Università degli Studi della Tuscia di Viterbo

DIPARTIMENTO DI SCIENZE E TECNOLOGIE PER L'AGRICOLTURA, LE FORESTE, LA NATURA E L'ENERGIA

CORSO DI DOTTORATO DI RICERCA
BIOTECNOLOGIE VEGETALI- XXIV° CICLO

IDENTIFICATION OF CHROMOSOME 5A ENCODED POLYPEPTIDES IN WHEAT KERNELS

AGR/07

Coordìnatore del corso: Prof.ssa Stefania Masci

Firma ……..

Tutor: Prof.ssa Stefania Masci / Prof. Domenico Lafiandra

Firma …… ……

Dottorando: Gaëtan Gombaud

Firma ……..

Abstract

This PhD work has been performed in the frame of the Italian project “Physical map of wheat chromosome 5A: Italian initiative for the sequencing of the whole genome”, that is part of the international initiatives ETGI & IWSC.

Wheat kernel proteins are represented for 80% of gluten and for 20% by soluble proteins. Whereas chromosome localization of the former is well known, that of soluble proteins, including polypeptides with structural and metabolic functions, still needs to be identified and mapped.

A proteomic approach has been performed to identify polypeptides encoded by genes on chromosome 5A. Intervarietal and interspecific chromosome substitution lines of durum and bread wheat have been used, leading to the identification of 48 soluble proteins in the two tetraploid genotypes taken into consideration, and a total of 86 proteins in the 5 bread wheat cultivars analysed.

The biological functions of most of the identified polypeptides are: stress/defense; carbohydrate metabolism, protein synthesis/assembly, storage, and a significative portion corresponds to unknown proteins.

The knowledge of polypeptides encoded by genes at chromosomes 5A will allow correlating their presence with specific physiological characteristics, along with quality properties.

INDEX

Abstract 2

I. Introduction 3

A. Wheat 4

1. Generalities 5

a) Phylogeny, origin and ploidy of wheat 5

b) Wheat in numbers 9

c) Wheat kernel: structure and composition 10

(1) The bran 11

(2) The endosperm 11

(3) The germ 14

2. Wheat proteins and their chromosome localization 15

a) Albumins / Globulins 17

b) Gluten 24

(1) Gliadins 24

(2) Glutenin 26

c) Chromosomal assignment 27

B. Proteomics 36

II. Aim of the work 38

III. Material and method 40

A. Plant Material 41

B. Methods 43

1. Metabolic and chloroform-methanol (CM) fractions 43

a) Protein extraction 43

b) Quantification 46

c) Two dimensional electrophoresis 46

d) Gel staining and acquisition 47

e) Image analysis 47

f) Mass spectrometry 49

2. Glutenin and gliadin fractions 51

a) Protein extraction and quantification 51

b) A-PAGE/SDS-PAGE for the gliadin fraction 51

3. Nuclear proteome 52

a) Isolation of leaf nuclei, protein extraction and quantification 52
b) One-dimensional electrophoresis 53
c) Image analysis 53
IV. Results and discussion 55
A. Tetraploid wheat 56
1. Set-up of the extraction procedure 56
2. Metabolic fraction 57
3. CM-like fraction 67
4. Gliadin and glutenin fraction 72
5. Nuclear proteome of leaves 74
B. Bread wheat 78
1. Identification of 5A polypeptides in the cultivar Hope 78
2. Identification of 5A polypeptides in the cultivar Thatcher 87
3. Identification of 5A polypeptides in the cultivar Timstein 95
4. Identification of 5A polypeptides in the cultivar Cheyenne 101
5. Identification of 5A polypeptides in the cultivar Chinese Spring 108
6. Comparison between various bread wheat cultivars 113
V. Conclusion 117
Acknowledgments 119
References 120

I. Introduction

A. Wheat

Wheat is one of the most cultivated cereal crops in the world. Its productivity, good nutritional value as well as the capacity to adapt itself to a varied environment has made this cereal one of the most important sources for the human and animal nutrition since ancient time. Additionally to its adaptability to grow over a wide range of climatic conditions and soil fertility, wheat has other advantages: it is easily transported and safely stored over long periods of time, and it can be easily transformed into a broad range of products. This ability of wheat depends mainly on gluten, the storage proteins conferring the capacity to form, after hydration, a coherent mass, insoluble and with visco-elastic properties. However, also others compounds, for example the starch, influence the qualitative characteristics and the possible final uses of the wheat flour (Zeng et al., 1997; Yasui et al., 1999; Lee et al., 2001; Merita et al., 2002; Bhattacharya et al., 2002). Moreover, wheat is not used only for human (65%) and animal (21%) feeding, but also in the textile industries, the preparation of glucose syrups, the production of adhesive and biodegradable plastic (Orth and Shellenberger, 1988).

Two principal types of wheat are cultivated: the durum wheat (*Triticum durum*) and the bread wheat *(Triticum aestivum*). Durum wheat is tetraploid (2n=28), with genome AABB, and is mostly used in the pasta and semolina industry. Bread wheat is hexaploid (2n=42), with genome AABBDD, and is used in the bread-making and cookies industry.

1. Generalities

a) Phylogeny, origin and ploidy of wheat

All *Triticum* species are native to the 'Fertile Crescent' (which encompasses the Eastern Mediterranean, South-Eastern Turkey, Northern Iraq and Western Iran, and its neighbouring regions of the Transcaucasus and Northern Iran). The cultivated wheat appeared 10,000 years ago. Bread wheat (*Triticum aestivum* L.) was one of the first domesticated food crops and has been the major staple in Europe, North Africa and Central Asia for 8,000 years. Einkorn (diploid, genome AA) and emmer (tetraploid, AABB) wheats were the first cultivated forms, and the only ones subjected to domestication selection. Domesticated emmer wheat spread widely across the Near East and beyond. Then emmer wheat cultivation expanded eastward through the Mesopotamian plain to India, and westward through Anatolia to the Mediterranean coastal region and Europe. In these regions, it was one of the most prominent crops for almost 6,000 years.

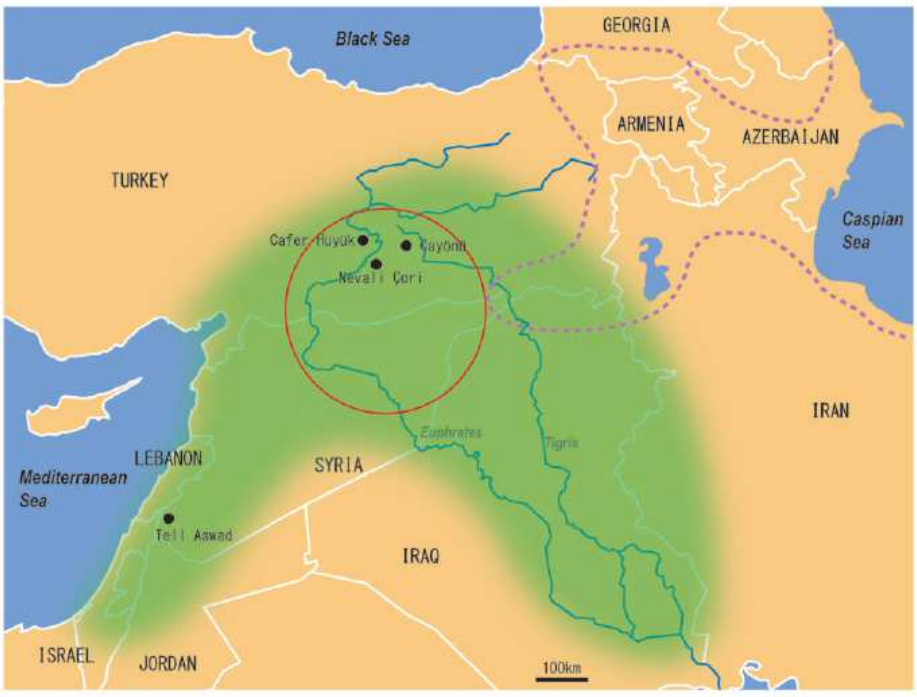


Figure 1: Map of the Near East that includes the Fertile Crescent (shaded in green). The red circle and dashed purple line respectively denote the 'core area' where agriculture is thought to have emerged and the western peripheral region of the natural range of *Ae. tauschii* where the allopolyploid speciation of *T. aestivum* supposedly took place (from Matsuoka, 2011).

Wheat belongs to the genus *Triticum* of the *Poaceae* family, which also includes the related crop species, rye and barley. The genus *Triticum* consists of six species which differ by their degree of ploidy, ranging from diploid to hexaploid with chromosome numbers 14, 28 or 42. The six species are *Triticum monococcum* L. (AA genome); *Triticum urartu* Tumanian ex Gandilyan (AA genome); *Triticum turgidum* L. (AABB genome); *Triticum timopheevii* (Zhuk.) Zhuk. (AAGG genome); *Triticum aestivum* L. (AABBDD genome); and *Triticum zhukovskyi* Menabde & Ericz. (AAAAGG genome). Moreover these species are grouped into three sections: Section Monococcon (diploid species); Section Dicoccoidea (tetraploid species); and Section Triticum (hexaploid species). Of these species, *T. urartu* exists only in its wild form, whereas *T. aestivum* and *T. zhukovskyi* exist only as cultivated forms. The other species, *T. monococcum, T. turgidum* and *T. timopheevii*, have both a wild and a domesticated form.

Bread wheat has the three genomes AA, BB and DD made up each of seven pairs of homeologous chromosomes numbered from 1 to 7 for a total of 42 chromosomes; durum wheat contains only the two genomes AA and BB and 28 chromosomes.

Table 1: The nomenclature of wild and cultivated *Triticum* wheats (according to Van Slageren 1994)

Section	Species and subspecies	Genome constitution	Example of common names
Monococcon	*Triticum monococcum* L.	AA	
	subsp. *aegilopoides*		Wild einkorn
	subsp. *monococcum*		Cultivated einkorn
	Triticum urartu Tumanian ex Gandilyan	AA	
Dicoccoidea	*Triticum turgidum* L.	AABB	
	subsp. *dicoccoides*		Wild emmer
	subsp. *dicoccon*		Cultivated emmer
	subsp. *durum*		Durum or macaroni wheat
	subsp. *polonicum*		Polish wheat
	subsp. *turanicum*		Khorassan wheat
	subsp. *turgidum*		Rivet wheat
	subsp. *carthlicum*		Persian wheat
	subsp. *paleocolchicum*		Georgian wheat
	Triticum timopheevii	AAGG	
	subsp. *Armeniacum*		Wild timopheevii
	subsp. *timopheevii*		Cultivated timopheevii
Triticum	*Triticum aestivum* L.	AABBDD	Common wheat
	subsp. *aestivum*		Bread wheat
	subsp. *compactum*		Club wheat
	subsp. *sphaerococcum*		Indian dwarf wheat
	subsp. *macha*		
	subsp. *spelta*		Spelt
	Triticum *zhukovskyi*	AAAAGG	

The phylogenesis of wheat is complex and not completely known. The genome A of tetraploid and hexaploid wheats are related to the A genome of wild and cultivated einkorn, while the genome B derived, probably, from the S genome of an unidentified diploid *Aegilops* species and the D genome comes from *Aegilops squarrosa* (also called *Triticum*).

The natural hybridization between *T. monococcum* and an *Aegilops* species carrying the B genome gave origin to a wild tetraploid wheat with the AABB genome (*T. turgidum* ssp. *dicoccoides*) which then gradually evolved to *T. turgidum* ssp. *dicoccum* and finally in *T. durum*, which is the currently cultivated durum wheat.

The cultivated bread wheat (with the AABBDD genome) is derived from a natural hybridization between *T. turgidum* ssp. *dicoccum* (AABB genome) and *Aegilops squarrosa* (DD genome).

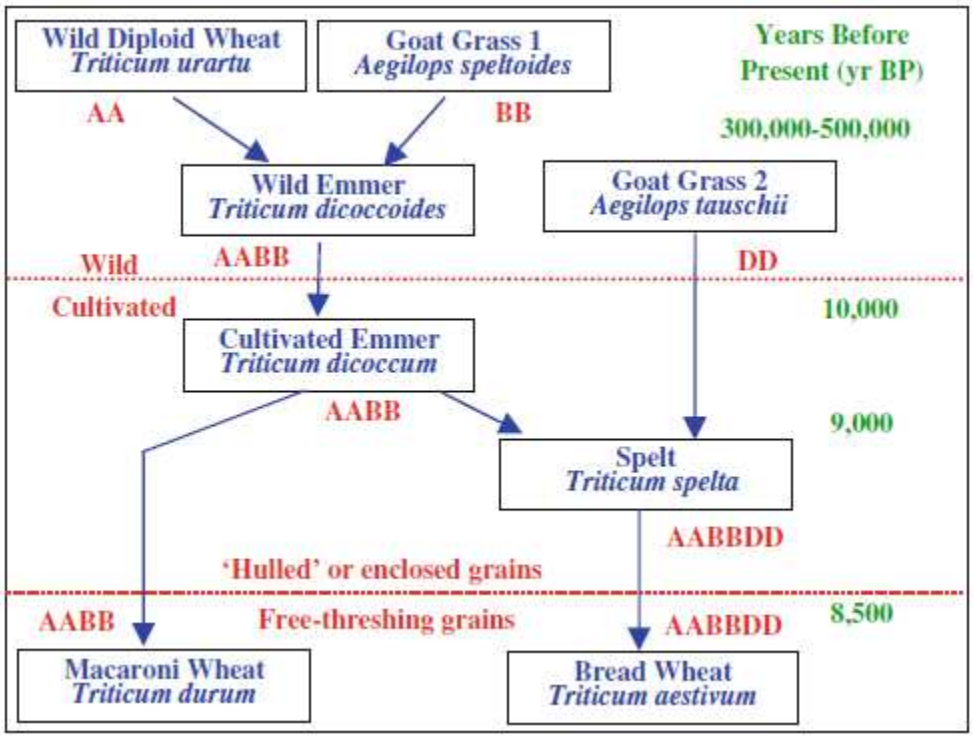


Figure 2: The evolution of wheat from the prehistoric Stone Age grasses to modern durum wheat and bread wheat (from Peng et al., 2011).

Plant domestication involves a range of complex morphological, physiological, and genetic changes referred to as the 'domestication syndrome' (Peleg et al., 2011). These changes, such as large fruits, increased apical dominance, loss of seed dormancy, and synchronized growth and flowering, distinguish crops from their wild progenitors.

The loss of natural seed dispersal, which results in the seeds being retained in the spike, facilitating harvesting, was a key event in the domestication of most cereals. Genetic studies have shown that the non-brittle rachis is controlled by recessive alleles at two major loci, *Brittle rachis-A1* (*Br-A1*) and *Brittle rachis-B1* (*Br-B1*) that are located respectively on the short arm of chromosomes 3A and 3B (Watanabe et al., 2002). It is thought that the spikes of non-brittle mutated plants were consciously selected by early farmers and that thus their frequency increased constantly in cultivated fields. But this process was slow and establishment of the non-brittle ancient cultivar took over one millennium.

A second important trait is the change from hulled forms, in which the glumes adhere tightly to the grain, to free-threshing naked forms. Indeed the early wheat varieties were characterized by hulled seeds that required drying to be liberated from the chaff, but harvesting became efficient when the farmers selected species with low degree of glume tenacity and with fragile rachis. Free-threshing wheats have thinner glumes and paleas that allow an early release of naked kernels. The free forms arose by a dominant mutant at the *Q* locus which modified the effects of recessive mutations at the *Tg* (tenacious glume) locus. Free-threshing genotypes, like *T. durum* and *T. aestivum*, represent the final steps of wheat domestication.

In addition to qualitative traits, also quantitative traits have been selected during domestication, as grain yield, seed size, plant height and heading date, flowering time, plant height, spike number/plant, spike weight/plant, single spike weight, kernel number/plant, kernel number/spike, kernel number/spikelet and spikelet number/spike (Table 2).

Table 2: Quantitative trait loci related to domestication in wheat (from Peng et al., 2011)

Trait	Number of QTL effects	Residing chromosome[a]
Seed size/weight	7	1A, 2A, 3A, 4A, 7A, 5B, 7B
	8	1B, 2A, 4A, 5A, 5B, 6B, 7A, 7B
	8	1B, 2A, 3A, 3B, 4B, 5A, 6A, 7A
Flowering time	4	2A, 4B, 5A, 6B
Grain yield	8	**1B**, **2A**, 3A, **5A**, 5B
	3	3A, 4A, 5A
Plant height	4	**5A**, **7B**
Spike number/plant	7	**1B**, 2A, 2B, 5A, **7A**
Spike weight/plant	10	**1B**, **2A**, 3A, **5A**, 5B, **7A**
Single spike weight	5	1B, 2A, 3A, **5A**
Kernel number/plant	9	**1B**, **2A**, 3A, **5A**, 5B, 7A
Kernel number/spike	7	1B, **2A**, 3A, **5A**, 6B
Kernel number/spikelet	7	1B, 2A, 3A, **5A**, 5B, 7B
Spikelet number/spike	6	1B, 2A, **5A**, **6B**

a The bolded chromosomes carry a pair of linked QTL

b) Wheat in numbers

Wheat is cultivated on more than 225 million hectares, representing the largest cultivated crop. In 2010, worldwide production of wheat has reached 651 million tons (http://faostat.fao.org/) (Figure 3 & 4). About 95% of the wheat grown worldwide is hexaploid bread wheat, and the 5% remaining is mostly tetraploid durum wheat. Others wheat species are still grown, in small amounts. China ranks first with 16,9% of the total world production, followed by India (11,8%), Russia (9,1%), the United States (8,8%) and France (5,6%), but the entire European Union is the largest producer with 143 million tonnes in 2010.

The most suitable latitudes for wheat cultivation are between 30° and 60° North and between 27° and 40° South (Nuttonson 1955). Wheat grows from 0 to 3000 m o.s.l. Wheat growing is optimal at 25°C, with a minimum of 3-4°C and a maximum of 30-32°C.

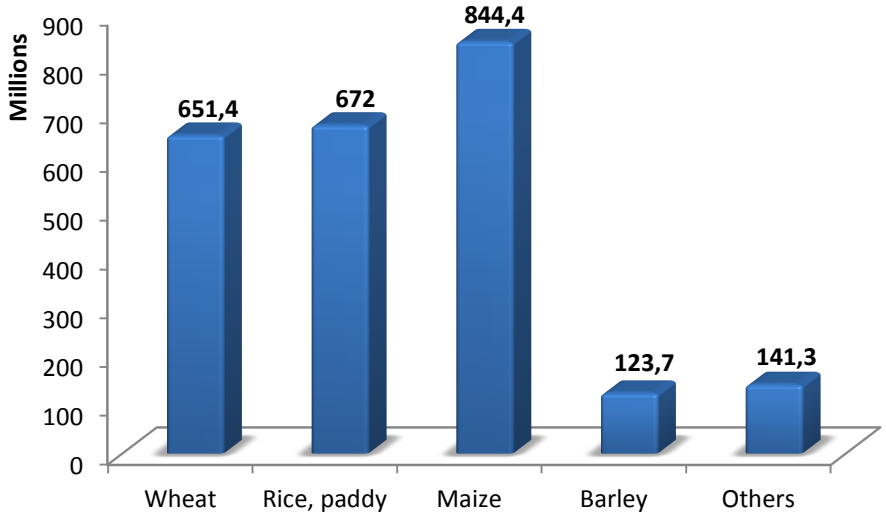


Figure 3: World cereals production (millions tons) in 2010 (http://faostat.fao.org)

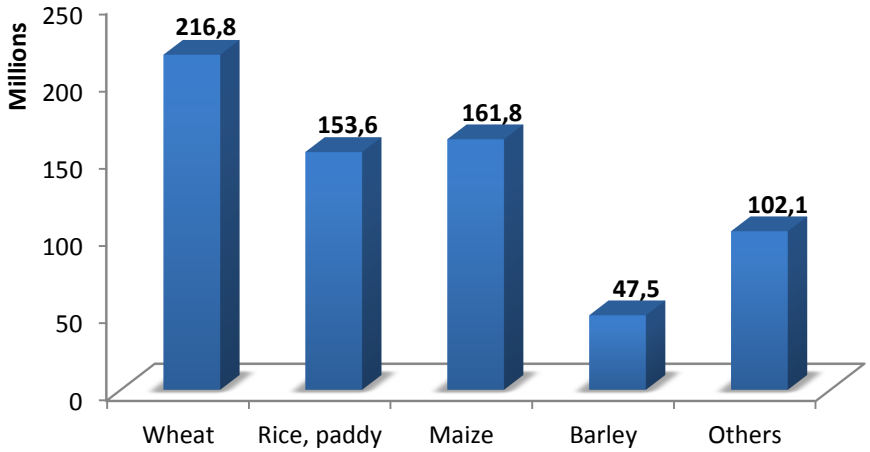


Figure 4: World cereals area harvested (millions ha) in 2010 (http://faostat.fao.org)

c) Wheat kernel: structure and composition

The edible part of wheat is the caryopsis (or kernel), a dry and indehiscent fruit, made up of a seed and teguments. The length of the grain is between 5 and 8 mm, its width between 2 and 4 mm, its thickness between 2,5 and 3,5 mm, its longitudinal section between 10 and 16 mm², its transverse section between 4 and 7,5 mm², its weight between 20 and 50 mg and its density between 1,3 and 1,4.

The grain is mainly constituted of starch (approximately 70 %), proteins (10-15 %) and of pentosans (8-10 %); other minors constituents (some % only) are lipids, cellulose, free sugars, minerals and vitamins (Feillet, 2000).

These constituents are not equally distributed in the wheat kernel. Table 3 reports the distribution of the major components in wheat kernel.

Table 3: Chemical composition of the whole wheat grain with its various parts. Average percentages converted on a dry matter basis (from Belderok, 2000)

	Whole grain	Endosperm	Bran	Germ
Proteins	16	13	16	22
Fats	2	1.5	5	7
Carbohydrates	68	82	16	40
Dietary fibres	11	1.5	53	25
Minerals	1.8	0.5	7.2	4.5

(1) The bran

The bran, envelope of the seed and the fruit, is formed by different cell layers. The pericarp, composed by the outer and the inner pericarp, surrounds the entire kernel. The epidermis, the hypodermis and the innermost layer, called the remnants of thin-walled cells compose the outer pericarp. The inner pericarp is composed by the cross cells and the tube cells. The testa, also called seedcoat, forms another layer firmly joined to the tube cells. The nucellar epidermis (hyaline layer) is tightly bound to the internal surface of the testa. Bran is particularly rich in dietary fiber, whereas starch, vitamins, proteins, and minerals are present in significant quantities.

(2) The endosperm

The endosperm is constituted by the starchy albumen and the aleurone layer (together 80-85 % of the grain mass). It is composed for 80% of carbohydrate, for 12-17% of proteins and for 1-2% of fats. The starchy albumen comprises cells filled with starch granules embedded in the middle of a protein matrix. The aleurone layer, typically only one cell layer thick at maturity, includes the starchy endosperm and part of the embryo. The aleurone cells are block-shaped (37-65 µm by 25-75 µm) within a single kernel and have thickened (6-8µm) double-layered cellulosic walls. The endosperm is mainly a source of nutrients for the embryo when the germination begins.

(a) The starch

Starch is the major component of the endosperm of the cereal grains and accounting for 65–75% of the grain dry weight and up to 80% of the endosperm dry weight. It is the stored form of energy: during germination, amylase enzymes (both synthesized and deposited during the period of grain filling or *de novo* synthesized upon germination) break the starch down to release glucose units for the developing embryo, roots, and shoots. Starch is also the most common carbohydrate and the major source of energy for human diet, providing it in a ‘‘slow-release’’ form that is well suited to our digestive systems.

Starch is composed only of D-glucose, linked by α-1-4 bonds to form linear chains, whereas branches are formed through the connection of α-1-4 linked chains via α-1-6 linkages.

Starch is made up of two types of polymers: amylose and amylopectin, which differ in degree of polymerization and branch frequency. Depending on the plant, starch generally contains 20 to 25% amylose and 75 to 80% amylopectin. In hexaploid and durum wheats, amylose content ranges from about 18 to 35%.

In all higher plants, starch is packaged into starch granules which are characterized by their density (1.5 – 1.6 g/cm3) and by their semi-crystalline nature, as indicated by their characteristic birefringence under polarized light (Buléon et al, 1998). The size distribution of wheat starch includes two, and sometimes three size classes of granules, a feature shared with starches from other *Triticeae*, notably barley and rye. The larger “A” granules have a diameter of 15-30 μm whereas “B” granules, initiated after the“A: granules, have a diameter that is typically below 10 μm. A third “C” class of very small granules is seen under cooler environmental conditions allowing a long grain fill period. Additionally wheat starch granules contain a defined set of prominent proteins located within the interior of the starch granule, including Granule-Bound Starch Synthase, Starch Synthase I, Branching Enzymes IIa and IIb, and Starch Synthase IIa. Starch granules extracted from mature grain are associated with a range of surface located proteins which become bound as the maturation and desiccation of the grain leads to the disruption of the amyloplast membrane, largely described as the purindolines, friabilins and grain softness proteins (Darlington et al. 2000; Baldwin 2001).

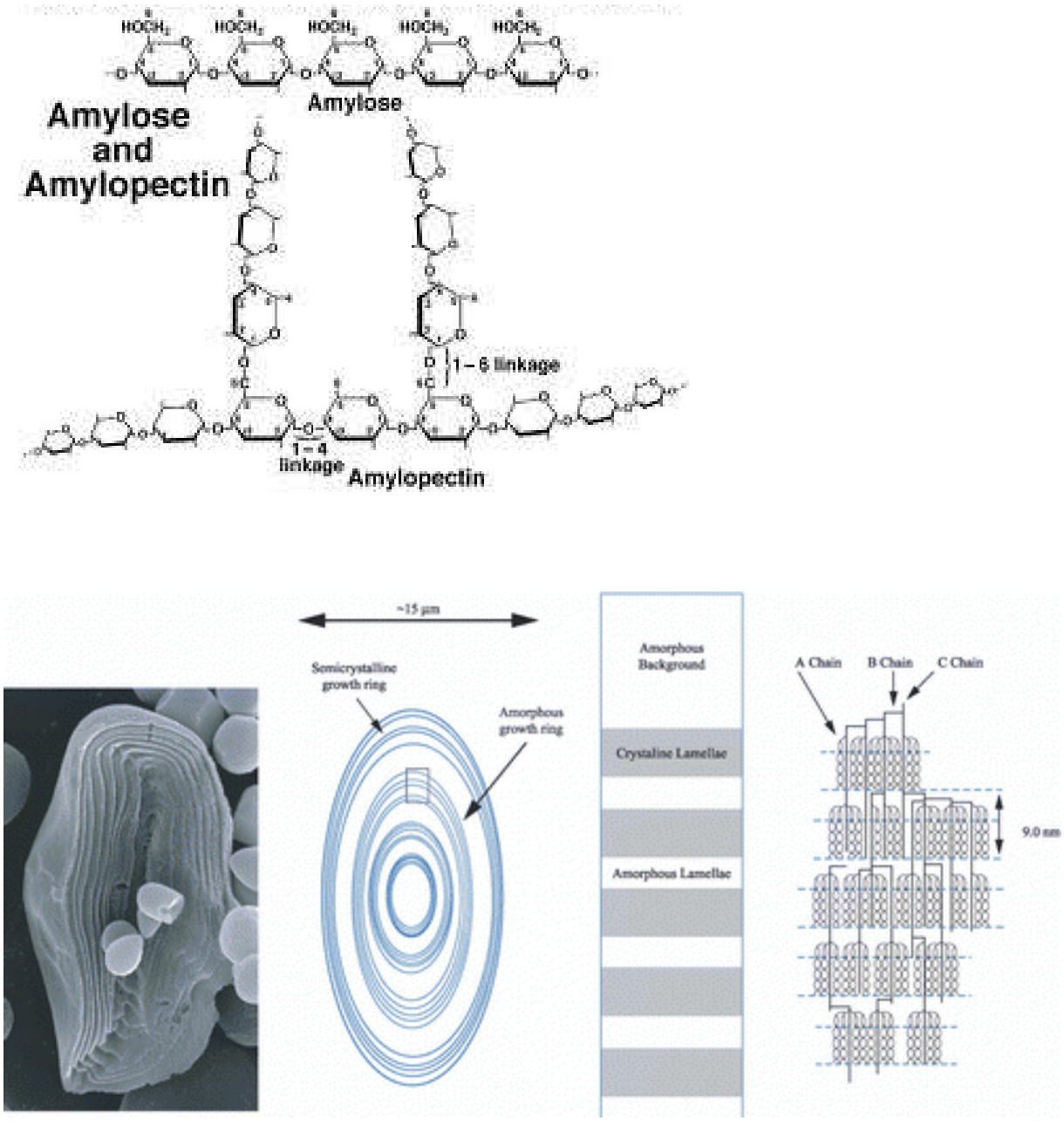


Figure 5: Structure of the amylose, the amylopectin and the starch granule (from Ball and Morell, 2003)

(3) The germ

The germ (3 %) consists of the embryo and the scutellum. This latter contains some storage proteins, is adjacent to the endosperm and serves to absorb nutrients from the endosperm during germination. The germ is a rich source of protein, sugar and oil. It is also quite rich in vitamin E and in B-vitamins.

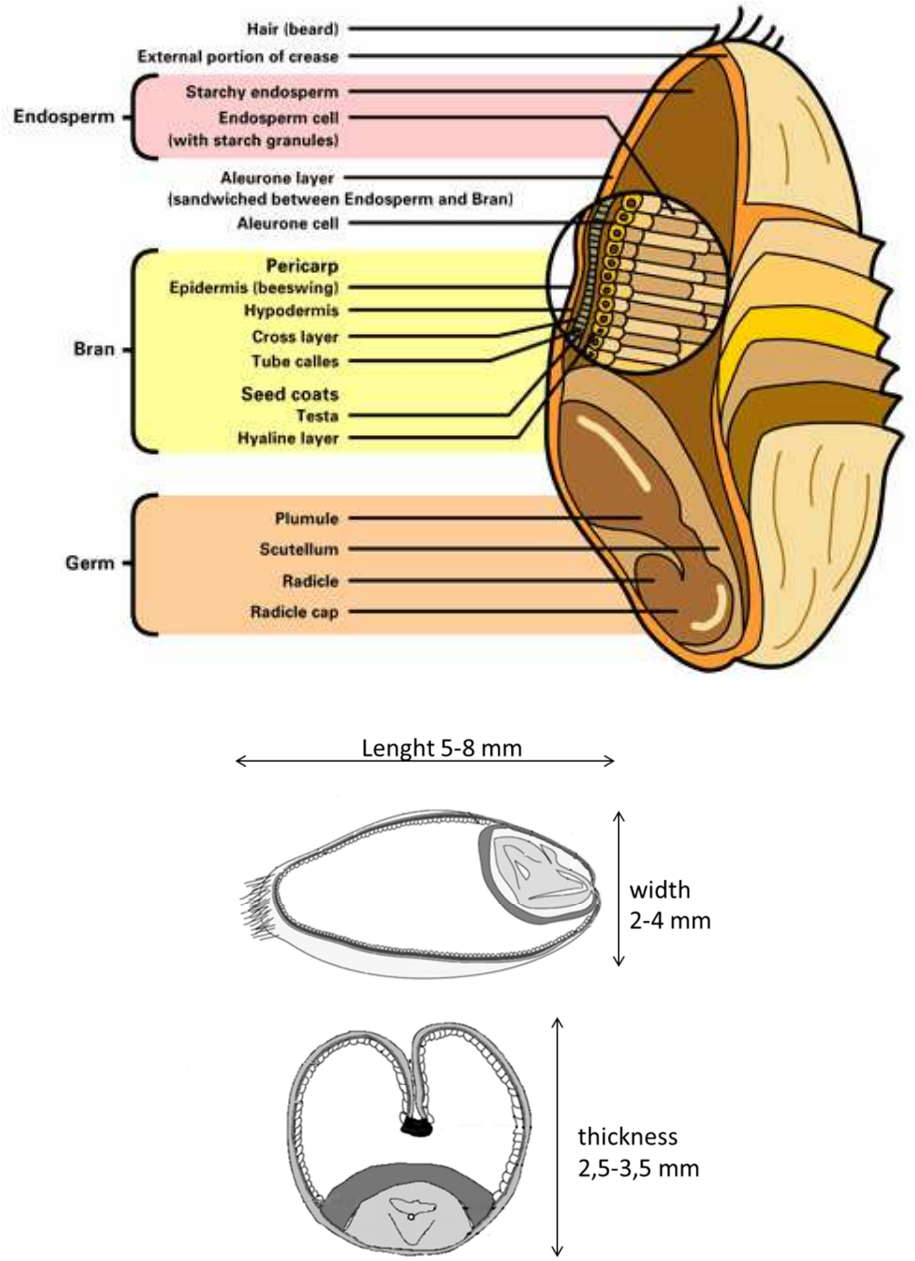


Figure 6: Structure and dimension of the wheat grain

2. Wheat proteins and their chromosome localization

At the beginning of the 20th century, a systematic study was conducted by Osborne (1907) to develop a classification for cereal-seed proteins based on their sequential extraction and differential solubility. The wheat proteins were classified into four different groups: albumins, soluble in water and dilute buffers; the globulins, soluble in saline solutions; gliadins, soluble in water/ethanol solution; and glutenins, partially soluble in the dilute acid solutions and in some detergents or dissociating agents. Albumins and globulins are generally proteins having biological functions essential to the physiology of the grain (enzymes, carriers, enzymes inhibitors …). The gliadins (monomeric) and the glutenins (polymeric) constitute the storage proteins of the grain. They are usually called prolamins because of their particular amino-acids composition in which a repetitive domain rich in **prol**ine and in glut**amin**e is present, that is partly responsible for the original rheological properties of wheat-based doughs (Shewry, et al., 1990). Prolamins are only localised in starch-based endosperm. The germ and the aleurone layer are the compartments richest in albumins and globulins (Table 4).

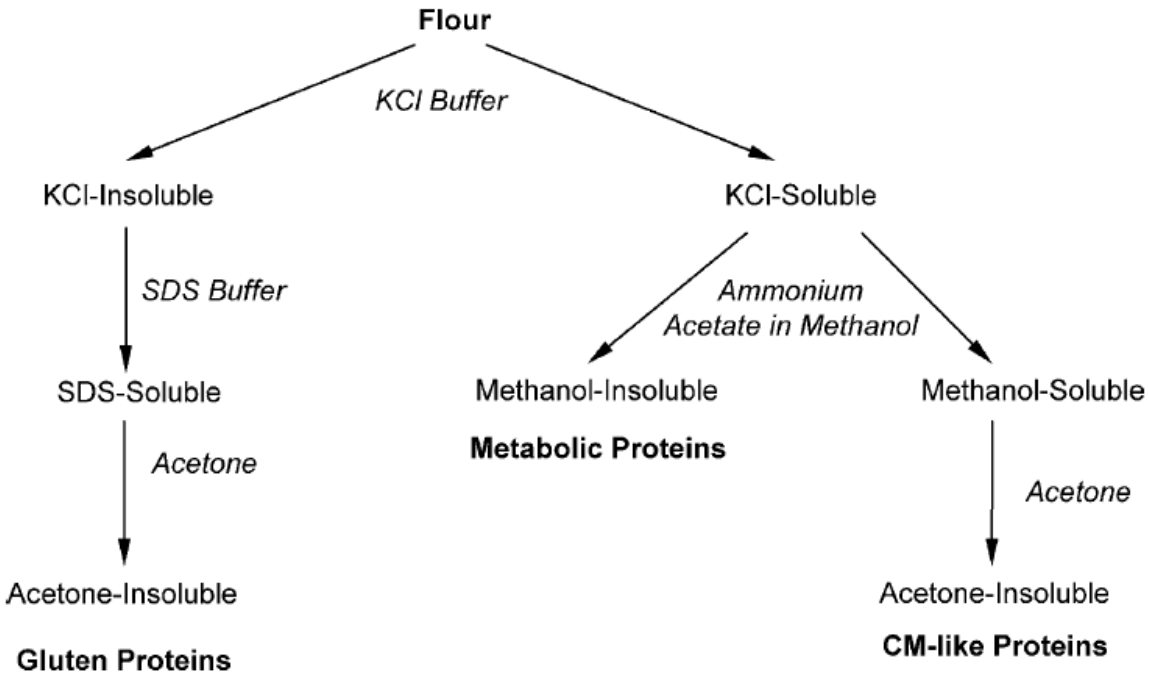


Figure 7: Separation of wheat flour proteins into gluten, metabolic and CM-like protein fractions (from Hurkman and Tanaka, 2004).

Table 4: Distribution of various proteins in the wheat grain (Popineau, 1988).

Grain part	Weight (% whole seed)	Proteins content (% dry content)	Proteins
Germ	3	35-40	Albumins Globulins
Aleurone layer	7-9	30-35	Albumins Globulins
Pericarp	4	6-7	
Starchy Albumen	80	9-14	Prolamins Albumins Globulins

The determination of the prolamins primary structures allowed Miflin et al. (1983) and Shewry, et al. (1986) to propose a more precise classification of wheat proteins based on the composition in sulfur amino-acids, their structure and their function in the grain. According to this classification, prolamins correspond to the storage proteins which are characterized by their composition in amino-acids very rich in proline and in glutamine. Their function is only the storage of nitrogen, in the form of amino-acids, usable during germination. This classification was recently detailed by Shewry and Halford (2002). Prolamins are separated in three groups: prolamins low in sulfur, prolamins rich in sulfur and prolamins of high molecular weight. However, to completely describe the complexity of prolamins family, it is also necessary to take into account the fact that they present different aggregation states at maturity: it is in fact possible to find monomeric gliadins and polymeric glutenins.

Table 5: Wheat proteins classification according to Shewry and Halford, 2002

Structural and functional proteins (30%)	Prolamins (70%)		
	Sulfur-poor	Sulfur-rich	High Molecular Weight
Albumins, globulins, others	ω-gliadins LMW-GS type D	α/β-, γ-gliadins LMW-GS type B LMW-GS type C	HMW-GS
	30-75 kDa (10-20%)	30-45 kDa (70-80 %)	65-90 kDa (6-10%)

a) Albumins / Globulins

Albumins and globulins account for approximately 20% of total proteins of the wheat flour. Indeed soluble proteins are complex mixtures containing metabolic enzymes, hydrolytic enzymes necessary for seed germination and enzyme inhibitors (Payne and Rhodes, 1982). Albumins contain mainly structural proteins, enzymes or enzymes inhibitors and storage proteins like 2S albumins. Globulins, present in the germ, the aleurone layer and the endosperm, contain mostly storage proteins and some proteins having physiological functions during the maturation of seed. They are rich in lysine and arginine but poor in tryptophan, asparagine and glutamine.

Also present, the triticins, proteins related to legumins (the seed storage proteins of legumes), account for 5% of the total seed proteins. When considering their solubility properties, they behave as globulins. They are located in protein bodies in the starchy endosperm of the wheat grain.

The genes of albumins and globulins of wheat have been assigned to chromosomes 3, 5, 6 and 7 (Garcia-Olmedo, et al, 1982).

Protocols developed to obtain protein fractions enriched in albumins and globulins (KCl-soluble) have allowed the analysis of this largely unexplored protein fraction (Vensel et al., 2005). By using a classical proteomic approach based on 2DE, 254 proteins, grouped in 13 different biochemical processes of the endosperm, were identified in this fraction: ATP interconversion reactions, carbohydrate metabolism, cell division, cytoskeleton, lipid metabolism, nitrogen metabolism, protein synthesis/assembly, protein turnover, signal transduction, protein storage, stress/defense, transcription/translation, and transport (Vensel et al., 2005). A complete list of the protein identifications by Vensel et al (2005) is reported in Table 6.

More recently, Dupont et al (2011) have analyzed the total wheat flour proteins and 233 spots were identified. Among these amylase/protease inhibitors represent 4.1%, triticins 1.6%, serpins 1.6%, avenin-like proteins 1.7%, beta-amylase 0.5%, globulins 0.4%, other enzymes and factors 1.9%, and all other 3% (the remaining is composed of α-gliadins for 20.4%, LMW-GS for 18.0%, HMW-GS for 17.1%, γ-gliadins for 12.2%, and ω-gliadins for 10.5%).

SwissProt no.	Protein	SwissProt no.	Protein
Q8S1A5	Carbamoyl phosphate synthetase	P55307	Catalase isozyme 1
Q8L5C2	4SNc-Tudor protein, (NTPase)	Q93YR3	Heat shock associated protein
Q8L5C2	4SNc-Tudor protein, (NTPase)	Q8RZW7	Selenium binding protein
Q42669	Aconitase	Q8LST6	Aldehyde dehydrogenase
Q9LZF6	Cell division cycle protein	Q9FPK6	Aldehyde dehydrogenase
P49608	Aconitase	Q8GU01	Globulin-2
O23927	Pyruvate Pi dikinase	Q9M4Z1	ADP-glucose PPase, SS
O23755	Elongation factor 2	O49218	Methylmalonate-semialdehyde dehydrogenase
O98447	ClpC protease	Q43772	UDP-glucose PPase
Q43638	Heat shock protein, 82K, precursor	Q43772	UDP-glucose PPase
Q9LF88	Late embryogenesis abundant protein-like	Q9M4Z1	ADP glucose PPase, SS
Q8W0Q7	Methionine synthase	Q9ASP4	Dihydrolipoamide dehydrogenase
Q9XGF1	Heat shock protein 80–2	Q93YR3	Heat shock associated protein
Q9M6E6	Poly(A)-binding protein	Q8W3W6	LMW glutenin subunit group 3 type II
Q9AT32	Poly(A)-binding protein	Q8GU18	LMW glutenin subunit
Q9M6E6	Poly(A)-binding protein	Q9ASP4	Dihydrolipoamide dehydrogenase
P93616	Poly(A)-binding protein	Q8GU18	LMW glutenin subunit group 3 type II
P93616	Poly(A)-binding protein	P52894	Alanine amino transferase 2
P93616	Poly(A)-binding protein	Q8GU01	Globulin-2
Q39641	Heat shock protein 70	P52894	Alanine amino transferase 2
P93616	Poly(A)-binding protein	Q42971	Enolase
P93616	Poly(A)-binding protein	Q42971	Enolase
Q01899	Heat shock protein 70	Q9ZRR5	Tubulin a-3 chain
Q9SPK5	10-Formyltetrahydrofolate synthetase	Q9ZRB0	Tubulin b-3 chain
Q40058	DNAK-type molecular chaperone HSP70	Q42971	Enolase
Q95I76	F23N19.10, putative stress-induced protein	O81237	6-Phosphogluconate dehydrogenase
P93616	Poly(A)-binding protein	Q8W3N9	26S Proteasome regulatory particle triple-A ATPase subunit 3
Q8L724	Stress-induced protein, sti1-like		
Q9FME2	RNA-binding protein, similarity	Q8VZ47	Argininosuccinate synthase-like protein
Q9FME2	RNA-binding protein, similarity	Q9AUV6	UDP-glucose dehydrogenase
Q95NX2	Phosphoglucomutase	Q08837	Triticin
Q944F5	Fructokinase	Q8W516	SGT1
O65305	Acetohydroxyacid synthase	Q8W3W4	LMW glutenin subunit group 4 type II
Q9ZR86	Protein disulfide isomerase-like protein	Q9FXT8	26S proteasome regulatory particle triple-A ATPase subunit 4
Q94DV7	Phosphoglycerate mutase, 2,3-bisphosphoglycerate-independent	P41378	Eukaryotic initiation factor 4A
		Q40058	DNAK-type molecular chaperone HSP70
Q94DV7	Phosphoglycerate mutase, 2,3-bisphosphoglycerate-independent	Q9FXT8	26S Proteasome regulatory particle triple-A ATPase subunit 4
Q8GWX8	2-Isopropylmalate synthase	P37833	Aspartate amino transferase
Q8W4M5	PPi-fructose-6-P 1-phosphotransferase	Q9SAU8	Heat shock protein 70
Q41141	PPi-fructose 6-P 1-phosphotransferase beta subunit	Q9FXT8	26S Proteasome regulatory particle triple-A ATPase subunit 4b (10 dpa)
Q9LWT6	Chaperonin 60 kDa beta subunit	Q9XGU8	Isocitrate dehydrogenase (NAD) (36 dpa)
P52589	Protein disulfide isomerase		
O49485	Phosphoglycerate dehydrogenase-like protein		

O82783	Importin a-2 subunit
P16098	β-Amylase
P12299	ADP-glucose PPase, LS
P12299	ADP-glucose PPase, LS
P12299	ADP-glucose PPase, LS
P30184	Leucine amino peptidase
P30184	Leucine amino peptidase
Q8RZF3	Ketol-reductoisomerase
Q9M7E0	Elongation factor 1-a
Q8RZW7	Selenium binding protein
Q8GU01	Globulin-2

Q8GU01	Globulin-2
Q9ZRI8	Formate dehydrogenase
P37833	Aspartate amino transferase
P93693	Serpin WZS2
P37833	Aspartate amino transferase
Q9ST58	Serpin
Q43492	Serpin homolog WZS3
P37833	Aspartate amino transferase
P93692	Serpin homolog WZS3
P04727	a/b-Gliadin clone PW8142
Q9FUS4	Actin

P93692	Serpin homolog WZS3
Q40676	Aldolase
P93692	Serpin homolog WZS3
Q41593	Serpin
Q9ZR33	Reversibly glycosylated polypeptide (10 dpa)
P93692	Serpin homolog WZS3 (36 dpa)
Q9ZRI8	Formate dehydrogenase
Q9ZR33	Reversibly glycosylated polypeptide
Q9ZRI8	Formate dehydrogenase
Q40676	Aldolase
Q8LK23	Peroxidase 1
O81221	Actin
Q41319	Acyl-acyl-carrier protein desaturase
Q93Y71	Protein disulfide-isomerase precursor
Q40676	Aldolase
Q8LK23	Peroxidase 1
Q9M4V4	Glyceraldehyde-3-P dehydrogenase (NAD)
Q40069	Peroxidase BP1
P25861	Glyceraldehyde 3-P dehydrogenase (NAD)
Q8LK23	Peroxidase 1
Q9ST58	Serpin
Q94CS6	Legumin-like protein
Q942N5	Auxin-induced protein (10 dpa)
Q9M4V4	Glyceraldehyde-3-P dehydrogenase (NAD) (36 dpa)
Q9M4V4	Glyceraldehyde-3-P dehydrogenase (NAD)
P26517	Glyceraldehyde-3-P dehydrogenase (NAD)
Q94KS2	TGF-b receptor-interacting protein 1
Q9XGC6	Adenosine kinase
Q03678	Globulin Beg 1
Q9FRV1	Chitinase-a
Q94KS2	TGF-b receptor-interacting protein 1
P41095	60S Acidic ribosomal protein P0
Q94CS6	Legumin-like protein
Q94CS6	Legumin-like protein

Q8LKV8	Seed globulin
Q04832	DNA-binding protein HEXBP
Q9FRV0	Chitinase-c
O23983	Ascorbate peroxidase
O23983	Ascorbate peroxidase
Q9FER4	20S Proteasome a-subunit
O23983	Ascorbate peroxidase
Q944C6	Small ras-related GTP-binding protein
P11955	Chitinase-a
P27919	Avenin
Q9FS79	Triosephosphate isomerase
Q8LKV8	Seed globulin
Q9FS79	Triosephosphate isomerase
P34937	Triosephosphate isomerase
Q8LK23	Peroxidase 1
Q8LKV8	Seed globulin
Q8S4P7	Thaumatin-like protein TLP7
Q9FER4	20S Proteasome a-subunit
P29546	Elongation factor 1-b
Q945R5	Ascorbate peroxidase
Q9LMK1	F10K1.21/F7A7_100 protein, similarity
Q9AXH7	1-Cys peroxiredoxin
Q09114	Avenin N9
P52572	Peroxiredoxin
Q9LSU2	20S Proteasome a-subunit B
Q96185	Superoxide dismutase [Mn]
Q9MB31	GSH-dependent dehydroascorbate reductase 1
Q93VQ6	Expressed protein
P16347	a-Amylase/subtilisin inhibitor
Q84UH6	Dehydroascorbate reductase
Q93XQ6	Cyclophilin A-2
Q09114	Avenin N9
Q9LMK1	F10K1.21/F7A7_100 protein, similarity
Q9MB31	GSH-dependent dehydroascorbate reductase 1

Q9C774	26S Proteasome regulatory subunit S12
Q09114	Avenin N9
Q9FT00	Malate dehydrogenase (NAD)
Q94JA2	Malate dehydrogenase (NAD)
P49027	Guanine nucleotide-binding protein b subunit-like protein
Q945R5	Ascorbate peroxidase
Q42988	PPi-fructose-6-P 1-phosphotransferase
Q9ZWJ2	Glyoxalase I
Q40676	Aldolase
Q07810	Tritin
T06212d)	Glucose and ribitol dehydrogenase
Q9C5Y9	Initiation factor 3g
Q8W5L9	Purple acid phosphatase
Q7X653	OSJNBb0118P14.5
Q8L5C6	Xylanase inhibitor protein I
Q7X653	OSJNBb0118P14.5
Q8L5C6	Xylanase inhibitor protein I
Q8LKV8	Seed globulin
P29305	14–3-3 Protein homolog
Q8LKV8	Seed globulin
Q8LKV8	Seed globulin
P01084	a-Amylase inhibitor 0.53
O49956	a-Amylase inhibitor Ima 1, monomeric
Q09114	Avenin N9
P35687	40S Ribosomal protein S21
Q40641	Polyubiquitin 6
Q40058	DNAK-type molecular chaperone HSP70
Q944R8	UDP-glucose dehydrogenase
P40978	40S Ribosomal protein S19
Q93W25	Cyclophilin A-1
Q43223	Sucrose synthase type 2
Q9LF88	Late embryogenesis abundant protein-like
Q8S7U3	Embryo-specific protein
Q8GU01	Globulin-2
Q8GU01	Globulin-2
Q41551	LMW glutenin (fragment)
P26517	Glyceraldehyde 3-P dehydrogenase (NAD)
T06212d)	Glucose and ribitol dehydrogenase
Q03678	Globulin Beg 1
Q8L8I0	Globulin-like protein
Q9LMK1	F10K1.21/F7A7_100 protein, similarity
Q9LST9	b-1 Subunit of 20S proteasome
Q9MB31	GSH-dependent dehydroascorbate reductase 1
Q9ZSU2	Translation initiation factor 5A
Q8LRM8	Translationally controlled tumor protein
Q93XQ8	Protein disulfide isomerase 2 precursor
Q41561	Heat shock protein 16.9C
Q41518	Glycine-rich RNA-binding protein
Q41518	Glycine-rich RNA-binding protein
Q42973	Ubiquitin-protein ligase
P35686	40S Ribosomal protein S20 (10 dpa)
Q43659	Grain softness protein 1b (36 dpa)
P17314	a-Amylase/trypsin inhibitor, CM3
P23345	Superoxide dismutase [Cu-Zn]
Q43472	Glycine-rich RNA-binding protein
Q02254	Nucleoside diphosphate kinase I
Q9XHS0	40S Ribosomal protein S12
Q95QG8	Pathogenesis-related protein 4
P01085	a-Amylase inhibitor 0.19
Q8S3L1	Glutaredoxin
P01084	a-Amylase inhibitor 0.53
Q9ST58	Serpin
P12783	Phosphoglycerate kinase
P93438	S-Adenosylmethionine synthetase 2
Q9FXT8	26S Proteasome regulatory A subunit
P50299	*S*-Adenosylmethionine synthetase 1
Q945R5	Ascorbate peroxidase
Q945R5	Ascorbate peroxidase
Q9ZR33	Reversibly glycosylated polypeptide
Q9LF88	Late embryogenesis abundant protein-like
Q8GU01	Globulin-2
Q8GU01	Globulin-2
P26517	Glyceraldehyde 3-P dehydrogenase (NAD)
Q43247	Glyceraldehyde 3-P dehydrogenase (NAD)
Q9FF52	60S Ribosomal protein L12
P28814	Barwin
P26517	Glyceraldehyde 3-P dehydrogenase (NAD)
Q8LK23	Peroxidase 1
Q947H4	Plasmodesmal receptor
P24296	Nonspecific lipid-transfer protein precursor

Table 6: Proteins of wheat endosperm identified by MS (redrawn from Vensel et al., 2005)

	Predominant protein	Predicted MW	pI
Farinins			
	Farinin Bu-1 full length	29 978	8,1
	Farinin Bu-1 C-terminus	18 832	8,4
	Farinin Bu-1 C-terminus	18 832	8,4
	Farinin Bu-2	30 567	7,5
	Farinin Bu-2	30 567	7,5
	Farinin Bu-2	30 567	7,5
	Farinin Bu-3	30 883	7,9
	Farinin Bu-3	30 883	7,9
Purinins			
	Purinin Bu-1	20 272	5,9
	Purinin Bu-1	20 272	5,9
	Purinin Bu-2	20 592	6,2
	Purinin Bu-2	20 592	6,2
	Purinin Bu-3	22 371	6,2
	Purinin Bu-3	22 371	6,2
Triticins (Tri-1)			
	Triticin [GenBank:DR736644] N-terminal subunit	inc	inc
	Triticin TC11_285558, N-terminal subunit	40 501	6,2
	Triticin TC11_285558, N-terminal subunit	40 501	6,2
	Triticin TC11_285558, N-terminal subunit	40 501	6,2
	Triticin TC11_285558, N-terminal subunit	40 501	6,2
	Triticin TC11_264477, C-terminal subunit	21 830	8,1
	Triticin TC11_264477, C-terminal subunit	21 830	8,1
Globulins			
	Globulin-1 [GenBank:ABG68030] (Glo-2)	22 941	8,6
	Globulin-2 Bu-17295	53 832	6,6
	Globulin-2 Bu-17295	53 832	6,6
	Globulin-2 Bu-17366	inc	7
	Globulin-2 Bu-18428	53 554	6,6
	Globulin-2 Bu-18428	53 554	6,6
	Globulin Glo-3-type TC234094	inc	inc
	Globulin Glo-3-type TC11_305389	inc	inc
	Globulin Glo-3-type TC11_ 305389	inc	inc
	Globulin Glo-3-type TC234094/WTAI-CM3 [SwissProt: P17314]	nd	nd
GSP and Puroindoline (Pin-D1)			
	Grain softness protein [GenBank:CAA56591]	16 157	8,1
	Grain softness protein [GenBank:CAA56586]	16 381	7,6
	Puroindoline-b [GenBank:AAT40244]	14 812	9
	Puroindoline-b [GenBank:AAT40244]	14 812	9
Alpha-amylase and protease inhibitors			
	CMx1/CMx3 TC11_308146	14 027	8,1

	CMx1/CMx3 TC11_309398	13 891	8
	WASI [SwissProt: P16347]	19 633	6,8
	WCI [GenBank:CAD19440]	12 943	7,4
	WDAI TC11_338524	13 239	5,7
	WDAI [GenBank:AAV91972]	13 191	5,2
	WDAI [SwissProt:P01085]	13 337	6,7
	WMAI [PRF:223520]	13 342	6,2
	WMAI [PRF:223520]	13 342	6,2
	WTAI-CM1 TC11_340510	13 096	6,7
	WTAI-CM2 [SwissProt:P16851]	13 034	6,2
	WTAI-CM2 [SwissProt:P16851]	13 034	6,2
	WTAI-CM3 [SwissProt:P17314]	15 832	6,7
	WTAI-CM3 [SwissProt:P17314]	15 832	6,7
	WTAI-CM16 [SwissProt:P16159]	13 437	5
	WTAI-CM16 [SwissProt:P16159]	13 437	5
	WTAI-CM17 [GenBank:CAA42453]	13 502	4,9
	WTAI-CM17 [GenBank:CAA42453]	13 502	4,9
	WCI [GenBank:CAD19440]/wheatwin-Bu-2/trypsin inhibitor factor TC11_315743	nd	nd
Serpins			
	Serpin Bu-1 Type 1b, like [GenBank:ACN59483]	37 667	5,4
	Serpin Bu-1 Type 1b, like [GenBank:ACN59483]	37 667	5,4
	Serpin Bu-1 Type 1b, like [GenBank:ACN59483]	37 667	5,4
	Serpin Bu-1 or Bu-4	inc	inc
	Serpin Bu-1 or Bu-4	inc	inc
	Serpin Bu-2 Serpin Z1c, like [SwissProt:Q9ST58]	42 882	5,6
	Serpin Bu-2 Serpin Z1c, like [SwissProt:Q9ST58]	42 882	5,6
	Serpin Bu-3, Z1a type [Swiss-Prot: P93693]	43 118	5,6
	Serpin Bu-3, Z1a type [Swiss-Prot: P93693]	43 118	5,6
	Serpin Bu-4 or Bu-5	inc	inc
	Serpin Bu-5, like [GenBank: CAA72274]	42 981	5,2
	Serpin Bu-5, like [GenBank: CAA72274]	42 981	5,2
	Serpin Bu-7, like [GenBank:ACN59484]	43 431	5,1
	Serpin Bu-7, like [GenBank:ACN59484]	43 431	5,1
Other Inhibitors			
	Tritin TC235992	29 653	9,8
	Xylanase inhibitor XIP-1 [PDB:1OM0]	30 285	8,3
Beta-amylase (b-Amy-A1, b-Amy-B1, b-Amy-D1)			
	Beta-amylase Bu-18	60 016	6,9
	Beta-amylase Bu-1	60 016	6,9
	Beta-amylase Bu2	54 481	5,9
	Beta-amylase Bu2	54 481	5,9
	Beta-amylase Bu3	54 319	5,8
	Beta-amylase Bu3	54 319	5,8
	Beta-amylase Bu3	54 319	5,8

Other enzymes			
	ADP-glucose PP lg subunit [GenBank:CAD98749]	53 030	5,6
	ADP-glucose PP lg subunit [GenBank:CAD98749]	53 030	5,6
	ADP-glucose PP sm subunit [GenBank:AAF61173]	52 061	5,5
	ADP-glucose PP sm subunit [GenBank:AAF61173]	52 061	5,5
	Alanine amino transferase TC11_282456	52 820	6,1
	ATP-synthase beta-subunit [GenBank:CAA52636]	58 562	5,4
	Chitinase, rye, [GenBank:BAB18520]	26 095	8,7
	Chitinase [GenBank:AAX83262]	26 022	8,3
	Dehydroascorbate reductase TC264934	23 358	5,9
	Enolase TC11_292359	48 033	5,4
	Glyoxalase I TC11_288238	32 568	5,4
	Glucose/ribitol dehydrogenase RS_UWI_14903	31 851	6,3
	Ketol-acid reducto isomerase TC234371	57 486	5,4
	Malate dehydrogenase [GenBank:AAT64932]	35 486	5,8
	Malate dehydrogenase [GenBank:AAT64932]	35 486	5,8
	Methionine synthase RS_UWI_10957	84 552	5,7
	Orthophosphate dikinase TC11_322894	73 501	5,8
	PDI3 [GenBank:AAK49425]	54 094	4,9
	Sucrose synthase 2 [GenBank:CAA03935]	92 608	6,2
	Thiamine biosynth enzyme TC11_308909	33 167	5,7
	Triose-phosphate isomerase [GenBank:CAC14917]	26 803	5,4
	27 K thiol reductase-like TC11_300123	23 642	6,1
	27 K thiol reductase like TC11_299048	23 788	6,1
Other			
	Elongation factor EF1A [Swiss-Prot: Q03033]	49 169	9,2
	HSP70 (Butte 86) [GenBank:AAB99745]	71 031	5,1
	Initiation factor Eif4A [Swiss-Prot: P41378.1]	46 928	5,3
	LTP Bu-2	9 606	8,2
	Thaumatin-like protein TC11_283136	21 408	7,9

Table 7: Proteins of wheat endosperm identified by MS (redrawn from Dupont et al., 2011)

b) Gluten

Gluten proteins (or prolamins) are the principal cereal storage proteins, except for oat and rice (Shewry et al., 1995), and can represent up to 80% of the total proteins.
Prolamins are divided into gliadin and glutenin according to their alcohol-solubility (Osborne, 1907).

(1) Gliadins

Gliadins are heterogeneous mixtures of single-chained polypeptides which are soluble in 70% aqueous alcohol. They are subdivided in four groups according to their electrophoretic mobility at acid pH: α-, β-, γ- and ω-gliadins, although α– and β-gliadins are actually grouped in α/β-gliadins, since they are structurally identical (Kasarda et al, 1987). Gliadins are monomeric proteins, with intramolecular disulfides bonds (for α/β- and γ-gliadins). They interact by non-covalent connection between them or with glutenin polymers. Their molecular weight is between 30 and 45 kDa for α/β- and γ-gliadins and between 45 and 75 kDa for the ω-gliadins (Masci et al., 2002).

Using one-dimensional electrophoresis, gliadins of a single wheat grain can be separated into 20–25 components (Bushuk and Zillman 1978; Autran et al 1979; Wrigley et al 1982; Metakovsky et al 1984). Two-dimensional electrophoresis allows better separation with a resolution of up to 50 components (Wrigley 1970; Payne et al 1982; Lafiandra and Kasarda 1985; Pogna et al 1990).

The γ-gliadins differ from α/β-gliadins in the amount of aspartic acid, proline, methionine, tyrosine, phenilalanine, and tryptophan (Bietz et al 1977). The ω-gliadins differ in amino acid composition from other gliadins and do not have cysteine. The ω-gliadins are characterized by high levels of glutamine (+glutamate) (40–50 mol %), proline (20–30 mol %), and phenylalanine (7– 9 mol %), which represent >80% of the total amino acid residues (Tatham and Shewry 1995). All gliadins are low in the ionic amino acids (histidine, arginine, lysine, and free carboxylic groups of aspartic acid and glutamic acid). Also, gliadins can be classified according to their N-terminal amino acid sequence.

They show a large polymorphism and a variety of wheat can contain from 20 to 50 different gliadins (Brown et al., 1981; Pogna et al., 1990). The gliadins are coded by genes located on chromosomes 1 and 6 of the three genomes A, B and D, at the loci *Gli-1* and *Gli-2*.

Gliadin content contribute to dough viscosity and extensibility (Shewry et al, 2003), whereas glutenins to elasticity.

Genetics and polymorphism:

Using starch gel electrophoresis, Boyd and Lee (1967) and Shepherd (1968) have analyzed the compensating nullisomic–tetrasomic and ditelosomic series developed by Sears (1954, 1966) in the bread wheat cv. Chinese Spring. This led them to report that gliadin proteins are controlled by genes present on the short arm of the homoeologous group 1 and 6 chromosomes. The poor resolution of the one-dimensional separations and the overlapping of many gliadin components meant that it was possible to assign only a few of them to specific chromosomes.

This chromosomal assignment was confirmed by two-dimensional electrophoresis in Chinese Spring as well as in different bread wheat cultivars (Brown et al., 1981; Payne et al., 1982; Lafiandra et al., 1984). Similar studies of durum wheat (Joppa et al., 1983; Lafiandra et al., 1987) and wild relatives of wheat have also confirmed these results (Lafiandra et al., 1993).

Extensive studies of the inheritance of gliadins in the progeny of specific crosses have indicated that the major gliadin genes occur in tightly linked clusters, termed blocks, with intrablock recombination being rare (Sozinov and Poperelya, 1980). It is now generally accepted that the ω- and γ-gliadins are controlled by clusters of tightly linked genes present at the *Gli-1* loci (*Gli- A1*, *Gli-B1*, and *Gli-D1*) on the short arms of the homoeologous group 1 chromosomes, whereas the α/β-gliadins are controlled by the *Gli-2* loci (*Gli-A2*, *Gli-B2*, and *Gli-D2*) present on the short arms of the group 6 chromosomes (Payne, 1987). Shewry et al., 1984 have attributed this spatial separation of gliadin genes on the group 1 and 6 chromosomes to an ancient interchromosomal translocation, with the *Gli-2* locus originating from the translocation of a γ-type gene from chromosome 1 to chromosome 6, followed by divergence of the coding sequence to give rise to the α-type sequence. The *Gli-1* loci have been shown to be present on the distal parts of the group 1 chromosomes, showing independent or loose linkage with their respective centromeres (Shepherd, 1988). The *Gli-2* loci have been studied in less detail, but telocentric mapping showed 35% recombination between the *Gli-A2* locus and the centromere (Payne, 1987). The individual *Gli-1* and *Gli-2* loci exhibit extensive polymorphism, as detected by electrophoretic techniques, with allelic blocks differing in

the numbers, proportions, and mobilities of different components. This results in a great diversity of gliadin patterns, providing the basis for distinguishing different wheat cultivars (Metakovsky, 1991).

(2) Glutenin

Glutenins, which are heterogeneous mixtures of polymers stabilized by inter- and intrachain disulfides bonds, represent 40% of the total protein content. The glutenin polymeric structure can reach very high molecular sizes (Wrigley, 1996). The glutenins can be divided according to their molecular weight after reduction of disulphide bonds, in High Molecular Weight (HMW) and Low Molecular Weight (LMW) glutenin subunits, with the latter being further divided into B-, C-, and D-type subunits according to size, isoelectric points, and composition (Payne and Corfield, 1979; Jackson et al., 1983). The C and D groups of LMW subunits are highly similar in sequence to gliadins, and are considered to be derived from these components by mutations resulting in the presence of odd cysteine residues, which are able to form interchain disulfide bonds. In contrast, the B-type LMW subunits form a discrete group (D'Ovidio and Masci, 2004).

High Molecular Weight Glutenin Subunits (HMW-GS)

Even if the HMW-GS are the minor components in terms of quantity, they are the key factors for breadmaking, because they are the major determinants of gluten elasticity. Their molecular weight, estimated by SDS-PAGE, is around 80-130kDa.

Using nullisomic-tetrasomic, nullisomic-trisomic, and ditelocentric lines of Chinese Spring, Bietz et al (1975) showed that HMW-GS were controlled by genes at loci *Glu-1* on the long arms of the chromosomes 1A, 1B and 1D. More detailed studies about the genetics of the HMW-GS and their relationship to breadmaking quality were conducted by Payne et al (1981) and Lawrence and Shepherd (1981).

Low Molecular Weight Glutenin Subunits (LMW-GS)

The LMW-GS (B-, C-, and D-subunits) represent about one-third of the total seed protein and ~60% of total glutenins. Because of the difficulty in identifying them in one-dimensional SDS-PAGE gels, they have been less studied than the HMW-GS.
The LMW-GS are controlled by genes at the *Glu-A3*, *Glu-B3*, and *Glu-D3* loci on the short arms of chromosome 1A, 1B, and 1D, respectively, but there is also evidence that

some LMW-GS (especially of the C-group) are controlled by genes on group-6 chromosomes (Masci et al, 2002). Two LMW-GS with molecular weights of ≈30–31,000 Da (*Glu-D4* locus) and 32,000 Da (*Glu-D5* locus) were reported to be coded at genes located on chromosomes 1D and 7D, respectively, although their exact location within the chromosome has not been established. (Gianibelli et al., 2001)

c) Chromosomal assignment

The wheat genome size is equivalent to five times that of the human genome and forty times that of rice (Vitulo et al., 2011) (Table 8). Moreover it has been shown that over 30,000 genes are expressed in the developing wheat grain (Wan et al., 2008). Even if the sequencing of the wheat genome is not yet completed, earlier studies have shown the chromosomal assignment of some genes and proteins.

Table 8: Genome size for some commercially important cereal crops

Cereal	Genome size (Mb)
Rice (*Oryza sativa*)	420
Maize (*Zea mays*)	2500
Barley (*Hordeum vulgare*)	4800
Wheat (*Triticum aestivum*)	16000

Group 1:

The triticin are encoded by the short arm of chromosome 1A (*Tri-A1* locus) and 1D (*Tri-D1* locus). Recently, genes coding for triticin were found on chromosome 1B (*Tri-B1* locus), but no protein expressed by this locus have been identified, suggesting that this locus is inactive (Dubcovsky, et al., 1997).

The LMW glutenin subunits are encoded by genes located on the short arm of the group 1 chromosomes at the loci *Glu-A3*, *Glu-B3* and *Glu-D3*. The genes encoding the HMW glutenin subunits are located on the long arm of chromosome group 1 at the loci *Glu-A1*, *Glu B1* and *Glu-D1* (Jackson, et al., 1983; Payne, et al., 1984; Singh, et al., 1988).

Also ω- and γ -gliadins genes have been located on the short arm of chromosome 1, at the *Gli-1* loci (*Gli-A1*, *Gli-A2*, and *Gli-A3*) (Payne et al., 1985)

Merlino et al., 2009 have mapped the thioredoxin H-type on the chromosomes 1AL and 1BL, a Heat Shock Protein (HSP70), an alpha-1-purothionin and a wheatwin1 on the 1BL. They have also mapped on the chromosome 1BS a glutathione-S-transferase. Also purothionin (apoprotein) are coded by genes on the chromosomes 1AL, 1BL and 1DL (Garcia-Olmedo et al., 1982).

In 2001, Singh et al., have assigned a 23kDa globulin to the chromosome 1D.

Glucose phosphate isomerase and malate dehydrogenase have been mapped on the 1A, 1B and 1D (Cox et al., 1987).

Group 2:

For the homologous group 2 Singh and Skerritt (2001) have found the α-amylase inhibitor to be encoded by genes on chromosomes 2A and 2B, and the α-amylase and a 42 kDa protein on chromosome 2D.

Merlino et al., 2009 have assigned to the chromosome 2BS the 1-cys peroxiredoxin.

Wu et al. (1999) have assigned to the homologous group 2 the mitochondrial manganese superoxide dismutase (MnSOD) gene.

Group 3:

Alpha-amylase inhibitors have been also assigned to the chromosomes 3BS and 3DS (Singh and Skerritt, 2001).They have also assigned a 32kDa protein and a 20kDa to the chromosome 3AL and 3Ds respectively.

A 16,9kDa Class I Heat shock protein was assigned to genes on the chromosomes 3AS, 3BS and 3DS. Moreover two others Small heat shock protein (HSP17,5 and HSP17,8) have been located on the 3BS and 3DS (Merlino et al.,2009). They have also assigned to the 3AS the ATP synthase beta subunit.

The analysis of grain esterase isozymes in Chinese Spring aneuploid genotypes by IEF indicated that genes on the long arms of chromosomes 3A, 3B and 3D control the production of 19 isozymes (Ainsworth et al., 1984).

Group 4:

Alpha-amylase inhibitors have been also assigned to the chromosome 4B and 4D (Carbonero, et al., 1999; Singh and Skerritt, 2001; Islam, et al., 2003).

The β-amylase enzymes are encoded by genes on the long arms of chromosomes 4A and 4D.

Three proteins with 51, 52, 56 kDa molecular weights have been located on the 4A chromosome. The serine carboxypeptidase II and a 66kDa protein have been assigned to the 4BS (Singh and Skerritt, 2001).

Payne et al (1985) have shown that globulins are controlled by genes on chromosome group 4, more precisely on the 4BL and 4DS. Moreover also two globulins, with molecular weights of 39 and 50kDa are encoded at the chromosome 4DS (Singh and Skerritt, 2001).

Also the CM-like proteins CM3 have been assigned to the chromosome 4A (Garcia-olmedo et al., 1982).

Group 5:

The UDP-glucose pyrophosphorylase, a serpin, a single-stranded nucleic acid binding protein, a 0.19 dimeric alpha-amylase inhibitor have been assigned to genes on the chromosome 5BL (Merlino et al., 2009).

In 2009 Chikmawati et al., have mapped the elongation Factor-1 alpha on the 5DL, a ribosomal protein S29-like protein on the 5AL and 5DL, a Lipid Transfer Protein 7a2b on the 5DL, a thioredoxin on the 5DL and the 5BS, a methionine synthase protein on the 5AS and 5BS. They have also confirmed the assignment of the UDP-glucose pyrophosphorylase made by Merlino et al. (2009) on the 5BL but have also assigned it to the 5AL and 5DL.

Singh and Skerritt (2001) have found a 38kDa albumin and two globulins to be encoded by genes on the chromosome 5DL.

Studies have shown that three orthologous loci for the Grain Softness Protein (GSP) exist in hexaploid wheat. They are located on the distal end of the homologous group 5. The GSP-1 locus on 5DS is closely linked to the puroindoline genes and thus is linked to grain texture. Clearly, GSP-1 is closely related to the puroindolines and is a member of the same protein ‘superfamily’ that includes α-amylase/trypsin inhibitors, the ‘CM’ proteins, and non-specific lipid transfer proteins (Morris, 2002).

In 2009, Pérez et al. have localized the Rad50 loci, responsible for a protein involved in the repair of breaks in double-stranded DNA, on the short arm of chromosomes 5A, 5B and 5D.

By using substitution lines, Veisz and Sutka (1998) have confirmed that chromosomes 5A, 5B and 5D carry genes responsible for frost resistance.

5A Chromosome:

The chromosome 5A with a predicted size of 827Mb represents 4,9% of the wheat genome and carries genes controlling important traits such as vernalization requirement (*TaVRT-1*), cold tolerance and abiotic stress tolerance, disease resistance (e.g. Fusarium head blight) and domestication traits (e.g. free threshing Q gene). Considering the predicted chromosome arm lengths of 295Mb and 532Mb for 5AS and 5AL, the coding fraction can be estimated as 1.08% and 1.30% in the short and the long arm, respectively. Considering an average coding sequence length of 2,000 bp, the number of genes was estimated at about 1,593 genes on the short arm and 3,495genes on the long arm, for a total of 5,088 genes for the whole chromosome. The gene ontology (Figure 6) allowed the association of 75,834 reads with 87 terms: 37 Biological Process, 25 Molecular Function and 25 Cellular Component. Biological Process terms were associated to 22,816 reads, Molecular Function categories annotated 44,196 reads, ‘‘binding’’ and ‘‘DNA binding’’ function accounting for 34% ; Cellular Component categories annotated 8,822 terms, 37% of them as ‘‘Membrane’’(Vitulo et al.,2011).

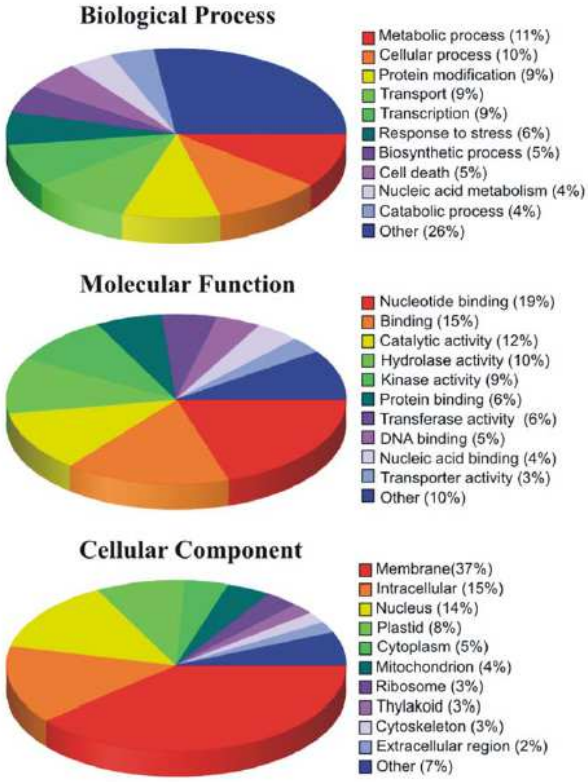


Figure 8: Percentage distribution of the GO entries relative to chromosome 5A (from Vitulo et al., 2011).

Tsujimoto and Noda (1990) have mapped, by cytological methods, a suppression speltoid gene (*Q*), which confers the free-threshing character of the spike and influences other important agronomic traits, and a β-amylase gene on the long arm of the 5A chromosome.

Chromosome 5A has been shown to carry alleles affecting important yield components and also at least two loci affecting height. Indeed Snape et al (1985), by using single-chromosome recombinant lines, have located *Vrn1*, *q* and *b1* genes on the long arm of the 5A. *Vrn1* is one of the main determinants of the winter/spring grown habit polymorphism; *q* determines ear morphology and *b1* gene determines the presence of awns. For the *Vrn1* also two others loci have been mapped on the chrosomome 5D *(Vrn-D1)* and on the chromosome 5B (*Vrn-B1).*

By the use of homozygous deletion lines of wheat for 5AL, generated in the variety Chinese Spring, the frost resistance gene *Fr1* has been mapped on the long arm of the 5A chromosome (Sutka et al., 1989; Snape et al., 2001).

The gene controlling the compact spike morphology was located on the chromosome 5AL (Kosuge et al., 2011).

In 2010, Burt et al have identified a single major QTL conferring resistance to eyespot on the long arm of chromosome 5A.

Others QTL have been localized on the 5A chromosomes by Jantasuriyarat et al (2004) by using recombinant inbred lines produced in the frame of the International Triticae mapping Initiative (ITMI). One QTL for glume tenacity which explains 15-22% of the phenotypic variance, one QTL affecting threshability responsible for 10% of the phenotypic variance, one QTL affecting the rachis fragility and explaining 16% of the phenotypic variance, and finally one QTL which affects the spike compactness and explains 14% of the phenotypic variance, were found.

Moreover it was showed that expression of genes localized in other chromosomes are regulated by the 5A chromosome during cold hardening. In fact, the transcription of 78 genes (39 up-regulated) proved to be chromosome 5A-dependent. These genes encoded proteins involved in transcriptional regulation, defence processes and carbohydrate metabolism (Kocsy et al., 2010).

Group 6:

For the homologous group 6 Merlino et al (2009) have localized the HSP70, the small HSP Hsp23.5 and the monomeric alpha-amylase inhibitor on the 6BS; the cytolosic glutathione reductase and the beta-amylase on the 6DL; and finally the EM4_WHEAT Em protein H5 on the 6AS. This confirms the localization of the α-amylase inhibitor made by Islam et al (2003) and Singh and Skerritt (2001) on the chromosome 6.

A 41kDa and a 48kDa albumins have been assigned to the chromosome 6BS; and a 42kDa albumins and a 38kDa globulin to the 6DS by Singh and Skerritt (2001).

Also gliadins are encoded by genes on the homologous group 6 (Lafiandra et al., 1984)

Group 7:

Concerning the homologous group 7, a serpin was assigned to the 7DS, a lactoylglutathione lyase to the 7DL and a putative glycyl-tRNA synthetase to the 7BS (Merlino et al., 2009).

The genes encoding for the Chloroplastic Copper/Zinc superoxide dismutase (Cu/ZnSOD) were found to be on the long arm of the group 7chromosomes by Wu et al. (1999).

Singh and Skerritt (2001) reported that 4 albumins (46kDa, 45kDa, 39kDa, 19kDa) are encoded by genes on the 7BS, one (56kDa) by the 7A and one 44kDa albumin by the 7D. They have also found 2 globulins to be encoded by genes on the 7DS and one on the 7BS.

Moreover the CM-1 and the CM-2 are encoded by genes on the group 7B and 7D respectively (Garcia-Olmedo et al., 1982)

Table 9: Chromosomal location of wheat proteins

	Protein Name	Chromosomal location	Biological process /molecular function	Reference
Group 1	Triticin	1A, 1B, 1D	Nutrient reservoir activity	Dubcovsky, et al., 1997
	Thioredoxin H-type	1AL, 1BL	Electron transport	Merlino et al., 2009
	Heat Shock protein (HSP70)	1BL	Protein Folding	Merlino et al., 2009
	Alpha-1-purothionin	1BL	Defence response	Merlino et al., 2009
	Wheatwin1	1BL	Defence response	Merlino et al., 2009
	Glutathione-S-transferase	1BS	Response to stress	Merlino et al., 2009
	23kDa globulin	1D	Nutrient reservoir activity	Singh and Skerritt, 2001
	Purothionin (apoprotein)	1AL, 1BL, 1DL	Defence response	Garcia-Olmedo et al., 1982
	LMW-glutenin	1AS, 1BS, 1DS	Nutrient reservoir activity	Payne et al., 1984
	HMW-glutenin	1AL, 1BL, 1DL	Nutrient reservoir activity	Payne et al., 1984
	Glucose phosphate isomerase	1A, 1B, 1D	glycolysis	Cox et al., 1987
	Malate dehydrogenase	1A, 1B, 1D	carbohydrate metabolism	Cox et al., 1987
	ω- γ-Gliadins	1AS, 1BS, 1DS	Nutrient reservoir activity	Payne et al., 1985
Group 2	Alpha-amylase inhibitor	2A, 2B	Response to stress	Singh and Skerritt, 2001
	α-amylase	2D	carbohydrate metabolism	Singh and Skerritt, 2001
	42kDa protein	2D	\	Singh and Skerritt, 2001
	1-cys-peroxiredoxin	2BS	Response to stress	Merlino et al., 2009
	Mitochondrial MnSOD	2	metal ion binding	Gusta et al., 1999
Group 3	Alpha-amylase inhibitor	3BS, 3DS	Response to stress	Singh and Skerritt, 2001
	32kDa protein	3AL	\	Singh and Skerritt, 2001

	20kDa protein	3DS	\	Singh and Skerritt, 2001
	ATP synthase beta subunit	3AS	ATP biosynthetic	Merlino et al., 2009
	GAD1	3BS	Glutamate metabolism	Merlino et al., 2009
	16,9kDa Class I Heat shock protein	3AS, 3BS, 3DS	Protein folding	Merlino et al., 2009
	Small Heat Shock Protein HSP 17,5	3BS	Protein folding	Merlino et al., 2009
	Small Heat Shock Protein HSP 17,8	3DS	Protein folding	Merlino et al., 2009
	Grain esterase	3AL, 3BL, 3DL	\	Ainsworth et al., 1984
Group 4	Alpha-amylase inhibitor	4B, 4D	Response to stress	Islam et al., 2003
	β-amylase	4AL, 4DL	carbohydrate metabolism	Singh and Skerritt, 2001
	Protein disulfide isomerase 2	4BS	electron transport	Merlino et al., 2009
	Serine carboxypeptidase II	4BS	proteolysis	Singh and Skerritt, 2001
	51 kDa, 52 kDa and 56 kDa albumins	4A	\	Singh and Skerritt, 2001
	66 kDa albumin	4BS	\	Singh and Skerritt, 2001
	Globulins	4BL, 4DS	Nutrient reservoir activity	Payne et al., 1985
	CM-3	4A	Response to stress	Garcia-Olmedo et al., 1982
	39 kDa and 50 kDa globulins	4DS	Nutrient reservoir activity	Singh and Skerritt, 2001
Group 5	UDP-glucose pyrophosphorylase	5BL	carbohydrate metabolism	Merlino et al., 2009
	Serpin	5AL	response to stress	Merlino et al., 2009
	Single-stranded nucleic acid binding protein	5BL	translation	Merlino et al., 2009
	0,19 dimeric α-amylase inhibitor	5BL	response to stress	Merlino et al., 2009
	Elongation Factor-1 alpha	5DL	GTP binding	Chikmawati et al., 2009
	Ribosomal protein S29-like protein	5AL, 5DL	translation	Chikmawati et al., 2009
	Lipid transfer protein 7a2b	5DL	lipid transport	Chikmawati et al., 2009
	Thioredoxin M	5BS, 5DL	electron carrier activity	Chikmawati et al., 2009
	Methionine synthase protein	5AS, 5BS	ATP binding	Chikmawati et al., 2009

	38kDa albumin	5DL	\	Singh and Skerritt, 2001
	HSP 70	6BS	Protein folding	Merlino et al., 2009
	Small heat shock protein Hsp 23,5	6BS	Protein folding	Merlino et al., 2009
	Monomeric alpha-amylase inhibitor	6BS	Response to stress	Merlino et al., 2009
	Cytolosic glutathione reductase	6DL	Electron transport	Merlino et al., 2009
Group 6	Beta-amylase	6DL	Carbohydrate metabolism	Merlino et al., 2009
	EM4_WHEAT Em protein H5	6AS	Response to stress	Merlino et al., 2009
	41kDa, 42kDa and 48Kda albumin	6BS	\	Singh and Skerritt, 2001
	38kDa globulin	6DS	\	Singh and Skerritt, 2001
	α- β- Gliadins	6AS, 6BS, 6DS	Nutrient reservoir activity	Lafiandra et al., 1984
	serpin	7DS	Response to stress	Merlino et al., 2009
	Lactoylgluttathione lyase	7DL	Carbohydrate metabolism	Merlino et al., 2009
	Putative glycyl-tRNA synthetase	7BS	Translation	Merlino et al., 2009
	Chloroplastic Copper/Zinc superoxide dismutase	7	metal ion binding	Wu et al., 1999
Group 7	46kDa, 45kDa, 39kDa, 19kDa albumins	7BS	\	Singh and Skerritt, 2001
	56kDa albumin	7A	\	Singh and Skerritt, 2001
	44kDa albumin	7D	\	Singh and Skerritt, 2001
	Globulins	7BS, 7DS	\	Singh and Skerritt, 2001
	CM-1	7B	Response to stress	Garcia-Olmedo et al,, 1982
	CM-2	7D	Response to stress	Garcia-Olmedo et al,, 1982

B. Proteomics

To exploit the genetic information from plant genome projects, it is necessary to have the ability to identify the multitude of polypeptides synthesized as a result of gene expression. This requires the technologies to isolate, fractionate, characterize and identify large numbers of proteins in parallel. Proteomics is an attractive approach to determine gene expression in biological systems because the proteome complement can change markedly as a result of developmental and environmental factors. Although proteomics is a relatively new science, it is based on previously established analytical techniques (Humphery-Smith et al., 1997). The different proteomic steps are illustrated in Figure 9. Originally, the core separation technology of proteomics was two-dimensional gel electrophoresis (2-DE), a system that is well suited to the separation of complex mixtures of proteins. It is significant that this methodology was first developed in relation to the separation of cereal grain proteins (Wrigley, 1968, 1970), with further developments of high-resolution 2-DE by O'Farrell (1975), Klose (1975), and Scheele (1975), demonstrating the enormous potential of this analytical technique for separating thousands of proteins in parallel. In the earlier applications of Wrigley (1968, 1970), it was possible for the first time to determine chromosomal locations, using aneuploid lines, for over 50 gliadin proteins.

Much more recently, 2-DE has been used to establish chromosomal locations for a wider range of genes for wheat grain proteins (Islam et al., 2002, 2003). The 2-DE methods currently used in proteomics offer much greater resolving power than the early methodology. The current state of 2-DE as a separation technology has previously been reviewed by Görg et al. (2000). 2-DE separates proteins using two independent physiochemical parameters. The 'first dimension', known as isoelectric focusing (IEF), separates the proteins in an immobilised pH gradient (IPG). Proteins migrate and stop when they reach their isoelectric point (p*I*). The 'second dimension' separates proteins according to their molecular mass alone, using sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE). (Görg et al., 2000; Herbert et al., 1997; Humphery-Smith et al., 1997).

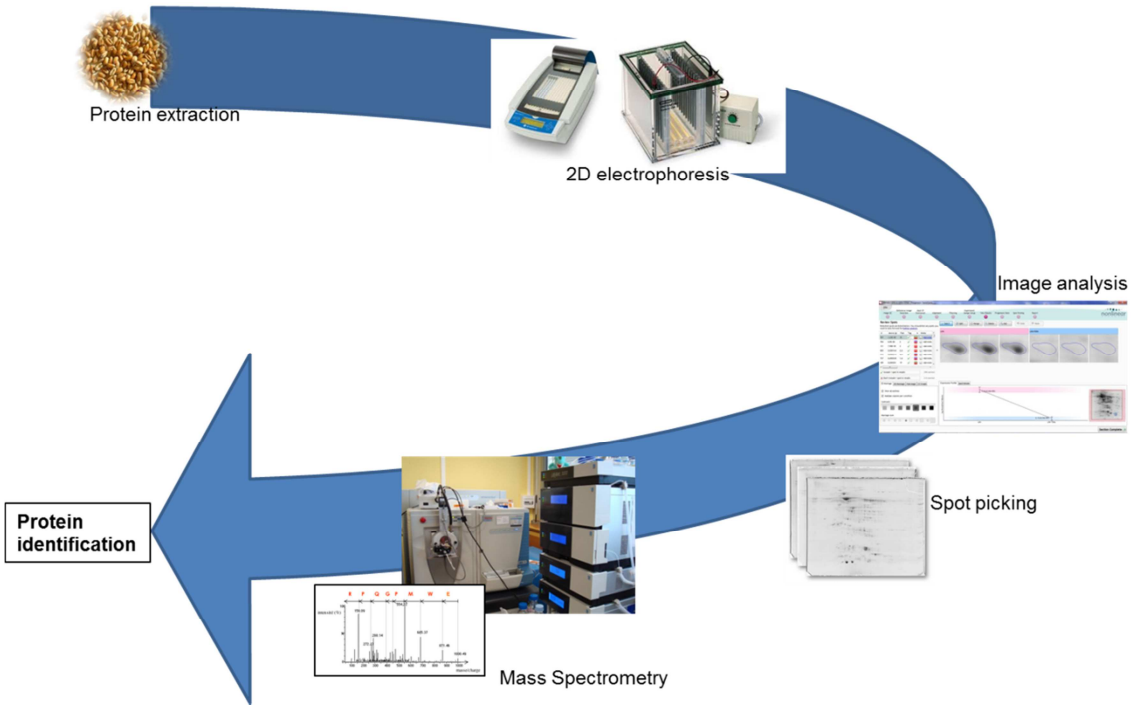


Figure 9: schematic representation of the proteomic study

II. Aim of the work

Among the species of agricultural interest, wheat is one of the most important, but the sequencing of its genome is very complex because its size is 16 billion base pairs, about 5 times the human genome. To handle the scientific challenge of analysing the wheat genome, two international consortia have been established: the IWGSC (International Wheat Genome Sequencing Consortium) and the ETGI (European Triticeae Genome Initiative). The strategy chosen, based on chromosome sorting, makes possible to divide the wheat genome in many libraries, each consisting of only one chromosome or chromosomal arm. The Italian initiative "Physical map of wheat chromosome 5A: Italian initiative for the sequencing of the whole genome", is responsible for the analysis of chromosome 5A and my PhD takes part of this project.

The aim of the project is to identify the polypeptides encoded by the genes on the chromosome 5A of tetraploid and hexaploid wheats. The knowledge of polypeptides encoded by genes at chromosome 5 will help in the understanding of the correlations between their presence and specific physiological characteristics, along with quality properties.

In order to identify these polypeptides, the study was performed with a proteomic approach based on 2D-Electrophoresis and Mass Spectrometry analyses.The first target was the identification of the 5A proteins of the durum wheat cv Langdon, by using the interspecific chromosome substitution lines with *T. dicoccoides* and this allowed also to characterize 5A encoded polypeptides of this latter. This study was done both on the metabolic fraction and the CM-like fraction. This work was also performed for the gliadin and the glutenin fractions, in order to checkif there are additional loci, besides those already known (present on chromosomes 1 and 6). The second objective was to identify 5A encoded polypeptides in bread wheats. The study was conducted in the same way, i.e. by means of the use of intervarietal 5A chromosome substitution lines, specifically five lines in which the 5A chromosome of the bread wheat cultivar Chinese Spring has been replaced by each of 5A chromosomes of cultivars Hope, Thatcher, Timstein and Cheyenne.

Besides the wheat kernel proteome, we were also interested in the leaf proteome, and more particularly in the nuclear proteome. Because the nuclear protein fraction is present in low quantity and thus it is difficult to have sufficient quantity to realize 2D electrophoresis, for this part the study we used 1D electrophoresis.

III. Materials and methods

A. Plant Material

For tetraploid wheat, the substitution line *T. turgidum* subsp. *durum* cv. Langdon-*T. dicoccoides* 5A or 5B (in which the pair of either chromosome 5A or 5B of Langdon is replaced by the pair of either chromosome 5A or 5B of *T. dicoccoides*) has been used, along with the parental lines (cv. Langdon, and *T. dicoccoides).*

For bread wheat, different intervarietal chromosome substitution lines have been used (Table 10), in which chromosome pair 5 (A, B, or D) of the bread wheat cultivar Chinese Spring (the recipient variety) has been replaced by the chromosome pair 5 (A, B or D) of the bread wheat cultivars Cheyenne, Hope, Thatcher, or Timstein (fig.10).

Because growing conditions affect protein expression, seeds of the parental lines along with their chromosomes 5 substitution lines, were first germinated at 4°C for 15 days, then were allowed to reach maturation in a climate chamber at ~20- 25 ° C until spike maturation.

In the case of the analysis of seed proteins, plants were grown until maturity, whereas, the analysis of nuclear proteins was performed on leaves collected three weeks after germination, when plants have around six leaves (Figure 11).

This latter analysis was performed on tetraploid wheat only.

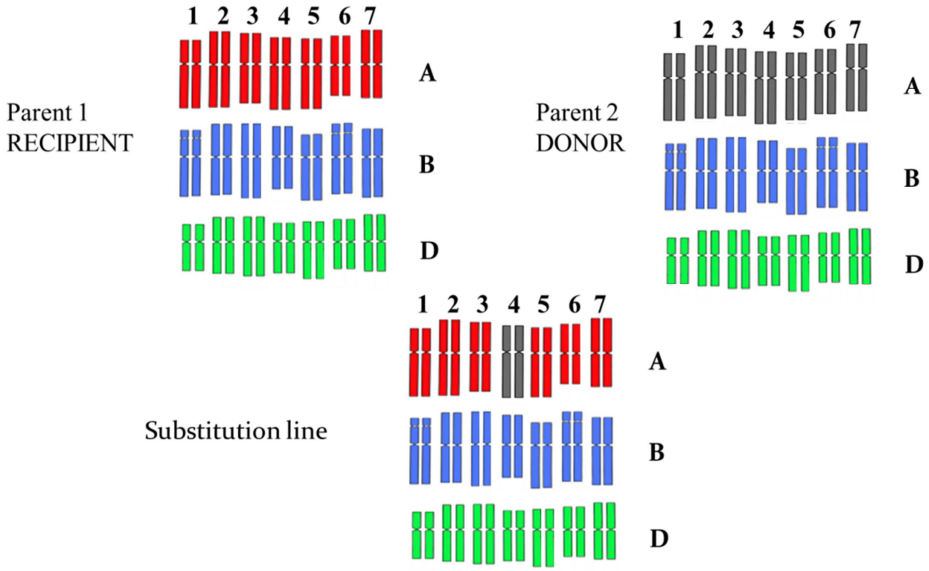


Figure 10: Schematic representation of the main steps used for obtaining a chromosome substitution line

Table 10: Substitution lines used in this study

	Name of the Substitution line	abbreviation
Hexaploid wheat	Chinese Spring (Hope5A); Chinese Spring (Hope5B);Chinese Spring (Hope5D)	CS-Hope5A; CS-Hope5B; CS-Hope5D
	Chinese Spring (Timstein5A); Chinese Spring (Timstein5B);Chinese Spring (Timstein5D)	CS-Timstein5A; CS-Timstein5B; CS-Timstein5D
	Chinese Spring (Thatcher5A); Chinese Spring (Thatcher5B);Chinese Spring (Thatcher5D)	CS-Thatcher5A; CS-Thatcher5B; CS-Thatcher5D
	Cheyenne (Chinese Spring5A);Cheyenne (Chinese Spring5B);Cheyenne (Chinese Spring5D)	CNN-CS5A; CNN-CS5B; CNN-CS5D
Tetraploid wheat	Langdon (Dicoccoides5A)	LDN-TD5A
	Langdon (Dicoccoides5B)	LDN-TD5B

Figure.11: Tetraploid wheat plants collected for nuclear proteome analysis

B. Methods

The ideal situation for performing proteomic comparisons would be to extract and separate total proteins, but this is not always possible, either for the enormous number of polypeptides that would be present, and for the biochemical characteristics of some polypeptides that hamper the possibility to extract or separate them by IEF. For this reason, it is more appropriate to perform pre-fractionation procedures that allow to obtain fractions enriched in specific protein classes.

1. Metabolic and chloroform-methanol (CM) fractions

a) Protein extraction

Because gluten proteins, that are the most abundant protein class in wheat endosperm, are mostly coded by chromosomes 1 and 6, we have first analysed the soluble protein fractions, in which it is presumed that polypeptides coded by genes at all chromosomes are present.

Proteins are extracted following a sequential extraction protocol. Wheat flour proteins are separated into gluten, metabolic, and chloroform/methanol-soluble proteins fractions based on solubility in KCl and methanol. (Hurkman et al., 2004).

Metabolic fraction:

First, about 30 seeds are milled to obtain 300 mg of flour. Then the flour is suspended in 1,18 mL of cold KCl-buffer (50 mM Tris-HCl, 100 mM KCL, 5 mM EDTA, pH 7,8) with 24 µL of protease inhibitor cocktail (ref 11 873 580 001, ROCHE). The suspension is incubated 1 hour on ice by vortexing every 15 minutes. The suspension is centrifuged at 8500 rpm for 25 minutes at 4°C.The supernatant or KCl-soluble fraction is collected and 5 volumes of 0,1M ammonium acetate in methanol are added, and incubated overnight at -20°C. Then the extract is centrifuged at 8500 rpm for 15 minutes at 4°C. The pellet, containing the metabolic fraction, is washed two times with cold acetone, and then dried at room temperature.

CM fraction:

Five volumes of cold acetone are added to the supernatant previously obtained, and incubated overnight at -20°C. Then it is centrifuged at 8500 rpm for 15 minutes at 4°C, and the pellet, containing the chloroform-methanol fraction, is washed two times with cold acetone, and dried at room temperature (Figure 12).

Finally the pellets containing the protein fractions of interest are resuspended in a rehydratation buffer (7M urea, 2M thiourea, 2% CHAPS, 2% Triton X-100).

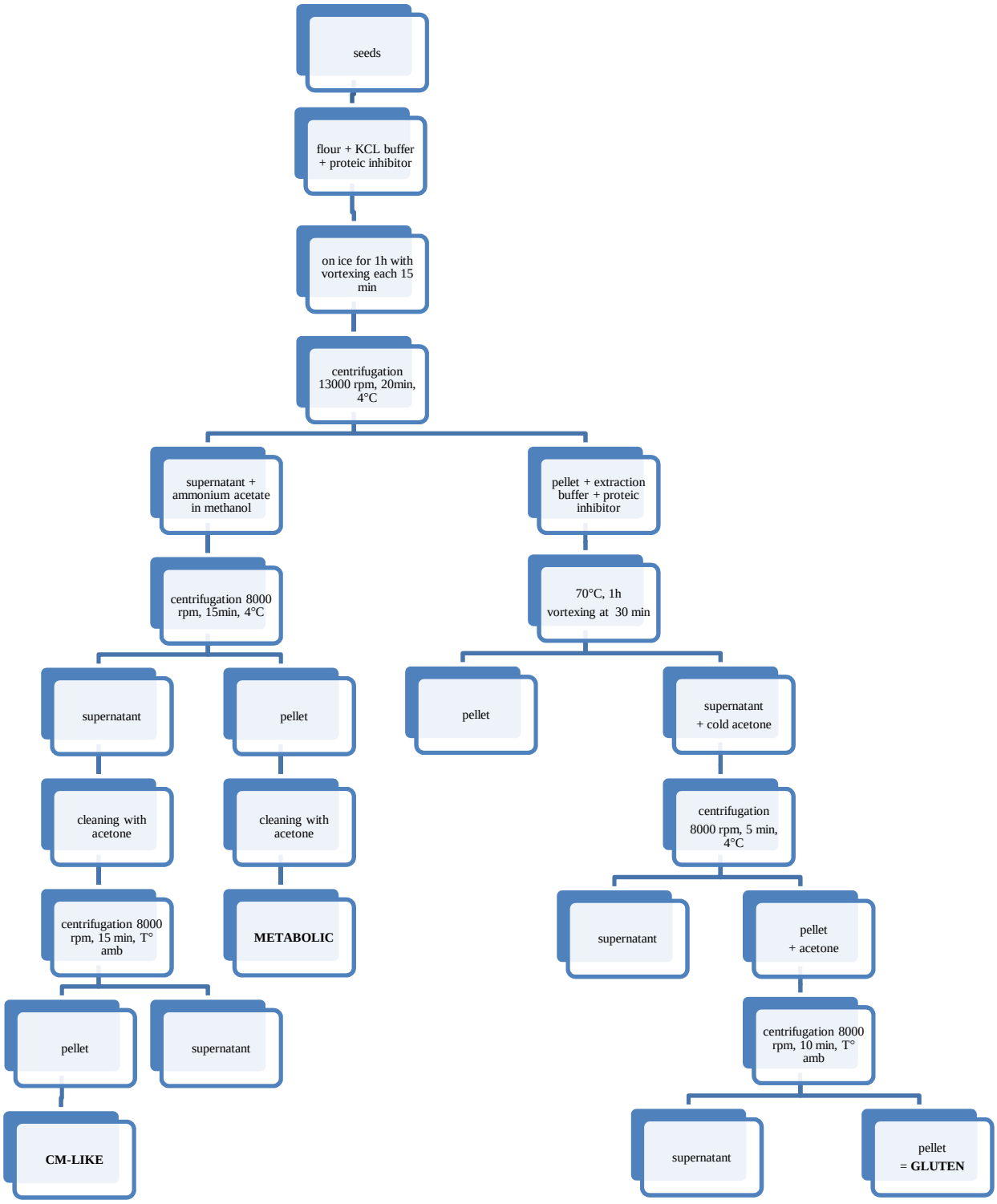


Figure 12: Schematic representation of the extraction protocol for obtaining the different wheat protein fractions

b) Quantification

2-D Quant Kit Assay (Ge Healthcare) is based on the specific binding of copper ions to protein. Precipitated proteins are resuspended in a copper-containing solution and unbound copper is measured with a colorimetric agent. The color density is inversely related to the protein concentration. The assay has a linear response to protein in the range of 0-50 µg. This Kit was used to have an accurate estimation of the protein extractions and thus be able to use always the same protein concentration prior the IEF. Six dilutions (0-50 µg) of BSA provided in the kit (2mg/mL) were used as standards. Five and ten µL of samples (containing the protein extractions dissolved in RB) have been assayed in each test.

c) Two dimensional electrophoresis

Eighteen centimetres immobiline pH gel strips (pH 3-10) are rehydrated for 12 hours, overnight, with 340 µL of rehydratation buffer added with 300 µg of proteins and 0,5% IPG buffer and 1,2% of Destreak. Then the strips are focused for a total of 95000 Volts (200V for 4h; 500V for 2h30; 1000V for 3h; 5000V for 2h; 8000V for 30 minutes in gradient step; 8000V for 9h30) using the IPGphor^{TM} Isoelectric Focusing System (Ge Healthcare).

After isoelectrofocalisation, strips are equilibrated for 15 minutes three times, first in the equilibration buffer (6M urea, 50mM Tris, 30% glycerol, 2% SDS) added with 1% DTT , then in the equilibration buffer added with 2,5% iodoacetamide, and finally with the equilibration buffer alone.

The strips are deposited on 20 centimetres SDS-PAGE gels (T=12%; C=2,67%), cast in the PROTEAN Plus Multi-Casting Chamber (Biorad). The gels are focused, in the PROTEAN Plus Dodeca Cell (Biorad), at 20 mA/gel for the first 20 minutes, then 40 mA/gel and the focusing is stopped 15 minutes after the dye front output.

Three technical replicates are made for each line studied. In this case, biological replicas were not performed, since we searched only for protein bands that were either present or absent, and not modulated in their expression levels.

d) Gel staining and acquisition

The 2DE gels were stained in a solution containing 100 gr/L of ammonium solphate, 1 gr/L of Coomassie G250, 11% of phosphoric acid and 20% methanol overnight and destained three times for 1 h with distilled water before image acquisition.
All the gels were scanned with EPSON Perfection V750 pro at 16 Bit and 300 dpi resolution in grayscale.

e) Image analysis

All the gels, three replicates for each studied line, were analysed with the software Progenesis Samespots ver 4.5 (Nonlinear Dynamics, UK) to detect the spots. After an alignment step in which gels are superimposed, the software detects all the spots present on the gel. After the software has performed spot measurement and background subtraction, the normalised volume of each spot in each replicate is compared to determine differences of expression. The spots are considered to be present/absent when the *p*-value is < 0.05 (generated by the application of analysis of variance ANOVA), the *q*-value≤ 0.05(False Discovery Rate), and the power ≥ 0.8, and by direct observation.

For the tetraploid wheat study, it was used a “master gel” made from a 50/50 mixture of *T. turgidum* ssp *durum* cv Langdon and *T. dicoccoides* in order to count spots that were specific to each of the two parents (Merlino et al., 2009). Master gel analysis allows to identify all the protein spots that are in common between the two genotypes, that thus cannot be attributed to any chromosome, whereas differential spots allow to perform chromosome localization by means of the appropriate genetic lines. This experiment was repeated two times in order to confirm the results.

In order to detect 5A encoded polypeptides in either the parental genotypes, the following scheme was applied (tab.11):

Table 11: Workflow for the image analysis done with Progenesis Samespots

Tetraploid wheat	*T. durum* cv Langdon vs *Triticum dicoccoides*	Identification of common spots[a]
	Triticum dicoccoides *vs* Langdon-*T.dicoccoides*5A	Identification of TD 5A spots
	T. durum cv Langdon vs Langdon-*T.dicoccoides*5A	Identification of TD 5A and LDN 5A spots
Bread wheat	*T.aestivum* cv Chinese Spring vs *T.aestivum* cv Hope	Identification of common spots[a]
	T.aestivum cv Hope *vs* CS/Hope5A	Identification of Hope 5A spots
	T.aestivum cv Chinese Spring *vs* CS/Hope5A	Identification of CS 5A and Hope 5A spots
	T.aestivum cv Chinese Spring vs *T.aestivum* cv Timstein	Identification of common spots[a]
	T.aestivum cv Timstein *vs* CS/Timstein5A	Identification of Timstein 5A spots
	T.aestivum cv Chinese Spring *vs* CS/Timstein5A	Identification of CS 5A and Timstein 5A spots
	T.aestivum cv Chinese Spring vs *T.aestivum* cv Thatcher	Identification of common spots[a]
	T.aestivum cv Thatcher *vs* CS/Thatcher5A	Identification of Thatcher 5A spots
	T.aestivum cv Chinese Spring *vs* CS/Thatcher5A	Identification of CS 5A and Thatcher 5A spots
	T.aestivum cv Chinese Spring vs *T.aestivum* cv Cheyenne	Identification of common spots[a]
	T.aestivum cv Cheyenne *vs* Cheyenne/CS5A	Identification of Cheyenne 5A spots
	T.aestivum cv Chinese Spring *vs* Cheyenne/CS5A	Identification of CS 5A and Cheyenne 5A spots

[a]: spots in common between the two parents are not taken into consideration for the analysis, since it is not possible to attribute them to any chromosome

f) Mass spectrometry

Mass spectrometry analyses were conducted at the IBiSA platform "Biopolymers-Structural Biology" located at the INRA Center of Angers-Nantes, during different stages I performed during my PhD activity (http://www.angers-nantes.inra.fr/plateformes_et_plateaux_techniques/plateforme_bibs).

Before being submitted to mass spectrometry, the spots of interest are picked and hydrolysed.

The spots are first washed to remove contaminants associated with coloring step. Then the disulfide bridges of proteins are broken by the use of 1,4-Dithiothreitol (DTT), and to avoid re-matching of the cysteine, proteins are alkylated with iodoacetamide. Subsequent steps of washing permit to remove the chemical reagents used in the steps of reduction / alkylation. Following these steps the proteins are then hydrolyzed with trypsin (which cleaves peptide chains mainly at the carboxyl side of the amino acids lysine or arginine, except when either is followed by proline) at a concentration of 15 ng/µL in a 25 mM NH_4HCO_3 buffer. Then 1 µL of trypsin is added in each eppendorf, which are incubated at 37 ° C overnight. The next day the reaction is stopped by adding 10 µL of 10% formic acid (Larré et al., 2010).

Nanoscale capillary liquid chromatography-tandem mass spectrometry (LC-MS/MS) analyses of the hydrolyzed proteins were performed using a U3000 RSLC system (Dionex, Amsterdam, the Netherlands) coupled to a LTQ-Orbitrap VELOS mass spectrometer (Thermo Scientific, San José, USA). Chromatographic separation was conducted on a reverse-phase capillary column (Acclaim Pepmap 100, C18, 3 µm particle size, 100 Å, 75 µm i.d., 15 cm length, Dionex) with a linear gradient from 3% to 40% of acetonitrile in 25 min, followed by an increase to 65% of acetonitrile within 5 min. Mass data acquisitions were piloted by the X-Calibur software (Thermo Scientific) using a typical parallelised "top 5 CID" experiment: MS data were recorded at 30,000 resolution in the Orbitrap analyzer, whilst the five most intense ions (with the exclusion of singly charged ions) were selected and fragmented in the LTQ ion trap (MS/MS measurements).

Proteins were identified by comparing the collected LC-MS/MS data with the Uniprot sequence databank, restricted to the *Viridiplantae* taxonomy (http://www.uniprot.org/) and with the Institute of Genomic Research (TIGR) expressed

sequence tag (EST) databank. Databank searches were performed using the Mascot server 2.2 program (Matrix Science), which is a powerful search engine that uses mass spectrometry data to identify proteins. First Mascot looks for the best peptide sequence match to each MS/MS spectrum in databank and then groups these peptide matches into protein matches. The mass tolerance was set at ±5ppm for parent ions (MS mode) and ±0.5 Da for fragment ions (MS/MS mode); one missed cut per peptide was allowed, and Carbamidomethylation of cysteine was specified as a static modification whereas the oxidation of Methionine was set as a variable modification.

Protein identification was based on a minimum of two peptides matching the databank sequence, with individual MASCOT ion scores above the significance threshold (threshold score of 33, $p<0.05$).

For the identfication, the results of the two databanks have been compared, and proteins have been selected following various criteria: Mascot score, peptides numbers, emPAI, sequence coverage and good correlation between MW/ pI theoric and MW/pI experimental. The Exponentially Modified Protein Abundance Index (emPAI) offers approximate, label-free, relative quantitation of the proteins in a mixture based on protein coverage by the peptide matches in a database search result, and can be used to represent the contribution of each protein in the spot.Moreover to verify that the identfication is the best possible for some set of peptides a supplementary blast on ncbi was realized.

The gene ontology (biological and molecular function) of each protein was search on the database of Uniprot and in literature.

Finally, information on the chromosomal assignement of the identified proteins was controlled in the GrainGenes EST database.

2. Glutenin and gliadin fractions

a) Protein extraction and quantification

For the gliadins, 100 mg of flour is suspended in 5 volumes of extraction buffer (1,04 mL of DimethylFormamide in 10 mL H_2O) and incubated for 1 hour. Then the solution is centrifuged for 15-20 minutes at 13000 rpm. The supernatant contains the gliadins. Proteins are precipitated in cold acetone at -20°C overnight, then resuspended in Acetonitrile 25%/ Trifluoroacetic acid 0,025%. After a centrifugation at 13000 rpm for 10 minutes, the protein concentration is calculated by spectrophotometer analysis (λ= 280 nm e ε=0.75). Finally, proteins are aliquoted in 30-40 µg fractions and vacuum-dried.

For the glutenins, 100 mg of flour is washed under agitation with 1 mL of propanol-50% for 30 minutes, centrifuged 15 minutes at 13000 rpm, and the supernatant is discarded. This step is repeated three times. Then the pellet is suspended in 10 volume of extraction buffer (50% propanol, 50 mM Tris pH 8.8, 1% DTT, 1 mM EDTA 10 mM Iodoacetamide) and incubated at 65°C for 1 hour. Finally, a centrifugation at 13000 rpm at 4°C for 15 minutes is performed. Four volume of cold acetone is added to the supernatant and incubated overnight at -20°C. After a centrifugation for 5 minutes at 13000 rpm at 4°C, the supernatant is discarded and the pellet is resuspended in rehydratation buffer (7M urea, 2M thiourea, 2% CHAPS, 2% Triton X-100).

b) A-PAGE/SDS-PAGE for the gliadin fraction

Since gliadins do not present a wide range of isoelectric points, IEF is not the method of choice for their separation. For this reason, we performed 2D electrophoresis, by using Acid PAGE (A-PAGE) as first dimension.

- A-PAGE

Gliadins are resuspended in 20-30 µl of sample buffer (5 mg DTT, 360 mg Urea, 1,4µL acetic acid and methyl violet as a tracking dye for a final volume of 1 mL).
The first-dimension electrophoresis was carried out in acidic solution with a Biorad apparatus for polyacrylamide gels (0.75 mm, T=6 and C=2.67) containing 4 M urea and 0.75% (v/v) acetic acid. A pre-run with normal polarity at constant current (27 mA/gel) was performed with 0.75% (v/v) acetic acid for 50 minutes. Then after one hour of

agitation, the gliadins are loaded on A-PAGE gel. The upper chamber contained 0.14% (v/v) acetic acid and the lower contained 0.25% acetic acid. The run was carried out at constant current (27 mA/gel) with reversed polarity (upper electrode positive), and was stopped 5 min after the dye reached the bottom of the gel.

- SDS-PAGE

The second-dimension electrophoresis was performed in the same apparatus. The main gel was T=12 and C=2. The stacking gel (T=3.75, C=2.67, 0.125 M Tris-HCl, pH 6.8, 0.1% SDS) was poured up to about 0.5 cm from the top of the glass plates. Bands of A-PAGE gel are cut and equilibrated in 15 mL of equilibration buffer (Tris 0,5 M, Glycerol 50%, SDS 10%, DTT 0,5%). After, bands are deposited on SDS-PAGE gels (T=12, C=2) and focused at 40 mA/gel up to 50 minutes after the exit of the front.

3. Nuclear proteome

a) Isolation of leaf nuclei, protein extraction and quantification

This was performed by using the Plant Nuclei Isolation/Extraction Kit **(**SIGMA CelLytic™PN)

- **Preparation of the cell lysate**

Eighteen grams of fresh leaves are crushed with a razor blade with the NIBA solution (Nuclei Isolation Buffer (NIB) + DTT 1M + protease inhibitor cocktail), and incubated for 10 minutes. Then the suspension is passed twice through the 20 µm filter mesh, and the filtrate is aliquoted into an even number of eppendorfs .The eppendorfs are centrifuged for 10 minutes at 13000 rpm at 4°C. The pellets are resuspended in the NIB buffer 1X (the supernatant is kept as a first control). 0,6%(v/v) of Triton X-100 is added to lyse the cell membrane, and incubated for 10 minutes on ice.

- **Isolation of nuclei and obtainment of the nuclear suspension**

500 µL of the lysate obtained is deposited on the top of eppendorfs containing a 0.8 mL sucrose 1.9 M cushion, which are centrifuged at 13000 rpm for 20 minutes.

The upper phase is kept as second control and the pellets are washed by re-suspending each of them in 400-500 µL of NIBA solution. A centrifugation for 5 min at

13000 rpm at 4°C is performed. Finally the nuclei pellets are pooled and re-suspended in Nuclei PURE Storage Buffer (Kit SIGMA) for a final volume of 200 µL.

The nuclei integrity is checked under fluorescence microscopy with 10 µL of nuclei suspension stained with propidium iodide.

- **Protein extraction from the nuclear suspension**

The nuclei suspension is centrifuged at 13000 rpm for 10 minutes. The pellet is dissolved with 120 µL of Extraction buffer 1M NaCl, and sonicated (2x2minutes; 3minutes; 2minutes). A centrifugation for 5 minutes at 13000 rpm is performed and the supernatant (protein extract) is kept and quantified by Bradford assay.

b) One-dimensional electrophoresis

18 µl of the protein suspension are deposited on NuPAGE® Novex® Bis-Tris Gels (SDS-PAGE) with MES SDS Running Buffer, and migrated at 200 V for around 50 minutes. Then the gels are stained with Blue Coomassie as described above.

Each nuclear extract was loaded in three lanes on six gels (ie 18 lanes in total for each extract) to avoid deformation due to the migration and in order to have significant results.

c) Image analysis

The one-dimensional gels were analysed, with LabImage (KAPELAN Bio-Imaging Solutions). LabImage detects lanes and bands on the gel, and thus allows comparing 1D profile in order to detect possible differences between various wheat lines.

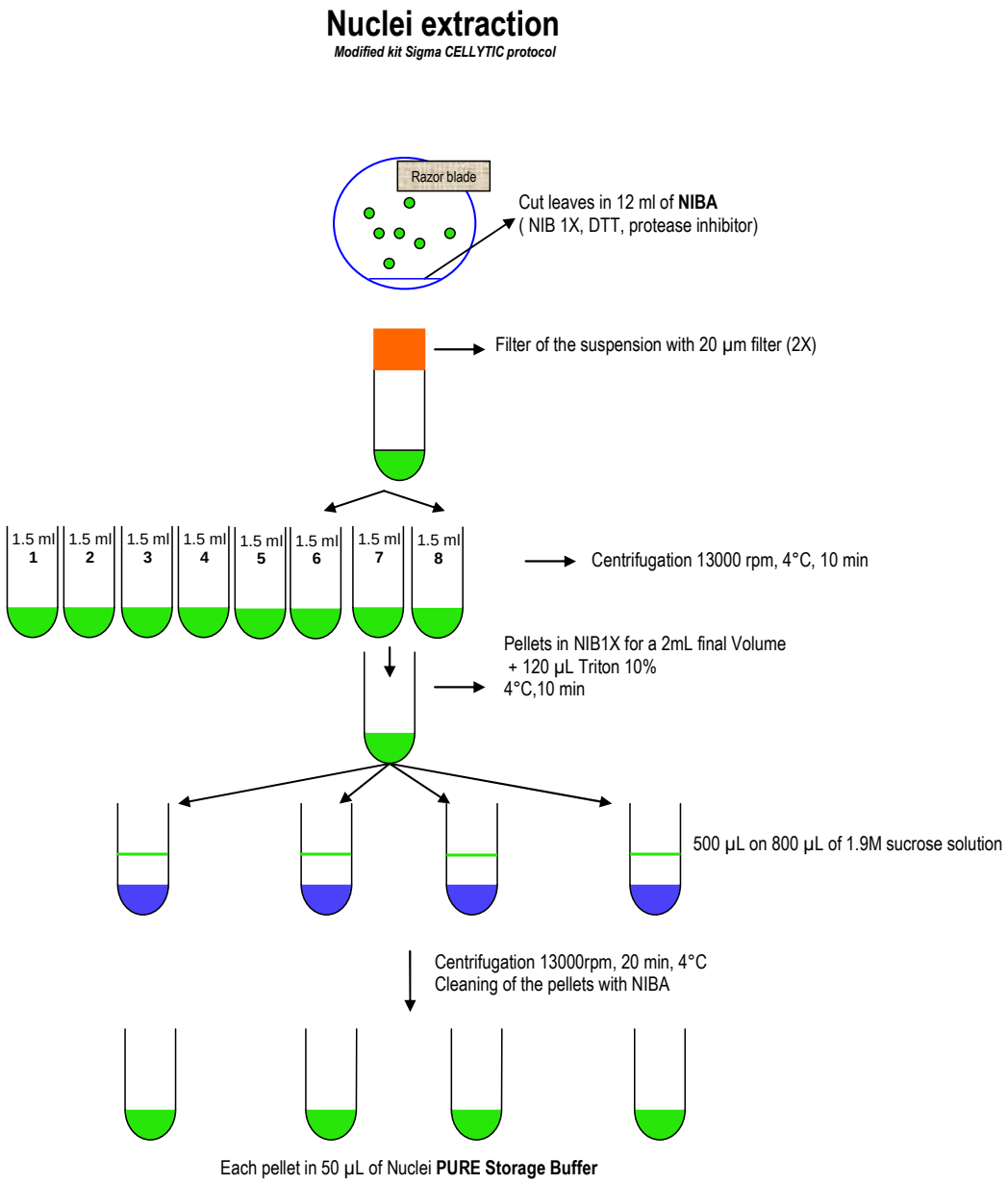


Figure 13: Scheme of the Isolation of leaf nuclei and their protein extraction

IV. Results and discussion

A. Tetraploid wheat

1. Set-up of the extraction procedure

Because chromosomes 5 encoded polypeptides should belong mainly to the soluble metabolic protein fraction, which are prevalently present in the outer layers of the wheat kernel, we first compared metabolic proteins extracted from the wheat flour, the sieved flour, and the remaining outer layers after sieving the flour of *Triticum dicoccoides*.

The Principal Components Analysis (PCA) performed by Progenesis Samespots revealed three distinct groups (fig.14) which confirmed that each extract type was different from the others. Moreover, image analysis revealed, by comparing all the spots volume of each extract, that sieved flour showed the best pattern. Since we expected that the outer layers were those containing most of the polypeptides, we suppose that the presence of polyphenols in the bran could affect protein solubilisation.

On the basis of these results, only the sieved flours were used for the analyses, by separating the different protein fractions on the basis of their solubility.

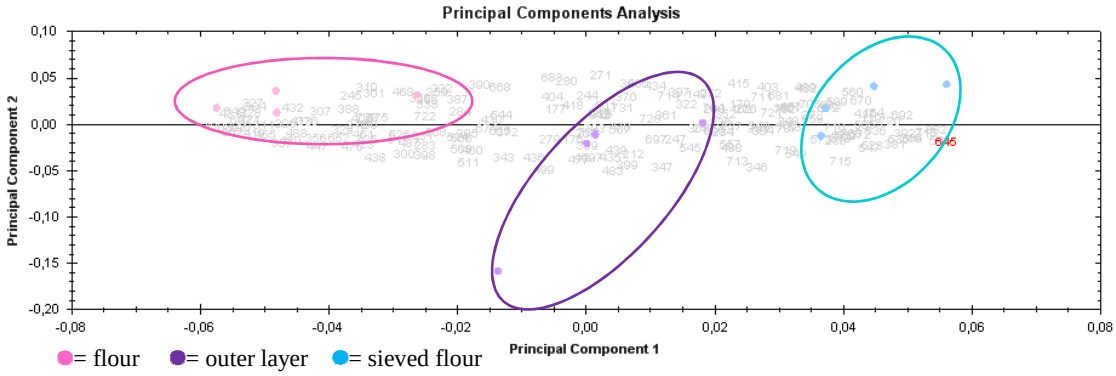
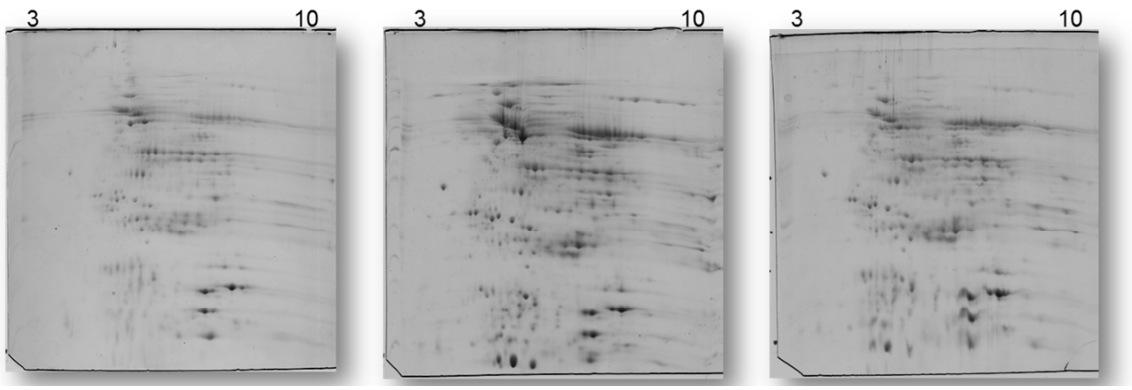


Figure 14: Principal Components Analysis of the metabolic fraction of wheat flour, the sieved flour and kernel outer layers of *T.dicoccoides*.



Wheat flour Sieved flour Outer layer

Figure 15: 2D gels of the wheat flour, the sieved flour and the outer layer of *T. dicoccoides*

2. Metabolic fraction

This analysis allowed to identify 453 spots in the master gel, 127 spots (28%) are specific of *T. dicoccoides* and 97 (21%) are specifics of durum wheat cultivar Langdon. The remaining spots were not assigned to either *T. dicoccoides* or *T. durum*.

The comparisons Master/LDN-*TD*5A, LDN/LDN-*TD5A* and *TD*/LDN-*TD5A*, confirmed that 26 spots were encoded by genes on the chromosome 5A of *T. dicoccoides*, whereas 10 spots were confirmed for Langdon (fig. 16 & 17). All the spots have a $p<0.05$, a q-value <0.05, a fold change $\geq 1,5$ and a power ≥ 0.8.

The mass spectrometry led to the identification of 20 proteins for *T.dicoccoides* and 8 proteins for *T. turgidum* ssp. *durum* cv Langdon. These proteins are involved in different biological processes and have various molecular functions (Figures 18 & 19 and subparagraphs below). In a few cases, two proteins were identified in the same spot.

5A polypeptides identified in *T.turgidum* ssp. *durum* cv Langdon:

Relatively to *T. durum*, gene ontology search was performed for each putative 5A polypeptide. Spot 1 is represented by Chitinase that is involved, among others, in the defense against chitin-containing phytopathogenic fungal infection. The acidic and basic isoforms of chitinases are induced in plants in response to pathogen attack, other environmental stimuli, but may be also expressed in plant tissues during normal development (Singh et al., 2007).

Spots 2 and 3 are represented by Putative Late Embryogenesis abundant proteins (LEA) that are known to be stress related proteins. LEA proteins are produced in abundance during seed development, comprising up to 4% of cellular proteins and even if their precise function is unknown, they are assumed to protect cellular or molecular structures from the damaging effects of water loss (Goyal et al, 2005).

Although the specific role of Globulin 3 (spot 3) is not known, globulins are known to be the principal storage proteins in the seeds of dicot species, and to be involved in the food and respiratory allergy to wheat.

Triticin (spot 4) are located in protein bodies in the starchy endosperm of the wheat grain. The triticins are disulfide-linked, hetero-tetramers made up of four subunits which have been designated as D (Mr 58,000 Da), d (Mr 22,000 Da), A (Mr 52,000 Da), and a (Mr 23,000 Da). The molecular weights of the native molecules are lower than other polymeric storage proteins. Triticins characterized so far are encoded by genes on the short arms of chromosome 1A (*Tri-A1* locus) and 1D (*Tri-D1* locus). Recently, triticin genes have also been observed on chromosome 1B (*Tri-B1* locus) but no expressed protein has been found, suggesting that this locus is not active (Dubcovsky et al., 1997). When considering their solubility properties, they behave as globulins. No important link with breadmaking quality has been established (Gianibelli et al., 2001).

Spot 6 is represented by alcohol dehydrogenase, which is an enzyme presumably required by plants for NADH metabolism, via reduction of acetaldehyde to ethanol, during periods of anaerobic stress. Genetic studies of the alcohol dehydrogenase of hexaploid wheat indicate that three structural genes, one located in each genome, encode three subunits which associate in all possible dimeric combinations. The available evidence indicates that in tetraploid wheat two alcohol dehydrogenase structural genes encode two subunits, designated α and β, which associate to form three isozymes, ADH-1 (αα subunits), ADH-2 (α/β), and ADH-3 (ββ) (Langston et al., 1980). The α-amylase inhibitor CM3 was also found in the spot 6, and this protein is described more in details in the paragraph IV.A.2.

Spot 7 is represented by glucose and ribitol dehydrogenase, which catalyses the oxidation of D-glucose without prior phosphorylation to D-β-gluconolactone using NAD or NADP as a coenzyme.

Xylanase inhibitor III is present in the spots 8 & 9. XIP-type inhibitors are glycosylated monomeric proteins with a molecular mass of 29 kDa and pI values of

8.7–8.9 and inhibit microbial xylanases from glycoside hydrolase families 10 and 11 (Elliott et al., 2009). In cereals, endogenous xylanase action is required for controlled remodeling of the plant cell wall during growth and development. In addition, cell wall breakdown during germination ensures that other hydrolytic enzymes, secreted from the aleurone or scutellum, can reach starch and protein (Simpson et al., 2003) and moreover xylanases are routinely used in bread making, wheat gluten–starch separation and as supplements in animal feed production. For these reasons XIP have an important role on quality properties.

By looking at the molecular function of this 8 proteins identified, the major part have a role in the binding (28%) and as nutrient reservoir (18%) (Fig.18). These results are consistent with those of Vitulo et al (2011), who have find molecular function of binding and DNA binding for 34% of the 2772 unigenes identified.

5A polypeptides identified in *T.dicoccoides*:

Spot 1 is represented by the tubulin β-3 chain. Tubulins are components of the cytoskeleton, both in microtubules and microfilaments, and play an important role in spindle formation and chromosome separation during cell division (Barcaccia et al., 2001).

Spot 2 was identified as alcohol dehydrogenase ADH1A, which was described previously.

Spot 4 is represented by the Glyceraldehyde-3-phosphate dehydrogenase (GADPH) which is involved in the glycolysis and glyconeogenesis. Secondary functions have been suggested: for example, Bustos and Iglesias (2003) reported that wheat endosperm and GAPDH undergoes posttranslational phosphorylation enabling interaction with 14-3-3 family proteins, thus exerting a regulation aimed at maintaining the levels of energy and reductants in the cytoplasm. Recently, using a proteomic approach it was established that GAPDH activity in *Arabidopsis* was inhibited by H_2O_2, suggesting that GAPDH is a direct target of H_2O_2 and might have a role in mediating ROS signaling in plants (Hancock et al, 2005).

Also the malate dehydrogenase is present in spots 4 and 5. It catalyzes the formation of oxaloacetate from malate in mitochondrial matrix, which leads via transamination to the formation of aspartate, a precursor of methionine biosynthesis.

Spot 6 was identified as Isoflavone reductase-like protein, which is an important enzyme in phenylpropanoid metabolism. This protein was recently described as allergen by Yang et al (2011).

Spot 7 was found as globulin 3, which was described before (spot 4 Langdon).

Spot 8 and 9 correspond to Pathogenesis-related 1b protein. The pathogenesis related (PR) proteins are low-molecular proteins (6-43 kDa), extractable and stable at low pH (< 3), thermostable, and highly resistant to proteases. They have dual cellular localisation – vacuolar and apoplastic, the apoplast being the main site of their accumulation, and have an important role in plant defense against pathogenic constraints and in general adaptation to stressful environment. Moreover it was recently found that among the identified plant allergens, 23% belong to the group of PRs (Aglika, 2005).

Spot 13 is represented by the Triosephosphate isomerase. Triosephosphate isomerase is an ubiquitous and highly conservative dimeric enzyme, consisting of subunits of Mr 26,000-27,000 that catalyzes the conversion of dihydroxyacetone phosphate (DHAP) and glyceraldehyde 3-phosphate (GAP), an important step in the glycolytic pathway (Cui and Karplus, 2001).

Spot 14 was identified as Manganese Superoxide dismutase (MnSOD) which is mainly located in the mitochondria. SODs are a group of metalloenzymes that protect cells from superoxide radicals by catalyzing the dismutation of the superoxide radical to molecular O_2 and H_2O_2 (Wu et al., 1999).

Spot 15 is represented by a mix of two proteins, a 27K protein and a Cold shock domain protein 3. This latter belongs into CSD (cold shock domain) proteins and it is implicated in regulation of transcription as RNA chaperone (Vıtamvas et al., 2012). Concerning the 27K protein, there are few informations. However, Kimoto et al (2009) have found and identified a “Tri a bd 27k protein” as an N-linked glycoprotein with mannose and fucose residues, which is a major wheat allergen.

Spot 20 was identified as Single-stranded nucleic acid binding (SSb) protein. The SSB protein binds to single-stranded regions of DNA in order to prevent premature annealing, to protect the single-stranded DNA from being digest by nucleases, and also to remove the DNA secondary structure allowing others enzymes to function effectively on it. The SSB protein is produced during all the steps of the DNA metabolism.This protein was found on the 5BL (Merlino et al., 2009).

Spot 21 is composed of two proteins, the globulin-like protein (describe previously) and the Pre-mRNA processing factor. The activities of several mRNA processing factors are coupled to transcription through binding to RNA polymerase II (Komarnitsky et al., 2000).

Spot 22 to 26 are represented by the α-amylase inhibitor in different aggregation state (monomeric, dimeric). It was already described more in details in the paragraph IV.A.2.

By looking at the molecular function of these proteins, also for *T. dicoccoides* the major part have a role in binding (57%). Among the proteins involved in binding, there are: the isoflavone reductase-like protein (nucleotide binding), the MnSOD (ion binding), the alcohol dehydrogenase ADH1A (nucleotide and ion binding), the Cold Shock domain protein 3 (DNA and ion binding), the single-stranded nucleic acid binding protein (nucleic acid and nucleotide binding) and the Pre-mRNA processing factor (nucleic acid and nucleotide binding).

Among these proteins, three are common between the two wheat lines: the globulin 3, the alcohol dehydrogenase and the α-amylase inhibitor CM3. The alcohol dehydrogenase was found to be encoded by genes on the 5A chromosome after interrogation of the GrainGenes database (http://wheat.pw.usda.gov/cgi-bin/westsql/map_locus.cgi). The globulin 3 was found associated to the 5B and the 5D chromosomes and we can assume that also genes encoding for this protein are present on the 5A because of the gene homoeology.

For the cultivar Langdon, among its identified 5A protein, the chitinase and the ADH1A were already reported to be encoded by genes on the chromosome 5A (GrainGenes), meanwhile the XIP-III was confirmed to be encoded by genes on the chromosome 5A by PCR (Kalunke, 2012).

After interrogation of the GrainGenes database, for *T.dicoccoides,* in addition to ADH1A previously cited, no other protein was found to be encoded by genes present on the chromosome 5A. The malate dehydrogenase and the globulin-like protein were found associated with the long arm of the chromosomes 5DL and 5BL, respectively.

Table 12: Proteins identified by LC-MS/MS in *T.dicoccoides*

Spot	Peptides	Protein	Chromosomal localization*
1	2	Tubulin beta-3 chain	\
2	2	Alcohol dehydrogenase ADH1A	1BL, 1DL, 2B, 2BS, 2DS, 3AL, 3BL, 3DL, 4A, 4BS, 4DS, **5AL**, 5BL, 5DL, 6AL, 6BL, 6DL, 7AL, 7BL, 7DL, 7DS
4	2	Malate dehydrogenase	1Al, 1BL, 1DL, 2AL, 2BL, 2DL, 2BS, 3AL, 3BL, 3DL, 5DL, 6AL, 6BL, 6DL
4	2	Glyceraldehyde-3-phosphate dehydrogenase	4DS, 6AL, 6BL
5	5	Malate dehydrogenase	1AL, 1BL, 1DL, 2AL, 2BL, 2DL, 2BS, 3AL, 3BL, 3DL, 5DL, 6AL, 6BL, 6DL
6	2	Isoflavone reductase-like protein	\
7	2	Globulin 3	1A, 1BL, 2AL, 2BS, 2DL, 3BL, 3DL, 5BS, 5BL, 5DS, 6AL, 6BL, 6DL, 6DS, 7BS, 7BL, 7D, 7DL
8	2	Pathogenesis-related 1b	\
9	2	Pathogenesis-related 1b	\
10, 11, 12	\	\	\
13	3	Triosephosphate isomerase	3AS, 3AL, 3BS, 3BL, 3DS
14	3	Manganese Superoxide dismutase (MnSOD)	2AL, 2BL, 2DL
15	2	27K protein	\
	2	Cold shock domain protein 3	7AL, 7BL
16	2	Gliadin/avenin-like seed protein	
17, 18, 19	\	\	\
20	3	Single-stranded nucleic acid binding protein	3DL
21	2	Pre-mRNA processing factor	\
	2	Globulin-like protein	2BS, 5BL
22 & 23	6 & 2	Alpha amylase inhibitor CM3	
24	4	Dimeric alpha-amylase inhibitor	4AL, 4BL, 4DL, 4BS, 4DS, 7AS, 7DS
25 & 26	5 & 3	Monomeric alpha-amylase inhibitor	

/ = no significative result
*= obtained by GrainGenes (http://wheat.pw.usda.gov/cgi-bin/westsql/map_locus.cgi)

Table 13: Proteins identified by LC-MS/MS in *T.turgidum* ssp. *durum* ***cv Langdon***

Spot	Peptides	Protein	Chromosomal localization*
1	1	chitinase	1AL, 1AS, 1BL, 1BS, 1D, 1DS, 2AL, 2AS, 2BL, 2BS, 2DL, 2DS, 3AL, 3BL, 3DL, 4AL, 4BL, **5AL, 5AS**, 6AL,
2	6	Putative late embryogenesis abundant protein	1AL, 1BL, 1DL, 3BS, 3DL, 4AL
3	6	Putative late embryogenesis abundant protein	
4	3	Globulin 3	1A, 1BL, 2AL, 2BS, 2DL, 3BL, 3DL, 5BS, 5BL, 5DS, 6AL, 6BL, 6DL, 6DS, 7BS, 7BL, 7D, 7DL
5	6	Triticin	1AS, 1BS, 1DS
6	4	Alcohol dehydrogenase ADH1A	1BL, 1DL, 2B, 2BS, 2DS, 3AL, 3BL, 3DL, 4A, 4BS, 4DS, **5AL**, 5BL, 5DL, 6AL, 6BL, 6DL, 7AL, 7BL, 7DL, 7DS
	4	Alpha amylase inhibitor CM3	4AL, 4BL, 4DL, 4BS, 4DS, 7AS, 7DS
7	10	Glucose and ribitol dehydrogenase	\
8	6	Xylanase inhibitor XIP-III	3DL
9	1	Xylanase inhibitor XIP-III	
10	4	Alpha amylase inhibitor CM3	4AL, 4BL, 4DL, 4BS, 4DS, 7AS, 7DS

/ = no significative result

*= obtained by GrainGenes (http://wheat.pw.usda.gov/cgi-bin/westsql/map_locus.cgi)

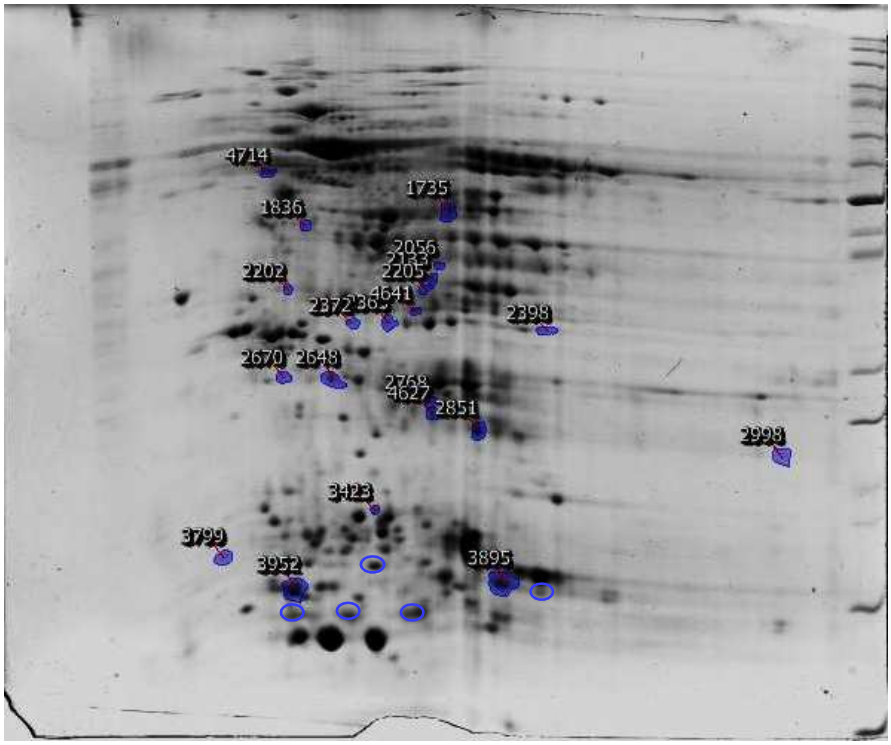


Figure 16: Polypeptides identified as encoded by genes on chromosome 5A of *T. dicoccoides*

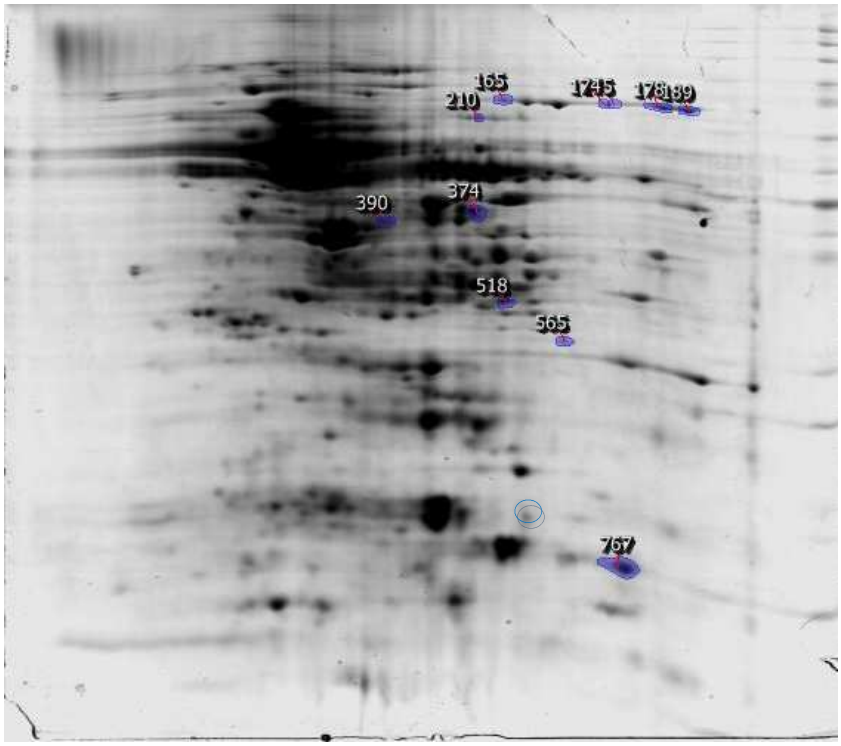


Figure 17: Polypeptides identified as encoded by genes on chromosome 5A of *T.turgidum* ssp. *durum*_cv Langdon

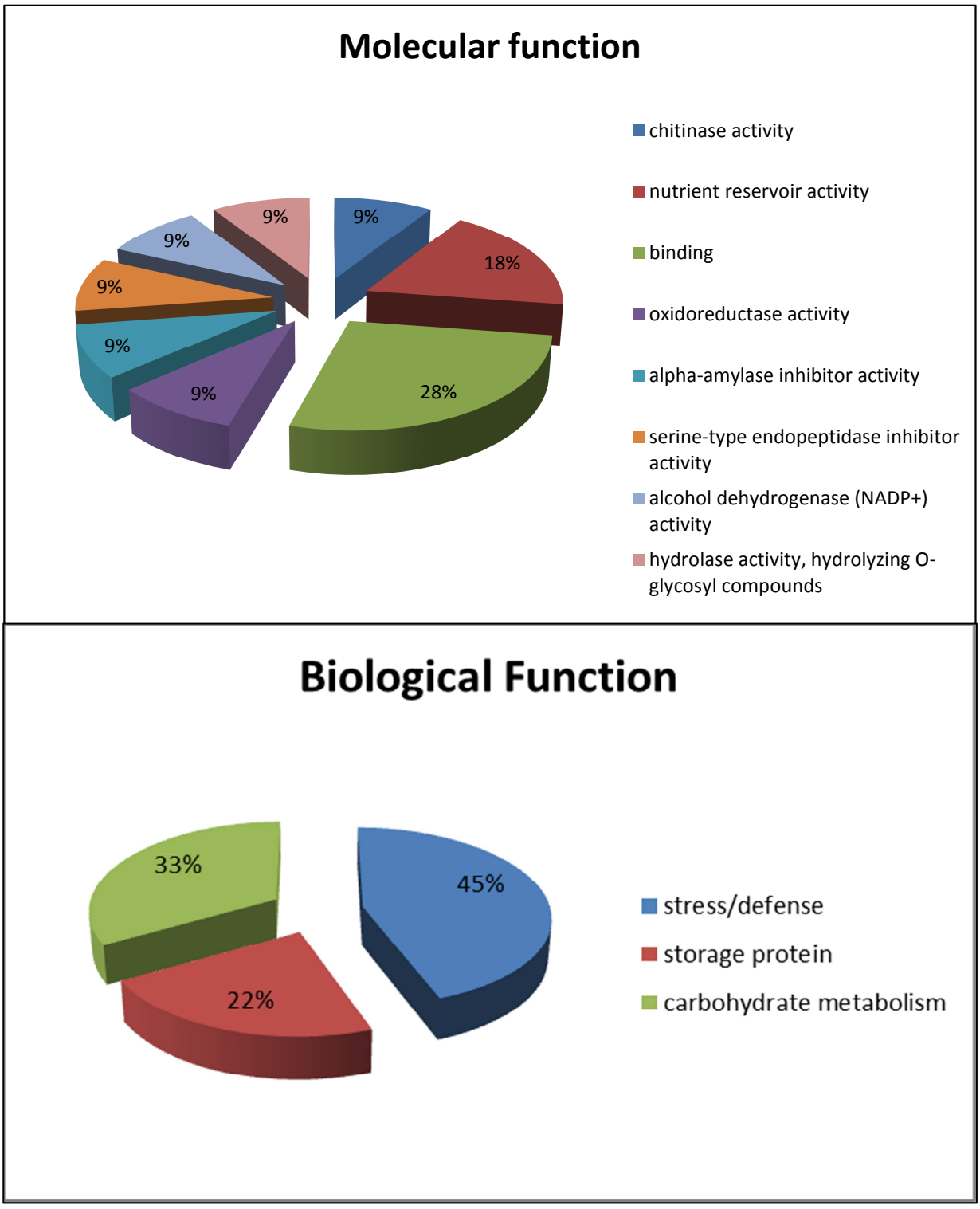


Figure 18: Percentage distribution of the biological process and the molecular function relative to the identified proteins in *T.turgidum* ssp. *durum* cv Langdon obtained by gene ontology (http://www.uniprot.org/uniprot) and by literature research

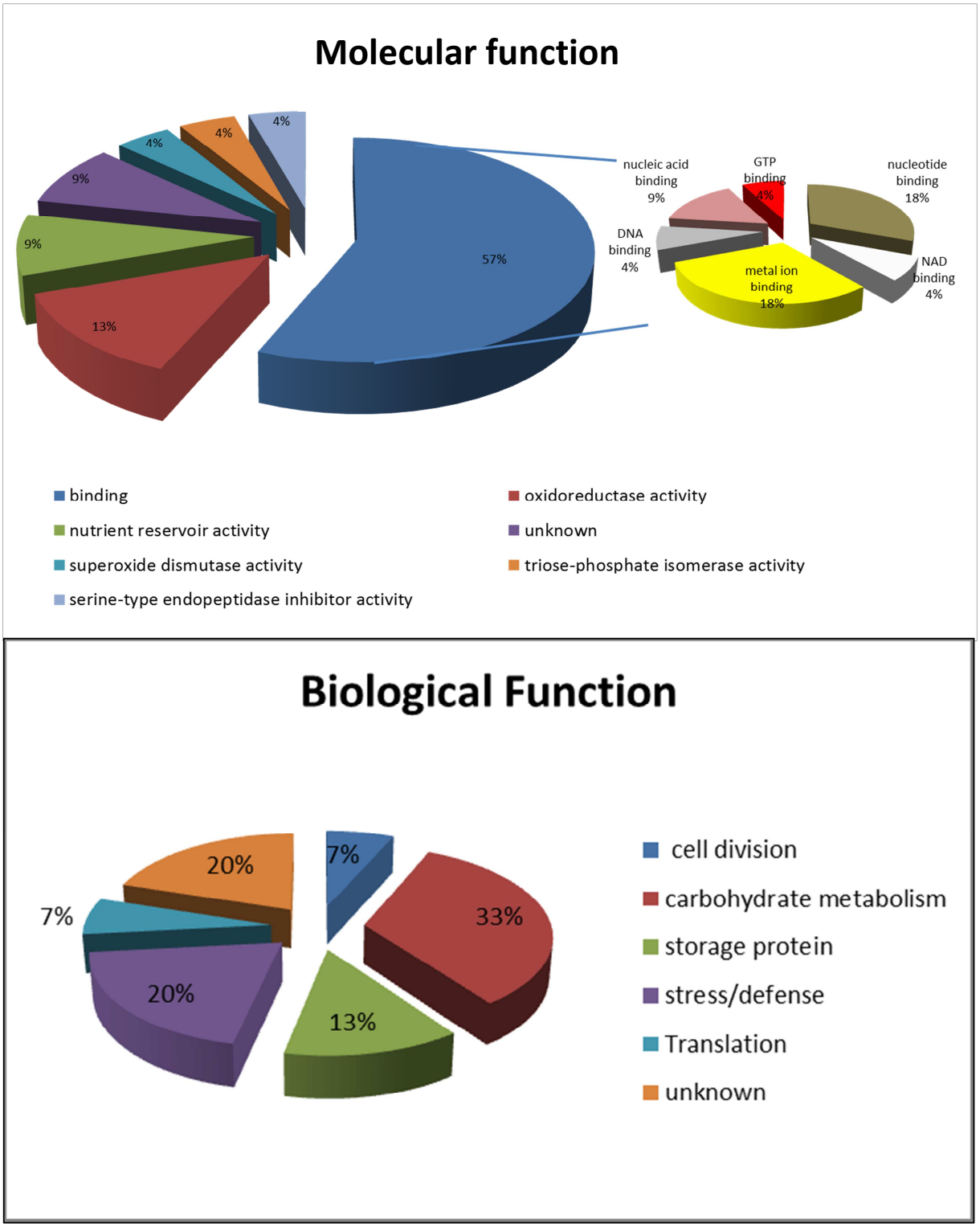


Figure 19: Percentage distribution of the biological function and the molecular function relative to the identified proteins in *T. diccocoides*, obtained by gene ontology (http://www.uniprot.org/uniprot) and by literature research

3. CM-like fraction

This fraction was taken into consideration because it contains most of the known allergenic proteins (Shewry, 2009; Salcedo et al., 2011; Mamone et al., 2011).

The proteomic comparison indicated that 6 spots were attributable to chromosome 5A of *T. dicoccoides* and 3 spots to Langdon (Figures 20 & 21).

These spots were picked and submitted to Mass Spectrometry analysis, whose results are reported in table 14. With the few exception of adenylate kinase (spot 1), Class III chitinase, vacuolar H+ pyrophosphatase (spot 2), and Superoxide dismutase [Cu-Zn] (spot 3), the great majority of polypetides identified belonged to the α-amylase/trypsin inhibitors (Hurkman and Tanaka, 2004). Even if others proteins have been identified in these spots, the α-amylase inhibitor was always the most abundant polypeptide. This is very likely an artifatct due to the problem of the great abundance of this protein type among CM proteins. Its abundance might cause fixation on the Liquid-Chromatography column used for MS analysis, consequently it is found in all the spots and may cover the signal of other proteins. The cereal α-amylase/trypsin inhibitor subunits are 12-16–kDa polypeptides with 4-5 intrachain disulphide bridges that are essential for their inhibitory activity. Members of the inhibitor family are restricted to the seed storage tissue (endosperm) (Salcedo et al., 2011). These inhibitors are encoded by a multigene family which is dispersed over several chromosomes; in particular they have been assigned to the short and long arms of the chromosomes 3, 4, 6 and 7 of B and D genome (Carbonero et al., 1999; Islam et al., 2003).

In order to try to solve this problem, the spots have been submitted to mass spectrometry twice, once with a mass exclusion list to remove the α-amylase/trypsin inhibitor of the results. However, using a mass exclusion list is risky because it could also eliminate other proteins than that referred. Thus identifications are made difficult, and the results presented need to be confirmed. For these reasons, we eventually performed this analysis on tetraploid wheats only, and relative results are reported below.

The spot 1 was identified, in addition to the α-amylase inhibitor, as an adenylate kinase. This protein catalyzes the phosphorylation of Adenosine-5’-monophosphate (AMP), using as phosphate donor Adenosine-5’-triphosphate (ATP) (Benito et al., 1989).

The spot 2 is represented by the chitinase, which was described previously (spot 1 of the metabolic fraction of Langdon). The vacuolar H+ pyrophosphatase was also identified in this spot. The vacuole is a dynamic organelle involved in several cellular processes (storage of metabolites and ions, regulation of cytosolic homeostasis, degradation and recycling of cellular components, and space filling) which are directly or indirectly related to either the transmembrane electrochemical gradient across the vacuolar membrane or acidic pH in the vacuole (Suzuki et al., 1999).

The spot 3 was identified as a Superoxide dismutase [Cu-Zn]. This group of proteins was described previously (spot 14 of the *T. dicoccoides* metabolic fraction).

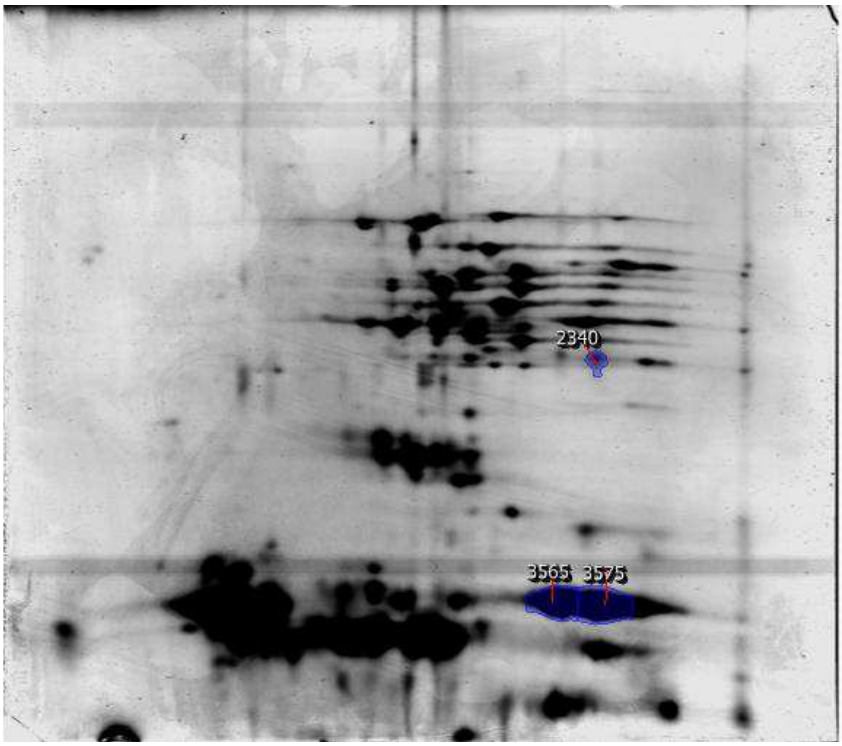


Figure 20: Polypeptides identified, in the CM-like fraction, as encoded by genes on chromosome 5A of *T. turgidum* ssp. *durum* cv. Langdon

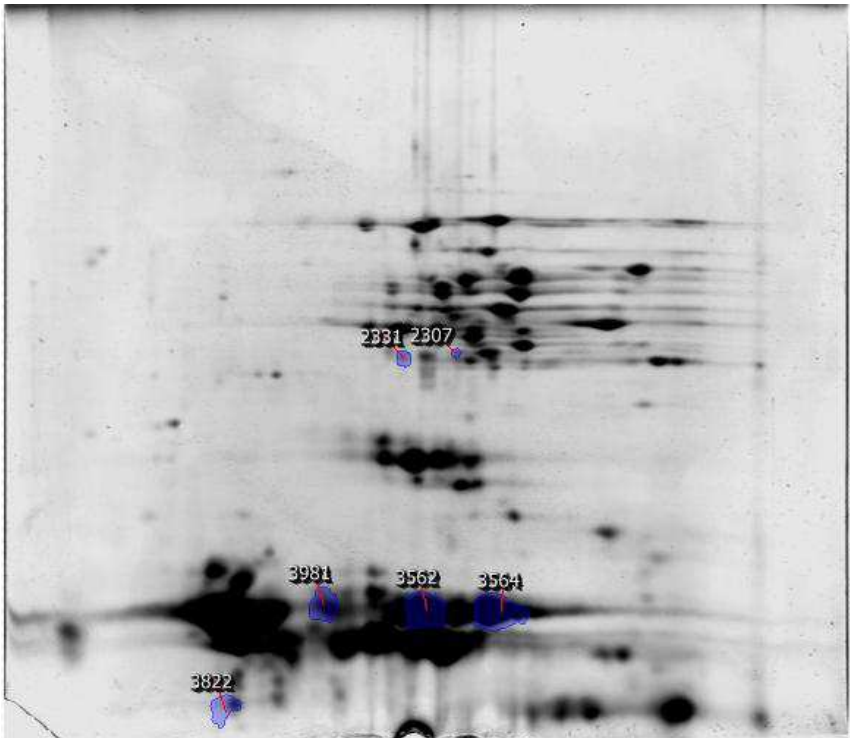


Figure 21: Polypeptides identified, in the CM-like fraction, as encoded by genes on chromosome 5A of *T. dicoccoides*

Table 14: Polypeptides identified in the CM fraction of T. *dicoccoides* and *T. turgidum* ssp. *durum* cv. Langdon

TIGR TC Triticum Release 12.0 (April 18, 2010)

	Serie 1				Serie 2			
n° Spot	Protein	Score Mascot	Peptides	emPAI	Protein	Score Mascot	Peptides	emPAI
1	Alpha-amylase/trypsin inhibitor CM3 precursor	704	12	17.93	Alpha-amylase/trypsin inhibitor CM3 precursor	589	9	9,91
					Dimeric alpha-amylase inhibitor	275	4	1,14
					Adenylate kinase A	134	2	0,28
2	Alpha-amylase/trypsin inhibitor CM3 precursor	673	11	17.93	Alpha-amylase/trypsin inhibitor CM3 precursor	576	9	9,91
	Dimeric alpha-amylase inhibitor	489	7	2,79	Dimeric alpha-amylase inhibitor	489	8	3,58
	Class III chitinase	81	2	0,25	Class III chitinase	114	2	0,25
	Vacuolar H+-pyrophosphatase	65	2	0,26				
3	Alpha-amylase/trypsin inhibitor CM3 precursor	678	11	17,93	Alpha-amylase/trypsin inhibitor CM3 precursor	680	11	26,35
	Dimeric alpha-amylase inhibitor	488	7	2,79	Dimeric alpha-amylase inhibitor	607	9	4,54
	Superoxide dismutase [Cu-Zn] 4A	257	3	0,94	Superoxide dismutase [Cu-Zn] 4A	183	3	0,5
	Alpha-amylase/trypsin inhibitor CM16 precursor	225	6	2,03	CM 17 protein precursor	96	3	0,63
	Alpha-amylase/trypsin inhibitor CM2 precursor	186	3	0,91				
	Alpha-amylase inhibitor 0.28 precursor	180	4	1,22				
	Vacuolar H+-pyrophosphatase	66	2	0,26				
4	Alpha-amylase/trypsin inhibitor CM3 precursor	723	12	46,47	Alpha-amylase/trypsin inhibitor CM3 precursor	722	11	17,93
	Dimeric alpha-amylase inhibitor	249	4	1,14	Dimeric alpha-amylase inhibitor	434	8	2,79
	Alpha-amylase/trypsin inhibitor CM2 precursor	89	2	0,54	0.19 alpha-amylase inhibitor	191	4	1,11
					Alpha-amylase/trypsin inhibitor CM16 precursor	136	3	0,45
					Alpha-amylase/trypsin inhibitor CM2 precursor	91	2	0,54
5	Alpha-amylase/trypsin inhibitor CM3 precursor	744	12	46,47	Alpha-amylase/trypsin inhibitor CM3 precursor	731	12	26,35
	Alpha-amylase inhibitor 0.53	190	2	1,79	Alpha-amylase/trypsin inhibitor CM3 precursor	607	9	20,03
					Dimeric alpha-amylase inhibitor	129	2	0,44
6	Alpha-amylase/trypsin inhibitor CM3 precursor	587	9	6,55	Alpha-amylase/trypsin inhibitor CM3 precursor	544	7	4,23
	Alpha-amylase inhibitor 0.28 precursor	383	8	5,02	Alpha-amylase inhibitor 0.28 precursor	367	8	5,02
1L	Alpha-amylase/trypsin inhibitor CM3 precursor	618	9	8,08	Alpha-amylase/trypsin inhibitor CM3 precursor	524	6	4,23
	Dimeric alpha-amylase inhibitor	148	2	0,46	Dimeric alpha-amylase inhibitor	160	2	0,46
2L	Alpha-amylase/trypsin inhibitor CM3 precursor	591	9	14,75	Alpha-amylase/trypsin inhibitor CM3 precursor	689	11	21,75
3L	Alpha-amylase/trypsin inhibitor CM3 precursor	594	9	14,75	Alpha-amylase/trypsin inhibitor CM3 precursor	619	10	21,75
					Alpha-amylase/trypsin inhibitor CM3 precursor	506	7	16,16

Uniprot-Viridiplantae release2010_09 (13/08/2010)

	Serie 1				Serie 2			
n° Spot	Protein	Score Mascot	Peptides	emPAI	Protein	Score Mascot	Peptides	emPAI
1	Alpha-amylase/trypsin inhibitor CM3 precursor	1762	12	18,57	Alpha-amylase/trypsin inhibitor CM3 precursor	2295	9	10,21
					Dimeric alpha-amylase inhibitor	193	4	1,45
2	Alpha-amylase/trypsin inhibitor CM3 precursor	2467	11	18,57	Alpha-amylase/trypsin inhibitor CM3 precursor	1894	9	10,21
	Dimeric alpha-amylase inhibitor	412	7	6,45	Dimeric alpha-amylase inhibitor	466	8	8,57
	Legumin A	51	2	0,14	Legumin A	95	2	0,3
3	Alpha-amylase/trypsin inhibitor CM3 precursor	3165	11	18,57	Alpha-amylase/trypsin inhibitor CM3 precursor	5177	11	40,17
	Dimeric alpha-amylase inhibitor	440	7	4,92	Dimeric alpha-amylase inhibitor	911	9	13,39
	Superoxide dismutase [Cu-Zn] 4A	200	3	0,97	Major allergen CM16	158	6	2,57
	Major allergen CM16	182	5	2,57	Superoxide dismutase [Cu-Zn] 4A	125	3	0,97
	Alpha-amylase/trypsin inhibitor CM2 precursor	132	3	0,92	Cereal-type amylase inhibitor	118	3	1,71
4	Alpha-amylase/trypsin inhibitor CM3 precursor	6826	12	48,58	Alpha-amylase/trypsin inhibitor CM3 precursor	5445	11	18,57
	Dimeric alpha-amylase inhibitor	223	3	1,45	Dimeric alpha-amylase inhibitor	320	8	3,74
					Alpha-amylase/trypsin inhibitor CM2 precursor	89	4	0,69
					Major allergen CM16	88	3	0,89
5	Alpha-amylase/trypsin inhibitor CM3 precursor	8730	12	48,58	Alpha-amylase/trypsin inhibitor CM3 precursor	6996	12	27,39
	Dimeric alpha-amylase inhibitor	244	3	1,43	Dimeric alpha-amylase inhibitor	99	9	0,56
6	Alpha-amylase/trypsin inhibitor CM3 precursor	6738	9	6,73	Alpha-amylase/trypsin inhibitor CM3 precursor	5683	7	4,33
	Monomeric alpha-amylase inhibitor	730	9	11,54	Monomeric alpha-amylase inhibitor	477	9	11,54
1L	Alpha-amylase/trypsin inhibitor CM3 precursor	3550	9	8,31	Alpha-amylase/trypsin inhibitor CM3 precursor	3228	6	4,33
	Dimeric alpha-amylase inhibitor	121	2	0,55	Dimeric alpha-amylase inhibitor	129	2	0,55
	Legumin A	116	2	0,3	Legumin A	103	2	0,14
2L	Alpha-amylase/trypsin inhibitor CM3 precursor	4696	9	15,25	Alpha-amylase/trypsin inhibitor CM3 precursor	5255	11	22,57
3L	Alpha-amylase/trypsin inhibitor CM3 precursor	5600	9	15,25	Alpha-amylase/trypsin inhibitor CM3 precursor	6524	10	22,57
					Dimeric alpha-amylase inhibitor	86	7	0,56

4. Gliadin and glutenin fraction

Regarding the glutenin and the gliadin fractions, even if previous studies have shown that gliadin and glutenin are mainly coded by genes at chromosomes 1 and 6 (Payne, et al., 1984; Lafiandra et al., 1984), our intention was to detect possible polypeptides encoded by genes on chromosome 5A, since Sreeramulu and Singh (1997) have assigned a low molecular weight glutenin subunits to the chromosomes 1D and 7D, suggesting that this possibility needs to be taken into consideration.

Because gliadins present a low variation in pI, the classical separation by IEF vs SDS-PAGE may not be resolutive for this protein fraction. For this reason, we used Acid-PAGE for the first dimension of gliadins, while we used the classical separation procedure for glutenin subunits.

For the gliadin fraction a total of 92 spots were detected in the comparison TD vs LDN-TD5A, while a total of 99 spots were detected in the comparison LDN vs LDN-TD5A, but no spots were found to be 5A specific (fig. 23).

For the glutenin fraction 191 spots were detected in the comparison TD vs LDN-TD5A and a total of 228 spots were found in the comparison LDN vs LDN-TD5A, however also for this fraction no spots were found to be 5A specific (fig.22).

In conclusion the analysis of the gels has excluded the hypothesis that, at least the major represented gliadins and glutenin subunits are encoded by genes on the chromosome 5A.

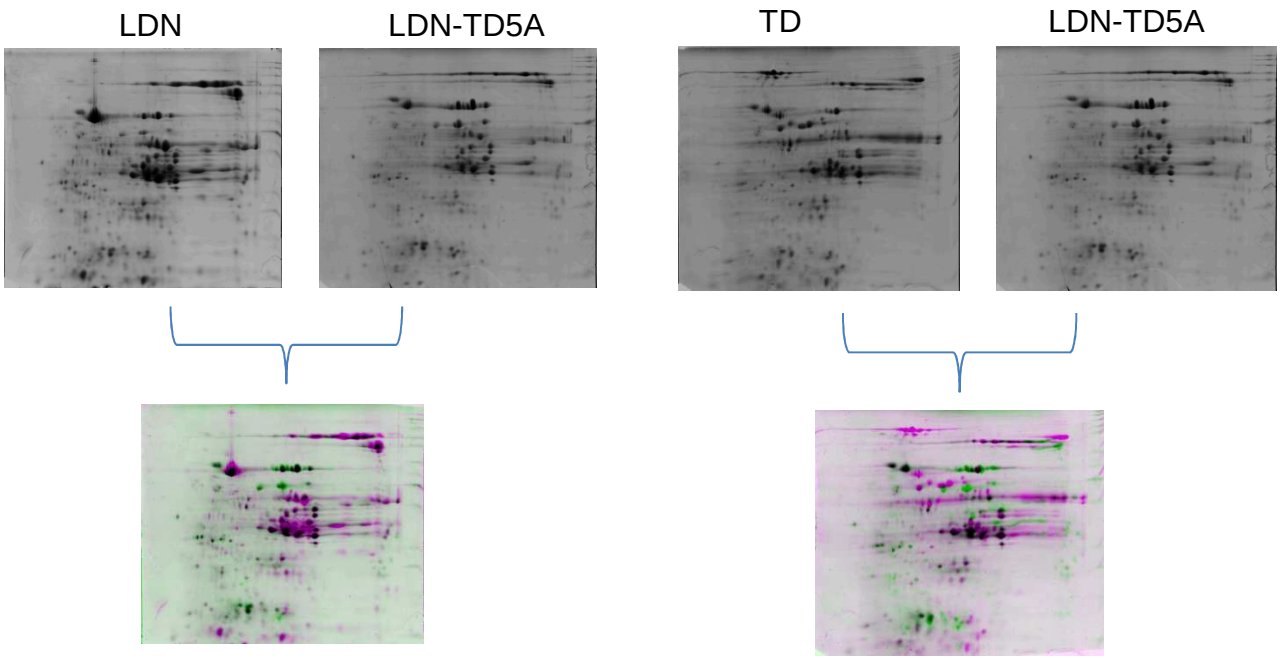


Figure 22: 2D gels relative to glutenin subunits. The lower part represents the alignment by Progenesis Samespots (Nonlinear Dynamics)

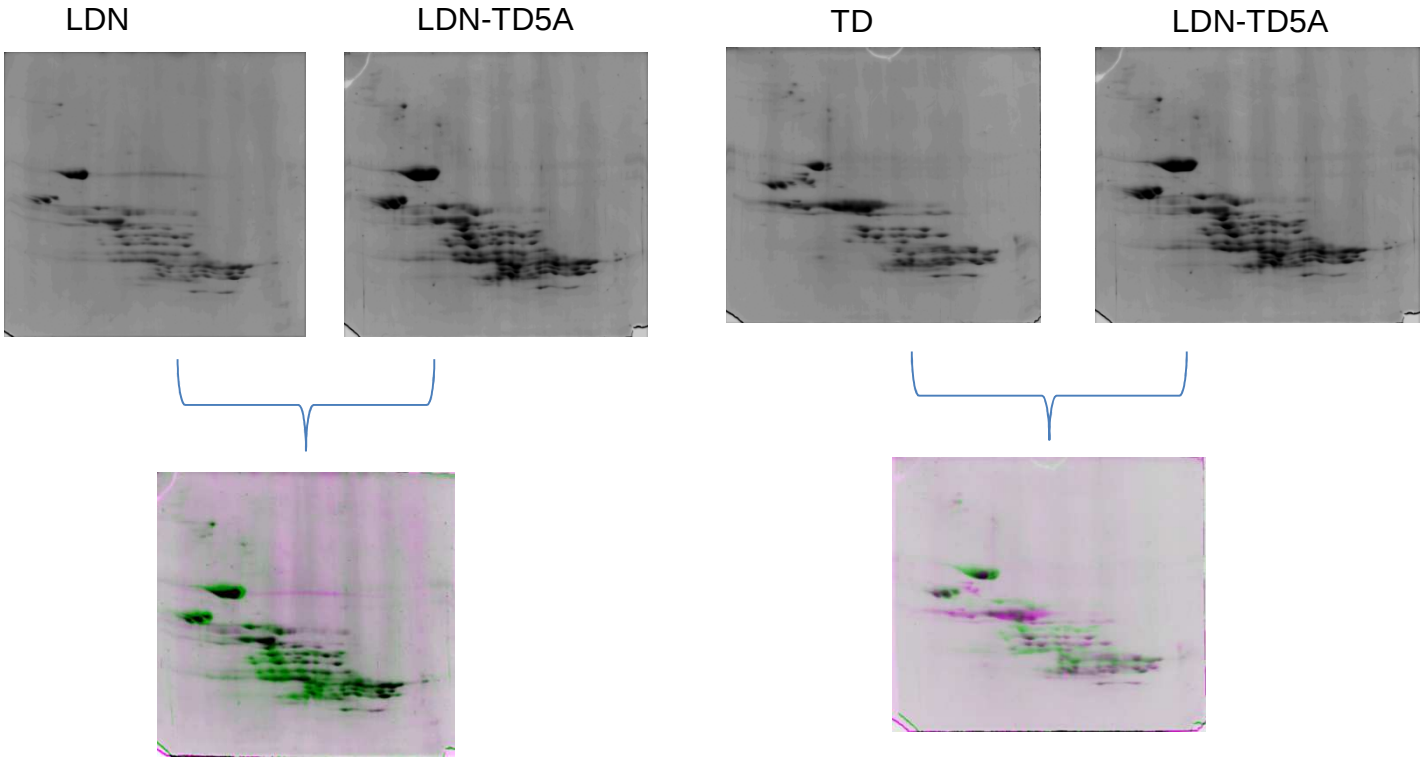


Figure 23: 2D gels relative to gliadins. The lower part represents the alignment by Progenesis Samespots (Nonlinear Dynamics)

5. Nuclear proteome of leaves

Although the main aim of the project is the identification of 5A encoded polypeptides expressed in seeds, we deemed it interesting to detect also polypeptides expressed in other plant tissues. In order to avoid the difficulties we encountered in studying the soluble protein fraction of seeds, due to the high number of polypeptides that make the comparisons rather complicate, we decided to focalize on the nuclear proteome of leaves, in which a low number of polypeptides should be present. We used leaves of seedlings collected three weeks after germination.

The nuclei were isolated as reported in the methods section, and the nuclei suspension was checked under fluorescence microscopy in order to check nuclei integrity (fig. 24).

We did not perform 2D gels because the amount of proteins was not enough, and thus decided to use SDS-PAGE in order to eventually perform 1D MS analyses on electrophoretic gel slices (Repetto et al, 2008).

Gel replicas were analysed with LabImage, in order to control their reproductibility. It was possible to detect 9 bands in Langdon, 8 bands in *T. dicoccoides* and 5 bands in LDN-TD5A (fig.25).

Although 1D SDS-PAGE usually does not allow to detect the presence/absence of protein bands, because the resolution power is obviously low compared to 2D separations, in this case the comparison allowed to identify a group of bands around 50 kDa in *T. turgidum* ssp. *durum* cv Langdon, that disappear in the substitution line LDN-TD5A (fig.26 & 27). Because a group of bands with similar molecular weight is instead present in *T. dicoccoides*, it is likely that bands present in Langdon are 5A encoded, but not those present in *T. dicoccoides*.

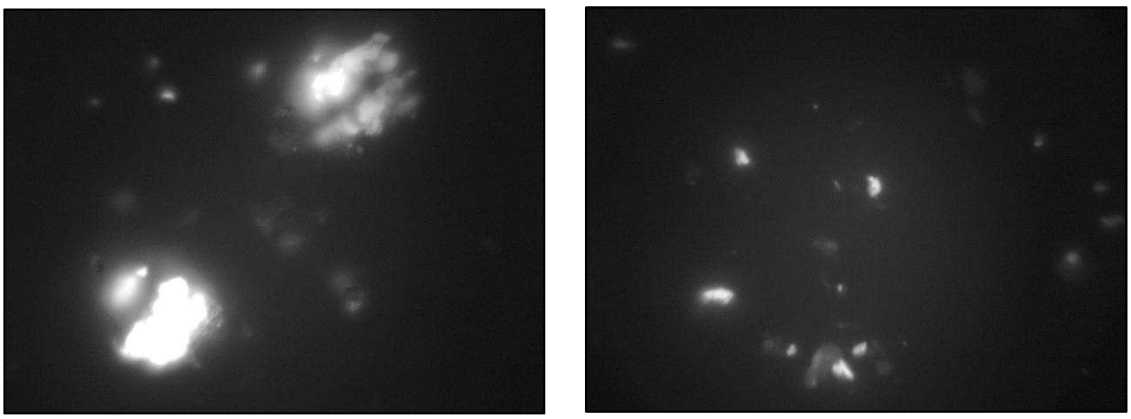


Figure 24: Nuclei suspension, stained with propidium iodide, observed under fluorescence microscopy.

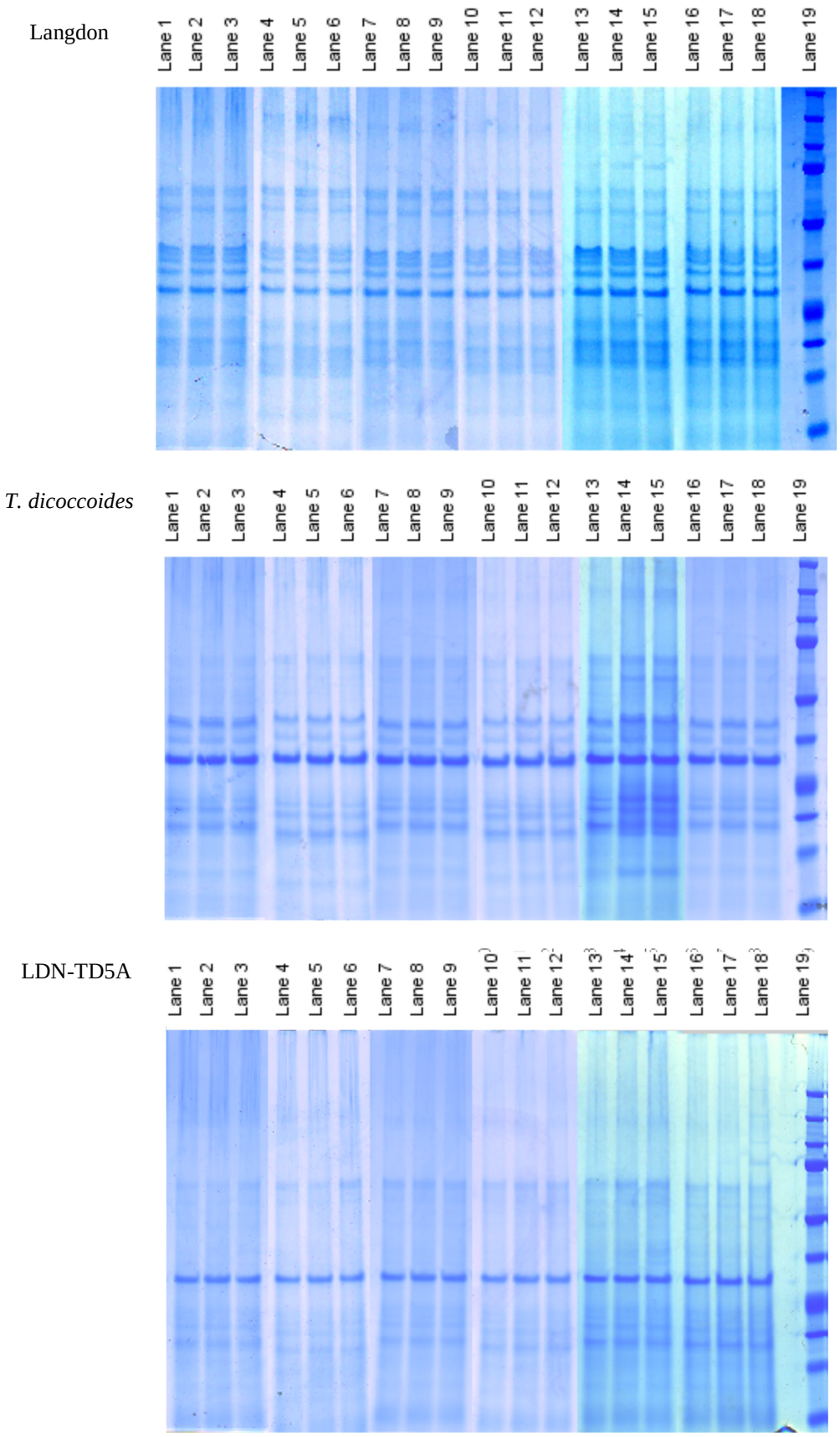


Figure 25: NuPAGE® Novex® Bis-Tris Gels of *T.durum* cv Langdon, *T. dicoccoides* and Langdon-*T. dicoccoides*5A nuclear protein extract, grouped by wheat lines. 18 replicates for each line.

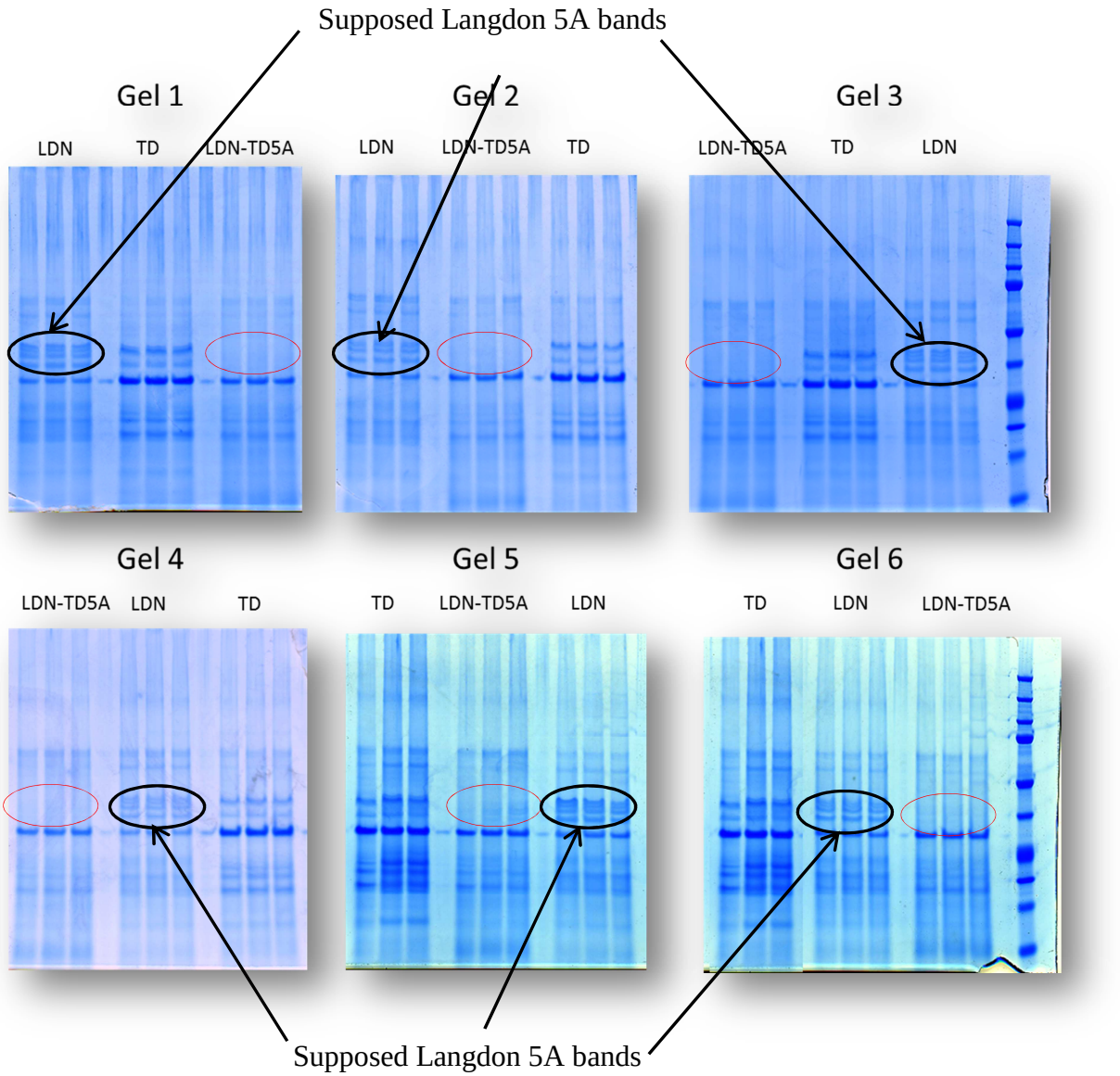


Figure 26: NuPAGE® Novex® Bis-Tris Gels of of nuclear proteins extracted from *T.durum* cv Langdon, *T. dicoccoides* and Langdon-*T. dicoccoides* 5A

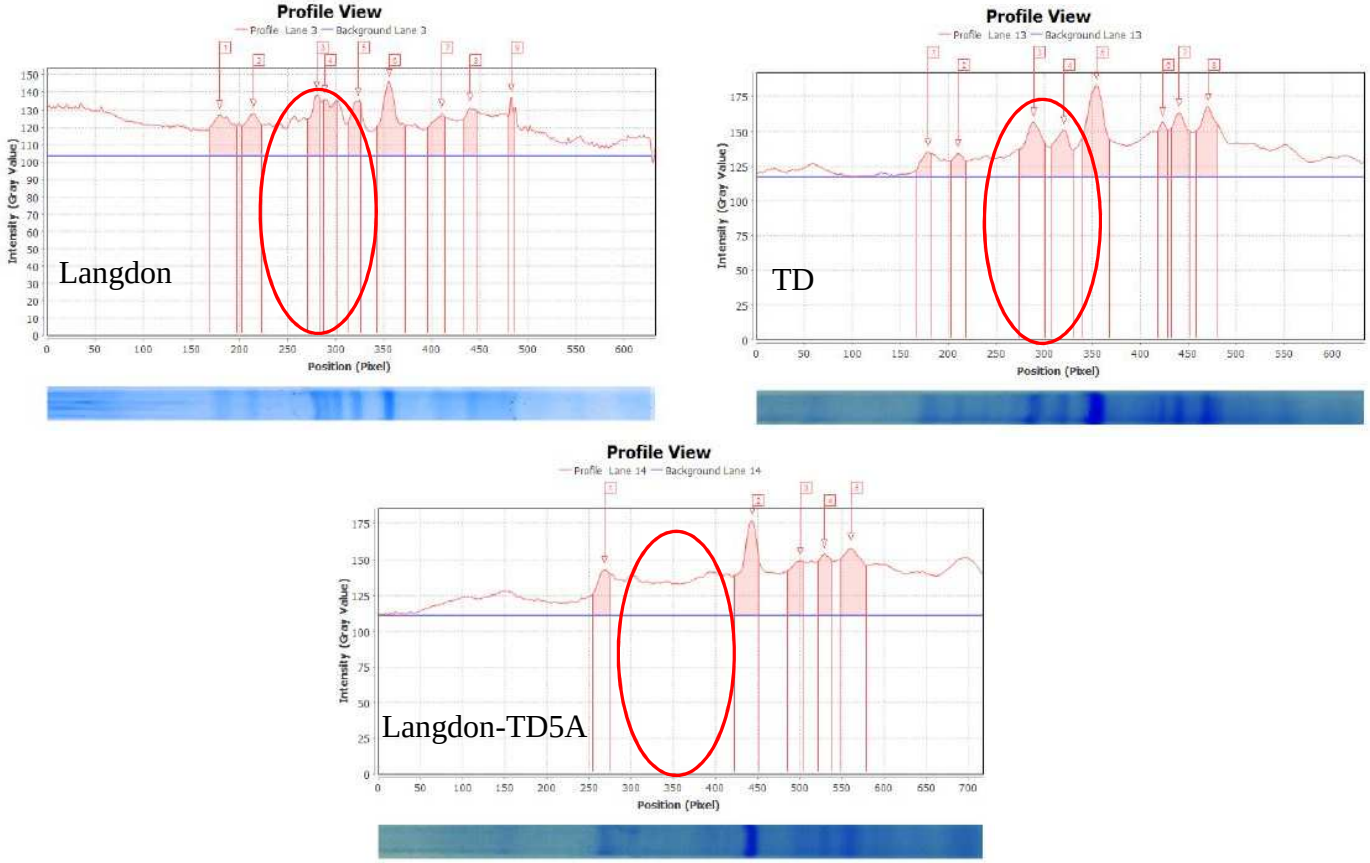


Figure 27: LabImage results on the Nupage bis-tris gels of *T.durum* cv Langdon, *T. dicoccoides* and Langdon-T. dicoccoides5A.

B. Bread wheat

In order to identify 5A polypeptides in different bread wheat cultivars, we used cultivars Cheyenne, Hope, Thatcher, and Timstein, for which 5A chromosome substitution lines are available in the Chinese Spring background. We first analysed 2D-gels, always comparing the pattern of the background cultivar Chinese Spring (CS) with that of the other parental line and the 5A chromosome substitution line. This allowed to attribute differential protein spots to chromosome 5A of either parental cultivars. After spot identification, MS analysis was performed, by using two databanks (UniProt and TIGR). The results are reported cultivar by cultivar.

1. Identification of 5A polypeptides in the cultivar Hope

According to the comparison CS vs Hope; CS vs CS-Hope5A; Hope vs CS-Hope5A, 19 spots correspond to polypeptides encoded by genes on the chromosome 5A of Hope. All the selected spots have a $p<0.05$, a q-value <0.05, a fold change $\geq 1,5$ and a power ≥ 0.8. The PCA (fig.31) confirmed the specificity of the 5A spots selected.

All the Hope 5A spots, their statistical values and their identification are presented in the next pages, but only the more abundant protein in the spot is described. The results obtained for Chinese Spring are presented separately, since the genetic background of this cultivar was in common to all the substitution lines, allowing a more detailed analysis.

The proteins identified in the cultivar Hope are mainly involved in the binding (38%) and in protease/hydrolase activity (17%) (fig.32).These results are comparable with these obtained in the tetraploid wheat, and with those reported by Vitulo et al (2011).

After interrogation of the GrainGenes EST database, on a total of 41 identfied proteins, 7 proteins (17%) (Protein disulfide isomerase; Elongation factor; 3ketoacyl-CoA thiolase like protein; xylanase inhibitor; Nascent polypeptide-associated complex; Fructose-biphosphate aldolase; chitinase) resulted to be encoded by genes on the chromosome 5A.

Among all the identified proteins, 5 proteins (12%) have been found in two or more spots.

Unfortunately it was not possible to attribute to either chromosome the great majority of the spots, since their mobility in 2D gel is overlapping (Fig. 30).

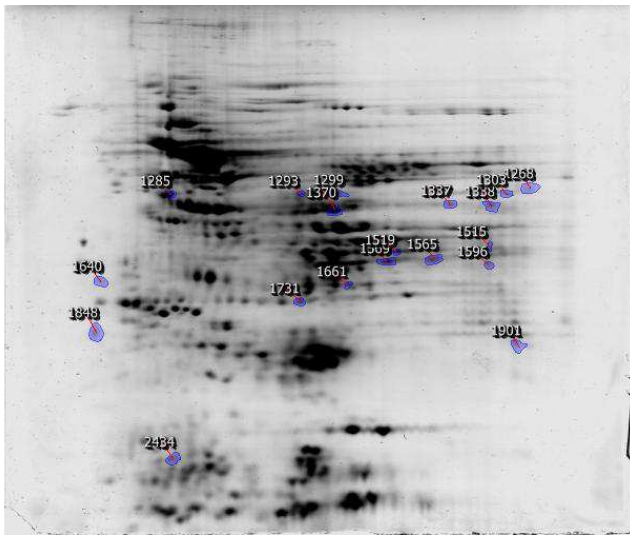


Figure 28: 5A specific spots for *T.aestivum* cv Hope

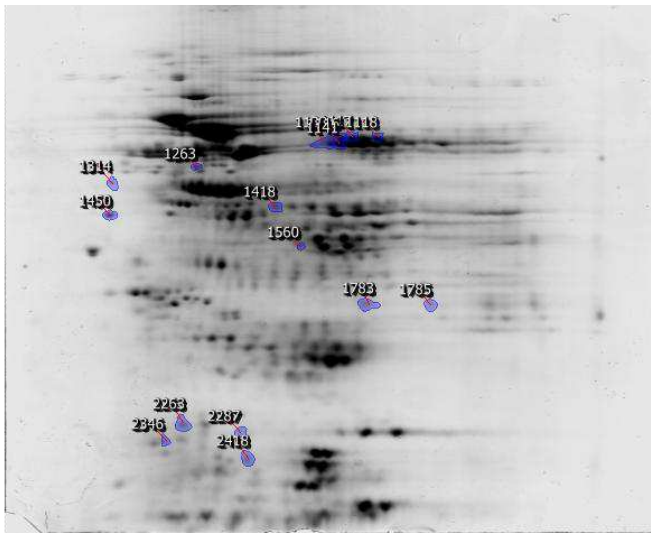


Figure 29: 5A specific spots for *T.aestivum* cv Chinese Spring

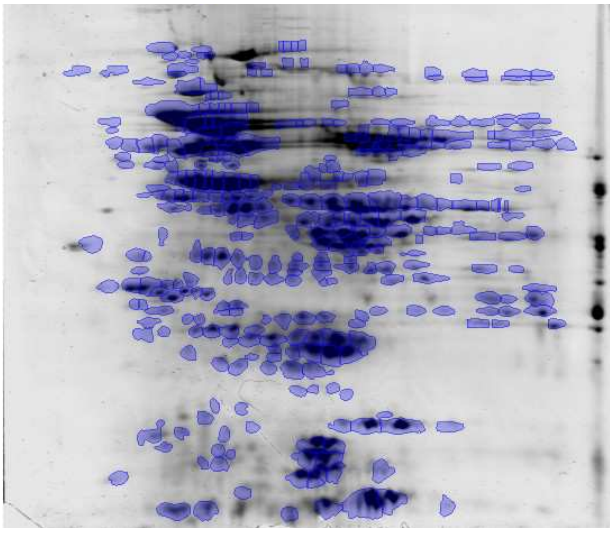


Figure 30: Common spots between *T.aestivum* cv Chinese Spring and *T.aestivum* cv Hope

Table 15: Identified polypeptides by LC MS/MS in the bread wheat *T. aestivum* cv Hope

N° spot	Protein	Score MASCOT	Peptides	emPAI
1	Glucose-1-phosphate adenylyltransferase	498	11	**2.19**
	Protein disulfide isomerase 3 precursor	319	8	0.73
	Aspartic proteinase	265	7	0.85
2	Elongation factor 1-gamma 2	482	11	**2.34**
	Triticin	313	7	1.36
	Citrate synthase	311	7	0.72
3	Alanine--glyoxylate aminotransferase 2	368	4	**1.39**
	Aspartate aminotransferase	255	8	0.96
4	HSP 70 precursor	623	11	4.12
	NADP-specific isocitrate dehydrogenase	534	15	**4.4**
	Alcohol dehydrogenase class 3	469	10	3.09
5	3-ketoacyl-CoA thiolase-like protein	457	8	**1.68**
6	Predicted protein	478	13	**2.32**
	3-ketoacyl-CoA thiolase-like protein	428	11	1.83
	Xylanase inhibitor	240	8	1.16
7	Predicted protein	630	17	**4.56**
	Globulin-2 precursor	182	6	0.66
8	Predicted protein	552	14	**3.01**
	Embryo globulin	145	6	0.74
	Globulin-like protein	143	6	0.47
9	Nucleosome chromatin assembly protein	309	5	**1.98**
10	Nascent polypeptide-associated complex NAC	118	2	0.49
	Predicted protein	373	10	**3.37**
11	Fructose-bisphosphate aldolase	209	4	**1.2**
12	Glucose and ribitol dehydrogenase	952	16	**7.07**
	Pyridoxine biosynthesis protein	331	7	5.8
	Short-chain dehydrogenase/reductase	501	9	1.49
13	Globulin-like protein	648	12	2.41
	Aldose reductase	267	7	**2.43**
14	rRNA N-glycosidase	201	7	1.44
	beta-amylase	283	9	**1.8**
15	Globulin-like protein	281	6	**2.55**
	Peroxidase 1	170	6	0.69
16	Peroxidase	174	5	**0.58**
17	Aldose reductase	220	9	**1.37**
	Malate dehydrogenase, glyoxysomal precursor	133	4	0.69
18	LEA1 protein	433	4	**4.49**
	Triosephosphate isomerase	251	5	1.18
	Class II chitinase	79	3	0.47
19	Superoxide dismutase [Cu-Zn]	100	3	0.5
	Alpha-amylase inhibitor	121	3	**1.11**

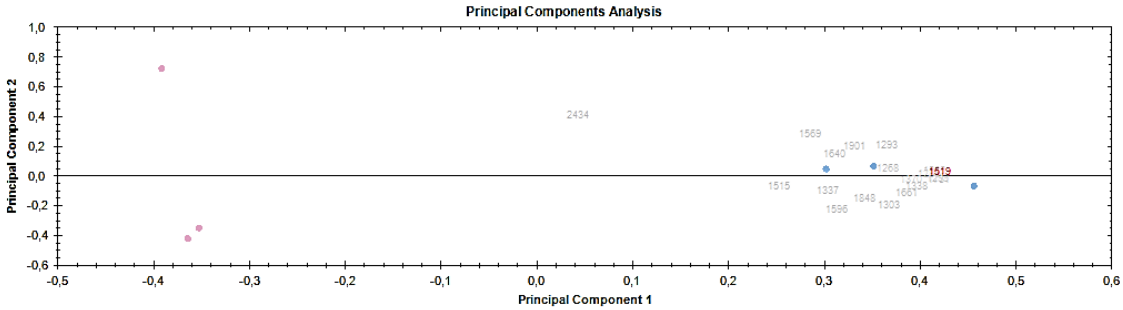


Figure 31: PCA representation in which the contribution of each spot of the two lines compared is reported. Blue points represent the three replicates of CS-Hope5A and the red points that of CS.

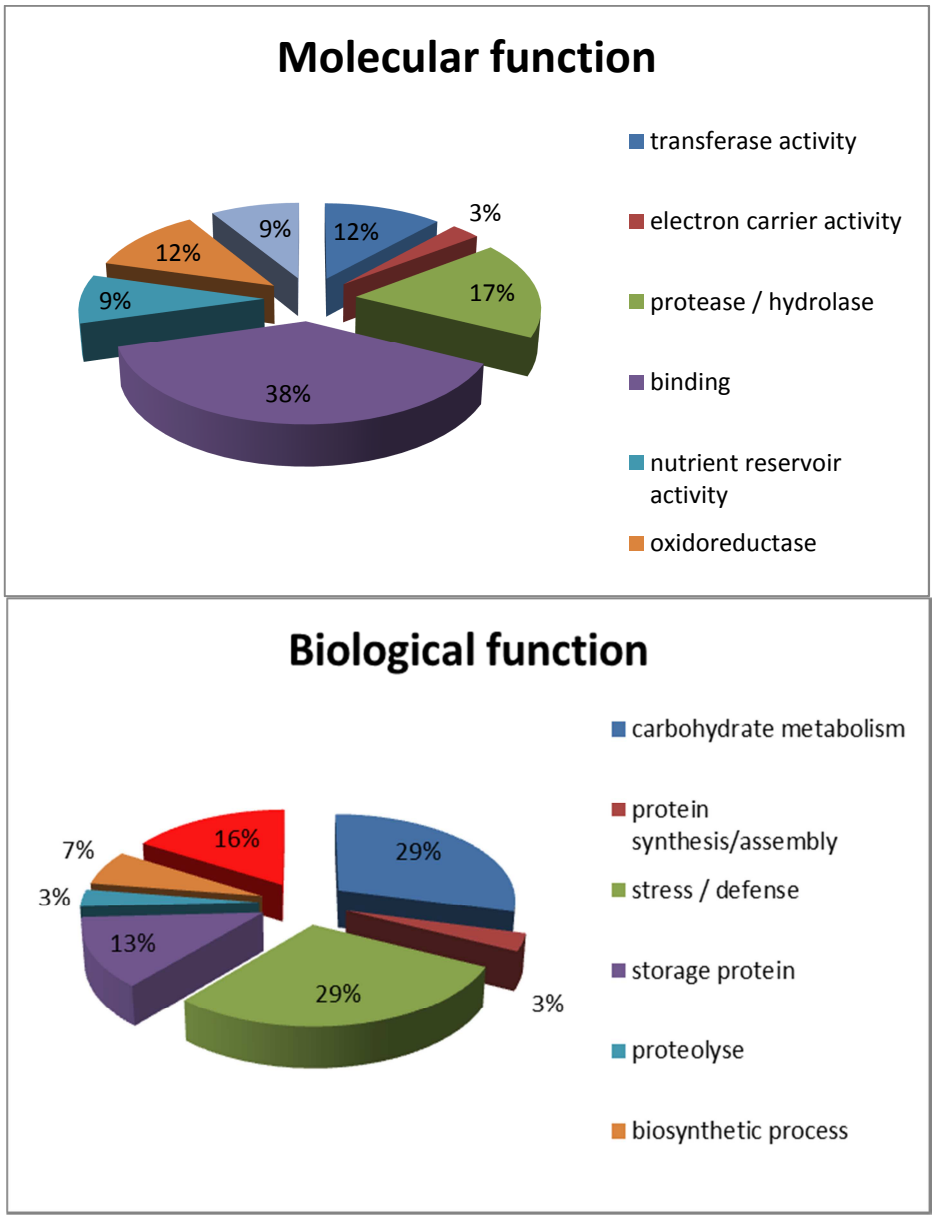
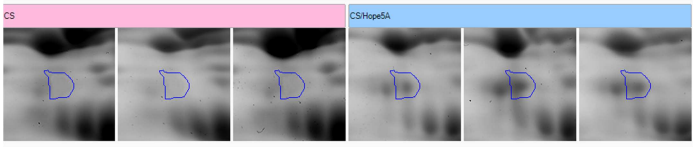


Figure 32: Percentage distribution of the biological process and the molecular function for the identified proteins in "Hope" obtained with the gene ontology (http://www.uniprot.org/uniprot/) and by literature research.

As an example, I report the information obtained by the Progenesis Samespots software, which show spot nr. 1, and the relative statistical analyses, along with MS results.

Spot 1



#	Anova (p)	Fold	q-value	Power
1285	0,00035	2,58169	0,01126	≥0,9995

TC	Protein	peptides	mass	Score	emPAI
TC368615_6	[2 - 1435] UniRef100_Q9M4Z1 Cluster: Glucose-1-phosphate adenylyltransferase; n=1; Triticum aestivum	11	52803	498	2,19
	Rep: Glucose-1-phosphate adenylyltransferase - Triticum aestivum (Wheat), complete				
TC401677_9	[3 - 1616] UniRef100_Q93XQ7 Cluster: Protein disulfide isomerase 3 precursor; n=3; Triticeae	8	59452	319	0,73
	Rep: Protein disulfide isomerase 3 precursor - Triticum aestivum (Wheat), complete				
TC390079_10	[3 - 1295] homologue to UniRef100_Q401N7 Cluster: Aspartic proteinase; n=1; Triticum aestivum	7	46684	265	0,85
	Rep: Aspartic proteinase - Triticum aestivum (Wheat), partial (95%)				

In the case of spot nr. 1, the most likely identification corresponds to glucose-1-phosphate adenylyltransferase that is involved in starch synthesis (Majoul et al., 2004). It is abundantly expressed in the whole grains, in particular in the endosperm.

The protein disulfide isomerase (PDI) was the second most abundant protein identified in spot 1. This protein is reported to be encoded by the 5A chromosome in the GrainGenes database. The PDI is one of the most abundant proteins in the endoplasmic reticulum (ER). The ability of PDI to bind to unfolded or partially folded proteins preventing their aggregation has also suggested its role as a chaperone. In cereals PDI may accomplish an important role in the folding of plant secretory proteins, particularly during the formation of endosperm protein bodies. The involvement of the typical PDI and probably of additional PDI-like proteins in the folding of endosperm storage proteins is especially important in wheat, because the processes occurring during protein synthesis and deposition may affect the functional properties of gluten, which play an integral role in determining the visco-elastic properties of wheat dough. Therefore, the genes encoding storage proteins, as well as factors that may affect their deposition, such as molecular chaperonesand foldase enzymes, are of particular interest to wheat industry (d'Aloisio et al., 2010).

Spot 2

Elongation factor 1-gamma 2 was described to participate in the polypeptide elongation cycle of protein synthesis (Lauer et al., 1984). Koonin et al (1994) have shown that Elongation factor 1-γ contains glutathione transferase (GST) domain, which is a widespread, conserved enzymatic module that may be covalently or noncovalently complexed with other proteins. Regulation of protein assembly and folding may be one of the functions of GST.

Spot 3

The Alanine--glyoxylate aminotransferase is reported in literature to function as a photorespiratory peroxisomal glutamate:glyoxylate aminotransferase (GGAT). GGAT, in photorespiration, catalyzes the reaction of glutamate and glyoxylate to produce 2-oxoglutarate and glycine. Photorespiratory transamination to glyoxylate, which is mediated by GGAT and serine glyoxylate aminotransferase (SGAT), is believed to play an important role in the biosynthesis and metabolism of major amino acids (Igarashi et al., 2006).

Spot 4

NADP-specific isocitrate dehydrogenase is an enzyme of the Krebs cycle. Citrate in the cytosol is converted to isocitrate, by the action of aconitase, and then isocitrate is converted to 2-oxoglutarate by the action of NADP-specific isocitrate dehydrogenase. The 2-oxoglutarate is a required input of carbon for amino acid biosynthesis and ammonia assimilation (Park and Kahn, 1999).

Spot 5

Metabolism of fatty acids requires β-oxidation, whereby acetyl-CoA is produced and is then processed via the glyoxylatecycle and gluconeogenesis, and 3-ketoacyl-CoA thiolase-like protein catalyzes the β-oxidation of fatty acids. Moreover, recently it was shown, in *Arabidopsis*, that this protein is also involved in the abscisic acid (ABA) Signal Transduction (Jiang et al., 2011).
Finally, 3-ketoacyl-CoA thiolase protein was already reported to be encoded by the 5A chromosome (GrainGenes database).

Spot 6

For this spot, the first result obtained is relative to an unknown protein of barley. This is common because wheat genome is not yet fully sequenced.

However, the two other results, 3-ketoacyl-CoA thiolase-like protein and the xylanase inhibitor, both described previously, were already reported to be encoded by genes on the chromosome 5A (GrainGenes database).

Spot 7

Also in this case, the most likely identification is relative to an unknown barley protein, whereas the second most likely is globulin-2 precursor.

Spot 8

The spot 8 is represented by an embryo-globulin, in addition to a predicted barley protein. Globulins are known to be the principal storage proteins in the seeds of dicot species, and to be involved in the food and respiratory allergy to wheat.

Spot 9

The spot 9 was identified as a nucleosome chromatin assembly protein. Nucleosome assembly protein is an integral component in the establishment, maintenance, and dynamics of eukaryotic chromatin. It shuttles histones into the nucleus, assembles nucleosomes, and promotes chromatin fluidity, thereby affecting the transcription of many genes (Park and Luger, 2006).

Spot 10

A predicted barley protein is the most abundant, whereas the second corresponds to a Nascent polypeptide-associated complex (NAC), likely performing a chaperone-like function (Raden and Gilmore, 1998).

Spot 11

The spot 11 was identified as Fructose-bisphosphate aldolase, which is involved in the glycolysis by catalyzing an aldol cleavage of fructose-1,6-bisphosphate to dihydroxyacetone-phosphate and glyceraldehyde 3-phosphate and a reversible aldol condensation (Wang et al., 2010).

By interrogation of the GrainGenes database, this protein is confirmed to be encoded by genes on the chromosome 5A.

Spot 12

The spot is mainly represented by the glucose and ribitol dehydrogenase, described previously (spot 7 of the cultivar Langdon, p.58).

Spot 13

In this spot, in addition to the globulin-like protein, it was identified an aldose reductase. Aldose reductases belong to the well-conserved aldo–keto reductase super family of enzymes in plants and animals. They are monomeric, cytosolic proteins that catalyze the NADPH dependent reduction of a variety of carbonyl metabolites.They are involved in stress response (Sree et al., 2000).

Spot 14

The spot 14 was identified as β-amylase. The β-amylases are water-soluble enzymes with a molecular weight of about 60 kDa. The β-amylases are responsible of the hydrolysis of (1->4)-alpha-D-glucosidic linkages in polysaccharides. They are encoded by genes on the long arms of chromosomes 4A, 4D and 5A (loci *β-Amy-1*). In the wheat kernel, they form polymers linked together by disulfide bonds (Gupta, et al., 1991) and may ling to glutenin subunits, thus contributing to glutenin polymers formation (Peruffo et al., 1996), although the amount of β-amylases seems inversely correlated to the size of glutenin macropolymers (Curioni et al., 1996).

Spot 15

The spot 15 is represented mainly by the globulin-like protein, previously described (spot 8 of the cultivar Hope p.84).

Spot 16

The spot 16 was identified as peroxidase. Peroxidases are able to utilize peroxide to oxidize a wide range of hydrogen donors including phenols, citocrome-c and nitrite. Depending on the plant tissue and isozyme/isoform, peroxidase can have others functions like oxidation of toxic reductants, biosynthesis and degradation of lignin, suberization, auxin catabolism, response to environmental stresses such as wounding,

pathogen attack and oxidative stress. Up to now, genes coding for peroxidases have been located on the chromosome 4BL, 7AS and 7BS for the wheat endosperm peroxidase; on the chromosome 3DS and 3DL for the embryo plus scutellum peroxidase and on the 1BS and 1DS for the leaf peroxidases (Bosch et al., 1987).

Spot 17

The two proteins identified in the spot 17 have been described previously (spot 13 of the cultivar Hope p.85 for the aldose reductase and spot 4 of *T. dicoccoides* p.59 for the malate dehydrogenase).

Spot 18

Also for the spot 18, identified proteins have been described previously (spots 2 & 3 of the cultivar Langdon p.58 for the LEA; spot 13 of *T. dicoccoides* p.60 for the triosephosphate isomerase ; spot 1 of the cultivar Langdon p.57 for the chitinase).

Spot 19

Proteins identified in the spot 19, α-amylase inhibitor and superoxide dismutase have been previously described (spot 14 of the *T. dicoccoides* p.60 and in the part relative to the CM-like fraction previously for the α-amylase inhibitor).

2. Identification of 5A polypeptides in the cultivar Thatcher

The proteins identified in the cultivar Thatcher are mainly involved in the binding (38%) and in protease/hydrolase activity (20%) (fig.37). Also for this cultivar, these results confirm those already described for the cultivar “Hope”, for tetraploid wheats here reported, and those reported by Vitulo et al (2011). It is also interesting to note that 28% of the proteins have a role in carbohydrate metabolism and 28% in stress / defense response.

After interrogation of the GrainGenes EST database, on a total of 45 identfied proteins, 9 proteins (20%) (Elongation factor; 3ketoacyl-CoA thiolase like protein; Phosphoenolpyruvate carboxylase; peroxidase; Fructose-biphosphate aldolase; Lipoprotein; adenylate kinase A; xylanase inhibitor; carboxypeptidase D) resulted to be encoded by genes on the chromosome 5A.

Among all the identified proteins, 12 proteins (26%) have been found in two or more spots.

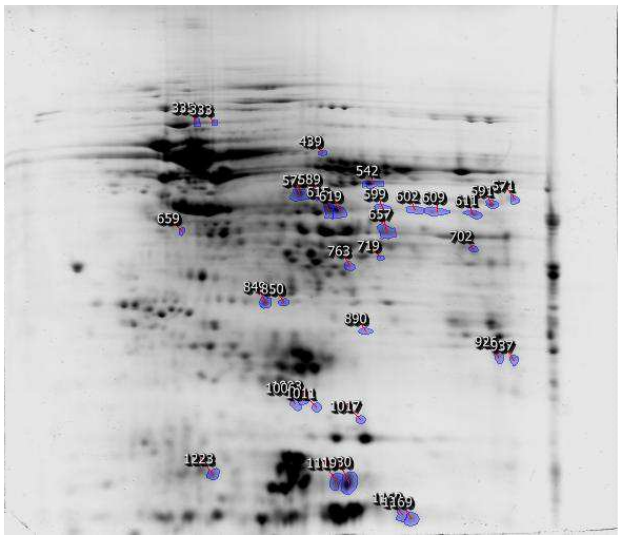


Figure 33: 5A specific spots found in *T. aestivum* cv Thatcher

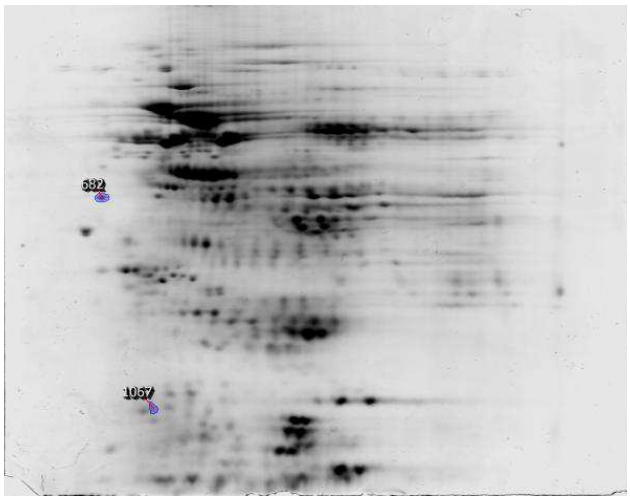


Figure 34: 5A specific spots found in *T. aestivum* cv Chinese Spring

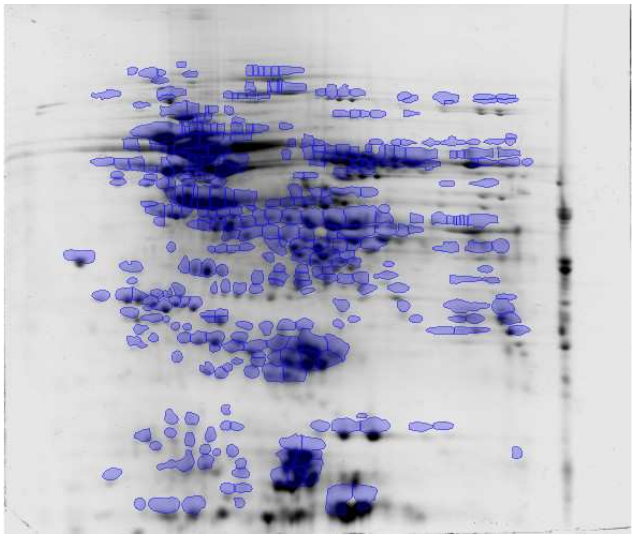


Figure 35: Common spots between *T.aestivum* cv Chinese Spring and *T.aestivum* cv Thatcher

Table 16: Identified polypeptides by LC MS/MS in the bread wheat *T. aestivum* cv Thatcher

N° spot		Score MASCOT	N° peptides	emPAI
1	Dihydrolipoamide dehydrogenase precursor	1231	18	**23.61**
	Beta amylase	798	11	6.83
	T-complex protein	537	9	6.25
2	Globulin-like protein	1058	12	2.81
	DNA-binding protein GBP16	524	15	**3.37**
3	Triticin	1542	8	**1.59**
4	Elongation factor 1-gamma 2	1023	17	**6.3**
	Triticin	652	8	1.44
	Serpin	311	10	1.4
5	Aspartate aminotransferase	878	16	**6.51**
	Alcohol dehydrogenase class 3	597	10	3.41
	NADP-specific isocitrate dehydrogenase	592	12	3.52
6	HSP70 precursor	720	14	**8.1**
	Serpin	276	8	1.05
7	3-ketoacyl-CoA thiolase-like protein	568	10	**2.32**
8	Phosphoenolpyruvate carboxylase	172	4	**0.14**
9	rRNA-N-glycosidase	227	5	**0.95**
	Peroxidase 1	179	6	0.69
10	Aspartate aminotransferase	1435	23	**21.4**
	Glyceraldehyde-3-phosphate dehydrogenase	462	11	2.27
11	Glyceraldehyde-3-phosphate dehydrogenase	736	15	**5.2**
	Fructose-bisphosphate aldolase	412	14	2.36
	Malate dehydrogenase	354	10	1.87
12	Malate dehydrogenase	309	5	**1.72**
	Fructose-bisphosphate aldolase	327	7	1.58
13	Peroxidase 1	470	11	2.72
	Aldose reductase	302	10	**2.96**
	NADP-specific isocitrate dehydrogenase	137	4	0.43
14	Lipoprotein-like	697	8	**23.89**
	(2R)-phospho-3-sulfolactate synthase-like	478	7	2.75
15	Adenylate kinase A	227	8	**1.98**
16	Xylanase inhibitor XIP-III	1005	13	**4.95**
17	Triosephosphate isomerase	200	5	**2.44**
18	LEA 1	389	6	**3.13**
	Elongation factor 1-alpha	122	5	1.03
19	Carboxypeptidase D	95	4	**3.23**
20	Ribosomal protein L18	231	5	1.4
	Endogenous alpha-amylase/subtilisin inhibitor	229	5	**1.43**
21	Endogenous alpha-amylase/subtilisin inhibitor	650	11	**18.4**
	Globulin 1	189	5	1.23
	Carboxypeptidase D	97	3	0.49
22	Alpha-2-purothionin precursor	363	6	**2.24**
23	USP family protein	203	4	**3.39**
	Peptidyl-prolyl cis-trans isomerase	137	4	1.13
	Glyceraldehyde-3-phosphate dehydrogenase	164	4	1.39

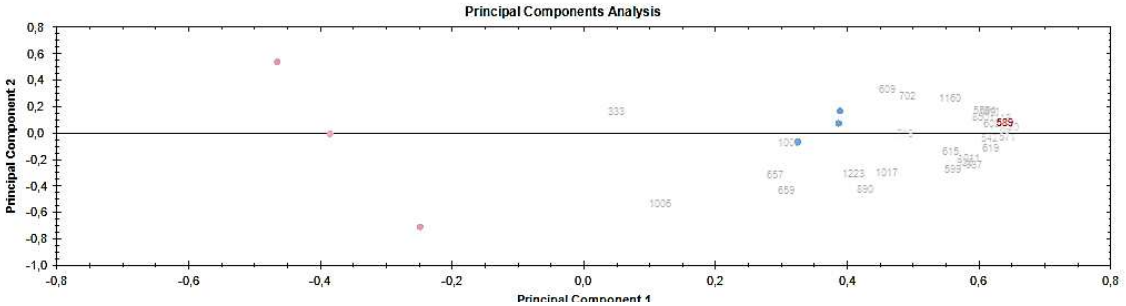


Figure 36: PCA representation in which the contribution of each spot, relative to the genotypes analysed is reported. Blue points represent the three replicates of CS-Thatcher5A and the red point that of CS.

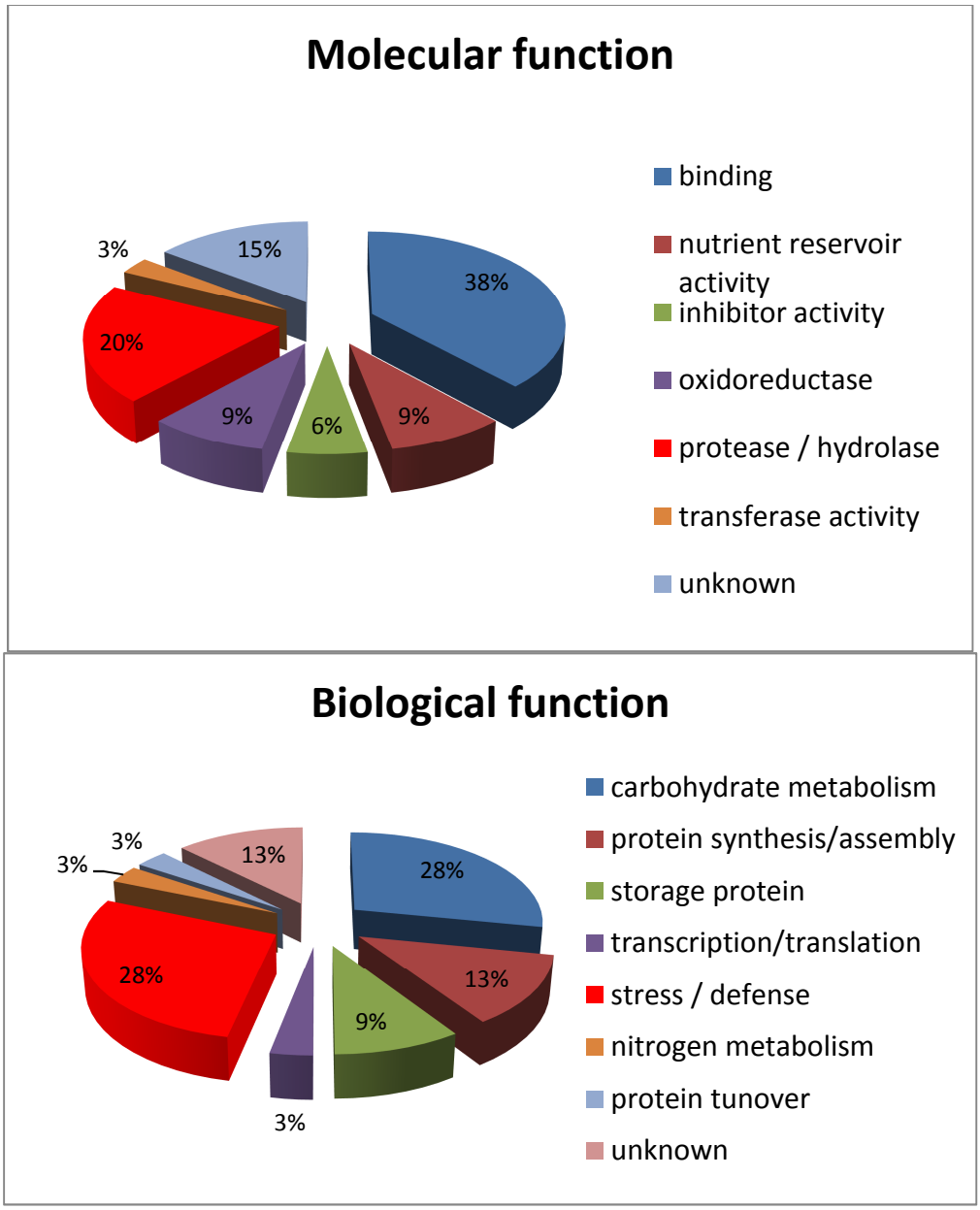


Figure 37: Percentage distribution of the biological processes and the molecular function relative to the identified proteins in "Thatcher" obtained with gene ontology (http://www.uniprot.org/uniprot/) and by literature research

Spot 1

The spot 1 is represented by the Dihydrolipoamide dehydrogenase. It is a flavoprotein which is a partner of pyruvate dehydrogenase in the pyruvate dehydrogenase complex which transforms pyruvate into acetyl-CoA and links cytosolic glycolytic metabolism with the tricarboxylic acid cycle (Manaa et al., 2011).

Spot 2

The spot is mainly represented by the DNA-binding protein GBP16. GBP16 is a member of a DNA-binding protein complex that specifically binds the single-stranded G-rich telomere sequence (Casati et al., 2005).

Spot 3

This spot was identified as a Triticin, which was described previously (spot 4 of Langdon p.58).

Spot 4

It is mainly represented by an Elongation factor, already described (spot 2 of Hope p.83).

Spot 5

The spot 5 is represented by the aspartate aminotransferase. Plant aspartate aminotransferase (AAT) catalyses the reversible transamination reaction between L-aspartate and 2-oxoglutarate to give oxaloacetate and L-glutamate (Maciga and Paszkowski, 2004).

Spot 6

The spot 6 was identified as HSP 70 precursor. HSPs have been known to protect cells against deleterious effects of stress and they have essential functions under no stressful conditions. It has been shown that all Hsps have structural and functional properties, and that they bind to ATP and to unfolded or partially denatured polypeptides (rev. in EFEOĞLU, 2009).

Spot 7

It is represented by the 3-ketoacyl-CoA thiolase-like protein, which was described previously (spot 5 of Hope p.83).

Spot 8

The spot 8 was identified as a phosphoenolpyruvate carboxylase. Phosphoenolpyruvate carboxylase is a ubiquitous cytosolic enzyme in higher plants and is also widely distributed in bacteria, cyanobacteria, and green algae. It catalyzes the irreversible β-carboxylation of phosphoenolpyruvate (PEP) in the presence of HCO_3^- and Me^{2+} to yield oxaloacetate (OAA) and Pi, and thus is involved intimately in C4-dicarboxylic acid metabolism in plants (Chollet et al., 1996).
The GrainGenes database interrogation reveals that the PEPC was already known to be encoded by genes on the chromosome 5A.

Spot 9

The spot 9 was identified as rRNA-glycosidase. They belong to the ribosome-inactivating family, which are widely distributed in higher plants (Massiah and Hartley, 1995).

Spot 10

This spot was identified as aspartate aminotransferase, described previously (spot 5 p.91).

Spot 11

The spot 11 is mainly represented by the glyceraldehyde-3-phosphate dehydrogenase, which was found and described in the spot 4 of *T. dicoccoides* p.59.

Spot 12

This spot was identified as a malate dehydrogenase, found and described in the spots 4 & 5 of *T. dicoccoides* p.59.

Spot 13

This spot is mainly represented by the peroxidase 1. This protein was described in the spot 16 of Hope p.85-86.

Spot 14

The spot 14 is represented mainly by the Lipoprotein-like. Some lipoproteins are involved in the formation of cell membranes and hydrophobic layer, while some of them play an important role in the transport of fatty acids or thir CoA derivatives, such as lipid transfer protein thionins (Deng et al., 2009).

Spot 15

The spot 15 was identified as Adenylate kinase A, which was already found and described (spot 1 of the CM-like fraction p.67).

Spot 16

This spot was identified as xylanase inhibitor XIP-III, which was previously described in the spot 8 & 9 of Langdon p.58-59.

Spot 17

The spot 17 was identified as triosephosphate isomerase, previously described in the spot 13 of *T. dicoccoides* p.60.

Spot 18

This spot is mainly represented by the Late Embryogenesis Abundant protein. This protein was described in spots 2 & 3 of the cultivar Langdon p.58.

Spot 19

The spot 19 was identified as Carboxypeptidase D. The carboxypeptidase is characterized by a broad substrate specificity and esterase and/or amidase activity, in addition to the intrinsic carboxypeptidase activity (Drzymała and Bielawski, 2009).

Spot 20

The spot 20 is mainly represented by the ribosomal protein L18. Ribosomal proteins are major components of ribosomes, and are regulated both developmentally and environmentally in plants (Yingyin et al., 2006).

Spot 21

The spot 21 is mainly represented by an endogenous α-amylase / subtilisin inhibitor. This protein family is described in the part relative to the CM-like fraction (p.67).

Spot 22

It contains an alpha-2-purothionin precursor that is involved in plant defence mechanisms.

Spot 23 and Spot 24:

The spot 23 and 24 are mainly represented by the USP family protein. Ubiquitin-Specific Proteases (USPs) are a family of unique hydrolases that specifically remove polypeptides covalently linked via peptide or isopeptide bonds to the C-terminal glycine of ubiquitin (Yan et al., 2000).

3. Identification of 5A polypeptides in the cultivar Timstein

The set of image analysis (CS vs Timstein; CS vs CS-Timstein5A; Hope vs CS-Timstein5A) revealed 13 spots to be polypeptides encoded by genes on the chromosome 5A of Timstein. The PCA (fig.41) confirmed the specificity of the 5A spots selected.

The Gene Ontology reveals that 31% of the identified proteins are involved in the carbohydrate metabolism, 23% in protein synthesis/assembly and 15% in transcription/translation. Also for this cultivar their molecular function is mainly “binding” (50%) (Fig.42).

In this cultivar, among the identified proteins, 6 of the 22 proteins (27%) were already known to be localized on the chromosome 5A by interrogation on GrainGenes database. These proteins are the β-D-glucan exohydrolase, the Hydroxyproline-rich glycoprotein, the pyruvate kinase, the alcohol dehydrogenase, the elongation factor and the peroxidase.

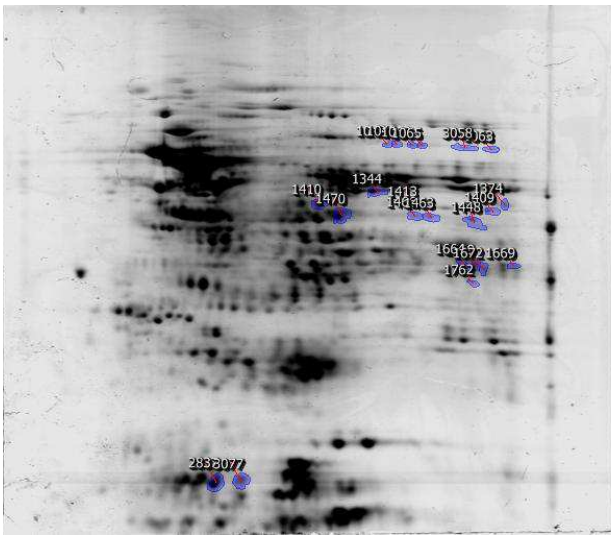


Figure 38: 5A specific spots of *T. aestivum* cv Timstein

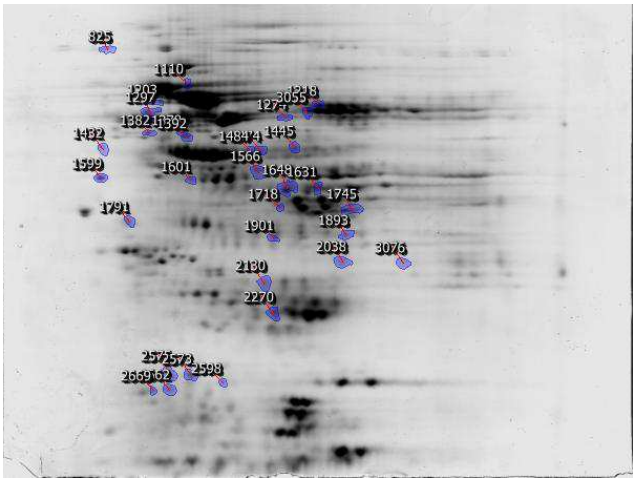


Figure 39: 5A specific spots of T. aestivum cv Chinese Spring

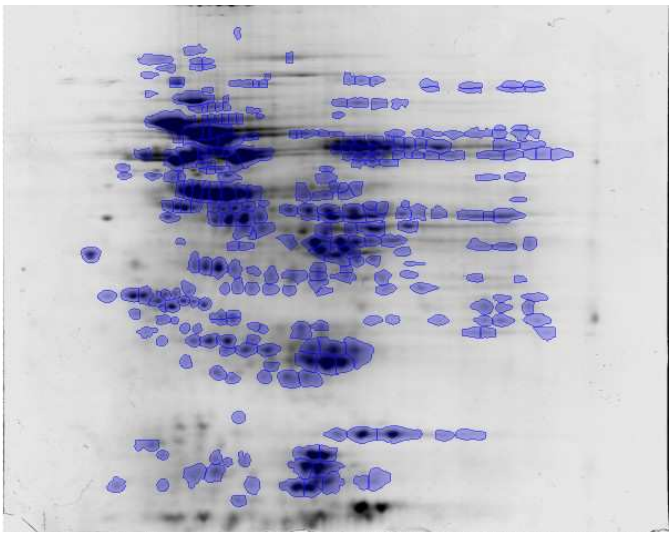


Figure 40: Common spots between *T.aestivum* cv Chinese Spring and *T.aestivum* cv Timstein

Table 17: Identified polypeptides by LC MS/MS in the bread wheat *T. aestivum* cv Timstein

N° spot		Score MASCOT	N° peptides	emPAI
1	Embryo globulin	230	11	**1.35**
	Hydroxyproline-rich glycoprotein	224	9	1.3
	Beta-D-glucan exohydrolase	197	7	0.44
2	Embryo globulin	961	17	**4.18**
	Beta-D-glucan exohydrolase	463	13	1.2
3	Embryo globulin	816	22	**9.82**
4	Beta-D-glucan exohydrolase	419	12	**0.88**
	Cytosolic NADP malic enzyme	299	10	0.81
5	Embryo globulin	313	12	**1.21**
6	SAR DNA binding protein	120	3	0.2
	Cytosolic NADP malic enzyme	115	4	**0.24**
7	SAR DNA binding protein	224	5	0.35
	Pyruvate kinase	139	4	**0.38**
8	Globulin-like protein	1047	14	**4.82**
	DNA-binding protein GBP16	420	14	2.15
9	Alcohol dehydrogenase ADH1A	111	6	**0.49**
10	Aspartate aminotransferase	506	12	**2.84**
	Elongation factor 1-gamma 2	498	12	2.11
11	Peroxidase 1	237	6	**1.58**
12	17.6kDa heat-shock protein	264	6	**2.38**
	Peptidyl-prolyl cis-trans isomerase	256	4	1.13
13	Small heat shock protein HSP17.8	204	8	**3.96**

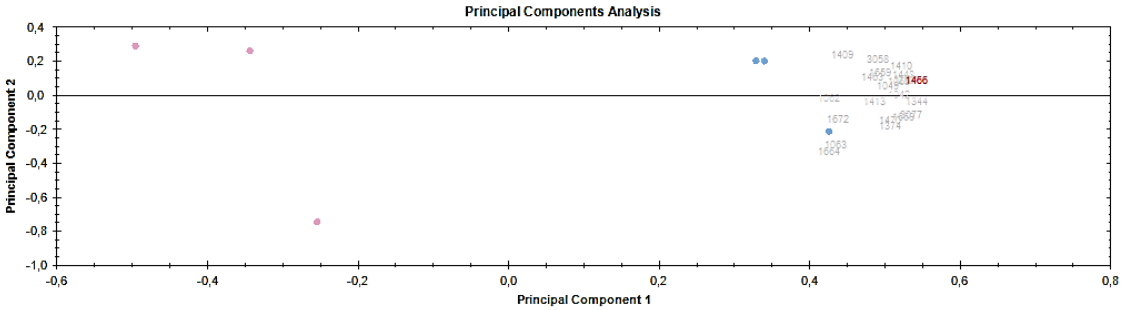
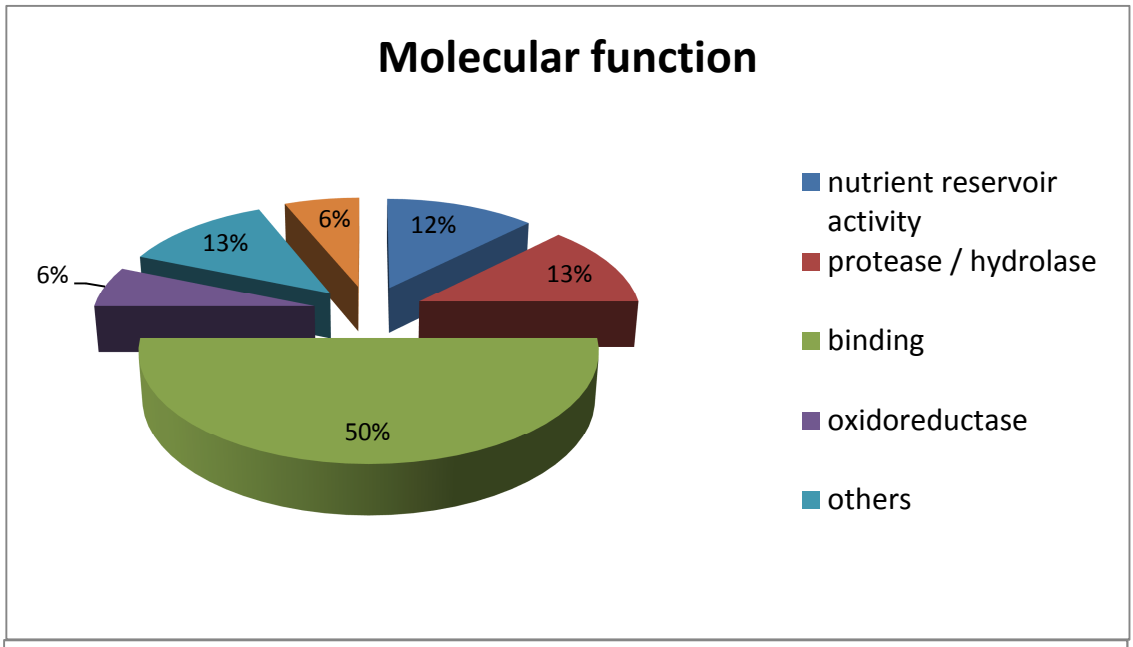


Figure 41: PCA representation in which the contribution of each spot is reported. Blue points represent the three replicates of CS-Timstein5A and the red points that of CS.



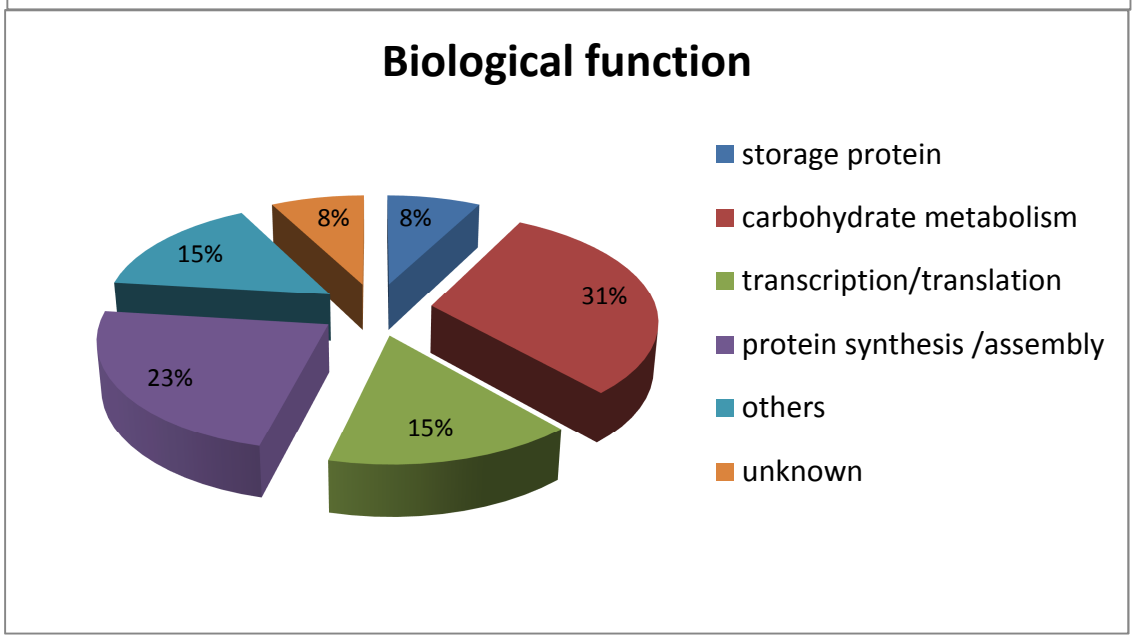


Figure 42: Percentage distribution of the biological process and the molecular function for the identified proteins in "Timstein" obtained with gene ontology (http://www.uniprot.org/uniprot/) and by literature research

Spot 1& 2& 3

These spot are mainly represented by the embryo globulin.Globulin are described in the spot 8 of the cultivar Hope (p.84).

Spot 4:

The spot 4 is represented by the β-d-glucan exohydrolase. β-d-glucan exohydrolase is involved, with two others hydrolases, in the depolymerization of (1→3),(1→4)-β-d-glucans in germinated grain or in the partial hydrolysis of the polysaccharide in elongating vegetative tissues. The (1→3),(1→4)-β-d-Glucans represent an important component of cell walls in the Poaceae family of higher plants, and are an important source of stored glucose for the developing seedling (Hrmova and Fincher, 2001).
In the GrainGenes database, ESTs of the β-d-glucan exohydrolase are found on the chromosome 5A.

Spot 5

Also this spot was identified as an embryo globulin, described in the spot 8 of Hope (p.84).

Spot 6

This spot is mainly represented by an NADP malic enzyme. The enzyme acts in many different metabolic pathways in plants (Casati et al., 1997).

Spot 7

The spot 7 is mainly represented by the pyruvate kinase. Pyruvate kinase (PK) is an important enzyme of glycolytic pathway that also functions in providing carbon skeleton for fatty acid biosynthesis (Ambasht and Kayastha, 2002).

Spot 8

The spot 8 is mainly represented by a globulin-like protein, which are described in the spot 8 of Hope (p.84).

Spot 9

The spot 9 was identified as alcohol dehydrogenase ADH1A. This protein was described in the spot 6 of Langdon (p.58).

Spot 10

The spot 10 was found to be mainly represented by the aspartate aminotransferase. This protein was found and described in the spot 5 of Thatcher (p.91).

Spot 11

The spot 11 was identified as peroxidase, which was already described in the spot 16 of Hope (p.85-86).

Spot 12 & Spot 13

This two spots are mainly represented by Heat Shock Protein (17.6kDa and 17.8kDa respectively). The HSPs have been described in the spot 6 of Thatcher (p.91).

4. Identification of 5A polypeptides in the cultivar Cheyenne

The set of image analysis (CS vs CNN; CNN vs CNN-CS5A; CNN vs CNN-CS5A) revealed 18 spots to be polypeptides encoded by genes on the chromosome 5A of Cheyenne. The PCA (fig.46) confirmed the specificity of the 5A spots selected.

Twenty-seven polypeptides were identified in the cultivar Cheyenne, as encoded by genes at chromosome 5A. They are mainly implicated in the metabolism of the carbohydrate (37%) and in the response to stress (23%). Their molecular function is binding for 30% of them, and 17% have protease/hydrolase functions (fig.47).

The ubiquitin-like modifier-activating enzyme 5, the enolase, the glucose-6-phosphate isomerase, the β-D-glucan exohydrolase, the β-glucosidase, the serpin, the ATP synthase, the chitinase and the glycine-rich RNA binding protein were already known to be encoded by genes on the chromosome 5A (GrainGenes EST database). In spot 12, the protein identified is a xylanase inhibitor, whose genes were already indicated as present on chromosome 5A. Moreover, the Calreticulin-like protein was reported on the chromosome 5B and 5D, and thus, on the basis oh chromosome homeology, it is likely that genes encoding for this protein are present in the chromosome 5A.

For this cultivar, only one protein (caleosin) was found in more than one spot.

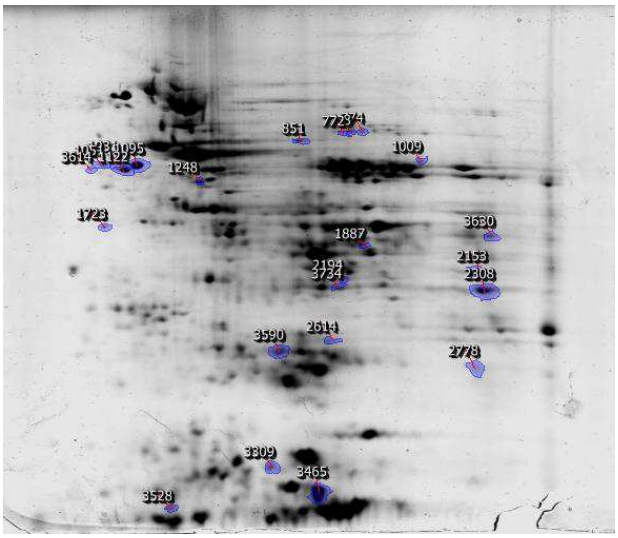


Figure 43: 5A specific spots for *T.aestivum* cv Cheyenne

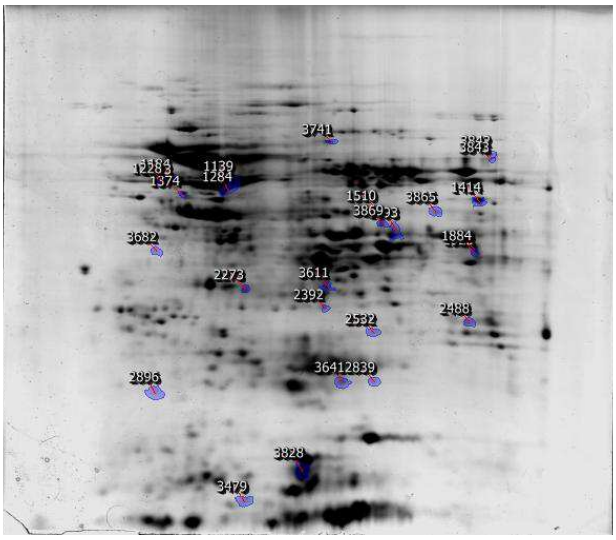


Figure 44: 5A specific spots for *T.aestivum* cv Chinese Spring

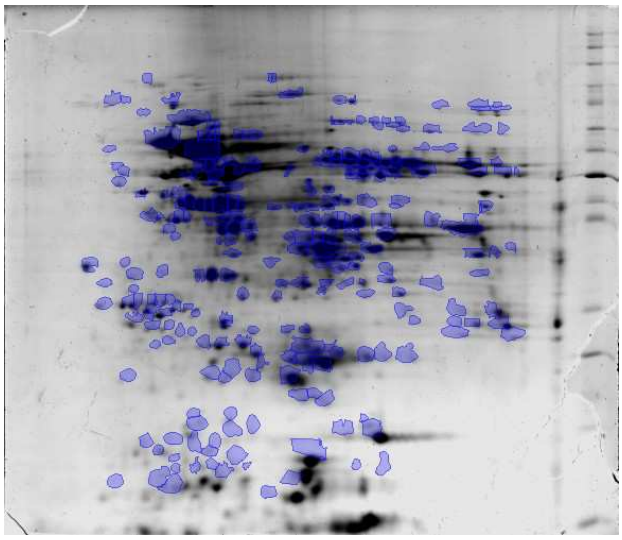


Figure 45: Common spots between *T.aestivum* cv Cheyenne and *T.aestivum* cv Chinese Spring

Table 18: Identified polypeptides by LC MS/MS in the bread wheat *T. aestivum* cv Cheyenne

N° spot		Score MASCOT	N° peptides	emPAI
1	Calcium-binding protein precursor	290	6	**0.53**
	Calreticulin-like protein	69	3	0.24
2	Ubiquitin like protein	575	7	**2.91**
	Enolase	311	9	0.99
3	Glucose-6-phosphate isomerase	156	4	**0.23**
4	UTP--glucose-1-phosphate uridylyltransferase	379	14	**1.83**
	Glucose-1-phosphate adenylyltransferase	371	12	1.78
5&6	Beta-D-glucan exohydrolase	269	7	0.52
	Beta-D-xylosidase	217	7	**0.53**
7	3-ketoacyl-coA thiolase like	568	10	**2.32**
8	Beta-glucosidase	414	9	**1.28**
9	Serpin	127	6	**0.77**
10	Glyceraldehyde-3-phosphate dehydrogenase	587	12	**7**
	Fructose-bisphosphate aldolase cytoplasmic isozyme	385	9	3.36
11	Glucose and ribitol dehydrogenase	1166	19	**17.96**
	Caleosin 1	463	8	2.9
	Vacuolar ATP synthase subunit E	217	8	1.49
12	Xylanase inhibitor	111	4	**0.53**
13	26 kDa endochitinase 1 precursor	265	4	0.96
	Xylanase inhibitor protein 1 precursor	186	7	1.23
	Aspartic proteinase	149	5	**1.25**
14	Triosephosphate isomerase	487	7	2.23
	1-Cys peroxiredoxin	476	8	**3.16**
	Caleosin 2	404	7	2.31
15	Thaumatin-like protein	388	10	**2.73**
	Manganese superoxide dismutase	113	3	0.48
16	16.9 kDa class I heat shock protei	382	6	**1.8**
17	USP family protein	307	6	**3.39**
18	Glycine-rich RNA-binding protein	407	6	**2.5**

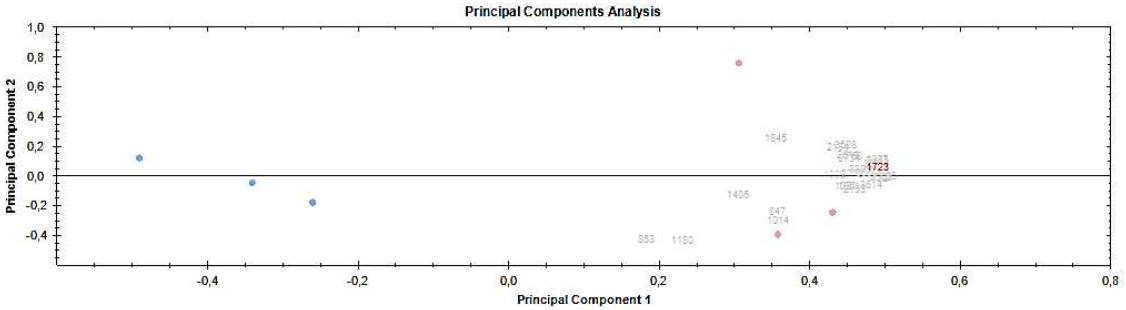
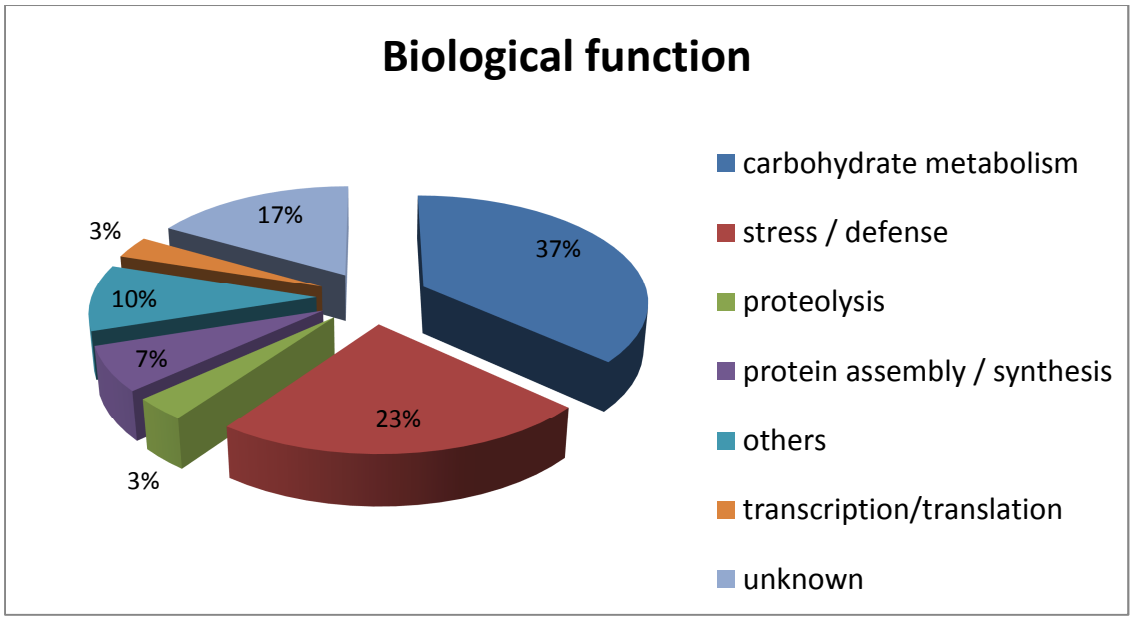


Figure 46: PCA representation in which the contribution of each spot is reported. Blue points represent the three replicates of CNN-CS5A and the red point that of CNN.



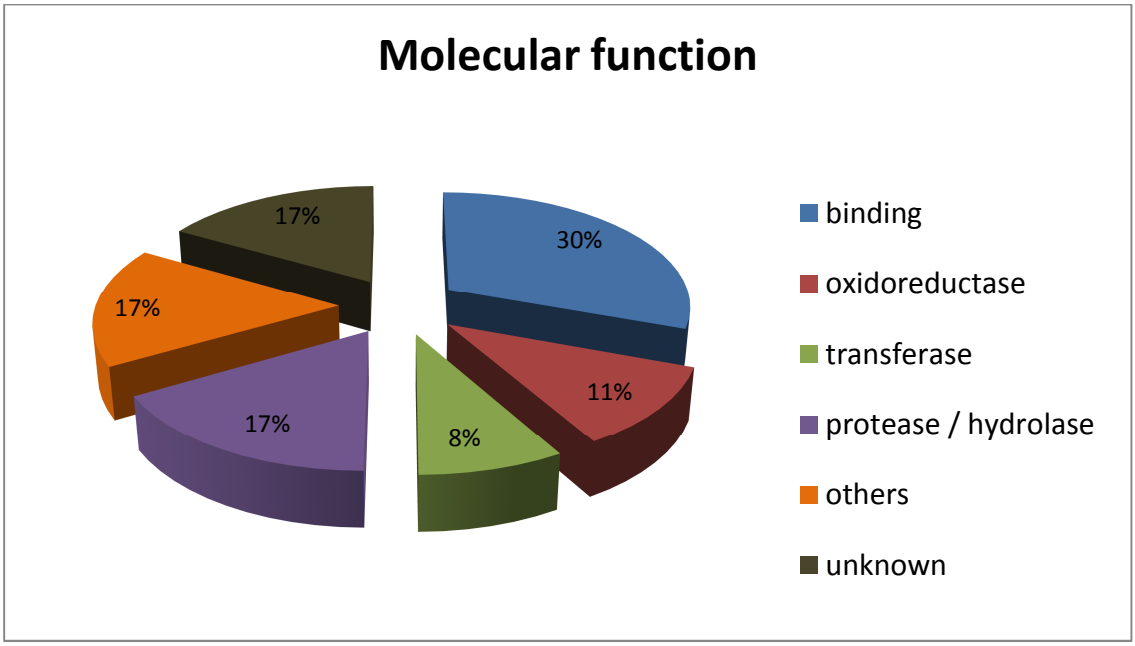


Figure 47: Percentage distribution of the biological process and the molecular function for the identified proteins in "Cheyenne" obtained with the gene ontology (http://www.uniprot.org/uniprot/) and by literature research.

Spot1

The spot 1 is mainly represented by a calcium-binding protein precursor. Calcium is required for many vital processes in fungi and plants. High levels of calcium are found in cell walls, vacuoles, and most organelles, while very low levels of calcium are present in the cytosol of plant cells. Calcium-binding proteins are involved in the events that accompany the action of calcium as a second messenger (Moreau, 1987).

Spot 2

The spot 2 was identified as ubiquitin-like-modifier-activating enzyme 5. In literature no informations about this specific protein are available, but information about ubiquitin-like modifier reveal is a member of the superfamily of ubiquitin-like polypeptides that become covalently attached to various intracellular target proteins as a way to alter their function, location, and/or half-life (Kurepa, 2002).

Spot 3:

The spot 3 was identified as glucose-6-phosphate isomerase (GPI). GPI catalyze reversible aldose-ketose reactions involving a histidine, a lysine, a glutamic acid residue and probably an arginine residue at the catalytic center.

Spot 4

The spot 4 is mainly represented by glucose-1-phosphate adenylyltransferase, which was also found and described in the spot 1 of Hope (p.82).

Spot 5 and Spot 6

These three spots are mainly represented by the beta-D-glucan exohydrolase, which was previously described in the spot 4 of Timstein (p.99).

Spot 7

The 3-ketoacyl-coA thiolase like identified in this spot was already found and described in the spot 5 of Hope (p.83).

Spot 8

The spot 8 was identified as β-glucosidase. In plants, β-D-glucosidases are involved in various functions, including lignification, regulation of the biological activity of

cytokinins, control of the biosynthesis of indole-3-acetic acid, and chemical defense against pathogens and herbivores (Sue et al., 2006).

Spot 9

The spot 9 was identified as serpin. Serpins constitute a large family of related proteins, the majority of which are Serine Protease Inhibitors. They are members of a large family of proteins that are structurally closely related, yet functionally diverse, most of which regulate proteolysis. They are also known to be allergens (Wu et al., 2012).

Spot 10

The spot 10 is mainly represented by the Glyceraldehyde-3-phosphate dehydrogenase. This protein was also identified and described in the spot 4 of *T. dicoccoides* (p.59).

Spot 11

This spot is mainly represented by the glucose and ribitol dehydrogenase, which was previously described (spot 7 of the cultivar Langdon, p.58).

Spot 12

The spot 12 was identified as a xylanase inhibitor, also find in the spots 8 & 9 of Langdon (p.58-59).

Spot 13

The spot 13 is mainly represented by the aspartic proteinase. Aspartic proteinases are widely distributed among plant species. The biological role of plant aspartic proteinases is not completely established. In general, plant APs have been implicated in protein processing and/or degradation in different plant organs, as well as in plant senescence, stress responses, programmed cell death and reproduction (rev. Simoes and Faro, 2004).

Spot 14

The spot 14 is mainly represented by the 1-Cys peroxiredoxin. Peroxiredoxins are thiol-based peroxidases. Peroxiredoxins are antioxidative enzymes that catalyze the reduction of alkyl hydroperoxides to alcohols and hydrogen peroxide to water. 1-Cys peroxiredoxins perform important roles during late seed development in plants (Kim et al., 2011).

Spot 15

The spot 15 was identified as thaumatin-like protein (TLP). They have antifungal activity and thus are involved in plant defense. TLP were recently identified wheat flour salt-soluble protein family to be associated with baker's respiratory allergy (Salcedo et al., 2011).

Spot 16

The spot 16 was identified as a 16.9kDa Heat Shock Protein. Heat Shock Protein have been found and described in the spot 6 of Thatcher (p.91).

Spot 17

It was identified as USP family protein. This protein was previously described (spot 23 of Thatcher p.94).

Spot 18

The spot 18 was identified as a Glycine-rich RNA-binding protein. Glycine-rich RNA-binding proteins (GR-RBPs) have been implicated to play roles in post-transcriptional regulation of gene expression in plants under various stress conditions. However, the functional roles of GR-RBPs in plant response to environmental stresses are largely unknown. It was been shown that glycine-rich RNA-binding protein contributes to the enhancement of freezing tolerance in *Arabidopsis thaliana* (Kim et al., 2005).

5. Identification of 5A polypeptides in the cultivar Chinese Spring

For the cultivar Chinese Spring, the attribution of polypeptides to chromosome 5A has been performed by taking into consideration the results obtained on the basis of each comparison with the different chromosomal substitution lines. Twelve spots have been selected and identified (fig.48).

In total, 25 proteins have been identified. Even if for 22% of them the gene ontology do not give any information on their biological function, the majority are involved in the carbohydrate metabolism (31%), which is consisten with results obtained on the others cultivars. Also for Chinese Spring, the identified proteins have mainly a binding function (33%) or a protease /hydrolase function (17%).

After GrainGenes EST database interrogation among these 25 identifications 6 (enolase, β-glucosidase; fructose-bisphosphate aldolase; xylanase inhibitor XIP-III; GTP-binding protein; ADP-ribosylation factor) have been confirmed to be encoded by 5A chromosome.

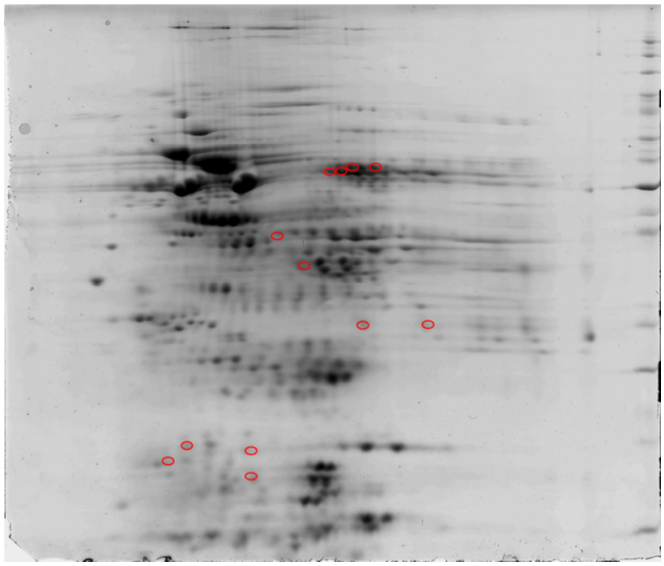
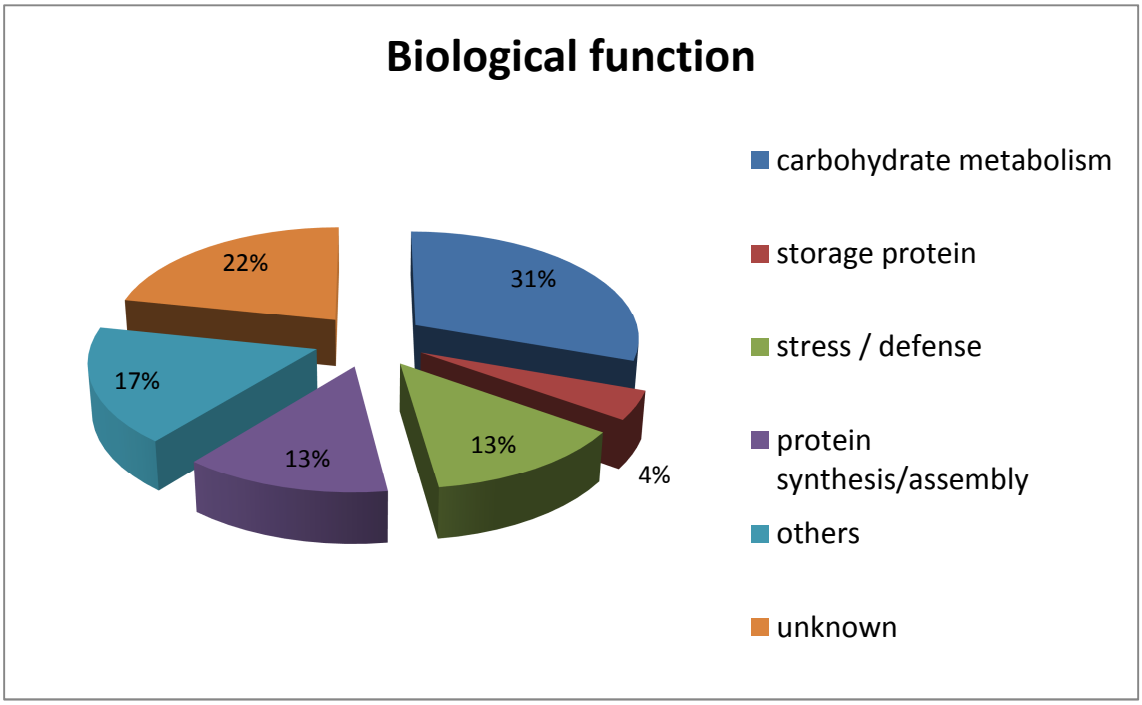


Figure 48: 5A identified spots in *T.aesitvum* cv Chinese Spring

Table 19: Identified polypeptides by LC MS/MS in the bread wheat *T. aestivum* cv Chinese Spring

N° spot		Score MASCOT	N° peptides	emPAI
1	Globulin-2 precursor	1137	22	**6.63**
	Enolase	385	12	1.28
	ATP synthase alpha subunit	313	11	1.04
2	Beta-glucosidase	337	8	**1.14**
3	Endo-1,4-beta-glucanase	350	5	0.81
	Adenosylhomocysteinase	292	12	**1.15**
4	Fasciclin-like protein FLA31	179	3	**0.57**
	Methylmalonate semi-aldehyde dehydrogenase	180	4	**0.28**
5	Glyceraldehyde-3-phosphate dehydrogenase	1868	12	**6.41**
	Phosphoglycerate kinase	587	16	3.57
6	R40g2 protein	350	10	**1.85**
	beta amylase	213	5	0.75
	Fructose-bisphosphate aldolase	206	6	0.85
7	Globulin 1	761	7	**6.44**
	Xylanase inhibitor XIP-III	250	8	1.57
8	Globulin 1	759	10	**7.51**
	Ribosomal protein S8	183	5	0.87
9	Alpha-2-purothionin precursor	621	4	**4.83**
	Peroxiredoxin	445	6	2.9
10	16.9 kDa class I heat shock protein	268	5	**2.59**
	18,9 kDa ABA-induced protein	251	4	2.5
	GTP-binding protein	160	6	1.4
11	Purothionin A-1 precursor	412	4	**7.1**
	Cold regulated protein	284	6	2.52
12	17.4 kDa class I heat shock protein 3	473	9	**7.49**
	Peptidyl-prolyl cis-trans isomerase	312	7	3.89
	ADP-ribosylation factor	144	4	0.92



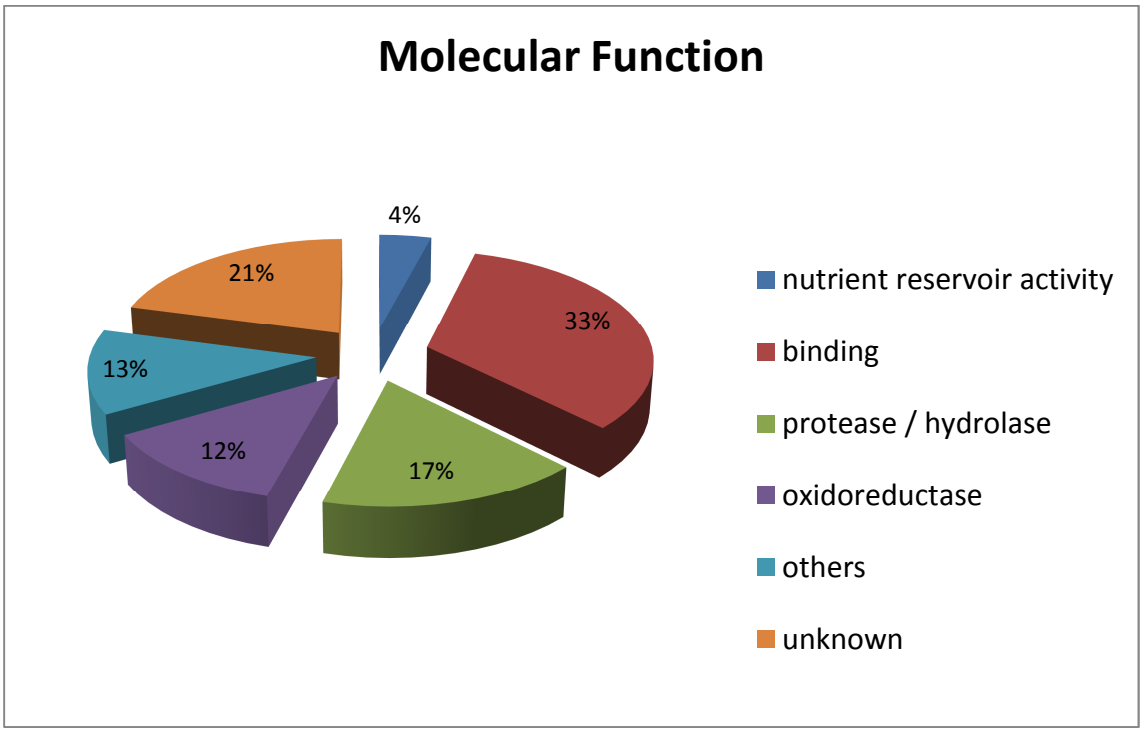


Figure 49: Percentage distribution of the biological process and the molecular function for the identified proteins in "Chinese Spring" obtained with the gene ontology of http://www.uniprot.org/uniprot/ and by literature research.

Spot 1

The spot 1 is mainly represented by the globulin-2 precursor. Globulins are described in the spot 3 of Langdon p.58.

Spot 2

The spot 2 was identified as Beta-glucosidase. This protein was previously found and described in the spot 8 of Cheyenne (p.105-106).

Spot 3

The spot 3 is mainly represented by the Adenosylhomocysteinase, also called S-Adenosyl-l-homocysteine hydrolase. S-Adenosyl-l-homocysteine hydrolase is one of the most highly conserved enzymes from bacteria to mammals. It plays a key role in the regulation of transmethylation reactions in all eukaryotic organisms Targets of AdoMet-dependent methyltransferases include a wide spectrum of cellular compounds, such as DNA, mRNA, histones H3 and H4, and other proteins as well as smaller metabolites, including lipids (Malanovic et al., 2008).

Spot 4

The spot 4 is mainly represented by the Fasciclin-like protein FLA31. The fasciclin-like arabinogalactan-proteins (FLAs) are a class of chimeric Arabinogalactan proteins (AGPs) which contain one or two AGP-like domains (rich in noncontinuous Pro residues) and one or two fasciclin-like domains besides the three features of the classical AGPs. AGPs comprise a family of hydroxyproline-rich glycoproteins that are implicated in plant growth and development (Liu et al., 2008).

Spot 5

This spot is mainly represented by the glyceraldehyde-3-phosphate dehydrogenase, which was found and described in the spot 4 of *T. dicoccoides* p.59.

Spot 6

The spot 6 is mainly represented by a protein similar to the R40g2 protein of Oryza sativa. This protein is reported to be ABA-inducible.

Spot 7 & 8

This spot is mainly represented by a globulin 1. Globulins are described in the spot 8 of the cultivar Hope (p.84).

Spot 9

The spot 9 is mainly represented by the alpha-2-purothionin precursor that is involved in plant defence mechanisms.

Spot 10 & 12

These two spot are mainly represented by a 16.9kDa and a 17.4kDa Heat shock protein respectively, previously described in the spot 6 of Thatcher (p.91).

Spot 11

The spot 11 is mainly represented by a purothionin, which are involved in plant defence mechanisms.

6. Comparison between various bread wheat cultivars

In total in the bread wheat, 86 proteins have been identified in five different cultivars. Those proteins have mainly a role of storage protein (26%), and are also involved in the carbohydrate metabolism (20%) and in stress /defense processes (19%). These 86 proteins have predominantly a function of binding (37%) and to lesser extent of protease /hydrolase (9%) and of oxidoreductase (9%) (fig.50).

Among these 86 proteins, 32 (37%) have been found in two or more cultivars (tab.20). Moreover, 22 of these proteins are common between two cultivars, 8 between 3 cultivars and only two (Fructose-bisphosphate aldolase and a inhibitor) are common of 4 cultivars. Thatcher and Hope are the two cultivars which share the more protein (7 proteins) and they also have 5 others proteins in common with Timstein (fig.51).

Among these proteins in common, the Elongation factor, the alcohol dehydrogenase, the 3-ketoacyl-CoA thiolase-like protein, the xylanase inhibitor, the beta-amylase, the peroxidase, the DNA-binding protein, the serpin, the β-D-glucan exohydrolase, the enolase and β-glucosidase were already known to be encoded by genes on the chromosome 5A (according to GrainGenes). The same database indicated that the globulin-2 precursor, the globulin-like protein, the Malate dehydrogenase and the Globulin 1 are encoded by genes on the chromosome 5B and/or 5D. Because of the homeology among these chromosomes, it is likely that they are encoded by genes on chromosome 5A as well.

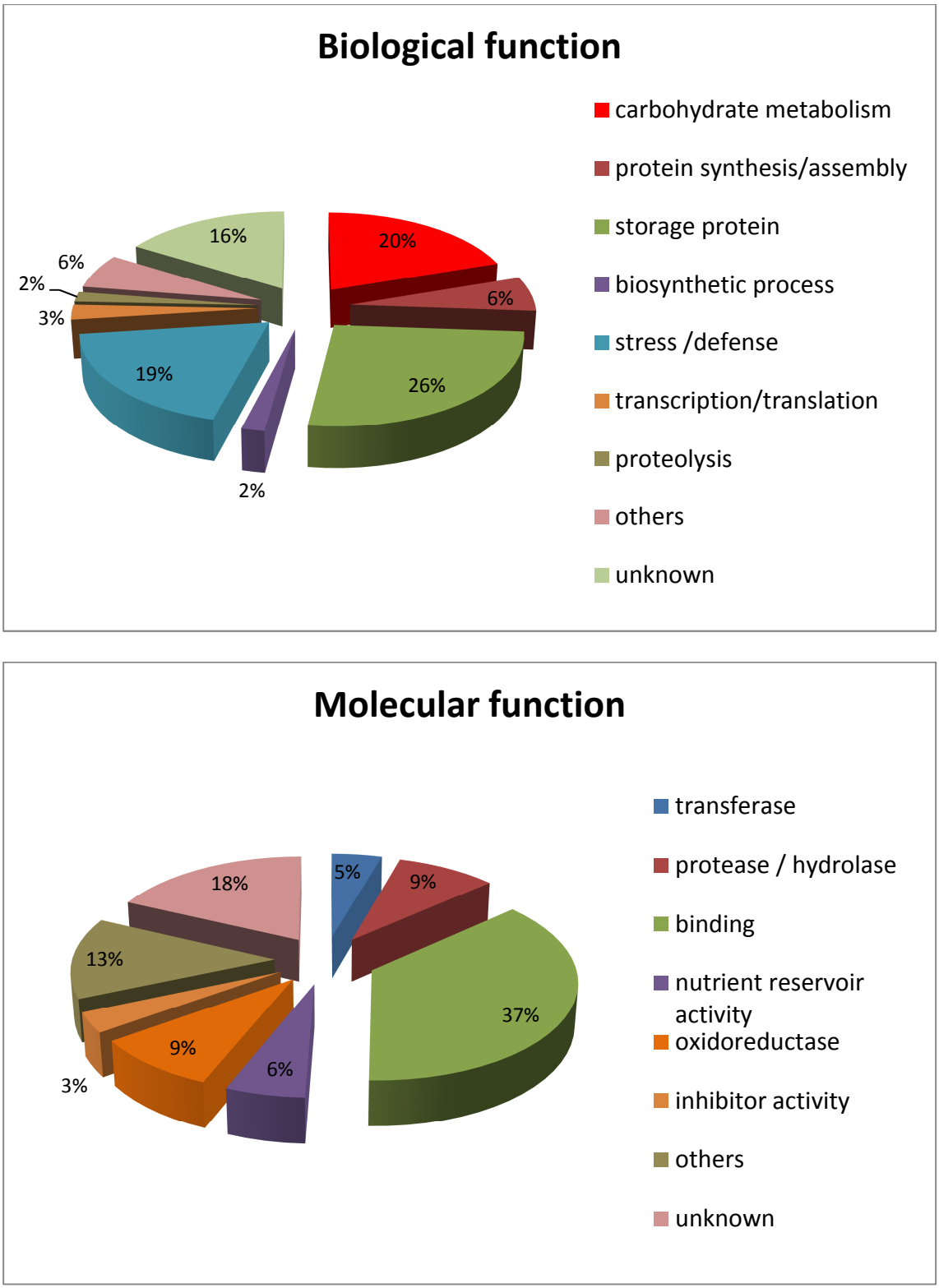


Figure 50: Percentage distribution of the biological process and the molecular function for the common identified proteins in bread wheat obtained with the gene ontology of http://www.uniprot.org/uniprot/ **and by literature research.**

Table 20: Identified proteins commons between two or more bread wheat cultivars.

Protein	Cultivars
Aspartic proteinase	Hope, Cheyenne
Elongation factor	Hope, Thatcher, Timstein
Triticin	Hope, Thatcher
Aspartate aminotransferase	Hope, Thatcher, Timstein
HSP 70 precursor	Hope, Thatcher
NADP-specific isocitrate dehydrogenase	Hope, Thatcher
Alcohol dehydrogenase	Hope, Thatcher, Timstein
3-ketoacyl-CoA thiolase-like protein	Hope, Thatcher
Xylanase inhibitor	Hope, Thatcher, Cheyenne, Chinese Spring
Globulin-2 precursor	Hope, Chinese Spring
Embryo globulin	Hope, Timstein
Globulin-like protein	Hope, Thatcher, Timstein
Fructose-bisphosphate aldolase	Hope, Thatcher, Cheyenne, Chinese Spring
Glucose and ribitol dehydrogenase	Hope, Cheyenne
Aldose reductase	Hope, Thatcher
beta-amylase	Hope, Thatcher
Peroxidase	Hope, Thatcher, Timstein
Malate dehydrogenase	Hope, Thatcher
Triosephosphate isomerase	Hope, Thatcher, Cheyenne,
DNA-binding protein	Thatcher, Timstein
Serpin	Thatcher, Cheyenne
Glyceraldehyde-3-phosphate dehydrogenase	Thatcher, Cheyenne
Globulin 1	Thatcher, Chinese Spring
Alpha-2-purothionin precursor	Thatcher, Chinese Spring
USP family protein	Thatcher, Cheyenne
Peptidyl-prolyl cis-trans isomerase	Thatcher, Timstein, Chinese Spring
Beta-D-glucan exohydrolase	Timstein, Cheyenne
heat-shock protein	Timstein, Chinese Spring, Cheyenne
Enolase	Cheyenne, Chinese Spring
Glucose-1-phosphate adenylyltransferase	Hope, Cheyenne
Beta-glucosidase	Cheyenne, Chinese Spring
Peroxiredoxin	Cheyenne, Chinese Spring

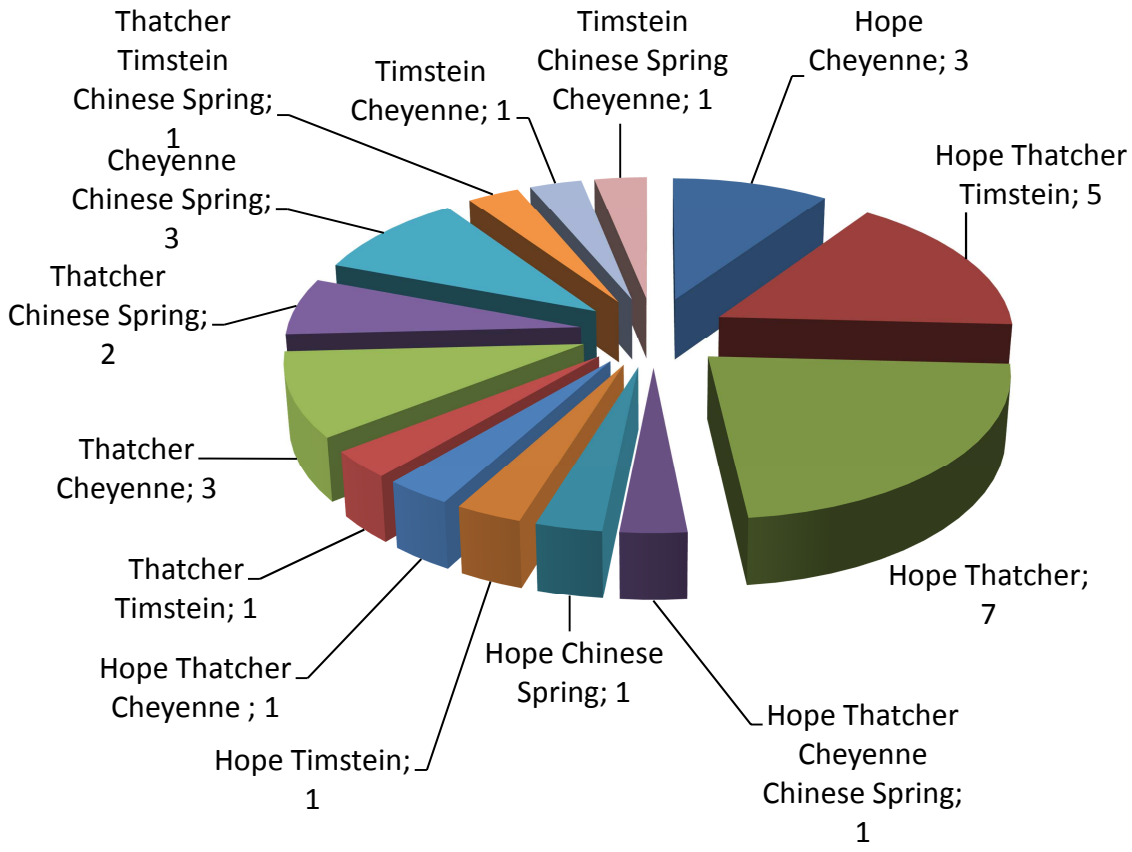


Figure 51: Representation of the number of proteins share between the various bread wheat cultivars.

V. Conclusion

Wheat is the dominant crop in temperate countries and is used for human food and livestock feed. Its success depends partly on its adaptability and high yield potential but also on the gluten protein fraction which confers the viscoelastic properties that allow dough to be processed into bread, pasta, noodles, and other food products. The wheat genome is about 5500 Mb for the tetraploid wheat and 16,000 Mb for the hexaploid wheat, and it has been shown that over 30 000 genes are expressed in the developing wheat grain (Wan et al., 2008). The genes present on the chromosome 5 have a role in the quantity of protein, in the frost resistance, in the hardness, in the compact spike morphology. Thus the knowledge of polypeptides encoded by genes at chromosomes 5 will allow correlating their presence with specific physiological characteristics, along with quality properties. Recently the virtual gene order of 392 genes for the 5A short arm and of 1,480 for the 5A long arm was performed (Vitulo et al., 2011).

By using a comparative proteomic study between parental lines and chromosome substitution lines, we were able to identify specific 5A protein both in tetraploid wheat and hexaploid wheat.

The first target was the identification of the 5A proteins of tetraploid wheat, more particularly the *T. turgidum* ssp. *durum* cv Langdon and *T. dicoccoides*, both on the metabolic fraction and the CM-like fraction. The gliadin and the glutenin fractions were also studied in order to check if there are additional loci, besides those already known present on chromosomes 1 and 6. Finally, mass spectrometry analysis led to the identification of 20 proteins of the metabolic fraction for *T. dicoccoides*, and 8 proteins for *T. turgidum* ssp. *durum* cv Langdon. The major part of the 5A identified proteins has a role in the binding (28% for Langdon and 57% for *T. dicoccoides*) and are involved mainly in the carbohydrate metabolism, in processes of stress / defense and as storage proteins. Concerning the CM-like fraction, we were able to identified 6 spots of *T. dicoccoides* and 3 for Langdon, although the great abundance of the α-amylase inhibitor in this fraction made difficult the validation of the identified proteins. However, in addition to the α-amylase inhibitor, the adenylate kinase, the chitinase, the vacuolar H+ pyrophosphatase, the Superoxide dismutase [Cu-Zn] have been identified.

We were also able to confirm that neither gliadins nor glutenins are encoded by genes on the chromosome 5A.

Preliminary results obtained relatively to the nuclear proteins of leaves of durum wheat, performed by 1D electrophoretic analyses, indicated the presence of a group of bands around 50 kDa in *T. turgidum* ssp. *durum* cv Langdon, that are very likely 5A encoded, and that will be submitted to 1D-MS analyses.

The second objective was to identify 5A encoded polypeptides in bread wheats. The 2D proteomic allowed us to identify specific cultivar 5A proteins in five cultivars,.We identified 39 proteins for cv. Hope, 45 proteins for cv Thatcher, 22 proteins for cv Timstein, 27 proteins for cv Cheyenne and 25 proteins for cv Chinese Spring. All these proteins have mainly a binding function, which is consistent with the results presented by Vitulo et al (2011). Among these proteins, 32 are common between two or more cultivars. After interrogation of the GrainGenes EST database, 15 proteins confirmed the chromosome localization.

Acknowledgments

First I want to thank my supervisors Prof. Stefania Masci and Prof. Domenico Lafiandra for giving me the opportunity to do my PhD research in their laboratory. I also thank them for the advice, the availability and the help for the conduct of my PhD. I especially acknowledge Prof. Stefania Masci for her patience and availability during the writing of this thesis.

I also want to thank all my laboratory colleagues Linda, Francesco, Eleonora, Lina, Paolo and Daniela.

A special thanks to Federica, who was able to support me as neighbor benchtop during this three years, and who has become much more than just a colleague.

I would like to thank all the INRA mass spectrometry team, and more particularly Hélène Rogniaux, for welcoming me into their laboratory and taking time to share their knowledge with me.

I express my acknowledgments to Ombretta Repetto for training and advices for the nuclear proteome study.

I want to give a special thanks to my life partner, Roberta, who was able to support me, to listen to me and to advise me each day which was essential to the success of these three years.

Finally, I want to thank all my family who has been a valuable support during my PhD, despite the distance.

References

Aglika E. (2005). Pathogenesis-Related proteins: research progress in the last 15 years. *Gen. Appl. Plant Physiology* **31**: 105-124.

Ainsworth C.C., Gale M.D. and Baird S. (1984). The genetic control of grain esterases in hexaploid wheat. *Theoretical and Applied Genetics* **68**(3): 219-226

Ambasht P.K. and Kayastha A.M. (2002). Plant Pyruvate Kinase. *Biologia Plantarum* **45**(1):1-10.

Autran J.-C., Lew E.J.-L., Nimmo C.C., and Kasarda, D.D. (1979). N-terminal amino sequencing of prolamins from wheat and related species. *Nature* **282**: 527 529.

Baldwin P.M. (2001). Starch granule-associated proteins and polypeptides: A review. *Starch* **53:** 475- 503.

Ball S.G. and Morell M.K. (2003). From bacterial glycogen to starch: Understanding the biogenesis of the plant starch granule. *Ann..Rev. Plant Biol.* **54:** 207-233.

Barcaccia G., Varotto S., et al. (2001). "Analysis of gene expression during flowering in apomeiotic mutants of Medicago spp.: cloning of ESTs and candidate genes for 2n eggs." *Sexual Plant Reproduction* **14**(4): 233-238.

Belderok, B. (2000). Developments in bread-making processes. *Plant Foods for Human Nutrition*.**55**: 15-20.

Benito C., Gallego F.J., Frade J.M., Zaragoza C., Figueiras A.M.(1989) Chromosomal location of adenylate kinase isozymes in Triticeae species. *Theoretical and Applied Genetics* **79** (2): 157-160.

Bhattacharya M., Erazo-Castrejón S.V., Doehlert D.C., McMullen M.S. (2002). Staling of bread as affected by waxy wheat flour blends. *Cereal Chem.* **79:** 178-182.

Bietz J.A., Huebner F.R., Sanderson J.E., and Wall J.S. (1977). Wheat gliadin homology revealed through N-terminal amino acid sequence analysis. *Cereal Chem*. **54**: 1070-1083.

Bietz J.A., Shepherd K.W. and Wall J.S. (1975). Single-kernel analysis of glutenin: Use in wheat genetics and breeding. *Cereal Chem.* **52**: 513- 532.

Bosch A., Vega C. and Benito C. (1987). The peroxidase isozymes of the wheat kernel: tissue and substrate specificity and their chromosomal location.*Theoretical and Applied Genetics* 73(5):701-706.

Boyd W.J.R. and Lee J.W. (1967). The control of wheat gluten synthesis at the genome and chromosome levels. *Experientia* **23**: 332–333.

Brown J.W.S., and Flavell R.B. (1981). Fractionation of wheat gliadin and glutenin subunits by two-dimensional electrophoresis and the role of group 6 and group 2 chromosomes in gliadins synthesis. *Theor*. *Appl. Genet*. **59**: 349-359.

Buléon A., Colonna P., Planchot V., Ball S. (1998) Starch granules: structure and biosynthesis. *International Journal of Biological Macromolecules* **23** : 85–112.

Burt C., Hollins T.W., Nicholson P. (2010) Identification of a QTL conferring seedling and adult plant resistance to eyespot on chromosome 5A of Cappelle Desprez. *Theoretical and Applied Genetics* **122**:119-128

Bushuk W., and Zillman R.R. (1978). Wheat cultivars identification by gliadin electrophoregrams. I. Apparatus method and nomenclature. *Can. J. Plant Sci*. *58*:505-515.

Bustos D.M., Iglesias A.A (2003). Phosphorylated Non-Phosphorylating Glyceraldehyde-3-Phosphate Dehydrogenase from Heterotrophic Cells of Wheat Interacts with 14-3-3 Proteins. *Plant Physiology* **133**: 2081-2088.

Carbonero P. and Garcia-Olmedo F. (1999). A multigene family of trypsin/α-amylase inhibitors from cereals., in Seed proteins, P.R. Shewry and R. Casey, Editors., Kluwer Academic Publishers. p. 617-633.

Casati P., Spampinato C.P., Andreo C.S. (1997). Characteristics and Physiological Function of NADP-Malic Enzyme from Wheat. *Plant Cell Physiol.* **38**: 928-934.

Casati P., Zhang X., Burlingame A.L. and Walbot V. (2005). Analysis of Leaf Proteome after UV-B Irradiation in Maize Lines Differing in Sensitivity. *Molecular & Cellular Proteomics* **4**:1673–1685.

Chikmawati T., Miftahudin, Gustafson J.P. (2009) PCR-Based EST Mapping in Wheat (Triticum aestivum L.). *Biodiversitas* **10**:55-62.

Chollet R., Vidal J., O'Leary M.H. (1996). Phosphoenolpyruvate : AUbiquitous, Highly Regulated Enzyme in Plants. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **47**:273-98.

Cox T.S., Murphy J.P. and Harrell G. (1987). Isoelectric focusing: patterns of kernel isozymes from 80 North American winter wheat cultivars. *Can. J. Plant Sci*. **68**: 65-72.

Cui Q. and Karplus M. (2001). Triosephosphate Isomerase a Theoretical Comparison of Alternative pathways. *J. Am. Chem. Soc* **123**: 2284-2290.

Curioni A., Pogna N.E. and Peruffo A.D. (1996). The quantity of bound β-amylases is related to the size of gluten polymers, in Gluten 96, e. C.W. Wrigley, Editor. RACI: Melbourne, Australia. p. 307-311.

D'Ovidio R. and Masci S. (2004). The low-molecular-weight glutenin subunits of wheat gluten. *J.Cereal Sci.* **39**: 321-339.

d'Aloisio E., Paolacci A.R., Dhanapal A.P., Tanzarella O.A., Porceddu E., Ciaffi M. (2010). The Protein Disulfide Isomerase gene family in bread wheat (T. aestivum L.). *BMC Plant Biology* **10**: 101.

Darlington H.F., Tecsi L., Harris,N., Griggs D.L., Cantrell I.C., Shewry P. R. (2000). Starch granule associated proteins in barley and wheat. *J. Cereal Sci.* **32:** 21-29.

Deng Z., Tian J., Zhang Y., Sun C. and Chen J. (2009). Accumulation of lipoproteins in Chinese winter wheat cultivars and their impact on dough mixing characteristics. *Frontiers of agriculture in China* **3**: 274-278.

Drzymała A. and Bielawski W. (2009). Isolation and characterization of carboxypeptidase III from germinating triticale grains. *Acta Biochim Biophys Sin* 69-78.

Dubcovsky J., Echaide M., Giancola S., Rousset M., Luo M.C., Joppa L.R. and Dvorak J. (1997). Seed-storage-protein loci in RFLP maps of diploid, tetraploid, and hexaploid wheat. *Theor. Appl. Genet*. **95**: 1169-1180.

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

Efeoglu B. (2009).Review: Heat Shock Proteins and Heat Shock Response in Plants. *Journal of Science* **22**(2):67-75.

Elliott G., Durand A., Hughes R.K., Kroon P.A., D'Ovidio R., Juge N. (2009) Isolation and characterisation of a xylanase inhibitor Xip-II gene from durum wheat. *Journal of Cereal Science* **50**:324-331.

Feillet P. (2000) Le grain de blé. Composition et utilisation. Ed. Inra-Quae, 308 p.

Garcia-Olmedo F., Carbonero P., Aragoncillo C., Salcedo G. (1982) Chromosomal control of wheat endosperm proteins: a critical review.

Gianibelli M.C., Larroque O.R., MacRitchie F., Wrigley C.W. (2001). Biochemical , Genetic, and Molecular characterization of wheat endosperm proteins. Online Review, p. 1-20.

Gorg A., Obermaier C., Boguth G., Harder A., Scheibe B., Wildgruber R., Weiss W. (2000). The current state of two-dimensional electrophoresis with immobilised pH gradients. *Electrophoresis* **21**:1037–1053.

Goyal K., Walton L.J., Tunnacliffe A. (2005) LEA proteins prevent protein aggregation due to water stress. *Biochem. J.* **388**:151-157.

Gupta R.B. and MacRitchie F. (1991). A rapid one-step one-dimensional SDS-PAGE procedure for analysis of LMW subunits of glutenin in wheat. *J. Cereal Sci.* **14**:105-109.

Hancock J.T., Henson D., Nvirenda M., Desikan R., Harrison J., Lewis M., Hughes J., Neill S.J. (2005). Proteomic identification of glyceraldehyde 3-phosphate dehydrogenase as an inhibitory target of hydrogen peroxide in Arabidopsis. *Plant Physiology and Biochemistry.*, **43**: 828–835.

Herbert B.R., Sanchez J.-C., Bini L. (1997). Two-dimensional electrophoresis: the state of the art and future directions, in: Wilkins, M.R., 177 Williams, K.L., Appel, R.D., Hochstrasser, D.F. (Eds.), Proteome Research: New Frontiers in Functional Genomics. Springer, Berlin, pp. 13–34.

Hrmova M. and Fincher G.B. (2001). Structure-function relationships of β- D-glucan endo- and exohydrolases from higher plants. *Plant Molecular Biology* **47**:73-91.

Humphery-Smith I., Cordwell S.J., Blackstock W.P. (1997). Proteome research: complementarity and limitations with respect to the RNA and DNA worlds. *Electrophoresis* **18**: 1217–1242.

Hurkman W., Tanaka C. (2004) Improved methods for separation of wheat endosperm proteins and analysis by two-dimensional gel electrophoresis. *Journal of Cereal Science* **40**:295-299.

Igarashi D., Tsuchida H., Miyao M., Ohsumi C. (2006). Glutamate:Glyoxylate Aminotransferase Modulates Amino Acid Content during Photorespiration. *Plant Physiology* **142**: 901-910.

Islam N., Tsujimoto H., Hirano H. (2003) Wheat proteomics: Relationship between fine chromosome deletion and protein expression. *Proteomics* **3**:307-316.

Islam N., Woo S.-H., Tsujimoto H., Kawasaki H., Hirano H. (2002) Proteome approaches to characterize seed storage proteins related to ditelocentric chromosomes in common wheat (Triticum aestivum L.). *Proteomics* **2**:1146-1155.

Jackson E.A., Holt L.M. and Payne P.I. (1983). Characterisation of high-molecular-weight gliadin and low-molecular-weight glutenin subunits of wheat endosperm by two-dimensional electrophoresis and chromosomal localisation of their controlling genes. *Theor. Appl. Genet*. **66**:29-37.

Jantasuriyarat C., Vales M.I., Watson C.J.W., Riera-Lizarazu O. (2004) Identification and mapping of genetic loci affecting the free-threshing habit and spike compactness in wheat (Triticum aestivum L.). *Theoretical and Applied Genetics* **108**:261-273.

Jiang T., Zhang X.-F., Wang X.-F., Zhang D.-P. (2011). Arabidopsis 3-Ketoacyl-CoA Thiolase-2 (KAT2), an Enzyme of Fatty Acid β-Oxidation, is Involved in ABA Signal Transduction *Plant and Cell Physiology* **52**(3):528-538.

Joppa L.R., Khan K. and Williams N.D. (1983). Chromosomal location of genes for gliadin polypeptides in durum wheat Triticum turgidum L. *Theor. Appl. Genet*. **64**: 289-293.

Kalunke R.M. (2012) Multi-transgene-stacking of glycosidase inhibitor genes to improve resistance against fungal pathogens in wheat. *PhD thesis*.

Kasarda D.D., Adalsteins A.E, Laird N.F (1987). γ-gliadins with α type structure coded on chromosome 6B of the wheat (*Triticum aestivum* L.) cultivar "Chinese Spring." *Proc 3rd Int Workshop Gluten Proteins*, 20-29.

Kim S.Y. et al. (2011). The 1-Cys peroxiredoxin, a regulator of seed dormancy, functions as a molecular chaperone under oxidative stress conditions. *Plant Science* **181**: 119-124.

Kim Y-O., Kim J.S., Kang H. (2005). Cold-inducible zinc finger-containing glycine-rich RNA-binding protein contributes to the enhancement of freezing tolerance in Arabidopsis thaliana. *The Plant Journal* **42**: 890-900.

Kimoto M., Suzuki M., et al. (2009). Isolation and Molecular Cloning of a Major Wheat Allergen, Tri a Bd 27K. *Bioscience, Biotechnology, and Biochemistry* **73**(1): 85-92.

Klose J. (1975). Protein mapping by combined isoelectric focussing and electrophoresis of mouse tissues. A novel approach to testing for induced point mutations in mammals. *Humangenetik* **26**: 231–243.

Kocsy G., Athmer B., Perovic D., Himmelbach A., Szűcs A., Vashegyi I., Schweizer P., Galiba G., Stein N. (2010) Regulation of gene expression by chromosome 5A during cold hardening in wheat. *Molecular Genetics and Genomics* **283**:351-363

Komarnitsky P, Cho E-J, Buratowski S. (2000). Different phosphorylated forms of RNA polymerase II and associated mRNA processing factors during transcription. *Genes & Development* **14**: 2452-2460.

Koonin E.V., Mushegian A.R., Tatusov R.L., Altschul S.F., Bryant S.H., Bork P., Valencia A. (1994). Eukaryotic translation elongation factor ly contains a glutathione transferase domain - Study of a diverse, ancient protein superfamily using motif search and structural modeling. *Protein Science* **3**: 2045-2054.

Kosuge K., Watanabe N., Melnik V.M., Laikova L.I., Goncharov N.P. (2011) New sources of compact spike morphology determined by the genes on chromosome 5A in hexaploid wheat. Genet. *Resour. Crop Evol.*

Kurepa J. (2002). The Small Ubiquitin-like Modifier (SUMO) Protein Modification System in Arabidopsis. Accumulation of SUMO1 and -2 conjugates is increased by stress. *Journal of Biological Chemistry* **278**: 6862-6872.

Lafiandra D., and Kasarda, D.D. (1985). One- and two-dimensional (two-pH) polyacrylamide gel electrophoresis in a single gel: Separation of wheat proteins. *Cereal Chem*. **62**:314–319.

Lafiandra D., Ciaffi M. and Benedettelli S. (1993). Seed storage proteins of wild wheat progenitors. In ‘‘Biodiversity and Wheat Improvement’’ (A. B. Damania, ed.), pp. 329–340. John Wiley & Sons, New York.Sozinov and Poperelya, 1980.

Lafiandra D., Colaprico G., Kasarda D.D. and Porceddu E. (1987). Null alleles for gliadin blocks in bread and durum wheat cultivars. *Theor. Appl. Genet*. **74**: 610–616.

Lafiandra D., Morris R. and Kasarda D.D. (1984). Chromosomal assignment of genes codingfor the wheat gliadin protein components of the cultivars Cheyenne and Chinese Spring by twodimensional (two-pH) electrophoresis. *Theor. Appl. Genet*. **68**: 531–539.

Langston P.J., Pace C.N., Hart G.E. (1980) Wheat Alcohol Dehydrogenase Isozymes Purification, Characterization, and Gene Expression. *Plant Physiology* **65**:518-522.

Larré C., Penninck S., Bouchet B., Lollier V., Tranquet O., Denery-Papini S., Guillon F., Rogniaux H. (2010) Brachypodium distachyon grain: identification and subcellular localization of storage proteins. *Journal of Experimental Botany* **61**(6):1771 -1783.

Lauer S.J., Burks E., Irvin J.D., Ravel J.M. (1984). Purification and Characterization of Three Elongation Factors, E F- lα, EF-lβγ, and EF-2, from Wheat Germ. *The Journal of Biological Chemistry* **259**(3):1644-1648.

Lawrence G. J. and Shepherd K.W. (1981). Inheritance of glutenin protein subunits of wheat. *Theor. Appl. Genet*. **60**:333-337.

M.R., Swanson B.G., Baik B.K. (2001). Influence of amylose content on properties of wheat starch and breadmaking quality of starch and gluten blends. *Cereal Chem.* **78:** 701-706.

Liu D., Tu L., Li Y., Wang L., Zhu L., Zhang X. (2008). Genes Encoding Fasciclin-Like Arabinogalactan Proteins are Specifically Expressed During Cotton Fiber Development. *Plant Molecular Biology Reporter* **26**: 98-113.

Maciga M and Paszkowski A. (2004). Genetic control of aspartate aminotransferase isoenzymes in Aegilops and Triticum species. *J. Appl. Genet*. **45**: 411-417.

Majoul T., Bancel E., Triboï E., Ben Hamida J., Branlard G. (2004). Proteomic analysis of the effect of heat stress on hexaploid wheat grain: Characterization of heat-responsive proteins from non-prolamins fraction. *Proteomics* **4**: 505-513.

Malanovic N., Streith I., Wolinski H., Rechberger G., Kohlwein S.D. and Tehlivets O.(2008). S-Adenosyl-L-homocysteine Hydrolase, Key Enzyme of Methylation Metabolism, Regulates Phosphatidylcholine Synthesis and Triacylglycerol Homeostasis in Yeast Implications for Homocysteine as a risk factor of atherosclerosis. *J Biol Chem.***283(**35):23989-23999.

Mamone G, Picariello G, Addeo F, Ferranti P. (2011). Proteomic analysis in allergy and intolerance to wheat products. *Expert Rev. Proteomics* **8**: 95–115.

Manaa A., Ahmed H.B., Valot B., Bouchet J-P., Aschi-Smiti S., Mathilde Causse M. and Faurobert M. (2011). Salt and genotype impact on plant physiology and root proteome variations in tomato. *J Exp Bot.* **62**(8):2797-813.

Masci S., Rovelli L., Kasarda D.D., Vensel W.H., Lafiandra D. (2002). Characterisation and chromosomal localisation of C-type low-molecular-weight glutenin subunits in the bread wheat cultivar Chinese Spring. *Theor. Appl. Genet*. **104**: 422–428.

Massiah A.J. and Hartley M.R. (1995). Wheat ribosome-inactivating proteins: Seed and leaf forms with different specificities and cofactor requirements. *Planta* **4**:633-640.

Matsuoka, Y. (2011). Evolution of Polyploid Triticum Wheats under Cultivation: The Role of Domestication, Natural Hybridization and Allopolyploid Speciation in their Diversification. *Plant and Cell Physiology* **52**(5): 750-764.

Merita N., Maeda T., Miyazaki M., Yamamori M., Miura H., Ohtsuka I. (2002). Dough and baking properties of high-amylose and waxy wheat flours. *Cereal Chem.* **79:** 491-495.

Merlino M., Leroy P., Chambon C., Branlard G. (2009) Mapping and proteomic analysis of albumin and globulin proteins in hexaploid wheat kernels (Triticum aestivum L.). *Theoretical and Applied Genetics* **118**:1321-1337.

Metakovsky E.V. (1991). Gliadin allele identification in common wheat II. Catalogue of gliadin alleles in common wheat. *J. Genet. Breed*. **45**:325-344.

Metakovsky E.V., Novoselskaya A.Y., and Sozinov A.A. (1984). Genetic analysis of gliadin components in winter wheat using two-dimensional polyacrylamide gel electrophoresis. *Theor. Appl. Genet*. **69**:31-37.

Miflin, B., Field J.M., and Shewry P.R. (1983). Cereal storage proteins and their effects on technological properties., in Seed Proteins, D.J.M.J.a.V. J., Editor. Academic Press: New-York.

Moreau R.A. (1987). Calcium-Binding Proteins in Fungi and Higher Plants. *Journal of Dairy Science* **70**(7):1504-1512.

Morris C.F. (2002) Puroindolines: the molecular genetic basis of wheat grain hardness. *Plant Molecular Biology* **48**:633-647.

O'Farrell P.H. (1975). High-resolution two-dimensional electrophoresis of proteins. *Journal of Biological Chemistry* **250:** 4007–4021.

Orth R.A., Shellenberger J.A. (1988). Origin, production, and utilization of wheat. In: Wheat Chemistry and Tecnology. Ed. Y. Pomeranz, American Association of Cereal Chemists, St Paul, Minnesota, pp. 1-14.

Osborne, T.B. (1907). The protein of the wheat kernel. Publication No. **84**. Carnegie Institute: Washington, DC.

Park Y.-J. and Luger K. (2006).The structure of nucleosome assembly protein 1 *Proc Natl Acad Sci U S A.* **103**(5): 1248–1253.

Park K.S. and Kahn M.L. (1999). Distribution of two isoforms of NADP-dependent isocitrate dehydrogenase in soybean (Glycine max). *Plant Mol Biol.* **40**(1):13-21.

Payne P I. (1987). Genetics of wheat storage protein and the effect of allelic variation on breadmaking quality. *Ann. Rev. Plant Physiol*. **38**:141-153.

Payne P.I. and Corfield K.G. (1979). Subunit composition of wheat glutenin proteins, isolated by gel filtration in a dissociating medium. *Planta* **145**:83-88.

Payne P.I., Corfield K.G., Holt L.M. and Blackman J.A. (1981). Correlation between the inheritance of certain high-molecular-weight subunits of glutenin and bread-making quality in progenies of six crosses of bread wheat. *J. Sci. Food Agric*. **32**:51-60.

Payne P.I., Holt L.M., Jarvis M.G. and Jackson E.A. (1985). Two-dimensional fractionation of the endosperm proteins of bread wheat (Triticum aestivum): Biochemical and genetic studies.*Cereal Chem*. **62**:319-326.

Payne P.I., Holt L.M., Lawrence G.J., and Law C.N. (1982). The genetics of gliadin and glutenin, the major storage proteins of the wheat endosperm. *Qual. Plant. Food Hum. Nutr*.**31**:229-241.

Payne P.I., Jackson E.A., Holt L.M. and Law C.N. (1984). The genetic linkage between endosperm protein genes on each of the short arms of chromosomes 1A and 1B in wheat. *Theor. Appl. Genet*. **67**:235–243.

Payne, P.I. and Rhodes, A.P. (1982). *Encyclopaedia of Plant Physiology*, New Series (Springer, Berlin), Vol. **14**A, 346-369.

Peleg Z., Fahima T., Korol AB., Abbo S., Saranga Y. (2011). Genetic analysis of wheat domestication and evolution under domestication. *Journal of Experimental Botany* **62**: 5051-5061.

Peng, J.H., D. Sun, and E. Nevo (2011), Domestication evolution, genetics and genomics in wheat. *Molecular Breeding* **28**(3): p. 281-301.

Pérez R., de Bustos A., Jouve N., Cuadrado A. (2009). Localization of Rad50, a Single-Copy Gene, on Group 5 Chromosomes of Wheat, Using a FISH Protocol Employing tyramide for Signal Amplification (Tyr-FISH). *Cytogenet. Genome Res*. **125**: 321-328

Peruffo A.D., Pogna N.E., and Curioni A. (1996). Evidence for the presence of disulphide bonds between β-amylase and low molecular weight glutenin subunits, in Gluten 96, e. C.W. Wrigley, Editor. RACI: Melbourne, Australia. p. 312-315.

Pogna N.E., Autran J.C., Mellini F., Lafiandra D., and Feillet, P. (1990). Chromosome 1B encoded gliadins and glutenin subunits in durum wheat: genetics and relationship to gluten strength. *J. Cereal Sci*. **11**:15-34.

Popineau, Y. (1988). Physiochimie du gluten, relation avec l'étiologie de la maladie coeliaque. In Symposium international sur l'intestin grêle : physiologie, physiopathologie. Amsterdam: Exepta Medica.

Raden D. and Gilmore R. (1998). Signal Recognition Particle-dependent Targeting of Ribosomes to the Rough Endoplasmic Reticulum in the Absence and Presence of the Nascent Polypeptide-associated Complex. *Mol Biol Cell.* **9**(1): 117–130.

Repetto O., Rogniaux H., Firnhaber C., Zuber H., Küster H., Larre C., Thompson R. and Gallardo K. (2008). Exploring the nuclear proteome of Medicago truncatula at the switch towards seed filling. *The Plant Journal* **56**:398-410.

Salcedo G., Quirce S., Diaz-Perales A. (2011). Wheat Allergens Associated With Baker's Asthma. *J Investig Allergol Clin Immunol* **21**: 81-92.

Scheele G.A. (1975). Two-dimensional gel analysis of soluble proteins. Characterisation of guinea pig exocrine pancreatic proteins. *Journal of Biological Chemistry* **250**: 5375–5385.

Sears E.R. (1954). The aneuploids of common wheat. Missouri Agric. Exp. Station Res. Bull.572, 1–58.

Sears E.R. (1966). Nullisomic–tetrasomic combinations in hexaploid wheat. In ''Chromosome Manipulations and Plant Genetics'' (R. Riley and K. R. Lewis, eds.), pp. 29–45. Oliver and Boyd,London.

Shepherd K.W. (1968). Chomosomal control of wheat endosperm proteins in wheat and rye. In‘‘Proceedings of the 3rd International Wheat Genetics Symposium’’ (K. W. Finlay and K. W.Shepherd, eds.), pp. 86–96. Australian Academy of Science, Canberra.

Shepherd K.W. (1988). Genetics of wheat endosperm proteins in retrospect and prospect. In‘‘Proceedings of the 7th International Wheat Genetics Symposium,’’ pp. 943–949. Institute of Plant Science Research, Cambridge.

Shewry P.R., Halford N.G., Lafiandra D. (2003). Genetics of wheat gluten proteins. *Adv.Genet.* **49**: 111-184.

Shewry P.R., Miflin B.J. and Kasarda D.D. (1984). The structural and evolutionary relationships of the prolamin storage proteins of barley, rye and wheat. *Philos. Trans. R. Soc.Lond. B Biol. Sci.* **304**: 333–339.

Shewry PR. (2009). Wheat. *Journal of Experimental Botany* **60**: 1537-1553.

Shewry, P., Tatham A., Forde J., Kreis M., and Miflin B. (1986). The classification and nomenclature of wheat gluten proteins, a reassessment. *J.Cereal Sci*., **4**: 97-106.

Shewry, P.R. and Halford N.G. (2002). Cereal seed storage proteins: structures, properties and role in grain utilization. *J Exp Bot* **53**(370): 947-958.

Shewry, P.R., and Tatham, A.S. (1990). The prolamin storage proteins of cereal seeds: Structure and evolution. *Biochem. J.* **267**:1-12.

Shewry, P.R., Napier J.A., Tatham A.S. (1995). Seed storage proteins: Structures and biosynthesis. *Plant Cell*, **7**: 945-956.

Simões I. and Faro C. (2004). Structure and function of plant aspartic proteinases. *European Journal of Biochemistry* **271**: 2067-2075.

Simpson D.J., Fincher G.B., Huang A.H.C., Cameron-Mills V. (2003) Structure and function of cereal and related higher plant (1->4)-beta-xylan endohydrolases *J Cereal Sci*, **37**:111–127

Singh A., Kirubakaran S.I., Sakthive N. (2007) Heterologous expression of new antifungal chitinase from wheat. *Protein Expression and Purification* **56**: 100-109.

Singh J. and Skerritt J.H. (2001) Chromosomal Control of Albumins and Globulins in Wheat Grain Assessed using Different Fractionation Procedures. *Journal of Cereal Science* **33**:163-181.

Singh N.K. and Shepherd K.W. (1988). Linkage mapping of the genes controlling endosperm storage proteins in wheat. I. Genes on the short arms of group-1 chromosomes. *Theor. Appl. Genet.* **75**: 628–641.

Slageren MW van. (1994). Wild wheats: a monograph of Aegilops L. and Amblyopyrum (Jaub. & Spach) Eig (Poaceae). *Wageningen Agriculture University Papers* 1994(7). 513 pp.

Snape J.W., Law C.N., Parker B.B., Worland A.J. (1985) Genetical analysis of chromosome 5A of wheat and its influence on important agronomics characters. *Theor Appl Genet* **71**: 518-526.

Snape J.W., Sarma R., Quarrie S.A., Fish L., Galiba G., Sutka J. (2001) Mapping genes for flowering time and frost tolerance in cereals using precise genetic stocks. *Euphytica* **120**:309-315.

Sozinov A.A. and Poperelya, F.A. (1980). Genetic classification of prolamines and its use for plant breeding. *Ann. Technol. Agric.* (Paris) **29**: 229–245.

Sree B.K., Rajendrakumar C.S.V., Reddy A.R. (2000). Aldose reductase in rice (Oryza sativa L.): stress response and developmental specificity. *Plant Science* **160**: 149-157.

Sreeramulu G. and Singh N.K. (1997) Genetic and biochemical characterization of novel low molecular weight glutenin subunits in wheat (Triticum aestivum L.) *Genome* **40**:41-48.

Sue M., Yamazaki K., Yajima S., Nomura T., Matsukawa T., Iwamura H. Toru Miyamoto T. (2006). Molecular and Structural Characterization of Hexameric β-D-Glucosidases in Wheat and Rye. *Plant Physiology* **141**:1237-1247.

Sutka J., Snape J.W. (1989) Location of a gene for frost resistance on chromosome 5A of wheat. *Euphytica* **42**:41-44.

Suzuki Y., Kanayama Y., Shiratake K., Yamaki S. (1999). Vacuolar H+ pyrophosphatase purifed from pear fruit. *Phytochemistry* **50**: 535-539.

Tatham A.S. and Shewry P.R. (1995). The S-poor prolamins of wheat, barley and rye. *J. Cereal Sci.* **22**: 1-16.

Tsujimoto H. and Noda K. (1990) Deletion mapping bt gametocidal genes in common wheat: position of speltoid suppression (Q) and β-amylase (β-amy-A2) genes on chromosome 5A. *Génome* **33**(6): 850-853.

Veisz O.B. and Sutka J. (1998) Frost resistance of "Chinese Spring"/ "Cheyenne" chromosome suhstitution lines under short- and long-day hardening conditions. *Plant Breeding* **17**:93-94.

Vensel, W.H., Tanaka C.K., Cai N., Wong J.H., Buchanan B.B., Hurkman W.J. (2005). Developmental changes in the metabolic protein profiles of wheat endosperm. *Proteomics* **5** :1594–1611.

Vítámvás P., Prášil I.T., Kosova K., Planchon S., Renaut J. (2012). Analysis of proteome and frost tolerance in chromosome 5A and 5B reciprocal substitution lines between two winter wheats during long-term cold acclimation. *Proteomics* **12**(1): 68-85.

Vitulo N., Albiero A., Forcato C., Campagna D., Dal Pero F., et al. (2011) First Survey of the Wheat Chromosome 5A Composition through a Next Generation Sequencing Approach. *PLoS ONE* 6(10): e26421.

Wan Y., Poole R.L., Huttly A.K., et al. (2008). Transcriptome analysis of grain development in hexaploid wheat. *BMC Genomics* **9**:121.

Wang J.R., Yan Z.H., Zheng Y.L., Cao W.G., We Y.M. (2010). Molecular cloning and phylogenetic analysis of fructose bisphosphate aldolase (cytoplasmic isozyme) in wheat, barley and rye. *Cereal Research Communications* **38**:489-496.

Watanabe, N., Sugiyama, K., Yamagishi, Y. and Sakata, Y. (2002) Comparative telosomic mapping of homoeologous genes for brittle rachis in tetraploid and hexaploid wheats. Hereditas **137**: 180–185.

Wrigley C.W. (1968). Gel electrofocusing a technique for analyzing multiple protein samples by isoelectric focusing. *Science Tools* **15:** 17–23.

Wrigley C.W. (1970). Protein mapping by combined gel electrofocusing and electrophoresis: Application to the study of genotypic variations in wheat gliadins. *Biochem. Genet*. **4**:509-516.

Wrigley C.W. (1996). Giant proteins with flour power. *Nature* **381**:738-739.

Wrigley C.W., Lawrence G. J., and Shepherd, K.W. (1982). Association of glutenin subunits with gliadin composition and grain quality in wheat. *Aust. J. Plant Physiol*. **9**:15-30.

Wu G., Wilen R.W., Robertson A.J., Gusta L.V. (1999) Isolation, Chromosomal Localization, and Differential Expression of Mitochondrial Manganese Superoxide Dismutase and Chloroplastic Copper/Zinc Superoxide Dismutase Genes in Wheat. *Plant Physiology* **120**:513-520.

Wu M.J., McKay S., Howes N., Chin J., Hegedus E. (2012). Identification of novel serpin isoforms and serpin polymorphisms among Australian wheat cultivars. *Journal of Cereal Science* **55**(2):202-209.

Yan N., Doelling J.H., Falbel T.G., Durski A.M.,Vierstra R.D.(2000). The Ubiquitin-Specific Protease Family from Arabidopsis. AtUBP1 and 2 Are Required for the Resistance to the Amino Acid Analog Canavanine. *Plant Physiology* **124**:1828–1843.

Yang, L., Tian D., et al. (2011). Proteomics-based identification of storage, metabolic, and allergenic proteins in wheat seed from 2-DE gels. *African Journal of Agricultural Research* **6**: 808-816.

Yasui T., Sasaki T., Matsuki J. (1999). Miling and flour pasting properties of waxy endosperm mutant lines of bread wheat *(Triticum aestivum* L.). *J. Sci. Food Agric*. **79:** 687-692.

Yingyin Y., Zhongfu N., Jinkun D., Xiuli W., Haiyan W., Qixin S.(2006). Isolation and characterization of 15 genes encoding ribosomalproteins in wheat (Triticum aestivum L.) *Plant Science* **170**(3):579-586.

Zeng M., Morris C.F., Batey I.L., Wrigley C.W., (1997). Sources of Variation for Starch Gelatinization, Pasting, and Gelation Properties in Wheat. *Cereal Chem.* **74:** 63-71.