Plant Biotechnology Journal

Plant Biotechnology Journal (2012) **10**, pp. 668–679 doi: 10.1111/j.1467-7652.2012.00705.x

Gene expression in the developing aleurone and starchy endosperm of wheat

Susan A. Gillies[1,*], Agnelo Futardo[1,2] and Robert J. Henry[2]

[1]*Southern Cross Plant Science, Southern Cross University, Lismore, NSW, Australia*

[2]*Queensland Alliance for Agriculture and Food Innovation, The University of Queensland, St Lucia, Qld, Australia*

Received 17 January 2012;
revised 1 April 2012;
accepted 5 April 2012.
**Correspondence* (Tel (61-2) 6620 3466; fax (61-2) 6622 2080; email susan.gillies@scu.edu.au)
Accession number for National Centre for Biotechnology Information/Gene Expression Omnibus No: X.

Summary

Wheat is a critical food source globally. Food security is an increasing concern; current production levels are not expected to keep pace with global demand. New technologies have provided a vast array of wheat genetic data; however, best use of this data requires placing it within a framework in which the various genes, pathways and interactions can be examined. Here we present the first systematic comparison of the global transcriptomes of the aleurone and starchy endosperm of the developing wheat seed (*Triticum aestivum*), at time points critical to the development of the aleurone layer; 6-, 9- and 14-day post-anthesis. Illumina sequencing gave 25—55 million sequence reads per tissue, of the trimmed reads, 70%—81% mapped to reference expressed sequence transcripts. Transcript abundance was analysed by performing RNA-Seq normalization to generate reads per kilobase of exon model per million mapped reads values, and these were used in comparative analyses between the tissues at each time point using Kal's Z-test. This identified 9414—13 202 highly differentially expressed transcripts that were categorized on the basis of tissue and time point expression and functionally analysed revealing two very distinct tissues. The results demonstrate the fundamental biological reprogramming of the two major biologically and economically significant tissues of the wheat seed over this time course. Understanding these changes in gene expression profiles is essential to mining the potential these tissues hold for human nutrition and contributing to the systems biology of this important crop plant.

Keywords: *Triticum aestivum*, aleurone, endosperm, transcriptome, next generation sequencing, grain.

Introduction

The most recent Food and Agriculture Organization and International Grain Council estimates give the global estimates for harvested wheat grain (*Triticum aestivum*) in 2010/11 at over 650 million tonnes (http://faostat.fao.org/; http://www.igc.int/). Although principally grown for human nutrition, wheat is also a source of food for livestock and is easily processed to provide the raw materials for a myriad of other foodstuffs and industrial commodities.

The economic and nutritional importance of wheat grain and the need for a better understanding of grain development to underpin any future potential manipulation of the seed have led to an abundance of studies addressing aspects of the grain transcriptome. These, however, examine the total grain or one tissue in isolation, or have followed the progress of specific genes and use a variety of sources for functional annotation (Laudencia- Chingcuanco *et al.*, 2007; Stamova *et al.*, 2009; Wan *et al.*, 2008, 2009). The work presented here follows the parallel development of the transcriptomes of two of the most important tissues of the wheat seed. Although our primary focus is in understanding the genetic mechanisms behind development of the aleurone layer, its derivation from endosperm makes interpretation of the data difficult if removed from the context of the endosperm. Consequently, both tissues have been included at each time point in our analysis to provide a direct comparison. Additionally, the controlled gene ontology vocabulary provided by the Gene Ontology project (http://www.geneontology.org/) was selected to provide a standardized vocabulary for functional analysis.

The starchy endosperm of the mature wheat seed is presently a primary source of nutrition, particularly in the western diet. It consists principally of carbohydrates in the form of simple starches, typically 55%—75% of total dry grain weight, with a storage protein content of 10%—20%. The aleurone layer develops at the perimeter of this tissue, where its primary biological role is in the digestion of the starchy endosperm to release nutrients in the form of free sugars and amino acids to the germinating embryo. A secondary role is in protection of the endosperm by storage of proteins which protect against stress and pathogens (Jerkovic *et al.*, 2010). In the wheat seed and in maize this layer is a single cell thick, whereas in other major cereals there are often multiple layers; in barley, there are typically three, and in rice one to several.

The aleurone layer is the most concentrated source of vitamins and minerals in the wheat seed and is additionally rich in proteins and lipids (Geisler-Lee and Gallie, 2005; Liu, 2011; Regvar *et al.*, 2011). Recent clinical trials have demonstrated its potential for beneficial health outcomes. Aleurone flour is a rich source of natural folate, a B-group vitamin critical in nucleotide synthesis, DNA methylation and gene expression, essential for the prevention of neurological deficits in the newborn. Increased consumption of aleurone has been shown to increase plasma folate (Fenech *et al.*, 2005) or increase plasma betaine and lower plasma homocysteine and low density lipoproteins (Price *et al.*, 2010), the latter three of which are associated with a lower risk of heart disease. Additionally, wheat aleurone is

© 2012 The Authors
Plant Biotechnology Journal © 2012 Society for Experimental Biology, Association of Applied Biologists and Blackwell Publishing Ltd

abundant in dietary fibre that is fermented in the gut, releasing short-chain fatty acids, which have potential roles in cancer prevention. Recent *in vitro* and animal studies indicate a possible role for aleurone intake in intestinal cancer prevention and treatment (McIntosh *et al.*, 2001; Stein *et al.*, 2010).

Aleurone development and maintenance is a dynamic process. This layer develops at the latter stages of endosperm specification. In a mechanism common amongst angiosperms, the endosperm is produced from a double fertilization event when a single sperm nuclei fuses with the two polar nuclei in the central cell. This is followed by a series of free nuclear divisions and directed migration of this nuclear material to the periphery of the central cell. Cell walls are formed from microtubular systems arising from the nuclear membranes. An initial anticlinal plane of cell division is followed by a periclinal plane of division, accompanied by cytokinesis which completes the first peripheral layer (Brown *et al.*, 1994). In wheat seed, this layer is complete by 6—7 days (S. Gillies, unpublished data, http://www.wheatbp.net).

Starchy endosperm cellularization is finalized by repetition of this process extending into the centre of the endosperm cavity, with progressive loss of the strict control of the planes of cell division (Becraft, 2001; Brown *et al.*, 1994; Olsen, 2001). Inner cells grow in size as they accumulate starch and protein bodies. The initial periclinal division signals the initiation of aleurone specification and gives rise to daughter cells which have distinct fates. As cellularization nears completion, the inner cells become starchy endosperm, the perimeter cells become aleurone (Becraft, 2001; Becraft and Yi, 2011; Olsen, 2001, 2004) which becomes resistant to desiccation, developing the capacity to remain living in the mature seed until germination of the embryo (Young and Gallie, 2000; Young *et al.*, 2004). In wheat, maturation of the aleurone is complete by 12 days after anthesis (personal observation, http://www.wheatbp.net).

The current understanding of the molecular mechanisms involved in aleurone development in cereals arises primarily from studies in maize. Aleurone fate is plastic and appears determined by positional cues, as it develops at the periphery of starchy endosperm masses of both a maize endosperm mutant and *in vitro* endosperm cultures (Geisler-Lee and Gallie, 2005; Gruis *et al.*, 2006; Olsen, 2004). Starchy endosperm, however, appears to be the default cell type as aleurone fails to develop at the periphery in the absence of appropriate cues (Becraft and Asuncion-Crabb, 2000; Becraft *et al.*, 1996; Gruis *et al.*, 2006; Lid *et al.*, 2002; Shen *et al.*, 2003).

Despite the importance of the wheat seed, the mechanism of development is far from established. Here we report the comparison of the transcriptomes of the developing aleurone and starchy endosperm of the wheat grain at time points critical in aleurone development using next generation sequencing. This work demonstrates two distinct tissues which undergo dynamic shifts over the time course examined and provides a genetic resource which may be used to promote further crop improvement.

Results

Experimental design

Six-, 9- and 14-day post-anthesis (DPA) tissues were selected as representative of pre/initial, mid and completed aleurone layer development. Six DPA was determined in preliminary experiments to be the first time point at which the initial peripheral layer of the endosperm could be isolated (Figures 1 and 2, http://www.wheatbp.net), which is enriched for the developing aleurone layer (Drea *et al.*, 2005). At 14 DPA, the aleurone layer in wheat is expected to be and was consistently found experimentally to be in its mature form 9 DPA represents a midpoint in development. Visibly the layer at 9 DPA resembled more the 6 DPA tissue, however, with a clearly more ordered and robust cell structure (Figure 2). During the time course studied starchy endosperm almost completes cellularization, it is complete by 16 DPA (Drea *et al.*, 2005; Wegel *et al.*, 2005). Seed was collected from appropriately aged wheat plants; the tissues were surgically separated, isolated and snap frozen; and cDNA was prepared and sequenced using an Illumina system. The level of cross contamination of tissues was low as demonstrated by the ability to detect genes with highly differential expression between the two tissues. A highly expressed gene originally detected in SAGE analysis (McIntosh *et al.*, 2007) has been the subject on ongoing analysis (Furtado unpublished) and was found to be expressed only in starchy endosperm at 6 and 9 DPA and at more than ten times the level in starchy endosperm at 14 DPA. Sequences were optimized by stringent removal of poor-quality reads and contaminating sequences.

RNA-Seq analysis using CLC Genomics Workbench (Version 4.0) (Katrinebjerg, Aarhus N, Denmark) was performed on the sequences after trimming. The Dana Farber Cancer Institute (DFCI) (Boston, MA) Wheat Gene Index was used as the reference database (Release 12.0, http://compbio.dfci.harvard.edu). This DFCI EST database is comprised of 93 508 tentative consensus sequences, 128 166 singleton ESTs and 251 mature transcripts for a total of 221 925 unique EST sequences. Of the trimmed sequences, 69.9%—81.4% mapped back to the DFCI ESTs. Of the mapped sequences, 42.1%—52.9% mapped specifically to one reference sequence, and 41.2%—57.9% nonspecifically (Table 1). Nonunique reads were discarded if matching more than ten distinct sequences, and if <10, reads were proportionately assigned. Sequenced transcripts were analysed after normalization of the counts to give reads per kilobase of exon model per million mapped reads (RPKM) (Mortazavi *et al.*, 2008).

Hierarchical cluster of sample analysis gives an overview of the total data. Cluster analysis of all RPKM values displays the close relationship between 6 and 9 DPA aleurone tissue (Figure 3). In contrast, in endosperm the 9 DPA and 14 DPA tissues are far more closely aligned. Of particular interest, 14 DPA aleurone tissue, after it has achieved its mature form, is

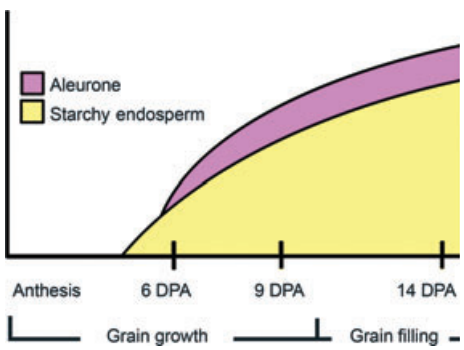

Figure 1 Schematic of development of the aleurone and endosperm layers. (Adapted from: Wheat: The big picture http://www.wheatbp.net).

© 2012 The Authors

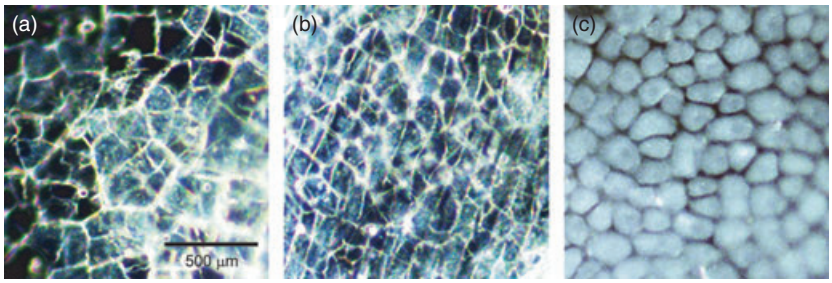

Figure 2 Development of the aleurone layer. Light microscopy at 10 × magnification demonstrates the increasing order and strength within this layer over time. (a) 6 day post-anthesis (DPA). (b) 9 DPA. (c) 14 DPA.

Table 1 Summary statistics for Illumina reads

	6 DPA		9 DPA		14 DPA	
	Al	En	Al	En	Al	En
Total reads	53 580 004	48 126 050	25 897 678	25 523 822	56 859 592	55 417 756
Mapped sequences (%)	27 150 859 (78.4)	30 581 013 (78.3)	4 549 034 (69.9)	4 104 682 (71.9)	38 911 327 (80.3)	37 502 901 (81.4)
Uniquely (%)	14 161 063 (52.1)	16 154 259 (52.9)	2 051 587 (45.1)	1 726 574 (42.1)	18 965 616 (50.0)	16 034 744 (42.8)
Non specifically (%)	12 989 796 (47.8)	14 426 759 (47.2)	2 497 447 (54.9)	2 378 108 (57.9)	18 945 711 (50.0)	21 466 157 (57.2)

The total reads produced by Illumina sequencing at each time point were trimmed and mapped against the Dana Farber Cancer Institute Wheat EST reference data base (Mapped sequences). Of the trimmed sequences 69.9–81.4% mapped back to the DFCI ESTs. Of these 42.1–52.9% mapped specifically to one reference sequence (Uniquely), and 41.2–57.9% non-specifically. Al, aleurone; En, starchy endosperm.

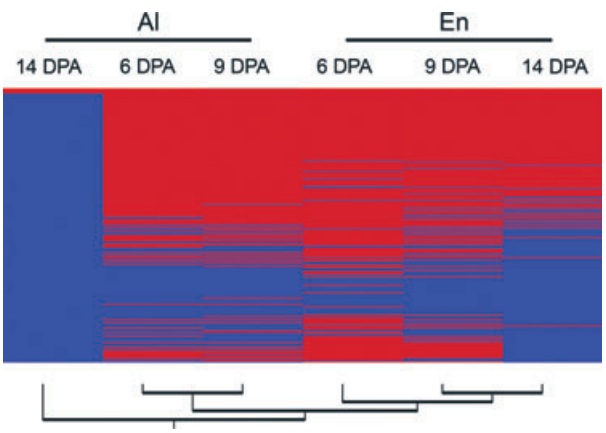

Figure 3 Hierarchical clustering of samples using all reads per kilobase of exon model values displays the relationship between each tissue and each time point. Clustering is calculated using the 1-Pearson correlation and single linkage. Al, aleurone; En, starchy endosperm.

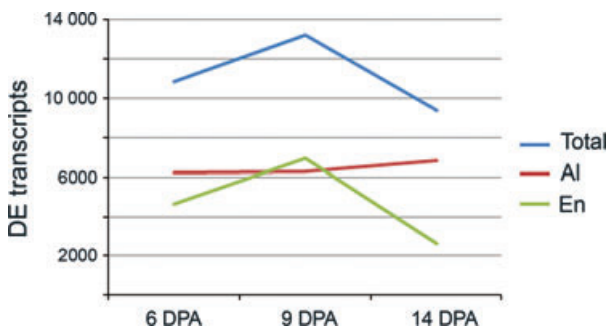

Figure 4 Differentially expressed transcripts over time. RNA-Seq reads per kilobase of exon model values were used in paired expression analysis between tissues at each time point using Kal's test. Multiple testing was corrected for by false discovery rate analysis of p-values and those with values less than or equal to 0.05 retained. Total, All differentially expressed transcripts at each time point. Al, up-regulated in aleurone; En, up-regulated in endosperm.

determined to be a very distinct tissue, as closely related to the endosperm tissues as aleurone tissues at earlier time points.

Differentially expressed (DE) transcripts

Aleurone (Al) was compared to starchy endosperm (En) at each time point. RPKM values were used in two-group comparisons using Kal's Z-test for statistical analysis, and false discovery rate (FDR) corrected p-values were generated to correct for multiple testing. ESTs that were expressed significantly differently (FDR-corrected *P*-value of 0.05 or less) were separated into those more highly expressed in the aleurone and those more highly expressed in the endosperm. As shown in Figure 4, the number of ESTs significantly up-regulated in the aleurone was relatively constant across the time points, in contrast the numbers of DE ESTs in the endosperm was highly variable.

These significantly differentially expressed ESTs were further divided into those expressed at one time point and those expressed at overlapping time points (Figure 5). Throughout the major stages of development of the aleurone layer, the numbers of these transcripts in aleurone remain relatively consistent (Figure 5a). Only an increase in the number of genes differentially expressed at 14 DPA exclusively is notable when the tissue has completed development, an observation

© 2012 The Authors

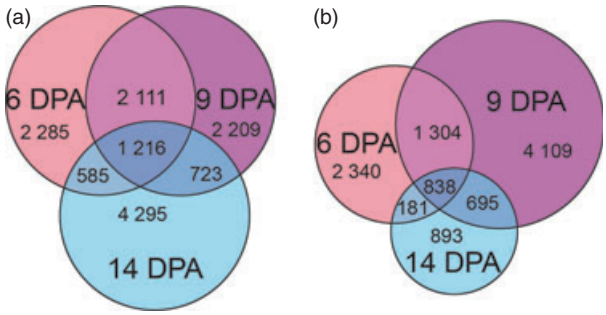

Figure 5 Differentially Expressed transcripts over time. Transcripts were categorized on the basis of those detected at one time point alone and those found at overlapping time points. (a) aleurone, (b) endosperm.

supported by Figure 3, where 14 DPA aleurone is displayed as a very distinct tissue from 6 to 9 DPA aluerone. As expected based on visual inspection of the tissues (Figure 2) and hierarchical clustering (Figure 3), the overlap between 6 and 9 DPA aleurone is far more extensive than the overlap between other time points. Over half (53%) of the DE transcripts identified in 6 DPA aleurone are also detected at 9 DPA. Endosperm, however, (Figure 5b) exhibits a greater fluctuation in numbers of highly differentially expressed genes as the seed moves from grain growth to grain filling. Interestingly, in both aleurone and endosperm the proportion of DE transcripts common to all time points is relatively low, comprising just 9.1% of the total in aleurone and 8.1% in starchy endosperm.

Functional annotation of differentially expressed transcripts identified all time points in aleurone or endosperm

The sequences represented in each sector of Figure 4 were subject to functional analysis. Annotation relied upon homology with distant species in many cases because of our relatively poor knowledge of the wheat genome. Repeated analysis of this data with more species-specific annotation will be worthwhile in the future. Analysis using the Mapman metabolic overview pathway (Usadel *et al.*, 2005, 2009) (Figure 6a) shows the expected results of high concentrations of transcripts binned into lipid metabolism (exotics, arrow 1; fatty acid synthesis and degradation, arrow 2) (Geisler-Lee and Gallie, 2005; Regvar *et al.*, 2011) and a high proportion of transcripts identified as associated with cell wall synthesis (arrow 3) in this tissue undergoing rapid cellularization. Less expected may be the significant concentration of transcripts identified as involved in the tricarboxylic acid cycle and mitochondrial electron transport (arrows 4, 5). These are not identified in DE expressed transcripts in endosperm and are indicative of the energy-intensive nature of aleurone development in comparison with endosperm.

Transcripts identified at all time points in endosperm (Figure 6b) are grouped primarily into the major carbohydrate bins of starch synthesis and degradation (arrows 1 and 2) (Morrison *et al.*, 1975a,b). Examination of secondary metabolites also displays distinct differences between aleurone and endosperm.

Analysis using level 2 GO terms gives an alternate overview of this DE data. Graphing each term as a proportion of the total demonstrates those DE transcripts common to all time points in either tissue (central region of Figure 5a,b) are, as expected for two closely related tissues at identical time points, broadly similar at level 2 (Figure 7a,c,e). However, although the GO terms identified are the same, the individual sequences within each category are unique to each tissue.

Biological process terms describe the operations or sets of molecular events with a defined beginning and end (http://www.geneontology.org/). Figure 6a demonstrates one of the few differences in terms is that aleurone has a greater proportion of transcripts identified as involved in localization (GO:0051179). These are the processes involved in transport to and maintenance of substances in a specific location. In endosperm, transcripts involved in response to stimulus (GO:0050896) are more frequently identified. These are described as transcripts involved in a change in state or activity of a cell as a result of a stimulus (http://www.geneontology.org/).

Analysis of biological process terms using the more detailed level 4 GO vocabulary, however, shows the emergence of clear differences between the tissues (Figure 7b). In aleurone, a higher proportion of DE transcripts are principally involved in processes which could be expected to be required for the development of specialized cells at a specific site. These processes include transport (transport (GO:0006810)) and definition of cell structure (regulation of anatomical structure size (GO:0090066); anatomical structure morphogenesis (GO:0009653); cell differentiation (GO:0030154); signal transmission (GO:0023060)). Additionally, as expected, transcripts identified as involved in lipid and protein metabolism (Geisler-Lee and Gallie, 2005; Regvar *et al.*, 2011) (lipid metabolic process (GO:0046493); cellular amino acid and derivative metabolic process (GO:0006519); cellular nitrogen compound metabolic process (GO:0034641)), are more prevalent in aleurone.

Conversely, transcripts that promote the rapid accumulation of carbohydrate and other storage macromolecules are more frequently detected in endosperm as predicted from previous analysis (Emes *et al.*, 2003; Morrison *et al.*, 1975b), including those involved in carbohydrate metabolic processes (GO:0005975); generation of precursor metabolites and energy (GO:0006091); cellular biosynthetic processes (GO:0044249); cellular macromolecule metabolic process (GO:0044260); and macromolecule biosynthetic process (GO:0009059).

Examining molecular function terms, which describe the elemental activities of a gene product, again at level 2 GO analysis the differences between the tissues are not marked (Figure 7c). However, consistent with above (Figure 7b), there are a higher proportion of sequences identified as being involved in transport (transporter activity (GO:0005215)) in aleurone. Endosperm has a higher proportion of transcripts having enzyme regulator activity (GO:0030234). These transcripts could be expected to assist in control of the rate of cellularization.

Level 4 molecular function analysis shows pronounced differences in activities between the tissues (Figure 7d). Aleurone has a higher ratio of sequences identified as involved in receptor activity (GO:0004872). This is defined as combining with an extracellular or intracellular messenger to initiate a change in cell activity and is consistent with this tissue also identified as having a higher proportion of sequences involved in signal transmission (Figure 7b). There is also a small proportion of transcripts involved in hydrolase activity, acting on acid anhydrides (GO:0016817), a category not found in starchy endosperm. This corresponds to a cluster of enzymes that hydrolyse the diphosphate bonds in nucleosidic phosphates and sulfonyl containing anhydrides (Enzyme Nomenclature, 1992, EC 3.6., http://enzyme.expasy.org/) such as those involved in hydrolysis

© 2012 The Authors

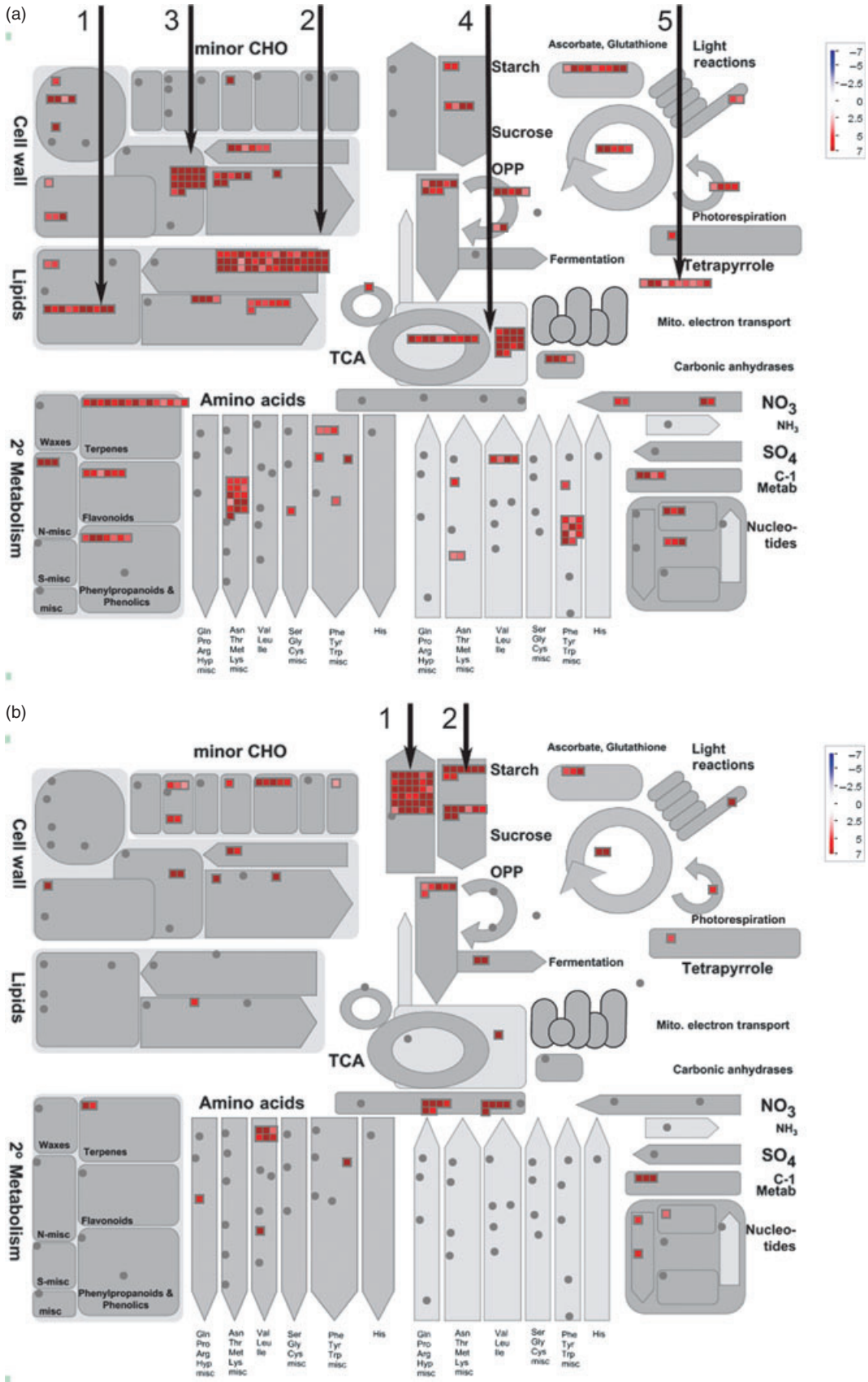

Figure 6 Mapman metabolic overview pathway of transcripts differentially expressed at all time points in either tissue. (a) aleurone, (b) endosperm.

© 2012 The Authors

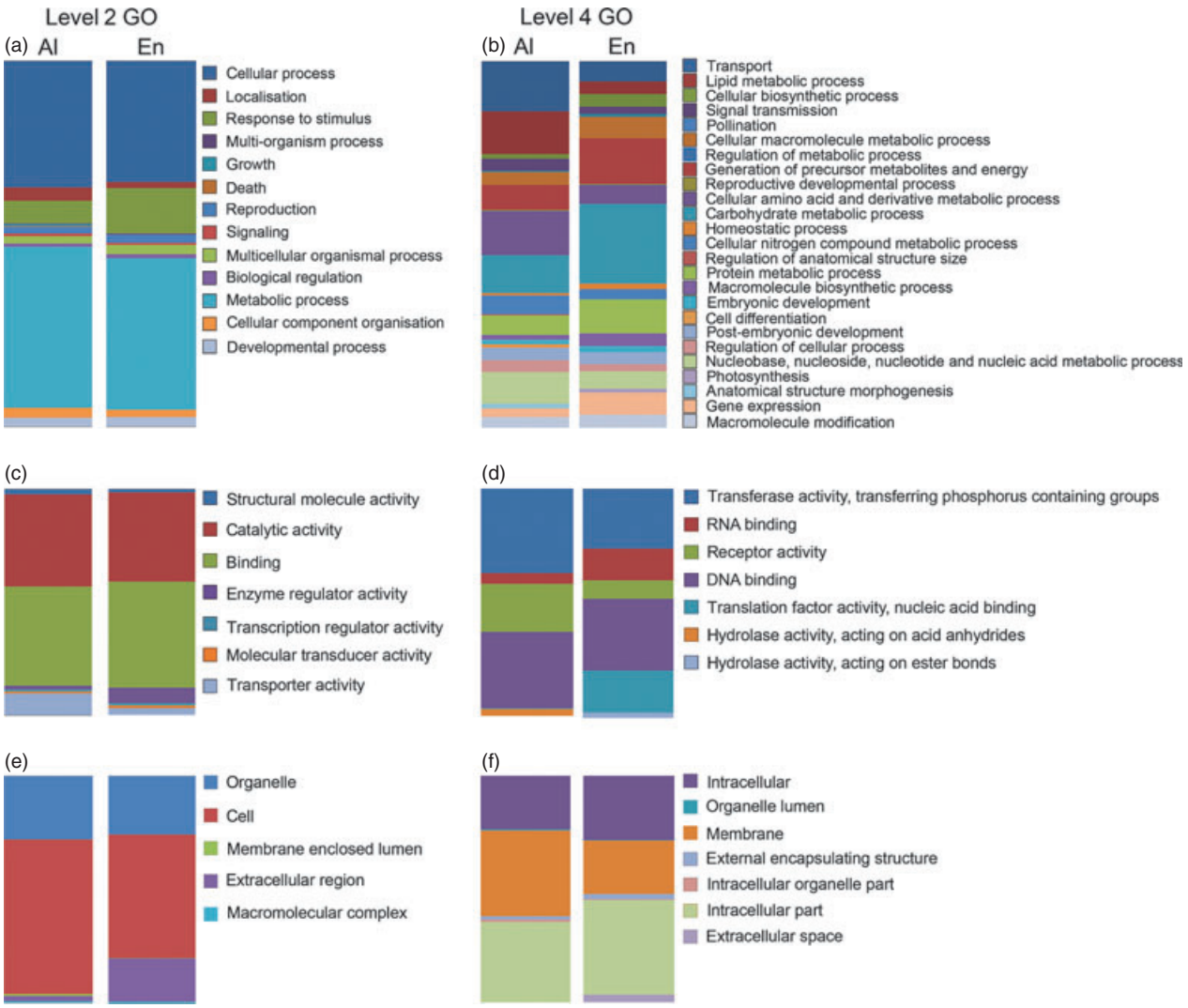

Figure 7 Gene Ontology (GO) classification of differentially expressed (DE) transcripts. The Level 4 and Level 2 GO vocabulary provided by the Gene Ontology project was used for functional annotation of DE transcripts detected at all time points in either tissue. (a, b) biological process terms. (c, d) molecular function terms. (e, f) cellular component terms. Al, aleurone; En, starchy endosperm.

of ATP and GTP to provide energy, which in agreement with Figure 6 indicates aleurone differentiation is a highly energy-intensive process.

In contrast, endosperm produces a small proportion of sequences identified as involved in hydrolase activity, acting on ester bonds (GO:0016788), a term not detected in DE transcripts up-regulated in aleurone. This group of hydrolases includes the deoxyribonucleases, ribonucleases, endonucleases, exonucleases and the phosphoric ester hydrolases (Enzyme Nomenclature, 1992, EC 3.1., http://enzyme.expasy.org/). This indicates rapid DNA/RNA turnover as phosphodiester bonds form the backbone of DNA/RNA molecules. Starchy endosperm additionally produces a far higher proportion of transcripts involved with RNA binding (GO:0003723) and translation factor activity, nucleic acid binding (GO:0008135).

Level 2 cellular component analysis (the parts of a cell or its extracellular environment) (Figure 7e) reflects the differing states of each tissue. Aleurone as a single layer tissue in wheat displays far fewer sequences identified as located in the extracellular region (GO:0005576) and has a small proportion of sequences identified as being targeted to the membrane enclosed lumen (GO:0031974), a category not detected in endosperm. Level 4 analysis (Figure 7f) shows only endosperm identifies sequences as being present in the extracellular space (GO:0005615), consistent with a tissue in the process of completing cellularization (Becraft, 2001; Drea *et al.*, 2005; Wegel *et al.*, 2005). In aleurone, a higher proportion of sequences are predicted to be found in membrane enclosed spaces identified by such terms as membrane (GO:0016020), organelle lumen (GO:0043233) and intracellular organelle part (GO:0044446).

Shifts in Level 4 GO categories in differentially expressed transcripts of aleurone and endosperm during aleurone development

Further analysis demonstrates the changes in each tissue over the time period critical to aleurone development. GO analyses use Level 4 terms. This level was selected as it was found to provide sufficiently detailed terms to clearly reveal the changes over time in each of the tissues. Those sequences that were

© 2012 The Authors

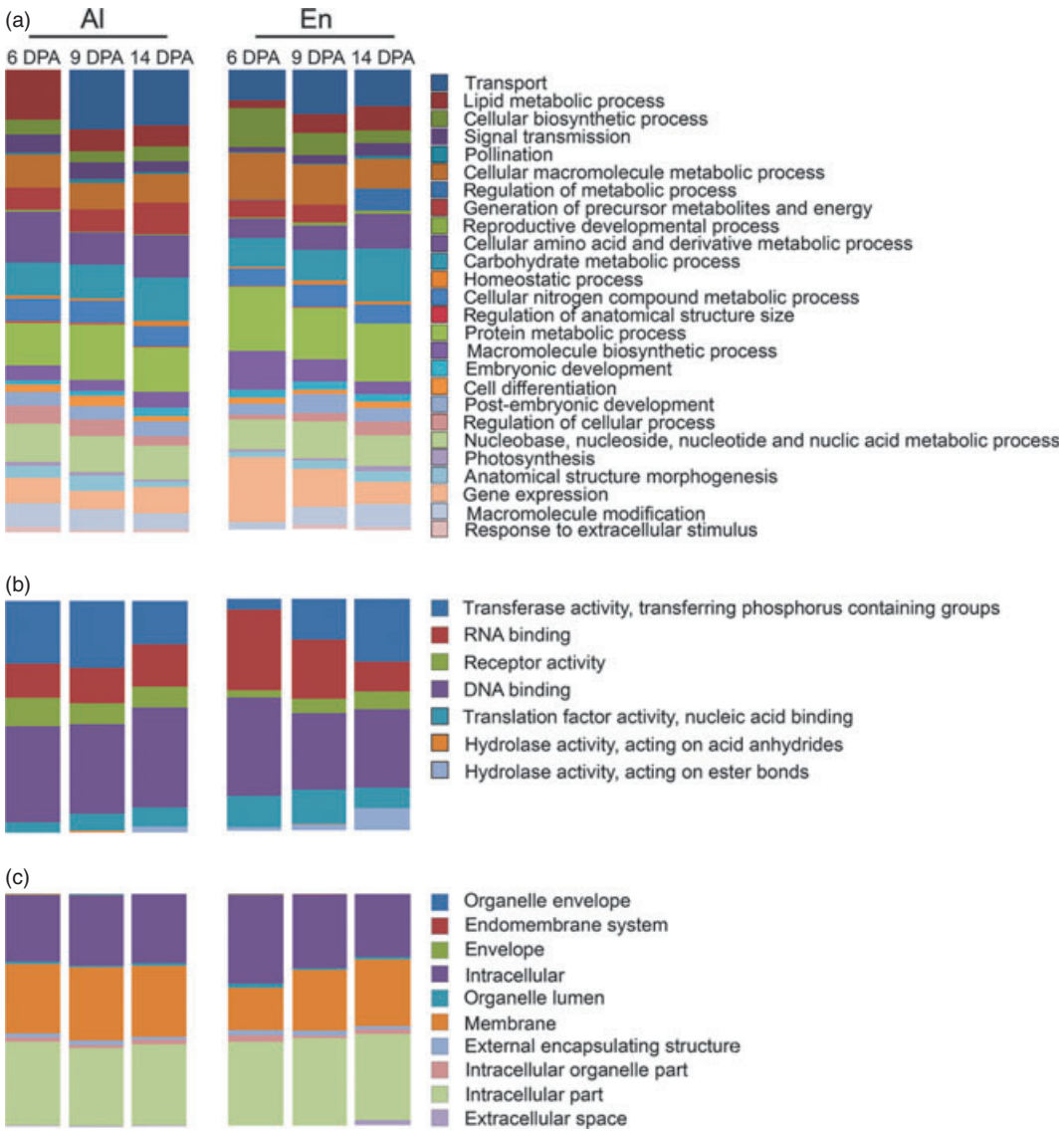

Figure 8 Functional Annotation of differentially expressed (DE) transcripts over time. Level 4 GO terms provided by the Gene Ontology project were used to annotate DE transcripts found exclusively at one time point in each tissue. (a) biological process terms; (b), molecular function terms; (c), cellular component terms. Al, aleurone; En, starchy endosperm.

found to be differentially expressed exclusively at one time point only (see Figure 5) were isolated and sent through the blast2GO program. Almost identical terms are retrieved for each tissue at each time point to those found for sequences common to all time points (Figure 6) and the total complement of sequences DE at any one time point (Figure S1). There are, however, clear changes in the proportion of transcripts in particular categories over time. Figure 8 shows the level 4 gene ontology terms retrieved at each time point, demonstrating the proportion of transcripts associated with that process/function/component over and above those produced at all time points. Figure S1 graphs the total data. This is the functional annotation of the total complement of all DE transcripts identified at any one time point. This is shown in comparison with DE transcripts identified at one time point exclusively (Only) and those transcripts differentially expressed at all time points (All). For almost every GO term, the total data produced proportions intermediate between 'Only' and 'All', demonstrating the 'Only' data on which the majority of the analysis rests represents a meaningful subset of the total data.

The biological process data (Figure 8a) show a sudden dramatic rise in aleurone in the proportion of transcripts associated with transport from 9 DPA. This is maintained at 14 DPA (transport (GO:0006810)), one of the few terms to display a proportional increase over time. This indicates that although by 14 DPA the tissue is in its final form, it remains very much an active tissue. Other increasing categories include generation of precursor metabolites and energy (GO:0006091), with a slight increase in carbohydrate metabolic process (GO:0005975) as the tissue matures and moves towards accumulation of storage molecules (Becraft, 2001; Becraft and Yi, 2011). Consistent with this completion of differentiation is the down-regulation of transcripts identified as involved in cell specification including cell differentiation (GO:0030154), anatomical structure morphogenesis (GO:0006810), signal transmission (GO:0023060) and regulation of cellular process (GO:0050794). Macromolecule

© 2012 The Authors

modification (GO:0043412) and lipid metabolic process (GO:0046493) are also reduced, indicating as expected, these constituents are nearing their final form.

Figure 8a demonstrates the endosperm is clearly in the process of reprogramming from rapid biosynthesis of a restricted range of storage macromolecules to a tissue requiring a broader range of molecules, with the requisite regulation required for the control of this diversified requirement. Cellular biosynthetic processes are down-regulated (GO:0044249) as are macromolecule biosynthetic process (GO:0009059) and cellular macromolecule metabolic process (GO:0044260), with an abrupt drop in transcripts involved in generation of precursor metabolites and energy (GO:0006139) after 9 DPA. There is a concurrent sharp rise in transcripts involved in regulation of metabolic process (GO:0019222) and macromolecule modification (GO:0043412), with an increase in regulation of cellular process (GO:0019222), signal transmission (GO:0023060), transport (GO:0006810) and anatomical structure morphogenesis (GO: 0009653). Additionally, there is an increase in lipid metabolism (lipid metabolic process GO:0046493) and cellular amino acid and derivative metabolic process (GO:0006519). However, the number of transcripts involved in carbohydrate metabolic processes (GO:0005975) continues to increase over time reflecting that endosperm remains principally a carbohydrate storage tissue. Interestingly, there is also a noticeable decline in the number of transcripts identified as involved in gene expression. This broad term describes the process by which a gene's sequence is converted into a mature gene product or products (proteins or RNA) (http://www.geneontology.org/). The decline in the proportion of these transcripts may be associated with a reduction in the rate of molecule accumulation.

Analysis of the molecular function data (Figure 8b) shows aleurone, as seen in previous analyses, changes little from 6 to 9 DPA. By 14 DPA, there is a drop in the number of sequences identified as involved in transferase activity, transferring phosphorus-containing groups (GO:0016772). This term covers a broad range of phosphotransferases and kinases. There is additionally a decline in receptor activity (GO:0004872). However, there is an increase in translation factor activity, nucleic acid binding (GO:0008135), again indicating at 14 DPA this tissue is still involved with the manufacture of new compounds.

Changes in endosperm are more marked. In contrast to aleurone, there is a large increase in the proportion of transcripts involved in the transfer of phosphorus-containing groups (GO:0016772), and additionally in transcripts associated with the hydrolysis of ester bonds (GO:0016788), again indicating rapid nucleotide turnover is necessary during starchy endosperm cellularization. Consistent with this is a large drop in transcripts involved in RNA binding (GO:0003723). There is also an increase in the proportion of terms identified as involved in receptor activity (GO:0004872).

Cellular component analysis (Figure 8c) shows the positional association of DE aleurone transcripts changing little over the time course examined, as expected as this is a single layer tissue. Again as endosperm becomes increasingly cellularized, this tissue exhibits more dynamic changes with transcripts moving from a simple intracellular location (intracellular GO:0005622) to being more commonly found at the membrane (GO:0016020) in proportions similar to that of aleurone. This finding is consistent with the finding of an increase in signal transmission over the time course examined (Figure 8a).

The Mapman cellular response overview visualization tool (Figure 9) displays the quite distinct stress responses of the two tissues. At 6 DPA, the majority of aleurone transcripts associated with abiotic stress are binned as touch/wounding transcripts (Figure 9a). At 9 DPA (Figure 9c), a greater proportion is identified as associated with drought/salt tolerance, while at 14 DPA (Figure 9e), heat and cold transcripts are commonly identified. In contrast, in endosperm the proportion of DE transcripts associated with heat reduces over the time course examined (Figure 9b,d,f), and touch/wounding transcripts increase over time. Both tissues display a dynamic stress response which alters rapidly over this short time course.

Of the antioxidant defence pathways, thioredoxin and ascorbate/glutathione become increasingly important in aleurone. Interestingly, the peroxiredoxins become suddenly apparent at 14 DPA, the importance of this group in redox regulatory reactions has recently been revealed (Dietz *et al.*, 2006; Tovar-Mendez *et al.*, 2011). In endosperm, few DE transcripts are identified as involved in redox reactions by 14 DPA, suggesting this stress response pathway has a narrow window for activation for wheat endosperm.

Aleurone displays a decreasing proportion of transcripts associated with cell division as expected as the tissue matures, whereas the proportion of transcripts identified as associated with cell cycling and cell development increase. Endosperm shows the proportion of DE transcripts binned as associated with cell division also reduces over time; however, in this tissue cell cycling also reduces, while development associated transcripts remains fairly static in proportions.

Discussion

Transcriptome sequencing provides a vast array of valuable data, in this analysis generating a total of over 250 million reads of approximately 75 base pairs. One of the challenges to be overcome is to reduce the complexity inherent in data sets of this size so as to provide an overview from which the tissue as a whole can be viewed and a framework in which the complex interplay of processes, functions and components can be integrated. To create this framework, the strategy of directly comparing each tissue at critical time points was selected, extracting the highly differentially expressed transcripts and annotating these. Here the GO vocabulary of the Gene Ontology project was utilized so that these standardized terms can be directly applied to other analyses. Further overviews of the data were examined using the Mapman tool. These data will provide a basis from which more detailed analysis can be begun in any number of experimental settings, particularly as wild-type wheat plants were examined. These can be used in comparisons with other wild-type varieties or treated/mutant plants. Here we provide a direct comparison of the most important biological and economic tissues of the wheat seed.

Endosperm is principally a carbohydrate storage tissue, and aleurone additionally accumulates lipids and proteins

These data validate findings from previous analyses. Figure 5b shows that of the DE transcripts detected at all time points, in starchy endosperm almost half are in the categories of carbohydrate metabolic process and generation of precursor metabolites and energy, as expected of what is principally a carbohydrate storage tissue (Emes *et al.*, 2003; Morrison *et al.*,

© 2012 The Authors

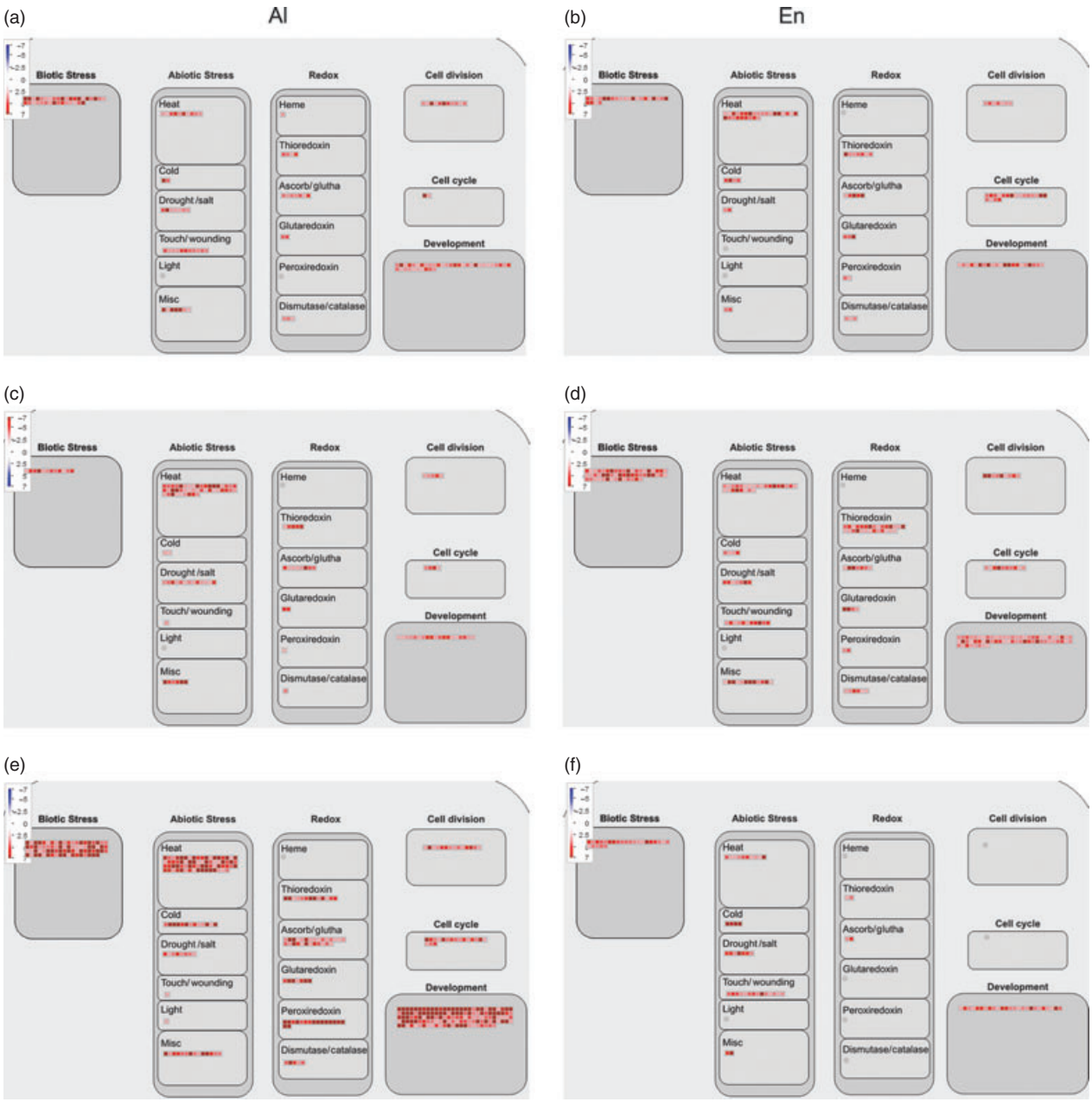

Figure 9 Mapman cellular response overview showing transcripts differentially expressed at each time point. 6 day post-anthesis (DPA) a, b, 9 DPA c, d, 14 DPA e, f. Al, aleurone; En, endosperm.

1975b; Stamova *et al.*, 2009). Additionally, as expected, aleurone shows a comparable proportion of DE transcripts for lipid metabolic process, carbohydrate metabolic process and cellular amino acid metabolic process (Liu, 2011; Morrison *et al.*, 1975a; Regvar *et al.*, 2011). Figure 7 demonstrates the consistent gradual changes in the proportions of the GO terms over the time course examined in each of the tissues, which when combined with the expected relationships shown by the hierarchical structure analysis demonstrates the validity of the approach taken. Additionally, this investigation demonstrates that although wheat currently does not have a completed annotated genomic sequence, an EST database can be used as a valid reference source.

It is essential to note, however, this analysis required the use of the more detailed Level 4 GO terms to clearly reveal the differences between the tissues, reinforcing the view that the selection of appropriate and consistent terms is critical in the meaningful interpretation of the data.

Aleurone and endosperm develop at significantly different rates

A combination of hierarchical sample clustering, DE transcript number and GO term analysis demonstrates the close yet distinct relationship between aleurone and starchy endosperm. Cluster analysis (Figure 3) demonstrates by 6 DPA, when aleurone is still being defined, it is already a very distinct tissue from

© 2012 The Authors

the endosperm from which it has been derived, far more closely related to 9 DPA aleurone than starchy endosperm. Clearly a large proportion of genes involved in aleurone development have already been activated by 6 DPA. Additionally, these data in combination with visual inspection (Figure 2) demonstrate 14 DPA aleurone is a very distinct tissue to aleurone at earlier time points. This observation is consistent with Figure 5, which shows the number of DE transcripts up-regulated in aleurone at 14 DPA alone is double that of 6 and 9 DPA. This finding, however, would not be detected by examining total numbers of DE transcripts alone (Figure 4) or by examining the changes in proportions of transcripts identified by GO terms over time (Figure 7).

Aleurone and endosperm reprogram over the time course examined

Examination of the functional annotation of DE transcripts emphasizes the marked differences between these tissues and reveals the reprogramming of each of these tissues over time. Starchy endosperm shifts from a tissue dominated by the rapid accumulation of storage macromolecules as discussed above (Figures 6 and 7) to one requiring a broader array of molecules, evidenced by the increasing proportion of transcripts associated with lipid metabolic process (GO:0005975), cellular amino acid and derivative metabolic process (GO:0006575) together with transcripts identified as involved in regulation of metabolic process (GO:0019222) and regulation of cellular process (GO:0050794). Aleurone, in contrast, as it finalizes cell specification, shows a slight shift towards accumulation of storage macromolecules as evidenced by the accumulation of the proportion of transcripts associated with generation of precursor metabolites and energy and carbohydrate metabolic process (Figure 7). Interestingly, however, the high and increasing proportion of transcripts identified as involved in transport indicate aleurone remains a very active tissue even after differentiation is complete.

Aleurone and endosperm display different molecular functions

An important finding is the fundamental differences in molecular function between the two tissues, which may point the way for manipulation of these tissues in the future. Although the proportion of sequences predicted to be involved in DNA binding is consistent between the tissues and across the time course (Figures 7b and 8b), there is significant differences and quite profound changes in all other categories, particularly in the starchy endosperm which has a far higher proportion of sequences involved with RNA binding and translation factor activity. This indicates that endosperm utilizes extensive post-transcriptional control of the cellular machinery, allowing the cell to rapidly manipulate protein production without new mRNA synthesis, processing or export (Dever, 2002; Mata *et al.*, 2005). Interestingly, the proportion of each of these transcripts drops significantly as the tissue approaches complete cellularization, suggesting post-transcriptional control is a particularly important process to facilitate rapid cellularization.

Aleurone and endosperm have distinct stress responses

Both aleurone and endosperm displayed rapidly changing stress response particularly in association with abiotic stress and redox reactions. Endosperm, however, identified few DE transcripts as involved in any redox pathway by 14 DPA. These data may help identify particular windows of vulnerability to particular stressors in early grain development.

Conclusion

By 6 DPA, aleurone and endosperm are distinct tissues which develop at different rates. In aleurone, highly differentially expressed transcripts are fairly evenly distributed between lipid, protein and carbohydrate biogenesis. In endosperm, carbohydrate accumulation is by far the dominant activity. However, the proportions of transcripts associated with this activity reduce over time, perhaps surprisingly as cellularization of the endosperm is not yet complete. Rapid development of the starchy endosperm appears facilitated by post-transcriptional factors, and these tissues have disparate and fast altering responses to stress.

This work provides a framework for future analysis of the developing cereal grain. It will be of particular value when the annotated wheat sequence becomes available so that the analysis of transcription start sites, polyadenylation signals, alternative splice sites and post-transcriptional modifications can be applied to these data. While this overview reinforces much current experimental data on the development of the aleurone and starchy endosperm layers, it additionally throws up new possible avenues for investigation. These data contribute to the understanding of the development of the aleurone and starchy endosperm layers in cereal grains and contribute in particular to the systems biology of the wheat seed, an essential food source.

Experimental procedures

Plant tissue collection

The *Triticum aestivum* cultivar Banks (Australian Spring wheat) was planted at four seeds per pot in 25-cm pots in soils containing equal parts Lithuanian peatmoss, perlite (size P500) and vermiculite (size 2) with 5 kg/m^3 long-term release fertilizer, 1 kg/m^3 trace element mix and 1 kg/m^3 dolomite and grown to maturity in a greenhouse during spring and early summer under natural light. Greenhouses were kept below 37°C by evaporative cooling. Plants were tagged at anthesis and seed was collected from appropriately aged (6, 9 and 14 DPA) plants from the central third of the spike. The stages of development in this study correspond to those analysed in an earlier SAGE study of whole seed transcripts (McIntosh *et al.*, 2007) with the 14 DPa point being at a midpoint of seed development prior to seed desiccation and with a wide diversity of genes being transcribed. Generally, aleurone was separated from endosperm by microdissection over ice by incision of the whole seed, prising off the pericarp and dissecting away the embryo/presumptive embryo. The testa was then gently scraped from the endosperm layers. The starchy endosperm was rolled from the aleurone/presumptive aleurone layer and the desired tissue quickly rinsed in ice-cold phosphate buffer saline, then snap frozen in liquid nitrogen. At 6 and 9 DPA, the endosperm is not completely cellularized, and therefore the contents of the central cavity were scraped into a cooled tube and snap frozen. At 6 DPA, these contents are principally liquid and at 9 DPA principally cellular. A minimum of five individual seeds from three separate spikes from three individual plants were pooled for each tissue preparation.

© 2012 The Authors

RNA isolation

Total RNA was prepared using a Trizol (Invitrogen, Carlsbad, CA) protocol and purified using an RNeasy Plant Kit (Qiagen, Hilden, Germany). Total RNA quality and concentration were determined using the RNA 6000 Pico kit (Agilent, Santa Clara, CA) on a 2100 Bioanalyzer (Agilent).

cDNA construction

Approximately 1.5 μg of purified cDNA was prepared from each tissue. Total cDNA was prepared using the Ambion MessageAmp II aRNA Amplification Kit (Invitrogen) essentially as per the manufacturer's instructions. However, the first-strand synthesis primer was replaced by one designed to contain a BpuE1 site ($CTTGAG(N)_{16/14}$) 19 bp upstream from a VN clamp adjacent to the polyA tail. This site was used to remove the polyA tail after cDNA synthesis following BpuE1 digestion (NEB, Ipswich, MA). The polyA tails were removed by spinning through Qiaquick PCR purification (Qiagen) size exclusion columns after enzymatic digestion.

Sequencing

Sequencing was performed in-house on the Illumina GAIIx system with a paired end 75-bp system.

Data analysis

Approximately 25—55 million reads were obtained per tissue. Sequences were initially trimmed using default quality parameters of the CLC Genomics Workbench Version 4.0. (CLC bio, Aarhus, Denmark). Remaining contaminating sequences were removed using an editing package (010 Editor, http://www.sweetscape.com/010editor/) and the reads returned to CLC for a further round of trimming based on quality scores to optimize the quality of the sequences used in subsequent analyses. The trimmed sequences were sent through the RNA-Seq analysis program of the CLC platform, mapping against the unannotated reference sequences derived from the DFCI Wheat Gene Index, Release 12.0 (The Computational Biology and Functional Genomics Laboratory, Dana Farber Cancer Institute and Harvard School of Public Health) using default parameters. The minimum length fraction was 0.9 and the minimum similarity fraction was 0.8. RPKM was selected as the normalized expression value. Dataset S1 provides the complete set of raw sequences and RNA-Seq data for each tissue.

Tissues that had undergone RNA-Seq analysis were used in paired expression analysis experiments between aleurone and starchy endosperm at 6, 9 and 14 DPA using a two-group comparison (Kal's test) and existing RPKM expression values. Multiple testing was corrected for by obtaining FDR analysis of p-values. Sequences with a FDR-corrected *P*-value of less than or equal to 0.05 were selected. These were separated into those differentially expressed highly in the aleurone and those expressed in the starchy endosperm. These were then categorized into those DE at one time point or at common time points.

Each category of ESTs was submitted to the BLAST2GO program (http://www.blast2go.org/, (Bioinformatics and Genomics Department, Centro de Investigación Príncipe Valencia, Valencia, Spain), (Conesa and Gotz, 2008; Conesa *et al.*, 2005) for functional annotation. This retrieves the gene ontology (GO) terms using the structured vocabulary provided by the Gene Ontology project (http://www.geneontology.org/) which separates the annotations into the three vocabularies of biological processes, molecular function and cellular components. To expand and refine the annotation data, the InterPro Scan, ANNEX and GOSlim functions were applied to the GO annotations from within the BLAST2GO program (BioBam Bioinformatics, Valencia, Spain). InterPro combines a number of protein signature searches (Zdobnov and Apweiler, 2001), the results of which were attached to the GO terms. ANNEX deduces new relationships between the GO terms (Botton *et al.*, 2008; Myhre *et al.*, 2006), and these results were further incorporated adding on average a further 18% of annotations. To refine the annotations to allow adequate visualization, GOslim was applied using the GOPlant database. GOSlim reduces the annotations used to those of the selected level of categories or nodes. This allows a higher-level simplified visualization of the data. The GOSlim function was applied in a species-specific context (plant.obo), which has been observed to increase retrieval of plant-specific terminology and reduce category fragmentation (Botton *et al.*, 2008).

Alternately, DE ESTs were visualized using the Mapman tools (Usadel *et al.*, 2005, 2009). Mapping files were created using the Mercator tool (http://mapman.gabipd.org/web/guest/mercator) which bins all DE transcripts according to hierarchical ontologies. Default parameters were retained and ORYZA TIGR5 rice protein and IPR Interpro scans selected. Experimental files were created using the test statistic calculated from the Kal's test described above.

Acknowledgement

This work was supported by the Grain Foods Cooperative Research Centre.

References

Becraft, P.W. (2001) Cell fate specification in the cereal endosperm. *Semin. Cell Dev. Biol.* **12**, 387–394.

Becraft, P.W. and Asuncion-Crabb, Y. (2000) Positional cues specify and maintain aleurone cell fate in maize endosperm development. *Development*, **127**, 4039–4048.

Becraft, P.W. and Yi, G. (2011) Regulation of aleurone development in cereal grains. *J. Exp. Bot.*, **62**, 1669–1675.

Becraft, P.W., Stinard, P.S. and McCarty, D.R. (1996) CRINKLY4: a TNFR-like receptor kinase involved in maize epidermal differentiation. *Science*, **273**, 1406–1409.

Botton, A., Galla, G., Conesa, A., Bachem, C., Ramina, A. and Barcaccia, G. (2008) Large-scale Gene Ontology analysis of plant transcriptome-derived sequences retrieved by AFLP technology. *BMC Genomics*, **9**, 347.

Brown, R.C., Lemmon, B.E. and Olsen, O.A. (1994) Endosperm development in barley: microtubule involvement in the morphogenetic pathway. *Plant Cell*, **6**, 1241–1252.

Conesa, A. and Gotz, S. (2008) Blast2GO: a comprehensive suite for functional analysis in plant genomics. *Int J Plant Genomics*, **2008**, 619832.

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.

Dever, T.E. (2002) Gene-specific regulation by general translation factors. *Cell*, **108**, 545–556.

Dietz, K.J., Jacob, S., Oelze, M.L., Laxa, M., Tognetti, V., de Miranda, S.M., Baier, M. and Finkemeier, I. (2006) The function of peroxiredoxins in plant organelle redox metabolism. *J. Exp. Bot.* **57**, 1697–1709.

© 2012 The Authors

Drea, S., Leader, D.J., Arnold, B.C., Shaw, P., Dolan, L. and Doonan, J.H. (2005) Systematic spatial analysis of gene expression during wheat caryopsis development. *Plant Cell*, **17**, 2172–2185.

Emes, M.J., Bowsher, C.G., Hedley, C., Burrell, M.M., Scrase- Field, E.S. and Tetlow, I.J. (2003) Starch synthesis and carbon partitioning in developing endosperm. *J. Exp. Bot.* **54**, 569–575.

Fenech, M., Noakes, M., Clifton, P. and Topping, D. (2005) Aleurone flour increases red-cell folate and lowers plasma homocyst(e)ine substantially in man. *Br. J. Nutr.* **93**, 353–360.

Geisler-Lee, J. and Gallie, D.R. (2005) Aleurone cell identity is suppressed following connation in maize kernels. *Plant Physiol.* **139**, 204–212.

Gruis, D.F., Guo, H., Selinger, D., Tian, Q. and Olsen, O.A. (2006) Surface position, not signaling from surrounding maternal tissues, specifies aleurone epidermal cell fate in maize. *Plant Physiol.* **141**, 898–909.

Jerkovic, A., Kriegel, A.M., Bradner, J.R., Atwell, B.J., Roberts, T.H. and Willows, R.D. (2010) Strategic distribution of protective proteins within bran layers of wheat protects the nutrient-rich endosperm. *Plant Physiol.* **152**, 1459–1470.

Laudencia- Chingcuanco, D.L., Stamova, B.S., You, F.M., Lazo, G.R., Beckles, D.M. and Anderson, O.D. (2007) Transcriptional profiling of wheat caryopsis development using cDNA microarrays. *Plant Mol. Biol.*, **63**, 651–668.

Lid, S.E., Gruis, D., Jung, R., Lorentzen, J.A., Ananiev, E., Chamberlin, M., Niu, X., Meeley, R., Nichols, S. and Olsen, O.A. (2002) The defective kernel 1 (dek1) gene required for aleurone cell development in the endosperm of maize grains encodes a membrane protein of the calpain gene superfamily. *Proc. Natl. Acad. Sci. USA*, **99**, 5460–5465.

Liu, K. (2011) Comparison of lipid content and fatty acid composition and their distribution within seeds of 5 small grain species. *J. Food Sci.* **76**, C334–C342.

Mata, J., Marguerat, S. and Bahler, J. (2005) Post-transcriptional control of gene expression: a genome-wide perspective. *Trends Biochem. Sci.* **30**, 506–514.

McIntosh, G.H., Royle, P.J. and Pointing, G. (2001) Wheat aleurone flour increases cecal beta-glucuronidase activity and butyrate concentration and reduces colon adenoma burden in azoxymethane-treated rats. *J. Nutr.* **131**, 127–131.

McIntosh, S., Watson, L., Bundock, P.C., Crawford, A.C., White, J., Cordeiro, G.M., Barbary, D., Rooke, L. and Henry, R.J. (2007) SAGE of the developing wheat Caryopsis. *Plant Biotechnol. J.* **5**, 69–83.

Morrison, I.N., Kuo, J. and O' Brien, T.P. (1975a) Histochemistry and fine structure of developing wheat aleurone cells. *Planta*, **123**, 105–116.

Morrison, W.R., Mann, D.L., Soon, W. and Coventry, A.M. (1975b) Selective extraction and quantitative analysis of non-starch and starch lipids from wheat flour. *J. Sci. Food Agric.* **26**, 507–521.

Mortazavi, A., Williams, B.A., McCue, K., Schaeffer, L. and Wold, B. (2008) Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nat. Methods*, **5**, 621–628.

Myhre, S., Tveit, H., Mollestad, T. and Laegreid, A. (2006) Additional gene ontology structure for improved biological reasoning. *Bioinformatics*, **22**, 2020–2027.

Olsen, O.A. (2001) ENDOSPERM DEVELOPMENT: cellularization and Cell Fate Specification. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **52**, 233–267.

Olsen, O.A. (2004) Nuclear endosperm development in cereals and Arabidopsis thaliana. *Plant Cell*, **16**(Suppl), S214–S227.

Price, R.K., Keaveney, E.M., Hamill, L.L., Wallace, J.M., Ward, M., Ueland, P.M., McNulty, H., Strain, J.J., Parker, M.J. and Welch, R.W. (2010) Consumption of wheat aleurone-rich foods increases fasting plasma betaine and modestly decreases fasting homocysteine and LDL-cholesterol in adults. *J. Nutr.* **140**, 2153–2157.

Regvar, M., Eichert, D., Kaulich, B., Gianoncelli, A., Pongrac, P., Vogel-Mikus, K. and Kreft, I. (2011) New insights into globoids of protein storage vacuoles in wheat aleurone using synchrotron soft X-ray microscopy. *J. Exp. Bot.* **62**, 3929–3939.

Shen, B., Li, C., Min, Z., Meeley, R.B., Tarczynski, M.C. and Olsen, O.A. (2003) sal1 determines the number of aleurone cell layers in maize endosperm and encodes a class E vacuolar sorting protein. *Proc. Natl. Acad. Sci. USA*, **100**, 6552–6557.

Stamova, B.S., Laudencia-Chingcuanco, D. and Beckles, D.M. (2009) Transcriptomic analysis of starch biosynthesis in the developing grain of hexaploid wheat. *Int. J. Plant Genomics*, **2009**, 407426.

Stein, K., Borowicki, A., Scharlau, D. and Glei, M. (2010) Fermented wheat aleurone induces enzymes involved in detoxification of carcinogens and in antioxidative defence in human colon cells. *Br. J. Nutr.* **104**, 1101–1111.

Tovar-Mendez, A., Matamoros, M.A., Bustos-Sanmamed, P., Dietz, K.J., Cejudo, F.J., Rouhier, N., Sato, S., Tabata, S. and Becana, M. (2011) Peroxiredoxins and NADPH-dependent thioredoxin systems in the model legume Lotus japonicus. *Plant Physiol.* **156**, 1535–1547.

Usadel, B., Nagel, A., Thimm, O., Redestig, H., Blaesing, O.E., Palacios-Rojas, N., Selbig, J., Hannemann, J., Piques, M.C., Steinhauser, D., Scheible, W.R., Gibon, Y., Morcuende, R., Weicht, D., Meyer, S. and Stitt, M. (2005) Extension of the visualization tool MapMan to allow statistical analysis of arrays, display of corresponding genes, and comparison with known responses. *Plant Physiol.* **138**, 1195–1204.

Usadel, B., Poree, F., Nagel, A., Lohse, M., Czedik-Eysenberg, A. and Stitt, M. (2009) A guide to using MapMan to visualize and compare Omics data in plants: a case study in the crop species, Maize. *Plant Cell Environ.* **32**, 1211–1229.

Wan, Y., Poole, R.L., Huttly, A.K., Toscano-Underwood, C., Feeney, K., Welham, S., Gooding, M.J., Mills, C., Edwards, K.J., Shewry, P.R. and Mitchell, R.A. (2008) Transcriptome analysis of grain development in hexaploid wheat. *BMC Genomics*, **9**, 121.

Wan, Y., Underwood, C., Toole, G., Skeggs, P., Zhu, T., Leverington, M., Griffiths, S., Wheeler, T., Gooding, M., Poole, R., Edwards, K.J., Gezan, S., Welham, S., Snape, J., Mills, E.N., Mitchell, R.A. and Shewry, P.R. (2009) A novel transcriptomic approach to identify candidate genes for grain quality traits in wheat. *Plant Biotechnol. J.* **7**, 401–410.

Wegel, E., Pilling, E., Calder, G., Drea, S., Doonan, J., Dolan, L. and Shaw, P. (2005) Three-dimensional modelling of wheat endosperm development. *New Phytol.* **168**, 253–262.

Young, T.E. and Gallie, D.R. (2000) Programmed cell death during endosperm development. *Plant Mol. Biol.* **44**, 283–301.

Young, T.E., Meeley, R.B. and Gallie, D.R. (2004) ACC synthase expression regulates leaf performance and drought tolerance in maize. *Plant J.* **40**, 813–825.

Zdobnov, E.M. and Apweiler, R. (2001) InterProScan – an integration platform for the signature-recognition methods in InterPro. *Bioinformatics*, **17**, 847–848.

Supporting information

Additional Supporting information may be found in the online version of this article:

Figure S1 Level 4 Gene Ontology analysis of all differentially expressed transcripts at each time point (Total) in comparison with transcripts DE at one time point (Only) and those common to all time points (All). A, Biological process terms; B, molecular function terms; C, cellular component terms. Al, aleurone; En, starchy endosperm.

Dataset S1 NCBI/GEO Submission. Raw Illumina sequencing data files and RNA-Seq analysis files are provided.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

© 2012 The Authors