Role of B3 domain transcription factors of the AFL family in maize kernel filling

Aurélie Grimault[a,b,c,1], Ghislaine Gendrot[a,b,c], Sandrine Chaignon[a,b,c], Françoise Gilard[d], Guillaume Tcherkez[d,2], Johanne Thévenin[e,f], Bertrand Dubreucq[e,f], Nathalie Depège-Fargeix[a,b,c], Peter M. Rogowsky[a,b,c,3]

[a] Université de Lyon, Ecole Normale Supérieure de Lyon, Université Lyon 1, Unité Reproduction et Développement des Plantes, F-69364 Lyon, France

[b] INRA, UMR879 Reproduction et Développement des Plantes, F-69364 Lyon, France

[c] CNRS, UMR5667 Reproduction et Développement des Plantes, F-69364 Lyon, France

[d] CNRS, UMR9213 Institute of Plant Sciences Paris-Saclay, F-91405 Orsay, France

[e] INRA, UMR1318, Institut Jean-Pierre Bourgin, Saclay Plant Sciences, Versailles, France

[f] AgroParisTech, Institut Jean-Pierre Bourgin, Saclay Plant Sciences, Versailles, France

running title: *ZmAFL* genes

[1] **present address:** Carnegie Institution of Science, 260 Panama Street, Stanford, CA 94305, U.S.A.

[2] **present address:** Research School of Biology, ANU College of Medicine, Biology and Environment, Australian National University, 2601 Canberra, ACT, Australia

[3] **corresponding author:** Dr. Peter M. Rogowsky, RDP, ENS de Lyon, 46 allée d'Italie, F-69364 Lyon, France, Tel +33 4 72 72 86 07, Fax +33 4 72 72, email address: peter.rogowsky@ens-lyon.fr.

Abstract

In the dicot *Arabidopsis thaliana*, the B3 transcription factors, ABA-INSENSITIVE 3 (ABI3), FUSCA 3 (FUS3) and LEAFY COTYLEDON 2 (LEC2) are key regulators of seed maturation. This raises the question of the role of ABI3/FUS3/LEC2 (AFL) proteins in cereals, where not only the embryo but also the persistent endosperm accumulates reserve substances. Among the five *ZmAFL* genes identified in the maize genome, *ZmAFL2* and *ZmAFL3/ZmVp1* closely resemble *FUS3* and *ABI3*, respectively, in terms of their sequences, domain structure and gene activity profiles. Of the three genes that fall into the *LEC2* phylogenetic sub-clade, *ZmAFL5* and *ZmAFL6* have constitutive gene activity, whereas *ZmAFL4,* like *LEC2,* has preferential gene activity in pollen and seed, although its seed gene activity is restricted to the endosperm during reserve accumulation. Knock down of *ZmAFL4* gene activity perturbs carbon metabolism and reduces starch content in the developing endosperm 20 DAP. ZmAFL4 and ZmAFL3/ZmVp1 *trans*-activate a maize oleosin promoter in a heterologous moss system. In conclusion our results suggest, based on gene activity profiles, that the functions of *FUS3* and *ABI3* could be conserved between dicot and monocot species. In contrast, LEC2 function may have partially diverged in cereals where our findings provide first evidence of the specialization of *ZmAFL4* for roles in the endosperm.

Keywords

AFL network; endosperm; maize; starch; transcription factor

Abbreviations

aba-insensitive 3 (ABI3); ABI3/FUS3/LEC2 (AFL); ADP-glucose pyrophosphorylase (AGPase); analysis of variance (ANOVA); basic leucine zipper (bZIP); days after pollination (DAP); fusca 3 (FUS3); glycerol-3-phosphate (G3P); leafy cotyledon (LEC); quantitative reverse transcription-PCR (qRT-PCR); RNA interference (RNAi); Zea mays viviparous1 (ZmVp1).

1. Introduction

In maize, as in many cereal crops, the reserve substances necessary for efficient germination of the embryo, and thus successful propagation, are stored both within the embryo (mainly oil and protein) and in a surrounding nourishing tissue, the endosperm (mainly starch and protein). Whereas the entire maize seed is cherished by mankind as a major source of food and animal feed, the non-food use of the maize seed is largely limited to the endosperm and more precisely to starch extraction for industrial products or hydrolysis for biofuels.

In maize, the accumulation of reserve substances in the embryo and endosperm occurs during the filling stage of seed development [1,2]. This stage is preceded by developmental events such as pattern formation, morphogenesis and differentiation, and followed by seed dehydration, which allows the seed to become quiescent. Distinct transcriptome profiles suggest that the three developmental stages (early, filling and desiccation) have dedicated genetic programs which are controlled mainly at the transcriptional level [3,4]. Despite passing through functionally equivalent stages, the embryo and endosperm clearly execute

different genetic programs to elaborate their characteristic morphology and to accumulate distinct reserve substances [5,6].

The control of seed filling has been well characterized in Arabidopsis, where the endosperm is largely transient and reserve substances therefore principally accumulate in the embryo. The three B3 domain transcription factors ABA-INSENSITIVE 3 (ABI3) [7], FUSCA3 (FUS3) [8] and LEAFY COTYLEDON 2 (LEC2) [9] form the so-called "AFL network" in association with LEAFY COTYLEDON 1 (LEC1) [10], which is homologous to the HAP3 subunits of the CAAT box-binding factor family [11,12]. Mutations in these genes cause pleiotropic but distinct effects on seed maturation, including a lack of both storage reserve accumulation and desiccation tolerance [13-15]. *AFL* genes have distinct temporal and spatial gene activity patterns during Arabidopsis seed development. All three genes are active in the embryo and *LEC2* and *FUS3* also are active in the endosperm [2,16]. Temporally, *LEC2* is the first gene to become active, with a peak at the heart stage. The activity of the *FUS3* gene peaks during early seed maturation, and finally *ABI3* is active during desiccation [17,18]. Despite a certain overlap in their gene activity patterns, each *AFL* gene has a distinct function. Networking is indicated by interactions among *AFL* genes. For example, LEC2 activates *ABI3* and *FUS3* gene activity whereas ABI3 and FUS3 auto-regulate themselves and interact through mutual activation [17,19,20]. In addition, interactions between *LEC1* and *AFL* genes have been shown by genetic and transcriptome analyses: LEC1 can activate *ABI3* and *FUS3* activity whereas *LEC1* activity is up-regulated by LEC2 [21-23].

AFL transcription factors are considered master regulators since they trigger a regulatory cascade by activating secondary transcription factors, which in turn govern multiple metabolic and developmental pathways. For example *WRI1*, which encodes the main regulator

of lipid biosynthesis in the seed, is a direct target of LEC2 [24]. However LEC2 also acts directly on genes involved in reserve accumulation including *OLE1*, encoding an oleosin [25] and *At2S1–S4* and *2S-like,* which encode seed storage proteins [17,26]. Direct targeting by LEC2 is mediated by its B3 domain, which binds specifically to RY-motifs such as CATGCA [19,26].

The functions of *AFL* genes have been extensively studied in Arabidopsis, and putative orthologs have been described in several monocot and dicot species [2]. However, it remains unclear to what extent their functions and targets are conserved. The best characterized *AFL* gene in cereals is *ZmVIVIPAROUS1 (ZmVp1)*, the maize ortholog of *ABI3* [27-29]. However, the extensive characterization of the *Zmvp1* mutant has focused largely on a single aspect, the role of *ZmVp1* in ABA-mediated regulation of kernel dormancy [30,31]. Indications for functional conservation also exist for *ZmLEC1* and *ZmWri1*, since their over-expression increases seed oil content in maize [32,33]. *ZmWri1a* and *ZmWri1b* both complement their putative co-ortholog *WRI1* in the Arabidopsis *wri1* mutant, despite minor qualitative changes in the oil of *wri1* mutants complemented with *ZmWri1a* and *ZmWri1b* [32,34]. However, since in maize oil accumulates principally in the scutellum of the embryo [35] and since dormancy also concerns primarily the embryo, these data do not shed light on the question of a potential control of endosperm reserve substances by members of the AFL network, which cannot be satisfactorily addressed in the exalbuminous Arabidopsis seed.

The identities of the transcriptional regulators of well-known structural genes needed for endosperm starch synthesis in cereals remain surprisingly elusive despite their potential as targets for the modification of important crop traits. Recent work has shown that the rice basic leucine zipper (bZIP) transcription factor bZIP58 directly regulates *Starch synthase IIa* and

Starch branching enzyme 1 gene activity [36]. In addition, seed storage protein gene activity in the maize endosperm is known to be regulated by the bZIP transcription factor OPAQUE2 and interacting proteins such as the zinc finger transcription factor PBF [37]. Here we address the important question of whether members of the AFL family could also participate in the regulation of seed storage product accumulation. We present the identification, phylogenetic analysis and gene activity analysis of putative maize orthologs of *ABI3*, *FUS3* and *LEC2*. Our results indicate a previously unexplored function for *ZmAFL4* in kernel metabolism and the regulation of starch accumulation in the maize endosperm.

2. Materials and methods

2.1. Plant material and plant culture

Maize plants were grown in a greenhouse with a 16-h photoperiod (400 µmol m^{-2} s^{-1}) at 24°C/19°C (day/night) and without control of the relative humidity, as described in Pouvreau *et al* [32]. All plants were propagated by hand pollination. Maize genotype B73 was used for temporal and spatial gene activity analyses of *ZmAFL* genes during kernel development, and genotype A188 for maize transformation and for gene activity analyses of *ZmAFL* genes in maize organs. Seed of the *Zmvp1* mutant allele *vp1-Mum1::Mu* was obtained from the Maize Genetics Cooperation Stock Center (stock 326BH).

2.2. T-DNA constructs and plant transformation

The coding sequences of *ZmAFL* genes were PCR amplified from cDNA from appropriate kernel stages (genotype B73), using primers reported in Supplementary Table S3. The PCR products were cloned into the vector pDONRZeo (Invitrogen), and the resulting

entry vectors were sequenced prior to LR recombination. The final plasmids L1238 (*ZmAFL2*-RNAi), L1240 (*ZmAFL4*-RNAi) and L1242 (*ZmAFL5/6*-RNAi) were used for maize transformation as described previously [32]. The plasmids contained the backbone of vector pSB11, a Basta resistance cassette (rice *Actin* promoter and intron, *Bar* gene and *Nos* terminator) next to the right border, a GFP cassette (*CsVMV* promoter and *FAD2* intron, *GFP* gene and *Nos* terminator) and the respective *ZmAFL* gene fragment separated by the rice *Tubulin* intron in a head to head configuration under the control of the constitutive rice *Actin* promoter and intron.

2.3. Sequence analysis

Protein sequences similar to *AtAFL* proteins were retrieved by using the BlastP program to query the maize genome database (http://blast.gramene.org/Multi/blastview) and the protein database of the National Center for Biotechnology Information (http://www.ncbi.nlm.nih.gov/blast/Blast.cgi). The protein sequence alignments were generated using the programs ClusalW2 or Omega available at http://www.ebi.ac.uk/Tools/msa/. Gene models were downloaded from release AGPv3 of the B73 maize (*Zea mays*) genome assembly (http://ensembl.gramene.org/Zea_mays/Info/Index). Functional B3 domains were identified using the Pfam HMM database (http://pfam.xfam.org/). A1, B1, B2 and C-terminal were defined as conserved blocks in multiple sequence alignments [7].

2.4. Phylogenetic analysis

Amino acid sequences were aligned using Clustal Omega (http://www.ebi.ac.uk/Tools/msa/clustalo). Conserved blocks were selected manually with the

Seaview program (http://doua.prabi.fr.fr/software/seaview.html) and phylogenetic trees were generated using PHYML with the WAG substitution model and 1000 bootstrap replicates.

2.5. Analysis of gene activity by qRT-PCR

Total RNA was extracted with the Tri-Reagent® (Molecular Research Center) and treated with DNase as described previously [38]. First strand cDNA was prepared using random hexamers (Amersham Biosciences) and Superscript III reverse transcriptase (Invitrogen) according to the manufacturer's protocol. The resulting cDNA was diluted 50 times. Quantitative RT-PCR was carried out using FastStart SYBR Green Master mix (Roche) on a StepOne Real-Time PCR System (Applied Biosystems) according to the manufacturer's protocol with the following program: 10 min at 95°C, followed by 40 cycles of 95°C for 10 sec, 60°C for 30 sec. The average threshold cycle (Ct) was calculated using StepOne software v2.3 (Applied Biosystems). The PCR efficiency (E) and the relative gene activity (R) of target genes were calculated as described [39] using *Actin* as a reference gene. The primers used are listed in Supplementary Table S3. The number of plants, seeds and biological replicates used for gene activity analyses are indicated in the respective figure legends. With regard to *ZmAFL4*-RNAi plants, the gene activity analyses of *ZmAFL4* (Figure 3) and of *Zein* genes (Supplementary Figure S5) were performed with the same biological material as for the metabolomics analyses and the determination of starch and protein content.

2.6. Metabolomic measurements and analysis

Metabolomics analyses were performed in biological triplicate on separated embryos and endosperm of homozygous *ZmAFL4*-RNAi and wild-type kernels at 20 DAP. Each replicate represented a different plant, from which at least 50 endosperms or 50 embryos of

the same ear were pooled. The extraction of metabolites from plant tissues and their analysis by gas chromatography coupled to time of flight mass spectrometry were performed as described [40]. Data were analyzed by principal component analysis.

2.7. Starch dosage

Three biological replicates representing individual homozygous transgenic or wild-type plants were analyzed. Starch was extracted from 10 or 20 mg fresh weight by three successive ethanol extractions at 80°C for 20 min each with 250 µl of 98% v/v ethanol, 80% v/v ethanol and 50% v/v ethanol. Starch content was measured by spectrophotometric quantification of NADPH at 340 nm, produced during glucose phosphorylation by hexokinase after enzymatic degradation of starch [41].

2.8. Physcomitrella patens *protoplast* trans-*activation assay*

Moss culture, protoplast preparation and transformation as well as flow cytometry were carried out as described [42]. The Zm*AFL* coding sequences were recombined into the destination vector pBS-TPp-A [42]. *Trans*-activation was monitored in four biological repetitions by GFP fluorescence quantification using flow cytometry. GFP fluorescence was detected with an FITC 527 nm/30 nm band-pass filter.

3. Results

3.1. The maize genome encodes five AFL *genes*

To identify *ZmAFL* (*ABI3, FUS3, LEC2*) candidate genes in the maize genome, a BLASTp search against the deduced amino acid sequences of LEC2, FUS3 and ABI3 was carried out in the maize protein data base (http://www.maizesequence.org). Six *ZmAFL* genes

were identified and named *ZmAFL1* to *ZmAFL6* (Table 1). *ZmAFL3* was identical to the *ZmVp1* gene originally described by McCarty *et al.* [43].

An initial analysis of domain structure revealed that the B3 domain of the deduced ZmAFL1 protein was associated with a CW-type zinc finger. Since this domain combination is characteristic of VAL/HSI but not AFL family members [44], and since a phylogenetic relationship between ZmAFL1 and VAL rather than AFL proteins was established (Supplementary Fig. S2), this gene was removed from our study.

The gene model (GRMZM2G405699) provided for *ZmAFL5* in the maize genome database (http://www.maizesequence.org) was truncated compared to genes from sorghum and rice. We extended the model using the EST sequence DN219455 to give the complete deduced amino acid sequence presented in Supplementary Fig. S1. This sequence shared more than 92% identity with the ZmAFL6 sequence indicating that the corresponding genes might be paralogs. However, further analysis showed that the genes were not located in duplicated blocks of the present maize genome, which arose during the most recent whole genome duplication 5 million year ago [45]. Consequently *ZmAFL5* and *ZmAFL6* were not nearly identical paralogs as defined by Emrich *et al.* [46] but merely two closely related genes without a precise phylogenetic history.

3.2. Phylogenetic relationship between maize and Arabidopsis AFL

To identify potential orthologs between maize and Arabidopsis *AFL* genes, and to clarify the phylogenetic position of the AFL subfamily, a partial phylogenetic tree of the B3 family was constructed for the monocots maize (*Zea mays*) and rice (*Oryza sativa*), the dicot *Arabidopsis thaliana* and the moss *Physcomitrella patens*. For maize 13 non redundant

sequences identified by Peng and Weselake [47] and named "ABI3/Vp1" by these authors were completed with 14 more maize protein sequences including ZmAFL2 and ZmAFL5 from the NCBI (Supplementary Table S1). The 28 rice and 37 Arabidopsis sequences were taken from Peng and Weselake [47] and the three Physcomitrella sequences from Marella *et al.* [48].

In the resulting tree the five ZmAFL proteins (ZmAFL2, ZmAFL3/ZmVp1, ZmAFL4, ZmAFL5 and ZmAFL6) clustered cleanly with the three Arabidopsis AFL proteins, founding an AFL clade, which could be further divided into three sub-clades (Fig. 1; Supplementary Fig. S2). A first, robust sub-clade was formed by ZmVp1 and ABI3, consistent with the observation that *ZmVp1* complements the Arabidopsis *abi3* mutant [30] and with multiple alignments confirming earlier observations that only ZmVp1 and ABI3 contained additional conserved domains named A1, B1 and B2 [7,49]. ZmAFL2 and FUS3 clustered in a second sub-clade consistent with the B3 domain sequence of ZmAFL2 showing stronger conservation with that of FUS3 (63% identity) than with those of ABI3 (62%) or LEC2 (54%). In addition only FUS3 and ZmAFL2 were found to contain a conserved C-terminal domain of approximately 60 amino acids. Together these results indicated that ZmAFL2 was the putative ortholog of FUS3. The remaining three maize AFL proteins, ZmAFL4, ZmAFL5 and ZmAFL6 were grouped in the LEC2 sub-clade. Consequently ZmAFL4, ZmAFL5 and ZmAFL6 were putative co-orthologs of LEC2.

3.3. Diversification of ZmAFL *gene activity patterns*

To investigate the potential conservation of *AFL* functions between maize and Arabidopsis, the activity of each *ZmAFL* gene in both vegetative and reproductive maize

organs was measured by quantitative reverse transcription-PCR (qRT-PCR) experiments. Whereas all 5 *ZmAFL* genes were active in the maize kernel (Fig. 2A), only *ZmVp1* gene activity was kernel specific, mirroring silique-specific gene activity of *ABI3* in Arabidopsis [7]. *ZmAFL2* and *ZmAFL4* showed strong additional gene activity in tassels and in pollen. Their gene activity patterns are reminiscent of those of *FUS3* and *LEC2*, characterized by a high level of transcripts in young siliques but also in stamens for *FUS3* or mature pollen for *LEC2* (http://bar.utoronto.ca/welcome.htm; [50-52]). *ZmAFL5* and *ZmAFL6* were active in almost all organs examined, mRNA levels being higher in reproductive organs.

To gain further insight into the role of *ZmAFL* genes in maize kernel development, more detailed temporal gene activity patterns were established (Fig. 2B). *ZmAFL2, ZmAFL5* and *ZmAFL6* gene activity was highest in ovules and young kernels with a peak at 3 days after pollination (DAP), although moderate gene activity, in particular of *ZmAFL2*, persisted during the filling stage between 9 and 30 DAP. *ZmAFL4* transcripts were most abundant at the filling stage with a peak of gene activity at 15 DAP. Finally, *ZmVp1* exhibited a high mRNA level from the end of the filling stage throughout the dehydration phase, reaching a maximum at 35 DAP. The temporal gene activity pattern of *ZmVp1* was very similar to its proposed ortholog in Arabidopsis, whereas the onset of *ZmAFL2* activity was earlier than that of *FUS3* [9,17,18,53]. In the LEC2 clade *ZmAFL5* and *ZmAFL6* were active very early in kernel development, resembling *LEC2* gene activity in Arabidopsis [9,17], whereas *ZmAFL4* peaked later during the filling stage. In summary, *ZmAFL* genes were sequentially expressed at different stages of kernel development, reflecting to a certain degree the situation in Arabidospis.

The spatial gene activity patterns of *ZmAFL* genes in the maize kernel were studied at 15 DAP in dissected embryo, endosperm and pericarp samples (Fig. 2C). Whereas *ZmAFL5* and *ZmAFL6* had gene activity in all three seed compartments, the other *ZmAFL* genes showed strong activity either in one single, or two compartments. *ZmAFL2* and *ZmVp1* were active exclusively in the embryo, whereas *ZmAFL4* transcripts were most abundant in the endosperm, present in the pericarp and nearly absent in the embryo from 9 to 30 DAP (Fig. 2C, Supplementary Fig. S3). In conclusion, the five *ZmAFL* genes presented both distinct spatial and distinct temporal gene activity patterns. Contrary to the situation in Arabidopsis, where *ABI3*, *LEC2* and *FUS3* are active in both the embryo and endosperm ([2,16], http://bar.utoronto.ca/welcome.htm), our results showed a dichotomy in spatial gene activity for *ZmAFL2* and *ZmVp1* (embryo) and *ZmAFL4* (endosperm and pericarp), suggesting differences in the regulation of seed development between the two species.

*3.4. Absence of macroscopic kernel phenotypes in Zm*AFL*-RNAi lines*

To elucidate the function of *ZmAFL* genes, three RNA interference (RNAi) constructs under the control of the constitutive rice *Actin* promoter were generated, targeting *ZmAFL2*, *ZmAFL4* or both *ZmAFL5* and *ZmAFL6.* From eight independent transformation events per construct, two were selected for analysis of T1 kernels based on the confirmation of complete T-DNA transfer, single transgene copy number and high gene activity level of the construct. Molecular analysis of T1 kernels indicated that none of the transgenic lines showed a complete suppression of target gene transcript levels. The most efficient silencing was found in *ZmAFL2*-RNAi and *ZmAFL4*-RNAi kernels, for which the strongest event showed a 61% and a 87% decrease in mRNA levels, respectively (Fig. 3A; Supplementary Fig. S4A). In

ZmAFL5/6-RNAi lines, no significant decrease in transcript level was detected. Despite the knockdown of *ZmAFL2* and *ZmAFL4* gene activity, the corresponding transgenic kernels did not show visible defects such as shriveled kernels, modification of pigmentation, developmental arrest or vivipary of the embryo.

The *ZmAFL2*-RNAi and *ZmAFL4*-RNAi kernels as well as *Zmvp1* mutant kernels were used to investigate potential cross regulation between the five *ZmAFL* genes (Supplementary Fig. S4). Whereas no significant changes in *ZmAFL* gene activity were observed in either *ZmAFL2*-RNAi (Supplementary Fig. S4A) or *ZmAFL4*-RNAi kernels at 15 DAP (Supplementary Fig. S4B), in *Zmvp1* kernels, *ZmAFL*2 mRNA levels were clearly increased at 30 DAP, suggesting a negative regulation of *ZmAFL2* by *ZmVp1* (Supplementary Fig. S4C). In addition, *ZmAFL4* gene activity was drastically reduced in *Zmvp1* kernels at 30 DAP (Supplementary Fig. S4C). Since *ZmAFL4* and *ZmVp1* are not expressed in the same compartment, this result indicated either an indirect regulatory mechanism, or movement of the ZmVp1 protein.

3.5. ZmAFL4, a transcription factor contributing to the regulation of starch accumulation

Among the five *ZmAFL* genes, only *ZmAFL4* displayed strong gene activity in the endosperm (Fig. 2C, Supplementary Fig. S3) and showed a peak of gene activity during the filling stage (Fig. 2B). ZmAFL4 was therefore an excellent candidate regulator for the accumulation of reserve substances in the maize endosperm. To verify this hypothesis and to assess the impact of knocking down *ZmAFL4* gene activity on carbon metabolism in the maize kernel, a metabolomics analysis was performed. Embryos and endosperms were dissected at 20 DAP from the T3 kernels from self-pollinated ears of three homozygous

ZmAFL4-RNAi plants from the strongest transformation event, and three wild-type siblings. Not surprisingly, 58 of the 81 metabolites tested showed significant differences between embryo (accumulating fatty acids) and endosperm (accumulating starch) samples, independent of the genotype. These included monosaccharides (glucose, fructose, galactose) and fatty acids (linoleic acid).

In the endosperm data analysis, the first principal component clearly separated wild-type and transgenic samples. Analysis of variance (ANOVA) revealed significant differences (p-value <0.05) between wild-type and *ZmAFL4*-RNAi endosperms in the content of short-chain organic acids (succinic, malic and glyceric acid), free amino acids (aspartic acid, cysteine and glycin) and monosaccharides (ribose, xylose, glucose and fructose). For all significantly different metabolites except cysteine, the content was found to decrease in transgenic endosperm (Table 2).

Within the embryo data set, the content of 10 metabolites was significantly lower (p-value <0.05) in transgenic *ZmAFL4*-RNAi embryos compared to wild-type embryos (Supplementary Table S2). These could be classed into short organic acids (succinic acid), free amino acids (leucine, phenylalanine, alanine, tyrosine and valine), monosacharides (ribose, arabinose) and glycerol-3-phosphate (G3P). Due to the absence of gene activity in the embryo, the effect of ZmAFL4 on embryo metabolites was likely indirect, possibly involving metabolic feedback.

To complement the metabolomics analysis, starch content was determined in the same 20 DAP endosperm samples, as well as in samples from a second independent transformation event, by a spectrophotometric method based on enzymatic degradation. A significant reduction in starch content (p-value <0.05) was revealed in transgenic kernels at 20 DAP (Fig.

3B). Since no morphological defects had been noted in mature *ZmAFL4*-RNAi kernels, starch content was also measured in mature kernels. No statistically significant decrease was detected in *ZmAFL4*-RNAi kernels (Fig. 3C), indicating that the deficit observed at 20 DAP had been compensated during later development. Taken together these findings suggest that *ZmAFL4* is involved in the regulation of starch metabolism, especially at the beginning of the filling stage.

3.6. ZmAFL4 is not involved in seed protein storage

In maize, large amounts of protein are deposited in the endosperm, zeins being the major class. Analysis of *Zein* gene activity showed only minor up or down regulation of individual genes in *ZmAFL4*-RNAi endosperm compared to wild-type endosperm (Supplementary Fig. S5A). In addition, the total protein content was not significantly different in transgenic endosperms or embryos compared to wild-type (Supplementary Fig. S5B). Therefore, in contrast to Arabidopsis *LEC2* [9,17,26], *ZmAFL4* does not appear to be involved in the regulation of seed storage protein accumulation.

3.7. ZmVp1 is involved in seed oil storage

To explore the implication of Zm*AFL* genes in maize seed oil storage, upstream sequences of the two putative maize *Oleosin* genes Zm*OLE2* (GRMZM2G096435) and Zm*OLE3* (GRMZM2G480954) were tested for *trans*-activation by ZmAFL transcription factors in a *Physcomitrella patens* protoplast system [42]. Whereas strong background activity rendered the use of the Zm*OLE3* promoter impracticable, statistically significant *trans*-activation of the Zm*OLE2* promoter was observed in the presence of Zm*Vp1* or Zm*AFL4* (Fig. 4). Since micro-array data indicated that within the kernel Zm*OLE2* is active exclusively in

the embryo [54], the rather weak *trans*-activation by ZmAFL4, which is not active in the embryo, was unlikely to be biologically relevant. On the other hand, the increase of Zm*OLE2* transcription in the presence of the embryo-specific Zm*Vp1* was supported by down-regulation of *ZmOLE2* activity in *Zmvp1* mutant kernels (Supplementary Fig. S6) and provided first evidence that ZmVp1 might be directly involved in the regulation of seed oil storage.

4. Discussion

Our results suggest that the gene activity and function of AFL (ABI3/FUS3/LEC2) transcription factors, key regulators of storage reserve accumulation in the Arabidopsis seed, are at least partially conserved in the maize kernel. The gene activity patterns of the five *ZmAFL* genes identified in the maize genome indicate that each gene is actively transcribed in at least one seed compartment. In particular *ZmAFL4,* a putative *LEC2* co-ortholog with preferential activity in endosperm during kernel development, is involved in carbon metabolism and contributes to the regulation of starch accumulation.

4.1. Divergence of AFL *gene activity and function between Arabidopsis and cereals*

The number of *AFL* genes increases from three in Arabidopsis, through four in rice, to five *ZmAFL* genes identified in the maize genome, raising the question of which cereal gene(s) were the putative (co)-orthologs of *ABI3*, *FUS3* and *LEC2*. A phylogenetic analysis together with the absence or presence of characteristic domains in addition to the common B3 domain confirmed the previously established orthologous relationship between *ZmAFL3/ZmVp1* and *ABI3* [30] and provided data suggesting that *ZmAFL2* is the putative

maize ortholog of *FUS3* and that *ZmAFL4*, *ZmAFL5* and *ZmAFL6* are potential co-orthologs of *LEC2*.

In Arabidopsis, *AFL* genes have distinct temporal and spatial gene activity patterns during seed development [17,18,53] and conserved gene activity patterns in cereals would be a first indication for functional equivalence. Interestingly, the gene activity patterns of two *ZmAFLs* mirrored that of their putative Arabidopsis orthologs. Firstly our data confirmed that *ZmVp1* has gene activity in embryo at the end of the filling stage, as described previously for *ABI3* [28] and *TaVp1* from wheat [55,56]. This result completes data from other studies showing gene activity of *ZmVp1* and *ABI3* in the outer endosperm layer [2,57]. Secondly, *ZmAFL2* gene activity peaks during early maize kernel development and continues at the beginning of the filling stage, as has previously been shown for *FUS3* in the Arabidopsis seed. In addition, both genes share strong gene activity in non-seed tissues, particularly in mature gametes. Similarly in rice, *OsLFL1* has gene activity in anthers, pollens and young developing embryos [58], whilst in wheat although *TaFUS3* activity has not been assessed in reproductive tissues it is active in young embryos [55]. However, at least in maize and rice seeds, the gene activity of the putative *FUS3* orthologs is restricted to the embryo whereas *FUS3* itself has gene activity in both embryo and endosperm.

The situation is more complex for *LEC2* and its putative co-orthologs. The fairly constitutive gene activity of *ZmAFL5, ZmAFL6* and rice *IDEF-1* [59] contrasts sharply with *LEC2* gene activity, leaving in a first instance *ZmAFL4* as its putative functional equivalent. *LEC2* gene activity is detected mainly in pollen, embryo and endosperm, reaching a maximum at the heart stage of embryogenesis. *ZmAFL4* is also expressed in pollen, but its gene activity pattern in the seed is spatially more restricted (absent from the embryo, strong in endosperm)

and temporally shifted (detected at filling stage only), representing a new gene activity pattern for an *AFL* gene. In wheat the gene activity of *TaL2LA*, which has been proposed to be the ortholog of *LEC2*, has not been assessed in reproductive tissues and shows only early gene activity in the embryo similar to *LEC2*, *ZmAFL5* and *ZmAFL6* [55]. The activity pattern of the closely related *Ta2L2B* has not been described in detail [55]. This data provides evidence for a certain divergence in the gene activity patterns and possibly functions of *LEC2* genes between monocot and dicot species, in contrast to the situation for *FUS3* and *ABI3*.

Similarly to the situation in Arabidopsis, where complex direct and indirect regulatory links and interactions between *AFL* genes have been demonstrated by molecular and genetic analyses [17,18,20,23], our results indicate that cross regulation does exist in maize. In *Zmvp1* mutant kernels *ZmAFL4* is down-regulated, whereas *ZmAFL2* is up-regulated. In the latter case the situation is at first sight different from Arabidopsis, where mutual activation loops between ABI3 and FUS3 have been established [18]. However, a better spatial and temporal resolution of the gene activity patterns in maize mutants or knockdowns is needed prior to in depth comparisons between the regulatory networks.

LEC2 and FUS3 are both known to be involved directly or indirectly in competence for somatic embryogenesis, storage reserve accumulation and the acquisition of desiccation tolerance [9,13-15]. Our *ZmAFL2*-RNAi, *ZmAFL4*-RNAi and *ZmAFL5/6*-RNAi lines seemed to be unaltered in their acquisition of desiccation tolerance and did not show any alteration in their capacity for organogenesis during the transformation process. Similarly in rice, RNAi lines for *OsLFL1*, the likely ortholog of *ZmAFL2* and *FUS3*, also exhibited normal grain development [58]. The lack of visible phenotypes in cereals may be due to (i) residual activity of the targeted AFL transcription factors due to incomplete knockdown by RNAi, (ii)

functional redundancy between AFL genes with overlapping gene activity patterns such as *ZmAFL4* and *ZmAFL6*, and/or (iii) modification of AFL function between Arabidopsis and cereals.

*4.2. Zm*AFL4 *is involved in starch metabolism*

Knocking down *ZmAFL4* gene activity causes considerable metabolic alterations in the maize kernel, highlighting ZmAFL4 as a regulator of carbon metabolism. However, whether the observed decrease in content of metabolites such as short organic acids, free amino acids and monosaccharides is linked to the transient decrease in the content of the reserve compound starch needs to be established. Starch biosynthesis initiates in the kernel at 10 DAP when starch grains become visible and the activity of AGPase starts to increase until it reaches a maximum at around 30 DAP [60,61]. The gene activity pattern of *ZmAFL4* mirrors this trend not only in the endosperm, it also in pollen, another tissue showing high levels of starch accumulation.

Although *ZmAFL4* is essentially active in the endosperm, knocking down its gene activity also causes metabolic changes in the embryo, and particularly a significant decrease in G3P, a key component for fatty acid biosynthesis. Since movement of the ZmAFL4 protein over several layers is improbable, the changes seen in the embryo are most likely adjustments to altered metabolite fluxes between the embryo and endosperm. In conclusion, in the maize kernel ZmAFL4 acts primarily in the endosperm where it may be exclusively linked with carbon metabolism.

*4.3. Role of Zm*AFL *genes in maize seed development*

In Arabidopsis at least one *AFL* gene, *LEC2*, is directly involved in the regulation of genes coding for seed storage proteins [17,26] and oleosins [25]. In maize, protein is mainly stored in endosperm where 70% of the protein content is composed of prolamins such as zeins [62]. No notable modification of total protein content or in the activity of *Zein* genes was observed in *ZmAFL4*-RNAi endosperm. Whereas these results seem to exclude ZmAFL4 from the regulation of this major class of seed protein genes, they do not allow the drawing of general conclusions regarding the regulation of embryo and endosperm seed protein genes by AFL transcription factors. With regard to oleosins, *trans*-activation of the Zm*Ole2* promoter by ZmVp1 and ZmAFL4 demonstrated that maize AFL transcription factors have the potential to directly regulate genes coding for these structural proteins of oil bodies, and indicated that ZmVp1 may play additional roles to those already established in ABA signaling [31].

The detailed functions of *ZmAFL4* as well as of *ZmAFL2*, *ZmVp1*, *ZmAFL5* and *ZmAFL6* remain to be fully elucidated. Since all RNAi transgenic lines used in this study had at least some residual gene activity, the identification of true knockout plants for each *ZmAFL* gene as well as the generation of multiple mutants will be essential for the investigation of ZmAFL function and the identification of direct target genes. In addition, in Arabidopsis LEC2, ABI3 and FUS3 associate with LEC1, which is also required for reserve accumulation and seed development. Together these proteins constitute the LAFL network [10,63]. Consequently the analysis of interactions between ZmAFLs and ZmLEC1 will also be of the utmost importance for a comprehensive understanding of the control of reserve accumulation in the maize kernel.

Acknowledgments

We are very grateful to Dr Gwyneth Ingram for her helpful comments. We thank Alexis Lacroix, Yannick Rasmus and Priscilla Angelot for maize culture, as well as Hervé Leyral and Isabelle Desbouchages for excellent technical assistance. We acknowledge the "Plateforme Analyse Génétique" (Bariza Blanquier) of the SFR Biosciences Gerland-Lyon Sud (UMS344/US8) and the "Plateforme Métabolisme-Métabolome" of the IFR87 Plant & Environment for metabolomics analysis. The work was supported in part by the project AMAIZING ANR-10-BTBR-03 to PR. A PhD fellowship of the "Conseil Régional Rhône-Alpes" was attributed to AG.

References

[1] V. Vernoud, M. Hajduch, A.-S. Khaled, N. Depège, P.M. Rogowsky, Maize embryogenesis, Maydica 50 (2005) 469-483.

[2] N. Sreenivasulu, U. Wobus, Seed-development programs: a systems biology-based comparison between dicots and monocots, Annu Rev Plant Biol 64 (2013) 189-217.

[3] X. Liu, J. Fu, D. Gu, W. Liu, T. Liu, Y. Peng, J. Wang, G. Wang, Genome-wide analysis of gene expression profiles during the kernel development of maize (Zea mays L.), Genomics 91 (2008) 378-387.

[4] R.S. Sekhon, C.N. Hirsch, K.L. Childs, M.W. Breitzman, P. Kell, S. Duvick, E.P. Spalding, C.R. Buell, N. de Leon, S.M. Kaeppler, Phenotypic and Transcriptional Analysis of Divergently Selected Maize Populations Reveals the Role of Developmental Timing in Seed Size Determination, Plant Physiol 165 (2014) 658-669.

[5] X. Lu, D. Chen, D. Shu, Z. Zhang, W. Wang, C. Klukas, L.L. Chen, Y. Fan, M. Chen, C. Zhang, The differential transcription network between embryo and endosperm in the early developing maize seed, Plant Physiol 162 (2013) 440-455.

[6] G. Li, D. Wang, R. Yang, K. Logan, H. Chen, S. Zhang, M.I. Skaggs, A. Lloyd, W.J. Burnett, J.D. Laurie, B.G. Hunter, J.M. Dannenhoffer, B.A. Larkins, G.N. Drews, X. Wang, R. Yadegari, Temporal patterns of gene expression in developing maize endosperm identified through transcriptome sequencing, Proc Natl Acad Sci U S A 111 (2014) 7582-7587.

[7] J. Giraudat, B.M. Hauge, C. Valon, J. Smalle, F. Parcy, H.M. Goodman, Isolation of the Arabidopsis ABI3 gene by positional cloning, Plant Cell 4 (1992) 1251-1261.

[8] H. Luerssen, V. Kirik, P. Herrmann, S. Misera, FUSCA3 encodes a protein with a conserved VP1/AB13-like B3 domain which is of functional importance for the regulation of seed maturation in Arabidopsis thaliana, Plant J 15 (1998) 755-764.

[9] S.L. Stone, L.W. Kwong, K.M. Yee, J. Pelletier, L. Lepiniec, R.L. Fischer, R.B. Goldberg, J.J. Harada, LEAFY COTYLEDON2 encodes a B3 domain transcription factor that induces embryo development, Proc Natl Acad Sci U S A 98 (2001) 11806-11811.

[10] R.W. Kwong, A.Q. Bui, H. Lee, L.W. Kwong, R.L. Fischer, R.B. Goldberg, J.J. Harada, LEAFY COTYLEDON1-LIKE defines a class of regulators essential for embryo development, Plant Cell 15 (2003) 5-18.

[11] M. Santos Mendoza, B. Dubreucq, M. Miquel, M. Caboche, L. Lepiniec, LEAFY COTYLEDON 2 activation is sufficient to trigger the accumulation of oil and seed specific mRNAs in Arabidopsis leaves, FEBS Lett 579 (2005) 4666-4670.

[12] S.A. Braybrook, J.J. Harada, LECs go crazy in embryo development, Trends Plant Sci 13 (2008) 624-630.

---

[13] D.W. Meinke, L.H. Franzmann, T.C. Nickle, E.C. Yeung, Leafy Cotyledon Mutants of Arabidopsis, Plant Cell 6 (1994) 1049-1064.

[14] K. Keith, M. Kraml, N.G. Dengler, P. McCourt, fusca3: A Heterochronic Mutation Affecting Late Embryo Development in Arabidopsis, Plant Cell 6 (1994) 589-600.

[15] E. Nambara, K. Keith, P. Mccourt, S. Naito, A Regulatory Role for the Abi3 Gene in the Establishment of Embryo Maturation in Arabidopsis-Thaliana, Development 121 (1995) 629-636.

[16] G. Barthole, A. To, C. Marchive, V. Brunaud, L. Soubigou-Taconnat, N. Berger, B. Dubreucq, L. Lepiniec, S. Baud, MYB118 represses endosperm maturation in seeds of Arabidopsis, Plant Cell 26 (2014) 3519-3537.

[17] T. Kroj, G. Savino, C. Valon, J. Giraudat, F. Parcy, Regulation of storage protein gene expression in Arabidopsis, Development 130 (2003) 6065-6073.

[18] A. To, C. Valon, G. Savino, J. Guilleminot, M. Devic, J. Giraudat, F. Parcy, A network of local and redundant gene regulation governs Arabidopsis seed maturation, Plant Cell 18 (2006) 1642-1651.

[19] G. Monke, L. Altschmied, A. Tewes, W. Reidt, H.P. Mock, H. Baumlein, U. Conrad, Seed-specific transcription factors ABI3 and FUS3: molecular interaction with DNA, Planta 219 (2004) 158-166.

[20] F. Wang, S.E. Perry, Identification of direct targets of FUSCA3, a key regulator of Arabidopsis seed development, Plant Physiol 161 (2013) 1251-1264.

[21] Y. Kagaya, R. Toyoshima, R. Okuda, H. Usui, A. Yamamoto, T. Hattori, LEAFY COTYLEDON1 controls seed storage protein genes through its regulation of FUSCA3 and ABSCISIC ACID INSENSITIVE3, Plant and Cell Physiology 46 (2005) 399-406.

[22] J. Mu, H. Tan, Q. Zheng, F. Fu, Y. Liang, J. Zhang, X. Yang, T. Wang, K. Chong, X.J. Wang, J. Zuo, LEAFY COTYLEDON1 is a key regulator of fatty acid biosynthesis in Arabidopsis, Plant Physiol 148 (2008) 1042-1054.

[23] S.L. Stone, S.A. Braybrook, S.L. Paula, L.W. Kwong, J. Meuser, J. Pelletier, T.F. Hsieh, R.L. Fischer, R.B. Goldberg, J.J. Harada, Arabidopsis LEAFY COTYLEDON2 induces maturation traits and auxin activity: Implications for somatic embryogenesis, Proc Natl Acad Sci U S A 105 (2008) 3151-3156.

[24] S. Baud, M.S. Mendoza, A. To, E. Harscoet, L. Lepiniec, B. Dubreucq, WRINKLED1 specifies the regulatory action of LEAFY COTYLEDON2 towards fatty acid metabolism during seed maturation in Arabidopsis, Plant J 50 (2007) 825-838.

[25] N.Y. Che, Y. Yang, Y.D. Li, L.L. Wang, P. Huang, Y. Gao, C.C. An, Efficient LEC2 activation of OLEOSIN expression requires two neighboring RY elements on its promoter, Sci China Ser C 52 (2009) 854-863.

[26] S.A. Braybrook, S.L. Stone, S. Park, A.Q. Bui, B.H. Le, R.L. Fischer, R.B. Goldberg, J.J. Harada, Genes directly regulated by LEAFY COTYLEDON2 provide insight into the control of embryo maturation and somatic embryogenesis, Proc Natl Acad Sci U S A 103 (2006) 3468-3473.

[27] D.R. McCarty, T. Hattori, C.B. Carson, V. Vasil, M. Lazar, I.K. Vasil, The Viviparous-1 developmental gene of maize encodes a novel transcriptional activator, Cell 66 (1991) 895-905.

[28] F. Parcy, C. Valon, M. Raynal, P. Gaubier-Comella, M. Delseny, J. Giraudat, Regulation of gene expression programs during Arabidopsis seed development: roles of the ABI3 locus and of endogenous abscisic acid, Plant Cell 6 (1994) 1567-1582.

[29] M. Suzuki, C.Y. Kao, D.R. McCarty, The conserved B3 domain of VIVIPAROUS1 has a cooperative DNA binding activity, Plant Cell 9 (1997) 799-807.

[30] M. Suzuki, C.Y. Kao, S. Cocciolone, D.R. McCarty, Maize VP1 complements Arabidopsis abi3 and confers a novel ABA/auxin interaction in roots, Plant J 28 (2001) 409-418.

[31] M. Suzuki, M.G. Ketterling, Q.B. Li, D.R. McCarty, Viviparous1 alters global gene expression patterns through regulation of abscisic acid signaling, Plant Physiol 132 (2003) 1664-1677.

[32] B. Pouvreau, S. Baud, V. Vernoud, V. Morin, C. Py, G. Gendrot, J.P. Pichon, J. Rouster, W. Paul, P.M. Rogowsky, Duplicate maize Wrinkled1 transcription factors activate target genes involved in seed oil biosynthesis, Plant Physiol 156 (2011) 674-686.

[33] B. Shen, W.B. Allen, P. Zheng, C. Li, K. Glassman, J. Ranch, D. Nubel, M.C. Tarczynski, Expression of ZmLEC1 and ZmWRI1 increases seed oil production in maize, Plant Physiol 153 (2010) 980-987.

[34] G. Barthole, L. Lepiniec, P.M. Rogowsky, S. Baud, Controlling lipid accumulation in cereal grains, Plant Sci 185-186 (2012) 33-39.

[35] H. Rolletschek, K. Koch, U. Wobus, L. Borisjuk, Positional cues for the starch/lipid balance in maize kernels and resource partitioning to the embryo, Plant J 42 (2005) 69-83.

[36] J.C. Wang, H. Xu, Y. Zhu, Q.Q. Liu, X.L. Cai, OsbZIP58, a basic leucine zipper transcription factor, regulates starch biosynthesis in rice endosperm, J Exp Bot 64 (2013) 3453-3466.

[37] J. Vicente-Carbajosa, S.P. Moose, R.L. Parsons, R.J. Schmidt, A maize zinc-finger protein binds the prolamin box in zein gene promoters and interacts with the basic leucine zipper transcriptional activator Opaque2, Proc Natl Acad Sci U S A 94 (1997) 7685-7690.

[38] M. Javelle, V. Vernoud, N. Depege-Fargeix, C. Arnould, D. Oursel, F. Domergue, X. Sarda, P.M. Rogowsky, Overexpression of the epidermis-specific homeodomain-leucine zipper IV transcription factor Outer Cell Layer1 in maize identifies target genes involved in lipid metabolism and cuticle biosynthesis, Plant Physiol 154 (2010) 273-286.

[39] Q. Xing, A. Creff, A. Waters, H. Tanaka, J. Goodrich, G.C. Ingram, ZHOUPI controls embryonic cuticle formation via a signalling pathway involving the subtilisin protease ABNORMAL LEAF-SHAPE1 and the receptor kinases GASSHO1 and GASSHO2, Development 140 (2013) 770-779.

[40] G. Tcherkez, A. Mahe, P. Gauthier, C. Mauve, E. Gout, R. Bligny, G. Cornic, M. Hodges, In folio respiratory fluxomics revealed by 13C isotopic labeling and H/D isotope effects highlight the noncyclic nature of the tricarboxylic acid "cycle" in illuminated leaves, Plant Physiol 151 (2009) 620-630.

[41] J.H. Hendriks, A. Kolbe, Y. Gibon, M. Stitt, P. Geigenberger, ADP-glucose pyrophosphorylase is activated by posttranslational redox-modification in response to light and to sugars in leaves of Arabidopsis and other plant species, Plant Physiol 133 (2003) 838-849.

[42] J. Thevenin, C. Dubos, W. Xu, J. Le Gourrierec, Z. Kelemen, F. Charlot, F. Nogue, L. Lepiniec, B. Dubreucq, A new system for fast and quantitative analysis of heterologous gene expression in plants, New Phytol 193 (2012) 504-512.

[43] D.R. McCarty, C.B. Carson, P.S. Stinard, D.S. Robertson, Molecular Analysis of viviparous-1: An Abscisic Acid-Insensitive Mutant of Maize, Plant Cell 1 (1989) 523-532.

[44] K. Swaminathan, K. Peterson, T. Jack, The plant B3 superfamily, Trends in Plant Science 13 (2008) 647-655.

[45] F. Murat, J.H. Xu, E. Tannier, M. Abrouk, N. Guilhot, C. Pont, J. Messing, J. Salse, Ancestral grass karyotype reconstruction unravels new mechanisms of genome shuffling as a source of plant evolution, Genome Res 20 (2010) 1545-1557.

[46] S.J. Emrich, L. Li', T.J. Wen, M.D. Yandeau-Nelson, Y. Fu, L. Guo, H.H. Chou, S. Aluru, D.A. Ashlock, P.S. Schnable, Nearly identical paralogs: Implications for maize (Zea mays L.) genome evolution, Genetics 175 (2007) 429-439.

[47] F.Y. Peng, R.J. Weselake, Genome-wide identification and analysis of the B3 superfamily of transcription factors in Brassicaceae and major crop plants, Theor Appl Genet 126 (2013) 1305-1319.

[48] H.H. Marella, Y. Sakata, R.S. Quatrano, Characterization and functional analysis of ABSCISIC ACID INSENSITIVE3-like genes from Physcomitrella patens, Plant J 46 (2006) 1032-1044.

[49] Y. Li, K. Jin, Z. Zhu, J. Yang, Stepwise origin and functional diversification of the AFL subfamily B3 genes during land plant evolution, J Bioinform Comput Biol 8 Suppl 1 (2010) 33-45.

[50] M. Schmid, T.S. Davison, S.R. Henz, U.J. Pape, M. Demar, M. Vingron, B. Scholkopf, D. Weigel, J.U. Lohmann, A gene expression map of Arabidopsis thaliana development, Nat Genet 37 (2005) 501-506.

[51] D. Winter, B. Vinegar, H. Nahal, R. Ammar, G.V. Wilson, N.J. Provart, An "Electronic Fluorescent Pictograph" browser for exploring and analyzing large-scale biological data sets, Plos One 2 (2007) e718.

[52] B.H. Le, C. Cheng, A.Q. Bui, J.A. Wagmaister, K.F. Henry, J. Pelletier, L. Kwong, M. Belmonte, R. Kirkbride, S. Horvath, G.N. Drews, R.L. Fischer, J.K. Okamuro, J.J. Harada, R.B. Goldberg, Global analysis of gene activity during Arabidopsis seed development and identification of seed-specific transcription factors, Proc Natl Acad Sci U S A 107 (2010) 8063-8070.

[53] H. Jia, D.R. McCarty, M. Suzuki, Distinct roles of LAFL network genes in promoting the embryonic seedling fate in the absence of VAL repression, Plant Physiol 163 (2013) 1293-1305.

[54] R.S. Sekhon, H. Lin, K.L. Childs, C.N. Hansey, C.R. Buell, N. de Leon, S.M. Kaeppler, Genome-wide atlas of transcription during maize development, Plant J 66 (2011) 553-563.

[55] K. Rikiishi, M. Maekawa, Seed maturation regulators are related to the control of seed dormancy in wheat (Triticum aestivum L.), Plos One 9 (2014) e107618.

[56] R.S. McKibbin, M.D. Wilkinson, P.C. Bailey, J.E. Flintham, L.M. Andrew, P.A. Lazzeri, M.D. Gale, J.R. Lenton, M.J. Holdsworth, Transcripts of Vp-1 homeologues are misspliced in modern wheat and ancestral species, Proc Natl Acad Sci U S A 99 (2002) 10203-10208.

[57] X. Cao, L.M. Costa, C. Biderre-Petit, B. Kbhaya, N. Dey, P. Perez, D.R. McCarty, J.F. Gutierrez-Marcos, P.W. Becraft, Abscisic acid and stress signals induce Viviparous1 expression in seed and vegetative tissues of maize, Plant Physiol 143 (2007) 720-731.

[58] L.T. Peng, Z.Y. Shi, L. Li, G.Z. Shen, J.L. Zhang, Overexpression of transcription factor OsLFL1 delays flowering time in Oryza sativa, J Plant Physiol 165 (2008) 876-885.

[59] T. Kobayashi, Y. Ogo, M.S. Aung, T. Nozoye, R.N. Itai, H. Nakanishi, T. Yamakawa, N.K. Nishizawa, The spatial expression and regulation of transcription factors IDEF1 and IDEF2, Ann Bot 105 (2010) 1109-1117.

[60] J.L. Prioul, E. Jeannette, A. Reyss, N. Gregory, M. Giroux, L.C. Hannah, M. Causse, Expression of ADP-glucose pyrophosphorylase in maize (Zea mays L.) grain and source leaf during grain filling, Plant Physiol 104 (1994) 179-187.

[61] W.L. Charlton, C.L. Keen, C. Merriman, P. Lynch, A.J. Greenland, H.G. Dickinson, Endosperm Development in Zea-Mays - Implication of Gametic Imprinting and Paternal Excess in Regulation of Transfer Layer Development, Development 121 (1995) 3089-3097.

[62] B.C. Gibbon, B.A. Larkins, Molecular genetic approaches to developing quality protein maize, Trends Genet 21 (2005) 227-233.

[63] T. Lotan, M. Ohto, K.M. Yee, M.A.L. West, R. Lo, R.W. Kwong, K. Yamagishi, R.L. Fischer, R.B. Goldberg, J.J. Harada, Arabidopsis LEAFY COTYLEDON1 is sufficient to induce embryo development in vegetative cells, Cell 93 (1998) 1195-1205.

Tables

Table 1: Zm*AFL* genes in the maize genome.

Gene Name	Gene model[1]	Protein size	Sorghum putative ortholog	Sorghum % identity[2]	Rice putative ortholog	Rice % identity[2]	Arabidopsis putative ortholog	Arabidopsis % identity[2]	% identity of B3 domain
ZmAFL1	GRMZM2G008356	957	Sb02g036430	91	LOC_Os07g37610	79	AT4G21550 (*VAL3*)	36	66
ZmAFL2	GRMZM2G035701	292	Sb03g032730	78	LOC_Os01g51610 (*LFL1*)	57	AT3G26790 (*FUS3*)	25	62
ZmAFL3 (ZmVp1)	GRMZM2G133398	691	Sb03g043480	81	LOC_Os01g68370 (*OsVP1*)	60	AT3G24650 (*ABI3*)	27	87
ZmAFL4	GRMZM2G149940	439	Sb06g032870	83	LOC_Os04g58000	37	AT1G28300 (*LEC2*)	17	54
ZmAFL5	GRMZM2G405699	313	Sb07g000220	81	LOC_Os08g01090 (*IDEF1*)	51		32	57
ZmAFL6	GRMZM2G125596	369		91		56		31	57

[1]Gene models refer to AGPv3 for maize and maize protein size is indicated in amino acids.

[2]% identity refers to protein identity between maize and other species.

Table 2: Metabolites with significantly altered content in *ZmAFL4*-RNAi endosperm from the strongest transformation event.

Metabolite	Trend	Average[1] *ZmAFL4*-RNAi endosperm	Average[1] WT endosperm	SD *ZmAFL4*-RNAi endosperm	SD WT endosperm	Ratio *ZmAFL4*-RNAi /WT	p-value[2] ANOVA
Glyceric acid	DOWN	0.13	0.26	0.00	0.01	0.50	0.0005613
2,4-dihydroxybutanoic acid	DOWN	0.69	1.10	0.00	0.05	0.62	0.003038
Phosphoric acid	DOWN	81.00	113.15	0.19	4.04	0.72	0.003185
Ribose	DOWN	2.27	3.27	0.07	0.14	0.69	0.005397
Aspartic acid	DOWN	51.39	79.57	1.11	4.32	0.65	0.006021
Malic acid	DOWN	36.84	49.12	0.91	2.06	0.75	0.00869
Glycine	DOWN	0.82	1.21	0.06	0.07	0.68	0.01651
Maleic acid	DOWN	0.57	1.00	0.08	0.08	0.57	0.02185
Succinic acid	DOWN	5.80	8.23	0.09	0.64	0.71	0.02579
Glucose	DOWN	35.29	61.50	2.81	6.97	0.57	0.02969
Cysteine	UP	1.11	0.62	0.14	0.10	1.78	0.04024
Fructose	DOWN	202.98	308.40	14.24	33.19	0.66	0.04628
Xylose	DOWN	0.18	0.31	0.04	0.03	0.59	0.04868

[1] Average of three biological replicates (each replicate representing a different plant, from which at least 50 endosperms and 50 embryos (see Supplemental Table 2) of the same ear were pooled) expressed in arbitrary units representing peak areas.

[2] In order of statistical significance of the difference between transgenic and wild-type endosperm established by one-factor ANOVA.

Figure captions

Fig. 1: Partial phylogenetic tree of the B3 family: the AFL clade.

(A) A maximum likelihood phylogenetic tree was generated for the "ABI3/VP1" branch of the B3 family as defined by Peng and Weselake (2013) with PHYML software using the WAG amino acids substitution model. Nod values are bootstrap values expressed as percentages and based on 1000 replicates. AFL proteins from *Arabidopsis thaliana* and *Zea mays* are colored in green and red, respectively. The entire phylogenetic tree is available in Supplementary Fig. S2. (B) Schematic view of conserved domains in maize and Arabidopsis AFL proteins as originally defined by Giraudat *et al.* (1992).

Fig. 2: Gene activity patterns of Zm*AFL* genes.

Relative mRNA levels were determined by qRT-PCR in different maize organs (A), during kernel development (B) and in dissected kernel compartments (C). Error bars correspond to the standard deviation calculated from technical triplicates on pools of tissues from more than two different plants (A, B) or to biological replicates (each representing pools of dissected tissues from a single plant) executed in technical triplicates (C).

Fig. 3: ZmAFL4 contributes to the regulation of starch accumulation.

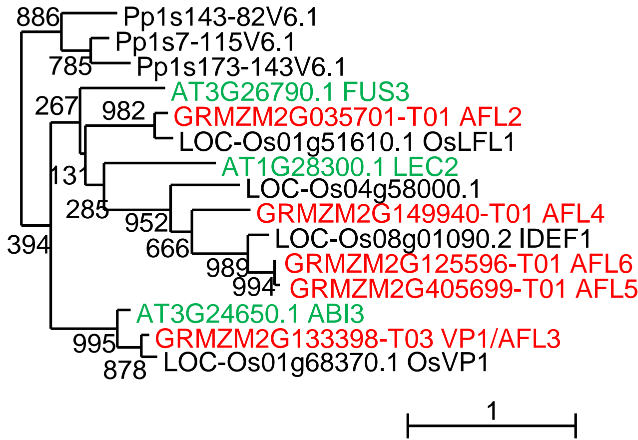
(A) The expression level of *ZmAFL4* in *ZmAFL4*-RNAi endosperm was established at 20 DAP by qRT-PCR in two independent *ZmAFL4*-RNAi transformation events and wild-type siblings. Error bars correspond to the standard deviation calculated from biological triplicates for each event. Each replicate represents a different plant, from which at least 50 endosperms of the same ear were pooled. (B, C) Starch content was determined in 20 DAP endosperm (B)

and in mature kernels (C) from the same plants as in (A). Error bars correspond to the standard deviation calculated from biological triplicates on pools of endosperm or mature kernels (Student's test value: * $p<0.05$). FW: Fresh weight.

Fig. 4: Transcriptional activation by ZmAFL transcription factors.

Trans-activation of the Zm*OLE2* promoter (GRMZM2G096435) measured in *Physcomitrella patens* protoplasts co-transfected with plasmids encoding the indicated AFL transcription factors alone or in combination. Mean activities were calculated from at least four biological repetitions. Error bars indicate standard deviation (Student's test significance: *, $p<0.05$; **, $p<0.01$; ***, $p <0.001$).

Figure 1



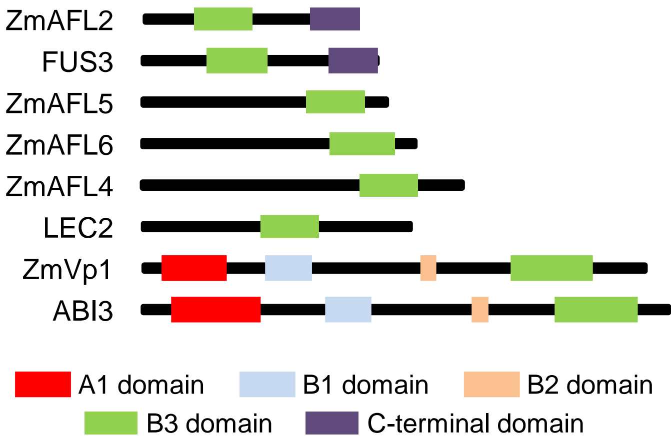
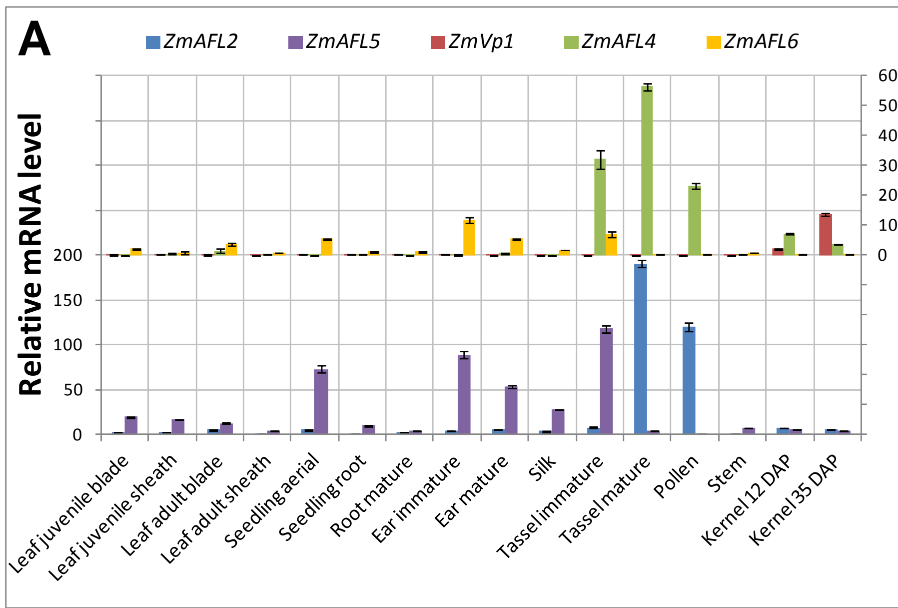
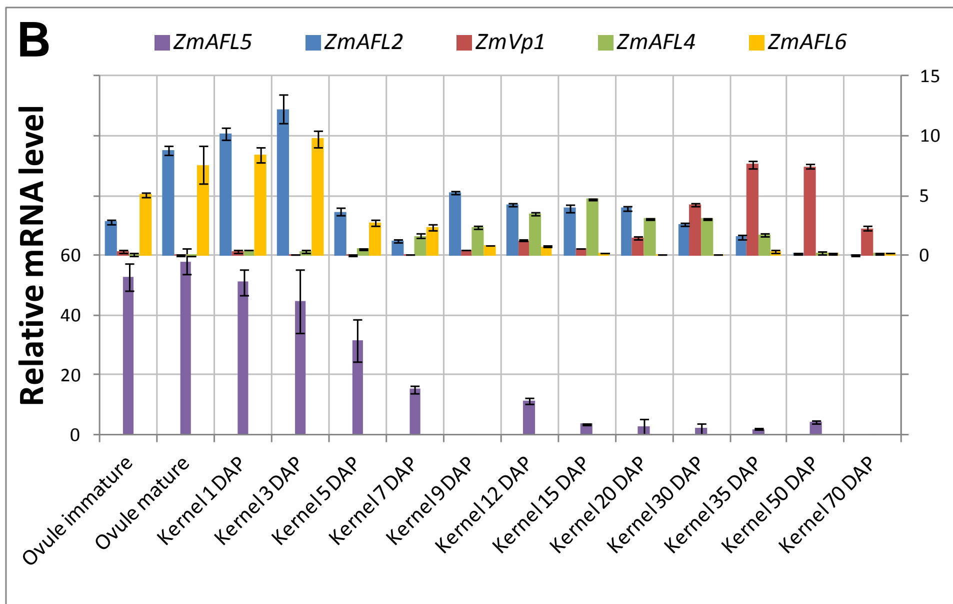


Fig. 1. Partial phylogenetic tree of the B3 family*:* the AFL clade.
(A) A maximum likelihood phylogenetic tree was generated for the "ABI3/VP1" branch of the B3 family as defined by Peng and Weselake (2013) with PHYML software using the WAG amino acids substitution model. Nod values are bootstrap values expressed as percentages and based on 1000 replicates. AFL proteins from *Arabidopsis thaliana* and *Zea mays* are colored in green and red, respectively. The entire phylogenetic tree is available in Supplementary Fig. S2. (B) Schematic view of conserved domains in maize and Arabidopsis AFL proteins as originally defined by Giraudat *et al.* (1992).

Figure 2





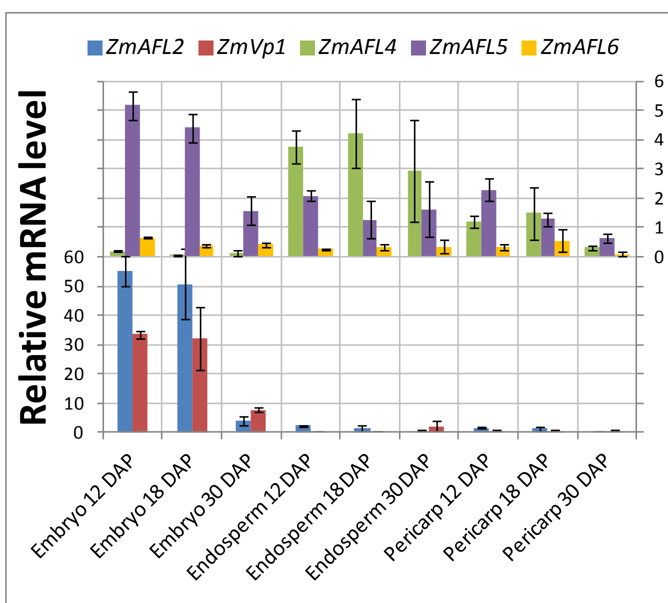


Fig. 2: Gene activity patterns of Zm*AFL* genes. Relative mRNA levels were determined by qRT-PCR in different maize organs (A), during kernel development (B) and in dissected kernel compartments (C). Error bars correspond to the standard deviation calculated from technical triplicates on pools of tissues from more than two different plants (A, B) and to biological replicates (each representing pools of dissected tissues from a single plant) executed in technical triplicates (C).

Figure 3

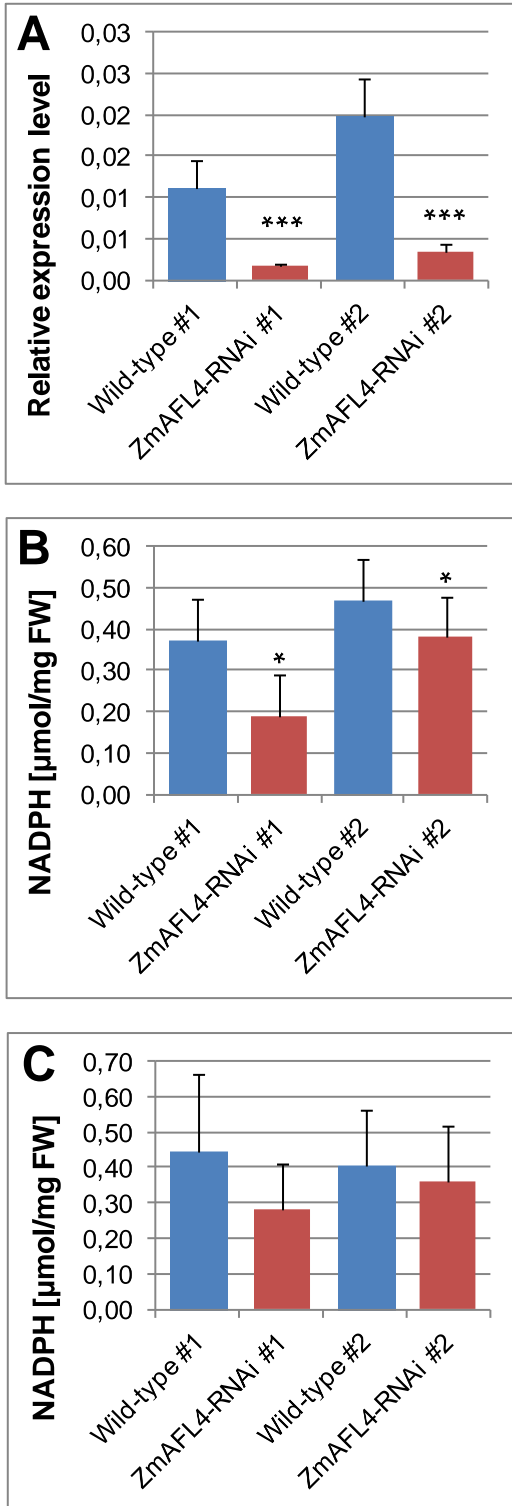


Fig. 3: ZmAFL4 contributes to the regulation of starch accumulation.
(A) The expression level of *ZmAFL4* in *ZmAFL4*-RNAi endosperm was established at 20 DAP by qRT-PCR in two independent *ZmAFL4*-RNAi transformation events and wild-type siblings. Error bars correspond to the standard deviation calculated from biological triplicates for each event. Each replicate represents a different plant, from which at least 50 endosperms of the same ear were pooled. (B, C) Starch content was determined in 20 DAP endosperm (B) and in mature kernels (C) from the same plants as in (A). Error bars correspond to the standard deviation calculated from biological triplicates on pools of endosperm or mature kernel (Student's test value: * $p<0.05$). FW: Fresh weight.

Figure 4

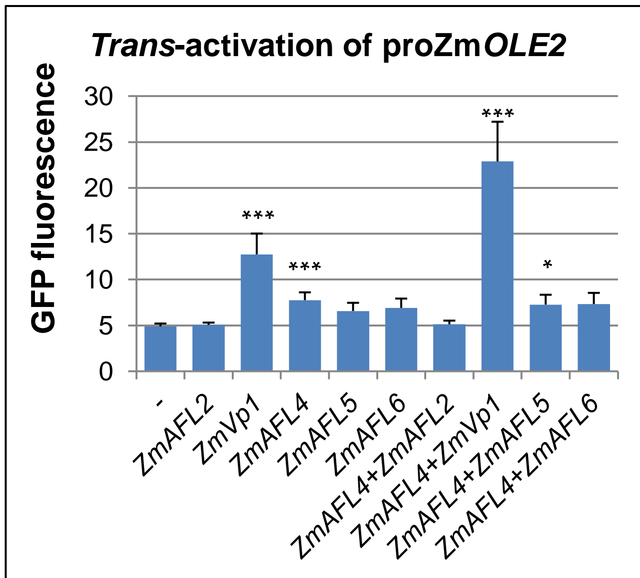


Fig. 4: Transcriptional activation by ZmAFL transcription factors**.**
Trans-activation of the Zm*OLE2* promoter (GRMZM2G096435) measured in *Physcomitrella patens* protoplasts co-transfected with plasmids coding for the indicated AFL transcription factors alone or in combination. Mean activities were calculated from at least four biological repetitions. Error bars indicate standard deviation (Student's test significance: *, p<0.05; **, p<0.01; ***, p <0.001).

Supplementary Figure S1

>ZmAFL2_(GRMZM2G035701_P01)
MAGITKRRTSPASTSSSSGDVLPQRVTRKRRSARRGPRSTARRPSAPPPMNELDLNTAALDPDHYATGL
RVLLQKELRNSDVSQLGRIVLPKKEAESYLPILMAKDGKSLCMHDLLNSQLWTFKYRYWFNNKSRMYVL
ENTGDYVKAHDLQQGDFIVIYKDDENNRFVIGAKKAGDEQTATVPQVHEHMHISAALPAPQAFHDYAGP
VAAEAGMLAIVPQGDEIFDGILNSLPEIPVANVRYSDFFDPFGDSMDMANPLSSSNNPSVNLATHFHDE
RIGSCSFPYPKSGPQM

>ZmAFL3_ZmVp1_(GRMZM2G133398_P03)
MEASAGSSPPHSQENPPEHGGDMGGAPAEEIGGEAADDFMFAEDTFPSLPDFPCLSSPSSSTFSSNSSS
NSSSAYTNTAGRAGGEPSEPASAGEGFDALDDIDQLLDFASLSMPWDSEPFPGVSMMLENAMSAPPQPV
GDGMSEEKAVPEGTTGGEEACMDASEGEELPRFFMEWLTSNRENISAEDLRGIRLRRSTIEAAAARLGG
GRQGTMQLLKLILTWVQNHHLQRKRPRDVMEEEAGLHVQLPSPVANPPGYEFPAGGQDMAAGGGTSWMP
HQQAFTPPAAYGGDAVYPSAAGQQYSFHQGPSTSSVVVNSQPFSPPPVGDMHGANMAWPQQYVPFPPPG
ASTGSYPMPQPFSPGFGGQYAGAGAGHLSVAPQRMAGVEASATKEARKKRMARQRRLSCLQQQRSQQLS
LGQIQASVHLQEPSPRSTHSGPVTPSAGGWGFWSPSSQQQVQNPLSKSNSSRAPPPSLEAAAVAPQTKP
APAGARQDDIHHRLAAASDKRQGAKADKNLRFLLQKVLKQSDVGSLGRIVLPKKEAEVHLPELKTRDGI
SIPMEDIGTSRVWNMRYRFWPNNKSRMYLLENTGEFVRSNELQEGDFIVIYSDVKSGKYLIRGVKVRPP
PAQEQGSGSSGGGKHRPLCPAGPGRAAAAGAPEDAVVDGVSGACKGRSPEGVRRVRQQGAGAMSQMAVS
I

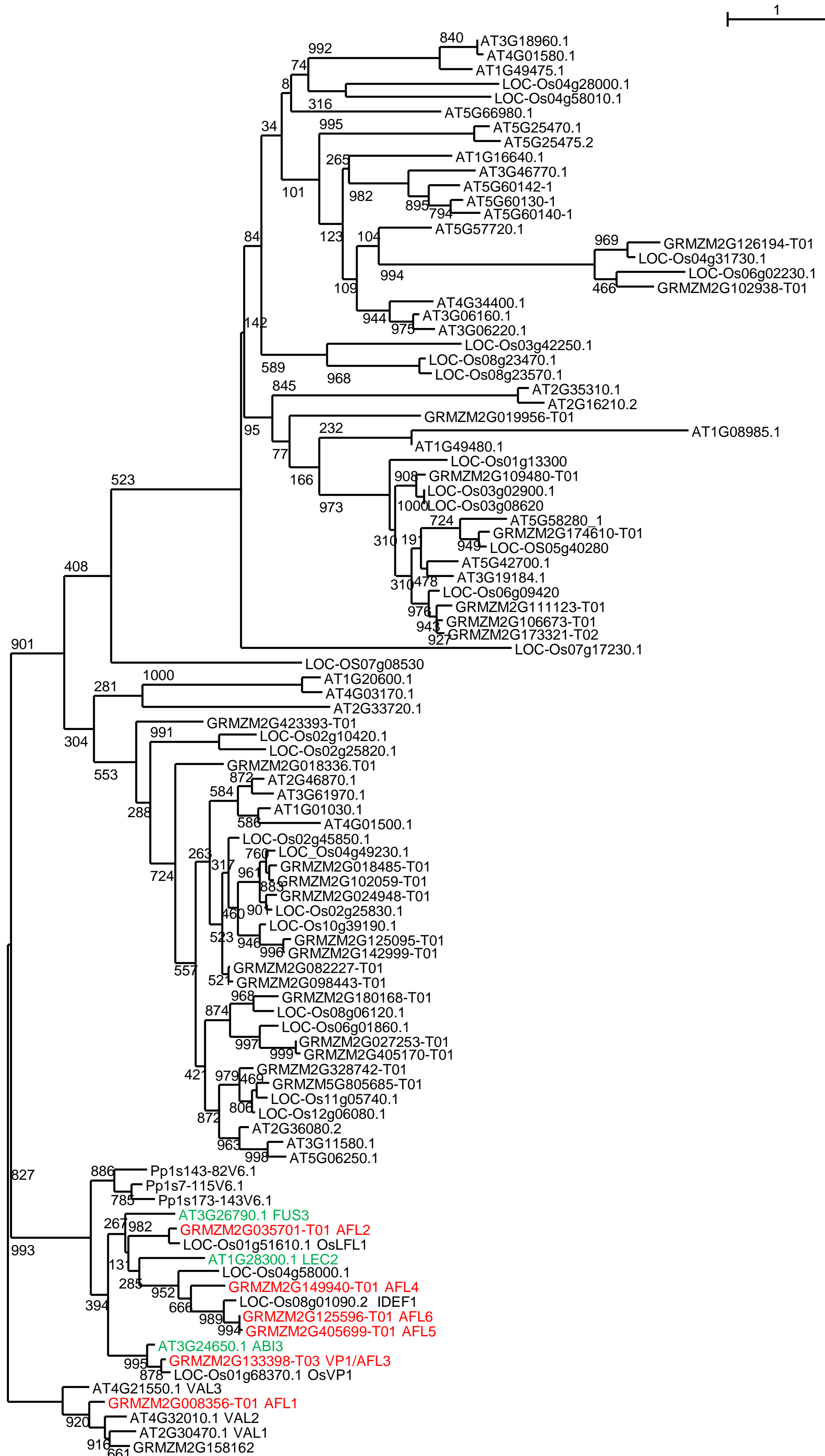
>ZmAFL4_(GRMZM2G149940_P01)
MANANGSSTGAGHSDLVRAISHEQHQAFMASVPRAAPGGVNVHHQQHFHQYPAGLIPAPVALPVHAPVS
SQTSPYSAQIAVPPPPPPLIASPDHRLHSLPPTGCYQLDYSPYGNAAAPSQQHTSAIRGFADWGTHSNA
LMSLAHATSFGNNGSSNINNNGLLHQNLSPYTTHTWTTTYVQRPYNTAVYAPATMNMLQTPPFHSNSHE
KESGAVFSNSFNMAPSVTPTSPFQLMSPSSTNYTSTQIFEETNNLEDTSRVFGGGDNESNNSEEPDPKP
AVEMEDLNQGNDHTSNKTANCQDYRMVLRKDLTNSDVGNIGRIVLPKKDAEPNLPILEDKDGLILEMDD
FELPVVWNFKYRYWPNNKSRMYILESTGEFVKRHGLQAKDILIIYRNKKSGRYVARAVKAEDIAPPECE
CVEAGNPREECGFSVSPSINKKIIT

>ZmAFL5_(GRMZM2G405699_P01mod)
MGQMGGPDGDGPHHQYHYQALLAAVQNPSQGLHVPLHAGAGAPAAGPGPRPGADADASSTHNANATPHS
QPPRAFTDWSASNSAFAAQPAPATTNTPFHYNLSQSYALWTHYMLNKNVSYSTYSTPHEPLRHTHIPDK
YSGCAFSLGFDSFTTMSLGPNICANMTPMERSISAKEPENSEDLPTVVRSSDEMDTRNSGDVRRDTVDT
LPESKQSHESCASVSNKFDSGEYQVILRKELTKSDVANSGRIVLPKKDAEAGLPPLVQGDPLILQMDDM
VLPIIWKFKYRFWPNNKSRMYILEAAGEFVKTHGLQAGDTLIIYKNSVPGKFIIRGEKSIQQTNP

>ZmAFL6_(GRMZM2G125596_P01)
MGQMGGPDGDGDGGAGPHHQYHYQALLAAVQNPSQGLHPFPLPFHVPLHAGAGAPAAGPGPGADADAST
HNANAAHHSQPPRGFTDWSASNSAFAAVASQPAPATTNTPFHYNLSQSYALWTHYMLNKNVSYSTYPTP
HEEHPHPLRHTHIQENPHPLRHTHIPDKDSGCASSLGFDSFTTMSLGPNICSHMTPMEGSISAKEPENS
EDLPAIVRSSDEMDTRNSGKVHRDTVGTLPESKQSHESCASVNNKFNSGEYQVILRKELTKSDVANSGR
IVLPKKDAEAGLPPLVQGDPLILQMDDMVLPIIWKFKYRFWPNNKSRMYILEAAGEFVKTHGLQAGDAL
IIYKNSVPGKFIIRGEKSIQQTNP

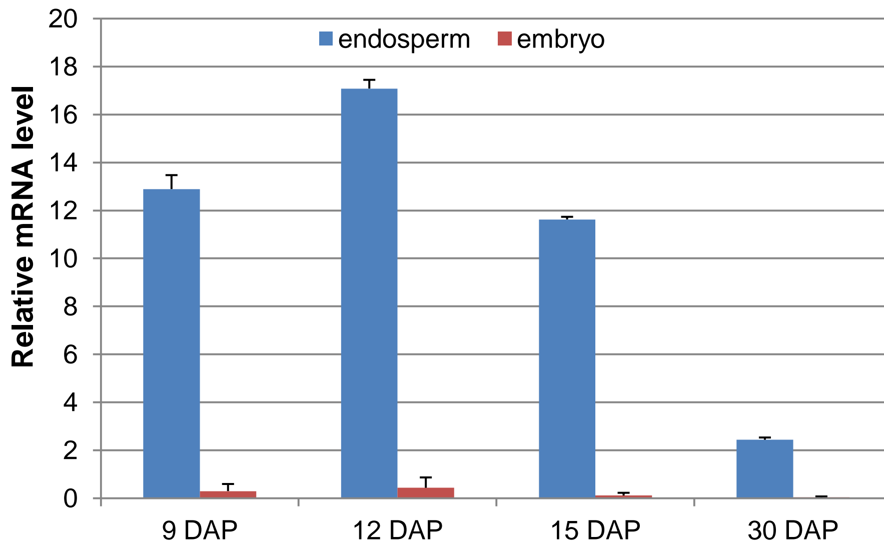
Supplementary Fig. S1. Deduced amino acid sequences of Zm*AFL* coding sequences.
Sequences were retrieved from http://www.maizesequence.org. In the case of multiple gene models, preference was given to the model best supported by EST coverage. The gene model of *ZmAFL5* was extended at its 3' end using the EST sequence DN219455.

Supplementary Figure S2



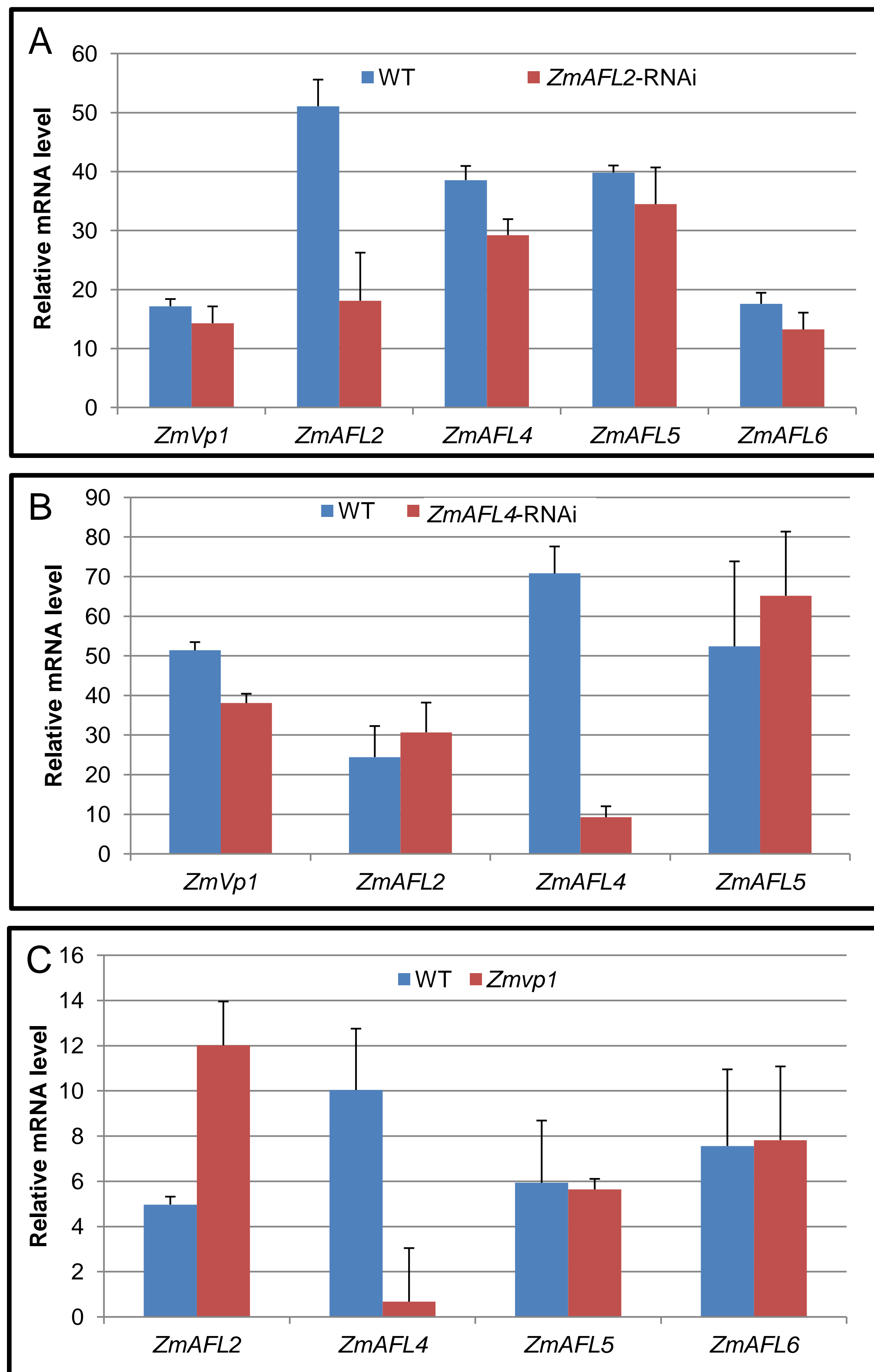
Supplementary Fig. S2. Partial phylogenetic tree of the B3 family of transcription factors.

Supplementary Figure S3



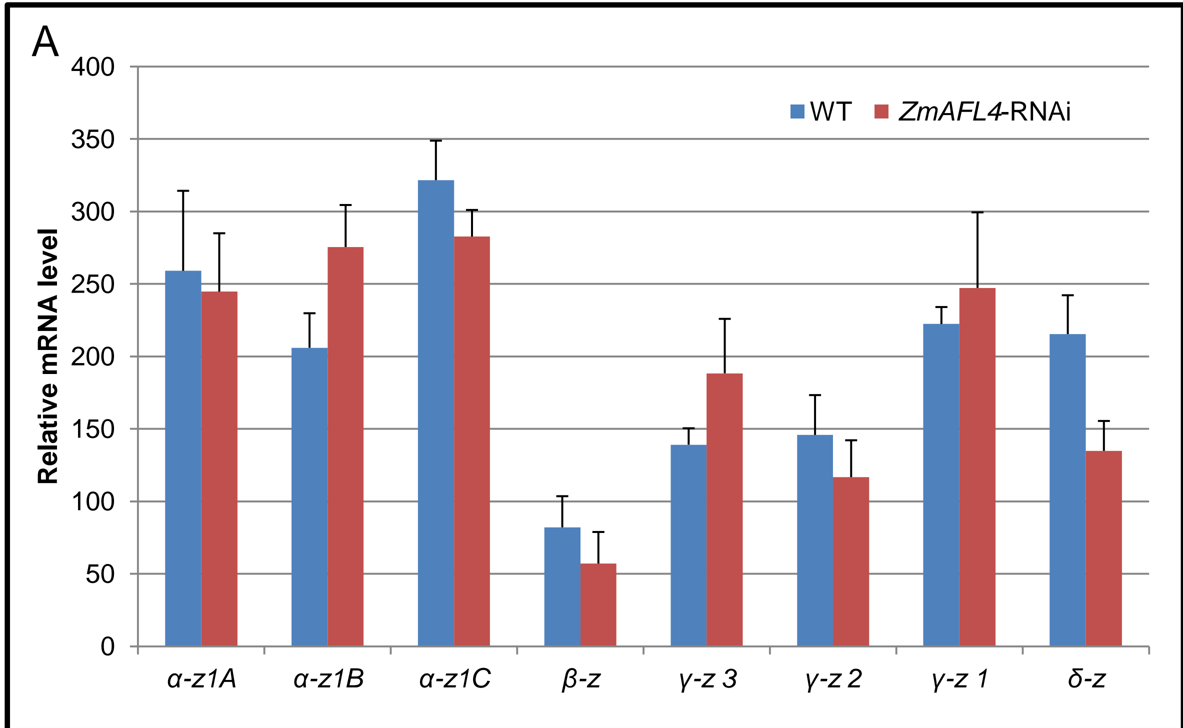
Supplementary Fig. S3. Zm*AFL4* gene activity in embryo and endosperm.
The gene activity pattern of *ZmAFL4* in dissected endosperms (blue) and embryos (red) during early development (9 and 12 DAP) and filling stage (15 and 30 DAP) was determined by qRT-PCR. Error bars correspond to the standard deviation calculated from technical triplicates on pools of tissues from more than two different plants.

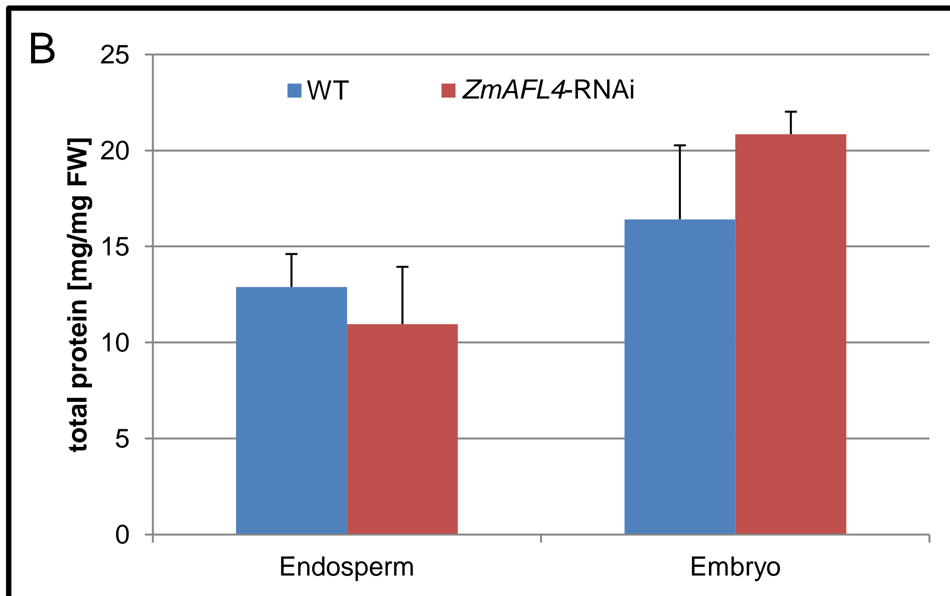
Supplementary Figure S4



Supplementary Fig. S4. Gene activity of *ZmAFL* genes in different genetic backgrounds. The mRNA levels of the indicated *ZmAFL* genes were determined by qRT-PCR at 12 DAP in *ZmAFL2*-RNAi and wild-type kernels (A), at 15 DAP in *ZmAFL4*-RNAi and wild-type kernels (B) and at 30 DAP in homozygous *Zmvp1* and wild-type kernels (C). Error bars correspond to the standard deviation calculated from biological replicates, each replicate consisting of 5 kernels from a single heterozygous plant (A,B). Error bars correspond to the standard deviation calculated from biological triplicates, each replicate consisting of a pool of 4 kernels from a single wild-type or homozygous *Zmvp1* plant (C).

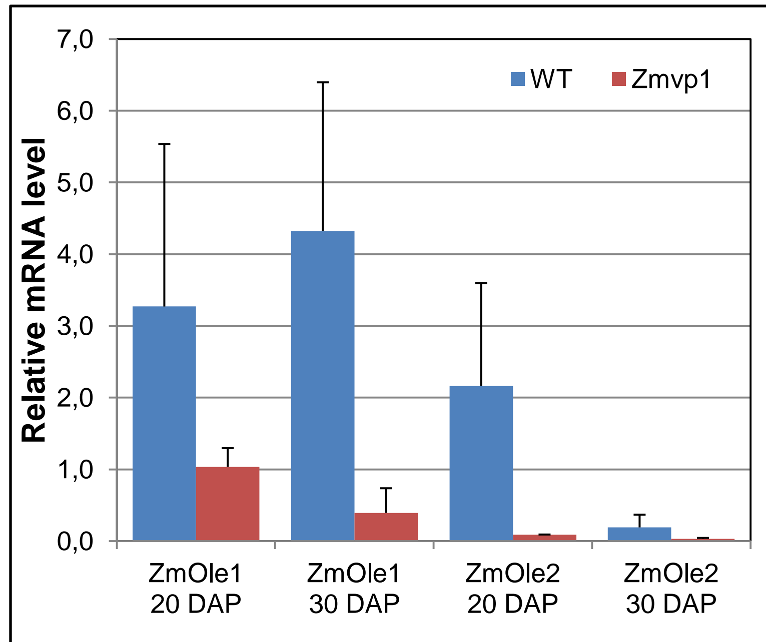
Supplementary Figure S5





Supplemental Figure S5. *Zein* gene activity and protein content in *ZmAFL4*-RNAi kernels.
(A) The mRNA level of *Zein* genes in *ZmAFL4*-RNAi endosperm at 20 DAP was established by qRT-PCR.
(B) Total protein content determined by Bradford dosage in *ZmAFL4*-RNAi endosperms and embryos. Error bars correspond to the standard deviation calculated from biological triplicates, each replicate representing a different plant, from which at least 50 endosperms or 50 embryos of the same ear were pooled.

Supplementary Figure S6



Supplementary Fig. S6. *ZmOle* gene activity in the *Zmvp1* mutant.
The gene activity pattern of *ZmOle1* and *ZmOle2* in wild-type (blue) and *Zmvp1* mutant kernels (red) was determined by qRT-PCR at 20 and 30 DAP. Error bars correspond to the standard deviation calculated from biological triplicates, each replicate consisting of a pool of 4 kernels from a single wild-type or homozygous *Zmvp1* plant.

Supplementary Table S1: Gene models encoding B3 domains in maize.

Gene model	% identity[1]	Source	Gene name
GRMZM2G133398_P02	89	NCBI, Peng et al,2013	*ZmAFL3 (ZmVp1)*
GRMZM2G035701_P01	59	NCBI	*ZmAFL2*
GRMZM2G405699_P01	59	NCBI	*ZmAFL5*
GRMZM2G149940_P01	58	NCBI, Peng et al,2013	*ZmAFL4*
GRMZM2G125596_P01	55	NCBI, Peng et al,2013	*ZmAFL6*
GRMZM2G018336_P01	40	NCBI	
GRMZM2G423393_P01	36	NCBI	
GRMZM2G082227_P01	35	NCBI	
GRMZM2G125095_P01	35	NCBI	
GRMZM2G142999_P01	35	NCBI	
GRMZM2G024948_P01	34	Peng et al,2013	
GRMZM2G098443_P01	34	NCBI	
GRMZM2G102059_P01	34	NCBI	
GRMZM5G805685_P01	34	NCBI, Peng et al,2013	
GRMZM2G027253_P01	33	NCBI	
GRMZM2G018485_P01	33	NCBI	
GRMZM2G180168_P01	33	NCBI	
GRMZM2G328742_P01	32	NCBI, Peng et al,2013	
GRMZM2G405170_P01	32	NCBI	
GRMZM2G102938_P01	30	Peng et al,2013	
GRMZM2G126194_P01	25	Peng et al,2013	
GRMZM2G106673_P01	24	Peng et al,2013	
GRMZM2G173321_P02	24	Peng et al,2013	
GRMZM2G111123_P01	23	Peng et al,2013	
GRMZM2G019956_P01	22	NCBI	
GRMZM2G174610_P01	21	Peng et al,2013	
GRMZM2G109480_P01	20	Peng et al,2013	

[1] identity with B3 domain of ABI3

Supplementary Table S2: Metabolites with significantly altered content in *ZmAFL4*-RNAi embryo from the strongest transformation event..

Metabolite	Trend	Average[1] *ZmAFL4*-RNAi embryo	Average[1] WT embryo	SD *ZmAFL4*-RNAi embryo	SD WT embryo	Ratio *ZmAFL4*-RNAi /WT	p-value[2] ANOVA
Glycerol-3-Phosphate	DOWN	4.12	5.63	0.78	0.66	0.73	0.001337
Valine	DOWN	47.33	58.87	1.23	0.92	0.80	0.002639
Phenylalanine	DOWN	1.32	2.12	0.00	0.19	0.62	0.01853
Arabinose	DOWN	0.34	0.40	0.02	0.10	0.84	0.02739
Leucine	DOWN	12.47	23.77	0.28	3.40	0.52	0.03604
Succinic acid	DOWN	0.88	1.48	0.02	0.19	0.59	0.04116
Erythritol	DOWN	69.06	117.22	6.85	23.17	0.59	0.04159
Tetradecanoic acid	DOWN	17.30	29.31	0.12	6.30	0.59	0.04641
Tyrosine	DOWN	8.96	14.48	0.31	1.83	0.62	0.04664
Ribose	DOWN	0.49	0.64	0.06	0.02	0.76	0.04887

[1] Average of three biological replicates (each replicate representing a different plant, from which at least 50 embryos and 50 endosperms (see Table 2) of the same ear were pooled) expressed in arbitrary units.

[2] In order of statistical significance of the difference between transgenic and wild-type embryo established by one-factor ANOVA.

.

Supplementary Table S3: Primer sequences used in this study.

Gene model	Gene name	Primer name	Primer sequence (5' to 3')	Use
GRMZM2G035701	*ZmAFL2*	ZmAFL2-RNAi-F	GGGGACAAGTTTGTACAAAAAAGCAGGCTCCTC AGCAAGGAGACTTCATCGTGAT	RNAi construct
		ZmAFL2-RNAi-R	GGGGACCACTTTGTACAAGAAAGCTGGGTCAGG TTGACCGAGGGGTTATTGG	
GRMZM2G149940	*ZmAFL4*	ZmAFL4-RNAi-F	GGGGACAAGTTTGTACAAAAAAGCAGGCTTCATC GCATCTCCAGACCACCGC	RNAi construct
		ZmAFL4-RNAi-R	GGGGACCACTTTGTACAAGAAAGCTGGGTGCAG TGTTATAAGGCCTCTGCACGT	
GRMZM2G405699	*ZmAFL5*	ZmAFL5-RNAi-F	GGGGACAAGTTTGTACAAAAAAGCAGGCTTAGG AGCAGTGATGAAATGGACACTAGAAAC	RNAi construct
		ZmAFL5-RNAi-R	GGGGACCACTTTGTACAAGAAAGCTGGGTGCCA TGTGTCTTCACGAATTCACCT	
GRMZM2G126010	*Actin*	actin-q-F	TACCCGATTGAGCATGGCA	Expression pattern
		actin-q-R	TCTTCAGGCGAAACACGGA	
GRMZM2G035701	*ZmAFL2*	ZmAFL2-q-F	CGACCCGTTCGGTGACTC	Expression pattern
		ZmAFL2-q-R	CACATCTGAGGCCCGGAT	
GRMZM2G133398	*ZmVp1*	VP1-q-F	AGGTCTCCGGAAGGCGT	Expression pattern
		VP1-q-R	AATATATGGCGGAGTCTGCTG	
GRMZM2G149940	*ZmAFL4*	ZmAFL4-q-F	GGAAACCCTAGAGAAGAGTGCGGCT	Expression pattern
		ZmAFL4-q-R	TTCTCGGTGTGCTCCTGCGC	
GRMZM2G405699	*ZmAFL5*	ZmAFL5-q-F	CCTTCTCGCGTAGTCCGTAG	Expression pattern
		ZmAFL5-q-R	GTCTCGACCCGTGGTAGC	
GRMZM2G125596	*ZmAFL6*	ZmAFL6-q-F	CAGGATCTATAAGGTAAAGGAAGTGG	Expression pattern
		ZmAFL6-q-R	AAATGTTGGTAAAATGAATGAGACAA	
GRMZ2G068506	*Brittle2*	BT2-q-F	CATACCTCAATCCTCAAGCTCA	Starch biosynthesis
		BT2-q-R	CGCTTCTTTGTCAAGGGGTA	
GRMZM2G348551	*Sugary 2*	SU2-q-F	GGCTGCTGAATGTTCTCCAT	Starch biosynthesis
		SU2-q-R	CCCATACCTTGGTACCACAAC	
GRMZM2G141399	*Dull1*	DU1-q-F	GAGGTTTGGTTTCGATGTTCA	Starch biosynthesis
		DU1-q-R	AGGCATCTCGTGGAACGTAA	
Not applicable [1]	*z1A*	αz1A-q-F	GCTCCTTGGTCTTTCTGCAA	Storage proteins
		αz1A-q-R	GGTAACTGCTGTAATAGGGCTGATG	

	z1B	αz1B-q-F	CCAGCCCTATCTTTGGTGCA	Storage proteins
		αz1B-q-R	TCAGTGCGGCCAATTGGTTA	
	z1C	αz1C-q-F	TTCCACAATGCTCACTTGCT	Storage proteins
		αz1C-q-R	GTTGTTGTAAGACGCTCGCC	
	z1D	αz1D-q-F	GTTGTTGTAAGACGCTCGCC	Storage proteins
		αz1D-q-R	AATGGTAGTAGCTGTTGTGC	
GRMZM2G138727[2]	*27-kD γ-zein*	γz1-q-F	AGTGTTGCCAGCAGCTCAG	Storage proteins
	(gz27)	γz1-q-R	TGGACTGGAGGACCAAGC	
GRMZM2G060429[2]	*16-kD γ-zein*	γz2-q-F	AGTGCGTCGAGTTCCTGAG	Storage proteins
	(gz16)	γz2-q-R	CACCGTATGTCGCCTGGTA	
GRMZM2G138689[2]	*50-kD γ-zein*	γz3-q-F	CACCAAGCAATCTACAACATGG	Storage proteins
	(gz50)	γz3-q-R	AAGATTGCCACCGCACTTT	
GRMZM2G086294[2]	*15-kD β-zein*	βz-q-F	TGTACGAGCCAGCTCTGATG	Storage proteins
	(bz15)	βz-q-R	GTAGCTGGGCAGCTGGTACT	
GRMZM2G100018[2]	*18-kD δ-zein*	δz-q-F	CAGCAACTGTTGGCCTCAC	Storage proteins
	(dz18)	δz-q-R	TGGCATCATATTCGGCATC	

[1]: Primers taken from Feng et al., 2009

[2]: Primers taken from Woo et al., 2001