



**HAL**  
open science

## Grain subproteome responses to nitrogen and sulfur supply in diploid wheat *Triticum monococcum* ssp. *monococcum*

Titouan Bonnot, Emmanuelle Bancel, David Alvarez, Marlène Davanture, Julie Boudet, Marie Pailloux, Michel Zivy, Catherine Ravel, Pierre Martre

### ► To cite this version:

Titouan Bonnot, Emmanuelle Bancel, David Alvarez, Marlène Davanture, Julie Boudet, et al.. Grain subproteome responses to nitrogen and sulfur supply in diploid wheat *Triticum monococcum* ssp. *monococcum*. *The Plant Journal*, 2017, 91 (5), pp.894-910. 10.1111/tpj.13615 . hal-01602778

HAL Id: hal-01602778

<https://hal.science/hal-01602778v1>

Submitted on 26 May 2020

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

Article type : Resource

## Grain subproteome responses to nitrogen and sulfur supply in diploid wheat *Triticum monococcum* ssp. *monococcum*

Titouan Bonnot<sup>1</sup>, Emmanuelle Bancel<sup>1</sup>, David Alvarez<sup>1</sup>, Marlène Davanture<sup>2</sup>, Julie Boudet<sup>1</sup>, Marie Pailloux<sup>3</sup>, Michel Zivy<sup>2</sup>, Catherine Ravel<sup>1,\*</sup> and Pierre Martre<sup>1,†</sup>

<sup>1</sup> UMR GDEC, INRA, Université Blaise Pascal, 5 chemin de Beaulieu, 63 177 Aubière, France

<sup>2</sup> UMR GQE, INRA, Université Paris-Sud, CNRS, AgroParisTech, Université Paris-Saclay, 91 190 Gif-sur-Yvette, France

<sup>3</sup> LIMOS, CNRS, Université Blaise Pascal, 63 173 Aubière, France

<sup>†</sup>Present address: UMR LEPSE, INRA, Montpellier SupAgro, 34060 Montpellier, France

\*For correspondence (email: catherine.grand-ravel@inra.fr)

**Running title:** Grain subproteome response to nutrition in einkorn

**Keywords:** *Triticum monococcum*, grain, nitrogen, sulfur, storage proteins, nuclear proteins, albumin-globulin, network.

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1111/tpj.13615

This article is protected by copyright. All rights reserved.

## SUMMARY

Wheat grain storage proteins (GSPs) make up most of the protein content of grain and determine flour end-use value. The synthesis and accumulation of GSPs depend highly on nitrogen (N) and sulfur (S) availability and it is important to understand the underlying control mechanisms. Here we studied how the einkorn (*Triticum monococcum* ssp. *monococcum*) grain proteome responds to different amounts of N and S supply during grain development. GSP composition at grain maturity was clearly impacted by nutrition treatments, due to early changes in the rate of GSP accumulation during grain filling. Large-scale analysis of the nuclear and albumin-globulin sub-proteomes during this key developmental phase revealed that the abundance of 203 proteins was significantly modified by the nutrition treatments. Our results showed that the grain proteome was highly impacted by perturbation in the N:S balance. S supply strongly increased the rate of accumulation of S-rich  $\alpha/\beta$ -gliadin and  $\gamma$ -gliadin, and the abundance of several other proteins involved in glutathione metabolism. Post-anthesis N supply resulted in the activation of amino acid metabolism at the expense of carbohydrate metabolism and the activation of transport processes including nucleocytoplasmic transit. Protein accumulation networks were analyzed. Several central actors in the response were identified whose variation in abundance was related to variation in the amounts of many other proteins and are thus potentially important for GSP accumulation. This detailed analysis of grain sub-proteomes provides information on how wheat GSP composition can possibly be controlled in low-level fertilization condition.

This article is protected by copyright. All rights reserved.

## INTRODUCTION

Bread wheat (*Triticum aestivum* L.) is the main stable crop in many regions of the world and globally provides more than 20% of calories and proteins in the human diet. In mature grain, 10-15% of the dry mass is protein. GSPs make up 60-80% of the total protein in grain and metabolic proteins (the albumins and globulins) 15-20% (Shewry and Halford, 2002). GSPs accumulate during the effective filling phase of grain development (Shewry *et al.*, 2012). The two main GSP fractions are glutenin and gliadin that when mixed together with water form gluten and are thus the main determinants of the rheological and bread-making properties of wheat dough. More precisely, glutenins form polymers that are mainly responsible for the viscoelastic properties of dough and gliadins are monomeric proteins which confer extensibility (Branlard *et al.*, 2001). The GSP composition of wheat grain thus determines its use-end value. Commonly, gliadins are subdivided into  $\omega$ 1,2-,  $\omega$ 5-,  $\alpha/\beta$ - and  $\gamma$ -gliadin classes and glutenins are classified as high-molecular-weight subunits (HMW-GS) or low-molecular-weight subunits (LMW-GS; Wieser, 2007). These different GSP subclasses and subunits differ in the proportions of cysteine and methionine residues they contain, and are thus classified as being S-rich ( $\alpha/\beta$ - and  $\gamma$ -gliadins, LMW-GS) or S-poor ( $\omega$ 1,2- and  $\omega$ 5-gliadins, HMW-GS; Shewry *et al.*, 1997; Shewry *et al.*, 2001).

The availability of N and S in the soil highly influences the GSP composition of grain. N supply increases the rate and duration of protein accumulation and so increases the proportion of S-poor GSPs in mature grain (Shewry *et al.*, 2001; Triboï *et al.*, 2003; Chope *et al.*, 2014). In S-deficient conditions, GSP synthesis and accumulation occurs earlier because the cellularization phase is shortened (Castle and Randall, 1987). Mature grain contains low concentrations of cysteine and methionine when S is deficient (Wrigley *et al.*, 1980), causing S-rich GSPs to accumulate more slowly, which is compensated by an increased rate of

This article is protected by copyright. All rights reserved.

accumulation of S-poor GSPs, particularly HMW-GS (Zhao *et al.*, 1999; Wieser *et al.*, 2004; Zörb *et al.*, 2010). The molecular mechanisms that control GSP synthesis and composition in response to N and S supply are not yet well understood (Dai *et al.*, 2015). A priority for wheat breeders is to improve grain yield while reducing the need for fertilizers, especially N input. There is a strongly negative correlation between grain yield and grain protein concentration traits, so increasing grain yield is generally detrimental to grain protein concentration and hence grain quality (Simmonds, 1995; Oury and Godin, 2007). In addition, S deficiency is more frequently observed in soils nowadays, partly because of reductions in industrial SO<sub>2</sub> emissions (Eriksen, 2009), hindering the synthesis of S-rich GSPs in wheat. Analyzing responses to N and S nutrition will reveal key actors that can be targeted in developing plants that produce high quality grain in low-level fertilization conditions.

Large-scale proteomics and other omics approaches are commonly used to gain an overview of changes taking place in an organism or tissue by identifying the proteins and genes involved in a given biological response (Das *et al.*, 2015). When proteomics is used to study the grain response to nutrition, bioactive proteins are expected to be found either in the albumin and globulin fraction or among regulatory proteins. Albumins and globulins are mainly enzymes or enzyme inhibitors involved in cell metabolism and development and they may thus influence the rheological properties of wheat flour (Hill *et al.*, 2008). Albumin-globulin synthesis depends on the nutritional status of the plant. Although the synthesis of such proteins is less affected by N than that of GSPs (Wieser and Seilmeier, 1998; Martre *et al.*, 2003), N and S nutrition has been reported to affect several enzymes in wheat grain. For example, a glyceraldehyde-3-phosphate dehydrogenase and a serpin were both increased in conditions combining high N with low S (Flæte *et al.*, 2005). As GSP synthesis is controlled at a transcriptional level and many regulatory factors are nuclear proteins, quality determinants were expected to be nuclear localized (Verdier and Thompson, 2008). The

This article is protected by copyright. All rights reserved.

transcriptional regulation of GSP synthesis was confirmed by Dai *et al.* (2015) who reported significantly linear relationships between regulatory gene expression and GSP accumulation rates. Several nuclear proteome studies in other plant species have identified protein actors central to various stress responses (Yin and Komatsu, 2015; Petrovská *et al.*, 2015). Albumin-globulin and nuclear proteins are thus two subproteomes that can be targeted as being involved in GSP synthesis affected by N and S nutrition.

Einkorn (*Triticum monococcum*) is well tolerant of low input cropping and is appreciated for its excellent nutritional properties due to its high protein, high carotenoid and high tocol contents (Corbellini *et al.*, 1999; Hidalgo *et al.*, 2006; Hidalgo and Brandolini, 2014). It is an ancestral wheat whose diploid genome is a sister to bread wheat (*Triticum aestivum*) genome A (Marcussen *et al.*, 2014). Genomic resources are now available for this species. Its genomic sequence is available at [https://urgi.versailles.inra.fr/download/iwggsc/TGAC\\_WGS\\_assemblies\\_of\\_other\\_wheat\\_species/](https://urgi.versailles.inra.fr/download/iwggsc/TGAC_WGS_assemblies_of_other_wheat_species/) and its transcriptome was analyzed and annotated by Fox *et al.* (2014). Therefore, einkorn appears to be a suitable species in which to explore proteome responses to N and S nutrition.

Our eventual aim is to identify key factors able to ensure high quality grain (high protein content, adequate GSP composition) under low fertilization levels. In this study we produced an overview of the *T. monococcum* grain proteome response to N and S nutrition. Four post-anthesis N and S treatments were applied to a *T. monococcum* genotype grown in a glasshouse and grains were harvested at different developmental stages. GSPs were quantified and N and/or S supply was found to have significant effects on the amounts of gliadin classes and glutenin subunits. In mature grains,  $\alpha/\beta$ - and  $\gamma$ -gliadins were increased by

This article is protected by copyright. All rights reserved.

S supply, resulting in an increased gliadin-to-glutenin ratio, while HMW-GS and  $\omega$ 1,2-gliadins were increased with N supply. By focusing on nuclear and albumin-globulin fractions during the effective grain filling phase, 203 proteins were highlighted as being significantly affected by nutrition. Protein data integration grouped proteins with similar accumulation behaviours, which may potentially act synergistically in the grain. A few of these nuclear and albumin-globulin proteins are described in more detail as they potentially play a central role in GSP synthesis control in response to N and S nutrition.

## RESULTS

### Grain storage protein composition is modified by N and S supply

To study the response of nuclear, albumin-globulin and storage protein subproteomes of developing einkorn wheat grains to N and S supply, four treatments were applied to plants during the effective grain filling period. Plants were grown in a greenhouse in conditions resembling those in the field. The four treatments were N-S-, N+S-, N-S+, N+S+ according to whether N and/or S were applied. Dry mass and protein, N and S content of grain harvested from ears every 100°Cd were measured (Figure 1). Temporal data were fitted with logistic function equations (eqn. 1) to illustrate the rate and duration of compound accumulation (Figure 1 b, d, f; Table S1) which determine the total quantity of storage compounds at maturity.

The treatments altering N and S availability strongly modified the duration during which N and S accumulated in grains, but the rates of accumulation were much less modified (Figure 1b, d; Table S1). Mature grain contained more N following high N treatment (N+S- and N+S+) compared with the low N and low S treatment (N-S-), mainly because N accumulation lasted longer when more N was supplied (Figure 1b; Tables S1, S2). Similarly, the higher

This article is protected by copyright. All rights reserved.

quantity of S per grain at maturity for the high S treatments (N-S+ and N+S+) compared with the low S treatments (N+S- and N-S-) was mainly explained by the longer duration of accumulation of grain S with high S supply (Figure 1d; Table S1). The duration of dry mass accumulation increased when more N and S was supplied, but lower rate of accumulation compensated for this, so the final quantity of dry mass per grain at maturity was not modified by the treatments (Figure 1f; Table S1). In addition, the treatments did not influence the grain morphology during the filling phase (Figure S1, Table S3). Therefore, at maturity concentrations of grain protein and S differed between the treatments and reflected the changes in the quantity of N and S per grain (Figure 1a, c, e; Table S2). In N+S- grain, the N-to-S ratio was higher than 30, an indication of severe S deficiency (Figure 1e).

The treatments also had significant effects on the rate and duration of GSP accumulation (Figure 2, Table S1) and the final quantity in mature grains (Figure S2). At maturity, there was 67% more  $\alpha/\beta$ - and 163% more  $\gamma$ -gliadin per grain for the high S treatments (N-S+ and N+S+) compared with the low S treatments (N+S- and N-S-). These S-rich proteins accumulated at higher rates (+29% and +53%) and for longer durations (+30% and +13%) under the higher S supply (Figure 2c, e; Table S1). The quantity per grain of the S-poor  $\omega$ 1,2-gliadin was 45% higher for N+S+ compared with N+S- because accumulation lasted longer (+242°Cd, ~ +11.7 d), while the rate of accumulation was similar (Figure 2d; Table S1). For S-poor HMW-GS, there was 52% more per grain for N+S- than for the other treatments because of a higher rate of accumulation (+106%), which more than compensated for the shorter duration of accumulation (-21%) for N+S- (Figure 2b; Table S1). No significant difference between the treatments was measured at grain maturity for the S-poor  $\omega$ 5-gliadin and the S-rich LMW-GS, probably because the effects of N and S supply on the rate of accumulation of these GSPs was compensated by opposing effects on the duration of

This article is protected by copyright. All rights reserved.



their accumulation (Figure 2a, f; Table S1). Therefore, except for  $\omega$ 5-gliadin and LMW-GS, the variations in S-rich and S-poor GSP due to the different treatments reflected grain N-to-S ratios at maturity. The gliadin-to-glutenin ratio was different between treatments and increased when grain N-to-S ratio decreased (Figure S2), modifying the GSP composition at maturity (Figure S2d).

### **Analyzing nuclear and albumin-globulin protein fractions gives high coverage of the grain proteome**

Some of the mechanisms by which N and S supply influence GSP accumulation during grain filling must be governed by proteins. To investigate the grain proteome response to nutrition, two subproteomes from the same biological material were analyzed by shotgun mass spectrometry: nuclear proteins from 100 to 500°Cd after anthesis and albumin-globulin proteins from 200 to 600°Cd after anthesis. In total 1677 and 2475 proteins were identified in nuclear and albumin-globulin protein extracts, respectively. These proteins were classified as participating in 24 biological processes and 20 molecular functions according to their gene ontology (GO) information (Tables S4, S5). A total of 978 (24%) different proteins were accurately quantified (466 nuclear proteins and 512 albumin-globulin proteins; Figure 3a, b). Overall 47% and 33% of all quantified nuclear proteins and albumin-globulin respectively were classified as “uncharacterized” as they could not be assigned to a particular biological process. In the albumin-globulin fraction, several proteins were involved in responses to stimulus (12%), protein metabolic processes (11%) and carbohydrate metabolic processes (8%). In nuclear extracts, many proteins were either ribosomal proteins or histones and were thus classified as having functions in protein metabolic processes (17%) and cellular component organization (11%), respectively. Several proteins involved in transport processes were also quantified (5% of the proteins from nuclear and 6% from albumin-globulin

This article is protected by copyright. All rights reserved.

subproteomes). A clear difference in GO term enrichment for molecular functions was found between albumin-globulin and nuclear proteins. Many albumin-globulin proteins possess catalytic activity (19%), hydrolase activity (9%) or nucleotide binding activity (12%). In the nuclear extract, an enrichment of proteins with DNA binding (10%), nucleic acid binding (13%) and nucleotide binding (16%) activities was observed as expected. These results highlighted the value of independently analyzing the two subproteomes from the same biological material as the information obtained was not redundant.

Only 69 (7%) of the 978 quantified proteins were present in both fractions, showing again the very low redundancy between the albumin-globulin and nuclear protein analyses (Figure 3c). Of these 69 proteins, 20 (29%) were ribosomal proteins, classified as functioning in protein metabolic processes and structural molecule activity (Figure 3d, e). Other proteins are involved in response to stimulus (11%) or transport processes (8%), with for example two importins (Figure 3d), and many possess nucleotide binding (12%) or RNA binding (9%) activities (Figure 3e).

### **Various functions and response patterns in proteins impacted by N and S supply**

The abundance of 209 proteins (i.e. 21% of the quantified proteins) was significantly affected by the nutrition: the six GSPs, 155 albumin-globulin proteins and 48 nuclear proteins; Table S5). A difference in GO term enrichment was observed between the 978 quantified proteins and the 209 nutrition-responsive proteins. Proteins acting in carbohydrate metabolic processes, generation of precursor metabolites and energy, or storage functions were more represented among the proteins whose abundance was modified by N and/or S supply (Figure 4a). Conversely, protein metabolic process (e.g. many ribosomal proteins) and cellular component organization process (e.g. many histones), which made up a high

proportion of the quantified proteome (22.7% and 11.8%, respectively) were less well represented among the proteins whose abundance was modified by N and/or S supply (11.1% and 4.9%, respectively). Albumin-globulin proteins were proportionally more affected by N and S supply (30.2%) than nuclear proteins (10.3%), with enrichment of GO terms for catalytic, transferase, hydrolase or enzyme regulator activities (Figure 4a). Consistent with this result, DNA and RNA binding activities were underrepresented among the proteins whose abundance was nutrition-responsive.

To analyze how the abundance of the 209 proteins responded to N and/or S supply, clustering analysis was performed and 4 clusters were established (Figure 4b). Clusters 1 and 4 grouped proteins with opposite response profiles. Cluster 1 included 45 proteins that were much less abundant in response to N+S+, none of which were GSPs. On the contrary, cluster 4, the largest cluster with 65 proteins, grouped proteins which became more abundant in response to N+S+ and included  $\alpha/\beta$ -gliadin,  $\omega 5$ -gliadin, LMW-GS and 35 (73%) nuclear proteins. Cluster 2 grouped 46 proteins that were highly abundant in the N+S- treated grain and included the S-poor HMW-GS and  $\omega 1,2$ -gliadin. Cluster 3 grouped 53 proteins that were more abundant in response to N-S+ and included the S-rich  $\gamma$ -gliadin. Clusters 2 and 3 indicate a clear difference between N and S imbalances on the grain proteome, as the 99 proteins grouped in these two clusters behaved similarly in response to N-S- and N+S+.

Generally, proteins with similar molecular functions or acting in the same biological process were found in several clusters. However, some processes/functions were preferentially detected in one, two or three clusters (Figure 4c). For example, proteins acting in carbohydrate metabolic processes or in response to stimulus were poorly represented in cluster 2, few of them being increased by high N supply, unlike proteins involved in transport

This article is protected by copyright. All rights reserved.

processes. Proteins acting in the generation of precursor metabolites and energy, also classified as acting in carbohydrate metabolic processes, were mainly members of cluster 1 and thus decreased with high N and S supply. Nucleotide binding, RNA binding and DNA binding proteins were mostly found in cluster 4, and were nuclear proteins that mostly increased by high N and S supply.

### **Network analysis reveals central actors in the grain proteome response to nutrition**

To provide an integrative view of the grain response to N and S supply and highlight proteins that co-accumulate during grain filling, a directed multilevel network was constructed using proteomic data for the 209 proteins whose abundance was modified by N and/or S supply from 300 to 500°Cd after anthesis. We used a data mining method based on the discovery of association rules between attributes (Agier *et al.*, 2007; Vincent *et al.*, 2015). Associations between proteins and comparisons between treatments were built according to individual protein amounts in the different samples. The resulting network comprised 206 significant linkages (edges) and 111 proteins (nodes), i.e. about 50% of the 209 input proteins, and included the six GSPs, 25 nuclear proteins, and 80 albumin-globulin proteins (Figure 5a; Tables S5, S6). Networked proteins showed significant and high differences in abundance between at least one pair of treatments. The comparison between N+S- and N-S+ treatments created the highest number of edges (61, 29.6%) suggesting that the effect of these treatments on the grain proteome differed the most. The comparison between N+S+ and N+S- was involved in 49 rules, highlighting the importance of N × S interactions on the grain proteome compared to N alone. Only 27 proteins differed in abundance between N+S+ and N-S-. This result demonstrates the importance of the balance between N and S supply. N+S-/N-S-, N-S+/N-S- and N+S+/N-S+ pairwise comparisons linked 35, 12 and 22 proteins, respectively.

This article is protected by copyright. All rights reserved.

In this network, many groups of proteins were formed. Grouped proteins had the same response to N and S supply and were mostly part of the same cluster, an indicator of the robustness of the network analysis. For example, the quantity of LMW-GS was higher for N+S+ compared with N-S- and four nuclear proteins and three albumin-globulin proteins were also grouped with LMW-GS: an RNA helicase 27, an alanine aminotransferase, a betaine dehydrogenase, a guanine nucleotide-binding 3 homolog, a monothiol glutaredoxin-S11, a putative NADP-dependent oxidoreductase P1 and a chitinase 1. Some proteins were linked because they had very different quantities in three or four treatment comparisons and were therefore strongly affected by N and S supply. For example, HMW-GS was abundant with N+S- treatment compared with the three other treatments. Several albumin-globulin with similar responses to N and S supply were thus grouped with HMW-GS: a zeta-carotene desaturase, a cysteine synthase, a glutathionyl-hydroquinone reductase, an isoflavone reductase, a methylthioribose kinase and a beta-D-glucan exohydrolase. Other proteins were highly connected in this network, for example,  $\gamma$ - and  $\omega$ 1,2-gliadin and several albumin-globulin proteins, like an importin subunit alpha, which was connected with four edges. This importin subunit was more abundant for N+S- and N+S+ than with N-S- and N-S+, showing that this protein is increased by high N availability independently of S supply.

Co-accumulation of proteins was also investigated in other networks, built within cluster (Figure 5b-e; Tables S5, S6). In cluster 1 co-accumulation network, there were 20 (44%) proteins (nodes) including four involved in carbohydrate metabolic process, two acting in response to stress, two ribosomal proteins, a serpin-ZX and a purple acid phosphatase (Figure 5b). In cluster 2 co-accumulation network, there were 13 (28%) proteins including HMW-GS and  $\omega$ 1,2-gliadins (Figure 5c). According to the network, the abundance of these GSPs was connected to that of two other storage proteins and a programmed cell death

protein 4. The network connected an importin subunit beta-1-like and a tyrosine-protein phosphatase with two edges (abundant in similar conditions and not abundant in other similar conditions), showing that these two proteins strongly co-accumulate. The co-accumulation network from cluster 3 contains 15 (28%) nodes, including two vicilin-like antimicrobial peptides 2-2, a cationic peroxidase SPC4-like, an ankyrin repeat domain-containing protein and a glutathione transferase, which were related to the accumulation of other proteins (Figure 5d). The co-accumulation network with cluster 4 proteins is the largest, meaning that proteins that increased with N+S+ co-accumulated in tight synchrony (Figure 5e). This network was composed of two modules. One module contained 19 nodes (i.e. 29% of the proteins of cluster 4) and 70 edges, the other one contained 24 nodes (37%) and 85 edges. Some proteins in the first module were highly connected to others, including a RNA helicase 27, a L-ascorbate peroxidase, a phosphoglycerate kinase, a RNA polymerase and a nucleolar complex protein 3. The second module comprised LMW-GS,  $\omega$ 5- and  $\alpha/\beta$ -gliadin and several other co-accumulating proteins: a betaine aldehyde dehydrogenase, a probable aldo-keto reductase 1, a phosphorylase and a multiprotein-bridging factor 1a.

In these two network approaches several proteins were revealed as being strongly affected by N and/or S supply and co-accumulating with other proteins (nodes with thick border line in Figure 5; Table S5). These proteins thus represent central actors of the proteome response to nutrition.

### **Focus on central actors of the einkorn grain proteome response to N and S supply**

Among proteins revealed in the network analysis, a DEAD-box ATP-dependent RNA helicase 27 (M7ZPU9) was found in cluster 4 (Figure 6a). This protein is known to be involved in many processes including RNA transcription, pre-mRNA splicing, mRNA export, ribosome biogenesis and translation initiation (Linder, 2006). This RNA helicase

This article is protected by copyright. All rights reserved.

Version postprint

responded like LMW-GS to N and S supply, and was connected in the co-accumulation network to 11 other proteins, including a trihelix transcription factor (Figure 6b) and a peter-pan like protein (Figure 6c). The trihelix transcription factor contains a SANT/Myb domain, like those found in chromatin-remodeling complexes (Boyer *et al.*, 2004), and its quantity was linked in network with that of four proteins (Figure 5e). The peter-pan like protein possesses a Brix domain, found in proteins involved in ribosomal RNA processing (Weis *et al.*, 2015) and was networked with six proteins (Figure 5e). These three central actors co-accumulated and all increased with N and S supply. Two other central actors were part of cluster 1 and thus decreased with N and S supply: a purple acid phosphatase (Figure 6d) acting in protein post-translational modification and a serpin-ZX (Figure 6e) that negatively regulates endopeptidase activity. The amounts of these two proteins were associated with those of seven and four other proteins, respectively, including a probable sarcosine oxidase. Like  $\gamma$ -gliadin, the abundance of two glutathione transferases increased markedly with high S, one of which was central in the network being connected to nine other proteins (Figures 5d and 6f).

## DISCUSSION

In wheat, N and S supply highly influences GSP accumulation, and is thus expected to impact the viscoelastic properties of gluten and flour end-use value. There is currently a concerted effort to reduce the use of N-based fertilizers and soil S availability is decreasing in many wheat growing regions. It is in this context that we describe grain GSP, nuclear and albumin-globulin subproteome responses to N and S supply in a diploid species closely related to bread wheat and durum wheat (*T. turgidum* L. ssp. *durum* (Desf.) Husn.). The 203 albumin-globulin and nuclear proteins whose abundance was modified by N and/or S supply have diverse functions. Clustering and network analyses allowed us to group proteins that had

This article is protected by copyright. All rights reserved.

similar abundance profiles in response to N and/or S supply during grain filling. We found that S supply highly impacted GSP accumulation while N and S supply had different effects on the grain proteome. Several proteins central to the proteome response and potentially involved in the regulation of GSP synthesis in response to nutrition were highlighted. Other proteins required in this response could be major actors of the plant cell response to multiple environmental factors. Main results are summarized in a scheme of grain cell illustrating which molecular pathways are affected by N and S supply during grain filling (Figure 7).

### **The strong influence of S on the quantity and composition of grain storage protein and other grain subproteomes**

S deficiency in soil leads to decreases in the amount of GSP in wheat grain (Shewry *et al.*, 2001). We obtained similar results with more total GSP per grain at maturity in the two S-containing treatments than in the low S treatments. This was due to an early increase in the rate of accumulation and longer duration of accumulation of gliadin, especially the S-rich classes  $\alpha/\beta$ - and  $\gamma$ -gliadin. This is in accordance with previous reports that the highest amount of grain gliadin was obtained with moderate N fertilization and high S supply (Zörb *et al.*, 2010).

In the present study, the abundance of three glutathione transferases was higher in response to N-S+ and N+S+ treatments than to N-S-. One of them was also identified as a central actor and had similar nutrition response profiles to those of  $\gamma$ -gliadin. This result suggests that glutathione may increase with S supply, as occurs in wheat (Dai *et al.*, 2015). Glutathione is one of the major pools of non-protein S in wheat grain (Rhazi *et al.*, 2003; Steinfurth *et al.*, 2012). In S-deficient conditions, lesser amounts of total cysteine and glutathione were found in wheat flour, which affected GSP composition and hence technological flour properties (Reinbold *et al.*, 2008). Glutathione is involved in the

This article is protected by copyright. All rights reserved.



glutathione-ascorbate cycle controlling the concentration of active oxygen in the cell (Noctor and Foyer, 1998). In this mechanism, ascorbate peroxidase is involved in the detoxification of H<sub>2</sub>O<sub>2</sub> (Caverzan *et al.*, 2012). One ascorbate peroxidase was also more abundant in einkorn grain treated with S. It is therefore possible that glutathione metabolism increases when S is freely available.

In network analysis, N+S<sup>-</sup> and N-S<sup>+</sup> had the most contrasting effects on the grain proteome, since three GSPs (HMW-GS,  $\omega$ 1,2- and  $\gamma$ -gliadin) and many nuclear and albumin-globulin proteins differentially accumulated between these two treatments. N+S<sup>+</sup> and N+S<sup>-</sup> also had very different effects, again demonstrating the strong effect of S on the proteome. Interestingly, even though many proteins were increased with N+S<sup>+</sup>, the network analysis of the N+S<sup>+</sup>/N-S<sup>-</sup> comparison connected fewer proteins than the N+S<sup>-</sup>/N-S<sup>+</sup> comparison (27 and 61 proteins, respectively). Taken together, these results suggest that the grain proteome responded more to the N/S balance than to the availability of either N or S individually. These results are in good agreement with previous studies showing the importance of N-to-S ratio in determining the rheological properties of wheat dough and GSP composition (Zhao *et al.*, 1999; Zörb *et al.*, 2010; Dai *et al.*, 2015).

### **Grain metabolism was influenced by N and S supply**

Carbohydrate metabolism was the biological process most represented in proteins with significant nutrition effects. Such proteins were found in all four of the clusters established, but a few of them were in cluster 2, that is they didn't increase with N only. Conversely, proteins involved in amino acid metabolism were highly abundant with high N supply. This is consistent with previous results where an increased concentration of free amino acids in wheat flour was measured in S-deficient conditions (Granvogl *et al.*, 2007).

This article is protected by copyright. All rights reserved.

We noted that the abundance of a cysteine synthase increased considerably with high N supply, which tallies with a previous observation in rice (*Oryza sativa* L.) shoots and roots where high levels of cysteine synthase gene transcripts were observed in S starvation and non-limiting N conditions (Nakamura *et al.*, 1999). The latter authors concluded that activation of this gene regulated the balance between S-containing amino acids (methionine and cysteine) and other amino acids. These results broadly indicate that amino acid metabolism is activated at the expense of sugar metabolism in response to N nutrition.

As mentioned above, changes in amount for several proteins involved in glutathione metabolism occurred in response to N and S supply. In our results, three glutaredoxins showed quantitative variations in response to nutrition, particularly in a condition of high N and S supply. Glutaredoxins, also known as thioltransferases, use glutathione and NADPH as cofactors and act in a wide range of processes such as DNA synthesis, signal transduction and defense against oxidative stress and in diverse subcellular compartments (Fernandes and Holmgren, 2004; Herrero and De La Torre-Ruiz, 2007; Li, 2014). A potential redox signaling involving glutaredoxins could happen in response to the modification of the N and S status in the grain cell. Since redox signaling is a common mechanism of many plant responses to environmental stimuli, proteins acting in this process could have a central role. Thus, it is conceivable that glutaredoxins and several other proteins may be recruited for multiple cellular responses, in addition to the response to nutrition. In this way, intracellular transport is often used to optimize metabolic functioning under certain environmental conditions (Geitmann and Nebenführ, 2015). In our results, intracellular transport was likely influenced by nutrition. Proteins participating in transport were more abundant with high N supply and less abundant with high S supply. Among transporter proteins significantly affected by nutrition were three importins, two of which were present in both the nuclear and albumin-globulin protein extracts. Importins are involved in the transport of NLS-containing proteins

This article is protected by copyright. All rights reserved.

from the cytoplasm to the nucleus through nuclear pore complexes localized in the nuclear envelope (Bednenko *et al.*, 2003). In plants, communication between these two cell compartments is important for growth, development and responses to environmental stimuli (Merkle, 2011; Tamura and Hara-Nishimura, 2013). Regarding the cellular response to nutrition, nucleocytoplasmic transport may be a regulatory mechanism activated by grain N/S status to modify transcription by activating import of transcriptional regulators.

### **Nutrition impact on DNA binding proteins suggests transcriptional reprogramming**

In wheat, GSP synthesis is regulated at the transcriptional level (Dai *et al.*, 2015). This regulation, first described in barley and conserved in other cereals, is governed by a network involving cis-regulatory elements located in promoters of GSP genes and their interacting transcription factors (Rubio-Somoza *et al.*, 2006; Verdier and Thompson, 2008). The eight transcription factors known to act in this regulation were not quantified in our experiment, which didn't allow us to say if their abundance was disturbed by N and S supply. Among the 48 nuclear proteins showing significant nutrition effects, 10 (21%) have a DNA binding function. Proportionally then nutrition had a strong effect on proteins with DNA binding domains as this molecular function represented only 10% of the quantified nuclear proteins. Among them, two DNA-directed RNA polymerases, responsible for RNA synthesis, were increased by N and S supply. A similar response was observed for a multiprotein-bridging factor 1a. In *Arabidopsis thaliana*, this type of factor is known to be a transcriptional co-activator (Tsuda *et al.*, 2004). These results suggest that transcription is activated under conditions of high N and S availabilities. As might be expected if more transcripts are being produced, several proteins involved in mRNA splicing also accumulated when N and/or S were provided.

This article is protected by copyright. All rights reserved.

In the DNA-binding domain class were three histone H1 proteins. Histones are responsible for DNA condensation, organization and regulation in the nucleus. In eukaryotes, assembly of two copies of each of the core histone, H2A, H2B, H3 and H4, results in an octamer protein that binds to superhelical DNA to form the basic structure of DNA compaction called the nucleosome. Histone H1 molecules are linkers acting to maintain chromatin structure (Harshman *et al.*, 2013). Three variants of histone H1 were part of clusters 1 and 2, i.e. they were less abundant when S was supplied. In *A. thaliana*, histone H1 mutations resulted in modifications of DNA methylation patterns, with more methylation found in some promoter sequences in endosperm (Wierzbicki and Jerzmanowski, 2005; Rea *et al.*, 2012). In plants as in animals, DNA methylation affects the binding of specific proteins to DNA and therefore the formation of the transcription machinery (Finnegan *et al.*, 1998; Vanyushin and Ashapkin, 2011). Interestingly in our study, there was one protein with a methyl-CpG DNA binding domain (MBD) in cluster 3, i.e. it increased with high S supply. MBD proteins potentially act in transcriptional repression by recruiting chromatin remodeling factors, histone methyltransferases and histone deacetylases, leading to chromatin compaction (Grafi *et al.*, 2007). These results suggest that an epigenetic response to S supply occurs in einkorn grain. A trihelix transcription factor that increased in the high N and high S condition was predicted to be a central actor in the grain proteome response. This uncharacterized protein contains a SANT/Myb domain, characteristic of some nuclear receptor co-repressors and in many chromatin-remodeling complexes (Boyer *et al.*, 2004). In barley, it has been reported that methylation of the promoters of storage protein genes could repress storage protein synthesis and control grain development (Sørensen *et al.*, 1996; Radchuk *et al.*, 2005). Thus the methylation state of GSP promoters could be a regulatory mechanism, modifying the DNA-binding capacity of transcription factors, as in maize Opaque 2 transcription (Rossi *et al.*, 1997; Sturaro and Viotti, 2001; Locatelli *et al.*, 2009).

This article is protected by copyright. All rights reserved.

Taken together, these different elements indicate that N and S alter transcription dynamics, with potential effects on chromatin compaction and accessibility governed by epigenetic mechanisms.

In summary, major changes occur in the proteome of developing einkorn wheat grain when the nutritional status is modified. Post-flowering N and S nutrition clearly influenced the N-to-S ratio in the grain, leading to significant changes in the GSP composition in mature grain, by modification of the rate and duration of GSP accumulation, especially of HMW-GS,  $\alpha/\beta$ - and  $\gamma$ -gliadin. Changes in N-to-S ratios were also related to changes in the nuclear and albumin-globulin proteomes. The balance of cell functions was probably disturbed by altering N and S supply, as albumin-globulin proteins involved in carbohydrate and amino acid metabolism and transport were differentially affected. Central to these major changes, several DNA-binding nuclear proteins could have roles in the early grain response. Several proteins highlighted in the present study will be targeted to investigate how grain composition is controlled.

## EXPERIMENTAL PROCEDURES

### Plant material and growth conditions

*Triticum monococcum* ssp. *monococcum* (accession ERGE 35821) seeds were germinated (5 February) at room temperature on filter paper moistened with demineralized water in Petri dishes. When the radicles were 0.5 cm to 1 cm long, seedlings were transplanted to 50 mL PVC columns (7.5 cm internal diameter (i.d.)  $\times$  50 cm deep; 2 seedlings per column) filled with a 2:1 (v/v) mixture of washed perlite and river sand. The columns were arranged in a greenhouse to form a homogenous stand with a population

This article is protected by copyright. All rights reserved.

density of 512 plants m<sup>-2</sup>. The experimental design was a randomized complete block design with four blocks and four treatments. The high plant density inhibited the development of axillary tillers which favored the synchronous development of the plants within and between containers. In the greenhouse, air temperature at the top of the plant stand was maintained at 20°C/15°C (16 h light/8 h dark) and air relative humidity at 55%/75%. During light periods plants received a mean total daily photosynthetic photon flux (PPF) of 152 and 161 μmol m<sup>-2</sup> d<sup>-1</sup> between transplantation and anthesis and between anthesis and grain ripeness maturity, respectively. Air temperature and relative humidity were measured in miniature forced-draft flues placed in each block in the center of the plant stands at ear height. Air temperature, relative humidity, and PPF were measured every min and 15-min averages were recorded with a data logger. Thermal time was calculated as cumulative degree-days (0°C base temperature) using average daily air temperature.

Until anthesis, each PVC column received 167 mL d<sup>-1</sup> of a modified Hoagland's nutrient solution (Castle and Randall, 1987) containing 3 mM N and 0.1 mM S prepared with demineralized water (1 mM KH<sub>2</sub>PO<sub>4</sub>, 1 mM KNO<sub>3</sub>, 0.5 mM Ca(NO<sub>3</sub>)<sub>2</sub>, 0.5 mM NH<sub>4</sub>NO<sub>3</sub>, 0.1 mM MgSO<sub>4</sub>, 1.9 mM MgCl<sub>2</sub>, 3.5 mM CaCl<sub>2</sub>, 4 mM KCl, 10 μM H<sub>3</sub>BO<sub>3</sub>, 0.7 μM ZnCl<sub>2</sub>, 0.4 μM CuCl<sub>2</sub>, 4.5 μM MnCl<sub>2</sub>, 0.22 μM MoO<sub>3</sub> and 50 μM EDFS-Fe). At anthesis when N and S demand is low, the nutrient solution was replaced with demineralized water to avoid excess build-up of N and S compounds in plants or potting substrate. Then from 200°Cd to 700°Cd after anthesis four combinations of N and S, which were determined after carrying out pre-experiments (Figure S3, Table S7), were supplied: N-S-, nutrient solution with no N or S; N+S-, 6 mM N with no S; N-S+, low N (0.5 mM) and high S (2 mM); or N+S+, high N (6 mM) and high S (2 mM). The nutrient solutions were modified as described in Dai *et al.* (2015).

This article is protected by copyright. All rights reserved.

## Grain sampling and processing

Main-stem ears were tagged when the anthers of the central florets appeared. Grains were collected from the central portion of ears every 100°Cd (approximately every 5 d) from 300°Cd after anthesis to ripeness. For each treatment and sampling date, four to ten main-stem ears were sampled per replicate, depending on the grain developmental stage. Except for those used to determine grain dry mass, grains were frozen in liquid N<sub>2</sub> just after harvesting then stored at -80°C. Four biological replicates of whole grain were analyzed unless indicated otherwise.

## Determination of total grain S and N concentration

Aliquots of 5 mg of wholemeal flour were weighed in tin capsules and the total N and S concentrations were determined with the Dumas combustion method using a FlashEA 1112 NC Analyzer (Thermo Electron) following the manufacturer's recommendations. Grain protein content was calculated by multiplying grain N concentration by 5.62 (Mossé *et al.*, 1985).

## Storage protein extraction and quantification

Gliadin and glutenin proteins were sequentially extracted from 100 mg of wholemeal flour milled from each replicate of grain sampled between 300°Cd and 1000°Cd after anthesis, as described by Plessis *et al.* (2013). Gliadin classes and glutenin subunits were separated and quantified by reverse phase high performance liquid chromatography (RP-HPLC) using an Agilent 1290 Infinity LC system (Agilent Technologies, <http://www.agilent.com>) as described by Dai *et al.* (2015). Briefly, gliadin and glutenin extracts were filtered through regenerated cellulose syringe filters (0.45 µm pore diameter,

This article is protected by copyright. All rights reserved.

UptiDisc, Interchim, <http://www.interchim.com>), then 4  $\mu\text{L}$  (for gliadins) or 2  $\mu\text{L}$  (for glutenins) of protein extract were injected into a C8 reversed-phase ZORBAX 300StableBound column (2.1  $\times$  100 mm, 3.5  $\mu\text{m}$ , 300  $\text{\AA}$ ; Agilent Technologies) maintained at 50°C. Proteins were separated at a flow rate of 1 mL min<sup>-1</sup> by using linear solvent gradients from 24% to 50% acetonitrile containing 0.1% (v/v) trifluoroacetic acid over 13 min for gliadins, and from 23% to 42% over 25 min for glutenins. Proteins were detected by UV absorbance at 214 nm. Chromatograms were processed with ChemStation 10.1 software (Agilent Technologies) and the HPLC peaks corresponding to each of the four gliadin classes and the two glutenin subunits were identified following the observations of Wieser *et al.* (1998; Figure S4).

#### **Nuclear and albumin-globulin protein extraction, digestion and desalting**

Nuclei were extracted and purified from 2 g of whole grains ground in extraction buffer (20 mM Hepes-KOH pH 7, 5 mM MgCl<sub>2</sub>, 10 mM 2-mercaptoethanol, 0.5 mM PMSF, 0.1% (v/v) phosphatase inhibitor cocktail (Sigma-Aldrich)) with a Polytron homogenizer (Kinematica POLYTRON® PT 10) according to Bancel *et al.* (2015). Nuclear proteins were then prepared using TRI Reagent® (Sigma-Aldrich) according to the manufacturer's instructions. Albumins and globulins were extracted from whole grains according to Marion *et al.* (1994) with the following modifications. Proteins were extracted for 2 h at 4°C in 10 mM sodium phosphate, 10 mM NaCl, pH 7.8. After centrifugation at 8000  $\times$  g for 20 min at 4°C, proteins in the supernatant were precipitated with ice-cold acetone for 2 h at -20°C. After centrifugation at 10 000  $\times$  g for 5 min at 4°C, the resulting pellets were washed three times in ice-cold acetone then dried at room temperature.

This article is protected by copyright. All rights reserved.



Nuclear and albumin-globulin protein pellets were solubilized in 30  $\mu\text{L}$  of a buffer containing 0.1% (v/v) ZALS I, 6 M urea, 2 M thiourea, 10 mM DTT, 30 mM Tris-HCl pH 8.8, and 5 mM  $\text{NH}_4\text{HCO}_3$ . Protein concentration was measured using the 2-D Quant Kit (GE Healthcare) with bovine serum albumin as standard. For each sample, 40  $\mu\text{g}$  of protein was alkylated with 58 mM iodoacetamide for 50 min in the dark at room temperature, then diluted 10-fold with 50 mM  $\text{NH}_4\text{HCO}_3$  and digested in solution overnight at 37°C with 800 ng of trypsin in 50 mM  $\text{NH}_4\text{HCO}_3$ . Digestion was stopped by adding 10% TFA.

The protein samples were desalted using a Strata-XL SPE column (Phenomenex) regenerated with 500  $\mu\text{L}$  of 100%  $\text{CH}_3\text{CN}$  and equilibrated three times with 500  $\mu\text{L}$  of buffer A (3%  $\text{CH}_3\text{CN}$ , 0.06% acetic acid). Digested protein (15  $\mu\text{g}$ ) mixed with buffer A was loaded onto the column. After three washes with 500  $\mu\text{L}$  of buffer B (40%  $\text{CH}_3\text{CN}$ , 0.06% acetic acid), peptides were eluted twice with 300  $\mu\text{L}$  of buffer B, then dried and stored at -20°C.

### LC-MS/MS analysis of nuclear and albumin-globulin proteins

Peptide samples were solubilized in a buffer of 3%  $\text{CH}_3\text{CN}$ , 0.05% TFA and 0.05% formic acid. Liquid chromatography was performed on a NanoLC Ultra system (Eksigent). Samples (1  $\mu\text{g}$ ) were loaded at 7.5  $\mu\text{L min}^{-1}$  on a C18 precolumn (5  $\mu\text{m}$ , 100  $\mu\text{m i.d.} \times 2$  cm length; NanoSeparations) connected to a separating BIOSPHERE C18 column (3  $\mu\text{m}$ , 75  $\mu\text{m i.d.} \times 150$  mm length or 300 mm length for nuclear and albumin-globulin proteins, respectively; NanoSeparations). Solvent A was 0.1% formic acid in water and solvent B was 0.1% formic acid in  $\text{CH}_3\text{CN}$ . Peptide separation was achieved using a linear gradient from 5% to 35% solvent B for 28 min (for nuclear proteins) or 60 min (for albumin-globulin) at 300  $\text{nL min}^{-1}$ . Including the regeneration and the equilibration steps, a single run took 45 min for nuclear proteins and 87 min for albumin-globulin proteins. Eluted peptides were analyzed with a Q-Exactive mass spectrometer (Thermo Electron) using a nanoelectrospray interface.

This article is protected by copyright. All rights reserved.

Ionization was performed with a 1.3 kV spray voltage applied to an uncoated capillary probe (10  $\mu\text{m}$  i.d., New Objective). The Xcalibur 2.3 SP1 interface was used to monitor data-dependent acquisition of peptide ions. This included a full MS scan covering a mass-to-charge ratio (m/z) of 400 to 1,400 with a resolution of 70,000 and an MS/MS step (normalized collision energy, 27%; resolution, 17,500). The MS/MS step was reiterated for the eight major ions detected during the full MS scan. Dynamic exclusion was set to 40 s.

### Peptide identification and quantification

Nuclear and albumin-globulin proteins were identified by matching peptides against the Uniprot protein database version 2014\_07 limited to the *Triticum* genus using X!Tandem version Sledgehammer 2013.09.01.1 ([www.thegpm.org/tandem](http://www.thegpm.org/tandem)). Enzymatic cleavage parameters were set as trypsin digestion with one possible missed cleavage. Cys carboxyamidomethylation was set as static modification, whereas Met oxidation, N-terminal deamidation, and N-terminal acetylation were set as variable modifications. Precursor mass tolerance was 10 ppm and fragment mass tolerance was 0.02. Identified proteins were filtered and grouped using X!TandemPipeline version 3.3.4 (<http://pappso.inra.fr/bioinfo/xtandempipeline/>). Peptide and protein e-value cut offs were set to 0.01 and  $10^{-4}$ , respectively, with at least two peptides per protein.

Relative quantification of all identified peptides was done using MassChroQ software version 2.1.3 (Valot *et al.*, 2011) and extracting ion chromatograms (XICs) for all identified peptides and integrating the area of XIC peaks at the corresponding retention time. Non repeated peptides were then removed. For the albumin-globulin fraction, these corresponded to peptides present in less than 90% of the samples. For nuclear proteins, this threshold was set at 30%, due to the high effect of the developmental stage on this subproteome, as previously reported on wheat grain (Bonnot *et al.*, 2015). To quantify proteins, only peptides

This article is protected by copyright. All rights reserved.

not shared by other proteins and correlated peptides were used. The quantity of the corresponding protein was calculated by summing the quantities of specific peptides for albumin-globulin, and by averaging the quantities of specific peptides for nuclear proteins. The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium (Vizcaíno *et al.*, 2014) via the PRIDE partner repository with the dataset identifier PXD006058.

### **Functional classification and subcellular localization**

Identified nuclear and albumin-globulin proteins were classified according to gene ontology (GO; Ashburner *et al.*, 2000). Proteins that differed in quantity between treatments but did not have any functional annotation or GO information were analyzed with Blast2GO tool version 3.2 (Conesa *et al.*, 2005) in order to assign potential function (Table S5).

Predicted subcellular localization using MultiLoc2 (Blum *et al.*, 2009), WoLF PSORT (Horton *et al.*, 2007) and LocTree3 (Goldberg *et al.*, 2014) programs were used to validate nuclear protein identification (Table S8). Results were collected using the integrative tool PSI (Liu *et al.*, 2013) which combines prediction results of different programs. The top three hits were considered for WoLF PSORT and MultiLoc2. Nuclear proteins were used for statistical analysis if they were predicted to be nuclear with at least one tool.

### **Statistical analysis**

Data were analyzed using the statistical software program R v3.2.2 (R Core Team, 2015). Differences in grain dry mass per grain and per ear, total quantity of N and S per grain, and the quantity of GSP per grain at maturity were analyzed using a one-way ANOVA with the factor nutrition (four levels). For the quantity of GSP, albumin-globulin and nuclear

proteins in immature grains, a two-way ANOVA was performed, using the factors of treatment and developmental stage (i.e. the thermal time after anthesis). Significant differences ( $\alpha = 0.05$ ) were then analyzed by Tukey's honestly significantly different (HSD) post-hoc test.

To determine the rate and duration of accumulation of grain dry mass, total N and S and protein fractions, data were fitted with a 3-parameter logistic function equation (Triboi *et al.*, 2003):

$$Q(t) = \frac{Q_{\max}}{1 + 0.05 \exp\left(\frac{-4r(t - t_{95})}{Q_{\max}}\right)} \quad (1)$$

where  $Q$  is the quantity of dry mass, N, S, or GSP per grain,  $t$  is the thermal time after anthesis in degree-days,  $Q_{\max}$  is the final value of  $Q$  approached as  $t \rightarrow \infty$ ,  $r$  is the maximum rate of accumulation defined as the derivative of the point of inflexion, and  $t_{95}$  is the duration of accumulation defined as the period from anthesis onwards during which 95% of  $Q_{\max}$  is accumulated.

A hierarchical clustering on principal components was computed on proteins with significant nutrition effects using the FactomineR package for R (Husson *et al.*, 2015) and its HCPC function. With this function, a principal component analysis (PCA) was firstly performed on the data set. The clustering was then applied on the result of the PCA, using the Ward's criterion, which is based on the multidimensional variance (i.e. inertia). The number of clusters was determined with the suggested partition, i.e. the one with the higher relative loss of inertia.

## Proteomic network inference and analysis

Relationships among GSP, albumin-globulin and nuclear protein dynamics were analyzed using the RulNet platform for network inference (<http://rulnet.isima.fr>; Vincent *et al.*, 2015). RulNet uses an algorithm describing rule semantics between attributes (Agier *et al.*, 2007) to generate rules among data. Protein data were scaled and semantics were written to find rules between proteins whose accumulation was significantly affected by nutrition and the four N/S treatments according to the relative abundance of proteins in the different samples (semantic 1, Method S1). Two quality measures were used to select the best rules (confidence and lift). Rules with a confidence > 0.6 and a lift > 1.5 were selected. A second semantic was used to find rules between co-accumulated proteins (semantic 2, Method S1). For this second semantic, thresholds for support, confidence and lift were set at 0.3, 0.9 and 1.5, respectively. Validated rules were visualized using Cytoscape software version 3.3.0 (Smoot *et al.*, 2011) and are presented in Table S6.

## ACKNOWLEDGEMENTS

We thank Richard Blanc (UMR GDEC, INRA, Blaise Pascal University) for help in the greenhouse and Mireille Dardevet, Annie Faye, Marielle Merlino, Sibille Perrochon and Isabelle Nadaud for help collecting grains (UMR GDEC). This work was supported by a Ph.D. grant from the French Ministry for Higher Education and Research to TB and funding from the French Government managed by the National Research Agency (ANR) in the framework of Investments for the Future (ANR-10-BTBR-03), France AgriMer and the French Fund Supporting Plant Breeding (FSOV). The authors declare no conflict of interest.

This article is protected by copyright. All rights reserved.

## SHORT SUPPORTING INFORMATION LEGENDS

**Figure S1.** *Triticum monococcum* grain pictures.

**Figure S2.** Effects of N and S supply on grain storage protein composition at maturity.

**Figure S3.** Effects of N and S supply on the total grain N and S and on the concentration of N, S and dry mass (DM) in einkorn: pre-experimental results.

**Figure S4.** HPLC chromatograms of gliadin (a) and glutenin (b) in mature einkorn grain.

**Table S1.** Estimated maximum rate and duration of accumulation of grain dry mass, total N, total S and GSP.

**Table S2.** Quantity of dry mass, N, S and GSP per grain at maturity.

**Table S3.** Grain sizes measured during grain filling.

**Table S4.** Gene Ontology terms associated to nuclear proteins and albumin-globulin proteins identified.

**Table S5.** Summary of the 209 proteins significantly impacted by N and S supply.

**Table S6.** RulNet rules for proteomic network inference.

**Table S7.** Pre-experimental results for the grain fresh mass, dry mass, N mass and concentration, S mass and concentration, N-to-S ratio, and protein concentration of developing einkorn grain.

**Table S8.** Subcellular localization for proteins identified in the nuclear fraction.

**Method S1.** Semantics used to infer regulatory network from proteomics data.

## REFERENCES

**Agier, M., Petit, J.M. and Suzuki, E.** (2007) Unifying framework for rule semantics: Application to gene expression data. *Fundam. Informaticae*, 543–559.

**Ashburner, M., Ball, C.A., Blake, J.A., et al.** (2000) Gene ontology: tool for the unification

This article is protected by copyright. All rights reserved.

of biology. The Gene Ontology Consortium. *Nat. Genet.*, **25**, 25–29.

**Bancel, E., Bonnot, T., Davanture, M., Branlard, G., Zivy, M. and Martre, P.** (2015)

Proteomic approach to identify nuclear proteins in wheat grain. *J. Proteome Res.*, **14**, 4432–4439.

**Bednenko, J., Cingolani, G. and Gerace, L.** (2003) Nucleocytoplasmic transport:

navigating the channel. *Traffic*, **4**, 127–135.

**Blum, T., Briesemeister, S. and Kohlbacher, O.** (2009) MultiLoc2: integrating phylogeny

and Gene Ontology terms improves subcellular protein localization prediction. *BMC*

*Bioinformatics*, **10**, 274.

**Bonnot, T., Bancel, E., Chambon, C., Boudet, J., Branlard, G. and Martre, P.** (2015)

Changes in the nuclear proteome of developing wheat (*Triticum aestivum* L.) grain.

*Front. Plant Sci.*, **6**.

**Boyer, L.A., Latek, R.R. and Peterson, C.L.** (2004) Opinion: The SANT domain: a unique

histone-tail-binding module? *Nat. Rev. Mol. Cell Biol.*, **5**, 158–163.

**Branlard, G., Dardevet, M., Saccomano, R., Lagoutte, F. and Gourdon, J.** (2001)

Genetic diversity of wheat storage proteins and bread wheat quality. *Euphytica*, **119**, 59–67.

**Castle, S. and Randall, P.** (1987) Effects of sulfur deficiency on the synthesis and

accumulation of proteins in the developing wheat seed. *Aust. J. Plant Physiol.*, **14**, 503–

516.

**Caverzan, A., Passaia, G., Rosa, S.B., Ribeiro, C.W., Lazzarotto, F. and Margis-**

**Pinheiro, M.** (2012) Plant responses to stresses: role of ascorbate peroxidase in the

antioxidant protection. *Genet. Mol. Biol.*, **35**, 1011–1019.

This article is protected by copyright. All rights reserved.

Bonnot, T., Bancel, E., Alvarez, D., Davanture, M., Boudet, J., Pailloux, M., Zivy, M., Ravel, C., Martre, P. (2017). Grain subproteome responses to nitrogen and sulfur supply in diploid wheat *Triticum monococcum* ssp. *monococcum*. *Plant Journal*, 91 (5), 894-910. DOI: 10.1111/tpj.13615

**Chope, G.A., Wan, Y., Penson, S.P., Bhandari, D.G., Powers, S.J., Shewry, P.R. and Hawkesford, M.J.** (2014) Effects of genotype, season, and nitrogen nutrition on gene expression and protein accumulation in wheat grain. *J. Agric. Food Chem.*, **62**, 4399–4407.

**Conesa, A., Gotz, S., Garcia-Gomez, J.M., Terol, J., Talon, M. and Robles, M.** (2005) Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics*, **21**, 3674–3676.

**Corbellini, M., Empilli, S., Vaccino, P., Brandolini, A., Borghi, B., Heun, M. and Salamini, F.** (1999) Einkorn characterization for bread and cookie production in relation to protein subunit composition. *Cereal Chem.*, **76**, 727–733.

**Dai, Z., Plessis, A., Vincent, J., et al.** (2015) Transcriptional and metabolic alternations rebalance wheat grain storage protein accumulation under variable nitrogen and sulfur supply. *Plant J.*, **83**, 326–343.

**Das, P., Nutan, K.K., Singla-Pareek, S.L. and Pareek, A.** (2015) Understanding salinity responses and adopting “omics-based” approaches to generate salinity tolerant cultivars of rice. *Front. Plant Sci.*, **6**, 712.

**Eriksen, J.** (2009) Soil sulfur cycling in temperate agricultural systems. In *Advances in Agronomy*. pp. 55–89.

**Fernandes, A.P. and Holmgren, A.** (2004) Glutaredoxins: Glutathione-Dependent Redox Enzymes with Functions Far Beyond a Simple Thioredoxin Backup System. *Antioxid. Redox Signal.*, **6**, 63–74.

**Finnegan, E.J., Genger, R.K., Peacock, W.J. and Dennis, E.S.** (1998) Dna methylation in plants. *Annu. Rev. Plant Physiol. Plant Mol. Biol.*, **49**, 223–247.

This article is protected by copyright. All rights reserved.



**Flæte, N.E.S., Hollung, K., Ruud, L., Sogn, T., Færgestad, E.M., Skarpeid, H.J.,**

**Magnus, E.M. and Uhlen, A.K.** (2005) Combined nitrogen and sulphur fertilisation and its effect on wheat quality and protein composition measured by SE-FPLC and proteomics. *J. Cereal Sci.*, **41**, 357–369.

**Fox, S.E., Geniza, M., Hanumappa, M., et al.** (2014) De novo transcriptome assembly and analyses of gene expression during photomorphogenesis in diploid wheat triticum monococcum G. K. Pandey, ed. *PLoS One*, **9**, e96855.

**Geitmann, A. and Nebenführ, A.** (2015) Navigating the plant cell: intracellular transport logistics in the green kingdom. *Mol. Biol. Cell*, **26**, 3373–8.

**Goldberg, T., Hecht, M., Hamp, T., et al.** (2014) LocTree3 prediction of localization. *Nucleic Acids Res.*, **42**, W350–W355.

**Grafi, G., Zemach, A. and Pitto, L.** (2007) Methyl-CpG-binding domain (MBD) proteins in plants. *Biochim. Biophys. Acta - Gene Struct. Expr.*, **1769**, 287–294.

**Granvogel, M., Wieser, H., Koehler, P., Tucher, S. Von and Schieberle, P.** (2007)

Influence of sulfur fertilization on the amounts of free amino acids in wheat. Correlation with baking properties as well as with 3-aminopropionamide and acrylamide generation during baking. *J. Agric. Food Chem.*, **55**, 4271–4277.

**Harshman, S.W., Young, N.L., Parthun, M.R. and Freitas, M.A.** (2013) H1 histones: current perspectives and challenges. *Nucleic Acids Res.*, **41**, 9593–9609.

**Herrero, E. and La Torre-Ruiz, M.A. De** (2007) Monothiol glutaredoxins: A common domain for multiple functions. *Cell. Mol. Life Sci.*, **64**, 1518–1530.

**Hidalgo, A. and Brandolini, A.** (2014) Nutritional properties of einkorn wheat (*Triticum monococcum* L.). *J. Sci. Food Agric.*, **94**, 601–612.

This article is protected by copyright. All rights reserved.

**Hidalgo, A., Brandolini, A., Pompei, C. and Piscozzi, R.** (2006) Carotenoids and tocopherols of einkorn wheat (*Triticum monococcum* ssp. *monococcum* L.). *J. Cereal Sci.*, **44**, 182–193.

**Hill, K., Horváth-Szancics, E., Hajós, G. and Kiss, É.** (2008) Surface and interfacial properties of water-soluble wheat proteins. *Colloids Surfaces A Physicochem. Eng. Asp.*, **319**, 180–187.

**Horton, P., Park, K.-J., Obayashi, T., Fujita, N., Harada, H., Adams-Collier, C.J. and Nakai, K.** (2007) WoLF PSORT: protein localization predictor. *Nucleic Acids Res.*, **35**, W585–W587.

**Husson, F., Josse, J., Le, S., Mazet, J. and Husson, M.F.** (2015) Package “FactoMineR”.

**Li, S.** (2014) Redox Modulation Matters: Emerging Functions for Glutaredoxins in Plant Development and Stress Responses. *Plants*, **3**, 559–582.

**Linder, P.** (2006) Dead-box proteins: a family affair-active and passive players in RNP-remodeling. *Nucleic Acids Res.*, **34**, 4168–4180.

**Liu, L., Zhang, Z., Mei, Q. and Chen, M.** (2013) PSI: A comprehensive and integrative approach for accurate plant subcellular localization prediction J. L. Heazlewood, ed. *PLoS One*, **8**, e75826.

**Locatelli, S., Piatti, P., Motto, M. and Rossi, V.** (2009) Chromatin and DNA modifications in the Opaque2-mediated regulation of gene transcription during maize endosperm development. *Plant Cell*, **21**, 1410–1427.

**Marcussen, T., Sandve, S.R., Heier, L., et al.** (2014) Ancient hybridizations among the ancestral genomes of bread wheat. *Science*, **345**, 1250092–1250092.

This article is protected by copyright. All rights reserved.

**Marion, D., Nicolas, Y., Popineau, Y., Branlard, G. and Landry, J.** (1994) *A new and improved sequential extraction procedure of wheat proteins*. Molecular and functional aspects (p.47), ed., Presented at Wheat kernel proteins, Viterbo, ITA (1994-09-28 - 1994-09-30).

**Martre, P., Porter, J.R., Jamieson, P.D. and Triboï, E.** (2003) Modeling grain nitrogen accumulation and protein composition to understand the sink/source regulations of nitrogen remobilization for wheat. *Plant Physiol.*, **133**, 1959–1967.

**Merkle, T.** (2011) Nucleo-cytoplasmic transport of proteins and RNA in plants. *Plant Cell Rep.*, **30**, 153–176.

**Mossé, J., Huet, J.C. and Baudet, J.** (1985) The amino acid composition of wheat grain as a function of nitrogen content. *J. Cereal Sci.*, **3**, 115–130.

**Nakamura, T., Yamaguchi, Y. and Sano, H.** (1999) Four rice genes encoding cysteine synthase: isolation and differential responses to sulfur, nitrogen and light. *Gene*, **229**, 155–161.

**Noctor, G. and Foyer, C.H.** (1998) Ascorbate and Glutathione: keeping active oxygen under control. *Annu. Rev. Plant Physiol. Plant Mol. Biol.*, **49**, 249–279.

**Oury, F.-X. and Godin, C.** (2007) Yield and grain protein concentration in bread wheat: how to use the negative relationship between the two characters to identify favourable genotypes? *Euphytica*, **157**, 45–57.

**Petrovská, B., Šebela, M. and Doležel, J.** (2015) Inside a plant nucleus: discovering the proteins. *J. Exp. Bot.*, **66**, 1627–1640.

**Plessis, A., Ravel, C., Bordes, J., Balfourier, F. and Martre, P.** (2013) Association study of wheat grain protein composition reveals that gliadin and glutenin composition are

This article is protected by copyright. All rights reserved.

trans-regulated by different chromosome regions. *J. Exp. Bot.*, **64**, 3627–3644.

**R Core Team** (2015) R: A language and environment for statistical computing. R foundation for statistical computing, Vienna, Austria.

**Radchuk, V. V., Sreenivasulu, N., Radchuk, R.I., Wobus, U. and Weschke, W.** (2005) The methylation cycle and its possible functions in barley endosperm development. *Plant Mol. Biol.*, **59**, 289–307.

**Rea, M., Zheng, W., Chen, M., Braud, C., Bhangu, D., Rognan, T.N. and Xiao, W.** (2012) Histone H1 affects gene imprinting and DNA methylation in Arabidopsis. *Plant J.*, **71**, 776–786.

**Reinbold, J., Rychlik, M., Asam, S., Wieser, H. and Koehler, P.** (2008) Concentrations of total glutathione and cysteine in wheat flour as affected by sulfur deficiency and correlation to quality parameters. *J. Agric. Food Chem.*, **56**, 6844–6850.

**Rhazi, L., Cazalis, R., Lemelin, E. and Aussenac, T.** (2003) Changes in the glutathione thiol–disulfide status during wheat grain development. *Plant Physiol. Biochem.*, **41**, 895–902.

**Rossi, V., Motto, M. and Pellegrini, L.** (1997) Analysis of the methylation pattern of the maize Opaque-2 (O2) promoter and in vitro binding studies indicate that the O2 B-Zip protein and other endosperm factors can bind to methylated target sequences. *J. Biol. Chem.*, **272**, 13758–13765.

**Rubio-Somoza, I., Martinez, M., Abraham, Z., Diaz, I. and Carbonero, P.** (2006) Ternary complex formation between HvMYBS3 and other factors involved in transcriptional control in barley seeds. *Plant J.*, **47**, 269–281.

**Shewry, P.R. and Halford, N.G.** (2002) Cereal seed storage proteins: structures, properties

This article is protected by copyright. All rights reserved.

and role in grain utilization. *J. Exp. Bot.*, **53**, 947–958.

**Shewry, P.R., Mitchell, R.A.C., Tosi, P., et al.** (2012) An integrated study of grain development of wheat (cv. Hereward). *J. Cereal Sci.*, **56**, 21–30.

**Shewry, P.R., Tatham, A.S. and Halford, N.G.** (2001) Nutritional control of storage protein synthesis in developing grain of wheat and barley. *Plant Growth Regul.*, **34**, 105–111.

**Shewry, P.R., Tatham, A.S. and Lazzeri, P.** (1997) Biotechnology of wheat quality. *J. Sci. Food Agric.*, **73**, 397–406.

**Simmonds, N.W.** (1995) The relation between yield and protein in cereal grain. *J. Sci. Food Agric.*, **67**, 309–315.

**Smoot, M.E., Ono, K., Ruschinski, J., Wang, P.-L. and Ideker, T.** (2011) Cytoscape 2.8: new features for data integration and network visualization. *Bioinformatics*, **27**, 431–432.

**Sørensen, M.B., Müller, M., Skerritt, J. and Simpson, D.** (1996) Hordein promoter methylation and transcriptional activity in wild-type and mutant barley endosperm. *Mol. Gen. Genet.*, **250**, 750–760.

**Steinfurth, D., Zörb, C., Braukmann, F. and Mühling, K.H.** (2012) Time-dependent distribution of sulphur, sulphate and glutathione in wheat tissues and grain as affected by three sulphur fertilization levels and late S fertilization. *J. Plant Physiol.*, **169**, 72–77.

**Sturaro, M. and Viotti, A.** (2001) Methylation of the Opaque2 box in zein genes is parent-dependent and affects O2 DNA binding activity in vitro. *Plant Mol. Biol.*, **46**, 549–560.

**Tamura, K. and Hara-Nishimura, I.** (2013) The molecular architecture of the plant nuclear pore complex. *J. Exp. Bot.*, **64**, 823–832.

This article is protected by copyright. All rights reserved.

**Triboï, E., Martre, P. and Triboï-Blondel, A.-M.** (2003) Environmentally-induced changes in protein composition in developing grains of wheat are related to changes in total protein content. *J. Exp. Bot.*, **54**, 1731–1742.

**Tsuda, K., Tsuji, T., Hirose, S. and Yamazaki, K.** (2004) Three Arabidopsis MBF1 homologs with distinct expression profiles play roles as transcriptional co-activators. *Plant Cell Physiol.*, **45**, 225–231.

**Valot, B., Langella, O., Nano, E. and Zivy, M.** (2011) MassChroQ: A versatile tool for mass spectrometry quantification. *Proteomics*, **11**, 3572–3577.

**Vanyushin, B.F. and Ashapkin, V. V** (2011) DNA methylation in higher plants: Past, present and future. *Biochim. Biophys. Acta*, **1809**, 360–368.

**Verdier, J. and Thompson, R.D.** (2008) Transcriptional regulation of storage protein synthesis during dicotyledon seed filling. *Plant Cell Physiol.*, **49**, 1263–1271.

**Vincent, J., Martre, P., Gouriou, B., Ravel, C., Dai, Z., Petit, J.-M. and Pailloux, M.** (2015) RulNet: A web-oriented platform for regulatory network inference, application to wheat –omics data. *PLoS One*, **10**, e0127127.

**Vizcaíno, J.A., Deutsch, E.W., Wang, R., et al.** (2014) ProteomeXchange provides globally coordinated proteomics data submission and dissemination. *Nat. Biotechnol.*, **32**, 223–6.

**Weis, B.L., Palm, D., Missbach, S., Bohnsack, M.T. and Schleiff, E.** (2015) atBRX1-1 and atBRX1-2 are involved in an alternative rRNA processing pathway in Arabidopsis thaliana. *RNA*, **21**, 415–425.

**Wierzbicki, A.T. and Jerzmanowski, A.** (2005) Suppression of histone H1 genes in Arabidopsis results in heritable developmental defects and stochastic changes in DNA methylation. *Genetics*, **169**, 997–1008.

This article is protected by copyright. All rights reserved.

**Wieser, H.** (2007) Chemistry of gluten proteins. *Food Microbiol.*, **24**, 115–119.

**Wieser, H., Antes, S. and Seilmeier, W.** (1998) Quantitative determination of gluten protein types in wheat flour by reversed-phase high-performance liquid chromatography. *Cereal Chem.*, **75**, 644–650.

**Wieser, H., Gutser, R. and Tucher, S. von** (2004) Influence of sulphur fertilisation on quantities and proportions of gluten protein types in wheat flour. *J. Cereal Sci.*, **40**, 239–244.

**Wieser, H. and Seilmeier, W.** (1998) The influence of nitrogen fertilisation on quantities and proportions of different protein types in wheat flour. *J. Sci. Food Agric.*, **76**, 49–55.

**Wrigley, C., Cros, D., Archer, M., Downie, P. and Roxburgh, C.** (1980) The sulfur content of wheat endosperm proteins and its relevance to grain quality. *Aust. J. Plant Physiol.*, **7**, 755–766.

**Yin, X. and Komatsu, S.** (2015) Quantitative proteomics of nuclear phosphoproteins in the root tip of soybean during the initial stages of flooding stress. *J. Proteomics*, **119**, 183–195.

**Zhao, F., Hawkesford, M. and McGrath, S.** (1999) Sulphur assimilation and effects on yield and quality of wheat. *J. Cereal Sci.*, **30**, 1–17.

**Zörb, C., Grover, C., Steinfurth, D. and Hermann Mühlhling, K.** (2010) Quantitative proteome analysis of wheat gluten as influenced by N and S nutrition. *Plant Soil*, **327**, 225–234.

## FIGURE LEGENDS

**Figure 1.** Effects of N and S supply on the accumulation of total grain N, S and dry mass (DM) in einkorn. Total protein concentration (a), mass of N per grain (b), S concentration (c), mass of S per grain (d), N-to-S ratio (e) and grain dry mass (f) were measured. The four treatments N-S-, N+S-, N-S+ and N+S+ were applied from 200 to 700 °Cd after anthesis. Data are means  $\pm$  1 s.e. for  $n = 4$  independent replicates. Lines in (b), (d) and (f) were fitted to the data using eqn. (1).

**Figure 2.** Effects of N and S supply on the accumulation of LMW-GS and HMW-GS and  $\alpha/\beta$ -,  $\gamma$ -,  $\omega 5$ - and  $\omega 1,2$ -gliadin in einkorn grain. The four treatments N-S-, N+S-, N-S+ and N+S+ were applied from 200 to 700 °Cd after anthesis. Data are means  $\pm$  1 s.e. for  $n = 4$  independent replicates. Lines were fitted to the data using eqn. (1).

**Figure 3.** Functional classification of quantified nuclear and albumin-globulin proteins in einkorn grain. Proteins were grouped as being involved in 24 biological processes (a) and 20 molecular functions (b) according to gene ontology annotation. In (a) and (b) nuclear proteins and albumin-globulin are represented in the inner and outer circles, respectively. (c) Venn diagram showing the numbers of quantified proteins in nuclear and albumin-globulin extracts. Biological processes (d) and molecular functions (e) of the 69 proteins found in both nuclear and albumin-globulin protein fractions.

**Figure 4.** Functions and response to nutrition profiles of proteins showing significant nutrition effect. (a) Percentage of proteins in each biological process and molecular function



classes for the 978 quantified proteins (open bars) and for the 209 proteins (155 albumin-globulin proteins, 48 nuclear proteins, and 6 grain storage proteins) which differed in abundance between the treatments (filled bars). (b) Boxplots of the normalized quantity of proteins in the four clusters for the four treatments. Boxes show the 25th to 75th percentile range, horizontal lines in boxes show medians, error bars outside boxes show the 10th to 90th percentile range, dots are outliers. The number of proteins in each cluster is indicated at the top of each panel. alg, albumin-globulin, nucl, nuclear protein. (c) Heatmaps of the number of proteins in different biological process and molecular function classes for the four clusters shown in (b).

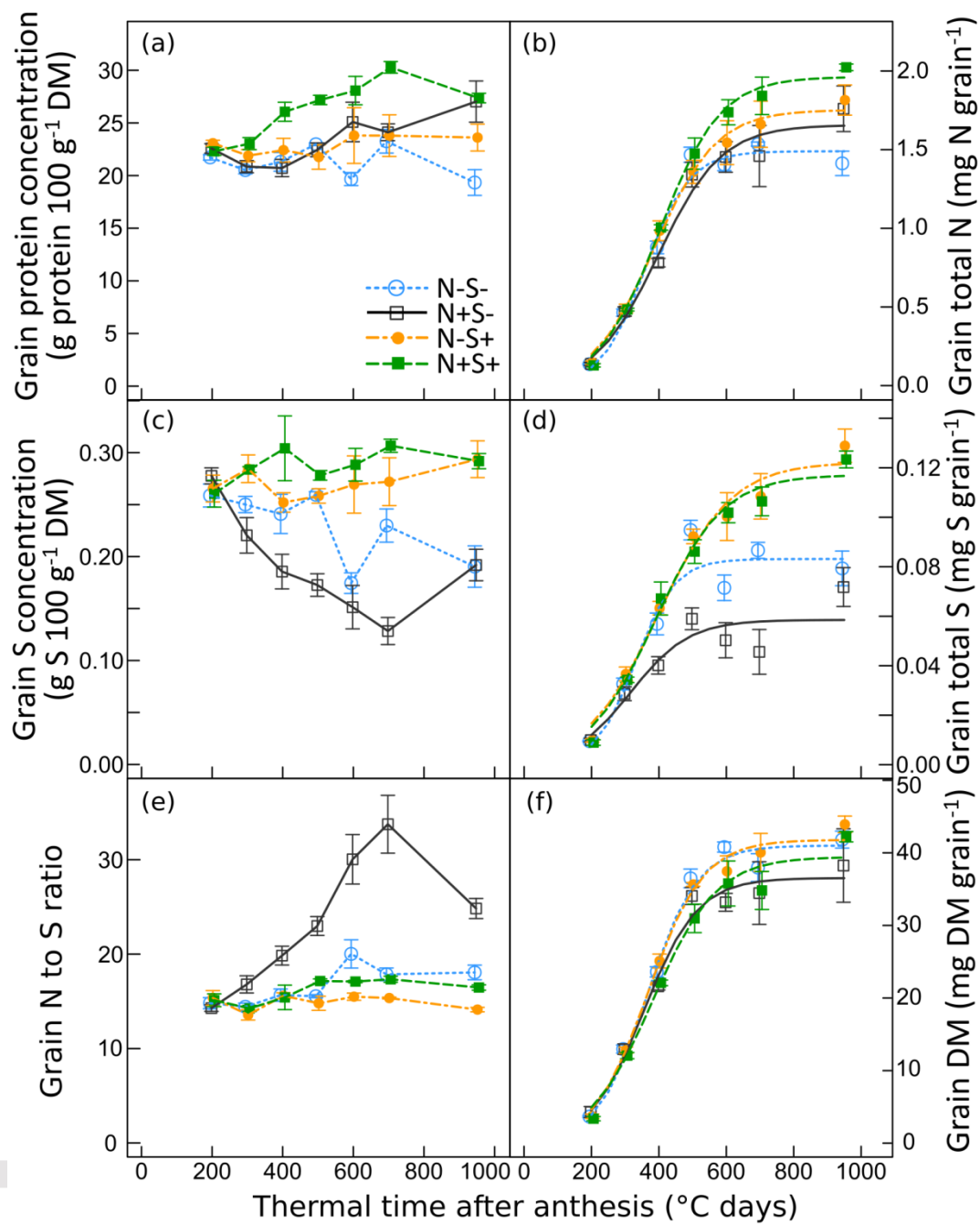
**Figure 5.** Network analysis of einkorn grain subproteome response to N and S supply. (a) Co-abundance network for proteins which differed in abundance between treatments and linkages with N and S treatments defined as central attributes. (b) to (e) Co-abundance networks for proteins in cluster 1 (b), cluster 2 (c), cluster 3 (d) and cluster 4 (e) as shown in Figure 4. Node shapes reflect attribute categories: squares, comparison between two treatments; circles, grain storage proteins (GSPs); hexagons, albumin-globulin proteins; and diamonds, nuclear protein. Node colours indicate which cluster the protein belongs to: purple, cluster 1; brown, cluster 2; red, cluster 3; green, cluster 4. In (b) to (e) the intensity of the node colour reflects the number of edges (connectivity). Edge colour indicates the biological significance between source and target: in (a) red when the protein was more abundant in the first treatment used in the comparison and blue when the protein was less abundant in the first treatment; in (b, c, d, e), red shows the abundance of proteins was high for the two nodes and blue when the abundance of proteins was low for the two nodes. Nodes with thick borders were found both in (a) and a corresponding cluster network in (b), (c), (d), or (e). Albumin-

This article is protected by copyright. All rights reserved.

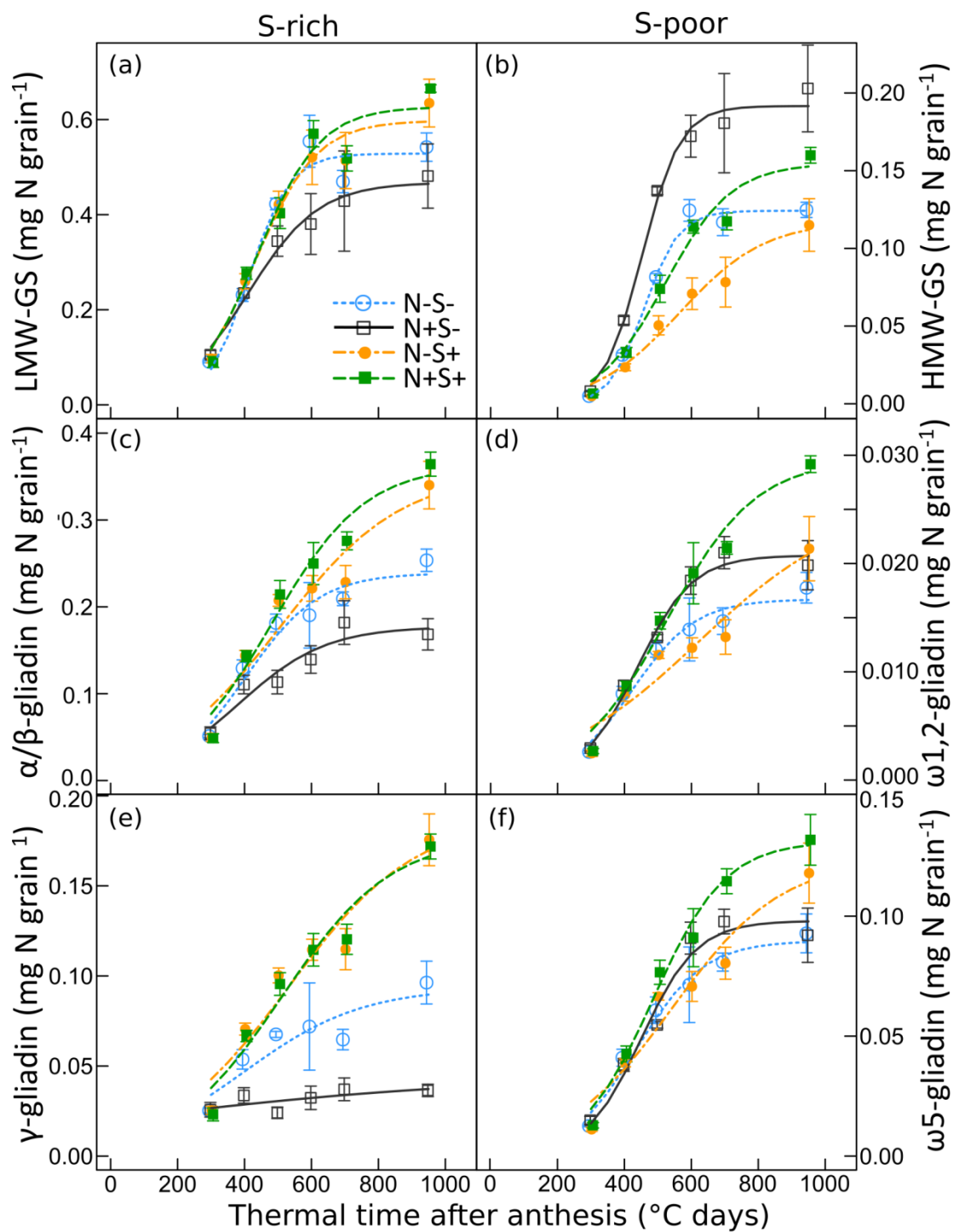
globulin and nuclear proteins are identified by their Uniprot accession reference (for details, see Table S5). Accessions written in red correspond to proteins highlighted in Figure 6.

**Figure 6.** Effects of N and S supply on the quantity of selected nuclear and albumin-globulin proteins during einkorn grain filling. (a) DEAD-box ATP-dependent RNA helicase 27, (b) trihelix transcription factor, (c) peter pan-like protein, (d) purple acid phosphatase, (e) serpin-ZX, (f) glutathione transferase. Data are scaled and are means  $\pm$  1 s.e. for  $n = 3$  independent replicates.

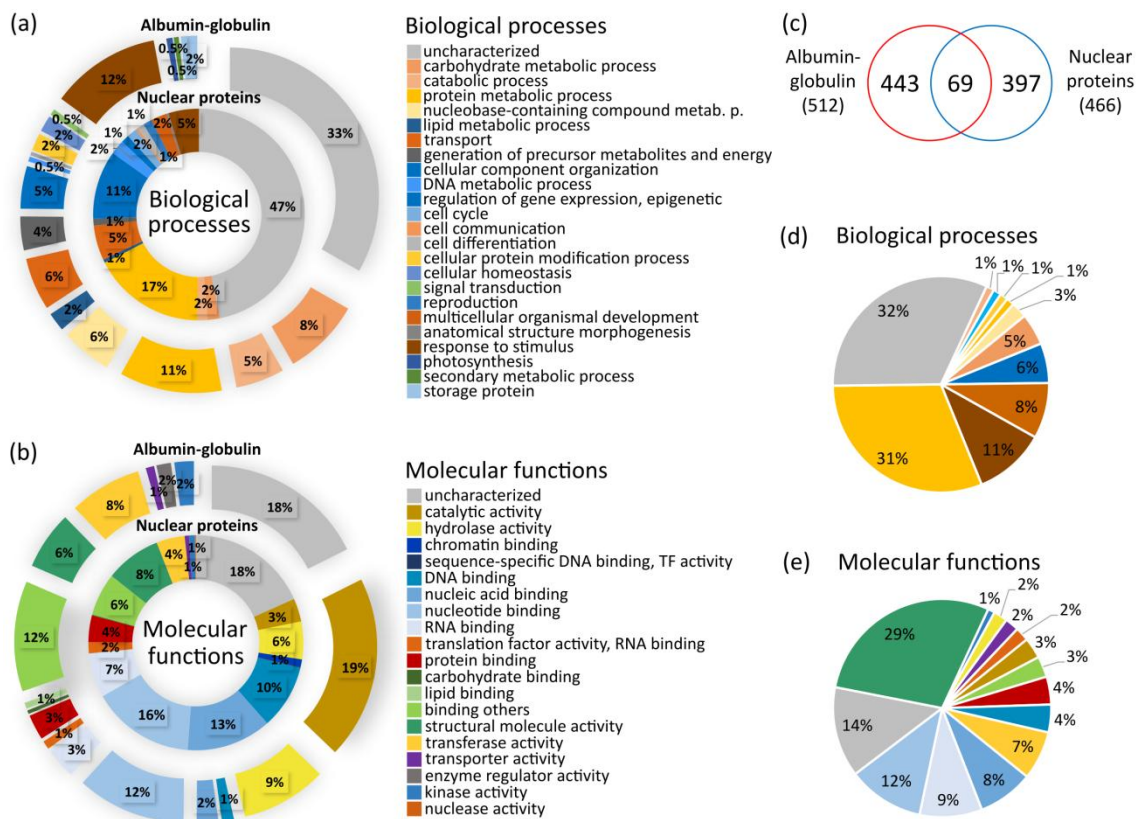
**Figure 7.** Schematic overview of einkorn grain subproteome responses to N and S supply. Proteins represented are some of the 209 proteins showing significant nutrition effects during grain filling. Relative protein quantities are scaled and means for  $n = 3$  independent replicates are represented by a color scale on a  $1 \times 4$  grid, each square representing one treatment. Functional classification was based on gene ontology annotation and subcellular localization was determined by prediction tools. Uniprot accessions corresponding to protein descriptions are given in parentheses. Accessions in red correspond to proteins highlighted in Figure 6.



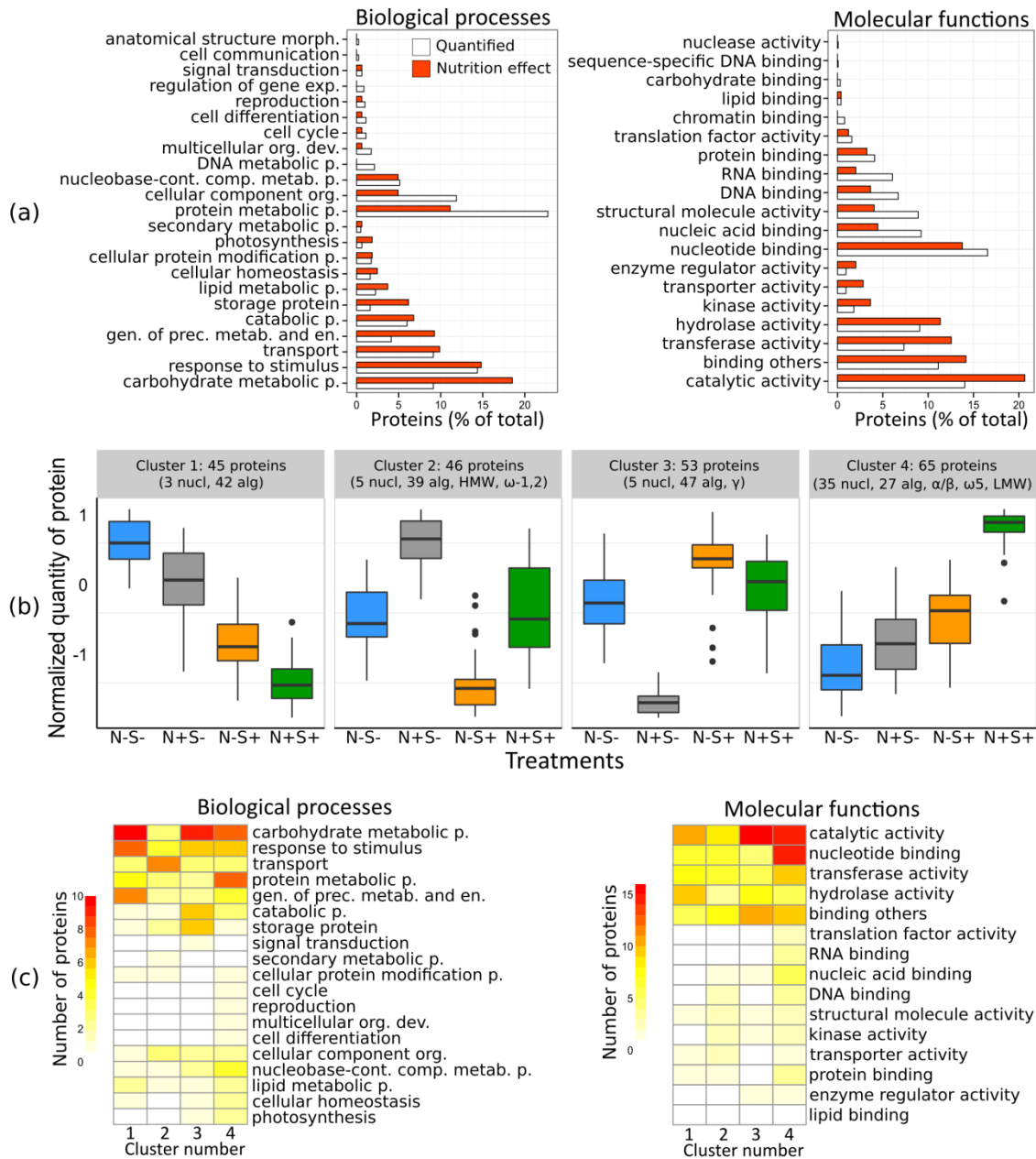
This article is protected by copyright. All rights reserved.



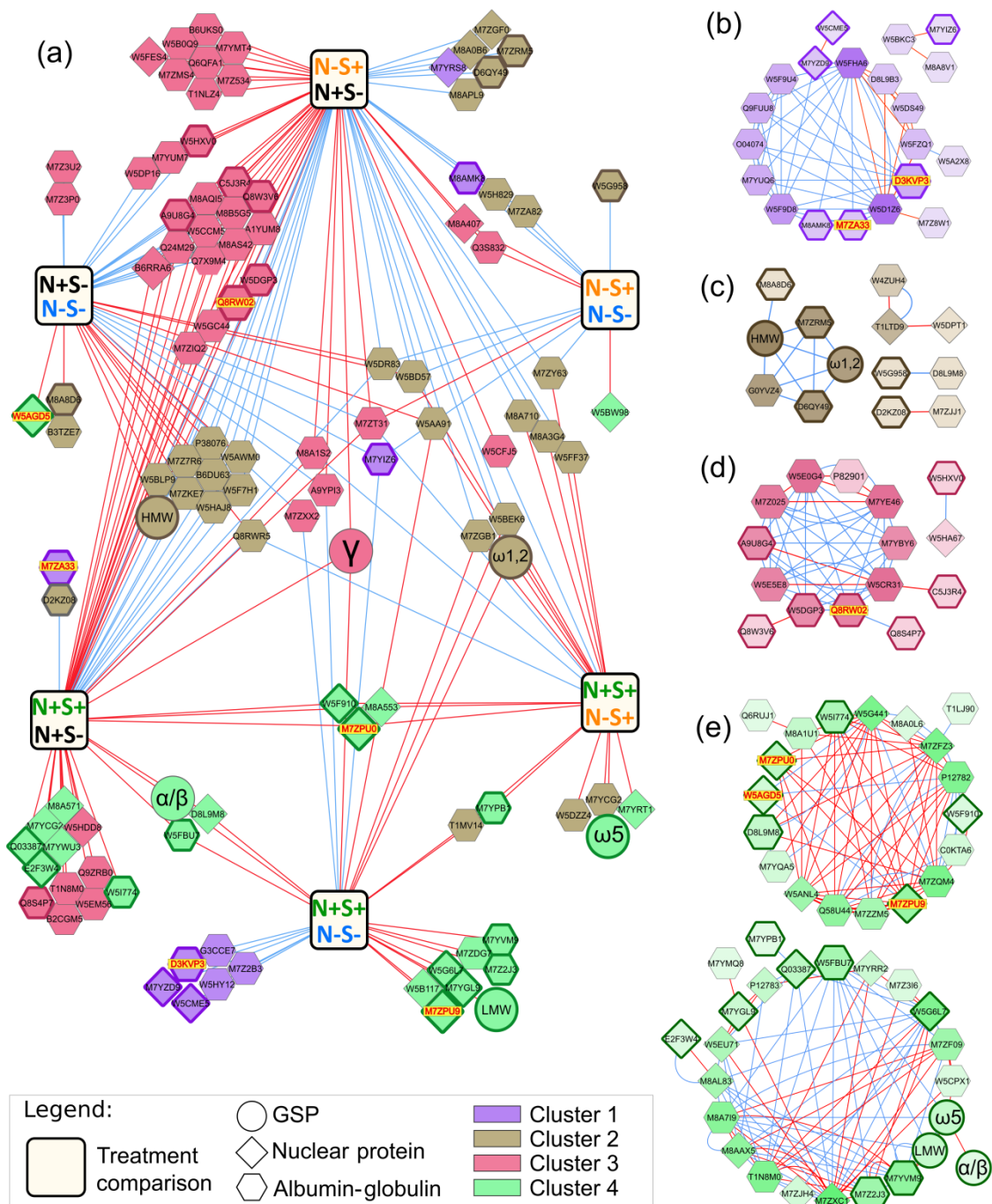
This article is protected by copyright. All rights reserved.



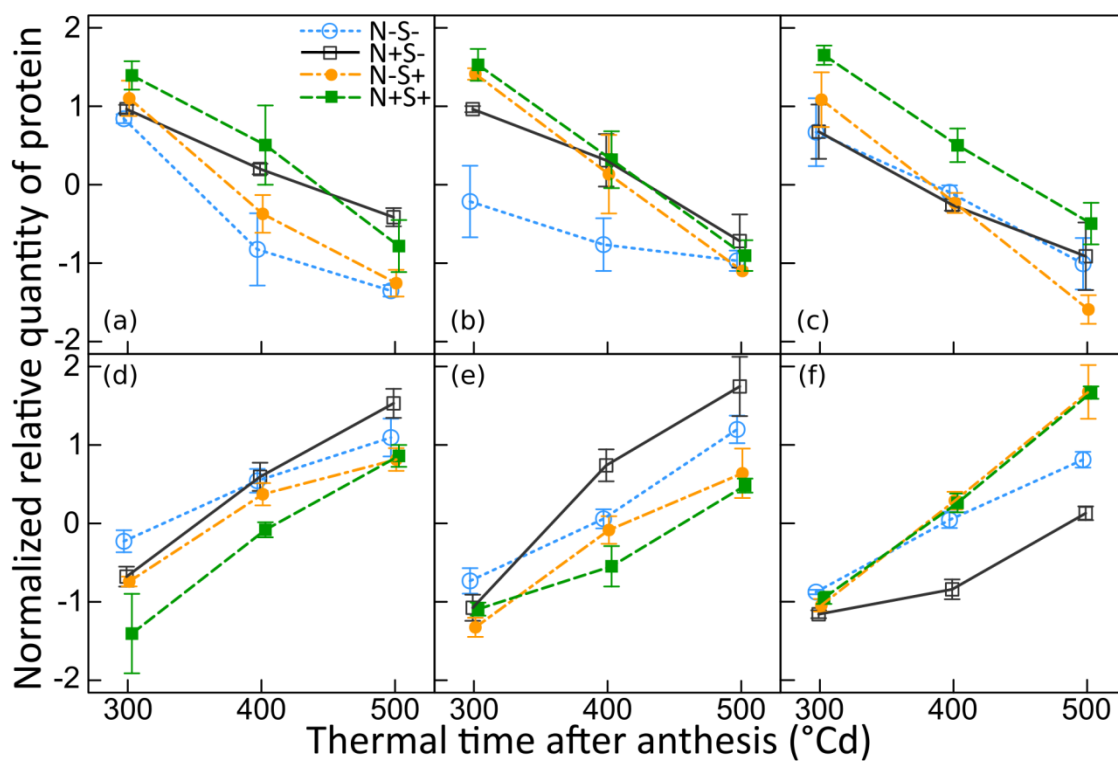
This article is protected by copyright. All rights reserved.



This article is protected by copyright. All rights reserved.

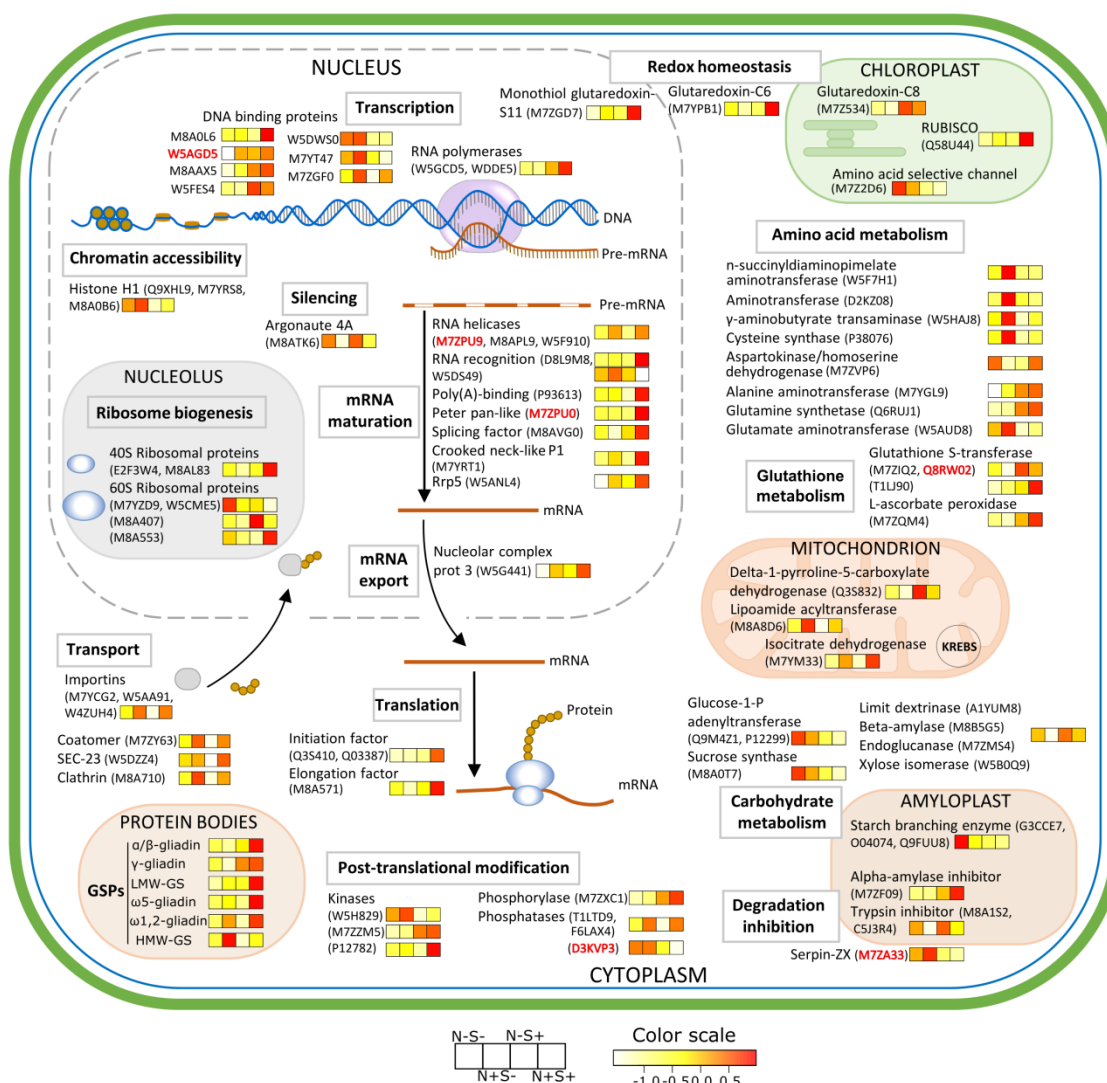


This article is protected by copyright. All rights reserved.



This article is protected by copyright. All rights reserved.





This article is protected by copyright. All rights reserved.