Transcriptome study of storage protein genes of field-grown barley in response to inorganic nitrogen fertilisers

Hansen, M.¹, Bowra, S.², Lange, M.¹, Holm, PB.¹, Vincze, E.¹
¹ University of Aarhus, Research Centre Flakkebjerg, Slagelse, 4200DK
² Verzyme, Plas Gogerddan, Aberystwyth, Wales SY23 3EB, UK

Abstract

The storage proteins of barley, both in terms of amino acid profile and quantity, are traits strongly influenced by the amount of nitrogen applied. Given this, we performed a developmental expression analysis of the genes from barley grains grown under field conditions to further our understanding the molecular and biochemical mechanisms underpinning nitrogen utilisation. A barley grain specific microarray, where a comprehensive set of genes involved in nitrogen mobilisation, storage protein synthesis and amino acid metabolism were assembled, was used to obtain a global but focused gene expression profile under different N regimes. Reviewing the expression of the storage protein homologues within the families revealed markedly different temporal profiles, for example some alleles were expressed very early in development. Furthermore the differential temporal expression of the homologues suggested that the genes of the different storage protein families were subject to different transcriptional regulation and responded differently to environmental stimuli. This finding may open the intriguing possibility of breeding selectively for specific alleles/homologues to confer enhanced amino acid profile of the barley storage proteins.

Introduction

Substantial increases in agricultural production has been realised during recent decades. Higher yields of field grown crops have been achieved through the development of superior varieties, improved disease management, and application of nitrogen fertilizers (Evenson and Gollin, 2003). However, while production has increased, the changes in agricultural practices have lead to undesired effects on the environment. In Denmark, measurements of nitrogen compounds in the ground water reservoirs have shown amounts above the permitted limits. Therefore, substantial efforts have in recent years been directed towards reducing the environmental nitrogen load by restrictions on the usage of fertilizers in inorganic and organic agriculture. To meet the demands for low input nitrogen fertilization regimes in Danish agriculture the nitrogen utilisation efficiency has to be improved. Therefore, it is crucial to obtain more information about the dry matter production,
nutrient uptake, plant component partitioning and processes of grain development. Especially grain filling is crucial to provide powerful resources for future breeding initiatives to have better seed quality and higher yield.

Barley (*Hordeum vulgare L.*) is the most widely grown cereal in Denmark and is mostly used as feed for monogastric livestock as pigs and chickens and for malt production. However, existing restrictions in Danish agriculture have affected the grain quality by reducing its protein content and affecting barley-malting quality (Madsen, 2006). Knowledge of the genetic basis of barley grain quality under low N input agricultural practice could facilitate a minimized nitrogen leach into the environment while maintaining productivity and quality.

Nitrogen is one of the most important macronutrients and is often related to limited plants growth and is employed for amino acid biosynthesis (See reviews, Stitt et al., 1999; Crawford et al., 1995). Both nitrogen and carbon act as signals regulating gene expression affecting many cellular processes (reviewed in: Gutiérrez et al., 2007; Palenchar et al., 2004; Rolland et al., 2001; Stitt et al., 2002). Few studies have investigated the nitrogen response in plants at the transcriptomic level e.g. in seedlings of Arabidopsis (Wang et al., 2000), shoots and roots (Scheible et al., 2004; Wang et al., 2003; 2004). In tomato roots several nitrate-induced genes were identified (Wang et al., 2001). Microarray studies have been carried out in various plant species during grain development (Drea et al., 2005; Ruuska et al., 2002; Sreenivasulu et al., 2006; Zhu et al., 2003). Recently, a microarray study defined the transcriptional differences in grains of organically and conventionally field-grown wheat (Lu et al., 2005). This was followed up by investigating the effect of low nitrogen stress on 10422 genes of root and leaf tissue in rice (Lian et al., 2006) and activity measurements of nitrate reductase, high-affinity ammonium transporters and glutamine synthetase both in root and shoot under different nitrogen forms and re-supply (Li et al., 2006). However, knowledge on the biochemical and molecular mechanisms controlling N uptake, assimilation, and recycling is still fragmentary (Mickelson et al., 2003).

Several other efforts have been made to elucidate genes that could improve nitrogen use efficiency (NUE). In maize and rice genes of the ammonium assimilation pathway have been suggested as candidate genes for future improvement of both nitrogen utilisation and thereby provide a withstanding biomass increase to meet the new demands of less nitrogen use in agricultural practices. In maize, studies have suggested that glutamine synthetase could enhance nitrogen utilisation (Hirel et al., 2005) as NAPH-glutamate synthase is the suggested candidate gene in rice (Obara et al., 2002). Furthermore protein turnover and the associated proteolytic pathways
appear to have a role within the complex web of nitrogen utilisation, which means that the activities of a range of enzymes affect N-metabolism in the plant (Yang et al., 2004; Donnison et al., 2007).

Barley (Hordeum vulgare L.), besides its importance as a crop, is an established model plant for agronomic, genetic, and physiological studies (Raun and Johnsson, 1998). The influence of N fertilizer levels and timing of application on grain yield and grain protein content was investigated in only a few studies (Penny et al., 1986; Bulman and Smith, 1993). In order to address the issue of changes in grain quality in relation to the effect of the decreased N input, we studied the developmental expression profile of genes affected by nitrogen of field-grown barley grains. For the simultaneous analysis of a great number of genes and their interactions in the complex patterns of the interconnecting metabolic pathways, a custom-made cDNA microarray was implemented. We constructed a cDNA microarray with 1035 genes primarily derived from the Clemson University cDNA library generated from developing barley grains. The elucidation of specific key candidate genes essential for grain filling, regulation of amino acid and prolamin metabolism could provide a powerful resource for future breeding initiatives for barley with improved nutritional composition and lower environmental impact.

Experimental procedures

Plant material
Spring barley (Hordeum vulgare L. cv. Barke) was grown in three field plots of 19.8 m² (12 m. * 1.65 m) during the summer of 2005, in Flakkebjerg, Denmark. Immediately after sowing the plots were fertiliser with NS24-7 (DLA Agro), which contains 12% ammonium, 12% nitrate and 7% sulphur, at a rate of 50 kg nitrogen/ha (low nitrogen dosage, LN), 120 kg nitrogen/ha (normal nitrogen dosage, NN) and 150 kg nitrogen/ha (high nitrogen dosage, HN). After one week the plots were fertilised again with PK 0-4-21 (DLA Agro) at a rate of 25 kg phosphor/ha and 60 kg potassium/ha. The plots were sprayed one month after sowing with a broad-spectrum herbicide mixture containing Express ST (Tribenuron-methyl 50%; E.L. du Pont de Nemours & Co), Oxitril CM (loxynil 17,32%; Bayer Crop Science) and Starane 180s (Fluroxypyr 180 g/liter; Dow Agrosciences) herbicides. Individual spikes were tagged and harvested in the morning (8 to 9 p.m) at 10, 15, 18 and 25 days after pollination (DAP) (see individual grains at different time-points and treatments in Figure 1). Samples were immediately frozen in liquid nitrogen, stored at -80°C until analysis. Mature grains were collected at 51 DAP for all three treatments and used for near-infrared
spectrometric analysis. The mature grains were harvested and used for the yield calculations of each nitrogen plot (16% water content of dry weight) at 51 DAP as well.

Figure 1. A phenotypic presentation of harvested grains at different nitrogen regimes during grains development. A) grains grown at 50 kgN/ha, B) 120 kgN/ha and C) 150 kgN/ha at 10 days after pollination (DAP), 15 DAP, 18 DAP and 25 DAP.

Near-infrared spectrophotometer
The mature grains were also harvested and analysed for water (%), starch (%) and protein content (%) in our laboratory using a near-infrared spectroscopy analyser (Foss Tecator, Infratec 1241, Grain Analyzer v.3.40). The near-infrared spectroscopy analyzer was calibrated and linked to the Danish NIT network (Buchmann et al., 2001).

Pre-processing of microarray data
Annotations of the probes can be found at: (URL: http://www.genome.clemson.edu/projects/barley/). The pre-processing is carried out according to Hansen et al., 2007.

Microarray design, data pre-procession, and identification of differential expression
The array contained 1035 genes. Each cDNA probe was spotted in triplicate in three subgrids across the slide to control for potential sources of variation in hybridisation across the area of the slide (technical replicates). The microarray experiments were performed using samples collected from field-grown barley subject to three different nitrogen regimes (50, 120, and 150kg/ha) at four time points (15, 18, 20 and 25 DAP). An interwoven loop experimental design was chosen (Altman and
Hua, 2006) in combination with three biological replicates per treatment resulting in 18 hybridisations while each hybridisation contained two different factors (time and treatment) for each slide (Fig. 2). This strengthens the statistical test of the two-way ANOVA identifying systematic error. Moreover, the ANOVA is based on well-developed statistical theories, which ensures that changes in gene expression as small as 1.2-fold can be detected as highly significant (Jin et al., 2001). The hybridisation protocol was performed according to Eisen and Brown (1999) with modifications according to Hansen et. al. (2007). Data acquisition and analysis was performed on an arrayWoRx microarray scanner (BioChipReader, Applied Precision, USA) using the arrayWoRx 2.0 software Suite.

Figure 2. An interwoven loop experimental design was used to assess the impact of applied nitrogen regimes on gene expression, during barley grain development, (Altman and Hua, 2006).

Annotation; The coloured circles represent the nitrogen regime (kg N/ha) Red=50 kg N/ha, Green=120 50 kg N/ha, Blue=150 50 kg N/ha. The first number in the circle corresponds to the nitrogen treatment and the second to the days after pollination (DAP). The RNA was prepared from three independent biological samples in total 36 independent RNA samples were hybridised. Within the interwoven loop figure the arrows indicate the specific pairwise combinations used for each hybridisation. The RNA sample at the head of the arrow was labelled with Cy5, while the RNA sample at the tail of the arrow was labelled with Cy3.

Clustering using Partitioning Around Medoids (PAM)

Co-regulated genes where identified by generating a distance matrix using a Pearson correlation between the expression values with the highest-confidence limits. The statistical package used was R (Becker et al., 1988) (http://www.r-project.org/). The distance matrix was subsequently clustered by the Partitioning Around Medoids method (PAM) (Kaufman and Rousseeuw, 1990) using the cluster package in R. The PAM algorithm is a robust version of k-means, and it searches for a specified number of medoids (representatives), k, around which clusters are constructed. The clusters were generated by minimizing the sum of the dissimilarities of all observations and assigning them to their closest medoid.
**Real-time RT-PCR expression analysis**

Total RNA was isolated from a pool of three individual grains from the middle of three independent barley spikes using FastRNA Pro Green Kit (Bio101, Systems, France) and resuspended in 50 µl DEPC-treated water according to manufacturers manual. The diluted RNA was quantified using a GeneQuant II DNA/RNA calculator (Pharmacia Biotech, Piscataway NJ, USA). First strand cDNA synthesis and real time RT-PCR were carried out as described Hansen et. al., (2007). Primer Express software (Applied Biosystems, Forster City, CA, USA) was used to design the primer to the same region as the microarray probe (Table S1 in Supplementary data available at JXB online). Primers for the hordein family were designed towards homologous regions identified using sequence alignments generated from accessions recovered from the genbank database. To investigate the specificity of each primer set (Table 1) a dissociation curve analysis was implemented. Expression level units of each gene of interest were calculated relatively to a calibrator and normalized to the housekeeping gene, actin in samples (Livak and Schmittgen, 2001). For the individual candidate genes, the expression at 18 DAP and 25 DAP was calibrated to the expression at 10 DAP (calibrator). The Ct value was obtained for each specific gene in the samples followed by a quality check of linear regression ($R^2$) and relative expression calculation for each gene using parameters of the software REST© according to Pfaffl et al. (2002).

<table>
<thead>
<tr>
<th>Gene product</th>
<th>Forward primer (5’ → 3’)</th>
<th>Reverse primer (5’ → 3’)</th>
</tr>
</thead>
<tbody>
<tr>
<td>C-hordein</td>
<td>ATAATCCCAGCAACCTCA</td>
<td>GATGGTGCACATCATTTATCCA</td>
</tr>
<tr>
<td>B-hordein</td>
<td>GCAAGGTATCCCTCCAGCAGC</td>
<td>TAAGTTGGGCAATCCGCACG</td>
</tr>
<tr>
<td>D-hordein</td>
<td>GACAGTCCACCGAGATGCTCA</td>
<td>CGATTACCACCAAAAGAGG</td>
</tr>
<tr>
<td>γ-hordein</td>
<td>TCCACCTAACTGCCTCCACC</td>
<td>CGGAGACTACACCAAATGCT</td>
</tr>
<tr>
<td>Glutelin</td>
<td>ACCATGGTCGAAAATGTGTCAAAT</td>
<td>TTAATAGTTGCCCGGCT</td>
</tr>
</tbody>
</table>

**Table 1.** Specific primers used for real time RT-PCR

**Results and discussion**

*Storage products of the harvested barley grain grown in different nitrogen regimes*

Yield was 41.6 hkg/ha for the low nitrogen treatment (LN), 55.5 hkg/ha for the normal nitrogen (NN) and 56.5 hkg/ha for the high nitrogen applications (HN). We observed a relatively higher proportion of developed tillers at normal and high nitrogen level compared to low nitrogen level. In the LN treatment the yield decrease was 25% when the fertiliser level was decreased 42%, while in the HN treatment 25% fertiliser increase corresponded to 2% yield increase only.
The harvested seeds from the three different nitrogen treatments were analysed for protein and starch contents. Grains grown at the NN contained 10.4% protein compared to LN and HN containing 8.3% and 11.1% protein, respectively. The starch content was highest in LN grains containing 64.8% whereas grains at NN and HN had 61.6% and 60.2%, respectively.

Phenotypic profiling
LN grains matured earlier (18 DAP) compared to NN and HN grains as measured by the colour of the spikes and the hardness of the endosperm according to Zadoks code 87 (Figure 1, Table 1; Zadoks et al., 1974). The flag leaves of LN barley showed increased rate of senescence at 15 DAP compared to the green and broad flag leaves of NN and HN barley. The senescence started out in the NN flag leaf at 18 DAP while HN stayed green until 22 DAP (data not shown).

Gene expression profiles affected by time, nitrogen and interaction
The two-way ANOVA resulted in genes affected by time, nitrogen treatment and the interaction of the two factors named T×N. The 145 most significantly affected by time were clustered into three clusters which correlated with an early, mid and late developmental stage (Figure 3). The value of k=3 was identified by manual inspection as the optimal number of clusters and it divided the data into three categories, one having the highest expression at day 10, another at 15 - 18 days, while the last cluster showed the highest expression at 25 days (Figure 3).
Figure 3. Cluster analysis. The gene expression profile of the 146 most significantly regulated genes representing an early-, mid- and late phase of the field-grown barley grain. The relative expression is depicted by the Z-score (transformed standard deviation) separated by the sampling time points. The coloured lines indicate the average gene expression during development.

The development transcriptional profile of storage proteins

It has long been recognised that there must be intricate/complex cross talk between the primary metabolism and storage product pathways. It is widely suggested that storage product accumulation occurs in the latter phase of grain or seed development in preparation for a period of dormancy before germinating and this would be fuelled by the utilisation of the storage products (for review see Shewry and Halford, 2002). However in our studies we observed storage product gene expression not only in the latter period of development but also the early stages. This is in line with a recent study, which reported early expression of hordeins in microspore-derived embryogenic development (Pulido et al. 2006). Pulido and colleagues (2006) suggested that these proteins might be synthesized and consumed according to the requirements of the embryogenic microspores and early embryos.

Barley storage protein is made up of glutelin, albumins, globulin and hordeins that are encoded by multi-gene families. To assess the expression of these families we have included homologues for barley storage proteins in the microarray and observed considerable variation in the temporal gene expression profile of the members of the family. For examples of the five B hordeins genes represented on the array, the expression profile of two falls within cluster 1, a further two exhibit an expression characteristic of cluster 2 while one gene appears to be expressed late in development and corresponds to cluster 3. Similar differences were observed for the expression patterns of the five γ-hordein genes, two were present in cluster 1 and three belonged to cluster 2, while the two D hordein genes were represented in cluster 2. Among the seven significantly expressed globulin genes two belonged to cluster 2 and five to cluster 3. The expression patterns of genes coding for a hordein C homologue, the albumins and the lysine-rich glutelin genes were all associated with cluster 3 where the respective mRNA levels increased late in development.

Reviewing the temporal expression profiles of the homologues within the storage protein families it is apparent and very striking that storage protein genes are expressed very early in development. Similar evidence has been reported by Rahman et al. (1984) using in vitro biochemical techniques, which demonstrated protein synthesis, thus the gene expression we observed seems to coincide with protein production.
Given this, the interplay as adjudged by the differential temporal expression of the homologues, suggests that the genes of each family of proteins is subject to different transcriptional regulation, implying that the regulatory units of the genes respond to different developmental of environmental stimuli, opening the intriguing possibility of breeding selectively for specific alleles/homologues to confer enhanced amino acid profile of the barley storage proteins.

**Validation by real-time RT-PCR**

The gene expression profiles obtained from the microarray experiments were validated by real-time RT-PCR for a selection of genes (Fig. 4). We used primers homologous to all members of the appropriate genes families present on the microarray so the real-time RT-PCR results represented an average expression level among the family members.

![Real-time PCR](image)

**Figure 4.** Genes of interest validated by real-time RT-PCR. Values of fold changes calculated relatively to 10 DAP are presented in logarithmic scale.

This was confirmed when we created, from the microarray absolute expression values, an average profile for the different hordein homologous (Fig. 5 shown examples of gamma- hordeins and glutelins). The profile of the C-hordein (data not shown) and the three glutelin homologues (Fig. 5) indicated that transcription of the respective genes continued to increase up to 25 DAP, which correlated with the results of the cDNA microarray and confirmed that the genes belong to cluster 3. Similarly the real-time PCR results matched the average profile pattern for the B-, D- (data not shown) and gamma-hordeins (Fig. 5).
Figure 5. The absolute expression profiles of the gamma hordein and glutelin genes mentioned in the real-time RT-PCR experiments: absolute expression values and the created average profile for the different homologous of the genes present on the microarray chips.

**Sulphur-rich- and sulphur-poor storage proteins were highly affected by N**

It is widely demonstrated that N nutrition increases the total protein fraction of the grain, and it correlates with the increase in the content of prolamins, named gliadins and glutelins in wheat (Wieser and Seilmeier, 1998; Triboi et al., 2000; Triboi et al., 2003) and hordeins in barley (Shewry et al., 2001). However, the metabolic/structural storage proteins albumins, globulins and glutelin are scarcely affected by increasing amounts of N input (Pechanek et al., 1997; Wieser and Seilmeier, 1998).

In our experiments several genes encoding storage proteins, both metabolic/structural and hordeins, were affected by nitrogen and/or TxN. Four metabolic/structural storage proteins coding genes, two glutelins and two globulins, was up-regulated by higher N and their developmental expression patterns belonged to Cluster 3, while one D-hordein and three γ-hordeins coding genes showed expression affected and they developmental profiles belonged to Cluster 2 (Figure 6 shows the example of D-hordein gene). The expression of the storage protein genes affected was higher in NN and HN grains than the LN grains (Figure 6). Interestingly, different homologues of the γ hordein family were affected by T&N as two γ1-hordeins and a single γ3-hordein were affected out of four present on the array. Furthermore, γ3 hordein expression indicated a different profile at different N levels (data not shown).
Figure 6. The absolute expression profiles of the D-hordein affected by the different N regimes: absolute expression values of the homologue (HVSMEi004I12).

Concluding remarks

The data we present provides transcriptomic analysis of cereal grain development of field grown material. It is based on a set of genes chosen from cDNA libraries of developing barley seeds. Although the available microarray data set deposited in the BarleyBase (http://www.plexdb.org) is very comprehensive it is limited to greenhouse material with 20DAP being the oldest developmental stage reported. Our study identified large number of genes responding to nitrogen and T&N from a range of different biological processes: genes of storage proteins accumulation; genes involved in the cell wall metabolism and starch biosynthesis, and cell cycle coding genes. Abiotic and biotic stress related genes were affected as well. We described and discussed in this paper the temporal expression profiles of a range of genes involved in storage protein accumulation. We conclude that the grain specific microarray coupled pathway specific analysis is a fast, reliable and cost-effective tool for monitoring temporal changes in the transcriptome of the gene families. The most intriguing aspect of this study was the observed differences in the expression patterns of the alleles/homologues of the different gene families coding genes influenced by time, nitrogen and T&N. Therefore microarray analysis could provide the knowledge required for rational design of breeding selectively for specific alleles/homologues to achieve better amino acid composition and increase the utility of the barley grain as food and feed.
Acknowledgement

We would like to thank K.B. Nellerup and O.B. Hansen for their excellent technical support and H. B. Rasmussen for the photographic work. This work was supported by a grant (93S-943-F07-00047) from The Danish Directorate for Food, Fisheries and Agri Business.

References


