Proteomics in food: quality, safety, microbes and allergens

Cristian Piras\textsuperscript{1}, Paola Roncada\textsuperscript{2*}, Pedro M Rodrigues\textsuperscript{3}, Luigi Bonizzi\textsuperscript{1}, Alessio Soggiu\textsuperscript{1}

\textsuperscript{1}Dipartimento di Scienze Veterinarie e Sanità Pubblica (DIVET), Università degli studi di Milano, Milano, Italy;

\textsuperscript{2}Istituto Sperimentale Italiano L. Spallanzani, Milano, Italy;

\textsuperscript{3}CCMAR, Centre of Marine Sciences, University of Algarve, Campus de Gambelas, 8005-139 Faro, Portugal.

*Correspondence to: Paola Roncada, Istituto Sperimentale Italiano L. Spallanzani, via Celoria 10, 20133 Milano, Italy; email paola.roncada@istitutospallanzani.it or paola.roncada@gmail.com

Abstract

Food safety and quality and their associated risks pose a major concern worldwide regarding not only the relative economical losses but also the potential danger to consumer’s health. Customer’s confidence in the integrity of the food supply could be hampered by inappropriate food safety measures. A lack of measures and reliable assays to evaluate and...
maintain a good control of food characteristics may affect the food industry economy and shatter consumer confidence. It is imperative to create and to establish fast and reliable analytical methods that allow a good and rapid analysis of food products during the whole food chain. Proteomics can represent a powerful tool to address this issue, due to its proven excellent quantitative and qualitative drawbacks in protein analysis. This review illustrates the applications of proteomics in the past few years in food science focusing on food of animal origin with some brief hints on other types. Aim of this review is to highlight the importance of this science as a valuable tool to assess food quality and safety. Emphasis is also posed in in food processing, allergies and possible contaminants like bacteria, fungi and other pathogens.

1. Introduction: The challenge of proteomics in the field of food quality and safety.

Population on Earth is getting close to 7.5 billion with at least 10% of malnourished (Food and Agriculture Organization of the United Nations) [1]. On the other side, metabolic diseases are dramatically increasing in developed countries. Another big concern is represented by the worldwide increasing amount of meat consumption especially in developing countries. Due to urbanization, industrialization and education it is expected an increase of at least 70% of meat consumption [2]. Food security, safety and quality are important for the rising population due to the increasing storage time of every food, from vegetable to meat and fruit. Together with the increased consumption of medium and long-term stored food, is more than mandatory with fast and feasible analytical methods for food analysis.
The aim of this review is to give an insight on how the advances in proteomics could be helpful in food science as a tool to evaluate food quality, traceability and safety, in the view of the improvement of public health.

2. Food quality

This chapter is organized in order to highlight some important characteristics that have to be considered when food quality is addressed, keeping in mind that, especially for some issues as nutritional properties and thermal treatment, the lines of separation with food safety is overlapped. Most important examples of proteomics application to food quality are focused on by the study of meat and dairy product quality. About these foods, proteomics has been used for the characterization of taste, flavor and consistency that represent the pure qualitative traits of food products.

2.1 Taste, flavor, consistency:
There are several applications of proteomics to food quality. One of the most representative example is the evaluation of meat quality because meat represents food of animal origin rich in protein content.

The contribution to meat quality is indeed influenced by several variables such as genetics, environment and post-mortem processing. If genetics can be investigated in order to evaluate the traits related to tenderness and quality, the same is not possible for the assessment of parameters modulated by environmental factors and processing procedures. Proteomics studies for meat quality evaluation have been applied to many types of meat such as beef, pork, lamb and chicken [3].

The synergic use of proteomics tool coupled with bioinformatics analysis is of high relevance. Data analysis through bioinformatics represents the real advance in this field and it is necessary in order to detect the key features of meat proteome that correlate with the quality index [4].

The major genetic traits that have been selected during breeding to improve meat production is related to muscle hypertrophy. Muscle hypertrophy in farm animals is able to rise the conversion index into a higher amount of meat production. However, this genetic variation is linked to some phenotypic changes detectable through proteomics techniques. Myostatin gene deletion is one of the causes of muscular hypertrophy in bulls. The deletion of this gene has been found to be linked to a differential expression of thirteen proteins in muscles of these animals [5]. The major variability in term of differential expression representation of proteins are closely linked to contractile apparatus such as troponin and myosin. Another study applied to ovine muscles that showed a quantitative trait loci (QTL) for muscle hypertrophy, highlighted the overexpression of proteins involved in glycolytic metabolism.
and chaperon proteins [6]. In both cases, in order to obtain these datasets, authors used 2D electrophoresis coupled with mass spectrometry.

As well as genetic factors, animal welfare could influence differences in meat quality. Compensatory growth in pigs that has been associated with meat tenderness [3, 7], it consists in the increased growth rate following a period of food restriction. It has been discovered that, after slaughter, pigs that underwent over a period of compensatory growth had a faster tenderization of meat. This trait was investigated from a proteomics point of view by Lametsch and colleagues [8] who discovered how it is linked to the differential protein expression. The down-regulation of stress and glycolytic proteins was documented using 2D electrophoresis and MALDI TOF MS coupled with a classic sample preparation. Briefly, whole tissue has been homogenized with an Ultra Turrax and the solubilized proteins are separated by 2D electrophoresis. In piglets, also birth weight can influence meat quality. Liu and colleagues demonstrated that low birth weight in response to high fat diet produces changes in the expression of stress proteins in muscles [9]. Meat quality in pig has been investigated using 2-DE coupled to MS. This study identified 27 candidate proteins that coincidently changed with meat quality traits during ageing. The majority of those proteins included cytoskeletal and metabolic proteins and relative degradation products [10]. Di Luca et al. using a gel based approach (2D-DIGE and WB) on pork meat exudate highlighted the lower abundance of stress response proteins in fresh pork meat with higher drip loss and water loss [11].

Pre-slaughter stress, as demonstrated by Franco and colleagues, is also influencing meat quality [12]. In the described work, authors described the differential protein expression in the longissimus thoracis bovine muscle of proteins involved in structural-contracitol and
metabolic functions. Interestingly, authors clustered these proteins with a visible qualitative parameter that is commonly used in beef (DARK FIRM DRY) (Fig.1).

Color of beef is important because it is a parameter that is associated with the consumer’s compliance. It has been described a positive correlation of glycolytic enzymes (phosphoglucomutase-1, glyceraldehyde-3-phosphate dehydrogenase, and pyruvate kinase M2) to redness and color stability [13, 14].

Processing procedures from slaughter to the market bench are as well strictly influencing product quality. The biochemical and enzymatic pathways that occur in this period are responsible for many quality traits such as tenderness that occurs through the degradation of contractile fibers [15]. Swine meat has been analyzed from 0 to 48 hours after slaughter in order to highlight the changes in the proteome pattern. Proteins like actin, myosin heavy chain and troponin T were found to be differentially expressed [16]. It has been demonstrated the correlation between actin fragments and tenderness. Another proteomic experiment important for meat tenderness is the study of the calpain-dependent myofibrillar degradation [15, 17]. The phosphoproteome is as well changing in 24 hours post-mortem, suggesting that glycolysis can play a key role in meat maturation process [18].

A recent metabolomics approach revealed that glutamate, serine and arginine could serve as good predictors of ultimate meat quality parameters [19].

Proteomics also improved the knowledge of milk protein composition. The complexity of milk proteome among species is related to the high evolutionary divergence that also reflects many differences in the proteome such as in post-translational modifications. This topic is particularly important, special if considering the probable formation of numerous proteoforms starting from a single gene product [20, 21]. The complete annotation of human
milk from a qualitative and quantitative point of view contributed to the improvement of better milk formulas as supplies for infants [22-24]. Large-scale analysis of proteome and its post-translational modifications provides the possibility to evaluate proteome changes in relation to environmental conditions or processing procedures. The first quality trait that can be investigated through proteomics is related to the evaluation of milk protein composition among species. This is important for assessment of the best milk protein characteristic for infant nutrition. An important proteomics study described, using 2D electrophoresis, the differential protein profile of human, horse, donkey, goat, sheep, cow and water buffalo. This study provided the basis for the comprehension of allergenic proteins in milk from non-human species [25] (Fig.1).

Proteomics advances have been recently considered highly relevant in the evaluation of milk quality for cheese-making. Seasonal quality of milk for cheese-making have been highlighted by Hinz and colleagues [26] who described, using 2D electrophoresis, how the different proteolysis of different lactation phases could shape the quality of cheddar cheese. The other very important quality of milk for cheese making in order to have a good quality cheese is related to its coagulating power. A milk with reduced coagulating power usually represents a burden for cheese quality. Jensen and colleagues [27] demonstrated that decreased levels of phosphorylated $\alpha_S1$-CN and glycosylated $\kappa$-CN are negatively related to milk coagulation processes. These advances make possible to experimentally evaluate the quality of milk for cheese-making before starting the caseification process [28].

Fish quality has been in the past often evaluated through 2D electrophoresis and successfully applied to product quality and to the determination of procedures applied to food processing [29]. Nowadays, several advances have been addressed in matter of food quality in fish. It has been demonstrated that, through the application of proteomics techniques, it is possible to
gain the overall information about wellness of farmed fish. Among the most important studies, the differential proteome profiling has been analyzed in farmed and wild fish reporting differences in the muscle proteome of sea bass, cod and *Sparus aurata* [30-33].

Flavor peptides have been identified by Zhang and colleagues in puffer fish. The probable peptides responsible for the flavor have been identified using an electronic tongue and MALDI-TOF/TOF MS/MS [34]. Also the ice storage can influence fish quality and change the texture and taste of fish muscle. Within 8 days of ice storage there are several changes that can occur that have been characterized through the proteomics point of view highlighting that there are several biochemical processes involved in protein changes during the post-mortem period [35].

Among the proteomics investigation applied to plants, a lot has been done about wheat. The most important advances in term of stress produced by abiotic response have been described by Komatsu and colleagues in a review article [36]. The manuscript describes the responses, at protein level, linked to the exposure to stressors agents such as heavy metals, salinity, flooding or drought. The proteomics response has been investigated through gel-based and mass spectrometry (MS)-based proteomic techniques and in several parts of wheat plant. As described, obtained results provide the necessary tools for the evaluation of protein profiling that, in turn, is linked to food quality (Fig.1).

### 2.2 Nutritional properties

The major components of a food are carbohydrates, proteins, lipids, sugars, vitamins, minerals and many other minor components. Ever component plays a key role in the nutritional properties of a food, including proteins and peptides. Proteins may have a
function in their entire form, may have enclosed the sequence of bioactive peptides or carry functional groups, cofactors and ions. They could also have a negative impact as allergenicity and toxicity.

The first characterization is related to the form: they can be functional, as whole protein or as a single peptide.

2.2.1 Proteins

Looking to the protein composition of a food, proteomics is a very effective tool for the food industry because it allows the understanding of quality, safety and nutritional requirements (figures 1 and 2). Food reaches the final consumer after a series of steps and, also a simple product, is subjected to a sequence of biological, physical and chemical stress. For example, the identification of key proteins allows to monitor the response of the food matrix to different physiological and non-physiological conditions as they occur during production and storage. This information can prevent damages reduce economic loss. Another important point is the function of the protein not only from a nutritional point of view (e.g. amino acid composition) but also the bioactivity of the protein as a potential functional part in food. Proteins are found in milk and its derivatives, in meat, fish, eggs, legumes, cereals. Their "biological value" and nutritional properties varies according to the amino acid composition. Among twenty amino acids that can constitute a protein, there are nine essentials. A protein rich in essential amino acids can be considered the “most valuable” since our organism is not able to manufacture them and can only access them through food. The proteins of higher biological value are those contained in meat, fish, milk and dairy products and eggs, because
they contain all the essential amino acids. Due to the increasing interest of consumers of a healthier diet to reduce diet-related diseases, the nutritional quality of the food is becoming the most important value of each product. Meat is considered to be a highly nutritious and valued food and in several years’ proteomics contributed to highlight quality traits in different meat types. Proteomics also contributed to define the impact of several steps (aging, mincing, cooking, curing) on protein bioavailability and digestibility and it is also possible to analyze protein complexes bigger than 200 kDa [37]. Cooking and other associated thermal treatments have a deep impact on the physico-chemical modification at the amino acidic level as described by Santanu Deb-Choudhury and co-workers using a redox proteomics approach. Soluble and insoluble fraction from minced meat boiled from 0 to 240 minutes has been analyzed by means of amino acid analysis, gel filtration, LC-MS/MS and fluorescence studies. The analytical procedures showed that essential amino acids were susceptible to heat in their soluble form, with a relative decrease of the levels after 15 minutes of heat exposure. Meat proteins, especially in their soluble form were affected by heat-induced aggregation. In particular soluble collagen was susceptible to oxidative modification while myosin to Maillard reaction with a potential problem on bioavailability of several key amino acids [38, 39]. In fish, the relevance of functional proteins has been studied by Sanmartín and colleagues. Authors successfully investigated the presence of functional proteins and peptides through SDS-PAGE, MALDI-TOF and MALDI-TOF/TOF [40] (Fig.1).

### 2.2.2 Peptides

Enclosed in the protein sequences there are smaller sequences (peptides) that can belong from protein cleavages or from the food microbiota. There are many examples of food that have bioactive peptides and, one of the most representative and studied from the proteomics point
of view is milk. Nutritional relevant peptides are called bioactive peptides. Their length can range from two to 100 amino acid residues and can exploit many biological functions from anti-oxidative, anti-hypertensive, anti-inflammatory or immune-boosting [41]. The screening of the peptidomic profile of food is now possible with high accuracy through the use of the last advances in mass spectrometry and bioinformatics tools. According to Lahrichi and colleagues, the most common range of bioactive peptides is comprised between 2 and 6 amino acids. This range of size is perfect to be studied through proteomics, however, there are some problems to overcome, such as the co-elution of many proteoforms in the same retention time that makes difficult to separately identify and quantify each specie [41] (Fig.1). The method for peptidomics screening in complex matrices has been already described and can be applied to several food matrices [41]. Another argument that is necessary to exploit is the one related to the enzymatic cleavage of food proteome that can produce considerable amounts of bioactive peptides. As stated before, one of the major examples is related to the analysis of milk proteome and the proteome-derived peptidome. Milk, more than bioactive proteins secreted in their final form such as lysozyme, lactoferrin, growth factors and many others [42], has several other proteins that are precursors of bioactive peptides produced by enzymatic lysis. A review from Panchaud and colleagues explains the method to retrieve the pattern of bioactive peptides that can be produced from proteolytic cleavage of existing proteins. The quantification of bio-functional peptides has been successfully applied and described by Holder and colleagues [43] starting from casein and b-casein hydrolysates. The evaluation of the quantity of target peptides has been performed through LC-ESI-MS/MS and a reverse phase high performance liquid chromatography method (RP-HPLC) after tryptic digestion.
Bioactive peptides have been also found in meat products such as ham and meat. It has been demonstrated that a meat-borne lactobacillus can produce a peptide with angiotensin I converting enzyme (ACE) inhibitory activity from porcine skeletal muscle proteins [44]. Anti-hypertensive peptides have been as well identified Spanish dry-cured ham [45]. There are available several databases and tools to evaluate the presence of or the putative formation of bioactive peptides, among them: BioPEP, PepBank, EROP or APD. Some other tools are specific for the detection of antimicrobial peptides, such as the Antimicrobial Peptide Database (APD) (http://aps.unmc.edu/AP/main.php) and BACTIBASE (http://bactibase.pfblab-tun.org/main.php), BAGEL2 (http://bagel2.molgenrug.nl/). A general resume of all knowledge about the peptides that have a biological function can be found in PepBank (http://pepbank.mgh.harvard.edu/), BioPep (http://www.uwm.edu.pl/biochemia/index.php/pl/biopep) and SwePep (http://www.swepep.org/). Among these databases, the last one, SwePep, represents an advance in the field, because it contains a repository of the MS/MS spectra relative to the identification of the peptides.

Top down peptidomics represents the election method for the large scale detection of bioactive peptides. Lahrichi and colleagues proposed a method based on LC-MSMS for the large scale analysis of this peptides including the bio-active ones. Authors used a mixture of peptides and successfully identified at least the 60% of them avoiding the problem of co-elution. The method was successfully applied even when the pool of peptides was identified in a complex matrix [41].
2.3 Product traceability

Food science is extremely important in terms of food quality, adulteration and legislation. The omics profiling of food through proteomics can play a key role in food traceability. This quality is considered of particular interest because of the increasing demand of cost and time-effective procedures able to discover the origin of food. MS-based approaches made possible the detection of differences related not only to quality, species or treatment, but also for geographical variation. There already exist several experimental evidences that document the successful application of proteomics in food traceability. One of the first examples is related to the evaluation of the presence in fish products, of fish with the same genus but of different species [46]. Authors applied a MALDI TOF/TOF approach with the screening of the intact proteome smaller than 10 kDa. Another interesting approach has been used by Wulf and colleagues [47] who successfully applied a Large-Scale Comparison of Tandem Mass Spectra to the characterization of samples of unknown origin. This approach is particularly promising because it is not linked to protein identification, but using a precise spectral library is as well able to detect differences among species or processing methods. Both these approaches could be extremely useful if applied to food traceability. To date, the most promising approach of MS to food traceability is related to the application of GC–MS and GC×GC–MS for the traceability of olive oils according to their volatile compounds profile [48]. The same author also successfully applied DART-TOF-MS technique for the traceability of beer [49]. The ICP–MS combined with specific statistical tools has been revealed extremely useful for the classification of products of protected geographical origin as honey or onion [50, 51].

Another approach has been used by Guo and colleagues in 2013 [52] (Fig.1). Authors analyzed four types of commercial marine species from the East China Sea through inductively coupled plasma mass spectrometry and atomic absorption spectrometry. Principal
component analysis applied was able to distinguish samples from different areas. The Isotope ratio mass spectrometry has also been successfully applied to regional geographical traceability of cattle. Authors, analyzing the ratios of Carbon and nitrogen stable isotopes were able to trace the diet and geographical origin of cattle in China [53]. It was recently demonstrated how the differential nutrition pattern can produce phenotypic differences in the proteome. Zheng and colleagues demonstrated how, a probiotic-improved diet may improve carcass characteristics and meat quality [54]. This provides the basis for future studies on the differences due to regional differences and diet difference to be applied to the product traceability procedures (Fig.1).

3. Food safety

Food safety aim is to prevent any risk of hazard in consumers in consumers. It is a very complex approach that includes, especially in the case of food from animal origin, management of animals (including welfare) up to the processing industries. The major problem in food safety is food poisoning related to foodborne bacteria, but is also true that accidental contamination with abiotic substances, change of proteins conformation or contents during processing, conservation and cooking of food including fraud and allergies are matter of food safety. This section is organized in order to explain how proteomics could help in this sense, with special attention to major (in term of incidence) foodborne pathogens, allergens, protein adulteration including thermal treatment and fraud.

3.1 Bacteria
Bacteria represent a key element in food processing, maturation and ripening. However, there are many cases where bacteria could represent a serious burden for food safety specially in the light of increasing antibiotic resistance. This section gives some information on how proteomics techniques could help in the detection of food-borne pathogens. Among the most common pathogens: *E. coli*, *S. aureus*, *Campylobacter* species, *L. monocytogenes* and other. Proteomics can provide huge support in the comprehension of the mechanisms of infection, antibiotic resistance and biofilm formation of these foodborne pathogens. There is a lot of literature describing these advances discovered through proteomics techniques. Aim of this section is to focus on last proteomics advances in the detection of foodborne pathogens in order to avoid outbreaks or health hazards. *E. coli* represents one of the most common foodborne pathogens responsible for health hazards and outbreaks. A direct method for enterohemorrhagic *Escherichia coli* O157:H7 has already been described by Ochoa and Harrington [55]. This method uses an enrichment step through magnetic beads coated with antibodies specific to antigens of serotype O157. After this enrichment step authors were able to detect $2 \times 10^6$ cells/mL through MALDI MS analysis. Recently, Fagerquist and colleagues successfully applied MALDI-TOF-TOF and tandem mass spectrometry (MS/MS) for the characterization, with a top-down approach, of different Shiga toxin-producing *Escherichia coli* (STEC) subtypes [56]. Remaining in the area of toxin-producers foodborne pathogens, another burden is represented by *Staphylococcus aureus* that can affect animal and human. Milk and dairy products could be influence by the presence of mastitis, especially in dairy cows. Subclinical mastitis represents one of the most relevant burdens for milk hygiene and quality worldwide [57]. The presence of this pathology negatively interferes with milk quality and usability for its transformation processes. High numbers of somatic cells (Somatic Cell Count) in milk represent an index of mastitis [58] that is also associated with a variable proteome and with different amounts of key important compounds such as
fatty acids and lactose [57]. In this field, proteomics provided many advances in the detection of subclinical mastitis in serum and milk of infected cows. Recently, Turk and colleagues [59] and Alonso-Fauste and colleagues [60] discovered both in serum and milk from mastitis cows, several good candidates as putative biomarkers for subclinical mastitis. In both studies, the technique applied was 2D electrophoresis and shotgun MS analysis. Among subclinical biomarkers of infection, several proteins such as serotransferrin, fibrinogen b-chain, and antimicrobial polypeptide (cathelicidin) have been identified [61]. Also the study of MFGM provided important knowledge in this field. Bovine MFGM analysis revealed the differential protein expression related to Staphylococcus aureus infection [62]. It has been estimated that S. aureus is responsible for about 185 000 food-related illnesses in the United States each year [63]. S. aureus toxin is responsible for several gastrointestinal disorders such as nausea, vomiting, and diarrhea within 12 hours from ingestion of few milligrams of toxin. Callahan and colleagues developed a method that uses mass spectrometry to identify a protein toxin (staphylococcal enterotoxin B), in food matrix [64]. The method is very sensitive and is based on the detection of tryptic fragments through tandem mass spectrometry (MS/MS). Authors state that staphylococcal enterotoxin B can be found in concentration up to 5 parts-per-billion in water-soluble food matrices. Another important advance in the absolute quantification of S.aureus toxins used a combination of immunocapture and Protein Standard Absolute Quantification (PSAQ) methods and is believed to play a key role in the counteraction of outbreaks [65].

Listeria monocytogenes is an ubiquitous foodborne pathogen responsible of illness and fatality in humans. Its detection has to go through the culture and takes from four to five days to have a result. Jadhav and colleagues described an easy and sensitive method to detect Listeria monocytogenes directly from selective enrichment broth [66]. This approach
demonstrated that it is possible to detect, through MALDI-TOF MS, up to 1 colony-forming unit (CFU) of *L. monocytogenes* per mL within 30h.

In this field of pathogens detection MALDI-TOF MS represents a powerful tool for the fast identification of foodborne pathogens. This technique is based on the profiling of the whole bacterial proteome providing a fingerprint specific of the analyzed microorganisms in that specific time and physiological condition. The fingerprint obtained through this method is specific of the analyzed microorganisms and has many applications such as the characterization of subspecies, strains and serovar. The whole procedure has been thoroughly described by Pavlovic and colleagues [67]. The examples described highlight how this new advances in proteomics could provide important efforts in the development of fast method for the detection of foodborne pathogens contamination in food.

Several experimental evidences have demonstrated that microorganisms can be identified through MALDI-TOF MS. This technique provides an economic and rapid method for bacterial species identification, from food, water or human specimens such as blood or urine [68]. In 2013, Nguyen and colleagues successfully applied this technique for the identification and classification of lactic acid bacteria in fermented food [69]. This method has also been applied to the detection of bacteria involved in spoilage of milk and pork [70, 71]. This application is very effective in terms of cost and time consumption, however, reproducibility still remains a problem (Fig.1).

Not only proteomics at the single-specie level is important to detect and investigate foodborne pathogens but also the study of the bacterial communities. Especially inside fermented food is of outstanding interest to understand the dynamics of bacterial growth and the mechanisms of interaction of hundreds or thousands of bacteria in the same environment. 

Despite the growing potential of mass spectrometric analysis in complex systems, currently,
the proteomics investigation of microbial communities (metaproteomics) is confined to activated sludge, acid mine drainage biofilms, freshwater and seawater microbial communities, soil, and human gut microbiota [72]. To the date, very few investigations have been done to study complex microbial systems in food using proteomics. Gagnaire et al., in a pre-metaproteomic work, described the release of bacterial protein inside Emmenthal cheese using a combination of protein fractionation, 2DE and mass spectrometry. Due to the interference of proteins belonging to the cheese matrix and technological limits also, authors reported few bacterial species and several metabolic processes linked to the ripening stage [73]. In a second attempt, the same authors, using an ITRAQ based approach, were able to identify and efficiently quantify thirty bacterial and bovine proteins in experimental Swiss-type cheeses [74]. Recently Cardenas et al. investigated the metaproteome of a fermented maize dough using a 1DE separation coupled to an LC-MS/MS of gel triptic extracts on a LTQ Orbitrap Velos. Authors identified roughly a thousand proteins mainly belonging to bacteria, fungi and plants. Several bacterial species like *Lactobacillus* and *Acetobacter* and fungal species like *Aspergillus* seemed dominant in this microbial ecosystem [75]. At the moment the investigation at the microbial community level using in particular metaproteomics require new technological and bioinformatic strategies. Despite the advancement of the NGS technology to sequence all microorganisms and the huge number of sequence databases currently available, complete and well annotated proteomic and genomic sequences remain a significant issue for the scientific community [76]. Moreover NGS opens new horizons for the advanced analysis of food composition, in particular of fermented and processed foods where microbial organisms play a key role. NGS technique is able to detect the taxa composition of the processed food and, coupled with mass spectrometry, is able to identify the processes performed by the microbial community at the time of sample collection. This information provides relevant insights about the transformation of processed
foods. Its use is limited in food science because it still too time and cost consuming. A lot of work still needs to be done in order to enhance process automation to cut costs and time consumption.

### 3.2 Fungi

In the food chain a big problem is represented by the contamination of food, especially cereals, by mycotoxins. In particular environmental conditions, when the temperature and humidity are favorable, molds proliferate and can produce mycotoxins. Mycotoxins are toxic secondary metabolites produced by several types of fungi, mainly belonging to the genera *Aspergillus*, *Penicillium* and *Fusarium*. The main classes of mycotoxins are aflatoxins, ochratoxins, trichothecenes, zearalenone, fumonisins, tremorgenic toxins and ergot alkaloids. Generally they enter the food chain through contaminated crops for the production of food and feed, mainly cereals. The presence of mycotoxins in food and feed can be harmful for human and animal health as it can cause adverse effects of various types, such as cancer and mutagenicity, gastrointestinal and renal failure. Some mycotoxins are also immunosuppressive and reduce resistance to infectious diseases. In this field proteomics [77] and in a greater extent metabolomics [78] has been used to characterize fungi metabolites and secreted proteins in order to build the most complete knowledgebase of secreted proteins and other metabolites [79, 80]. Immunochemical and chromatographic techniques are routinely used for mycotoxin detection. Currently, liquid chromatography coupled to tandem mass spectrometry is the method of choice for multiple detection of mycotoxins and confirmation.
purposes [78, 81]. In the last ten years proteomics has been used either to build a knowledgebase [82] or investigate how different environmental or artificial conditions during fungal growth can modulate fungi stress response [83-85] or proteome changes during mycotoxin producing conditions [86-89]. Recently, several proteomic-based studies have been made to investigate how mycotoxins can modulate immunity in vitro [90, 91], mitochondrial dynamics [92] or induce selective toxicity [93-96].

3.3 Other pathogens

More than bacterial and fungi pathogens, there are also several protozoa and helminths responsible for food hazards. Among them are the *Giardia intestinalis*, *Toxoplasma gondii* and *Taenia spp*. These organisms can be transmitted with water, soil, or contact with persons or animals. Nevertheless, it is increasing the incidence of their transmission with food such as undercooked fish, crabs, and mollusks, undercooked meat, raw vegetables that have been contaminated by human or animal feces. Cryptosporidiosis is a human gastrointestinal infection that could represent a real burden for immunocompromised patients. There is actually no current drug to counteract its growth and, to prevent its infection. In order to unravel the mechanisms behind the sporozoite transmission, Snelling and colleagues performed a study of LC-MS/MS coupled with a stable isotope N-terminal labeling of both inside and outside oocysts sporozoites [97].

3.4 Detection of allergens
The real challenge is the definition of allergens. Very often, proteins or peptides that are not allergenic, become allergens when its use increases. Allergens are a matter of public health, mainly because of the dramatically increasing incidence in the last 20 years. [98]. Since there is no cure for allergy, the only way is to avoid anaphylaxis with a correct dietary education, with responsible information that involves all the food chain, from farm to fork.

Proteomics represents the major tool to study the presence, composition and nature of food allergens (figure 3). Risk assessment now is realized only through the measurements of IgE to define a low, medium or high potential allergen. Clinically is also important to highlight with new methods the difference between sensitization and proved elicitation. Proteomics could provide new approaches and new approach methods for improved assays. There are two types of epitopes that can bind IgE; the linear and the conformational ones. The linear ones are strictly related to the primary structure of the protein, are thermostable and can be recognized by antibodies even after the reduction process. The conformational ones are dependent on the secondary and tertiary protein structure, are thermolabile and their binding with antibodies is not possible after protein reduction [99]. Proteomics can contribute in two ways to determine food allergens; a gel based approach or a gel free approach. The gel-based approach includes in the workflow 2D electrophoresis, 2D immunoblotting and MS approach for protein spots identification. The gel-free approach is characterized by an HPLC-MS/MS approach and the IgE binding assay of the trypsinized proteome. Bioinformatics analysis for the detection of specific immunoreactive epitopes is necessary at the end of experiments. Some examples of studies that have been recently performed using 2D electrophoresis for the detection of allergens in several foods are related to beer [100], beef [101], milk [102, 103], rice [104], and fish [105]. IgE immune reaction is not always directed to protein epitopes, but can also be directed to glycans. α-Gal represents one of the major reasons of allergy to
meat. This theory was successfully demonstrated by Apostolovic and colleagues who identified, through 2D electrophoresis, novel $\alpha$-Gal-containing proteins responsible for allergy reaction [101].

The fast detection of fish allergen parvalbumin has been described by Carrera and colleagues who proposed a Selected MS/MS Ion Monitoring (SMIM) in a linear ion trap (LIT) mass spectrometer successfully feasible in less than 2 hours [106]. Resuming these described advances, it has to be underlined that the proteomic analysis of food allergens has become a key issue in the food safety field in recent years [107]. This represents a key point, especially in the light of the dramatically increasing prevalence of food allergies coupled with the increased public awareness of anaphylaxis due to food allergy [98]. The limit of this technique is that it is restricted to the detection of linear epitopes. Indeed, both gel-free and gel based approach are linked to the previous reduction of the proteome prior to immunoassay with patient sera. This approach is selective for the detection of linear allergenic epitopes. As previously described, proteomics advances, in particular 2D electrophoresis coupled with MS, represent the best technique available for the detection and the discovery of new potential allergens. However, important advances have also been made in the field of absolute quantification of discovered antigens. From the analytical point of view it is really important to be able to detect allergens in complex food matrices or even in traces. Even a small allergen contamination, not indicated in the food label could represent a serious burden for allergic patients. For this reasons, sensitivity and specificity represent a major target for the analysis of the presence of an allergen. Because of the advantage of specificity of IgE against its specific allergen, the most diffuse methods for their detection are represented by ELISA or immunoblotting. However, even if their sensitivity can reach the amount of 5 ppm, there is the possibility of false positives because of the possible interaction
of the antibodies with the food matrix. Regarding allergens quantification the MRM targeted quantification is very promising. Houston and colleagues [108] successfully evaluated the concentration of 10 allergens in commercial soybean varieties using a label free proteomics approach. The absolute quantification was evaluated through MRM measuring amounts of antigens in the range between 0.5 to 0.7 µg/mg. The advance of this study is represent by the use of BSA as internal standard that allowed the reduction of technical variance up to the 7% of the measure. Koeberl and colleagues developed a method for the quantification always using MRM [109]. Authors reported an interesting comparison of pro and cons of both immunological and mass spectrometry methods and described the importance of a good selection of each antigen of signature peptides and related transitions.

3.5 Food processing procedures (thermal treatments)

Large-scale food production and processing includes the application of mechanical, chemical and physical treatments to preserve foods by slowing down or stopping the natural processes of decay and augmenting the conservation time. Several processing procedures can be applied singularly or in combination depending on the food type. Freezing, heating, drying, fermentation, salting and the use of chemicals are the most common and classical procedures. Other processing treatments are less common (e.g. microwaves, ultra-high pressure and Pulsed Electric Fields) or very specific for several foods (irradiation). Depending on the type of processing treatments these can lead to improvement or depauperation of the nutritional value of food. Over the last ten years proteomics have been successfully applied to the study of protein and protein modification in food before and after the transformation to obtain valuable information about the molecular changes at the protein level linked to each type of treatment (figure 4). One of the most used and studied food transformation is the thermal treatment. With this type of treatment, we can obtain a food microbiologically safe and
prolong the shelf life. Obviously, the nutritional value of each food is influenced by heat treatments. Milk and dairy products are subjected to different types of thermal treatments from pasteurization (72 °C for 15 s) to sterilization by ultra-high-temperature treatment (UHT; 135–150 °C for 2–6 s). These procedures lead to the Maillard reaction that is the non-enzymatic glycation of amino groups (mainly lysine residues in milk proteins) by reducing sugars (lactose is the main reducing sugar in milk) [110]. The products of this complex reaction can be different depending on the duration of the heating. In milk, lactulosyllysine (bound to several milk protein) is the main product in the early stage of the thermal treatment and many other reaction products are formed during the advanced stage (longer thermal treatment) of Maillard reaction. In the late nineties, before the advent of proteomics, antibody-based methods have been used to detect lactosylated caseins [111, 112] and lactosylated proteins in pasteurized and UHT milk [113]. However, these approaches were not applicable to the characterization of the lactose-binding site. For this reason, several strategies based on mass spectrometric methodologies have been developed for the structural analysis of milk proteins. In 1997, using the recently developed LC-ESI-MS technique Leonil and colleagues demonstrated that beta-lactoglobulin (β-Lg) in milk whey protein concentrate (WPC) was specifically modified by a covalent binding of a lactose residue on Lys47 under mild heat treatments due to the early Maillard reaction [114]. Using roughly the same approach Fogliano and colleagues analyzed purified β-Lg isolated from several milk samples subjected to three different thermal treatments: pasteurized (72–85°C for 15–30 s), ultra-high temperature (UHT 142–145°C for 2–5 s), and sterilized (115–120°C for 10–30 min). Lys-100 was identified by a combined mass spectrometric and structural analysis as a preferential lactosylation site of β-Lg during industrial thermal treatments [115]. Other attempts have been made to characterize the thermal induced modification in the milk whey fraction and to develop new analytical strategies for the rapid monitoring of structural
modification during the food processing. In this way, Siciliano et al., using ESI-MS and MALDI-MS, demonstrated that as β-Lg also alfa-lactalbumin (α-La) undergoes to lactosylation, preferentially in Lys98, during thermal treatment. The degree of lactosylation for both proteins was proportional to the thermal treatment used (sterilization>UHT>pasteurization) [116] (Fig.1). Moreover, authors reported that, during thermal treatment, the heavy denaturation of β-Lg caused the formation of aggregates with caseins. This lead to the depletion of whey protein from milk and a further reduction of nutritional value of thermal treated foods due to the limited bioavailability of proteins and amino acids. A relative quantification of β-Lg modification by MALDI-TOF MS has been reported by Meltretter et al. and as expected the relative level of modified residues was proportional to the intensity of the treatment [117]. Non-enzymatic glycosylation and oxidative modifications have been investigated at caseins level by means of immunochemical and mass spectrometric techniques originally by Scaloni and colleagues. Authors identified by ESI-MS and MALDI-MS several lactosylation sites in αs1- and β-CN correlated to the severity of the treatment applied, moreover a parallel carbonylation of caseins has been observed using anti- 2,4-dinitrophenylhydrazine antibodies [118]. The structural localization of protein-bound carbonyls was investigated in various thermal treated milk and milk powder by Fenaille and al. applying a combined immunochemical detection. Authors used anti 2,4-dinitrophenylhydrazine antibodies of modified milk protein and the identification of tryptic peptides was performed by MALDI-TOF MS and nanoESI-MS/MS [119]. Recently, Arena et al., to improve the systematic identification of lactosylation in whey proteins, applied an enrichment step by Proteominer followed by affinity chromatography and nLC-ESI-(LIT)-MS/MS analysis with CID and ETD fragmentation that allowed the identification of 271 non-redundant modification sites in 33 milk proteins [120] and 310 lactosylation sites in 56
proteins from milk fat globule (MFG) [121]. Non-enzymatic post-translational modification (nePTM) intermediate and advanced glycation end-products (AGEs) derived from Maillard reaction in milk and dairy products have been characterized by combined gel and mass spectrometric approaches [122, 123]. A modified peptide analyzed by MRM has been proposed for the detection of thermal treatment in milk and dairy products [124]. Not only thermal treatment is responsible of non-enzymatic post-translational modification (nePTM) in milk proteome, but also the temperature of storage of milk, as demonstrated by Holland using a classical 2-DE MALDI-TOF approach [125]. Recently, the proteomic investigation has been integrated with experiments in animal models to give a functional significance, at the physiological level of modifications generated by each type of thermal treatment. In this field, Lonnerdal et al. investigated the biological effects of site specific modification on the digestibility of milk proteins both in vivo and in vitro providing semi-quantitative data on modified peptide abundance after digestion for each thermal treatment[126]. As previously reported, thermal treatment like cooking are also very common during meat processing and these treatments in most cases lead to oxidative modification and Maillard reaction of specific meat proteins with a parallel decrease of the bioavailability of several aminoacids[38]. A 2-DE/MS proteomic approach coupled to multivariate statistics was also applied to the investigation of storage time and freezing temperature in fish meat. Authors reported that frozen storage time have major influence on protein profile compared to different freezing temperatures. In particular the abundance of fragments of several glycolytic and cytoskeletal proteins was directly correlated to the storage time [127] (Fig.1).

3.6 Product adulteration
Food fraud is a worldwide problem today. All the valuable ingredients in each food are susceptible of adulteration. Generally, food is adulterated when a valuable constituent has been omitted or substituted in whole or part with other low quality constituent, but the definition can be broader and more complex. The use of fast analysis at high sensitivity and specificity are such a critical for the verification of the quality and safety of the food and to ensure the health of the consumer. A common adulteration in the food industry is characterized by the cross-species contamination in processed foods (Fig.1). Recently von Bargen and colleagues [128] applied a targeted proteomics approach for the detection of specific peptides from horse and pork meat in beef products. Authors were able to detect down to 0.24% horse or pork in a beef meat matrix using a MRM/MRM^3 strategy coupled to an optimized fast extraction strategy. To quantitatively detect chicken meat into a mixed meat food with high reproducibility and sensitivity, Santandreu and coworkers implemented a method based on an off-gel fractionation step coupled to the AQUA labelling and a MS detection of myosin 3 selected peptides on a conventional LC-ion trap MS/MS. With this experimental procedure they were able to detect as low as 0.5% w/v contaminating chicken in pork meat with high confidence [129]. 2D-gel based strategy was applied by Montowska and co-workers, analyzed the differences in the amount of myosin light chain (MLC) in different meat products made from cattle, pig, chicken, turkey, duck and goose [130]. With this approach it was possible to detect as low as 10% of different meat analyzing at least 3 isoforms of MLC. On the same samples the authors searched also for other protein biomarkers suitable for the use in the authentication of meat products. Several Blood plasma proteins, metabolic enzymes and regulatory proteins were found as potential target to build specific test [131]. The adulteration of meat does not only involve the fraudulent use of mixtures of meat of different species. Frequently, soybean proteins are added to meat as emulsifiers to improve their functional properties, moreover the low cost of those proteins
promotes their use in doses exceeding the permitted. In 2006, Leitner using a 2D-LC-MS/MS approach confidently identified five high-abundance major variants of glycinin and of all three chains of alpha-conglycinin as marker of soybean proteins in processed meat [132]. Fish-based foods are affected by similar problems including the fraudulent substitution of high quality fish with the low quality ones to obtain a higher gain. Mazzeo and co-workers in 2008 developed and successfully applied a MALDI-TOF based method for the fish authentication. Analyzing protein samples from 25 different fish muscle tissues it was able to establish, in few minutes, a strict discrimination among the analyzed species based on characteristic features of parvalbumins in the MALDI linear spectrum after PMF [133]. Contrarily to several IEF and SDS-PAGE based approaches [134, 135], MALDI-TOF is faster and easily discriminates very close species based if coupled with bioinformatics analysis. Other investigations applied both 2-DE and MALDI-TOF to analyze the parvalbumin isoforms in closely related species of the family Meluccidae [136] or different types of tuna fish. Recently Wulff presented an interesting approach for the authentication of fish products [137]. Using a reference spectral library made from 22 different fish species it was possible to correctly classify, without any genome or protein sequence database, more than 90% of the unknown spectra deriving from unknown and also heavy processed samples. As in meat and fish, authenticity of dairy products is a very important point in the food market worldwide. Nowadays especially milk and typical cheeses with PDO label are adulterated by the use of low cost dairy-products (powdered milk, mixtures of milk from different species, low-quality milk, etc). The evaluation of the quality of dairy product is mainly based on the traditional procedures (genetic[138], chromatographic [138, 139], electrophoretic [138, 139], and immunoenzymatic [140] methods). In addition several proteomic-based techniques have been implemented to assess the authenticity of dairy products and for a fast and accurate detection of the fraud. To highlight differences in the
protein profile of milk from 5 different species (Bovine, Caprine, Buffalo, Equine and Camel) Hinz and co-workers applied a classical 2-DE / MS based approach[141]. Other similar attempts have been made using gel based [142] or gel free approaches. An ITraq based approach was applied by Yang [143] to obtain a quantitative differential and functional expression pattern of 211 proteins from the milk whey fraction of Cow, Yak, Buffalo, Goat and Camel. As suggested by the authors, the results constitute a knowledgebase for the evaluation of the adulteration of expensive milks with bovine milk or low quality milk. The MALDI-TOF MS approach proved to be a rapid, simply and accurate analytical method for the evaluation of cow milk presence in sheep or water buffalo milk or for the detection of powdered milk in fresh milk [144], down to a 1% of adulteration level [145]. Also adulteration in donkey milk was detected by monitoring the protein profiles of the most abundant whey proteins as α-lactalbumin (α-LA) and β-lactoglobulin [146] down to a 0.5% level [147]. The application of multivariate techniques such as linear PLS (Partial least squares) regression and non-linear Kernel PLS coupled to the MALDI-TOF whole spectra information has been proposed and successfully applied for the analysis of binary and tertiary mixtures of milk. It has been as well successfully applied for the predictions of the levels of milk species adulteration achieving high accuracy levels with typical errors between 2–10% for cow’s milk [148]. Other authors reported the successful use of CE-MS technique to monitor milk adulteration in a concentration range between 5 and 95% [149]. Similar results may be achieved with an HPLC/ESI-MS approach using β-lactoglobulin whey protein as the molecular marker [150] or the MRM technique coupled to caseinomacropeptide (CMP) as a biomarker to fluid milk adulteration through whey addition [151]. Recently, a complementary peptidomic and proteomic approach based on MALDI and ClinProt technology for biomarker recognition was able to recognize adulteration up to 5% associated with thermal treatment [152]. In the dairy products market a common adulteration is the addition of sheep milk to
goat cheeses, similarly it is frequent that sheep cheeses contain cow milk. Several proteomics methods have been developed to evaluate the authenticity of cheese. Guarino and co-workers, using a gel free method based on the SRM analysis of typical sheep's peptide produced by plasmin hydrolysis of caseins, were able to detect up to 2% of sheep's milk in cheese [153]. Typical PDO soft cheeses, like Italian buffalo mozzarella, are frequently adulterated with low quality milk and powder milk. Few years ago, MALDI-TOF in linear and reflectron mode has been used to detect mixtures of cow and ewe milk in water buffalo mozzarella using species-specific mass features of α-lactalbumin and β-lactoglobulins, as molecular markers [154]. Unfortunately, despite the speed and simplicity of analysis with MALDI-TOF, this technique is not suitable for a quantitative analysis. To achieve quantitative results, a MRM-based UPLC/QqQ-MS/MS approach has been applied to the adulteration of buffalo mozzarella. This technique, looking to the phosphorylated β-casein f33-48 tryptic peptide as a novel species-specific proteotypic marker, is able to detect up to 0.001% of bovine milk in buffalo milk with a linearity over four orders of magnitude [155]. As demonstrated by Claydon and colleagues, the peptidomics approach is useful for the detection of meat species also in highly processed foods [156]. Authors performed the horse meat identification using heat-stable peptides as markers. Samples were analyzed through nLC-MS/MS and data analysis was performed through the use of a species-specific peptides database (Fig.1).

3.7 Chemicals and other contaminants

Proteomics is not only useful for the detection of biological hazards, but, as described below, it can also provide reliable indirect index of contamination with xenobiotics. One example is represented by the documented differential protein expression of oyster in relation to HgCl₂ contamination. Zhang and colleagues described the differential protein expression of 13
proteins and, 4 of them showed interesting features as possible biomarkers to be applied for the detection of Hg contamination in food [157]. Illicit corticosteroid treatment represents as well a burden for food safety. Guglielmetti and colleagues reported bovine paraoxonase/arylesterase 1 precursor (PON1) as a specific and reliable biomarker of corticosteroids treatment [158]. In both of this cases, the approach used is 2D electrophoresis coupled with mass spectrometry and, the proposed protein, represents an indirect biomarker to detect chemical hazardous contamination.

4. Conclusions

There is no doubt that food safety and quality is of global importance, especially because it affects health, economy and trade. Food safety is essential for food security and food quality, and alerts are daily issued. One of the key words is prevention and it is mandatory to support industries to produce safe and quality food. Proteomics represents a real challenge in this field, because it is able to produce rapid methods to investigate the modification or the presence or absence of targeted proteins in complex food including raw materials and matrices. Proteomics can give a valuable add-on in building food safety intervention.

References


Figure legends
Figure 1. General approach for proteomics application in food science. Image references:

**Figure 2.** Schematic proteomics workflow for food analysis. From left to right: Proteins, naturally occurring peptides and DNA from food matrix are extracted. DNA can be used for PCR and/or NGS, proteins and peptides can be separated by different techniques. Separated proteins and related enzymatic or non enzymatic post-translational modification (ne)PTM can be detected and identified using several approaches: antibody-based techniques, bottom-up and/or top-down proteomics. Finally experimental data can be integrated and analyzed using bioinformatic software.
**Figure 3.** Image resuming the last advances for allergens quantification using a proteomics approach [55, 56].

![Diagram of allergens quantification process]

- Protein extraction
- Tryptic digestion
- Mass spectrometry
- Detection of signature peptides
- Mass spectrometry of food for allergens signature peptides quantification [55, 56]
**Figure 4.** thermal treatment and detection of modified proteins. From left to right. Foods are subjected to thermal treatment. Thermal treated Proteins and related peptides undergoing several non enzymatic post-translational modification (ne)PTM depending on the reaction condition (Maillard reaction, oxidation). Modified proteins and related ne(PTM) can be detected, identified and quantitated using several approaches: antibody-based techniques, 2-DE/MS, MS based techniques and MRM.