Netherlands (Brugman et al. 1998) and 6.9% from Spain (Martínez-Gimeno et al. 2000).

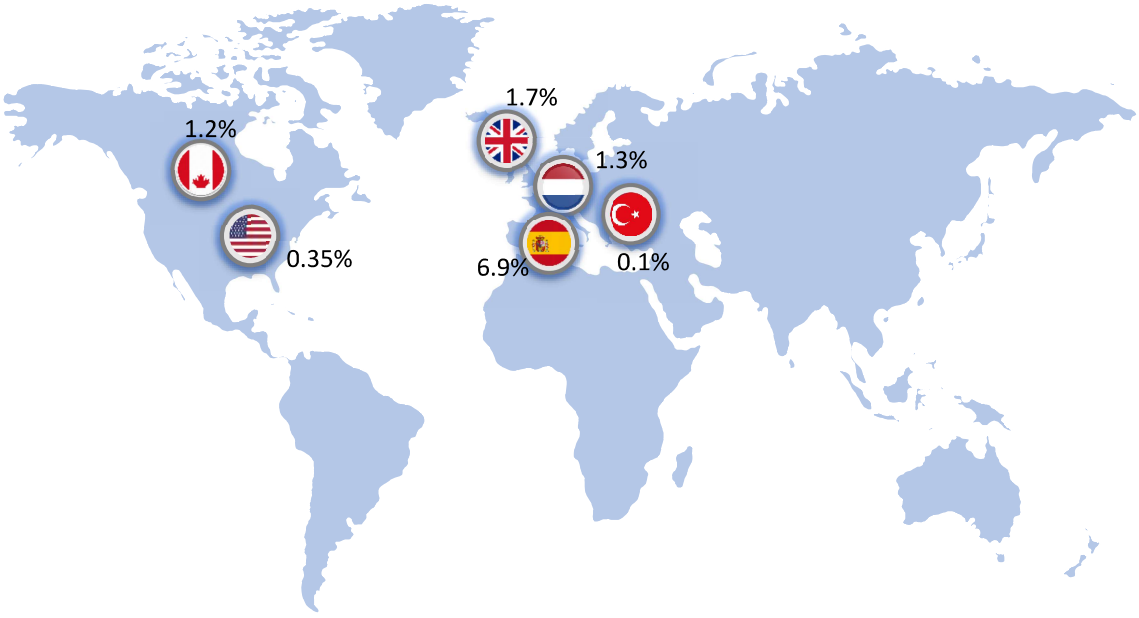


Figure 1 Tree nuts allergy prevalence data from the general population in the United States (0,35%), Canada (1,2%), United Kingdom (1,7%) and from childreen in Turkey (0,1%), Netherlands (1,3%), and Spain (6,9%). Adapted from (Mandalari and Mackie 2018).

All these studies show that prevalence data is very limited and due to the cross-reactivity between different species in the tree nut group, almond allergy prevalence in specific is very difficult to calculate (Mandalari and Mackie 2018).

1.1.1. Molecular pathway of Immunoglobulin E-mediated food reaction

Food allergy can occur by several immunological mechanisms that lead to a reaction to food allergens. The most common mechanism of food allergy expression is a hypersensitivity manifestation where specific Immunoglobulin E (IgE) antibodies interact with mast cells and basophils leading to a rapid physiological response (Renz et al. 2018). Normally, food allergy symptoms, after a first exposure and consequent IgE production, appear near immediate or a few minutes later of food ingestion but, in more rare cases, it could take several hours for the symptoms to manifest (Burton et al. 2013).

In people suffering from food allergy disorders, the absorption process of allergens in the intestinal epithelium and consequent access to bloodstream and mucosa is increased (Perrier and Corthesy 2011). When food allergens are ingested occurs an interaction between them and IgE and its high-affinity fragment crystallizable receptor (FCER1) on basophils in circulation or mast cells present in mucosal tissues leading to their activation (Figure 2). FCER1 crosslinking leads to a signaling cascade where tyrosine protein kinase SYK will promote exocytose of granules containing mediators of hypersensitivity like histamine, chymase and tryptase (Renz et al. 2018). This process together with synthesis of lipid metabolites like prostaglandins, leukotrienes and platelet-activating factor (PAF) (Vadas et al. 2008) will result in physiological responses such as the activation of nociceptive nerves that promote itch and soft muscle constriction, vasodilation, higher vascular permeability and in the most severe cases anaphylaxis (Williams and Sharma 2015).

Although this is the generic mechanism after food ingestion, non-IgE mediated reactions (Mishra et al. 2007; Clayton et al. 2014) like the inflammatory process subjacent to eosinophilic esophagitis (Mishra et al. 2007), can also occur. The physiological response is dependent of the kind of mediators released by mast cells and basophils but is also dependent of tissue location where those mediators would act. This two factors combined will directly influence the physiological response. (Renz et al. 2018).

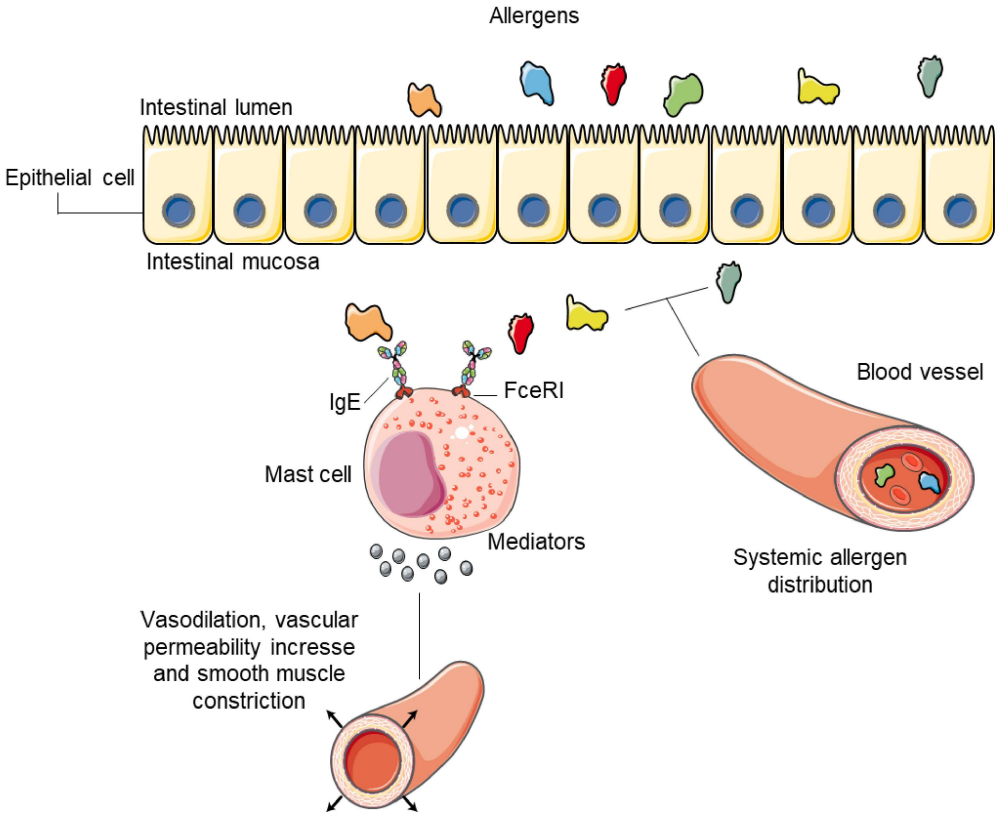


Figure 2 General mechanisms of IgE mediated response to food allergens. Interaction between food allergens and IgE and its high-affinity FC receptor (FCER1) on basophils in circulation or mast cells present in mucosal tissues leading to their activation and consequent physiological response (Lee 2016).

1.1.2. Legal framework

There are several regulatory frameworks for food allergen labeling according to countries or regions that differ significantly around the world due to the priority level that each jurisdiction applies to specific allergens. The criteria for the development of the allergen's priority list and the standards for the addition or removal of allergens from the regulations differ and they are often unclear (Gendel 2012).

The Regulation (EU) No. 116/2011 sets the regulation on food labelling forbidding misleading consumers and any claims that a certain food, like almonds, can prevent, treat or cure human diseases cannot be made. Moreover, nutritional and allergens information must be highlighted in the list of ingredients and included in non-packed foods or any product where they are used as ingredient with the punishment of being withdrawn from the market.

Regulation (EU) No 1169/2011 of the European Parliament and of the Council of 25 October 2011 on the provision of food information to consumers states that shall be indicated in the list of ingredients with a clear reference to the name of the substance or

product causing allergies or intolerances and shall be emphasized through a typeset that clearly distinguishes it from the rest of the list of ingredients, for example by means of the font, style or background color. In this list of substance or product causing allergies or intolerances are included nuts with a clear reference to almonds, hazelnuts, walnuts and others, cereals containing gluten, crustaceans, eggs, fish, peanuts, soybeans, milk, celery, mustard, sesame, lupin, mollusks.

In the United States, the food labelling requirements are quite similar to the ones applied in the European Union where the Food Allergen Labeling and Consumer Protection Act of 2004 states that in any food source containing a major food allergen, or protein derived from them, should be printed right next to the ingredient list and specifically have the word "contains" before it. The term "major food allergen" refers to milk, egg fish, crustacean shellfish, tree nuts (like almonds, pecans or walnuts), wheat, peanuts and soybeans however, any highly refined oil derived from any of the previous foods and products derived from those oils are considered exceptions.

For the appliance of the food labelling requirements is important to defined threshold values which correspond to the minimal concentration of a specific food allergen in a food able to trigger any reaction in a sensitized individual. Though, is very difficult to establish a threshold since they vary according to the individual/population, the allergens itself and the consequent food processing (Derr 2006). In order to get there, wide population tests and data are needed. For almond, currently, no thresholds are established (Mandalari and Mackie 2018) which shows a clear sign that further investigations and regulations are imperative.

1.2. Almond (Prunus dulcis)

One of the most important foods in human nutrition are tree nuts, namely due to their excellence in terms of taste as well as their versatility to be used combined with other foods and more recently by their potential health benefits. All these characteristics make tree nuts consumed all around the world in the most various forms according to the availability in the region and the populational habits (Costa et al. 2012; Rehm and Drewnowski 2017).

Almond (*Prunus dulcis* Mill.) is a member of the Rosaceae family and is considered a native plant from Minor Asia (Bottone et al. 2018) being one of the oldest nut trees cultivated worldwide with special relevance in the Mediterranean warm-arid countries

(Nanos et al. 2002; Piscopo et al. 2010), namely the Apulia region on southern Italy (De Giorgio et al. 2007). Among tree nuts, almond presents as one of the most important nuts which is very noticeable in tree nut production data (Table 1) around the world (Walnut 3663, Almond 3183 and Hazelnut 864 ktons/year; (FAOSTAT 2020)). Furthermore, its nutritional properties should be highlighted; high levels of mono and polyunsaturated fatty acids, phytosterols and low glycemic index are associated with reduction of some risk factors for cardiovascular disease and diabetes (Griel and Kris-Etherton 2006; Jenkins et al. 2008; Richardson et al. 2009; Mandalari et al. 2010). It has also been described antioxidant and inflammatory activities due to its polyphenol content, including flavonoids, hepato and neuroprotective potential and perhaps, the most known, cholesterol-lowering properties (Berryman et al. 2011; Mandalari et al. 2011; Kozlowska and Szostak-Wegierek 2014; Seo et al. 2015). Also, almond derived products like their oils have demonstrated both antibacterial and antifungal capabilities (Tian et al. 2011) which make almond a product of great interest both to the consumer and producer.

Table 1 Production (ktons) of almond, hazelnut and walnut in Portugal, all the European Union and in the whole world between 2016 and 2018 (FAOSTAT).

	Production (kton/year)		
	2016	2017	2018
Portugal			
Almond	8,70	23,14	14,30
Walnut	4,32	4,59	4,58
Hazelnut	0,32	0,31	0,24
European Union			
Almond	323,52	403,46	474,38
Walnut	176,48	186,78	206,03
Hazelnut	151,40	160,84	166,98
World			
Almond	2 493,30	2 687,81	3 182,90
Walnut	3 447,74	3 510,24	3 662,51
Hazelnut	739,95	996,72	863,89

Production data shows that Portuguese productors have a preference to produce almond in opposition to hazelnut and walnut. This scenario is the same for the European

Union but different when compared to the production worldwide; walnut production slightly exceeds that of almond. Nevertheless, in 2018, global almond production had an increase of about 18% from 2017, which is noteworthy as hazelnut and walnut production maintains steady since 2016. In Europe, that boost was even more noticeable, but occurred a year sooner, between 2016 and 2017 where the production grew up about 24%. In particular, Portugal saw its production grow by 65%, in the same period of time, to decrease in 2018, which shows an irregular production.

With the almond nutritional value comes the agronomical properties of different cultivars. For example, Bolling et al. (2010) described that the individual polyphenols synthesis was only due to the cultivar itself however, total polyphenols and antioxidant activity were significantly dependent on both genotype and environmental growing conditions. Pursuing this point of view, Summo et al. (2018) performed a study aiming to determine if either the cultivar or harvest time influence the chemical composition of the fruit. From that, the team concluded that, in fact, harvest time and genotype both have a strong influence on the fruit nutritional value.

1.2.1. Almond allergy

Nuts allergy is associated with clinical symptoms that can range in severity from mild to life-threatening and in this sense when a patient is diagnosed with allergy to a certain nut is often advised to avoid the consumption of the entire group (Ewan 1996; Sicherer et al. 2003). Nevertheless, it is still unclear if the taxonomic proximity between tree nuts groups and peanuts is a key factor for the cross-reactivity between those two or it comes from the high structural homology of IgE-binding epitopes (Wallowitz et al. 2004; Maleki et al. 2011). In general, tree nuts allergy is caused by non-pollen-mediated food sensitization however, in cases like almond and hazelnut, sensitization to plane tree pollen, birch pollen or mugwort pollen may induce allergy (Vieths et al. 2002; Flinterman et al. 2006) like schematically represented in Figure 3. On other hand, tree nuts allergy cross reaction is highly related to botanical family associations which, for almond, is common to see cross-reactivity between other members of the Rosaceae family (Hasegawa et al. 2009; Noorbakhsh et al. 2011). Furthermore, within the Rosaceae family a strong source of cross-reaction lies in the structural homology between allergic lipid-transfer proteins (LTP's). Specifically, in the tree nut group, almond Pru du 3, chestnut Cas s 8, hazelnut Cor a 8 and walnut Jug r 3 are the most predisposed to show cross-

reactivity. Besides those, peach Pru p 3 holds higher IgE-binding affinity and a higher number of epitopes compared to other LTP's which results in the fact that peach is a primary sensitizer to LTP's (Egger et al. 2010) and makes it a strong cause for cross-reactivity to other plants, including nuts like almond (Asero 2011). Other study performed by Kewalramani et al. (2006) showed extensive IgE cross-reactivity between almond and apricot seeds and that may exist some cross-reactive proteins with pine nut, pecan, walnut and sunflower.

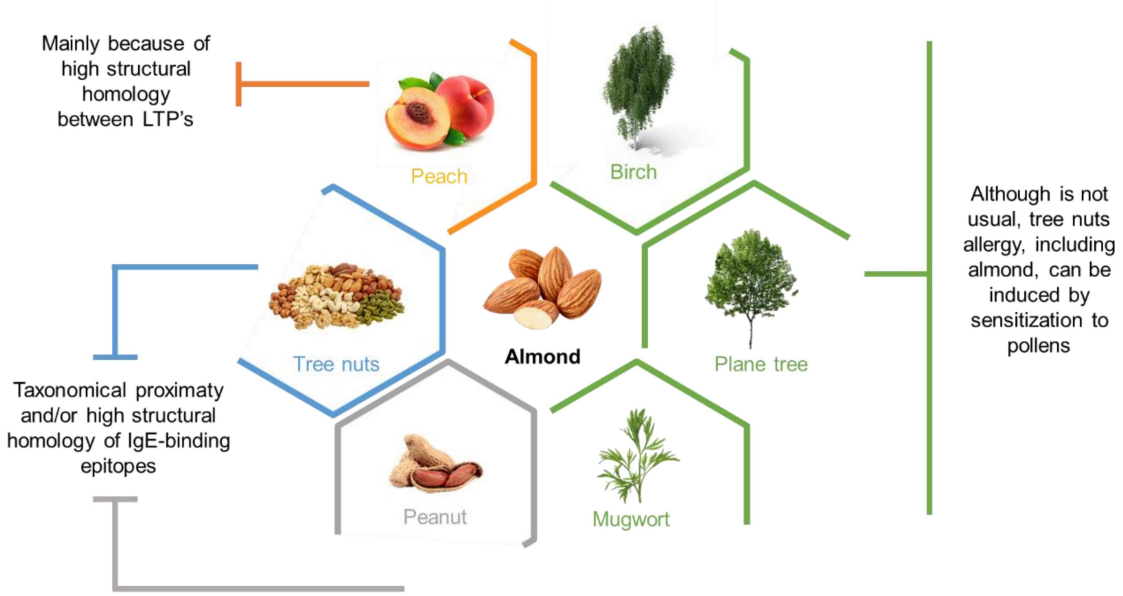


Figure 3 Most associated allergic cross-reactions with almond. In orange, allergic cross reactions between almond and peach are most commonly due to high structural homology between allergic LTP's present in the Rosasceae family that both belong; In blue and grey, is still unclear if cross reactivity between almond and other tree nuts groups and peanut is a consequence of taxonomical proximity and/or high structural homology of IgE-binding epitopes; Finally, in green are represented three different pollens which, although it is not usual, when sensitized to them, allergy to tree nuts like almond could be induced.

1.2.2. Almond allergens

To date, ten groups of almond allergens have been identified namely: Pru du 1, Pru du 2, Pru du 2S albumin, Pru du 3, Pru du 4, Pru du 5, Pru du 6 (amandin), Pru du γ-conglutin, Pru du 8 and Pru du 10. From these groups, only Pru du 1, Pru du 2, Pru du 2S albumin and Pru du γ-conglutin are not included in the WHO-IUIS list of allergens. Their corresponding biochemical name, biological function, GenBank nucleotide and UniProt annotations, molecular weight, food processing effects and clinical relevance are summarized in Table 2.

1.2.2.1. WHO/IUIS designated almond allergens

1.2.2.1.1. Pru du 6 (amandin)

Pru du 6 or amandin is the most well and widely studied almond allergen according to its biochemical function and molecular structure (Roux et al. 2001; Sathe et al. 2002; Albillos et al. 2008; Albillos et al. 2009). It was first reported as an allergen in 1999 (Kewalramani et al. 2006) but only in 2010 was recognized and added to the WHO-IUIS database.

Biochemically, amandin, also known as almond major protein (AMP), is a member of the cupin superfamily, namely the 11S seed storage globulin family (Ewan 1996; Sicherer et al. 2003). Globulins are very abundant proteins in legumes and tree nuts and in almond they correspond to roughly 65% of total almond protein content (Costa et al. 2012).

As an allergen, Pru du 6 have been associated with severe allergic reactions (Roux et al. 2001). Studies on the Pru du 6 isoforms, Pru du 6.01 and Pru du 6.02, showed that the 6.01 isoform is more broadly recognized than the 6.02 isoform. In addition, its denaturation had only slightly effects on IgE-binding intensity in sensitive subjects (Willison et al. 2011). In fact, Pru du 6 polypeptides are considered as highly resistant to heat treatment which is one of the most common strategies to decrease or even eliminate the allergenic potential of foods. Due to its heat resistance, contamination of food with Pru du 6 polypeptides presents a serious threat to sensitized patients (Venkatachalam et al. 2002). On the other hand, some experiments using *in vitro* models of gastrointestinal digestion suggested that this allergen is sensitive to pepsin but, interestingly, when almond flour is added to other foods, pepsin's action on Pru du 6 is a lot less effective (Mandalari et al. 2014). Holden et al. (2008) suggested that the reaction between Pru du 6 and α-conglutin from lupine, another 11S globulin, may be the cause of it.

1.2.2.1.2 Pru du 5 (60S acidic ribossomal protein P2)

Pru du 5, also known as 60S acidic ribosomal protein P2, is encoded by *P. dulcis* 60S acidic ribosomal protein gene and was included in the WHO/IUIS allergen list in 2007. This name comes from the fact that this allergen is an 11 kDa protein which is a member of the 60S large subunit of the eukaryotic 80S ribosomes (Zhang and Jin 2020) and its biological function is related to protein biosynthesis. Pru du 5 is considered a major

almond allergen due to the presence of specific IgE antibodies in 50% of sensitized patients' sera (Abou Alhasani and Roux 2009).

This allergen can exist as a complex with other ribosomal components/proteins or in is free state (Vieths et al. 2002) with the ability to form homodimers and oligomers (Flinterman et al. 2006; Hasegawa et al. 2009). On the allergenicity front, this data is very important because oligomerization give the allergen the capability of cross-linking IgE antibodies on mast cells and/or basophils surfaces even if the recognition is made from a single epitope of the allergen (Zhang and Jin 2020).

Although being considered a major allergen and present in the WHO-IUIS allergen list, many authors believe that this classification must be supported by more studies concerning the IgE reactivity of allergic patients' sera to this allergen (Costa et al. 2012; Mandalari and Mackie 2018). Also, studies regarding the biochemical and immunological properties of Pru du 5 on its natural state as an allergen are lacking (Zhang and Jin 2020) leading to the conclusion that newer and further studies are needed.

1.2.2.1.3. Pru du 3 (nsLTP)

Added to the WHO/IUIS database in 2009, Pru du 3 is a non-specific lipid transfer protein 1 (nsLTP1) belonging to the subfamily of nonspecific lipid transfer proteins (nsLTPs) (Noorbakhsh et al. 2011). This family includes proteins constituted by a hydrophobic core to ease lipid transference like phospholipids, steroids, fatty acids and glycolipids between membranes. Besides that, nsLTPs are also known as pathogenesis-related 14 (PR-14) proteins, a member of the prolamin superfamily (Egger et al. 2010; Costa et al. 2012), which actively participate in plant-defense mechanisms against fungal and bacterial pathogens and other environmental stresses (Asero 2011).

In almond were identified and characterized three nsLTP (Yao et al. 2018) with identical molecular weight (9kDa) and similar amino acid length:117, 123 and 116 amino acids for Pru du 3.01, 3.02 and 3.03, respectively. In the three isoallergens, there are eight cysteine conserved residues, which allow the formation of four disulfide bonds (Costa et al. 2012).

Due to the typical accumulation of this protein family in outer epidermal layers, the peels are associated to stronger allergenicity compared with the pulps of the fruits in the Rosaceae family. Regarding allergenicity, this protein family is quite concerning because of its resistance to abrupt pH changes, pepsin digestion, thermal treatments, and the ability

of restore folding structures and the consequent proprieties after cooling (Mills et al. 2007). Cross-reactivity is also a major concern once the nsLTP family is characterized by a high level of conserved sequences and tridimensional structures allowing IgE recognition, which in turn results in cross-reactivity between species (Asero 2011). Furthermore, the Rosaceae fruits and seeds normally present nsLTP proteins and with that comes a high probability of cross-reactivity between for example apple, peach, cherry, apricot and almond (De Angelis et al. 2018). These latest evidences are the main reasons why nsLTPs are included in the panallergens group - allergens ubiquitously spread throughout nature – showing a high level of conservation besides being from different and unrelated organisms (Zhang and Jin 2020).

1.2.2.1.4. Pru du 4 (profilins)

Pru du 4 proteins are included in the profilin family and are encoded by the putative genes *Pru du 4.01* and *Pru du 4.02* (Tawde et al. 2006) which, although present different size fragments (1041 and 754 bp, respectively) encode two proteins with similar sequences (131 aa), molecular weights (roughly 14 kDa) and acidic properties (*p*I near 4.6) (Costa et al. 2012).

These proteins can establish high-affinity complexes with monomeric actin leading to its polymerization into filaments (Staiger et al. 1993). Once they are associated with actin, it is not surprising that profilin allergens are included in the panallergens group with Pru p 4.01 and Pru av 4 from peach and sweet cherry, respectively, being the most similar and identical proteins (99 and 98%, respectively) in relation to almond profilins. In general, profilins seem to present moderate structural stability and harsh conditions contribute to their denaturation and consequent loss of conformational structure. In almond, Pru du 4 profilins are very difficult to detect by immunoblot screens because of the low levels and their labile character. Because of almond profilins antibodies are detected in 44% of patients' sera, they are classified as minor allergens (Tawde et al. 2006).

1.2.2.1.5. Pru du 8

Pru du 8 is one of the latest allergens included in the WHO-IUIS database. This allergen was reactive in six of eighteen sera of almond allergic patients (Che et al. 2019;

Zhang and Jin 2020). Biochemically speaking, Pru du 8 is characterized by a signature repeat of a $CX_3CX_{10-12}CX_3C$ (X being any amino acid) motif which is also related to the N-terminal or the signal peptide of some vicilins (Che et al. 2019) and it was also reported to maintain antimicrobial function of some peptides derived from macadamia vicilin (Marcus et al. 1999).

The first nomenclature attempt for this allergen was based on the sequencing of two short peptides of this allergen to reveal the identity of an IgE-reacting protein several years ago. Nevertheless, the result was a misidentification of this allergen as an almond 2S albumin because of the sequence alignments of the two peptides sequences and those in other 2S albumin proteins (Poltronieri et al. 2002). More recently, *in silico* investigations and bioinformatic analyses reopen the debate, naming this allergen as Pru du vicilin (almond 7S vicilin) although some authors believe in a second misidentification (Garino et al. 2015; Zhang and Jin 2020). In fact, the authors claim that this misidentification is due to the similarity between the signal peptides of vicilins of other species and Pru du 8. Besides that, it is argued that some Pru du 8 orthologs present in the NCBI database, most of them predicted by automatic genome annotations, are incorrectly named as vicilin-like proteins due to the absence of the cupin signature domains of 7S vicilins (Che et al. 2019; Zhang and Jin 2020).

All this controversy shows that further studies are needed in order to better elucidate about the actual protein family of Pru du 8.

1.2.2.1.6. Pru du 10

To date, this allergen was the last one to be added to the WHO-IUIS database. This allergen corresponds to mandelonitrile lyase 2 (formerly hydroxynitrile lyase 2) which is an highly effective catalytic enzyme (Yao et al. 2018). His allergenicity was recognized after allergic response to almond ingestion where thirteen of eighteen almond allergic patients were sensitized. Also, Pru du 10.0101 isoallergen was identified and added to the WHO-IUIS allergen information.

Besides being identified in raw almond samples, this protein was also identified in digested samples which may indicate that this allergen is able to overcome the digestion process (De Angelis et al. 2018). Still, there is still a lack of information regarding this allergen which clearly shows that more studies should address this issue.

1.2.2.2. Allergens not included in the WHO/IUIS allergen list

There are two main processes to classify a protein as a food allergen, based on immunological data as the IgE reactivity or based on sequence similarity with proteins of other species already considered allergens. For an allergen to be included in the WHO-IUIS database, immunological data is required and because of that, some authors defend that those which cannot be supported by it should hardly be assumed as an allergen. However, bioinformatic-based investigation is very important to promote further investigation and aware the scientific and industrial community for the dangers of food allergens.

1.2.2.2.1. Pru du γ-conglutin

The IgE and serological reactivity to Pru du γ-conglutins were not associated with any clinical symptoms and because of that, they are not recognized into standard clinical nomenclature (Mandalari and Mackie 2018).

After the report and characterization of conglutins in other fruits and seeds like lupine (Kolivas and Gayler 1993), peanut (Burks et al. 1991), soybean (Burks Jr et al. 1988) or cashew (Wang et al. 2002), in almond was also identified a N-terminal peptide sequence of 25 aa belonging to a IgE binding protein with a molecular weight of 45 kDa presenting around 40% identity rate between the mature forms of γ-conglutin from withe and narrow-leafed lupine (Poltronieri et al. 2002). Moreover, with a high similarity, approximately 50%, between this almond protein and 7S globulin from soybean, this allergen was considered a vicilin (7S globulins) of the cupin superfamily (Costa et al. 2012; Zhang and Jin 2020). Nevertheless, some authors do not agree with this classification stating that γ-conglutin is not a vicilin due to their biochemical properties (Zhang and Jin 2020). In particular, γ-conglutin presents sequence and structural similarities with xyloglucan-specific endo-beta 1,4-glucanase inhibitors, however, such glucanase inhibition properties are not related to the natural γ-conglutin due to is peptidase cleavage susceptibility (Scarafoni et al. 2016).

The same authors believe that more studies regarding immunological and biochemical properties of this protein are needed and the confirmation of this assumptions would make this protein the first food allergen from this supposed protein family.

1.2.2.2.2. Pru du 1- PR-10 protein (Pathogenesis Related-10 protein)

Pathogenesis related proteins are a common group of proteins, generally upregulated in plants to promote defense mechanisms against pathogens like viruses, bacteria or fungi and environmental factors (Zhang and Jin 2020). The PR-10 family is related to the intracellular defense processes and the response to fungal and bacterial infections. Due to its function, there are numerous isoforms which promote different IgE-binding capabilities (De Angelis et al. 2018). Furthermore, PR-10 proteins are constitutively expressed in different plant parts and usually are not related to other PR proteins (Fernandes et al. 2013). They are commonly seen as pollen or food allergens (Mittag et al. 2004; Scala et al. 2011) and because of that they can be considered as panallergens, being responsible for cross-reaction events (Asero 2011).

Although there is no immunological data to support their classification as an allergen, the high similarity and identity between almond PR-10 proteins and the peach counterparts, which are known allergens (Pru p 1), almond PR-10 proteins are assumed as an allergen and named as Pru du 1 (Asero 2011).

1.2.2.2.3. Pru du 2 (PR-5/thaumatin-like protein)

This allergen group is also known as PR-5 or thaumatin-like proteins (TLPs) and are responsible for the biological response to pathogen infection, fungal proteins, and osmotic stress. The TLP's group is known to be very resistant to proteases, heat-induced denaturation, and pH variations possibly because of sixteen conserved cysteine residues which form eight disulfide bonds (De Angelis et al. 2018). Several isoallergen genes have been identified which code for TLP ranging in molecular weight from 23 to 27 kDa. Also, the isoallergens aminoacidic sequence length ranges from 246 aa to 330 (Chen et al. 2008).

Like PR-10 proteins, no immunological characterization of PR-5 almond proteins exists. Although, it is believed that these proteins are almond allergens due to the high sequence identity with Pru p 2, a peach allergen (Palacin et al. 2010). Moreover, due to their biochemical properties, traditional food-processing practices do not significantly influence these protein's structure and characteristics so, they could affect sensitized patients (Costa et al. 2012).

1.2.2.2.4. Pru du 2S albumin

Included in the prolamin superfamily, 2S albumins are an important group of seed storage proteins involved in seed growth and in defense related mechanisms (Shewry et al. 1995; Roux et al. 2003). Besides 2S albumin, the prolamin superfamily also includes other protein groups like the nonspecific lipid transfer proteins (nsLTPs), prolamin storage proteins and α-amylase/trypsin inhibitors which may indicate several cross-reactions (Shewry and Halford 2002).

2S albumins are thought to be somehow resistant to acidic pH, enzyme digestion particularly the albumins with proteolytic activity and surfactant denaturation effects. These conclusions come from the fact that is believed to this proteins cause sensitization along the intestinal tract which only be possible if the previous resistances were actually accurate (Moreno and Clemente 2008).

As an allergen, the strongest data that lead to the classification of almond 2S albumins as almond allergens is the two short partial peptide sequences with high similarity with 2S albumins of others species (Clemente et al. 2004) that, as discussed in section 3.1.5, some authors believe to be a misidentification and really correspond to Pru du 8 proteins (Zhang and Jin 2020). In fact, 2S albumins of other species, like Ara h 2 (peanut 2S albumins) for example, are very potent allergens (Koppelman et al. 2004; Palmer et al. 2005; Nicolaou et al. 2010) and for this reason the assessment of whether these almond proteins are allergens or not is required and imperative.

Table 2 Almond allergens and their biological function, molecular weight, food processing effects and clinical relevance.

Allergen	Biochemical name	WHO-IUIS	Isoallergen and variants	GenBank nucleotide	UniProt	Biological Function	MW (kDa)	Processing	Clinical Relevance	References
Pru du 3	non-specific Lipid Transfer Protein 1 nsLTP1	Yes (2009)	Pru du 3.0101	FJ652103	C0L0I5	Non-specific lipid transfer protein (nslTP1) and plant defense proteins against pathogens	9	Very resistant to pH, thermal and enzyme treatments	Systemic and life-threatening symptoms; cross reactivity among Rosaceae fruit	(Buhler et al. 2015)
Pru du 4	Profilin	Yes (2006)	Pru du 4.0101 Pru du 4.0102	AY081850 AY081852	Q8GSL5 Q8GSL5	Actin-binding protein for cellular function	14	Unstable during heat processing	Mild symptoms and mainly in oral cavity	(Tawde et al. 2006)

Pru du 5	60S acidic ribossomal prot. P2	Yes (2007)	Pru du 5.0101	DQ836316	Q8H2B9	Protein synthesis	10	Unknown	Unknown	(Abou Alhasani and Roux 2009)
Pru du 6	Amandin, 11S globulin legumin-like protein	Yes (2010)	Pru du 6.0101 Pru du 6.0201	GU059260 GU059261	E3SH28 E3SH29	Major storage protein	360	Stable to dry heat but can be denatured by boiling	Severe IgE allergic reactions	(Willison et al. 2011)
Pru du 8	Antimicrobial seed storage protein	Yes (2018)	Pru du 8.0101	MH922028	A0A516F3L2	Antimicrobial and seed storage function	31	Unknown	Unknown	(Che et al. 2019)
Pru du 10	Mandelonitrile lyase 2	Yes (2019)	Pru du 10.0101	AF412329.1	Q945K2	Highly efficient catalytical enzyme	60	Resistant to enzyme digestion	Unknown	(De Angelis et al. 2018; Yao et al. 2018)

Pru du γ-conglutin	Cupin superfamily	No	_____	_____	_____	7S vicilin storage protein	45 for each subunit	Unknown	Unknown	(Poltronieri et al. 2002)
Pru du 1	PR-10 protein	No	_____	_____	_____	Plant pathogenic and stress response	17	Wet heat processing reduces IgE reactivity	Unknown	(Fernandes et al. 2013)
Pru du 2	PR-5/thaumatin-like protein	No	_____	_____	_____	Pathogenic response	23-27	Resistant to protease, pH or heat treatment	Unknown	(Liu et al. 2010)
Pru 2S albumin	Prolamin super family	No	_____	_____	_____	Seed storage protein	12	Stable to heat treatment	Unknown	(Poltronieri et al. 2002)

1.2.3. Methods for almond allergens detection

Most of the methods used for the detection of almond allergens are based in immunochemical properties, DNA techniques and lately, in Mass Spectrometry (MS) approaches (Costa et al. 2012).

The immunochemical methods are based on the interaction between immunoglobulins and epitopes present in the target protein. For almond allergens detection, lateral flow devices, immunoblotting and especially Enzyme-Linked Immunosorbent Assay (ELISA), are very standard methods and the usual techniques for quantitative and qualitative detection of food allergens (Jackson et al. 2008; Wang et al. 2010). This comes from the fact that, ELISA tests, for example, have enough sensitiveness for protein detection (in the orders of ppm), being the main advantage the fast assessment which is a great advantage for clinical purposes. (Wang et al. 2010). Several immunological commercial kits, like the ones exemplified in Table 3, have been developed with the objective of delivering the most sensitive result in the short amount of time. As seen in the kit's characteristics, the assay type should be taken in serious consideration according to the situation that are supposed to be used.

Table 3 Example of commercial immunological kits for almond detection and/or quantification and their main characteristics: time for results including extraction times, assay type, limit of detection (LOD), limit of quantification (LOQ) and their manufacturers.

Kit [1]	Assay Time	Assay Type	LOD (ppm)	LOQ (ppm)	Company
ELISA-based					
MonoTrace ELISA kit	40 min	Monoclonal antibody-based ELISA	0.15	1	BioFront Technologies, Tallahassee, FL, USA
SENSISpec ELISA almond	75 min	Sandwich enzyme immunoassay	0.2	0.4	Eurofins Technologies, Budapest, Hungary
RIDASCREEN FAST Mandel/Almond	50 min	Polyclonal antibody specifically for almond protein detection, sandwich ELISA	0.1	2.5	R-Biopharm AG, Madrid, Spain
AgraQuant® Plus Almond	30 min	Sandwich enzyme-linked immunosorbent assay	0.5	1	Romer Labs®, Getzersdorf, Austria
LFD-based					
AgraStrip® Almond	11 min	Lateral flow device	2	———	Romer Labs®, Getzersdorf, Austria

Reveal 3-D Almond Test	10 min	Lateral flow device	5	_____	Neogen Corp., Lansing, MI, USA
Lateral Flow Almond incl. Hook Line [2]	10 min	Lateral flow device	1	_____	R-Biopharm AG, Madrid, Spain

[1] Mention of commercial kits and trade names is only for exemplification purposes and the authors declare no competing financial interest. [2] The hook line is included with the purpose of overcoming the hook effect—very high amounts of an analyte in the sample can lead to falsely lowered or negative results.

Another possible approach, instead of looking directly for the protein itself, is the DNA-based method where it is performed an amplification of the gene fragment responsible for encoding the allergen by Polymerase Chain Reaction (PCR) allowing quantitative and qualitative measurement using real-time PCR or endpoint PCR assays, respectively (Mandalari and Mackie 2018). One of the advantage of these methods rely on the detection of low quantity of almond DNA even after food processing which could promote the degradation of some allergen proteins and therefore not be detected by immunological approaches (Pafundo et al. 2009). However, the presence of the gene encoding the allergens does not imply its expression, although for almond allergens case, being essential proteins in the normal development of almond the expression of the respective genes is very like to occur. Because of that, the synergistically use of DNA-based techniques and ELISA could overcome some of the drawbacks of both techniques (van Hengel 2007).

Proteomics play a very important role in the food allergy problematic, firstly on a fundamental investigation basis to characterize allergens and further to their application to the diagnostic routines. Namely, a variety of tests and methods must be applied to characterize allergens according to their allergenic activities, purity and folding properties. Following this line of thought, SDS-Polyacrylamide Gel Electrophoresis (SDS-PAGE) is a reliable technique to determine purity and following 2 Dimension (2D) electrophoresis, capillary electrophoresis or High Performance Liquid Chromatography (HPLC) are great techniques to access individual isoforms and obtain more additional information in general. Further, MS techniques are the powerful tolls to determine proteins molecular masses being Matrix Assisted Laser Desorption Ionization (MALDI) and ElectroSpray Ionization (ESI) the most commonly used (Hoffmann-Sommergruber 2016; Piras et al. 2016). MS techniques have been the most recent methods to be explored for qualitative and quantitative purposes (Johnson et al. 2011; Monaci et al. 2018). For example, the isolation and characterization of Pru du 3 allergen was done using MS techniques where the full sequence was obtained by Liquid Chromatography ElectroSpray Ionization Orbitrap Mass Spectrometry (LC-ESI-Orbitrap-MS) (Buhler et al. 2015). Mass

spectrometry has the advantage of ELISA tests which can directly identify proteins with quite sensitivity and therefore could provide a direct risk evaluation and, besides that, can be used for the detection of multiple allergens simultaneously (Bignardi et al. 2010). MS could be the chosen technique for a standard test; however, it is a relatively recent approach which demands expensive equipment and specialized personnel. At this standpoint, further improvements are required in order to allow easier access and profitable use by clinical facilities (Mandalari and Mackie 2018).

Another methodology under development is based on microarrays. Namely, allergen microarrays like the MeDALL allergen-chip have been explored for the diagnosis and monitoring of allergies. The main advantages rely on the simultaneously detection of several allergens with a minimal amount of sera in reduced time. The development of this chip has the propose of monitoring IgE and IgG reactivity profiles against 170 allergens in sera collected from European birth cohorts. With that information, it would be possible to make a geographical association of clinical important allergens in different populations and track the progress of food allergy itself and allow clinical therapies to act in a prophylactic and more personalized manner (Lupinek et al. 2014).

Chapter 2: Objectives

Due to almond varietal diversity, differences in the IgE-binding properties of each variety may be suspected, which could allow the development of therapeutic strategies based on the use of hypoallergenic varieties to induce desensitization. With this work, we intended:

- To assess the genetic diversity among different almond varieties including popular modern varieties as well as traditional Portuguese ones;
- To evaluate the allergenic potential of a genetically diverse set of almonds (n = x varieties);
- To characterize the almond allergome and to identify major and minor allergens;
- To develop a DNA-based approach to detect the presence of important almond allergens.

Chapter 3: Materials and methods

3.1. Plant Material

Almond varieties (Table 4) were collected in 2007, 2014, 2015, and 2016 from four different locations from the northern area of Portugal specifically, the Trás-of-Montes region as shown in Figure 4.

Figure 4 Map of Portugal's north region highlighting the geographical location of the four regions (Murça, Mirandela, Moncorvo and Alfandega da Fé) from which almond samples were collected.

Besides that, commercial crops (meaning, varieties not acquired directly from the producer) from the same varieties were added to the sample group with the objective of secure a wider study as possible. All the details (code, variety, year of collection and geographical location) about each sample used to perform this study can be found in Table 4. To facilitate sample identification in the text, samples will be referenced as 'variety name' (year and/or region).

Table 4 Detailed information of all the samples used in proteomic studies, including sample code, variety, year of collection and geographical location.

Sample Code	Variety	Year	Type of crop/Region
A	Amendoão	2016	Murça
B	Bonita	2016	Moncorvo
C	Casanova	2016	Mirandela
D	Casanova	2016	Moncorvo
E	Casanova	2007	Comercial crop
F	Ferraduel	2014	Moncorvo
G	Ferraduel	2007	Comercial crop
H	Ferragnés	2015	Alfandega
I	Ferragnés	2016	Alfandega
J	Ferragnés	2016	Mirandela
K	Ferragnés	2007	Comercial crop
L	Ferrastar	2007	Comercial crop

M	Gloriette	2007	Comercial crop
N	Gloriette	2015	Moncorvo
O	Gloriette	2016	Alfandega
P	Gloriette	2016	Mirandela
Q	Masbovera	2015	Alfandega
R	Masbovera	2016	Mirandela
S	Marcelina	2007	Comercial crop
T	Marcona	2007	Comercial crop
U	Orelha de mula	2007	Comercial crop
V	Pegarinhos	2016	Mirandela
W	Pegarinhos	2016	Moncorvo
X	Pegarinhos	2016	Murça
Y	Refego	2007	Comercial crop
Z	Refego	2016	Moncorvo

Genomic studies were performed with a different group of samples due the amount of material available at the time these techniques were performed. The commercial crops were common with the proteomic-studied group: C – 'Casanova' 2016, E – 'Casanova' (2007), H – 'Ferragnés' (2015), K – 'Ferragnés' (2007), L – 'Ferrastar' (2007), M – 'Gloriette' (2007), N – 'Gloriette' (2015), Q – 'Masbosvera' (2015), S – 'Marcelina' (2007), T – 'Marcona' (2007), U – 'Orelha de Mula' (2007), X – 'Pegarinhos' (2016) and Y – 'Refego' (2007); however, other varieties were added to diversify the sample group. Specific information about the samples specifically used in the genomic studies is presented in Table 5.

Table 5 Sample code, variety, year and type of crop of the samples used for the realization of genomic techniques.

Sample code	Variety	Year	Type of crop/Region
1	Orelha de Mula	2007	Comercial
2	Pegarinhos	2007	Regional
3	Pegarinhos	2007	Regional
4	Ferrastar	2007	Comercial
5	Ferragnes	2007	Comercial
6	Ferraduel	2007	Comercial
7	Casanova	2007	Comercial
8	Gloriette	2007	Comercial
9	Refego	2007	Comercial
10	Marcelina	2007	Comercial
11	Marcona	2007	Comercial
12	Duro Italiano	2007	Regional

13	Gloriette	2015	Moncorvo
14	Masbosvera	2015	Alfandega
15	Ferragnés	2015	Alfandega
16	Casanova	2016	Mirandela
17	Pegarinhos	2016	Murça

3.2. Protein extraction and quantification

Each sample was subjected to a defatting process to remove excess lipid content to facilitate the next steps of protocols. To do that, acetone was added several times to almond flours following a paper filtration process and dried overnight as described by Pastorello et al. (2002).

Almond proteins were extracted using a saline solution. Briefly, defatted almond samples were suspended (1:20, w/v) in 0.1 mol/L triphosphate buffered saline (TBS, pH 7.4). The mixture was incubated at 20 °C with stirring for 1 h followed by a centrifugation step at 10.000 g for 10 min and the supernatant was collected. In the case of two-dimensional electrophoresis (2-DE), defatted hazelnut samples were suspended (1:10, w/v) in extraction solution containing 4% (w/v) CHAPS (7 mol/L urea, 2 mol/L thiourea, 1% (v/v) immobilized pH gradient (IPG) buffer), 20 mmol/L of DTT in milliQ ultrapure water, as described elsewhere (Ribeiro et al. 2015). The mixture was stirred for 1 h at 4 °C, disrupted by sonication with an ultrasonic homogenizer (Vibra-CellTM VCX130, Sonics & Materials Inc., Newtown, United States) in three bursts of 10 s at 30% of full power (Portugal et al. 2015), centrifuged at 30,000xg for 20 min at 4 °C and the supernatant collected. Protein concentration was in both cases determined using the 2D Quant kit (GE Healthcare, Buckinghamshire, UK) following the manufacturer's instructions. For the both cases, protein concentration was determined using the 2D Quant kit (GE Healthcare, Chicago, IL, USA), following the manufacturer's instructions. Briefly, the kit is based on the precipitation of proteins while leaving interfering substances behind, and consequent specific binding of copper ions to protein. A standard curve for the semiquantitative procedure was done using increasing bovine serum albumin (BSA) concentrations. The concentration is determined by reading the absorbance level and matching those results with the ones of the standard curve.

3.3. One-dimensional electrophoresis (1-DE)

The almond protein extracts were separated by electrophoresis using a big gel system (160 x 180 x 1 mm) in order to obtain maximum protein separation for profiling and band analysis.

Briefly, almond proteins were resuspended in sample buffer (2% SDS, 40% glycerol, 0.02% bromophenol blue, 0.08 M Tris-HCl pH 8.0) and separated in a resolving gel using 12.52% T (total monomer concentration) and 0.97% C (weight percentage of crosslinker) and tris-glycine buffer. Also, when reducing conditions were applied, to the sample buffer was added 1% (w/v) dithiothreitol (DTT). For each sample, sample buffer was added to a total volume of 30 µL and protein quantity per well was 60 µg. The gels were stained with Coomassie Blue R-250 for 24 h and then washed in distilled water overnight. Coomassie-stained gels were scanned with a flatbed scanner (Umax PowerLook 1100, Fremont, CA, USA) and the digitized images were analyzed using CLIQS 1D Pro software (TotalLab, UK) to establish a binary matrix indicating the absence/presence of specific bands. Besides that, was also performed a small gel system (80 x 80 x 1 mm) with the intention to perform immunoblotting assays using almond-allergic patients' sera. In this case, following the manufacturer instructions, proteins were resuspended in Bolt™ LDS sample buffer with 10% (v/v) Bolt™ Sample Reducing Agent (Thermo Fisher Scientific, Massachusetts, USA) and separated in precast Bolt Bis-Tris Plus gels using 4-12% gradient polyacrylamide concentration and MES SDS Running Buffer (Thermo Fisher Scientific, Massachusetts, USA) (Ribeiro et al. 2020).

3.4. Immunoblotting assays

All gels, containing 20 µg of protein per well, were electroblotted to a nitrocellulose membrane using the iBlot 2 Dry Blotting Systems (Invitrogen, California, USA) with a template program: 20 V for 1 min, 23 V for 4 min and 25 V for 3 min. After that, the membrane was blocked with 5% dry milk in 1x TBST (0,01M Tris-Based Saline with 0,1% Tween 20) for 1 h and then incubated with the patient sera described in Table 6 (PlasmaLab International Washington, USA) or negative controls overnight at 4ºC with agitation following a 1 h incubation at room temperature with agitation with the anti-human IgE (Ɛ-chain specific) – peroxidase antibody (Sigma Aldrich, Missouri, USA). Between incubation times, the membranes were washed using a 1% dry milk diluted in 1x TBST solution and after the last incubation, membranes were washed 2 times with 1x TBST and 1 time with 1x TBS for 10 min each. The results were visualized with chemiluminescent horseradish peroxidase (HRP) substrate (WesternBright™ ECL kit, Advansta Inc., CA, USA) and X-ray films (LucentBlue X-ray film, Advansta Inc., CA, USA). The X-ray films were then let to rest until completely dry and scanned to be analyzed (Ribeiro et al. 2020) .

Table 6 Clinical information about the patients which sera was used to perform immunoblotting assays. Here is summarized data about age, sex, diagnosis (known reactions linked to allergies) and units (kU/l) of almond allergen (code F20) of each patient.

Code	Age	Sex	Diagnosis	F20 Allergen (kU/l)
#A	59	M	Allergic rhinitis (sneenzing, watery eyes, nose running and skin itching)	7.25
#B	71	F	Allergic rhinitis	18.9
#C	38	F	Allergic rhinitis (itchy tongue/mouth)	3.3
#D	34	M	Allergic rhinitis, asthma	10.6
#E	50	F	Allergic rhinitis	12.0
#F	26	M	Eczema, swelling throat, rash/hives, dissey, itchiness, asthma, anaphylaxis in response to all nuts	_____
#G	30	M	Allergic rhinitis, asthma sometimes triggered by allergies, anaphylaxis to peanuts and severe reaction to most tree nuts, eczema, itchy dry skin	29.0
#H	37	F	Eczema, itchy eys, nose bleeds, asthma, weezing	_____
#I	45	F	Allergic rhinitis, stomach cramping, diahrea, vommiting, anaphylaxis, body hives	13.9
#J	51	M	Asthma, eczema, elevated IgE, weezing, throat itches	14.7

3.5. Two-dimensional electrophoresis (2-DE)

2DE sample extracts were first resuspended in a rehydration solution, consisting of an (1:10, w/v) in extraction solution containing 4% (w/v) CHAPS (7 mol/L urea, 2 mol/L thiourea, 1% (v/v) immobilized pH gradient (IPG) buffer), 20 mmol/L of DTT in milliQ ultrapure water and bromophenol blue up to the volume recommended by the manufacturer for the immobilized pH (3-10) gradient (IPG) strips (7 cm) (Ribeiro et al. 2013; Ribeiro et al. 2015). Isoelectric focusing (IEF) of 80µg of protein was performed on an EttanTM IPGPhor IITM system (Amersham Biosciences, Uppsala, Sweden) using the following conditions: a first step of active rehydration was performed at 50 V for 12 h, followed by a linear gradient up to 300 V for 30 min; a step at 300 V for 30 min; a linear gradient up to 1,000 V for 30 min; a linear gradient up to 5,000 V for 1.5 h and finally, a step at 5,000 V for 45 min. Focused IPG strips were equilibrated twice for 15 min in equilibration buffer [(6 mol/L urea, 30% (w/v) glycerol, 2% (w/v) sodium dodecyl sulphate (SDS) in 0.05 mol/L Tris–HCl buffer pH 8.8)]. In the first equilibration step, 1% DTT was added to the original equilibration buffer, and 4%

iodoacetamide to the second step (Ribeiro et al. 2020). Bromophenol blue was added to both solutions. The equilibrated IPG strips were gently rinsed with MES electrophoresis buffer, softly dried to remove excessive buffer, and then applied to NuPAGETM 4-12% Bis-Tris ZOOMTM Protein Gel (Thermo Fisher Scientific; 80 x 80 x 1 mm; 4-12%; IPG-well). After electrophoresis, the 2-D gels were used for immunoblotting or fixed in 10% (v/v) acetic acid, 40% (v/v) methanol solution for 1 h, followed by Coomassie Brilliant Blue G-250 staining overnight. Excess staining was removed by rinsing the gels with 20% (v/v) methanol solution. Coomassie-stained gels were scanned with a flatbed scanner (Umax PowerLook 1100; Fremont, CA, USA) and the digitized images were analyzed using Lab Scanner Image Master 5.0 software (Amersham Biosciences; GE Healthcare) and Progenesis SameSpots v4.5 (Non-linear Dynamics Limited, Newcastle, UK).

3.6. DNA extraction

DNA was extracted from 100 mg of defatted almonds using the commercial Nucleospin Food kit (Macherey-Nagel, Düren, Germany), according to the manufacturer's instructions with minor alterations, which included the incubation with 2 µL of RNase A (2 mg/mL) at room temperature for 5 min, after the lysis step. Yield and purity of the DNA extracts were assessed by UV spectrophotometric DNA quantification on a Synergy HT multi-mode microplate reader (BioTek Instruments, Inc., Vermont, USA), using a Take3 micro-volume plate accessory and using the nucleic acid quantification protocol with sample type defined for double-strand DNA in the Gen5 data analysis software version 2.01 (BioTek Instruments, Inc., Vermont, USA). All the extracts were kept at −20 ºC until further analysis (Costa et al. 2013).

3.7. End-point PCR

All samples were previously diluted to obtain an equal 10 ng/µL concentration and PCR amplifications were carried out in 25µL of total reaction volume containing 2µL of DNA extract, 670 mM of Tris–HCl (pH 8.8), 160 mM of (NH4)2SO4, 0.1% of Tween 20, 200µM of each dNTP, 1.0 U of SuperHot Taq DNA Polymerase (GenaxxonBioscience, Ulm, Germany), 3.0 mM of MgCl2and 200 nM of each primer (described in Table 7) using a MJ Mini thermal cycler (BioRad, Hercules, CA, USA) with the next program: initial denaturation at 95 °C for 5 min; 40 cycles at 95 °C for 30 s, 65 °C for 30 s and 72 °C for 45 s; and final extension at 72 °C for 5 min (Costa et al. 2013).

The Prd6 set was used with the goal of confirming the presence of Pru du 6 protein in all the sample group and to confirm the usefulness of this technique as standard protocol for the detection of almond as described before (Costa et al. 2013). For Prd3.1/Prd3.2 primer sets, sample 13 – Gloriette 2015, was diluted (v:v 1/10) five times and each concentration was tested using temperatures of 59ºC, 61ºC and 63ºC and both primer sets (Prd3.1 and Prd3.2) to optimize reaction conditions. After each reaction, electrophoresis was performed in a 1.0% agarose gel containing 1× Gel Red (Biotium, Hayward, CA, USA) for staining and carried out in 1× STGB (GRISP, Porto, Portugal) for 25 min at 200 V. Using the UV light tray Gel Doc™ EZ System (Bio-Rad Laboratories, Hercules, CA, USA), a digital image of the agarose gel was obtained with Image Lab software version 5.2.1 (Bio-Rad Laboratories, Hercules, CA, USA).

Table 7 Detailed information about primers sets Prd6, Prd3.1 and Prd3.2 Specifically, their sequences, reference in the case of Prd6 set and GenBank information about Prd3.1 and Prd3.2.

Primer	Sequence (5'-3')	Reference/Amplicon
Prd6		
1-F Forward	CCG CAG AAC CAG TGC CAG CT	(Costa et al. 2013) / 121 bp
1-R Reverse	CCC CGG CAC ACT GGA AGT CCT	
Prd3.1		Prunus dulcis lipid transfer protein 3 precursor, mRNA, complete cds GenBank: FJ652103.1 / 121 bp
Forward	ATG GCG GTT GGT GGT CCC AA	
Reverse	CAG TGC AGC AAC TAG GGT TCA	
Prd3.2		Prunus dulcis lipid transfer protein 3 precursor, mRNA, complete cds GenBank: FJ652103.1 / 109 bp
Forward	TCT CTA CAG CTT GGC TCA GAC C	
Reverse	AGC CGC AAG CCC AGC ATT TG	

3.8. Statistical analysis

The binary matrix produced by the CLIQS 1D Pro software (TotalLab, UK) was used for the clustering of the varieties (Peakall and Smouse 2006). An unweighted pair group method with arithmetical averages (UPGMA) dendrogram of genetic similarity was generated using the Jaccard's similarity coefficient in NTSYS software, v. 2. (Rolph 2000). The adjustment of the UPGMA dendrogram to the binary matrix was evaluated by the cophenetic matrix correlation with COPH and MXCOMP modules of NTSYS software.

The results are expressed as mean ± standard deviation (SD) unless otherwise stated. Differences among the different samples were determined by one-way analysis of variance

(ANOVA). Multiple comparisons were performed using Tukey's post-hoc test, and the criterion for significance was $p < 0.05$ (GraphPad Prism v6.03, GraphPad Software, La Jolla, California, USA).

Chapter 4: Results and discussion

4.1. Genetic Diversity

The protein profile obtained for the varieties under study, both for native and reduced conditions, was similar as can be seen in Figure 5a and 5b, and 5c and 5d, respectively. The analysis was performed using a large SDS-PAGE system (160 x 180 x 1mm) to promote better separation and enhance resolution of protein profiles. Some differences were spotted by bioinformatic analyses using CLIQS 1D Pro software. However, in the perspective of profiling almond varieties it looks clear that reduced conditions promote a more representative band profile allowing a better and easier profiling pattern and consequent variety classifying. In global analyses of the reduced profiles, two major groups can be distinguished between 16 to 36 kDa and 36 to 64 kDa, respectively. This fact is not surprising due to the fact that most proteins and/or their subunits molecular weight is comprehended in this range (Bezerra et al. 2021).

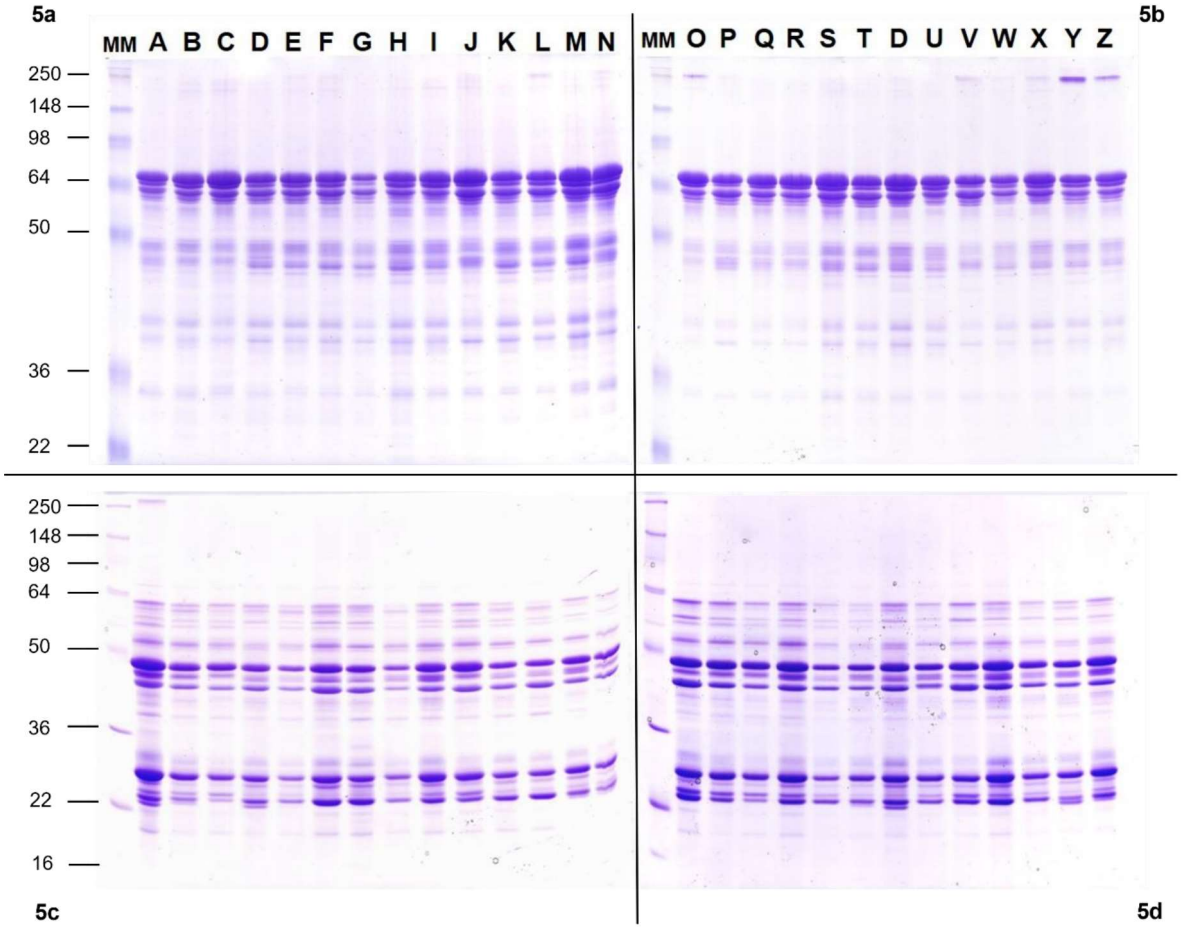


Figure 5 Protein profile of the samples tested in native conditions – Figure 5a and 5b, and in reduced conditions achieved by the addition of DTT to sample extract – Figure 5c and 5d. For all cases sample D was used as inter-gel control to ensure that the same conditions were applied in both gels. From MM to Z, in each well was deposited:

MM - Molecular Marker, A - Amendoão, B - Bonita, C - Casanova, D - Casanova, E - Casanova, F - Ferraduel, G - Ferraduel, H - Ferragnés, I - Ferragnés, J - Ferragnés, K - Ferragnés, L - Ferrastar, M - Gloriette, N - Gloriette, O - Gloriette, P - Gloriette, Q - Masbovera, R - Masbovera, S - Marcelina, T - Marcona, U - Orelha de mula, V - Pegarinhos, W - Pegarinhos, X - Pegarinhos, Y - Refego, Z – Refego. See Table 4 for more information about each variety.

Specifically, native conditions showed the most representative proteins have a molecular weight above 50 and slightly under 64 kDa. Overall, reducing conditions promoted more intense regions at lower molecular weight areas most likely because by breaking bisulfide bounds, smaller subunits of larger and representative proteins were released (Ribeiro et al. 2020). For example, amandin (Pru du 6), one of the proteins most represented in almond proteome and a major allergen, is known to have two main polypeptides with approximately 61 and 63 kDa. Moreover, each of those subunits is characterized for having an acidic (within 42 and 46 kDa) and a basic (20 to 22 kDa) subunits linked by bisulfide bounds (Garcia-Mas et al. 1995). Those results are coherent with our profiles once that in native conditions, two major bands are represented in the 60 kDa range and, when reducing conditions were applied, the two most intense regions appeared in the 40 and 20 kDa areas. Besides that, in both profiles, regions under 20 kDa are very faint. Nevertheless, it was possible to observe a band most likely corresponding to Pru du 5 and Pru du 4 proteins which the literature report them to have molecular weights around 11 and 14 kDa, respectively (De Angelis et al. 2018). However, SDS-Page resolution for these molecular weights is considered poor and may vary from case to case (Schägger 2006), which can explain the fact of in our case, only one band is visible instead of the two.

The reducing conditions allowed to disclose differences regarding the presence/absence and the intensity of protein bands within the collection, especially in the 25 to 40 kDa range. Proper variety identification was achieved and different algorithm-based cluster analyses based on the absence/presence of protein alleles using the unweighted pairwise group method with arithmetical average (UPGMA) were performed. With these results, dendrograms representing hierarchical associations within the studied group were generated. First, different tree-building algorithms were performed; the dendrogram based on Jaccard's similarity coefficient (Figure 6) was preferred based on the higher cophonetic correlation coefficient for this matrix (r=0,855) when compared to the others, which reveals a more representative clustering (Rohlf and Fisher 1968).

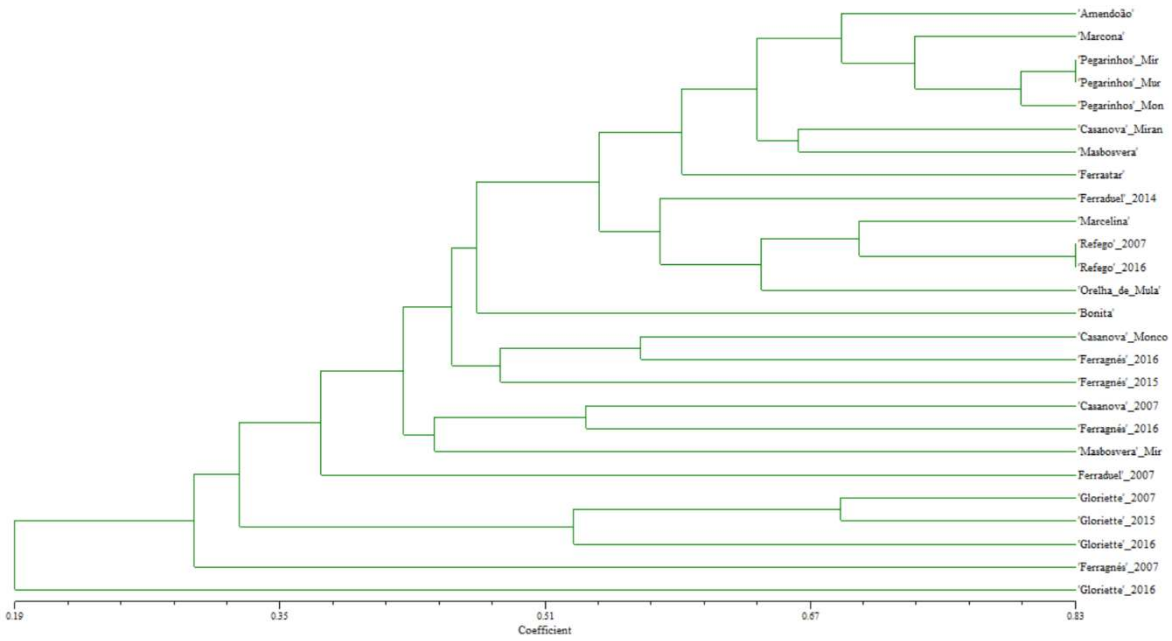


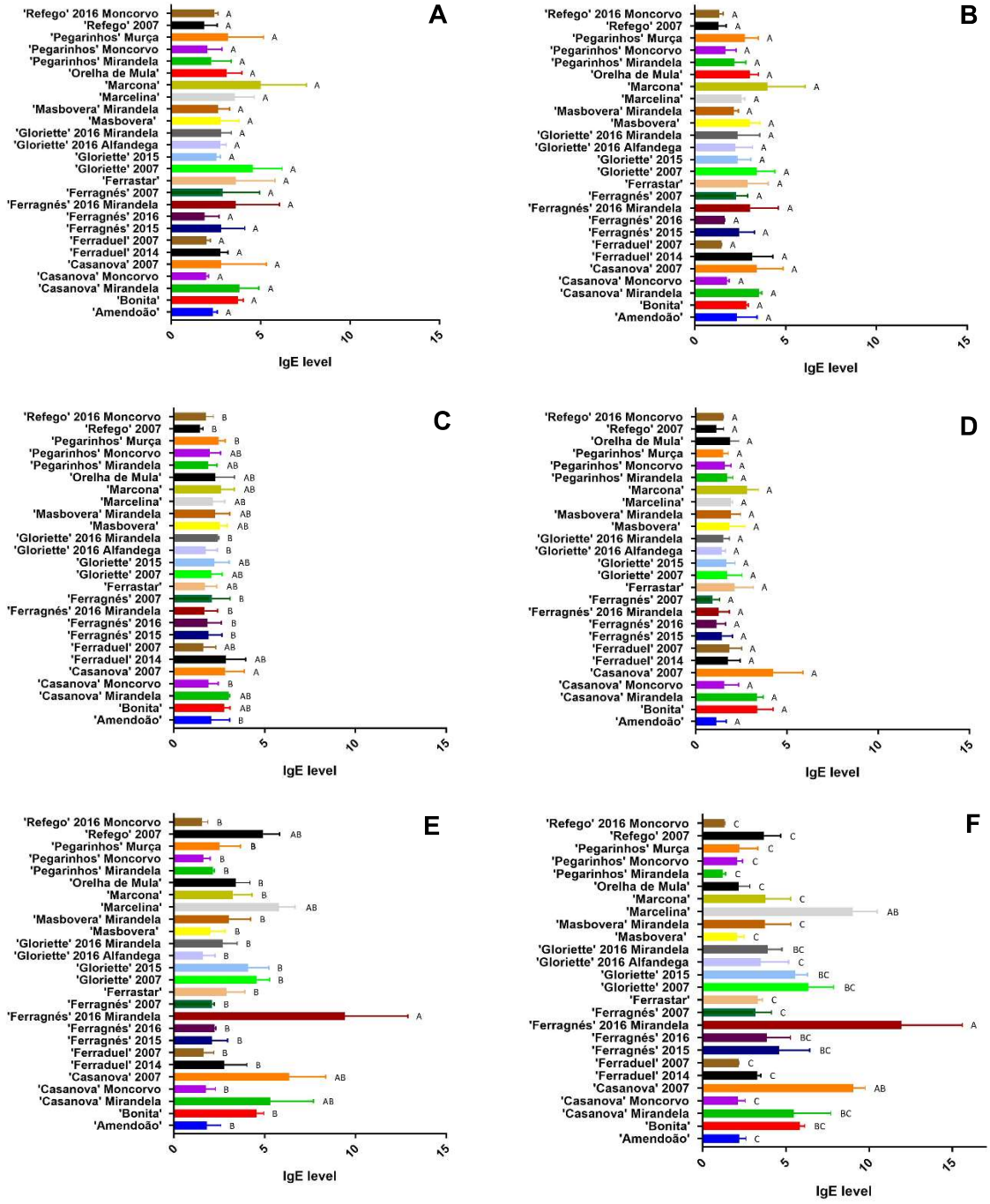
Figure 6 Dendrogram based on Jaccard's similarity coefficient with a cophonetic correlation for the matrix (r=0,855) revealing the most representative algorithm among those tested.

The results showed that among all the studied varieties, both 'Pegarinhos' (Mirandela)/'Pegarinhos' (Murça), and 'Refego' (2007)/'Refego' (2016) were the pairs with higher similarity index (0.83). Moreover, these results were surprisingly accurate since the fact that variety identification is often compromised by homonymy and synonymy problems (Pérez-Sánchez and Morales-Corts 2021). Some examples of that, besides the ones already described are the 'Gloriette', 'Ferraduel' samples that most of them were paired together and the 'Masbovera' case that was isolated from the others. Besides that, it was not found a clear relationship between year/geographical locations and the clustering of the different varieties. Finally, the varieties under study proven to be a genetically diverse group and therefore appropriate for this study.

4.2. Allergenic potential

The IgE-binding properties of each variety for each allergic patient sera analyzed are shown in Figure 7. This analysis focused on the total IgE-reactivity of individual varieties in comparison to the respective loading controls (Romero-Calvo et al. 2010). Patients #H and #J showed the lowest reactivity when compared with the rest of the group. On the other hand, patients #E and #F presented the higher reactivity. However, and with the results of the statistical analyses, no significant differences were found among the varieties tested in the patients #A, #B, #D, #I and #J cases. In the rest of the cases, significant differences were found;

the respective statistical group for each variety after the statistical analyses are summarized in Figure 7.



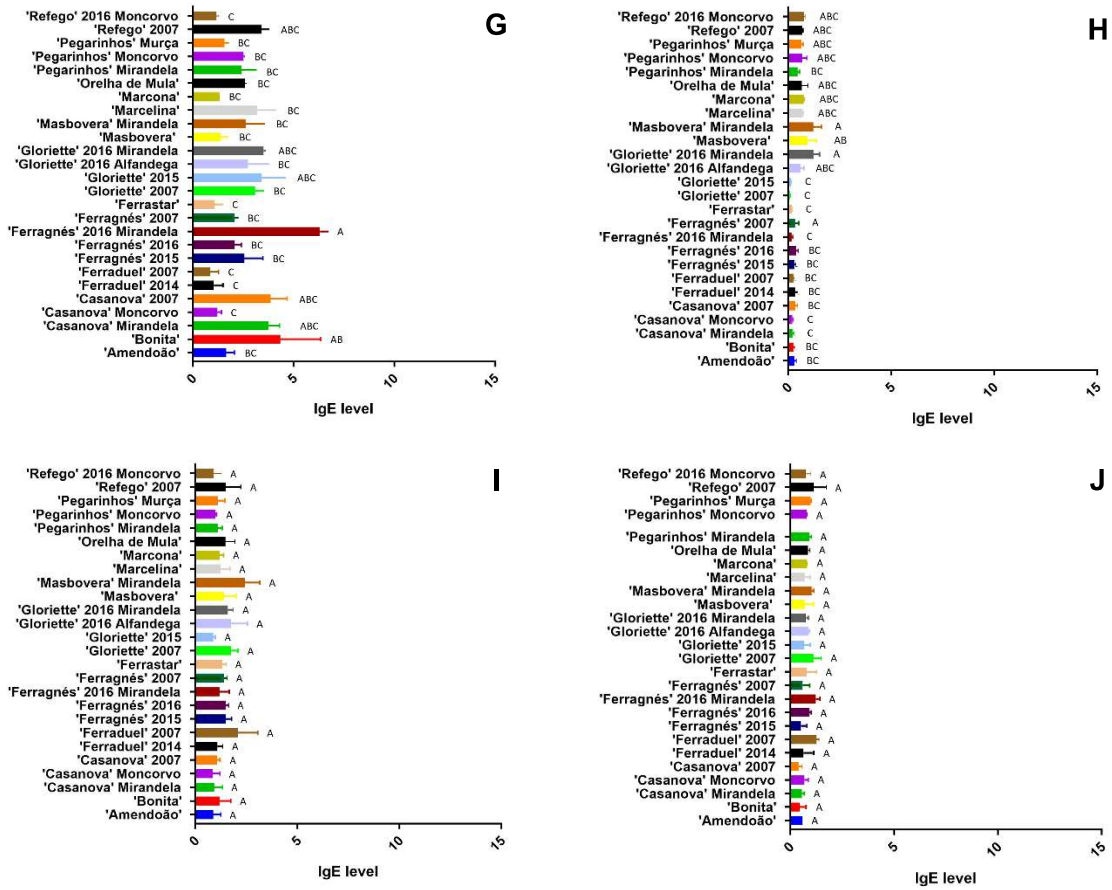


Figure 7 Allergenic potential of each variety: , A - Amendoão, B - Bonita, C - Casanova, D - Casanova, E - Casanova, F - Ferraduel, G - Ferraduel, H - Ferragnés, I - Ferragnés, J - Ferragnés, K - Ferragnés, L - Ferrastar, M - Gloriette, N - Gloriette, O - Gloriette, P - Gloriette, Q - Masbovera, R - Masbovera, S - Marcelina, T - Marcona, U - Orelha de mula, V - Pegarinhos, W - Pegarinhos, X - Pegarinhos, Y - Refego, Z – Refego, see Table 4 for more information about each variety. Arbitrary units of IgE levels of each Patients #A to #J sera to all varieties when compared to loading controls. Each variety was given a statistical group after anova tests and the respective group is represented with letters A, B and C being A≥AB≥ABC≥BC≥C regarding IgE levels.

Considering these results, for the patients #C and #E, three groups of varieties emerged, group A, AB and B. For both cases, only one variety was included in group A, namely, 'Casanova' (2007) in the case of Patient #C and 'Ferragnés' (2016 Mirandela) in the case of Patient #E. It is worth noting that in Patient #E case, the 'Ferragnés' (2016 Mirandela) sample presented near 10 arbitrary units of IgE levels in contrast with the rest of the samples (~5 arbitrary units). This contrast is even greater when compared with Patient #C results where even the samples with higher IgE levels do not even reach the 5 units benchmark. This clearly shows that Patient #E reaction to almond proteins is more intense which agrees with the data presented in Table 6, where is possible to see that Patient #C has one of the lowest levels of the IgE (3,3 kU/I) when compared with Patient #E (12,0 kU/I) as well as with the rest of the group.

Patient #H results were, in part, similar with the ones of Patient #C regarding the general low reaction levels. However, four statistical groups have emerged A, AB, ABC and C because of the near neutral reaction to varieties 'Casanova' (Mirandela), 'Casanova' (Moncorvo),

‘Ferragnés’ (2016 Mirandela), ‘Ferrastar’, ‘Gloriette’ (2007) and ‘Gloriette’ (2015) that were, consequently, included in group C. Even though IgE levels were low, ‘Gloriette’ (2016 Mirandela) and ‘Masbosvera’ (Mirandela) varieties presented enough higher levels to be statistically significant compared to the rest of the group and therefore were included in group A. Despite not having previous information about the IgE levels, it would be plausible to assume that those levels would be low.

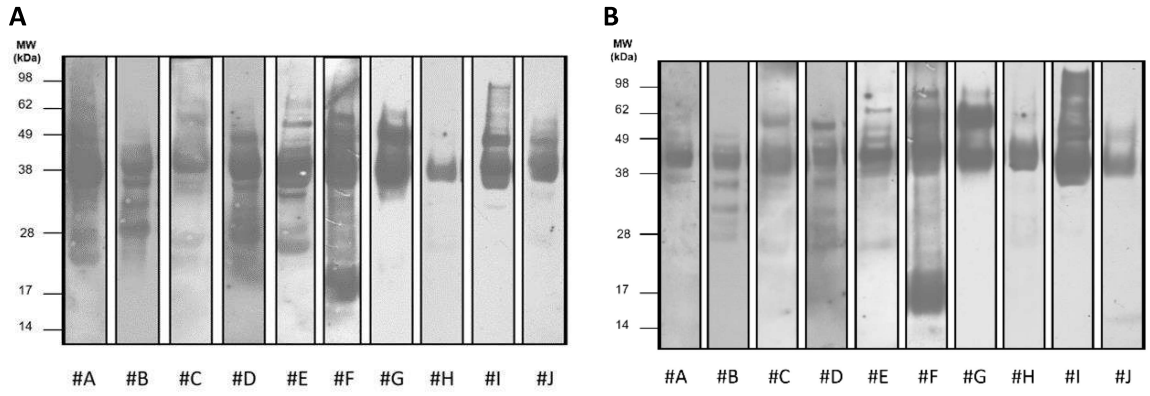
Besides that, Patients #F and #G were the ones that showed highest reaction levels. Starting with the #F case, it was clear the massive reaction to the variety ‘Ferragnés’ (2016 Mirandela) with over 10 arbitrary units regarding IgE levels, being the only one present in group A. The rest of the varieties presented relatively lower ranges and therefore classified in group C with the exception of ‘Casanova’ (2007) and ‘Marcelina’ varieties in the AB group with levels close to 10, and ‘Bonita’, ‘Casanova’ (Mirandela), ‘Ferragnés’ (2015), ‘Ferragnés’ (2016), ‘Gloriette’ (2007) and ‘Gloriette’ (2016 Mirandela) with levels closer or higher than 5 in the BC group.

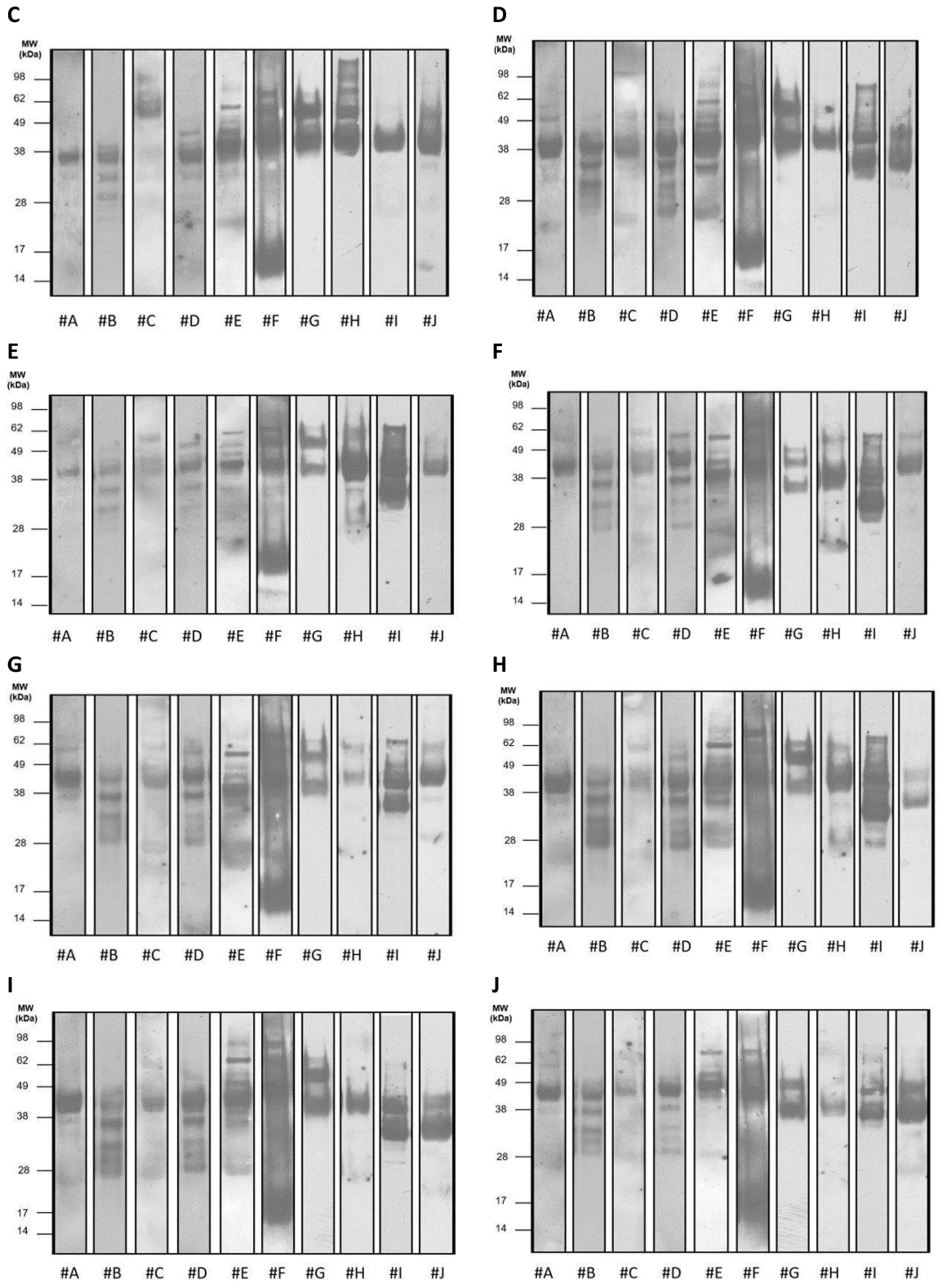
For the Patient #G IgE levels were in the range of the rest of the group, which was surprising once that this was the one with the highest levels of the IgE (29,0 kU/l), however five statiscal groups were found: A, AB, ABC, BC and C although variety ‘Ferragnéss’ (2016 Mirandela) stands out in group A with over 5 units with the rest of the varieties ranging well under that value. Among these, ‘Casanova’ (Moncorvo), ‘Ferraduel’ (2014), ‘Ferraduel’ (2007), ‘Ferrastar’ and ‘Refego’ (2016 Moncorvo) varieties presented the lowest levels and end up in group C. The rest of the varieties, although being distributed in three different groups, presented quite similar values.

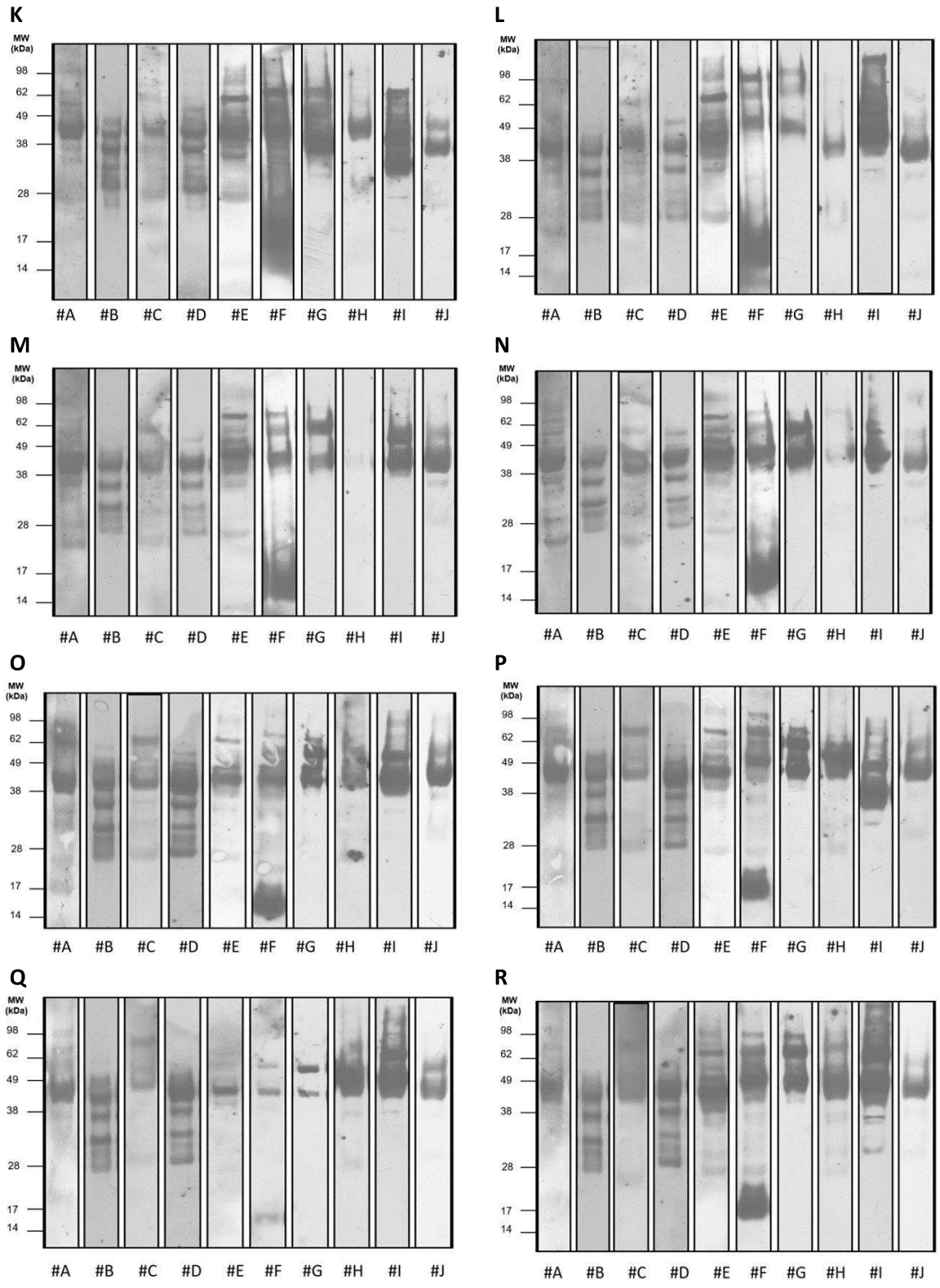
With that being said, ‘Ferragnés’ (2016 Mirandela) presented higher relative IgE levels that all the other tested varieties in 30% of the test group. This allows to conclude that from our sample group, this variety could have the greatest allergenic potential, nevertheless further studies are required as for example Basophil Activation Tests (BAT) (Duan et al. 2021). However, other samples of the ‘Ferragnés’ variety did not show such values which could lead us to the conclusion that the allergenic potential of this specific variety could be associated with environmental or production factors since that other Mirandela location samples (‘Gloriette’ (2016 Mirandela) and ‘Masbovera’ (Mirandela)) also scored higher allergenicity values in Patient #H. On the opposite side of the scale, it is very difficult to highlight specific varieties with patterns of low allergenicity since that, in the cases where significant differences were found, the bottom groups vary a lot from case to case. Yet is fair to point out that some consistency was found in ‘Amendoão’ that had low to medium-low range in all the patients and

'Refego' (2016 Moncorvo) that was included in the group C in Patients #C, #E, #F and #G cases, representing four of the five patients where significant differences were found.

Differences in the sera' immune responses to specific proteins can be noticed (Figure 8). First, is common among all sera two major proteins/allergens around the 25-30 and 40 kDa, most probably correspondent to Pru du 6 subunits, a major almond allergen (De Angelis et al. 2018). Looking to general results, is clear that the reaction profiles are more dependent of the patient sera than the variety itself. This means that, the same serum reacts to the same proteins across all varieties despite some of them triggered stronger reaction than others. In contrast, the same sample hardly presented the same profile across all patients. Two-dimensional electrophoresis (2DE) analysis was performed to further disclose some proteins that may not be distinguished in the one-dimensional analysis concerning the identification of specific allergens. Pru du 8 and Pru du 10 were only recently considered as almond allergens and there is a lack of information about these allergens in the protein profile of almond and how common reaction to them is. The presence of protein spots around 31 kDa and 60 kDa could represent the reaction to these two proteins, respectively. Moreover, according to the literature, Pru du 1 is the only protein with a molecular weight of 17 kDa and Patient #F constantly presents a band around that value (Fernandes et al. 2013). This could indicate a reaction to Pru du 1 that is not included in the WHO-IUIS list and could be recognized as potential allergen, besides the fact that one reaction in a ten individuals' group could not be considered as significant. Just for context, for an allergen be recognized by the WHO/IUS Allergen Nomenclature Sub-committee, the supposed allergen must react to at least five sera of patients allergic to the food source and not to those without allergy (WHO-IIUIS Allergen Nomenclature Sub-Committee). Anyway, it should be noted that studies with a larger allergic population are already planned in order to better elucidate this potential almond allergen not yet described in the WHO-IUIS list.




C
MW (kDa)
98
62
49
38
28
17
14
#A #B #C #D #E #F #G #H #I #J
D
MW (kDa)
98
62
49
38
28
17
14
#A #B #C #D #E #F #G #H #I #J
E
MW (kDa)
98
62
49
38
28
17
14
#A #B #C #D #E #F #G #H #I #J
F
MW (kDa)
98
62
49
38
28
17
14
#A #B #C #D #E #F #G #H #I #J
G
MW (kDa)
98
62
49
38
28
17
14
#A #B #C #D #E #F #G #H #I #J
H
MW (kDa)
98
62
49
38
28
17
14
#A #B #C #D #E #F #G #H #I #J
I
MW (kDa)
98
62
49
38
28
17
14
#A #B #C #D #E #F #G #H #I #J
J
MW (kDa)
98
62
49
38
28
17
14
#A #B #C #D #E #F #G #H #I #J


K
L
M
N
O
P
Q
R
MW (kDa)
98
62
49
38
28
17
14
#A
#B
#C
#D
#E
#F
#G
#H
#I
#J

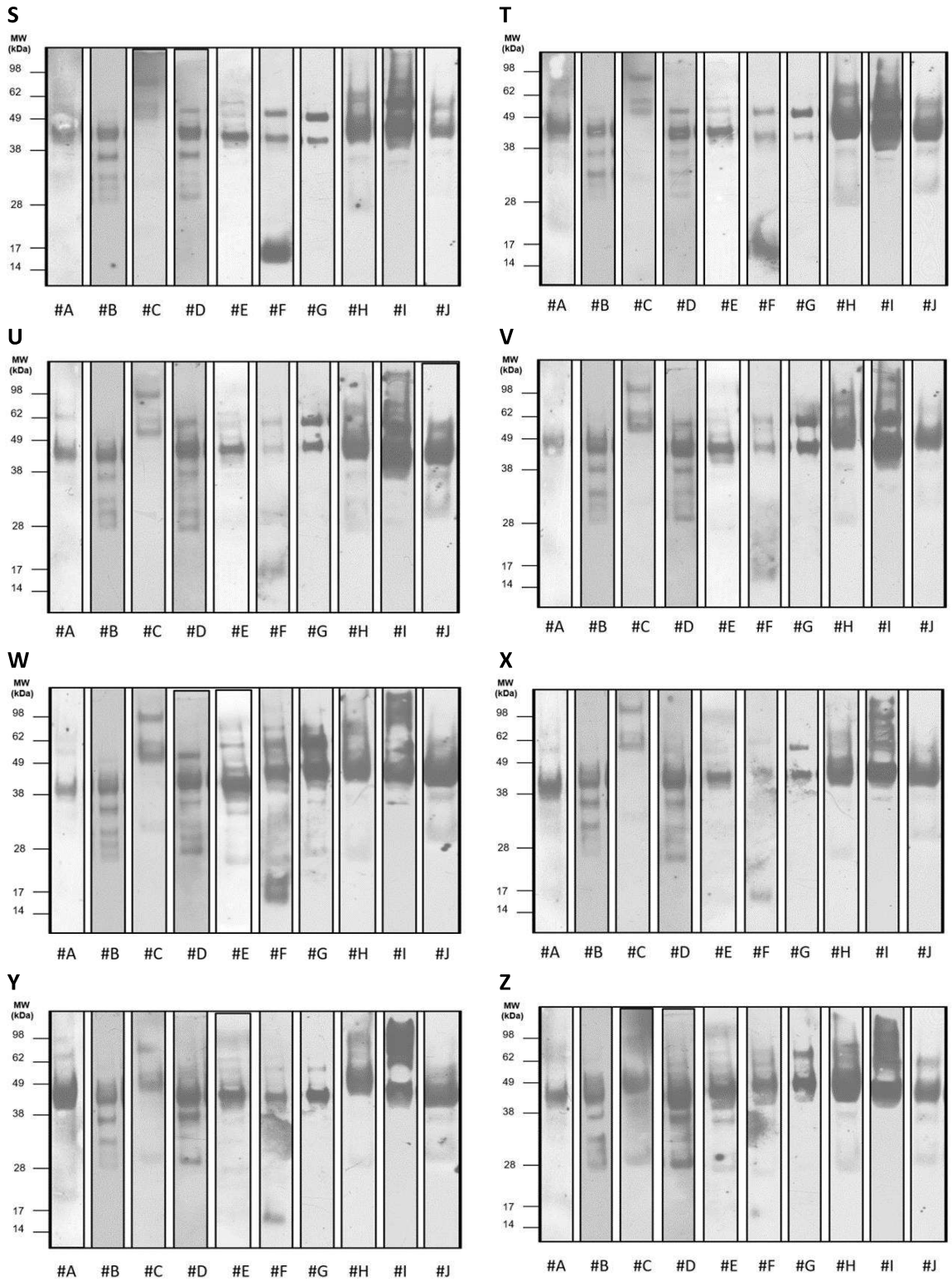


Figure 8 Immune response of patients #A to #J to each variety tested, A - Amendoão, B - Bonita, C - Casanova, D - Casanova, E - Casanova, F - Ferraduel, G - Ferraduel, H - Ferragnés, I - Ferragnés, J - Ferragnés, K - Ferragnés, L - Ferrastar, M - Gloriette, N - Gloriette, O - Gloriette, P - Gloriette, Q - Masbovera, R - Masbovera, S - Marcelina, T - Marcona, U - Orelha de mula, V - Pegarinhos, W - Pegarinhos, X - Pegarinhos, Y - Refego, Z – Refego, see Table 4.

A 2DE-based immunoblotting using the variety 'Casanova' (Moncorvo) protein extract and a serum pool of the ten allergic patients enrolled in the studied was performed. The choice

of this variety is due to the fact that it showed an important nutritional profile in a previous study, and therefore be a great almond variety for the food industry (Oliveira et al. 2018). As seen in Figure 9, almond has a complex proteome where some signs of the presence of proteins with identical isoelectric points (pI) and molecular weights were noticed as previously described by (Ribeiro et al. 2020). As seen in the X-Ray film, most of the proteins reacted with the pooled serum. More specifically, the reactive proteins were located in the 36-120 kDa range and alongside all the pI range (3-10). This comparison shows the complexity of the almond allergy because almond allergens represent most of its total proteome and there is still a long way to fully characterize all these proteins so that new strategies and therapies could be developed in order to properly fight this problematic. The spots obtained in the 2D analyses were extracted from the gel for protein identification, however, due to the COVID-19 pandemic, those results were delayed and were not available at the time of the master's thesis writing.

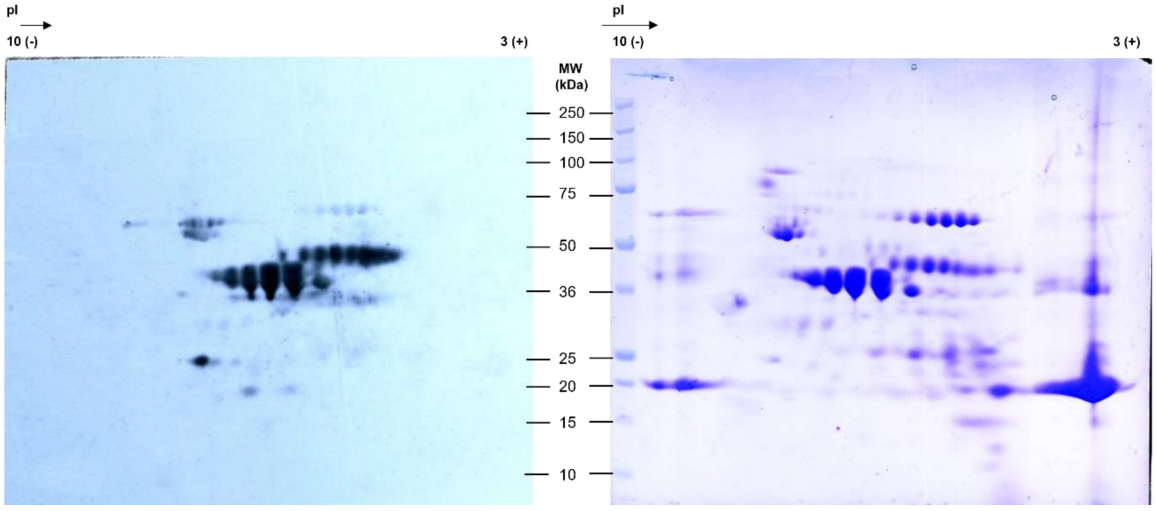


Figure 9 Immune response of patients #A to #J to variety 'Casanova' (Moncorvo) protein extract, see Table 4 for more information about this variety.

4.3. Genomic approach

PCR assays were performed with the intention of developing a genomic approach for the detection and quantification of specific almond allergens, namely for Pru du 6 and Pru du 3 allergens.

As seen in table 8, DNA was successful extracted from almond flour and it was obtained with considerable yield and purity, which showed that the extraction protocol is well suited for almond DNA.

Table 8 Quantification results of samples DNA extraction. 1 - Orelha de Mula, 2 - Pegarinhos 1 grão, 3 - Pegarinhos 2grão, 4 - Ferrastar, 5 - Ferragnés, 6 - Ferraduel, 7 - Casanova, 8 - Gloriette, 9 - Refego, 10 - Marcelina, 11 - Marcona, 12 - Duro Italiano, 13 - Gloriette , 14 - Masbosvera, 15 - Ferragnés, 16 - Casanova, 17 – Pegarinhos. See Table 5 for more information on each samples.

Sample	260	280	260/280	ng/µL	Mean ng/µL
1 - Orelha de Mula	0,018	0,01	1,816	17,995	17,69±0,42
	0,017	0,011	1,557	17,39	
2 - Pegarinhos	0,024	0,012	2,016	24,019	23,32±0,98
	0,023	0,012	1,847	22,625	
3 – Pegarinhos	0,016	0,008	2,024	16,024	16,26±0,34
	0,017	0,008	2,024	16,512	
4 - Ferrastar	0,022	0,011	2,017	21,977	21,33±0,90
	0,021	0,011	1,833	20,699	
5 - Ferragnés	0,02	0,01	2,019	19,882	21,31±2,01
	0,023	0,013	1,704	22,738	
6 - Ferraduel	0,027	0,012	2,268	26,745	26,71±0,04
	0,027	0,013	2,015	26,685	
7 - Casanova	0,023	0,012	1,932	22,581	23,49±1,28
	0,024	0,016	1,508	24,401	
8 - Gloriette	0,021	0,011	1,925	20,523	19,85±0,93
	0,019	0,009	2,134	19,195	
9 - Refego	0,024	0,013	1,859	24,014	25,32±1,84
	0,027	0,016	1,634	26,63	
10 - Marcelina	0,008	0,004	2,048	7,963	7,54±0,59
	0,007	0,007	1,009	7,123	
11 - Marcona	0,008	0,003	2,761	7,965	7,04±1,30
	0,006	0,002	3,167	6,124	
12 - Duro Italiano	0,014	0,008	1,77	13,963	13,17±1,11
	0,012	0,006	2,032	12,394	
13 - Gloriette	0,016	0,007	2,318	15,893	16,20±0,44
	0,017	0,01	1,614	16,519	
15 - Ferragnés	0,019	0,01	1,918	18,802	20,68±2,66
	0,023	0,009	2,472	22,569	
15 - Ferragnés	0,04	0,02	2,06	40,301	40,50±0,28
	0,041	0,018	2,234	40,711	
16 - Casanova	0,01	0,004	2,565	9,741	9,40±0,48
	0,009	0,005	1,833	9,059	
17 – Pegarinhos	0,014	0,007	2,027	14,013	14,74±1,03
	0,015	0,008	1,897	15,476	

In figure 10 is represented the electrophoretic run in agarose gel using Prd6 primer set for the detection of Pru du 6 gene. The fragment presented the expected size of 121 bp corresponding to *P. dulcis* clone 276NPL prunin gene encoding for the Pru du 6 allergen (Costa et al. 2013). As expected Pru du 6 was present across all of almond samples. Moreover, this protocol revealed to be well suited for the detection of this important almond allergen.

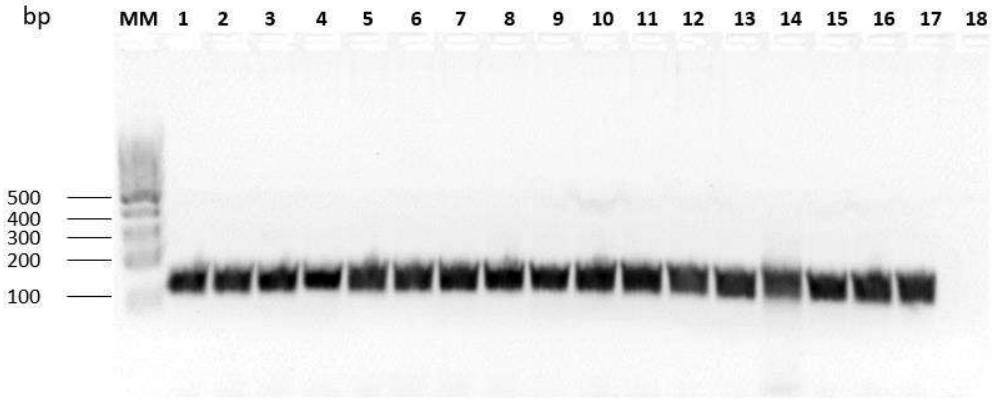


Figure 10 Digital image of the PCR assay for Prd6 primer set. From left to right, the sample that was depoisited: MM – Molecular Marker, 1 - Orelha de Mula, 2 - Pegarinhos 1 grão, 3 - Pegarinhos 2grão, 4 - Ferrastar, 5 - Ferragnés, 6 - Ferraduel, 7 - Casanova, 8 - Gloriette, 9 - Refego, 10 - Marcelina, 11 - Marcona, 12 - Duro Italiano, 13 - Gloriette , 14 - Masbosvera, 15 - Ferragnés, 16 - Casanova, 17 – Pegarinhos and 18 – Non template control. See Table 5 for more information on each samples.

For the optimization of PCR conditions for the detection of Pru du 3 gene, both sets of primers (Prd3.1 in Figure 11A and Prd3.2 in Figure 11B) amplified the pretended fragment of 121 bp. The Prd3.1 set performed better across the different dilutions, especially using 63ºC in the reaction conditions (Figure 11A). Therefore, this set was preferred for the development of detection protocols targeting this gene.

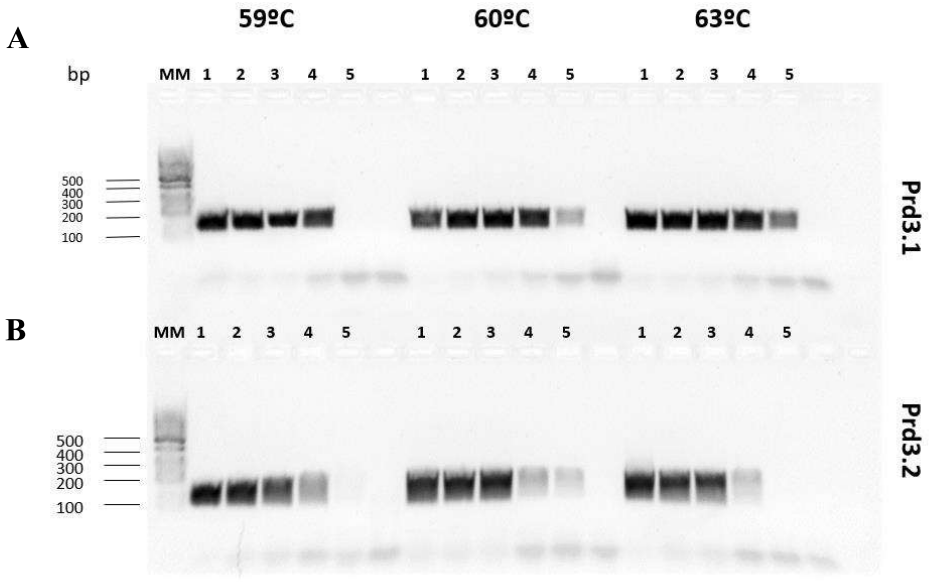


Figure 11 Digital image of the PCR assay for the conditions optimization for Prd3.1 (Figure 11A) and Prd3.2 (Figure 11B) primers sets. MM represents the Molecular Marker and in both sets was used sample 13 – 'Gloriette' diluted five times with 1 being the most concentrated sample and 5 the lowest one. All temperatures and both sets were performed using a non-template control for proper validation

With these results we showed that basic molecular techniques are useful for the detection of almond DNA encoding Pru du 6 and Pru du 3 allergens. These proteins, among all the almond allergens are known for developing the most harmful consequences to almond allergic patients. Therefore, studies approaching the detection of allergens in foods by DNA methods are needed and the development of protocols to assess their presence, even indirectly are of great importance. This genomic part, which is still preliminary, serves as a complement to the proteomic results already obtained, and subsequent studies in the field of transcriptomics will still be conducted with a view to detecting and quantifying important transcripts for the onset of almond allergy.

Chapter 5: Conclusion

Tree nuts are getting a lot of attention in the last decade and, in the last few years, almond has been a preference among the group either for raw production or almond-based food products, most of them to be consumed as alternatives to classical foods like milk, for example. With that growing interest, the allergenicity concern has been rising among producers, consumers, regulation institutions and the scientific community.

In this work, the protein profiles of different almond varieties presented very slight differences, nevertheless with a bioinformatic-aided image analysis was possible to achieve a significant clustering of the population under study. On the other hand, different IgE-binding properties were noticed. For instance, 'Ferragnés' variety of the Mirandela region presented the highest IgE reactivity value for the patients sera #E, #F and #G, in contrast with 'Amendoão' and 'Refego' of the Moncorvo region that presented the lowest values for the patients sera #C, #E and #F, and also #G just for the 'Moncorvo' variety. This kind of analysis allowed a good, wide perspective about the allergenic potential of several almond varieties; however, further studies are required.

Furthermore, environmental factors could take some influence into these results, once that 'Ferragnés' samples from different regions presented relative lower IgE reactiveness values when compared to the Mirandela. This seems to indicate that the allergenic potential is affected by edaphoclimatic factors. In order to confirm this conclusion and properly assess this kind of influence, it would be interesting to perform wider studies using samples of the same varieties from different regions/years.

In general, taking these results into consideration, it does not seem that some variety presents hypoallergenic properties. In the meantime, it sure looks like that dietary avoidance continues to be the key strategy to properly avoid the risks of allergic responses to almond. So, as further perspectives to these results, individual protein identification would be useful to determine if there is any protein causing immune responses that has not been described before. On the other hand, genomic and transcriptomic data, allied with the proteomics data, could lead to a better understanding of the metabolic pathways and intrinsic characteristics of each variety and direct field research to understand if the environmental conditions and the production habits of the regions studied have, in fact, any significant impact on the allergenic potential of the different varieties.

Chapter 6 : References

Abou Alhasani M, Roux KH. 2009. cDNA Cloning, expression and characterization of an allergenic 60s ribosomal protein of almond (Prunus dulcis).

Alasalvar C, Shahidi F. 2008. *Tree nuts: composition, phytochemicals, and health effects*. CRC press.

Albillos SM, Jin T, Howard A, Zhang Y, Kothary MH, Fu T-J. 2008. Purification, crystallization and preliminary X-ray characterization of prunin-1, a major component of the almond (Prunus dulcis) allergen amandin. *Journal of agricultural and food chemistry* **56**: 5352-5358.

Albillos SM, Menhart N, Fu T-J. 2009. Structural stability of amandin, a major allergen from almond (Prunus dulcis), and its acidic and basic polypeptides. *Journal of agricultural and food chemistry* **57**: 4698-4705.

Albin S, Nowak-Węgrzyn A. 2015. Potential treatments for food allergy. *Immunology and allergy clinics of North America* **35**: 77.

Asero R. 2011. 7 Lipid Transfer Protein Cross-reactivity Assessed In Vivo and In Vitro in the Office: Pros and Cons. *Journal of Investigational Allergology and Clinical Immunology* **21**: 129.

Ben-Shoshan M, Turnbull E, Clarke A. 2012. Food allergy: temporal trends and determinants. *Current allergy and asthma reports* **12**: 346-372.

Berryman CE, Preston AG, Karmally W, Deckelbaum RJ, Kris-Etherton PM. 2011. Effects of almond consumption on the reduction of LDL-cholesterol: a discussion of potential mechanisms and future research directions. *Nutrition reviews* **69**: 171-185.

Bezerra M, Ribeiro M, Igrejas G. 2021. An Updated Overview of Almond Allergens. Nutrients 13.

Bignardi C, Elviri L, Penna A, Careri M, Mangia A. 2010. Particle-packed column versus silica-based monolithic column for liquid chromatography–electrospray-linear ion trap-tandem mass spectrometry multiallergen trace analysis in foods. *Journal of Chromatography A* **1217**: 7579-7585.

Bolling BW, Dolnikowski G, Blumberg JB, Chen C-YO. 2010. Polyphenol content and antioxidant activity of California almonds depend on cultivar and harvest year. *Food Chemistry* **122**: 819-825.

Bottone A, Montoro P, Masullo M, Pizza C, Piacente S. 2018. Metabolomics and antioxidant activity of the leaves of Prunus dulcis Mill.(Italian cvs. Toritto and Avola). *Journal of Pharmaceutical and Biomedical Analysis* **158**: 54-65.

Boyce JA, Assa'ad A, Burks AW, Jones SM, Sampson HA, Wood RA, Plaut M, Cooper SF, Fenton MJ, Arshad SH. 2011. Guidelines for the diagnosis and management of food allergy in the United States: summary of the NIAID-sponsored expert panel report. *Journal of the American Academy of Dermatology* **64**: 175-192.

Brugman E, Meulmeester J, Spee-Van der Wekke A, Beuker R, Radder J, Verloove-Vanhorick S. 1998. Prevalence of self-reported food hypersensitivity among school children in The Netherlands. *European journal of clinical nutrition* **52**: 577-581.

Buchanan R, Dennis S, Gendel S, Acheson D, Assimon SA, Beru N, Bolger P, Carlson D, Carvajal R, Copp C. 2008. Approaches to establish thresholds for major food allergens and for gluten in food. *Journal of Food Protection* **71**: 1043.

Buhler S, Tedeschi T, Faccini A, Garino C, Arlorio M, Dossena A, Sforza S. 2015. Isolation and full characterisation of a potentially allergenic lipid transfer protein (LTP) in almond. *Food Additives & Contaminants: Part A* **32**: 648-656.

Burks AW, Williams LW, Helm RM, Connaughton C, Cockrell G, O'Brien T. 1991. Identification of a major peanut allergen, Ara h I, in patients with atopic dermatitis and positive peanut challenges. *Journal of Allergy and Clinical Immunology* **88**: 172-179.

Burks Jr AW, Brooks JR, Sampson HA. 1988. Allergenicity of major component proteins of soybean determined by enzyme-linked immunosorbent assay (ELISA) and immunoblotting in children with atopic dermatitis and positive soy challenges. *Journal of Allergy and Clinical Immunology* **81**: 1135-1142.

Burton OT, Darling AR, Zhou JS, Noval-Rivas M, Jones TG, Gurish MF, Chatila TA, Oettgen HC. 2013. Direct effects of IL-4 on mast cells drive their intestinal expansion and increase susceptibility to anaphylaxis in a murine model of food allergy. *Mucosal immunology* **6**: 740-750.

Che H, Zhang Y, Jiang S, Jin T, Lyu S-C, Nadeau KC, McHugh T. 2019. Almond (Prunus dulcis) allergen Pru du 8, the first member of a new family of food allergens. *Journal of agricultural and food chemistry* **67**: 8626-8631.

Chen L, Zhang S, Illa E, Song L, Wu S, Howad W, Arús P, Van de Weg E, Chen K, Gao Z. 2008. Genomic characterization of putative allergen genes in peach/almond and their synteny with apple. *BMC genomics* **9**: 543.

Clayton F, Fang JC, Gleich GJ, Lucendo AJ, Olalla JM, Vinson LA, Lowichik A, Chen X, Emerson L, Cox K. 2014. Eosinophilic esophagitis in adults is associated with IgG4 and not mediated by IgE. *Gastroenterology* **147**: 602-609.

Clemente A, Chambers SJ, Lodi F, Nicoletti C, Brett GM. 2004. Use of the indirect competitive ELISA for the detection of Brazil nut in food products. *Food Control* **15**: 65-69.

Costa J, Mafra I, Carrapatoso I, Oliveira MBP. 2012. Almond allergens: molecular characterization, detection, and clinical relevance. *Journal of agricultural and food chemistry* **60**: 1337-1349.

Costa J, Oliveira M, Mafra IJC, Allergy T. 2013. Applicability of a real-time PCR system to verify labelling compliance of nut allergens in chocolates. **3**: 1-1.

Costa J, Oliveira MBP, Mafra IJFri. 2013. Novel approach based on single-tube nested real-time PCR to detect almond allergens in foods. 51: 228-235.

De Angelis E, Bavaro SL, Forte G, Pilolli R, Monaci L. 2018. Heat and pressure treatments on almond protein stability and change in immunoreactivity after simulated human digestion. *Nutrients* **10**: 1679.

De Giorgio D, Leo L, Zacheo G, Lamascese N. 2007. Evaluation of 52 almond (Prunus amygdalus Batsch) cultivars from the Apulia region in Southern Italy. *The Journal of Horticultural Science and Biotechnology* **82**: 541-546.

Derr LE. 2006. When food is poison: the history, consequences, and limitations of the Food Allergen Labeling and Consumer Protection Act of 2004. *Food and drug law journal* **61**: 65-165.

Duan L, Celik A, Hoang JA, Schmidthaler K, So D, Yin X, Ditlof CM, Ponce M, Upton JE, Lee JSJA. 2021. Basophil activation test shows high accuracy in the diagnosis of peanut and tree nut allergy: the markers of nut allergy study. 76: 1800-1812.

du Toit G, Tsakok T, Lack S, Lack G. 2016. Prevention of food allergy. *Journal of Allergy and Clinical Immunology* **137**: 998-1010.

Egger M, Hauser M, Mari A, Ferreira F, Gadermaier G. 2010. The role of lipid transfer proteins in allergic diseases. *Current allergy and asthma reports* **10**: 326-335.

Ewan PW. 1996. Clinical study of peanut and nut allergy in 62 consecutive patients: new features and associations. *Bmj* **312**: 1074-1078.

FAOSTAT. 2020. Food and Agriculture Organization of the United Nations Acessed at september 2020; Online available at http://www.fao.org/faostat/en/#home.

Fernandes H, Michalska K, Sikorski M, Jaskolski M. 2013. Structural and functional aspects of PR-10 proteins. *The FEBS journal* **280**: 1169-1199.

Flinterman AE, Hoekstra MO, Meijer Y, van Ree R, Akkerdaas JH, Bruijnzeel-Koomen CA, Knulst AC, Pasmans SG. 2006. Clinical reactivity to hazelnut in children: association with sensitization to birch pollen or nuts? *Journal of allergy and clinical immunology* **118**: 1186.

Garcia-Mas J, Messeguer R, Arús P, Puigdomènech PJPmb. 1995. Molecular characterization of cDNAs corresponding to genes expressed during almond (Prunus amygdalus Batsch) seed development. **27**: 205-210.

Garino C, De Paolis A, Coïsson JD, Arlorio M. 2015. Pru du 2S albumin or Pru du vicilin? *Computational Biology and Chemistry* **56**: 30-32.

Gendel SM. 2012. Comparison of international food allergen labeling regulations. *Regulatory Toxicology and Pharmacology* **63**: 279-285.

Green PH, Lebwohl B, Greywoode R. 2015. Celiac disease. *Journal of Allergy and Clinical Immunology* **135**: 1099-1106.

Griel AE, Kris-Etherton PM. 2006. Tree nuts and the lipid profile: a review of clinical studies. *British Journal of Nutrition* **96**: S68-S78.

Gupta RS, Springston EE, Warrier MR, Smith B, Kumar R, Pongracic J, Holl JL. 2011. The prevalence, severity, and distribution of childhood food allergy in the United States. *Pediatrics* **128**: e9-e17.

Gupta RS, Warren CM, Smith BM, Jiang J, Blumenstock JA, Davis MM, Schleimer RP, Nadeau KC. 2019. Prevalence and severity of food allergies among US adults. *JAMA network open* **2**: e185630-e185630.

Hasegawa M, Inomata N, Yamazaki H, Morita A, Kirino M, Ikezawa Z. 2009. Clinical features of four cases with cashew nut allergy and cross-reactivity between cashew nut and pistachio. *Allergology International* **58**: 209-215.

Hoffmann-Sommergruber K. 2016. Proteomics and its impact on food allergy diagnosis. *EuPA open proteomics* **12**: 10-12.

Holden L, Sletten GB, Lindvik H, Fæste CK, Dooper MM. 2008. Characterization of IgE binding to lupin, peanut and almond with sera from lupin-allergic patients. *International archives of allergy and immunology* **146**: 267-276.

Jackson LS, Al-Taher FM, Moorman M, DeVRIES JW, Tippett R, Swanson KM, Fu T-J, Salter R, Dunaif G, Estes S. 2008. Cleaning and other control and validation strategies to prevent allergen cross-contact in food-processing operations. *Journal of Food Protection* **71**: 445-458.

Jenkins DJ, Hu FB, Tapsell LC, Josse AR, Kendall CW. 2008. Possible benefit of nuts in type 2 diabetes. *The Journal of nutrition* **138**: 1752S-1756S.

Johnson PE, Baumgartner S, Aldick T, Bessant C, Giosafatto V, Heick J, Mamone G, O'connor G, Poms R, Popping B. 2011. Current perspectives and recommendations for the development of mass spectrometry methods for the determination of allergens in foods. *Journal of AOAC International* **94**: 1026-1033.

Kewalramani A, Maleki S, Cheng H, Teuber S. 2006. Cross-reactivity among almond, peanut and other tree nuts in almond allergic patients. *Journal of Allergy and Clinical Immunology* **117**: S32.

Kolivas S, Gayler KR. 1993. Structure of the cDNA coding for conglutin γ, a sulphur-rich protein from Lupinus angustifolius. *Plant molecular biology* **21**: 397-401.

Koppelman S, Wensing M, Ertmann M, Knulst A, Knol E. 2004. Relevance of Ara h1, Ara h2 and Ara h3 in peanut-allergic patients, as determined by immunoglobulin E Western blotting, basophil–histamine release and intracutaneous testing: Ara h2 is the most important peanut allergen. *Clinical & Experimental Allergy* **34**: 583-590.

Kozlowska A, Szostak-Wegierek D. 2014. Flavonoids-food sources and health benefits. *Roczniki Państwowego Zakładu Higieny* **65**.

Kulis M, Vickery BP, Burks AW. 2011. Pioneering immunotherapy for food allergy: clinical outcomes and modulation of the immune response. *Immunologic research* **49**: 216-226.

Kumfer AM, Commins SP. 2019. Primary prevention of food allergy. *Current allergy and asthma reports* **19**: 7.

Lee J-BJIn. 2016. Regulation of IgE-mediated food allergy by IL-9 producing mucosal mast cells and type 2 innate lymphoid cells. **16**: 211-218.

Liu J-J, Sturrock R, Ekramoddoullah AK. 2010. The superfamily of thaumatin-like proteins: its origin, evolution, and expression towards biological function. *Plant cell reports* **29**: 419-436.

Lupinek C, Wollmann E, Baar A, Banerjee S, Breiteneder H, Broecker BM, Bublin M, Curin M, Flicker S, Garmatiuk T. 2014. Advances in allergen-microarray technology for diagnosis and monitoring of allergy: the MeDALL allergen-chip. *Methods* **66**: 106-119.

Luyt DK, Vaughan D, Oyewole E, Stiefel G. 2016. Ethnic differences in prevalence of cashew nut, pistachio nut and almond allergy. *Pediatric allergy and immunology: official publication of the European Society of Pediatric Allergy and Immunology* **27**: 651-654.

Maleki SJ, Teuber SS, Cheng H, Chen D, Comstock SS, Ruan S, Schein CH. 2011. Computationally predicted IgE epitopes of walnut allergens contribute to cross-reactivity with peanuts. *Allergy* **66**: 1522-1529.

Mandalari G, Genovese T, Bisignano C, Mazzon E, Wickham M, Di Paola R, Bisignano G, Cuzzocrea S. 2011. Neuroprotective effects of almond skins in experimental spinal cord injury. *Clinical Nutrition* **30**: 221-233.

Mandalari G, Mackie AR. 2018. Almond allergy: an overview on prevalence, thresholds, regulations and allergen detection. *Nutrients* **10**: 1706.

Mandalari G, Rigby NM, Bisignano C, Curto RBL, Mulholland F, Su M, Venkatachalam M, Robotham JM, Willison LN, Lapsley K. 2014. Effect of food matrix and processing on release of almond protein during simulated digestion. *LWT-Food Science and Technology* **59**: 439-447.

Mandalari G, Tomaino A, Arcoraci T, Martorana M, Turco VL, Cacciola F, Rich G, Bisignano C, Saija A, Dugo P. 2010. Characterization of polyphenols, lipids and dietary fibre from almond skins (Amygdalus communis L.). *Journal of Food Composition and Analysis* **23**: 166-174.

Marcus JP, Green JL, Goulter KC, Manners JM. 1999. A family of antimicrobial peptides is produced by processing of a 7S globulin protein in Macadamia integrifolia kernels. *The Plant Journal* **19**: 699-710.

Martínez-Gimeno A, Del Castillo P, García-Hernández G, Luna-Paredes C, García-Sánchez JA, Nogales-Espert A. 2000. 388 Prevalence of food allergy/intolerance in children: Results from a population based survey. *Journal of Allergy and Clinical Immunology* **105**: S130.

McWilliam V, Koplin J, Lodge C, Tang M, Dharmage S, Allen K. 2015. The prevalence of tree nut allergy: a systematic review. *Current allergy and asthma reports* **15**: 54.

Mills EC, Sancho A, Moreno J, Kostyra H. 2007. The effects of food processing on allergens. In *Managing allergens in food*, pp. 117-133. Elsevier.

Mishra A, Schlotman J, Wang M, Rothenberg ME. 2007. Critical role for adaptive T cell immunity in experimental eosinophilic esophagitis in mice. *Journal of leukocyte biology* **81**: 916-924.

Mittag D, Akkerdaas J, Ballmer-Weber BK, Vogel L, Wensing M, Becker W-M, Koppelman SJ, Knulst AC, Helbling A, Hefle SL. 2004. Ara h 8, a Bet v 1–homologous allergen

from peanut, is a major allergen in patients with combined birch pollen and peanut allergy. *Journal of Allergy and Clinical Immunology* **114**: 1410-1417.

Monaci L, De Angelis E, Montemurro N, Pilolli R. 2018. Comprehensive overview and recent advances in proteomics MS based methods for food allergens analysis. *TrAC Trends in Analytical Chemistry* **106**: 21-36.

Moore LE, Stewart PH, deShazo RD. 2017. Food allergy: what we know now. *The American journal of the medical sciences* **353**: 353-366.

Moreno FJ, Clemente A. 2008. 2S albumin storage proteins: what makes them food allergens? *The open biochemistry journal* **2**: 16.

Nanos GD, Kazantzis I, Kefalas P, Petrakis C, Stavroulakis GG. 2002. Irrigation and harvest time affect almond kernel quality and composition. *Scientia Horticulturae* **96**: 249-256.

Nicolaou N, Poorafshar M, Murray C, Simpson A, Winell H, Kerry G, Härlin A, Woodcock A, Ahlstedt S, Custovic A. 2010. Allergy or tolerance in children sensitized to peanut: prevalence and differentiation using component-resolved diagnostics. *Journal of Allergy and Clinical Immunology* **125**: 191-197. e113.

Nollet LM, Van Hengel AJ. 2016. *Food allergens: analysis instrumentation and methods*. CRC press.

Noorbakhsh R, Mortazavi SA, Sankian M, Shahidi F, Tehrani M, Azad FJ, Behmanesh F, Varasteh A. 2011. Pistachio allergy-prevalence and in vitro cross-reactivity with other nuts. *Allergology International* **60**: 425-432.

Nowak-Węgrzyn A, Sampson HA. 2011. Future therapies for food allergies. *Journal of allergy and clinical immunology* **127**: 558-573.

Oliveira I, Meyer A, Afonso S, Ribeiro C, Gonçalves BJJofs, technology. 2018. Morphological, mechanical and antioxidant properties of Portuguese almond cultivars. **55**: 467-478.

Orhan F, Karakas T, Cakir M, Aksoy A, Baki A, Gedik Y. 2009. Prevalence of immunoglobulin E-mediated food allergy in 6–9-year-old urban schoolchildren in the eastern Black Sea region of Turkey. *Clinical & Experimental Allergy* **39**: 1027-1035.

Pafundo S, Gulli M, Marmiroli N. 2009. SYBR® GreenER™ Real-Time PCR to detect almond in traces in processed food. *Food Chemistry* **116**: 811-815.

Palacin A, Tordesillas L, Gamboa P, Sanchez-Monge R, Cuesta-Herranz J, Sanz M, Barber D, Salcedo G, Díaz-Perales A. 2010. Characterization of peach thaumatin-like proteins and their identification as major peach allergens. *Clinical & Experimental Allergy* **40**: 1422-1430.

Palmer GW, Dibbern Jr DA, Burks AW, Bannon GA, Bock SA, Porterfield HS, McDermott RA, Dreskin SC. 2005. Comparative potency of Ara h 1 and Ara h 2 in immunochemical and functional assays of allergenicity. *Clinical Immunology* **115**: 302-312.

Panel N-SE. 2010. Guidelines for the diagnosis and management of food allergy in the United States: report of the NIAID-sponsored expert panel. *Journal of Allergy and Clinical Immunology* **126**: S1-S58.

Panjari M, Koplin J, Dharmage S, Peters R, Gurrin L, Sawyer S, McWilliam V, Eckert J, Vicendese D, Erbas B. 2016. Nut allergy prevalence and differences between Asian-born children and Australian-born children of A sian descent: a state-wide survey of children at primary school entry in V ictoria, A ustralia. *Clinical & Experimental Allergy* **46**: 602-609.

Pastorello EA, Vieths S, Pravettoni V, Farioli L, Trambaioli C, Fortunato D, Lüttkopf D, Calamari M, Ansaloni R, Scibilia JJJoa et al. 2002. Identification of hazelnut major allergens in sensitive patients with positive double-blind, placebo-controlled food challenge results. **109**: 563-570.

Pérez-Sánchez R, Morales-Corts MRJA. 2021. Agromorphological Characterization and Nutritional Value of Traditional Almond Cultivars Grown in the Central-Western Iberian Peninsula. 11: 1238.

Perrier C, Corthesy B. 2011. Gut permeability and food allergies. *Clinical & Experimental Allergy* **41**: 20-28.

Piras C, Roncada P, Rodrigues PM, Bonizzi L, Soggiu A. 2016. Proteomics in food: quality, safety, microbes, and allergens. *Proteomics* **16**: 799-815.

Piscopo A, Romeo F, Petrovicova B, Poiana M. 2010. Effect of the harvest time on kernel quality of several almond varieties (Prunus dulcis (Mill.) DA Webb). *Scientia Horticulturae* **125**: 41-46.

Poltronieri P, Cappello M, Dohmae N, Conti A, Fortunato D, Pastorello E, Ortolani C, Zacheo G. 2002. Identification and characterisation of the IgE-binding proteins 2S albumin and conglutin γ in almond (Prunus dulcis) seeds. *International archives of allergy and immunology* **128**: 97-104.

Portugal C, Pinto L, Ribeiro M, Tenorio C, Igrejas G, Ruiz-Larrea FJIjofm. 2015. Potential spoilage yeasts in winery environments: Characterization and proteomic analysis of Trigonopsis cantarellii. **210**: 113-120.

Prescott S, Allen KJ. 2011. Food allergy: riding the second wave of the allergy epidemic. *Pediatric allergy and immunology* **22**: 155-160.

Rehm CD, Drewnowski A. 2017. Replacing American snacks with tree nuts increases consumption of key nutrients among US children and adults: results of an NHANES modeling study. *Nutrition journal* **16**: 17.

Renz H, Allen KJ, Sicherer SH, Sampson HA, Lack G, Beyer K, Oettgen HC. 2018. Food allergy. *Nature reviews Disease primers* **4**: 1-20.

Ribeiro M, Bancel E, Faye A, Dardevet M, Ravel C, Branlard G, Igrejas GJIjoms. 2013. Proteogenomic characterization of novel x-type high molecular weight glutenin subunit 1Ax1. 1. **14**: 5650-5667.

Ribeiro M, Costa J, Mafra I, Cabo S, Silva AP, Gonçalves B, Hillion M, Hébraud M, Igrejas GJN. 2020. Natural Variation of Hazelnut Allergenicity: Is There Any Potential for Selecting Hypoallergenic Varieties? **12**: 2100.

Ribeiro M, Nunes FM, Guedes S, Domingues P, Silva AM, Carrillo JM, Rodriguez-Quijano M, Branlard G, Igrejas GJSr. 2015. Efficient chemo-enzymatic gluten detoxification: reducing toxic epitopes for celiac patients improving functional properties. **5**: 1-17.

Richardson DP, Astrup A, Cocaul A, Ellis P. 2009. The nutritional and health benefits of almonds: a healthy food choice. *Food Science and Technology Bulletin: Functional Foods* **6**: 41-50.

Rohlf FJ, Fisher DRJSB. 1968. Tests for hierarchical structure in random data sets. **17**: 407-412.

Romero-Calvo I, Ocón B, Martínez-Moya P, Suárez MD, Zarzuelo A, Martínez-Augustin O, de Medina FSJAb. 2010. Reversible Ponceau staining as a loading control alternative to actin in Western blots. **401**: 318-320.

Roux KH, Teuber SS, Robotham JM, Sathe SK. 2001. Detection and stability of the major almond allergen in foods. *Journal of agricultural and food chemistry* **49**: 2131-2136.

Roux KH, Teuber SS, Sathe SK. 2003. Tree nut allergens. *International archives of allergy and immunology* **131**: 234-244.

Sampson HA, Mendelson L, Rosen JP. 1992. Fatal and near-fatal anaphylactic reactions to food in children and adolescents. *New England Journal of Medicine* **327**: 380-384.

Sánchez-García S, del Río PR, Escudero C, Martínez-Gómez MJ, Ibáñez MD. 2012. Possible eosinophilic esophagitis induced by milk oral immunotherapy. *Journal of allergy and clinical immunology* **129**: 1155-1157.

Sathe SK, Wolf WJ, Roux KH, Teuber SS, Venkatachalam M, Sze-Tao KWC. 2002. Biochemical characterization of amandin, the major storage protein in almond (Prunus dulcis L.). *Journal of agricultural and food chemistry* **50**: 4333-4341.

Scala E, Alessandri C, Palazzo P, Pomponi D, Liso M, Bernardi ML, Ferrara R, Zennaro D, Santoro M, Rasi C. 2011. IgE recognition patterns of profilin, PR-10, and tropomyosin panallergens tested in 3,113 allergic patients by allergen microarray-based technology. *PloS one* **6**: e24912.

Scarafoni A, Consonni A, Pessina S, Balzaretti S, Capraro J, Galanti E, Duranti M. 2016. Structural basis of the lack of endo-glucanase inhibitory activity of Lupinus albus γ-conglutin. *Plant Physiology and Biochemistry* **99**: 79-85.

Schägger HJNp. 2006. Tricine–sds-page. **1**: 16-22.

Seo K, Lee D, Lee D, Kang H, Kim H, Kim Y, Baek N. 2015. Hepatoprotective and neuroprotective tocopherol analogues isolated from the peels of Citrus unshiu Marcovich. *Natural product research* **29**: 571-573.

Shewry PR, Halford NG. 2002. Cereal seed storage proteins: structures, properties and role in grain utilization. *Journal of experimental botany* **53**: 947-958.

Shewry PR, Napier JA, Tatham AS. 1995. Seed storage proteins: structures and biosynthesis. *The plant cell* **7**: 945.

Sicherer SH, Allen K, Lack G, Taylor SL, Donovan SM, Oria M. 2017. Critical issues in food allergy: a national academies consensus report. *Pediatrics* **140**: e20170194.

Sicherer SH, Muñoz-Furlong A, Sampson HA. 2003. Prevalence of peanut and tree nut allergy in the United States determined by means of a random digit dial telephone survey: a 5-year follow-up study. *Journal of allergy and clinical immunology* **112**: 1203-1207.

Sicherer SH, Sampson HA. 2018. Food allergy: a review and update on epidemiology, pathogenesis, diagnosis, prevention, and management. *Journal of Allergy and Clinical Immunology* **141**: 41-58.

Stemeseder T, Klinglmayr E, Moser S, Lang R, Himly M, Oostingh GJ, Zumbach J, Bathke AC, Hawranek T, Gadermaier GJIaoa et al. 2017. Influence of intrinsic and lifestyle factors on the development of IgE sensitization. **173**: 99-104.

Summo C, Palasciano M, De Angelis D, Paradiso VM, Caponio F, Pasqualone A. 2018. Evaluation of the chemical and nutritional characteristics of almonds (Prunus dulcis (Mill). DA Webb) as influenced by harvest time and cultivar. *Journal of the Science of Food and Agriculture* **98**: 5647-5655.

Tawde P, Venkatesh YP, Wang F, Teuber SS, Sathe SK, Roux KH. 2006. Cloning and characterization of profilin (Pru du 4), a cross-reactive almond (Prunus dulcis) allergen. *Journal of allergy and clinical immunology* **118**: 915-922.

Tian H, Zhang H, Zhan P, Tian F. 2011. Composition and antioxidant and antimicrobial activities of white apricot almond (Amygdalus communis L.) oil. *European Journal of Lipid Science and Technology* **113**: 1138-1144.

Vadas P, Gold M, Perelman B, Liss GM, Lack G, Blyth T, Simons FER, Simons KJ, Cass D, Yeung J. 2008. Platelet-activating factor, PAF acetylhydrolase, and severe anaphylaxis. *New England Journal of Medicine* **358**: 28-35.

van Hengel AJ. 2007. Food allergen detection methods and the challenge to protect food-allergic consumers. *Analytical and bioanalytical chemistry* **389**: 111-118.

Varshney P, Steele PH, Vickery BP, Bird JA, Thyagarajan A, Scurlock AM, Perry TT, Jones SM, Burks AW. 2009. Adverse reactions during peanut oral immunotherapy home dosing. *Journal of Allergy and Clinical Immunology* **124**: 1351-1352.

Venkatachalam M, Teuber SS, Roux K, Sathe S. 2002. Effects of roasting, blanching, autoclaving, and microwave heating on antigenicity of almond (Prunus dulcis L.) proteins. *Journal of agricultural and food chemistry* **50**: 3544-3548.

Vieths S, Scheurer S, BALLMER-WEBER B. 2002. Current understanding of cross-reactivity of food allergens and pollen. *Annals of the New York Academy of Sciences* **964**: 47-68.
Wallowitz M, Teuber S, Beyer K, Sampson H, Roux K, Sathe S, Wang F, Robotham J. 2004. Cross-reactivity of walnut, cashew, and hazelnut legumin proteins in tree nut allergic patients. *Journal of Allergy and Clinical Immunology* **113**: S156.
Wang F, Robotham JM, Teuber SS, Tawde P, Sathe SK, Roux KH. 2002. Ana o 1, a cashew (Anacardium occidental) allergen of the vicilin seed storage protein family. *Journal of Allergy and Clinical Immunology* **110**: 160-166.
Wang X, Young O, Karl D. 2010. Evaluation of cleaning procedures for allergen control in a food industry environment. *Journal of food science* **75**: T149-T155.
WHO-IIUIS Allergen Nomenclature Sub-Committee. Acesses at september 2021. Online available at http://www.allergen.org/
Williams KW, Sharma HP. 2015. Anaphylaxis and urticaria. *Immunology and Allergy Clinics* **35**: 199-219.
Willison LN, Tripathi P, Sharma G, Teuber SS, Sathe SK, Roux KH. 2011. Cloning, expression and patient IgE reactivity of recombinant Pru du 6, an 11S globulin from almond. *International archives of allergy and immunology* **156**: 267-281.
Yao L, Li H, Yang J, Li C, Shen Y. 2018. Purification and characterization of a hydroxynitrile lyase from Amygdalus pedunculata Pall. *International journal of biological macromolecules* **118**: 189-194.
Young E, Stoneham MD, Petruckevitch A, Barton J, Rona R. 1994. A population study of food intolerance. *The lancet* **343**: 1127-1130.
Yu W, Freeland DMH, Nadeau KC. 2016. Food allergy: immune mechanisms, diagnosis and immunotherapy. *Nature Reviews Immunology* **16**: 751.
Yunginger JW, Sweeney KG, Sturner WQ, Giannandrea LA, Teigland JD, Bray M, Benson PA, York JA, Biedrzycki L, Squillace DL. 1988. Fatal food-induced anaphylaxis. *Jama* **260**: 1450-1452.
Zhang Y, Jin T. 2020. Almond allergens: update and perspective on identification and characterization. *Journal of the Science of Food and Agriculture*.