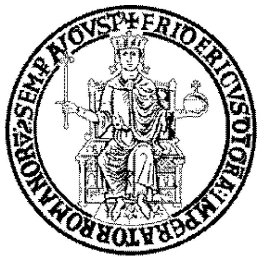
UNIVERSITÁ DI NAPOLI FEDERICO II

DIPARTIMENTO DI AGRARIA



DOTTORATO IN FOOD SCIENCE

XXXIV CICLO

Proteomic characterization and definition of the potential allergenicity of sustainable novel food

Tutor:
Ch.mo Prof. Pasquale Ferranti

Candidato:
Giovanni D'Auria

Co-tutor:
Dr. Gianfranco Mamone

Coordinatore: Ch.ma Prof.ssa Amalia Barone

2020-2021

To my family and Giulia

“The role of the infinitely small in nature is infinitely great.”
— Louis Pasteur

“An experiment is a question which science poses to Nature and a measurement is the recording of Nature’s answer.”

— Max Planck

INDEX

Overview of complete PhD training activities and list of publication

List of publications

Fully related to PhD thesis

D'Auria, G., Nitride, C., Ferranti, P. 2023. Moringa oleifera Lam. proteins-properties and food applications. Chapter in *Reference Module in Food Science*; Elsevier, ISBN 9780081005965, https://doi.org/10.1016/B978-0-12-823960-5.00005-6.

D'Auria, G., Nitride, C., Nicolai, M.A., Mamone, G., Montesano, D., Mills, E.N.C., Ferranti, P. Identification of potential allergens in a novel food ingredient from Moringa oleifera leaves. *Submitted to Food Chemistry 2022.*

D'Auria, G.*, Wang, K.*, Nitride, C., Mamone, G., Montesano, D., Ferranti, P., Mills, E.N.C. A follow-up allergenicity risk assessment using specific sera screening on Moringa oleifera leaf proteins after bioinformatic prediction. *In preparation.*

Landolfi, V.*, **D'Auria, G.***, Nicolai, M. A., Nitride, C., Blandino, M., Ferranti, P. 2021. The effect of nitrogen fertilization on the expression of proteins in wheat and tritordeum varieties using a proteomic approach, *Food Research International* 148, 110617.

Nitride, C., **D'Auria, G.§,** Dente, A., Landolfi, V., Picariello, G., Mamone, G., Blandino, M., Romano, R., Ferranti, P. 2022. Tritordeum as an Innovative Alternative to Wheat: A Comparative Digestion Study on Bread. *Molecules* 27(4), 1308. https://doi.org/10.3390/molecules27041308.

Pourjoula, M., Picariello, G., Garro, G., **D'Auria, G.**, Nitride, C., Rheza Ghaisari, A., Ferranti, P. 2020. The protein and peptide fractions of kashk, a traditional Middle East fermented dairy product, *Food Research International* 132, 109107.

Folliero, V., Lama, S., Franci, G., Giugliano, R., **D'Auria, G.**, Ferranti, P., Pourjula, M., Galdiero, M., Stiuso, P. 2022. Casein-derived peptides from the dairy product kashk exhibit wound healing and antibacterial properties: structural and functional characterization. *Food Research International* 153, 110949.

* These authors contributed equally

§ Corresponding author

Partly related to the PhD thesis

D'Auria, G., 2019. Microalgal Biomass Recycling: From Filter to Feed. Chapter in "Reference Module in Food Science", Elsevier, http://dx.doi.org/10.1016/B978-0-08-100596-5.22658-9

D'Auria, G., Nitride, C., Ferranti, P., Microalgae to contrast the climate change: a novel food and feed ingredient with technological applications. Chapter in Reference Module in Food Science; Elsevier. https://doi.org/10.1016/B978-0-12-823960-5.00024-X

Communications in conferences and posters

D'Auria G. A full proteomic profiling of Moringa oleifera leaves for application as a novel food. First Virtual Workshop on the Developments in the Italian PhD Research on Food Science Technology and Biotechnology, University of Palermo, September 14th-15th, 2021.

D'Auria, G., Mamone, G., Nicolai, M.A., Nitride, C., Montesano, D., Ferranti, P. Moringa oleifera leaves as a novel food ingredient: digestibility and allergenicity, 6Th Foodomics Conference, 14-16 October 2020, Cesena, Italy.

Scientific mission

The University of Manchester: Division of Infection, Immunity & Respiratory Medicine| School of Biological Sciences I Faculty of Biology, Medicine and Health, Manchester Institute of Biotechnology, John Garside Building 131 Princess Street, Manchester M1 7DN, England.

Project: In vitro allergenicity assessment of novel plant-based protein ingredients

Supervisor: Prof. Clare Mills

From 27/08/2021 to 27/11/2021

Participation in research project

PRIN: PROGETTI DI RICERCA DI RILEVANTE INTERESSE NAZIONALE – Bando 2015 Prot. 2015SSEKFL Research Project Title PROCESSING FOR HEALTHY CEREAL FOODS (end May 2020)

This page intentionally left blank

ABSTRACT

The world population is foreseen to grow up to 9.8 billion of people by the next thirty years, reflecting in an increasing demand of alternative and sustainable sources of food proteins. The European Union (EU) requires a complete assessment of safety use of novel foods before their placement in the market. This assessment incudes the evaluation of the allergenic potential of the ingredient using a weight-of-evidence approach. To develop and validate the approach in this thesis three models of novel food were chosen as case studies: two of vegetal origin and one of animal origin.

SECTION 1 – A novel food of plant origin: *Moringa oleifera* leaf

Moringa oleifera leaves are a source of proteins with high biological value, considered a novel food in the EU and UK. In this section, the allergenic potential of *Moringa oleifera* leaves has been addressed. The leaves are a traditional medicine in Asia where is used as a *panacea* to treat and prevent several diseases, alongside being part of the oriental diet both as a raw and cooking ingredient. The ability of the plant to resist unfavorable environmental conditions, such as drought and high temperatures, makes it a sustainable agricultural choice for the future. However, the high content of protein trips the alarm on their possible allergenicity. C**hapter 1** is a comprehensive review of the literature, to provide nutritional, technical and safety information on the proteinaceous tissues of Moringa (*i.e.,* seeds and leaves). **Chapter 2** is the proteomic-based characterization of the Moringa leaf powder ingredient. The majority of the identified proteins belonged to photosynthetic and metabolic pathways. *In-silico* analysis of the leaf proteome highlighted Moritides (mO1 and mO2) as potential cross-allergens of the hevein (Hev b 6.02), a latex allergen implicated in the latex-fruit syndrome. The non-specific lipid transfer protein

(nsLTP), a major panallergen in food, was also identified in the leaf, and its identification was confirmed by *de-novo* sequencing.

Following up on the proteomic and *in-silico* characterization, **chapter 3** is a confirmatory *in-vitro* study of the cross-allergenicity predicted in chapter 2. Moringa protein extracts were immune-based assayed using animal IgG anti-Pru p 3 (the nsLTP of peach's skin) and using human sera IgE either from latex or nsLTP allergic patients. The immunoblotting performed with latex allergic sera showed no specific immunoreactivity to Hev b 6.02. The removal of protein glycosylations in Moringa extracts reduced spurious IgE response. Animal IgG did not show Pru p 3 homologues. Conversely, human sera IgE analyses reproducibly highlighted an immunoreactive band in patients allergic to nsLTP. Immune-based assays revealed the imperfection of the current approaches for allergenicity risk assessment and the importance of a weight-of-evidence approach.

SECTION 2 – An innovative hybrid cereal: *Tritordeum*

Tritordeum martinii is a hexaploid (AABBHchHch) hybrid derived from crossbreeding *Hordeum chilense* with durum wheat. The grain derived flour has protein contents and bread-making properties comparable to soft wheat.

In chapter 4 the proteomic characterization of tritordeum Bulel cv was obtained for the first time in comparison with two controls flours of *Triticum aestivum*: a modern grain, Bologna cv, and a landrace grain Andriolo cv. The INFOGEST model was used to simulate i*n-vitro* digestion of the flours. The digestion products were analysed by ELISA using the R5 antibody, targeting the celiac toxic motifs "QQPFP". The digestion resistant peptides were sequenced by mass spectrometry and *in-silico* assessed to identify the celiac (CD) and allergenic sequences. Tritordeum Bulel cv exhibited a lower immunoreactivity than landrace (-51%) or modern (-58%) cvs. *In-silico* analysis showed that tritordeum has fewer CD epitopes belonging to the ω-gliadins than the landrace or

modern *cv*. Tritordeum presented also fewer α-gliadin allergenic epitopes than the modern wheat *cv*. The lower frequency of celiac epitopes in tritordeum, compared to the old and the modern wheat, is probably due to the absence of a D genome.

In chapter 5 the refined flour of Bulel cv was also used to produce model bread, along with another tritordeum flour, Aucan cv, taking as reference semolina and flour derived from durum and soft wheat cvs, respectively. To evaluate protein expression differences among the flour samples, we applied a gel-based proteomic approach. Breads were *in-vitro* digested using the INFOGEST model and the digestion products were analysed to compare bio-accessibility of nutrients, the potential R5 immunoreactivity and to map the digestion resistant peptides of tritordeum breads. The protein characterisation of flours showed very different and complex proteomes among the samples and within the tritordeum cvs. The amino acid bio-accessibility and the reducing sugars of tritordeum and wheat breads were comparable. Tritordeum cvs had about 15% higher alpha-amino nitrogen released at the end of the duodenal simulated digestion than soft wheat ($p < 0.05$). Bulel tritordeum flour, bread and digested bread had about 55% fewer R5-epitopes compared to soft wheat. The differences in protein expression found between the two tritordeum cvs reflected in different digestion products, allergenic and celiacogenic potential of the duodenal peptides. Our studies highlight that attention should be paid to considering all the tritordeum cvs as a *unicum* in terms of protein expression since sometimes the protein contribution can vary along with the genomic characteristics of the *Hordeum* and *Triticum* parents.

SECTION 3 – A novel food from a fermented dairy by-product: Kashk

This section is related to a fermented novel food made by wastes of dairy productions: kashk. This product is traditionally made by Iranian nomad populations using the waste of homemade yoghurts

and butter made with the milk of goats. This product is spontaneously fermented and dried under the sun. To date, this product is also manufactured at an industrial level in a liquid form.

Chapter 6 aimed at characterizing the kashk proteome and peptidome, comparing a traditional product with an industrial counterpart using a combination of proteomic approaches. Kashk products were *in-vitro* digested using a static model of food digestion (INFOGEST) to which was added the jejunal compartment implementing the use of brush border membrane enzymes (BBM). The phosphorylated casein-derived peptides (CPP) resistant to digestion were mapped alongside precursor peptides of bioactive sequences. The fermentation appeared to have hydrolysed peptide sequences knowingly responsible for milk-allergic reactions.

The presence of peptide with potential bioactivity, led to the characterization of the kashk water-soluble peptide extract (KE) **in chapter 7** by liquid chromatography-mass spectrometry. Furthermore, it was also evaluated the antibacterial and wound healing activity of KE. Significant antibacterial activity against clinical isolates of *Staphylococcus aureus* from patients with atopic dermatitis was confirmed with growth inhibition by approximately 45% (500 μg/mL). Interestingly, this activity was associated to the presence of antimicrobial peptides like Caseicidin 15 and 17 and was unique to traditional kashk. Its potential skin repair activity on an inflamed model of keratinocytes through scratch tests was also evaluated. The traditional kashk extracts sped up wound closure *in-vitro* in the presence of TNF-α, by approximately 44% (500 μg/mL), compared to control cells. This phenomenon was associated with the abundance of hydrophobic residues and wound healing bioactive peptide precursor.

ABSTRACT (ITA)

La popolazione mondiale è destinata a crescere fino a 9.8 miliardi di persone nei prossimi trent'anni generando una crescente domanda di fonti alternative e sostenibili di proteine. L'Unione Europea (UE) richiede una valutazione completa relativa sicurezza di un nuovo alimento (novel food) prima della sua immissione sul mercato. Il processo di valutazione comprende anche lo studio della potenziale allergenicità mediante un approccio basato sul peso dell'evidenza. Per sviluppare e convalidare l'approccio in questa tesi sono stati scelti come casi studio tre modelli di novel food: due di origine vegetale e uno di origine animale.

SEZIONE 1 – Un nuovo alimento di origine vegetale: la foglia di *Moringa oleifera*

Le foglie di *Moringa oleifera* sono una fonte di proteine ad alto valore biologico, considerate un novel food nell'UE e nel Regno Unito. In questa sezione è stato affrontato il potenziale allergenico delle foglie di *Moringa oleifera*. Le foglie sono una medicina tradizionale in Asia dove vengono utilizzate come panacea per curare e prevenire diverse malattie, oltre a far parte della dieta orientale come ingrediente da consumare sia crudo che cotto. La capacità della pianta di resistere a condizioni ambientali sfavorevoli, come siccità e alte temperature, la rende una scelta agricola sostenibile per il futuro. Tuttavia, l'alto contenuto di proteine fa scattare l'allarme sulla loro possibile allergenicità. Il **capitolo 1** è una revisione completa della letteratura, per fornire informazioni nutrizionali, tecniche e di sicurezza sui tessuti proteici della Moringa (cioè semi e foglie). Il **capitolo 2** è la caratterizzazione proteomica dell'ingrediente foglie di Moringa in polvere. La maggior parte delle proteine identificate apparteneva a vie fotosintetiche e metaboliche. L'analisi *in-silico* del proteoma fogliare ha evidenziato le Moritidi (mO1 e mO2) come potenziali cross-allergeni dell'eveina (Hev b 6.02), un allergene del lattice implicato

nell'allergia lattice-frutta. Nella foglia di Moringa è stata anche identificata la proteina di trasferimento lipidico (nsLTP), uno dei principali pan-allergeni negli alimenti, la cui identificazione è stata confermata mediante sequenziamento *de-novo.*

Facendo seguito alla caratterizzazione proteomica e *in-silico*, il **capitolo 3** è uno studio *in-vitro* atto a confermare la cross-allergenicità predetta nel capitolo 2. Gli estratti proteici di Moringa sono stati saggiati con test immunologici utilizzando IgG animali anti-Pru p 3 (la nsLTP della buccia di pesca) e utilizzando sieri umani IgE da pazienti allergici al lattice o alla nsLTP. L'immunoblotting eseguito con sieri allergici al lattice non ha mostrato immunoreattività specifica a Hev b 6.02. La rimozione delle glicosilazioni sulle proteine estratte dalla Moringa ha ridotto la risposta spuria delle IgE. Le IgG animali non hanno mostrato omologhi Pru p 3. Al contrario, le analisi delle IgE sieriche umane hanno evidenziato in modo riproducibile una banda immunoreattiva nei pazienti allergici alla nsLTP. I test immunologici hanno rivelato l'imperfezione degli attuali approcci per la valutazione del rischio di allergenicità e l'importanza di un approccio basato sul peso dell'evidenza.

SEZIONE 2 – Un cereale ibrido innovativo: il Tritordeum

Il *Tritordeum martinii* è un ibrido esaploide (AABBHchHch) derivato dall'incrocio di *Hordeum chilense* con grano duro. La farina derivata dai grani di tritordeum ha contenuto proteico e proprietà panificabili paragonabili al grano tenero.

Nel **capitolo 4** è stata ottenuta per la prima volta la caratterizzazione proteomica del tritordeum cultivar (cv) Bulel rispetto a due farine di controllo di *Triticum aestivum*: un grano moderno, cv Bologna, e un grano antico cv Andriolo. Per simulare la digestione *in-vitro* delle farine è stato utilizzato il modello INFOGEST. I prodotti di digestione sono stati analizzati mediante ELISA utilizzando l'anticorpo R5, specifico per i motivi celiaco-tossici "QQPFP". I peptidi resistenti alla

digestione sono stati sequenziati mediante spettrometria di massa e analizzati *in-silico* per identificare le sequenze celiache (CD) e allergeniche. Il tritordeum cv Bulel ha mostrato un'immunoreattività inferiore rispetto alle varietà antiche (-51%) o moderne (-58%). L'analisi *in-silico* ha rivelato che il tritordeum mostrava meno epitopi CD appartenenti alle ω-gliadine rispetto alla razza autoctona o moderna cv. Il tritordeum presentava anche meno epitopi allergenici dell'α-gliadina rispetto al moderno grano cv. La minore frequenza di epitopi celiaci nel tritordeum, rispetto al grano antico e moderno, è probabilmente dovuta all'assenza del genoma D.

Nel **capitolo 5** le farine raffinate di due cultivar di tritordeum, Bulel e Aucan venivano utilizzate per produrre modelli di pane, prendendo come riferimento il pane ottenuto con la semola e la farina derivate da cultivar di grano duro e grano tenero. Per valutare le differenze di espressione proteica tra i campioni di farina, abbiamo applicato un approccio proteomico basato su gel. I pani sono stati digeriti *in-vitro* utilizzando il modello INFOGEST e i prodotti di digestione sono stati analizzati per confrontare la bioaccessibilità dei nutrienti, la potenziale immunoreattività R5 e per mappare i peptidi resistenti alla digestione dei pani tritordeum. La caratterizzazione proteica delle farine ha mostrato proteomi molto diversi e complessi tra i campioni e all'interno delle stesse cultivar di tritordeum. La bioaccessibilità degli aminoacidi e gli zuccheri riducenti del tritordeum e del pane di frumento erano comparabili. Le cultivar di tritordeum avevano circa il 15% in più di azoto alfa-amminico rilasciato alla fine della digestione simulata duodenale rispetto al grano tenero ($p < 0,05$). La farina, il pane e il pane digerito del tritordeum Bulel avevano circa il 55% in meno di epitopi R5 rispetto al grano tenero. Le differenze nell'espressione proteica riscontrate tra i due tritordeum si riflettono in diversi prodotti di digestione, potenziale allergenico e celiaco dei peptidi duodenali. I nostri studi evidenziano che occorre prestare attenzione a considerare tutte le cultivar di

tritordeum come un *unicum* in termini di espressione proteica poiché a volte il contributo proteico può variare insieme alle caratteristiche genomiche dei genitori *Hordeum* e *Triticum.*

SEZIONE 3 – Un novel food ottenuto da un sottoprodotto lattiero-caseario fermentato: il Kashk

Questa sezione è relativa ad un novel food fermentato, prodotto dagli scarti delle produzioni casearie: il kashk. Questo prodotto è tradizionalmente preparato dalle popolazioni nomadi iraniane utilizzando gli scarti di yogurt e burro fatti con il latte di capra. Questo prodotto viene fermentato spontaneamente ed essiccato al sole. Ad oggi, questo prodotto è fabbricato anche a livello industriale sotto forma liquida.

Il capitolo 6 mirava a caratterizzare il proteoma e il peptidoma del kashk confrontando il prodotto tradizionale con quello industriale utilizzando una combinazione di approcci proteomici. Il kashk è stato digerito *in-vitro* utilizzando un modello statico di digestione (INFOGEST) a cui è stato aggiunto il compartimento del digiuno, implementando l'uso degli enzimi dell'orletto a spazzola (BBM). I peptidi fosforilati derivati dalla caseina (CPP) e resistenti alla digestione sono stati mappati insieme ai peptidi precursori di sequenze bioattive. La fermentazione sembrava aver idrolizzato le sequenze peptidiche notoriamente responsabili di reazioni allergiche al latte. La presenza di peptidi con potenziale bioattività ha spinto la caratterizzazione dell'estratto peptidico idrosolubile del kashk (KE) nel **capitolo 7** mediante cromatografia liquida accoppiata a spettrometria di massa. Inoltre, è stata anche valutata l'attività antibatterica e cicatrizzante del KE che ha mostrato una significativa attività antibatterica su isolati clinici di *Staphylococcus aureus* da pazienti con dermatite atopica, con un tasso di inibizione di circa il 45% (500 μg/mL). È interessante notare che l'attività antimicrobica, riscontrata unicamente negli estratti peptidici del kashk tradizionale, era associata alla presenza di peptidi antibatterici come la caseicidina 15 e 17.

È stata anche valutata potenziale attività di riparazione cutanea su un modello infiammato di cheratinociti attraverso un test di graffio. Gli estratti di kashk tradizionale hanno accelerato *in-vitro* la chiusura delle ferite in presenza di TNF-α, di circa il 44% (500 μg/mL), rispetto alle cellule di controllo. Questo fenomeno era associato all'abbondanza di residui idrofobici e alla presenza del precursore di un peptide bioattivo cicatrizzante.

GENERAL INTRODUCTION

1. Novel Food ingredients

Our world is facing a critical crossroads with the world's population foreseen to reach 9.8 billion people by 2050 (FAO, 2018). A direct consequence is the rising demand for valuable food proteins. To date, these proteins derive for about 33% from animal derived foods and this value is predicted to double in the next fifty years (Henchion et al., 2017). Livestock farming and extensive agriculture are the primary cause of biodiversity loss and pollution, therefore, severely contributing to climate change (Leip et al., 2015). In addition, the reserves that the Earth can renew do no longer meet the entire population's requirements. The "Earth Overshoot Day", which represents the point when humanity's demand for environmental resources and services in a year exceeds what the Earth can regenerate in that year, (https://www.overshootday.org/about/) is slowly but steadily shortening over the centuries (**Figure 1**). Green biotech and novel food ingredients could represent an important strategy to get alternative protein sources that can be produced sustainably. Sustainable food consumption plays an fundamental role in the Farm to Fork Strategy (F2F) at the heart of the European Green Deal, with the overall goal of achieving "net-zero" emissions in Europe by 2050, with a reduction to 50% by 2030 (D'Amato & Akdis, 2020). The availability of "novel foods" providing more sustainably produced proteins is thus becoming increasingly important to support this goal.

European Union (EU) defines as "novel" an ingredient that has not been consumed significantly before 15 May 1997 (European Parliament, 1997). In addition to agricultural products from third countries (e.g., chia seeds, noni fruit juice, etc.), novel foods include foods and food ingredients made from microorganisms, fungi, or algae (for example, oil rich microalgae and Quorn made from fungal proteins) (Pali-Schöll et al., 2019). In the definition of novel foods are also included

ingredients prepared from natural substances such as rapeseed protein isolates or those ingredient production base on new technologies and processes (e.g., UV-milk).

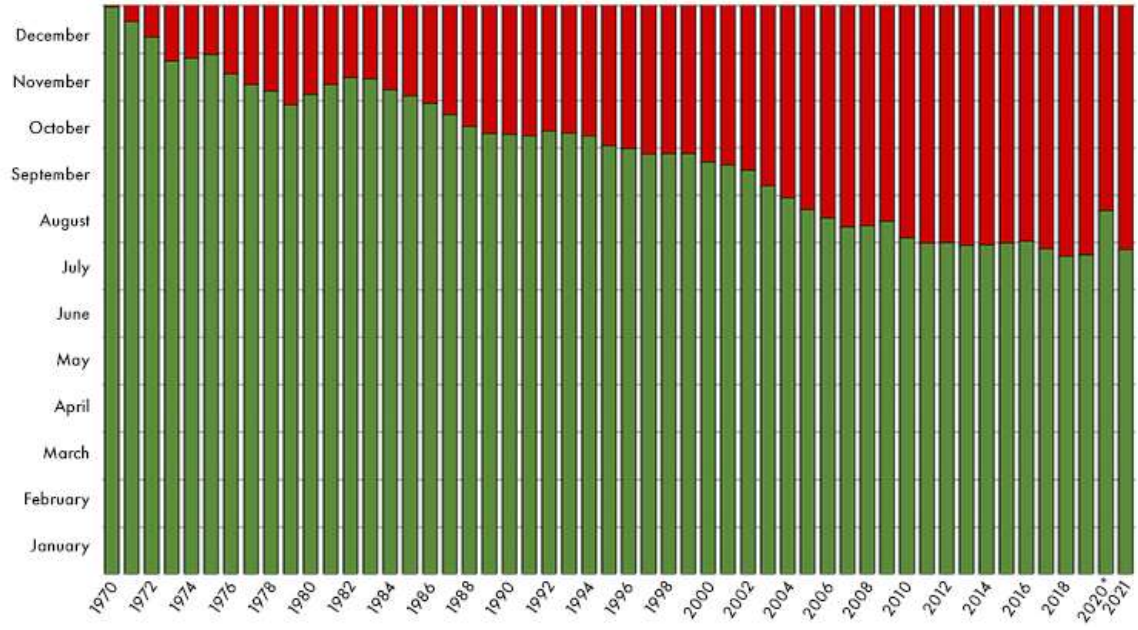


Figure 1. The Earth overshoot day from 1970 to 2021. In red are highlighted the months where we overused our resources. It is noteworthy that during the pandemic of SARS-CoV19, the Earth overshoot day had a delay of about three weeks due to the lower energy consumed.

A wide variety of viable *"novel food"* (e.g., mealworms, seitan and Quorn) is already available on the market, while others are still facing important nutritional and health evaluations by food safety bodies like the European Food Safety Authority (EFSA). These evaluations are focused on the nutritional value, the safety assessment including the potential toxicity and allergenicity of the ingredient. The allergenicity risk assessment of novel food is rather challenging for a series or reasons: i) the lack of knowledge of the proteome; ii) what makes a protein allergenic for some

people is not well understood clinically; iii) the lack of methods capable of predicting allergenic proteins, particularly de novo allergens; iv) the need for a consensus approach.

2. Food allergy

Food allergies, as well as food intolerances, are classified as a non-toxic reaction to food within the broader category of adverse reactions. Individuals with these disorders are susceptible to specific foods that are harmless for the population (Valenta et al., 2015). A food allergy is an immune response to a protein(s) food (Boyce et al., 2010; Burks et al., 2011). Conversely, an intolerance is a reaction not involving the immune system, which may involve metabolic, toxic, physiological, or undefined mechanisms. Food allergy can be classified into IgE-mediated reaction, non-IgE-mediated reaction, and mixed reaction according to the specific immunological mechanism involved (**Figure 2**). FA is associated with different exposure routs and organs, including those of the gastrointestinal tract, the skin, and the respiratory system. A variety of clinical manifestations may occur, ranging from mild and localized symptoms, such as fleeting oral itching, to severe and sometimes fatal reactions, such as anaphylactic shock (De Martinis et al., 2020). The principal foods involved in triggering the reaction are known as the "big eight". These foods are milk, eggs, fish, shellfish, tree nuts, peanuts, wheat, and soybean. However, in April 2021, an act of the Food Allergy Safety, Treatment, Education, and Research (FASTER) became law, recognizing sesame as the 9th major food allergen in the United States (https://www.fda.gov/food/food-labeling-nutrition/food-allergies). Basing on a recent FAO/ WHO consultation, considering the lack of data on prevalence, severity, and potency, or because some foods are consumed differently in different regions, the Committee considered some allergens

(buckwheat, celery, lupin, mustard, oats, soybeans, tree nuts, such as Brazil nuts, macadamia, pine nuts) not suitable for inclusion on global priority allergen lists but even to include in some national priority allergen lists. As dietary trends are characterized by increased consumption of plant-based foods and diets with alternative protein sources, it has been recommended to include pulses, insects, and other foods like kiwi fruit on a "watch list" and evaluate them for potential inclusion on the priority allergen list as data on prevalence, severity, and potency become available (FAO/WHO, 2021).

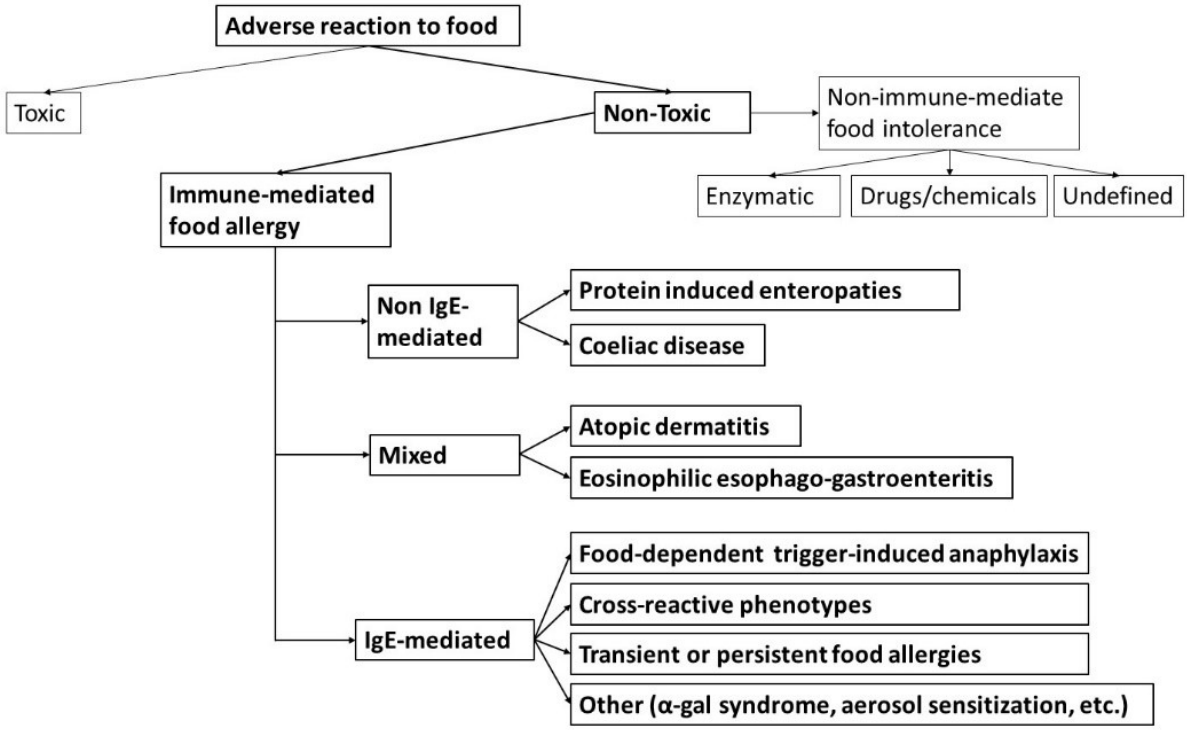


***Figure 2**. Scheme of adverse reaction to food; in bold is highlighted the non-toxic immune-mediated adverse reactions to food. (Adapted from De Martinis et al., (2020)).*

Although no consensus exists on the allergenicity parameters and criteria that should be used to evaluate and make decisions, currently, allergenic proteins are identified by a few specific

biochemical characteristics. Protein abundance, multiple linear IgE-binding epitopes, and allergen structure are all characteristics that are considered likely to make a food protein an allergen. The resistance to digestion and processing was considered among these characteristics, however there is still no evidence about the reliability of this method and some authors even suggested to exclude it from the allergenicity assessment for now, until knowledge gaps will be filled (K. Verhoeckx et al., 2019).

The correlation between the abundance of a protein in a matrix and its predisposition to be an allergen is supposed basing on the likelihood to encounter the immune system rather than those representing only a small percentage of the total protein ingested. However, there are exceptions to this rule: Gal d 4 (lysozyme) an important allergen in hen's egg, is an example of a major allergen that does not constitute a significant portion of the food (around 3.4%) (Caubet & Wang, 2011).

The biochemical structure also influences the overall likelihood of a protein to be a potential allergen. It is accepted that an allergen triggers an allergic response and subsequent inflammation if it cross-links at least two IgE antibodies on the receptor of pre-sensitized effector cells, such as mast cells and basophils (Bufe, 2001). This cross-linking is possible thanks to the so-called epitopes, which are structural segments of antigens that are recognized by either antibodies or antigen receptors (e.g., T-cell receptors). Epitopes are recognized by antibodies and T-cell receptors in two different ways: conformationally (secondary and tertiary structure) and linearly (primary structure). Each arm of an antibody present an antigen-binding site, called paratope. Paratopes comprise six hypervariable loops, forming a surface structure that is complementary to that of an epitope (C. Liu & Sathe, 2018). The paratope-epitope complex interacts with non-covalent forces highly specifically. Linear epitopes are composed of several amino acid residues

≥ 8 which composition seems to contain not predictable motifs. As determined in synthetic allergen model systems, the optimal suggested distance among peptides to link the paratopes of at least two IgE antibodies is attested around 6 nm (Bucaite et al., 2019). This observation suggests that one epitope alone is not enough for triggering an allergic response.

Despite the predominance of conformational epitopes (especially in aeroallergies), it has been suggested that these kinds of epitopes are less protagonist in food allergies, because food proteins usually undergo various treatments before exposure (e.g., cooking, denaturation) and throughout the human gastrointestinal digestions (e.g., proteolysis), which destabilise protein three-dimensional structures (Vila et al., 2001).

The same authors found that milk-allergic children with persistent symptoms were disproportionately more likely to have specific IgE antibodies to linear epitopes rather than to native alpha and beta-casein proteins (Vila et al., 2001).

Considering the resistance to denaturation and digestion another aspect that can help to predict the allergenicity of a protein. It should be noted that some allergenic proteins are sensible to enzymatic digestion (i.e., aeroallergens). Basing on the different route of exposure (respiratory tract), their allergenicity can be explained basing on the fact that in these tissues proteolytic enzymes are not present (Pekar et al., 2018). Another factor that contributes to the overall allergenicity of a protein is the presence of disulphide bonds, which contribute to their resistance to denaturation and proteolysis. Besides these factors, glycosylation and enzymatic activity may also contribute to the allergenic power (e.g., chitinase, lysozyme, etc.) (Huby et al., 2000). Interestingly, antibodies have been found to recognize non-peptidic structures, such as glycans. Because of their universal presence in plant and invertebrate foods, two non-mammal carbohydrate epitopes, α-1,3-fucose and β-1,2-xylose, are associated with IgE-mediated cross-reactivity (Aalberse, 2000).

Traditionally, carbohydrate epitopes were considered having little clinical significance (Ronald van Ree, 2002). In contrast, galactose-α-1,3-galactose, another carbohydrate epitope found in red meat, was found to cause delayed anaphylaxis (Apostolovic et al., 2014). Notably, some factor such as the presence of epitopes, are more essential than other elements such as glycosylation, the resistance to proteolysis and enzymatic activity which, although important, are just subsidiary (Huby et al., 2000). Recently, some authors have suggested to evaluate the allergenicity risk assessments of novel food in a risk-based manner, which is based on two fundamental parameters: the percentage of incidence and the potency of the allergen (ED50 in mg of proteins) (Houben et al., 2019). This serves to put into perspective the potency of an allergen with its public health relevance. It should be noted that a food that contains allergens may still be approved by authorities (see mealworm).

2.1 Mechanism of IgE-mediated food allergy reaction

The development of IgE-mediated allergic reactions can be divided into two phases: sensitization and elicitation (**Figure 3**). Pekar et al., (2018) stated that sensitization to food antigens can take place in various tracts of the human body (*e.g.,* oro-gastrointestinal (GI) tract, skin and the respiratory tissues) and for this reason, a person can be sensitised to an allergen in other ways than ingestion. A subject can therefore be sensitised to a specific allergen, present in a matrix (i.e., pollen, latex) that is phylogenetically related with a well described allergenic matrix/food, having homologous proteomes. This phenomenon is known as cross-reactivity and it is observed especially in contact and aeroallergen sensitisations (Popescu, 2015). This highlights the significance of the identification of cross-reactive proteins, especially in the allergenicity risk assessment of novel food.

After ingestion, the proteolytic enzymes present in the stomach and intestine break down food proteins. Afterwards, partially digested proteins and peptides can access the gut mucosa after being absorbed through the intestinal epithelium. Antigen-presenting cells (APCs) such as dendritic cells absorb allergens through endocytosis from the tight junctions between the enterocytes of the intestinal mucosa (Kumar et al., 2012; Morelli et al., 2004). Protein hydrolysis and other enzymatic reactions inside the APC vesicles lead to the association of an antigen with the major histocompatibility complex (MHC) molecules. After binding to class II MHC molecule, the allergen is presented by APCs through its surface to naïve CD4+ T cells that carry the T cell receptor (TCR) (Kumar et al., 2012). These cells are activated along with costimulatory signals such as monocyte-derived interleukin-1 (IL-1) and other cytokines, including autocrine stimulation by IL-2 (Van Den Eeckhout et al., 2021). These specific T cells are stimulated to transform mostly into T helper 2 (Th2) cells, which response, induced by IL-4, IL-5, and IL-13, is responsible for the stimulation of B cells activation, resulting in elevated levels of allergen-specific immunoglobulins E (IgE) (Barni et al., 2020; Moens & Tangye, 2014). These classes of molecules present high compatibility with IgE receptors (FcƐRI) on the surface of mast cells (tissues) or basophils (blood) (**Figure 3**) (Anvari et al., 2019). If the antigenic protein recurs, it will form cross-links with the IgE present on mast cells and basophils (elicitation pathway). In response to this crosslinking, mast cells and basophils degranulate and release vesicles containing mediators of hypersensitivity, including histamine, leukotrienes, prostaglandins, and inflammatory cytokines (Iweala & Burks, 2016). This reflects in a macroscopic series of mild to severe symptoms ranging from cutaneous (hives, itching, swelling, erythema) to respiratory (hoarseness, respiratory distress), gastrointestinal (nausea, vomiting, diarrhoea) and cardiovascular (blood pressure decrease and cardiac arrest) (Ring et al., 2014).

However, the correlation between the intrinsic properties of a protein (structure and function) and its allergenic potential (failure of oral tolerance) remains not well understood. As stated before, protein stability to denaturation and proteolysis, as well as the impact of food processing, can affect its immunogenicity and allergenicity potential (Huby et al., 2000; Pekar et al., 2018).

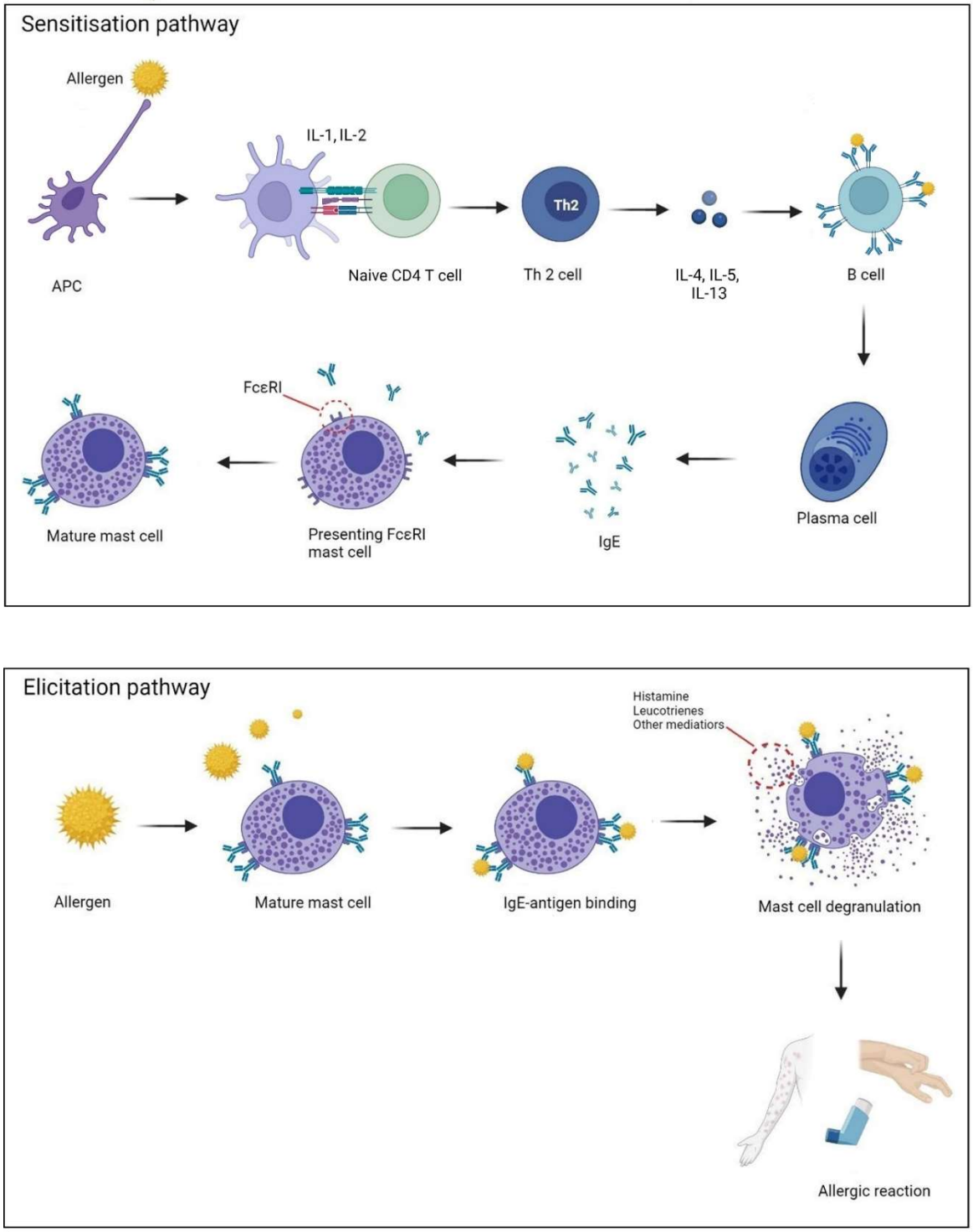


Figure 3. *IgE mediated allergy pathways. (The image was created with BioRender.com).*

2.2 Epidemiology of food allergies

In the last 20 years, FA appears to have been increasing in prevalence (up to 10% of the population affected (Gupta et al., 2019), mainly affecting people living in industrialized and westernised countries (Sicherer & Sampson, 2018). Notably, in the Western world, the incidence of FA in children is higher (5%–8%) than in adults (2%–4%) (Sicherer & Sampson, 2014).

The prevalence of food allergies may differ geographically depending on several factors such as i) the diet or lifestyle habits (peanut allergy has increased because of the widespread use of peanuts in the US), ii) cooking methods (some preparation can increase or diminish the allergenic potential of some ingredients), iii) the age at which a child is weaned (earlier peanut protein consumption can lower the prevalence of peanut allergy), iv) aeroallergen exposure, which determines regional patterns of food allergy, v) genetic factors (genetic variants in the HLA-DR and HLA-DQ genes might increase the risk of peanut allergy in children of European descent) (Bartra et al., 2016).

In Europe, hazelnuts were the most common tree nut allergens, while cashews and walnuts were most common in the United States. It is estimated that approximately the 80% of all reported allergies in children are due to milk, egg and peanuts. Even though many children's allergies will disappear during their growth, peanut, tree nuts, and fish allergies rarely recede in the adulthood, being the most common allergies among adults (Bannon, 2004). The United Kingdom shows the highest incidence of egg allergies (2.18%) while Greece the lowest (0.07%). In The Netherlands and the United Kingdom, the rate for milk allergy is the highest (1%) whereas in Lithuania, Germany, and Greece, the lowest (<0.3%) (Schoemaker et al., 2015; Xepapadaki et al., 2016). The prevalence of milk, egg, wheat, fish, shellfish, and tree nut allergies in Northern Europe appears to be higher than in the rest of Europe. In Western Europe, however, the prevalence of soy allergy seems to be higher than elsewhere (Bartra et al., 2016).

Once again, data suggested that in the USA, food allergies varied among different ethnicity. Gupta et al., (2019) found that FA were prevalent among African American children, compared with their Caucasian counterparts. Food allergy rates were also lower among White adults compared to Black, Hispanic, Asian, and multi-racial counterparts.

In addition, studies indicate that first-generation immigrants are less prone to suffer from allergies than native-born individuals (Panjari et al., 2016).They are likely to develop the same type of sensitization and allergies of their host countries' citizens (Warren et al., 2020). These findings show that genes-environment interactions may play a role in food allergy disease aetiology in these populations. Additionally, as global migration continues to urbanized regions with a high prevalence of food allergies, the burden of food allergies will probably become more severe (Allen & Koplin, 2019).

3. Food allergens

To realize the importance of the cross-reactivity evaluation in the allergenicity risk assessment of novel food, it is crucial to understand the biochemistry and the structure of the protein families that are mainly involved in the most common allergies. Throughout the chapter, we will refer to allergens using the WHO/IUIS nomenclature. The nomenclature is based on the Linnean system and applies to all known allergens. For example, the allergens from the latex tree *Hevea brasiliensis* will start with the first three letters of the genus (Hev), followed by the first letter of the species (b) and a number which change basing on the allergenic protein.

In the introduction to each section, detailed information is provided for each potential allergen involved with the investigated matrix.

4. Allergenicity risk assessment

Toxicological and allergenic safety assessments must be carried out before a novel food or ingredient can be released on the European market (EC Regulation No 258/97, EU Regulation 2015/2283 and EU recommendation 97/618 EC, http://eur-lex.europa.eu/). This allergenicity assessment is a "weight-of-evidence" approach based on EFSA GMO (Genetically Modified Organisms) guidance: "Allergenicity assessment of GM plants" published in 2011 and implementing Regulation EU (No) 503/2013 (EFSA, 2011). EFSA recently published a document on allergenicity assessment of GM plants (Mackie et al., 2019). This document describes the new scientific and regulatory developments on protein digestibility *in-vitro*. EFSA therefore evaluated the implementation of *in-vitro* tests of protein digestibility in allergenicity assessment. However, panel stated that additional research must be conducted before any additional guidelines can be provided. A model should be able to reflect the human situation, but it must also be predictive, and this can only be achieved when the method is physiologically relevant than the currently used methods (i.e., resistance to pepsin assay). On the one hand, pepsin digestion of proteins based on a static model is not physiologically relevant, because it considers factors such as matrix and kinetics. Allergenicity RA digestion experiments, closer to the actual physiology (e.g., including brush border membrane enzymes - BBM) may enhance the predictive ability of the assays, but meantime increase the cost, the complexity, reducing their relevance as a screening method (Verhoeckx et al., 2019). For these reasons, all the parameters involved in the in-vitro digestion make extremely hard to design an assay or strategy that is simple, accurate and predictive, especially since it is yet unclear exactly how these factors influence digestion and how they can be included (Verhoeckx et al., 2019). Another and more recent EFSA report (Naegeli et al., 2021)

stressed out the necessity to clarify and identify the aspects of the digestion and absorption of dietary proteins that can apply to assessing potential risks of allergenicity posing some questions:

- Which in-vitro digestion model is the best?
- Which item would be the most appropriate to test in such a model?
- In order to assess the relevant proteins/fragments identified in previous steps, what actions are needed?
- What can be done to integrate this information into a weight-of-evidence approach?
- What kind of test material would be most appropriate and workable? (i.e. purified protein vs. whole-plant extracts and considerations of food matrix and food-processing conditions); what are the most reliable analytical techniques to evaluate the allergenicity of the digestion products (e.g. immunoblotting, Tandem mass spectrometry, MALDI-ToF, Chromatography and/or SDS–PAGE)?
- How should fragments be evaluated (i.e., size, persistence, abundance, etc.)? Is it possible to set cut-off values for risk assessment and fit them into the sensitisation and elicitation scenarios?
- Is it possible to set acceptability limits for digestibility to rank potential allergenic proteins?

Because of its complexity, the prediction of allergenicity represents the most challenging aspects of the safety assessment of novel food ingredients. For food safety authorities like EFSA, the allergenicity risk assessment (RA) of novel proteins is essential since they can cause immune-mediated adverse reactions, sometimes life-threatening, in susceptible individuals.

4.1 Background analysis

The allergenicity RA is based on a "four-pillars" approach. The first pillar is represented by an in-depth and systematic study of the literature and of the background information (source, production, and transformation processes) about the *novel food* investigated. At the outset evaluation is performed of the history of allergic reaction following the consumption of the novel food in those countries outside the EU where it is commonly consumed by at least 25 years. During this step is necessary to consider the epidemiology and the geographic area where this novel product is consumed, since the allergy reflects the population. Interestingly, in the Mediterranean areas, where there is a lack of birches, the main hazelnut severe life-threatening allergies in adults are associated with sensitization to allergen Cor a 8, a non-specific Lipid Transfer Protein (ns-LTP) (Datema et al., 2015). In the background's context investigation, the phylogeny analysis plays a pivotal role in order to locate the examined *novel food* into a particular family that could comprise known allergens, thus discriminating potential cross-allergens on a proteomic or genetic basis.

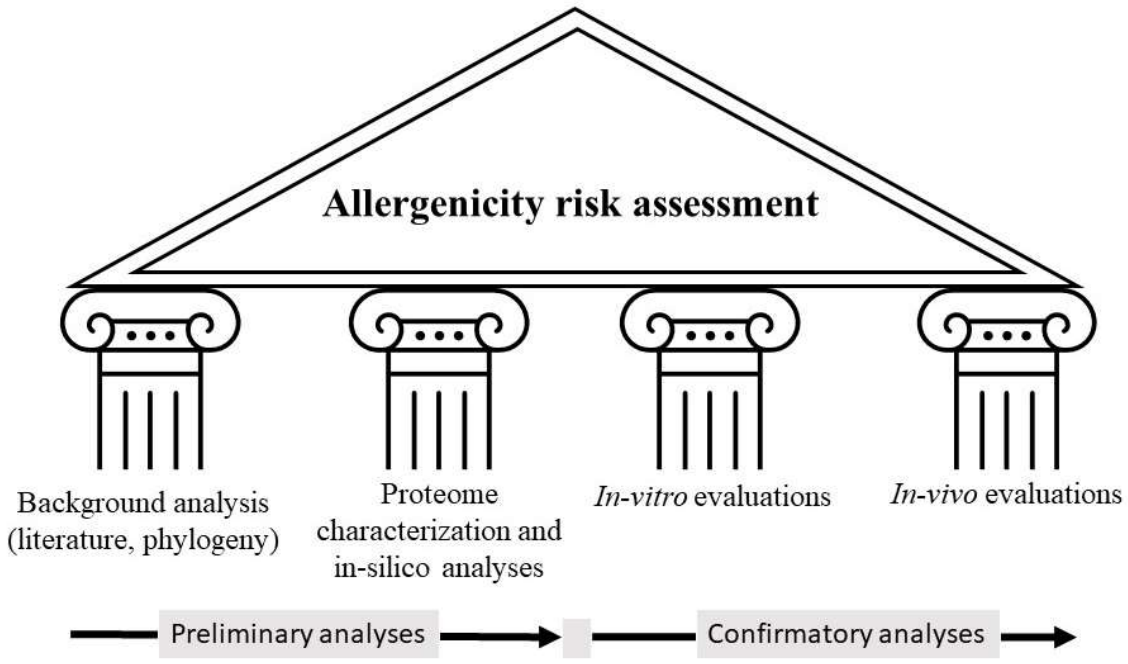


Figure 4. The four pillars of the allergenicity risk assessment.

4.2 Proteome characterisation and in-silico analyses

The second pillar of allergenicity RA is the molecular characterisation of the novel food/ingredient representing a preliminary analysis that lays the basis and directs the confirmatory *in-vitro* and *in-vivo* assays that are an essential step during the allergenicity risk assessment.

The proteome characterisation begins with the extraction of proteins from the matrix. The protein extraction techniques can be divided either in "analytical-grade", which provides highly purified protein fractions but involve the use of toxic extractants (i.e., 2-sulfanylethan-1-ol, tris saturated phenol, trichloroacetic acid, etc.), or in "food grade" extracts, which relies on soluble protein extracts, less complex but suitable for human consumption. Both characterisations are useful, and the choice depends on the type of food that is going to be assessed (whole ingredient, fraction thereof, protein hydrolysate).

A complete proteome's characterisation can rely on gel-based (electrophoresis), and gel-free (liquid chromatography) techniques coupled with advanced tandem mass spectrometry (MS/MS) (Scherp et al., 2011). The most common proteomics approach is called "bottom-up" since the identification of a protein (from protein isolates or purified proteins) starts from the sequencing of the peptides generated from their digestion by specific proteolytic enzymes (i.e., trypsin, chymotrypsin, etc.) (Zhang et al., 2013).

The gel-based analyses involve the use of polymeric gels (a mixture of acrylamide/bis-acrylamide) in which, thanks to an induced electric field, negative charged proteins (thanks to a pre-treatment with sodium dodecyl sulphate – SDS) migrate to the anode. This migration is faster or slower basing on protein molecular weight. In order to make them visible, proteins are usually stained with colouring agents like Blue Brilliant Coomassie, which binds to the basic residues of the

proteins (Brunelle & Green, 2014; Kruger, 1996). However, not all proteins can be revealed with this method, that is less sensitive than other methodologies involving the use of chemiluminescent agents. The protein of interest is excised from the gel and, after some operations like reduction and carbamidomethylation (to disrupt potential disulphide bonds and avoid potential refolding), it can be digested with specific proteolytic enzymes (e.g., trypsin with cleaves Arg and Lys) and can be identified by high resolution mass spectrometry (Shevchenko et al., 2006). The previous gel-based techniques can be also enriched of another information: the proteins' isoelectric point (bi-dimensional electrophoresis). The separation of proteins in a pH gradient is fundamental in enhancing the resolution of the SDS-PAGE to corroborate the identification (Issaq & Veenstra, 2008).

The most common gel-free proteomic techniques involve the use of chromatographic systems like HPLC coupled with MS/MS, which separate proteins or peptides basing on the stationary phase of the HPLC column (e.g., a C4 or C18 reverse phase column–RP) (Abdallah et al., 2012). This method is more sensitive than gel-based techniques and is capable of detecting very low abundant proteins. The output of this analysis are sequenced peptides that, through bioinformatic tools based on protein databases, are inferred to a parent protein. Many bioinformatic tools of peptide inferring, perform *in-silico* a theoretical digestion of the protein database, producing a list of hypothetical mass spectra of the generated peptides, which is then matched with the experimental mass spectrometry data, searching for potential homologies (Chen et al., 2020). This operation is based on a 100% of sequence identity match between the identified peptide and a segment of a protein present in the protein database. This method excludes identification of protein in which the identified peptide match for less than 100% (for example one single amino acid difference). Notably, the presence in the database of the parent protein is important because, if absent, the

peptide is associated by homology to another protein which belongs to same protein family (but from a different organism!) but could present various modifications or isoforms far from the sequence associated. On the other hand, if in the peptidome are present peptides whose inferring does not lead to any protein (because it does not return a 100% match), this could suggest the presence of isoforms or point modifications, which can be crucial in the allergenicity risk assessment of a novel food. Although it is possible to infer these proteins manually (by using the BLAST tool on UniprotKB), this operation is time consuming. To date there is no software able to associate these modified peptides with a protein family, since they cannot discriminate which are the amino acids that are changed. Unfortunately this important information remains thus unused. This approach provide evidence of protein expression in the ingredient, and it is a unique approach for acquiring the protein fingerprint to guide the in-silico allergenicity evaluation.

Within *in-silico* allergenicity RA, the most common bioinformatic evaluation tools compare the sequence of a novel protein with known allergens seeking for sequence homology greater than a specific value to define potential cross-allergens. This assessment assigns to each protein a score that relies on the evaluation of a threshold value (i.e., 35% homology over a sliding window of 80 amino acids) or on more advanced machine learning calculations (Fernandez et al., 2021). Bioinformatics is developing quickly and in the last few years, a rapid food allergen identification using artificial intelligence (AI) has been proposed as a new auxiliary tool for guiding *in-vivo* and *in-vitro* analyses. The new AI-based method operates random forest computing to generate an *in-silico* model that uses physicochemical and biochemical properties (e.g., molecular weight, taxonomy, total amino acid count, percentage of tryptophan and glutamic acids, etc.) of the novel proteins for predicting both their cross- and de novo allergenic potential (Westerhout et al., 2019). The proteomic characterisation of the matrix is fundamental to individuate proteins, whose amino

acid structure may share high sequence homology with known allergens (cross-reactivity), and to give to AI the information required for predicting *de-novo* allergens. However, due to the scarce availability of curated protein sequences in databases, a complete allergenicity risk assessment on the complete proteome of a novel food is hard to perform. This often reflects on arbitrary databases choices that negatively affects the accuracy of the allergenicity RA.

In fact, bioinformatic approaches have a big bottleneck represented by protein databases: the most curated is the database employed in the AI predictions, the most reliable is the output. Artificial intelligence algorithms are fed with both protein sequence/structure databases (e.g., UniprotKB, Swissprot, etc.) and allergen databases (e.g., Compare, Allergome, WHO/IUIS, etc.) that lack of curated and reliable information also because often these databases are "user access level" that means that they are updated with unreviewed sequences provided by users. To date, just a few allergen databases and proteome databases are considerable as reliable. For example, well characterised proteome such as cow's milk and hen's egg can be reasonably considered as curated, whereas green plants databases are very poor, especially those of minor and tropical plants. In this respect, a lot of sequencing work still needs to be done since there is a big hole in the knowledge of novel food protein sequences. For this reason, to date, I cannot imagine that an allergenicity RA, based only on bioinformatic predictions, would be a reliable tool for the prediction of the allergenicity of *novel food* proteins. Thus, there is and there will still be the need for a complete allergenicity risk assessment workflow, also based on *in-vitro* and *in-vivo* assays. However, *in-silico* analysis can anyway have a crucial role in the preliminary phases of the allergenicity risk assessment.

4.3 In-vitro evaluations

The third pillar of the allergenicity risk assessment is represented by *in-vitro* (and *ex-vivo*) assays, performed to evaluate specific IgE binding of the proteins in the novel food. An example of the importance of evaluation is the allergenicity risk assessment of mealworm. The presence of IgE binding to tropomyosin, arginine kinase, sarcoplasmic calcium-binding protein, and myosin light chain (minor shellfish allergens) in shrimp-allergic patients suggests they are at risk for mealworm allergy (Verhoeckx et al., 2014).

Currently, the most common analyses for semi-quantitative allergen detection in foods are immune based assays. The immune-based assays performed to evaluate the presence or potential immunoreactivity of a specific protein (potential cross-allergen) with an antibody (called primary antibody). This antibody could arise from animals (IgG) or from the human sera (IgE) of patients allergic to a known allergen. If in the novel food is present the suspected cross-allergen, which structure was found by *in-silico* analyses to be homologous the known allergen, the primary antibodies are likely to bind to that. After this cross-linking, there is the addition of a secondary antibody, which is specific for binding the human antibody and is labelled with an enzyme, which is the horseradish peroxidase (HRP) (Lacy & O'Kennedy, 2005). When HRP interacts with chromogenic substrates (e.g., 3,3',5,5'-Tetramethylbenzidine, 3,3′-Diaminobenzidine, etc.) it produces coloured products, and when it reacts with chemiluminescent substrates, it emits light (e.g., luminol). If the primary antibody has recognised the antigen, the binding between the secondary antibody and the primary antibody in presence of HRP and of a chromogenic substrate will produce a coloured product. The spectrophotometric absorption is directly proportional to the analyte concentration. Owing to their reliability and accuracy, enzyme-linked immunosorbent

assays (ELISAs) are the most used technique in the presence's evaluation and immunoreactivity of potential cross-allergens (Konstantinou, 2017; Sakamoto et al., 2018).

In allergy diagnosis, cross-reactive carbohydrate determinants (CCDs) play an important role. This phenomenon describes a protein linked with carbohydrate complex responsible for the phenomenon of cross-reactivity of sera from allergic patients. This occurs especially with vegetal proteins, which use glycosylation in several metabolic pathways, enhancing however the occurrence of false-positive results during IgE based *in-vitro* assays (Van Ree & Aalberse, 1999). For this reason, cellular assays such as basophil activation tests (BAT) are often coupled to these typologies of immune- based *in-vitro* assays to evaluate besides the IgE binding, there could be a further effect on the immune system. Based on flow cytometry, the basophil activation test (BAT) measures degranulation that takes place following stimulation by allergens, providing results much closer to reality (Santos et al., 2021).

4.4 In-vivo evaluations

If the previous assays show the likelihood that a protein (s) of the novel food represents a cross-allergen, it is possible to proceed with *in-vivo* assays. As a first test for the detection of specific IgE to foods, skin prick tests (SPTs) are the most likely to be used. Besides being comfortable with patients, SPTs are easy to perform, quick, cheap, and sensitive (Bignardi et al., 2019). Hence, they provide the best method of demonstrating IgE responses to foods. SPTs can detect tissue bound IgE and an atopic state in patients with type 1 allergies. SPTs can induce an immediate hypersensitivity response in the skin (Fatteh et al., 2014). Currently, major plant food allergens are being used in in-vitro and in-vivo diagnostics as purified, natural, and recombinant sources to improve sensitivity, specificity, and reproducibility. However, unlike aeroallergens, food allergen

extracts have not been standardized. SPTs to eggs, milk, peanuts, and fish have a very high negative predictive accuracy (most >90%), but very poor specificity (50%–85%) demonstrating to be effective to rule out an IgE-mediated food allergy (Fernandez-Riva & Miles, 2003). Conversely, a positive test could show a clinically relevant food allergy, in which case an oral challenge should confirm the diagnosis. Therefore, SPTs and in vitro IgE tests are valuable for demonstrating food specific IgE antibodies, but they cannot establish the diagnosis of clinical food allergy. Food allergies can be diagnosed conclusively with the oral food challenge test. There are three main types of oral food challenges: open (both patient and physician know what food was ingested) single-blind (only the physician can see what food was consumed) or double-blind (neither patient nor physician knows what food was consumed) (Fernandez-Riva & Miles, 2003). It is also possible to conduct a placebo-controlled double-blind challenge to increase the effectiveness of the assessment, especially in those patients who appear to have multiple food allergies (Bock et al., 1988). Food challenges should be conducted in hospitals where emergency care is readily available. As a first step, a dose unlikely to elicit symptoms is administered to the patient under fasting conditions, based on the eliciting dose reported in his medical history or his last positive provocation test. Afterwards, incremental (double) amounts of food are given to the patient at periods slightly longer than expected for a positive reaction until the patient consumes normal amounts of food. Foods can be served in capsules (dehydrated) or in a vehicle designed to conceal taste, consistency, colour, or odour. It is usually necessary to confirm a negative blind challenge with an open feeding (Huijbers et al., 1994; Noè et al., 1998).

5. Impact of food processing and digestion on the allergenic properties of food allergens

It is well established that food processing may alter the allergenicity of proteins by inducing structural changes in their scaffold. Because of these biochemical and physical changes, processed proteins are broken down differently during digestion and thus may be presented to the immune system differently (Mills et al., 2009). These types of process-induced modifications include protein denaturation, aggregation, and chemical modifications such as glycosylation induced by Maillard's reaction or interaction with lipids and polyphenols. All these alterations that food proteins undergo during processing may affect their allergenic potential, which is often pre-determined by the structure of the protein. In addition, these chemical and physical modifications may destroy existing allergens, or may encourage interactions between proteins and other components in the food matrix, resulting in the formation of novel allergens (neo-allergens) (Sathe et al., 2005; Teodorowicz et al., 2015).

The agglomeration of unfolded proteins may generate weak or strong aggregates (depending on the formation of disulphide bridges), which negatively influence their solubility. Beyer et al., (2001) showed that IgE from human peanut allergic sera bound peanut allergens more strongly in roasted peanuts than in boiled or fried peanuts. This may suggest that different heat treatments can modulate IgE binding being affected by protein aggregation and solubility. This is clear in Mondoulet et al., (2005) who found out that boiled peanuts were less allergenic because allergens leached out of the matrix during boiling. In roasted peanuts, where the processing temperature is higher (140 °C), Maillard reaction was found to cross-link peanut allergens Ara h 1 and Ara h 2 to form high-Mr aggregates that bind IgE more effectively than unmodified allergens, resulting also more resistant to gastric digestion (Maleki et al., 2000). It was also discovered that the

modifications occurring because of the Maillard's reaction increased the IgE-binding ability of allergenic shellfish tropomyosin (Nakamura et al., 2005).

Structural and biochemical modification of proteins can also affect their digestion by proteases (i.e., pepsin, trypsin, etc.) in the gastrointestinal (GI) tract. Although there is no evidence tis generally accepted that proteins that show a high resistance to digestion are more likely to be allergens. In fact, an intact or not fully proteolysed protein is more likely to interact and being absorbed by with cells of the immune system like APCs, reflecting in possible harmful reactions. For this reason, the resistance to pepsin, which is one of the proteolytic enzymes of the GI tract, was used as a model in the allergenicity's evaluation of proteins. In fact, this assay was developed to determine how resistant a protein is to the extremes of pH and pepsin proteolysis in the mammalian stomach (Bannon, 2004). To date, the relationship between allergenic potential and resistance to pepsin digestion is not yet understood. Considerable evidence following the use of this method suggests that allergenicity is not strictly associated with resistance to pepsin digestion (Verhoeckx et al., 2019). This is mainly because this model was not designed to mimic the entire conditions of human digestion, and therefore it cannot predict the half-life of a protein in-vivo. During their transit in the gastrointestinal tract, beyond the pepsin hydrolysis in the stomach, food proteins and peptides undergo other cleavages by pancreatic enzymes in the duodenum (i.e., trypsin, chymotrypsin, elastase and carboxypeptidase) and in the jejune (by brush border membrane enzymes). Before they are absorbed, peptides are hydrolyzed by proteolytic enzymes in the mucus layer (which contains additional eso- and endo-hydrolases) covering the epithelium. In addition, further intracellular degradation of proteins and peptides may take place after these luminal events (Wickham et al., 2009). Although during this process, most food proteins are

broken down into immunologically inactive fragments, a small amount may escape digestion and maintain their partial or intact structure and immunostimulatory properties.

In order to evaluate the fate and the potential immunoreactivity of the proteins and polypeptide escaped the digestion, several in-vitro digestion models have been proposed. However many of these methods contradict each other in terms of digestion parameters (e.g., enzyme unities, salts concentration, pH, etc) making difficult to compare results across research groups in order to reach a consensus (Di Stasio, 2018). To overcome this issue, the Infogest COST action has developed and is constantly upgrading a standardised static digestion protocol based on appropriate physiological conditions that may be applied for several types of food matrices (liquid or solid), and can be adapted depending on experimental requirements (depending on food matrix composition) (Brodkorb et al., 2019).

6. PhD project aims

This thesis aimed to carry out the preliminary allergenicity risk assessment of three models of sustainable and protein-rich novel food/ingredients. These matrices were represented by two vegetal-based (*Moringa oleifera* and *Tritordeum*) and one animal-based (Kashk) products. This preliminary allergenicity risk assessments paved the bases for more directed in-*vitro* and *in-vivo* assessment procedures meant to evaluate the use of these matrices in the production of novel foods. Once the protein matrices have been characterized by proteomics analysis, the protein sequences were compared *in-silico* with known allergen/epitope sequences on specific allergen databases (such as WHO/IUIS) or literature data to obtain a prediction of their cross-allergenic potential. However, given the low predictive value of the cross-allergenic potential based only on the bioinformatics approach, further *in-vitro* assessments were also carried out.

To this aim, two of the model matrices (Tritordeum and Kashk) were also selected to evaluate the impact of processing and digestion on the overall allergenic potential. These products were *in-vitro* digested using the static Infogest model and the potential allergenicity/celiacogenicity was evaluated on the digestome.

The information obtained will allow the development of more direct *in-vitro* (e.g., histamine release, cellular assays, etc.) and *in-vivo* analytical protocols for the complete allergenicity assessment of the novel foods investigated in this thesis.

7. References

Aalberse, R. C. (2000). Structural biology of allergens. *Journal of Allergy and Clinical Immunology*, *106*(2), 228–238. https://doi.org/10.1067/mai.2000.108434

Abdallah, C., Dumas-Gaudot, E., Renaut, J., & Sergeant, K. (2012). Gel-Based and Gel-Free Quantitative Proteomics Approaches at a Glance. *International Journal of Plant Genomics*, *2012*, 1–17. https://doi.org/10.1155/2012/494572

Allen, K. J., & Koplin, J. J. (2019). What can urban/rural differences in food allergy prevalence tell us about the drivers of food allergy? *Journal of Allergy and Clinical Immunology*, *143*(2), 554–556. https://doi.org/10.1016/j.jaci.2018.10.055

Anvari, S., Miller, J., Yeh, C. Y., & Davis, C. M. (2019). IgE-Mediated Food Allergy. *Clinical Reviews in Allergy and Immunology*, *57*(2), 244–260. https://doi.org/10.1007/s12016-018-8710-3

Apostolovic, D., Tran, T. A. T., Hamsten, C., Starkhammar, M., Cirkovic Velickovic, T., & van Hage, M. (2014). Immunoproteomics of processed beef proteins reveal novel galactose-α-

1,3-galactose-containing allergens. *Allergy*, *69*(10), 1308–1315. https://doi.org/10.1111/all.12462

Bannon, G. A. (2004). What makes a food protein an allergen? *Current Allergy and Asthma Reports*, *4*(1), 43–46. https://doi.org/10.1007/s11882-004-0042-0

Barni, S., Liccioli, G., Sarti, L., Giovannini, M., Novembre, E., & Mori, F. (2020). Immunoglobulin E (IgE)-mediated food allergy in children: Epidemiology, pathogenesis, diagnosis, prevention, and management. *Medicina (Lithuania)*, *56*(3). https://doi.org/10.3390/medicina56030111

Bartra, J., García-Moral, A., & Enrique, E. (2016). Geographical differences in food allergy. *Bundesgesundheitsbl*, *59*(6), 755–763. https://doi.org/10.1007/s00103-016-2357-0

Beyer, K., Morrowa, E., Li, X.-M., Bardina, L., Bannon, G. A., Burks, A. W., & Sampson, H. A. (2001). Effects of cooking methods on peanut allergenicity. *Journal of Allergy and Clinical Immunology*, *107*(6), 1077–1081. https://doi.org/10.1067/mai.2001.115480

Bignardi, D., Comite, P., Mori, I., Ferrero, F., Fontana, V., Bruzzone, M., Mussap, M., & Ciprandi, G. (2019). Allergen-specific IgE: comparison between skin prick test and serum assay in real life. *Allergologie Select*, *3*(01), 9–14. https://doi.org/10.5414/ALX01891E

Bock, S. A., Sampson, H. A., Atkins, F. M., Zeiger, R. S., Lehrer, S., Sachs, M., Bush, R. K., & Metcalfe, D. D. (1988). Double-blind, placebo-controlled food challenge (DBPCFC) as an office procedure: a manual. *The Journal of Allergy and Clinical Immunology*, *82*(6), 986–997. https://doi.org/10.1016/0091-6749(88)90135-2

Boyce, J. A., Assa'ad, A., Burks, A. W., Jones, S. M., Sampson, H. A., Wood, R. A., Plaut, M.,

Cooper, S. F., Fenton, M. J., Arshad, S. H., Bahna, S. L., Beck, L. A., Byrd-Bredbenner, C., Camargo, C. A. J., Eichenfield, L., Furuta, G. T., Hanifin, J. M., Jones, C., Kraft, M., … Schwaninger, J. M. (2010). Guidelines for the diagnosis and management of food allergy in the United States: report of the NIAID-sponsored expert panel. *The Journal of Allergy and Clinical Immunology*, *126*(6 Suppl), S1-58. https://doi.org/10.1016/j.jaci.2010.10.007

Brodkorb, A., Egger, L., Alminger, M., Alvito, P., Assunção, R., Ballance, S., Bohn, T., Bourlieu-Lacanal, C., Boutrou, R., Carrière, F., Clemente, A., Corredig, M., Dupont, D., Dufour, C., Edwards, C., Golding, M., Karakaya, S., Kirkhus, B., Le Feunteun, S., … Recio, I. (2019). INFOGEST static in vitro simulation of gastrointestinal food digestion. *Nature Protocols*, *14*(4), 991–1014. https://doi.org/10.1038/s41596-018-0119-1

Brunelle, J. L., & Green, R. (2014). Coomassie blue staining. In *Methods in Enzymology* (1st ed., Vol. 541). Elsevier Inc. https://doi.org/10.1016/B978-0-12-420119-4.00013-6

Bucaite, G., Kang-Pettinger, T., Moreira, J., Gould, H. J., James, L. K., Sutton, B. J., & McDonnell, J. M. (2019). Interplay between Affinity and Valency in Effector Cell Degranulation: A Model System with Polcalcin Allergens and Human Patient–Derived IgE Antibodies. *The Journal of Immunology*, *203*(7), 1693–1700. https://doi.org/10.4049/jimmunol.1900509

Bufe, A. (2001). Significance of IgE-binding epitopes in allergic disease. *Journal of Allergy and Clinical Immunology*, *107*(2), 219–221. https://doi.org/10.1067/mai.2001.112850

Burks, A. W., Jones, S. M., Boyce, J. A., Sicherer, S. H., Wood, R. A., Assa'ad, A., & Sampson, H. A. (2011). NIAID-sponsored 2010 guidelines for managing food allergy: applications in the pediatric population. *Pediatrics*, *128*(5), 955–965. https://doi.org/10.1542/peds.2011-

0539

Caubet, J.-C., & Wang, J. (2011). Current Understanding of Egg Allergy. *Pediatric Clinics of North America*, *58*(2), 427–443. https://doi.org/10.1016/j.pcl.2011.02.014

Chen, C., Hou, J., Tanner, J. J., & Cheng, J. (2020). Bioinformatics Methods for Mass Spectrometry-Based Proteomics Data Analysis. *International Journal of Molecular Sciences*, *21*(8), 2873. https://doi.org/10.3390/ijms21082873

Clare Mills, E. N., Sancho, A. I., Rigby, N. M., Jenkins, J. A., & Mackie, A. R. (2009). Impact of food processing on the structural and allergenic properties of food allergens. *Molecular Nutrition and Food Research*, *53*(8), 963–969. https://doi.org/10.1002/mnfr.200800236

D'Amato, G., & Akdis, C. A. (2020). Global warming, climate change, air pollution and allergies. *Allergy: European Journal of Allergy and Clinical Immunology*, *75*(9), 2158–2160. https://doi.org/10.1111/all.14527

Datema, M. R., Zuidmeer-Jongejan, L., Asero, R., Barreales, L., Belohlavkova, S., de Blay, F., Bures, P., Clausen, M., Dubakiene, R., Gislason, D., Jedrzejczak-Czechowicz, M., Kowalski, M. L., Knulst, A. C., Kralimarkova, T., Le, T.-M., Lovegrove, A., Marsh, J., Papadopoulos, N. G., Popov, T., … van Ree, R. (2015). Hazelnut allergy across Europe dissected molecularly: A EuroPrevall outpatient clinic survey. *Journal of Allergy and Clinical Immunology*, *136*(2), 382–391. https://doi.org/10.1016/j.jaci.2014.12.1949

De Martinis, M., Sirufo, M. M., Viscido, A., & Ginaldi, L. (2020). Food Allergy Insights: A Changing Landscape. *Archivum Immunologiae et Therapiae Experimentalis*, *68*(2), 1–15. https://doi.org/10.1007/s00005-020-00574-6

Di Stasio, L. (2018). Intolerances and food allergies : assessment of the stability of allergenic proteins to gastrointestinal digestion. *FedOA - University of Naples Federico II Open Archive*, 1–158.

FAO/WHO. (2021). *Ad hoc Joint FAO/WHO Expert Consultation on Risk Assessment of Food Allergens Part 1: Review and validation of Codex priority allergen list through risk assessment*.

Fatteh, S., Rekkerth, D. J., & Hadley, J. A. (2014). Skin prick/puncture testing in North America: a call for standards and consistency. *Allergy, Asthma & Clinical Immunology*, *10*(1), 44. https://doi.org/10.1186/1710-1492-10-44

Fernandez-Riva, M., & Miles, S. (2003). Chapter 1: Food Allergies: Clinical and Psychosocial Perspectives. *Plant Food Allergens*, 1–23. https://doi.org/10.1002/9780470995174

Fernandez, A., Mills, E. N. C., Koning, F., & Moreno, F. J. (2021). Allergenicity Assessment of Novel Food Proteins: What Should Be Improved? *Trends in Biotechnology*, *39*(1), 4–8. https://doi.org/10.1016/j.tibtech.2020.05.011

Gupta, R. S., Warren, C. M., Smith, B. M., Jiang, J., Blumenstock, J. A., Davis, M. M., Schleimer, R. P., & Nadeau, K. C. (2019). Prevalence and Severity of Food Allergies Among US Adults. *JAMA Network Open*, *2*(1), e185630. https://doi.org/10.1001/jamanetworkopen.2018.5630

Henchion, M., Hayes, M., Mullen, A., Fenelon, M., & Tiwari, B. (2017). Future Protein Supply and Demand: Strategies and Factors Influencing a Sustainable Equilibrium. *Foods*, *6*(7), 53. https://doi.org/10.3390/foods6070053

Houben, G., Blom, M., Alvito, P., Assunção, R., Crevel, R., Fæste, C. K., Le, T. M., Madsen, C. B., Remington, B., Stroheker, T., Vassilopoulou, E., Verhoeckx, K., Žiarovská, J., & Constable, A. (2019). Defining the targets for the assessment of IgE-mediated allergenicity of new or modified food proteins. *Food and Chemical Toxicology*, *127*(March), 61–69. https://doi.org/10.1016/j.fct.2019.02.036

Huby, R. D. J., Dearman, R. J., & Kimber, I. (2000). Why are some proteins allergens? *Toxicological Sciences*, *55*(2), 235–246. https://doi.org/10.1093/toxsci/55.2.235

Huijbers, G. B., Colen, A. A., Jansen, J. J., Kardinaal, A. F., Vlieg-Boerstra, B. J., & Martens, B. P. (1994). Masking foods for food challenge: practical aspects of masking foods for a double-blind, placebo-controlled food challenge. *Journal of the American Dietetic Association*, *94*(6), 645–649. https://doi.org/10.1016/0002-8223(94)90163-5

Issaq, H. J., & Veenstra, T. D. (2008). Two-dimensional polyacrylamide gel electrophoresis (2D-PAGE): advances and perspectives. *BioTechniques*, *44*(5), 697–700. https://doi.org/10.2144/000112823

Iweala, O. I., & Burks, A. W. (2016). Food Allergy: Our Evolving Understanding of Its Pathogenesis, Prevention, and Treatment. *Current Allergy and Asthma Reports*, *16*(5), 37. https://doi.org/10.1007/s11882-016-0616-7

Konstantinou, G. N. (2017). Enzyme-Linked Immunosorbent Assay (ELISA). In *Food Allergens* (pp. 79–94). https://doi.org/10.1007/978-1-4939-6925-8_7

Kruger, N. J. (1996). The Bradford Method for Protein Quantitation. *The Protein Protocols Handbook*, *32*, 15–20. https://doi.org/10.1007/978-1-60327-259-9_4

Kumar, S., Verma, A. K., Das, M., & Dwivedi, P. D. (2012). Molecular mechanisms of IgE mediated food allergy. *International Immunopharmacology*, *13*(4), 432–439. https://doi.org/10.1016/j.intimp.2012.05.018

Lacy, A., & O'Kennedy, R. (2005). IMMUNOASSAYS, TECHNIQUES | Radioimmunoassays. In *Encyclopedia of Analytical Science* (pp. 335–344). Elsevier. https://doi.org/10.1016/B0-12-369397-7/00269-7

Leip, A., Billen, G., Garnier, J., Grizzetti, B., Lassaletta, L., Reis, S., Simpson, D., Sutton, M. A., de Vries, W., Weiss, F., & Westhoek, H. (2015). Impacts of European livestock production: nitrogen, sulphur, phosphorus and greenhouse gas emissions, land-use, water eutrophication and biodiversity. *Environmental Research Letters*, *10*(11), 115004. https://doi.org/10.1088/1748-9326/10/11/115004

Liu, C., & Sathe, S. K. (2018). Food Allergen Epitope Mapping. *Journal of Agricultural and Food Chemistry*, *66*(28), 7238–7248. https://doi.org/10.1021/acs.jafc.8b01967

Mackie, A., Dupont, D., Torcello-Gómez, A., Jardin, J., & Deglaire, A. (2019). Report on EFSA project OC/EFSA/GMO/2017/01 "In vitro protein digestibility" (Allergestion). *EFSA Supporting Publications*, *16*(12), 1–82. https://doi.org/10.2903/sp.efsa.2019.en-1765

Maleki, S. J., Chung, S.-Y., Champagne, E. T., & Raufman, J.-P. (2000). The effects of roasting on the allergenic properties of peanut proteins. *Journal of Allergy and Clinical Immunology*, *106*(4), 763–768. https://doi.org/10.1067/mai.2000.109620

Moens, L., & Tangye, S. G. (2014). Cytokine-Mediated Regulation of Plasma Cell Generation: IL-21 Takes Center Stage. *Frontiers in Immunology*, *5*. https://doi.org/10.3389/fimmu.2014.00065

Mondoulet, L., Paty, E., Drumare, M. F., Ah-Leung, S., Scheinmann, P., Willemot, R. M., Wal, J. M., & Bernard, H. (2005). Influence of Thermal Processing on the Allergenicity of Peanut Proteins. *Journal of Agricultural and Food Chemistry*, *53*(11), 4547–4553. https://doi.org/10.1021/jf050091p

Morelli, A. E., Larregina, A. T., Shufesky, W. J., Sullivan, M. L. G., Stolz, D. B., Papworth, G. D., Zahorchak, A. F., Logar, A. J., Wang, Z., Watkins, S. C., Falo, L. D., & Thomson, A. W. (2004). Endocytosis, intracellular sorting, and processing of exosomes by dendritic cells. *Blood*, *104*(10), 3257–3266. https://doi.org/10.1182/blood-2004-03-0824

Naegeli, H., Bresson, J., Dalmay, T., Dewhurst, I. C., Epstein, M. M., Firbank, L. G., Guerche, P., Hejatko, J., Moreno, F. J., Mullins, E., Nogué, F., Rostoks, N., Sánchez Serrano, J. J., Savoini, G., Veromann, E., Veronesi, F., & Dumont, A. F. (2021). Statement on in vitro protein digestibility tests in allergenicity and protein safety assessment of genetically modified plants. *EFSA Journal*, *19*(1). https://doi.org/10.2903/j.efsa.2021.6350

Nakamura, A., Watanabe, K., Ojima, T., Ahn, D.-H., & Saeki, H. (2005). Effect of Maillard Reaction on Allergenicity of Scallop Tropomyosin. *Journal of Agricultural and Food Chemistry*, *53*(19), 7559–7564. https://doi.org/10.1021/jf0502045

Noè, D., Bartemucci, L., Mariani, N., & Cantari, D. (1998). Practical aspects of preparation of foods for double-blind, placebo-controlled food challenge. *Allergy*, *53*(46 Suppl), 75–77. https://doi.org/10.1111/j.1398-9995.1998.tb04967.x

Panjari, M., Koplin, J. J., Dharmage, S. C., Peters, R. L., Gurrin, L. C., Sawyer, S. M., McWilliam, V., Eckert, J. K., Vicendese, D., Erbas, B., Matheson, M. C., Tang, M. L. K., Douglass, J., Ponsonby, A.-L., Dwyer, T., Goldfeld, S., & Allen, K. J. (2016). Nut allergy

prevalence and differences between Asian-born children and Australian-born children of Asian descent: a state-wide survey of children at primary school entry in Victoria, Australia. *Clinical & Experimental Allergy*, *46*(4), 602–609. https://doi.org/10.1111/cea.12699

Pekar, J., Ret, D., & Untersmayr, E. (2018). Stability of allergens. *Molecular Immunology*, *100*(March), 14–20. https://doi.org/10.1016/j.molimm.2018.03.017

Popescu, F.-D. (2015). Cross-reactivity between aeroallergens and food allergens. *World Journal of Methodology*, *5*(2), 31. https://doi.org/10.5662/wjm.v5.i2.31

Ring, J., Beyer, K., Biedermann, T., Bircher, A., Duda, D., Fischer, J., Friedrichs, F., Fuchs, T., Gieler, U., Jakob, T., Klimek, L., Lange, L., Merk, H. F., Niggemann, B., Pfaar, O., Przybilla, B., Ruëff, F., Rietschel, E., Schnadt, S., … Brockow, K. (2014). Guideline for acute therapy and management of anaphylaxis. *Allergo Journal International*, *23*(3), 96–112. https://doi.org/10.1007/s40629-014-0009-1

Sakamoto, S., Putalun, W., Vimolmangkang, S., Phoolcharoen, W., Shoyama, Y., Tanaka, H., & Morimoto, S. (2018). Enzyme-linked immunosorbent assay for the quantitative/qualitative analysis of plant secondary metabolites. *Journal of Natural Medicines*, *72*(1), 32–42. https://doi.org/10.1007/s11418-017-1144-z

Santos, A. F., Alpan, O., & Hoffmann, H. (2021). Basophil activation test: Mechanisms and considerations for use in clinical trials and clinical practice. *Allergy*, *76*(8), 2420–2432. https://doi.org/10.1111/all.14747

Sathe, S. K., Teuber, S. S., & Roux, K. H. (2005). Effects of food processing on the stability of food allergens. *Biotechnology Advances*, *23*(6), 423–429. https://doi.org/10.1016/j.biotechadv.2005.05.008

Scherp, P., Ku, G., Coleman, L., & Kheterpal, I. (2011). *Gel-Based and Gel-Free Proteomic Technologies* (pp. 163–190). https://doi.org/10.1007/978-1-61737-960-4_13

Schoemaker, A. A., Sprikkelman, A. B., Grimshaw, K. E., Roberts, G., Grabenhenrich, L., Rosenfeld, L., Siegert, S., Dubakiene, R., Rudzeviciene, O., Reche, M., Fiandor, A., Papadopoulos, N. G., Malamitsi-Puchner, A., Fiocchi, A., Dahdah, L., Sigurdardottir, S. T., Clausen, M., Stańczyk-Przyłuska, A., Zeman, K., … Beyer, K. (2015). Incidence and natural history of challenge-proven cow's milk allergy in European children - EuroPrevall birth cohort. *Allergy*, *70*(8), 963–972. https://doi.org/10.1111/all.12630

Shevchenko, A., Tomas, H., Havli, J., Olsen, J. V, & Mann, M. (2006). In-gel digestion for mass spectrometric characterization of proteins and proteomes. *Nature Protocols*, *1*(6), 2856–2860. https://doi.org/10.1038/nprot.2006.468

Sicherer, S. H., & Sampson, H. A. (2014). Food allergy: Epidemiology, pathogenesis, diagnosis, and treatment. *Journal of Allergy and Clinical Immunology*, *133*(2), 291-307.e5. https://doi.org/10.1016/j.jaci.2013.11.020

Sicherer, S. H., & Sampson, H. A. (2018). Food allergy: A review and update on epidemiology, pathogenesis, diagnosis, prevention, and management. *Journal of Allergy and Clinical Immunology*, *141*(1), 41–58. https://doi.org/10.1016/j.jaci.2017.11.003

Teodorowicz, M., Jansen, A. P. H., Roovers, M. H. W. M., Ruinemans-Koerts, J., Wichers, H. J., & Savelkoul, H. F. J. (2015). Maillard-type neoallergens present in processed soy extract may cause an allergic reaction in soy allergic patients. *Clinical and Translational Allergy*, *5*(S3). https://doi.org/10.1186/2045-7022-5-S3-P21

Tomé, D. (2013). Digestibility Issues of Vegetable versus Animal Proteins: Protein and Amino

Acid Requirements—Functional Aspects. *Food and Nutrition Bulletin*, *34*(2), 272–274. https://doi.org/10.1177/156482651303400225

Valenta, R., Hochwallner, H., Linhart, B., & Pahr, S. (2015). Food allergies: The basics. *Gastroenterology*, *148*(6), 1120-1131.e4. https://doi.org/10.1053/j.gastro.2015.02.006

Van Den Eeckhout, B., Tavernier, J., & Gerlo, S. (2021). Interleukin-1 as Innate Mediator of T Cell Immunity. *Frontiers in Immunology*, *11*. https://doi.org/10.3389/fimmu.2020.621931

Van Ree, R., & Aalberse, R. C. (1999). Specific IgE without clinical allergy. *Journal of Allergy and Clinical Immunology*, *103*(6), 1000–1001. https://doi.org/10.1016/S0091-6749(99)70169-7

van Ree, Ronald. (2002). Carbohydrate Epitopes and Their Relevance for the Diagnosis and Treatment of Allergic Diseases. *International Archives of Allergy and Immunology*, *129*(3), 189–197. https://doi.org/10.1159/000066770

Verhoeckx, K., Bøgh, K. L., Dupont, D., Egger, L., Gadermaier, G., Larré, C., Mackie, A., Menard, O., Adel-Patient, K., Picariello, G., Portmann, R., Smit, J., Turner, P., Untersmayr, E., & Epstein, M. M. (2019). The relevance of a digestibility evaluation in the allergenicity risk assessment of novel proteins. Opinion of a joint initiative of COST action ImpARAS and COST action INFOGEST. *Food and Chemical Toxicology*, *129*, 405–423. https://doi.org/10.1016/j.fct.2019.04.052

Verhoeckx, K. C. M., van Broekhoven, S., den Hartog-Jager, C. F., Gaspari, M., de Jong, G. A. H., Wichers, H. J., van Hoffen, E., Houben, G. F., & Knulst, A. C. (2014). House dust mite (Der p 10) and crustacean allergic patients may react to food containing Yellow mealworm proteins. *Food and Chemical Toxicology*, *65*, 364–373.

https://doi.org/10.1016/j.fct.2013.12.049

Vila, L., Beyer, K., Järvinen, K.-M., Chatchatee, P., Bardina, L., & Sampson, H. A. (2001). Role of conformational and linear epitopes in the achievement of tolerance in cow's milk allergy. *Clinical & Experimental Allergy*, *31*(10), 1599–1606. https://doi.org/10.1046/j.1365-2222.2001.01218.x

Warren, C. M., Jiang, J., & Gupta, R. S. (2020). Epidemiology and Burden of Food Allergy. *Current Allergy and Asthma Reports*, *20*(2), 1–9. https://doi.org/10.1007/s11882-020-0898-7

Westerhout, J., Krone, T., Snippe, A., Babé, L., McClain, S. U., Ladics, G. S., Houben, G. G., & Verhoeckx, K. C. (2019). Allergenicity prediction of novel and modified proteins: Not a mission impossible! Development of a Random Forest allergenicity prediction model. *Regulatory Toxicology and Pharmacology*, *107*(July), 104422. https://doi.org/10.1016/j.yrtph.2019.104422

Wickham, M., Faulks, R., & Mills, C. (2009). In vitro digestion methods for assessing the effect of food structure on allergen breakdown. *Molecular Nutrition & Food Research*, *53*(8), 952–958. https://doi.org/10.1002/mnfr.200800193

Xepapadaki, P., Fiocchi, A., Grabenhenrich, L., Roberts, G., Grimshaw, K. E. C., Fiandor, A., Larco, J. I., Sigurdardottir, S., Clausen, M., Papadopoulos, N. G., Dahdah, L., Mackie, A., Sprikkelman, A. B., Schoemaker, A. A., Dubakiene, R., Butiene, I., Kowalski, M. L., Zeman, K., Gavrili, S., … Beyer, K. (2016). Incidence and natural history of hen's egg allergy in the first 2 years of life-the EuroPrevall birth cohort study. *Allergy*, *71*(3), 350–357. https://doi.org/10.1111/all.12801

General Introduction

Zhang, Y., Fonslow, B. R., Shan, B., Baek, M.-C., & Yates, J. R. (2013). Protein Analysis by Shotgun/Bottom-up Proteomics. *Chemical Reviews*, *113*(4), 2343–2394. https://doi.org/10.1021/cr3003533

Web references

http://allergen.org/index.php

http://allergen.org/search.php?allergenname=tri+a+&searchname=Search

http://eur-lex.europa.eu/

https://www.fda.gov/food/food-labeling-nutrition/food-allergies

https://www.overshootday.org/about/

SECTION 1

A novel food of plant origin: *Moringa oleifera* leaf

Moringa oleifera leaf represents a clear example of a novel food already consumed by populations outside EU being widely used in developing and poor countries for the weaning of malnourished infants (Arise et al., 2014). The following *M. oleifera* leaf allergenicity assessment considered the main plant food allergens.

Plant food allergens

Vegetal sources of proteins such as leaves, ancient and hybrid cereals, minor pulses are expanding the catalogue of "novel foods". However, the high content in proteins of these matrices generates concerns in terms of the allergenic potential. Below are reported the main plant food allergens that can be found in leaves.

The prolamin superfamily

Prolamins owe their name to the exorbitant amount of glutamine and proline in their structures, along with their high solubility in alcohol/water mixtures. The prolamin superfamily comprises three types of plant-derived food allergen proteins, namely nonspecific lipid transfer proteins (nsLTPs), α-amylase/trypsin inhibitors, and 2S albumin.

Despite the low sequence homology degree, every protein of the prolamin superfamily shows stability to heat and proteolysis. Their resistance is because of a common and well-conserved eight

cysteine skeleton (Nguyen et al., 2014) (**Figure 1**), and a very similar α-helical structure linked by disulphide bonds.

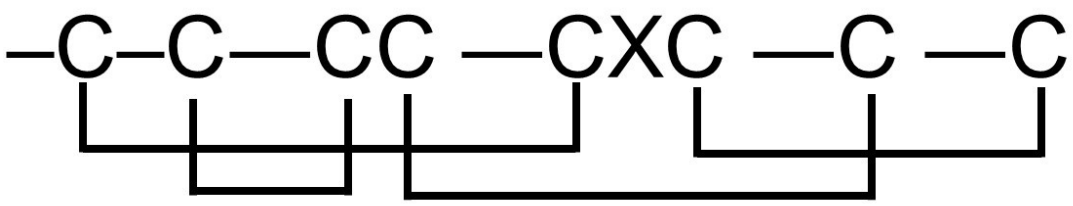


Figure 1. Eight cysteine skeleton of nsLTP

Essential for the transportation of membrane lipids and phospholipids, nsLTPs make up 4% of all soluble proteins in higher plants (Liu et al., 2015).

Peaches (Pru 7), apples (Mal 3) and apricots (Pru ar 3) have been identified as the major food allergens in Rosaceae fruits (Egger et al., 2010). NsLTPs are widespread in other plant-derived foods, such as vegetables, grains, and nuts, sharing a high sequence of identity and similarity. There are two types of nsLTPs: nsLTP1 and nsLTP2, which have molecular weights of 9 kDa and 7 kDa, respectively. The spatial structure of nsLTPs is stabilised by four internal disulphide bonds between four α-helices that form an internal cavity, which is covered by the side chains of hydrophobic residues of the amphipathic helices and the C-terminal region. Notably, these hydrophobic residues that represent potent cleavage sites for pepsin, are hidden inside the nsLTP cavity. This particular conformation, makes these amino acids inaccessible to the proteolytic enzyme (Asero et al., 2000). Structural analyses showed that the cavity of nsLTP can adapt its volume to bind one or two monoacyl lipids. As a result, the folding of nsLTP gives rise to proteins with lipid-binding properties that make them more stable and resistant to proteolysis by digestive

enzymes. This, along with the high thermal stability, are very important characteristics of small allergenic proteins.

Recently, the discovery that nsLTPs and structurally related plant proteins are considered by the scientific community as panallergen in plant-derived foods, raised new concerns about the application of plant biotechnology and plant breeding programmes devoted to this specific family of food proteins (Díaz-Perales et al., 2000; McKenna et al., 2016). As a consequence of their extracellular localization and ability to bind lipids and fatty acids, nsLTPs are thought to play a role in the transport of lipophilic compounds for epicuticular wax formation (Sterk et al., 1991). For this reason, besides seeds and peels, nsLTP has been found also in leaves (Hartz et al., 2007; Pyee et al., 1994).

Chitinases

Class I chitinases found in plants have been implicated as major allergens responsible for the latex-plant food cross-reactivity known as latex-fruit syndrome (Diaz-Perales et al., 1999). IgE from patients with latex allergies appears to recognize mainly epitopes in their N-terminal regions, which are homologous to hevein, the primary allergen in latex. Based on similarities in their amino acid structures and consensus sequences, plant chitinases have been classified into six classes (I-VI) (Patil et al., 2000). These enzymes have a molecular weight of approximately 33 kDa and share an N-terminal domain, the hevein-like domain, of approximately 40 amino acid residues with almost 70% sequence identity with latex hevein (Mills & Shewry, 2003). A hinge rich in glycine and proline or hydroxyproline residues links this N-terminal domain to the catalytic domain. Other chitinases with one or two hevein domains are placed in classes IV and V, respectively. Catalytic domains of class II chitinases and class IV chitinases are homologous to

that of class I chitinases (except for a different number of amino acids). Owing to its catalytic activity against the N-Acetylglucosamine (GlcNAc)-containing oligosaccharides and the β-1,4 linkages of chitin, plant chitinases serve primarily as a defence mechanism against predators, in which chitin represents the main constituent (e.g., insect exoskeleton, fungi cell wall and secta). For this reason, chitinases were categorised as pathogen-related (PR) proteins that are induced by wounding, infection, or pest attack and are toxic to pests and pathogens. Notably, several well-characterized classes of allergens seem to have protective roles (Hoffmann-Sommergruber, 2002; Sinha et al., 2014). Several fruit chitinases were cross-reactive in latex sensitised patients. In fact, IgE anti hevein (Hev b 6.02), the major allergen of natural rubber latex, reacts with fruit class I chitinases containing hevein-like domains (Wagner & Breiteneder, 2002). This phenomenon, also known as fruit-latex syndrome, has been linked to an increasing number of plant sources, including avocados, bananas, chestnuts, kiwis, peaches, tomatoes, potatoes, and bell peppers (Salcedo et al., 2001).

References

Arise, A. K., Arise, R. O., Sanusi, M. O., Esan, O. T., & Oyeyinka, S. A. (2014). Effect of Moringa oleifera flower fortification on the nutritional quality and sensory properties of weaning food. *Croatian Journal of Food Science and Technology*, *6*(2), 65–71. https://doi.org/10.17508/cjfst.2014.6.2.01

Asero, R., Mistrello, G., Roncarolo, D., de Vries, S. C., Gautier, M.-F., Ciurana, C. L. F., Verbeek, E., Mohammadi, T., Knul-Brettlova, V., Akkerdaas, J. H., Bulder, I., Aalberse, R. C., & van Ree, R. (2000). Lipid Transfer Protein: A Pan-Allergen in Plant-Derived Foods That Is Highly Resistant to Pepsin Digestion. *International Archives of Allergy and*

Immunology, *122*(1), 20–32. https://doi.org/10.1159/000024355

Diaz-Perales, A., Collada, C., Blanco, C., Sanchez-Monge, R., Carrillo, T., Aragoncillo, C., & Salcedo, G. (1999). Cross-reactions in the latex-fruit syndrome: A relevant role of chitinases but not of complex asparagine-linked glycans. *Journal of Allergy and Clinical Immunology*, *104*(3), 681–687. https://doi.org/10.1016/S0091-6749(99)70342-8

DÍaz-Perales, A., Lombardero, M., SÁnchez-Monge, R., GarcÍa-Selles, F. J., Pernas, M., FernÁndez-Rivas, M., Barber, D., & Salcedo, G. (2000). Lipid-transfer proteins as potential plant panallergens: cross-reactivity among proteins of Artemisia pollen, Castanea nut and Rosaceae fruits, with different IgE-binding capacities. *Clinical & Experimental Allergy*, *30*(10), 1403–1410. https://doi.org/10.1046/j.1365-2222.2000.00909.x

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*(5), 326–335. https://doi.org/10.1007/s11882-010-0128-9

Hartz, C., del Mar San Miguel-Moncín, M., Cisteró-Bahíma, A., Fötisch, K., Metzner, K. J., Fortunato, D., Lidholm, J., Vieths, S., & Scheurer, S. (2007). Molecular characterisation of Lac s 1, the major allergen from lettuce (Lactuca sativa). *Molecular Immunology*, *44*(11), 2820–2830. https://doi.org/10.1016/j.molimm.2007.01.030

Hoffmann-Sommergruber, K. (2002). Pathogenesis-related (PR)-proteins identified as allergens. *Biochemical Society Transactions*, *30*(6), 930–935. https://doi.org/10.1042/bst0300930

Liu, F., Zhang, X., Lu, C., Zeng, X., Li, Y., Fu, D., & Wu, G. (2015). Non-specific lipid transfer proteins in plants: presenting new advances and an integrated functional analysis. *Journal of Experimental Botany*, *66*(19), 5663–5681. https://doi.org/10.1093/jxb/erv313

McKenna, O. E., Asam, C., Araujo, G. R., Roulias, A., Goulart, L. R., & Ferreira, F. (2016). How relevant is panallergen sensitization in the development of allergies? *Pediatric Allergy and Immunology*, *27*(6), 560–568. https://doi.org/10.1111/pai.12589

Mills, E. N. C., & Shewry, P. R. (Eds.). (2003). *Plant Food Allergens*. Wiley. https://doi.org/10.1002/9780470995174

Nguyen, P. Q. T., Wang, S., Kumar, A., Yap, L. J., Luu, T. T., Lescar, J., & Tam, J. P. (2014). Discovery and characterization of pseudocyclic cystine-knot α-amylase inhibitors with high resistance to heat and proteolytic degradation. *FEBS Journal*, *281*(19), 4351–4366. https://doi.org/10.1111/febs.12939

Patil, R. S., Ghormade, V., & Deshpande, M. V. (2000). Chitinolytic enzymes: an exploration. *Enzyme and Microbial Technology*, *26*(7), 473–483. https://doi.org/10.1016/S0141-0229(00)00134-4

Pyee, J., Yu, H., & Kolattukudy, P. E. (1994). Identification of a Lipid Transfer Protein as the Major Protein in the Surface Wax of Broccoli (Brassica oleracea) Leaves. *Archives of Biochemistry and Biophysics*, *311*(2), 460–468.

Salcedo, G., Diaz-Perales, A., & Sanchez-Monge, R. (2001). The role of plant panallergens in sensitization to natural rubber latex. *Current Opinion in Allergy and Clinical Immunology*, *1*(2), 177–183. https://doi.org/10.1097/01.all.0000011004.49250.bc

Sinha, M., Singh, R. P., Kushwaha, G. S., Iqbal, N., Singh, A., Kaushik, S., Kaur, P., Sharma, S., & Singh, T. P. (2014). Current Overview of Allergens of Plant Pathogenesis Related Protein Families. *The Scientific World Journal*, *2014*, 1–19. https://doi.org/10.1155/2014/543195

Sterk, P., Booij, H., Schellekens, G. A., Van Kammen, A., & De Vries, S. C. (1991). Cell-specific expression of the carrot EP2 lipid transfer protein gene. *The Plant Cell*, *3*(9), 907–921. https://doi.org/10.1105/tpc.3.9.907

Wagner, S., & Breiteneder, H. (2002). The latex-fruit syndrome. *Biochemical Society Transactions*, *30*(6), 935–940. https://doi.org/10.1042/bst0300935

Chapter 1

Moringa oleifera Lam. proteins: properties and food applications

Moringa oleifera Lam. Proteins: Properties and Food Applications

Giovanni D'Auria, Chiara Nitride, and Pasquale Ferranti, Department of Agricultural Sciences, University of Naples Federico II, Naples, Italy

Key Points

- *Moringa oleifera* seeds and leaves are nutritious tissues as they are rich in proteins
- Moringa's seed and leaf powders are being used in feed and food production
- Moringa seeds proteins are useful in purifying water
- Moringa leaves proteins are related to primary metabolism and defense

Abstract

Moringa oleifera has been and is used in Asian traditional medicine as a *panacea* to treat and prevent several diseases and is part of the oriental diet both as a raw and cooked ingredient. The ability of the plant to resist unfavorable environmental conditions, such as drought and high temperatures, makes it a sustainable agricultural choice for the future. The high amounts of phytochemicals such as polyphenols, glucosinolates, isothiocyanates, which have been extensively characterized, make Moringa leaf an ideal ingredient in cosmetics, and phytopharmacology, etc. In addition, Moringa leaves have an abundant protein content, making them a nutritious food. The characterization of *M. oleifera*'s proteome has been neglected despite the relatively high abundance of proteins, which is unusual for a plant leaf. This chapter aims at a comprehensive review of the information available in literature on the proteinaceous tissues of Moringa (i.e., seeds and leaves). Nutritional and food safety information about Moringa leaves and seed proteins were collected together with research gaps, particularly those related to their technological application in the food industry.

Chapter 2

Identification of potential allergens in a novel food ingredient from Moringa oleifera leaves

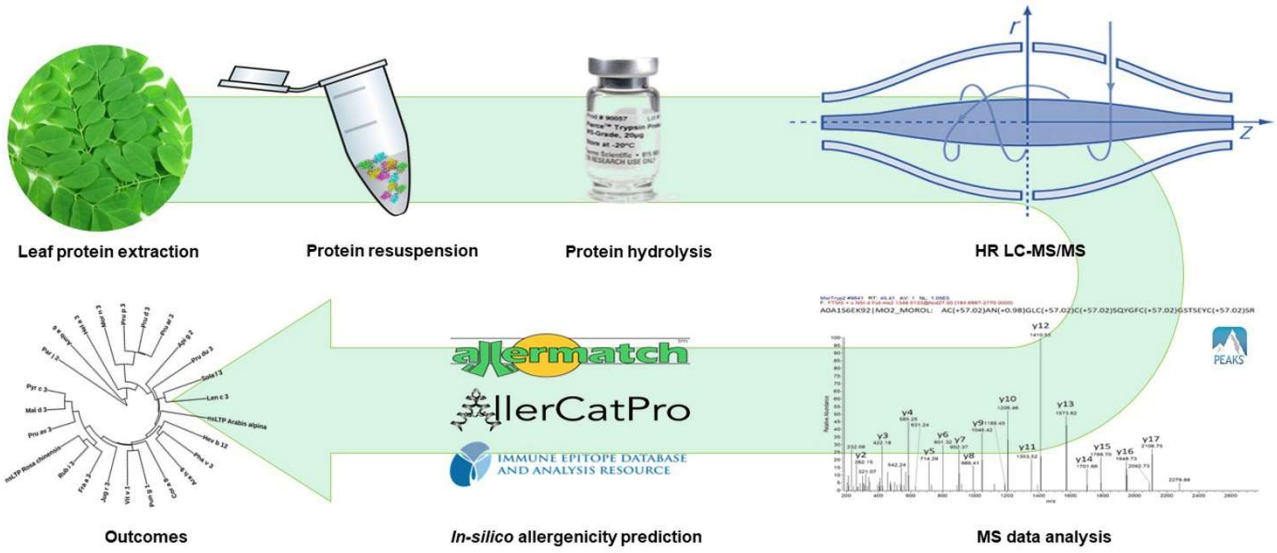
Submitted to Food Chemistry

Food Chemistry

Identification of potential allergens in a novel food ingredient from Moringa oleifera leaves

--Manuscript Draft--

Manuscript Number:	FOODCHEM-D-21-10187
Article Type:	Research Article (max 7,500 words)
Keywords:	De-novo sequencing; Hev b 6.02; leaf proteome; Morintides; nsLTPs; Proteomics
Corresponding Author:	Chiara Nitride University of Naples Federico II: Universita degli Studi di Napoli Federico II Portici , Naples ITALY
First Author:	Giovanni D'Auria
Order of Authors:	Giovanni D'Auria
	Chiara Nitride
	Maria Adalgisa Nicolai
	Gianfranco Mamone
	Domenico Montesano
	E. N. Clare Mills
	Pasquale Ferranti
Abstract:	Background: Alternative sources of edible proteins are required to feed the world's growing population. Moringa oleifera leaves are a source of proteins with high biological value, considered as a novel food in the EU and UK. An assessment of the potential allergenicity of Moringa leaf proteins is required. Methods: Proteins from Moringa leaf powder were characterised using traditional proteomic approaches. The proteins identified were evaluated for their allergenic potential using in-silico tools. Results: The main proteins identified belonged to photosynthetic and metabolic pathways. In-silico analysis of the leaf proteome identified Moritides as potential allergens by homology with a latex allergen implicated in fruit-latex-fruit allergy. This analysis also identified a nsLTP, a major panallergen in food. The presence of these putative allergens was confirmed by de-novo sequencing. Conclusions: Our study allowed identification of putative allergens, Morintides and nsLTP. Further in-vitro and in-vivo investigations are required to confirm their allergenic potential.
Suggested Reviewers:	Richard Goodman University of Nebraska-Lincoln rgoodman2@unl.edu
	Jens Brockmeyer University of Stuttgart: Universitat Stuttgart jens.brockmeyer@lc.uni-stuttgart.de
	Kitty Verhoeckx Utrecht University: Universiteit Utrecht kverhoe2@umcutrecht.nl
	Karin Hoffmann-Sommergruber Medical University of Vienna: Medizinische Universitat Wien karin.hoffmann@meduniwien.ac.at
	Michelle Colgrave Edith Cowan University m.colgrave@ecu.edu.au



Graphical abstract - Identification of potential allergens in a novel food ingredient from Moringa oleifera leaves

Highlights

- Phenol extraction with Urea solubilisation is effective for leaf proteomics.
- Moringa leaves proteome consists mainly of photosynthetic and metabolic proteins.
- Morintides share high sequence homology with Hev b 6, a contact allergen of latex.
- The NsLTP, a food panallergen, was identified by homology with *Rosa chinensis*.

Identification of potential allergens in a novel food ingredient from *Moringa oleifera* leaves.

Giovanni D'Auria[1], Chiara Nitride[1*], Maria Adalgisa Nicolai[1], Gianfranco Mamone[2], Domenico Montesano[3], E. N. Clare Mills[4], Pasquale Ferranti[1,2]

Affiliations:

[1]Department of Agricultural Sciences, University of Naples Federico II, Portici, Italy.

[2]Institute of Food Science and Technology (ISA), Avellino, Italy.

[3]Department of Pharmacy, University of Naples Federico II, Naples, Italy.

[4]Division of Infection, Immunity and Respiratory Medicine, School of Biological Sciences, Manchester Academic Health Sciences Centre, Manchester Institute of Biotechnology, University of Manchester, Manchester, UK

*Corresponding author. Email address: chiara.nitride@unina.it

Keywords: *De-novo sequencing; Hev b 6.02; leaf proteome; Morintides; nsLTPs; proteomics.*

Abbreviations

2-DE: two-dimensional electrophoresis

ACN: Acetonitrile

CCDs: Cross-reactive carbohydrate determinants

CHAPS: 3-[(3-Cholamidopropyl) dimethylammonium]-1-propanesulfonate

CHTRY: Chymotrypsin

DTT: dithiothreitol

EDTA: ethylenediaminetetraacetic acid

GdnHCl_b: guanidine hydrochloride buffer

IAA: iodoacetamide

IEF: Isoelectric focusing

IPG: immobilized pH gradient

MPF: Moringa protein fraction

MS: Mass Spectrometry

ns-LTPs: nonspecific lipid transfer proteins

RuBisCO: ribulose biphosphate carboxylase/oxygenase

SDS: sodium dodecyl sulphate

TCA: Trichloroacetic acid

TFA: trifluoroacetic acid

Tris-HCl: tris(hydroxymethyl) aminomethane hydrochloride

TRY: Trypsin

U/T_b: Urea/Thiourea buffer

Abstract

Background: Alternative sources of edible proteins are required to feed the world's growing population. *Moringa oleifera* leaves are a source of proteins with high biological value, considered as a novel food in the EU and UK. An assessment of the potential allergenicity of Moringa leaf proteins is required.

Methods: Proteins from *Moringa* leaf powder were characterised using traditional proteomic approaches. The proteins identified were evaluated for their allergenic potential using *in-silico* tools.

Results: The main proteins identified belonged to photosynthetic and metabolic pathways. In-silico analysis of the leaf proteome identified Moritides as potential allergens by homology with a latex allergen implicated in fruit-latex-fruit allergy. This analysis also identified a nsLTP, a major panallergen in food. The presence of these putative allergens was confirmed by *de-novo* sequencing.

Conclusions: Our study allowed identification of putative allergens, Morintides and nsLTP. Further *in-vitro* and *in-vivo* investigations are required to confirm their allergenic potential.

1 Introduction

Moringa oleifera is believed to be a *panacea* against many diseases and many parts of the plant, including leaves and seeds, have long been used as ingredients in Ayurvedic medicine and supplements for both pregnant women and babies during weaning (Mishra *et al.,* 2012). A fast-growing perennial plant, *M. oleifera* can resist extreme conditions such as drought and high temperatures (Alhakmani *et al.*, 2013). In a world where climate change is having a tremendous impact on food security, this vigorous plant, which is amenable to low-input agriculture (Sujatha & Poonam, 2017), has potential as an economically viable and environmentally beneficial source of novel food ingredients (Montesano *et al.*, 2018). *M. oleifera* leaves and seeds have been explored as innovative and more sustainable protein sources for human and animal nutrition (Abd El-Hack *et al.,* 2018) as they are rich in proteins of high biological value, whose content varies from 14 to 31.4% on a dry basis (Nouman *et al.,* 2016) with an appropriate amino acid content for human nutrition (Olaofe *et al.*, 2013). As it happens for other agricultural products, the composition of Moringa leaves depends on the season and on environmental factors such as temperatures, soil composition and the latitude/ altitude (Shih *et al.,* 2011). *M. oleifera* has a long history of usage in developing countries for the weaning of malnourished infants (Zhou *et al.*, 2012). It is also reported that about 50 g of leaf powder could satisfy a woman's daily iron and calcium requirements during pregnancy and nursing (Mishra *et al.*, 2012), providing essential amino acids, minerals, and vitamins (Rocchetti *et al.*, 2020).

M. oleifera leaves are also a rich source of phytochemicals such as polyphenols, glucosinolates, isothiocyanates (Rocchetti *et al.*, 2020) and is relatively low in anti-nutrients such as saponins, phytate and tannins (Su & Chen, 2020). However, phytochemicals can also pose safety issues with

toxic effects, including cytotoxicity and genotoxicity having been identified (Asare *et al.,* 2012). Consequently, the introduction of alternative protein ingredients like *Moringa* leaf protein powder, into new markets, such as North America and Europe, requires an assessment of its safety. In the USA these novel ingredients require safety evidence in order to achieve the GRAS status, whilst in the EU and UK is required a complete chemical characterisation of the ingredient, alongside the food safety assessment, including toxicological and allergenic risks (EC regulation No 258/97 and EU recommendation 97/618 EC; http://eur-lex.europa.eu/; EFSA 2019). This kind of assessment shares the same approach developed for newly expressed protein in GMOs through transgenesis (EFSA 2010, Fernandez *et al.,* 2021). The allergenicity risk assessment relies on a weight-of-evidence approach starting with a phylogenetic study of the new food ingredient to identify any relationship with known allergenic foods and an evaluation of its capacity to cause allergic disease. Such an assessment can be complemented by other data to characterise any components present in the novel protein ingredient that may be similar to those found in allergenic foods. To date, the only reported case of anaphylactic reaction following consumption of *M. oleifera* was in Australia and associated to the ingestion of pods (Berglund, 2018). Currently, there is no evidence of food allergic reactions caused by the leaves or leaf powder.

Further allergenicity assessment requires an analysis of the *M. oleifera* proteome. However, there are few data and no systematic studies about the protein characterization of the *Moringa* leaf and there is no sequenced genome for the organism. In order to address this our knowledge gap, the first complete protein profiling of *M. oleifera* leaves has been undertaken using advanced MS analysis. These data have been used to undertake an *in-silico* assessment of the allergenic potential of Moringa leaves protein to identify potential cross-reactive allergens and to support the safety assessment of the novel ingredient.

2 Materials and methods

2.1 Materials

The *M. oleifera* leaves were supplied in a dried form by Sud Rienergy S.r.l. - Favella Group, Italy (Corigliano Calabro, Cosenza, Italy – 39°129 13′ 27,69″ N and 9° 01′ 29,69″ E). Moringa leaves were randomly collected in the orchard from different plants from June to September 2019. Intact leaves were dried under the sun until a constant weight was reached. Dried leaves were powdered in a blender and passed through 250 μm sieve to give a powder of 10 ± 1% moisture. Samples were stored in amber glass containers, in a dry place at room temperature, until required. Unless otherwise specified, all the reagents and solvents employed in the present analyses were purchased from Sigma-Aldrich (St. Louis, MO, USA). The solvents employed in this research were of mass spectrometry grade.

2.2 Production and solubilisation of Moringa protein fraction (MPF)

The Moringa protein fraction (MPF) was obtained by extracting the leaf powder following the procedure of Rocco *et al.* (2006), with some modifications (**Figure 1**). To increase the mechanical cell wall disruption, 0.5 g of dried Moringa leaf powder was further ground using a pestle and mortar before suspending in 4.5 ml of extraction buffer (700 mM sucrose, 500 mM Tris-HCl, pH 8.0, 50 mM EDTA, 100 mM KCl, 2% (w/v) 2-Sulfanylethan-1-ol) by vortexing for 15 min. The sample was placed in an ice-bath every 5 min of mixing to avoid overheating. After the addition of an equal volume of phenol equilibrated with 10 mM Tris-HCl pH 8.0 and 1 mM EDTA, the mixture was shaken vigorously for 10 min and centrifuged at 3433 g for 30 min (Centrifuge XS R-8D, Remi Instruments, New Delhi, India) at room temperature (20° C). The supernatant

containing the proteins solubilised in the phenol phase was removed and precipitated by the addition of five volumes of 0.1M CH_3COONH_4 in methanol and left at -20° C for 12 h. The precipitate was collected by centrifugation at 7900 g for 10 min (Microcentrifuge Multispin 12, Steroglass, Perugia, Italy) and the resulting pellet washed once with cold methanol (-20° C) and twice with cold acetone (-20° C). The pellet (MPF) was then dried under gaseous nitrogen and stored at -20 °C until required. Two different buffers were used to re-solubilise the MPF (1:20 w/v): i) chaotropic-charged buffer (GndHCl_b) containing 6M CH_6ClN_3 (guanidinium hydrochloride), 100 mM NH_4HCO_3 pH 8.3 and 10 mM DTT; ii) chaotropic-zwitterionic buffer (U/T_b) including 7M urea, 2M thiourea, 2% (w/v) CHAPS, 50 mM DTT pH 8.8. Proteins were extracted by heating at 60° C for 30 min with sonication in an ultrasonic bath. Protein extracts were alkylated by adding 55 mM IAA and incubated at room temperature (20° C) in the dark for 30 min. Proteins were diluted 20 fold with 50 mM NH_4HCO_3 and 1mM $CaCl_2$ pH 7.8, to reduce the concentration of Guanidine and Urea to < 1M and CHAPS < 0.1% and reduce protease denaturation. Extracts were centrifugated for 15 minutes at 3433 g (Centrifuge XS R-8D, Remi Instruments, New Delhi, India) and the supernatants collected and stored at -20 °C, until required.

2.3 Protein content determination

Two grams of powdered Moringa leaves and 0.2 g of MPF were added to a Kjeldahl tube (VELP scientifica - Italy), containing 0.5 g Cu_2SO_4 and 12 g K_2SO_4 and 20 ml of 96% H_2SO_4. The mineralisation was performed on a heating digester (DK6 VELP Scientifica–Italy) following a thermal ramp: 230° C for 20 min, 290° C for 45 min, 320° C for 35 min, and 420° C for 60 min. The sample was diluted with 50 ml of deionised water, and 90 ml of 45% NaOH were added to each tube. The solution of the ammonia was distilled over on an automatic distillation system

(UDK 132 VELP Scientifica – Italy) with steam and collected in a flask containing 50 ml of 4% (w/v) H_3BO_3. The total nitrogen was determined by titration with 0.1M HCl, after adding a mixed indicator (methyl red 0,1% (w/v) and bromocresol green 0,2% (w/v) in ethanol). According to Yeoh & Wee, (1994), 4.45 was used as total nitrogen to total protein conversion factor and the results were expressed as g of protein per 100 g of starting material. Samples were analysed in biological triplicates. Protein quantification of U/T_b and GndHCl_b extracts was carried out using the "high-sensitive proteins" kit (BEN S.r.l., Milan, Italy REF. PTP375), which is based on the reaction of pyrogallol with the protein's basic residues. The entire analysis was accomplished on the Icubio Imagic-M9 (R-Biopharm, Milan) following the manufacturer instructions.

2.4 SDS-PAGE analysis

Three milligrams of MPF powder were suspended in 1 ml of Laemmli buffer 1X [Tris-HCl 50mM pH 6.8, containing 2% (w/v) SDS, 10% (w/v) glycerol, 5% (v/v) 2-sulfanylethan-1-ol and traces of bromophenol blue]. Alternatively, either U/T_b, GdnHCl_b extracts or digested protein solutions were mixed 1:1 (v/v) with Laemmli buffer 1X. Samples were boiled for 10 min and 15 µg of protein per well loaded into a 12% polyacrylamide gel. Broad-Range SDS-PAGE Standards (Bio-Rad Laboratories, Inc., Hercules, CA) were used as molecular markers. The separation was carried out at room temperature and constant voltage (120 V). Gels were fixed in 40% (v/v) methanol containing 10% (v/v) acetic acid for 1 h and rinsed three times in Milli-Q water with gentle shaking for 5 min. The gels were stained for 16 h in Coomassie Brilliant Blue R-250 staining solution (Bio-Rad Laboratories, Inc., Hercules, CA) and de-stained in Milli-Q water. Protein stained gels were imaged and processed using GelAnalyzer version 19.1 [GelAnalyzer 19.1 (www.gelanalyzer.com) by Istvan Lazar Jr., PhD and Istvan Lazar Sr., PhD, CSc]. The apparent

molecular mobility of the unknown proteins was estimated with a cubic spline curve using the SRS1 Cubic Spline for Excel software (https://www.srs1software.com/SRS1CubicSplineForExcel.aspx).

2.5 2D- Electrophoresis of MPF

MPF was dissolved in IPG strip rehydration buffer containing 8 M urea, 2% (w/v) CHAPS, 2% (v/v) Pharmalytes pH 4.0–7.0, 20 mM DTT and traces of bromophenol blue to a concentration of 3 mg/ml. Immobiline Dry Strips (pH 4–7 linear gradient,11 cm, Bio-Rad Laboratories, Inc., Hercules, CA) were rehydrated overnight in the Immobiline Dry-Strip Reswelling Tray (Amersham Pharmacia) and isoelectric focusing was performed using a Multiphor II system (Pharmacia Biotech, Uppsala, Sweden). Proteins were focused up to 15,000 Vh at a maximum voltage of 6000 V at 20° C. IPG strips were then placed in equilibration buffer [6 M urea, 30% (v/v) glycerol, 2% (w/v) SDS, 2% (w/v) DTT] for 15 min followed by incubation with 2.5% (w/v) IAA for 15 min. IPG strips were placed on a 10% polyacrylamide gel for the second dimension using a Protean II system (BioRad Laboratories, California). Separation was carried out at 220 V constant voltage, after which gels were fixed overnight (16 h) with TCA (24%) and stained with Coomassie Brilliant Blue R-250 staining solution (Bio-Rad Laboratories, Inc., Hercules, CA). The gel was imaged and analysed as previously described (section 2.4).

2.6 In-gel trypsin digestion

Selected protein spots were manually excised from stained 2D gel and de-stained by repeated washes with 50% (v/v) ACN in 25mM NH_4HCO_3. Gel spots were placed in 10mM DTT in 25 mM NH_4HCO_3 for 1 h at 56° C and then alkylated with 55 mM IAA in 25 mM NH_4HCO_3 for 45 min

at room temperature in the dark. Gel spots were then washed with 25 mM NH_4HCO_3, dehydrated with 0.1 ml of ACN and then dried in a vacuum centrifugal concentrator (VR-maxi St.a.-1, Heto, Denmark). Gel pieces were rehydrated the with 10–20 µl of a trypsin solution (proteomics grade 12.5 ng/ml in 25 mM NH_4HCO_3) for 16 h at 37° C. The resulting peptides were extracted from the gel pieces with 40 µl of 50% (v/v) acetonitrile/ 5% (v/v) aqueous formic acid for 20 min. Gel spots were then centrifuged at 7900 g for 5 min (Microcentrifuge Multispin 12, Steroglass, Perugia, Italy) to collect the supernatant. The extraction/centrifugation was repeated three times and, for each sample, the extracts were combined and dried in the vacuum centrifuge before MS analysis.

2.7 In-solution digestion

Diluted MPF extracts (section 2.2) were digested using either trypsin (proteomics grade) or chymotrypsin (proteomics grade) using an enzyme: substrate ratio of 1:100 (w/w) and incubated at 37° C for 16 h. All the digestions were performed on biological duplicates. To interrupt the hydrolytic processes, 100% formic acid was added to the samples to a final concentration of 0.1% (v/v). The peptide mixtures were desalted on C-18 Sep-Pak cartridges (Waters, Milford, Massachusetts, USA) following the instruction of the manufacturer and then concentrated with gaseous nitrogen and resuspended in 0.1% (v/v) formic acid solution. The hydrolysed samples were finally analysed by high resolution MS.

2.8 RP-HPLC analysis of MPF.

HPLC separation of *Moringa* proteins was performed using a C4 Grace-Vydac column (250x4.6 mm 5 µm, 214TP52) (Hesperia, CA, USA) attached to an HPLC chromatograph (HP 1100 Agilent - Palo Alto, CA, USA) modular system equipped with a diode array detector. Solvent A was 0.1%

of TFA (v/v) in water; solvent B was 0.1% (v/v) of TFA in ACN. MPF (1 mg) was dissolved in 1 ml of solvent A and 200 µl were the injected on column. After 5 min of isocratic elution using 5% (v/v) of solvent B a 5-75% (v/v) gradient ramp was applied over 60 min at a flow rate of 0.200 ml/min. The absorbance of the eluate was monitored at λ 220 and 280 nm. Peptide digests (section 2.7) were analysed using the same chromatography set up but using a C18 Grace-Vydac column (250x4.6mm 5µm, 218TP52) (Hesperia, CA, USA) and an injection volume of 50 µl.

2.9 HR LC-MS/MS analyses

MS analysis was performed by using a Q-Exactive Orbitrap mass spectrometer (Thermo Scientific, San Jose, CA, USA), online coupled with an Ultimate 3000 ultra-high-performance liquid chromatography instrument (Thermo Scientific). Samples were loaded through a 5 mm long, 300 µm id pre-column (LC Packings, USA) and separated by an EASY-Spray™ PepMap C18 column (2 µm, 15 cm×75 µm) 3 µm particles, 100 Å pore size (Thermo Scientific). Eluent A was 0.1% formic acid (v/v) in Milli-Q water; eluent B was 0.1% formic acid (v/v) in acetonitrile. The column was equilibrated at 5% B. The peptides generated after the hydrolysis of the spots were separated applying a 5–40% gradient of B over 60 min whereas the shotgun analyses were performed using a 5-70% gradient of B over 90min. The flow rate was 300 nl/min. The mass spectrometer operated in data-dependent mode and all MS1 spectra were acquired in the positive ionisation mode with an m/z scan range of 350 to 1600. Up to 10 most intense ions in MS1 were selected for fragmentation in MS/MS mode. A resolving power of 70,000 full width at half maximum (FWHM), an automatic gain control (AGC) target of 1×10^6 ions and a maximum ion injection time (IT) of 256 ms were set to generate precursor spectra. MS/MS fragmentation spectra were obtained at a resolving power of 17,500 FWHM. To prevent repeated fragmentation of the most abundant

ions, a dynamic exclusion of 10s was applied. Ions with one or over six charges were excluded. Spectra were processed by using the Xcalibur Software 3.1 version (Thermo Scientific) and then analysed using PEAKS Studio (version 6.0, Bioinformatics Solution Inc., Waterloo, Canada). The analysis of the mass spectra generated from in-gel digested proteins was performed using as reference proteome all the *Viridiplantae* plant sequence data (downloaded from https://www.uniprot.org/ in October 2020). Redundant sequences were removed using dbtoolkit-4.2.5 (Martens *et al.*, 2005). The database searching parameters were: i) mass tolerance value of 8 ppm for the precursor ion and 0.01 Da for MS/MS fragments; ii) enzyme: trypsin, maximum of 2 missed cleavages; iii) fixed modification: carbammidomethyl-cysteine; variable modifications: oxidation of the methionine and deamidation (NQ). The false discovery rate was set at 1%, and to improve the confidence protein identification, only proteins with -10lgP value > 50 were retained. Also, for the mass spectra generated from in-solution digested proteins, the background database used for the MS analysis was the whole *Viridiplantae* (downloaded from https://www.uniprot.org/ in June 2021). For in-solution *shotgun* analyses, *de-novo* searching peptides were selected based on the ALC (%) (Average Local Confidence), using a threshold of 70%. *De-novo* sequenced peptides were manually blasted in Uniprot in order to corroborate potential allergens. *De-novo* peptides were processed using PEAKS Studio, Jalview and Uniprot (https://www.uniprot.org/) to identify conserved regions of protein families and to perform the alignments. The first 20 top scored *de-novo* peptides were blasted searching for point modifications, as this aspect is not part of the study. For each duplicate analysis, the resulting identified protein databases were merged to obtain a unique dataset for trypsin and for chymotrypsin, and the redundancies were removed after a manual investigation.

2.10 Bioinformatic analysis

Common proteins were identified among the samples using the InteractiVenn software (http://www.interactivenn.net/) and were represented as a proportional Venn diagram with R through the *eulerr* package (eulerr: Area-Proportional Euler and Venn Diagrams with Ellipses. R package version 6.1.1, https://CRAN.R-project.org/package=eulerr). The identified proteins from the 2DE and shotgun analyses were examined again for their biological function using Gene Ontology and GO Annotations software (https://www.ebi.ac.uk/QuickGO/). The results of the GO analysis were reported as pie charts created with Meta-chart (https://www.meta-chart.com/). The percentage of identity and similarity among proteins was calculated using the Sequence Manipulation Suite (https://www.bioinformatics.org/sms2/ident_sim.html).

Only full lengths protein sequences, identified by MS and belonging to *M. oleifera* were evaluated *in-silico* per their potential allergenicity using Allermatch^{tm} (http://www.allermatch.org/, March 2021) and the AllerCatPro (https://allercatpro.bii.a-star.edu.sg/) (Maurer-Stroh *et al.*, 2019). They are based on the criteria of homology as defined by EFSA "with known allergens of at least 35% on a sliding window of 80 amino acids" (EFSA Panel on Genetically Modified Organisms (GMO), 2010) and on the search of 6-mer and 8-mer allergenic epitopes. *In-silico* analysis was performed also to identify potential epitopes using IEDB (https://www.iedb.org/) (Vita *et al.,* 2018). Potential allergens identified were processed through a multiple protein sequence alignment (MSA) with other known allergens and proteins of interest by using the Clustal Omega tool (https://www.ebi.ac.uk/Tools/msa/clustalo/) (Sievers *et al.,* 2011). The alignment was processed in Jalview to obtain phylogenetic trees using the BLOSUM62 algorithm. The resulting calculation was then visualised and edited in iTOL v6 (https://itol.embl.de/).

3 Results and discussion

3.1 Production of MPF and protein amount determination

The protein content of the Moringa leaves was determined by Kjeldahl analysis and corresponded 22.59±0.18 g/100 g dw. This was calculated using a conversion factor of 4.45, calculated by Yeoh & Wee, (1994) based on the amino acid composition of the proteins of Carica papaya leaves, a species taxonomically correlated with *M. oleifera* (Lin *et al.,* 2019). This value is consistent with recently published data for Moringa leaves from "Pakistan black" variety (Nouman *et al.,* 2016) although we applied a conversion coefficient of 4.45 rather than the canonical 6.25 conversion factor from nitrogen, which may have overestimated the amount of total protein being 31%.

The extraction of proteins from leaf tissue is challenging, as leaves contain high levels of substances (*e.g.,* polyphenols, anthocyanins, polysaccharides, cell wall particles) that bind to proteins and reduce their solubility. These also impair protein separation by electrophoresis, resulting in horizontal and vertical streaking, smearing, and reduced numbers of distinctly resolved protein bands. For these reasons, proteins were extracted from leaves using the phenol-based extraction aiming at the highest yield for exhaustive characterisation of the proteome (Abdullah *et al.,* 2017). The phenol solubilisation can effectively enhance the quantity of protein extracted from recalcitrant tissues, lowering the level of contamination, and minimising the proteolysis phenomena (Rose *et al.,* 2004). The phenol extraction was performed with a phenol buffered to pH 8.0 to ensure that nucleic acids were present into the lower buffer phase and not into the phenol upper phase (Faurobert *et al.,* 2007). The concentration of protein in the powdered MPF determined by Kjeldahl analysis corresponded to 57.77±0.45 g/100g dw.

3.2 M. oleifera protein profiling using RP-HPLC and electrophoretic methods

The RP-HPLC chromatogram of MPF showed several proteins in the hydrophobic region (>50% Solvent B) and was consistent with leaf protein profiles (Zolla *et al.*, 1999) (**Figure S1**). Subsequently, the protein profile of the MPF was determined by one-dimensional and two-dimensional SDS-PAGE. The 1D electrophoresis showed two intense bands with apparent molecular mobility of 51.66 KDa accounted for about 35% of the stained proteins (**Figure 2**). The profile obtained is similar to Wang *et al.*, (2016), who observed a predominant band around 55 kDa, which also accounted for 30% of *M. oleifera* leaves protein staining per track and seems likely to correspond to the Mr 51.66 band observed here. An additional 13.12 Mr band was observed and contributed to about 17% of the total protein stained per track (**Figure 2- Table S1**). Based on abundance and Mr, these bands are likely to correspond to the large and small subunit of RuBisCO, respectively (Krishnan & Natarajan, 2009). The denaturation due to the phenol extraction may have caused the "blurry" profile of the proteins with high background, impairing the visualisation of less abundant proteins.

Two-dimensional electrophoresis showed the complexity of the proteome in the presence of several spots in the region of 20-50 Mr with the 1D bands resolved into a complex of "spot trains", typical of proteins highly modified by post-translational modifications (PTMs) such as glycosylation (**Figure 2**). The Mr 13.12 band, observed in 1D was not resolved, suggesting it either has an extreme IP or was lost during the gel fixation process.

Among the 25 excised spots, only 3 inferred proteins belonged to *M. oleifera* annotated proteome in Uniprot, but over 70 were identified by homology with closely related species (**Table S2**). The gene ontology (GO) analysis showed these proteins belonged to fundamental metabolic pathways like protein metabolism and photosynthesis as expected for leaf tissues (**Figure 2**). Some spots in

a broad range of molecular weight (66 and 21 Mr) were found to correspond to RuBisCO. The enzyme is the primary player of the photosynthesis and is constantly turned over in the leaf's chloroplasts, even at a late stage of senescence (dead or dying leaves) (Hirel & Gallais, 2006). Since the leaves under analysis had been slowly sun dried, the fragments of RuBisCO may reflect proteolytic events occurring during drying/storage of the leaves. This aspect should be considered when evaluating the processing and the storage conditions of the ingredient.

3.3 Proteomic profiling of M. olifera leaves using shotgun proteomics

The 2D electrophoresis provides a map of the leaf proteome, but it is limited in terms of its ability to resolve hydrophobic proteins, small size, and low abundancy proteins. Visualisation of the proteins on the gel depends on the staining method applied and may not be equally effective for all proteins. Therefore, complementary shotgun proteomics was performed. Initially, the performance of the buffers used for the solubilisation of MPF was evaluated (*i.e.,* GdnHCl_b and U/T_b).

The direct protein quantification using the pyrogallol assay demonstrated the U/T_b was 2.8 fold more efficient than the GndHCl_b at solubilising the proteins (**Table S3**). These data were further confirmed by SDS-PAGE (**Figure S2**) showing a more complex profile for U/T_b. These results can be attributed to the phenol extraction, which may have increased the aggregation of the leaf proteins, which may be more soluble in a zwitterionic buffer than in a charged one.

The U/T_b and GndHCl_b solubilized proteins were then digested with trypsin prior to performing *shotgun* MS analysis. In order to optimise the protease activity, 1mM $CaCl_2$ was previously added in the protein dilution buffer (section 2.2) since Ca^{2+} ions possess a well known stabilising effect on some proteases (Kotormán *et al.,* 2003). The MS analyses highlighted that U/T_b was a more

effective extractant (**Figure S3**) allowing 14-fold more proteins to be identified compared to the GndHCl_b extract. The MS data of the trypsin hydrolysed samples dissolved in GndHCl_b allowed the most abundant proteins to be identified, the majority of which were also found in the U/T_b. Proteins identified were primarily involved in photosynthesis, like RuBisCO, photosystems I and II, and ATP-synthase, or protein families involved in the glycolysis such as glyceraldehyde-3-phosphate dehydrogenase, phosphoglycerate kinase, phosphoglycerate mutase (**Table S4 and S7**). Tubulin or actin family proteins, associated with the cellular structure of the leaf, were also identified. Among the leaf proteins, Morintides mO2 and mO1, "truncated chitinases" with a short C-terminal tail (also known as 8C-hevein-like peptides) were identified (Kini *et al.,* 2017). The molecular weight of the mature mO1 and mO2 proteins is around 4 kDa, with the isoforms differing by a point mutation at residue 15, that is Gln (mO1) → Gly (mO2). The unique peptides, carrying this mutation, were identified confirming the coexistence of both isoforms [mO1-AC̲AN**Q**LC̲C̲SQYGFC̲GSTSEYC̲SR and mO2-AC̲AN**G**LC̲C̲SQYGFC̲GSTSEYC̲SR (**Table S7**)].

Notably, the non-specific lipid transfer proteins (ns-LTP) could be observed only in guanidine extracted proteins (**Table S5** and **Table S8**). Since the U/T_b extract gave a better protein coverage, it was selected for further analysis by using chymotrypsin to further improve the coverage of *M. oleifera* leaf proteome.

As expected, the chromatographic profiles of chymotryptic digests were distinctly different from those obtained by trypsin (**Figure S4**). MS analysis showed chymotrypsin provided a smaller number of identifications (243 compared to 383) but were complementary with only 134 proteins identified in common. Chymotrypsin identified a greater proportion of proteins involved in metabolism and photosynthesis. Trypsin digestion, on the other hand, identified a greater

proportion of proteins involved in protein metabolism, such as the ribosomal subunits (i.e., 30S, 40S and 50S), peptidases (M16, M41) and molecular chaperonins and 90 kDa heat shock proteins (Hsp90). The common proteins are mainly involved in photosynthesis and metabolic process of nitrogen compounds such as glutamate synthase and alanine aminotransferase (**Figure 3–Table S10**). Comparison of *shotgun* and in-gel digestion showed that, notably, 28 different proteins mainly related to general metabolism and photosynthesis could be identified only with 2DE (**Figure 3 - Table S10**). Antioxidant proteins like catalase, ascorbate peroxidase and thioredoxin were commonly identified. Morintide mO2 was also identified in both trypsin and chymotrypsin. However, mO1 and nsLTP were not detected in chymotrypsin hydrolysed samples.

Only *de novo* peptides with an ALC (%) > 70% were analysed in order to increase the confidence of identifications with *de-novo* sequencing (Ma & Lajoie, 2009). The peptide sequences identified in the two technical MS analyses were merged, and duplicates removed. A total of 9913, 131 and 2985 peptides were identified by *de novo* sequencing for U/T_b trypsin, GndHCl_b trypsin and U/T_b chymotrypsin, respectively (**Tables S11-13**). The 20 ALC top scoring *de-novo* sequenced peptides of tryptic and chymotryptic protein hydrolysed were blasted in Uniprot searching for homologous sequences. These peptides were not inferred to any protein by the PEAKS software because they carried punctual modifications, deletion, and point modifications, which are not contained in the proteins of the *Viridiplantae* database, thus impeding an exact match. These *de-novo* peptides shared a high sequence identity (from 100% to 65%) with segments of metabolic proteins such as ATP synthase β, photosystem related proteins, RuBisCO large subunit and oxygen-evolving enhancer protein **(Tables S11-13)**.

3.5 In-silico allergenicity risk assessment of MPF using AllerCatPro and Allermatch™

In-silico analysis of the MPF proteome identified four different classes of putative allergens (**Table 1)** (*in-silico* allergenicity risk prediction flowchart and evaluations are available for consultation **Figure S5, Table S14-S17**). Some proteins, such as RuBisCO from *M. oleifera* (Uniprot ID: A0A482EC08), were identified with a score of 98.8% as a probable allergen ("Spi o Rubisco"- Uniprot ID: P00875) by the algorithm AllerCatPro (Maurer-Stroh *et al.*, 2019) but not by Allermatch™. The latter tool uses the highly curated allergen sequence database, Allergen-online (http://www.allergenonline.org/) available through the Food Allergy Research and Resource Program (FARRP), which excludes RuBisCO as an allergen. In contrast, AllerCatPro uses several databases, including Allergome database (https://www.allergome.org/), which has a less stringent approach to curation (Kadam *et al.,* 2016). This emphasises one of the shortcoming of in-silico prediction of allergens: the quality of the allergen designations assigned in allergen sequence database (Fernandez *et al.,* 2021). The database of allergens should rank the probable allergens by applying a score to data supporting their designation as allergens (Javed *et al.,* 2017). Notably, Spi o Rubisco is not a WHO/IUIS recognized allergen (http://www.allergen.org/index.php) and is present only on the Allergome database (Allergome ID: 3814). This entry is based on one study that described a non-specific and irreproducible IgE binding to RuBisCO from spinach using sera from people allergic to soybean (Hoff *et al.,* 2007) and an IgE binding study using serum from a single patient reporting a reaction after eating spinach (Foti *et al.,* 2012).

Another putative allergen identified uniquely with AllerCatPro is the mitochondrial beta subunit of ATP synthase from *Penicillum glabrum* (93.6% - Uniprot ID P85446). This is described as an

allergen in Uniprot, but with no supporting evidence. However, a serological study of *Cannabis sativa* allergic subjects showed that ATP synthase β could bind IgE (Nayak *et al.,* 2013).

Morintides (mO1 and mO2- Uniprot ID: A0A1S6EK91 and A0A1S6EK92, respectively) were also flagged as possible allergenic proteins both by the allergenicity prediction tools using the 3x6 overlaps principle before described in Maurer-Stroh *et al.,* (2019) for AllerCatPro and the simple 6 amino acid word-match for Allermatch[tm]. Morintides are too short to be assessed according to the 80 AA linear window and therefore we also evaluated the full-length homology using the Sequence Manipulation Suite alignment tool and cross-reactive epitope sequences through IEDB search. According to the Sequence Manipulation Suite, mO1, mO2 and the latex allergen Hevein-Hev b 6.02 (Uniprot ID: P02877- IEDB epitope ID: 13808) share a full length homology of 62.8% and a similarity of 81.40% (mO1) and 79.07% (mO2). TheAllerCatPro evaluation counted 13 and 18 3x6 overlaps between P02877 and mO1 and mO2, respectively (**Table S14**). According to Allermatch[TM], besides the overall sequence homology with Hev b 6.02, Morintides show an exact match for a 6-mer (CGRQAG – **Figure 4, Table S16** and **Table S17**), which is present in the first portion of the longer Hev b 6 epitope (EQCGRQAGGKLCPNNLCCSQ - IEDB epitope ID: 13807). Notably, Morintides portion 1-20 (QN**CGRQAG**NRA**C**A**N**Q(G)**LCCSQ**) shares 70% of identity with the entire epitope. According to Allermatch[tm] Morintides also share a high sequence homology with the chitin binding domain of the allergenic class I chitinase - of *Musa acuminata* (Mus a 2.0101 – Uniprot ID: Q8VXF1), *Hevea brasiliensis* (Hev b 11–Uniprot ID: Q949H3) and the class IV chitinase of *Cryptomeria japonica* (Uniprot ID: Q5NTA4) (**Table S16-17** and **Figure 4**). Another portion of Morintides (CGSTSEYCSR) was found to share a high sequence identity (80.0%) with a 10-mer epitope of Hev b 6.02: **CGST**D**EYCS**P (IEDB ID: 104739). However, the exact 6-mer homology search can give high false positive allergen prediction rate (EFSA 2010).

Therefore, a further Allermatch [tm] search of 8-mer sequences was performed. This identified only the mO2 as a potential allergen because of a match with an 8-mer peptide (GLCCSQYG) present in the mature domain of class IV chitinase of *Cryptomeria japonica*, Hev b 11 (class I chitinase) and Mus a 2.0101 (class I chitinase) (**Figure 4**, **Table S18** and **Table S19**). The first residue of this 8-mer peptide contains the punctual modification Gln (mO1) → Gly (mO2) that distinguishes mO1 from mO2. The dendrogram built on the multiple sequence alignment among Morintides and fruit chitinases, showed a close relationship between Morintides, Hev b 6.02 and class I chitinases (**Figure 5**). A high sequence homology and similarity, and conserved regions between Morintides (and particularly mO2) and the main allergens involved in the fruit-latex allergy (*i.e.,* Prs a 1, Cas a 5 and Mus a 2.0101) can also be observed (**Figure 4**).

3.6 Direct identification of nsLTP

nsLTPs are a panallergen family of cross-reactive allergens responsible for cause severe reactions (Dìaz-Perales *et al.,* 2000; McKenna *et al.,* 2016). Therefore, a direct analysis was done for this important allergen. The tryptic digest of the guanidine buffer (GndHCl_b, **Table S5**) allowed the identification by homology of the *ns*-LTP proteins (A0A087H3P3 and A0A2P6R8F4) belonging to *Arabis alpina* and *Rosa chinensis*, respectively. GndHCl_b may be more efficient in solubilising these small cationic proteins or the fact that this buffer solubilised fewer proteins could have enhanced the ionisation of those low in abundance. Interestingly, *Rosa chinensis* belongs to Rosaceae family that comprises well-known allergens such as Pru p 3, a nsLTP from peach (*Prunus persica*). The phylogenetic tree built on the MSA among A0A087H3P3, A0A2P6R8F4 and the best characterised allergenic ns-LTPs (**Figure 4)** highlights an interesting sequence homology between A0A2P6R8F4 and Rub i 3 (Uniprot ID: Q0Z8V0), a food allergen from red

raspberry (*Rubus ideaus* - *Rosaceae* family). Basing on **Figure 5,** we analysed the sequence alignment of A0A2P6R8F4 and the most closely branched nsLTP allergens and discovered that three main areas (**Figure 4**) in the amino acid sequence of *Rosa chinensis* nsLTP (GSVPPACCNGIRSLN; ASGLPGKCGVSVPYK and PYKISTSTNCNNVK) share a high sequence homology degree (87%, 80% and 85.7%) with two pentadecameric and one tetradecameric epitopes of Mal d 3: AAGLPGKCGVNVPYK (IEDB epitope ID: 167) and GAVPPACCNGIRTIN (IEDB epitope ID: 18816) and PYKISTSTNCATVK (IEDB epitope ID: 50121), respectively. The full-length sequences of mature A0A2P6R8F4 and Mal d 3 allergen share a sequence identity and similarity of 75.82% and 82.42%, respectively. It is also noteworthy the presence of interesting non-epitopic homologies between *Rosa chinensis* nsLTP and Rub i 3, which share a sequence identity and similarity of 82.42% and 85.71%, respectively. Basing on above mentioned evidences, a series of *in-vitro* experiments to evaluate a possible cross-reactivity in nsLTP allergic patients will also be required.

3.7 Confirmation of potential allergens by de-novo sequencing

To further corroborate the presence of these potential allergens we performed an in-depth research for *de-novo* peptides inferable either to Morintides or nsLTP. We found 3 peptides in the tryptic and 1 peptide in chymotryptic (U/T_b) which share a high homology with some segments of Morintides (**Figure S6**). In tryptic digest of the GndHCl_b solubilised proteins, we found 9 *de-novo* peptides, which structure share similar sequence with some conserved regions of nsLTPs. Since nsLTP of *M. oleifera* leaf has not been sequenced, these peptides were aligned on the 2 nsLTPs sequences A0A087H3P3 and A0A2P6R8F4. Five of these peptides were more structurally close to A0A087H3P3 (**Figure S7**) and four to A0A2P6R8F4 (**Figure S8**). Analysis of the U/T_b (trypsin) sample allowed four peptides to be identified, which were more similar to A0A087H3P3

(**Figure S9**) and a further four, which showed similarity to segments of A0A2P6R8F4 (**Figure S10**). Even though these peptides presented single amino acid variations, they showed a high score of sequencing, especially in the conserved regions (**Figure 4**). These data suggest the existence of other isoforms of Morintides and can give an information on the sequence of the authentic *Moringa* leaf nsLTPs.

4 Conclusions

Alternative sources of edible proteins are required to face the ongoing growth of world population. New dietary proteins may pose a potential risk of allergic reaction in consumers and thus require a risk assessment before their commercialisation. We evaluated the potential allergenicity of *M. oleifera* leaf powder as a novel source of proteins with high biological value.

The most abundant proteins, the RuBisCO and the subunit beta of ATP synthase, whose allergenicity has not been unequivocally demonstrated in literature, were identified as potential allergens only by AllerCatPro, via Allergome. Spurious IgE binding can be observed towards plant extracts, which can result from non-specific binding or cross-reactive carbohydrate determinants (CCDs). In allergy diagnosis, CCDs play an important role in enhancing the occurrence of false-positive results during IgE based in-vitro assays, although this does not have any clinical implications. Consequently, such immunoblotting data needs confirmation through blot inhibition experiments using purified protein or inhibition ELISAs.

Morintides (mO1 and mO2), sharing a high homology of sequence with Hev b 6.02, a contact allergen from *Hevea brasiliensis,* indicate a potential allergenic power of this protein family. By homology mass spectrometry, we identified also the *ns*LTP, a known food panallergen, whose sequence is not available in *Moringa* protein databases.

This lack of information, especially for the novel food, represents the biggest bottleneck in the *in-silico* allergenicity assessment. In this context, extensive genomic, transcriptomic and *de-novo* peptide sequencing is crucial for confirming and supporting allergenicity risk assessment of novel food. However, the manual inferring of *de-novo* sequenced peptides is rather difficult and time-consuming: proteomics software should improve in this respect.

Our *in-silico* allergenicity prediction of *M. oleifera* leaf proteins indicates the importance in designing further *in-vitro* and *in-vivo* assays to further assess the allergenic potential of this novel ingredient before placing it on the wider European market.

5 Acknowledgements

We want to thank Sud Rienergy S.r.l. - Favella Group, which provide us with *Moringa* leaves.

The authors also thanks Dr. Tiziana Granato (Project Manager Enzymatic & Automation at R-Biopharm Italia S.R.L.) for kindly supporting with the spectrophotometric testing.

This paper is dedicated to Dr. M.A. Nicolai who tragically passed away on July 27th, 2020. We want to thank her for her mastery as professor and scientist. We will never forget her.

6 Funding

This research did not receive any specific grant from funding agencies in the public, commercial,

or

not-for-profit sectors.

7 References

Abd El-Hack, M.E., Alagawany, M., Elrys, A.S., Desoky, E.M., Tolba, H.M.N., Elnahal, A.S.M., Elnesr, S.S., & Swelum, A.A. (2018). Effect of Forage Moringa oleifera L. (moringa) on Animal Health and Nutrition and Its Beneficial Applications in Soil, Plants and Water Purification. *Agriculture*, 8, 145. https://doi.org/10.3390/agriculture8090145.

Abdullah, F. I., Chua, L. S., & Rahmat, Z. (2017). Comparison of protein extraction methods for the leaves of ficus deltoidea. *Journal of Fundamental and Applied Sciences,* 9(2), 908-924. https://doi.org/10.4314/jfas.v9i2.19.

Alhakmani, F., Kumar, S., & Alam Khan, S. (2013). Estimation of total phenolic content, in–vitro antioxidant and anti–inflammatory activity of flowers of Moringa oleifera. *Asian Pacific Journal of Tropical Biomedicine,* 8, 623-627. https://doi.org/10.1016/S2221-1691(13)60126-4.

Asare, G.A., Gyan, B., Bugyei, K., Adjei, S., Mahama, R., Addo, P., Otu-Nyarko, L., Kwame Wierdu, E., & Nyarko, A. (2012). Toxicity potentials of thenutraceutical Moringa oleifera at supra-supplementationlevels. *Journal of Ethnopharmacology,* 139, 265–272. https://doi.org/10.1016/j.jep.2011.11.009.

Berglund, L.J. (2018). Anaphylaxis to Moringa oleifera: First description. *Australian Medical Journal*, 11 (3), 176–177. https://doi.org/10.21767/AMJ.2018.3344.

Díaz-Perales, A., Lombardero, M., Sánchez-Monge, R., García-Selles, F. J., Pernas, M., Fernández-Rivas, M., Barber, D., & Salcedo, G. (2000). Lipid-transfer proteins as potential plant panallergens: cross-reactivity among proteins of Artemisia pollen, Castanea nut and Rosaceae fruits, with different IgE-binding capacities. *Clinical and experimental allergy : journal of the*

British Society for Allergy and Clinical Immunology, 30(10), 1403–1410. https://doi.org/10.1046/j.1365-2222.2000.00909.x.

EFSA Panel on Genetically Modified Organisms (GMO). (2010). Draft Scientific Opinion on the 1762 assessment of allergenicity of GM plants and microorganisms and derived food and feed. EFSA 1763 Journal 8 (7): 1700, 168 pp.

EFSA (European Food Safety Authority), 2019. Technical Report on the notification of leaf powder of Moringa stenopetala as a traditional food from a third country pursuant to Article 14 of Regulation (EU) 2015/2283. EFSA supporting publication 2019:EN-1672. 14 pp. doi:10.2903/sp.efsa.2019.EN-1672.

Faurobert, M., Pelpoir, E., & Chaïb J. (2007). Phenol Extraction of Proteins for Proteomic Studies of Recalcitrant Plant Tissues. *Plant Proteomics*. *Methods in Molecular Biology*, 355, 9-14. https://doi.org/10.1385/1-59745-227-0:9.

Fernandez, A., Mills, E.N.C., Koning, F., & Moreno, F.J. (2021). Allergenicity Assessment of Novel Food Proteins: What Should Be Improved? *Trends in Biotechnology,* 39 (1), 4-8. https://doi.org/10.1016/j.tibtech.2020.05.011.

Foti, C., Damiani, E., Zambonin, C. G., Cassano, N., Nettis, E., Ferrannini, A., Calvano, C. D., Aresta, A., Romita, P., Aloia, A. M., & Vena, G. A. (2012). Urticaria and angioedema to rubisco allergen in spinach and tomato. *Annals of allergy, asthma & immunology : official publication of the American College of Allergy, Asthma, & Immunology*, 108 (1), 60–61. https://doi.org/10.1016/j.anai.2011.09.011.

Hirel, B., & Gallais, A. (2006). RuBisCO synthesis, turnover and degradation: some new thoughts on an old problem. *New Phytologist*, 169 (3), 445–448. https://doi.org/10.1111/j.1469-8137.2006.01641.x.

Hoff, M., Son, D.Y., Gubesch, M., Ahn, K., Lee, S.I., Vieths, S., Goodman, R.E., Ballmer-Weber, B.K., & Bannon, G. A. (2007). Serum testing of genetically modified soybeans with special emphasis on potential allergenicity of the heterologous protein CP4 EPSPS. *Molecular Nutrition & Food Research*, 51 (8), 946–955. https://doi.org/10.1002/mnfr.200600285.

Javed, B., Padfield, P., Sperrin, M., Simpson, A., & Mills, E. N. C. (2017). A protocol for a systematic review to identify allergenic tree nuts and the molecules responsible for their allergenic properties. *Food and Chemical Toxicology*, 106, 411–416. https://doi.org/10.1016/j.fct.2017.06.004.

Kadam, K., Sawant, S., Jayaraman, V.K., & Kulkarni-Kale, U. (2016). Databases and Algorithms in Allergen Informatics. *Bioinformatics-Updated Features and Applications*, 53. https://doi.org/10.5772/63083.

Kini, S.G., Wong, K.H., Tan, W.L., Xiao, T., & Tam, J.P. (2017). Morintides: Cargo-free chitin-binding peptides from Moringa oleifera. *BMC Plant Biology*, 17 (1), 1-13. https://doi.org/10.1186/s12870-017-1014-6.

Kotormán, M., Laczkó, I., Szabó, A., & Simon, L. (2003). Effects of Ca2+ on catalytic activity and conformation of trypsin and α-chymotrypsin in aqueous ethanol. *Biochemical and Biophysical Research Communications*, 304(1), 18–21. https://doi.org/10.1016/s0006-291x(03)00534-5.

Krishnan, H.B., & Natarajan, S.S. (2009). A rapid method for depletion of Rubisco from soybean (Glycine max) leaf for proteomic analysis of lower abundance proteins. *Phytochemistry,* 70, 1958–1964. https://doi.org/10.1016/j.phytochem.2009.08.020.

Lin, W., Dai, S., Chen, Y., Zhou, Y. & Liu, X. (2019). The complete chloroplast genome sequence of Moringa oleifera Lam. (Moringaceae). Mitochondrial DNA B, 4, 4094-4095. https://doi.org/10.1080/23802359.2019.1627922.

Ma, B., & Lajoie, G. (2009). De Novo Interpretation of Tandem Mass Spectra. *Current Protocols in Bioinformatics*, 25 (1), 13-10. doi:10.1002/0471250953.bi1310s25.

Martens, L., Vandekerckhove, J., & Gevaert, K. (2005). DBToolkit: processing protein databases for peptide-centric proteomics. *Bioinformatics*, 21 (17) 3584-3585 https://doi.org/10.1093/bioinformatics/bti588.

Maurer-Stroh, S., Krutz, N. L., Kern, P. S., Gunalan, V., Nguyen, M. N., Limviphuvadh, V., Eisenhaber, F., & Gerberick, G. F. (2019). AllerCatPro-prediction of protein allergenicity potential from the protein sequence. *Bioinformatics*, 35(17), 3020–3027. https://doi.org/10.1093/bioinformatics/btz029.

McKenna, O. E., Asam, C., Araujo, G. R., Roulias, A., Goulart, L. R., & Ferreira, F. (2016). How relevant is panallergen sensitization in the development of allergies? *Pediatric Allergy and Immunology,* 27(6), 560–568. https://doi.org/10.1111/pai.12589.

Mishra, S.P., Singh, P., & Singh, S. (2012). Processing of Moringa oleifera Leaves for Human Consumption. *Bulletin of Environment, Pharmacology and Life Sciences*, 2 (1), 28-31.

Montesano, D., Cossignani, L., & Blasi, F. (2018). Sustainable crops for food security: Moringa (Moringa oleifera Lam). In book: Reference Module in Food Science. *Encyclopedia of Food Security and Sustainability*, 409-415. https://doi.org/10.1016/B978-0-08-100596-5.22574-2

Nayak, A. P., Green, B. J., Sussman, G., Berlin, N., Lata, H., Chandra, S., ElSohly, M.A., Hettick, J.M., & Beezhold, D. H. (2013). Characterization of Cannabis sativa allergens. *Annals of Allergy, Asthma & Immunology,* 111, 32–37. https://doi.org/10.3923/ijds.2016.69.74

Nouman, W., Anwar, F., Gull, T., Newton, A., Rosa, E., & Domínguez-Perles, R. (2016). Profiling of polyphenolics, nutrients and antioxidant potential of germplasm's leaves from seven cultivars of Moringa oleifera Lam. *Industrial Crops and Products*, 83, 166–176. https://doi.org/ 10.1016/j.indcrop.2015.12.032.

Olaofe, O., Adeyeye, E., & Ojugbo, S. (2013). Comparative study of proximate, amino acids and fatty acids of Moringa oleifera tree. *Elixir Applied Chemistry*, 54, 12543-12554.

Rocchetti, G., Pamplona Pagnossa, J., Blasi, F., Cossignani, L., Hilsdorf Piccoli, R., Zengin, G., Montesano, D., Cocconcelli, P.S., & Lucini, L. (2020). Phenolic profiling and in vitro bioactivity of Moringa oleifera leaves as affected by different extraction solvents. *Food Research International,* 127, 108712. https://doi.org/10.1016/j.foodres.2019.108712.

Rocco, M., D'Ambrosio, C., Arena, S., Faurobert, M., Scaloni, A., & Marra, M. (2006). Proteomic analysis of tomato fruits from two ecotypes during ripening. *Proteomics,* 6, 3781- 3791. https://doi.org/ 10.1002/pmic.200600128.

Rose, J. K. C., Bashir, S., Giovannoni, J. J., Jahn, M. M., & Saravanan, R. S. (2004). Tackling the plant proteome: practical approaches, hurdles and experimental tools. *The Plant Journal,* 39(5), 715–733. https://doi.org/10.1111/j.1365-313x.2004.02182.x.

Shih, M.C., Chang, C.M., Kang, S.M., & Tsai, M.L. (2011). Effect of Different Parts (Leaf, Stem and Stalk) and Seasons (Summer and Winter) on the Chemical Compositions and Antioxidant Activity of Moringa oleifera. *International Journal of Molecular Sciences*, 12, 6077-6088. https://doi.org/10.3390/ijms12096077.

Sievers, F., Wilm, A., Dineen, D., Gibson, T.J., Karplus, K., Li, W., Lopez, R., McWilliam, H., Remmert, M., Söding, J., Thompson, J.D., & Higgins, D.G. (2011). Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Molecular systems biology*, 7(1), 539. https://doi.org/10.1038/msb.2011.75.

Su, B., & Chen, X. (2020). Current Status and Potential of Moringa oleifera Leaf as an Alternative Protein Source for Animal Feeds. *Frontiers in Veterinary Science*, 7 (53). https://doi.org/10.3389/fvets.2020.00053.

Sujatha, B. K., & Poonam, P. (2017). Moringa oleifera- Nature'' Gold. *Imperial Journal of Interdisciplinary Research*. 3 (5), 1175-1179.

Vita, R., Mahajan, S., Overton, J.A., Dhanda, S.K., Martini, S., Cantrell, J.R., Wheeler, D.K., Sette, A., & Peters, B. (2018). The Immune Epitope Database (IEDB): 2018 update. *Nucleic Acids Research,* Oct 24. https://doi.org/10.1093/nar/gky1006.

Wang, L., Zou, Q., Wang, J., Zhang, J., Liu, Z., & Chen, X. (2016). Proteomic Profiles Reveal the Function of Different Vegetative Tissues of Moringa oleifera. *The Protein Journal*, 35, 440–447. https://doi.org/10.1007/s10930-016-9690-x.

Yeoh, H.H., & Wee, Y.-C. (1994). Leaf protein contents and nitrogen-to-protein conversion factors for 90 plant species. *Food Chemistry*, 49(3), 245–250. https://doi.org/10.1016/0308-8146(94)90167-8.

Zhou, H., Wang, X., Ye, F., Zeng, X.L., & Wang, Y. (2012). Relationship between child feeding practices and malnutrition in 7 remote and poor counties, P R China. *Asia Pacific Journal of Clinical Nutrition*, 21 (2), 234-240.

Zolla, L., Timperio, A.M., Testi, M.G., Bianchetti, M., Bassi, R., Manera, F., & Corradini, D. (1999). Isolation and characterization of chloroplast Photosystem II antenna of spinach by reversed-phase liquid chromatography. *Photosynthesis Research* 61, 281–290. https://doi.org/10.1023/A:1006397419473.

Web references

http://www.allergen.org/index.php

http://www.allergenonline.org/

http://www.allermatch.org/

http://www.eur-lex.europa.eu/

http://www.gelanalyzer.com/

https://www.allercatpro.bii.a-star.edu.sg/

https://www.allergome.org/

https://www.bioinformatics.org/sms2/ident_sim.html

https://biorender.com/

https://CRAN.R-project.org/package=eulerr

https://www.ebi.ac.uk/QuickGO/

https://www.ebi.ac.uk/Tools/msa/clustalo/

https://www.iedb.org/

http://www.interactivenn.net/

https://itol.embl.de/

https://www.meta-chart.com/

https://www.srs1software.com/SRS1CubicSplineForExcel.aspx

https://www.uniprot.org/

Figures

Figure 1 – Protein extraction methodologies from *M. oleifera* leaf. The figure was created with BioRender.com (https://biorender.com/).

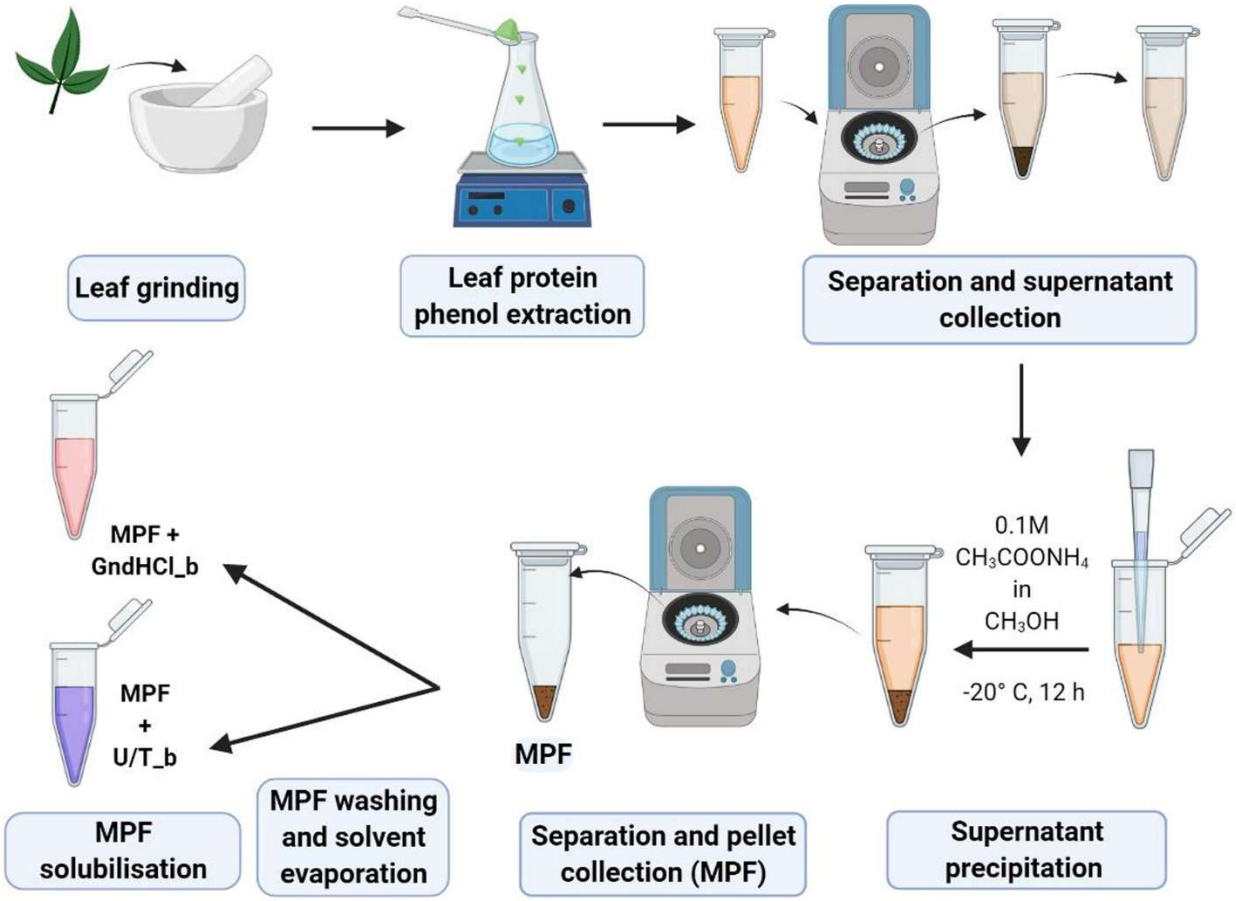


Figure 2 – *In-gel* based protein profiling of MPF by SDS PAGE (Panel A), 2D Electrophoresis (Panel B), and GO analysis of selected 2DE spots (Panel C).

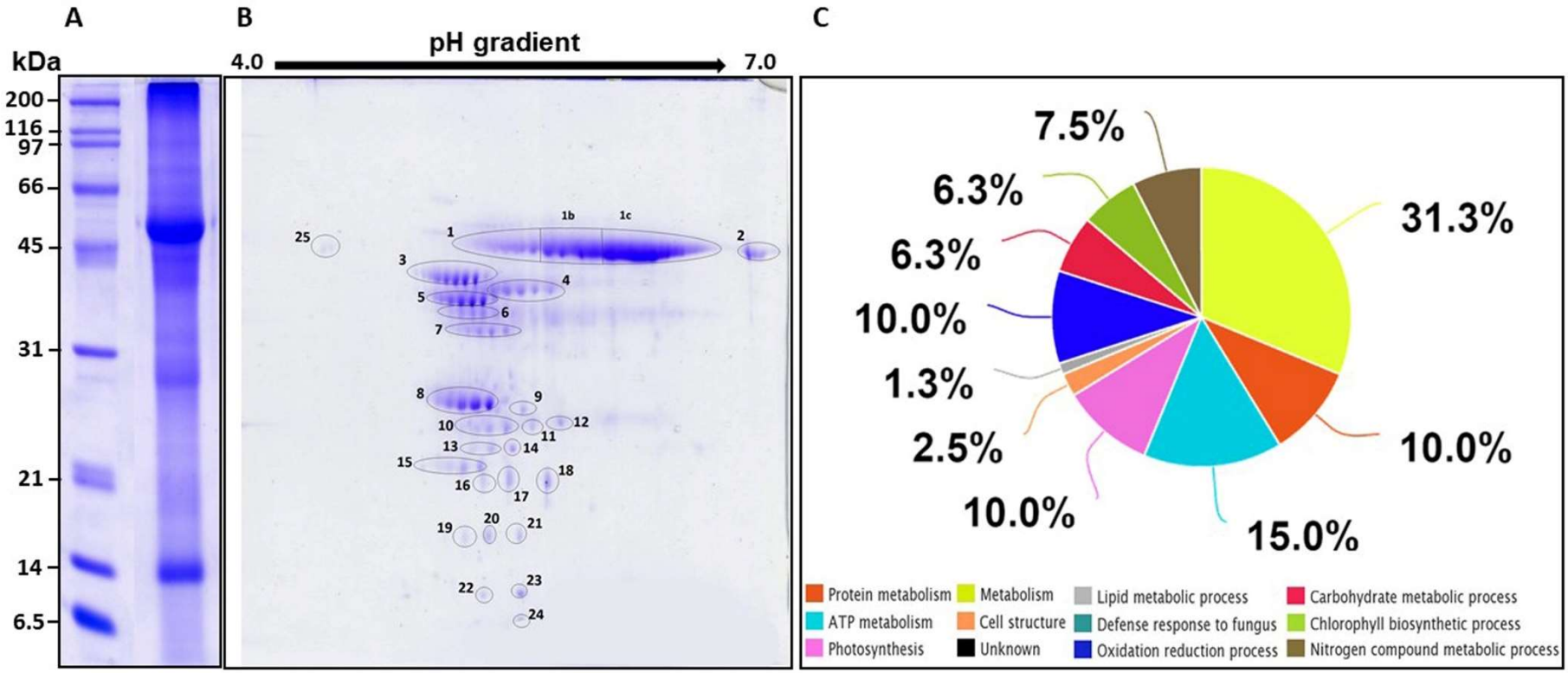


Figure 3 – Venn diagram of 2DE and shotgun identified proteins (trypsin and chymotrypsin). Graphical pie charts and color coding indicate the metabolic pathways of the identified proteins. The Venn diagram was created using the R eulerr package. The pie charts were created using meta-chart (https://www.meta-chart.com/).

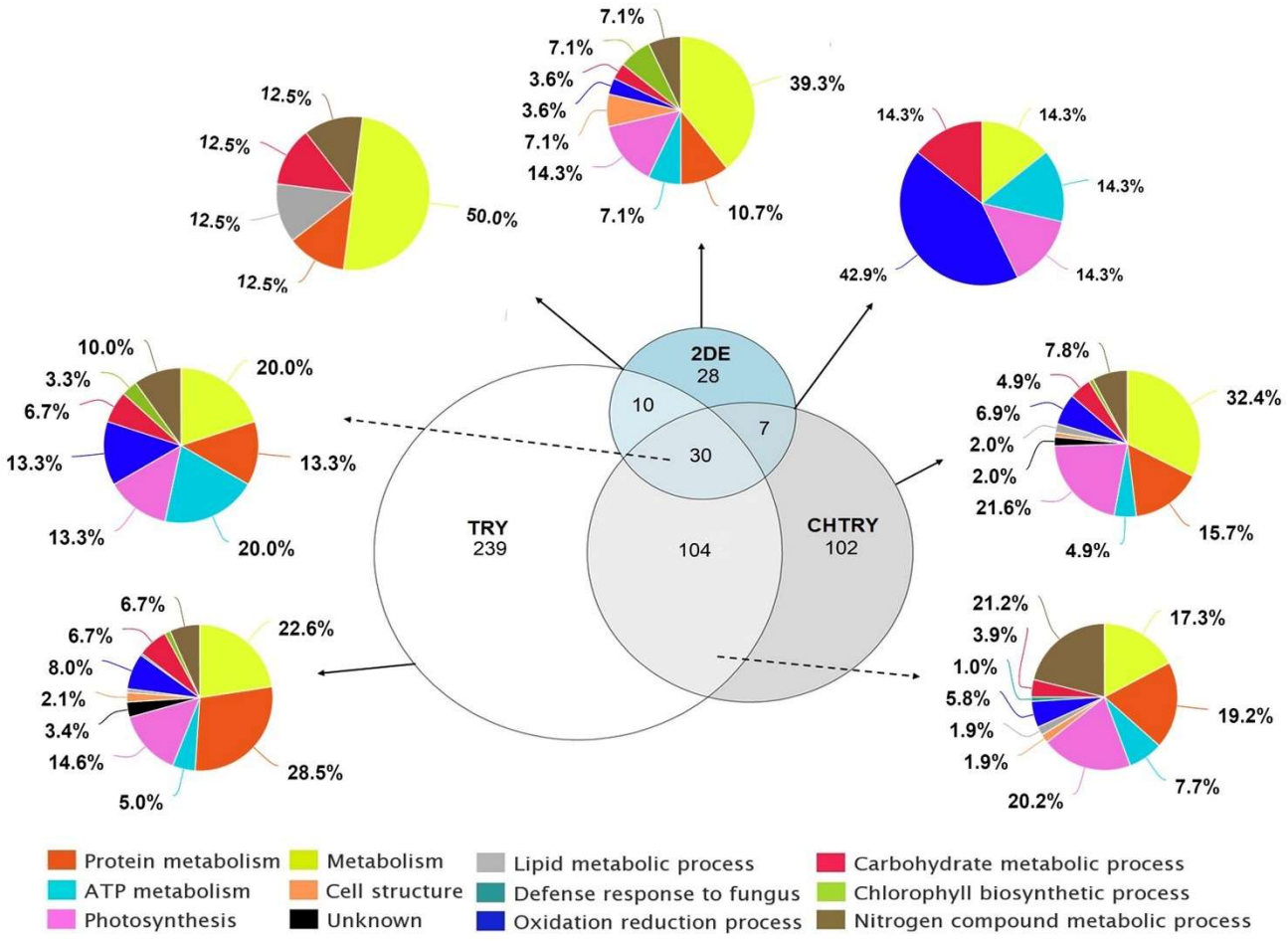


Figure 4 - Amino acid sequence alignment of: A) hevein domains of latex, fruit allergenic class I chitinases and Morintides; B) allergenic nsLTP and the nsLTP sequences of *Rosa chinensis* and *Arabis alpina,* identified by homogy with in *Moringa oleifera* leaves. The common sequences are showed in blue gradient range according to the homology degree. The effective epitopes are represented in continuous line boxes, whilst interesting homologies (≥ 6 residues) are represented as dashed boxes.

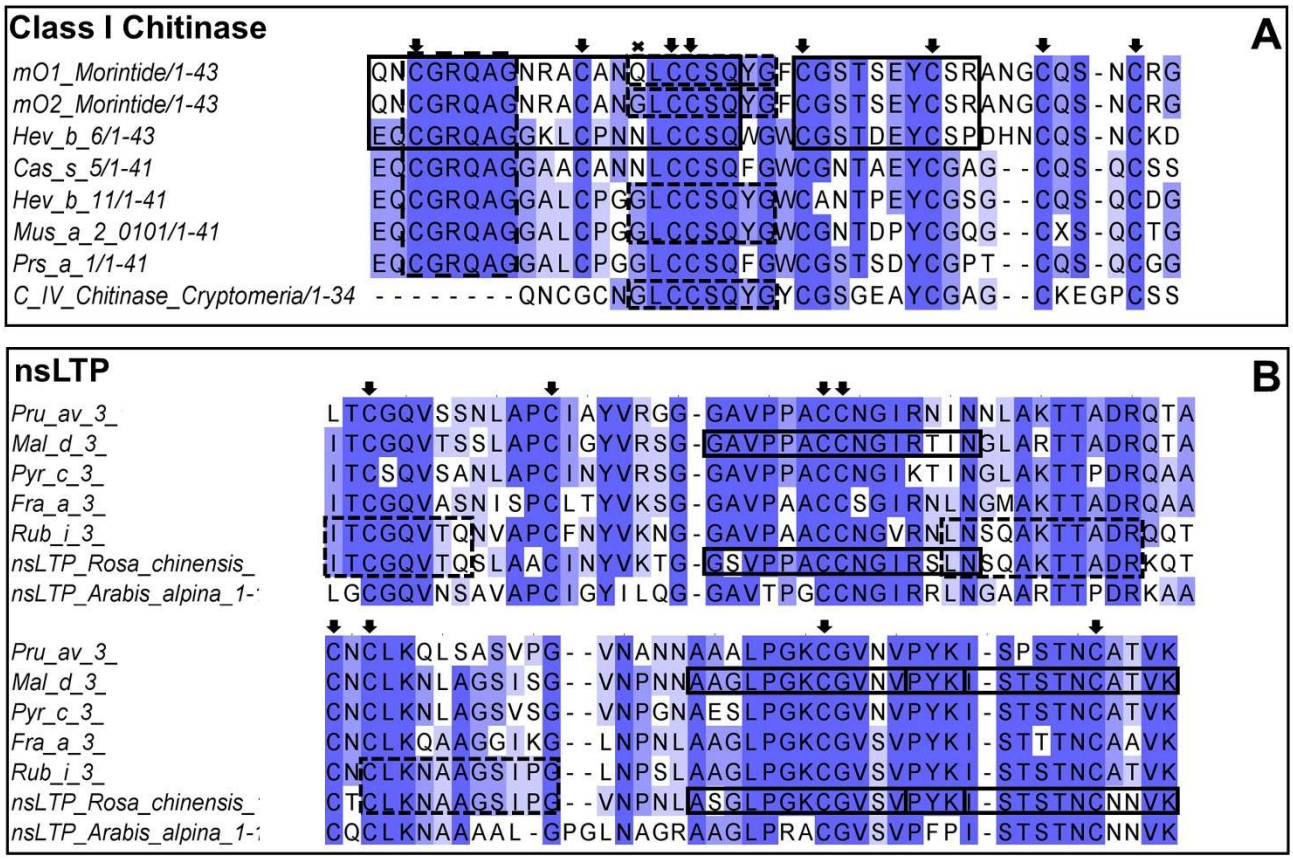


Figure 5 – Phylogenetic trees are representations of multi sequence alignment between proteins of interest and well-known allergens. Panel A) Chitinase and Morintides. Morintide mO2 shows the highest sequence homology (H) and similarity (S) with the main fruit-latex allergens (Prs a 1: H= 60.5%, S= 76.7%; Cas a 5: H= 60.5%, S= 74.4%; Mus a 2.0101: H= 55.8%, S= 69.8%); Panel B) ns-LTPs.

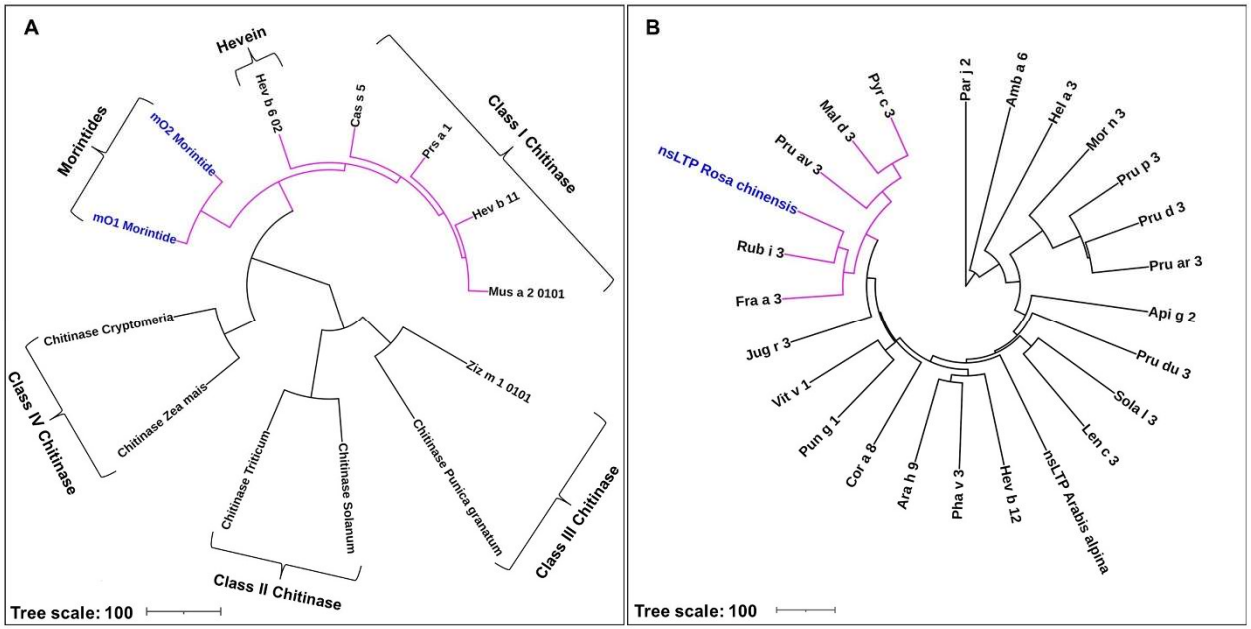


Table caption

Table 1 – Putative allergens recognized through Allermatch and AllerCatPro softwares during *in-silico* analyses.

Putative Allergen	In-Silico Tool		Comments
	AllerCatPro	Allermatch™	
RuBisCO	✓	×	Spi o Rubisco (Allergome)
ATP Synthase β	✓	×	P85446 (Uniprot – Unpublished data)
Morintides (mO1 and mO2)	✓	✓	Hev b 6.02 (WHO/IUIS)
nsLTPs	✓	✓	Rub i 3 (WHO/IUIS)

Supplementary tables captions

Table S1 – Densitometry data of SDS-PAGE of MPF (Figure 1) obtained using the Gel Analyzer software (v.19.1).

Table S2 – List of proteins identified by HR LC-MS/MS analysis of selected spots in 2DE gel.

Table S3 – Pyrogallol assay for the high-sensitive protein quantification for the different methods of solubilization of MPF.

Table S4- List of proteins identified by HR LC-MS/MS analysis of the tryptic hydrolysed MPF solubilised in Urea/Thiourea buffer.

Table S5 – List of proteins identified by HR LC-MS/MS analysis of the tryptic hydrolysed MPF solubilised in guanidine hydrochloride buffer.

Table S6 - List of proteins identified by HR LC-MS/MS analysis of chymotryptic hydrolysed MPF solubilised in Urea/Thiourea buffer.

Table S7 - List of peptides identified by HR LC-MS/MS analysis of the tryptic hydrolysed MPF solubilised in Urea/Thiourea buffer.

Table S8 - List of peptides identified by HR LC-MS/MS analysis of the tryptic hydrolysed MPF solubilised in guanidine hydrochloride buffer.

Table S9 - List of peptides identified by HR LC-MS/MS analysis of the chymotryptic hydrolysed MPF solubilised in Urea/Thiourea buffer.

Table S10 – List of protein/protein families rappresented in the Venn diagram in Figure 3 (obtained using http://www.interactivenn.net/)

Table S11 - List of *de novo* peptides identified by shotgun analysis of tryptic hydrolysis of MPF solubilised in Urea/Thiourea buffer. The 20 ALC (indicated in green) top-scoring *de novo* peptides were BLAST searched using Uniprot for identification with homologous sequences. Peptides coloured blue are the morintides, peptides in yellow are the nsLTP.

Table S12 - List of *de novo* peptides identified by shotgun analysis of tryptic hydrolysis of MPF solubilised in guanidine hydrochloride buffer. The 20 ALC (indicated in green) top-scoring *de novo* peptides were BLAST searched using Uniprot for identification with homologous sequences. Peptides coloured blue are the morintides, peptides in yellow are the nsLTP.

Table S13 - List of *de novo* peptides identified by shotgun analysis of chymotryptic hydrolysis of MPF dissolved in Urea/Thiourea buffer. The 20 ALC (indicated in green) top-scoring *de novo* peptides were BLAST searched using Uniprot for identification with homologous sequences. Peptides coloured blue are the morintides, peptides in yellow are the nsLTP.

Table S14 – Protein allergenicity potential prediction through the AllerCatPro software (v 1.7) of trypsin hydrolysate.

Table S15 - Protein allergenicity potential prediction through the AllerCatPro software (v 1.7) of chymotrypsin hydrolysed.

Table S16 – Protein allergenicity potential prediction through the Allermatch^{tm} software of mO1.

Table S17 - Protein allergenicity potential prediction through the Allermatch^{tm} software of mO2.

Table S18 – 8-mer search in mO2 through the Allermatch^{tm} software.

Table S19 - 8-mer search in mO2 through the FARRP software.

Supplementary figures captions

Figure S1 – RP-HPLC of MPF soluble in acidified water.

Figure S2 – SDS-PAGE of (1) urea/thiourea solubilised proteins and (2) guanidinium hydrochloride solubilised proteins.

Figure S3 – Protein/protein families identified after tryptic hydrolysis of MPF urea/thiourea (TRY U/T_b) and guanidinium hydrochloride (TRY GndHCl_b) solubilised proteins.

Figure S4 -. HPLC chromatrographic traces of (A) trypsin and (B) chymotrypsin digested proteins solubilised in urea/thiourea. SDS-PAGE 1) solubilised proteins; 2) trypsin or chymotrypsin digests.

Figure S5 - Decisional tree for the optimization of MPF's solubilization.

Figure S6- *De-novo* sequenced peptides aligned with Morintides. The confidence level of peptide fragment assignment is reporded in the histogram on the right of the figure.

Figure S7- *De-novo* sequenced peptides aligned with nsLTP (A0A087H3P3) identified in the GndHCl_b/trypsin. The confidence level of peptide fragment assignment is reporded in the histogram on the right of the figure.

Figure S8- *De-novo* sequenced peptides aligned with nsLTP (A0A2P6R8F4) identified in the GndHCl_b/trypsin. The confidence level of peptide fragment assignment is reporded in the histogram on the right of the figure.

Figure S9- *De-novo* sequenced peptides aligned with nsLTP (A0A087H3P3) identified in the U/T_b/trypsin. The confidence level of peptide fragment assignment is reporded in the histogram on the right of the figure.

Figure S10- *De-novo* sequenced peptides aligned with nsLTP (A0A2P6R8F4) identified in the U/T_b/trypsin. The confidence level of peptide fragment assignment is reporded in the histogram on the right of the figure.

Declarations of interest: none

Chapter 3

A follow-up allergenicity risk assessment using specific sera screening on Moringa oleifera leaf proteins after bioinformatic prediction.

In preparation

A follow-up allergenicity risk assessment using specific sera screening on Moringa oleifera leaf proteins after bioinformatic prediction.

Giovanni D'Auria[1§], Kai Wang[2§], Chiara Nitride[1], Maria Ruano-Zaragoza[3], Pasquale Ferranti[1], E. N. Clare Mills*[2].

Affiliations:

[1]Department of Agricultural Sciences, University of Naples Federico II, Portici, Italy.

[2]Division of Infection, Immunity and Respiratory Medicine, School of Biological Sciences, Manchester Academic Health Sciences Centre, Manchester Institute of Biotechnology, University of Manchester, Manchester, UK

[3] Hospital Clinic of Barcelona, Barcelona, 08036, Spain

*Corresponding author: clare.mills@manchester.ac.uk

§ These authors contributed equally

Keywords: immunoblotting; leaf allergens; Morintides; nsLTP; ELISA.

Abstract

Background: This study represents an *in-vitro* follow-up of a previous bioinformatic prediction of the allergenicity of *Moringa oleifera* leaf. The *in-silico* analysis labelled two proteins as potential allergens: i) morintides, because of a high sequence homology with the latex allergen Hev b 6.02; ii) and non-specific lipid transfer protein (nsLTP), inferred by homology to the *Rosaceae* family.

Methods: We performed immune-based assays on Moringa protein extracts, using animal IgG anti-Pru p 3 (the nsLTP of peach's skin); and using human sera IgE either from latex or nsLTP allergic patients. The protein glycan cleavage was additionally performed to remove IgE reactive carbohydrate cross determinants (CCDs).

Results: Animal IgG did not show the presence of Pru p 3 homologues. Conversely, human sera IgE analyses reproducibly highlighted an immunoreactive band in patients allergic to nsLTP. On the other hand, the immunoblotting performed with latex allergic sera showed no specific immunoreactivity to Hev b 6.02. The removal of CCDs reduced spurious IgE response.

Conclusions: Immune-based assays revealed the imperfection of the current approaches for allergenicity risk assessment. The nature of the IgE binding observed in this study represents the subject for future research.

Abbreviations

CCD: carbohydrate cross determinant

GndHCl_b: Guanidinium hydrochloride buffer

MFG-: Moringa food grade protein isolate extracted under non-reducing conditions

MFG+: Moringa food grade protein isolate extracted under reducing conditions

NR: non reduced

PBS: Phosphate-buffered saline

R: reduced

U/T_b: Urea/Thiourea buffer

1. Introduction

Moringa oleifera leaf has been explored as a novel sustainable nutrition source, benefiting from its high protein content (14-31.4% on a dry basis) (Nouman et al., 2016), high stress resistance to extreme conditions (Alhakmani et al., 2013), and low input costs (Sujatha & Patel, 2017). As part of the safety assessment, it is important to evaluate its allergenic potential. The preliminary evaluation performed in **Chapter 2** of this thesis using bioinformatics is part of the current "weight-of-evidence" approach applied to the allergenicity risk assessment of novel dietary protein sources (EFSA, 2010; K. Verhoeckx et al., 2016). The *in-silico* assessment identified two proteins, namely morintides and *Rosa chinensis* nsLTP with high sequence homology with Hev b 6.02 and *Rosaceae* nsLTP, respectively. Within *in-silico* analyses, bioinformatic tools compare the sequence of the protein with known allergens. This assessment assigns to each protein a score that relies on the evaluation of a threshold value (i.e., 35% homology over a sliding window of 80 amino acids, presence of allergens 6-mers or 8-mers) or on more advanced machine learning calculations (Fernandez et al., 2021).

However, these bioinformatic tools are based on allergen sequence databases (e.g., WHO/IUIS, Compare, Allergome, etc.) and the protein databases (i.e., UniprotKB), which present strengths (e.g., open source, peer reviewing processes, etc.), but also some weaknesses. Some of the existing allergen databases are not frequently updated, differ from each other in terms of nomenclature and inclusion criteria, and some of these offer little systematic information about the allergens (Radauer & Breiteneder, 2019). Besides, poor gene sequencing of novel food and the consequent lack of *de novo* protein sequences limits the *in-silico* allergenicity prediction. Consequently, even though bioinformatics generates important

information, the election methods in the preliminary evaluation of the allergenic potential of a novel food remain the *in-vitro* approach.

As an important step of the *in-vitro* assessment, specific IgE screening is recommended by EFSA guidelines aiming to test the possible IgE-cross reactivity of novel proteins with known allergens when the source of the gene/protein is considered allergenic or there is any indication of relationship or structure similarity with known allergens (EFSA, 2010). Considerations include the choices of different IgE binding tests (e.g. Radio or Enzyme Allergosorbent Assay, Enzyme-Linked Immunosorbent Assay and electrophoresis followed by immunoblotting with specific IgE-containing sera) and the criteria for serum selection. Since the sera from clinically well characterised allergic individuals by skin prick test (SPT) and/or oral food challenge (OFC) are limited in number and quantity, the antibodies raised in animals are always used for the pre-screening phase. It is noteworthy that the specificity and affinity of IgE response vary among allergic individuals and a positive IgE binding result does not necessarily indicate an allergic reaction will occur (K. Verhoeckx et al., 2016). Additionally, the cross-reactive carbohydrate determinants (CCDs) (mostly found in plants), binding to IgE, can obscure *in-vitro* detection of true allergens (Kaulfürst-Soboll et al., 2011; Mari et al., 2008; Mari, 2002; Mari et al., 1999). As a result, a previous cleavage of protein glycans can be considered as a way to minimize the interference from spurious IgE binding in routine serum tests. Besides, additional assays must be taken into consideration to test not only the IgE binding capability but also the functionality of novel proteins to act as allergens, such as the cell-based assays (e.g. basophil degranulation test), and the *in-vivo* SPT and the 'gold standard' for food allergy diagnosis, double-blind placebo-controlled food challenge (DBPCFC) (EFSA, 2010).

This paper follows up the *in-silico* allergenicity predictions made in the previous work by D'Auria et al., (**Chapter 2**), by performing an *in-vitro* pre-screening using allergen-specific animal antibodies and a confirmatory IgE binding assays using sera of allergic patients. This study correlated the *in-silico* results with the *in-vitro* observations and paved the way for cellular based assays and preliminary *in-vivo* evaluations.

2. Materials and methods

2.1. Moringa leaves powder, Pru p 3 and other reagents samples

The M. oleifera leaves were supplied in a dried form by Sud Rienergy S.r.l. – Favella Group, Italy (Corigliano Calabro, Cosenza, Italy). Moringa leaves were randomly collected in the orchard from different plants in the period going from June to September 2019. Intact leaves were dried under the sun until a constant weight was reached. Afterwards, dried leaves were powdered in a blender and passed through 250 μm sieve to get a thin powder having a moisture 10 ± 1%. Samples were stored in amber glass containers, in a dry place at room temperature, until required. Unless otherwise stated, chemicals were purchased from the commercial origin (Sigma-Aldrich, Dorset, UK). Ultrapure water was purified using a Milli-Q ultrapure water system (Millipore, Darmstadt, Germany). Pru p 3 was purified from peach peel extract as described previously (Gaier et al., 2008).

2.2. Patient sera

Sera from Spanish peach allergic patients was kindly provided by Dr. María Ruano Zaragoza from Hospital General Universitario de Alicante (Valencia) with the approval from the

research committee ISABIAL (Instituto de Investigación Sanitaria y Biomédica de Alicante). The diagnosis of LTP allergy was previously confirmed by assessment of clinical manifestations, skin prick test and LTP specific IgE. From the patient pool, 7 patients with Pru p 3 specific IgE levels higher than 2.5 ISU were selected for this project. These patients were 3 females and 4 males, ranging from 22 to 44 years old (for details see **Supporting information, Table S1**). Sera from natural rubber latex allergic patients were got from PlasmaLab International (Everett WA, USA). The diagnosis of latex allergy was previously confirmed by immunocap FEIA. From the patient pool, only those with the highest latex specific IgE levels were selected for the analyses (patient 1: 17.9 kU/l, male; patient 2: 16.5 kU/l, female) (for details see **Supporting information, Table S2**).

2.3. Production and solubilisation of Moringa protein fraction (MPF) and of Moringa food grade protein isolate (MFG)

The Moringa protein fraction (MPF) was extracted from Moringa leaves powder and solubilised following the procedure of D'Auria et al., (**Chapter 2**).

Moringa leaves powder was also treated with a double extraction methodology in order to produce a food-grade protein isolate (MFG). Moringa leaves were extracted in a 1:5 ratio (w/v) either with 50 mM Tris-HCl pH 8.8 (MFG-) or with 50 mM Tris-HCl pH 8.8 + 10 mM DTT (MFG+). The mixtures were stirred for 2 h at room temperature and centrifugated at 4000 g for 15 min. After the centrifugation, the supernatants were recovered, and their protein content was evaluated.

2.4. Natural rubber latex protein extraction

The protein extraction from liquid natural rubber latex was performed following the procedure of Yeang, et al., (2002). Briefly, 50 ml of commercially available liquid latex (purchased in a local market) was ultracentrifuged at 44,000 g for 1 h (Beckman, Indianapolis, USA). The ultracentrifugation resulted in 4 main fractions (from top to down): i) rubber particles, ii) lipid phase, iii) cytosolic protein fraction (C-serum), and iv) the lutoid-body (bottom) fraction (B-serum). Due to the highest content of proteins, only B and C-sera were collected for the following analyses. Notably, literature data reported that B-serum is the latex fraction that contain the hevein allergen Hev b 6.02 (Subroto et al., 1996; Yeang et al., 2002). These fractions were mixed by vortexing (10 min) and ultrasonicated (20 °C, 10 min) to break the lutoid-bodies and release the proteins. The latex protein extract was stored at -20 °C until required.

2.5. Protein quantification

Protein quantification of MPF extracts was carried out following D'Auria et al., (**Chapter 2**). The protein content of MFG and latex extract was determined using the Biorad RC DC^{TM} kit (Biorad, Hertfordshire, UK), following the guidelines of the manufacturer. The absorbance was measured at 750 nm with a spectrophotometer (Oasys UVM 340 – Biochrom Ltd, Cambridge, UK). The calibration curve was prepared using various concentrations of BSA (0-1.5 mg/ml). The calibration curve followed the linear equation.

2.6. Protein glycan cleavage by PNGase F

Moringa leaf food grade protein extracts, latex proteins and purified Pru p 3 were deglycated by using the PNGase F glycan cleavage kit (Gibco, ThermoFisher Scientific, Hertfordshire, UK), following the manufacturer's instructions.

2.7. SDS-PAGE analysis

Samples were prepared for SDS-PAGE by mixing 65 µl of either Moringa or latex extracts or purified Pru p 3 with 25 µl NuPAGE (LDS) buffer (Thermo Scientific, Hertfordshire, UK), and either 10 µl 0.5M dithiothreitol (DTT) or 10 µl ultrapure water for non-reduced samples. Reduced samples (R) were boiled for 10 min while for those non-reduced (NR) this operation was avoided. The SDS-PAGE was performed on a NuPAGE system equipped with 4-12% Bis-Tris gels (1 mm × 26 wells, Invitrogen, Groningen, Netherlands), following the manufacturer's instruction. Before the separation, reduced and non-reduced samples of Moringa leaves (10 µg), latex (10 µg) and purified Pru p 3 (1 µg) were loaded in each corresponding well. The SDS-PAGE was performed both before and after PNGase F glycan cleavage of samples. SeeBlue^{TM} Prestained Standard (Invitrogen) was used as molecular weight marker. The separation was conducted at a constant voltage (200 V,350 mA and 100 W) for 35 minutes. SDS-PAGE resolved proteins were fixed for 1 h in fixing buffer (50% v/v Methanol, 10% v/v Acetic acid), rinsed three times in deionized water for 5 min and stained for 16 h using Coomassie G-250 stain (Simply Blue Safe Stain, Invitrogen, Paisley, UK). After being de-stained by rinsing with MilliQ water a few times, the gels were imaged using Typhoon gel Scanner (Amersham, UK).

2.8. Animal IgG and human sera IgE immunoblotting

Electrophoresis resolved proteins were electroblotted onto nitrocellulose membrane at 15 V, 350 mA for 15 min, using a Trans-blot SD semi-dry transfer cell (Biorad, Hertfordshire, UK) as previously described (Smith et al., 2015). The nitrocellulose membrane was blocked for 1 h at room temperature with 5% (w/v) skim milk powder in washing buffer (1× PBS, 0.05% Tween 20, pH 7.4). After extensive washings with washing buffer (4 × 5) min the membrane was incubated with anti–Pru p 3 rabbit polyclonal IgG (produced in UNIMAN – Manchester, UK) diluted 1:5000 (v/v) in the antibody (Ab) dilution buffer [2.5% (w/v) skim milk powder in washing buffer] under mechanic agitation at 4 °C for 16 h. The membrane was rinsed 4 × 5 min with washing buffer and incubated 1 h with monoclonal goat anti-rabbit IgG conjugated with alkaline phosphatase (AP) (ThermoFisher Scientific, Hertfordshire, UK), diluted 1:5000 (v/v) with antibody dilution buffer. The membrane was rinsed with washing buffer (4 × 5 min) before staining. The staining was obtained with the addition of the substrate (1-StepTM NBT/BCIP, Thermo scientific, USA, cod. 34042) for 10 min.

Blots were also developed using human sera [diluted 1:10 (v/v) in antibody dilution buffer] followed by several washing steps and by reaction with horseradish peroxidase (HRP)-conjugated mouse anti-human IgE antibody (SouthernBiotech, Alabama, USA) as previously described (Downs et al., 2016). The bound antibodies were revelated by using a chemiluminescent substrate (Pierce SuperSignal^{TM} West Dura Extended Duration Substrate). The membranes were imaged with equal exposure time using GeneGnome (Syngene, Cambridgeshire, UK). For the secondary antibody only control, antibody dilution buffer instead of the primary antibody was used in both IgG and IgE immunoblotting.

2.9. ELISA Immunoassays

The indirect and inhibition ELISA were performed using rabbit anti-Pru p 3 antibodies. The assay is described in detail in the **Supporting information 1.1.**

3. Results and discussion

As a novel protein source with high biological value, in our previous work we assessed *in-silico* the potential allergenicity of *Moringa oleifera* leaf. Among identified proteins, Morintides (mO1 and mO2) share a high degree of homology with Hev b 6.02, a major contact allergen from *Hevea brasiliensis*, suggesting their possible allergenicity. In addition, through homology mass spectrometry (MS) we identified the food panallergen nsLTP family. Although nsLTP of Moringa leaf has not been sequenced so far, the peptides we identified through MS, mapped a segment of an nsLTP inferred to *Rosaceae* family, which among nsLTPs are the most represented allergens (e.g., Pru p 3, Mal d 3, Rub i 3, etc.) (Borges et al., 2006). For these reasons, we decided to perform *in-vitro* evaluation involving nsLTP and latex allergic patients. In addition to providing other information about the allergenicity potential of Moringa leaf, these data will also serve to rank the reliability of the previous *in-silico* allergenicity predictions.

3.1. Production of protein isolates from Moringa oleifera leaves powder and natural rubber Latex

Two typologies of Moringa protein isolate were assayed in this study. The first was a phenol-based extract while the second was an extract obtained using 50 mM Tris-HCl (with and without DTT as the reducing agent). In view of producing a novel ingredient from Moringa leaves, the second

extraction, defined as "food grade" mimicked the industrial conditions of isolated protein preparation close to reality. Before the analysis, the phenol-based extract was solubilised in urea/thiourea buffer (U/T_b) or guanidine hydrochloride (GndHCl_b) to reproduce the condition described in D'Auria et al., (**Chapter 2**).

The protein concentration determined by D'Auria et al., was used to prepare the solutions used for the immune-based analysis to work at equal concentration of protein sample. The protein quantification of food-grade extract obtained under reducing and non-reducing conditions (MFG+ and -) reported a protein extraction of 7.2 mg/g of dry leaves powder and 5.2 mg/g dry leaves powder, respectively. The addition of a reducing agent such as DTT increased the extraction of proteins from Moringa leaves powder by about 40%. The combination of the C and B serum fractions of latex extracts showed a protein concentration of 13 mg/ml of extract corresponding to 1.3 mg/ml of latex. This result is far below literature values of about 36 mg/ml of C and B serum mixture (Yeang et al., 2002). This loss of proteins is possibly due to the nature of the latex; in fact, unlike the Yeang study, which analysed raw fresh latex, our investigation was based on commercially available latex, which generally undergoes processing steps like ammonia addition and partial vulcanisation that may cause to protein losses.

3.2. SDS-PAGE analysis of MPF and MFG

As previously stated in D'Auria et al., (**Chapter 2**), MPF is better solubilised in the Urea/Thiourea buffer than in the guanidinium hydrochloride (**Figure 1, panel A**) and showed the typical electrophoretic profile of leaf proteins extracted with phenol (Wang et al., 2016). Although the food-grade protein isolate (MFG) was blurrier due to the presence of co-extracted polyphenols and polysaccharides, the protein profile appeared more complex than the MPF (**Figure 1, panel D**).

The MFG profile showed greater number of electrophoretic bands below Mr ~ 50 kDa than the MPF. On the other hand, the MPF appeared to be more selective for some proteins with the presence of aggregated proteins (Mr >200 kDa). The electrophoretic profile of the proteins extracted with reducing agent was like the profile of proteins reduced after the extraction (**Figure 1, panel D**). New bands can be seen in the region between Mr ~ 60-40 kDa and Mr ~ 6 kDa. These bands may be polypeptides arose from higher molecular weight proteins after the reduction of the disulphide bonds. Interestingly, the MFG-R showed a unique band at Mr ~ 9.4 kDa, while MFG+R had a unique faint band at Mr ~ 12.3 kDa (**Figure 1, panel D**). Interestingly, the band at Mr ~ 9.4 kDa after glycan cleavage was more pronounced (**Figure 2, panel C**) suggesting the possible presence of additional deglycosylated polypeptides with similar molecular mobility.

3.3. Anti-Pru p 3 animal IgG immunoblotting and ELISA assays

The results of immunoblotting using polyclonal rabbit anti-Pru p 3 IgG, showed no response for MPF and MFG (**Figure 1, panel B and E**).

The presence of Pru p 3 analogous in MFG extracts was also investigated using ELISA assays. MPF samples were excluded since the presence of urea or guanidinium hydrochloride in the extracts could produce specious results. Both indirect and competitive ELISA did not show any significant response attributable to the presence of Pru p 3 analogues in Moringa extracts. Even if in the competitive ELISA, MFG showed lower absorption than the blank, this cannot be attributed to the presence of a reactive protein since the absorbance did not decrease with the increasing concentration. This feature indicates a higher background of the moringa extract compared to the blank buffer.

3.4. Human sera IgE immunoblotting of MFG before and after PNGase F glycan cleavage

3.4.1. Sera of patients allergic to nsLTP

Confirmatory human IgE assays were performed using from sera of nsLTP allergic patients. The results of immunoblotting using sera of nsLTP allergic subjects showed IgE binding in five patients out of seven. Particularly, a generalized IgE binding in the region of high and medium molecular weights (Mr > 198 to 20 kDa) was observed in patients 13 and 9 (**Figure 2, panel B**). This phenomenon was also observed in the positive control (purified Pru p 3) even if to a lower extent. A more specific IgE binding was evident for patients 12, 14 and 15, with a single band was recognised. This band was located at approximately Mr ~ 9.4 kDa in the profile of the MFG- R. Even if this region is typical of nsLTPs (Skypala et al., 2021), the presence of other IgE reactive polypeptides, different from nsLTP, cannot be ruled out. Curiously, this IgE response could be observed only in MFG-R, suggesting the presence of a reactive polypeptide released after reduction of disulphide bonds. Further investigations are required to confirm the identity of this polypeptide and unveil its IgE reactivity.

The IgE binding found in patients 9 and 13 could be non-specificities, attributable to CCDs or to a high level of total IgE derived from possible poly sensitisation to inhalant allergens (Jensen-Jarolim et al., 1998). For this reason, immunoblots with sera of patients 9 and 13 were repeated after deglycosylation with PNGase F (**Figure 2, panel D**). Despite the PNGase F glycan cleavage did show an impact in reducing non-specific IgE bindings in patients 9 and 13, it did not abolish the overall trend. Newly bands could be detected in the region of 28-17 kDa in patent 9. This was likely due to the remotion of steric hindrances due to glycans, which could have impeded the binding of the immunoglobulin before. Interestingly, after the PNGase F treatment, the specific

reactive band at Mr ~ 9.4 kDa observed in the other patients could be detected also in patient 9 (**Figure 2, panel D**).

3.4.2. Sera of patients allergic to Latex

Chapter 2 highlighted a high structural similarity between Morintides and hevein (Hev b 6.02). For this reason we expected an IgE reactivity around Mr ~ 4.7 kDa, which represent the characteristic molecular mobility of the hevein (Yeang et al., 2002). Remarkably, this reactivity has not been shown neither before nor after glycans cleavage (**Figure 3**). However, for the latex allergic patients, the overall western blotting exhibited a broad range of IgE reactivity. Even in this case, this could suggest the presence of CCDs or inhalant allergy among the patients.

The IgE binding profile differed significantly across the two patient sera. Patient 1 showed faint reactive bands after reduction in the region of Mr 49-62 kDa. The profile changed when proteins were reduced, as expected. After the glycans cleavage, IgE binding in patient 1 changed with an overall fainter immunoreactivity in the region of Mr 49-62 kDa and a newly faint band around Mr ~ 11 kDa corresponding to the latex IgE band (**Figure 3, panel B**). The positive control exhibited a specific IgE binding at Mr ~ 11.4 kDa under reducing and non-reducing conditions (**Figure 3, panel B**).

Patient 2 had a more complex and intense IgE binding profile with several bands reacting before and after reduction (**Figure 3, panel A**). Particularly, IgE bands could be observed after reduction around 28 kDa and above 62 kDa (**Figure 3, panel A**). The deglycosylation did not significantly change the immunoreactivity profile except for a band in the region of 38-49 kDa that was absent in the proteins extracted under reducing conditions (**Figure 3, panel B**). Likewise patient 1, the PNGase F treatment released an immunoreactive band at Mr ~ 9.6 kDa suggesting a specific IgE

binding to a potentially glycosylated polypeptide. Differently from Patient 1, with a specific IgE binding reactivity at 11 kDa, the positive control in patient 2 exhibited a broad IgE binding profile raging from Mr > 6 to 49 kDa.

Despite the collected evidence, the analyses suffer from two limitations. The sera were purchased with no indication about the specific latex allergen to which patients reacted. The latex extract used as a positive control was obtained from partially processed commercial latex, therefore may not contained Hev b 6, challenging the research of this specific allergen in MFG extracts. Further analyses using either fresh raw latex or purified hevein are required to refine the analysis.

4. Conclusions

This study highlighted two important aspects that should be considered during the preliminary allergenicity risk assessment: i) the use of animal IgG alone to pre-screen the presence of a potential allergen in complex matrices is not useful as a stand-alone; ii) bioinformatics prediction, although important, cannot be more than subsidiary in the overall preliminary phase of allergenicity risk assessment. Screening with target serum the protein extracts is essential to discover the presence of IgE reactive proteins. Besides, the inclusion of different types of controls in the experimental plan is crucial in preventing the specious assignation of allergenicity to non-allergic proteins. In this case, non-specific binding between IgE and CCDs could be found in Moringa proteins, indicating spurious IgE responses. Moreover, the analysis of proteins reduced and non-reduced is important to understand about the protein structure and the type of epitope involved in the IgE binding.

Our study demonstrated that CCDs may contribute to the overall IgE reactivity, although not to a great extent. Future investigations should confirm the nature of the peculiar polypeptide found in

the MFG-R profile and of its IgE reactivity. The next step is a further *in-vitro* investigation of the allergenicity potential of MFGs through IgE ELISA using patient sera and histamine release assays. These analyses should be coupled with clinical *in-vivo* assays such as SPT, or preferable, DBPCFC.

5. References

Alhakmani, F., Kumar, S., & Khan, S. A. (2013). Estimation of total phenolic content, in–vitro antioxidant and anti–inflammatory activity of flowers of Moringa oleifera. *Asian Pacific Journal of Tropical Biomedicine*, *3*(8), 623–627. https://doi.org/10.1016/S2221-1691(13)60126-4

Borges, J. P., Jauneau, A., Brulé, C., Culerrier, R., Barre, A., Didier, A., & Rougé, P. (2006). The lipid transfer proteins (LTP) essentially concentrate in the skin of Rosaceae fruits as cell surface exposed allergens. *Plant Physiology and Biochemistry*, *44*(10), 535–542. https://doi.org/10.1016/j.plaphy.2006.09.018

Downs, M. L., Simpson, A., Custovic, A., Semic-Jusufagic, A., Bartra, J., Fernandez-Rivas, M., Taylor, S. L., Baumert, J. L., & Mills, E. N. C. (2016). Insoluble and soluble roasted walnut proteins retain antibody reactivity. *Food Chemistry*, *194*, 1013–1021. https://doi.org/10.1016/j.foodchem.2015.08.119

EFSA. (2010). Scientific Opinion on the assessment of allergenicity of GM plants and microorganisms and derived food and feed. *EFSA Journal*, *8*(7), 1–168. https://doi.org/10.2903/j.efsa.2010.1700

Fernandez, A., Mills, E. N. C., Koning, F., & Moreno, F. J. (2021). Allergenicity Assessment of Novel Food Proteins: What Should Be Improved? *Trends in Biotechnology*, *39*(1), 4–8. https://doi.org/10.1016/j.tibtech.2020.05.011

Gaier, S., Marsh, J., Oberhuber, C., Rigby, N. M., Lovegrove, A., Alessandri, S., Briza, P., Radauer, C., Zuidmeer, L., Van Ree, R., Hemmer, W., Sancho, A. I., Mills, C., Hoffmann-Sommergruber, K., & Shewry, P. R. (2008). Purification and structural stability of the peach allergens Pru p 1 and Pru p 3. *Molecular Nutrition and Food Research*, *52*(SUPPL. 2), 220–229. https://doi.org/10.1002/mnfr.200700274

Jensen-Jarolim, E., Vogel, M., Zavazal, V., & Stadier, B. M. (1998). Nonspecific binding of IgE to allergens. *American Journal of Rhinology*, *12*(2), 150–151.

Kaulfürst-Soboll, H., Mertens, M., Brehler, R., & von Schaewen, A. (2011). Reduction of cross-reactive carbohydrate determinants in plant foodstuff: Elucidation of clinical relevance and implications for allergy diagnosis. *PLoS ONE*, *6*(3), 10–13. https://doi.org/10.1371/journal.pone.0017800

Mari, A., Ooievaar-de Heer, P., Scala, E., Giani, M., Pirrotta, L., Zuidmeer, L., Bethell, D., & Van Ree, R. (2008). Evaluation by double-blind placebo-controlled oral challenge of the clinical relevance of IgE antibodies against plant glycans. *Allergy*, *63*(7), 891–896. https://doi.org/10.1111/j.1398-9995.2008.01703.x

Mari, Adriano. (2002). IgE to cross-reactive carbohydrate determinants: analysis of the distribution and appraisal of the in vivo and in vitro reactivity. *International Archives of Allergy and Immunology*, *129*(4), 286–295. https://doi.org/10.1159/000067591

Mari, Adriano, Iacovacci, P., Afferni, C., Barletta, B., Tinghino, R., Di Felice, G., & Pini, C.

(1999). Specific IgE to cross-reactive carbohydrate determinants strongly affect the in vitro diagnosis of allergic diseases. *Journal of Allergy and Clinical Immunology*, *103*(6), 1005–1011. https://doi.org/10.1016/S0091-6749(99)70171-5

Nouman, W., Anwar, F., Gull, T., Newton, A., Rosa, E., & Domínguez-Perles, R. (2016). Profiling of polyphenolics, nutrients and antioxidant potential of germplasm's leaves from seven cultivars of Moringa oleifera Lam. *Industrial Crops and Products*, *83*, 166–176. https://doi.org/10.1016/j.indcrop.2015.12.032

Radauer, C., & Breiteneder, H. (2019). Allergen databases—A critical evaluation. *Allergy: European Journal of Allergy and Clinical Immunology*, *74*(11), 2057–2060. https://doi.org/10.1111/all.13841

Skypala, I. J., Asero, R., Barber, D., Cecchi, L., Diaz Perales, A., Hoffmann-Sommergruber, K., Pastorello, E. A., Swoboda, I., Bartra, J., Ebo, D. G., Faber, M. A., Fernández-Rivas, M., Gomez, F., Konstantinopoulos, A. P., Luengo, O., van Ree, R., Scala, E., & Till, S. J. (2021). Non-specific lipid-transfer proteins: Allergen structure and function, cross-reactivity, sensitization, and epidemiology. *Clinical and Translational Allergy*, *11*(3). https://doi.org/10.1002/clt2.12010

Smith, F., Pan, X., Bellido, V., Toole, G. A., Gates, F. K., Wickham, M. S. J., Shewry, P. R., Bakalis, S., Padfield, P., & Mills, E. N. C. (2015). Digestibility of gluten proteins is reduced by baking and enhanced by starch digestion. *Molecular Nutrition and Food Research*, *59*(10), 2034–2043. https://doi.org/10.1002/mnfr.201500262

Subroto, T., Van Koningsveld, G. A., Schreuder, H. A., Soedjanaatmadja, U. M. S., & Beintema, J. J. (1996). Chitinase and β-1,3-glucanase in the lutoid-body fraction of Hevea latex.

Phytochemistry, *43*(1), 29–37. https://doi.org/10.1016/0031-9422(96)00196-3

Sujatha, B. K., & Patel, P. (2017). Moringa Oleifera – Nature ' s Gold. *Imperial Journal of Interdisciplinary Research*, *3*(5), 1175–1179.

Verhoeckx, K., Broekman, H., Knulst, A., & Houben, G. (2016). Allergenicity assessment strategy for novel food proteins and protein sources. *Regulatory Toxicology and Pharmacology*, *79*, 118–124. https://doi.org/10.1016/j.yrtph.2016.03.016

Wang, L., Zou, Q., Wang, J., Zhang, J., Liu, Z., & Chen, X. (2016). Proteomic Profiles Reveal the Function of Different Vegetative Tissues of Moringa oleifera. *Protein Journal*, *35*(6), 440–447. https://doi.org/10.1007/s10930-016-9690-x

Yeang, H. Y., Arif, S. A. M., Yusof, F., & Sunderasan, E. (2002). Allergenic proteins of natural rubber latex. *Methods*, *27*(1), 32–45. https://doi.org/10.1016/S1046-2023(02)00049-X

Figures

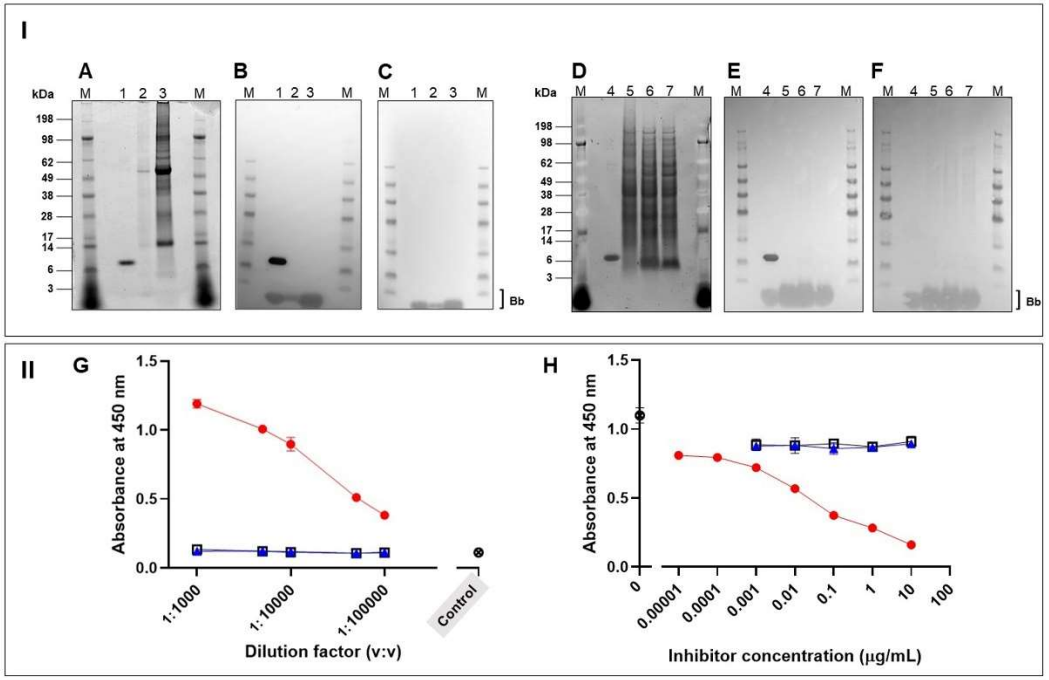


Figure 1. Animal IgG Immunoblot and ELISA of moringa leaf extracts developed with anti-Pru p 3 antibody preparation. I) SDS-PAGE and immunoblotting. Moringa proteins were extracted in either phenol (A, B, C) or in 50 mM Tris pH 8.8 (D, E, F). Samples of the SDS-PAGE (A and D), the immunoblots using rabbit anti-Pru p 3 IgG (B and E) and 2nd antibody only control (C and F) were loaded as following: Lane 1 and 4: Purified Pru p 3; Lane 2: Moringa protein extracted in phenol and solubilised in GndHCl_b; Lane 3: Moringa protein extracted in phenol and solubilised in U/T_b; Lane 5: Moringa protein extracted in 50 mM Tris pH 8.8 without DTT, non-reduced; Lane 6: Moringa protein extracted in 50 mM Tris pH 8.8 without DTT, reduced; Lane 7: Moringa protein extracted in 50 mM Tris pH 8.8 with DTT. Bb: Bromophenol blue

II) ELISA for rabbit anti-Pru p 3 IgG reactivity to Moringa extracts. G) Indirect ELISA using plates coated with purified Pru p 3 (●), Moringa extracted with DTT (Moringa DTT+) (▲) and without DTT (Moringa DTT-) (□). Control (⊗) used PBST instead of primary antibody; **H**) Inhibition ELISA using plate coated with purified Pru p 3. Control (⊗) was done as 0 µg/mL inhibitor.

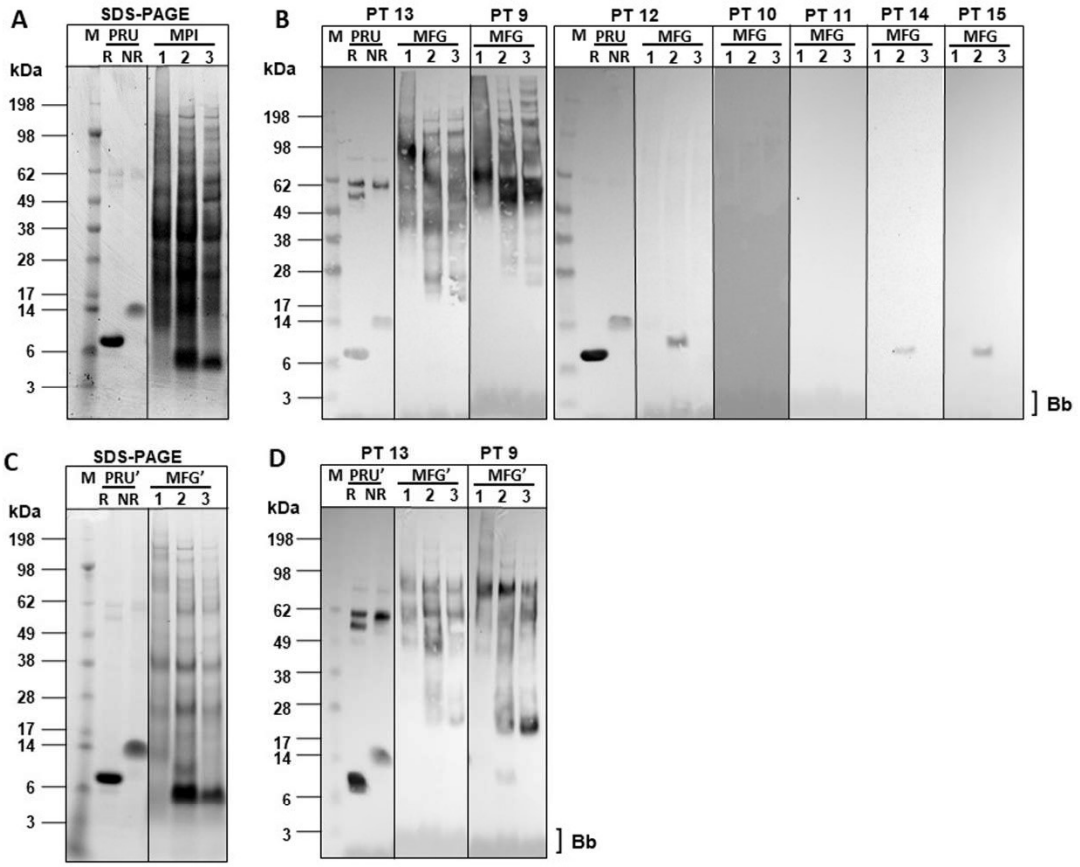


Figure 2. Immunoreactivity of Moringa food grade isolated proteins (A, B) and those proteins after PNGase f deglycosylation (C, D) using sera from patients (PT) allergic to nsLTP. M: molecular standards, PRU: Pru p 3, MFG: Moringa food grade protein isolate, PRU': Pru p 3 after PNGase F deglycosylation; MFG': Moringa food grade protein isolate after PNGaseF deglycosylation. R: reduced, NR: non-reduced. The samples of moringa were loaded as following:
Lane 1: Moringa protein extracted in 50 mM Tris pH 8.8 without DTT, non-reduced;
Lane 2: Moringa protein extracted in 50 mM Tris pH 8.8 without DTT, reduced;
Lane 3: Moringa protein extracted in 50 mM Tris pH 8.8 with DTT, reduced.
Bb: Bromophenol blue

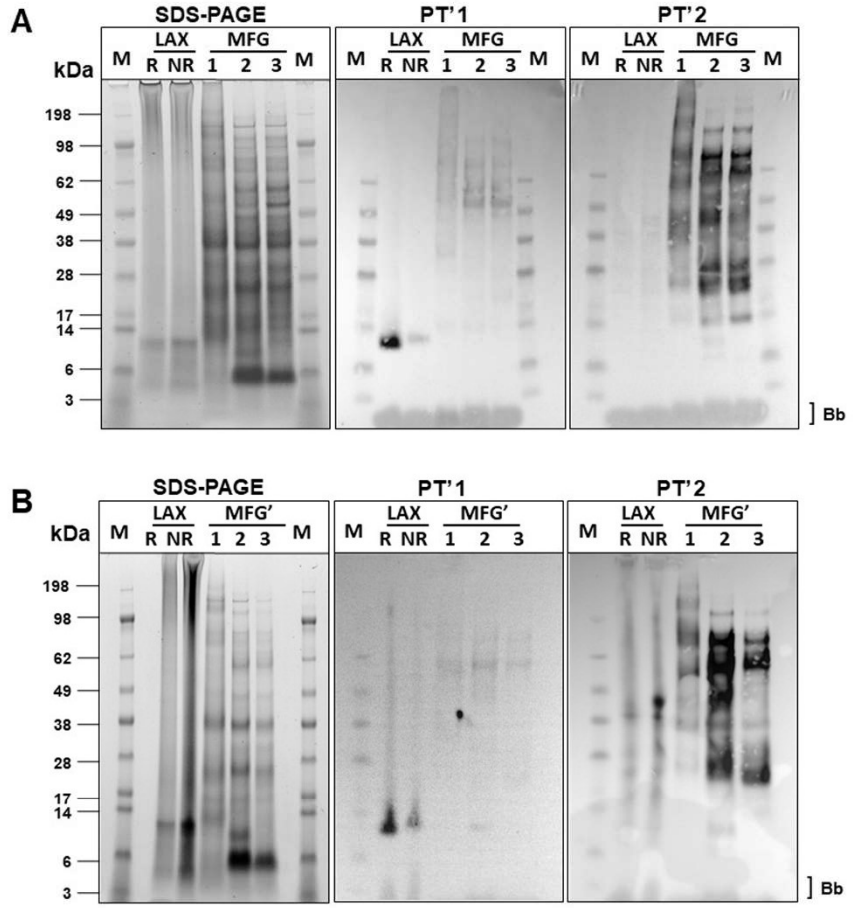


Figure 3. Immunoreactivity of Moringa food grade isolated proteins (A) and those proteins after PNGase f deglycosylation (B) using sera from patients (PT') allergic to latex. M: molecular standards, LAX: Latex, MFG: Moringa food grade protein isolate, LAX': Latex after PNGase F deglycosylation; MFG': Moringa food grade protein isolate after PNGaseF deglycosylation. R: reduced, NR: non-reduced. The samples of moringa were loaded as following:
Lane 1: Moringa protein extracted in 50 mM Tris pH 8.8 without DTT, non-reduced;
Lane 2: Moringa protein extracted in 50 mM Tris pH 8.8 without DTT, reduced;
Lane 3: Moringa protein extracted in 50 mM Tris pH 8.8 with DTT, reduced.
Bb: Bromophenol blue

Supporting information

1. Methods

1.1 Immunoassay

Inhibition ELISA was carried out similarly as described in Toda et al., (2011) with a few modifications. In brief, flat-bottom 96 well plates (Nunc MaxiSorp, Thermo Scientific, Waltham, USA) were coated with purified intact Pru p 3 (10 µg/mL) diluted in 1X PBS (pH 7.4) at 4 °C for 16 h. Plates were washed with washing buffer (0.05% Tween 20 in 0.01 M PBS, pH 7.4) and blocked with 0.1% BSA in 0.01 M PBS, pH 7.4 for 2 hours at 37 °C. A volume of 50 µl of diluted animal IgG rabbit anti Pru p 3 (1:5000 v/v) was incubated at 37 °C for 1 hour with serial dilutions of intact Pru p 3 and MFG extracts (0.00001 to 10 µg/ml). Inhibition with 0.01 M PBS buffer only served as a negative control. To plates were added 100 µl/well of the Ab-protein extracts mixture and incubated overnight at 4 °C. IgE binding was determined by incubating for 1 hour at 37 °C with 1:5,000 goat anti-rabbit IgG antibody conjugated to horseradish peroxidase (HRP) (Southern Biotech, Cambridge, UK) (100 µL). The substrate used was 3,3′,5,5′-tetramethylbenzidine (TMB) (ThermoFisher Scientific, Hertfordshire, UK).

For the indirect ELISA assay, a 96-wells plate were coated overnight (16 h) at 4 °C with 10 µg/ml of purified Pru p 3 (positive control) and of MFG (+ and -) dissolved in 1X PBS buffer (test samples). The wells were washed four times with the ELISA washing solution (0.05% Tween 20 in 0.01 M PBS, pH 7.4) and blocked in 0.1% BSA in 0.01 M PBS, pH 7.4 for 2 hours at 37 °C. Afterwards, dilution from 1:1000 to 1:100000 (v/v) of the first animal antibody (IgG Rabbit anti-Pru p 3) were loaded into the wells and made react overnight (16 h) at 4 °C. The wells were washed four times with the washing solution and then a dilution of 1:5000 (v/v) of the second

antibody (goat anti-rabbit HRP conjugated) was added to each well. The reaction was conducted for 1 h at 37 °C. The development was performed adding to each well 100 µl of 1 StepTM Turbo TMB-ELISA reactive (Thermo scientific, USA). The plate was incubated in dark for 30 minutes at room temperature (20°C).

For both inhibition and indirect ELISAs, the reaction was stopped with 1 M hydrochloric acid and absorbance was measured at 450 nm on a spectrophotometer (Oasys UVM 340 - Biochrom Ltd, Cambridge, UK). Results were analysed with GraphPad Prism 9 adjusting the curve to a sigmoidal model from.

Supplementary tables

Table S1. Summarised information for patients allergic to LTP. Patient data includes gender, age, grade of symptom severity according to Hugh Sampson's classification, LTP sIgE concentration and sIgE levels upon Pru p 3 exposure. (ISU= ISAC standardised units, M= male, F= female). R, rhinitis; A, asthma.

Patient No.	Gender	Age (Years)	LTP sIgE (KU/l)	sIgE Pru p 3 (ISU)	Aero allergy related symptoms
9	M	34	13	5.10	R; A
10	F	31	6	4.58	R
11	M	22	6	4.16	R
12	M	28	0	3.85	R
13	M	30	9	3.73	R; A
14	F	44	6	2.96	No
15	F	26	0	2.56	R

Table S2. Summarised information for patients allergic to latex. Patient data includes gender, age, sIgE concentration and sIgE levels upon latex exposure. M= male, F= female. n.a. not available.

Patient No.	**1**	**2**
Gender	M	F
Age (Years)	52	41
Allergen sIgE (KU/l)		
Latex	17.9	16.5
Almond	12.2	n.a.
Cherry	0.89	n.a.
Soybean	4.14	n.a.
Tomato	4.13	n.a.
Shrimp	n.a.	24.7
Carrot	n.a.	31.7
Common ragweed	n.a.	5.54
Celery	n.a.	15.9
Total IgE (IU/mL)	5662	n.a.
Self-reported allergies	Avocado; Raw vegetables and nuts	Banana; melons; avocado; papaya; tomatoes

SECTION 2

An innovative hybrid cereal: Tritordeum

Tritordeum is a hexaploid (AABBHchHch) hybrid derived from crossbreeding *Hordeum chilense*, with durum wheat. This grain represents an innovative crop and has been described to be resistant to abiotic stress and to present good bread-making flour comparable to those of soft wheat (Martín et al., 1999). The allergenicity and celiacogenic assessment of tritordeum flour is challenged by the contribution availability of several cultivars, hybridised with diverse *Hordeum chilense* and Triticum durum parents. Little proteomic characterisation of the cultivars is available in literature.

Seeds allergens

The major seeds' allergens belong to the prolamin superfamily (See Introduction to Section 1). Among the proteins belonging to the prolamin superfamily, trypsin and α-amylase inhibitory proteins (i.e., Tri a 28 and Tri a 29) are among the most important allergen in seeds, especially in wheat, barley and rice grains and in mung and soybean seeds (Nakase et al., 1996). This type of proteins are implicated in baker's asthma (wheat, barley, and rye), sometimes of wheat-dependent exercise-induced anaphylaxis (WDEIA) or in food anaphylaxis (Pastorello et al., 2007; Salcedo et al., 2003; Strobl et al., 1995). Wheat prolamins such as nsLTP (Tri a 14), α-gliadins (Tri a 21), γ-gliadins (Tri a 20), fast ω-gliadins (especially Tri a 19, Glia-ω5), and glutenins (Tri a 26 and 36), may also be responsible for baker's asthma, cereal hypersensitivity, anaphylaxis and WDEIA

(http://allergen.org/search.php?allergenname=tri+a+&searchname=Search) (Cianferoni, 2016; Morita et al., 2003; Sandiford et al., 1997).

Coeliac disease

It is possible for food intolerances to be immune-mediated as well as non-immune-mediated. As a non-immune response after ingestion of food, food intolerance manifests allergy-like symptoms (Ortolani & Pastorello, 2006). The condition is caused by enzyme deficiencies, pharmacological reactions, and responses to toxic or irritating components in foods. Globally, experts estimate that 5% - 6% of adults suffer from food intolerance, and that percentage ranges from 0.3 % to 20 % among infants and young children. Coeliac disease (CD), an enteropathy not mediated by IgE, is the most widely documented immune-mediated food intolerance, believed to affect 0.5 to 1% of people in Western countries (Choung et al., 2017). CD results from a complex interplay between a strong genetic component, immunological, and environmental factors. Almost all CD patients carry specific variants of the human leukocyte antigen (HLA) class II genes HLA-DQA1 and HLA-DQB1. Together, these two genes encode the two chains (α and β) of the coeliac-associated heterodimer proteins DQ2 and DQ8, which are expressed on antigen-presenting cells. About 90% of coeliac disease patients carry DQ2 and most others carry DQ8. Even if patients with coeliac disease usually carry HLA-DQ2, HLA-DQ8, the 40% of healthy people from the Americas, Europe, and southeast Asia possess these alleles, indicating that these genes are not sufficient for the development of coeliac disease (Lebwohl et al., 2018).

Rich in glutamines and prolines, gluten is incompletely digested by gastric and duodenal proteases and by brush border membrane peptidases. This reflects in the presence of long peptides, which enter the lamina propria of the small intestine, triggering the response of the adaptive immunity.

This response depends on deamidation of gliadin fraction by tissue transglutaminase (tTGase) that represents the main autoantigen in CD (Dieterich et al., 1997). Deamidation increases the immunogenicity of gliadin, facilitating the binding to the HLA-DQ2 or HLA-DQ8 molecules on the antigen presenting cell (Volta et al., 2008). Gliadins are then presented to gliadin reactive CD4+ T cells. Once again, some studies have shown that specific gliadin peptides activate innate immune cells, such as macrophages, dendritic cells, and cytotoxic intraepithelial lymphocytes, or to induce a direct toxic effect on enterocytes (Gianfrani et al., 2005; Sollid, 2000). In celiac patients, T cells recognized mostly α-gliadins peptides, whereas γ-gliadins and glutenins are much less recognized (Arentz-Hansen et al., 2000). Probably, this occurs because α-gliadin contains a 33-mer (Glia-α 57–89) fragment resistant to digestion, which contains a cluster of epitopes (Ozuna et al., 2015). Binding to chemokine receptor 3, alpha-gliadin motifs can modulate intestinal barrier functions by provoking the release of zonulin, which disassembly the interepithelial tight junction. This disassembling allows toxic peptides to penetrate the lamina propria, triggering CD enteropathy (Paterson et al., 2007).

The symptomatology around coeliac disease is divided into intestinal and extra intestinal. The most common intestinal symptoms comprise diarrhoea, loss of appetite, abdominal pain, and vomit. This symptomatology is, however, quite rare (Lebwohl et al., 2018). Extra-intestinal symptoms are more frequent; they include iron deficiency microcytic anaemia, deficiency of folic acid, osteoporosis, growth retardation (in children). The clinical strategies to diagnose CD comprise the evaluation of the presence of HLA-DQ2/8, the titration of anti tTGase IgGs, anti endomysial antibodies, and IgG-deamidated gliadin peptide (Caio et al., 2019). However, the intestinal damage (i.e., villous atrophy) can only be detected with duodenal biopsy, which represents the gold standard in CD diagnosis so far (Glissen Brown & Singh, 2019).

***Section 2* –An innovative hybrid cereal: Tritordeum**

To date, the only solution for patients with CD is a strictly gluten-free diet.

In vivo studies suggested the suitability of tritordeum bread in the diet of subjects affected by non-celiac gluten sensitivity (Vaquero et al., 2018). We demonstrated substantial differences in protein expression and immunoreactivity across two tritordeum cultivars. This suggests the need for a wider proteomic study including more varieties of tritordeum, aiming at selecting the cultivars with best technological and nutritional characteristics.

References

Arentz-Hansen, H., Körner, R., Molberg, Ø., Quarsten, H., Vader, W., Kooy, Y. M. C., Lundin, K. E. A., Koning, F., Roepstorff, P., Sollid, L. M., & McAdam, S. N. (2000). The Intestinal T Cell Response to α-Gliadin in Adult Celiac Disease Is Focused on a Single Deamidated Glutamine Targeted by Tissue Transglutaminase. *Journal of Experimental Medicine*, *191*(4), 603–612. https://doi.org/10.1084/jem.191.4.603

Caio, G., Volta, U., Sapone, A., Leffler, D. A., De Giorgio, R., Catassi, C., & Fasano, A. (2019). Celiac disease: A comprehensive current review. *BMC Medicine*, *17*(1), 1–20. https://doi.org/10.1186/s12916-019-1380-z

Choung, R. S., Larson, S. A., Khaleghi, S., Rubio-Tapia, A., Ovsyannikova, I. G., King, K. S., Larson, J. J., Lahr, B. D., Poland, G. A., Camilleri, M. J., & Murray, J. A. (2017). Prevalence and Morbidity of Undiagnosed Celiac Disease From a Community-Based Study. *Gastroenterology*, *152*(4), 830-839.e5. https://doi.org/10.1053/j.gastro.2016.11.043

Cianferoni, A. (2016). Wheat allergy: diagnosis and management. *Journal of Asthma and Allergy*, *9*, 13–25. https://doi.org/10.2147/JAA.S81550

Dieterich, W., Ehnis, T., Bauer, M., Donner, P., Volta, U., Riecken, E. O., & Schuppan, D. (1997). Identification of tissue transglutaminase as the autoantigen of celiac disease. *Nature Medicine*, *3*(7), 797–801. https://doi.org/10.1038/nm0797-797

Gianfrani, C., Auricchio, S., & Troncone, R. (2005). Adaptive and innate immune responses in celiac disease. *Immunology Letters*, *99*(2), 141–145. https://doi.org/10.1016/j.imlet.2005.02.017

Glissen Brown, J. R., & Singh, P. (2019). Coeliac disease. *Paediatrics and International Child Health*, *39*(1), 23–31. https://doi.org/10.1080/20469047.2018.1504431

Lebwohl, B., Sanders, D. S., & Green, P. H. R. (2018). Coeliac disease. *The Lancet*, *391*(10115), 70–81. https://doi.org/10.1016/S0140-6736(17)31796-8

Martín, A., Alvarez, J. ., Martín, L. ., Barro, F., & Ballesteros, J. (1999). The Development of Tritordeum: A Novel Cereal for Food Processing. *Journal of Cereal Science*, *30*(2), 85–95. https://doi.org/10.1006/jcrs.1998.0235

Morita, E., Matsuo, H., Mihara, S., Morimoto, K., Savage, A. W. J., & Tatham, A. S. (2003). Fast ω-gliadin is a major allergen in wheat-dependent exercise-induced anaphylaxis. *Journal of Dermatological Science*, *33*(2), 99–104. https://doi.org/10.1016/S0923-1811(03)00156-7

Nakase, M., Hotta, H., Adachi, T., Aoki, N., Nakamura, R., Masumura, T., Tanaka, K., & Matsuda, T. (1996). Cloning of the rice seed α-globulin-encoding gene: sequence similarity of the 5'-flanking region to those of the genes encoding wheat high-molecular-weight glutenin and barley D hordein. *Gene*, *170*(2), 223–226. https://doi.org/10.1016/0378-1119(95)00887-X

Ortolani, C., & Pastorello, E. A. (2006). Food allergies and food intolerances. *Best Practice & Research Clinical Gastroenterology*, *20*(3), 467–483. https://doi.org/10.1016/j.bpg.2005.11.010

Ozuna, C. V., Iehisa, J. C. M., Giménez, M. J., Alvarez, J. B., Sousa, C., & Barro, F. (2015). Diversification of the celiac disease α-gliadin complex in wheat: a 33-mer peptide with six overlapping epitopes, evolved following polyploidization. *The Plant Journal*, *82*(5), 794–805. https://doi.org/10.1111/tpj.12851

Pastorello, E. A., Farioli, L., Conti, A., Pravettoni, V., Bonomi, S., Iametti, S., Fortunato, D., Scibilia, J., Bindslev-Jensen, C., Ballmer-Weber, B., Robino, A. M., & Ortolani, C. (2007). Wheat IgE-Mediated Food Allergy in European Patients: α-Amylase Inhibitors, Lipid Transfer Proteins and Low-Molecular-Weight Glutenins. *International Archives of Allergy and Immunology*, *144*(1), 10–22. https://doi.org/10.1159/000102609

Paterson, B. M., Lammers, K. M., Arrieta, M. C., Fasano, A., & Meddings, J. B. (2007). The safety, tolerance, pharmacokinetic and pharmacodynamic effects of single doses of AT-1001 in coeliac disease subjects: a proof of concept study. *Alimentary Pharmacology & Therapeutics*, *26*(5), 757–766. https://doi.org/10.1111/j.1365-2036.2007.03413.x

Salcedo, G., Snchez-Monge, R., Garca-Casado, G., Armentia, A., Gomez, L., & Barber, D. (2003). The Cerealα-Amylase/Trypsin Inhibitor Family Associated with Bakers' asthma and Food Allergy. In *Plant Food Allergens* (pp. 70–86). Blackwell Publishing Ltd. https://doi.org/10.1002/9780470995174.ch5

Sandiford, C. P., Tatham, A. S., Fido, R., Welch, J. A., Jones, M. G., Tee, R. D., Shewry, P. R., & Newman Taylor, A. J. (1997). Identification of the major water/salt insoluble wheat

proteins involved in cereal hypersensitivity. *Clinical and Experimental Allergy : Journal of the British Society for Allergy and Clinical Immunology*, *27*(10), 1120–1129. http://www.ncbi.nlm.nih.gov/pubmed/9383251

Sollid, L. M. (2000). Molecular Basis of Celiac Disease. *Annual Review of Immunology*, *18*(1), 53–81. https://doi.org/10.1146/annurev.immunol.18.1.53

Strobl, S., Muehlhahn, P., Bernstein, R., Wiltscheck, R., Maskos, K., Wunderlich, M., Huber, R., Glockshuber, R., & Holak, T. A. (1995). Determination of the Three-Dimensional Structure of the Bifunctional .alpha.-Amylase/Trypsin Inhibitor from Ragi Seeds by NMR Spectroscopy. *Biochemistry*, *34*(26), 8281–8293. https://doi.org/10.1021/bi00026a009

Vaquero, L., Comino, I., Vivas, S., Rodríguez-Martín, L., Giménez, M. J., Pastor, J., Sousa, C., & Barro, F. (2018). Tritordeum: a novel cereal for food processing with good acceptability and significant reduction in gluten immunogenic peptides in comparison with wheat. *Journal of the Science of Food and Agriculture*, *98*(6), 2201–2209. https://doi.org/10.1002/jsfa.8705

Volta, U., Granito, A., Fiorini, E., Parisi, C., Piscaglia, M., Pappas, G., Muratori, P., & Bianchi, F. B. (2008). Usefulness of Antibodies to Deamidated Gliadin Peptides in Celiac Disease Diagnosis and Follow-up. *Digestive Diseases and Sciences*, *53*(6), 1582–1588. https://doi.org/10.1007/s10620-007-0058-0

Chapter 4

The effect of nitrogen fertilization on the expression of protein in wheat and tritordeum varieties using a proteomic approach.

Food Research International 148 (2021) 110617

Contents lists available at ScienceDirect

Food Research International

journal homepage: www.elsevier.com/locate/foodres

The effect of nitrogen fertilization on the expression of protein in wheat and tritordeum varieties using a proteomic approach

Viola Landolfi [a,1], Giovanni D'Auria [b,1], Maria Adalgisa Nicolai [b], Chiara Nitride [b], Massimo Blandino [a,*], Pasquale Ferranti [b]

[a] Università degli Studi di Torino, Dipartimento di Scienze Agrarie, Forestali e Alimentari Largo Braccini 2, Grugliasco, TO 10095, Italy
[b] Università di Napoli Federico II, Dipartimento di Agraria, Parco Gussone, Portici, NA 80055, Italy

ARTICLE INFO

Keywords:
Soft wheat
Tritordeum
N rate
Gliadins
Glutenins
Toxic epitopes
Allergens

ABSTRACT

Wheat, an essential ingredient for several bakery preparations, is also responsible for gluten-related diseases in sensitive subjects. The effect of the N fertilization rate (80 vs 160 kg N ha^{-1}) on gluten protein expression profile has been evaluated considering two soft wheats (landrace and modern) and one tritordeum cultivar (*cv*), grown in the same experimental field in North Italy. The proteins of refined flour were characterized through advanced proteomic approaches, including chromatography (RP-HPLC) and electrophoresis. A static model system was used to simulate *in vitro* digestion and the digestome peptides were examined by mass spectrometry and *in silico* approaches, to investigate the celiac and allergenic sequences. The CD-toxic epitopes in the digested samples were quantified by means of a R5 ELISA assay.

The N fertilization rate increased the grain protein content, but it did not lead to any difference in gluten composition, with exception of glu/glia ratio in the modern wheat *cv*. Moreover, the gluten composition and the occurrence of toxic/allergenic epitopes varied to a great extent, according mostly to the genotype. A lower immunoreactivity, determined using R5 ELISA, was detected for the digested tritordeum flours than for the landrace (-51%) or modern (-58%) cvs, while no significant difference was observed for the N rates between each genotype. *In silico* analysis showed that tritordeum has fewer CD epitopes belonging to the ω-gliadins and a lower LMW-GS than the landrace or modern *cv*. Tritordeum presented fewer α-gliadin allergenic epitopes than the modern wheat *cv*. The lower frequency of celiac epitopes in tritordeum, compared to the old and the modern wheat, is probably due to the absence of a D genome.

1. Introduction

The quality of soft wheat (*Triticum aestivum* L. subsp. *aestivum*) flour influences the organoleptic and structural properties of the final products to a great extent. Modern hexaploid wheats (AABBDD) are products of breeding processes that were aimed at producing cultivars (*cvs*) with an improved bread-making potential, a more balanced glutenin/gliadin (glu/glia) ratio and a higher high molecular weight glutenin (HMW-GS) / low molecular weight glutenin (LMW-GS) ratio (Dhaka & Khatkar, 2015) than wheat landraces (AABBDD) (local ecotype) or old wheat varieties, which have been cultivated since before the second half of the XX° century (Migliorini et al., 2016). A new cereal, tritordeum (x

Abbreviations: cvs, cultivars; glu, glutenins; glia, gliadins; HMW-GS, high molecular weight glutenins; LMW-GS, low molecular weight glutenins; GS, glutenin subunits; GPC, grain protein content; CD, celiac disease; NCGS, Non-Celiac Gluten Sensitivity; CPVO, Community Plant Variety Office; TW, test weight; TKW, thousand kernel weight; NIR, Near Infrared Reflectance; ACN, acetonitrile; 1-PrOH, propan-1-ol; EtOH, ethanol; AmBic, ammonium bicarbonate; SSF, simulated salivary fluid; SGF, simulated gastric fluid; SIF, simulated intestinal fluid; TCA, trichloroacetic acid; DTT, dithiothreitol; IAA, iodoacetamide; SDS, sodium dodecyl sulfate; Tris-HCl, tris(hydroxymethyl) aminomethane hydrochloride; EDTA, ethylenediaminetetraacetic acid; TFA, trifluoroacetic acid; TAME, *p*-toluenesulfonyl-L-arginine methyl ester; TPI, total protein isolates; 2-DE, 2-sulfanylethanol, two-dimensional electrophoresis; IPG, immobilized pH gradient; IEF, Isoelectrofic focusing; LIT, nano-ESI-linear ion trap; GID, gastrointestinal digested; Gln, glutamine; MW, molecular weight; OGD, oral, gastric and duodenal phases; WDEIA, wheat-dependent exercise-induced anaphylaxis.

* Corresponding author.
E-mail address: massimo.blandino@unito.it (M. Blandino).
[1] These authors contributed equally: Viola Landolfi and Giovanni D'Auria.

https://doi.org/10.1016/j.foodres.2021.110617
Received 6 April 2021; Received in revised form 15 July 2021; Accepted 19 July 2021
Available online 21 July 2021

Tritordeum martinii A. Pujadas, nothosp. nov.) (AABBHchHch), has recently received a great deal of scientific and commercial attention (Ribeiro et al., 2016; Vaquero et al., 2018). This amphiploid species, which is derived from the crossing of a South American wild barley (*Hordeum chilense* Roem. et Schultz.) and a durum wheat (*Triticum turgidum* L. subsp. *durum* Desf.), has been described as a good bread-making flour, with similar performances to those of soft wheat, and it has found its way into the baking industry (Martín et al., 1999).

Gluten, from a food technology perspective, is the main player in bread making, as it represents around 60% of the wheat proteome (Fiedler et al., 2014). Gluten can be subdivided, according to the Osborne classification, into 2 main fractions with the following subfractions : water-soluble non-prolamins, albumins and globulins; and prolamins, alcohol-soluble gliadins and polymeric alcohol-insoluble glutenins (Mamone et al., 2000). Monomeric gliadins are classified, on the basis of their electrophoretic mobility, into 3 fractions: α-, γ- and ω-gliadins. The polymeric glutenin subunits (GS) cover a wide range of molecular weights, and are conventionally classified as LMW-GS (12–60 kDa) or HMW-GS (up to 120 kDa) (Mamone et al., 2000). The rheological properties of the flour depend on the quality and the quantity of the gluten and in particular on the glu/glia and HMW-GS/LMW-GS ratios. A large amount of gliadins increases the viscosity of a dough, thereby influencing its extensibility, while a greater quantity of glutenins enhances the strength of the dough, thereby influencing its elasticity (Plessis et al., 2013). Gluten strength is defined by considering the composition of the gluten proteins, the quantity of the HMW glutenin subunits and the glu/glia ratio (Johansson et al., 2001).

The meteorological and soil conditions, the type of *cv* and the different nitrogen rates are all factors that qualitatively and quantitatively modulate the flour protein profile (Godfrey et al., 2010). The environmental conditions can influence the grain protein content (GPC) to a greater extent than the *cv* (Wan et al., 2013). Moreover, N fertilization is a crop technique that is expected to quantitatively change the GPC and the rheological properties of the derived flour (Garrido-Lestache et al., 2004).

The mapping of gluten expressed proteins is fundamental to assess the technological quality of different cereal species and cvs (Cho et al., 2018). Gluten proteins can cause inflammation of the small intestine in subjects affected by celiac disease (CD), which is a common autoimmune enteropathy of the small intestine, whose prevalence has risen in the last 50 years and today affects about 1% of Western populations (Lebwohl et al., 2015). The pathogenesis of CD begins with the gluten digestion derived peptides, which are transported throughout the microvilli and trigger an immune response of the T-cells (Dunne et al., 2020). CD disease has multifactorial etiologies, including the interaction of genetic factors in susceptible individuals (HLA-DQ2 or -DQ8 haplotype), environmental components and immunological mechanisms (Kagnoff, 2007). Non-Celiac Gluten Sensitivity (NCGS) is a "new" condition that has emerged in recent years, and it affects up to 6% of the world's population (Casella et al., 2018). Its diagnosis is a challenging task as it requires the clinical exclusion of CD and wheat allergies, and the symptoms are hard to distinguish from those of the irritable bowel syndrome (Casella et al., 2018). The triggering causes are not clear and require further investigation. Tritordeum bread was recently found to be tolerated by subjects with non-coeliac wheat sensitivity (NCWS) to a similar extent to gluten free bread (Sánchez-León et al., 2021). Extensive characterization studies, including proteomics and immunobased approaches, aimed at identifying the epitopes and toxic motifs responsible for triggering allergic and toxic reactions, are necessary. An extensive collection of known epitopes is available in the ProPepper™ database (Juhász et al., 2015).

A deep understanding of the role of the environmental and agronomical conditions, specifically the choice of *cv* and N fertilization, on the gluten composition and expression of specific epitopes, could provide useful technological information for bread making. In this respect, the type of selected wheat *cv* can also make a great difference in the consumers' acceptance of bakery products. In this paper, we present a proteomic characterization of three types of cereal grains: landrace wheat, modern wheat and tritordeum. Proteins isolated from grain flours were profiled using advanced proteomic approaches, including chromatography and electrophoresis. The flours were digested using an in vitro batch model to simulate the oral, gastric and duodenal compartments of humans. Mass spectrometry was used to create a picture of the obtained digestome. The collected peptide sequences were analyzed *in silico* for the presence of toxic and allergenic sequences. Additionally, a competitive ELISA analysis was performed on the digestome to evaluate any diversification of the immunoresponse related to fertilization and to the genotype.

2. Materials and methods

2.1. Experimental site and treatments

The effect of the N fertilization rate was evaluated by means of a full factorial experimental design, considering 2 soft wheat and 1 tritordeum *cvs*, characterized by different gluten compositions.

Two N fertilization treatments were compared: a low N rate (80 kg N ha^{-1}, N80) and a high N rate (160 kg N ha^{-1}, N160). The total N rate for each treatment was top-dressed applied as a granular ammonium nitrate fertilizer, split equally between tillering (growth stage 23) and the beginning of stem elongation (growth stage 32).

The considered genotypes were:

- Landrace, *cv* Andriolo, which was mainly cultivated in mountain areas in the Tuscany region (Italy) in the XIX° century, and is a soft wheat *cv* (hexaploid AABBDD) (Migliorini et al., 2016);
- Modern, *cv* Bologna (S.I.S., San Lazzaro di Savena, Bologna, Italy; genealogy (H89092 X H89136) X Soissons), is a soft wheat (hexaploid AABBDD) *cv*, classified as improver wheat according to the Synthetic Index of Quality (Foca et al., 2007), which was registered in the Italian varietal list in 2002 (https://www.sian.it/mivmPubb/listeVarieta.do);
- Tritordeum, *cv* Bulel (Agrasys S.L., Barcelona, Spain), which was registered in the CPVO List (Community Plant Variety Office) in 2011 (hexaploid AABBHchHch).

A field study was carried out in the Po plain at Carmagnola in North West Italy (44° 50′ N, 7° 40′ E; elevation 245 m), in the 2016–17 growing season. The experiment was performed on the University of Turin experimental farm in a deep silty-loam soil (Typic Udifluvents), characterized by a medium cation-exchange capacity and organic matter content.

The same agronomic techniques were adopted for all the *cvs*. Briefly, the previous crop was soft wheat, and planting was performed in 12 cm wide rows at a seeding rate of 400 seeds/m^2 on 27 October 2016, following an autumn plowing (30 cm) and disk harrowing to prepare a proper seedbed. The weed control was in accordance with the agronomic management practices usually carried out in the North of Italy for the cultivation of soft wheat, while no fungicide or insecticide was applied. Harvesting was carried out with a combine-harvester on 13 July 2017. The N rates (in three replicates) were assigned to experimental units, using a completely randomized block design. The plot size was 7 × 1.5 m.

2.2. Grain yield and kernel quality traits

The grain yields were obtained by harvesting the whole plot with a Walter Wintersteiger cereal plot combine-harvester. Grain moisture was analyzed using a Dickey-John GAC2100 grain analyzer (Auburn, IL, The USA). The grain yield results were adjusted to a 13% moisture content. Three kg grain samples, which were milled (Bona 4RB mill) to obtain refined flour, were obtained from each plot.

The test weight (TW), thousand kernel weight (TKW) and grain protein content (GPC; N × 5.7, on a dry matter basis) were determined according to Blandino et al. (2015). GPC was determined on whole grains (1-mm-sieve Cyclotec mill), by means of NIR (Near Infrared Reflectance) spectroscopy, according to AACC 39–10 (AACC, 2000).

2.3. Materials and chemical reagents

All the reagents used in this study were of analytical or higher grade. The HPLC-grade solvents: water, acetonitrile (ACN), propan-1-ol (1-PrOH) and ethanol (EtOH) were from Carlo Erba (Milan, Italy). Sodium phosphate, ammonium bicarbonate (AmBic), acetic acid and the other chemicals, used to produce the simulated salivary fluid (SSF), the simulated gastric fluid (SGF) and the simulated intestinal fluid (SIF), were also provided by Carlo Erba. The enzymes used for in vitro human digestion were purchased from Sigma (St Louis, MO, USA), in line with those indicated by the Infogest protocol (Brodkorb et al., 2019). Trichloroacetic acid (TCA), dithiothreitol (DTT), iodoacetamide (IAA), sodium dodecyl sulfate (SDS), glycerol, tris(hydroxymethyl) aminomethane hydrochloride (Tris-HCl), ethylenediaminetetraacetic acid (EDTA), guanidine chloride, trifluoroacetic acid (TFA), 2-vinylpyridine monomer and *p*-toluenesulfonyl-L-arginine methyl ester (TAME) were purchased from Sigma-Aldrich (St Louis, MO, USA). Egg lecithin was purchased from Lipid Products (Redhill, UK). The electrophoresis reagents were all from Bio-Rad (Milan, Italy).

2.4. Protein extraction

2.4.1. Production of the total protein isolate

The total protein isolate (TPI) of the wheat samples was produced according to the procedure of (Dupont et al., 2011). An aliquot of 1 ml of SDS buffer (2% SDS, 10% glycerol, 50 mM DTT, 40 mM Tris-HCl, pH 6.8) was added to the flour samples (50 mg), which were then incubated at room temperature with discontinuous mixing for 1 h. The obtained pellets were removed by centrifugation at 7,900 g for 10 min (Microcentrifuge Multispin 12, Steroglass, Perugia, Italy). The proteins in the supernatant were precipitated with 4 vol cold (-20 °C) propan-2-one, following incubation overnight at −20 °C and centrifugation at 7,900 g for 10 min. The pellets were dried under a nitrogen stream and stored at −20 °C.

2.4.2. Selective extraction of the gliadins and glutenins

The protein fractions were extracted according to Mamone et al. (2000), with slight modifications. The albumins and globulins were removed from non-defatted flour (1 g) by adding a solution (10 ml) containing 100 mM KCl, 50 mM Tris-HCl pH 7.8 and 5 mM EDTA, and centrifuging for 15 min at 3,433 g. After removal of the supernatant, the gliadins were extracted with 10 ml of 70% (v/v) ethanol (twice) and the glutenins with 50% v/v 1-PrOH; 50 mM Tris-HCl (pH 8.5); 1% (w/v) DTT. Glutenin extraction was performed at 60 °C for 30 min. The solutions were mixed for 2 min and stirred on a magnetic plate for 10 min. The cysteine residues of the glutenin extracts were pyridylethylated (VP glutenins) at 60 °C for 15 min with 2-vinylpyridine.

2.5. HPLC analysis

Reversed phase (RP)-HPLC analysis of the gliadins and glutenins was carried out in an HPLC chromatograph (HP 1100 Agilent, Palo Alto, CA, USA) modular system equipped with a diode array detector. The column effluent was monitored, by means of UV detection, at λ = 220 and 280 nm. The separation was performed using a C8 Vydac, 208TP52, 2.1 × 250 mm, 5 μm column (Hesperia, CA, The USA). Solvent A was 0.1% TFA v/v in water, while solvent B was 0.1% TFA in ACN. After 5 min of isocratic elution, using 25% solvent B (0.1% TFA in ACN, v/v), a 25–60% gradient ramp was applied for 60 min at a flow rate of 0.200 ml/min. This procedure was performed according to the work of Mamone et al. (2000) with some modifications. The peaks were then integrated using Agilent ChemStation software to obtain the relative quantity of both the gliadins and glutenins in each sample.

2.6. SDS–PAGE of the gliadins and glutenins

The glutenins and gliadins obtained from the selective extraction (section 2.4.2.), were precipitated in cold propan-2-one (1:4 v:v), suspended in an SDS-PAGE Laemmli Buffer [0.125 M Tris–HCl pH 6.8, 5% SDS, 20% glycerol, 5% (w/v) 2-sulfanylethanol, 0.02% bromophenol blue and boiled in a water bath for 5 min. The proteins (10 mg/ml) were loaded onto a 12% polyacrylamide gel in a final volume of 3 μL. Analysis was carried out at room temperature and constant voltage (100 V). After migration, the proteins were fixed overnight with TCA (24%) and stained with Coomassie G-250. The electrophoretic profiles were analyzed using GelAnalyzer 19.1 software (www.gelanalyzer.com by Istvan Lazar Jr., PhD and Istvan Lazar Sr., PhD, CSc).

2.7. Two-dimensional electrophoresis analysis of the total protein isolate (TPI)

The TPI was dissolved (3 mg/ml) in an immobilized pH gradient (IPG) strip rehydration buffer, containing 8 M urea, 2% (w/v) CHAPS, 2% (v/v) Pharmalytes pH 3.0–10.0, 20 mM DTT and traces of bromophenol blue, for two-dimensional electrophoresis (2-DE). The isoelectric focalisation was performed using pH 3–10 Immobiline Dry Strips (11 cm) linear gradient, from Bio-Rad Laboratories. The strips were rehydrated overnight in an Immobiline Dry-Strip Reswelling Tray (Amersham Pharmacia). Isoelectrofic focusing (IEF) was carried out using the Multiphor II system (Pharmacia Biotech, Uppsala, Sweden). The proteins were focused on up to 15,000 Vh at a maximum voltage of 6000 V at 20 °C. After focusing, the proteins were reduced for 15 min in an equilibration buffer (6 M urea, 30% glycerol, 2% SDS, 2% DTT), and alkylated for 15 min with 2.5% IAA. SDS-PAGE was performed in the second dimension on a 10% polyacrylamide gel, using a Protean II system (BioRad Laboratories, Hercules, California, USA). The run was performed at a constant voltage of 220 V. After migration, the proteins were fixed overnight with TCA (24%) and stained with Coomassie G-250.

2.8. In batch gastroduodenal digestion of flour

Flour samples were subjected to in vitro simulated human digestion, using the static model system optimized in the framework of the Infogest COST Action project (Brodkorb et al., 2019). Simulated salivary fluid (SSF), simulated gastric fluid (SGF), and simulated intestinal fluid (SIF) were prepared according to the harmonized conditions (Brodkorb et al., 2019). Mastication was simulated using a manual mincer. One g of flour was mixed with 1 ml of SSF and minced for 2 min at 37 °C. The SSF included 1500 U ml^{-1} human salivary amylase. SGF was added to the bolus (50:50 v/v); a concentration of up to a 0.17 mmol/L of gastric liposome (egg lecithin prepared in a vesicular form) and 2000 U/ml of SGF porcine pepsin was added to this mixture. The pH of the digesta was adjusted to 2.7 and the incubation was carried out for 2 h at 37 °C under magnetic stirring. After adding SIF (50:50 v/v) to the gastric digesta, the pH was adjusted to 7.0, using NaOH 1 M. Ten mmol/L of bile salts (5 mmol/L sodium taurocholate, 5 mmol/L sodium glycodeoxycholate and 1.8 mmol/L egg lecithin in the final concentration), and pancreatin were added to the intestinal mixture. The amount of pancreatin was calculated on the basis of the measured trypsin activity, and a final volume of up to 100 U of trypsin/ml was added. Trypsin activity was determined by means of a TAME assay (Brodkorb et al., 2019). Each sample was incubated for 2 h at 37 °C under magnetic stirring. The enzymatic activity was stopped by boiling the samples for 5 min in a water bath. The samples were then centrifuged at 3,433 g for 15 min to collect the supernatants, which were then stored at −20 °C until use. The digestion

was monitored by means of SDS-PAGE analysis. A part of the supernatant of each sample (1 ml) was desalted using a Phenomenex (Torrance, CA, USA) Strata-X SPE cartridge following the manufacturer instruction and then concentrated with gaseous nitrogen before the LC-MS/MS analysis was performed.

2.9. R5 competitive ELISA assay on the digested samples

RIDASCREEN® Gliadin competitive (Art. No. R7021, R-BIOPHARM AG, Darmstadt, Germany) was used according to the manufacturer's instructions. This Competitive ELISA relies on the use of the R5 monoclonal antibody (Di Stasio et al., 2020). The gastroduodenal digestion products were analyzed in duplicate.

2.10. LC-MS/MS analysis

LC-MS analysis was performed using a Dionex UltiMate 3000 nano-UHPLC system coupled with a nano-ESI-linear ion trap (LIT) Thermo XL mass spectrometer (Thermo Fisher Scientific, Waltham, MA, The USA). Samples were resuspended in a 0.1% (v/v) formic acid solution, loaded through a 5 mm long, 300 µm id pre-column (LC Packings, The USA) and separated in an Acclaim™ PepMap™ C18 column (150 mm × 75 µm, 3 µm) at a flow rate of 0.200 µL/min. Eluent A was 0.1% formic acid (v/v) in Milli-Q water; eluent B was 0.1% formic acid (v/v) in ACN. The column was equilibrated at 5% B. Peptides were separated by applying a 5–40% gradient of B over 40 min. MS data were obtained over the 200 to 2000 m/z mass range. Data-dependent MS/MS spectra were collected from the five most abundant precursor ions upon fragmentation (charge state ≥ 2; isolated width: 2 Da; min. signal required: 500), using CID activation with 35.0% normalized collision energy, an activation Q of 0.25, and an activation time of 30 ms. The spectra were processed using Xcalibur Software, 3.1 version (Thermo Scientific). The mass spectra were then analyzed using Protein Prospector software. The GluPro v 1.2 database of wheat gluten protein sequences (Daly et al., 2020) was used as background database for the analysis of the mass spectrometry data. The database searching parameters used for the identification of the peptidomes of the simulated gastrointestinal digested (GID) flour included "no enzyme" specification, and Met-oxidation and pyroglutamic acid for N-terminus glutamine (Gln) as variable protein modifications, with a mass tolerance value of 1 Da for the precursor ion and 0.6 Da for the MS/MS fragments. Analyses were carried out in triplicate.

2.11. Searching criteria for the celiac disease related epitopes, allergic epitopes and biopeptides

The identified peptides were searched manually for the presence of described toxic (CD epitopes), allergenic (IgE-binding epitopes) and bioactive amino acid sequences. The applied bioinformatic pipeline is described in Fig. 1. The Propepper database (Juhász et al., 2015) and the CD epitope table presented in Mamone et al. (2011) were used as the sources of the CD-toxic motifs. Only toxic motifs that appeared in at least 3 identified peptides were selected for comparison.

The IgE binding sequences were retrieved from the Immune Epitope Database (IEDB) (Vita et al., 2019), a repository of peer-reviewed epitopic sequences that have shown proved ability to trigger immune responses. The following search parameters were applied to extract IgE epitopes from the database: substring, allergy disease, and Triticum: ID

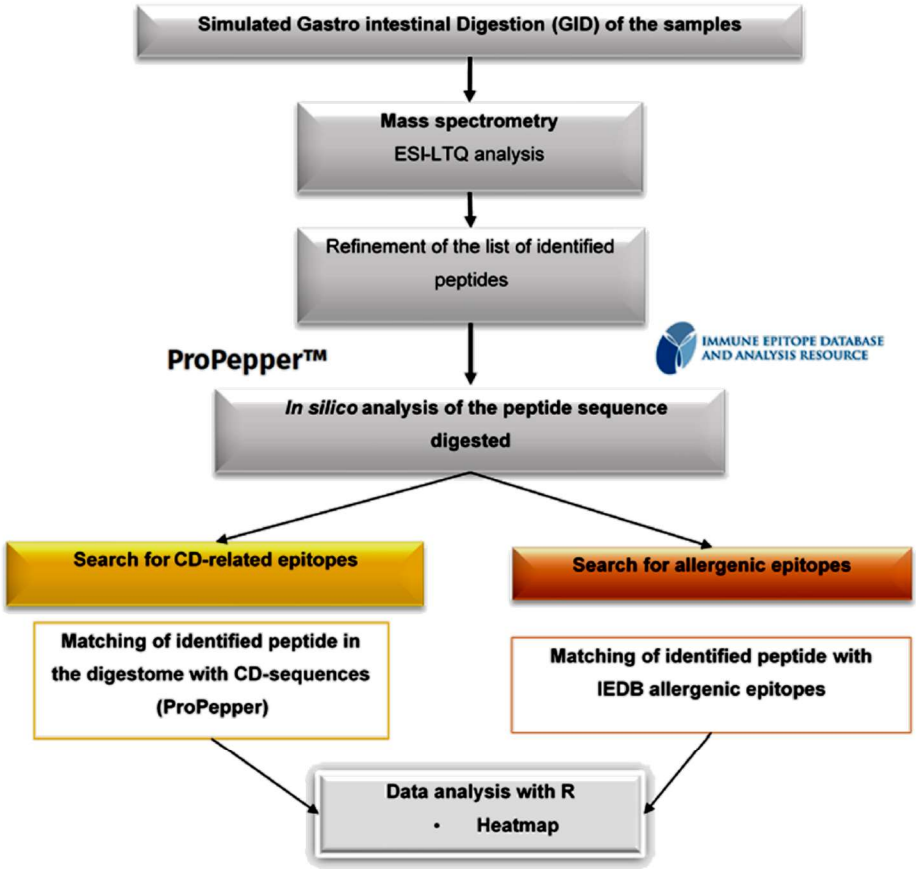


Fig. 1. Flow chart of the in silico analysis after in vitro digestion and Mass spectrometry analysis (ESI-LTQ). The identification of the CD and allergen epitopes obtained thanks to the already indexed ones. *Peptides were defined unique within the given database. Uniqueness was assessed using Skyline software and matching the list of identified peptides with the reference database as background in the "import peptide list" window.

4564. RStudio® (64-Bit-R-3.6.3) was used to create a heatmap to represent the density of the verified epitopes per *cv*/N rate. The use of grayscale heatmaps permitted the distribution of the analyzed samples of the CD and IgE-binding epitopes to be visualized.

The bioactive peptides were retrieved from a literature review (Babini et al., 2017; Liu & Udenigwe, 2019; Suetsuna & Chen, 2002).

2.12. Statistical analysis

The agronomic parameters (grain yield, TW, TKW), the GPC, the glu/glia and HMW/LMW-GS ratios and the percentage of single gliadin and glutenin fractions were compared by means of an analysis of variance (ANOVA), in which the combination of genotypes and N fertilization were the independent variables. Multiple comparison tests were performed, according to the Ryan-Einot-Gabriel-Welsh F (REGW-F) test, on the treatment means. Statistical data analysis was carried out with the SPSS software package, version 26.0.

3. Result and discussion

3.1. Agronomical and productive parameters

Due to the good soil fertility of the site where the experiment was carried out, the N rate did not result in any significant increase in grain yield, TW or TKW for any of the compared *cvs* (Table 1). Clear differences were observed in the productive parameters and grain traits of the compared *cvs*. The grain yield of modern *cv* was 57% and 33% higher than that of landrace and tritordeum, respectively. Modern wheat showed the highest test weight value and the lowest TKW, the opposite of the landrace *cv*, which showed the highest TKW and lower values of TW, as a consequence of strong lodging under both N fertilizations.

3.2. Protein content

Although in the considered productive conditions, N fertilization did not have a significant impact on the grain yield, the GPC was clearly influenced by the combination of *cv* and the N application (Table 1).

On average, the highest GPC was recorded for landrace (17.7%) and tritordeum (16.6%), followed by the modern wheat *cv* (14.9%). A significant increase in the protein content was recorded in the landrace (+1.1%) and tritordeum (+1.5%) *cv* kernels for with a double N rate. The GPC increased in N160 for the modern wheat, albeit not significantly (+0.7%). This increase confirms previous results in literature (Garrido-Lestache et al., 2004; Godfrey et al., 2010; Johansson et al., 2001). To better understand the nutritional value of tritordeum, it will be interesting in future works to verify the increase in essential amino acids related to the N fertilization (Zhang et al., 2016).

Table 1
Effect of the cultivar and N rate on the grain yield, test weight (TW), thousand kernel weight (TKW) and grain protein content (GPC).

Cultivar	N rate	Grain yield	TW	TKW	GPC
	(kg N/ha)	(t ha^{-1})	(kg hl^{-1})	(g)	(%)
landrace	N80	2.9 c	73.8 b	45.0 a	17.1 b
	N160	3.0 c	74.5 b	45.6 a	18.2 a
modern	N80	6.8 a	77.3 a	33.0 c	14.5 d
	N160	7.0 a	76.2 a	31.9 c	15.2 cd
tritordeum	N80	4.8 b	70.3 c	36.8 b	15.8 c
	N160	4.2 b	69.4 c	36.1 b	17.3 b
p-value		< 0.001	< 0.001	< 0.001	< 0.001

Means followed by different letters are significantly different (the level of significance, *p*-value, is reported in the table), according to the REGW-F test.

3.3. Proteomic analysis of the whole grain protein and gluten fractions

3.3.1. DE of the whole grain proteins

The 2-DE separation of crude extracted proteins for all the *cvs*/N rates allowed maps to be generated to show the typical migration patterns of the main cereal protein families (**Fig. S1**). Distinctive spots were observed for each subfraction with corresponding molecular weight (MW) ranges of 80–120 kDa for HMW-GS, 60–68 kDa for the ω5-gliadins, 43–60 kDa for the ω1,2-gliadins and 32–45 kDa for the α- and γ–gliadins and LMW-GS (Mamone et al., 2005).

The analysis revealed differences in the protein expression across the considered *cvs* but did not show macroscopic qualitative differences that could be attributed to the N rates. Tritordeum flour appeared to have the proteome with the largest qualitative differences in the presence of protein spots in the acid region around 31 kDa, corresponding to the LMW glutenins and the gliadin region (α, γ). These differences were also clear in the SDS-PAGE profile of the isolated gliadin fraction (**Fig. S2**) and can be attributed to the nature of the tritordeum *cv*, which is an amphiploid of *T. turgidum* and *H. chilense* (Alvarez et al., 1995, 1999; Sillero et al., 1999).

3.3.2. SDS-PAGE and analysis of the osborn fractions

The gliadin and glutenin Osborne extracts were separated by means of SDS-PAGE. The glutenins (**Fig. S2_B**) showed the typical separation of the HMW-GS (80–140 kDa) and LMW-GS (10–70 kDa) families. The glutenin expression in the analyzed *cvs* clearly differed. The gel analysis highlighted 8 bands in the landrace samples, while only 6 bands were observed in the modern wheat and tritordeum. Unlike the modern and landrace samples, tritordeum presented a less complex profile in the HMW glutenin region. This observation is in line with the calculation of HMW/LMW ratios retrieved from the HPLC profile (Table 2). The tritordeum gliadins showed a more complex profile, which results in a higher number of bands than the modern and landrace *cvs* (**Fig. S2_A**). Interestingly no qualitative diversification, as induced by the N effect, was observed among the samples.

3.3.3. HPLC analysis

The RP-HPLC gradient separation of the gliadin and glutenin fractions was optimized on a RP C8 column to achieve high resolution profiles. The chromatographic profile of the gliadins (**Fig. S3**) showed the typical peaks attributed to ω5, ω1,2, α- and γ-sub fractions, and allowed the relative amounts of the expressed gliadin subfractions to be determined (Table 2) (Mamone et al., 2000). In all the genotypes, the α- and γ-fractions had similar values and were more abundant than the ω-gliadins. Both *cv* and N rate factors did not influence the relative amount of α- and γ-gliadins fractions. Also Vaquero et al. (2018) reported a stable behaviour of these gliadin fractions across different tritordeum and wheat genotypes. Conversely, the ω5 fraction was significantly higher in the wheat landrace compared to the tritordeum, while the modern cv had an intermediate content. This evidence is interesting, considering the potential IgE-binding capacity associated with this gliadin fraction (Matsuo et al., 2004).

The same approach was used to analyze the glutenin isolated fraction, and the determined HMW/LMW-GS ratios for each cv/N rate are reported in Table 2. Unlike the gliadin fraction, the glutenin fraction in tritordeum differed significantly from the other two wheat *cvs* for both glutenin subunits, leading to a HMW/LMW-GS ratio significantly lower than both the landrace and modern wheat *cvs*.

As far as the gluten fraction composition, the N rate significantly influenced only the glu/glia ratio in the modern wheat *cv*. Otherwise, N fertilization did not affect the prolamin ratio in landrace and tritordeum *cvs*, confirming results of Johansson et al. (2001). For these cvs, it is supposed that the higher N rate could have increase to a similar extent all the single individual storage protein components (Zhen et al., 2020), and they have equally contributed to the rise of GPC (Martre et al., 2003).

Table 2
Effect of the cultivar and N rate on the glutenin/gliadin (glu/glia) and HMW/LMW-GS ratios and the percentage of the single fractions of gliadin and glutenin, as obtained from the HPLC-MS analysis.

Cultivars	N rate	glu/glia	HMW/LMW-GS	Gliadins			Glutenins		
	(kg N/ha)			ω5	ω1,2	α	γ	HMW-GS	LMW-GS
landrace	N80	0.84 b	0.70 a	2.8 a	5.8 a	43.9 a	47.5 a	41.2 a	58.8 b
	N160	0.88 b	0.83 a	2.7 a	4.2 a	45.4 a	47.7 a	45.5 a	54.5 b
modern	N80	0.89 b	0.79 a	1.4 bc	5.8 a	44.7 a	48.1 a	44.1 a	55.9 b
	N160	1.16 a	0.82 a	2.1 ab	5.5 a	42.3 a	50.1 a	45.0 a	55.0 b
tritordeum	N80	0.96 b	0.29 b	1.1 c	4.3 a	53.4 a	41.2 a	22.6 b	77.4 a
	N160	0.99 b	0.30 b	0.9 c	5.5 a	52.3 a	41.3 a	22.9 b	77.1 a
p-value		0.001	<0.001	0.001	0.251	0.119	0.063	<0.001	0.002

Means followed by different letters are significantly different (the level of significance, *p*-value, is reported in the table), according to the REGW-F test.

3.4. Residual immunoreactivity of the digested flours by means of ELISA

The immunoreactivity of the peptides generated during the gastro-duodenal digestion of flours was tested using a competitive enzyme-linked immunosorbent assay (Di Stasio et al., 2020). The assay uses a monoclonal antibody to target the "QQPFP" repeated sequence found in prolamins.

Digested tritordeum flour showed a lower immunoreactivity than that of the landrace (-50%) and modern (-58%) *cvs*. No significant differences were identified for the samples with different N rates for any of the *cvs*. The lower expression of the above-mentioned celiac-toxic epitopes in tritordeum, which could be attributed to the absence of genome D, is shown in Fig. 2.

This evidence prompted us to conduct a more detailed structural analysis of the protein gastrointestinal digests of the flour by means of mass spectrometry.

3.5. Mass spectrometry profiling of the gastrointestinal digested flours

In CD, the peptides derived from the digestion of food transmigrate through the intestinal epithelium. They are then deamidated in the subepithelial lamina propria by the tissue transglutaminase and presented to the T-cells. An already validated static and multi-phasic in vitro digestion model (Brodkorb et al., 2019), comprising oral, gastric and duodenal sequentially simulated phases (OGD), was applied to collect sequence level information on the peptides resistant to OGD proteolysis. The mass spectrometry analysis allowed several flour derived peptides to be identified which indicated a resistance to digestive enzymes, as already described in literature (Mamone et al., 2015). The natural presence of protease and amylase inhibitors in the grain flours may have contributed to the reduced digestibility. However, being pepsin the first proteolytic enzyme in the stomach environment where the pH is below 3.0, their inactivation is possible. The analysis of metabolic proteins has not been objective of the study and may require a dedicated investigation.

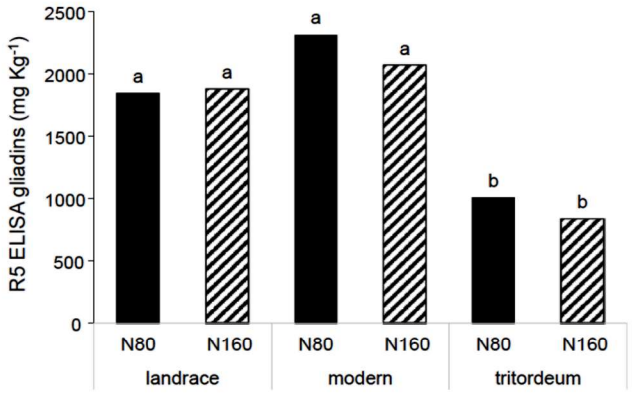


Fig. 2. Quantification of the "QQPFP" celiac toxic motif recognized by the R5 monoclonal antibody in the soft wheat (landrace and modern) and tritordeum cultivars for the N80 (80 kg N ha^{-1}) and N160 (160 kg N ha^{-1}) fertilization rates. Bars with different letters are significantly different (*p*-value < 0.05), according to the REGW-F test.

The main peptides that were identified belonged to the digestive enzymes that were undergoing a natural and expected autolysis (Mamone et al., 2015). The grain derived identified peptides are listed in tables S1-6. The peptides were inferred to a single protein or to protein families when no unique peptide could be identified. We identified CD peptide sequences, derived from the γ-hordein and B1-hordein fractions, in the digestome of the tritordeum flour. Although the *p*-value of the identified peptides, which reflects the probability of finding a random peptide, was acceptable (*p*-value < 0.05), the high number of repeated regions and the high homology sequences, which are typical of wheat gluten proteins, led to high E-values, a parameter that describes the number of hits one can "expect" to see by chance when searching a database of a particular size. The E-value is obtained by multiplying the *p*-value by the total number of qualified peptides in the searched database, the masses of which fall into the precursor ion mass range, plus/minus the specified tolerance (Alves & Yu, 2015). Since the LTQ mass spectrometer works at a low resolution, a tolerance of 1 Da was applied, and this caused the E value to be large. The peptides were validated manually, to confirm the software identifications, and exemplar spectra are provided in the supplementary material section (**Fig. S4**).

3.6. In silico evaluation of the in vitro gastrointestinal identified peptides

The digestion-derived peptides were *in silico* assessed to evaluate sequences known to potentially trigger CD, and allergic reactions. We also evaluated positive sequences with potential bioactivity.

3.7. Celiac toxic motifs

The toxic motifs of CD predominantly arise from α-gliadins (Ozuna et al., 2015), which were found to be the most abundant gliadin fraction in our experiment (about 60%), followed by γ-type gliadins (Fig. 3).

The same number of epitopes belonging to the α-gliadins was observed for modern and tritordeum, while landrace showed 15% more. Landrace also showed the highest numbers of identified epitopes for the γ-gliadins, followed by tritordeum (-18%) and then modern (-27%). Tritordeum showed 50% fewer CD epitopes belonging to the ω-gliadins and 23% and 44% lower epitope numbers of the LMW-GS class than the landrace and modern *cvs*.

The greatest number of CD epitopes was found in the α-gliadin class in all three varieties. On the basis of our observations, it can be seen that the most frequently identified toxic motifs (*i.e.* IPEQ, PQQLPQ, QPQQPF, QPQPFP and VRVPVPQL) are similar to those identified in previous studies and they were identified in all the studied samples (**Fig. S5**) (Osman et al., 2000). For example, the presence of the CD epitope IPEQ was observed in all the digested flours, and it has already been described in the consensus sequences recognized by IgG human

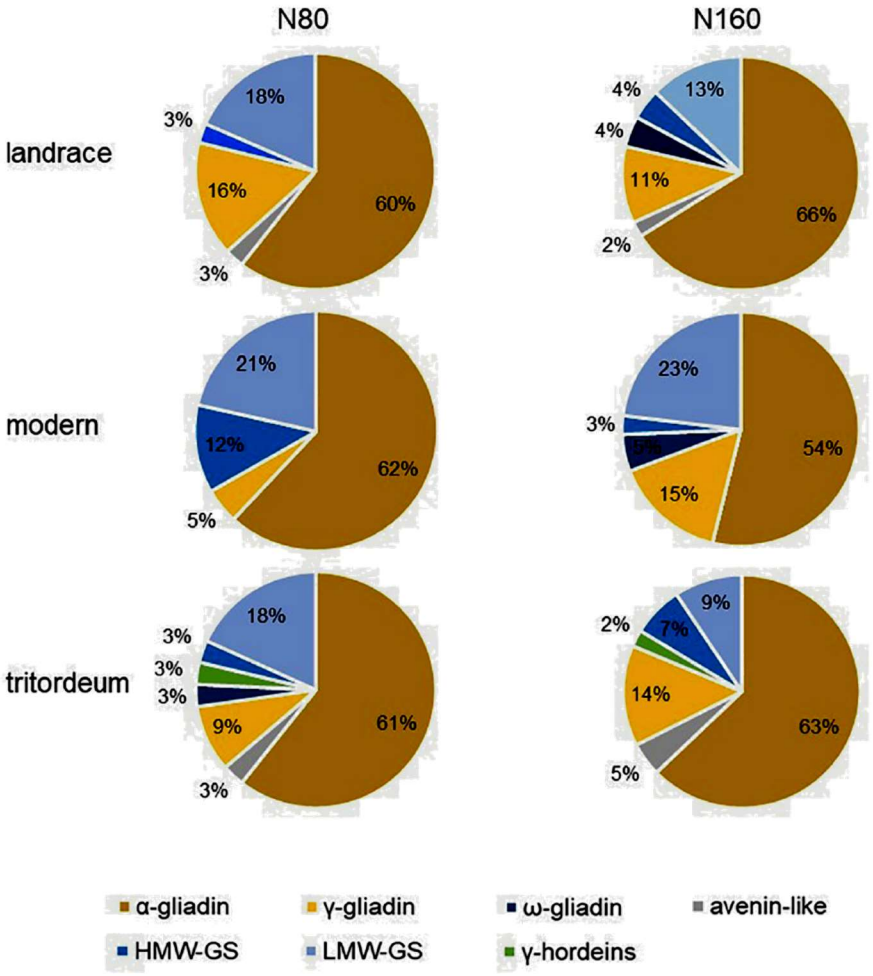


Fig. 3. Gluten sources of the celiac epitopes in the soft wheat (landrace and modern) and tritordeum cultivars for the N80 (80 kg N ha^{-1}) and N160 (160 kg N ha^{-1}) fertilization rates.

antibodies in a Pepscan experiment (Osman et al., 2000). Similarly, QQQPFP and PQQLPQ, which were also identified in all the flours, were recognized by human IgG and IgA antibodies. The QGSFQP sequence was only identified in landrace and modern at a high N rate and in both the tritordeum samples (Osman et al., 2000). Truncated versions of the 33-mer (Glia-α 57–89, LQLQPFPQPQLPYPQPQLPYPQPQLPYPQPQPF) and the 25-mer (Glia-α 31–55, LGQQQPFPPQQPYPQPQPFPSQQPY) gliadin derived peptides, which have been described as being resistent to digestion, and are known to be strong stimulators of the T-cell response, were identified in all the analyzed samples (**Table S7**) (Ozuna et al., 2015) . The full length 33-mer peptide contains three repeated regions (p62–67, PQPQLPY) and is exclusively present in a D-genome encoded α-gliadin (Camarca et al., 2009). The D-genome in tritordeum is replaced by the Hch genome of *Hordeum chilense*. A single fragment of 33-mer (LQPFPQPQLPYPQPH) was unexpectedly identified in the tritordeum digested flour. Although the breeding program developed crosses between soft wheat and tritordeum in order to improve its yield and bread-making quality (Ávila et al., 2021), and a small introgressions from D genome into tritordeum cannot be ruled out, our experiment was carried out using a tritordeum *cv* carefully selected for the absence of any soft wheat chromosomes. However, the *Hordeum chilense* genome has not yet been sequenced, and the natural presence of the peptide in the tritordeum proteome cannot therefore be excluded. Alternatively, a cross-contamination might also have occurred. A targeted approach would be required to confirm the presence of partial 33-*mer*-like sequences in tritordeum.

3.8. *Allergenic epitopes*

Wheat-dependent, exercise-induced anaphylaxis (WDEIA) is a serious allergy in which the combination of wheat ingestion and physical exercise leads to anaphylaxis. Patients with WDEIA have IgE antibodies against ω5-gliadin, one of the major gluten allergens (Lehto et al., 2003). Together with the evaluation of the CD sequences, the presence of IgE-binding sequences was also evaluated (Vita et al., 2019). The IEDB epitopic sequence was manually searched in the pool of digested peptides. An epitope was considered as present when at least one precursor peptide was identified in the digestome. In order to compare the samples, we only took into consideration the presence of the epitopic sequence, and not the number of times the sequence was identified. A graphic example of the search for GID sequences with a known allergen IEDB sequence is shown in supplementary **Table S8**.

The digestive derived peptides were only considered for a comparative evaluation when 100% sequence homology was shared with the described IgE binding epitope. The relative abundance of the identified epitopes (**Fig. S6**) was used to qualitatively compare the different cultivars and N rates. The largest number of allergenic epitopes belonged to the γ-gliadin fraction (about 40%), followed by epitopes of HMW-GS and α-gliadins prolamins for all the digested samples (Fig. 4). The modern *cv* was distinguished by a greater percentage (30%) of α-gliadin allergenic

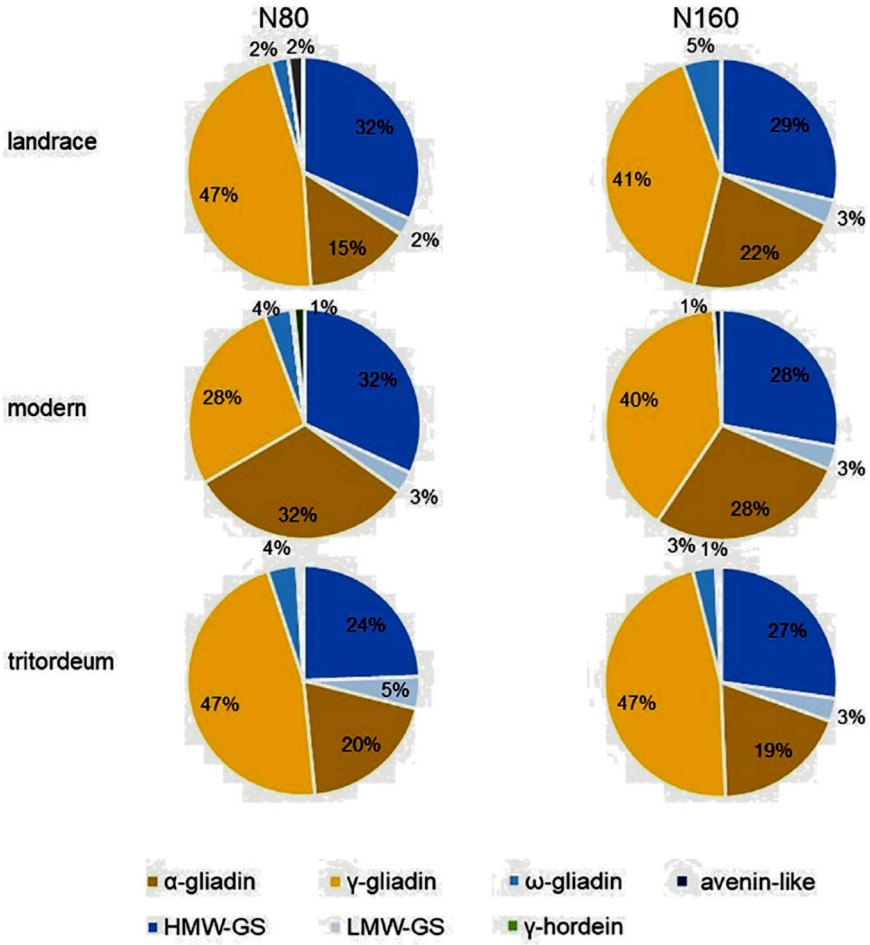


Fig. 4. Gluten sources of the allergenic epitopes in the soft wheat (landrace and modern) and tritordeum cultivars for the N80 (80 kg N ha^{-1}) and N160 (160 kg N ha^{-1}) fertilization rates.

epitopes, but showed a lower percentage for HMW-GS, LMW-GS and the ω-gliadins than landrace and tritordeum.

A limited number of the identified epitope sequences which belong to the 5 main gluten allergens, according to the World Health Organization and the International Union of Immunological Societies (WHO/IUIS), were identified in all the *cvs* and at both N rates. This allergenic wheat gluten proteins list includes (http://www.allergen.org): Tri a 19 (ω5), Tri a 20 (γ), Tri a 21 (α), Tri a 26 (HMW-GS) and

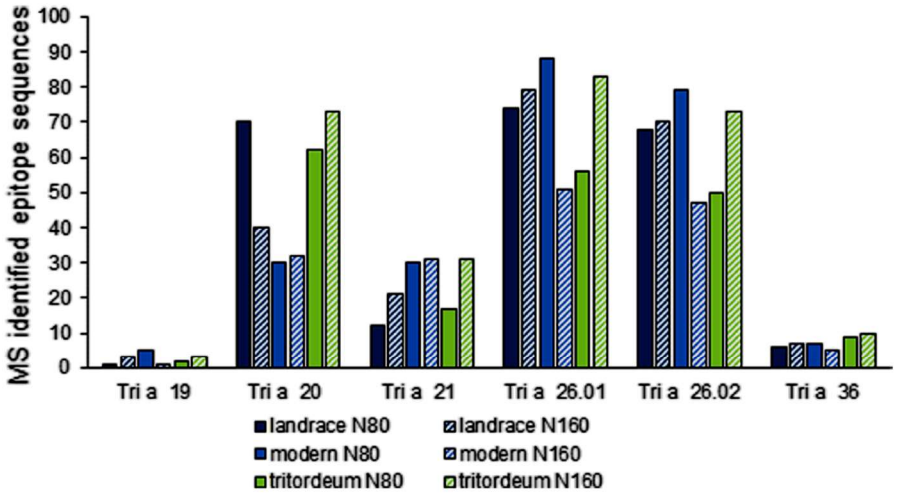


Fig. 5. A graphical qualitative rapresentation of the MS identified epitope sequences inferred to the five main 5 main wheat allergens (tri a 19, tri a 20, tri a 21; tri a 26.01; tri a 26.02; tri a 36) in the soft wheat (landrace and modern) and tritordeum cultivars for the N80 (80 kg N ha^{-1}) and N160 (160 kg N ha^{-1}) fertilization rates.

Tri a 36 (LMW-GS) (Fig. 5). Tri a 26 (HMW-GS) was the protein with the highest number of identified epitope sequences in all the *cv* and N rate combinations. HMW-GS presents disulfide-bridge linked aggregates, which induce a high resistance to in vitro enzymatic digestion (Anderson et al., 1984).

The second allergen with the highest number of epitope sequences identified was Tri a 20, belonging to the γ-gliadins, which has been identified as a strong trigger of WDEIA symptoms after ingestion of wheat protein (Yokooji et al. 2013). For the Tri a 21, the α-gliadin allergen, only a few peptide sequences were identified in the digested samples. This protein is responsible for the bakers' allergy (Sander et al., 2015).

Epitopes with a lower frequency matched with Tri a 19 and Tri a 36, show an occurrence that may be ascribed to a high susceptibility to enzymatic digestion. Although, Tri a 36 allergen (LMW-GS) is known to retain IgE reactivity, since it survives the extensive oral, gastric and duodenal in vitro digestion (Baar et al., 2012), a very small number of allergenic epitopes was found in all the considered cvs and N rates, thus suggesting that the amount of Tri a 36 that forms and accumulates in mature seeds was probably quite low for each cv at the time of the collection (Baar et al., 2012). The occurrence of epitope sequences, determined by means of in silico analysis, should be confirmed by means of an appropriate immunological test using the sera of wheat allergic subjects.

3.9. Bioactive peptides

Together with the negative effects on human health, we evaluated any potential positive effects derived from the ingestion of the considered wheat and tritordeum *cvs*. Numerous digestion derived peptides are known to be precursors of peptides with described antioxidant and opioid effects. The bioactive peptides found in all the digested samples are shown in Fig. 6. The "GYYPT" and "YPQPQPF" opioid sequences were identified in all the samples. Gliadorphin 7 (also known as gluteomorphin), a δ-exorphin with a YPQPQPF sequence, which is formed during the digestion of α-gliadin, deserves particular attention because its presence could be related to neurodevelopmental disease and psychotic disorders (Liu and Udenigwe, 2019).

"GYYP" was identified in digested flour from the modern *cv*. For both N rates, and in landrace for the high N application. Peptides containing the "PYPQ" antioxidant sequence were identified in all the samples, while "LQPGQGQQG" was only identified in the modern *cv*.

4. Conclusions

We have investigated the effect of N fertilization on the gluten profile of three genotypes with different gluten compositions. The study has shown that increasing the N rate influenced the GPC to a great extent, while it was found to have less impact on the gluten composition and on the type of celiac and allergenic epitopes after in vitro digestion. On the other hand, these parameters resulted in a greater variation only according to the genotype.

A strong genetic effect was observed with regard to the composition of gluten, especially for glutenins. Tritordeum showed the lowest levels of HMW-GS/LMW-GS and was also distinguished by lower values of ω5, compared to wheat *cvs*. Competitive R5 ELISA analysis highlighted that tritordeum may be regarded as a variety with a lower presence of highly CD epitopes than the old and modern wheat varieties that were here considered. The absence of the D genome in tritordeum could be one of the reasons for the the lower immunodominant toxicity of this amphiploid.

The study on the expression of celiacogenic and allergenic sequences in different genomes remains of fundamental importance to provide scientific information. This proteomic study paves the way toward a more inclusive study with a larger number of genotypes, grown under the same environmental and agronomic conditions, to better understand their potential immunotoxicity. The qualitative analysis of the epitopes made it possible to observe that in all the studies *cv* and agronomic conditions (e.g. fertilization rate) the largest number of sequences identified belonging to the class of α- and γ-gliadins, respectively for celiac and allergenic epitopes. The CD epitopes showed a greater presence in the α-gliadin fraction, where a 33-mer is present as strong stimulators of the T-cell response resistent to digestion. All the compared *cv* and N rate combinations showed quite concordantly a high presence of epitopes belonging to the allergens Tri a 26 and Tri a 21, of HMW-GS and γ-gliadins respectively, and very low for Tri a 19 and Tri a 36, belonging to ω5 and LMW-gs respectively. These results showed an interesting in-depth proteomic profile of the conditions studied. To verify the quantitative effect of a N rate and its interaction with the genotype, strictly at the epitope level, it would require to build a method to quantify robust information considering the heterogeneity of the genome and the presence of repeated regions in the starting matrix. This will be a future goal of great interest.

The protein profile of tritordeum, compared to the two soft wheat cvs, supports its application as an interesting diet ingredient, given its lower R5 immunogenicity and high GPC content. However, the use of

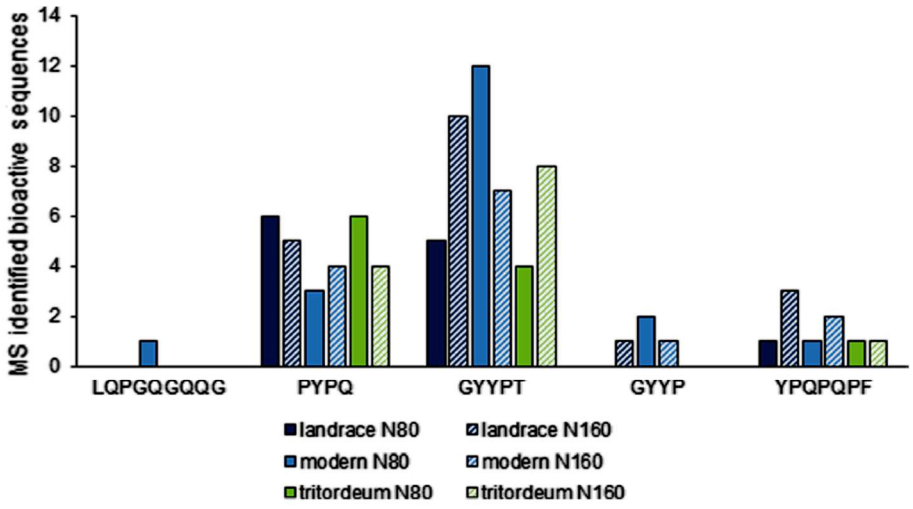


Fig. 6. A graphical qualitative rapresentation of the MS identified epitope sequences inferred to a selection of bioactive peptide in the soft wheat (landrace and modern) and tritordeum cultivars for the N80 (80 kg N ha^{-1}) and N160 (160 kg N ha^{-1}) fertilization rates.

this new amphiploid species, as a promising alternative to soft wheat, will require specific breeding programs to enhance the end-use quality and immunogenic, nutritional and agronomic traits of such a crop.

CRediT authorship contribution statement

Landolfi Viola: Data curation, Formal analysis, Investigation, Methodology, Writing – original draft. **D'Auria Giovanni:** Data curation, Formal analysis, Investigation, Methodology. **Nicolai Maria Adalgisa:** Formal analysis, Methodology. **Chiara Nitride:** Writing – review & editing. **Blandino Massimo:** Conceptualization, Funding acquisition, Project administration, Supervision. **Ferranti Pasquale:** Conceptualization, Funding acquisition, Project administration, Supervision.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was supported by the Regione Piemonte (POR FESR 2014-2020), as a part of the DEUM Project.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.foodres.2021.110617.

References

Alvarez, J. B., Campos, L. A. C., Martín, A., Sillero, J. A., & Martín, L. M. (1999). Genetic analysis of prolamins synthesised by the Hch genome and their effects on gluten strength in hexaploid tritordeum. *Euphytica, 107*(3), 177–184. https://doi.org/10.1023/A:1003574213125.

Alvarez, J. B., Martin, L. M., Rogers, W. J., & Martin, A. (1995). Relationship between the endosperm storage proteins composition and the breadmaking quality in hexaploid tritordeum. *Cereal Research Communications, 23*(1/2), 167–172.

Alves, G., & Yu, Y. K. (2015). Mass spectrometry-based protein identification with accurate statistical significance assignment. *Bioinformatics, 31*(5), 699–706. https://doi.org/10.1093/bioinformatics/btu717.

Anderson, O. D., Litts, J. C., Gautier, M.-F., & Greene, F. C. (1984). Nucleic acid sequence and chromosome assignment of a wheat storage protein gene. *Nucleic Acids Research, 12*(21), 8129–8144. https://doi.org/10.1093/nar/12.21.8129.

Ávila, C. M., Rodríguez-Suárez, C., & Atienza, S. G. (2021). Tritordeum: Creating a New Crop Species—The Successful Use of Plant Genetic Resources. *Plants, 10*(5), 1029. https://doi.org/10.3390/plants10051029.

Baar, A., Pahr, S., Constantin, C., Scheiblhofer, S., Thalhamer, J., Giavi, S., … Valenta, R. (2012). Molecular and immunological characterization of Tri a 36, a low molecular weight glutenin, as a novel major wheat food allergen. *Journal of Immunology, 189*(6), 3018–3025. https://doi.org/10.4049/jimmunol.1200438.

Babini, E., Tagliazucchi, D., Martini, S., Dei Più, L., & Gianotti, A. (2017). LC-ESI-QTOF-MS identification of novel antioxidant peptides obtained by enzymatic and microbial hydrolysis of vegetable proteins. *Food Chemistry, 228*, 186–196. https://doi.org/10.1016/j.foodchem.2017.01.143.

Blandino, M., Vaccino, P., & Reyneri, A. (2015). *Late-season N increases improver common and durum wheat quality., 107*(2), 680–690. https://doi.org/10.2134/agronj14.0405.

Brodkorb, A., Egger, L., & Alminger, M. (2019). INFOGEST static in vitro simulation of gastrointestinal food digestion. *Nature Protocols, 14*(4), 991–1014. https://doi.org/10.1038/s41596-018-0119-1T.

Camarca, A., Anderson, R. P., Mamone, G., Fierro, O., Facchiano, A., Costantini, S., … Gianfrani, C. (2009). Intestinal T Cell Responses to Gluten Peptides Are Largely Heterogeneous: Implications for a Peptide-Based Therapy in Celiac Disease. *The Journal of Immunology, 182*(7), 4158–4166. https://doi.org/10.4049/jimmunol.0803181.

Casella, G., Villanacci, V., Di Bella, C., Bassotti, G., Bold, J., & Rostami, K. (2018). Non celiac gluten sensitivity and diagnostic challenges. *Gastroenterology and Hepatology From Bed to Bench, 11*(3), 197–202.

Cho, K., Beom, H.-R., Jang, Y.-R., Altenbach, S. B., Vensel, W. H., Simon-Buss, A., … Lee, J.-Y. (2018). Proteomic Profiling and Epitope Analysis of the Complex α-, γ-, and ω-Gliadin Families in a Commercial Bread Wheat. *Frontiers in Plant Science, 9*. https://doi.org/10.3389/fpls.2018.0081810.3389/fpls.2018.00818.s00110.3389/fpls.2018.00818.s002.

Daly, M., Bromilow, S. N., Nitride, C., Shewry, P. R., Gethings, L. A., & Mills, E. N. C. (2020). Mapping Coeliac Toxic Motifs in the Prolamin Seed Storage Proteins of Barley, Rye, and Oats Using a Curated Sequence Database. *Frontiers in Nutrition, 7*.

Dhaka, V., & Khatkar, B. S. (2015). Effects of Gliadin/Glutenin and HMW-GS/LMW-GS Ratio on Dough Rheological Properties and Bread-Making Potential of Wheat Varieties: Gluten Proteins, Dough Rheology and Bread Quality. *Journal of Food Quality, 38*(2), 71–82. https://doi.org/10.1111/jfq.12122.

Di Stasio, L., Picascia, S., Auricchio, R., Vitale, S., Gazza, L., Picariello, G., … Mamone, G. (2020). Comparative Analysis of in vitro Digestibility and Immunogenicity of Gliadin Proteins From Durum and Einkorn Wheat. *Frontiers in Nutrition, 7*. https://doi.org/10.3389/fnut.2020.0005610.3389/fnut.2020.00056.s00110.3389/fnut.2020.00056.s002.

Dunne, M. R., Byrne, G., Chirdo, F. G., & Feighery, C. (2020). Coeliac Disease Pathogenesis: The Uncertainties of a Well-Known Immune Mediated Disorder. *Frontiers in Immunology, 11*, 1374. https://doi.org/10.3389/fimmu.2020.01374.

Dupont, F. M., Vensel, W. H., Tanaka, C. K., Hurkman, W. J., & Altenbach, S. B. (2011). Deciphering the complexities of the wheat flour proteome using quantitative two-dimensional electrophoresis, three proteases and tandem mass spectrometry. *Proteome Science, 9*(1), 10. https://doi.org/10.1186/1477-5956-9-10.

Fiedler, K. L., McGrath, S. C., Callahan, J. H., & Ross, M. M. (2014). Characterization of Grain-Specific Peptide Markers for the Detection of Gluten by Mass Spectrometry. *Journal of Agricultural and Food Chemistry, 62*(25), 5835–5844. https://doi.org/10.1021/jf500997j.

Foca, G., Ulrici, A., Corbellini, M., Pagani, M. A., Lucisano, M., Franchini, G. C., & Tassi, L. (2007). Reproducibility of the Italian ISQ method for quality classification of bread wheats: An evaluation by expert assessors. *Journal of the Science of Food and Agriculture, 87*(5), 839–846. https://doi.org/10.1002/(ISSN)1097-001010.1002/jsfa.v87:510.1002/jsfa.2785.

Garrido-Lestache, E., López-Bellido, R. J., & López-Bellido, L. (2004). Effect of N rate, timing and splitting and N type on bread-making quality in hard red spring wheat under rainfed Mediterranean conditions. *Field Crops Research, 85*(2), 213–236. https://doi.org/10.1016/S0378-4290(03)00167-9.

Godfrey, D., Hawkesford, M. J., Powers, S. J., Millar, S., & Shewry, P. R. (2010). Effects of Crop Nutrition on Wheat Grain Composition and End Use Quality. *Journal of Agricultural and Food Chemistry, 58*(5), 3012–3021. https://doi.org/10.1021/jf9040645.

Johansson, E., Prieto-Linde, M. L., & Jönsson, J.Ö. (2001). Effects of Wheat Cultivar and Nitrogen Application on Storage Protein Composition and Breadmaking Quality. *Cereal Chemistry, 78*(1), 19–25. https://doi.org/10.1094/CCHEM.2001.78.1.19.

Juhász, A., Haraszi, R., & Maulis, C. (2015). ProPepper: A curated database for identification and analysis of peptide and immune-responsive epitope composition of cereal grain protein families. Database: The Journal of Biological Databases and Curation, 2015, 1758-0463. https://doi.org/10.1093/database/bav100.

Kagnoff, M. F. (2007). Celiac disease: Pathogenesis of a model immunogenetic disease. *The Journal of Clinical Investigation, 117*(1), 41–49. https://doi.org/10.1172/JCI30253.

Lebwohl, B., Ludvigsson, J. F., & Green, P. H. R. (2015). Celiac disease and non-celiac gluten sensitivity. *The BMJ, 351*, Article h4347. https://doi.org/10.1136/bmj.h4347.

Lehto, M., Palosuo, K., Varjonen, E., Majuri, M.-L., Andersson, U., Reunala, T., & Alenius, H. (2003). Humoral and cellular responses to gliadin in wheat-dependent, exercise-induced anaphylaxis. *Clinical & Experimental Allergy, 33*(1), 90–95. https://doi.org/10.1046/j.1365-2222.2003.01568.x.

Liu, Z., & Udenigwe, C. C. (2019). Role of food-derived opioid peptides in the central nervous and gastrointestinal systems. *Journal of Food Biochemistry, 43*(1), e12629. https://doi.org/10.1111/jfbc.12629.

Mamone, G., Addeo, F., Chianese, L., Di Luccia, A., De Martino, A., Nappo, A., … Ferranti, P. (2005). Characterization of wheat gliadin proteins by combined two-dimensional gel electrophoresis and tandem mass spectrometry. *Proteomics, 5*(11), 2859–2865. https://doi.org/10.1002/(ISSN)1615-986110.1002/pmic.v5:1110.1002/pmic.200401168.

Mamone, G., Ferranti, P., Chianese, L., Scafuri, L., & Addeo, F. (2000). Qualitative and quantitative analysis of wheat gluten proteins by liquid chromatography and electrospray mass spectrometry. *Rapid Communications in Mass Spectrometry, 14*(10), 897–904. https://doi.org/10.1002/(SICI)1097-0231(20000530)14:10<897::AID-RCM962>3.0.CO;2-Z.

Mamone, G., Nitride, C., Picariello, G., Addeo, F., Ferranti, P., & Mackie, A. (2015). Tracking the Fate of Pasta (*T. durum* Semolina) Immunogenic Proteins by in Vitro Simulated Digestion. *Journal of Agricultural and Food Chemistry, 63*(10), 2660–2667. https://doi.org/10.1021/jf505461x.

Mamone, G., Picariello, G., Addeo, F., & Ferranti, P. (2011). Proteomic analysis in allergy and intolerance to wheat products. *Expert Review of Proteomics, 8*(1), 95–115. https://doi.org/10.1586/epr.10.98.

Martín, A., Alvarez, J. B., Martín, L. M., Barro, F., & Ballesteros, J. (1999). The Development of Tritordeum: A Novel Cereal for Food Processing. *Journal of Cereal Science, 30*(2), 85–95. https://doi.org/10.1006/jcrs.1998.0235.

Martre, P., Porter, J. R., Jamieson, P. D., & Triboï, E. (2003). Modeling Grain Nitrogen Accumulation and Protein Composition to Understand the Sink/Source Regulations of Nitrogen Remobilization for Wheat. *Plant Physiology, 133*(4), 1959–1967.

Matsuo, H., Morita, E., Tatham, A. S., Morimoto, K., Horikawa, T., Osuna, H., … Dekio, S. (2004). Identification of the IgE-binding Epitope in ω-5 Gliadin, a Major Allergen in Wheat-dependent Exercise-induced Anaphylaxis. *Journal of Biological Chemistry, 279*(13), 12135–12140. https://doi.org/10.1074/jbc.M311340200.

Migliorini, P., Spagnolo, S., Torri, L., Arnoulet, M., Lazzerini, G., & Ceccarelli, S. (2016). Agronomic and quality characteristics of old, modern and mixtures wheat varieties and landraces for organic bread chain in diverse environments of northern Italy.

European Journal of Agronomy, 79, 131–141. https://doi.org/10.1016/j.eja.2016.05.011.

Osman, A. A., Günnel, T., Dietl, A., Uhlig, H. H., Amin, M., Fleckenstein, B., … Mothes, T. (2000). B cell epitopes of gliadin. *Clinical and Experimental Immunology, 121*(2), 248–254. https://doi.org/10.1046/j.1365-2249.2000.01312.x.

Ozuna, C. V., Iehisa, J. C. M., Giménez, M. J., Alvarez, J. B., Sousa, C., & Barro, F. (2015). Diversification of the celiac disease α-gliadin complex in wheat: A 33-mer peptide with six overlapping epitopes, evolved following polyploidization. *The Plant Journal: For Cell and Molecular Biology, 82*(5), 794–805. https://doi.org/10.1111/tpj.2015.82.issue-510.1111/tpj.12851.

Plessis, A., Ravel, C., Bordes, J., Balfourier, F., & Martre, P. (2013). Association study of wheat grain protein composition reveals that gliadin and glutenin composition are trans-regulated by different chromosome regions. *Journal of Experimental Botany, 64*(12), 3627–3644. https://doi.org/10.1093/jxb/ert188.

Ribeiro, M., Rodriguez-Quijano, M., Nunes, F. M., Carrillo, J. M., Branlard, G., & Igrejas, G. (2016). New insights into wheat toxicity: Breeding did not seem to contribute to a prevalence of potential celiac disease's immunostimulatory epitopes. *Food Chemistry, 213*, 8–18. https://doi.org/10.1016/j.foodchem.2016.06.043.

Sánchez-León, S., Haro, C., Villatoro, M., Vaquero, L., Comino, I., González-Amigo, A. B., … Barro, F. (2021). Tritordeum breads are well tolerated with preference over gluten-free breads in non-celiac wheat-sensitive patients and its consumption induce changes in gut bacteria. *Journal of the Science of Food and Agriculture, 101*(8), 3508–3517. https://doi.org/10.1002/jsfa.10982.

Sander, I., Rihs, H.-P., Doekes, G., Quirce, S., Krop, E., Rozynek, P., … Raulf, M. (2015). Component-resolved diagnosis of baker's allergy based on specific IgE to recombinant wheat flour proteins. *The Journal of Allergy and Clinical Immunology, 135*(6), 1529–1537. https://doi.org/10.1016/j.jaci.2014.11.021.

Sillero, J. C., Alvarez, J. B., & Martín, L. M. (1999). Analysis of the endosperm proteins synthesised by the Hch genome in tritordeum. *Cereal Research Communications, 27*(4), 455–461. https://doi.org/10.1007/BF03543563.

Suetsuna, K., & Chen, J.-R. (2002). Isolation and Characterization of Peptides with Antioxidant Activity Derived from Wheat Gluten. *Food Science and Technology Research, 8*(3), 227–230. https://doi.org/10.3136/fstr.8.227.

Vaquero, L., Comino, I., Vivas, S., Rodríguez-Martín, L., Giménez, M. J., Pastor, J., … Barro, F. (2018). Tritordeum: A novel cereal for food processing with good acceptability and significant reduction in gluten immunogenic peptides in comparison with wheat. *Journal of the Science of Food and Agriculture, 98*(6), 2201–2209. https://doi.org/10.1002/jsfa.2018.98.issue-610.1002/jsfa.8705.

Vita, R., Mahajan, S., Overton, J. A., Dhanda, S. K., Martini, S., Cantrell, J. R., Wheeler, D. K., Sette, A., & Peters, B. (2019). The Immune Epitope Database (IEDB): 2018 update. Nucleic Acids Research, 47(D1), D339–D343. https://doi.org/10.1093/nar/gky1006.

Wan, Y., Shewry, P. R., & Hawkesford, M. J. (2013). A novel family of γ-gliadin genes are highly regulated by nitrogen supply in developing wheat grain. *Journal of Experimental Botany, 64*(1), 161–168. https://doi.org/10.1093/jxb/ers318.

Yokooji, T., Kurihara, S., Murakami, T., Chinuki, Y., Takahashi, H., Morita, E., … Matsuo, H. (2013). Characterization of causative allergens for wheat-dependent exercise-induced anaphylaxis sensitized with hydrolyzed wheat proteins in facial soap. *Allergology International: Official Journal of the Japanese Society of Allergology, 62*(4), 435–445. https://doi.org/10.2332/allergolint.13-OA-0561.

Zhang, M., Ma, D., Wang, C., Zhao, H., Zhu, Y., & Guo, T. (2016). Responses of Amino Acid Composition to Nitrogen Application in High- and Low-Protein Wheat Cultivars at Two Planting Environments. *Crop Science, 56*(3), 1277–1287. https://doi.org/10.2135/cropsci2015.08.0504.

Zhen, S., Deng, X., Xu, X., Liu, N., Zhu, D., Wang, Z., & Yan, Y. (2020). Effect of high-nitrogen fertilizer on gliadin and glutenin subproteomes during kernel development in wheat (Triticum aestivum L.). *The Crop Journal, 8*(1), 38–52. https://doi.org/10.1016/j.cj.2019.06.002.

Chapter 5

Tritordeum as an innovative alternative to wheat: a comparative digestion study on bread

molecules

MDPI

Article

Tritordeum as an Innovative Alternative to Wheat: A Comparative Digestion Study on Bread

Chiara Nitride [1], Giovanni D'Auria [1,*], Andrea Dente [1], Viola Landolfi [2], Gianluca Picariello [3], Gianfranco Mamone [3], Massimo Blandino [2], Raffaele Romano [1] and Pasquale Ferranti [1,3]

[1] Department of Agricultural Sciences, University of Naples Federico II, Via Università 100, 80055 Portici, Italy; chiara.nitride@unina.it (C.N.); and.dente@studenti.unina.it (A.D.); raffaele.romano@unina.it (R.R.); ferranti@unina.it (P.F.)
[2] Department of Agricultural, Forest and Food Sciences, University of Turin, Largo Braccini 2, 10095 Grugliasco, Italy; viola.landolfi@unito.it (V.L.); massimo.blandino@unito.it (M.B.)
[3] Institute of Food Science, National Research Council, 83100 Avellino, Italy; picariello@isa.cnr.it (G.P.); mamone@isa.cnr.it (G.M.)
* Correspondence: giovanni.dauria@unina.it; Tel.: +39-81253-9346

Abstract: Tritordeum results from the crossbreeding of a wild barley (*Hordeum chilense*) species with durum wheat (*Triticum turgidum* spp. *turgidum*). This hexaploid crop exhibits agronomic and rheological characteristics like soft wheat, resulting in an innovative raw material to produce baked goods. We applied a gel-based proteomic approach on refined flours to evaluate protein expression differences among two widespread tritordeum cultivars (Aucan and Bulel) taking as the reference semolina and flour derived from a durum and a soft wheat cvs, respectively. The products of in vitro digestion of model breads were analyzed to compare bio-accessibility of nutrients and mapping tritordeum bread resistant peptides. Significant differences among the protein profiles of the four flours were highlighted by electrophoresis. The amino acid bio-accessibility and the reducing sugars of tritordeum and wheat breads were comparable. Tritordeum cvs had about 15% higher alpha-amino nitrogen released at the end of the duodenal simulated digestion than soft wheat ($p < 0.05$). Bulel tritordeum flour, bread and digested bread had about 55% less R5-epitopes compared to the soft wheat. Differences in protein expression found between the two tritordeum cvs reflected in diverse digestion products and allergenic and celiacogenic potential of the duodenal peptides. Proteomic studies of a larger number of tritordeum cvs may be successful in selecting those with good agronomical performances and nutritional advantages.

Keywords: tritordeum; in vitro digestion; peptidomics; alpha amino nitrogen; R5; wheat allergy; celiac disease

Citation: Nitride, C.; D'Auria, G.; Dente, A.; Landolfi, V.; Picariello, G.; Mamone, G.; Blandino, M.; Romano, R.; Ferranti, P. Tritordeum as an Innovative Alternative to Wheat: A Comparative Digestion Study on Bread. *Molecules* **2022**, *27*, 1308. https://doi.org/10.3390/molecules27041308

Academic Editor: Sascha Rohn

Received: 7 January 2022
Accepted: 13 February 2022
Published: 15 February 2022

1. Introduction

Wheat grains are the world's most important staple food crop. The derived flour is a key ingredient in the preparation of bakery and pasta products, accounting for 20% of the total dietary calories and proteins in the human diet [1]. Throughout the centuries, the natural selection and hybridization among different wheat varieties, aimed at obtaining species easy to harvest and high in yield, have led to the modern tetraploid durum (*Triticum turgidum* spp. *durum*, AABB) and hexaploid bread wheat (*Triticum aestivum* spp. *aestivum*, AABBDD) [2]. The fast global changes of the last decade have made agricultural productivity more uncertain. Particularly, rising temperatures and decreased water availability are primary reasons for crop yield losses and reductions in the area harvested [3–5]. Barley is a crop that is adapted to a wide range of environmental landscapes, including high altitude and high latitude regions and to saline and dry conditions. Furthermore, barley flour has low machinability and bread-making performance compared to wheat. It was found that barley breads have an increased viscosity of the bolus, due to the presence

of resistant starch and fibers, that reduced enzymes' accessibility and therefore slowed the in vitro static and dynamic starch digestion [6]. Since the beginning of the twentieth century, cereal breeders have focused their efforts on the development of interspecific wheat hybrids to obtain new cereals with increased phytochemical content, improved agronomic performances and technological qualities. The hexaploid hybrid tritordeum (x *tritordeum martini*, AABBHchHch) is the product of cross-breeding *Hordeum chilense*, a South American wild barley species, and durum wheat. This hybridization aimed to combine the excellent traits of the *Hordeum*, such as high endosperm carotenoid content and higher tolerance to biotic and abiotic stress, with the technological qualities of wheat [7]. Tritordeum is today commercialized as an innovative alternative to conventional small cereal crops (www.tritordeum.com) [8,9], with rheological and baking performances similar to bread wheat [10]. Interestingly, in a clinical study involving subjects with non-celiac gluten sensitivity (NCGS), tritordeum breads were sensorially more appreciated than the gluten-free counterpart, showing good gastrointestinal tolerance [8,11].

A few recent studies in Europe have been focused on the agronomical traits, looking at yield performance of tritordeum cultivars (cvs) over conventional soft and durum wheats [7,12,13]. Scientific works have so far mainly focused on the bioactive compounds' content: tritordeum, in fact, has higher levels of carotenoids and arabinoxylans than wheat, and these result in a greater total antioxidant activity. Tritordeum has twice the amount of β-glucans compared to durum wheat, although a similar amount to soft wheat, but significantly higher arabinoxylans [14] with prebiotic, immunomodulatory, antitumor, and anti-inflammatory activities [15,16]. Grain protein content (GPC) and gluten composition, in addition to the nutritional traits, play a major role in conferring the technological properties of wheat and other small cereals to the dough. By comparing the GPC of tritordeum with that of durum and soft wheat in different climatic conditions, tritordeum results in a higher GPC, ranging between 11% and 17%. Tritordeum produced by organic farming showed higher GPC than durum wheat with a larger number of high-molecular-weight glutenin subunits [13]. Although one of the parental lines of tritordeum is a durum wheat cultivar (cv), the similarity, in terms of derived flour quality, is much closer to that of hexaploid soft wheat, with specific interest for bread-making and baking processes [13].

Furthermore, little is known about the protein level differences across cvs or regarding the tritordeum protein digestibility.

This study aims at comparing the protein profile of two tritordeum cvs, Bulel and Aucan, with soft wheat cv Altamira and a durum wheat cv, Antalis. Model breads were used to compare the digestion products (free-amino nitrogen and free glucose) using an in vitro digestion that included a standardized oral, gastric, and duodenal model simulating the physiological conditions of a healthy adult [17].

This work represents the first molecular characterization by advanced mass spectrometry of the peptides resistant to digestion of tritordeum bread, which was prepared with the two most common cvs, namely Aucan and Bulel. We mapped the contribution of the parent *H. chilense* to the tritordeum proteome and evaluated in silico the presence of peptides related to celiac toxicity and allergenicity. The immunoreactivity of the R5 monoclonal antibody targeting the celiacogenic sequence "QQPFP" was studied in the flours, as well as in undigested and digested bread.

2. Results and Discussion

2.1. Flour Protein Characterization

Flour from two tritordeum cvs, Aucan and Bulel, which are the most widely cultivated in Europe and agronomically characterized, were compared to flour obtained from two wheat cvs, Altamira (soft wheat) and Antalis (durum wheat), for mapping differences in protein content and quality.

The durum wheat, Antalis cv, and tritordeum, Aucan and Bulel cvs, had a total protein content (TPC) higher than the soft wheat, Altamira cv ($p < 0.05$) (Table 1).

Aucan also had the highest ash amount. These data are in line with a previous study looking at the adaptability of tritordeum cvs in the east Mediterranean region, showing Aucan with a higher protein content compared to Bulel and to the soft wheat cv Falado [11].

Table 1. Ash and grain protein content (GPC) and total protein content of the flour (TPC).

Species	Cv	Ashes (%)	GPC (%)	TPC (%)
Soft wheat	Altamira	1.89 ± 0.04^{b}	11.17 ± 0.23^{a}	8.65 ± 0.01^{a}
Durum wheat	Antalis	1.82 ± 0.02^{a}	11.66 ± 0.35^{b}	9.19 ± 0.21^{ab}
Tritordeum	Aucan	1.96 ± 0.02^{c}	13.43 ± 0.15^{c}	10.63 ± 0.33^{c}
Tritordeum	Bulel	1.86 ± 0.04^{ab}	11.97 ± 0.07^{b}	9.84 ± 0.3^{bc}

Values followed by different letters are significantly different ($p < 0.05$).

The concentration of gliadins in the flours was determined using a commercial sandwich ELISA test kit with the R5 monoclonal antibody to target the QQPFP celiac toxic motif (Figure 1) and the homologous LQPFP, QLPYP, QQSFP, QQTFP, PQPFPF, QQPYP, and PQPFP to a lower degree [18]. The toxic sequence appears repeatedly in the ω-, γ-, and α/β-gliadins [19]. The data are expressed as mg of R5 gliadin per kg of flour and the mg of gluten can be extrapolated using a conversion factor of two as suggested by the ELISA manufacturer. Despite the fact that the Altamira cv showed the lowest TPC, this soft wheat had the highest R5-gliadin concentration, which was comparable to Aucan, which had a significantly higher protein content. The durum wheat cv Antalis had a 40% lower content of R5-gliadin concentration per kg of flour compared to soft wheat cv Altamira. These results were in line with literature data showing the gluten content in tritordeum to be comparable or even higher to bread wheats, with Aucan higher in gluten compared with Bulel by four percentage points [11]. Interestingly, Bulel showed a 66% lower R5 immunoreactivity per kg of flour compared to Aucan. The reduced immunoreactivity of Bulel cv underlies important differences in terms of R5-gliadin sequences between the two tritordeum cvs. The analytical methodologies available for gluten determination in wheat may not be appropriate for all tritordeum cvs, since they may lead to an underestimation of the gluten content due to structural differences of the protein sequences.

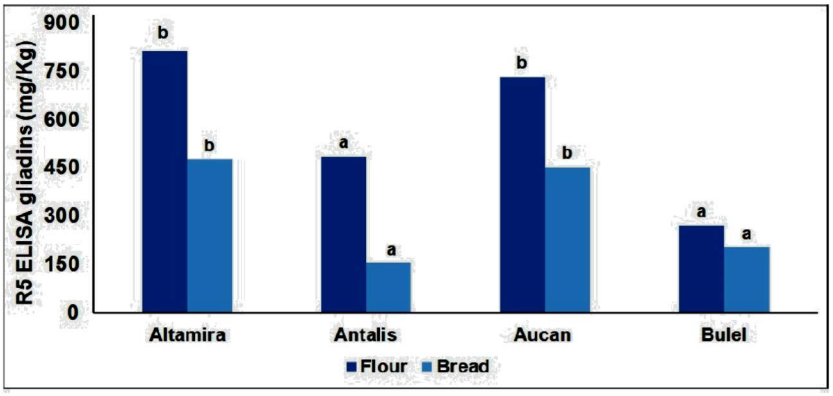


Figure 1. Quantification of the "QQPFP" celiac toxic motif recognized by the R5 monoclonal antibody of proteins extracted from the flours and the model bread obtained from different cereals. Within each product, flour and bread, bars with different letters are significantly different ($p < 0.05$) and the REGW-F test.

The protein profile under reducing conditions of the Osborne fractionated proteins is presented in Figure 2. Tritordeum cv Bulel showed an electrophoretic profile of all fractions being less complex compared to both Aucan and durum wheat. The salt soluble protein profile (albumins and globulins) (Figure 2A) of both the tritordeum cvs appeared compara-

ble to that of soft wheat, with a higher number of bands than the durum wheat flour. These differences were more pronounced in the lower-molecular-weight region (Mr < 30 kDa) and may be attributed to proteins encoded by the *H. chilense* inherited genome.

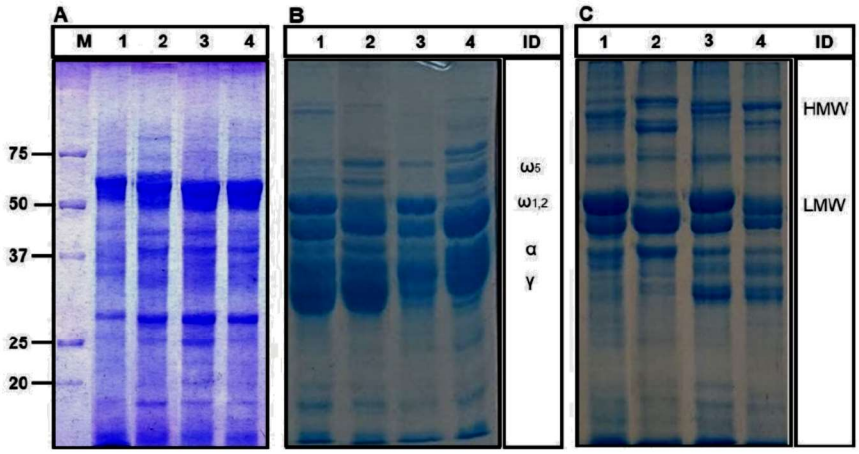


Figure 2. Electrophoresis of Osborne fractions: albumins and globulins (Panel **A**); gliadins (Panel **B**) and glutenins (Panel **C**). Panel **A** was stained in Brilliant Blue Coomassie R250; Panel **B**,**C** with G250. M: Molecular markers (Precision plus Protein—Biorad); Lane 1: durum wheat cv Antalis; Lane 2: soft wheat cv Altamira; Lane 3: tritordeum cv Bulel; Lane 4: tritordeum cv Aucan; ID: Identification based on Landolfi et al., 2021 [20]. ω5: omega 5 gliadins; ω1,2: omega 1,2 gliadins; α: alpha gliadins; γ: gamma gliadins; HMW: high molecular weight glutenins; LMW: low molecular weight glutenins.

The electrophoretic profile of gliadins can be divided into four zones representing the typical regions of ω-, γ-, β-and α-gliadins (Figure 2B). A greater protein variability was detected in Bulel compared to Aucan and the two wheat flours. The lowest number of bands were identified in the Bulel's gliadins, particularly in the high molecular mobility region (Mr > 50 kDa) where the ω-gliadins migrate. This profile is consistent with the ELISA data showing Bulel characterized by the lowest concentration of gliadin detectable with the R5 antibody (Figures 1 and 2). Aucan showed a greater complexity in the same region even when compared to the two reference wheat flours.

The electrophoretic profile of glutenin fractions varied across the four cvs (Figure 2C) in terms of the number of electrophoretic bands detected and electrophoresis mobility. The high-molecular-weight glutenin subunits (HMW-GSs) are responsible for the gluten supramolecular structure, providing the cysteines involved in the formation of the disulfide-bonded backbone in gluten network, affecting the rheological properties of dough [21]. The presence of HMW-GS in tritordeum is due to the contribution of *H. chilense* locus "Glu-Hch1" gene expression on the chromosome 1Hch [22]. This locus is homologous of the wheat Glu-1 locus and to the barley Hor-3 locus [23]. The *H. chilense* genome promotes a similar effect on gluten strength as the D genome inherited by the wheat species from *Aegilops tauschii* [24].

The region of the low-molecular weight glutenin subunits (LMW-GSs) in the two tritordeum cvs appeared similar between each other and to the durum wheat with a higher number of bands with faster molecular mobility compared to soft wheat. Unlike the control wheat, both the tritordeum cvs showed the presence of two main bands in the region of Mr < 30 kDa, which likely are expressed by the *H. chilense* inherited genome (mother).

Once again, the electrophoresis showed substantial differences in terms of the overall protein expression between the two tritordeum cvs under evaluation. Since the two tritordeum cultivars share the same *H. chilense* line as mother, while they differ in the line of *T. turgidum* spp. *durum* used as father, these differences should be attributed primarily to the durum wheat inherited genome (Arcadia S.p.A., personal communication).

2.2. *Digestomics*

The simulated gastroduodenal digestion was performed on model breads prepared using refined flours to a 35% starting hydration (Figure S1, Supplementary Materials). Breads were subjected to simulated digestion within a few hours from cooking to avoid any alteration of the starch that would have impaired (affected) the digestion. The R5 immunoreactivity of the four kinds of bread was measured and reported in Figure 1. In all cases, the R5 gliadin content of the bread samples was lower than the respective flour, although being comparable in terms of order of magnitude. This was somewhat expected, due to both the formation of the gluten network and baking-induced protein modifications, which may have partly impaired the protein extraction.

2.2.1. Quantitative Analysis of the End Products of Digestion

The digestion products of bread were quantitatively evaluated. The reducing sugar release (RSR) over duodenal digestion was measured by the enzymatic-spectrophotometric method (Figure S2). The RSR curves of the four bread samples were comparable ($p > 0.05$). Total starch content in tritordeum is knowingly higher than barley and lower than wheat, with a content in resistant starch similar to barley [25]. In the stomach and in the intestine, resistant starch, together with the higher viscosity (due to the presence of soluble fibers), makes the chyme of the barley bread less accessible to the enzymes, reducing the glycemic index compared with the reference wheat bread [6]. The content of β-glucans in tritordeum was found to be five times lower than barley [13]. The quantification of reducing sugars at the end of digestion has only been a side part of this study. We are planning a forthcoming investigation aimed at quantifying the end products of starch digestion to confirm a different behavior for tritordeum bread than from its wheat counterpart.

Free amino acids, di- and tripeptides are the products of protein digestion that can be transported across the intestinal barrier. The starting content of free amino acids released during bread preparation was measured in the undigested cooked breads (Figure S3). Bread made with the two tritordeum cvs and the durum wheat cv had a starting content of α-amino nitrogen of 0.35% (w/w), the soft bread was 0.25% and the protein-free bread was 0.15%. The commercial protein-free bread, used as background reference to quantify the endogenous amino acids products of natural gastroduodenal enzymes turnover, had an α-amino nitrogen content < 0.1%, in line with what was declared on the food package label.

The level of α-amino nitrogen determined in the digested protein-free bread accounted for half of the content of the analyzed bread samples on average. This background level is most likely due to the autoproteolysis of the digestive enzymes. This underlines the importance of having protein-free reference matrices to be used as background reference samples. The breads baked with the soft and the durum wheat showed a comparable digestibility (Figure 3). The two tritordeum cvs showed the highest release ($p > 0.05$) of α-amino nitrogen related to the protein concentration of the flour determined by Kjeldahl.

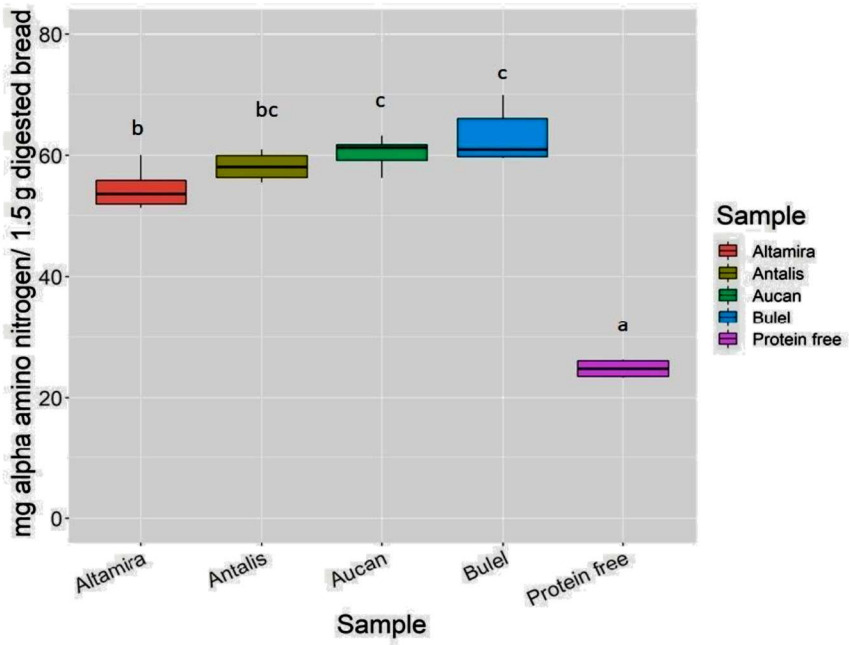


Figure 3. Alpha-amino nitrogen released at the end of the duodenal simulated digestion of 1.5 g of bread obtained from the different cereals. Bars with different letters are significantly different ($p < 0.05$) and the Tukey-test.

2.2.2. Qualitative Evaluation of the Peptides Resistant to Digestion

The availability of well annotated and curated protein sequence databases is essential for inferring relevant information from mass spectrometry data. The analysis of cereal seed storage proteins is challenging because of the natural polymorphism, with a high number of protein isoforms differing by point mutations, and the homology across cvs and species [26]. Tritordeum, being a novel crop, lacks a protein database. Therefore, the identification of the proteins was performed using a combined database of the two parent proteomes, *H. chilense* and *Triticum turgidum* spp. *durum*, and of *H. vulgare*.

The tritordeum bread derived peptides, resistant to gastroduodenal digestion primarily belonging to α-amylase inhibitors (AAI) and to the glutenin family (Tables 2, S1 and S3). The AAI are knowingly resistant to gastroduodenal digestion, mainly due to the presence of disulphide bridges that stabilize the polypeptide chain [27,28], and are involved in IgE-mediated wheat (Tri a 28–39) and barley food allergies (Hor v 15) [29,30].

Several peptides were identified as derived from *Triticum* proteins, and fewer were associated with the *Hordeum* (Tables S1 and S3). Peptides belonging to the γ-3-ordeins (Uniprot ID: Q6EEY5) and the D-hordein (Uniprot ID: B0L965) from *H. chilense* could be identified in both digests of tritordeum bread. Unique proteins to tritordeum cv Aucan and Bulel were identified. AAI and glutenin subunits may be suggested as suited species markers for discriminating between the Bulel and the Aucan varieties. As expected, the unique proteins were expressed from the *Triticum* father line differing between the two tritordeum. Interestingly, despite the common *Hordeum* mother, a gamma prolamin of the *Hordeum brachyantherum* subsp. *brachyantherum* (mother) could be uniquely identified by homology in cv Bulel.

Table 2. LCMSMS identified proteins in tritordeum bread digests. Only proteins identified in both technical replicates were taken into consideration to increase the confidence in identification. Isoforms were removed and the extensive list of identified proteins is available as Supplementary Materials Tables S1 and S3.

	Accession	Species	−10LgP	Coverage (%)	Peptides	Description
PROTEINS IDENTIFIED IN BOTH TRITORDEUM DIGESTS	Q9XGF0	TRITD	74.72	20	12	LMW-GS
	A0A446W0B5	TRITD	72.20	14	7	AAI
	K4N1X7	TRITD	74.67	10	8	HMW-GS
	A0A446W0A1	TRITD	76.67	12	9	AAI
	H8Y0D1	TRITD	68.82	15	9	Alpha prolamin
	A0A446W0B4	TRITD	63.66	12	4	UNP
	A0A446W085	TRITD	71.15	11	7	AAI
	A0A446TL77	TRITD	39.57	5	2	rRNA N-glycosidase
	A0A446W0C7	TRITD	51.01	9	3	AAI
	A0A446V2J2	TRITD	42.65	4	2	AAI
	A0A446V2Q9	TRITD	45.34	8	3	AAI
	Q6EEY5	HORCH	40.78	8	3	Gamma 3 hordein
	B0L965	HORCH	31.52	2	1	D-hordein
	A0A446YMF0/ M0WF36	TRITD/HORVV	21.54	4	1	UNP
	A0A287EEX5	TRITD	40.07	6	3	UNP
PROTEIN IDENTIFIED ONLY IN TRITORDEUM CV BULEL	A0A446JGR8	TRITD	63.12	8	5	AAI
	A0A0E4G9A4	TRITD	48.57	6	5	HMW-GS
	H8Y0M9	HORBR	37.19	12	3	Gamma prolamin
	A0A7H1K1W3	TRITD	31.27	7	2	AAI
	A0A446IHD3	TRITD	20.67	6	1	AAI
	A0A446IHC0	TRITD	31.50	4	1	AAI
PROTEINS IDENTIFIED IN TRITORDEUM CV AUCAN	A0A2L1K3K6	TRITD	77.43	12	11	HMW-GS
	Q41603	TRITD	44.31	9	3	LMW-GS

TRITD = *Triticum turgidum* subsp. *durum*; HORCH = *Hordeum chilense*; HORBR = *Hordeum brachyantherum* subsp. *brachyantherum*; HORVV = *Hordeum vulgare*; UNP = uncharacterized protein; LMW-GSs = low molecular weight-glutenin subunits; HMW-GSs = high molecular weight-glutenin subunits; AAI = α-amylase inhibitors.

The immunoreactivity of duodenal digests determined using a competitive R5-competitive ELISA was 50% lower in digested tritordeum and durum wheat bread samples compared with the soft wheat bread ($p < 0.05$) (Figure 4).

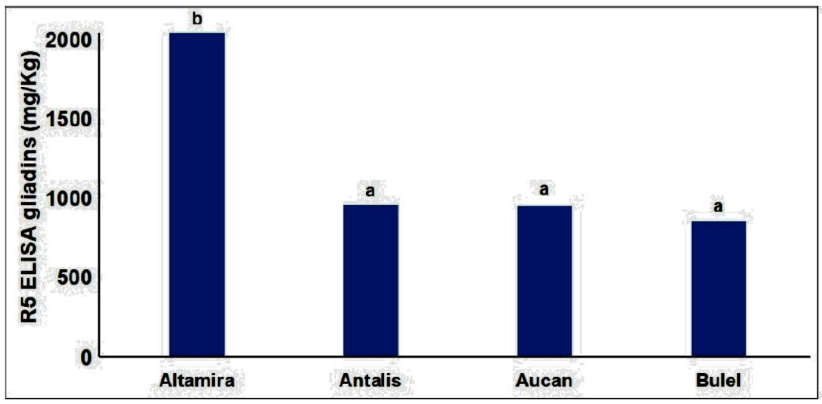


Figure 4. Quantification of the "QQPFP" celiac toxic motif recognized by the R5 monoclonal antibody in duodenal digests (mg of gliadins per kg of the soluble duodenal digest) of bread obtained from different cereals. Bars with different letters are significantly different ($p < 0.05$) and the REGW-F test.

While the R5 immunoreactivity of Aucan flour and bread was comparable with the soft wheat flour, the digests of tritordeum cv Aucan bread had an immunoreactivity comparable with that of tritordeum cv Bulel and the durum wheat digests. The analyses

were performed on the soluble digest, which is likely the fraction to be uptaken in the gut. The lower R5-immunoreactivity of Bulel digest is in line with previous literature data [20]. The reduced immunoreactivity of the Aucan bread duodenal digest may be explained by a low digestion level that could have spared large protein fragments carrying the R5-epitope(s) trapped in the insoluble fraction. This fraction is not uptaken by enterocytes and represents the primary fermentation substrate of gut microbiota. A recent in-vivo study, showed a significant decrease of gluten intestinal peptides, determined by ELISA, in the stool of subjects fed with tritordeum bread compared with wheat bread-fed subjects [8]. The bread produced with Bulel flour, under our analytical conditions, showed similar results as the in vivo study presented by Vaquero et al., 2018 [8]. The Aucan bread instead behaved in a completely different way suggesting future in vivo studies may need to be designed to include different tritordeum cvs to confirm their suitability for subjects affected by non-celiac gluten sensitivity, especially in consideration of the relative stability of AAI.

Due to the complexity of the mass spectrometry data an in-silico evaluation was carried out only on those peptides identified in both technical replicates, to enhance confidence. Overall, 93 peptides resistant to gastroduodenal digestion identified by mass spectrometry were common to the digests of Aucan and Bulel bread (Tables S3 and S4); 38 and 59 peptides were uniquely identified in the Bulel and Aucan bread digests, respectively (Table S5). Many of the unique peptides were inferred to the unique proteins previously listed in Table 1. Interestingly, one HMW-GS protein (Uniprot accession K4N1X7) was common to the two tritordeum cvs. The majority of the peptides identified in both digests mapped to the same protein regions, however, they had different N- and C-terminal trimming, therefore were assigned as unique (Figure S5). Two peptides with sequences, 130-QSGQGQQPGQGQQP-143 and 213-QSGQGQQPGQGQPG-226 were uniquely identified in Aucan and were located in the N-terminal region of the protein. No peptides were identified in the same protein region among the Bulel-derived peptides. In contrast, the peptide with sequence 342-SLQQPGQGQQPGQGQPG-358 was identified only in Bulel. These misidentifications may be due to the bioinformatic protein inferring process, that would only list peptides with 100% identity. Wheat proteins are characterized by high polymorphism and the presence of several protein sequences differing by few amino acids [26]. The protein assignment informs about the presence of a protein family rather than a specific protein, especially for gluten proteins. In this case, it may indicate the presence of two isoforms of the HMW-GS (K4N1X7) expressed in the two tritordeum cvs, carrying mutations in the two identified regions. Two studies previously attempted to map the products of tritordeum cvs that had undergone simulated digestion, working either on isolated proteins [8] or the flour [20]. This is the first study mapping the digestion products of model bread prepared with tritordeum flour, using the INFOGEST standardized model [17].

The in-silico epitope analysis showed a larger number of peptide precursors of celiac toxic motifs and IgE binding peptides for the Aucan bread than for the Bulel counterpart (Figure 5). The analysis of epitopes also showed the prevalent contribution of *Triticum* in the overall allergenicity/celiacogenic potential of tritordeum bread.

The sequence analysis (Figure 5 and Figure S4) of digestion-resistant peptides showed the high frequency among others of QQPFP, QQPYP, PQPFP sequences, which are targets of the R5 competitive ELISA. The mapping of the R5-epitopes within the protein sequences from the two tritordeum cvs highlighted the prevalent contribution of *Triticum*-derived sequences in Aucan and *Hordeum*-derived sequences in Bulel.

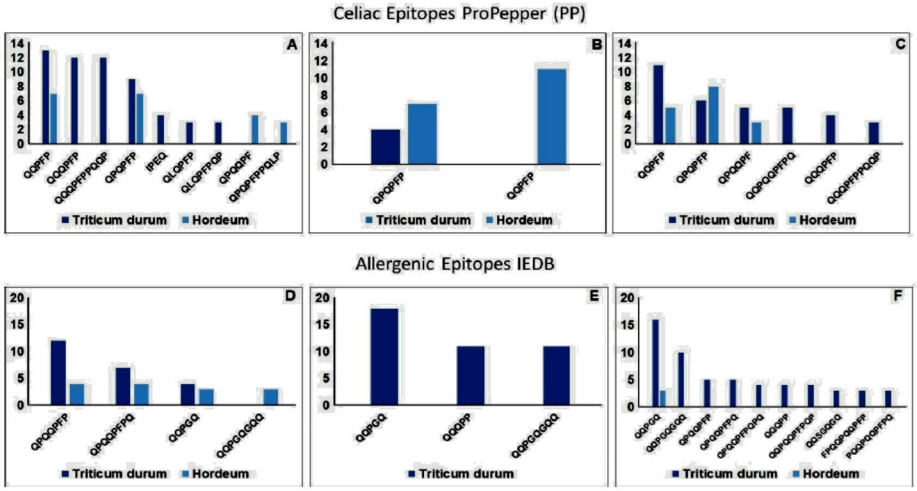


Figure 5. Graphical representation of the peptides surviving the digestion with potential adverse effects on human health. (**A**) number of CD epitopes (ProPepper) common epitopes found in digests of bread baked with the two tritordeum cvs); (**B**) number of CD epitopes (ProPepper) found uniquely in digests of tritordeum cv Bulel bread; (**C**) number of CD epitopes (ProPepper) found uniquely in digests of tritordeum cv Aucan bread; (**D**) number of allergenic epitopes (IEDB) common epitopes found in digests of bread baked with the two tritordeum cvs; (**E**) number of allergenic epitopes (IEDB) found uniquely in digests of tritordeum cv Bulel bread; (**F**) number of allergenic epitopes (IEDB) found uniquely in digests of tritordeum cv Aucan bread. Only epitopes identified in at least 3 precursor peptides were reported.

3. Materials and Methods

3.1. Grains and Flours Production

A field study was carried out on the north-west Italian plain at Cigliano (45°18′ N, 08°1′ E; elevation 237 m), in the 2019–2020 growing season. The experiment was performed on a silty-loam soil sub acid, characterized by a medium cation-exchange capacity and organic matter content. In the same experimental field, the following genotypes have been cultivated side by side:

- A soft wheat (hexaploid AABBDD), cv named Altamira (seeds provided by Limagrain Italia S.p.A., Busseto, Italy) classified as ordinary bread-making wheat [31] registered in the Italian varietal list in 2009 (https://www.sian.it/mivmPubb/listeVarieta.do; Sian code: 11239; consulted on the 20 December 2021) and widely cultivated in Italy;
- A durum wheat (tetraploid AABB), cv named Antalis (seeds provided by CGS Sementi S.p.A., Acquasparta, Italy), characterized by medium-high GPC and gluten index; registered in the Italian varietal list in 2014 and widely cultivated in Italy;
- Tritordeum (hexaploid AABBHchHch), cv named Bulel (seeds provided by Arcadia S.p.A., Pamplona, Spain), which was registered in the CPVO (Community Plant Variety Office) List in 2015;
- Tritordeum (hexaploid AABBHchHch), cv named Aucan (seeds provided by Arcadia S.p.A., Pamplona, Spain), which was registered in the CPVO List in 2013.

The treatments were assigned to experimental units using a completely randomized block design with four replicates. The plot size was 7 × 1.5 m.

The same agronomic techniques have been adopted for all cvs. Briefly, the previous crop was maize and planting was performed in 12 cm wide rows at a seeding rate of 400 seeds m^{-2} on November 6th 2019, following an autumn plowing (30 cm) and disk harrowing to prepare a proper seedbed. A N fertilization treatment of 130 kg N ha^{-1} was used on all the cultivated samples. The total N rate for each treatment was top-

dressed applied as a granular ammonium nitrate fertilizer, split 50 kg N ha^{-1} at tillering (growth stage, GS23) and 80 kg N ha^{-1} at the beginning of stem elongation (GS32). The foliar diseases were controlled by applying a fungicide (pyraclostrobin 150 g ha^{-1} and fluxapyroxad 75 g ha^{-1}, Priaxor®, BASF Agricultural Solutions) at booting stage (GS45). Harvesting was carried out with a Walter Wintersteiger cereal plot combine-harvester on 29 June 2020.

Grains (2 kg) from each plot and cv were milled using the Bona 4RB mill (Bona, Monza, Italy) to obtain refined flour, (tritordeum and soft wheat) and semolina (durum wheat). GPC (Kjeldahl N × 5.7, on a dry matter basis) and ash content were determined according to Blandino et al., 2015 [32] on grains collected at the commercial maturity stage. Grains (2 kg) from each plot and cv were milled using the Bona 4RB mill (Bona, Monza, Italy) to obtain refined flour.

3.2. Materials

All the reagents used in this study were of analytical or higher grade. Sodium phosphate, ammonium bicarbonate (AmBic), acetic acid and the other chemicals used to produce the simulated salivary fluid (SSF), the simulated gastric fluid (SGF) and the simulated intestinal fluid (SIF), were provided by Carlo Erba (Milan, Italy). The enzymes used for in vitro human digestion were purchased from Sigma (St Louis, MO, USA), in line with those recommended by the INFOGEST protocol [33]. Trichloroacetic acid (TCA), sodium dodecyl sulfate (SDS), glycerol, tris(hydroxymethyl) aminomethane hydrochloride (Tris-HCl), ethylenediaminetetraacetic acid (EDTA), guanidine chloride, trifluoroacetic acid (TFA), formic acid (FA), acetonitrile (ACN), 2-vinylpyridine monomer, and *p*-toluenesulfonyl-L-arginine methyl ester (TAME) were also from Sigma-Aldrich. Egg lecithin was purchased from Lipid Products (Redhill, UK). The electrophoresis reagents were all obtained from Bio-Rad (Milan, Italy).

The protein-free bread (<0.1%, *w*/*w*) (Amino' pane le rosette produced by Antica Farmacia Orlandi) was purchased from a local pharmacy.

3.3. Quantification of Protein in Flour and Semolina

The Kjeldahl analysis was performed as described by Abrams et al., 2014 [34], with some modifications. Two grams of each flour were weighed in a Kjeldahl tube, in which a mixture of copper and potassium sulphate (0.5 g and 12 g, respectively) and 20 mL of 96% sulfuric acid were added. The mineralization was performed following a thermal ramp: 230 °C for 20 min, 290 °C for 45 min, 320 °C for 35 min, and 420 °C for 60 min. The sample was diluted with 50 mL of deionized water, and 90 mL of 45% NaOH were added. The solution of the ammonia was distilled over steam and collected in a flask containing 50 mL of 4% boric acid. The total nitrogen was determined by titration with 0.1 N HCl, after adding a mixed indicator (methyl red 0.1% and bromocresol green 0.2% in ethanol). A conversion factor of 5.7 was used to convert the total nitrogen to total protein and the results were expressed as g of total protein over 100 g of sample. Samples were analyzed in biological duplicates.

3.4. Gliadin Quantification with R5 Commercial ELISA

Flour and bread samples were analyzed with the RIDASCREEN® Gliadin (Art. No. R7001, R-BIOPHARM AG, Darmstadt, Germany), which is a sandwich enzyme immunoassay (ELISA) based on R5 monoclonal antibody recognising the "QQPFP" celiac toxic motif. Proteins were extracted in the Cocktail (patented) recommended by Codex Alimentarius for the optimized extraction of gliadin from heat-processed and non-heated food samples (Art. No.: R7006/R7016, patent WO 02/092633, R-BIOPHARM AG, Darmstadt, Germany), which was used according to the manufacturer's instructions, and according to the AOAC Official Method of Analysis for gluten detection (OMA 2012.01).

RIDASCREEN® Gliadin competitive (Art. No. R7021, R-BIOPHARM AG, Darmstadt, Germany) was used to analyse the products of in vitro bread digestion, according to the manufacturer's instructions.

The gastroduodenal digestion products were analyzed in duplicate.

3.5. Osborne Fractionation

The Osborne fractionation was performed as previously described in Landolfi et al. 2021 [20]. The albumins and globulins were solubilized from the non-defatted flour (1:10, w/v) in 100 mM KCl, 50 mM Tris-HCl pH 7.8, and 5 mM EDTA for 4 h at room temperature (−20 °C). The solution was centrifugated for 15 min at 3500× g and the supernatants from two consecutive extractions were pooled. The gliadins were extracted 1:10 w/v with 70% (v/v) ethanol for twelve hours at room temperature (−20 °C). Glutenin extraction was performed at 60 °C for 30 min, in 50% v/v 1-PrOH + 50 mM Tris-HCl (pH 8.5) + 1% (w/v) 1,4-Dithio-D-threitol (DTT). The cysteine residues of the glutenin extracts were pyridylethylated at 60 °C for 15 min with 2-vinylpyridine.

3.6. 1-Dimensional Electrophoresis (SDS-PAGE)

Purified protein fractions were separated by SDS-PAGE under reducing conditions, using a Mini-PROTEAN cell systems (Bio-Rad). To this purpose, proteins were precipitated in cold (−20 °C) propan-2-one (1:4, v/v), suspended in the SDS-PAGE Laemmli Buffer (0.125 M Tris–HCl pH 6.8, 5% SDS, 20% glycerol, 5% (w/v) 2-sulfanylethanol, 0.02% bromophenol Blue) and boiled in a water bath for 5 min. After quantification with a micro-Lowry kit (Sigma-Aldrich, Saint Loius, MO, USA), 25 µg of gliadins, 75 µg of glutenins, and 75 µg of albumins/globulins were loaded onto a 12% acrylamide gel (Bio-Rad). Migration of proteins was conducted at 120 V for 10 min and 220 V for 35 min. Afterwards, gels were fixed with TCA (24%) overnight (16 h) and stained with Coomassie® Brilliant Blue R-250 (gliadins and glutenins) and G-250 (albumins and globulins).

3.7. Preparation of the Model Breads

Model breads were prepared using commercial baker's yeast (*Saccharomyces cerevisiae*). The same recipe was used for all the flours from different cvs. Flour (50 g) was mixed with 33 g of water and 5 g of yeast and 7 g of salt were added. The mixture was allowed to stand at 20 °C for 12 h. Loaves were baked at 230 °C for 40 min. After cooling, loaves were cut into slices and subjected to in vitro digestion within a few hours to avoid any alteration of the digestibility due to storage conditions (e.g., starch retrogradation due to freezing) [35].

3.8. Static Oral-Gastric-Duodenal Digestion of Model Breads

In vitro oral and gastroduodenal digestion was carried out using the harmonized and standardized INFOGEST method [33]. The trypsin activity of the porcine pancreatin was determined using the *p*-toluene-sulfonyl-L-arginine methyl ester (TAME) as the substrate according to Brodkorb [17] and measuring the absorbance at 247 nm for 10 min. The trypsin activity was found to be 9.5 U/mg of powder. At the end of the duodenal digestion, samples were boiled for 5 min to interrupt the enzymatic digestion and centrifuged at 7900× g, for 30 min. The supernatant containing digestion products, which are likely to be absorbed by enterocytes, was collected, and processed for further analysis, including peptidomics and α-amino nitrogen determination. Aliquots of duodenal digests were collected every 30 min.

3.9. Preparation of Samples for the Alpha Amino Nitrogen Determination

The solubilization of proteins from the cooked bread samples and protein-free bread was performed in the SSF without enzymes (1:12, w/v) for 3 h at 37 °C. Prior to α-amino nitrogen determination, all sample were de-proteinized TCA up to a final concentration of 20% (w/v). After the protein precipitation was conducted for 30 min at 20 °C, the solution

was centrifuged at 4000× *g* for 30 min, 4 °C and neutralised to a pH 7 with 1 N NaOH prior to analysis.

The content of free α-amino nitrogen in the samples was determined using the EnzytecTM Alpha-amino Nitrogen kit by R-Biopharm (E2500 R-Biopharm AG, 64297 Darmstadt, Germany) following the manufacturer's instructions. The iCubio i-Magic M9 (Origlia S.r.L, 20007 Cornaredo, Italy) was set to perform the enzymatic reaction in full automatization and the absorbance was read at 340 nm. All samples were assayed in triplicate and absorbance values were averaged.

3.10. Free Glucose Quantitative Determination

D-glucose was quantified directly in the soluble digest using the Enzytec™ Liquid D-Glucose kit by R-Biopharm (E8140 R-Biopharm AG, 64297 Darmstadt, Germany), following the manufacturer's instructions. The analyses were performed on the iCubio i-Magic M9 (Origlia S.r.L, 20007 Cornaredo, Italy) as described for the α-amino nitrogen (Section 3.9). All samples were assayed in triplicate and absorbance values were averaged.

3.11. Preparation of Peptides for Mass Spectrometry Analysis

Peptide digests were desalted using C18 Sep-Pak 360 mg sorbent weight (WAT051910, particle size 55–105 µm, pore size 125 Å) (Waters Co., Milford, MA, USA). The equilibration and cleaning phases were carried with a 0.1% TFA in water. Peptides were eluted with 70% acetonitrile (*v*/*v*) containing 0.1% TFA (*v*/*v*).

3.12. Liquid Chromatography-Tandem Mass Spectrometry (LC/MSMS) Analysis

LC-MS/MS analysis was performed by using a Q Exactive Orbitrap mass spectrometer (Thermo Scientific, San Jose, CA, USA), online coupled with Ultimate 3000 ultra-high performance liquid chromatography equipment (Thermo Scientific, 95134 San Jose, CA, USA)). Samples were loaded through a 5mm long 300 µm id pre-column (LC Packings, San Jose, CA, USA) and separated by an EASYSpray™ PepMap C18 column (2 µm, 25 cm × 75 µm) 3 µm particles, 100 Å pore size (Thermo Scientific, San Jose, CA, USA). Eluent A was 0.1% formic acid (FA) (*v*/*v*) in water; eluent B was 0.1% FA (*v*/*v*) in 80% (*v*/*v*) ACN. The column was equilibrated at 5% B. Peptides were separated applying a 5–40% gradient of B over 60 min. The flow rate was 300 nL/min. The mass spectrometer operated in data-dependent mode and all MS1 spectra were acquired in the positive ionization mode with an m/z scan range of 350 to 1600. Up to 10 of the most intense ions in MS1 were selected for fragmentation in MS/MS mode. A resolving power of 70,000 full width at half maximum (FWHM), an automatic gain control (AGC) target of 1×10^6 ions and a maximum ion injection time (IT) of 120 ms were set to generate precursor spectra. MS/MS fragmentation spectra were obtained at a resolving power of 17,500 FWHM. To prevent repeated fragmentation of the most abundant ions, a dynamic exclusion of 10s was applied. Ions with one or over six charges were excluded.

A specific database was generated for the analysis of the MS/MS data. The database included UniprotKB entries for *Triticum turgidum* spp. *durum*, *Hordeum chilense* and *Hordeum spontaneum*, downloaded on the 01/06/2021. The DB Toolkit was used to customize the database and remove redundant sequences [36,37]. The *Sus scrofa* protein sequences, downloaded on the 16/10/2019 from UniprotKB, were also included in the database to detect contaminants, thus increasing the confidence of identification of the Tritordeum-derived peptides.

PEAKS Studio (version 6.0, Bioinformatics Solution Inc., 202-140 Columbia St W, Waterloo, ON, Canada) was used for database searching, applying the following parameters: oxidation on methionine, deamidation on the glutamine and asparagine, and pyroglutamic for N-terminus glutamine as variable modifications; mass tolerance value of 8 ppm and 0.02 Da for precursor and MS/MS fragment ions, respectively; no cleavage specificity. The peptide-level false discovery rate (FDR) was set at 0.1%. Proteins with score −10LgP > 20 were accepted.

3.13. In Silico Analysis of Peptides Resistant to Digestion

Peptides identified at the end of the gastroduodenal digestion were in silico evaluated for their celiacogenic potential and IgE capacity. IgE binding sequences were retrieved from the free Immune Epitope Database (IEDB) (https://www.iedb.org/, downloaded on the 20 September 2021).

The celiac toxic motif was retrieved from the ProPepper database (https://www.propepper.net/, downloaded on the 20 September 2021) [38]. These epitopic/celiacogenic sequences were manually searched in the pool of resistant peptides identified by MS in the duodenal digests of Bulel and Aucan breads.

To increase the confidence of identification and the strength of the *in-silico* analysis, the epitopic/celiacogenic sequences were searched, only considering the peptides identified in both technical replicates. The analyses were performed with the peptides common to the tritordeum cvs and those uniquely identified in Bulel and Aucan. Peptides resulting from digestion that belonged to the same protein were aligned with Clustal Omega (https://www.ebi.ac.uk/Tools/msa/clustalo/, accessed on 20 December 2021) and graphically evaluated using the WebLogo software (https://weblogo.berkeley.edu/logo.cgi, accessed on 20 December 2021) to highlight the recurring regions. This analysis was aimed at aligning in silico the surviving peptides to their toxic-allergenic potential.

3.14. Statistical Analysis

Statistical analysis was performed using SPSS statistic v.27 Chicago: SPSS Inc. Data were compared by one-way analysis of variance (ANOVA), followed by the Tukey–Kramer post hoc test ($\alpha = 0.05$) for all analysis and Ryan–Eino–Gabriel–Welsch F test ($\alpha = 0.05$) for the ELISA analysis.

4. Conclusions

This study examined for the first time the complex proteome of the tritordeum, highlighting the subtle but technologically and nutritionally relevant differences in the protein set of two commercially mainstream tritordeum cvs, namely Aucan and Bulel. The inter-cvs differences observed may be attributable to the different contributions of the *Triticum turgidum* spp. *durum* genome. Our results suggest that attention should be paid in considering all the tritordeum cvs as a *unicum* in terms of protein expression, since in some cases the protein contribution can vary along the genomic characteristics of the *Hordeum* and *Triticum* parents.

In the same way, the first in vitro digestomic analysis carried out on bread baked with tritordeum flour in the present study evidenced that the process of digestion produced different peptidomes, with possible different outcomes in terms of immunoreactivity and allergenicity.

Supplementary Materials: The following are available online, Figure S1: (A) Breads prepared with 100% reference flour; (B) Sliced bread. Tritordeum bread appeared yellower than soft wheat. Tritordeum Baked breads showed a comparable alveolation to soft wheat bread, cv Altamira. Figure S2: Reducing sugar release (expressed as mg of glucose) from 1.5 g of digested breads. Error bars represent the variability over two breads digested in two days and two technical replicates. (Panel A) kinetic of breads duodenal digestion (0–30–60–90 and 120 min); (Panel B) mg of reducing sugar released after 4 h of gastroduodenal digestion of 1.5 g of bread. Figure S3: mg of alpha amino nitrogen determined in 1.5 g of the cooked breads. Bars with different letters are significantly different (p-value < 0.05) and the REGW-F test. Figure S4: Graphical representation of the peptides surviving digestion. Peptides belonging to the same protein region were aligned and the height of the amino acid reflects its abundance in the sequences. The sequence R5-QQPFP sequence was highly repeated in Aucan bread *Triticum* derived peptides, while in Bulel was found highly repeated in *Hordeum*-derived peptides. Sequence logo of unique peptides from Aucan (Panel A,C) and Bulel (Panel B,D) bread duodenal digestomes. The frequency of the sequences is expressed in bits [39]. Figure S5: Coverage of the *Triticum tugidum* spp. *durum* HMW GS (K4N1X7) by digestion-resistant peptides from Aucan bread (Panel A) and Bulel bread (Panel B). The alignment highlights the uniqueness

in several cases is due to the hydrolysis of peptides differing by few amino acids (Supplementary Materials Table S5). Table S1: Nano-LC MS/MS identification of the proteins resistant to the in vitro gastroduodenal digestion of Aucan bread. Table S2: Nano-LC MS/MS identification of the proteins resistant to the in vitro gastroduodenal digestion of Bulel bread. Table S3: list of LCMS/MS identified peptides resistant to in vitro digestion of Aucan bread. Table S4: list of LCMS/MS identified peptides resistant to in-vitro digestion of Bulel bread. Table S5: list of LCMS/MS identified peptides resistant to digestion uniquely identified in Aucan and Bulel breads. This list of peptides was used for the in silico evaluation.

Author Contributions: Conceptualization, C.N., G.D.; methodology, A.D., G.D. and C.N.; validation, C.N. and G.D.; formal analysis, C.N., G.D., A.D., V.L., R.R., G.M. and G.P.; investigation, C.N., A.D. and G.D.; resources, P.F., G.M. and G.P.; data curation, C.N., G.D., A.D. and V.L.; writing—original draft preparation, C.N., G.D.; writing—review and editing, C.N., G.D., M.B., G.P., G.M., R.R., P.F. and V.L.; supervision, C.N., P.F. and M.B.; funding acquisition, P.F. and M.B. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the Regione Piemonte (POR FESR 2014-2020), as a part of the D.E.U.M. Project.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Mass spectrometry raw data are available from the corresponding author upon request. The identifications and elaborations are provided as Supplementary Supporting Materials.

Acknowledgments: The authors thank Tiziana Granato (Project Manager Enzymatic and Automation at R-Biopharm Italia S.R.L.) for kindly supporting with the enzymatic testing. The sponsor did not have any role in study design, data analysis, or reporting of results. The research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors. Chiara Nitride thanks the European Union and MIUR for funding through PON-AIM project code E71G19000110001, proposal code AIM1850344-1.

Conflicts of Interest: The authors declare no conflict of interest.

Sample Availability: Bread samples are available in frozen form upon request to the corresponding author. The grains and flours of compared cultivars can be requested to prof. Massimo Blandino at the University of Turin.

Abbreviations

AAI	alpha amylase inhibitors
ACN	acetonitrile
AGC	automatic gain control
AmBic	ammonium bicarbonate
ANOVA	analysis of variance
cv	cultivar
DTT	1,4-Dithio-D-threitol
ELISA	enzyme-linked immunosorbent assay
EDTA	ethylenediaminetetraacetic acid
FA	formic acid
FDR	false discovery rate
GPC	grain protein content
GS	growth stage
Hch	*Hordeum chilense*
HMW-GS	High-molecular-weight glutenin subunit
IEDB	Immune Epitope Database
IT	injection time
LC-MS/MS	liquid chromatograph-mass spectrometer/mass spectrometry
LMW-GS	Low-molecular-weight glutenin subunit
RSR	reducing sugar release

LMW-GS	Low-molecular-weight glutenin subunit
RSR	reducing sugar release
SDS-PAGE	Sodium dodecyl sulfate-polyacrylamide gel electrophoresis
SGF	simulated gastric fluid
SIF	simulated intestinal fluid
SSF	simulated salivary fluid
TAME	p-toluenesulfonyl-L-arginine methyl ester
TCA	trichloroacetic acid
TFA	trifluoroacetic acid
TPC	total protein content
Tris-HCl	tris(hydroxymethyl) aminomethane hydrochloride

References

1. Poole, N.; Donovan, J.; Erenstein, O. Viewpoint: Agri-nutrition research: Revisiting the contribution of maize and wheat to human nutrition and health. *Food Policy* **2021**, *100*, 101976. [CrossRef] [PubMed]
2. Örgeç, M.; Doğan, C.N.; Ağıl, F.; Şahin, G.; Zencirci, N. Introduction (Chapter 1). In *Book Wheat Landraces*, 1st ed.; Zencirci, N., Baloch, F.S., Habyarimana, E., Chung, G., Eds.; Springer Nature: Cham, Switzerland, 2021; pp. 1–12.
3. Obembe, O.S.; Hendricks, N.P.; Tack, J. Decreased wheat production in the USA from climate change driven by yield losses rather than crop abandonment. *PLoS ONE* **2021**, *16*, e0252067. [CrossRef] [PubMed]
4. Daloz, A.S.; Rydsaa, J.H.; Hodnebrog, Ø.; Sillmann, J.; van Oort, B.; Mohr, C.W.; Agrawal, M.; Emberson, L.; Stordal, F.; Zhang, T. Direct and indirect impacts of climate change on wheat yield in the Indo-Gangetic plain in India. *J. Agric. Food Res.* **2021**, *4*, 100132. [CrossRef]
5. Bento, V.A.; Ribeiro, A.F.S.; Russo, A.; Gouveia, C.M.; Cardoso, R.M.; Soares, P.M.M. The impact of climate change in wheat and barley yields in the Iberian Peninsula. *Sci. Rep.* **2021**, *11*, 15484. [CrossRef] [PubMed]
6. Sagnelli, D.; Chessa, S.; Mandalari, G.; Di Martino, M.; Sorndech, W.; Mamone, G.; Vincze, E.; Buillon, G.; Nielsen, D.S.; Wiese, M.; et al. Low glycaemic index foods from wild barley and amylose-only barley lines. *J. Funct. Foods* **2018**, *40*, 408–416. [CrossRef]
7. Martín, A.; Cabrera, A.; Hernández, P.; Ramírez, M.C.; Rubiales, D.; Ballesteros, J. Prospect for the use of Hordeum chilense in durum wheat breeding. In *Durum Wheat Improvement in the Mediterranean Region: New Challenges*; Royo, C., Nachit, M., Di Fonzo, N., Araus, J.L., Eds.; CIHEAM (Options Méditerranéennes: Série A. Séminaires Méditerranéens; n. 40); Seminar on Durum Wheat Improvement in the Mediterranean Region: New Challenges, 12–14 April 2000; Instituto Agronómico Mediterráneo de Zaragoza: Foggia, Spain, 2002; pp. 111–115.
8. Vaquero, L.; Comino, I.; Vivas, S.; Rodríguez-Martín, L.; Giménez, M.J.; Pastor, J.; Sousa, C.; Barro, F. Tritordeum: A novel cereal for food processing with good acceptability and significant reduction in gluten immunogenic peptides in comparison with wheat. *J. Sci. Food Agric.* **2018**, *98*, 2201–2209. [CrossRef]
9. Alvarez, J.B.; Ballesteros, J.; Sillero, J.A.; Martin, L.M. Tritordeum: A new crop of potential importance in the food industry. *Hereditas* **1992**, *116*, 193–197. [CrossRef]
10. Martín, A.; Alvarez, J.B.; Martín, L.M.; Barro, F.; Ballesteros, J. The development of tritordeum: A novel cereal for food processing. *J. Cereal Sci.* **1999**, *30*, 85–95. [CrossRef]
11. Sánchez-León, S.; Haro, C.; Villatoro, M.; Vaquero, L.; Comino, I.; González-Amigo, A.B.; Vivas, S.; Pastor, J.; Sousa, C.; Landa, B.B.; et al. Tritordeum breads are well tolerated with preference over gluten-free breads in non-celiac wheat-sensitive patients and its consumption induce changes in gut bacteria. *J. Sci. Food Agric.* **2021**, *101*, 3508–3517. [CrossRef]
12. Kakabouki, I.; Beslemes, D.F.; Tigka, E.L.; Folina, A.; Karydogianni, S.; Zisi, C.; Papastylianou, P. Performance of six genotypes of tritordeum compare to bread wheat under east mediterranean condition. *Sustainability* **2020**, *12*, 9700. [CrossRef]
13. Visioli, G.; Lauro, M.; Vamerali, T.; Dal Cortivo, C.; Panozzo, A.; Folloni, S.; Piazza, C.; Ranieri, R. A comparative study of organic and conventional management on the rhizosphere microbiome, growth and grain quality traits of tritordeum. *Agronomy* **2020**, *10*, 1717. [CrossRef]
14. Giordano, D.; Reyneri, A.; Locatelli, M.; Coïsson, J.D.; Blandino, M. Distribution of bioactive compounds in pearled fractions of tritordeum. *Food Chem.* **2019**, *301*, 125228. [CrossRef]
15. Mendis, M.; Simsek, S. Arabinoxylans and human health. *Food Hydrocoll.* **2014**, *42*, 239–243. [CrossRef]
16. Jayachandran, M.; Chen, J.; Chung, S.S.M.; Xu, B. A critical review on the impacts of β-glucans on gut microbiota and human health. *J. Nutr. Biochem.* **2018**, *61*, 101–110. [CrossRef] [PubMed]
17. Brodkorb, A.; Egger, L.; Alminger, M.; Alvito, P.; Assunção, R.; Ballance, S.; Bohn, T.; Bourlieu-Lacanal, C.; Boutrou, R.; Carrière, F.; et al. INFOGEST static in vitro simulation of gastrointestinal food digestion. *Nat. Protoc.* **2019**, *14*, 991–1014. [CrossRef] [PubMed]
18. Kahlenberg, F.; Sanchez, D.; Lachmann, I.; Tuckova, L.; Tlaskalova, H.; Méndez, E.; Mothes, T. Monoclonal antibody R5 for detection of putatively coeliac-toxic gliadin peptides. *Eur. Food Res. Technol.* **2006**, *222*, 78–82. [CrossRef]
19. Immer, U.; Haas-Lauterbach, S. Gluten Detection. In *Molecular Biological and Immunological Techniques and Applications for Food Chemists*; Popping, B., Diaz-Amigo, C., Hoenicke, K., Eds.; John Wiley & Sons, Inc.: Hoboken, NJ, USA, 2009; pp. 359–376, ISBN 9780470637685.

20. Landolfi, V.; D'Auria, G.; Nicolai, M.A.; Nitride, C.; Blandino, M.; Ferranti, P. The effect of nitrogen fertilization on the expression of protein in wheat and tritordeum varieties using a proteomic approach. *Food Res. Int.* **2021**, *148*, 110617. [CrossRef]
21. Li, Y.; Fu, J.; Shen, Q.; Yang, D. High-Molecular-Weight Glutenin Subunits: Genetics, Structures, and Relation to End Use Qualities. *Int. J. Mol. Sci.* **2020**, *22*, 184. [CrossRef]
22. Alvarez, J.B.; Campos, L.A.C.; Martín, A.; Sillero, J.A.; Martín, L.M. Genetic analysis of prolamins synthesised by the H(ch) genome and their effects on gluten strength in hexaploid tritordeum. *Euphytica* **1999**, *107*, 177–184. [CrossRef]
23. Tercero, J.A.; Bernardo, A.; Jouve, N. Encoding genes for endosperm proteins in Hordeum chilense. *Theor. Appl. Genet.* **1991**, *81*, 127–132. [CrossRef]
24. Alvarez, J.B.; Martín, A.; Martín, L.M. Variation in the high-molecular-weight glutenin subunits coded at the Glu-Hch1 locus in Hordeum chilense. *Theor. Appl. Genet.* **2001**, *102*, 134–137. [CrossRef]
25. Mikulíková, D.; Benková, M.; Kraic, J. The Potential of Common Cereals to form Retrograded Resistant Starch. *Czech. J. Genet. Plant. Breed.* **2018**, *42*, 95–102. [CrossRef]
26. Daly, M.; Bromilow, S.N.; Nitride, C.; Shewry, P.R.; Gethings, L.A.; Mills, E.N.C. Mapping Coeliac Toxic Motifs in the Prolamin Seed Storage Proteins of Barley, Rye, and Oats Using a Curated Sequence Database. *Front. Nutr.* **2020**, *7*, 87. [CrossRef] [PubMed]
27. Nguyen, P.Q.T.; Wang, S.; Kumar, A.; Yap, L.J.; Luu, T.T.; Lescar, J.; Tam, J.P. Discovery and characterization of pseudocyclic cystine-knot α-amylase inhibitors with high resistance to heat and proteolytic degradation. *FEBS J.* **2014**, *281*, 4351–4366. [CrossRef] [PubMed]
28. Salcedo, G.; Snchez-Monge, R.; Garca-Casado, G.; Armentia, A.; Gomez, L.; Barber, D. The Cerealα-Amylase/Trypsin Inhibitor Family Associated with Bakers' asthma and Food Allergy. In *Plant Food Allergens*; Blackwell Publishing Ltd.: Oxford, UK, 2003; pp. 70–86.
29. Koehler, P.; Wieser, H.; Konitzer, K. Celiac Disease—A Complex Disorder (Chapter 1). In *Celiac Disease and Gluten*; Koehler, P., Wieser, H., Konitzer, K., Eds.; Academic Press: Cambridge, MA, USA; Elsevier: London, UK, 2014; pp. 1–96, ISBN 9780124202207.
30. Armentia, A.; Sanchez-Monge, R.; Gomez, L.; Baber, D.; Salcedo, G. In vivo allergenic activities of eleven purified members of a major allergen family from wheat and barley flour. *Clin. Exp. Allergy* **1993**, *23*, 410–415. [CrossRef] [PubMed]
31. Foca, G.; Ulrici, A.; Corbellini, M.; Pagani, M.A.; Lucisano, M.; Franchini, G.C.; Tassi, L. Reproducibility of the Italian ISQ method for quality classification of bread wheats: An evaluation by expert assessors. *J. Sci. Food Agric.* **2007**, *87*, 839–846. [CrossRef]
32. Blandino, M.; Vaccino, P.; Reyneri, A. Late-Season Nitrogen Increases Improver Common and Durum Wheat Quality. *Agron. J.* **2015**, *107*, 680–690. [CrossRef]
33. Minekus, M.; Alminger, M.; Alvito, P.; Ballance, S.; Bohn, T.; Bourlieu, C.; Carrière, F.; Boutrou, R.; Corredig, M.; Dupont, D.; et al. A standardised static in vitro digestion method suitable for food – an international consensus. *Food Funct.* **2014**, *5*, 1113–1124. [CrossRef]
34. Abrams, D.; Metcalf, D.; Hojjatie, M. Determination of Kjeldahl Nitrogen in Fertilizers by AOAC Official MethodSM 978.02: Effect of Copper Sulfate as a Catalyst. *J. AOAC Int.* **2014**, *97*, 764–767. [CrossRef]
35. Burton, P.; Lightowler, H.J. The impact of freezing and toasting on the glycaemic response of white bread. *Eur. J. Clin. Nutr.* **2008**, *62*, 594–599. [CrossRef]
36. Bromilow, S.; Gethings, L.A.; Buckley, M.; Bromley, M.; Shewry, P.R.; Langridge, J.I.; Clare Mills, E.N. A curated gluten protein sequence database to support development of proteomics methods for determination of gluten in gluten-free foods. *J. Proteom.* **2017**, *163*, 67–75. [CrossRef] [PubMed]
37. Martens, L.; Vandekerckhove, J.; Gevaert, K. DBToolkit: Processing protein databases for peptide-centric proteomics. *Bioinformatics* **2005**, *21*, 3584–3585. [CrossRef] [PubMed]
38. Juhász, A.; Haraszi, R.; Maulis, C. ProPepper: A curated database for identification and analysis of peptide and immune-responsive epitope composition of cereal grain protein families. *Database* **2015**, *2015*, bav100. [CrossRef] [PubMed]
39. Schneider, T.D.; Stephens, R.M. Sequence logos: A new way to display consensus sequences. *Nucleic Acids Res.* **1990**, *18*, 6097–6100. [CrossRef] [PubMed]

SECTION 3

A novel food from a fermented dairy by-product: Kashk

The purpose of this section is the characterisation and the preliminary allergenicity risk assessment of the Kashk, the product of the fermentation of dairy by-product.

Kashk has a very ancient origin. In fact, is thought to have been invented by Iranian nomads in response to the need of a durable food during their long movements. Traditionally, this dairy product is made from the production waste of home-made yogurt or sour buttermilk, from goat's milk. These dairy wastes are spontaneously fermented and dried under the sun. Today, kashk is also manufactured at an industrial level in a liquid form using as a starting ingredient cow's and sheep's milk.

Milk-derived food allergens

Cow milk allergy affects 0.6% to 3% of children under the age of 6 years, 0.3% of elder children and teens, and less than 0.5% of adults. It is interesting to note that milk allergies typically vanish with infants outgrow. Although most children can consume milk and dairy products as adults, 15% of the affected children remain allergic to their adult lives (Villa et al., 2018).

Caseins

Caseins accounts for approximately 80% of the total protein in milk. In the official WHO/IUIS list of allergens, caseins are all categorized as Bos d 8 even if, basing on the family, they are also categorised as individual allergens (http://allergen.org/index.php; EFSA Panel on Dietetic

Products, 2014). Different casein families are encoded by different genes on the same chromosome: αs1-casein (Bos d 9), β-casein (Bos d 11) representing the predominant caseins (40% and 35%, respectively), while αs2-casein (Bos d 10) and κ-casein (Bos d 12) are less represented (12.5%, each one). In addition, the hydrolysis of β-casein generates three γ-caseins (γ1, γ2, and γ3), which represent the segments 29–209, 106–209, and 108–209 of β-casein, respectively.

The lack of a clearly defined tertiary structure suggests that their allergenic epitopes are linear rather than conformational. This makes their epitope more resistant than those conformational, since not affected by denaturation.

Although these proteins are not affected by denaturing agents, such as urea or heat treatment, in the digestive process, they are extensively degraded by proteolytic enzymes (Villa et al., 2018). In fact, because of their non-compact and flexible structures, caseins are well digested. Interestingly, those patients allergic to linear epitope of milk proteins are more likely to retain their allergenicity with outgrowth than those allergic to conformational ones (Vila et al., 2001). This is probably due to the development of the digestive tract, which in adulthood generate more extreme environment (lower pH).

Interestingly, among milk allergic patients, about 50% of them react with αs1-casein (Natale et al., 2004).

Beside the possible negative effect associated with the casein consumption, positive effect related to the high level of phosphorylation of casein are worthy to be cited. Gastrointestinal digestion of milk protein casein may produce casein phosphopeptides (CPPs). These peptides, containing the amino acid sequence -Pse-Pse-Pse-Glu-Glu- where Pse is a phosphoseryl residue can stabilise calcium and phosphate ions enhancing their bioavailability (Cross et al., 2007). This positively

affects the health of bones and teeth acting as anticariogenic as observed in hard cheese (Ferranti et al., 1997).

Whey proteins

Whey proteins represent 20% of cow's milk protein. According to Monaci et al., (2006), the primary whey allergens are the globular proteins Bos d 5 (β-Lg) and Bos d 4 (α-La), which account for 50% and 25% of the whey protein fraction, followed by other minor components such as Bos d 6 (BSA), immunoglobulins (Ig), and lactoferrin (LF).

The beta-lactoglobulin belongs to the apolipoprotein family and is made up of 162 amino acids, presenting an overall molecular mass of 18.3 kDa. Its tertiary structure is composed of nine antiparallel β -sheets stabilized by two disulphide bridges forming a structure known as a β -barrel (Stanic-Vucinic et al., 2012). The three-dimensional structure of this protein is well-conserved in the apolipoprotein family, even if they have low sequence homology (Chapman & Wood, 2001). Its compact fold makes β-Lg resistant to denaturation and proteolytic hydrolysis (Lönnerdal & Lien, 2003).

Bos d 4 is a single-chain polypeptide of 123 amino acids with a molecular weight of approximately 14 kDa. This protein shows calcium-binding properties and is involved in milk lactose synthesis. Milk allergy patients have a prevalence of α-La - specific IgE ranging from 27.6% to 62.8% depending on the study population (Matsuo et al., 2015).

Notably, the segment 125 to 135 of Bos d 5 (TPEVDDEALEK), is considered the major antigenic site, with an exceptional capacity to survive simulated milk protein digestion and to bind to specific IgE from human sera (Picariello et al., 2010). Despite being present in milk at low levels, Bos d 6 reacted with IgE in 50% of milk allergic patients, which led to its classification as a major allergen.

The protein comprises 582 amino acids with a molecular weight of 66.3 kDa and shows a stable tertiary structure.

Currently, bovine Ig is under study for its potential allergenicity, and their IgE-binding epitopes are not yet identified. Some studies reported that individuals with milk allergies have LF-specific IgE. However, the relevance of this protein's allergenicity is still in question, since these patients also have IgE against the other major milk allergens. To date, no data has been reported on identifying LF IgE-binding epitopes (Villa et al., 2018).

References

Adams, S. L., Barnett, D., Walsh, B. J., Pearce, R. J., HillL, D. J., & Howden, M. E. H. (1991). Human IgE-binding synthetic peptides of bovine β-lactoglobulin and α-lactalbumin. In vitro cross-reactivity of the allergens. *Immunology and Cell Biology*, *69*(3), 191–197. https://doi.org/10.1038/icb.1991.28

Chapman, M. D., & Wood, R. A. (2001). The role and remediation of animal allergens in allergic diseases. *Journal of Allergy and Clinical Immunology*, *107*(3), S414–S421. https://doi.org/10.1067/mai.2001.113672

Cross, K., Huq, N., & Reynolds, E. (2007). Casein Phosphopeptides in Oral Health - Chemistry and Clinical Applications. *Current Pharmaceutical Design*, *13*(8), 793–800. https://doi.org/10.2174/138161207780363086

EFSA Panel on Dietetic Products, N. and A. (NDA). (2014). Scientific Opinion on the evaluation of allergenic foods and food ingredients for labelling purposes. *EFSA Journal*, *12*(11). https://doi.org/10.2903/j.efsa.2014.3894

Ferranti, P., Barone, F., Chianese, L., Addeo, F., Scaloni, A., Pellegrino, L., & Resmini, P. (1997). Phosphopeptides from Grana Padano cheese: nature, origin and changes during ripening. *Journal of Dairy Research*, *64*(4), 601–615. https://doi.org/10.1017/S0022029997002392

Lönnerdal, B., & Lien, E. L. (2003). Nutritional and Physiologic Significance of α-Lactalbumin in Infants. *Nutrition Reviews*, *61*(9), 295–305. https://doi.org/10.1301/nr.2003.sept.295-305

Matsuo, H., Yokooji, T., & Taogoshi, T. (2015). Common food allergens and their IgE-binding epitopes. *Allergology International*, *64*(4), 332–343. https://doi.org/10.1016/j.alit.2015.06.009

Monaci, L., Tregoat, V., van Hengel, A. J., & Anklam, E. (2006). Milk allergens, their characteristics and their detection in food: A review. *European Food Research and Technology*, *223*(2), 149–179. https://doi.org/10.1007/s00217-005-0178-8

Natale, M., Bisson, C., Monti, G., Peltran, A., Perono Garoffo, L., Valentini, S., Fabris, C., Bertino, E., Coscia, A., & Conti, A. (2004). Cow's milk allergens identification by two-dimensional immunoblotting and mass spectrometry. *Molecular Nutrition & Food Research*, *48*(5), 363–369. https://doi.org/10.1002/mnfr.200400011

Picariello, G., Ferranti, P., Fierro, O., Mamone, G., Caira, S., Di Luccia, A., Monica, S., & Addeo, F. (2010). Peptides surviving the simulated gastrointestinal digestion of milk proteins: Biological and toxicological implications. *Journal of Chromatography B*, *878*(3–4), 295–308. https://doi.org/10.1016/j.jchromb.2009.11.033

Stanic-Vucinic, D., Stojadinovic, M., Atanaskovic-Markovic, M., Ognjenovic, J., Grönlund, H.,

van Hage, M., Lantto, R., Sancho, A. I., & Velickovic, T. C. (2012). Structural changes and allergenic properties of β-lactoglobulin upon exposure to high-intensity ultrasound. *Molecular Nutrition & Food Research*, *56*(12), 1894–1905. https://doi.org/10.1002/mnfr.201200179

Vila, L., Beyer, K., Järvinen, K.-M., Chatchatee, P., Bardina, L., & Sampson, H. A. (2001). Role of conformational and linear epitopes in the achievement of tolerance in cow's milk allergy. *Clinical & Experimental Allergy*, *31*(10), 1599–1606. https://doi.org/10.1046/j.1365-2222.2001.01218.x

Villa, C., Costa, J., Oliveira, M. B. P. P., & Mafra, I. (2018). Bovine Milk Allergens: A Comprehensive Review. *Comprehensive Reviews in Food Science and Food Safety*, *17*(1), 137–164. https://doi.org/10.1111/1541-4337.12318

Chapter 6

The protein and peptide fractions of kashk: a traditional Middle East fermented dairy product

Food Research International 132 (2020) 109107

Contents lists available at ScienceDirect

Food Research International

journal homepage: www.elsevier.com/locate/foodres

The protein and peptide fractions of kashk, a traditional Middle East fermented dairy product

Mina Pourjoula[a], Gianluca Picariello[b], Giuseppina Garro[c], Giovanni D'Auria[c], Chiara Nitride[c], Amid Rheza Ghaisari[a], Pasquale Ferranti[b,c,*]

[a] Department of Food Hygiene and Quality Control, Faculty of Veterinary Medicine, Shiraz University, Shiraz, Iran
[b] Institute of Food Science and Technology (ISA), Avellino, Italy
[c] Department of Agricultural Sciences, University of Naples Federico II, Portici, Italy

ARTICLE INFO

Keywords:
Kashk
Fermented milk
Protein digestion
Milk bioactive peptides
Proteomics
Peptidomics
Mass spectrometry

ABSTRACT

Kashk is a typical dairy product of Iran, made from sour milk. It is traditionally produced from buttermilk in a dry, round-shaped form. Today, it is also produced at industrial level in a liquid form starting from fermented milk. We aimed to characterise the kashk proteome and peptidome comparing a traditional product with the industrial using a combination of proteomic approaches including advanced chromatographic and electrophoretic separation technique coupled to tandem mass spectrometry. We identified also phosphorylated casein-derived peptides (CPP) and investigated kashk protein digestibility using a static model of food protein digestion. The molecular characterization, coupled with bioinformatic *in silico* analysis, allowed the identification of potential bioactive peptides.

1. Introduction

Kashk is a fermented dairy product manufactured in Iran, which used to be the main protein source for Middle East nomad populations (Çakır, Coşkun, Akoğlu, İşleyen, Kıralan, & Bayrak, 2009; Kurmann, Rasic, & Kroger, 1992). Today, it is available in dried or liquid form.

Traditionally kashk is produced using by-products from the cheese-making production, like the buttermilk, which are fermented, salted, and dried (ISIRI 2005; Kamber, 2008; Oghbaei & Prakash, 2008). The product is shaped in small balls and used for domestic meal preparation or sold in local markets.

Nowadays, kashk is also produced industrially as a whitish dense liquid similar to sour cream, which is used as an ingredient in traditional Middle East cuisine (ISIRI, 2009). While the traditional product is primarily produced using sheep milk, the industrial kashk is produced starting from yogurt (Noori, Keshavarzian, Mahmoudi, Yousefi, & Nateghi, 2013; Sevgi Kirdar, 2012) from bovine or sheep milk. The effect of additives on the stability of non-fat kashk has been investigated (Ghorban Shiroodi, Mohammadifar, Ghorbani Gorji, Ezzatpanah, & Zohouri, 2012; Ghorbani Gorji, Mohammadifar, & Ezzatpanah, 2010).

From the nutritional standpoint, kashk is rich in protein: its average protein content is 13.7% (dry traditional type) and 8.6% (liquid industrial type). Depending on the geographical region of production, the content of salt can vary from 0.8 to 1.5%. Mineral composition is related to content of calcium and phosphorus, which are essential micronutrients in the human diet. The calcium and phosphorus content of kashk varies in the 0.19–0.25% and 0.18–0.31% ranges, respectively (Jafari, Ghaisari, Khaniki, & Shariatifar, 2015; Soltani & Güzeler, 2012).

The production of traditional kashk from buttermilk can have a technological interest in recovering otherwise wasted dairy by-products. Waste whey from diary industry is a well described source of bioactive milk-derived peptides (De Simone et al., 2009, 2011). These peptides are naturally present in milk as product of protein degradation during processing. Peptides represent a higher-bioavailability form of (essential) amino acids compared to proteins and may exert a multitude of health-promoting actions, including mineral transport, immune-modulatory, anti-hypertensive, antibacterial and antioxidant activity. A detailed characterization of the protein and peptide composition of kashk has never been undertaken so far.

For these reasons, we characterized by 'omic' analysis representative samples of traditional and industrial kashk to generate protein and peptide maps. This study represent the first molecular characterization of kashk. We also investigate the resistance of the kashk bioactive peptides to simulated gastrointestinal digestion as well as monitored the generation of new potentially functional peptide sequences.

* Corresponding author at: Department of Agricultural Sciences, University of Naples Federico II, Parco Gussone, I-80055 Portici (NA), Italy.
E-mail address: ferranti@unina.it (P. Ferranti).

https://doi.org/10.1016/j.foodres.2020.109107
Received 5 October 2019; Received in revised form 13 February 2020; Accepted 15 February 2020
Available online 18 February 2020

2. Materials & methods

2.1. Materials

Three samples of traditional dry and three samples of industrial liquid kashk types were collected at local markets in Shiraz (Iran). Comparative sample of cheese and yogurt were purchased from Italian supermarkets. All the reagents used in the present study were of analytical or higher grade. To prevent undesired peptide hydrolysis, immediately after collection, samples were added with a protease inhibitor cocktail Sigma-Aldrich (St Louis, MO, USA), aliquoted (50 mL), refrigerated, and stored at −20 °C until used. HPLC-grade solvents, water and acetonitrile (ACN), were from Merck (Whitehouse Station, NJ, USA). Pepsin, trypsin, chymotrypsin, dithiothreitol (DTT), iodoacetamide (IAA), Tris-HCl, urea, guanidine chloride, sodium citrate, ammonium bicarbonate, phospholipid preparations, trichloroacetic acid (TCA), trifluoroacetic acid (TFA), p-toluene-sulfonyl-L-arginine methyl ester (TAME), sodium tetraborate were purchased from Sigma-Aldrich (St Louis, MO, USA). Plasmin was purchased from Boehringer (Mannheim, West. Germany). Reagents for electrophoresis analysis were from Bio-Rad (Milan, Italy). Brush border membrane (BBM) enzymes were purified from porcine jejunum, as previously detailed (Picariello et al., 2015). The casein fraction was precipitated from skimmed bovine and ovine milk at pH 4.6 with sodium acetate/acid acetic buffer as described by Aschaffenburg and Drewry (1959). The sodium caseinate precipitate was washed twice with the buffer, freeze-dried and stored at −20 °C before use. The supernatant, containing the whey proteins, was dialysed in tubing with molecular cut off 12,000–14,000 Da and subsequently freeze-dried and stored at −20 °C before use. The aminopeptidase activity was determined by colorimetric assay using p-nitroaniline as the substrate, while the total activity of BBM peptidases was assayed by HPLC using angiotensin I as a substrate (Picariello et al., 2015).

2.2. Preparation of the protein and peptide fractions

Peptide extraction was carried out solubilizing 1 g of kashk sample in either 20 mL of water or in 20 mL of sodium citrate (0.2 M, pH 8.0). Casein was isoelectrically precipitated at pH 4.6 from both water and citrate soluble fractions, according to the Aschaffenburg's procedure (Aschaffenburg & Drewry, 1959). The precipitated casein fraction and the soluble whey proteins/peptides fraction were stored at −20 °C until analysis.

The pH 4.6-soluble fraction was lyophilised and re-suspended in 6 mL of 0.1% (v/v) aqueous TFA. The samples were desalted by solid phase extraction using a Sep-pak C18 cartridge (Waters, Milford, MA), previously equilibrated in 0.1% TFA and eluted with 70/30/0.1 ACN/water/TFA (v/v/v). The peptide isolates were dried using a Savant concentrator (Speed-Vac, Milan, Italy) and stored at −20 °C for structural analysis. The soluble phosphopeptide fraction was prepared from the kashk water-soluble extract following the procedure described by Ferranti, Barone, et al. (1997).

2.3. Sodium dodecyl sulphate - polyacrylamide gel electrophoresis (SDS–PAGE) analysis

Protein samples were dissolved in a 9 M urea solution (0.2:1, w:v) containing 1% β-mercaptoethanol. The solubilised proteins were diluted (1:1, v:v) with a denaturing solution (0.062 M Tris-HCl, 2% (w/v) SDS, 5% (v/v) β-mercaptoethanol, pH 6.8). Samples were heated at 100 °C for 5 min.

Protein solution (10 μl) was loaded onto a precast 12% polyacrylamide gel (Bio-Rad, Richmond, CA, USA). The running buffer was 192 mM glycine, 25 mM Tris and 0.1% SDS. Analysis was carried out at room temperature and constant voltage (100 V). Proteins were visualized using Coomassie blue staining (Bio-Safe Coomassie Stain G-250, Bio-Rad). An Amersham Low MW Calibration Kit (14.4–97 kDa, GE Healthcare UK Limited, UK) was used as the molecular weight (MW) standard. The gels were scanned and processed using the LABScan software 3.00 (Amersham Bioscience, Uppsala, Sweden).

2.4. Analysis by isoelectric focusing and immunoblotting

Proteins were solubilized in 9 M urea as described in Section 2.3. Isoelectric focusing electrophoretic analysis of the casein fraction was carried out on ultra-thin layer polyacrylamide gel (0.25 mm), prepared as described in the EU Official Journal No. 37/2001. The pH gradient was in the range of 2.5–6.5 and was obtained by mixing Ampholine (GE Healthcare, UK) 2.5–5.0, 4.5–5.4, and 4.0–6.5 in the ratio 1.6:1.4:1 (v/v/v). The gels were stained using Coomassie Brilliant Blue (CBB) G-250 as described by Neuhoff, Arold, Taube, and Ehrhardt (1988). Proteins separated by UTLIEF, were transferred by capillary diffusion from the gels onto a nitrocellulose membrane (0.45 μm, Trans-Blot, Bio-Rad, Richmond, CA, USA). Immunostaining was carried out using polyclonal antibodies against bovine β-casein (β-CN), β-lactoglobulin (β–Lg) and α-lactalbumin (α–La) produced by Primm (Milan, Italy), as primary antibodies and horseradish peroxidase-labelled goat anti-rabbit IgG polyclonal antibodies as secondary reagents (BioRad, Hercules, CA), according to the procedure described by Chianese, Mauriello, Moio, Intorcia, and Addeo (1992).

2.5. Assay of plasmin activity on whole casein fraction

Whole casein isolated from the kashk samples by isoelectric precipitation (70 mg) was dissolved in 1 mL of 0.05 M sodium tetraborate buffer, pH 8. The suspension was treated with 25 μl plasmin suspension (5U/mL) (EC 3.4.21.7; Boeringer, Germany) and incubated at 37 °C. Samples (200 μl) were collected after 60 min and the reaction stopped by addition of an equal volume of 24% trichloroacetic acid in water. Samples were centrifugation at 3000 g for 5 min and the pellet was prepared for electrophoretic analysis as described in Section 2.3.

2.6. Protein and peptide MALDI TOF MS analysis

Protein and peptide extracts were analyzed by Matrix Assisted Laser Desorption Ionisation-Time of Flight mass spectrometry (MALDI TOF MS) using a Voyager DE PRO MALDI-TOF mass spectrometer (PerSeptive BioSystems, Framingham, MA) equipped with a N_2 laser (λ = 337 nm), using α-cyano-4-hydroxy-cinnamic acid (for peptides) or sinapinic acid (for proteins) as the matrix (10 mg mL^{-1} in 50% ACN, v/v, containing 0.1% TFA). Mass spectra were acquired in the reflector or linear positive ion mode, for peptides or proteins, respectively, using the Delay Extraction (DE) technology. The accelerating voltage was 20 kV. External mass calibration was performed using a commercial mixture of standard peptides (PerSeptive Biosystems, Framingham, MA). A resolution of ≥8.000 was calculated in the working mass range. Raw data were analyzed using the Data Explorer 4.0 software furnished with the spectrometer. Post-source decay (PSD) MS analysis was carried out after isolation of the precursor ions using a timed ion selector set at an ion gate width of 1 Da. The PSD mass spectra were divided into seven segments; the laser power and the guide wire voltage were varied for each segment to optimize fragmentation and data collection. Approximately 200 laser shots were acquired for each segment. Fragmented ions were refocused onto the final detector by stepping down the voltage applied to the reflector. Finally, the individual segments were stitched together using the software purchased with the instrument. Peptide mass fingerprinting (PMF)-based identifications were carried out interrogating the National Center for Biotechnology Information (nrNCBI) and Swiss-Prot/TrEMBL databases with Mascot (Matrix Science, London, UK) and Protein Prospector MS-FIT (http://prospector.ucsf.edu/) search engines. Mass tolerance of 0.3 Da, variable pyro-glutamic acid formation at N-terminal Gln and possible

methionine oxidation were set as search parameters. Searches were taxonomically restricted to Mammalian. Probability MOWSE scores were automatically calculated by the search engines; only protein candidates with score higher than the random match region ($p < 0.05\%$) were considered. The identification of top scores of protein candidates was validated by manual peptide mass mapping. BIOPEP database of bioactive peptides (www.uwm.edu.pl/biochemia/index.php/pl/biopep) was used to screen potential bioactive sequences.

2.7. Reverse phase- high performance liquid chromatography (RP–HPLC) analysis of kashk proteins

Reversed phase (RP)-HPLC analysis of casein was carried out using a C_4 Vydac 2.1 mm i.d. column (Hesperia, CA, USA). Solvent A was 0.1% in water (v/v). After 5 min of isocratic elution using 30% solvent B (0.1% TFA in ACN, v/v), a 30–50% gradient ramp was applied over 60 min at a flow rate of 0.200 mL/min. The HPLC chromatograph was an HP 1100 Agilent (Palo Alto, CA, USA) modular system equipped with a diode array detector. The column effluents were monitored by UV detection at $\lambda = 220$ and 280 nm.

2.8. In vitro gastrointestinal-BBM digestion of kashk

Gastro-intestinal digestion (GID) of kashk was carried out on 0.1 g sample according to the Infogest harmonized protocol (Minekus et al., 2014), integrated with an additional step simulating the action of intestinal BBM peptidases (Picariello et al., 2015). The aliquots for the subsequent intestinal reactions were directly supplemented with BBM, omitting the 5 min boiling. Peptidase activity was estimated determining the aminopeptidase N (APN) and dipeptidyl IV (DPP-IV) activity by spectrophotometric assays using L-leucine p-nitroanilide and Gly-Pro-4 nitroanilide as the substrates, respectively (Claude et al., 2019). The APN and DPP-IV activity was 879 μU/μL and 152 μU/μL, respectively. These values were roughly consistent with the value of 1018 μU/μL total peptidase activity, determined by RP-HPLC using angiotensin-I (Sigma) as the substrate (Picariello et al., 2015). The peptide digests arising from gastro-duodenal digestion were two-fold diluted with 0.1 M sodium phosphate buffer pH 7.2, supplemented with 100 μU BBM/100 μg peptides and incubated at 37 °C up to 6 h.

The amount of the peptide substrate was estimated according to the original protein material, determined with the modified micro-Lowry assay (kit from Sigma), hence containing at this stage free amino acids too. At present, there is not an established consensus about the conditions for digestion with intestinal BBM peptidases. Thus, the peptidase-to-substrate ratio was chosen according to conditions with previously assessed physiological consistence (Shan et al., 2002). Aliquots were collected and analyzed at intermediate incubation times. The peptides arising from the gastrointestinal digestion were separated by RP-HPLC performed using a 2.0 mm i.d. × 250 mm, C18, 5 μm reverse-phase column (Phenomenex, Torrance, CA, USA) with a flow rate of 0.2 mL/min using the same chromatographic system as above. The column was equilibrated at 5% solvent B. Peptide separation (25 μg for each analysis) was carried out with a 5–60% solvent B gradient over 60 min. The column effluent was monitored by UV detection (220 and 280 nm) and each peak was manually collected.

2.9. Liquid chromatography- high resolution – tandem mass spectrometry (LC-HR-MS/MS analysis)

LC-MS analysis was performed using a Q-Exactive Orbitrap mass spectrometer (Thermo Scientific, San Jose, CA, USA), online coupled with an Ultimate 3000 ultra-high performance liquid chromatography instrument (Thermo Scientific). Samples were resuspended in 0.1% (v/v) formic acid solution, loaded through a 5 mm long, 300 μm id pre-column (LC Packings, USA) and separated by an EASY-Spray™ PepMap C18 column (2 μm, 15 cm × 75 μm) 3 μm particles, 100 Å pore size (Thermo Scientific). Eluent A was 0.1% formic acid (v/v) in Milli-Q water; eluent B was 0.1% formic acid (v/v) in ACN. The column was equilibrated at 5% B. Peptides were separated applying a 4–40% gradient of B over 60 min. The flow rate was 300 nL/min. The mass spectrometer was operated in data-dependent mode and all MS1 spectra were acquired in the positive ionization mode with an m/z scan range of 350 to 1600. Up to 10 most intense ions in MS1 were selected for fragmentation in MS/MS mode, with 10 s dynamic exclusion. A resolving power of 70,000 full width at half maximum (FWHM), an automatic gain control (AGC) target of 1×10^6 ions and a maximum ion injection time (IT) of 256 ms were set to generate precursor spectra. MS/MS fragmentation spectra were obtained at a resolving power of 17,500 FWHM. Ions with one or more than six charges were excluded. Spectra were processed using the Xcalibur Software 3.1 version (Thermo Scientific). Mass spectra were elaborated using the Proteome Discoverer 2.1 software (Thermo Scientific), restricting the search to *Bos taurus, Ovis aries* and *Capra hircus* databases (November 2017). The search was also expanded to mammalian database, extracted from the NCBI (downloaded on November 2017). Database searching parameters for identification of the identification of peptides in simulated gastrointestinal digests and peptides isolated from kashk samples were: Met oxidation and pyroglutamic for N-terminus Gln as variable protein modifications; a mass tolerance value of 8 ppm for precursor ion and 0.01 Da for MS/MS fragments; no modification of cysteine residues included and no proteolytic enzyme selected. The false discovery rate and protein probabilities were calculated by Target Decoy PSM Validator working between 0.01 and 0.05 for strict and relaxed searches, respectively.

3. Results

3.1. The proteome of kashk

3.1.1. Electrophoretic analysis

Fig. 1 shows the electrophoretic profile of the total proteins isolated from exemplificative samples of traditional or industrial kashk. We compared kashk proteins to isolated bovine and ovine casein and whey proteins as well as proteins isolated from Grana Padano, Pecorino cheese and yogurt. The traditional and the industrial kashk samples showed the presence of intact casein fractions with the αs2-casein and

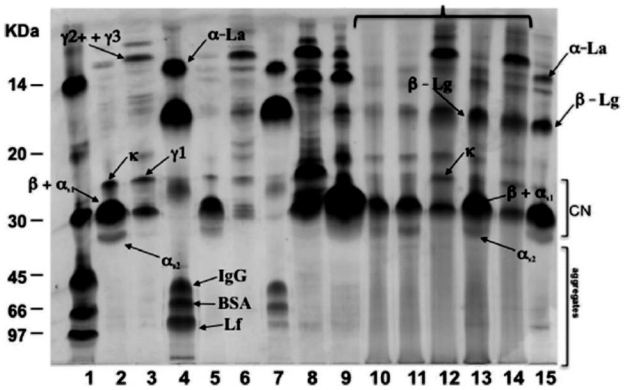


Fig. 1. SDS-PAGE analysis of the total proteins isolated from traditional and industrial kashk compared to yogurt, Pecorino and Grana cheese. Bovine and ovine casein and whey proteins, in house isolated, were used as reference. Kashk samples were also subjected to plasmin (PL) action, as indicated. (1) Molecular weight markers; (2) Bovine CN; (3) Bovine CN + PL; (4) Bovine whey; (5) Ovine CN; (6) Ovine CN + PL; (7) Ovine whey; (8) Grana Padano cheese; (9) Pecorino cheese; (10) Traditional kashk; (11) Traditional kashk fraction insoluble at pH 4.6; (12) Traditional kashk fraction insoluble at pH 4.6 + PL; (13) Industrial kashk; (14) Industrial kash + PL; (15) Yogurt. Gel was stained using Coomassie R-G250.

the κ-casein having respectively the highest and the lowest (about 19 kDa) molecular mobility, as expected.

The N-terminal region (1–105) of κ-casein, arising from chymosin-mediated splitting of κ -casein, which is expected at an apparent molecular mobility of ~15 kDa, was absent in both traditional and industrial Kashk samples, confirming that these products do not derive from rennet coagulation, while they are more similar to buttermilk and yogurt.

The β-Lg (18 kDa) was detected in both kashk types analysed appearing relatively more abundant in industrial one. Several faint bands of α-La (14 kDa) were detected in both products. The presence of denatured and aggregated species of whey proteins is due, most likely, to intense heat treatments used during kashk production.

It was noteworthy the absence of immunoglobulin (Ig), bovine serum albumin (BSA) and lactoferrin (Lf) bands in the pH 4.6-soluble fractions of either traditional or industrial samples (data not shown). Whey proteins have different thermal sensitivity, in the order: Ig > BSA > β-Lg > α-La (Jelen & Rattray, 1995). Thermal treatment up to 90 °C might have caused the complete denaturation of Ig and BSA while the β-Lg and α-La, although in lesser amounts, were still detectable in the nitrogen fraction insoluble at pH 4.6. Likely, these proteins were involved in the formation of high molecular weight aggregates held together by disulphide bonds, then reduced under electrophoretic conditions.

The protein profile of kashk samples was quite similar to the typical profile of a yogurt (Fig. 1, lane 15). Analysis of the same samples subjected to *in vitro* plasmin hydrolysis confirmed that kashk is a product of acid coagulation in the absence of γ-caseins. The profiles of the other kashk samples examined in this study did not exhibit appreciable differences (data not shown).

3.1.2. Isoelectric focusing (IEF) with immunodetection analysis of pH 4.6-insoluble nitrogen fraction from kashk

To further investigate the composition of the pH 4.6-insoluble protein fraction, samples were analyzed by IEF with either Blue Coomassie staining or immunodetection with specific polyclonal antibodies against the single β-CN, β-Lg and α-La (Fig. 2). By comparison with whole standard casein and whey proteins, it was possible to clearly detect β-CN and its related species at both higher and lower isoelectric point (Chianese et al., 1995; Mamone et al., 2013).

Notably, immunodetection revealed the presence of bovine proteins in industrial sample only. In contrast, exclusively ovine/goat milk proteins occurred in the traditional product. Bovine β-Lg A and B were identified in both bovine or ovine whey with the use of specific antibodies; in kashk samples (particularly in industrial type) components at higher isoelectric point (pI) than native proteins were detected, strengthening the hypothesis of the formation of aggregates between denatured caseins and whey proteins in consequence of intense heat treatment. Formation of these aggregates in kashk would also explain the detection of whey proteins in the casein fraction.

3.1.3. HPLC and MALDI TOF MS analysis of the insoluble fraction

The HPLC separation of the 4.6-insoluble fraction of kashk in comparison with standard milk caseins confirmed the presence of intact casein fractions (data not shown). The presence of large casein fragments was also indicated by the presence of characteristic γ-casein peaks. HPLC analysis of the complementary pH 4.6-soluble fraction in comparison with sheep whey proteins revealed an almost identical profile for both traditional and industrial kashk types with the almost total disappearance of native whey proteins, which were possibly made insoluble by heat denaturation. Structural protein analysis carried out by MALDI-TOF MS (data not shown) confirmed electrophoretic and HPLC data further substantiating the presence of both bovine and ovine proteins in the industrial sample.

β-Lg was found in both samples, more abundant in industrial khask. Adducts arising from non-enzymatic glycosylation (+324, lactose) were also present, due to the intense heat treatment (Maillard reaction) producing adducts at lower pH, thus confirming electrophoretic evidence.

3.2. The peptidome of kashk

3.2.1. Structural analysis of the peptide fraction

The peptide fraction of kashk samples was characterized using advanced peptidomic procedures. In particular, MALDI MS analysis of the citrate extract of kashk (Fig. 3 for representative sample spectra) allowed clearly differentiating the industrial from the traditional types. As shown in the figure, where only the peptides giving the prominent signals are annotated, the main peptides derived from the N- and C-terminal regions of the most abundant proteins αs1-CN and β-CN. These peptides also included phosphorylated sequences, such as those of the N-terminal peptides 1–28 and 1–32 of mature β-CN.

Interestingly, peptide sequences detected in the traditional kashk sample corresponded to the ovine/caprine milk proteins, whereas in the industrial sample the peptides from bovine counterparts dominated, identifying the diversity of the species of the milk employed in the two product types. Exclusively in traditional kashk, intense signals of the peptide α_{s1}-CN (1–23) were detected. This peptide is normally originated by chymosin action on casein during cheese making, and remains in the whey fraction after milk coagulation. Therefore, this finding suggests the use of residual cheese whey as one of the ingredients. Alternatively, the peptide could have been produced by the enzymatic action of some bacterial protease.

In any case, the enormous complexity disclosed by MALDI MS analysis of the kashk samples peptidome required a high-throughput analysis by HPLC-ESI MS/MS to be fully defined. Water and citrate extracts exhibited a very different peptide composition, due to the different specificity of the extraction systems, with pure water being a more effective solvent than citrate. The Venn diagram in Fig. 4 provides a graphical represent of the total number of peptides identified by high-resolution tandem mass spectrometry analysis in the traditional and

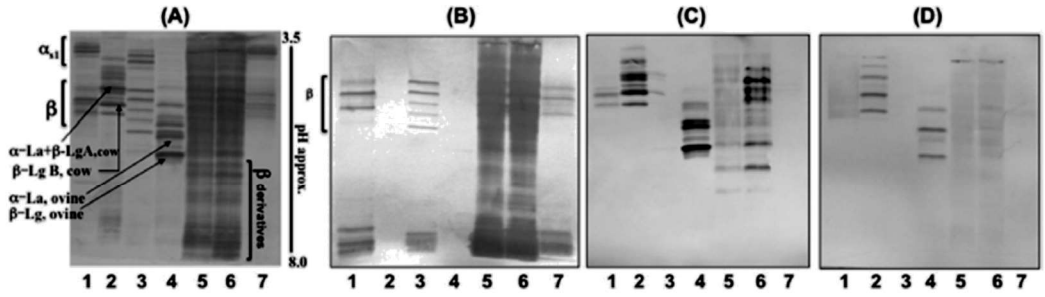


Fig. 2. Ultra-Thin Layer Polyacrylamide Gel Isoelectric Focusing (UTLIEF) analysis of pH 4.6-insoluble nitrogen fraction from traditional and industrial kashk. Coomassie brilliant blue staining (panel *A*) and immunoblotting with polyclonal antibodies against β-CN (panel *B*), β-Lg (panel *C*), α-La (panel *D*). (1) Bovine casein; (2) Bovine whey; (3) Ovine casein; (4) Ovine whey; (5) Traditional kashk; (6) Industrial kashk; (7) Yogurt.

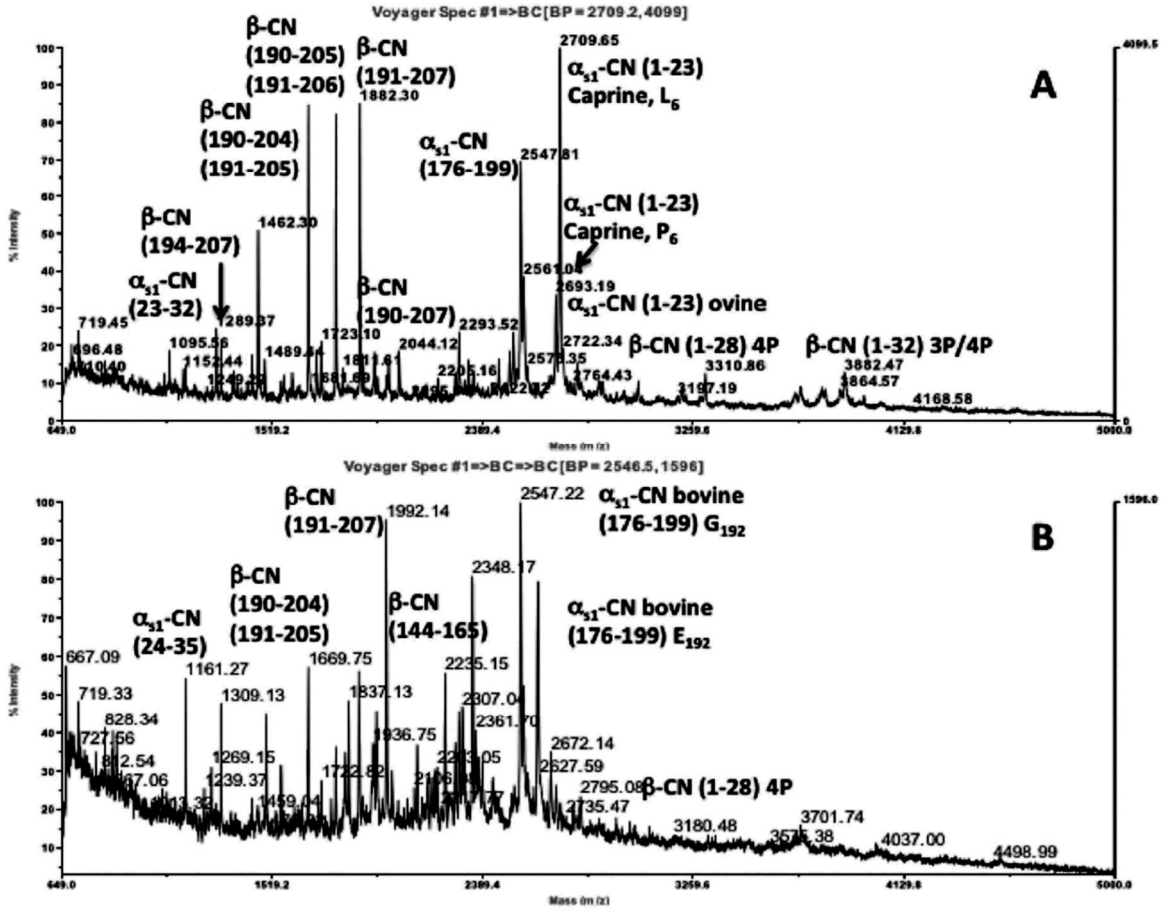


Fig. 3. Peptides detected in the citrate soluble peptide fraction of traditional (A) and industrial (B) kashk samples.

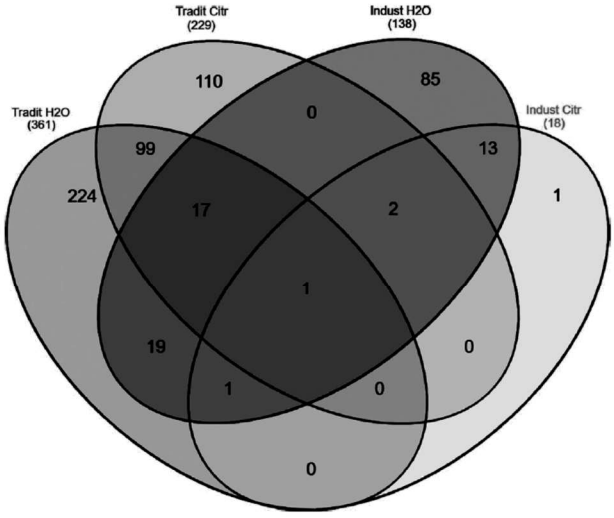


Fig. 4. Venn diagram showing the overlapping peptides identified by high-resolution tandem mass spectrometry analysis in industrial and traditional kashk extracted using water and citrate. The Venn was generated using InteractiVenn (Heberle, Meirelles, da Silva, Telles, & Minghim, 2015). Peptides identified in both natural and modified forms, carrying oxidation of the methionine or the formation of the pyroglutamic acid (variable modifications) were counted only once by the software.

industrial kashk, solubilized using the two extraction methods. Results are summarized in Supplementary material S1-S2 (citrate extract) and S3-S4 (water extract). The analysis of traditional kashk samples allowed the identification of a total number of **414** and **241** peptides in water and citrate extracts, respectively. A minor number of peptides was identified in industrial kashk samples. These peptides arise from the proteolytic process occurring during kashk production and maturation. Primarily are casein-derived peptides, mainly from the two most abundant α_{s1}-CN and β-CN, and from several whey proteins. The list includes peptides with potential bioactivity, for instance the N-terminal region of β-casein possessing Ca^{2+} transport, immune-stimulating and antimicrobial activity (Pizzano et al., 2000). Similarly, a high number of peptides from the β-CN C-terminus has been found, associated with immomodulating activity (Bonomi et al., 2011). In Supplementary material S5-8, the map of the water and citrate soluble peptides detected for the two main proteins α_{s1}-CN and β-CN is aligned along the respective parent protein sequence. In the figures, it can be observed that the protein degradation proceeds from production of large peptides to small sized ones. Interrogation of the database BIOPEP allowed to predict potential bioactivity of peptides. In fact, in the peptide fraction extracted in water, a total of **195** (traditional) and **67** (industrial) peptides were identified with reported potential ACE- inhibitor, opioid agonist or immune-modulating activities (supplementary material S5-8, Tables 1 and 2).

3.2.2. Structural analysis of kashk phosphopeptides

A particularly important class of bioactive peptides identified in various kind of dairy products is constituted by casein-derived

Table 1
CPPs isolated from industrial kashk by selective precipitation as barium salts (S: phosphoserine residue; P: phosphate group).

CPP industrial kashk				
MW (measured)	MW (expected)	Peptide	Species	Sequence
979.09	979.23	β(16–22)3P	bovine	LSSSEES
1481.22	1481.21	αs1(64–74)4P	bovine	SISSSEEIVPN
1681.12	1682.32	αs1(62–74)4P	bovine	AESISSSEEIVPN
1787.33	1788.12	β(15–28)3P	bovine	SLSSSEESITRINK
2069.45	2069,81	αs1(61–78)5P	bovine	EAESISSSEEIVPNSVEQK
2132.54	2131.43	αs1(64–79)5P	bovine	SISSSEEIVPNSVEQKH
2461.28	2461,54	αs1(61–79)5P	bovine	EAESISSSEEIVPNSVEQKH
3076.222	3074.99	β(4–25)4P	bovine	EELNVPGEIVESLSSSEESITR
3450.20	3451.34	β(2–29)4P	bovine	ELEELNVPGEIVESLSSSEESITRINKK
3478.55	3478.99	β(1–28)4P	bovine	RELEELNVPGEIVESLSSSEESITRINK
3602.20	3601.34	β(1–29)4P	bovine	RELEELNVPGEIVESLSSSEESITRINK
3824.51	3824.99	β(1–31)4P	ovine/caprine	REQEELNVVGETVESLSSSEESITHINKKIE
4542.33	4541.34	αs1(35–73)7P	ovine/caprine	ENINELSKDIGSESIEDQAMEDAKQMKAGSSSSSEEIV
4669.12	4671.99	αs1(34–73)7P	ovine/caprine	NINELSKDIGSESIEDQAMEDAKQMKAGSSSSSEEIV
4766.43	4766.34	αs1(34–74)7P	ovine/caprine	NINELSKDIGSESIEDQAMEDAKQMKAGSSSSSEEIVP
5422.98	5422.54	αs1(34–77)7P	ovine/caprine	NINELSKDIGSESIEDQAMEDAKQMKAGSSSSSEEIVPNSA
5552.02	5551.97	αs1(34–78)7P	ovine/caprine	NINELSKDIGSESIEDQAMEDAKQMKAGSSSSSEEIVPNSAE
5680.55	5679.99	αs1(34–79)7P	ovine/caprine	NINELSKDIGSESIEDQAMEDAKQMKAGSSSSSEEIVPNSAEQ
5767.20	5767.34	αs1(34–80)7P	ovine/caprine	NINELSKDIGSESIEDQAMEDAKQMKAGSSSSSEEIVPNSAEQK

phosphopeptides (CPPs), which have been associated with various biological and nutritional properties, including calcium ion transport and enhancement of mineral bioavailability (FitzGerald, 1998; Koide, Itoyama, Fukushima, Miyazawa, & Kuwata, 1991; Liu et al., 2017).

To investigate the presence of CPPs in kashk, CPPs were prepared by selective precipitation with barium salts. MALDI-TOF MS analysis performed in positive and in negative ion mode, revealed the presence of numerous CPPs (Tables 1 and 2) in the isolated casein fraction of both traditional and industrial kashk. The industrial kashk appeared more complex that the traditional in the presence of a larger number of CPPs generated from both bovine and ovine caseins. Further *in vitro* bioactivity screenings are required to confirm functional and the potential nutritional properties related to the identified components.

3.3. Evaluation of kashk protein digestibility using an in vitro static model

A static and multi-phasic *in vitro* model of the gastro intestinal digestion (GID), consisting of sequential simulated gastric, duodenal and intestinal (peptidases from human intestinal brush border membrane enzymes, BBM) phases (Picariello et al., 2013, 2015), has been adapted and applied to identify both the casein- and whey protein-derived peptides particularly resistant to the GID proteolysis. In this study, we aimed at evaluating the process of digestion of kashk and the formation of resistant peptides with possible biological activity. To identify the resistant domains of the milk proteins, the digesta derived from any step of digestion, including the complete gastroduodenal–intestinal BBM degradation, were characterized by complementary MS-based techniques.

In Fig. 5 the MALDI spectra of the gastric (A), pancreatic (B) and intestinal (C) steps of GID of both traditional (1) and industrial (2) kashk are shown. In the first two phases, low molecular weight peptides were generated, which were different between the two kashk types. After the BBM digestion phase, only a few resistant peptides were identified, which were virtually the same in the two sample types of kashk, indicating for both kashk types high levels of digestibility.

HPLC-ESI MS analysis allowed detection of minor amounts of intact α-La at the end of the simulated GID process including BBM enzymes, while the β-Lg was almost completely hydrolyzed. However, several peptides mainly arising from casein and β-Lg hydrolysis were also found after BBM digestion.

4. Discussion

During the various stages of kashk production, as for other fermented dairy products, peptides of different size are released from the milk proteins by combined action of indigenous milk enzymes, such as plasmin, and microbial enzymes (Hartmann & Meisel, 2007). This family of components forms the so-called peptidome, already described for many types of cheese and other fermented products (Addeo et al., 1992, 1994; De Filippis, Genovese, Ferranti, Gilbert, & Ercolini, 2016; Ferranti, Barone, et al., 1997; Gobbetti, Ferranti, Smacchi, Goffredi, & Addeo, 2000; Hebert et al., 2008). The present study is the first molecular characterization of the complex proteome and of the peptidome of kashk, a regional dairy product of Middle East and Iran. Proteomic and peptidomic patterns of kashk reflected, as expected, the nature of the raw material and the process-induced modifications. Thus, this study assumes particular relevance considered the very peculiar manufacturing of kashk in the use of byproducts/waste as main ingredients in the circular economy idea.

Electrophoretic data suggest that the proteolytic process of both the

Table 2
CPPs isolated from traditional kashk by selective precipitation as barium salts (S: phosphoserine residue; P: phosphate group).

CPP traditional kashk				
MW (measured)	MW (expected)	Peptide	Species	Sequence
1279.71	1279.20	β(15–24)3P)	ovine/caprine	SLSSSEESIT
1359.78	1359.44	β(15–24)4P	ovine/caprine	SLSSSEESIT
1569.32	1569.33	α_{s2}(126–137)2P	ovine/caprine	EQLSTSEENSKK
1725.63	1725.44	α_{s2}(125–137)2P	ovine/caprine	REQLSTSEENSKK
3398.44	339,921	β(1–28)4P	ovine/caprine	RELEE LNVPGETVESLSSSEESITHINK
3478.44	3478.21	β(1–28)3P	ovine/caprine	RELEE LNVPGETVESLSSSEESITHINK

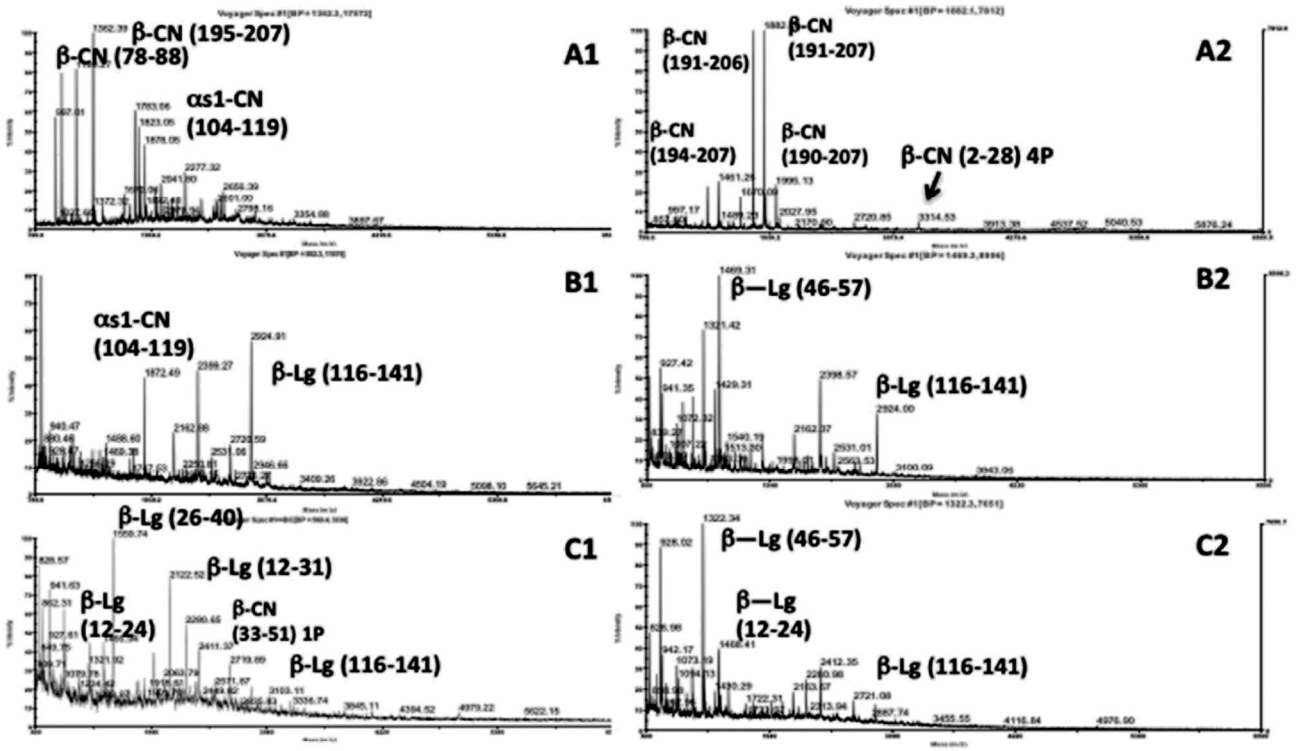


Fig. 5. MALDI TOF mass spectra of the gastric (panel A), pancreatic (panel B) and intestinal (panel C) steps of GID of both traditional (1) and industrial (2) kashk samples.

traditional or industrial kashk follows a pathway similar to that shown for several cheeses, where caseins and whey proteins are degraded by primary action of endopeptidases producing large-sized fragments and shorter peptides primarilly from the N- and C-terminals of the major proteins (De Filippis et al., 2016; Ferranti, Itolli, et al., 1997; Hebert et al., 2008). This evidence has been fully confirmed by tandem mass spectrometry analysis. Large fragments are in turn reduced to an enormous number of shorter peptides by the secondary action of proteases and by-peptidases, probably of microbial origin. They also include amino- and carboxy-peptidases, as indicated by the numerous peptides differing only by a single amino acid residue at the N- or C-terminal side, identified by MS data (supplementary material S5-8). The α_{s1}-CN and β-CN-derived peptides, with molecular mass up to 5 kDa, dominated over those derived from other CNs and whey proteins in both the traditional and industrial products. This was consistent with the results of previous studies concerning other dairy products including yogurt and cheese (Chianese et al., 1997; Ferranti, Itolli, et al., 1997).

The different source of milk used, the other ingredients and the process employed for the production of traditional and industrial kashk led to products with very diverse proteome and peptidome (different proteolytic patterns, animal species peptides, protein Maillard reaction adducts). The peptidomic approach enables differentiating products with similar denomination but very different in terms of composition and nutritional characteristics (Saavedra, Hebert, Minahk, & Ferranti, 2013). Although only a limited number of samples of different product types (traditional vs. industrial one) were analyzed in detail, due to a limited availability of samples, the changes occurring in the oligopeptide fraction of the kashk samples gave precise indications on the production of kashk peptides. The traditional kashk, made from by-products of cheese and yogurt production, is an extremely rich source of milk proteins and milk derived peptides and particulary CCPs. These peptides have been shown to have anticariogenic properties and to prevent enamel demineralisation (Aimutis, 2004; Grenby et al., 2001). While it is debated whether they are effectively relased by drinking milk, it has been largely proved that they are abundant in ripened cheese (Ferranti, Itolli, et al., 1997) and they have been also recently identified in blood of subjects after consumption of Parmigiano Reggiano cheese (Caira et al., 2016). This study has evidenced in both kashk types, a variety of naturally occurring CPPs, which are ready bioavailabilee for consumers.

Interestingly, deep peptidomic analysis showed the presence in kashk of a number of peptides neither derived from casein nor from the major whey proteins, α-La and β-Lg. These peptides derive from proteolysis of the so-defined 'minor' or 'hidden' milk proteome (Liao, Alvarado, Phinney, & Lönnerdal, 2011a, 2011b; Picariello et al., 2012). This is among the first reports of peptides from minor milk proteins in a dairy product. The potential biological activity of these peptides is yet be explored (Mukhopadhy & Sweeney, 2016). The information obtained could drive research in the utilisation of kashk for the development of dietary supplements for functional foods, food integrators and of drugs for pharmaceutical industry.

In this study, the process of gastro-duodenal and intestinal degradation was investigated using a static model of the human digestion. The production of potential bioactive peptides or related precursor was evidenced. Among the limited number of large-sized peptides surviving digestion, the β-Lg (116–141) (MW 2922 Da) was detected. This peptide, as well as other milk protein-derived ones, had been already identified in a previous study of simulated GID of milk proteins based on the use of a Caco-2 monolayer model (Picariello et al., 2013).

Noteworthy, in both final peptide digests, we did not observe the inner shorter peptide 125–135 of β–Lg, previously reported as the main peptides surviving simulated milk protein digestion and considered a probable epitopic components related to food allergenic reaction (Picariello et al., 2010). The absence of this peptide is not fully surprising considering that kashk is produced from fermented rather than from native milk, that is in turn subjected to a quite harsh technological process involving intense thermal treatments (and dessication in the traditional case). Kashk may be a food, or a food ingredient, better

tolerated in cases of milk allergy. Further studies are needed to clarify this very important issue. The use of gastrointestinal enzymes such as pepsin or chymotrypsin may be useful to generate *in vitro* set of peptides, with different functional or bioactive properties, which may find applications in functionalized foods. Other peptide sequences with potential bioactivity have been found after BBM treatment, and their nutritional effects deserve to be carefully investigated.

CRediT authorship contribution statement

Mina Pourjoula: Investigation. **Gianluca Picariello:** Data curation, Writing - original draft. **Giuseppina Garro:** Visualization, Investigation. **Giovanni D'Auria:** Investigation. **Chiara Nitride:** Investigation, Data curation. **Amid Rheza Ghaisari:** Conceptualization. **Pasquale Ferranti:** Supervision, Writing - review & editing.

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.foodres.2020.109107.

References

Addeo, F., Chianese, L., Salzano, A., Sacchi, R., Cappuccio, U., Ferranti, P., & Malorni, A. (1992). Characterization of the 12% trichloroacetic acid-insoluble oligopeptides of Parmigiano-Reggiano cheese. *Journal of Dairy Research, 59*(3), 401–411.

Addeo, F., Chianese, L., Sacchi, R., Spagna Musso, S. S., Ferranti, P., & Malorni, A. (1994). Characterization of the oligopeptides of Parmigiano-Reggiano cheese soluble in 120 g trichloroacetic acid/l. *Journal of Dairy Research, 61*(3), 365–374.

Aimutis, W. R. (2004). Bioactive properties of milk proteins with particular focus on anticariogenesis. *Journal Nutrition, 134*, 989S–995S.

Aschaffenburg, R. & Drewry, J. (1959). Proceedings of 15th International Dairy Congress (Vol. 3, p. 1631), London.

Bonomi, F., Brandt, R., Favalli, S., Ferranti, P., Fierro, O., Frøkiær, H., ... Iametti, S. (2011). Structural determinants of the immunomodulatory properties of the C-terminal region of bovine β-casein. *International Dairy Journal, 21*(19), 770–776.

Çakır, I., Coşkun, H., Akoğlu, I., İşleyen, M., Kıralan, M. & Bayrak, A. (2009) Journal of Food, Agriculture and Environment 17 (3-4), 116–119.

Caira, S., Pinto, G., Vitaglione, P., Dal Piaz, F., Ferranti, P., & Addeo, F. (2016). Identification of casein peptides in plasma of subjects after a cheese-enriched diet. *Food Research International, 84*, 108–112.

Chianese, L., Mauriello, R., Moio, L., Intorcia, N., & Addeo, F. (1992). Determination of ovine casein heterogeneity using gel electrophoresis and immunochemical techniques. *Journal of Dairy Science, 59*, 39–47.

Chianese, L., Garro, G., Ferranti, P., Malorni, A., Addeo, F., Rabasco, A., & Molina Pons, P. (1995). Discrete phosphorylation generates the electrophoretic heterogeneity of ovine β-casein. *Journal of Dairy Research, 62*(1), 89–100.

Chianese, L., Caira, S., Ferranti, P., Laezza, P., Malorni, A., Mucchetti, G., ... Addeo, F. (1997). The oligopeptides of sweet and acid cheese whey. *Lait, 77*, 699–715.

Claude, M., Lupi, R., Picariello, G., Drouet, M., Larré, C., Denery-Papini, S., ... Brossard, C. (2019). Digestion differently affects the ability of native and thermally aggregated ovalbumin to trigger basophil activation. *Food Research International, 118*, 108–114.

De Filippis, F., Genovese, A., Ferranti, P., Gilbert, J. A., & Ercolini, D. (2016). Metatranscriptomics reveals temperature-driven functional changes in microbiome impacting cheese maturation rate. *Scientific Reports, 25*(6), 21871. https://doi.org/10.1038/srep21871.

De Simone, C., Picariello, G., Mamone, G., Stiuso, P., Dicitore, A., Vanacore, D., ... Ferranti, P. (2009). Characterisation and cytomodulatory properties of peptides from Mozzarella di Bufala Campana cheese whey. *Journal of Peptide Science, 15*, 251–258.

De Simone, C., Ferranti, P., Picariello, G., Scognamiglio, I., Dicitore, A., Addeo, F., ... Stiuso, P. (2011). Peptides from water buffalo cheese whey induced senescence cell death via ceramide secretion in human colon adenocarcinoma cell line. *Molecular Nutrition & Food Research, 55*, 229–238.

Ferranti, P., Barone, F., Chianese, L., Addeo, F., Scaloni, A., Pellegrino, L., & Resmini, P. (1997a). Phosphopeptides from Grana Padano cheese: Nature, origin and changes during ripening. *Journal of Dairy Research, 64*(4), 601–615.

Ferranti, P., Itolli, E., Barone, F., Malorni, A., Garro, G., Laezza, P., ... Addeo, F. (1997b). Combined high resolution separation techniques (FPLC and HPLC) and mass spectrometry-based identification of peptides and proteins in Grana Padano cheese. *Lait, 77*, 683–697.

FitzGerald, R. J. (1998). Potential uses of caseinophosphopeptides. *International Dairy Journal, 8*, 451–457.

Ghorbani Gorji, E., Mohammadifar, M. A., & Ezzatpanah, H. (2010). Influence of gum tragacanth, *Astragalus gossypinus*, addition on stability of non fat Doogh, an Iranian fermented milk drink. *International Journal of Dairy Technology, 64*, 262–268.

Ghorban Shiroodi, S., Mohammadifar, M. A., Ghorbani Gorji, E., Ezzatpanah, H., & Zohouri, N. (2012). Influence of gum tragacanth on the physicochemical and rheological properties of Kashk. *Journal of Dairy Research, 79*(1), 93–101.

Gobbetti, M., Ferranti, P., Smacchi, E., Goffredi, F., & Addeo, F. (2000). Production of angiotensin-I-converting-enzyme-inhibitory peptides in fermented milks started by *Lactobacillus delbrueckii* subsp. *bulgaricus* SS1 and *Lactococcus lactis* subsp. *cremoris* FT4. *Applied and Environment Microbiology, 66*(9), 3898–3904.

Grenby, T. H., Andrews, A. T., Mistry, M., & Williams, R. J. H. (2001). Dental caries-protective agents in milk and milk products: investigations in vitro. *Journal of Dentistry, 29*(2), 83–92.

Hartmann, R., & Meisel, H. (2007). Food-derived peptides with biological activity: From research to food applications. *Current Opinion in Biotechnology, 18*, 163–169.

Heberle, H., Meirelles, G. V., da Silva, F. R., Telles, G. P., & Minghim, R. (2015). InteractiVenn: A web-based tool for the analysis of sets through Venn diagrams. *BMC Bioinformatics, 16*, 169–176.

Hebert, E. M., Mamone, G., Picariello, G., Raya, R. R., Savoy, G., Ferranti, P., & Addeo, F. (2008). Characterization of the pattern of alphas1- and beta-casein breakdown and release of a bioactive peptide by a cell envelope proteinase from Lactobacillus delbrueckii subsp. lactis CRL 581. *Applied and Environment Microbiology, 74*(12), 3682–3689.

ISIRI. (2005). Institute of Standards and Industrial Research of Iran (ISIRI). Dried Kashk-Specification. No, 1188, ICS: 67.100.99, Teheeran, Iran.

ISIRI. (2009). Institute of Standards and Industrial Research of Iran (ISIRI). Industrial Liquid Kashk- Specification. No, 6127, ICS: 67.100.99, Teheran, Iran.

Jafari, M., Ghaisari, H. R., Khaniki, G. J., & Shariatifar, N. (2015). Comparison of chemical and microbiological characteristics between traditional and industrial Kashks in the Fars province (IRAN). *Annals Food Science and Technology, 16*(2), 462–470 available on-line at www.afst.valahia.ro 462.

Jelen, P., & Rattray, W. (1995) Thermal denaturation of whey proteins. In: Heat-induced changes in milk. Second edition. International dairy federation edited by P.F. Fox (pp. 66–85). ISBN 9290980179.

Kamber, U. (2008). The manufacture and some quality properties of kurut, a dried dairy product. *International Journal of Dairy Technology, 61*(2), 146–150.

Koide, K., Itoyama, K., Fukushima, T., Miyazawa, F., & Kuwata, T. (1991). Method for separation and concentration of phosphopeptides. *European Patent Application EP, 0 443 718*, A2.

Kurmann, J. A., Rasic, J. L., & Kroger, M. (1992). *Encyclopedia of Fermented fresh milk products: An international inventory of fermented milk, cream, buttermilk, whey, and related products.* The Netherlands: Kluwer Academic Publishers Group.

Liao, Y., Alvarado, R., Phinney, B., & Lönnerdal, B. (2011a). Proteomic characterization of specific minor proteins in the human milk casein fraction. *Journal of Proteome Research, 10*(12), 5409–5415.

Liao, Y., Alvarado, R., Phinney, B., & Lönnerdal, B. (2011b). Proteomic characterization of human milk fat globule membrane proteins during a 12 month lactation period. *Journal of Proteome Research, 10*(8), 3530–3541.

Liu, G., Miao, J., Sun, S., Luo, Z., Xia, Z., Guo, B., ... Cao, Y. (2017). The effect of dietary magnesium and caseinphosphopeptides on bone metabolism in rats. *Food & Function, 8*(12), 4487–4495.

Mamone, G., Caira, S., Garro, G., Mauriello, R., Nicolai, M. A., Picariello, G., ... Addeo, F. (2013). Challenging the heterogeneity of casein by an IEF/MALDI-TOF "virtual 2D-like" approach. *Food Research International, 54*(1), 1263–1272.

Minekus, M., Alminger, M., Alvito, P., Balance, S., Bohn, T., Bourlieu, C., ... Brodkorb, A. (2014). A standardised static in vitro digestion method suitable for food - an international consensus. *Food & Function, 5*(6), 1113–1124.

Mukhopadhy, A., & Sweeney, T. (2016). Milk proteins: Processing of bioactive fractions and effects on gut health. In I. Gigli (Ed.). *Milk proteins - from structure to biological properties and health aspects* (pp. 83–100). London, U.K.: Intech Open.

Neuhoff, V., Arold, N., Taube, D., & Ehrhardt, W. (1988). Improved staining of proteins in polyacrylamide gels including isoelectric focusing gels with clear background at nanogram sensitivity using Coomassie Brilliant Blue G-250 and R-250. *Electrophoresis, 9*, 255–262.

Noori, A., Keshavarzian, F., Mahmoudi, S., Yousefi, M., & Nateghi, L. (2013). Comparison of traditional Doogh (yogurt drinking) and Kashk characteristics (Two traditional Iranian dairy products). *European Journal of Experimental Biology, 3*(6), 252–255.

Oghbaei, M., & Prakash, J. (2008). Nutritional and functional properties of kashk: Fermented sheep milk. *Journal of Food Science and Technology Nepal, 4*, 38–42.

Picariello, G., Ferranti, P., Fierro, O., Mamone, G., Caira, S., Di Luccia, A., ... Addeo, F. (2010). Peptides surviving the simulated gastrointestinal digestion of milk proteins: Biological and toxicological implications. *Journal of Chromatography B: Analytical Technologies in the Biomedical and Life Sciences, 878*(3–4), 295–308.

Picariello, G., Ferranti, P., Mamone, G., Klouckova, I., Mechref, Y., Novotny, M. V., & Addeo, F. (2012). Gel-free shotgun proteomic analysis of human milk. *Journal of Chromatography A, 1227*, 219–233.

Picariello, G., Iacomino, G., Mamone, G., Ferranti, P., Fierro, O., Gianfrani, C., ... Addeo, F. (2013). Transport across Caco-2 monolayers of peptides arising from in vitro digestion of bovine milk proteins. *Food Chemistry, 139*, 203–212.

Picariello, G., Miralles, B., Mamone, G., Sánchez-Rivera, L., Recio, I., Addeo, F., & Ferranti, P. (2015). Role of intestinal brush border peptidases in the simulated digestion of milk proteins. *Molecular Nutrition & Food Research, 59*(5), 948–956.

Pizzano, R., Nicolai, M. A., Padovano, P., Ferranti, P., Barone, F., & Addeo, F. (2000). Immunochemical evaluation of bovine β-casein and its 1–28 phosphopeptide in cheese during ripening. *Journal of Agriculture and Food Chemistry, 48*(10), 4555–4560.

Saavedra, L., Hebert, E. M., Minahk, C., & Ferranti, P. (2013). An overview of "omic" analytical methods applied in bioactive peptide studies. *Food Research International, 54*(1), 925–934.

Sevgi Kirdar, S. (2012). A survey on Chemical, Biochemical and Microbiological characteristics of a Traditional Dairy product in Mediterrean Region. *Kes Journal Animal Veterinary, 11*(3), 330–334.

Shan, L., Molberg, O., Parrot, I., Hausch, F., Filiz, F., Gray, G. M., ... Khosla, C. (2002). Structural basis for gluten intolerance in celiac sprue. *Science, 297*, 2275–2279.

Soltani, M., & Güzeler, N. (2012). The production and quality properties of liquid kashks. *Gida, 38*, 1–7.

Chapter 7

Casein-derived peptides from the dairy product kashk exhibit wound healing properties and antibacterial activity against Staphylococcus aureus: structural and functional characterization

Food Research International 153 (2022) 110949

Contents lists available at ScienceDirect

Food Research International

journal homepage: www.elsevier.com/locate/foodres

Casein-derived peptides from the dairy product kashk exhibit wound healing properties and antibacterial activity against *Staphylococcus aureus*: Structural and functional characterization

Veronica Folliero [a,1], Stefania Lama [b,1], Gianluigi Franci [c], Rosa Giugliano [a], Giovanni D'Auria [d], Pasquale Ferranti [d,*], Mina Pourjula [e], Massimiliano Galdiero [a], Paola Stiuso [b]

[a] Department of Experimental Medicine, University of Campania Luigi Vanvitelli, Naples, Italy
[b] Department of Precision Medicine, University of Campania Luigi Vanvitelli, 80138 Naples, Italy
[c] Department of Medicine, Surgery and Dentistry "Scuola Medica Salernitana", University of Salerno, 84081 Baronissi, Italy
[d] Department of Agricultural Sciences, University of Naples Federico II, 80138 Portici, Italy
[e] Department of Food Hygiene and Quality Control, Faculty of Veterinary Medicine, Shiraz University, Shiraz, Iran

ARTICLE INFO

Keywords:
Khask
Fermented milk
Peptidome
Bioactive peptides
Atopic dermatitis
Antibacterial activity
Wound-healing activity

ABSTRACT

Kashk is a fermented dairy product typical of the Middle East, traditionally produced with sour milk and/or dairy waste. The kashk water-soluble peptide fraction was characterized at the molecular level by liquid chromatography–mass spectrometry and its antibacterial and skin healing activity was evaluated. Antibacterial assays showed a significant antibacterial activity against clinical isolates of *Staphylococcus aureus* (*S. aureus*) from patients with atopic dermatitis, inhibiting bacterial growth by approximately 45% (500 μg/mL). Skin repair activity was evaluated on keratinocytes through scratch tests showing accelerated wound closure *in vitro* in the presence of TNF-α, by approximately 44% (500 μg/mL), compared to control cells. Furthermore, based on the MTT assay, the kashk peptide fraction did not show toxicity on keratinocytes. The results suggested that the peptide kashk extract may be useful in skin care for patients with atopic dermatitis.

1. Introduction

In recent decades, there is greater evidence that milk proteins are a potential source of biologically active peptides (Chakrabarti et al., 2018). These peptides are inactive within the parent protein sequence and can be released during proteolytic processes. Bioactive peptides are short sequences spanning 3–20 amino acid residues. After food digestion, these peptides can be absorbed in the intestine and enter the bloodstream, ensuring their bioavailability and a physiological effect at the target site. These peptide sequences can influence different physiological responses including cardiovascular, digestive, endocrine, immune and other activities (Karami & Akbari-adergani, 2019). Several studies have reported the antibacterial activity of various peptides derived from milk proteins on different pathogenic strains (Guha et al., 2021). Among them, lactoferricin, a peptide isolated from a peptic hydrolysate of bovine lactoferrin, carries out its antibacterial activity through the destruction of the bacterial wall of *Escherichia coli*, *Bacillus subtilis* and *Staphylococcus aureus* (*S. aureus*), (Wakabayashi et al., 2003). Moreover, Isracidin, the N-terminal segment of αs1-casein, has been proposed for therapeutic treatment of infections caused by *S. aureus*, *Streptococcus pyogenes* and *Listeria monocytogenes* (Mohanty et al., 2016).

Over the years, several reports on milk and fermented dairy products as potential sources of bioactive peptides are accumulating (Punia et al., 2020; Tulipano, 2020; Beverly et al., 2019). Recent studies have observed the ability of some bioactive peptides to promote wound healing in the skin, through the modulation of cytokine production, cell migration and proliferation (Song, et al., 2019). Two related studies investigated how the supplementation with camel milk protein hydrolysates influenced wound healing in diabetic rats. Diabetic rats fed with camel milk peptides had faster healing of skin lesions than control

* Corresponding author.
E-mail addresses: veronica.folliero@unicampania.it (V. Folliero), stefania.lama@unicampania.it (S. Lama), gfranci@unisa.it (G. Franci), rosa.giugliano@unicampania.it (R. Giugliano), giovanni.dauria@unina.it (G. D'Auria), ferranti@unina.it (P. Ferranti), massimiliano.galdiero@unicampania.it (M. Galdiero), paola.stiuso@unicampania.it (P. Stiuso).
[1] These authors have contributed equally to this work.

https://doi.org/10.1016/j.foodres.2022.110949
Received 22 May 2021; Received in revised form 31 December 2021; Accepted 8 January 2022
Available online 24 January 2022

(Abdel-Salam, 2014; Ebaid et al., 2015), although the identity of the active peptides was not investigated. As an indirect support to these findings, molecular analysis of bovine colostral milk identified over ninety proteins related to the various stages of wound repair (Altomare et al., 2016).

Kashk is a fermented dairy product of the Middle East (Turkey, Iran), traditionally produced with cheese-making wastes such as buttermilk, which used to be the main protein source for the nomad population in Iran. Today, it is an ingredient of Iranian meals and is produced also industrially in liquid form. Despite its importance as a cheap and nutritious ingredient, kashk has been the subject of few scientific investigations. Recently, the analysis of the protein composition of both industrial and traditional kashk showed that kashk undergoes extensive proteolysis during production and maturation (Pourjoula et al., 2020). The detailed molecular characterization of kashk water peptide extracts (KE), coupled with bioinformatic *in silico* analysis, also detected the presence of potential bioactive peptides, that included calcium transporters, immunomodulators, ACE inhibitors, opioids, antimicrobials, although this was based only by sequence comparison with known bioactive ones. Therefore, the supposed active functions remain to be directly confirmed yet. Moving from this background, the purpose of this study was to evaluate the antibacterial and healing activity of the peptide obtained from kashk for perspective topical treatment of lesions in atopic dermatitis (AD) patients.

AD is a common chronic inflammatory skin disease which manifests itself through a variety of symptoms including skin itching and dryness and skin lesions. About 90% AD patients have lesions colonized *by S. aureus* (Ogonowska, et al. 2020), that contributes to the severity of the disease (Lee et al., 2018). *S. aureus* performs an important role in the pathogenesis of AD through the secretion of toxins, antigens and exoenzymes which interact with keratinocytes and inflammatory cells, causing unbalance of skin hemostasis and damage of the epidermal barrier (Iwamoto et al., 2019). The association between skin colonization by *S. aureus* and disease severity is supported by empirical antibiotic treatments (Carrera et al., 2019; Serretiello, et al., 2020). Incorrect use of antibiotics involves development of antimicrobial resistance (AMR) among *S. aureus* strains (Seiti Yamada Yoshikawa et al., 2019). Hence, the conventional antibiotics are nowadays facing limited effectiveness. In fact, resistance of *S. aureus* to antibiotics represents a major cause of therapeutic complication (Blicharz et al., 2019). Moving from this background, this study was aimed to assess the presence of bioactive peptides in KE, looking for alternative molecules able either to improve skin lesions or to contrast *S. aureus* colonization. The peptidome characterization by mass spectrometry (MS) allowed us to identify the sequences of the bioactive peptides, that could represent novel tools in the treatment of AD.

2. Materials and methods

2.1. Materials

Representative samples of either traditional dry (two samples) or industrial liquid (two samples) kashk were collected from local producers in Shiraz (Iran). All the reagents used in this study were of analytical or higher grade. HPLC-grade solvents, water and acetonitrile (ACN) were purchased from Carlo Erba (Milan, Italy). Diethyl ether and trifluoroacetic acid (TFA) were from Sigma-Aldrich (St Louis, MO, USA).

2.2. Preparation of water-soluble peptide extracts

The water-soluble extract was prepared, according to Ferranti et al., 1997a; Ferranti et al., 1997b; Rizzello et al., 2005 and Pourjoula et al., 2020, with the following modifications. Before peptide extraction, the kashk samples were finely grated and defatted with diethyl ether in a 1:30 ratio (w/v) under mechanical stirring for 30 min (twice). The samples were then centrifuged (4000 rpm, 15 min) and the supernatant was discarded. Peptide extraction was carried out by solubilizing defatted samples in ultra-pure water in a 1:20 ratio (w/v). The mixtures were sonicated for 15 min (DU-45, Argo Lab, Modena, Italy), shaken for additional 30 min and then left at 4 °C for 30 min. Finally, the samples were centrifuged (4000 rpm, 20 min) to collect the supernatants, which were ultra-filtered on a Centriprep cartridge having a 3-KDa cut-off membrane (Millipore, Bedford, MA). The peptide extracts were dried using a Savant concentrator (Speed-Vac, Milan, Italy) and stored at −20 °C either for structural analysis or for biological assays. Part of the resulting dried water-soluble peptides was suspended in aqueous TFA (0,1% v/v) before performing semi-preparative RP-HPLC separation.

2.3. Reverse phase HPLC separation

Reverse phase (RP)-HPLC analysis of water-soluble peptides was carried out injecting 100 µL of kashk peptide solution prepared in 0,1% (v/v) TFA on a HPLC chromatogram (HP 1100 Agilent - Palo Alto, CA, USA) modular system equipped with a diode array detector. The column effluents were monitored by UV detection at $\lambda = 220$ and 280 nm. The separation was performed using Vydac C18 RP 218TP52 (2.1x250 mm, 5 µm, 300A). Solvent A was 0.1% (v/v) TFA in water; solvent B was 0.1% (v/v) TFA in ACN. After 5 min of isocratic elution using 5% solvent B, a 5–60% gradient ramp was applied over 60 min at a flow rate of 0.2 mL/min. The aliquots, taken at 10-min intervals, were collected by performing several chromatographic runs. Subsequently, samples concentrated in a vacuum evaporator to remove residual ACN before freeze-drying.

2.4. LC-MS/MS of kashk active fractions

The active HPLC peptide fractions were concentrated under vacuum and analysed through LC-MS/MS analysis, using a Dionex UltiMate 3000 nano-UHPLC system coupled with nano-ESI-linear ion trap (LIT) Thermo XL mass spectrometer (Thermo Fisher Scientific, Waltham, MA, USA). Samples dissolved in 0.1% (v/v) formic acid were loaded through a 5 mm long, 300 µm id pre-column (LC Packings, USA) and separated by an Acclaim™ PepMap™ C18 column (150 mm × 75 µm, 3 µm). Flow rate: 0.200 µL/min. Eluent A was 0.1% formic acid (v/v) in Milli-Q water; eluent B was 0.1% formic acid (v/v) in acetonitrile. The column was equilibrated at 5% B. Peptides were separated applying a 5–40% gradient of eluent B over 40 min. MS data were recorded in the 200–2000 m/z range. Data-dependent MS/MS spectra were collected from the five most abundant precursor ions upon fragmentation (charge state $\geq$2; isolated width 2 Da; min. signal required: 500) using CID activation with 35.0% normalized collision energy, activation Q of 0.25, and activation time of 30 s. Spectra were processed using the Xcalibur Software 3.1 version (Thermo Scientific). Mass spectra were then analyzed using the Peaks 6.0 software, restricting the search to *Bos taurus*, *Ovis aries* and *Capra hircus* database (UniProt, October 2020). Database searching parameters for the identification of the peptidome included "no enzyme" specification, Met-oxidation, phosphorylation of serine and threonine and lactose adduct on lysine as variable protein modifications. The parameters for the identification of the water-soluble peptides considered a mass tolerance value of 1.0 Da for the precursor ion and 0.6 Da for MS/MS fragments. An FDR value of 1% was applied to the results and only peptides with a $-10\lg P > 30$ were selected. The identified peptides were further processed on MilkAMP, a free on-line database of milk antimicrobial peptides (AMPs) (Nielsen et al., 2017) to identify those with antimicrobial activity inside the bioactive fractions.

2.5. Bacterial strains

The bacteria used in this study were the standard *S. aureus* strain (ATCC 6538) and their clinical isolates, including multisensitive (MS), resistant to methicillin (MR), resistant to macrolides (Ma.Res), resistant

Table 1
Bacterial strains and related resistance phenotypes, used in this study.

	Resistance phenotypes
S. aureus ATCC 6538	Reference strains: multi-sensitive strain
S. aureus MS	Clinical isolate: multi-sensitive strain
S. aureus MR	Clinical isolate: methicillin resistant strain
S. aureus Ma.Res	Clinical isolate: macrolide resistant strain
S. aureus Quin.Res	Clinical isolate: quinolone resistant strain
S. B.LP	Clinical isolate: producer of Beta-lactamase

to quinolones (Quin.Res) and producer of beta-lactamase (B.LP) strains (Table 1) (Supplementary Fig. 1). Clinical isolates of *S. aureus* were collected from skin lesions of AD patients. The swabs performed were plated on the Columbia agar with 5% sheep blood and Chapman's medium (bioMerieux, United Stated). After incubation of the media at 37 °C, each plate was analysed for the presence of S. aureus. The bacterial identifications were performed by Matrix Assisted Laser Desorption Ionization - Time of Flight MS (MALDI-TOF) (Bruker Daltonics, Germany). Briefly, the suspected bacterial colony was spread over a MALDI-TOF 96 MSP (Bruker Dal-tonics, Germany). Samples were treated with 1 µL of matrix solution (saturated solution of alpha-cyano-4-hydroxycinnamic acid with 50% (v/v) acetonitrile and 2.5% (v/v) trifluoroacetic acid) (Bruker Daltonics, Germany) and finally dried for 2 min. The obtained spectra were imported into MALDI BioTyper 3.0 software (Bruker Dal-tonics, Germany) and evaluated by standard pattern matching with respect to the reference spectra. A valid identification of the bacterial species was assumed by a score higher than 2.

Antibiotic sensitivity testing was performed by a Phoenix BD (Becton Dickinson, United States) system. Briefly, a 0.5 McFarland (McF) (ID) bacterial inoculum was prepared using a Phoenix spectrophotometer. A 25 µL volume of standardized bacterial suspension was added to Phoenix AST Broth, previously supplemented with a drop of Phoenix AST Indicator. The ID and AST stock were loaded into the Phoenix panels, which were subsequently sealed, registered and deposited in the Phoenix device. Antibiograms were interpreted using Epicenter software version 7.22A (Becton Dickinson Diagnostic Systems, USA) after 16 h of incubation (Folliero et al., 2020).

To normalize the bacterial suspensions for the antibacterial activity test, colonies of each strain, plated on Mueller-Hinton agar plates, were inoculated in the Mueller-Hinton broth medium (MHB) and grown at 37 °C under shaking (180 rpm) over night (pre-inoculum). Following incubation, the pre-inoculum was transferred to fresh MHB and further incubated at 37 °C until the exponential growth phase (6×10^8 CFU/mL) (inoculum). Serial dilution of inoculum was carried out to determine the concentration of bacteria required for the test (1×10^6 CFU /mL) (Pironti et al., 2021).

2.6. Cell cultures

The immortalized human keratinocytes (HaCaT cells) were obtained from the American Type Culture Collection (ATCC) and were cultured in 5% CO2 at 37 °C in regular Dulbecco's Modified Eagle's Medium (DMEM) (Euroclone S.P.A., Milan, Italy), supplied with 10% heat-inactivated fetal bovine serum (Sigma-Aldrich, Germany), 1% L-glutamine and 1% penicillin–streptomycin solution.

2.7. Antibacterial activity assays

The antimicrobial activity tests were performed in accordance with the broth microdilution method outlined by the National Committee on Clinical Laboratory Standards (NCCLS). The assays were carried out in sterile 96-well microplates. The concentrations (from 500 to 2 µg/mL) of tarditional and industrial KE fractions (250 µg/mL) were prepared in distilled sterilized water for a final volume of 100 µL/well. A volume of 50 µL of bacterial inoculum (6×10^8 CFU/mL) was added to each well, obtaining a final bacterial concentration of 5×10^5 CFU/mL (Buonocore et al., 2020). The positive control used was vancomycin. Antimicrobial activities were expressed in growth inhibition percentages after 20 h of incubation at 37 °C. The percentage of growth inhibition was evaluated according to the following equation:

$$\% \text{ Growth inhibition} = (A_c - A_s)/A_c \times 100.$$

in which A_c is the mean of the absorbance values (600 nm) of the negative control while A_s is the mean of the absorbance values (600 nm) of the treated bacteria. All experiments were carried out in triplicate and standard ± mean deviations were reported (Franci et al., 2018).

2.8. Cell proliferation assay

The effect of KE on HaCaT cells was evaluated by the MTT viability assay. Untreated and treated HaCaT cells with 20 ng/mL of TNF-α for 3 h were seeded in 96-well plates at the appropriate density of 3×10^3 per well and were exposed for 24 h with increasing concentrations of KE (250–0 µg/mL). Then, 50 µL of 1 mg/mL 3-(4,5-Dimethylthiazol-2-yl)-2,5-Diphenyltetrazolium bromide (MTT) were mixed with 200 µL of medium and added to the well. After a 4 h incubation at 37 °C, the medium was removed, then the formazan crystals were solubilized by adding 100 µL of DMSO and by mixing it in an orbital shaker for 5 min. Absorbance at 570 nm was measured using a plate reader (Tecan Infinite F200 Microplate Reader, New York, USA). Experiments were carried out by triplicate determination on at least three separate experiments. All data were expressed as mean ± standard deviation.

2.9. In vitro scratch assay

HaCaT cells were seeded in 6-well plates under normal culture conditions at the density of 4.5×10^4 cells/well. At ~90% confluence, the medium was removed and the cell monolayer was scratched across the center using a sterile pipet tip (200 µL) to produce a uniform wound area. Through gently washing with sterile PBS1X, cellular debris was removed from the well. To simulate the inflammation, the cells were pretreated for 3 h with 20 ng/mL of TNF-α (R&D Systems, Space Import-Export SRL, Milan, Italy). The cells were incubated with KE at 500 µg protein/mL final concentration. Cell migration was monitored by microscopy using an inverted light microscope (Eclipse TE 300 Nikon, New York, USA) and images were captured (10X magnification) at different time points after the scratch (0–18 h). The images acquired were analyzed quantitatively, using a specific wound healing tool of ImageJ software (Maryland, USA). The percentage of relative wound closure was assessed according to the following equation (Moghadam et al., 2019):

$$\text{Wound closure } \% = (W_0 - W_n)/W_0 \times 100$$

in which W_0 and W_n are the width of the scratch before and after treatment.

2.10. Statistical analysis

The experiments were performed in triplicate and the results expressed as mean ± standard deviation. Data analysis was conducted, using Graphpad Prism software (Version 7.05 for Windows, California, USA). Statistical significance was referred to the negative control and was calculated through one way ANOVA with Dunnett's test as *post hoc*. *P-values* lower than 0.05 were considered significant (Pironti et al., 2021).

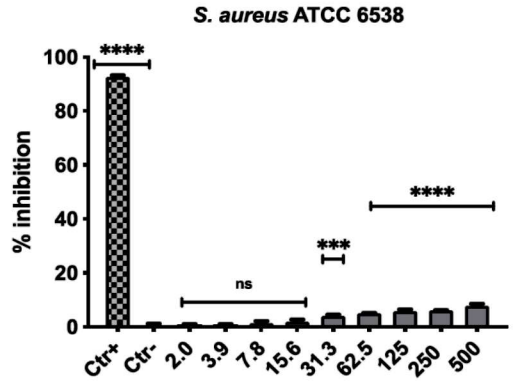


Fig. 1. Antibacterial effect of peptide extract from industrial KE on *S. aureus* strains ATCC, **** $p < 0.0001$, *** $p < 0.0002$, ** $p < 0.0021$, *$p < 0.0332$, ns (not significant).

3. Results

3.1. Antibacterial activity of peptide fraction from kashk

The antibacterial activity of traditional and industrial KE was studied against standard *S. aureus* and relative clinical strains, isolated from AD patients. The data of the antibacterial activity of these extracts are shown in Figs. 1 and 2.

Industrial KE did not exhibit antibacterial activity on the standard strain (*S. aureus* ATCC 6538), differently from the traditional KE (Figs. 1–2). The results are given in the histogram as percentage of growth inhibition. As shown in Fig. 2, a dose-dependent effect was observed for traditional KE on all tested *S. aureus* strains. In particular, the growth of all *S. aureus* strains was decreased by about 45%, compared to the untreated control. The interaction effect of KE with bacterial strains was highly significant at $P < 0.05$.

3.2. RP-HPLC separation of KE peptides

The peptidome of kashk has been recently characterized and the presence of potential bioactive sequences has been reported, including opioid, immunomodulatory and antimicrobial sequences (Pourjoula et al., 2020). However, no direct evidence of the occurrence of bioactive peptides has been reported in kashk so far. Therefore, as the antibacterial assay established that the traditional KE was the sample with the strongest activity, we fractionated the total KE using RP-HPLC. The chromatogram showed the presence of a high number of peptides in the hydrophilic region (<50% eluent B, 0–40 min elution time) (Fig. 3).

3.3. Antibacterial activity of KE HPLC fractions

The RP-HPLC fractions of KE eluted at 10-min intervals were assayed for antimicrobial activity and the results are presented in Table 2. Comparing the bacterial growth inhibition rate of the fractions, a low inhibitory effect was found for the 10–20 min and 40–50 min fractions. Interesting results were instead observed for the fractions eluted at 20–30 min and 30–40 min. In particular, the former inhibited the growth of all *S. aureus* strains studied on average by 35.4%, while the latter led to inhibit the growth of all strains by about 24.2%. The strong antibacterial activity of these fractions, similar to whole KE, was possibly related to the presence of antibacterial peptides in their composition. The effect of the interaction between the single fractions and the bacterial strains was significant ($P < 0.05$). These findings prompted us to undertake detailed molecular characterization of peptide composition.

3.4. Structural mass spectrometryanalysis of peptides

LC-MS/MS analysis of the active RP-HPLC fractions of KE allowed to obtain the detailed structural identification of the main peptides and to identify antimicrobial peptide sequences therein. The software identification of the peptides was also validated by manual inspection. Representative MS/MS spectra of bioactive peptides are presented in

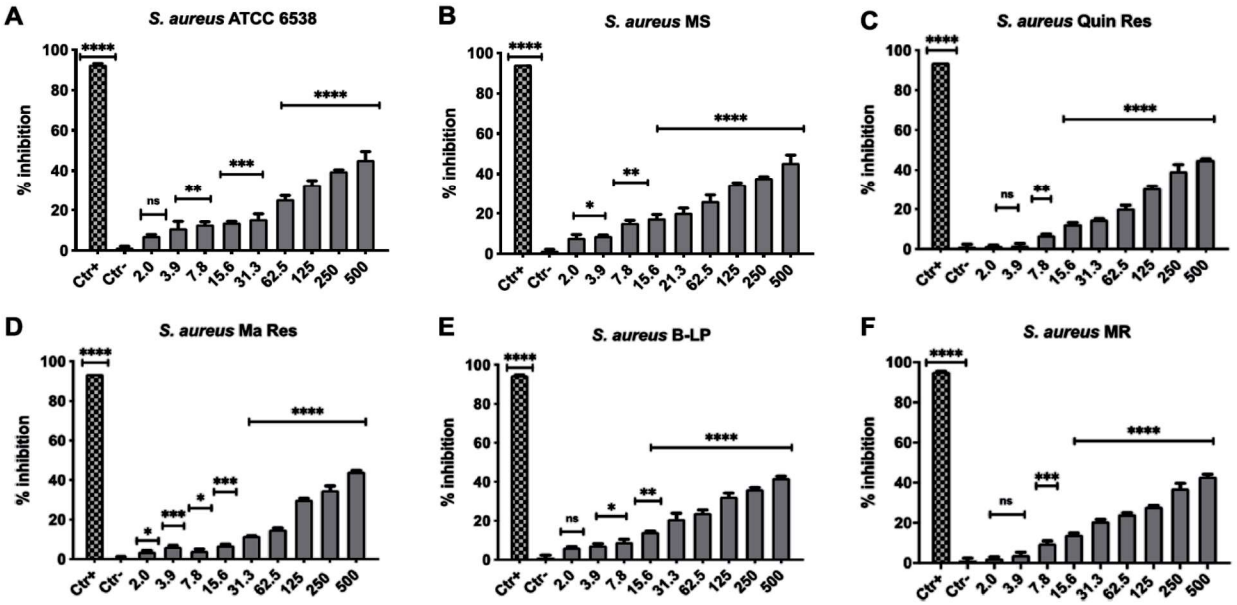


Fig. 2. Antibacterial effect of peptide extract from traditional KE on *S. aureus* strains: ATCC, ** $p < 0.0059$, *** $p < 0.0008$, **** $p < 0.0001$, ns (A); MS, * $p < 0.0445$, *** $p < 0.0002$, **** $p < 0.0001$, ns (B); Quin.Res, ** $p < 0.0096$, **** $p < 0.0001$, ns (C); Ma.Res, * $p < 0.0248$, ** $p < 0.0096$, *** $p < 0.0003$, **** $p < 0.0001$, ns (D); B-LP, * $p < 0.0255$, ** $p < 0.0074$, **** $p < 0.0001$, ns (E); and MR, *** $p < 0.0005$, **** $p < 0.0001$, ns (F). All absorbance values were recorded and converted to % of inhibition compared to controls. The results are representative of three independent experiments, and stringent statistical filters were applied.

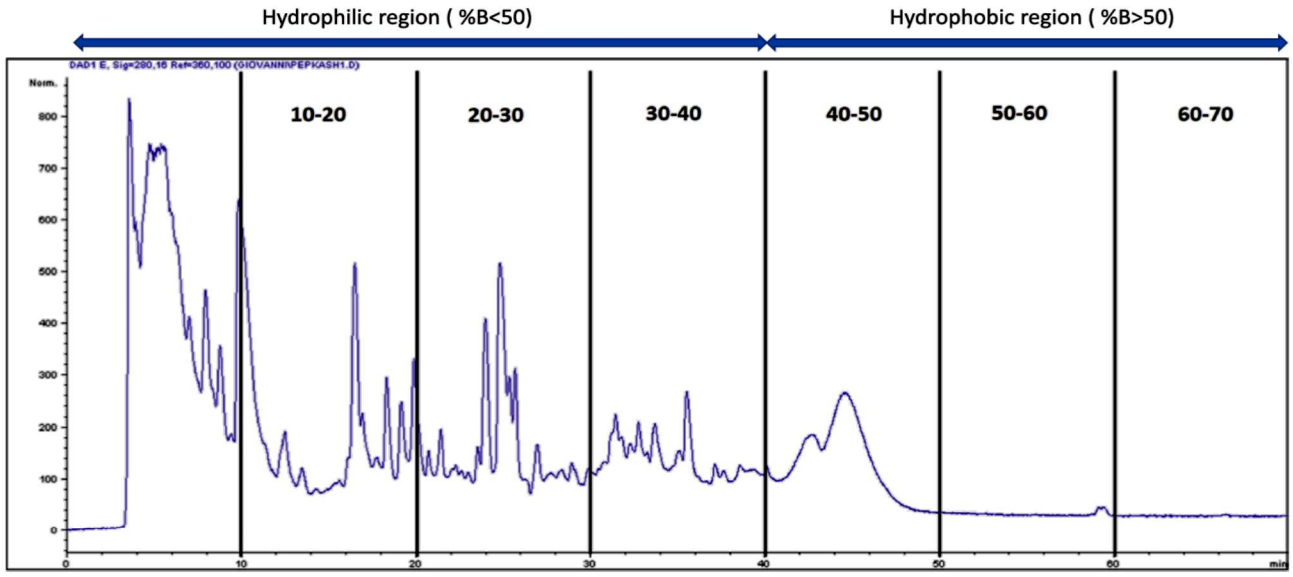


Fig. 3. RP-HPLC separation of traditional KE peptide extract.

Table 2
Antibacterial activity of traditional KE extract and fractions, $p < 0.05$.

Bacteria	Bacterial growth inhibition (% ± SD) – 250 μg/ml				
	Total	Retention time			
	KE	Fraction (10–20 min)	Fraction (20–30 min)	Fraction (30–40 min)	Fraction (40–50 min)
S. aureus ATCC 6538	39.7 ± 0.06	10.9 ± 1.34	37.5 ± 2.21	25.8 ± 0.09	10.3 ± 0.87
S. aureus MS	38.4 ± 0.09	9.7 ± 0.97	36.3 ± 1.56	24.8 ± 1.32	9.8 ± 1.70
S. aureus MR	37.1 ± 0.98	8.9 ± 1.76	35.6 ± 0.76	22.1 ± 0.98	10.5 ± 2.02
S. aureus Ma.Res	33.2 ± 1.03	9,1 ± 1.32	31.5 ± 1.23	21.9 ± 1.39	8.7 ± 1.82
S. aureus Quin.Res	39.3 ± 1.20	10.1 ± 0.09	37.3 ± 1.98	24.3 ± 1.49	10.1 ± 0.56
S. aureus B. LP	36.2 ± 0.10	10.4 ± 2.12	34.1 ± 1.76	26.1 ± 1.23	10.8 ± 1.01

Fig. 4. Only peptides with a −10lgP > 30 were considered as identified after manual inspection. Overall, the peptide composition of the two fractions was much less complex than that of whole KE (Pourjoula et al, 2020), and was dominated, expectedly, by casein-derived fragments (Table 3). Interestingly, among the 30 peptides observed, there were four immediate precursors and nine almost complete fragments of the known antimicrobial peptide caseicin C. This peptide, *i.e.* the fragment 180–193 of bovine α_{s1}-casein (SDIPNPIGSENSEK), first identified in casein hydrolysates of *Lactobacillus casei* and *Lactobacillus acidophilus* DPC6026, has been reported to be active against Gram-positive (Gram+) bacteria. The peptides found here derived from the same portion of the caprine/ovine homologous of the bovine α_{s1}-casein, which is due to the fact that traditional kashk is produced using caprine and ovine dairy residues. Their sequence (SDIPNPIGSENSGK) is therefore identical to that of bovine caseicin C except for a substitution G for E on the second-last residue. Notably, we also found a precursor of caseicin B, that is the fragment 15–22 of αs1-casein (VLNENLLR). Also in this case the fact that traditional kashk is produced with residues of caprine/ovine milk was confirmed by a peptide point modification at position 2 in the caseicin B sequence, where a P residue replaces L (GLSPEV**P**NENLLR – Table 3). Hayes et al. (2006) previously described the production of antimicrobial casein-derived peptide caseicin B from the fermentation of sodium caseinate by *Lactobacillus johnsonii* DPC 6026. This peptide, derived from the fermentation of αs1-casein, is a fragment of isracidin, a previously characterized AMP (Hill et al., 1974). Caseicin B has been reported to be active against both Gram-negative (Gram−) (*i.e. Escherichia coli* and *Salmonella* spp.) and Gram+ (*Listeria* spp., *S. aureus* and *Streptococcus mutans*) bacteria (Kent et al., 2012; McDonnell et al., 2012; Norberg et al., 2011).

In the fraction eluted between 30 and 40 min, in addition to the caseicin C, we also found precursors and fragments of caseicidin 15 and 17. Casecidin 15 (YQEPVLGPVRGPFPI) and casecidin 17 (YQEPVLGPVRGPFPIIV), correspond to the C-terminal region of bovine β-casein, more specifically to residues193–207 and 193–209, respectively (Fig. 5). These two AMPs were both able to inhibit identically *Escherichia coli* DPC6053 (Birkemo et al., 2009). Also in this case we found a single modification on the second-last residue of casecidin 17 where a leucine replaces the isoleucine of the bovine counterpart. AMPs interact electrostatically with the anionic groups of the bacterial cell walls.

In Gram− bacteria, most AMPs bind to the surface of lipopolysaccharides and to teichoic acids of Gram+ bacterial membrane. After the initial electrostatic binding to the bacterial surface, some AMPs insert themselves into bacterial cell membranes in a perpendicular orientation to form toroidal-shaped transmembrane pores (Mohanty et al., 2016). Leakage of cellular components through these pores causes cell death. This action is mainly related to the amino acid sequence and to the structure of the AMPs themselves. The presence of hydrophobic residues alternating with hydrophilic amino acid residues originates an amphipathic structure capable of permeabilizing the bacterial membrane (Clare et al., 2003). The presence of helix stabilising residues (Ala and Leu) and of positively charged residues (Lys and Arg) is correlated with the antimicrobial activity of peptides (Hancock and Lehrer, 1998). Milk AMPs found in this study share these characteristics, that contribute to explain their antimicrobial effect.

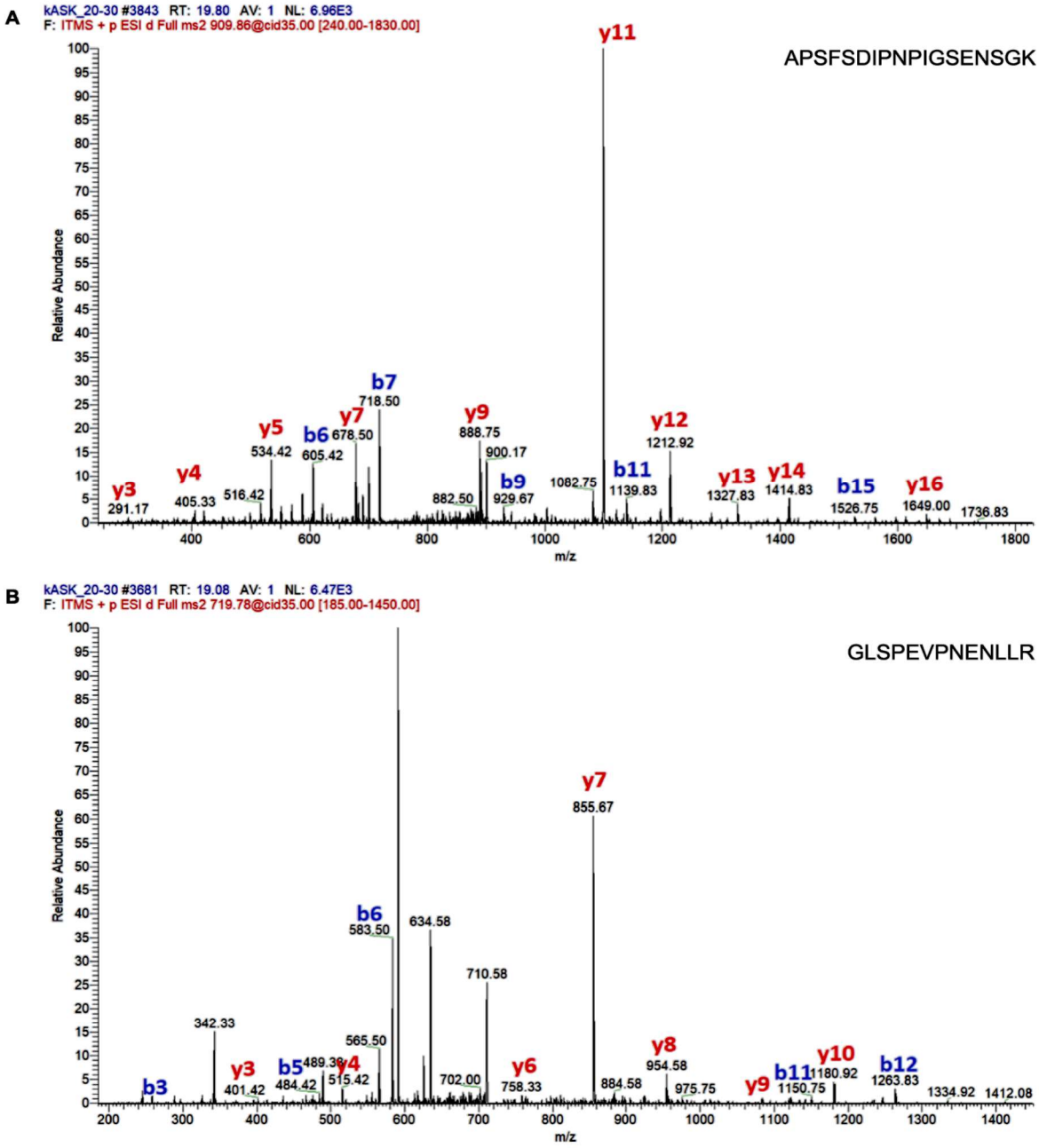


Fig. 4. Representative MS/MS fragmentation spectra of caseicin C precursor peptide (APSFSDIPNPIGSENSGK) (A) and Isracidin peptide fragment (GLSPEVPNENLLR) (B).

3.5. KE promotes HaCaT cells proliferation and migration

The proliferation and migration of human keratinocytes are critical factors contributing to efficient wound healing. Several growth factors and cytokines have been shown to influence these processes and can contribute to the healing response (Kim et al., 2018).

Scratch assays and MTT tests were conducted to evaluate the effect of the KE peptide extract in the healing process *in vitro*. Human keratinocyte monolayers were scratched and stimulated with TNF-α for 3 h to induce inflammation. After 6 h of treatment with KE peptide extract, a significant rate of cell migration was observed. The migration percentages indicated in Fig. 6 were calculated taking as reference the migration rate of keratinocytes not exposed to the extract. HaCaT cells treated with TNF-α had a lower recovery in the scratch area after 6 h (% open wound 75%) than control cells (% open wound 30%). In contrast, cells treated with KE in the presence of TNF-α showed a greater recharge of the scratch area of approximately 50% compared to cells treated with TNFα. After 18 h of exposure, wound healing rates were 60% and 75% for for TNF-α stimulated keratinocytes and and inflamed cells treated with KE peptide extract, respectively.

Furthermore, proliferation of keratinocytes stimulated with TNF-α and treated with KE was evaluated. KE promoted a significant increase

Table 3
Antimicrobial peptides identified in 20–30 and 30–40 fractions of traditional KE.

Fraction (min)	Peptide	Uniprot ID	−10lgP	Mw (Da)	Protein name	Bioactive compound	Active against
20–30	D.APSFSDIPNPIGSENSGKT.T	Q69EZ6_CAPHI	45,88	1.916,9115	Alpha-S1-casein	caseicin C (precursor)	Gram +
	D.APSFSDIPNPIGSEN.S	Q69EZ6_CAPHI	42,78	1.543,7155	Alpha-S1-casein	caseicin C (fragment)	Gram +
	D.IPNPIGSENSGKTTMPL.W	Q69EZ6_CAPHI	42,11	1.754,8872	Alpha-S1-casein	caseicin C (fragment)	Gram +
	D.APSFSDIPNPIGSENSGK.T	Q69EZ6_CAPHI	39,88	1.815,8639	Alpha-S1-casein	caseicin C (precursor)	Gram +
	D.APSFSDIPNPIGSE.N	Q69EZ6_CAPHI	38,47	1.429,6725	Alpha-S1-casein	caseicin C (fragment)	Gram +
	F.LPYPYYAKPV.A	A0A452G9D9_CAPHI	38,46	1.209,6433	Kappa-casein		
	F.LPYPYYAKPIA.V	K0P739_CAPHI	37,56	1.294,6960	Kappa-casein (Fragment)		
	D.APSFSDIPNPIGSENSG.K	Q69EZ6_CAPHI	34,74	1.687,7689	Alpha-S1-casein	caseicin C (fragment)	Gram +
	Q.GLSPEVPNENLLR.F	Q69EZ6_CAPHI	34,22	1.436,7623	Alpha-S1-casein	caseicin B (precursor)	Gram+ and Gram −
	F.LPYPYYAKPI.A	K0P739_CAPHI	33,58	1.223,6589	Kappa-casein (Fragment)		
30–40	L.YQEPVLGPVRGPFPILV	Q712N8_CAPHI	50,29	1.880,0559	Beta-casein	casecidin 17	Gram −
	D.IPNPIGSENSGKTTMPLW	I6WY32_CAPHI	49,47	1.940,9666	Alpha-S1-casein	caseicin C (fragment)	Gram +
	L.YQEPVLGPVRGPFPIL.V	Q712N8_CAPHI	48,26	1.780,9875	Beta-casein	casecidin 17 (fragment)	Gram −
	L.YQEPVLGPVRGPFP.I	Q712N8_CAPHI	47,82	1.554,8193	Beta-casein	casecidin 15	Gram −
	D.IPNPIGSENSGKITMPLW	D3TU01_SHEEP	47,19	1.953,0029	Alpha-S1-casein	caseicin C (fragment)	Gram +
	D.IPNPIGSENSGKTTM(+15.99) PLW	I6WY32_CAPHI	46,98	1.956,9615	Alpha-S1-casein	caseicin C (fragment)	Gram +
	D.APSFSDIPNPIGSENSGKTTMPLW	I6WY32_CAPHI	45,99	2.545,2158	Alpha-S1-casein	caseicin C (precursor)	Gram +
	L.YQEPVLGPVRGPFPI.L	Q712N8_CAPHI	45	1.667,9034	Beta-casein	casecidin 15	Gram −
	L.LYQEPVLGPVRGPFPILV	Q712N8_CAPHI	43,96	1.993,1400	Beta-casein	casecidin 17 (precursor)	Gram −
	N.NQFLPYPYYAKPV.A	A0A452G9D9_CAPHI	43,77	1.598,8132	Kappa-casein		
	N.PIGSENSGKTTMPLW	I6WY32_CAPHI	41,86	1.616,7869	Alpha-S1-casein	caseicin C (fragment)	Gram +
	E.PVLGPVRGPFPILV	Q712N8_CAPHI	41,04	1.459,8915	Beta-casein	casecidin 17 (fragment)	Gram −
	D.APSFSDIPNPIGSENSGKITMPLW	D3TU01_SHEEP	40,9	2.557,2522	Alpha-S1-casein	caseicin C (precursor)	Gram +
	Q.FLPYPYYAKPV.A	A0A452G9D9_CAPHI	37,26	1.356,7118	Kappa-casein		
	D.APSFSDIPNPIGSE.N	I6WY32_CAPHI	36,14	1.429,6725	Alpha-S1-casein	caseicin C (fragment)	Gram +
	L.GPVRGPFPIL.V	Q712N8_CAPHI	34,15	1.051,6178	Beta-casein	casecidin 17 (fragment)	Gram −
	F.VVAPFPEVFR.K	I6WY32_CAPHI	32,34	1.159,6389	Alpha-S1-casein		
	G.PVRGPFPILV	Q712N8_CAPHI	32	1.093,6648	Beta-casein	casecidin 17 (fragment)	Gram −
	N.SGKTTMPLW	I6WY32_CAPHI	31,95	1.019,5110	Alpha-S1-casein		
	L.GPVRGPFPILV	Q712N8_CAPHI	30,76	1.150,6862	Beta-casein	casecidin 17 (fragment)	Gram −

in viability: there was a 1.9-fold increase when cells were treated with 250 μg/mL under normal growth conditions; a 2.3-fold increase occurred when cells were stimulated with TNF-α and treated with 250 μg/ml. Taken together, the scratch and vitality results suggested that KE-treated keratinocytes would have major wound healing performance in damaged epithelia and would positively contribute to tissue recovery (Fig. 6).

4. Discussion

AD is a chronic inflammatory skin disease that affects a large proportion of the world's population. This condition causes significant morbidity, having a negative effect on the life quality of patients (Kapur et al., 2018). AD pathogenesis is the result of complex interactions between the alteration of the skin barrier and immune dysregulation (Yang et al., 2020). These defective conditions contribute to increased colonization by *S. aureus*. The increased adhesion of *S. aureus* is due to cutaneous fibronectin and fibrinogen, associated with AD skin lesions. Moreover, endogenous antimicrobial peptides characteristic of innate immunity are less abundant in atopic skin, explaining the increased susceptibility of AD patients to *S. aureus* infections (Nowick & Grywalska, 2018). These bacteria pass through the subcutaneous tissues, causing abscesses and bloodstream infection. Colonization promotes flare-up of AD, increasing the inflammatory state and severity of the lesions (Alexander et al., 2020). Long-term use of topical or systemic antibiotics does not entail benefits in AD patients, posing risks of adverse effects, such as contact dermatitis and the emergence of resistant bacteria (Frazier & Bhardwaj, 2020). Due to limited access to hospital treatments for AD patients, novel therapies that accelerate wound closure and as well as prevent infections trigger a great interest. Novel approaches in AD management are therefore required. The growing interest in finding new strategies for the AD treatment has stimulated the research for bioactive compounds from natural sources (Wang et al., 2021). Dairy products are considered a rich source of bioactive peptides. Endogenous milk proteases, bacterial proteases and product processing conditions act in synergy to form bioactive peptides from the degradation of milk proteins. Theese peptides possess a wide variety of biological activities, such as antimicrobial, antioxidative, healing, anticytotoxic and antihypertensive (Jakubczyk, et al. 2020).

In this scenario, the peptides obtained from kashk, a Middle East fermented dairy product, were studied. Samples of traditional dry and industrial liquid kashk produced in Iran were collected at local markets. The peptides were extracted with water since this methodology has proven the most effective for oligopeptide extraction from kashk (Pourjoula et al., 2020). We first evaluated the traditional and industrial KE ability to directly inhibit the growth of *S. aureus* strains, using the broth microdilution method to define the rate of bacterial growth inhibition. Most of the strains used were isolated from skin swabs from AD patients and selected based on the most common mechanisms of antibiotic resistance. Industrial KE did not exhibit antibacterial activity, unlike the traditional one. Instead, traditional KE inhibits the growth of *S. aureus* strains by approximately 45%, indicating a generalized action. The antibacterial potential of several fermented milk products is well described. Théolier et al. (2014) assessed the antibacterial properties of

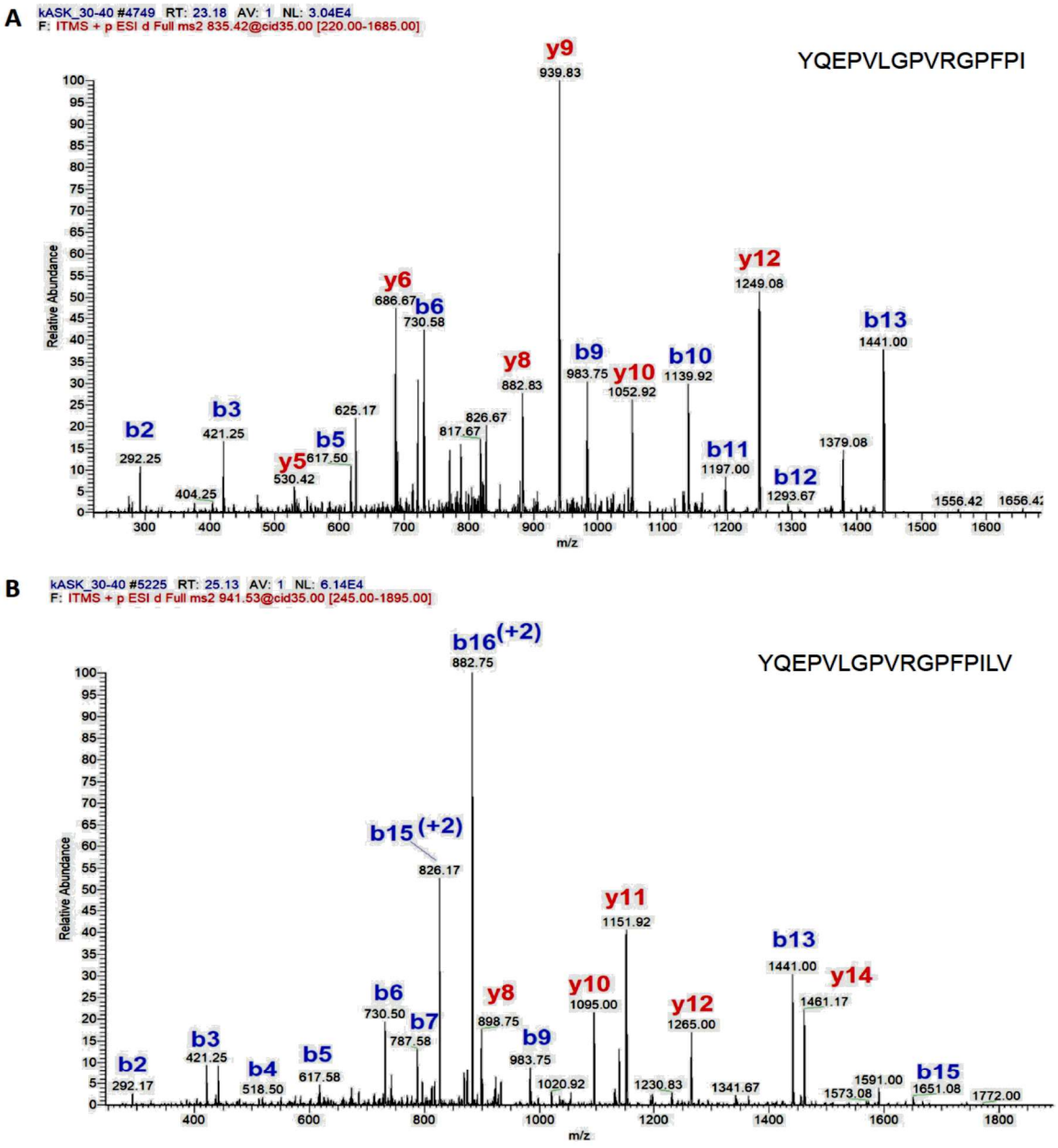


Fig. 5. Representative MS/MS fragmentation spectra of caseicidin 15 (YQEPVLGPVRGPFPI) (A) and Caseicidin 17 (YQEPVLGPVRGPFPILV) (B).

gouda cheese. Gouda cheese peptides totally inhibited the growth of *Listeria ivanovii* HPB28, *Listeria monocytogenes* Scott A3, *Escherichia coli* MC4100 and *Escherichia coli* O157:H7 (ATCC 35150) at 8.5, 34, 17 and 34 mg/mL, respectively (Théolier et al., 2014). De Lima et al. reported that peptide extract from sheep's milk fermented in Brazilian kefir induced growth inhibition >87.80% for *Escherichia coli* (ATCC 25922), *Klebsiella pneumoniae* (ATCC 29665), *Pseudomonas aeruginosa* (ATCC 27853), *Bacillus cereus* (ATCC 33019), *Bacillus subtilis* (ATCC 6633) and *S. aureus* (ATCC 6538) (de Lima et al., 2018). Moreover, antibacterial potentials of bioactive peptides in buffalo yoghurt were evaluated by Taha et al. (2017). The peptide fraction at a concentration of 50 mg/mL generated inhibited the growth of *Escherichia coli* (ATCC 2592), *S. aureus* (ATCC 6538), and *Salmonella typhimurium* (ATCC 9027), *Bacillus cereus* (ATCC 33019) and *Pseudomonas aeruginosa* (ATCC 27853) (Taha et al., 2017). To our best knowledge, this study is the first evidence of the antibacterial potential of the peptides of kashk

Recently, Pourjula et al. characterized the kashk peptidome predicting *in-silico* the presence of potential bioactive sequences (Pourjoula et al., 2020). The study remarked that bacterial proteolysis production of kashk represented the key step in production of bioactive peptides. In this work, we firstly demonstrated that the kashk type that exhibits antimicrobial activity is the traditional one. Notably, in the Middle-East, kashk is traditionally produced by inducing a spontaneous fermentation of dairy waste from caprine and ovine milk origin. The fermentation and the harsh sun-drying undoubtedly affected the development of an extremely complex peptidome that is much more abundant and heterogeneous than that of the industrial KE (Pourjoula et al., 2020). This extreme proteolysis likely led to the formation of the bioactive peptides

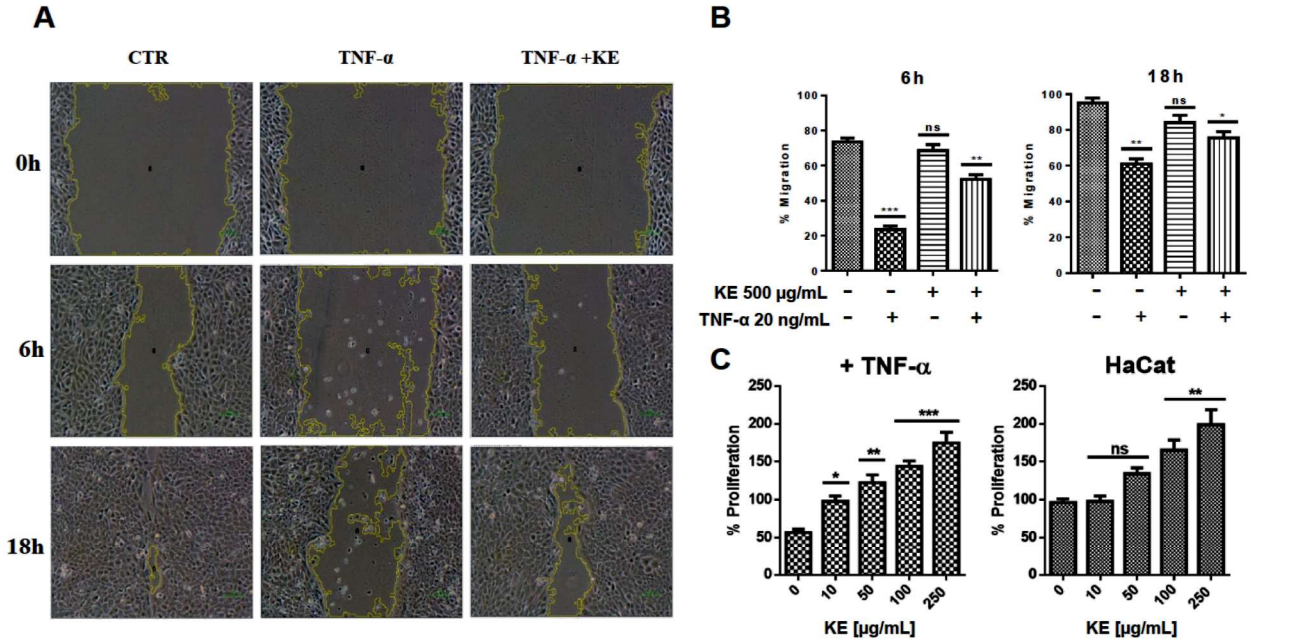


Fig. 6. Wound-healing activity of the traditional KE. The yellow line denoted the margin of the gap (A); Migration rate (%) of HaCaT cell in response to treatment with traditional KE extract and TNF-α for 6 h, ** $p < 0.0048$, *** $p < 0.0007$, ns, and 18 h, * $p < 0.0237$, ** $p < 0.0069$, ns (B); Proliferation rate (%) of HaCaT cell after 24 h of exposure with traditional KE extract in the presence of TNF-α, * $p < 0.0241$, ** $p < 0.0051$, *** $p < 0.0006$, and in the absence of TNF-α, ** $p < 0.0069$, ns (C). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

we observed in our study. On the other hand, industrial liquid kashk is commonly produced by industrial yogurt from cow milk, with use of selected starter strains, which led to a milder proteolysis (Say et al., 2015). Therefore, traditional and industrial kashk are products with very different composition, including the peptidome. The different production approaches, leading to a very different peptidome, determine the observed differences in functional properties.

We firstly fractionated the total water-soluble extract from traditional KE by RP-HPLC, to characterize at the molecular level the chromatographic fractions with the highest antibacterial activity. The resulting fractions were evaluated for their antibacterial activity (250 µg/mL). The antibacterial potential of the 20–30 and 30–40 min fraction was comparable to that of the total extract. These fractions inhibited the growth of all *S. aureus* strains at a average values of 35.4 and 24.2%, respectively.

These results on the bioactivity of the KEs prompted us to undertake a structural characterization study of peptides in the active fractions in order to identify the possible structure/function relationship. In this respect, the above mentioned heterogeneity of the fractions themselves was a challenge, for which a MS-based peptidomic approach was necessary. For decades, LC-MS/MS has been employed in the identification of peptides produced in dairy products (Ali et al., 2019; Savastano et al., 2019; Sforza et al., 2012; Ferranti et al., 2004), supporting our understanding of the structural basis of the observed bioactivity of dairy products. MS analysis also allows the comparison of different fermented products predicting in advance their potential bioactivity. In this respect, Pourjoula et al., (2020) performed an extensive characterization of both traditional dried and liquid industrial KEs, predicting some potential bioactivities of the peptide fraction. In this study, MS analysis allowed us to identify the bioactive peptides.

Once identified the peptides in the active fractions of traditional kashk extract, we carried out an *in silico* analysis to predict their antimicrobial activity, Bioactive precursors or portion fragments of casein, known to exert antimicrobial activity, were identified using a database of antimicrobial milk peptides (MilkAMP) (Nielsen et al., 2017). The active fractions contained peptides derived from the N- and C-terminal part of αS1-casein (caseicin B and C) and from the C-terminus of β-casein (caseicidin 15 and 17). The analysis identified peptides derived from caseicidin 15 and 17, active against Gram–, others from caseicin C, active against Gram+, and caseicin B derived peptides with action against both bacterial types. Hayes et al. reported that caseicin B-derived peptides exhibited greater antibacterial action against Gram− strains and lower against Gram+ strains (Hayes et al., 2006). In contrast, caseicin C was reported to have efficient antibacterial activity against Gram+ strains. Baranyi et al. indicated that caseicin C-derived peptides inhibited the growth of two Gram+ strains (*Bacillus subtilis* and *Staphylococcus lentus*) at a concentration of 3 mM but no inhibition against Gram− strains was observed (Baranyi et al., 2003). Birkemo et al. reported the antibacterial profile of caseicidin 15 and 17 limited to Gram− strains, recording a MIC value of 0.4 mg/mL on *Escherichia coli* (Birkemo et al., 2009). The increased activity recorded in the KE fraction eluted between 20 and 30 min in HPLC could be explained by the presence of a peptide population uniformly active against Gram+ bacteria, in contrast to the peptide fraction 30–40 min, also containing peptides active on Gram– strains. Regarding the structural basis of the antimicrobial activity, the structural characterization provided interesting indications. The high presence of hydrophobic amino acids, such as Pro and Iso, justifies the antibacterial efficacy of the KE peptide extract. In fact, Wang et al. (2018) argued the importance of hydrophobic amino acid residues for the design of novel peptides against *S. aureus* (Wang et al., 2018).

The wound repair process is divided in three sequential and overlapping steps: the inflammatory phase, the proliferative phase, and the remodeling phase (Cañedo-Dorantes & Cañedo-Ayala, 2019). Accumulating evidence reveals that in AD occurs a prolonged increase of inflammatory mediators leading an impaired would healing and an increased risk of bacterial colonization. Our results showed that KE accelerated skin repair in presence of exacerbated TNFα and promoted the transition of keratinocytes from a quiescent to a hyperproliferative phenotype.

Furthermore, we carried out keratinocyte stimulation with the KE

peptide extract to evaluate healing and proliferative ability in the presence of TNF-α. HaCat cell line treated with KE in the presence of TNF-α showed a healing potential of approximately 88% compared to cells stimulated only with TNF-α after 18 h of treatment. Moreover, an increase in cell viability of approximately 90% was observed following treatment with KE compared to the untreated control and in addition to TNF-α. Also this stimulating effect of KE on keratinocytes could be associated to the KE peptide composition. Huseini et al. reported the positive influence of kefir extract in a process of wound healing tested in experimental burns infected with *Pseudomonas aeruginosa* (ATCC 27853), although no structural evidence was proposed (Huseini et al., 2012). In the present study, the abundance of hydrophobic residues like Leu, Val, Pro and Tyr in the amino acid sequence of the identified peptides may be responsible for the proliferation of keratinocytes. Indeed, Amiot et al., (2004) found that hydrophobicity is an important property of milk protein hydrolysates to stimulate keratinocyte growth. Moreover, we have identified a peptide in the 30–40 fraction (LYQEPVLGPVRGPFPILV) that is a precursor of a bioactive peptide from beta-casein, which improves the growth of human keratinocytes in culture (LYQEPVL) (Fan et al., 2019).

5. Conclusions

AD is an inflammatory skin disease that often predisposes to colonization or infection by microbial organisms, particularly *S. aureus*. Our preliminary results showed that the KE water-soluble peptide fraction was able either to accelerate wound closure in an inflamed keratinocyte cellular model or to reduce the colonization rate of *S. aureus*. The data here presented form the basis for further investigation to fully clarify the mode of action of KE water-soluble peptide fraction both in the microbiological and tissue regenerative fields.

Author contribution

Veronica Folliero: Methodology, Writing. **Stefania Lama:** Methodology. **Gianluigi Franci:** Supervision. **Rosa Giugliano:** Methodology, Writing – original draft. **Pasquale Ferranti:** Supervision, Writing. **Paola Stiuso:** Supervision.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

Funding: This research received no external funding.

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.foodres.2022.110949.

References

Abdel-Salam, B. K. (2014). Modulatory effect of why proteins in some cytokines involved in wound healing in male diabetic albino rats. *Inflammation, 37*, 1616–1622. https://doi.org/10.1007/s10753-014-9888-z

Alexander, H., Paller, A. S., Traidl-Hoffmann, C., Beck, L. A., De Benedetto, A., Dhar, S., Girolomoni, G., Irvine, A. D., Spuls, P., Su, J., Thyssen, J. P., Vestergaard, C., Werfel, T., Wollenberg, A., Deleuran, M., & Flohr, C. (2020). The role of bacterial skin infections in atopic dermatitis: Expert statement and review from the International Eczema Council Skin Infection Group. *British Journal of Dermatology, 182*, 1331–1342. https://doi.org/10.1111/bjd.18643

Ali, E., Nielsen, S. D., Aal, A. E., El-Leboudy, A., Saleh, E., & LaPointe, G. (2019). Use of mass spectrometry to profile peptides in whey protein isolate medium fermented by Lactobacillus helveticus LH-2 and Lactobacillus acidophilus La-5. *Frontiers in nutrition, 6*, 152. https://doi.org/10.3389/fnut.2019.00152

Altomare, A., Fasoli, E., Colzani, M., Paredes Parra, X. M., Ferrari, M., Cilurzo, F., Rumio, C., Cannizzaro, L., Carini, M., Righetti, P. G., & Aldini, G. (2016). An in depth proteomic analysis based on ProteoMiner, affinity chromatography and nano-HPLC-MS/MS to explain the potential health benefits of bovine colostrum. *Journal of Pharmaceutical and Biomedical Analysis, 121*, 297–306. https://doi.org/10.1016/j.jpba.2016.01.013

Amiot, J., Germain, L., Turgeon, S., Lemay, M., Ory-Salam, C., & Auger, F. A. (2004). Peptides from milk protein hydrolysates to improve the growth of human keratinocytes in culture. *International Dairy Journal, 14*, 619–626. https://doi.org/10.1016/j.idairyj.2003.11.007

Baranyi, M., Thomas, U., & Pellegrini, A. (2003). Antibacterial activity of casein-derived peptides isolated from rabbit (Oryctolagus cuniculus) milk. *Journal of Dairy Research, 70*, 189–197. https://doi.org/10.1017/s0022029903006150

Beverly, R. L., Underwood, M. A., & Dallas, D. C. (2019). Peptidomics analysis of milk protein-derived peptides released over time in the preterm infant stomach. *Journal of Proteome Research, 18*, 912–922. https://doi.org/10.1021/acs.jproteome.8b00604

Birkemo, G. A., O'Sullivan, O., Ross, R. P., & Hill, C. (2009). Antimicrobial activity of two peptides casecidin 15 and 17, found naturally in bovine colostrum. *Journal of Applied Microbiology, 106*, 233–240. https://doi.org/10.1111/j.1365-2672.2008.03996.x

Blicharz, L., Rudnicka, L., & Samochocki, Z. (2019). Staphylococcus aureus: An underestimated factor in the pathogenesis of atopic dermatitis?*Advances in. Dermatology and Allergology, 36*, 11–17. https://doi.org/10.5114/ada.2019.82821

Buonocore, C., Tedesco, P., Vitale, G. A., Esposito, F. P., Giugliano, R., Monti, M. C., D'Auria, M. V., & de Pascale, D. (2020). Characterization of a New Mixture of Mono-Rhamnolipids Produced by *Pseudomonas gessardii* Isolated from Edmonson Point (Antarctica). *Marine Drugs, 18*, 269. https://doi.org/10.3390/md18050269

Cañedo-Dorantes, L., & Cañedo-Ayala, M. (2019). Skin acute wound healing: A comprehensive review. *eCollection*. https://doi.org/10.1155/2019/3706315

Carrera, Y. I. L., Hammadi, A. A., Huang, Y. H., Llamado, L. J., Mahgoub, E., & Tallman, A. M. (2019). Epidemiology, diagnosis, and treatment of atopic dermatitis in the developing countries of Asia, Africa, Latin America, and the middle east: A review. *Dermatology and Therapy, 9*, 685–705. https://doi.org/10.1007/s13555-019-00332-3

Chakrabarti, S., Guha, S., & Majumder, K. (2018). Food-derived bioactive peptides in human health: Challenges and opportunities. *Nutrients, 10*, 1738. https://doi.org/10.3390/nu10111738

Clare, D. A., Catignani, G. L., & Swaisgood, H. E. (2003). Biodefense properties of milk: The role of antimicrobial proteins and peptides. *Current Pharmaceutical Design, 9*, 1239–1255. https://doi.org/10.2174/1381612033454874

de Lima, M. D. S. F., da Silva, R. A., da Silva, M. F., da Silva, P. A. B., Costa, R. M. P. B., Teixeira, J. A. C., Porto, A. L. F., & Cavalcanti, M. T. H. (2018). Brazilian kefir-fermented sheep's milk, a source of antimicrobial and antioxidant peptides. *Probiotics Antimicrobial Proteins, 10*, 446–455. https://doi.org/10.1007/s12602-017-9365-8

Ebaid, H., Abdel-Salam, B. K., Hassan, I., Al-Tamimi, J., Metwalli, A., & Alhazza, I. (2015). Camel milk peptide improves wound healing in diabetic rats by orchestrating the redox status and immune response. *Lipids in Health and Disease, 14*, 132. https://doi.org/10.1186/s12944-015-0136-9

Fan, M., Guo, T., Li, W., Chen, J., Li, F., Wang, C., Shi, Y., Li, D. Y., & Zhang, S. (2019). Isolation and identification of novel casein-derived bioactive peptides and potential functions in fermented casein with Lactobacillus helveticus. *Food Science and Human Wellness, 8*, 156–176. https://doi.org/10.1016/j.fshw.2019.03.010

Ferranti, P., Barone, F., Chianese, L., Addeo, F., Scaloni, A., Pellegrino, L., & Resmini, P. (1997). Phosphopeptides from Grana Padano cheese: Nature, origin and changes during ripening. *Journal of Dairy Research, 64*, 601–615. https://doi.org/10.1017/s0022029997002392

Ferranti, P., Itolli, E., Barone, F., Malorni, A., Garro, G., Laezza, P., Chianese, L., Migliaccio, F., Stingo, V., & Addeo, F. (1997). Combined high resolution separation techniques (FPLC and HPLC) and mass spectrometry- based identification of peptides and proteins in Grana Padano cheese. *Le Lait, 77*, 683–697. https://doi.org/10.1051/lait:1997649

Ferranti, P., Traisci, M. V., Picariello, G., Nasi, A., Boschi, V., Siervo, M., Falconi, C., Chianese, L., & Addeo, F. (2004). Casein proteolysis in human milk: Tracing the pattern of casein breakdown and the formation of potential bioactive peptides. *The Journal of Dairy Research, 71*, 74–87. https://doi.org/10.1017/s0022029903006599

Folliero, V., Caputo, P., Della Rocca, M. T., Chianese, A., Galdiero, M., Iovene, M. R., Hay, C., Franci, G., & Galdiero, M. (2020). Prevalence and antimicrobial susceptibility patterns of bacterial pathogens in urinary tract infections in University Hospital of Campania "Luigi Vanvitelli" between 2017 and 2018. *Antibiotics (Basel), 9*, 2015. https://doi.org/10.3390/antibiotics9050215

Franci, G., Folliero, V., Cammarota, M., Zannella, C., Sarno, F., Schiraldi, C., de Lera, A. R., Altucci, L., & Galdiero, M. (2018). Epigenetic modulator UVI5008 inhibits MRSA by interfering with bacterial gyrase. *Scientific Reports, 8*, 13117. https://doi.org/10.1038/s41598-018-31135-9

Frazier, W., & Bhardwaj, N. (2020). Atopic dermatitis: Diagnosis and treatment. *American Family Physician, 101*, 590–598.

Guha, S., Sharma, H., Deshwal, G. K., & Rao, P. S. (2021). A comprehensive review on bioactive peptides derived from milk and milk products of minor dairy species. *Food Production, Processing and Nutrition, 3*, 2. https://doi.org/10.1186/s43014-020-00045-7

Hayes, M., Ross, R. P., Fitzgerald, G. F., Hill, C., & Stanton, C. (2006). Casein-derived antimicrobial peptides generated by Lactobacillus acidophilus DPC6026. *Applied and Environmental Microbiology Journal, 72*, 2260–2264. https://doi.org/10.1128/AEM.72.3.2260-2264.2006

Hancock, R. E., & Lehrer, R. (1998). Cationic peptides: A new source of antibiotics. *Trends in Biotechnology, 16*, 82–88. https://doi.org/10.1016/s0167-7799(97)01156-6

Hill, R. D., Lahav, E., & Givol, D. (1974). A rennin-sensitive bond in alpha-s1 B-casein. *Journal of Dairy Research, 41*, 147–153. https://doi.org/10.1017/s0022029900015028

Huseini, F. H., Rahimzadeh, G., Fazeli, M. R., Mehrazma, M., & Salehi, M. (2012). Evaluation of wound healing activities of kefir products. *Burns, 38*, 719–723. https://doi.org/10.1016/j.burns.2011.12.005

Iwamoto, K., Moriwaki, M., Miyake, R., & Hide, M. (2019). Staphylococcus aureus in atopic dermatitis: Strain-specific cell wall proteins and skin immunity. *Allergology International, 3*, 309–315. https://doi.org/10.1016/j.alit.2019.02.006

Jakubczyk, A., Karáś, M., Kamila Rybczýnska-Tkaczyk, K., Zieli´nska, E., & Zieli´nski, D. (2020). Current trends of bioactive peptides—New Sourcesand therapeutic effect. *Food, 9*, 846. https://doi.org/10.3390/foods9070846

Kapur, S., Watson, W., & Carr, S. (2018). Atopic dermatitis. *Allergy, Asthma & Clinical Immunology, 14*, 52. https://doi.org/10.1186/s13223-018-0281-6

Karami, Z., & Akbari-adergani, B. (2019). Bioactive food derived peptides: A review on correlation between structure of bioactive peptides and their functional properties. *Journal of Food Science and Technology, 56*, 535–547. https://doi.org/10.1007/s13197-018-3549-4

Kent, R. M., Guinane, C. M., O'Connor, P. M., Fitzgerald, G. F., Hill, C., Stanton, C., & Ross, R. P. (2012). Production of the antimicrobial peptides Caseicin A and B by Bacillus isolates growing on sodium caseinate. *Letters in Applied Microbiology, 55*, 141–148. https://doi.org/10.1111/j.1472-765X.2012.03271.x

Kim, S., Lee, S. K., Kim, H., & Kim, T. M. (2018). Exosomes secreted from induced pluripotent stem cell-derived mesenchymal stem cells accelerate skin cell proliferation. *International Journal of Molecular Science, 19*, 3119. https://doi.org/10.3390/ijms19103119

Lee, S. Y., Lee, E., Park, Y. M., & Hong, S. J. (2018). Microbiome in the gut-skin axis in atopic dermatitis. *Allergy, Asthma & Immunology Research, 10*, 354–362. https://doi.org/10.4168/aair.2018.10.4.354

McDonnell, M. J., Rivas, L., Burgess, C. M., Fanning, S., & Duffy, G. (2012). Inhibition of verocytotoxigenic Escherichia coli by antimicrobial peptides caseicin A and B and the factors affecting their antimicrobial activities. *International Journal of Food Microbiology, 153*, 260–268. https://doi.org/10.1016/j.ijfoodmicro.2011.10.001

Moghadam, S. E., Moridi Farimani, M., Soroury, S., Ebrahimi, S. N., & Jabbarzadeh, E. (2019). Hypermongone C accelerates wound healing through the modulation of inflammatory factors and promotion of fibroblast migration. *Molecules (Basel, Switzerland), 24*, 2022. https://doi.org/10.3390/molecules24102022

Mohanty, D., Jena, R., Choudhury, P. K., Pattnaik, R., Mohapatra, S., & Saini, M. R. (2016). Milk derived antimicrobial bioactive peptides: A review. *International Journal of Food Properties, 12*, 837–846. https://doi.org/10.1080/10942912.2015.1048356

Nielsen, S. D., Beverly, R. L., Qu, Y., & Dallas, D. C. (2017). Milk bioactive peptide database: A comprehensive database of milk protein-derived bioactive peptides and novel visualization. *Food Chemistry, 232*, 673–682. https://doi.org/10.1016/j.foodchem.2017.04.056

Norberg, S., O'Connor, P. M., Stanton, C., Ross, R. P., Hill, C., Fitzgerald, G. F., & Cotter, P. D. (2011). Altering the composition of caseicins A and B as a means of determining the contribution of specific residues to antimicrobial activity. *Applied and Environmental Microbiology Journal, 77*, 2496–2501. https://doi.org/10.1128/AEM.02450-10

Nowick, D., & Grywalska, E. (2018). The role of immune defects and colonization of staphylococcus aureus in the pathogenesis of atopic dermatitis. *Analytical Cellular Pathology (Amsterdam), 2018*, 1956403. https://doi.org/10.1155/2018/1956403

Ogonowska, P., Gilaberte, Y., Barańska-Rybak, W., & Nakonieczna, J. (2020). Colonization with staphylococcus aureus in atopic dermatitis patients: Attempts to reveal the unknown. *Frontiers in Microbiology, 11*. https://doi.org/10.3389/fmicb.2020.567090

Pironti, C., Dell'Annunziata, F., Giugliano, R., Folliero, V., Galdiero, M., Riccardi, M., Motta, O., Proto, A., & Franci, G. (2021). Comparative analysis of peracetic acid (PAA) and permaleic acid (PMA) in disinfection processes. *Science of the Total Environment, 797*, 149206. https://doi.org/10.1016/j.scitotenv.2021.149206

Pourjoula, M., Picariello, G., Garro, G., D'Auria, G., Nitride, C., Rheza Ghaisari, A., & Ferranti, P. (2020). The protein and peptide fractions of kashk, a traditional Middle East fermented dairy product. *Food Research International, 132*, 109107. https://doi.org/10.1016/j.foodres.2020.109107

Punia, H., Tokas, J., Malik, A., Sangwan, S., Baloda, S., Singh, N., Singh, S., Bhuker, A., Singh, P., Yashveer, S., Agarwal, S., & Mor, V. S. (2020). Identification and detection of bioactive peptides in milk and dairy products: Remarks about agro-foods. *Molecules, 25*, 3328. https://doi.org/10.3390/molecules25153328

Rizzello, C. G., Losito, I., Gobbetti, M., Carbonara, T., De Bari, M. D., & Zambonin, P. G. (2005). Antibacterial activities of peptides from the water-soluble extracts of Italian cheese varieties. *Journal of Dairy Science, 88*, 2348–2360. https://doi.org/10.3168/jds.S0022-0302(05)72913-1

Savastano, M. L., Liu, Y., Mels, J., Dittrich, D., Haus, S., Gensberger-Reigl, S., & Pischetsrieder, M. (2019). Profiling of multiphosphorylated peptides in kefir and their release during simulated gastrointestinal digestion. *ACS omega, 4*, 7963–7970. https://doi.org/10.1021/acsomega.8b03105

Say, D., Soltani, M., & Güzeler, N. (2015). Dried yoghurts: Kurut and kashk. Pamukkale University. *Journal of Engineering Sciences, 21*(9), 428–432. https://doi.org/10.5505/pajes.2015.47048

Seiti Yamada Yoshikawa, F., Feitosa de Lima, J., Notomi Sato, M., Alefe Leuzzi Ramos, Y., Aoki, V., & Leao Orfali, R. (2019). Exploring the role of staphylococcus aureus toxins in atopic dermatitis. *Toxins (Basel), 11*, 408. https://doi.org/10.3390/toxins11060321

Serretiello, E., Astorri, R., Chianese, A., Stelitano, D., Zannella, C., Folliero, V., Santella, B., Galdiero, M., Franci, G., & Galdiero, M. (2020). The emerging tick-borne Crimean-Congo haemorrhagic fever virus: A narrative review. *Travel Medicine and Infectious Disease, 37*, 101871. https://doi.org/10.3390/toxins11060321

Sforza, S., Cavatorta, V., Lambertini, F., Galaverna, G., Dossena, A., & Marchelli, R. (2012). Cheese peptidomics: A detailed study on the evolution of the oligopeptide fraction in Parmigiano-Reggiano cheese from curd to 24 months of aging. *Journal of Dairy Science, 95*, 3514–3526. https://doi.org/10.3168/jds.2011-5046

Song, Y. W., Zhang, X., Bian, W., Liu, N., Yin, S., Yang, M., Luo, M., Tang, J., & Yang, X. (2019). A short peptide potentially promotes the healing of skin wound. *Bioscience Reports, 39*, 20181734. https://doi.org/10.1042/BSR20181734

Taha, S., El Abd, M., De Gobba, C., Abdel-Hamid, M., Khalil, E., & Hassan, D. (2017). Antioxidant and antibacterial activities of bioactive peptides in buffalo's yoghurt fermented with different starter cultures. *Food Science And Biotechnology, 26*, 1325–1332. https://doi.org/10.1007/s10068-017-0160-9

Théolier, J., Hammami, R., Fliss, I., & Jean, J. (2014). Antibacterial and antifungal activity of water-soluble extracts from Mozzarella, Gouda, Swiss, and Cheddar commercial cheeses produced in Canada. *Dairy Science & Technology, 94*, 427–438. https://doi.org/10.1007/s13594-014-0170-9

Tulipano, G. (2020). Role of bioactive peptide sequences in the potential impact of dairy protein intake on metabolic health. *International Journal of Molecular Science, 21*, 8881. https://doi.org/10.3390/ijms21228881

Wakabayashi, H., Takase, M., & Tomita, M. (2003). Lactoferricin derived from milk protein lactoferrin. *Current Pharmaceutical Design, 9*, 1277–1287. https://doi.org/10.2174/1381612033454829

Wang, X., Mishra, B., Lushnikova, T., Narayana, J. L., & Wang, G. (2018). Amino acid composition determines peptide activity spectrum and hot-spot-based design of merecidin. *Advanced Biosystems, 2*, 1700259. https://doi.org/10.1002/adbi.201700259

Wang, V., Boguniewicz, J., Boguniewicz, M., & Ong, P. Y. (2021). The infectious complications of atopic dermatitis. *Annals of Allergy, Asthma & Immunology, 126*, 3–12. https://doi.org/10.1016/j.anai.2020.08.002

Yang, G., Seok, J. K., Kang, H. C., Cho, Y. Y., Lee, H. S., & Lee, J. Y. (2020). Skin barrier abnormalities and immune dysfunction in atopic dermatitis. *International Journal of Molecular Science, 21*, 2867. https://doi.org/10.3390/ijms21082867

SECTION 4

Less relevant PhD publications

This section aims to give an overview of the other research activities performed during the three years of the PhD school, which are not very related to the main purpose of this thesis.

Following are presented published or submitted manuscripts which aimed at giving general information about other typologies of novel food (e.g., micro-algae), or in discovering their potential application and bioactivities.

Chapter 8

Microalgal Biomass Recycling: From Filter to Feed

Microalgal Biomass Recycling: From Filter to Feed

Giovanni D'Auria, Department of Agricultural Sciences, University of Naples Federico II, Naples, Italy

Aquaculture and Its Environmental Impact

Aquaculture is an ancient activity probably born in China, where it was tradition to bred various species of carp in rice fields (Sachs, 2007). So far China represents two-thirds of the total amount of aquaculture production in the world in terms of weight and market value. Aquaculture is an ever-growing field of food production. Indeed, while in 1970s aquaculture provided only 7% of fish for human consumption, this share was increased to 26% in 1994 and to 39% in 2004, with superior performance compared to both fish capture and animal farming for meat production. This way, aquaculture has guaranteed the increase in fish consumption by the second half of last century, from 9.9 kg *per capita* in the 1960s to 14.4 kg in the 1990s and 19.7 kg in 2013. This value is destined to increase, crossing the line of 20 kg *per capita* (FAO, 2016).

Unlike fishing, in aquaculture man just picks up the necessary quantity of fish requested by the market out of real natural tanks that can be found either in the sea, not so far from the coasts, or on the dry land. Those tanks can be considered as real fish-farms where all the breeding phases before the "harvest" are overseen (i.e. insemination of fish eggs, breeding of juvenile fish, etc.). However, the intensive development of aquaculture industry has contributed to increase its environmental impact. In the aquaculture systems, the residues of uneaten feed and the fish protein catabolism result in accumulation of nitrogen compounds such as ammonia (expelled by gills and urine), nitrites, nitrates and phosphorous. The accumulation of these compounds represents a real hazard for the health of farmed fish as well as an important pollutant. Furthermore, a high concentration of faeces in the water could easily increase the percentage of pathogenic microorganisms in the aquaculture systems. More than half of the nitrogen and phosphorous amount contained in feeds is not used and is released as a waste in the water. Because of the restrictions in the digestive capacity of fish, part of the feed eaten remains undigested and therefore is excreted in an unaltered state. In water, ammonia (NH_3) and ammonium ion (NH_4^+) are in equilibrium with each other, depending on pH and temperature. The sum of the two forms is called "Total Ammonia Nitrogen" (TAN). Ammonia (NH_3), thanks to its lipo-solubility, is the most toxic of the two forms. Indeed, it can pass through biological membranes and thus infiltrate in tissues more easily than ammonium (NH_4^+). The maximum permitted level of ammonia in aquaculture wastewater is 0.0125 ppm. Indeed, when ammonia exceeds this level, the fish might present damage to the kidneys, gills and liver tissue (Mugg, J. et al., 2000).

That said, in order to cleanse wastewater from organic and inorganic residues, it is important to develop efficient recycling methods for aquaculture systems, as they may offer an efficient strategy to treat wastewaters before these are reutilized or discharged. The present chapter deals with the disposal strategies for the major aquaculture sewage pollutants: the nitrogen compounds.

Background of the Present Systems for Water Depuration

Aquaculture sewage disposal is not only a serious and urgent issue but can be also considered as a precious opportunity to be exploited.

Traditionally, removal of nitrogen compounds is achieved by either using chemical-physical or biological treatments. Chemical processes, such as air stripping or selective ion exchange filters, have more advantages than biological treatments (e.g. lack of sensitivity to toxins and to fast temperature changes, etc.). Chemical processes also have some drawbacks like atmospheric pollution and the accumulation of high concentrated and toxic by-products. Moreover, these technologies are expensive and require continuous maintenance. Since their depuration yield rate is near to 95%–96%, biological filters represent the core of the aquaculture sewage disposal so far. The amount of nitrogen into aquaculture wastewater is regularized through a process called "nitrification" (Reaction 1). It is defined as the biological oxidation of the inorganic nitrogen compounds (NH_4^+ and NO_2^-) by nitrifying bacteria (nitrosants and nitricants). Whether the nitrifying activities are promoted by high temperature, thus at low temperature the reaction is decelerated, even tough not stopped. The useful thermal range for the best growth rate of nitrifying bacteria has an optimum between 25–35 °C. Therefore, under the same general conditions, the yield of bacterial species that are present in the filter will

be greater in thermophilic conditions. The pH hugely influences the nitrifying efficiency: the range of tolerance is between 6–9, but it can even stand at more extreme pH values (5–10); the optimum range is however achieved when the pH value is between 7–8.

$$2NH_4^+ + 3O_2 \rightarrow 2NO_2^- + 4H^+ + 2H_2O$$

$$2NO_2^- + O_2 \leftrightarrows 2NO_3^-$$

Reaction 1- General reaction of nitrification.

The nitrification process can be expressed in two steps:

- The first step, called nitrosation (as shown in Reaction 2), favours the ammonium oxidation and leads to the nitrite formation passing for the intermediate hydroxylamine (NH_2OH). Such process happens by the action of some bacteria called nitrosants. Among these, *Nitrosomonas* is the most present one.

$$NH_4^+ + \frac{1}{2}O_2 \rightarrow NH_2OH + H^+$$

$$NH_2OH + O_2 \rightarrow H^+ + H_2O + NO_2^-$$

Reaction 2- The first step of the nitrification reaction (nitrosation)

- The second oxidative step (Reaction 3), called nitration, leads to the transformation of nitrite to nitrate. The principal involved bacteria are known as nitricants. Among these, *Nitrospira* are the most representative bacteria (Smith et al., 2001).

The second step also takes place in an aerobic environment, but contrarily to the first one, this reaction is reversible. In fact, in anaerobic environment, it could happen that the reaction proceeds in reverse.

$$2NO_2^- + O_2 \leftrightarrows 2NO_3^-$$

Reaction 3- The second step of the Nitrification reaction (Nitrication)

To date, in the majority of installed systems, the nitrogen removal is done by biologic nitrification processes. Among these, the most famous methods are:

- Percolating filters;
- Disc-type biofilters;
- Fluidized bed filters.

Percolating filters. The percolation filter is a gravitational open filter where the flow is in a perfect contact with the filtering medium and the air. In this system, a spray rotative system, from which the water is distributed homogeneously throughout the filter section, is positioned on the top. During their fall, the water drops are in contact with the air that oxygenates them. The presence of the oxygen allows a high multiplication of nitrifying bacteria. These bacteria are situated in natural supports suitable for bacterial growth like porous stone particles, sand, gravel, or in synthesized material like plastic balls or polyurethane sponges. Water is distributed homogeneously through the filter section, wetting the surface, and constitutes the most part of the volume of filtering medium. The filter presents lateral aerators that contribute to aeration.

Disc-type biofilters. A device that often follows other simpler filtration typologies (e.g. decantation) is the disc-type biofilter (**Fig. 1**). As the name suggests, it is constituted by a number of bio-filters, in the shape of a disk, that are fixed on a rotating shaft, whose rotation speed can be varied. This filtration system is immersed for around half its diameter in the wastewater. This arrangement has been set to guarantee an appropriate aeration of bacterial biofilm. The "bio-disc" can be considered as an autonomous bioreactor on which nitrifying bacteria can thrive: the colonies together form a biofilm. The wastewater in which the bio-discs are embedded is stirred by the presence of aerators that prevent sedimentation of particulate. Moreover, this agitation allows the homogeneous distribution in the filtration tank of the oxygen accumulated both by the discs in the phase of emersion and by the air which gurgles in it.

Fluidized bed filters. The fluidized bed filters have the advantage of not getting clogged. These filters are made of a light substrate particles (sand, plastic, granular coal) that are kept in suspension by an ascendant continuous flow. A sand filter become fluidized when the speed of the water that flows through the "bed" is enough both to raise the grains of sand and to separate the grains each other. The flowing rate is function of the geometry, the dimensions and of the density of the particles. The fluidized bed filter has many excellent advantages. In comparison to other types of biofilters, it presents a large specific area where the nitrifying bacteria can implant themselves. Furthermore, owing to fact that they are shaped like a tall column, these filters require less space than others. The fluidized bed filters are self-cleaning and, moreover, are able to filter efficiently wastewaters that present high loads of toxic compounds. However, are also present several disadvantages like the relatively high energy demand of the system due to the high pressure drops across the filter: high pressures are needed for fluidizing the sand. Another main problem is related to the "age" of the bed, that is linked to the amount of bacterial biofilm present on the sand particles. Indeed, when the biofilm accumulates on the grains, the particle size increases and consequently the limiting factors previously observed change (i.e. geometry, size, density, etc.). In order to prevent the sand from coming out the tank, its capacity should be oversized and/or the flow of the water should be adjusted.

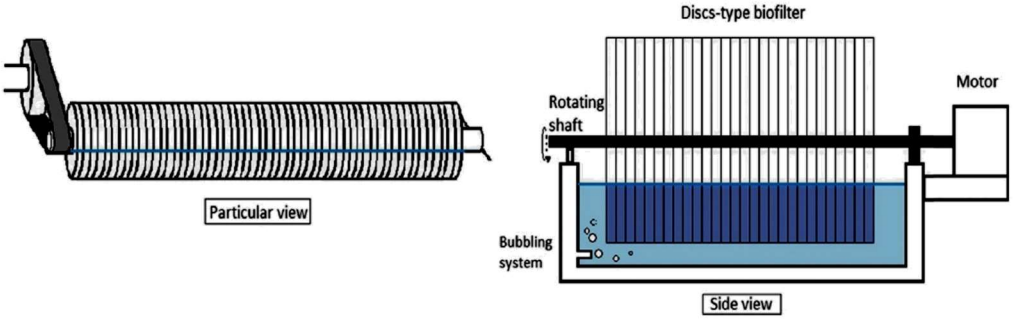


Figure 1 Disc-type biofilter.

The high number studies on the use of urban wastewater as a growing medium for the microalgal cultivation, have also stimulated another research area: filtration of the fish farms wastewater using microalgae. This idea includes the use of microalgae either in form of biofilm on special filters or in a free, in order to develop an integrate system where it is possible to see their coexistence with nitrifying bacteria. The nitrification, performed by nitrifying bacteria, as already mentioned, leads to the nitrate formation whose concentration in water can easily be lowered by microalgal action.

The Filtering Activity of Microalgae

As shown in Fig. 2, microalgae may be easily applied to reduce water nitrate concentration produced by nitrifying bacteria activity. This process occurs in their chloroplast, more specifically in the stroma (Fig. 3) where, through the catalytic action of various enzymes (i.e. nitrate reductase, nitrite reductase, glutamine synthetase, amino transferase), the synthesis of almost all of amino acids takes place. Thus, the utilization of microalgae in the wastewater recycle system may represent a winning bet both for the purification efficiency and from an economic point of view.

Microalgae, being autotrophic organisms, are able to use inorganic nutrients and to efficiently transform them in organic compounds (i.e. amino acids, proteins, carbohydrates, lipids, etc.). The studies of Gonzales et al. (1997), Martinez (2000), and Wang et al. (2009), have shown the high filtering and depurative capacity of microalgae for the purification of urban wastewaters. Consequently, this technology may easily translate in aquaculture wastewater treatments. Furthermore, in addition to depurative action, microalgae also show the power of lowering CO_2 and the oxygenating ability of the water that have to be re-sent to the fish breeding tanks. Indeed, several studies have already shown that microalgal growth in the aquaculture wastewater is possible (Guo et al., 2013; Van Den Hende et al., 2014). Microalgal species as *Chlorella* sp. and *Scenedesmus obliquus* have been widely utilized for the aquaculture wastewaters treatment showing an exceptional adaptability and an extraordinary ability to remove nitrogen and phosphorous. Chevalier and de la Noue (1985) have used two *Scenedesmus* species that have been immobilized in k-carrageenan for

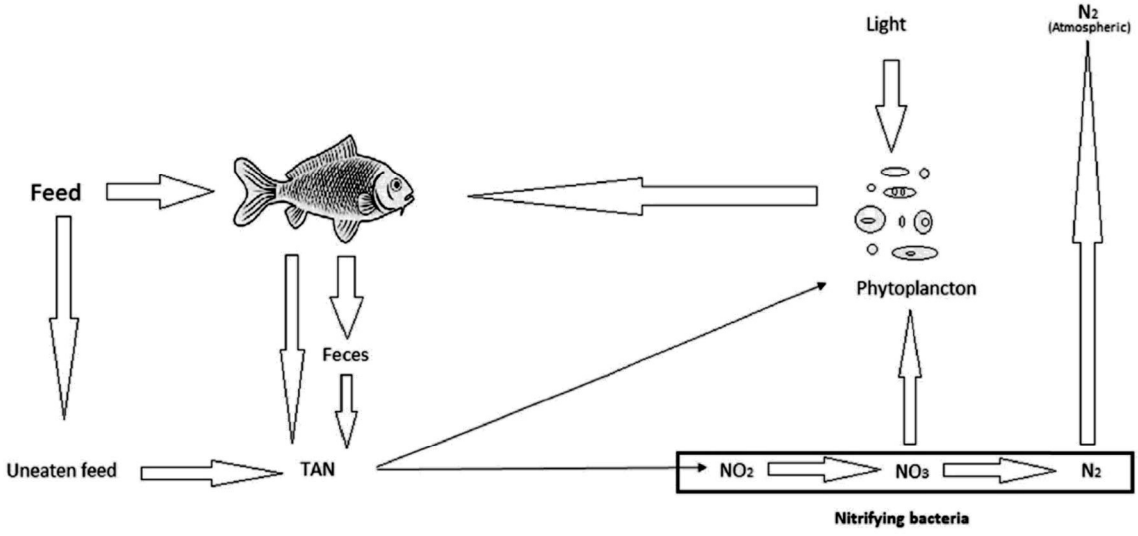


Figure 2 Nitrogen cycle in an aquatic system.

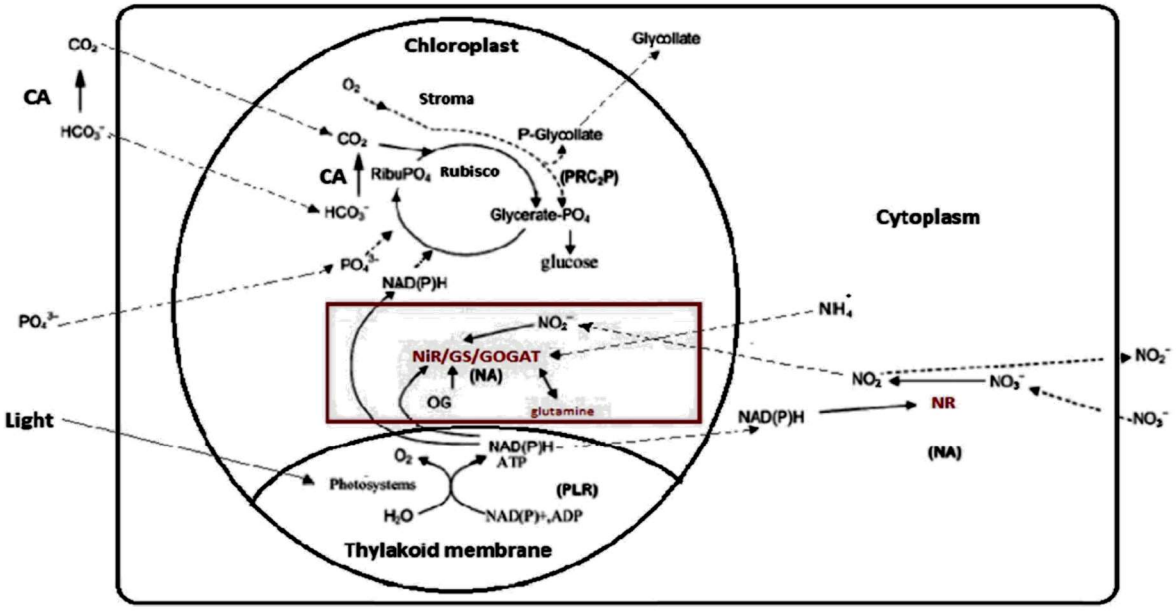


Figure 3 Biochemical pathways related to the using of nitrate.

Table 1 Changes in nitrogen compounds of *Sc. dimorphus* and *B. braunii* in aquaculture wastewater during the treatment period – Velichkova et al. (2018)

		Hydrochemical parameters (mean ± SEM)				
Hours	*Microalgae species*	*Ammonia*	*Nitrite – nitrogen*	*Nitrate - nitrogen*	*Total nitrogen*	*Phosphate*
0	*Sc. dimorphus*	0.301 ± 0.01^{ns}	0.08 ± 0.02^{ns}	1.8 ± 0.2^{ns}	18.7 ± 0.2^{ns}	1.4 ± 0.02^{ns}
	B. braunii	0.302 ± 0.02^{ns}	0.09 ± 0.01^{ns}	1.9 ± 0.1^{ns}	17.7 ± 0.9^{ns}	1.3 ± 0.01^{ns}
24	*Sc. dimorphus*	0.241 ± 0.01^{ns}	0.0241 ± 0.02^{ns}	1.75 ± 0.07^{ns}	12.25 ± 0.35^{ns}	1.36 ± 0.11^{ns}
	B. braunii	0.228 ± 0.08^{ns}	0.0218 ± 0.007^{ns}	1.6 ± 0.1^{ns}	11.75 ± 0.35^{ns}	1.17 ± 0.28^{ns}
96	*Sc. dimorphus*	0.202 ± 0.09^{ns}	0.0232 ± 0.007^{ns}	$1.3 \pm 0.14^{*}$	$11.1 \pm 0.28^{*}$	0.75 ± 0.71^{ns}
	B. braunii	0.172 ± 0.1^{ns}	0.0162 ± 0.007^{ns}	$0.95 \pm 0.07^{*}$	$9.3 \pm 0.74^{*}$	0.95 ± 0.03^{ns}
168	*Sc. dimorphus*	0.20 ± 0.05^{ns}	$0.0226 \pm 0.007^{*}$	$1.00 \pm 0.014^{*}$	$8.5 \pm 0.98^{*}$	0.31 ± 0.75^{ns}
	B. braunii	0.158 ± 0.12^{ns}	$0.0147 \pm 0.02^{*}$	$0.7 \pm 0.04^{*}$	$6.8 \pm 0.4^{*}$	0.41 ± 0.02^{ns}

Note: *-show statistically significant differences ($p < 0.05$), ns = no significant difference.

the depuration of wastewaters. The growth curve of microalgal cells has shown no significant differences between immobilized and free cells except that a latency phase one day longer in immobilized cells than those free. This delay may be due to the thermal stress that microalgal cells have undergone during inclusion in carrageenan. No significant differences have been shown for the absorption of nitrogen and the phosphorous between free and trapped microalgal cells. Moreover, the use of other more expensive microalgal harvest technologies (like centrifugation, flocculation, etc.) could be avoided using these static filters on which are entrapped microalgae.

Confirming these good results, more recently, in order to lowering the ammonium compounds in the aquaculture wastewater, Velichkova et al. (2018) have used two other species of microalgae: *Scenedesmus dimorphus* and *Botryococcus braunii*. The research has shown a better removal efficiency of nitrogen compounds (i.e. nitrite, nitrate and TAN) for both the species in 96–168 h (Table 1).

Feeds and Feeding of Farmed Fish: A Little Focus

A good planning of filtration and depuration system in aquaculture is fundamental, but it is also essential to improve the nutritional formulation of the used feeds. To guarantee a fast and satisfying growth of farmed fish it is essential to ensure a right nutritional regime to farmed species. An efficient diet for fish is based on the quantity and quality of nutrients and on the specific needs, to optimize the growth and the characteristics of the farmed species. The raw materials that are destined to the formulation of the feeds are chosen by their digestibility grade, their eco-friendly and their nutritive power. The most used feeds are composed by vegetal oils,

flours and fish oils. This last raw material derives from fish products of low commercial value but of high nutritional value. For feed formulation other vegetal raw materials like soy flour are also used. These diet regimes, to provide a correct nutritional intake, must contain proteins, lipids, carbohydrates in the right quantities in order to promote a fast growth in the shortest possible time, favouring, therefore, an improvement of the aspect and of the flavour of the farmed fish. A correct diet is also designed to prevent the arise of diseases. Moreover, in order to allow the market to be supplied with non-toxic feeds, also the one destined to fish farms (similarly to that destined to humans) is strictly controlled.

As already described, a part contamination of aquaculture wastewater, comes from the degradation of undigested and uneaten feed. Nevertheless, the aquaculture technology has designed modern systems of feed distribution that make this operation more homogenous, reducing the waste. Diluting the feed with microalgal extracts could improve the problem related to uneaten food. Indeed, these "enriched" feeds would provide more nutrients and energy for the same weight, allowing farmers to use less quantities of them.

Design of "Enriched" Flours for Aquaculture

Utilizing microalgae for the depuration of wastewaters, as already mentioned, has shown many advantages including the possibility of an efficient recycling and exploiting of microalgal biomass once extracted. Microalgae represent a needful food source for all growth stages of larvae of some crustaceans and for some fish species in aquaculture. Furthermore, microalgae are able to synthetize essential amino acids, polyunsaturated fatty acids (PUFA) and more other typologies of high value-added compounds (**Table 2**). In fact, microalgal cells represent the first ring of the aquatic food chain and this implies that many essential high value-added compounds product, are transferred to higher trophic levels through intermediate zooplankton reaching the humans as well. Adding, for example, microalgae to animal feed would bring a real improvement to the final foodstuff quality and healthiness. Actually, the microalgal biomass is rich of various useful compounds like β-carotene, astaxanthin, folic acid, eicosapentaenoic acid (EPA), docosahexaenoic acid (DHA) and polysaccharides like β-glucans. Among the high value-added compounds contained in microalgae, there are several compounds with antioxidant activity (e.g. bioactive peptides, anthocyanins, polyphenols, tocopherols). Recently, microalgae have received attention due to their important group of proteins called phycobilin. Indeed, these proteins have showed important hepatoprotective, anti-inflammatory, and antioxidant activities (Bhat and Madyastha, 2000). In addition to the well-known properties of whole proteins also some peptides have shown a great bioactivity as antioxidant (Byun et al., 2009). In general, these peptides show their bioactivity only after their release from the intact protein during the digestion. Therefore, enriching feed with microalgae or with their extracts, would improve, in the first place, the health and then the nutritional quality of farmed fish making them comparable to wild ones, which are known to show better nutritional characteristics (Lenas et al., 2011). In fact, a more varied diet is typical of wild fish, which, in their daily diet, eat more phyto and zooplankton than the farmed ones. Nevertheless, especially among prokaryotic microalgae (cyanobacteria) it is also observed the production of hazardous compounds called cyanotoxins. These toxic compounds (e.g. microcystin, nodularin, anatoxin-a/-a(s), saxitoxin, ciguatoxin, cylindrospermopsin, etc.) show hepatotoxic, neurotoxic, cytotoxic and dermatotoxic activity. Other minor compounds, like okadaic acid and dinophysistoxins, also produced by some unicellular algae may accumulate in shellfish and thus cause a gastrointestinal syndrome in humans after consuming bivalve molluscs (Katircioğlu et al., 2004). This is why also this type of feedstuff should be strictly controlled before its use in aquaculture.

Indeed, farming fish with feeds enriched of microalgae, would favour a greater intake of these substances in order to increase their well-known healthiness towards to human health. Therefore, it is useful to remember some of the benefits that a diet rich in these substances would bring:

- Lowering the risk of hypertension and consequently of heart attacks and strokes;
- Improvement of retina and brain development;
- Improvement of Crohn's disease symptomatology;
- Reduction of the incidence of breast, colon and prostate tumours;
- Improvement of the painful manifestations of the rheumatoid arthritis;
- Slowing down of inflammations and cellular aging.

Hence, it is clear that a diet based on "enriched" fish can provide numerous benefits in order to prevent many diseases.

Table 2 High value-added substances present in the microalgal biomass

Compounds	*Typology*
Vitamins	A, B1, B6, B12, C, E, biotin, nicotinic acid, riboflavin, pantothenate, folic acid
Pigments	B-carotene, astaxanthin, lutein, zeaxanthin, canthaxanthin, chlorophyll, phycocyanin, fucoxanthin
Polyunsaturated fatty acids	DHA, EPA, ARA
Antioxidants	Polyphenols, catalases, superoxide dismutase, tocopherols
Other	Antifungal, antiviral agents, amino acids, proteins, phytosterols, antimicrobial.

Conclusions

The use of phytoplankton in an integrate system of filtration would allow to make the most of the microalgae potential. Some closed-loop systems of aquaculture already use microalgae-enriched feeds but unfortunately, integrated systems, as the ones above described, are not much used yet. To date, in these breeding typologies, more complex and expensive systems are preferred, both for the wastewater depuration and for the feeds. Nevertheless, because of their eclecticism, microalgae are successful in many fields. In the last years, as previously mentioned, microalgae have been successfully used for the urban wastewater depuration as an active compound in biologic filters. Luckily, depurative and nutritional value of phytoplankton has been also noted in other fields. For example, during space missions, microalgae cultivation would offer a potential recycling of CO_2 produced, so as to generate oxygen and food for the astronauts. Oils and high value-added compounds produced from microalgae, can surely add a basic nutrition to the ability to limit the damage resulting from microgravity and from the cosmic radiation exposure. Microalgae would also offer, on the space stations, a further method of wastewater purification and recycling. Lastly, they would produce useful biomass for nutrient extraction allowing an integrated cycle in the space as well as on the Earth.

References

Bhat, B.V., Madyastha, K.M., 2000. C-Phycocyanin: a potent peroxyl radical scavenger in vivo and in vitro. Biochem. Biophys. Res. Commun. 275, 20–25.

Byun, H.G., Lee, J.K., Park, H.G., Jeon, J.K., Kim, S.K., 2009. Antioxidant peptides isolated from the marine rotifer, Brachionus rotundiformis. Process Biochem. 44, 842–846.

Chevalier, P., de la Noue, J., 1985. Wastewater nutrient removal with microalgae immobilized in carrageenan. Enz. Microb. Technol. 7, 621–624.

FAO, 2016. The State of World Fisheries and Aquaculture 2016. Contributing to Food Security and Nutrition for All. Rome. 200 pp.

González, L.E., Cañizares, R.O., Baena, S., 1997. Efficiency of ammonia and phosphorus removal from a colombian agroindustrial wastewater by the microalgae Chlorella vulgaris and Scenedesmus dimorphus. Bioresour. Technol. 60 (3), 259–262.

Guo, Z., Liu, Y., Guo, H., Yan, S., Mu, J., 2013. Microalgae cultivation using an aquaculture wastewater as growth medium for biomass and biofuel production. J. Environ. Sci. 25, 85–88.

Katircioğlu, H., Akin, B.S., Atici, T., 2004. Microalgal toxin(s): characteristics and importance. Afr. J. Biotechnol. 3 (12), 667–674.

Lenas, D., Chatziantoniou, S., Nathanailides, C., Triantafillou, D., 2011. Comparison of wild and farmed sea bass (*Dicentrarchus labrax* L) lipid quality. Procedia Food Sci. 1, 1139–1145.

Martínez, M., 2000. Nitrogen and phosphorus removal from urban wastewater by the microalga Scenedesmus obliquus. Bioresour. Technol. 73 (3), 263–272.

Mugg, J., Serrano, A., Liberti, A., Rice, M., 2000. Aquaculture Effluents: A Guide for Water Quality Regulators and Aqua Culturists No. NRAC Publication No. 00-003. University of Massachusetts, North Dartmouth, MA; NRAC and Connecticut, Sea Grant.

Sachs, J.D., 2007. The Promise of the Blue Revolution. Scientific American.

Smith, Z., Mccaig, A.E., Stephen, J.R., Embley, T.M., Prosser, J.I., 2001. Species diversity of uncultured and cultured population of soil and marine ammonia oxidizing bacteria. Microb. Ecol. 42, 228–237.

Van Den Hende, S., Beelen, V., Bore, G., Boon, N., Vervaeren, H., 2014. Up-scaling aquaculture wastewater treatment by microalgal bacterial flocs: from lab reactors to an outdoor raceway pond. Bioresour. Technol. 159, 342–354.

Velichkova, K., Ivaylo, S., Yordan, S., 2018. Integrated use of two microalgal species for the treatment of aquaculture effluent and biomass production. Environ. Eng. Manag. J. 17 (7), 1575–1581.

Wang, L., Min, M., Li, Y., Chen, P., Chen, Y., Liu, Y., Ruan, R., 2009. Cultivation of green algae Chlorella sp. in different wastewaters from municipal wastewater treatment plant. Appl. Biochem. Biotechnol. 162 (4), 1174–1186.

Chapter 9

Microalgae to contrast the climate change: a novel food and feed ingredient with technological applications.

 https://doi.org/10.1016/B978-0-12-823960-5.00024-X

Microalgae to Contrast the Climate Change: A Novel Food and Feed Ingredient With Technological Applications

Giovanni D'Auria, Chiara Nitride, and Paquale Ferranti, Department of Agricultural Sciences, University of Naples "Federico II", Naples, Italy

Abstract

The overpopulation and the negative impact of conventional crops and livestock on the environment are forcing the introduction in the food market of sustainable alternative ingredients. Microalgae may be an alternative source of proteins thanks to their easy cultivation, low land-use intensity and resources. In addition, this matrix can be exploited for the treatment of urban, agricultural and industrial wastewaters and in aquaculture and hydroponic, representing an important component in the perspective of a waste-saving circular process. This chapter aims at describing the chemical composition of microalgae and their current applications in food industry as novel functional and technological ingredients. It is also reported the state-of-art of current cultivation technologies of microalgal biomasses. Further efforts are still required to reduce process, energy and water usage during microalgae growth as well as to enhance the texture, taste and appeal of microalgae-based food products, thus increasing the number of consumers willing to buy them.

Key Points

- Microalgae are a sustainable alternative as source of proteins
- Wastewaters can be sustainably treated using microalgae-based systems
- Microalgae can improve technological and nutritional properties of feed and food
- Hydroponic cultures can be positively influenced by microalgae-based systems

Introduction

SECTION 5

General discussion and conclusions

It is still unclear how certain dietary proteins can sensitize and elicit allergies in specific individuals. In this regard, the allergenicity risk assessment of novel dietary protein can be a useful tool in supporting their introduction in food market and in preventing their accidental consumption by allergic subjects in novel food formulation. To date, the allergenicity assessment is carried out predominantly on raw ingredients or thereof purified proteins, neglecting the effect of food matrix and processing (e.g., biochemical changes in proteins, Maillard's reactions, protein-protein aggregation, protein interaction with polyphenols and/or lipids). The applicability of this approach to novel foods is challenged by the lack of comprehensive proteome information. This issue leads to hard identifications of potential *de-novo* and cross-allergens.

Bioinformatics plays a significant role in the current weight-of-evidence approach applied to the risk assessment of allergenicity in novel dietary protein sources. In the bioinformatic allergenicity prediction process, the upstream analysis is usually performed at genomic level, looking for allergens phylogenetically related with the investigated novel food. Conversely, the downstream analysis corresponds to the evaluation of the expressed proteins, which returns a proper picture of the ingredient. Although these analyses are important for designing targeted confirmatory *in-vitro* experiments, both approaches suffer from the lack of genomic and proteomic sequences. This thesis aimed at a preliminary evaluation of the potential allergenicity of sustainable and protein-rich novel food/ingredients. Strengths and weaknesses of the bioinformatic approach in the allergenicity risk assessment of novel ingredients have been highlighted.

Section 5 - General discussion and conclusions

The analysis of *Moringa oleifera* leaf is a good example of how important are curated proteomic and allergens database for the allergenicity assessment of a novel food. Moringa is phylogenetically related to *Carica papaya*, whose fruit allergy has a low prevalence related to the presence of cysteine proteases (Cari p 2) and endopolygalacturonases (PGs, Cari p 1). Papaya fruit can be also related to the latex-fruit syndrome provoked by fruit chitinases. However, the previous phylogenetic evaluation provided limited information, as the allergenicity information about papaya is restricted to fruit, while in this thesis the subject of the investigation was the leaf of *M. oleifera*.

The proteomic analyses showed that the most expressed protein in the Moringa leaf was the RuBisCO, which is one of the few proteins present in *Moringa oleifera* database in UniprotKB. The *in-silico* allergenicity predictions are mainly based on an EFSA threshold, which represents the value of sequence identity percentage with known allergens. In potential allergenic proteins this value is greater than 35% over a sliding window of 80 amino acids. However, these tools rely on protein and allergen databases whose accuracy is pivotal for the outcome's reliability. In fact, using some allergenic prediction tools like Allergome.org, the RuBisCO found in Moringa leaf was flagged as a potential allergen. This assignation was based only on a few and weak literature information. For this reason, to support the bioinformatic evaluations, a standardised ranking of allergen, and curation of data is crucial to avoid possible misinterpretations. Another bottleneck of the bioinformatic approach is that usually the interpretation of the mass spectrometry data about the protein expressed in the investigated novel food is carried out in software only considering the protein database that is available about the investigated matrix, such as the *Moringa oleifera* database on UniprotKB, which only contains 198 protein sequences (11 curated and 187 unreviewed, on January 2022). The previous is a good example of how proteome databases of food

ingredients, and the more novel food, can lack in information. To date, the use of bigger database (that should also comprise the novel food itself), such as *Viridiplantae* on UniProtKB, which comprises 13060199 protein entries (40990 curated and 13019209 unreviewed, in January 2022) is encouraged. In fact, the discovery of proteins like the nsLTP in the Moringa leaf and the subsequent allergenic assessment of this pan-allergen (whose sequences are not available in Moringa protein databases on UniProtKB), would be neglected by using *M. oleifera* database alone. Its identification by homology with the *Rosacea*e family was only possible via proteomic mass spectrometry using the *Viridiplantae* database as background. These findings have been important in directing the following *in-vitro* assays aimed at showing its cross reactivity in subjects with nsLTP allergy. The collection of information about the protein expression in the investigated ingredient is crucial to drive more practical and solid *in-silico* assessment, which in any case require confirmatory *in-vitro* assays. The collected evidence may suggest that Moringa leaf may be potentially allergenic for nsLTP allergic patients, who should be properly warned.

When analysing a more complex genome such as the hybrid hexaploid tritordeum grains, bioinformatic approaches are more limited due to the lack of protein information related to the hybridisation parents, particularly the *Hordeum chilense*. The generation of protein maps for tritordeum cultivars (cvs) is important in assessing their allergenicity. Interestingly, the R5 immunoreactivity of tritordeum flours changes across the two main tritordeum cvs being lower in Bulel cv compared to Aucan cv. This reflects a great difference is terms of abundance of the well described toxic sequence “QQPFP” targeted by the R5 antibody, which is the reference antibody for gluten determination in wheat. This is strong evidence that map the protein expression is essential for assessing allergenicity and celiacogenicity. In literature, the gluten content of Bulel

cv has been showed in line with the Aucan cv; therefore, the R5 immunoreactivity cannot be related to the gluten quantity in all tritordeum cvs. In this respect, further analyses using wheat polyclonal antibody may be more appropriate. However, despite the lower specific R5 immunogenicity of tritordeum cv Bulel, the overall celiacogenic potential of tritordeum cannot be ruled out, as other expressed toxic sequences can be involved.

The availability of a well curated milk protein database makes the assessment of the Kashk straightforward. Being the product of a harsh fermentation, the proteomic analysis is essential in mapping the recurrent sequences, both resistant to processing and gastroduodenal digestion. Based on our evidence, the fermentation process generated short peptide sequences that can be associated with a potential toleration of Kashk by subjects affected by milk allergy.

The *in-vitro* and *in-vivo* confirmation of the above statements is essential prior to the commercialisation of the food product for a rigorous and trustful labelling.

Final remarks

This thesis showed how complex can be the preliminary phases of the allergenicity risk assessment of novel foods. The difficulties are mainly because of the lack and the scarce reliability of the database in terms of curated protein sequences and of allergens, respectively. These issues are even worse for novel food, which genome is not always sequenced, thus invalidating the allergenicity predictions reflecting in probable misinterpretation of data. Further sequencing work is worthy to be performed in order to expand the protein databases, which on their side should increase the reviewing work in order to provide even more curated protein sequences. Allergens databases are

not excluded from these duties. It is important to reach a consensus among allergen databases, which must be built on systematic and well-defined ranking criteria.

The data obtained provide fundamental evidence to direct future *in-vitro* and preliminary *in-vivo* assays in a further evaluation of the allergenicity potential and the consumption safety of the novel foods investigated in this thesis.

ACKNOWLEDGEMENTS

Laboratorio

In primo luogo, vorrei ringraziare il mio relatore, il prof. Pasquale Ferranti, la cui competenza è stata indispensabile nel formulare domande e metodi di ricerca. I suoi riscontri mi hanno spinto ad affinare il mio modo di ragionare e hanno portato il mio lavoro ad un livello superiore.
A lui voglio dedicare queste parole che Dante, attraverso Stazio, consacrò al suo maestro Virgilio:

"[…] Facesti come quei che va di notte,
che porta il lume dietro e sé non giova,
ma dopo sé fa le persone dotte."
(Purgatorio canto XXII, vv. 64-73)

Un ringraziamento speciale va anche a Chiara, che con la sua esperienza è stata un punto di riferimento che mi ha aiutato ad orientarmi e destreggiarmi nel mare della ricerca scientifica.

Grazie anche a Clare che con il suo rigore e il suo amore per la scienza ha ricalcato nella mia mente il giusto modo di approcciarsi ad essa.

Infine, un enorme grazie e un caro ricordo vanno alla prof.ssa Nicolai che con grande amore ed entusiasmo ha dedicato tutta la sua vita alla ricerca scientifica. Grazie per aver sempre avuto (e speso) una buona parola per me, che il suo esempio mi guidi in questo mondo che tanto amava.

Famiglia

Per anni sono stato soprannominato dalla maestra delle elementari "Giovanni dai mille perché". Chi mi conosce sa molto bene che fin da piccolo la curiosità è stata una mia caratteristica fondamentale.
Si può dire che tutto sia iniziato con le prime osservazioni della natura celata sotto i vasi e i sassi del giardino. Quella natura composta dagli insetti più disparati, dal muschio che cresce solo dove è umido, dalle piante che puzzano se sfregate o dalla combustione di un fiammifero, ha fatto sorgere in me un'implacabile sete di conoscenza. Una sete che, con le prime esperienze di laboratorio era diventata vera e propria arsura. A pensarci bene nulla era cambiato: gli insetti sotto il vaso erano diventati una colonia di Lattobacilli in una piastra Petri, il muschio e le piante erano tornate alla loro forma unicellulare, le microalghe, e il fiammifero si era trasformato in un becco Bunsen. La voglia di scoprire e di migliorare mi aveva reso insopportabile il lavoro da tecnico in

azienda, così decisi di darci un taglio, di smettere di guardare ciò che amavo dal buco della serratura e provai a cercare la chiave attraverso il percorso di dottorato. Un viaggio che ha avuto tanto da insegnarmi, ma che oggi volge al capolinea lasciandomi arricchito della tanto agognata chiave: il rigore e il metodo. Perché attraverso il metodo e il rigore la natura si rende meno schiva e comincia a farsi conoscere.

Il primo grazie va ai miei genitori i cui sforzi e sacrifici mi hanno permesso una vita agiata che mi ha consentito di dedicarmi alle mie passioni; è un grazie accompagnato dalle mie più sentite scuse per tutte le volte che ho risposto sgarbatamente prima di un esame o di un esperimento importante. Il vostro supporto è stato il motore potente con cui ho affrontato questa scalata. Grazie anche a mio fratello e a mia sorella, che hanno sempre creduto in me dimostrandomi grande stima e appoggio: siete stati la forte corazza che mi ha fatto sentire sempre al sicuro. Vorrei ringraziare anche mio zio Nunzio: figura che ho sempre stimato per l'estrema intelligenza e curiosità e che ha avuto un forte influenza sulla mia.

Un ringraziamento speciale va a Giulia, la persona senza la quale non avrei mai intrapreso questo percorso, la mia isola felice, il "bicchiere d'acqua nelle notti dei miei guai". Il suo supporto e la stima sono stati fondamentali nei momenti decisivi, e costruttivi in quelli di confronto. Se la mia vita fosse esprimibile in un'equazione lei sarebbe il termine quadratico che rende tutto positivo.

Ringrazio inoltre Roberta che, silenziosa e delicata, tra una cena da Melius e una passeggiata a Sorrento, mi ha sempre dato incoraggiamenti e buoni consigli.
Che figuracce avrei fatto con i miei primi paper senza l'aiuto di Giovanni che, con dovizia e maestria nella scrittura e nella lingua inglese mi ha aiutato a dar voce alle mie ricerche. Il confronto costante con lui è sempre stato fondamentale. Mi ha insegnato l'importanza di rincorrere i propri sogni, anche, a volte, andando contro le aspettative degli altri.

E, in fine, non possono non dire grazie ai miei suoceri e cognatini (tutti) che con attenzioni, sorrisi e gentilezza hanno sempre dimostrato di credere in me, ascoltando i miei soliloqui su esperimenti andati male o sulle vicissitudini lavorative e non solo.