Theor Appl Genet (2012) 125:1433–1448
DOI 10.1007/s00122-012-1924-5

ORIGINAL PAPER

Proteomic and genetic analysis of wheat endosperm albumins and globulins using deletion lines of cultivar Chinese Spring

Marielle Merlino · Sabrina Bousbata · Birte Svensson · Gérard Branlard

Received: 27 March 2012 / Accepted: 11 June 2012 / Published online: 30 June 2012

Abstract Albumins and globulins from the endosperm of *Triticum aestivum* L. cv Chinese Spring (CS) were analysed to establish a proteome reference map for this standard wheat cultivar. Approximately, 1,145 Coomassie-stained spots were detected by two-dimensional gel electrophoresis (2DE), 410 of which were identified using mass spectrometry and data mining. Salt-soluble endosperm proteins from 67 CS deletion lines were also separated by 2DE (four gels per line). Image analysis of the 268 2DE gels as compared to the CS reference proteome allowed the detection of qualitative and quantitative variations in endosperm proteins due to chromosomal deletions. This differential analysis of spots allowed structural or regulatory genes, encoding 211 proteins, to be located on segments of the 21 wheat chromosomes. In addition, variance analysis of quantitative variations in spot volume showed that the expression of 391 proteins is controlled by one or more chromosome bins with 262 significant increases and 196 significant decreases in spot volume. The spot volume of several proteins was increased or decreased by numerous chromosomal regions and homoeologous-like regulation was revealed for some proteins. Quantitative or qualitative variation in a total of 386 proteins was influenced by genes assigned to at least one chromosomal region, while 66 % of all stained proteins were not found to be influenced by chromosome bins. Proteomics of deletion lines can, therefore, be used to simultaneously analyse the composition and genetics of a complex tissue, such as the wheat endosperm.

Abbreviations

Agl Albumins and globulins
CS Chinese Spring
DL Deletion line

Communicated by X. Xia.

Electronic supplementary material The online version of this article (doi:10.1007/s00122-012-1924-5) contains supplementary material, which is available to authorized users.

M. Merlino · G. Branlard (✉)
INRA, UMR 1095 GDEC-UBP, 5 chemin de Beaulieu,
63100 Clermont Ferrand, France
e-mail: branlard@clermont.inra.fr

S. Bousbata · B. Svensson
Enzyme and Protein Chemistry, Department of Systems Biology,
Technical University of Denmark, 2800 Kgs Lyngby, Denmark

S. Bousbata
Laboratoire de Parasitologie Moléculaire,
Plate-forme de Protéomique, IBMM,
Université Libre de Bruxelles,
12 rue des Professeurs Jeener et Brachet,
6041 Gosselies, Belgium

Introduction

Albumins and globulins (agl) of wheat endosperm each account for approximately 10 % of the total protein content of flour. These salt-soluble proteins are mainly enzymes or other proteins involved in cell functions that determine the characteristics of kernel storage components, flour composition and dough properties. Some of the major enzymes present in the wheat endosperm have been identified using both classical biochemical and molecular methods, and more recently, by proteomics (Salt et al. 2005; Skylas et al. 2000; Vensel et al. 2005). Similar studies have been conducted on barley (Finnie et al. 2002) and rice (Komatsu

et al. 2004). Many of these soluble proteins, however, remain to be identified and mapped.

The allohexaploid nature of common wheat *Triticum aestivum* L. (2n = 6x = 42) was used to create aneuploid lines (Sears 1954), which made it possible to assign structural genes encoding specific wheat proteins to chromosome arms. Payne et al. (1985) and Lafiandra et al. (1989) first used two-dimensional gel electrophoresis (2DE) to study the genetic control of storage proteins in the nullisomic–tetrasomic and ditelosomic lines of Chinese Spring (CS). Comparison of the 2DE patterns of ditelosomic lines with the euploid CS pattern also allowed the localisation of structural and regulatory leaf genes (Desfrancs and Thiellement 1985; Thiellement et al. 1986). Chromosomal control of seed proteins has been investigated using 2DE as well as isoelectric focusing, reversed-phase high-performance liquid chromatography (HPLC) (Singh and Skerritt 2001), and isotope coded affinity tagging electrospray ionisation techniques (Islam et al. 2003). The latter technique was used to improve protein separation and quantification when analysing fine deletion lines of chromosome 1B (Tsujimoto et al. 2001). More recently, some enzymes were identified and mapped in a proteomics analysis of the International Triticeae Mapping Initiative segregating progeny (Merlino et al. 2009), but many agl which are not segregating remain unknown and unassigned to chromosomes.

In the present study, a proteome reference map of the parental wheat cultivar CS was established. In parallel, comparison of the endosperm proteomes of 67 CS deletion lines (Endo and Gill 1996) were used to locate structural genes encoding agl found in the mature endosperm and their regulatory genes on the *Triticum aestivum* map.

Materials and methods

Plant material

Sixty-seven deletion lines (DLs) (Table 1) derived from the common wheat *Triticum aestivum* L. cv Chinese Spring (Endo and Gill 1996) and the parental cultivar were used in this study. These DLs are maintained in the Wheat Genetic and Genomic Resources Centre (Kansas State University, Department of Plant Pathology, Manhattan, KS 66506-5502, USA). Plants were grown from seed in 2005 with full fungicide protection in a greenhouse at the INRA station of Clermont-Ferrand (France). Seeds were harvested from bagged ears. The C-banding technique and molecular markers were used to verify each DL carrying the expected deletion. The nomenclature employed for the DLs is as in the following example: 1AL-1 indicates the plant lacking bin number 1 on the long arm of chromosome 1A. The choice of the deletion lines was defined by the chromosomal position of the deletion (one or two lines per chromosome arm) and their availability at the research station.

Outside layers and embryos were manually removed from 5 to 6 seeds giving approximately 100 mg of starchy endosperm per biological replicate per line. Two biological replicates and two technical replicates were analysed by 2DE.

Protein extraction and two-dimensional gel electrophoresis

Agl proteins were extracted essentially as described by Marion et al. (1994) using a less concentrated extraction buffer (10 mM sodium phosphate, 10 mM NaCl, pH 7.8 at

Table 1 Number of deletion lines per group of homoeology

Group 1	Group 2	Group 3	Group 4	Group 5	Group 6	Group 7
1AL-1	2AL-1	3AL-2	4AL-12	5AL-10	6AS-1	7AL-1
1AS-1	2AS-5	3AL-5	4AL-13	5AL-12	6BL-5	7AS-5
1BL-1	2BL-1	3AS-2	4AS-1	5AL-17	6BS-2	7BL-2
1BS-10	2BL-3	3AS-4	4BL-1	5AL-23	6BS-5	7BS-1
1BS-9	2BS-1	3BL-10	4BS-1	5AS-3	6DL-6	7DL-2
1DL-4	2BS-3	3BL-2	4DL-13	5BL-6	6DS-2	7DL-5
1DS-1	2DL-9	3BL-7	4DS-1	5BS-5		7DS-4
1DS-5	2DS-1	3BL-8	4DS-3	5BS-6		7DS-5
	2DS-5	3BS-1		5BS-8		
		3BS-2		5DL-1		
		3BS-3		5DL-5		
		3BS-4		5DS-1		
		3BS-5		5DS-2		
		3BS-7				
		3DL-3				
8	9	15	8	13	6	8

4 °C) supplemented with a protease inhibitor cocktail (P9599, Sigma, St. Louis, MO, USA). Agl were precipitated from the extract supernatant with ice-cold acetone overnight at 20 °C. The dried protein pellet was dissolved in the 2DE sample buffer containing 4 % CHAPS, 7 M urea, 2 M thiourea, 1 % DTT and 1 % IPG buffer 3-11 NL (GE Healthcare, Uppsala, Sweden). The protein concentration was determined using the Bradford protein assay (B6916, Sigma, St. Louis, MO, USA). For each 2DE, 150 µg of protein extract was loaded on 24-cm long immobiline strips (pH 3–11 NL) (GE Healthcare, Uppsala, Sweden). The first dimension and SDS-PAGE were performed as previously described by Debiton et al. (2011).

Qualitative and quantitative image analysis

Four (two biological and two technical) replicates of 2DE were performed for each of the 67 DLs plus the parental cultivar CS. Gels were scanned with a GS800 scanner (Biorad, Richmond, VA, USA) and analysed using SameSpots software v4.1 (Nonlinear Dynamics Ltd, UK). Images were aligned with respect to the CS reference images. Images were exported for spot detection, matching, normalisation and quantification. Spots were quantified on the basis of their normalised volume, that is, the spot volume divided by the total volume of the whole set of gel spots. Spots were scored as variable if they fulfilled the following stringent statistical criteria: ANOVA $p \leq 0.05$, fold difference ≥ 2, $q < 0.05$ and power > 0.8.

The CS reference gel was qualitatively compared with the gels (four 2DE replicates) of each DL relative to each chromosomal arm or chromosome. The statistical criteria used for chromosomal bin assignment were ANOVA $p \leq 0.01$, fold difference ≥ 2, and $q < 0.05$. The ANOVA p value was decreased to be more stringent on validation of the absence of the spot. Although there were no missing values for any gel area when using SameSpots software, a spot was considered as absent if its expression profile was similar to the gel background in all four gel replicates.

Normalised spot volumes for each DL were analysed quantitatively according to the following statistical criteria: ANOVA $p < 0.05$, fold difference ≥ 2, and $q < 0.05$. These criteria revealed numerous spots that varied significantly, including those already found to vary qualitatively above.

Spots showing false matches were not included in both qualitative and quantitative analysis.

Protein identification

Major spots (562), revealed using Coomassie Brilliant Blue (CBB) staining, were subjected to in-gel trypsin digestion. Samples were prepared for mass spectrometry (MS) according to Bak-Jensen et al. (2007). Tryptic digests were analysed on an Ultraflex ToF/ToF mass spectrometer (Bruker, Bremen, Germany). Spectra were processed using FlexAnalysis 2.2 and Biotools 2.2 software (Bruker). MSMS spectra were obtained on the three most abundant peaks of the PMF spectrum. The protein mass list was first searched against the National Center for Biotechnology Information (NCBI) non-redundant protein (nr) database using in-house Mascot software (http://www.matrixscience.com). If the search did not result in identification of a protein, the mass list was searched against the Institute of Genomic Research (TIGR) wheat tentative consensus (TC) database (http://www.tigr.org) or the database of ESTs other than those from human and mouse. The search criteria were monoisotopic mass accuracy <50 ppm, one missed cleavage, and complete carbamidomethylation of cysteine, partial oxidation of methionine and pyroglutamic acid as allowed modifications.

Results

Identification of Chinese Spring endosperm proteins

The first objective to the study was to establish a proteome reference map of the parental wheat cultivar CS. The agl fraction of wheat endosperm was extracted from *Triticum aestivum* cv. Chinese Spring and constituent proteins were separated by 2DE. A total of 562 of the major spots were subjected to MALDI-ToF–ToF MS peptide mass fingerprinting allowing the unambiguous identification of 427 proteins (410 spots) (Fig. 1 and Table in electronic supplementary material). According to the NCBInr and TIGR databases that were searched, 327 of the proteins were wheat specific. The proteins in 152 spots were not identified either because the spectra were of poor quality (50 spots) or because the protein was not present in the databases searched (102 spots). Most proteins (284) were matched by two or more peptides while 143 were identified by a single peptide match. Approximately, 34 % of the 427 identified proteins matched EST sequences and 7 proteins had either an unknown function or no homology to any sequence in the databases.

Based on the database annotations, proteins and their isoforms fell into different families/classes defined by the Gene Ontology project (http://www.geneontology.org). As expected, proteins involved in carbohydrate metabolism (166 proteins), proteases (58 proteins), globulins (50 proteins) and proteins involved in amino acid synthesis and protein metabolism/catabolism (49 proteins) were more abundant than proteins involved in stress (27 proteins), defence (21 proteins), fatty acid metabolism (5 proteins) and pumps, channels and transport proteins (4 proteins).

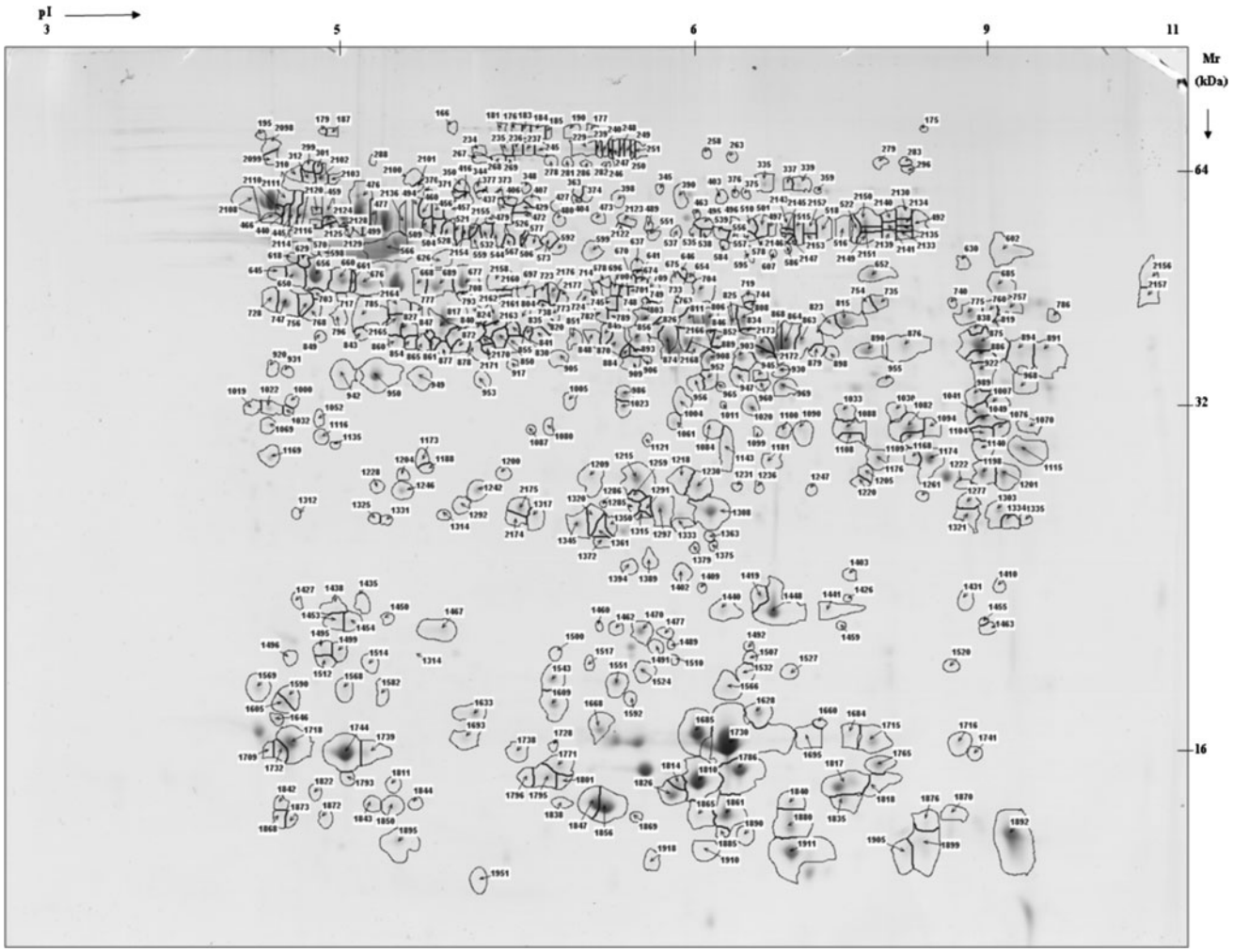

Fig. 1 Two-dimensional proteome map of salt soluble protein extract of wheat cv Chinese Spring. First dimension at a pH gradient from 3 to 11 and the second vertical dimension on 14 % SDS-Page. Spots used for chromosomal bins location and those identified by mass spectrometry are annotated

Variation in agl proteins in CS deletion lines

The gel of endosperm proteins from the parental CS line used as reference to align all gel images had a total of 1,145 CBB-stained agl spots. The number of spots that were significantly different (absent/present, increased/decreased) in each CS DL are listed in Table 2. The B genome produced the highest number of varying agl spots (738) probably because many DLs were available for that genome (28 DLs including 10 for chromosome 3B) as compared to the A and D genomes (19 and 20 DLs respectively) which each produced 453 varying agl spots. The total number of quantitatively varying agl, i.e., spots that either increased (i) or decreased (d), was very similar for genome A (182 i and 208 d) and D (162 i and 200 d) deletions, unlike genome B deletions (244 i and 378 d) that revealed more decreased spots. For some DLs, a similar number of spots increased and decreased like 3AS-4 (26 i and 29 d), 6AS-1 (34 i and 31 d), 1DL-4 (34 i and 36 d), and 6DS-2 (18 i and 17 d), but for the majority of DLs, a number of spots with increased and decreased spot volume were different.

In some DLs very many spots varied like 5BS-5 (86 spots), 1DL-4 (80), 7AS-5 (75), 6AS-1 (72), and 5DS-2 (70). In contrast, some chromosomal regions affected the volume of very few spots like 1AS-1, 5BS-6, and 7DL-2 with only 2 spots affected, 5AL-23 and 4DS-1 with only 1 spot affected, and 2BL-3 with no varying agl spots.

Qualitative variations in agl proteins from CS deletion lines

Image analysis of the 2DE gels of endosperm proteins from the 67 DLs revealed that the absence of 211 agl spots (Fig. 1) correlated with one or more deletion bin and 101 of these proteins were identified. 175 of the disappearing

Table 2 Number of spot variations observed relative to each DL

DL	Number of variations di-i-de-to	DL	Number of variations di-i-de-to	DL	Number of variations di-i-de-to
1AL-1	6-8-20-34	1BS-9	6-2-5-13	1DS-5	4-3-8-15
1AS-1	2-0-0-2	1BS-10	2-13-16-31	1DS-1	2-5-3-10
2AS-5	4-5-11-20	1BL-1	8-12-17-37	1DL-4	10-34-36-80
2AL-1	0-2-6-8	2BS-3	10-2-4-16	2DS-5	5-6-5-16
3AS-4	2-26-29-57	2BS-1	11-3-7-21	2DS-1	0-4-7-11
3AS-2	0-4-9-13	2BL-3	0-0-0-0	2DL-9	0-1-5-6
3AL-2	5-3-10-18	2BL-1	1-18-13-32	3DL-3	20-7-17-44
3AL-5	3-5-8-16	3BS-3	2-0-3-5	4DS-3	8-4-3-15
4AS-1	6-6-11-23	3BS-7	0-1-2-3	4DS-1	0-0-1-1
4AL-12	0-10-16-26	3BS-4	9-6-17-32	4DL-13	1-2-7-10
4AL-13	4-8-16-28	3BS-2	0-5-14-19	5DS-2	1-47-22-70
5AS-3	3-0-1-4	3BS-1	0-11-25-36	5DS-1	3-8-7-18
5AL-12	1-21-16-38	3BS-5	0-10-17-27	5DL-1	2-11-16-29
5AL-10	9-0-0-9	3BL-2	0-5-13-18	5DL-5	12-6-10-28
5AL-17	0-4-0-4	3BL-8	1-22-32-55	6DS-2	11-18-17-46
5AL-23	1-0-0-1	3BL-10	0-12-25-37	6DL-6	0-3-1-4
6AS-1	7-34-31-72	3BL-7	6-12-24-42	7DS-4	10-2-21-33
7AS-5	8-45-22-75	4BS-1	19-4-8-31	7DS-5	0-1-14-15
7AL-1	2-1-2-5	4BL-1	15-14-24-53	7DL-5	0-0-0-0
		5BS-6	2-0-0-2	7DL-2	2-0-0-2
		5BS-5	3-46-37-86		
		5BS-8	1-2-8-11		
		5BL-6	4-1-1-6		
		6BS-2	9-18-20-47		
		6BS-5	0-3-9-12		
		6BL-5	1-8-3-12		
		7BS-1	3-0-4-7		
		7BL-2	3-14-30-47		
Total A	63-182-208-453	Total B	116-244-378-738	Total D	91-162-200-453

di spot disappeared, *i* spot increased, *de* spot decreased, *to* total

spots are unique to one DL (Table 3), many of which were assigned to the same chromosome bin, so they probably correspond to multiple protein isoforms produced by posttranslational modifications or from a gene cluster. For example, among the spots 1082, 1084, 1088, 1104, 1108, and 1109 which were absent in 1DL-4 (Table 3), five were identified as seed globulins (Table in electronic supplementary material). Three spots that migrated close together on the gel (516, 518 and 2140) were absent from DL 2DS-5 and identified as β-glucosidase. Three out of five spots (1793, 1795, 1796, 1801 and 1810) absent from DL 3BS-4 were identified as a dimeric α-amylase inhibitor. Four spots (875, 876, 886, and 922) absent from DL 7AS-5 were identified as peroxidases. Some likely homoeologous spots were also revealed. For example, 2130, 2140 and 2143, identified as β-glucosidase, were absent from DL 2BS-1, 2DS-5 and 2AS-5, respectively. Spots 1739 and 1741, matching dimeric α-amylase inhibitor 0.19, were absent from DLs 3AS-4 and 3BS-4, respectively.

Another 36 agl spots disappeared in several (from two to seven) DLs (Table 4), indicating the complexity of the mechanisms involved in the regulation of the expression of a given agl protein in the hexaploid wheat genome. For example, the four spots 677, 793 (serpin), 854 (serpin-Z2B) and 872 were absent from both DLs 4BS-1 and 5AL-10 and the four spots 1033, 1231, 1236 and 1459 were absent from both DL 2BS-3 and 3DL-3. The absence of some spots from DLs may be related to homoeologous genes e.g. spots 646, 960 and 1510 (eukaryotic small heat shock proteins) absent from DLs 4AS-1 and 4BS-1, spots 1499 (purothionin A-1) and 1510 absent from DLs 4AL-13 and 4BL-1 and spot 1080 absent from DLs 6AS-1 and 6DS-2. The spot 1460 was not only absent from DLs 1BL-1 and 1DL-4 but also from 7AL-1 and 7BL-2.

Table 3 List of 175 protein spots assigned to a unique chromosome bin

DL	Spots that disappeared	Total
1AL-1	283, 1041, 1049, 1116, 1325	5
1AS-1	1394, 1402	2
1BL-1	263, 947, 1140, 1261, 1409	5
1BS-9	629, 1005, 1023, 1315	4
1DL-4	654, 1082, 1084, 1088, 1104, 1108,1109, 1331, 1363	9
1DS-1	637, 641	2
1DS-5	1361, 1372, 1389	3
2AS-5	2143, 2145, 2146, 2147	4
2BL-1	1350	1
2BS-1	492, 1076, 1334, 1660, 2130, 2133, 2134, 2135	8
2BS-3	279, 607, 1143, 1375	4
2DS-5	516, 518, 2139, 2140, 2141	5
3AL-2	363, 719, 1099, 1431, 1435	5
3AL-5	841, 1826	2
3AS-4	1741	1
3BL-7	538, 586, 670, 701, 819, 898	6
3BL-8	825	1
3BS-3	1646	1
3BS-4	1568, 1739, 1744, 1793, 1795, 1796, 1801, 1810	8
3DL-3	345, 584, 595, 700, 760, 847, 1011, 1204	8
4AL-13	1628	1
4AS-1	1317, 1345	2
4BL-1	630, 1291, 1590, 1605, 1668, 1685, 1716, 1718, 1732	9
4BS-1	496, 510, 528, 804, 824, 945, 1168	7
4DL-13	1569	1
4DS-3	282, 1462, 1470, 1491, 1551, 1592	6
5AL-10	668, 708, 2158, 2160, 2161	5
5AL-12	1173	1
5BL-6	840, 1520, 1899	3
5BS-5	1477	1
5DL-1	661	1
5DL-5	535, 537, 539, 557, 650, 656, 703, 717, 740, 768, 796, 1087	12
5DS-2	337	1
5DS-1	573	1
6AS-1	969	1
6BS-2	652, 735, 846, 864, 868, 1455, 1869	7
6BL-5	835	1
6DS-2	890, 1061, 1835, 1838, 1843, 1844, 1847, 1850, 1856, 1951	10
7AL-1	1811	1
7AS-5	745, 875, 876, 879, 886, 922	6
7BL-2	1121, 1410	2
7BS-1	1135, 1890	2
7DL-2	1895	1
7DS-4	390, 398, 849, 1030, 1765, 1786, 1840, 1865, 1880	9
Total		175

The loss of some spots in DLs correlated with deletion bins located on the long and the short arm of the same chromosome, e.g., spots 1426 and 1460 for DLs 1BL-1 and 1BS-9, spot 359 for DLs 3AL-5 and 3AS-4, spots 375, 618, 889, 1510 for 4BL-1 and 4BS-1, and spot 1303 (thaumatin-like protein) for 5BS-6 and 5BL-6 (Table 4).

Quantitative variations in agl proteins in CS deletion lines

Of the 1,145 spots analysed, 391 spot volumes varied significantly between DLs and the CS parental line. The expression of each of these proteins, either activation (the spot volume decreased in DLs) or repression (the spot volume increased in DLs), was under the control of one or several regulators located in missing chromosome bins. A similar explanation can be used to account for the only three unidentified spots to appear, spots 883 (in 3AS-4), 1009 (in 5BS-5) and 1418 (in 5DS-2), i.e., the structural genes or regulators are activated in DLs though silent or expressed below our level of detection in CS (Desfrancs and Thiellement 1985).

Among these 391 spots, 150 spots varied (100 i and 50 d) in only one DL, 76 spots in two DLs, 42 spots in three DLs (Table 5) and 123 spots in at least four DLs. We noted that the regulatory influence on spot volume was not equally distributed for the deletion bins analysed: 262 chromosome bins had regulators that decreased spot volume while 196 chromosome bins had regulators that increased spot volume. Numerous proteins had their expression augmented by the same deletion bin, e.g. 5BS-5 and 5DS-2 significantly influenced the spot volume of 13 and 22 spots respectively (Table 5). Some quantitative variations were in opposite directions, increased for some proteins and decreased for others. For example, spots 847 (reductase 1), 855 (alcohol dehydrogenase) and 861 (allyl alcohol dehydrogenase) were increased, while spots 825 (aspartate aminotransferase) and 828 were reduced when the chromosome bin 1DL-4 was absent. This suggests the presence of an inhibitor and an activator in that CS chromosome bin. The inhibitory effect in the normal CS bin was also deduced for spots 913 and 914, which increased when 3AS-4 was missing, spots 187 (cell division cycle protein 48) and 189, which increased when 5DS-2 was absent, spots 267 and 268 (two methionine synthases), which increased when 3BL-7 was absent.

Table 4 List of 36 protein spots assigned to several chromosome bins

Spot no.	Deletion lines
359	3AS-4, 3AL-5, 5AL-23
375	3DL-3, 4BL-1, 4BS-1
578	3DL-3, 4BS-1
618	1DS-5, 4BL-1, 4BS-1
646	4AS-1, 4BS-1
677	4BS-1, 5AL-10
793	4BS-1, 5AL-10
838	2BS-1, 7AS-5
854	4BS-1, 5AL-10
872	4BS-1, 5AL-10
884	5DS-1, 6AS-1
889	4BL-1, 4BS-1
955	3DL-3, 6AS-1
960	1BS-10, 4AL-13, 4AS-1, 4BS-1
965	3DL-3, 7AS-5
986	3BS-4, 5BS-8
1033	2BS-3, 3DL-3, 7DS-4
1080	3DL-3, 6AS-1, 6DS-2
1100	5DS-1, 6AS-1
1201	2BS-1, 7DL-2
1231	2BS-3, 3DL-3, 6AS-1
1236	2BS-3, 3DL-3
1247	2BS-1, 3DL-3
1303	5BL-6, 5BS-6
1321	4BS-1, 5AS-3
1403	2BS-3, 5DL-1
1426	1BL-1, 1BS-9, 2BS-3
1438	1BL-1, 3DL-3, 4BL-1, 6AS-1
1459	2BS-3, 3DL-3
1460	1BL-1, 1BS-9, 1DL-4, 3DL-3, 7AL-1, 7BL-2, 7BS-1
1489	4DS-3, 5BS-6
1499	1AL-1, 4AL-13, 4BL-1, 4AS-1
1507	4DS-3, 5BS-5
1510	3BS-3, 4AS-1, 4AL-13, 4BL-1, 4AS-1, 5BS-5
1814	1BS-10, 5AS-3, 6BS-2
1861	5AS-3, 6BS-2

Among the 391 significantly varying spots, 108 had coding sequences (structural genes), responsible for their presence/absence assigned on at least one DL (Table 6). The volume of 40.7 % of these spots always increased and the volume of 43.5 % always decreased however many deletion bins were involved. Only 15.7 % had spot volumes that either increased or decreased according to the deletion bin concerned. Spot 1402, which has its structural gene on DL 1AS-1, was decreased when the 1AL-1 bin was absent suggesting the presence of an activator in the 1AL-1 bin. Spot 1741 (dimeric α-amylase inhibitor 0.19), which was absent from DL 3AS-4 had a significantly larger spot volume when the "homoeologous" chromosome bin 3BS-1 was absent, indicating the presence of an inhibitor in that bin. Similarly, spot 879 in CS was absent in DL 7AS-5 yet augmented when 7DS-5 was absent. Quantitative variations were also found between "homoeologous" DLs like 5BS-5 and 5DS-2 (Table 5) which lowered the expression of seven proteins (407, 450, 456, 612, 1204, 1292 and 1514), two of which were identified as enzymes involved in carbohydrate metabolism. The expression of some proteins was always increased or always decreased by numerous deletion lines. Spots 793, 824 and 840 (all serpins) were always augmented (i.e., the expression is inhibited in CS) by 26, 12 and 11 DLs, respectively. In contrast, spots 890, 930, 2172, and 2173 (gliadin) and spot 1020 (a glucose and ribitol dehydrogenase) were always reduced (i.e., expression is increased in CS) by 17, 22, 20, 28 and 46 DLs, respectively.

Discussion

There are 11875 *Triticum aestivum* protein entries in NCBInr (version February 2012) of which only 320 are from the endosperm. To assess the level of proteome coverage achieved in our approach, a virtual 2DE of these 320 predicted endosperm protein was run, revealing that they span a pH region from 4 to 10 and molecular weights from 10 to 90 kDa, which fall within the ranges of our experimental parameters. However, proteome coverage can be limited by the dynamic range of the expressed protein, i.e. the difference between the most abundant and least abundant proteins. This is due to the technological difficulty of detecting very low abundance proteins in the presence of very high ones. The low number of endosperm proteins in NCBInr to some extent explains the low success rate in protein identification and the necessity of using alternative databases. Indeed, 34 % of the proteins identified here matched EST sequences and 102 spots did not match any sequence in the databases searched despite producing clear peptide mass fingerprints (PMF) and de novo sequencing spectra. Moreover, 143 proteins were identified based on a single peptide despite generating several good quality tandem MS spectra. Proteins identified with a single peptide were manually validated after the observation of at least three *b* and *y* ions in a row in the MSMS spectrum. To investigate this further, the identified proteins were digested in silico and the resultant peptides were searched against the acquired PMF spectra. Surprisingly, most of the predicted peptides were absent from our PMF spectra explaining the difficulty of identifying the proteins by PMF. We suggest two possible reasons for this: (1) proteins present in our gels are absent from the searched

Table 5 Quantitative spot variations affected by one, two or three DLs

Variations	Spots affected	Variations	Spots affected	Variations	Spots affected	Variations	Spots affected	Variations	Spots affected	Variations	Spots affected
↓ 1AL-1	1205, 1374, 1402	↑ 5BS-5	417, 479, 528, 577, 626, 635, 841, 961, 964, 1132, 1158, 1200, 1391	↑ 1AL-1 ↑ 3BL-7	2162	↑ 7AS-5 ↑ 7BL-2	1320	↑ 1AL-1 ↑ 3BL-8 ↑ 6BS-2	1489	↓ 3AL-2 ↑ 5BS-5 ↑ 5DS-2	850
↑ 1BS-10	717, 729	↓ 5BS-5	557, 1222, 1668, 2102	↑ 1BS-10 ↑ 3BS-2	2160	↓ 1BL-1 ↓ 1DL-4	1176	↑ 1BS-10 ↑ 5BS-5 ↑ 5DS-2	959	↓ 3AS-4 ↓ 7AS-5 ↓ 7DS-5	1022
↓ 1BS-10	945, 1333, 1715, 1795	↑ 5DL-1	277	↑ 1DL-4 ↑ 6AS-1	494	↓ 1DL-4 ↓ 3BL-8	282	↑ 1BS-9 ↑ 1BS-10 ↑ 6AS-1	849	↓ 3AS-4 ↓ 5DL-1 ↓ 5DL-5	339, 660
↑ 1BL-1	2154	↑ 5DS-1	1183	↑ 1DL-4 ↑ 7AS-5	1277	↓ 1DL-4 ↓ 4BL-1	240	↑ 1DL-4 ↑ 3BL-10 ↑ 3BL-7	2165	↓ 3BL-10 ↓ 3BL-7 ↓ 5BS-5	754
↓ 1BL-1	960	↓ 5DS-1	897, 2169	↑ 2BL-1 ↓ 3BS-1	1534	↓ 1DL-4 ↓ 6AS-1	902, 903	↑ 1DL-4 ↑ 6AS-1 ↑ 6BL-6	2120	↓ 3BL-10 ↓ 5BS-5 ↑ 7AS-5	263
↑ 1DL-4	847, 855, 861, 1340	↑ 5DS-2	179, 187, 189, 326, 366, 370, 489, 514, 529, 599, 616, 628, 631, 638, 678, 821, 834, 1133, 1404, 1413, 1418 (new), 1515	↑ 2BL-1 ↓ 3AS-4	1716	↓ 2BL-1 ↓ 3AL-5	1002	↑ 1DL-4 ↑ 4BL-1 ↑ 5AL-12	1910	↓ 3BL-2 ↓ 5DS-2 ↓ 6BS-5	1499
↓ 1DL-4	241, 825, 828	↓ 5DS-2	765, 1116, 1850, 1892, 1898	↑ 2BL-1 ↑ 6AS-1	881	↓ 3AS-4 ↓ 6DS-2	656	↑ 1DL-4 ↑ 3BL-8 ↑ 7AS-5	1544	↓ 3BS-4 ↓ 3BS-2 ↓ 5BS-5	1341
↑ 2AS-5	1895	↑ 6AS-1	477, 528, 843, 1335, 1723, 2136, 2137, 2146	↑ 2BS-3 ↑ 5AL-12	1233	↓ 3AS-4 ↓ 3BL-10	2130	↑ 2BL-1 ↓ 3AS-4 ↑ 6BL-6	522	↓ 4AL-13 ↓ 5BS-5 ↓ 7AL-1	1212
↑ 2BL-1	679, 1141, 1191, 1272, 1298, 1532, 1541, 1542	↓ 6AS-1	823, 2166	↑ 3AS-4 (new) ↑ 3AL-5	883	↓ 3BL-10 ↑ 5AL-12	827	↑ 2BL-1 ↓ 3AS-4 ↓ 6BS-2	763	↓ 5AL-12 ↑ 6DS-2 ↑ 7AS-5	495
↓ 2BL-1	344, 629	↓ 6BL-5	907	↑ 3AS-4 ↑ 3BL-8	1055, 1491	↓ 3BL-10 ↓ 3BL-7	539	↑ 2BS-1 ↑ 3BS-1 ↑ 7BL-2	1833	↓ 7AS-5 ↓ 7BL-2 ↓ 7DS-4	835
↓ 2BS-1	1507	↑ 6BS-2	1519, 1537	↑ 3AS-4 ↑ 5DS-2	749, 1182	↓ 3BL-7 ↑ 7AS-5	535	↑ 3BL-10 ↑ 3BL-7 ↑ 4AS-1	802		
↑ 2DS-1	798	↓ 6BS-2	865	↑ 3AS-4 ↑ 6DS-2	1870	↓ 3BS-1 ↑ 7AL-1	838	↑ 3BL-2 ↑ 3BL-10 ↑ 3BL-7	817		
↑ 3AL-2	950, 1084	↑ 6BS-5	1028	↑ 3AS-4 ↑ 7AS-5	1004, 1592	↓ 3BS-1 ↓ 3BL-7	884	↑ 3BL-7 ↑ 4DS-3 ↑ 5DS-2	1288		

Table 5 continued

Variations	Spots affected	Variations	Spots affected	Variations	Spots affected	Variations	Spots affected	Variations	Spots affected	Variations	Spots affected
↑ 3AS-4	168, 350, 371, 597, 913, 914, 1886	↑ 6DL-6	1134	↑ 3BL-8 ↑ 5BS-5 (new)	1009	↓ 3BS-1 ↓ 3BL-8	2135	↑ 3BL-8 ↑ 6BS-2 ↑ 7AS-5	1477		
↓ 3AS-4	650, 989, 1104	↑ 6DS-2	751, 1704	↑ 4AL-12 ↑ 6BS-2	429	↓ 3BS-2 ↑ 7AS-5	646	↑ 3BS-2 ↑ 5BS-5 ↑ 5DS-2	472		
↑ 3BL-10	782	↓ 6DS-2	703	↑ 4AL-12 ↓ 4DL-13	2111	↓ 3BS-4 ↓ 5BS-5	296	↑ 3BS-2 ↓ 7DS-4 ↓ 7DS-5	756		
↑ 3BL-2	1174	↑ 7AS-5	705, 1099, 1188, 1286, 1450, 1468, 1521	↑ 4AL-13 ↑ 5BL-6	1282	↓ 3DL-3 ↓ 5BS-5	1817	↑ 4BL-1 ↑ 6DS-2 ↑ 7BL-2	1905		
↑ 3BL-7	267, 268, 2164	↓ 7AS-5	1109, 1847	↑ 4BS-1 ↑ 5BS-8	1332	↓ 3DL-3 ↓ 5DS-2	1334	↑ 4BL-1 ↑ 6BL-3 ↑ 7BL-2	1303		
↑ 3BL-8	1175, 1463, 1488, 1528	↑ 7BL-2	648, 1088	↑ 4DS-3 ↑ 5BS-5	1317	↓ 4AL-13 ↓ 5BS-5	1220	↑ 5AL-12 ↑ 5BS-5 ↓ 6BS-2	636		
↓ 3BL-8	403, 2145	↓ 7BL-2	1822	↑ 5AL-12 ↓ 6BS-2	953	↓ 4DL-13 ↑ 6AS-1	2114, 2115	↑ 5DL-1 ↑ 5DL-5 ↑ 6BS-5	1030		
↑ 3BS-1	947	↓ 7BS-1	860, 1787	↑ 5BS-5 ↑ 5DS-2	407, 450, 456, 612, 1204, 1292, 1514	↓ 5AL-12 ↑ 5DS-2	777	↑ 5DL-5 ↑ 6AS-1 ↑ 6DS-2	2116		
↓ 3BS-1	1741	↓ 7DS-4	532	↑ 5BS-5 ↑ 6BL-5	695	↓ 5AL-12 ↓ 6AS-1	1485	↑ 5DS-1 ↑ 6DS-2 ↑ 7AS-5	1008		
↓ 3BS-2	1738	↑ 7DS-5	879	↑ 5BS-8 ↑ 6BL-5	886	↓ 5BS-5 ↓ 5DS-2	1467, 1495	↓ 1AL-1 ↓ 1BL-1 ↓ 2BL-1	906		
↑ 3BS-4	1930	↓ 7DS-5	728	↑ 5DL-1 ↑ 7AS-5	451	↓ 5BS-5 ↑ 7AS-5	1258	↓ 1AL-1 ↓ 3BL-7 ↓ 6DS-2	874		
↓ 3BS-4	283			↑ 5DL-1 ↑ 7BL-2	437	↓ 5DS-2 ↓ 6AS-1	2168	↓ 1BL-1 ↓ 1DL-4 ↓ 7BL-2	922		
↓4AL-13	852			↑ 5DS-2 ↓ 6AS-1	222, 224	↓ 5DS-2 ↓ 6BS-5	1872	↓ 1BS-10 ↓ 5DS-2 ↑ 7DS-4	1801		

Table 5 continued

Variations	Spots affected	Variations	Spots affected								
				Variations	Spots affected	Variations	Spots affected	Variations	Spots affected	Variations	Spots affected
↓ 4AS-1	367			↑ 6AS-1 ↓ 7BL-2	1431	↓ 6BS-2 ↓ 6BS-5	775	↓ 1DL-4 ↑ 3BL-10 ↑ 5BS-5	856		
↓4DL-13	2107, 2110			↑ 6AS-1 ↑ 7DS-4	1403	↓ 6BS-2 ↓ 7AS-5	1843, 2133	↓ 1DL-4 ↓ 3BL-8 ↓ 4BL-1	239		
↑5AL-12	443, 830, 979			↑ 6BS-2 ↑ 7AS-5	1510	↓ 7AS-5 ↓ 7BL-2	693	↓ 1DL-4 ↓ 4BL-1 ↓ 7BL-2	250		
↓5AL-12	301			↑ 6DS-2 ↑ 7AS-5	488, 496, 657	↓ 7DS-4 ↓ 7DS-5	747	↓ 2DS-5 ↓ 2DS-1 ↑ 6BL-6	2149, 2150		

↓ decreased spot volume,
↑ increased spot volume

databases and only very weak homologs are matched; (2) the majority of the matched proteins are ESTs or TCs which are known to have sequence errors or frame shifts.

A total of 1145 agl spots from mature starchy endosperm were revealed by CBB staining, and only 211 were absent due to a missing chromosome bin. This means that the structural gene coding for the protein or a regulator necessary for the expression of the gene could be located on that particular chromosome bin (Desfrancs and Thiellement 1985). This low number (18.6 %) of assigned proteins is similar to previous reports (Islam et al. 2002; Thiellement et al. 1986). The allo-hexaploid structure of the wheat genome with numerous homoeologous coding sequences between the three genomes, and the possibility of multiple locations of regulators controlling a given protein, are major factors hampering the localisation of structural genes (Islam et al. 2002; Zivy et al. 1984). Of the 391 agl proteins differing in spot volume and the 211 proteins qualitatively varying between DLs and CS, a total of 386 proteins were influenced by at least one DL. The remaining 759 spots (66 % of the CBB-stained agl) did not vary qualitatively or quantitatively. First, this is likely to be related to the fact that only 67 DLs were used which do not cover all the (homoeologous) chromosomes. Secondly, the stringency of statistical analyses and particularly the fold difference ≥ 2 rejected spots that may have differed to a lesser extent between a DL and CS. In addition, the low percentage of varying proteins observed could also be explained by gene dosage compensation (activation or repression), which regulates the level of protein expression independently of the number of active structural genes present (Bottley et al. 2006; Galili et al. 1986).

The influence of each genome on the proteome is often discussed. The B genome is responsible for the greatest variations in spots (Desfrancs and Thiellement 1985), as we found here, where 26.3 % of spot variations per DL were due to the B genome as compared to 23.8 and 22.2 % for the A and D genomes respectively. Some regions were responsible for many varying spots. Some like 5BS-5 and 5DS-2 were located in distal regions composed of clusters of gene-rich regions (Moore 2000). Inversely, the expression of some proteins was influenced by numerous DL bins (e.g. serpin, alpha- or beta-gliadin, glucose and ribitol dehydrogenase) distributed over almost all the chromosomes. This cannot be fully explained by gene redundancy or complex mechanisms involving numerous chromosome sites. Possibly some metabolic disturbances occurring during seed development and storage compound accumulation, resulting from incomplete compensation due to the lacking bin, could cause such specific variations, e.g., serpin may increase for stress defence and glucose and

Table 6 Quantitative variations of chromosome assigned spots and their identification

108 spots	Chromosomal assignment	Quantitative variations	Identifications
1116	1AL-1	↓ 5DS-2	Class III chitinase
283	1AL-1	↓ 3BS-4	
1325	1AL-1	↑ 3AS-4 ↑ 3BS-1 ↑ 4DL-13 ↑ 6BS-2	
1499	1AL-1, 4AL-13, 4AS-1, 4BL-1	↓ 3BL-2 ↓ 5DS-2 ↓ 6BS-5	Purothionin A-1
1402	1AS-1	↓ 1AL-1	
947	1BL-1	↑ 3BS-1	Aldose reductase
263	1BL-1	↓ 3BL-10 ↓ 5BS-5 ↑ 7AS-5	
960	1BS-10, 4AL-13, 4AS-1, 4BS-1	↓ 1BL-1	
1814	1BS-10, 5AS-3, 6BS-2, 6BL-3	↓ 1AL-1 ↓ 1DL-4 ↓ 3BS-5 ↓ 3BL-8 ↓ 3BL-7 ↓ 5BS-5 ↓ 6AS-1	Pathogenesis-related protein
629	1BS-9	↓ 2BL-1	
1023	1BS-9	↑ 3BS-1 ↑ 3BL-8 ↓ 5BS-5 ↑ 6DS-2	
1084	1DL-4	↑ 3AL-2	
1088	1DL-4	↑ 7BL-2	Seed globulin
1104	1DL-4	↓ 3AS-4	Seed globulin
1331	1DL-4	↑ 1BL-1 ↑ 3AS-4 ↑ 3BS-1 ↑ 4DL-13	
1108	1DL-4	↑ 3BL-8 ↑ 5DL-1 ↑ 6BS-2 ↑ 6BS-5 ↑ 7BL-2	Seed globulin
1109	1DL-4	↓ 7AS-5	Seed globulin
637	1DS-1	↓ 3 BS-1 ↓ 4AL-12 ↓ 4BL-1 ↓ 5AL-12 ↓ 5DL-1 ↓ 7AS-5 ↓ 7BL-2 ↓ 7DS-5	Triticin precursor
641	1DS-1	↓ 3BS-1 ↓ 3BL-10 ↓ 4AL-12 ↓ 5DL-1 ↑ 6AS-1 ↓ 6BS-5 ↓ 7BL-2 ↓ 7DS-5	
2146	2AS-5	↑ 6AS-1	
2145	2AS-5	↓ 3BL-8	
1334	2BS-1	↓ 3DL-3 ↓ 5DS-2	Secretory protein
2130	2BS-1	↓ 3AS-4 ↓ 3BL-10	Beta-glucosidase
2135	2BS-1	↓ 3BS-1 ↓ 3BL-8	Beta-glucosidase
2133	2BS-1	↓ 6BS-2 ↓ 7AS-5	Beta-glucosidase
2134	2BS-1	↓ 3AS-4 ↓ 3BL-8 ↓ 3BL-10 ↓ 4AL-13	Beta-glucosidase
838	2BS-1, 7AS-5	↓ 3BS-1 ↑ 7AL-1	Peroxidase 1
1403	2BS-3, 5DL-1	↑ 6AS-1 ↑ 7DS-4	
1431	3AL-2	↑ 6AS-1 ↓ 7BL-2	
1435	3AL-2	↑ 5BS-5 ↑ 5DS-2 ↓ 6BS-2 ↓ 7AS-5 ↓ 7BL-2	
719	3AL-2	↓ 2BS-3 ↓ 2BS-1 ↓ 3BS-4 ↓ 3BL-10 ↓ 4AS-1 ↑ 5AL-12 ↑ 5BS-5	
1099	3AL-2	↑ 7AS-5	

Table 6 continued

108 spots	Chromosomal assignment	Quantitative variations	Identifications
841	3AL-5	↑ 5BS-5	Fructose 1,6-bisphosphate aldolase
1741	3AS-4	↓ 3BS-1	
701	3BL-7	↑ 1DS-5 ↑ 2AL-1 ↑ 2BL-1 ↑ 2DL-9 ↑ 6AS-1	EST with homology to putative pectinacetylesterase precursor
825	3BL-8	↓ 1DL-4	Aspartate aminotransferase
1801	3BS-4	↓ 1BS-10 ↓ 5DS-2 ↑ 7DS-4	0.19 dimeric alpha-amylase inhibitor
1795	3BS-4	↓ 1BS-10	0.19 dimeric alpha-amylase inhibitor
986	3BS-4, 5BS-8	↓ 1AL-1 ↓ 1BS-10 ↓ 1BL-1 ↓ 1DS-1 ↓ 2AS-5 ↓ 2BL-1 ↓ 3BL-8 ↓ 5DS-1 ↓ 6AS-1	
847	3DL-3	↑ 1DL-4	Reductase 1
1204	3DL-3	↑ 5BS-5 ↑ 5DS-2	Triosephosphate isomerase
1438	3DL-3, 1BL-1, 4BL-1, 6AS-1	↓ 1DL-4 ↑ 3BL-10 ↓ 5AL-12 ↓ 5DL-1	Alpha-1-purothionin
955	3DL-3, 6AS-1	↓ 1AL-1 ↓ 1BS-10 ↓ 1BL-1↓ 1DL-4 ↓ 2AL-1 ↓ 2BS-3 ↓ 2BL-1 ↓ 2DL-9 ↓ 3AS-4 ↓ 3AS-2 ↓ 3BS-3 ↓ 3BL-10 ↓ 5BS-5	
375	3DL-3, 4BS-1, 4BL-1	↓ 1AL-1 ↓ 1DL-4 ↓ 3BS-5 ↓ 3BL-2 ↓ 3BL-8 ↓ 5DL-1	Pyrophosphate-fructose 6-phosphate 1-phosphotransferase alpha subunit
1628	4AL-13	↑ 3BL-8 ↓ 5BS-5 ↓ 5BS-8 ↑ 7AS-5	
1317	4AS-1	↑ 4DS-3 ↑ 5BS-5	
646	4AS-1, 4BS-1	↓ 3BS-2 ↑ 7AS-5	
1716	4BL-1	↑ 2BL-1 ↓ 3AS-4	
1668	4BL-1	↓ 5BS-5	Alpha-amylase/trypsin inhibitor CM3
824	4BS-1	↑ 1BL-1 ↑ 1DS-1 ↑ 1DL-4 ↑ 2DS-5 ↑ 2DS-1 ↑ 3BS-5 ↑ 4AS-1 ↑ 4AL-12 ↑ 4BL-1 ↑ 5AL-17 ↑ 6AS-1 ↑ 6BL-5	Serpin
804	4BS-1	↑ 1DL-4 ↑ 3DL-3 ↑ 4AL-12 ↑ 4AL-13 ↑ 4BL-1 ↑ 6AS-1	
528	4BS-1	↑ 6AS-1	Beta-amylase precursor
496	4BS-1	↑ 6DS-2 ↑ 7AS-5	Globulin-like protein
510	4BS-1	↓ 3BS-2 ↓ 3BL-10 ↓ 3BL-7 ↓ 5AL-12 ↑ 6DS-2	Globulin-like protein
945	4BS-1	↓ 1BS-10	
889	4BS-1, 4BL-1	↓ 3AS-4 ↓ 3AS-2 ↓ 3AL-5 ↓ 3BS-5 ↓ 3BL-2 ↓ 3BL-8 ↓ 3BL-10 ↓ 3BL-7 ↓ 4AS-1 ↓ 5BS-8 ↓ 5DS-1 ↓ 5DL-1 ↓ 6AS-1	
1592	4DS-3	↑ 3AS-4 ↑ 7AS-5	
1491	4DS-3	↑ 3AS-4 ↑ 3BL-8	
282	4DS-3	↓ 1DL-4 ↓ 3BL-8	
1462	4DS-3	↑ 1DL-4 ↑ 3BL-8 ↑ 6BS-2 ↑ 7AS-5	

Table 6 continued

108 spots	Chromosomal assignment	Quantitative variations	Identifications
1551	4DS-3	↑ 1DL-4 ↑ 3AS-4 ↑ 3BL-8 ↓ 4BS-1 ↓ 5BS-5 ↑ 7AS-5 ↓ 7DS-5	
1470	4DS-3	↑ 3BL-8 ↓ 4BS-1 ↓ 5AL-12 ↓ 5BS-5 ↓ 5BS-8 ↑ 7AS-5 ↓ 7DS-5	
2160	5AL-10	↑ 1BS-10 ↑ 3BS-2	Aspartate aminotransferase
2161	5AL-10	↑ 1BS-10 ↑ 3BS-4 ↑ 3BS-2 ↑ 3DL-3 ↓ 5AL-12 ↑ 5BS-5	
2158	5AL-10	↑ 1BS-10 ↑ 3BS-5 ↑ 3DL-3 ↑ 4AL-12 ↑ 4AL-13 ↑ 5BS-5	Serpin
668	5AL-10	↑ 3AL-5 ↑ 3BL-10 ↑ 3BL-7 ↑ 4AS-1 ↑ 4AL-12 ↑ 4AL-13	Serpin
872	5AL-10, 4BS-1	↑ 1AL-1 ↑ 1BS-10 ↑ 1BL-1 ↑ 1DS-1 ↑ 1DL-4 ↑ 2DS-5 ↑ 3AL-5 ↑ 3BL-7 ↑ 3DL-3 ↑ 4AS-1 ↑ 4BL-1 ↑ 5AL-17	
793	5AL-10, 4BS-1	↑ 1AL-1 ↑ 1BS-10 ↑ 1DS-5 ↑ 1DS-1 ↑ 1DL-4 ↑ 2AS-5 ↑ 2BS-1 ↑ 2DS-5 ↑ 2DS-1 ↑ 3AS-2 ↑ 3AL-5 ↑ 3BS-7 ↑ 3BS-4 ↑ 3BS-1 ↑ 3BS-5 ↑ 3BL-2 ↑ 3BL-10 ↑ 3DL-3 ↑ 4AS-1 ↑ 4AL-12 ↑ 5AL-17 ↓ 5DS-2 ↑ 6AS-1 ↑ 6BL-5	Serpin
854	5AL-10, 4BS-1	↑ 1DS-1 ↑ 1DL-4 ↑ 2BS-3 ↑ 4AS-1 ↑ 5DL-5 ↑ 6BL-3	Serpin-Z2B
677	5AL-10, 4BS-1	↑ 2DS-5 ↑ 3AL-5 ↑ 3BL-7 ↑ 4AL-13 ↑ 5BS-5	
1321	5AS-3, 4BS-1	↑ 1DL-4 ↑ 5DL-1 ↑ 6BL-3 ↑ 7AS-5 ↑ 7BL-2	Permatin precursor
840	5BL-6	↑ 1BS-10 ↑ 1BL-1 ↑ 1DL-4 ↑ 3BL-2 ↑ 3BL-10 ↑ 4BL-1 ↑ 5AL-12 ↑ 5AL-17 ↑ 6AS-1 ↑ 6BL-5 ↑ 6DS-2	Serpin
1899	5BL-6	↑ 1DL-4 ↑ 4BL-1 ↑ 5AL-12 ↓ 5BS-5 ↓ 5DS-2	Nonspecific lipid-transfer protein precursor (LTP)
1303	5BL-6, 5BS-6	↑ 4BL-1 ↑ 6BL-3 ↑ 7BL-2	Thaumatin-like protein
1477	5BS-5	↑ 3BL-8 ↑ 6BS-2 ↑ 7AS-5	
1510	5BS-5, 4AS-1, 4AL-13, 4BS-1, 4BL-1, 3BS-3	↑ 6BS-2 ↑ 7AS-5	Eukaryotic Small Heat Shock Protein
1507	5BS-5, 4DS-3	↓ 2BS-1	
1489	5BS-6, 4DS-3	↑ 1AL-1 ↑ 3BL-8 ↑ 6BS-2	
557	5DL-5	↓ 5BS-5	EST with no homologue
717	5DL-5	↑ 1BS-10	
703	5DL-5	↓ 6DS-2	Serpin
656	5DL-5	↓ 3AS-4 ↓ 6DS-2	
539	5DL-5	↓ 3BL-10 ↓ 3BL-7	

Table 6 continued

108 spots	Chromosomal assignment	Quantitative variations	Identifications
535	5DL-5	↓ 3BL-7 ↑ 7AS-5	Globulin-like protein
796	5DL-5	↑ 1BS-10 ↑ 4BS-1 ↑ 5AL-12 ↑ 6AS-1	Serpin
768	5DL-5	↑ 1BS-10 ↑ 4BS-1 ↑ 5AL-12 ↑ 6AS-1	Serpin
1087	5DL-5	↓ 3BS-1 ↓ 6DS-2 ↓ 7AS-5 ↓ 7BL-2 ↓ 7DS-4	
650	5DL-5	↓ 3AS-4	Serpin
573	5DS-1	↑ 3AS-2 ↑ 3BS-4 ↑ 3BS-2 ↑ 3BS-5 ↑ 5AL-12	
884	5DS-1, 6AS-1	↓ 3BS-1 ↓ 3BL-7	
969	6AS-1	↓ 1BL-1 ↓ 1DS-5 ↓ 2AS-5 ↓ 2BL-1 ↓ 3BS-1 ↓ 3BS-5 ↓ 3BL-7 ↓ 4AL-13	
835	6BL-5	↓ 7AS-5 ↓ 7BL-2 ↓ 7DS-4	tRNA uridine 5-carboxymethylaminomethyl modification enzyme GidA
735	6BS-2	↓ 1AL-1 ↓ 3BS-5 ↓ 3BL-10 ↓ 3BL-7 ↓ 5BS-5	Alpha–beta-gliadin precursor
652	6BS-2	↓ 1AL-1 ↓ 1BL-1 ↓ 1DS-5 ↓ 1DL-4 ↓ 2AS-5 ↓ 3AL-5 ↓ 3BS-4 ↓ 3BS-1 ↓ 3BS-5 ↓ 3BL-10 ↓ 4AL-13	Alpha-gliadin storage protein
1850	6DS-2	↓ 5DS-2	Alpha amylase
1835	6DS-2	↓ 3BS-5 ↓ 4AS-1 ↓ 4AL-12 ↓ 4AL-13 ↓ 4BL-1 ↓ 6BL-5	
1843	6DS-2	↓ 6BS-2 ↓ 7AS-5	
1847	6DS-2	↓ 7AS-5	Alpha-amylase inhibitor
890	6DS-2	↓ 1AL-1 ↓ 1BL-1 ↓ 1DS-5 ↓ 1DL-4 ↓ 2AS-5 ↓ 2DS-1 ↓ 3AS-4 ↓ 3AL-5 ↓ 3BS-4 ↓ 3BS-1 ↓ 3BL-2 ↓ 3BL-8 ↓ 3BL-10 ↓ 3BL-7 ↓ 3DL-3 ↓ 4AL-12 ↓ 4AL-13	Alpha–beta-gliadin MM1 precursor
745	7AS-5	↓ 3BS-1 ↓ 3DL-3 ↓ 5BS-8 ↓ 6BS-2 ↓ 6DS-2 ↓ 7DS-4	HSP70
879	7AS-5	↑ 7DS-5	
886	7AS-5	↑ 5BS-8 ↑ 6BL-5	Peroxidase 1
922	7AS-5	↓ 1BL-1 ↓ 1DL-4 ↓ 7BL-2	Peroxidase
1890	7BS-1	↑ 1DL-4 ↑ 4BL-1 ↓ 5DS-2 ↑ 5DL-1 ↑ 6AS-1 ↑ 7AS-5	
1895	7DL-2	↑ 2AS-5	
849	7DS-4	↑ 1BS-9 ↑ 1BS-10 ↑ 6AS-1	Serpin 4
1030	7DS-4	↑ 5DL-1 ↑ 5DL-5 ↑ 6BS-5	Seed storage protein (Fragment)
1840	7DS-4	↑ 2AS-5 ↑ 2BL-1 ↑ 4DS-3 ↑ 6AS-1	

↓ decreased spot volume, ↑ increased spot volume

ribitol dehydrogenase for energy metabolism, while alpha- and beta-gliadin may have been rendered more soluble as a consequence of altered storage protein polymerisation and aggregation.

Several qualitative and quantitative spot variations were associated with the homoeologous structure of the wheat genome, but additional molecular markers would be required to confirm possible homoeoallelism. Some spots, close to each other on the gel, disappeared with homoeologous deletion bins (e.g., spots 2130, 2140 and 2143 or spots 1739 and 1741). The quantity of some protein spots varied in the same sense when the protein was under the control of these homoeologous chromosome bins, e.g., expression was reduced for 1BL-1 and 1DL-4 or augmented for 5BS-5 and 5DS-2. This was not the case for all quantitative variations, since a protein can be regulated by different chromosome bins, not necessarily homoeologous. In general the localisation of the structural gene was different from these regulators. Spot 1317 had its structural gene on 4AS-1 and its expression was decreased by regulators on 4DS-3 (a possible homoeologous locus) and on 5BS-5. This kind of heterologous regulation was found for 108 spots (Table 6).

The two homoeologous deletion bins, 5BS-5 and 5DS-2, were revealed to encode regulators repressing four proteins involved in starch and sucrose metabolism. Spot 407, pyrophosphate-fructose 6-phosphate 1-phosphotransferase, and spot 1204, triose-phosphate isomerase, are involved in glycolysis. Spot 1292 is dehydroascorbate reductase, an enzyme that acts in ascorbate and aldarate metabolism in which glucose is transformed to ascorbic acid that has antioxidant properties. Spot 456 had homology with a ketol-acid reductoisomerase involved in amino acid (leucine, isoleucine or valine) biosynthesis from pyruvate produced in glycolysis. Other proteins varied similarly in quantity for these two DLs, particularly spot 1514, identified as a barley embryo globulin having homologies with a globulin 3 in wheat.

Among the 211 proteins assigned to one or more chromosome bins, some were previously identified using mapped segregating progenies (Merlino et al. 2009). For example, several serpins assigned on 5AL, 5BL and 7DS were more precisely localised on 5AL-10, 5BL-6, and 7DS-4 respectively. Other serpins were also localised on 4BS-1, 5DL-1, 5DL-5. The α-1-purothionin, previously assigned on 1BL, was localised on 1BL-1 (spot 1438), and also assigned on 3DL-3, 4BL-1 and 6AS-1. In general, the chromosomal localisation of structural genes was more precisely confirmed or newly reported here. These chromosomal localisations will be useful to precisely locate genes encoding, in particular, stress defence proteins which were increased in response to high temperature (Hurkman et al. 2009).

Conclusion

In this proteomic approach, the chromosomal regions of genes controlling 386 endosperm proteins were localised for the standard wheat cultivar Chinese Spring. These precise locations will be important for further genetic analyses focused on understanding endosperm composition and for further studies of seed enzymatic activities and flour properties. Quantitative variations are evidence that interaction occurs between different chromosome bins either to up or down regulate protein expression. In parallel, the proteomic analysis helped us to establish a CS proteome reference map with 410 identified agl proteins. This proteomic map could be used in the future to analyse endosperm gene expression in particularly in response to biotic or abiotic stresses.

Acknowledgments This work was financed by the European Union 6th Framework Program Healthgrain Project (Reference ID: FOOD-CT-2005-514008; 2005-2010) with the objectives of identifying sources of variation in bioactive compounds and developing 'omics' technologies and a biotechnology toolkit for accelerated breeding of improved cultivars. We are grateful to G. Gay and A. Loussert for providing seeds of the cultivar Chinese Spring and the deletion lines. B. Andersen is gratefully acknowledged for assistance with mass spectrometry and data retrieval. The Danish Centre for Advanced Food Studies (LMC) is thanked for financing part of the Ultraflex Tof/Tof mass spectrometer.

References

Bak-Jensen KS, Laugesen S, Ostergaard O, Finnie C, Roepstorff P, Svensson B (2007) Spatio-temporal profiling and degradation of ot-amylase isozymes during barley seed germination (vol 274, pg 2552, 2007). FEBS J 274:4009

Bottley A, Xia GM, Koebner RMD (2006) Homoeologous gene silencing in hexaploid wheat. Plant J 47:897–906

Debiton C, Merlino M, Chambon C, Bancel E, Decourteix M, Planchot V, Branlard G (2011) Analyses of albumins, globulins and amphiphilic proteins by proteomic approach give new insights on waxy wheat starch metabolism. J Cereal Sci 53:160–169

Desfrancs CC, Thiellement H (1985) Chromosomal localization of structural genes and regulators in wheat by 2D electrophoresis of ditelosomics lines. Theor Appl Genet 71:31–38

Endo TR, Gill BS (1996) The deletion stocks of common wheat. J Hered 87:295–307

Finnie C, Melchior S, Roepstorff P, Svensson B (2002) Proteome analysis of grain filling and seed maturation in barley. Plant Physiol 129:1308–1319

Galili G, Levy AA, Feldman M (1986) Gene-dosage compensation of endosperm proteins in hexaploid wheat *Triticum aestivum*. Proc Natl Acad Sci USA 83:6524–6528

Hurkman WJ, Vensel WH, Tanaka CK, Whitehand L, Altenbach SB (2009) Effect of high temperature on albumin and globulin accumulation in the endosperm proteome of the developing wheat grain. J Cereal Sci 49:12–23

Islam N, Woo SH, 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

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

Komatsu S, Kojima K, Suzuki K, Ozaki K, Higo K (2004) Rice Proteome Database based on two-dimensional polyacrylamide gel electrophoresis: its status in 2003. Nucleic Acids Res 32:D388–D392

Lafiandra D, Benedettelli S, Margiotta B, Porceddu E (1989) Chromosomal location of gliadin coding genes in *Triticum-aestivum ssp. spelta* and evidence on the lack of components controlled by *Gli-2* loci in wheat aneuploids. Theor Appl Genet 78:177–183

Marion D, Nicolas Y, Popineau Y, Branlard G, Landry J (1994) A new and improved sequential extraction procedure of wheat proteins. In: Wheat kernel proteins. Viterbo, Italy, pp 197–199

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.). Theor Appl Genet 118:1321–1337

Moore G (2000) Cereal chromosome structure, evolution, and pairing. Ann Rev Plant Physiol Plant Mol Biol 51:195–222

Payne PI, Holt LM, Jarvis MG, Jackson EA (1985) Two-dimensional fractionation of the endosperm proteins of bread wheat (*Triticum aestivum*): biochemical and genetic studies. Cereal Chem 62:319–326

Salt LJ, Robertson JA, Jenkins JA, Mulholland F, Mills ENC (2005) The identification of foam-forming soluble proteins from wheat (*Triticum aestivum*) dough. Proteomics 5:1612–1623

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

Singh J, Skerritt JH (2001) Chromosomal control of albumins and globulins in wheat grain assessed using different fractionation procedures. J Cereal Sci 33:163–181

Skylas DJ, Mackintosh JA, Cordwell SJ, Basseal DJ, Walsh BJ, Harry J, Blumenthal C, Copeland L, Wrigley CW, Rathmell W (2000) Proteome approach to the characterisation of protein composition in the developing and mature wheat-grain endosperm. J Cereal Sci 32:169–188

Thiellement H, Bahrman N, Desfrancs CC (1986) Regulatory effects of homoeologous chromosome arms on wheat proteins at two developmental stages. Theor Appl Genet 73:246–251

Tsujimoto H, Yamada T, Hasegawa K, Usami N, Kojima T, Endo TR, Ogihara Y, Sasakuma T (2001) Large-scale selection of lines with deletions in chromosome 1B in wheat and applications for fine deletion mapping. Genome 44:501–508

Vensel WH, Tanaka CK, Cai N, Wong JH, Buchanan BB, Hurkman WJ (2005) Developmental changes in the metabolic protein profiles of wheat endosperm. Proteomics 5:1594–1611

Zivy M, Thiellement H, Devienne D, Hofmann JP (1984) Study on nuclear and cytoplasmic genome expression in wheat by two-dimensional gel electrophoresis. 2. Genetic differences between two lines and two groups of cytoplasms at five developmental stages or organs. Theor Appl Genet 68:335–345