

available at www.sciencedirect.comwww.elsevier.com/locate/jprot

Searching for allergens in maize kernels via proteomic tools

Elisa Fasoli^a, Elide A. Pastorello^b, Laura Farioli^c, Joseph Scibilia^b, Giancarlo Aldini^d, Marina Carini^d, Adriano Marocco^e, Egisto Boschetti^f, Pier Giorgio Righetti^{a,*}

^aDepartment of Chemistry, Materials and Chemical Engineering “Giulio Natta”, Politecnico di Milano, Via Mancinelli 7, 20131 Milano, Italy

^bUnit of Allergology and Immunology, Niguarda Ca’ Granda Hospital, Milano, Italy

^cUnit of Laboratory Medicine, Niguarda Ca’ Granda Hospital, Milano, Italy

^dDipartimento di Scienze Farmaceutiche “Pietro Pratesi”, Faculty of Pharmacy, University of Milano, via Mangiagalli 25, 20133 Milano, Italy

^eIstituto di Agronomia Generale e Coltivazioni Erbacee, Università Cattolica del Sacro Cuore, 29100 Piacenza, Italy

^fBio-Rad Laboratories, C/o CEA-Saclay 91181 Gif-sur-Yvette, France

ARTICLE DATA

Keywords:

Zea-mays

Hexapeptide ligands

Allergens

Vicilin

Globulin-2

Thioredoxin

Lipid transfer protein

50 kDa γ -zein

Trypsin inhibitor

ABSTRACT

Up to the present, only one major allergen had been univocally identified by chemical analysis (N-terminal sequencing) in the salt-extractable (cytoplasmic) protein fraction of maize kernels (*Zea mays*): the lipid transfer protein (Pastorello et al., *Allergy Clin. Immunol.* 2000;106:744–751). In the present study, two-dimensional maps of kernel flour have been set up, the proteins transferred to nitrocellulose membranes and confronted with sera of various patients allergic to maize proteins. Via spot excision and Orbitrap mass analysis, the following new allergens have been identified: vicilin, globulin-2, 50 kDa gamma-zein, endo-chitinase, thioredoxin and trypsin inhibitor. Vicilin was found to be composed of a string of six spots, all of them allergenic; also globulin-2 was composed of a string of five spots, exhibiting equivalent allergenicity. The 50 kDa gamma-zein, found here in the maize flour soluble fraction, is identical to the 50 kDa allergen reported by Pasini et al. (*Allergy* 2002; 57:98–106), but present in the insoluble fraction and solubilized via -S-S- reducing agents. However, the form here described might be a truncated species, since it exhibits an apparent Mr, in SDS-PAGE, of ca. 35 kDa. The homology of three of them (vicilin, globulin-2 and thioredoxin) with other vegetable systems has been investigated via BLAST analysis, the ones with highest homology belonging to rice, wheat and barley. The present data add a non-negligible amount of previously unreported allergens to the scanty panorama of maize proteins.

© 2009 Elsevier B.V. All rights reserved.

1. Introduction

Maize (*Zea mays* L.) is a plant belonging to the family of the grasses (*Poaceae*) and is cultivated globally, being one of the most important cereal crops world-wide. Maize is not only an important human nutrient, but also a basic element of animal feed and a common ingredient for industrial products. Maize

is found all over the world, mainly due to its easy processing, ready digestion, and cheaper cost than other cereals. It is also a versatile crop, allowing it to be grown across a range of agro-ecological zones. Today, the major maize production areas are located in temperate regions of the globe [1]. The United States, China, Brazil, and Mexico account for 70% of the global production. In Europe, maize production is concentrated in the

* Corresponding author. Tel.: +39 02 23993045; fax: +39 02 239 930 80.

E-mail address: piergiorgio.righetti@polimi.it (P.G. Righetti).

Mediterranean area (France, Italy and Spain) and in Eastern Europe (Romania and Hungary). According to FAO data, 589 million tons of maize was produced worldwide in 2000.

The maize seed is called the kernel, which contains starch, protein, oil and other nutritionally valuable substances such as carbohydrate, protein, iron, vitamin B, and minerals. Maize oil contains polyunsaturated fatty acid and vitamin E. Maize derived ingredients are found in a broad variety of processed foods. About 15% of maize harvested is transformed into derivatives, like alcohol, oil, starch, maize syrup and dextrose. Maize oil is commonly used as salad or frying oil and in margarine [2].

Up to recent times, very little was known about the occurrence of maize allergy. There were no studies that investigated how often maize allergy would occur. In fact, on the basis of a retrospective study on patients with histories of food allergies, Moneret-Vautrin et al. [3] concluded that food allergy to maize is rare. The first comprehensive study on maize allergens was performed by Pastorello's group [4] in the year 2000. By screening sera from 22 patients with systemic symptoms after maize ingestion, it was found that 19 of them (86%) recognized a 9 kDa protein, with $pI > 9$, that was identified, via N-terminal sequencing, as lipid transfer protein (LTP), that was thus confirmed to be the major maize allergen. LTP had first been identified as an important allergen in fruits, but it is also present in nuts, various vegetables and in cereals. In a second study by this group [5], it was found that LTP is an extremely stable protein, and maintains IgE-binding activity even after cooking at 100 °C. LTP appears to be resistant also to gastrointestinal digestion [6]. These properties make LTP a strong food allergen that can cause severe reactions. Another characteristic of allergenic LTPs is the geographical distribution of sensitization to these proteins inasmuch as they have been described as important allergens only in Southern Europe. Interestingly, however, maize LTP has been found to be a relevant allergen also in a group of patients from the US [7], suggesting that sensitization to LTP could be a worldwide phenomenon. Other proteins have been described as important allergens in maize. For example in a third article of the series [8], by exploiting a phage display, enabling enrichment for IgE-binding proteins via three rounds of affinity selections, a second important allergen was recognized in maize, namely thioredoxin. This list could hardly be exhaustive; in fact, in 2002, Pasini et al. [9] reported another major allergen in maize, namely the 50 kDa protein, a protein not found in the salt-extractable fraction (albumins and globulins), as in Refs. [4–6], but in the insoluble fraction after reduction with 20 mM dithiothreitol. Interestingly, in a 2D proteomic study of maize pollen, Petersen et al. [10] found two major classes of allergens, belonging to the family of profilins and expansins (so-called because they are involved in the stretching of cell walls and in the growth of pollen tubes), which, as previously known, represent the group 1 of cross-reactive allergens of grass pollens. A detailed characterization of one group of pollen allergens (β -expansins) has also been given by Li et al. [11].

On the assumption that these findings could be considerably expanded, we have undertaken a proteomic study of maize salt-extractable proteins, aimed at a positive identification of allergens via spot excision and mass spectrometry analysis of all spots found to bind IgEs from known allergic

patients. This study differs considerably from previous ones on a few important aspects. To start with, instead of using maize flour bought in supermarkets (thus of unknown genetic background), as performed in previous works, we are here using the historically important public US inbred line B73. This is because B73 plays an important role in maize improvement in many different breeding programs all over the world. The inbred B73 line has been adopted for sequencing the maize genome; recently the sequencing of 15,200 BAC clones was completed. Assembled together, these clones represent a high-quality draft sequence of the maize genome, comprised of approximately 2 billion non-redundant bases [12]. Secondly, the allergens were uniquely identified via two-dimensional (2D) maps [13], rather than via SDS-PAGE as customarily done in most studies of allergens. Thirdly, we have adopted pre-treatment of the salt-extractable proteins, for eliminating lipids and polysaccharides (starch), so as to have clear and non-smeared 2D maps and further treatment with combinatorial ligand libraries, which have given brilliant results so far in augmenting the visibility of low-abundance proteins [14,15]. This has resulted in discovery of a number of previously unreported allergens. Such allergens are only those found to be in common among all patients investigated.

2. Materials and methods

2.1. Materials

The solid-phase combinatorial peptide Library-1 (ProteoMiner™), its carboxylated version (Library-2) as well as materials for electrophoresis such as gel plaques and reagents were from Bio-Rad Laboratories (Hercules, CA, USA). *N*-ethylmaleimide, urea, thiourea, 3-[3-cholamidopropyl dimethylammonio]-1-propanesulfonate (CHAPS), isopropanol, acetonitrile, trifluoroacetic acid and sodium dodecyl sulphate were all from Sigma-Aldrich (St Louis, Mo, USA). Complete protease inhibitor cocktail tablets was from Roche Diagnostics, (Basel, CH). Sequencing grade trypsin was from Promega (Madison, WI, USA). All other chemical were also from Aldrich and were of analytical grade.

2.2. Plant material

The B73 inbred line of maize from Iowa Stiff Stalk Synthetic (BSSS) was grown in the nursery at the Istituto di Agronomia, Catholic University, Piacenza, Italy. Plots consisted of double rows that were 5 m long, with 0.75 m between rows and contained 15 plants per row. Recommended crop management techniques were used [33]. Plants were hand self pollinated, hand harvested and mature ears were dried at 15.5% moisture and stored at 4 °C. For proteome analysis, samples of 100 g were taken from an individual ear, dried at 40 °C and ground with a laboratory mill Cyclotec (Foss, Padova, Italy) to pass a 1.0 mm sieve.

2.3. Maize treatment

Fifty g of *Zea mays* B73 flour were treated with a mixture of CHCl_3 : CH_3OH (70%: 30%, v/v) for delipidation. By centrifugation, the solid phase was separated from organic solvent enriched in lipids and it was dried with a mechanic vacuum pump. Maize proteins were extracted three times (120 mL each time) in physiological saline solution (PBS, 20 mM phosphate buffer, pH 7.2, containing 150 mM NaCl) under constant agitation with a magnetic stirrer at 40 °C. The total liquid recovery was 300 mL, with a total protein content of 750 mg. Tablets of complete protease inhibitor cocktail were added to the extraction buffer in order to prevent protease action.

2.4. ProteoMiner treatment

We used a combinatorial library of hexameric peptides as capturing agent for all extracted proteins. We adopted two types of hexapeptide library beads: amino-terminal hexapeptide beads (Library-1) and beads in which the peptides have been derivatized by carboxylation (Library-2). The initial, pre-treated and clarified protein solution was loaded on a column containing 1 mL of Library-1 and shaken for 2.5 h at room temperature. After centrifugation at 4000 rpm for 10 min, the clean supernatant was transferred to a second column containing 1 mL of Library-2 and shaken for 2.5 h at room temperature. After centrifuging at 4000 rpm for 10 min, the two columns were washed with PBS, and then each column was subjected to two distinct elutions using a TUC solution (2 M Thiourea, 7 M Urea, 2% CHAPS) and a UCA solution (9 M Urea, 2% CHAPS, 50 mM citric acid to pH 3.3). Finally the eluates were neutralized and desalted by dialysis at 4 °C against 100 mM ammonium carbonate (cut off of dialysis membrane was 2000 Da).

2.5. One- and two-dimensional electrophoresis

One (SDS-PAGE) and two-(2D-PAGE) dimensional gel analyses were performed exactly as described in D'Ambrosio et al. [16], to which the readers are referred for further details.

2.6. Immuno-detection

After electrophoresis (either in SDS-PAGE or 2D-PAGE), the maize components were electroblotted to a nitrocellulose (NC) membrane (Amersham 0.45 μm) by using a trans-blot cell from Bio-Rad at 4 °C, 0.45 A and 100 V for 4 h. The NC membrane was blocked by incubation with PBS and 0.1% Tween-20. In the case of SDS-PAGE, the membrane was cut into individual strips that were incubated overnight at room temperature with serum of individual patients. Negative sera were used as well as controls. In the case of 2D maps, the membrane was cut to a size that would precisely match the dimensions of the 2D gel slab and marked with a pencil along the contour. Binding by specific antibodies was detected by incubation with ^{125}I -labeled anti-human IgE diluted 1:4 with blocking solution for 6 h and exposure on X-ray film at –70 °C for 6 days. For spot cutting, the Coomassie-stained duplicate gels were overlaid onto the X-ray film on a light-box so as to precisely match the IgE reacted spots with the blue spots in the control gel.

2.7. Protein digestion for MS analysis

The 2D-gel spots were cut in an aseptic condition and destained with a solution composed by 50% acetonitrile (AcN) and 50% 5 mM Tris, pH 8.8. Proteins were then digested overnight at 37 °C with 0.2 μg of modified sequencing grade trypsin (Promega, Madison, WI, USA) in 40 mM ammonium bicarbonate. The resulting peptides were extracted from the gel by three steps: two incubations in 50% formic acid and 50% acetonitrile and a last incubation in acetonitrile, each ones followed by sonication for 10 min. The three collected extractions were dried in speed-vac and resuspended with 10 μL of 0.1% trifluoroacetic acid (TFA). Six μL of each sample were loaded on C18 Zip-tips and the peptides were eluted by 50% acetonitrile solution. The organic peptide solution was dried in speed-vac.

2.8. NanoLC-MS/MS analysis

All digested peptide mixtures were separated by online reversed-phase (RP) nanoscale capillary liquid chromatography (nanoLC) and analyzed by electrospray tandem mass spectrometry (ESI-MS/MS). Chromatography was performed using a Surveyor LC system (ThermoFinnigan, Milan, Italy) on a 180 $\mu\text{m} \times 10\text{ cm}$ column packed with 5 μm , Biobasic-18 stationary phase (Thermo, Superchrom, Milan, Italy). The pump flow rate was split 1:75 for a column flow rate of 2 $\mu\text{L}/\text{min}$. The column effluent was directly electrosprayed using the silica emitter source without further splitting. Mobile phases A and B were 0.1% formic acid in water and in acetonitrile, respectively. The separation of peptides obtained by enzymatic digestion was achieved with a gradient of 0–60% B over 60 min. Before the next analysis, both the precolumn and the column were first washed with 90% solvent B for 10 min and then equilibrated with 100% solvent A for 20 min. For the identification of peptides, an LTQ XL-Orbitrap mass spectrometer was used (Thermo Scientific, Milan, Italy) and the electrospray interface (dynamic nanospray probe, Thermo Scientific, Milan, Italy) was set as follows: spray voltage 1.5 Kv; capillary temperature 180 °C, capillary voltage 30 V; tube lens offset 110 V, no sheath or auxiliary gas flow. During analysis, the mass spectrometer continuously performed scan cycles in which first a high-resolution (resolution $r = 30,000$ at m/z 400) full scan (300–2000 m/z) in profile mode was made by the Orbitrap, after which MS2 spectra were recorded in centroid mode for the 3 most intense ions (isolation width, 3 m/z ; normalized collision energy, 30 CID arbitrary units). Protonated phthalates [dibutylphthalate (plasticizer), m/z 279.159086; bis(2-ethylhexyl)phthalate, m/z 391.284286] and polydimethylcyclodioxane ions $[(\text{Si}(\text{CH}_3)_2\text{O})_6 + \text{H}]^+$; m/z 445.120025] were used for real time internal mass calibration [17]. Dynamic exclusion was enabled (repeat count, 3; repeat duration, 10 s; exclusion list size, 25; exclusion duration, 120 s; relative exclusion mass width, 5 ppm). Charge state screening and monoisotopic precursor selection was enabled, singly and unassigned charged ions were not fragmented.

2.9. Protein identification

The acquired MS/MS spectra were automatically searched against protein databases using the Bioworks software (rev.

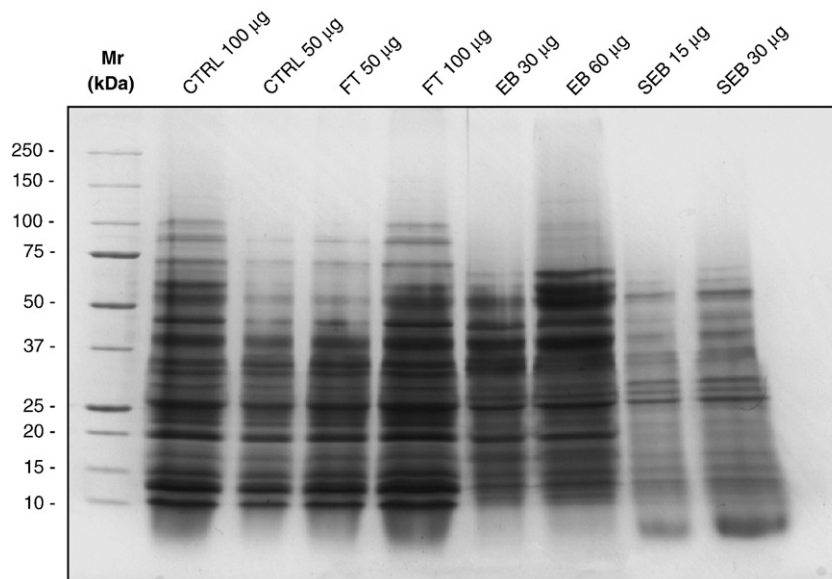


Fig. 1 – SDS-PAGE in 8–18% T polyacrylamide gel slab of (from left to right): molecular mass ladder; control (CTRL), salt-extractable maize proteins in 100 and 50 µg loads; flow through (FT) from the combinatorial ligand libraries, in 50 and 100 µg loads; TUC eluates from ProteoMiner (EB) in 30 and 60 µg loads and TUC eluates from carboxylated library (SEB) in 15 and 30 µg loads. Staining with colloidal Coomassie blue.

3.3.1 sp1, Thermo Scientific). The database of *Zea mays* proteins was built using the FASTA database utilities, the uniprot_trembl.fasta.gz (from ftp.expasy.org) as original database and *Zea mays* as string to include (the target database contained 4051 entries). The following settings were used for the search engine: mass type, mono-isotopic precursor, and fragments; enzyme, trypsin (KR); missed cleavage sites, 2; peptide tolerance, 20 ppm; fragment ion tolerance, 1 AMU; number of results scored, 250; ions and ion series calculated, B and Y; peptide matches reported, 10. Carbamidomethylation of cysteine was set as a fixed modification. Peptides were identified on the basis of the probability and the cross-correlation (Xcorr) values, which assure the reliability of the amino acid sequences. Peptides with a probability lower than 1.0×10^{-3} were selected. Xcorr values greater than 1.5 (+1), 2.0 (+2), and 2.5 (+3) were considered for identifying the peptides. The Xcorr function measures the similarity between the mass-to-charge ratios (m/z) for the fragment ions predicted from amino acid sequences obtained from the database, and the fragment ions observed in the MS/MS spectrum.

3. Results

Fig. 1 shows the SDS-PAGE patterns of control, salt-extractable proteins as well as of the eluates from the two ligand libraries, ProteoMiner (EB) and carboxylated ProteoMiner (SEB). This approach has given brilliant results in a number of applications, such as the analysis of human sera [18] and urines [19], of rDNA proteins [20], of cytoplasm of red blood cells [21] and chicken egg white [16]. Although, as expected, there appears to be the reduction of some high-abundance proteins and a concomitant enhancement of low-abundance proteins, espe-

cially in the low-Mr region (see the two SEB tracings), by reconstituting the two eluates and performing both SDS-PAGE and 2D mapping (not shown), no large differences appeared in the number of spots but only in the topology of a relatively limited number of spots. Due to these reduced differences the

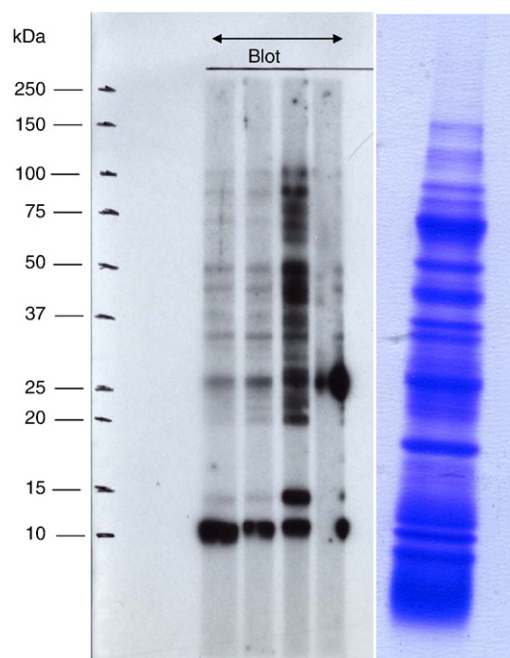


Fig. 2 – Immunoblots from single SDS-PAGE strips with four different maize-allergic patients (first four lanes to the left). Far right: Coomassie-stained control strip of soluble cytoplasmic proteins. First lane to the left: molecular mass ladder.

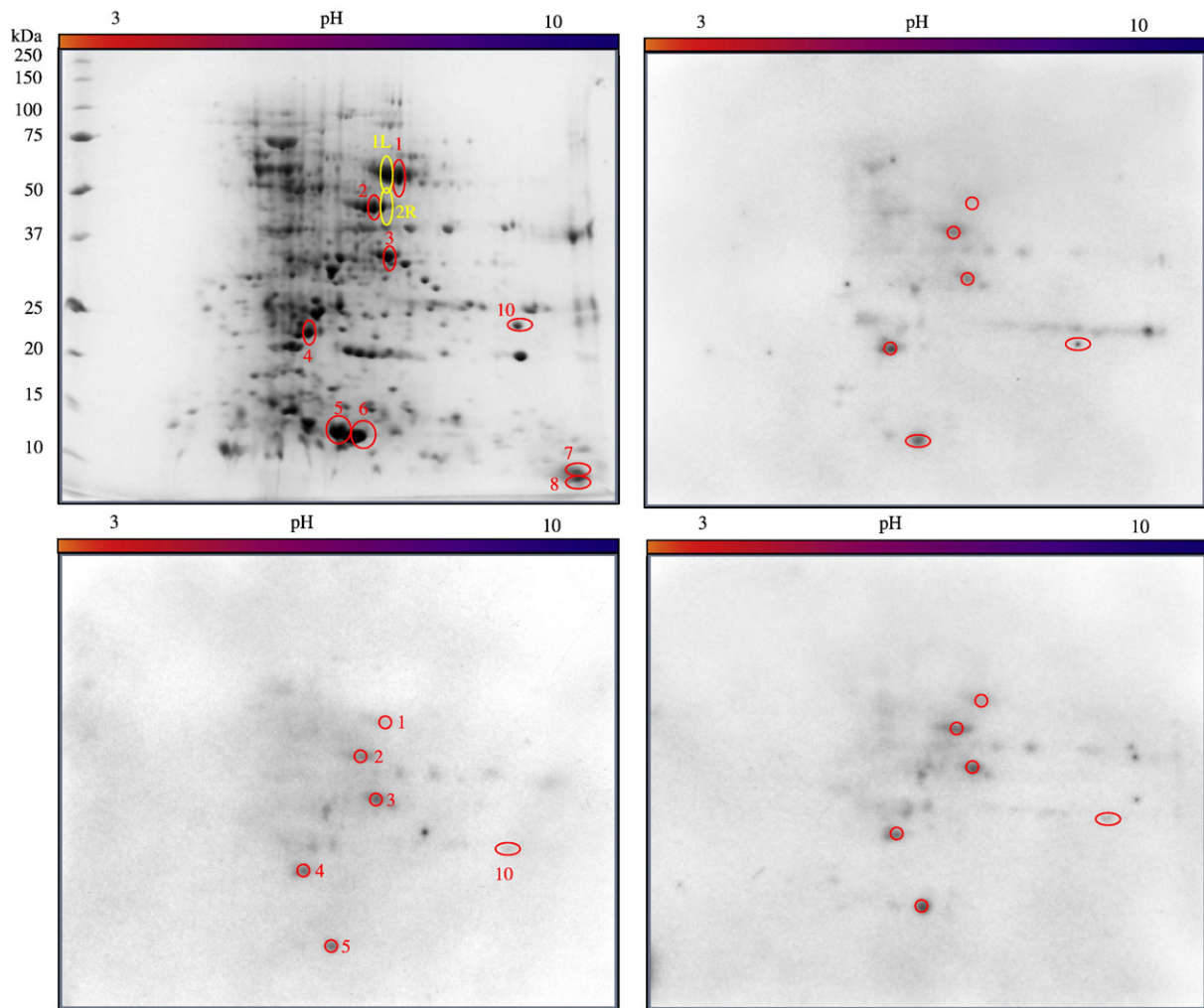


Fig. 3 – Two dimensional maps of salt-extractable proteins of maize flour. First dimension: IPG pH 3–10 strips, 7 cm long; second dimension: SDS-gel slab in the 7.5–18% T porosity gradient. Upper left panel: control gel stained with Coomassie blue. The other three panels represent autoradiography of blotted proteins revealed with radio-iodinated antibodies (IgE) against sera of 3 different maize-allergic patients. The circles and numbers 1–8 refer to spots excised and subjected to MS analyses with an Orbitrap equipment. 1L indicates an isoform to the left of the major spot (identified as vicilin); 2R indicates an isoform to the right of the major spot No. 2 (identified as globulin 2). In both cases these isoforms most likely represent PTMs of the same gene products.

study was restricted to the non-treated corn extract. These underoptimized results from ProteoMiner treatment might be due to the presence of extract components that partially annihilate the effect of hexapeptide library treatment as recently reported [22]. As described in [22], the primary amine terminal group in fact could be a strong interacting moiety for macromolecules that are largely present in plant extract such as polyphenols, polysaccharides and others, thus dramatically reducing the effect of protein capture according to the inverse of their abundance. To prevent these phenomena, the recommended specific treatments of plant extract were not applied in the present study on maize seeds extracts.

Fig. 2 shows a typical SDS-PAGE profile of cytoplasmic, soluble proteins of maize, as well as four immunoblots confronted with sera of four different patients selected at the Niguarda Hospital on the basis of a positive DBPCFC for maize and positive maize specific IgE (Phadia-ImmunoCAP –

Italy). The strong band at around 10 kDa is the well-known lipid transfer protein (LTP), previously identified by Pastorello et al. [4]. However, a number of additional reacting zones appear, distributed in the Mr 10 to ca. 100 kDa range. Attempts at cutting some of these bands and subjecting them to MS analysis failed, due to the multitude of proteins detected in each single zone, as well demonstrated in [21], where as many as 300 unique gene products could be identified by Orbitrap analysis in a single gel slice from SDS-PAGE. Thus, this approach, although much in vogue in typical studies on allergens, was abandoned too, in favour of 2D map analysis, in which the numerous proteins would be scattered in the pI/Mr plane, thus increasing the chance of finding individual species in each excised spot.

Fig. 3 (upper left panel) shows a typical 2D map of the soluble, cytoplasmic proteins of maize, analyzed in an IPG pH 3–10 strip in the first dimension, followed by a 7.5–18% T

Table 1 – List of proteins identified in spots nos. 1–5 and 6 in Fig. 3.

Spot n.	Accession number	Protein name	Mr (Da)	pI	Number peptides ^a	Coverage (%) ^b	Score-P-PRO ^c
1	Q03865	Vicilin-like embryo storage protein	66162	6.23	24	39.6	1E-30
1L	Q03865	Vicilin-like embryo storage protein	66162	6.23	21	32.7	1E-30
2	Q7M1Z8	Globulin-2 precursor	49892	6.16	34	69.5	1E-30
2R	Q7M1Z8	Globulin-2 precursor	49892	6.16	27	46.7	2E-11
3	Q946W1	50 kD gamma zein	34545	5.30	3	18.3	2.12E-11
4	Q5XML0	Proteasome subunit beta type	23072		5	24.8	1.5E-13
5	Q09HU3	Trypsin inhibitor	13570	6.65	8	61.4	2.02E-12
5	Q4W1F6	Thioredoxin h2 protein	13030	6.22	5	27.9	2.10E-11
9	Q2XX13	Lipid-transfer protein	11270	8.48	1	17.2	1.2E-7
10	Q6JBK3	Endo-chitinase	28983	8,16	9	24,5	1.3E-13

^a Total number of peptides assigned to the protein; ^b Percentage of protein sequence covered by the matching peptides; ^c Displays the probability of finding a match as good or better than the observed match by chance. The value displayed for the protein is the probability of the best peptide match (the peptide with the lowest score).

1L indicates an isoform to the left of the major spot (identified as vicilin); 2R indicates an isoform to the right of the major spot No. 2 (identified as globulin 2).

porosity gradient in the second dimension. About 300 spots, stained with Coomassie blue, could be counted, most of them crowded in the pH 4–7 region and covering the 10–150 kDa Mr range. The other three maps are immunoblots confronted with sera of three different patients (lanes 1, 2 and 4 in Fig. 2, the sample of lane 3 having been left out, due to hyper-reactivity to most proteins in the maize extract; for the sake of brevity, only 3 maps are here shown, but seven total patients were screened, as in general requested in these studies). Several immuno-reactive zones can be seen; however, we have concentrated our efforts on a set of six common spots (labelled 1–5 and 10) which appear to be common to all

patients analyzed. By overlaying the Coomassie-stained gel to the X-ray film, it was possible to excise accurately the corresponding protein zones and send them to MS analyses. The MS identifications (done in triplicate on three different patients) are listed in Table 1. Spot No. 1 has been identified as vicilin-like embryo storage protein; spot No. 2 is globulin-2; spot No. 3 is 50 kDa gamma-zein; spot No. 4 is proteasome subunit beta; spot No. 5 was found to be an envelope of two distinct proteins, trypsin inhibitor and thioredoxin; finally, spot No. 10 was identified as endo-chitinase. Other spots, marked 6–8, and gel regions devoid of proteins were also analyzed as negative controls. In both 1D SDS-PAGE (Fig. 1)

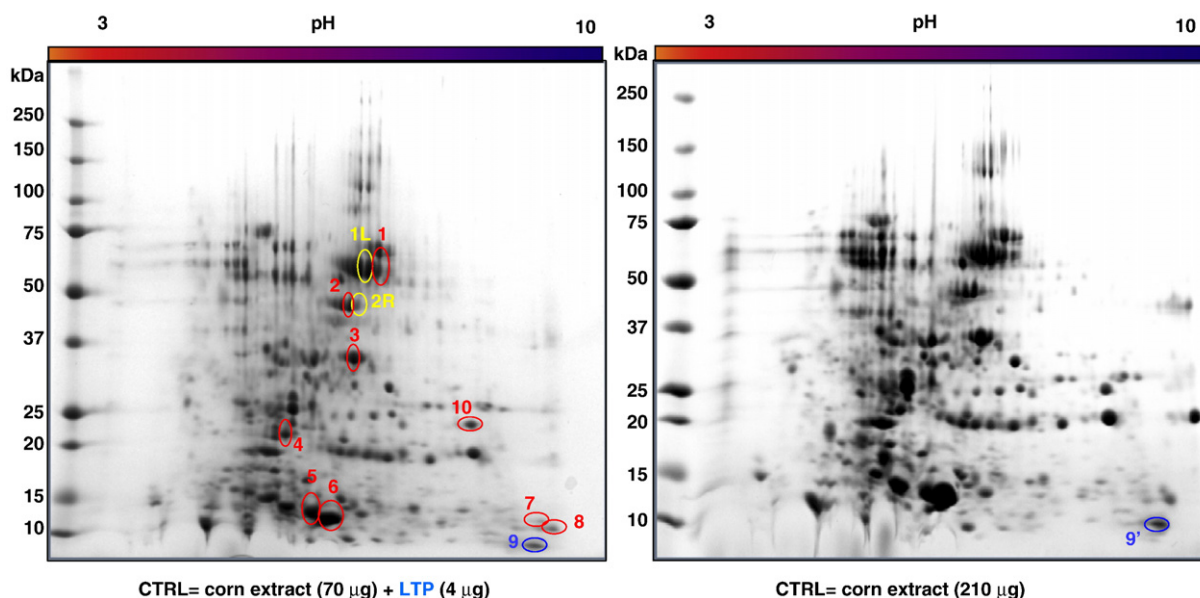


Fig. 4 – Two dimensional maps of salt-extractable proteins of maize flour. First dimension: IPG pH 3–10 strips, 7 cm long; second dimension: SDS-gel slab in the 7.5–20% T porosity gradient (left panel) or 7.5–22% T porosity gradient (right panel). The gel to the left side was spiked with 4 µg of LTP purified to near-homogeneity; the gel to the right was loaded with 3 times higher protein levels (210 µg) as compared to the left one (70 µg). All gels stained with colloidal Coomassie blue. Spot No. 9 in both gels has been identified as LTP by nanoLC-MS/MS analysis.

Table 2–Sequence and m/z values of all peptides identified in the proteins listed in Table 1.

Sequence	MH+	z
Spot n. 1 –		
Accession number Q03865 (Vicilin-like embryo storage protein)		
NPEFSLSSFSKSIQR	1726.87	2
TSSDRLER	963.48	2
HGQDKGIIVR	1122.64	3
HASEGGHGPWPLPPFGESR	2152.01	3
GPYSLLDQR	1048.54	2
GPYSLLDQRPSIANQHGLYEADAR	2799.38	4
PSIANQHGLYEADAR	1769.86	3
IAYVPNGK	861.48	2
GYAEIVCPHR	1144.55	2
RSEEEEEESSEEQEEAGQGYHTIR	2838.19	4
SEEEEEESSEEQEEAGQGYHTIR	2682.09	3
LSPGTAFVVPAGHPFVAVASR	2080.13	3
DSNLQIVCFEVHADRR	1745.83	3
DSNLQIVCFEVHADRNK	2117.00	3
VFLAGADNVLQK	1274.71	2
VFLAGADNVLQKLR	1658.92	2
VAKALSFASK	1021.60	2
ALSFASK	723.40	2
ALSFASKAEVDEVLSRR	2064.07	3
AEVDEVLSRR	1203.58	2
AEVDEVLSRR	1359.69	2
GERERHGR	996.51	2
HGREEREK	1040.52	2
EEREKEER	1233.57	2
Spot n. 2 –		
Accession number Q7M1Z8 (Globulin-2 precursor)		
PYHYGEESFR	1284.56	2
FTHELLEDAVGNYR	1663.81	2
VAELEAAPR	955.52	2
AFLQPSHYDADEVMFVK	1996.95	3
EGEGVIVLLR	1084.64	2
RESFCVR	896.44	2
ESFCVR	740.34	2
EGDVMVIPAGAVVYSANTHQSEWFR	2763.32	3
VVMLLSPPVSTSGR	1444.82	3
FEFFPIGGESPESFLSVFSDDIVQASFNTR	3498.64	4
REEWEK	876.42	2
EEWEK	720.32	2
VFEKQSK	865.48	2
QSKGEITTASEEQIR	1676.84	2
GEITTASEEQIR	1333.66	2
SCSRGGRSSR	1052.50	2
WEIKPSSLTDK	1303.69	2
PSSLTDK	747.39	2
HYEITGDECPHLR	1569.71	3
LLDMDVGLANIR	1400.76	3
IAIVLK	656.47	2
GHGREEEER	1227.53	2
GHGREEEEREEQGGGGGQK	2283.98	4
EEEEEREEQGGGGGQK	1876.78	4
IREGSVIVIPAGHPTALVAGEDK	2329.29	4
IREGSVIVIPAGHPTALVAGEDKNLAVLCFEVNASFDDK	4095.12	4
EGSVIVIPAGHPTALVAGEDK	2060.10	4
EGSVIVIPAGHPTALVAGEDKNLAVLCFEVNASFDDK	3825.94	4
NLAVLCFEVNASFDDK	1784.85	3
NLAVLCFEVNASFDDKVFLAGTNSALQK	3014.53	4
VFLAGTNSALQK	1248.69	3
LLAFGADEEQVDR	1590.77	3
VIGAQK	615.38	2
VIGAQKDAVFLR	1316.77	3
DAVFLR	720.40	2

Table 2 (continued)

Sequence	MH+	z
Spot n. 3 –		
Accession number Q946W1 (50 kD gamma zein)		
QQQPQQPQQYQQGQEK	1970.93	3
SQQQQCHCQEQQTTR	1960.83	3
CSYNYSSSNLK	1515.64	3
Spot n. 4 –		
Accession number Q5XML0 (Proteasome subunit beta type)		
MECVLGVVGR	1062.54	2
NLHLYQFR	1090.58	2
NTIPLSTAAANFTR	1547.82	3
CIKEIRLR	1030.62	2
LVVAPQNFVIK	1227.74	2
Spot n. 5 –		
Accession number Q09HU3 (Trypsin inhibitor)		
SAGTSCVPGWAIPHNPLPSR	2150.03	4
WYVTSR	811.41	2
TCGIGPR	703.35	2
LPWPELK	882.51	2
LPWPELKR	1038.61	2
ELADIPAYCR	1150.55	2
CTALSILMDGAIPPGPDAQLEGR	2325.16	4
LEDLPGCPR	999.49	3
Accession number Q4W1F6 (Thioredoxin h2 protein)		
AEFDAHMTK	1049.47	2
AIAPLFVEHAK	1195.68	2
KFTQVVFLK	1109.67	2
FTQVVFLK	981.58	2
VDVDEVK	803.41	2
Spot n. 9 –		
Accession number Q2XX13 (lipid-transfer protein)		
NAAAGVSGLNAGNAASIPSK	1769.91	2
Spot n. 10 –		
Accession number Q6JBK3 (chitinase)		
SNAYCDASNR	1177.25	2
QWPCAAGQK	1065.21	2
VAQDAVIAFK	1062.24	2
NQAGSGCEGK	1027.12	2
VDPGPNLTC	992.15	2
QYCQQLR	1015.19	2
VMPQGFQATIR	1177.45	2
MANAPR	659.78	2

and in 2D experiments, sera of non-allergic, healthy individuals, were also analyzed as negative controls and they did not exhibit any reactive zone (not shown). Additionally, since the vicilin spot (No. 1) appeared as a string of six spots in the Coomassie stained gel (and gave immuno-reactivity in the three immuno-blots) an additionally spot (labelled 1L) to the left of the originally excised spot, was further analyzed and identified as vicilin. The same applies to spot 2 (globulin-2) which appears as a string of five spots. In this case an additional zone (marked 2R), to the right of the originally excised spot, was also analyzed and identified as globulin 2 (see also Table 1). It is reasonable to assume that, in both cases, these strings of spots represent PTMs (post-translational modifications) of the same gene product, although their precise nature has not been ascertained here.

Table 3 – BLAST sequence homology search of Vicilin-like embryo storage protein, Globulin-2 precursor and Thioredoxin h2 protein from *Zea mays*.

Protein	Accession number	Identities (%)	Positivities (%)
Vicilin-like embryo storage protein (<i>Zea Mays</i>)			
Putative globulin <i>Oryza sativa</i> subsp. <i>japonica</i> (Rice)	Q75GX9	60	73
Putative uncharacterized protein <i>Oryza sativa</i> subsp. <i>indica</i> (Rice)	A2XKC1	54	66
Storage protein [<i>Triticum aestivum</i> (Wheat)]	Q7DMU0	52	66
Embryo globulin [Bgl1] [<i>Hordeum vulgare</i> (Barley)]	Q03678	52	66
Globulin-like protein (Fragment) [<i>Oryza sativa</i> (Rice)]	Q9ZRG9	49	61
Globulin-2 precursor			
Putative uncharacterized protein [OsI_013387] [<i>Oryza sativa</i> subsp. <i>Indica</i> (Rice)]	A2XMU4	60	71
Putative globulin [<i>Oryza sativa</i> subsp. <i>japonica</i> (Rice)]	Q852L2	60	71
Globulin-like protein [<i>Oryza sativa</i> subsp. <i>japonica</i> (Rice)]	Q8L8I0	59	70
7S globulin [<i>Sesamum indicum</i> (Oriental sesame) (Gingelly)]	Q9AUD0	38	57
48-kDa glycoprotein precursor [<i>Corylus avellana</i> (European hazel)]	Q8S4P9	38	53
Thioredoxin h2 protein Q4W1F6			
Thioredoxin H-type [<i>Oryza sativa</i> subsp. <i>japonica</i> (Rice)]	Q0D840	67	81
Thioredoxin H [<i>Triticum aestivum</i> (Wheat)]	Q8GVD3	64	80
Thioredoxin h isoform 1 [<i>Hordeum vulgare</i> var. <i>distichum</i> (Two-rowed barley)]	Q7XZK3	62	81
Thioredoxin h1 [<i>Glycine max</i> (Soybean)]	A8HTK0	57	74
Thioredoxin [<i>Ipomoea batatas</i> (Sweet potato) (Batate)]	Q6VBI4	56	72

The best five scoring sequences are reported in the table for each protein. The comparison of the protein query sequence against UniProt Knowledgebase (Swiss-Prot + TrEMBL) protein database was performed at the SIB using the BLAST network service. Identity: The extent to which two amino acid sequences are invariant. Positivity: The number and fraction of residues for which the alignment scores have positive values.

Although these data add a considerable amount of new proteins to the allergen list of maize, curiously 2D maps did not allow us to detect the major allergen, namely LTP [4], that was well visible in the SDS-PAGE profile of Fig. 1. It occurred to us that, perhaps, due to its low Mr value, LTP had been lost with the dye front in the 7.5–18%T second dimension gel. The experiments were therefore repeated by using a 7.5–20% T second dimension gel slab, and spiking the maize extract with purified (to near-homogeneity) LTP, provided by Pastorello's lab. Fig. 4 (left panel) shows the 2D profiling, this time with a new spot visible in the low-Mr, high pI region (marked with No. 9). Upon spot excision and immuno-blotting, this spot was univocally identified as LTP (see Table 1). As an additional proof, the cytoplasmic maize extract was further analyzed by loading to the first dimension IPG strip three-times as much protein (210 µg) and exploiting a 7.5–22% T porosity gradient in the SDS-gel slab. As shown in the right panel of Fig. 4, an intense spot (marked 9') was now visible; by MS analysis and immunoblot it was univocally identified as LTP. As a summary of all data here reported, Table 2 lists the sequences of all peptides identified in all proteins reported in Table 1, together with their m/z values.

4. Discussion

Our data add a non-negligible number of allergens to the meager list of maize allergens reported so far, namely: vicilin, globulin-2, 50 kDa gamma-zein, endo-chitinase, thioredoxin and trypsin inhibitor, all of them known allergens in a number of other vegetable systems. It must be emphasized, in fact, that the single allergen so far identified by direct chemical

analysis (N-terminal sequencing) has been only LTP [4,5]. In addition, although the 50 kDa gamma-zein here reported is identical to the 50 kDa allergen reported by Pasini et al. [9] as a major allergen in maize, there is a fundamental difference between the two sets of data. In the case of Pasini et al. [9], this protein was identified not in the cytoplasmic fraction, but in the insoluble fraction after solubilization via reduction of -S-S-bridges with dithiothreitol. In our case, instead, this same protein is found in the salt-soluble fraction, in the absence of reducing agents. The discrepancy, though, is only apparent: as visible in our 2D maps (Figs. 3 and 4) the 50 kDa γ-zein is located at an apparent Mr value of 35 kDa, rather than at 50 kDa. We hypothesize that this species has been released via a proteolytic attack, cleaving off a residue of ca. 15 kDa and rendering it soluble in the cytoplasm. This process most likely occurred *in vivo*, not *in vitro*, since, prior to flour treatment, we add a cocktail of five protease inhibitors. We have in purpose left out, from the total count, spot No. 4, identified as the proteasome subunit beta type, since this protein, to our knowledge, has not been listed before as an allergen in any other food stuff. We are now trying other systems (notably peach, peanuts and hazelnuts) in order to prove its existence in different species as well. The novelty of the present data, thus, is not so much in the identification of these additional six allergens that are already known in other food stuff, but in having found them, for the first time, in maize flour, a system that has been thought to be a non-allergenic food par excellence [3]. Having thus set out the panorama of standard allergens in maize, this study lays out the foundation for screening GMO maize in search of potential new allergens, a matter of major concern in present-day GMOs (genetically modified organisms).

It might be asked if the allergens here reported bear a resemblance to other allergens found in a variety of different species. In order to prove that, we have run a BLAST sequence homology study [23] in the three most important new allergens here found, namely vicilin, globulin-2 and thioredoxin. The best five scoring sequence for each of these proteins is reported in Table 3. It can be appreciated that, in the case of vicilin, the three best scoring sequences are those found in rice, wheat and barley, with homologies between 50 and 60%, as one might expect, given the fact that they belong to related families of cereals. Vicilin may thus be considered as a newly discovered food allergen belonging to this important family of foods. In the case of globulin-2, the best scoring sequences are just those of rice (60%), the *Sesamum indicum* and European hazel scoring with much lower figures (only 38%). For thioredoxin, the data resemble those of vicilin, with rice, wheat and barley scoring from 62 to 67%; however also sweet potato and soybean score very well (60% and 57%, respectively).

As an additional remark, it should be noted that, in one case, two allergens are found in the same spot (No. 5 in Figs. 3 and 4), containing a mixture of thioredoxin and trypsin inhibitor. This might most probably due to the fact that these two proteins have very similar Mr values (13030 and 13570, respectively) and thus might co-elute in the same spot, considering that in the low Mr region SDS-PAGE remarkably loses resolution [24]. It might be argued that their pI values are distinct enough so as to prevent fusion in the same zone (pI 6.22 for thioredoxin and 6.65 for trypsin inhibitor), but these are only theoretical pIs and cannot possibly take into account post-translational modifications, such as deamidation or other charge-modifying variations along the polypeptide chain. Additionally, however, it must be emphasized that, in 2D maps of cell lysates, that could potentially contain thousands of protein species, singlets (i.e., spots contains just a single protein) are the minority, doublet and triplets being the most abundant population. This has been theoretically predicted [25] and experimentally verified [26]. In this last case, in fact, even a sextet has been reported.

As a final remark, one might wonder why the signal of the allergens is so strong in the mono-dimensional SDS-PAGE profile (see Fig. 2) and rather weak in the 2D maps (see Fig. 3). Surprisingly, in fact, LTP, that gives a strongly reactive band in SDS-PAGE (see the zone at ca. 10 kDa in Fig. 2) gives a very faint zone in the immunoblots after 2D mapping. There is no doubt to our mind that LTP is indeed the major allergen in maize, as clearly identified in [4,5]. It should be emphasized that only recently scientists have begun to use 2D maps for allergen profiling, most of the work done in the past having been executed in mono-dimensional gels. It would appear that the signal in 2D maps is in general weak, as also witnessed by the study of Herndl et al. [27], who have adopted 2D mapping in the screening of allergens in *Malus domestica* (apple): as evinced from their 2D blots, the reactive bands are faint, including that of LTP, which is barely visible in the same map region as in our Fig. 4. The reason might rest in the two different types of denaturants adopted in SDS-PAGE as opposed to 2D maps. Quite a few authors believe that in the mixed SDS-protein micelle a substantial portion of the protein remains in a structured form. Thus Tanford [28] suggests that

these complexes can be described as “short, rigid, rod-like segments, with intervening regions possessing some flexibility”. In their “necklace model”, Shirahama et al. [29] and Takagi et al. [30] suggest that in the spherical micelles the polypeptide forms α -helices mostly in the hydrophobic region of the SDS micelles. According to Ibel et al. [31] the polypeptide chain is helically coiled around SDS micelles (flexible helix model). Also Mattice et al. [32] bring forward a flexible α -helix/random coil structure. Conversely, in 2D maps, due to the strong denaturing action of urea/thiourea (classical hydrogen and hydrophobic bond breaking agents), the polypeptide chains assume a random coil conformation. It is reasonable to hypothesize that in mono-dimensional SDS-PAGE proteins, due to partial maintenance of their structure, can give a much stronger reaction with antibodies than in 2D maps, where structural motifs have been largely lost.

5. Conclusions

The present study adds a non-negligible amount of previously unreported allergens to the panorama of maize proteins (vicilin, globulin-2, 50 kDa γ -zein, endo-chitinase, thioredoxin and trypsin inhibitor), considering that up to the present only LTP [4,5] was properly identified by direct chemical analysis and that the 50 kDa protein reported by Pasini et al. [8] as a major allergen had been detected in the insoluble, not in the cytoplasmic, salt-extractable proteins as in the present case. While the present list might not be exhaustive, nevertheless it adds substantial knowledge to the uncharted grounds of maize kernel allergens.

Acknowledgements

PGR was supported by Fondazione Cariplo (Milano), by PRIN 2009 (MURST, Rome) and by the European Community, grant “Allergy card”. We thank Dr. Angela Bachi (S. Raffaele Hosp., Milan) for identification of endo-chitinase.

REFERENCES

- [1] Hallauer AR, Russell WA, Lamkey KR. Corn breeding. In: Sprague GF, Dudley JW, editors. Corn and Corn Improvement. Number 18 in the series Agronomy, Third Edition Wisconsin, USA: American Society of Agronomy, Inc., Crop Science Society of America, Inc., and Soil Science Society of America, Inc.; 1988. p. 463–564.
- [2] Wright KN. Nutritional properties and feeding value of corn and its by-products. In: Watson SA, Ramstad PE, editors. Corn Chemistry and Technology. St. Paul, Minnesota: American Association of Cereal Chemists, Inc.; 1988. p. 447–78.
- [3] Moneret-Vautrin DA, Kanny G, Beaudouin E. L'allergie alimentaire au maïs existe-t-elle? Allerg Immunol (Paris) 1998;30:230.
- [4] Pastorello EA, Farioli L, Pravettoni V, Ispano M, Scibola E, Trambaioli C, et al. The major maize allergen, which is responsible for food-induced allergic reactions, is a lipid transfer protein. Allergy Clin Immunol 2000;106:744–51.

- [5] Pastorello EA, Pompei C, Pravettoni V, Farioli L, Calamari AM, Scibilia J, et al. Lipid-transfer protein is the major maize allergen maintaining IgE-binding activity after cooking at 100 °C, as demonstrated in anaphylactic patients and patients with positive double-blind, placebo-controlled food challenge results. *J Allergy Clin Immunol* 2003;112:775–83.
- [6] Maleki SJ. Food processing: effects on allergenicity. *Curr Opin Allergy Clin Immunol* 2004;4:241–5.
- [7] Nakajima O, Teshima R, Takagi K, Okunuki H, Sawada J. ELISA method for monitoring human serum IgE specific for Cry1Ab introduced into genetically modified corn. *Regul Toxicol Pharmacol* 2007;47:90–5.
- [8] Weichel M, Vergoossen NJ, Bonomi S, Scibilia J, Ortolani C, Ballmer-Weber BK, et al. Screening the allergenic repertoires of wheat and maize with sera from double-blind, placebo-controlled food challenge positive patients. *Allergy* 2006;61:128–35.
- [9] Pasini G, Simonato B, Curioni A, Vincenzi S, Cristaudo A, Santucci B, et al. IgE-mediated allergy to corn: a 50 kDa protein, belonging to the Reduced Soluble Proteins, is a major allergen. *Allergy* 2002;57:98–106.
- [10] Petersen A, Dresselhaus T, Grobe K, Becker WM. Proteome analysis of maize pollen for allergy-relevant components. *Proteomics* 2006;6:6317–25.
- [11] Li LC, Bedinger PA, Volk C, Jones AD, Cosgrove DJ. Purification and characterization of four beta-expansins (Zea m 1 isoforms) from maize pollen. *Plant Physiol* 2003;132:2073–85.
- [12] Wilson R. Sequence and assembly of the maize B73 genome. 50th Annual Maize Genetics Conference, Washington D.C., USA, February 27 – March 1; 2008. p. 18.
- [13] Hamdan M, Righetti PG. *Proteomics Today: Protein Assessment and Biomarkers Using Mass Spectrometry, 2-D Electrophoresis and Microarray Technology*. Hoboken: Wiley-VCH; 2005. p. 341–403.
- [14] Righetti PG, Boschetti E, Lomas L, Citterio A. Protein equalizer technology: the quest for a “democratic proteome”. *Proteomics* 2006;6:3980–92.
- [15] Boschetti E, Lomas L, Citterio A, Righetti PG. Romancing the “hidden proteome”, Anno Domini two zero zero seven. *J Chromatogr A* 2007;1153:277–90.
- [16] D'Ambrosio C, Arena S, Scaloni A, Guerrier L, Boschetti E, Mendieta ME, et al. Exploring the chicken egg white proteome with combinatorial peptide ligand libraries. *J Proteome Res* 2008;7:3461–74.
- [17] Olsen JV, de Godoy LM, Li G, Macek B, Mortensen P, Pesch R, et al. Parts per million mass accuracy on an Orbitrap mass spectrometer via lock mass injection into a C-trap. *Mol Cell Proteomics* 2005;4:2010–21.
- [18] Sennels L, Salek M, Lomas L, Boschetti E, Righetti PG, Rappsilber J. Proteomic analysis of human blood serum using peptide library beads. *J Proteome Res* 2007;6:4055–62.
- [19] Castagna A, Cecconi D, Sennels L, Rappsilber J, Guerrier L, Fortis F, et al. Exploring the hidden human urinary proteome via ligand library beads. *J Proteome Res* 2005;4:1917–30.
- [20] Antoniolli P, Fortis F, Guerrier L, Rinalducci S, Zolla L, Righetti PG, et al. Capturing and amplifying impurities from purified recombinant monoclonal antibodies via peptide library beads: a proteomic study. *Proteomics* 2007;7:1624–33.
- [21] Roux Dalvai F, Gonzalez de Peredo A, Simó C, Guerrier L, Bouyssie D, Zanella A, et al. Extensive analysis of the cytoplasmic proteome of human erythrocytes using the Peptide ligand library technology and advanced mass spectrometry. *Mol Cell Proteomics* 2008;7:2254–69.
- [22] E. Boschetti, L.V. Bindschedler, C. Tang, E. Fasoli, P.G. Righetti. Combinatorial peptide ligand libraries and plant proteomics: A winning strategy at a price. *J Chromatogr A* 2009;1216:1215–22.
- [23] Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, et al. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* 1997;25:3389–402.
- [24] Neville DM. Molecular weight determination of protein-dodecyl sulphate complexes by gel electrophoresis in a discontinuous buffer system. *J Biochem* 1971;246:6328–38.
- [25] Pietrogrande MC, Marchetti N, Dondi F, Righetti PG. Spot overlapping in two-dimensional polyacrylamide gel electrophoresis separations: a statistical study of complex protein maps. *Electrophoresis* 2002;23:283–91.
- [26] Campostrini N, Areces LB, Rappsilber J, Pietrogrande MC, Dondi F, Pastorino F, et al. Spot overlapping in two-dimensional maps: a serious problem ignored for much too long. *Proteomics* 2005;5:2385–95.
- [27] Herndl A, Marzban G, Kolarich D, Hahn R, Boscia D, Hemmer W, et al. Mapping of *Malus domestica* allergens by 2D electrophoresis and IgE-reactivity. *Electrophoresis* 2007;28:437–48.
- [28] Tanford C. *The hydrophobic effect. Formation of Micelles and Biological Membranes*. 2nd edition. New York: Wiley; 1980. p. 159–64.
- [29] Shirahama K, Tsujii K, Takagi T. Free-boundary electrophoresis of sodium dodecyl sulphate-protein polypeptide complexes with special reference to SDS-polyacrylamide gel electrophoresis. *J Biochem (Tokyo)* 1974;75:309–19.
- [30] Takagi T, Tsujii K, Shirahama K. Binding isotherms of sodium dodecyl sulphate to protein polypeptides with special reference to SDS-polyacrylamide gel electrophoresis. *J Biochem (Tokyo)* 1975;77:939–47.
- [31] Ibel K, May RP, Sandberg M, Mascher E, Greijer E, Lundahl P. Structure of dodecyl sulphate-protein complexes at subsaturating concentrations of free detergent. *Biophys Chem* 1994;53:77–83.
- [32] Mattice WL, Riser JM, Clark DS. Conformational properties of the complexes formed by proteins and sodium dodecyl sulphate. *Biochemistry* 1976;15:4264–72.
- [33] Neuffer MG. Growing maize for genetic studies. In: Freeling M, Walbot V, editors. *The Maize Handbook*. New York: Springer Verlag; 1994. p. 197–209.