Gene expression profile of the developing endosperm in durum wheat provides insight into starch biosynthesis

by Chen, J., Watson-Lazowski, A., Uttam Kamble, N., Vickers, M. and Seung, D.

Harper Adams University

RESEARCH

Gene expression profile of the developing endosperm in durum wheat provides insight into starch biosynthesis

Jiawen Chen[1†], Alexander Watson-Lazowski[1,2†], Nitin Uttam Kamble[1], Martin Vickers[1] and David Seung[1*]

Abstract

Background Durum wheat (*Triticum turgidum* subsp. *durum*) is widely grown for pasta production, and more recently, is gaining additional interest due to its resilience to warm, dry climates and its use as an experimental model for wheat research. Like in bread wheat, the starch and protein accumulated in the endosperm during grain development are the primary contributors to the calorific value of durum grains.

Results To enable further research into endosperm development and storage reserve synthesis, we generated a high-quality transcriptomics dataset from developing endosperms of variety Kronos, to complement the extensive mutant resources available for this variety. Endosperms were dissected from grains harvested at eight timepoints during grain development (6 to 30 days post anthesis (dpa)), then RNA sequencing was used to profile the transcriptome at each stage. The largest changes in gene expression profile were observed between the earlier timepoints, prior to 15 dpa. We detected a total of 29,925 genes that were significantly differentially expressed between at least two timepoints, and clustering analysis revealed nine distinct expression patterns. We demonstrate the potential of our dataset to provide new insights into key processes that occur during endosperm development, using starch metabolism as an example.

Conclusion We provide a valuable resource for studying endosperm development in this increasingly important crop species.

Keywords Durum, Endosperm, Gene expression, Grain development, RNAseq, Starch, Wheat

Background

The endosperm constitutes about 80% of a wheat grain by weight, and is rich in storage reserves that are mobilised during seed germination to nourish the developing embryo [1]. Starch is the main storage carbohydrate in the endosperm, providing much of the calories in wheat-based foods; and storage protein (particularly gluten) is not only important for nutrition, but also for the functional properties of flour for bread and pasta-making [2, 3]. It is therefore important to understand the gene regulation that orchestrates the development of the endosperm and the synthesis of storage reserves.

†Jiawen Chen and Alexander Watson-Lazowski are co-first authors.

*Correspondence:
David Seung
David.Seung@jic.ac.uk
[1]John Innes Centre, Norwich Research Park, Norwich NR4 7UH, UK
[2]Harper Adams University, Newport, Shropshire TF10 8NB, UK

Chen *et al. BMC Plant Biology*   (2023) 23:363

Page 2 of 13

The endosperm originates from double fertilisation, where a pollen nucleus fuses with the two polar nuclei of the central cell [2, 4]. The triploid nucleus then undergoes numerous mitotic divisions without cell wall formation, forming a multinucleate syncytium, which then becomes cellularised into alveoli by approx. 4 days post anthesis (dpa) [2]. The alveoli cells then differentiate into both aleurone and starchy endosperm cells, and starch synthesis subsequently begins in the latter [5]. The basic structure of the starchy endosperm is formed via further cell divisions and expansion by approx. 10 dpa [2]. It is only after this stage that most of the starch and protein is synthesised: for example, in bread wheat grown in the United Kingdom, grain filling is most active between 14 and 28 dpa [2]. Grain filling is followed by grain maturation and desiccation.

Starch is synthesised in amyloplasts of starchy endosperm cells from sucrose imported into the developing grain from source tissue. It exists as semi-crystalline, insoluble starch granules that are composed of two distinct glucose polymers - amylopectin and amylose. However, wheat (and other species of the Triticeae) is unique among cereals in that it has a "bimodal" distribution of starch granules, where two types of starch granules arise from a distinct temporal pattern of granule initiation [6]. A-type granules are large (18–20 μm in diameter), are flattened in shape, and they start to initiate at approx. 4 days post anthesis (dpa) and reach their final number by approx. 8 dpa. B-type granules are small (<10 μm in diameter) and near spherical, and appear at approx. 15–20 dpa [5–7]. The size distribution of A- and B-type granules in wheat has implications on both pasta and breadmaking quality [8–10], and on milling and processing [11–13]. However, we are only beginning to understand the temporal control that coordinates the formation of these two granule types.

While many studies have focused on grain development in hexaploid bread wheat (*Triticum aestivum*), relatively fewer resources exist for tetraploid durum wheat (*Triticum turgidum* ssp. *durum*). Durum wheat is widely grown in semi-arid climates, primarily for pasta production. Worldwide, it represents ≈5% of wheat grown, but it is still the 10th most widely cultivated cereal crop [14]. It is also gaining interest due to its potential for yield and quality improvement [15, 16], its adaptability to warm and dry climates [16–18], and the effects of climate change creating new suitable regions for growing the crop [17]. Further, its tetraploid genome and shorter life cycle compared to bread wheat makes it an ideal experimental model for wheat research. The wheat *in silico* TILLING mutant collection in variety Kronos allows easy isolation of mutants for functional genomics, and mutations in only two homeologs need to be combined for most genes [19, 20]. Durum wheat has now been fully

sequenced with a high-quality reference genome [21]. To enable further research into important processes occurring during endosperm development in durum wheat, we generated a high-quality transcriptomics dataset using RNA sequencing of the developing endosperm of variety Kronos, comparable to similar resources available for bread wheat [22]. We demonstrate the potential of our dataset to reveal new insights into grain development, using the process of starch biosynthesis as an example.

Results and Discussion
Distinct patterns of gene expression observed during grain development

To explore gene expression patterns that underpin key processes of endosperm development in durum wheat (*Triticum turgidum* ssp. *durum*), we generated a transcriptomics dataset from endosperm tissue extracted throughout grain development, in variety Kronos. Developing grains were harvested from plants grown in controlled environment chambers - starting from the earliest timepoint in which the starchy endosperm could be reliably dissected from the pericarp (6 dpa) and until the grain turned yellow, indicating maturation (30 dpa) (Fig. 1A).

We measured the starch content of these grains throughout endosperm development, which follows a sigmoidal pattern (Fig. 1B). A logistic growth model fit the data, and this model indicates that the starch accumulation rate is highest at 13 dpa. This is consistent with the onset of the grain filling period at 14 dpa in bread wheat grown in the United Kingdom [2].

RNA was extracted from endosperm tissue dissected from the grains (n=3), with each replicate using multiple grains (3–6 depending on the timepoint) harvested from an independent plant. RNA sequencing (RNA-Seq) using Illumina technology yielded between 78,762,223 and 169,888,168 reads per sample. These data were pseudoaligned using *kallisto* [23] to the annotated transcripts from the Svevo v1 durum wheat transcriptome (using Ensembl canonical transcripts only) [21]. A range of 66.356–77.040% of reads per sample were successfully pseudoaligned. We then calculated transcripts per million (TPM) values for each detected transcript (Supplemental File 1). Out of a total of 66,559 annotated canonical transcripts, we detected an average expression over one TPM (for at least one timepoint) for 31,879 transcripts.

We also calculated TPM values for all annotated transcripts in the Svevo v1 transcriptome, rather than using the Ensembl canonical transcripts only, as a resource for future analysis (Supplemental File 2). These data were not used in the analysis described in the remainder of this paper.

We used a principal component analysis (PCA) to visualise and identify samples that had similar gene expression patterns (Fig. 1C). Replicates for each of the early timepoints (6, 8, 10 and 13 dpa) formed distinct clusters that were well separated from each other. However, the later timepoints (15, 18, 20 dpa) appeared to largely cluster together with some overlap between replicates being apparent. This suggests additional factors may contribute to the variation over these timepoints, whereas temporal variation is limited. The latest timepoint (30 dpa) again separated out into a distinct cluster. It is plausible that the largest changes in gene expression are observed during early grain development when the basic endosperm structure is forming [2], and that the expression pattern stabilises as the grain as undergoes filling. The distinct shift in expression pattern at 30 dpa is likely due to processes such as maturation and senescence being initiated.

Significantly differentially expressed (DE) genes (DEGs) were calculated using DESeq2 (FDR≤0.05 and $\log_2$fold change>1.5) for all pairwise comparisons of the sampled timepoints (Supplemental File 3). This resulted in a total of 29,925 DEGs. When comparing early timepoints (6, 8,

10 and 13 dpa) sequentially, a large number of DEGs were identified (between 2,185 and 6,407 per comparison). However, when comparing within the middle timepoints (15 and 18 dpa) only 34 DEGs were identified. Comparing the last two timepoints (20 and 30 dpa) again showed high numbers of DEGs (5,250). These numbers are consistent with the PCA (Fig. 1A) in suggesting that most changes in the endosperm transcriptome happen during the formation of the endosperm rather than during grain filling, with an additional spike in the number of DEGs as maturation and senescence initiates. All DEGs (genes which were identified as DE in at least one comparison) were then clustered by their TPM values over endosperm development using *Clust* [24], and 14,102 of the DEGs were found to cluster into nine distinct gene expression patterns (Fig. 2; Supplemental File 4). Those DEGs that did not cluster did not meet the cut-off criteria of *Clust*, either due to not meeting the minimum TPM, minimum cluster size, or due to having an expression pattern that was perceived by *Clust* as being too flat over time. These nine clusters (cluster C0-C8) will be referred to as the 'global clusters' hereafter. Several global
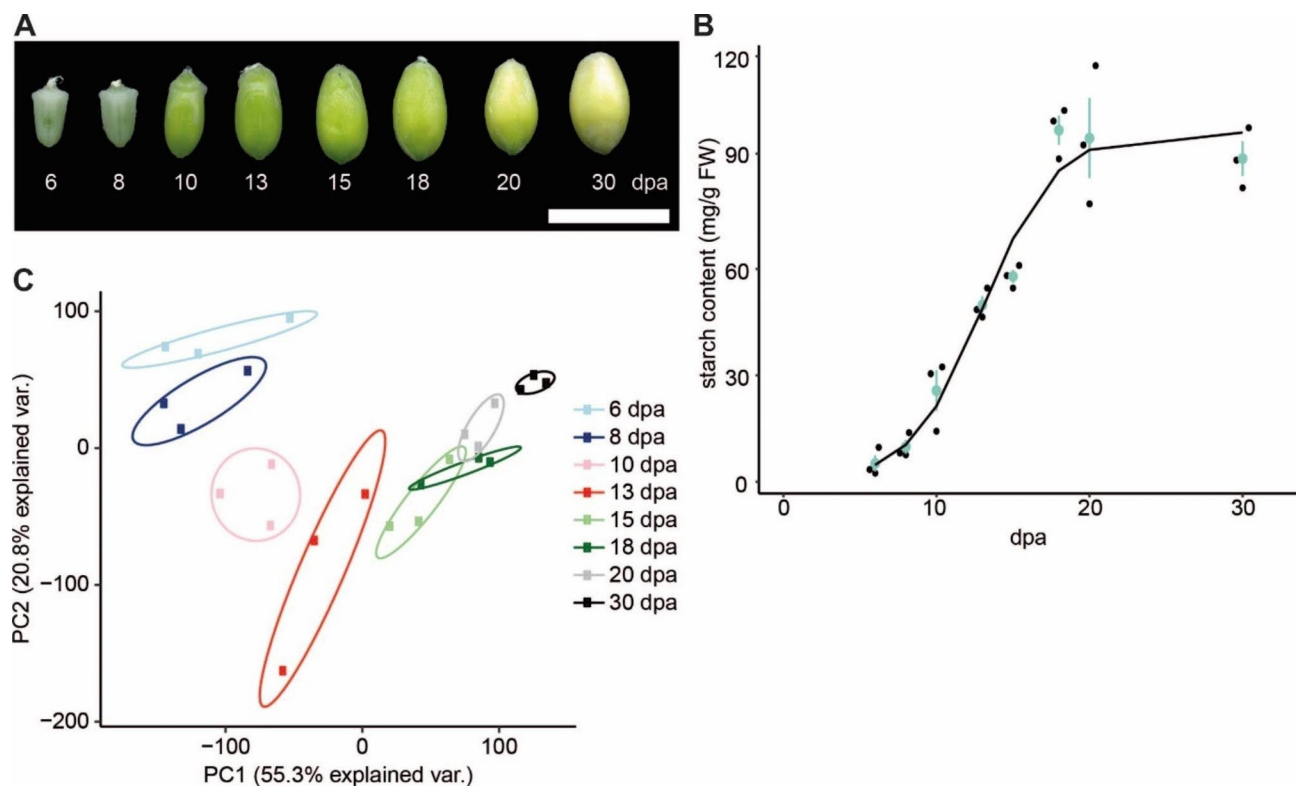

Fig. 1 Distinct transcriptome profiles in the endosperm through grain development. **(A)** Morphology of grains harvested at each timepoint. Bar = 1 cm, dpa = days post anthesis. **(B)** Total starch content of developing endosperm in grains as shown in A, in units of mg/g fresh weight (FW) of grain material. The black dots are biological replicates and the blue dots are the mean of three biological replicates per time point, with error bars representing standard error of the mean. The black line is a fitted logistic growth curve based on all the biological replicates. **(C)** Principal component analysis (PCA) of calculated TPM values. Genes included had a TPM over one for all replicates within our study (15,826 genes included in the analysis). Principal components 1 and 2 (PC1 and PC2) could represent 55.3% and 20.8% of the variation, respectively. Ellipses enclose the three replicates for each timepoint for visualising the spread
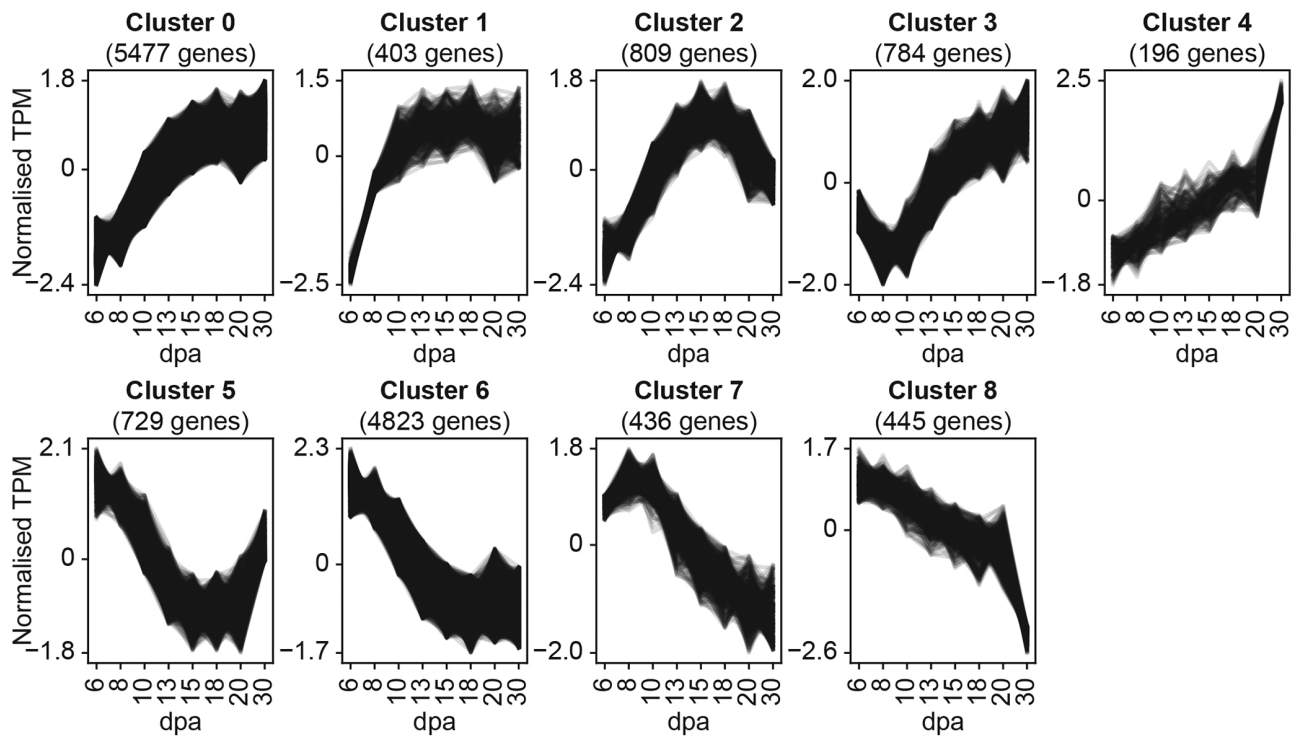
Fig. 2 Identification of distinct expression profiles during endosperm development. Out of 29,925 significantly differentially expressed genes (FDR ≤ 0.05 and $\log_2$ fold change > 1.5), 14,102 clustered into nine distinct expression patterns over endosperm development using *Clust* (tightness value of 5). TPM values were normalised using inbuilt normalisation methods within *Clust*. The number of genes corresponds to the number of genes identified within that cluster. Clustering was performed over endosperm development (days post anthesis (dpa))

clusters represented general up- or down-regulation over endosperm development (C0, C1 and C6). However, the remaining global clusters were more dynamic and may be associated with processes that act during a specific period of development. Most of the genes were grouped into cluster 0 (5477 genes), which follows an increasing pattern similar to the starch accumulation curve (Fig. 1B).

Analysis of gene expression patterns related to starch metabolism

We analysed the expression patterns of known genes related to starch metabolism, as an example of how a specific process can be studied using our dataset. These genes were subcategorised by their role in starch metabolism: ADP-glucose metabolism, amylopectin synthesis, amylose synthesis, starch phosphorylation, starch degradation, maltooligosaccharide (MOS) metabolism, plastidial sugar transporters and starch granule initiation (Supplemental File 5). Starting from all detected genes (including those that were removed for the DEG/clustering analysis above), we applied a minimum expression filter to remove genes that had very low expression, keeping only those that had an average expression (among the three replicates) of ≥1 TPM for at least one timepoint (Supplemental File 5). Expression patterns for genes

within each category were visualised through heatmaps, and hierarchal clustering (HC) was performed to identify genes with similar expression levels and patterns (Fig. 3). Almost all starch-related genes that remained after the minimum expression filter had been applied also met the criteria of being a DEG, indicating that they have (to some degree) dynamic expression patterns over development. The implications of this dataset on each category is discussed below.

We also checked the expression pattern of selected starch metabolism genes using qRT-PCR, which showed that the patterns were consistent with our RNAseq results (Supplemental Fig. 1). We chose genes from different categories as specified above, including ones which were previously used for qRT-PCR validation of RNA-seq data exploring starch metabolism in hexaploid wheat [22]. Our data is overall also consistent with the RNA-seq data of hexaploid wheat [22], except for BAM5.3-A1 and BGC1-A1 which show an increase then decrease in expression in hexaploid wheat, whereas our dataset shows consistent increase in expression towards 30 dpa in both the qRT-PCR and RNA-seq. This could be due to biological differences between durum wheat and bread wheat, differences in growth conditions between the two studies, and/or technical factors (such as differences in sequencing depth).
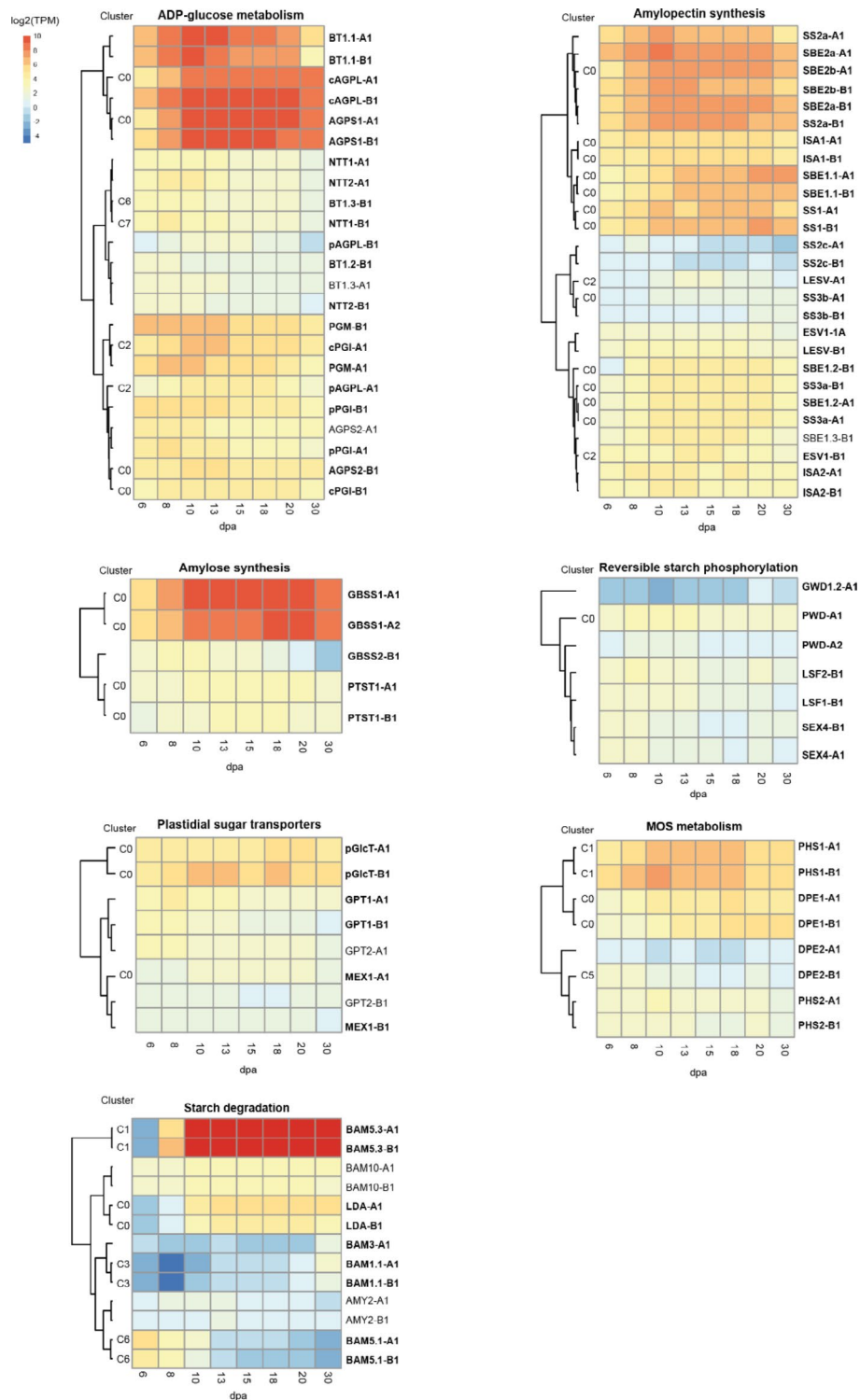
Fig. 3 Expression patterns of genes involved in starch metabolism. Average expression over endosperm development and hierarchal clustering of genes within seven categories related to starch metabolism. A total of 91 genes were visualised after filtering out genes with low expression (keeping only genes with average expression of ≥1 TPM for at least one timepoint). Average log$_2$ TPM expression values (n = 3) were used to illustrate expression, with the same scale used for all categories. Genes in bold were identified as DEGs, and those that were grouped into a global cluster in Fig. 2 are annotated with their cluster number (C0-C8). Details of all starch metabolism genes, including average TPM values, can be found in Supplemental File 5. TPM = transcripts per million, dpa = days post anthesis

Chen *et al. BMC Plant Biology*   (2023) 23:363

Page 6 of 13

ADP-glucose metabolism ADP-glucose is the substrate for starch synthesis as it serves as the glucosyl donor for starch synthase-mediated polymer extension. In cereal endosperm, the final step in the conversion of imported sucrose to ADP-glucose takes place in the cytosol and is catalysed by ADP-Glucose Pyrophosphorylase (AGPase) [25, 26]. The ADP-glucose is subsequently imported into the amyloplast by the Brittle1 (BT1) transporter [27]. AGPase is a multimeric enzyme consisting of small (AGPS) and large (AGPL) subunits. AGPS1, which gives rise to the cytosolic small subunit [25, 28], and cAGPL encoding the cytosolic large subunit, are the major subunits expressed in the endosperm – and their expression patterns clustered together with BT1.1. The co-expression of these genes is consistent with these components working together to synthesise and import ADP-glucose into the plastid. The genes encoding plastidial AGPase subunits (AGPS2 and pAGPL) were only lowly expressed (relative to AGPS1 and cAGPL) – suggesting that they do not play a major role in the endosperm. Similarly, other BT1 homologs (BT1.2 and BT1.3), as well as nucleoside triphosphate transporters (NTTs), had low expression compared to BT1.1. The import of ATP into plastids through the NTTs is important for plastidial ADP-glucose synthesis in potato tubers [29] and Arabidopsis guard cells [30], and their low expression in the durum wheat endosperm may reflect the cytosolic synthesis of ADP-glucose in this tissue.

We also observed that durum wheat has only one set of homeologous phosphoglucomutase (PGM) genes, similar to what was shown for hexaploid wheat [22]. Both homeologs were expressed in the endosperm, most strongly during early grain development. We ran a phylogenetic tree analysis to better understand the origin of wheat PGM. Closely related barley also had a single PGM gene, which together with the wheat PGMs grouped into the same clade as the cytosolic PGMs of Arabidopsis (PGM2 and PGM3), rather than the clade with the plastidial PGM1 (Supplemental Fig. 2). However, both wheat PGM homeologs and most cereal PGM sequences within the PGM2/PGM3 clade were predicted to be plastid localised by TargetP [31]. In the developing bread wheat endosperm, there is measurable PGM activity in both the cytosol and amyloplasts [32], but bread wheat PGM localised to the cytosol of wheat protoplasts [22], and it remains to be determined how a single PGM paralog can give rise to both plastidial and cytosolic activities.

Amylopectin synthesis Amylopectin is the primary constituent of starch and is a highly branched polymer with α-1,4-linked glucose chains and frequent α-1,6-linked branches. Its synthesis requires coordinated action of Starch Synthases (SS) that elongate the polymer chains, Starch Branching Enzymes (SBE) that introduce

the branch points, and Debranching enzymes (primarily Isoamylases - ISA) that ensure correct branch pattern formation [33]. Most DEGs within this category that were assigned to a global cluster appeared in C0. Under HC, genes with relatively high expression clustered together – where one branch contained homeologs of SS2a and of the two SBE2 paralogs (SBE2a and SBE2b), which peak at 10 dpa and remain highly expressed until 20 dpa. Their co-expression is interesting because in wheat and maize, these two proteins form a multienzyme complex [34, 35]. The other branch was formed by SS1, SBE1.1 and ISA1 homeologs. SBE1.1 expression levels peaked later than SBE2 (at 20–30 dpa). This is consistent with previous observations in bread wheat endosperm, where SBE2 protein was already visible at 13 dpa and remained present at a similar abundance until 34 dpa, whereas SBE1 was only detectable from 18 dpa and increased in relative abundance towards later grain development [36]. Interestingly, ISA1 and ISA2 did not cluster together under global clustering or HC. In Arabidopsis leaves, ISA1 and ISA2 act exclusively in a heteromeric form [37], but in rice and maize endosperm, the major ISA activity is contributed by ISA1 homomers [38, 39]. The distinct expression patterns between ISA1 and ISA2 may reflect a similar mechanism in durum wheat endosperm. The paralogs of SS2 (SS2b and SS2c) and BE1 (BE1.2 and BE1.3) were detectable but were not as highly expressed as SS2a and BE1.1 in the endosperm. SS3a had higher expression levels than SS3b, but they both had relatively low expression levels compared with SS1 and SS2a.

Homeologs of Early Starvation1 (ESV1) and Like Early Starvation (LESV), ESV1-B and LESV-A, were unique within this category in that they clustered in C2, peaking in expression during mid-grain development (13–15 dpa). ESV1 and LESV were discovered in Arabidopsis and are hypothesised to organise amylopectin within the starch granule [40]. Their expression levels in the durum wheat endosperm were relatively weak compared to the core amylopectin biosynthesis enzymes. However, the recent implication of ESV1 in starch synthesis in the rice endosperm warrants further investigation of these genes in wheat [41].

Amylose synthesis Amylose is the minor polymer of starch and consists of α-1,4-linked linear chains with very few branches, and is synthesised by a specialised SS isoform, the Granule Bound Starch Synthase (GBSS) [42]. Wheat starch contains approx. 25–29% amylose [43]. All cereals have two isoforms of GBSS – GBSS1 which is mainly expressed in the endosperm and pollen grains, and GBSS2 which is expressed in vegetative organs and the pericarp [44]. In our dataset, we only observed strong expression for GBSS1, which clustered into C0, plateauing in expression towards 18–20 dpa. The very low levels

of GBSS2 expression demonstrate that our RNA samples were not substantially contaminated with the surrounding pericarp tissue. Interestingly, we also detected relatively low levels of Protein Targeting to Starch 1 (PTST1) expression. PTST1 in Arabidopsis leaves acts together with GBSS in amylose biosynthesis by facilitating GBSS localisation on starch granules, and *ptst1* mutants produce amylose-free starch [45]. However, in rice endosperm, eliminating PTST1 has only minor reductions on amylose content [46]. The low expression levels of PTST1 in durum wheat makes it unlikely to play a major role in amylose synthesis in the durum wheat endosperm.

Reversible starch phosphorylation Cereal starch is almost devoid of covalently bound phosphate [47]. Consistent with this, we observed very low expression levels of the dikinases that phosphorylate starch (GWD and PWD), and the glucan phosphatases that remove phosphates (SEX4 and LSF2). The non-active phosphatase homolog, LSF1, that targets β-amylases to starch granules in Arabidopsis [48], also had very low expression.

MOS metabolism Two plastidial enzymes involved in processing MOS were expressed in the endosperm: α-glucan phosphorylase (PHS1) and disproportionating enzyme (DPE1) - with PHS1 peaking at around 10 dpa (and appearing in C1) while DPE1 peaked later (18–20 dpa; appearing in C0). PHS1 activity has been observed in the developing wheat endosperm, where it was proposed to participate in starch synthesis either by trimming glucan chains at the granule surface, or by degrading MOS released by ISA during starch synthesis [49]. It is also possible that PHS1 participates in granule initiation. In Arabidopsis, PHS1 is not required for normal starch synthesis or degradation [50], but *phs1* mutations can affect starch granule numbers per chloroplast if other components of MOS metabolism are simultaneously abolished [51]. In rice, the grains of *phs1* knockout mutants are severely shrivelled, particularly when grown at low temperature [52]. The enzyme's role in granule initiation remains unclear due to the conditional nature of these phenotypes (upon additional mutations and temperature) and should be further investigated in wheat. DPE1 in Arabidopsis leaves participates in starch degradation, where it processes MOS to allow complete degradation to maltose [53]. Since there is very little starch degradation occurring in the developing endosperm, it is not clear what this enzyme does in this tissue. However, there is evidence in vitro that the rice PHS1 and DPE1 can physically interact and elongate MOS simultaneously [54].

By contrast to the plastidial PHS1 and DPE1, their cytosolic counterparts, DPE2 and PHS2, showed very low expression. It is plausible that the activity of these enzymes is not important in the developing endosperm, given that cytosolic MOS processing is typically associated with the further metabolism of maltose generated during starch degradation in leaves [55, 56].

Plastidial sugar transporters We observed very low expression of the plastidial maltose exporter, MEX1, consistent with there being low starch degradation to generate maltose for cytosolic processing. Interestingly, we observed substantial expression of the glucose transporter pGlcT. In Arabidopsis leaves, pGlcT works together with MEX1 to export products of starch degradation [57]. Its role beyond Arabidopsis leaves remains to be determined. Also, we observed generally low expression for all glucose-6-phosphate (Glc6P)/phosphate translocators (GPT), which are important in heterotrophic or part-heterotrophic tissues such as the embryo and guard cells of Arabidopsis [58, 59] and rice pollen [60], where they import glucose-6-phosphate (Glc-6P) into the plastid to support metabolism, including ADP-glucose synthesis. Their low expression levels are aligned with the cytosolic synthesis of ADP-Glucose in wheat endosperm.

Starch degradation The low expression of most starch-degrading enzymes is consistent with the endosperm being primarily a site of starch synthesis. The major plastidial α-amylase and isoamylase activities, AMY3 and ISA3, were filtered out due to very low expression at all timepoints (Supplemental File 5). Expression of most β-amylase isoforms was generally low, except the two homeologs of BAM5.3 (encoded on chromosomes 5A and 4B), which had extremely high levels of expression during grain development from 10 dpa onwards (and appeared in C1). Interestingly, BAM5.3 is an ortholog of Arabidopsis BAM5 and BAM6, and is localised to the cytosol [61]. The role of β-amylases is not understood in cereal grains as even during grain germination, starch degradation is primarily carried out by the AMY1 family of α-amylases [62, 63]. It has even been postulated that the β-amylase accumulated during grain filling may act as a storage protein [64]. BAM5.3 appears to be the primary isoform accumulated during grain development in durum wheat, and its function in the cytosol deserves further investigation. Future studies should check whether this protein is enzymatically active. Notably, we also observed substantial expression of limit dextrinase (LDA), which in barley has been proposed to be involved starch granule initiation, since a mutant in LDA has altered distribution of A- and B-type granules [65]. It is possible that it may have a similar role in starch biosynthesis in wheat.

Although the AMY1 family of α-amylases is widely known to be involved in starch degradation during germination, we examined the expression patterns of AMY1 homologs in the developing durum wheat endosperm. This family is characterised by having an N-terminal

signal peptide, which allows enzymes synthesised *de novo* to be secreted into the endosperm from the aleurone during germination [63]. We identified 15 AMY1 genes in durum wheat. Surprisingly, one homolog of AMY1 (encoded by homeologs TRITD5Av1G228350 and TRITD5Bv1G227200) was strongly expressed in the endosperm during early grain development, peaking at 8 dpa, before decreasing to almost undetectable levels after 15 dpa (Fig. 4A). The role of this secreted α-amylase in the endosperm during grain development is yet to be investigated. The other AMY1 homologs, as expected, had very low or almost no expression in the developing endosperm (Fig. 4B).

It is possible that the starch degradation enzymes that are present in developing endosperm (BAM5.3, the AMY1 isoform and LDA) may serve indirect functions in starch synthesis, for example by supplying MOS substrates to starch biosynthesis enzymes (see MOS metabolism above). One hypothesis is that they could release glucans from A-type granules, providing substrates for the initiation of B-type granules.

Granule initiation A distinct set of proteins involved in starch granule initiation in Arabidopsis has recently been discovered. However, research investigating their roles in mediating the distinct spatiotemporal pattern of A- and B-type granule initiation in wheat is still on-going [66, 67]. The genes encoding these proteins showed distinct expression patterns that are consistent with their proposed roles in A- and B-type granule initiation (Fig. 5A). Starch Synthase 4 (SS4) was primarily expressed during early grain development, consistent with a role in establishing correct granule number per amyloplast in early grain development [68]. B-Granule Content1 (BGC1) not only limits granule initiations to a single A-type granule per amyloplast during early grain development, but it also has a critical role in promoting B-type granule initia-

tion during later grain development [6]. Consistent with this, BGC1 expression increases during grain development, and is highly expressed at 15–20 dpa when B-type granules initiate. For both SS4 and BGC1, the patterns of transcript accumulation are consistent with patterns of protein accumulation observed through immunoblotting [68]. Myosin Resembling Chloroplast Protein (MRC) was expressed primarily during early grain development, showing essentially an inverse expression pattern to BGC1. This distinct expression pattern appeared in C6, one of the few starch-related genes to do so (alongside BAM5.1). We demonstrated that MRC is a repressor of B-type granule initiation specifically during early grain development, as mutants of durum wheat defective in MRC show the early initiation of B-type granules [69].

So far, the Mar1 Filament-like Protein (MFP1) is the only known granule initiation protein identified in Arabidopsis that has not been explored in wheat. Interestingly, all cereals have two paralogs of MFP1 (MFP1.1 and MFP1.2) [70], which were both expressed in the durum endosperm (Fig. 5A). Given that MFP1 is a thylakoid-associated protein in Arabidopsis leaves [70], the expression of this protein in the endosperm (where amyloplasts lack true thylakoid membranes) is noteworthy, and its localisation and role in the endosperm should be investigated.

When comparing the expression pattern of granule initiation genes from our study with that of hexaploid bread wheat *(Triticum aestivum)* [22], we consistently observed the opposite expression pattern of *Ta*BGC1 compared to the other granule initiation genes (Fig. 5B). However, in the hexaploid, *Ta*BGC1 is still expressed highly at 15–20 dpa, but it then decreases in expression towards 30 dpa; whereas it seems to plateau in our durum wheat dataset. *Ta*SS4 still has an inverse expression pattern compared to *Ta*BGC1, dipping in expression between 20 and 30 dpa and then increasing again. It is interesting that the
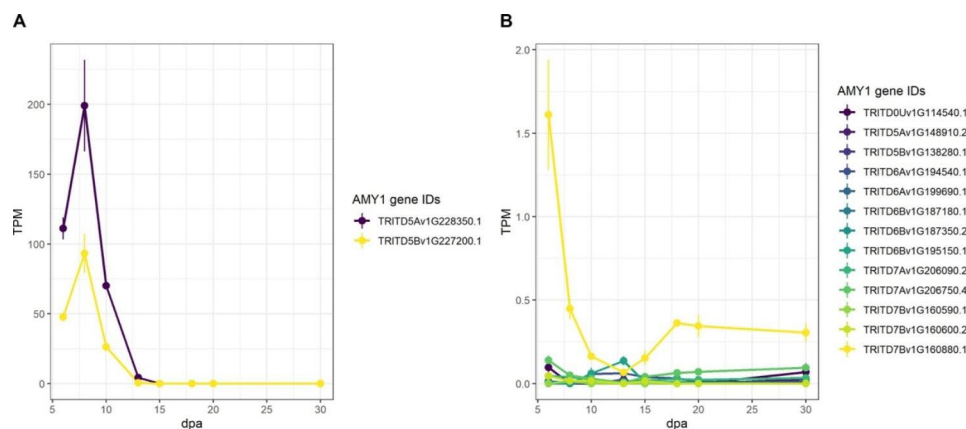

Fig. 4 Expression patterns of AMY1 genes in developing wheat endosperm. Transcript accession numbers are the canonical transcripts from Ensembl Plants. **(A)** Average TPM values of the two homeologs with high expression in the endosperm. Values are the mean ± SEM of the *n* = 3 replicates. **(B)** As for A, but for AMY1 genes with low expression in the endosperm

Chen *et al. BMC Plant Biology*   (2023) 23:363

Page 9 of 13

expression patterns of these genes in hexaploid wheat look different to those in durum wheat, which could be due to biological differences between the species, or in the experimental set-up and sequencing depths. These discrepancies highlight the importance of our dataset as an additional resource for the study of durum wheat.

To assess how specific these expression patterns are to species that have A- and B-type granules, we examined the expression patterns of these genes in the endosperm of rice and maize using publicly available data. Rice produces compound granules, where granules are initiated during early grain development and there is no second wave of granule initiation in later grain development that leads to B-type granules; and maize produces simple granules without distinct populations of different granule sizes [67]. Among the granule initiation genes, BGC1 had the most different expression patterns when comparing durum wheat with rice and maize. Rather than peaking during mid-late grain development, BGC1/ Floury Endosperm 6 (FLO6) expression in rice was highest at the beginning of endosperm development and then decreased, increasing again only slightly in late grain development (after 20 dpa) (Fig. 5C). PTST2b expression (PTST2 is the ortholog of BGC1) in maize was also high during early grain development, dipping and then plateauing in later grain development (Fig. 5D). It is therefore tempting to speculate that change in the BGC1 expression pattern is one of the key factors that facilitates the formation of A- and B-type granules.

Future outlook

We have generated a high-quality transcriptomics dataset for endosperm development in durum wheat, and we demonstrated its potential to provide new insights and hypotheses into key processes during grain development, using starch metabolism as an example. Other processes to be investigated may include protein and lipid biosynthesis, cell wall biosynthesis, micronutrient transport, as well as development of endosperm structure and subsequent maturation in general.

Materials and methods
Plant growth and harvest

Wheat plants were grown in controlled environment rooms (CERs) at 60% relative humidity with 16 h light at 20 °C and 8 h dark at 16 °C. The CER light intensity was 400 µmol photons m^{-2} s^{-1}. Each spike was individually marked for anthesis when an anther was visible from the spike, and this was taken as day 0. Checking for anthesis and harvesting were performed at 1–2 h after midday. Four spikes from each individual plant were used, each for a different developmental time point. Three biological replicates per time point were collected from independent plants. Therefore, some of the biological replicates

for separate time points were spread over several of the same plants. We ensured that replicates for every time point were spread across different spikes in each plant, so that variability between spikes would not be biased towards one time point. Although there were slight differences in the day of flowering between spikes, many plants had spikes that flowered within a few days of each other and in our optimal growth conditions, the first 3–4 tillers look similar. Any variability between tillers and spikes is not likely be a huge variable for our experiment, as the quality control steps demonstrated that the principal difference separating the samples was their developmental stage (Fig. 1C). For sample collection, the whole spike was cut at the base, frozen in liquid nitrogen and stored at -80 °C. Then, grains were threshed from the spikes on dry ice, using the middle section of each spike, avoiding the middle grain of each spikelet, and stored at -80 °C again. Endosperms were subsequently dissected on dry ice, frozen in liquid nitrogen and stored at -80 °C until RNA extraction.

RNA extraction and sequencing

Total RNA was extracted (n=3) from the frozen endosperm tissue using the RNeasy PowerPlant kit (Qiagen). An on-column DNase digest was incorporated during the extraction. Either 6 (for 6, 8 and 10 dpa), 4 (for 13, 15, and 18 dpa) or 3 (for 20 and 30 dpa) endosperms were pooled from each spike for each RNA extraction. Quality control, poly-A selection library preparation and RNA sequencing (RNA-Seq) using a NovaSeq 6000 machine (Illumina) was carried out at Novogene (Cambridge, UK). Each sample returned between 78,762,223 and 169,888,168 million clean 150 bp paired end reads. These clean RNA-Seq reads can be downloaded from the NCBI Gene Expression Omnibus repository (Accession: GSE216253).

TPM value calculation and differential expression

Clean reads from above were pseudoaligned to the Svevo v1 Durum wheat transcriptome downloaded from Ensembl Plants, either using canonical transcripts only (Supplemental File 1) or all transcripts (Supplemental File 2), using *kallisto* [23]. The default settings were used for mapping. A range of 66.356–77.040% of reads per sample aligned to the transcriptome. Both estimated count values and normalised expression as transcript per million (TPM) for each gene were retrieved from the pseudoalignment. *DESeq2* [73], run though the *DEapp* shiny [74], was then used to identify significantly differentially expressed genes (DEGs; significance cut-offs used - FDR≤0.05, log$_2$ fold change>1.5 and a minimum count cut-off of 10 counts per million in at least three samples).
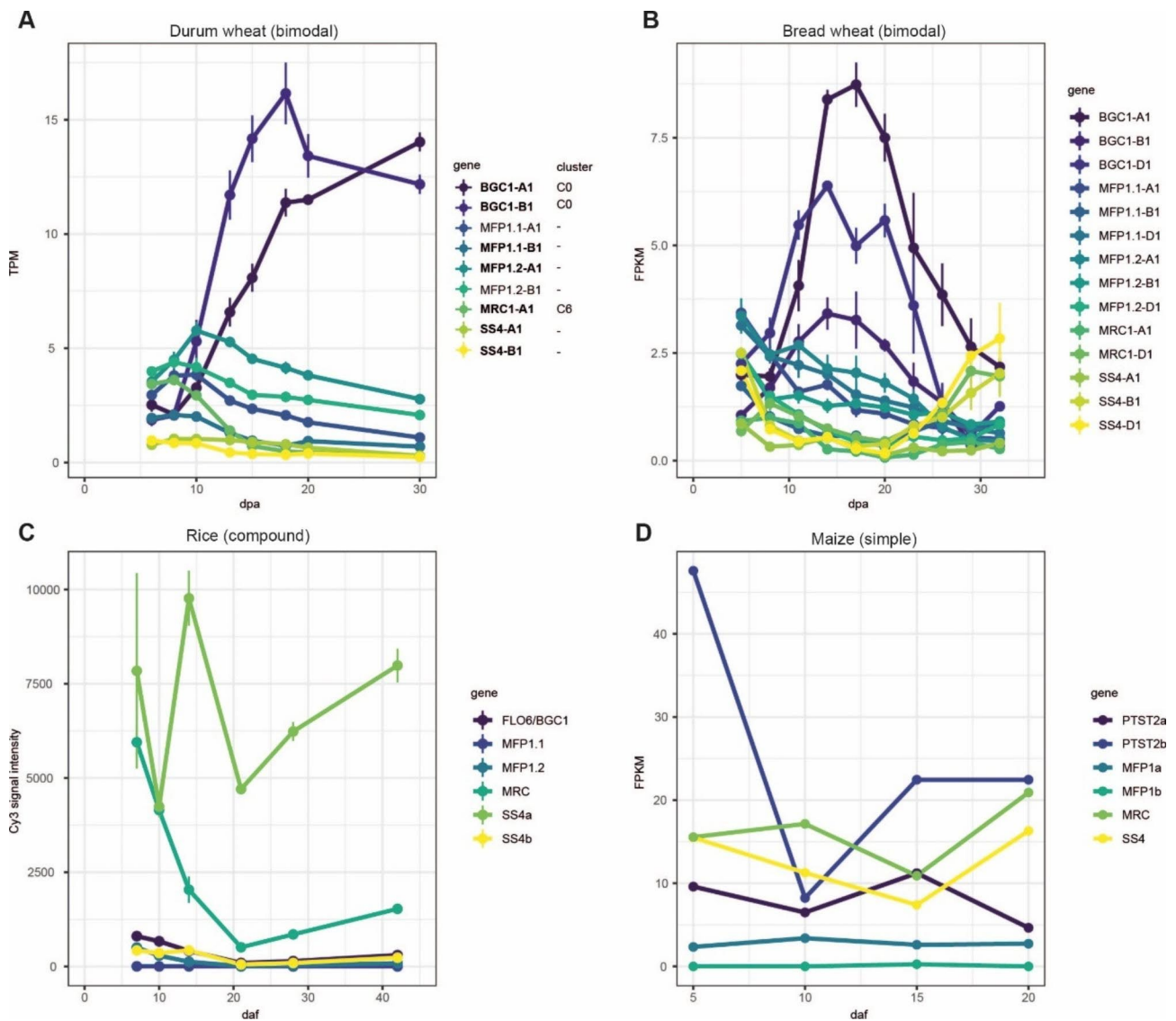
Fig. 5 Expression patterns of starch granule initiation genes in species with different endosperm starch granule morphology. **(A)** Average TPM values (n = 3) of starch granule initiation genes in durum wheat endosperm. Genes in bold were identified as DEGs, and numbered cluster annotations (beneath 'cluster') refer to the corresponding global cluster from Fig. 2 if the gene was included. **(B)** Average FPKM values (n = 3) of starch granule initiation genes in bread wheat endosperm, using data from Gu et al. 2021 [22]. **(C)** Average expression values (n = 3) from microarray analysis of starch granule initiation genes from rice developing endosperm, extracted from a public database (https://ricexpro.dna.affrc.go.jp/) [71]. FLO6/BGC1 = LOC_Os03g48170, MFP1.1 = LOC_Os05g08790, MFP1.2 = LOC_Os01g08510, MRC = LOC_Os02g09340, SS4a = LOC_Os01g52260, SS4b = LOC_Os05g45720 **(D)** Expression values (n = 3) from RNA sequencing analysis of starch granule initiation genes from maize developing endosperm, using data from Qu et al. 2016 [72]. PTST2a = GRMZM2G104501, PTST2b = GRMZM2G122274, MFP1a = GRMZM2G106233, MFP1b = GRMZM2G142413, MRC = GRMZM2G104357, SS4 = GRMZM2G044744. For A - C, error bars represent the standard error of the mean. For D, no data was available on the error values. dpa = days post anthesis. daf = days after flowering

Global clustering analysis

Clustering analyses were performed using *Clust* [24], on all of the DEGs identified. The DEGs were first filtered by setting a minimum threshold of 10 TPM for the sum of the TPM values of the three biological replicates, in at least one time point. The DEG TPM values (as separate biological replicates) were then used as the input for the analysis, which implemented the suggested in-built normalisation methods of *Clust* ($\log_2$, Z-score and quartile

normalisation) [24]. Parameters used included a tightness value of 5, a minimum cluster size > 10, and the removal of clusters with flat expression profiles over time. Several tightness values were tested, and the number of clusters identified plateaued at a tightness value of 5.

Principal component analysis and data visualisation

Principal component analysis (PCA) was carried out in R first using the 'prcomp' function from the 'stats' package

to generate the components of the PCA, then plotted using the 'ggbiplot' function from the 'ggbiplot' package. TPM values were used as the input for the PCA, with genes with a TPM over one for all replicates being included. For heatmap visualisation and hierarchal clustering, the R function 'pheatmap' from the 'pheatmap' package was used. Only the rows (genes) were clustered and not the columns (timepoints), otherwise default settings were used. Average $\log_2$ TPM values for the three replicates at each time point were used as the input for heatmap and hierarchal clustering analyses. Analyses were carried out in version 1.0.12 of the pheatmap package, running on R version 4 (https://cran.rstudio.com/web/packages/pheatmap/index.html).

Identification of durum wheat starch genes

Genes related to starch synthesis in wheat were manually curated: Orthologs in durum wheat were identified using BLASTp against the Svevo v1 Durum wheat proteome on Ensembl Plants [75], as well as using the "Orthologues" tool Ensembl Plants. In both cases, a known homolog from Arabidopsis, bread wheat, rice, or other species was used as a query sequence. The full list of genes analysed can be found in Supplemental File 5.

qRT-PCR

For qRT-PCR, we generated cDNA using the same RNA samples described above for RNA-seq, using the Go-Script Reverse Transcription System (Promega). The amount of RNA in each sample was measured using a Nanodrop, then 500 ng of each sample was used for cDNA synthesis.

Five starch metabolism genes (AGPS1, SBE2b, GBSS1, BAM5.3, BGC1) were selected for a qRT-PCR assay to check whether this was consistent with our RNAseq expression patterns. We chose four genes which were also selected in the hexaploid wheat RNAseq study [22]: AGPS1, SBE2b, GBSS1, and the reference gene ATG8d [76]. For AGPS1, SBE2b and ATG8d, we used primers reported in previous studies [22, 76] and these were not homeolog-specific. We designed new primers for GBSS1-A2, BAM5.3-A1 and BGC1-A1, which were homeolog-specific. For primers used, see Supplemental File 6.

We performed qRT-PCR using the SYBR Green Jumpstart Taq ReadyMix (Sigma-Aldrich), using 12 µL reactions. First, we performed primer efficiency tests using cDNA from two time points, 13 dpa and 18 dpa. This was done using 5 dilutions of the cDNA samples. Using the 13 dpa sample, the primer efficiencies were all between 95 and 105%, except for the BGC1-A1 primers (Supplemental File 6). This varied slightly using 18 dpa samples, indicating some difference in primer efficiency depending on the cDNA sample.

We performed the qRT-PCR for all time points using 12 µL reactions containing 2.5 ng of cDNA. Three technical replicates of three biological replicates per time point were used. The PCR programme was: preincubation at 95 °C for 5 min, amplification for 45 cycles at [95 °C for 10 s, annealing at 60 °C for 30 s, elongation at 72 °C for 15 s], melting curve generation at 95 °C for 5 s, 65 °C for 1 min and increasing to 97 °C at 0.11 °C/s, and cooling at 40 °C for 30 s. For data analysis, we took the average values of three technical replicates for each biological replicate. We used the ΔΔCt method to calculate relative expression, using the average value of three biological replicates at 6 dpa as reference sample and ATG8d as a reference gene.

Starch extraction and quantification

The same grains as used for RNA-seq were used for endosperm starch quantification, as described in [68]. Briefly, three to five dissected endosperms of known fresh weight were used for each sample. These were homogenised in 0.7 M perchloric acid. Insoluble material was collected by centrifugation, washed three times in 80% ethanol, then resuspended in water. Starch was digested using α-amylase/amyloglucosidase (Roche, Basel), and the released glucose was assayed using the hexokinase/glucose-6-phosphate dehydrogenase assay (Roche).

For curve fitting to the data, the nls() function with the SSlogis model in R (version 4.2.0) was used. This was plotted with the 'growthcurve' package (https://github.com/briandconnelly/growthcurve) The parameters of the model were extracted with the coef() function.

Chen *et al. BMC Plant Biology*   (2023) 23:363

Page 12 of 13

designed the study and significantly contributed to writing the manuscript. All authors read and approved the final manuscript.

Data Availability
The full dataset generated and analysed during the current study is available in the NCBI Gene Expression Omnibus repository (https://www.ncbi.nlm.nih.gov/geo/; Accession: GSE216253).

Declarations

Ethics approval and consent to participate
We declare that experimental research on plants were done in compliance of institutional, national, and international guidelines and legislation. Seeds of durum wheat cultivar Kronos were obtained from the Germplasm Resources Unit (GRU) of the John Innes Centre, which distributes them in accordance with the guidelines set by the International Treaty on Plant Genetic Resources for Food and Agriculture.

Consent for publication
Not applicable.

Competing interests
The authors declare that they have no competing interests.

References

1. Barron C, Surget A, Rouau X. Relative amounts of tissues in mature wheat (Triticum aestivum L.) grain and their carbohydrate and phenolic acid composition. J Cereal Sci. 2007;45(1):88–96.
2. Shewry PR, Mitchell RAC, Tosi P, Wan Y, Underwood C, Lovegrove A, Freeman J, Toole GA, Mills ENC, Ward JL. An integrated study of grain development of wheat (cv. Hereward). J Cereal Sci. 2012;56(1):21–30.
3. Shewry PR, Hey SJ. The contribution of wheat to human diet and health. Food Energy Secur. 2015;4(3):178–202.
4. Olsen OA. Endosperm development: cellularization and cell fate specification. Annu Rev Plant Physiol Plant Mol Biol. 2001;52:233–67.
5. Bechtel DB, Zayas I, Kaleikau L, Pomeranz Y. Size-distribution of wheat starch granules during endosperm development. Cereal Chem. 1990;67:59–63.
6. Chia T, Chirico M, King R, Ramirez-Gonzalez R, Saccomanno B, Seung D, Simmonds J, Trick M, Uauy C, Verhoeven T, et al. A carbohydrate-binding protein, B-GRANULE CONTENT 1, influences starch granule size distribution in a dose-dependent manner in polyploid wheat. J Exp Bot. 2020;71:105–15.
7. Chia T, Adamski NM, Saccomanno B, Greenland A, Nash A, Uauy C, Trafford K. Transfer of a starch phenotype from wild wheat to bread wheat by deletion of a locus controlling B-type starch granule content. J Exp Bot. 2017;68:5497–509.
8. Soh HN, Sissons MJ, Turner MA. Effect of starch granule size distribution and elevated amylose content on durum dough rheology and spaghetti cooking quality. Cereal Chem. 2006;83:513–9.
9. Park SH, Wilson JD, Seabourn BW. Starch granule size distribution of hard red winter and hard red spring wheat: its effects on mixing and breadmaking quality. J Cereal Sci. 2009;49:98–105.
10. Saccomanno B, Berbezy P, Findlay K, Shoesmith J, Uauy C, Viallis B, Trafford K. Characterization of wheat lacking B-type starch granules. J Cereal Sci 2022, 104.
11. Stoddard FL, Sarker R. Characterization of starch in Aegilops species. Cereal Chem. 2000;77:445–7.
12. Lindeboom N, Chang PR, Tyler RT. Analytical, biochemical and physicochemical aspects of starch granule size, with emphasis on small granule starches: a review. Starch - Stärke. 2004;56:89–99.
13. Edwards MA, Osborne BG, Henry RJ. Effect of endosperm starch granule size distribution on milling yield in hard wheat. J Cereal Sci. 2008;48(1):180–92.
14. Beres BL, Rahmani E, Clarke JM, Grassini P, Pozniak CJ, Geddes CM, Porker KD, May WE, Ransom JK. A systematic review of durum wheat: enhancing production systems by exploring genotype, environment, and management (G x E x M) synergies. Front Plant Sci. 2020;11:568657.
15. Lafiandra D, Sestili F, Sissons M, Kiszonas A, Morris CF. Increasing the versatility of durum wheat through modifications of protein and starch composition and grain hardness. Foods 2022, 11(11).
16. De Vita P, Taranto F. Durum wheat (Triticum turgidum ssp. durum) breeding to meet the challenge of climate change. In: *Advances in Plant Breeding Strategies: Cereals* 2019: 471–524.
17. Martínez-Moreno F, Ammar K, Solís I. Global changes in cultivated area and breeding activities of durum wheat from 1800 to date: a historical review. Agronomy 2022, 12(5).
18. Dias AS, Semedo J, Ramalho JC, Lidon FC. Bread and durum wheat under heat stress: a comparative study on the photosynthetic performance. J Agron Crop Sci. 2011;197(1):50–6.
19. Krasileva KV, Vasquez-Gross HA, Howell T, Bailey P, Paraiso F, Clissold L, Simmonds J, Ramirez-Gonzalez RH, Wang X, Borrill P, et al. Uncovering hidden variation in polyploid wheat. Proc Natl Acad Sci. 2017;114:E913–21.
20. Uauy C, Wulff BBH, Dubcovsky J. Combining traditional mutagenesis with new high-throughput sequencing and genome editing to reveal hidden variation in polyploid wheat. Annu Rev Genet. 2017;51:435–54.
21. Maccaferri M, Harris NS, Twardziok SO, Pasam RK, Gundlach H, Spannagl M, Ormanbekova D, Lux T, Prade VM, Milner SG, et al. Durum wheat genome highlights past domestication signatures and future improvement targets. Nat Genet. 2019;51:885–95.
22. Gu Y, Han S, Chen L, Mu J, Duan L, Li Y, Yan Y, Li X. Expression and regulation of genes involved in the reserve starch biosynthesis pathway in hexaploid wheat (Triticum aestivum L). Crop J. 2021;9:440–55.
23. Bray NL, Pimentel H, Melsted P, Pachter L. Near-optimal probabilistic RNA-seq quantification. Nat Biotechnol. 2016;34(5):525–7.
24. Abu-Jamous B, Kelly S. Clust: automatic extraction of optimal co-expressed gene clusters from gene expression data. Genome Biol. 2018;19(1):172.
25. Burton RA, Johnson PE, Beckles DM, Fincher GB, Jenner HL, Naldrett MJ, Denyer K. Characterization of the genes encoding the cytosolic and plastidial forms of ADP-glucose pyrophosphorylase in wheat endosperm. Plant Physiol. 2002;130(3):1464–75.
26. Tetlow IJ, Davies EJ, Vardy KA, Bowsher CG, Burrell MM, Emes MJ. Subcellular localization of ADPglucose pyrophosphorylase in developing wheat endosperm and analysis of the properties of a plastidial isoform. J Exp Bot. 2003;54:715–25.
27. Wang Y, Hou J, Liu H, Li T, Wang K, Hao C, Liu H, Zhang X. TaBT1, affecting starch synthesis and thousand kernel weight, underwent strong selection during wheat improvement. J Exp Bot. 2019;70(5):1497–511.
28. Hou J, Li T, Wang Y, Hao C, Liu H, Zhang X. ADP-glucose pyrophosphorylase genes, associated with kernel weight, underwent selection during wheat domestication and breeding. Plant Biotechnol J. 2017;15:1533–43.
29. Tjaden J, Möhlmann T, Kampfenkel K, Henrichs G, Neuhaus HE. Altered plastidic ATP/ADP-transporter activity influences potato (Solanum tuberosum L.) tuber morphology, yield and composition of tuber starch. Plant J. 1998;16:531–40.
30. Lim SL, Flutsch S, Liu J, Distefano L, Santelia D, Lim BL. Arabidopsis guard cell chloroplasts import cytosolic ATP for starch turnover and stomatal opening. Nat Commun. 2022;13(1):652.
31. Emanuelsson O, Brunak S, von Heijne G, Nielsen H. Locating proteins in the cell using TargetP, SignalP and related tools. Nat Protoc. 2007;2:953–71.
32. Esposito S, Bowsher CG, Emes MJ, Tetlow IJ. Phosphoglucomutase activity during development of wheat grains. J Plant Physiol. 1999;154(1):24–9.
33. Pfister B, Zeeman SC. Formation of starch in plant cells. Cell Mol Life Sci. 2016;73:2781–807.
34. Tetlow IJ, Beisel KG, Cameron S, Makhmoudova A, Liu F, Bresolin NS, Wait R, Morell MK, Emes MJ. Analysis of protein complexes in wheat amyloplasts reveals functional interactions among starch biosynthetic enzymes. Plant Physiol. 2008;146(4):1878–91.
35. Hennen-Bierwagen TA, Lin Q, Grimaud F, Planchot V, Keeling PL, James MG, Myers AM. Proteins from multiple metabolic pathways associate with starch biosynthetic enzymes in high molecular weight complexes: a model

for regulation of carbon allocation in maize amyloplasts. Plant Physiol. 2009;149:1541–59.

36. Morell MK, Blennow A, Kosar-Hashemi B, Samuel MS. Differential expression and properties of starch branching enzyme isoforms in developing wheat endosperm. Plant Physiol. 1997;113:201–8.

37. Sundberg M, Pfister B, Fulton D, Bischof S, Delatte T, Eicke S, Stettler M, Smith SM, Streb S, Zeeman SC. The heteromultimeric debranching enzyme involved in starch synthesis in Arabidopsis requires both Isoamylase1 and Isoamylase2 subunits for complex stability and activity. PLoS ONE. 2013;8:e75223.

38. Kubo A, Colleoni C, Dinges JR, Lin Q, Lappe RR, Rivenbark JG, Meyer AJ, Ball SG, James MG, Hennen-Bierwagen TA, et al. Functions of heteromeric and homomeric isoamylase-type starch-debranching enzymes in developing maize endosperm. Plant Physiol. 2010;153:956–69.

39. Utsumi Y, Utsumi C, Sawada T, Fujita N, Nakamura Y. Functional diversity of isoamylase oligomers: the ISA1 homo-oligomer is essential for amylopectin biosynthesis in rice endosperm. Plant Physiol. 2011;156:61–77.

40. Feike D, Seung D, Graf A, Bischof S, Ellick T, Coiro M, Soyk S, Eicke S, Mettler-Altmann T, Lu K-J, et al. The starch granule-associated protein EARLY STARVATION1 is required for the control of starch degradation in Arabidopsis thaliana leaves. Plant Cell. 2016;28:1472–89.

41. Song K, Lee DW, Kim J, Kim J, Guim H, Kim K, Jeon JS, Choi G. EARLY STARVA-TION 1 is a functionally conserved protein promoting gravitropic responses in plants by forming starch granules. Front Plant Sci. 2021;12:628948.

42. Seung D. Amylose in starch: towards an understanding of biosynthesis, structure and function. New Phytol. 2020;228:1490–504.

43. Jane JL, Chen YY, Lee LF, Mcpherson AE, Wong KS, Radosavljevic M, Kasemsuwan T. Effects of amylopectin branch chain length and amylose content on the gelatinization and pasting properties of starch. Cereal Chem. 1999;76:629–37.

44. Vrinten PL, Nakamura T. Wheat granule-bound starch synthase I and II are encoded by separate genes that are expressed in different tissues. Plant Physiol. 2000;122:255–64.

45. Seung D, Soyk S, Coiro M, Maier BA, Eicke S, Zeeman SC. PROTEIN TARGETING TO STARCH is required for localising GRANULE-BOUND STARCH SYNTHASE to starch granules and for normal amylose synthesis in Arabidopsis. PLoS Biol. 2015;13:e1002080.

46. Wang W, Wei X, Jiao G, Chen W, Wu Y, Sheng Z, Hu S, Xie L, Wang J, Tang S, et al. GBSS-BINDING PROTEIN, encoding a CBM48 domain-containing protein, affects rice quality and yield. J Integr Plant Biol. 2020;62:948–66.

47. Blennow A. Phosphorylation of the starch granule. In: Starch: Metabolism and Structure 2015: 399–424.

48. Schreier TB, Umhang M, Lee SK, Lue WL, Shen Z, Silver D, Graf A, Müller A, Eicke S, Stadler-Waibel M, et al. LIKE SEX4 1 acts as a beta-amylase-binding scaffold on starch granules during starch degradation. Plant Cell. 2019;31:2169–86.

49. Tickle P, Burrell MM, Coates SA, Emes MJ, Tetlow IJ, Bowsher CG. Characterization of plastidial starch phosphorylase in Triticum aestivum L. endosperm. J Plant Physiol. 2009;166:1465–78.

50. Zeeman SC, Thorneycroft D, Schupp N, Chapple A, Weck M, Dunstan H, Haldimann P, Bechtold N, Smith AM, Smith SM. Plastidial alpha-glucan phosphorylase is not required for starch degradation in Arabidopsis leaves but has a role in the tolerance of abiotic stress. Plant Physiol. 2004;135:849–58.

51. Malinova I, Mahlow S, Alseekh S, Orawetz T, Fernie AR, Baumann O, Steup M, Fettke J. Double knockout mutants of Arabidopsis grown under normal conditions reveal that the plastidial phosphorylase isozyme participates in transitory starch metabolism. Plant Physiol. 2014;164:907–21.

52. Satoh H, Shibahara K, Tokunaga T, Nishi A, Tasaki M, Hwang S-K, Okita TW, Kaneko N, Fujita N, Yoshida M, et al. Mutation of the plastidial alpha-glucan phosphorylase gene in rice affects the synthesis and structure of starch in the endosperm. Plant Cell. 2008;20:1833–49.

53. Critchley JH, Zeeman SC, Takaha T, Smith AM, Smith SM. A critical role for disproportionating enzyme in starch breakdown is revealed by a knock-out mutation in Arabidopsis. Plant J. 2001;26:89–100.

54. Hwang S-K, Koper K, Satoh H, Okita TW. Rice endosperm starch phosphorylase (Pho1) assembles with disproportionating enzyme (Dpe1) to form a protein complex that enhances synthesis of malto-oligosaccharides. J Biol Chem. 2016;291:19994–20007.

55. Chia T, Thorneycroft D, Chapple A, Messerli G, Chen J, Zeeman SC, Smith SM, Smith AM. A cytosolic glucosyltransferase is required for conversion of starch to sucrose in Arabidopsis leaves at night. Plant J. 2004;37:853–63.

56. Fettke J, Chia T, Eckermann N, Smith A, Steup M. A transglucosidase necessary for starch degradation and maltose metabolism in leaves at night acts on cytosolic heteroglycans (SHG). Plant J. 2006;46:668–84.

57. Cho MH, Lim H, Shin DH, Jeon JS, Bhoo SH, Park YI, Hahn TR. Role of the plastidic glucose translocator in the export of starch degradation products from the chloroplasts in Arabidopsis thaliana. New Phytol. 2011;190(1):101–12.

58. Flutsch S, Horrer D, Santelia D. Starch biosynthesis in guard cells has features of both autotrophic and heterotrophic tissues. Plant Physiol. 2022;189(2):541–56.

59. Andriotis VM, Pike MJ, Bunnewell S, Hills MJ, Smith AM. The plastidial glucose-6-phosphate/phosphate antiporter GPT1 is essential for morphogenesis in Arabidopsis embryos. Plant J. 2010;64(1):128–39.

60. Qu A, Xu Y, Yu X, Si Q, Xu X, Liu C, Yang L, Zheng Y, Zhang M, Zhang S, et al. Sporophytic control of anther development and male fertility by glucose-6-phosphate/phosphate translocator 1 (OsGPT1) in rice. J Genet Genomics. 2021;48(8):695–705.

61. Schreier TB, Fahy B, David LC, Siddiqui H, Castells-Graells R, Smith AM. Introduction of glucan synthase into the cytosol in wheat endosperm causes massive maltose accumulation and represses starch synthesis. Plant J. 2021;106(5):1431–42.

62. Ziegler P. Cereal beta-amylases. J Cereal Sci. 1999;29:195–204.

63. Ju L, Deng G, Liang J, Zhang H, Li Q, Pan Z, Yu M, Long H. Structural organization and functional divergence of high isoelectric point alpha-amylase genes in bread wheat (Triticum aestivum L.) and barley (Hordeum vulgare L). BMC Genet. 2019;20(1):25.

64. Wilkes MAMA, Seung D, Levavasseur G, Trethowan RMRM, Copeland L. Effects of soil type and tillage on protein and starch quality in three related wheat genotypes. Cereal Chem. 2010;87:95–9.

65. Sparla F, Falini G, Botticella E, Pirone C, Talamè V, Bovina R, Salvi S, Tuberosa R, Sestili F, Trost P. New starch phenotypes produced by TILLING in barley. PLoS ONE. 2014;9:e107779.

66. Seung D, Smith AM. Starch granule initiation and morphogenesis – progress in Arabidopsis and cereals. J Exp Bot. 2019;70:771–84.

67. Chen J, Hawkins E, Seung D. Towards targeted starch modification in plants. Curr Opin Plant Biol. 2021;60:102013.

68. Hawkins E, Chen J, Watson-Lazowski A, Ahn-Jarvis J, Barclay JE, Fahy B, Hartley M, Warren FJ, Seung D. STARCH SYNTHASE 4 is required for normal starch granule initiation in amyloplasts of wheat endosperm. New Phytol. 2021;230:2371–86.

69. Chen J, Chen Y, Watson-Lazowski A, Hawkins E, Barclay J, Fahy B, Denley-Bowers R, Corbin K, Warren F, Blennow A et al. The plastidial protein MRC promotes starch granule initiation in wheat leaves but delays B-type granule initiation in the endosperm *BioRxiv* 2022, https://doi.org/10.1101/2022.10.07.511297.

70. Seung D, Schreier TB, Bürgy L, Eicke S, Zeeman SC. Two plastidial coiled-coil proteins are essential for normal starch granule initiation in Arabidopsis. Plant Cell. 2018;30:1523–42.

71. Sato Y, Takehisa H, Kamatsuki K, Minami H, Namiki N, Ikawa H, Ohyanagi H, Sugimoto K, Antonio BA, Nagamura Y. RiceXPro version 3.0: expanding the informatics resource for rice transcriptome. Nucleic Acids Res. 2013;41(Database issue):D1206–1213.

72. Qu J, Ma C, Feng J, Xu S, Wang L, Li F, Li Y, Zhang R, Zhang X, Xue J, et al. Transcriptome dynamics during maize endosperm development. PLoS ONE. 2016;11(10):e0163814.

73. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 2014;15(12):550.

74. Li Y, Andrade J. DEApp: an interactive web interface for differential expression analysis of next generation sequence data. Source Code Biol Med. 2017;12:2.

75. Yates AD, Allen J, Amode RM, Azov AG, Barba M, Becerra A, Bhai J, Campbell LI, Carbajo Martinez M, Chakiachvili M, et al. Ensembl Genomes 2022: an expanding genome resource for non-vertebrates. Nucleic Acids Res. 2022;50(D1):D996–D1003.

76. Mu J, Chen L, Gu Y, Duan L, Han S, Li Y, Yan Y, Li X. Genome-wide identification of internal reference genes for normalization of gene expression values during endosperm development in wheat. J Appl Genetics. 2019;60:233–41.

Publisher's Note