Duplicate maize Wrinkled1 transcription factors activate target genes involved in seed oil biosynthesis

Benjamin Pouvreau\textsuperscript{123}, Sébastien Baud\textsuperscript{4}, Vanessa Vernoud\textsuperscript{123}, Valérie Morin\textsuperscript{123}, Cyrille Py\textsuperscript{123}, Ghislaine Gendrot\textsuperscript{123}, Jean-Philippe Pichon\textsuperscript{5}, Jacques Rouster\textsuperscript{5}, Wyatt Paul\textsuperscript{5} and Peter M. Rogowsky\textsuperscript{123}\* 

\textsuperscript{1} Université de Lyon, Ecole Normale Supérieure de Lyon, Université Lyon 1, IFR128 BioSciences Lyon Gerland, Unité Reproduction et Développement des Plantes, F-69364 Lyon, France
\textsuperscript{2} INRA, UMR879 Reproduction et Développement des Plantes, F-69364 Lyon, France
\textsuperscript{3} CNRS, UMR5667 Reproduction et Développement des Plantes, F-69364 Lyon, France
\textsuperscript{4} INRA, UMR1318 Institut Jean-Pierre Bourgin, Route de Saint-Cyr, F-78026 Versailles, France
\textsuperscript{5} Biogemma SAS, Laboratoire de Biologie Cellulaire et Moléculaire, 8 rue des frères Lumière, F-63028 Clermont-Ferrand, France

running title: \textit{ZmWri genes}

word count (including abstract, data displays and references): 9208 words

* corresponding author peter.rogowsky@ens-lyon.fr, Tel +33 4 72 72 86 07, FAX +33 4 72 72 86 01
ABSTRACT

WRI1, a key regulator of seed oil biosynthesis in Arabidopsis was duplicated during the genome amplification of the cereal ancestor genome 90 million years ago. Both maize co-orthologs ZmWri1a and ZmWri1b show a strong transcriptional induction during the early filling stage of the embryo and complement the reduced fatty acid content of Arabidopsis wri1-4 seeds, suggesting conservation of molecular function. Over-expression of ZmWri1a not only increases the fatty acid content of the mature maize grain but also the content of certain amino acids, of several compounds involved in amino acid biosynthesis and of two intermediates of the tricarboxylic acid cycle. Transcriptomic experiments identified 18 putative target genes of this transcription factor, 12 of which contain in their upstream regions an AW-box, the *cis* element bound by AtWRI1. In addition to functions related to late glycolysis and fatty acid biosynthesis in plastids, the target genes also have functions related to coenzyme A (CoA) biosynthesis in mitochondria and the production of glycerol backbones for triacylglycerol (TAG) biosynthesis in the cytoplasm. Interestingly, the higher seed oil content in ZmWri1a OE is not accompanied by a reduction in starch, thus opening possibilities for the use of the transgenic maize lines in breeding programs.

KEYWORDS

AP2 transcription factor family, development, lipid, wri1 mutation, Zea mays

INTRODUCTION

Whilst intensive agricultural and industrial use of maize kernel is largely due to its high starch content, the oil stored in the maize kernel also has considerable importance. In compound feeds for poultry, pigs and ruminant animals the oil is appreciated for its calorific value and its fatty acid composition, mainly oleic and linoleic acids. Maize oil is also a value added by-product of the starch industry, reflected by a worldwide production of 2,294,470 tons, used mainly for human food purposes, in 2009 (http://faostat.fao.org). Maize oil is highly regarded for its low level of saturated fatty acids, on average 11% palmitic acid and 2% stearic acid, and its relatively high levels of polyunsaturated fatty acids such as linoleic
Oil and starch are accumulated in different compartments of the maize kernel; 85% of the oil is stored in the embryo, whereas 98% of the starch is located in the endosperm (Watson, 1987). Consequently the relative amounts of oil and starch are correlated with the relative sizes of the embryo and endosperm, and successful breeding for high oil content in the Illinois High Oil strains has mainly been achieved through an increase in embryo size (Moose et al., 2004). Whereas the embryo represents less than 10% of the kernel weight in normal or high protein lines, it can contribute more than 20% in high oil lines (Bressani and Mertz, 1958). However, genetic components may also modulate oil content in the embryo, independently of its size, as shown by the cloning of a high-oil QTL in maize, which is caused by an amino acid insertion in an acyl-CoA:diacylglycerol acyltransferase catalyzing the last step of oil biosynthesis (Zheng et al., 2008).

The oil stored in most seeds is composed of triacylglycerols (TAGs). An extensive genetic knowledge of storage oil biosynthetic pathways, and their regulation, has been generated in the model plant Arabidopsis (Beisson et al., 2003; Baud and Lepiniec, 2010). Fatty acids are synthesized from acetyl-CoA in plastids and then exported to the cytoplasm as acyl-CoA (Ohlrogge and Browse, 1995). At the endoplasmic reticulum they are used for the acylation of the glycerol-3-phosphate backbone either by the relatively straightforward Kennedy pathway, or by acyl-exchange between lipids (Napier, 2007; Bates et al., 2009). The resulting TAGs are stored in specialized structures called oil bodies. Late steps of plastidial glycolysis necessary for the production of acetyl-CoA and several steps of fatty acid synthesis are regulated by WRINKLED1 (WRI1), (Ruuska et al., 2002; Baud et al., 2007; Baud et al., 2009; Maeo et al., 2009). WRI1 encodes a transcription factor of the large APETALA2/ethylene-responsive element binding protein (AP2/EREBP) family (Cernac and Benning, 2004). Loss-of-function mutants have no obvious phenotype during vegetative development but produce wrinkled, incompletely filled seeds with an 80% reduction in seed oil content (Focks and Benning, 1998). In addition, a delay in embryo elongation and a modification of fatty acid composition towards longer and more desaturated fatty acids were observed in wri1 mutants (Baud et al., 2007). Expression of WRI1 under the control of the
35S-promoter can lead to a slight increase of seed oil content and cause the ectopic accumulation of TAGs in developing seedlings, which show aberrant development consistent with a prolonged embryonic state (Cernac and Benning, 2004).

Expression of the *WRI1* gene has been shown to be under the direct control of the transcription factor LEAFY COTYLEDON2 (LEC2) (Baud et al., 2007), which is considered, together with LEC1, FUSCA3 (FUS3) and ABA INSENSITIVE3 (ABI3), as a master regulator of seed development (Braybrook and Harada, 2008; Suzuki and McCarty, 2008; North et al., 2010). The pleiotropic phenotypes of the respective mutants suggest that the four gene products act on separate developmental or metabolic pathways and, in so-doing, coordinate the timing of developmental events or the fluxes through particular pathways (Santos-Mendoza et al., 2008). In maize the knowledge of the regulatory network formed by LEC1, LEC2, ABI3 and FUS3 remains incomplete. The best characterized gene is *Viviparous1 (VP1)* (McCarty et al., 1991), the maize ortholog of *ABI3*. However, the extensive characterization of the *vp1* mutant focused largely on a single aspect, the role of *Vp1* in ABA-mediated regulation of seed dormancy (Suzuki et al., 2003). After the cloning and molecular characterization of *ZmLec1* in maize (Zhang et al., 2002), recent work demonstrated that its over-expression increases seed oil production but reduces seed germination and plant growth. In contrast over-expression of the putative downstream factor *ZmWri1*, stimulates oil accumulation without undesirable side effects (Shen et al., 2010).

Published data therefore suggest that the influence of WRI1 on seed oil content may be conserved between Arabidopsis and maize. However, it remains to be demonstrated whether the underlying mechanisms are conserved and whether ZmWri1 activates the same target genes as AtWRI1 in Arabidopsis (Ruuska et al., 2002; Baud et al., 2007; Maeo et al., 2009). More generally, clarification of the extent to which knowledge gained in the model species Arabidopsis can be generalized to other plants, and more precisely, whether the regulatory networks active in the exalbuminate seed of Arabidopsis are conserved in cereal grains with persistent endosperms, is required. This question is complicated by independent whole genome duplications that happened in the monocot and dicot lineages (Abrouk et al., 2010). In monocots a first duplication occurred in the common ancestor of all cereals and was followed by multiple chromosome fusions, translocations and other rearrangements (Salse et
al., 2008; Devos, 2010). Later on maize was concerned by an allo-duplication, which probably resulted from hybridization of two maize progenitors, and which occurred after its split from sorghum (Wei et al., 2007). Here we present the molecular characterization of duplicate *ZmWri1* genes in maize, their functionality in the Arabidopsis *wri1-4* mutant and the effect of the over-expression of *ZmWri1a* in maize, on the kernel transcriptome and metabolome.

**RESULTS**

*Wri1* is duplicated in maize

To identify the maize ortholog of *AtWRI1*, we performed a tblastn search of the maize genome with the amino acid sequence of *AtWRI1*. The highest scores were obtained for the two closely related gene models GRMZM2G124524 and GRMZM2G174834 (release 4a.53), which we termed *ZmWri1a* and *ZmWri1b*, respectively. To clarify the phylogenetic relationship between *AtWRI1* and the two maize genes, we enlarged the search to the related Arabidopsis proteins *AtWRI2* (*At2g41710*), *AtWRI3* (*At1g16060*) and *AtWRI4* (*At1g79700*) and included sequences from the dicot grape (*Vitis vinifera*), and from the monocots rice (*Oryza sativa*) and sorghum (*Sorghum bicolor*). The construction of a phylogenetic tree using the whole amino acid sequence of the transcription factors clearly demonstrated that *ZmWri1a* and *ZmWri1b* fell into the same clade as *AtWRI1* and that they were the closest maize relatives. Other maize proteins were present in the clades defined by *AtWRI2* or *AtWRI3/AtWRI4* (Fig. 1). Consequently *ZmWri1a* and *ZmWri1b* appeared to be co-orthologs of *AtWRI1*.

To determine the evolutionary timing of the *AtWRI1* duplication we evaluated synteny between the respective chromosome regions in maize, sorghum and rice with the SynMap and GEvo tools (http://synteny.cnr.berkeley.edu/CoGe). There was not only substantial synteny between the regions around *ZmWri1a* and *ZmWri1b* in maize but also with the regions around the sorghum gene (Sb05g001790) and the two rice genes (Os11g03540 and Os12g03290) present in the *AtWRI1* clade. In addition all 5 regions showed synteny with a region at the top of chromosome 8 of sorghum, which did not contain an *AtWRI1*-related gene. The most parsimonious interpretation of these data is that the
duplication did not occur during the last whole genome duplication of maize 5 Mya, but that it was already present in the common ancestor of maize and rice 60 Mya and that one copy was subsequently lost in sorghum after the split from maize 12 Mya.

**ZmWri1a and ZmWri1b exhibit highest expression in the embryo**

To assess functional conservation between Arabidopsis and maize *WRI1* genes, we first established expression profiles of both maize genes in major organs of the maize plant by qRT-PCR experiments using gene specific primers. Both *ZmWri1a* and *ZmWri1b* showed strongest expression in young maize kernels, although they were expressed to various extents in all organs tested (Fig. 2A). *ZmWri1a* relative mRNA levels were higher than *ZmWri1b* mRNA levels in reproductive tissues, whereas the opposite was true in leaves, where both genes and, in particular *ZmWri1b*, showed considerably higher expression in sheaths than in blades. The preferential expression of *ZmWri1a* and *ZmWri1b* in maize kernels mirrored the preferential expression of *AtWRI1* in Arabidopsis siliques (Cernac and Benning, 2004).

During the development of the maize caryopsis, *ZmWri1a* and *ZmWri1b* were expressed at low levels in immature or mature ovules (data not shown) and in kernels younger than 7 days after pollination (DAP, Fig. 2B). A first expression maximum between 7 and 12 DAP was shared by both genes, although with different kinetics. Finally *ZmWri1a* showed a second expression peak between 30 and 50 DAP (Fig. 2B).

To gain further insight into the spatial expression pattern of *ZmWri1* genes at the onset of the filling stage, which lasts approximately from 12 to 30 DAP in our material, qRT-PCR was performed on dissected embryos and endosperms. Both genes showed an almost constant expression level in the endosperm and a sharp increase in expression between 9 and 12 DAP in the embryo (Fig. 2C). At 12 DAP, expression was predominant in the embryo, *ZmWri1a* and *ZmWri1b* mRNA levels being more than 50-fold and more than 5-fold higher than in the endosperm, respectively. The relative expression levels of the two genes somewhat differed between embryo and endosperm, *ZmWri1a* being stronger in the embryo and *ZmWri1b* stronger in the endosperm.
Both ZmWri1a and ZmWri1b complement the wri1 mutation in Arabidopsis

To further investigate the functionality of ZmWri1a and ZmWri1b in planta, complementation experiments were carried out in Arabidopsis. Homozygous Arabidopsis wri1-4 mutants (Baud et al., 2007) were transformed with maize ZmWri1a or ZmWri1b cDNA, the expression of which was driven by the seed-specific AT2S2 (At4g27150) promoter. For each construct seven independent primary transformants were selected and propagated; the progenies of T3 lines were subjected to detailed analyses. A microscopic observation of mature dry seeds showed a complete reversion of the wrinkled seed phenotype usually observed in the wri1-4 mutant background (Fig. 3A). Fatty acid analyses confirmed the ability of both ZmWri1a and ZmWri1b to restore the defect in fatty acid accumulation previously described in wri1-4 seeds (Fig. 3B, C).

It is noteworthy that the fatty acid profile of wri1-4 seeds complemented either with ZmWri1a or ZmWri1b differed from the profile in wild-type seeds. In particular, complemented seeds exhibited an increased content of 18:3 fatty acid species, and a concomitant reduction of 18:1 and 18:2 fatty acid species (Fig. S1). This might be the consequence of a delay in AT2S2 promoter activity compared to AtWRI1 promoter activity. Experiments using a Pro\textsubscript{AT2S2}:uidA transgene have shown that Pro\textsubscript{AT2S2} activity was specifically detected in the embryonic tissues of maturing seeds from the early-bent stage onward (data not shown), whereas the Pro\textsubscript{AtWRI1} activity was already detected in torpedo-shaped embryos (Baud et al., 2007). The delayed induction of maize ZmWri1 cDNAs in complemented seeds compared to that of AtWRI1 in wild-type seeds may postpone the production of de novo fatty acids until mid-maturation, when FATTY ACID DESATURASE 3 (FAD3) is already highly active (Zimmermann et al., 2004).

Over-expression of ZmWri1a increases fatty acid content of the maize kernel

Since over-expression of AtWRI1 under the control of a strong and constitutive promoter had limited (Cernac and Benning, 2004) or no effect (Baud et al., 2009) on the oil content of oleaginous Arabidopsis seeds, we wanted to test whether a similar approach could impact more significantly on the composition of non-oleaginous maize kernels. To this end we placed the ZmWri1a cDNA under the control of the constitutive Cassava vein mosaic
virus (CsVMV) promoter and generated 11 independent transformation events in maize. In the T1 generation three events were selected for further analysis on the basis of complete T-DNA transfer, single transgene copy number and high expression level of the transgene.

RT-PCR experiments on T1 leaves established comparable expression levels of the transgene in the three selected events and a near infrared spectrum (NIRS) analysis indicated that the transgenic, hemizygous T2 kernels of all three events had a higher content of certain fatty acids (linoleic acid, oleic acid, total fatty acids) and certain amino acids (proline, serine, tyrosine) than wild-type kernels from the same ears (data not shown). The transgenic kernels did not show any macroscopically visible defects such as wrinkled or plump phenotypes and there was no significant difference in kernel weight between transgenic and wild-type kernels (data not shown).

To further characterize the impact of the ZmWri1a transgene on the carbon metabolism of the maize kernel, we performed a GC-TOF-MS metabolomics analysis on T4 kernels from self-pollinated ears of three homozygous ZmWri1a-OE plants and three wild-type siblings. Out of 152 metabolites analyzed, 79 were actually detected, and the relative abundance of 22 of these was significantly different between mature ZmWri1a-OE and wild-type kernels (Table 1). In parallel, principal component analysis confirmed the strong difference between the ZmWri1a-OE and wild-type samples, the first three components explaining 69% of the variability. Among the 10 metabolites with the most strongly significant differences (p-value of Student's t-test < 0.01) were four fatty acids (palmitic, linolenic, oleic and stearic acid), three short organic acids (succinic, citric and glyceric acid) and two free amino acids (lysine and glutamic acid), all of which showed an increased content in transgenic kernels. An independent GC analysis of total fatty acid content (from C16 to C22 in their saturated and unsaturated forms, up to 3 double bonds) allowed the quantification of the increase between transgenic (71.09 µg/mg DW) and wild-type kernels (64.68 µg/mg DW).

Since starch, the major storage product of the maize kernel, is a polymer and consequently not captured by GC-MS based techniques, we used spectroscopic methods to measure starch content as well as the relative amounts of amylose and amylopectin. No statistically significant difference was detected between mature ZmWri1a-OE and wild-type
kernels despite a consistent trend for slightly lower starch, amyllose and amylopectin contents in transgenic kernels (data not shown).

Finally, the microscopic observation of cytological sections of 24 DAP kernels did not reveal any morphological defect in ZmWri1a-OE kernels, and size measurements of the embryo at 24 DAP and at maturity did not show any significant size differences between ZmWri1a-OE and wild-type embryos. In summary, the increased content in oil and certain free amino acids in ZmWri1a-OE kernels impacted neither on the deposition of other storage compounds, nor on kernel morphology.

**Transcriptome analysis identifies 18 putative target genes of ZmWri1a**

To identify genes regulated by the transcription factor ZmWri1a, a transcriptome comparison between transgenic ZmWri1a-OE maize plants over-expressing ZmWri1a and wild-type sister plants was carried out. RNA was extracted from the aerial parts of 18 DAS (days after sowing) plantlets and used to hybridize a genome wide 46 K micro-array. A first gene list of only 5 differentially expressed genes was established based on a p-value <0.05 for the biological triplicate and strong expression differences (logR>0.3 or <-0.3). Using lower stringency for the second parameters (logR>0.2 or <-0.2) but including high spot intensity (logI>2.3) as an additional criterion, the list was extended to 25 candidates. The differential expression was confirmed for 18 of the 25 candidates by qRT-PCR experiments based on the same samples that had been used for the initial micro-array analysis (Table 2). In four cases the up-regulation observed after array hybridization was not confirmed by qRT-PCR. Two candidates were eliminated due to redundancy because two distinct oligonucleotides on the array, which had been designed based on EST assemblies, corresponded in reality to a single gene in the more recently established maize genome sequence (MZ00042163 and MZ00016943 as well as MZ00044044 and MZ00026553). Among the confirmed candidates, two genes corresponding to MZ00024552 (chr 1) and MZ00043500 (chr 9) were near isogenic paralogs so that it was not possible to design gene-specific primers and test whether the up-regulation in ZmWri1a-OE plantlets concerned one or both genes.

Since the function of ZmWri1a has been attributed to the kernel, the up-regulation of the 18 confirmed target genes was also assessed in developing ZmWri1a-OE and wild-type
kernels at 16, 24, 32 and 40 DAP (Table 2 and Fig. 4). At 24 DAP all 18 genes were significantly up-regulated and for 15 of them this was also true at all other developmental stages tested. However, in two cases (MZ00031529 and MZ00039375) the up-regulation in ZmWri1a-OE kernels was limited to the stages 16 and 24 DAP, whereas expression was much stronger in wild-type kernels at 32 and particularly at 40 DAP. The corresponding genes may therefore be controlled by two distinct regulatory pathways with ZmWri1a either promoting the first or inhibiting the second. Similar reasoning would also explain the up-regulation of the gene corresponding to MZ00042142 in wild-type kernels at 16 DAP but not in transgenic kernels at later stages.

**Putative targets of ZmWri1a act in the glycolysis, the fatty acid or TAG-related biosynthetic pathways**

To determine whether the 18 putative target genes acted in common metabolic or developmental pathways, complete protein sequences were assembled starting from the 70 nt oligonucleotide deposited on the microarray, exploiting the very rich maize EST data available (Alexandrov et al., 2009; Soderlund et al., 2009), as well as the recently established maize genome sequence (Schnable et al., 2009). Based on sequence similarities most of the 18 genes encoded enzymes of the glycolysis and the fatty acid or TAG-related biosynthetic pathways (Table 2).

The deduced amino acid sequence of the first gene showed some homology to the PEP utilizing domain (pfam00391) of a bacterial phosphoenolpyruvate-protein phosphotransferase which has been shown to produce pyruvate from PEP and use the phosphate, via carrier proteins, for various processes such as sugar transport (Rabus et al., 1999). Two near isogenic genes (MZ00024552, MZ00043500) coded for closely related plastidial pyruvate kinases catalyzing the irreversible synthesis of pyruvate and ATP from PEP and ADP. The first step of fatty acid synthesis is executed by the pyruvate dehydrogenase multienzyme complex (PDHC), which catalyzes the decarboxylation of pyruvate into acetyl-CoA. Among the targets isolated, three genes encoded either \( \alpha \) (MZ00024718) or \( \beta \) (MZ00016632, MZ00014741) components of the pyruvate dehydrogenase (E1 subunit of PDHC) and two genes (MZ00017663, MZ00043050) encoded dihydrolipoyl acyltransferase (E2 subunit of PDHC).
In addition, the lipoyltransferase encoded by MZ00017651 likely catalyzes the covalent attachment of lipoic acid to the PDHC-E2 apo-enzyme, thereby creating the active holo-enzyme.

The first gene clearly involved in acyl chain production (MZ00056535) encodes 3-ketoacyl-acyl carrier protein synthase III, the isoform responsible for the initial condensation of acetyl-CoA with malonyl-ACP. This reaction, like the subsequent cycles responsible for fatty acid elongation, requires acyl carrier proteins (ACP), one of which was also encoded by a confirmed candidate gene (MZ00016866). Two other candidates encoded enzymes involved downstream in the plastidial fatty acid biosynthetic process: oleoyl-acyl carrier protein thioesterase (MZ00041636) releases free fatty acid from ACP at the inner plastid membrane, whereas the long chain acyl-CoA synthetase MZ00017355 activates free fatty acids to CoA-esters in the outer plastid envelope (Shockey et al., 2002).

Despite its suggestive annotation, the function of the acetyl-CoA acetyltransferase encoded by MZ00040095 in fatty acid synthesis remains unclear. This cytoplasmic enzyme is thought to be involved either in isoprenoid biosynthesis (mevalonate pathway), isoleucine degradation or lysine degradation and may influence the CoA pool (Carrie et al., 2007). The ketopantoate hydroxymethyltransferase encoded by MZ00042163/MZ00016943 has a more direct link to CoA biosynthesis, since it catalyzes the reaction from valine to 2-keto-isovalerate, a precursor of pantoate, the universal precursor for the synthesis of the 4'-phosphopantetheine moiety of CoA and ACP. The CoA biosynthetic pathway also contains a decarboxylation reaction, which may be facilitated by the action of a β-type carbonic anhydrase regulating the balance between carbon-dioxide and carbonic acid and possibly encoded by MZ00044044/MZ00026553.

Finally, MZ00015977 encodes a cytosolic NAD-dependent glycerol-3-phosphate dehydrogenase providing glycerol backbones necessary for TAG biosynthesis. No precise function could be attributed to the outer membrane lipoprotein MZ00031529, while the last confirmed candidate has no informative annotation (MZ00039375).
DISCUSSION

WRI1, a key regulator of oil biosynthesis discovered in Arabidopsis, is duplicated in maize. Our expression and complementation data indicate that both ZmWri1a and ZmWri1b fulfill a role similar to that of AtWRI1 in the developing embryo. We have shown that at least ZmWri1a acts via the transcriptional regulation of enzymes involved in glycolysis or in the fatty acid and TAG biosynthetic pathways, and that over-expression of ZmWri1a is sufficient to increase the oil content in maize kernels.

Conserved function and expression between ZmWri1 genes and AtWRI1

The duplication of WRI1 is not specific to maize but is also found in rice and probably occurred 90 million years ago during the genome duplication of the common ancestor of cereals. Both maize genes encode functional proteins, since they both complement the wrinkled phenotype and the low fatty acid content of Arabidopsis wri1-4 mutant seeds. These data suggest that the overall molecular function, i.e., the transcriptional activation of multiple target genes involved in fatty acid biosynthesis, has been conserved over more than 150 million years between AtWRI1, ZmWri1a and ZmWri1b. The qualitative differences in fatty acid composition between wri1-4 mutants complemented with maize genes and wild-type controls might be explained by a heterochronic effect of the AT2S2 promoter used for complementation and/or by ongoing co-evolution between each individual transcription factor and its respective cis elements/targets within species.

Despite the overall conservation of their molecular function the two maize genes ZmWri1a and ZmWri1b may not play the same role in the maize plant. Although both ZmWri1 genes mirror the expression pattern of AtWRI1 with a preferential expression in the kernel and a strong induction in the embryo at the onset of the maturation phase, the expression patterns of ZmWri1a and ZmWri1b are clearly distinct. Contrary to ZmWri1a, ZmWri1b is barely expressed in reproductive tissues prior to pollination, whereas the opposite is true in leaf sheaths. In the kernel, ZmWri1a has a bi-phasic expression pattern with a second peak at the end of the filling stage, which is missing for ZmWri1b. In the embryo ZmWri1a is the predominant gene, while ZmWri1b but not ZmWri1a shows substantial expression in the endosperm. All these expression differences likely reflect an ongoing
specialization on an evolutionary time scale, which may imply subtle changes in the sets of target genes and/or altered responses to regulators of carbon partitioning. One may speculate that ZmWri1b is evolving towards a house keeping gene maintaining TAG biosynthesis for various cellular processes, while ZmWri1a remains mainly devoted to storage fatty acid biosynthesis during grain filling.

**Conserved regulation of oil biosynthesis between maize and Arabidopsis**

The over-expression of ZmWri1a under the control of the constitutive CsVMV promoter was sufficient to increase the fatty acid content of the maize kernel, and confirmed recent results obtained by over-expression of the same gene under the control of the embryo-preferred OLE promoter (Shen et al., 2010). The underlying mechanism involves the transcriptional up-regulation of at least 18 target genes identified by a comparative transcriptome analysis between ZmWri1a-OE and wild-type leaves and confirmed by qRT-PCR experiments in leaves and kernels. The manually improved annotations of the deduced amino acid sequences clearly indicate that the majority of these genes have functions related to late glycolysis and fatty acid or TAG biosynthesis. This reflects the situation in Arabidopsis, where a list of putative target genes of AtWRI1 has been previously established by a combination of microarray analysis, targeted qRT-PCR, yeast-one-hybrid and EMSA experiments (Ruuska et al., 2002; Baud et al., 2007; Baud et al., 2009; Maeo et al., 2009). Common targets between maize and Arabidopsis include subunits of the pyruvate kinase and pyruvate dehydrogenase complexes, ketoacyl-ACP synthase, acyl-ACP thio-esterase and acyl carrier proteins (Fig. 5). While our confirmed candidates did not include all putative targets isolated in Arabidopsis (Fig. 5), such as G-6P/phosphate translocator, PEP/phosphate translocator, acetyl-CoA carboxylase, transacylase, hydroxyacyl-ACP dehydrogenase, enoyl-ACP reductase or acyl-ACP desaturase, genes encoding these enzymes may be up-regulated in ZmWri1a-OE plants but have escaped classification as differentially expressed for a variety of technical reasons, such as expression levels close to background, or low spot quality of one of the replicas. Similarly, novel target genes with differential expression detected only in maize may also be targets of AtWRI1 in Arabidopsis. These include genes encoding glycerol-3-phosphate dehydrogenase, PEP-protein phosphotransferase, lipoyltransferase and long
chain acyl-CoA synthase (Fig. 5). It is noteworthy that two of these enzymes are predicted to be located at the plastid membrane (long chain acyl-CoA synthase) or in the cytoplasm (glycerol-3-phosphate dehydrogenase) and have functions related to TAG biosynthesis.

Among the remaining four ZmWri1a targets one has no similarity to proteins or domains with known function, whereas the other three have roles related to CoA biosynthesis and/or homeostasis (Fig. 5). At first sight, their predicted localization in mitochondria (ketopantoate hydroxymethyltransferase, carbonic anhydrase) or the cytoplasm (acetyl-CoA acetyltransferase) makes it difficult to establish a direct link to fatty acid biosynthesis localized in plastids. However, it is known that CoA is present in all compartments and that precursors are shuttled between compartments (Rubio et al., 2008). In addition the over-expression of a CoA biosynthetic enzyme has been shown to lead to increased seed oil content in Arabidopsis (Rubio et al., 2008). The maize acetyl-CoA acetyltransferase shows similarity to ACAT1 and ACAT2 from Arabidopsis, suggesting a function in the mevalonate pathway leading to isoprenoid biosynthesis, rather than in the degradation of lysine of isoleucine (Carrie et al., 2007). The increase rather than decrease in lysine content in ZmWri1a-OE kernels further substantiates this hypothesis. While the mevalonate pathway consumes rather than produces acetyl-CoA, the reaction releases CoA, which may be transported to the plastid and have a beneficial effect on fatty acid synthesis.

AW-boxes are present in the upstream regions of ZmWri1a target genes

In Arabidopsis the AW-box 5' CnTnG(n)7CG 3' has been identified as the AtWRI1 binding site in the promoter sequences of three target genes and 28 AW-boxes have been found in the upstream regions of 19 genes involved in fatty acid biosynthesis out of 46 examined (Maeo et al., 2009). A survey of the putative promoter sequences - 3 kb upstream of the predicted ATG start codons - of the 18 ZmWri1a target genes isolated in this study demonstrated the presence of 33 AW-boxes in 12 of the 18 sequences analyzed (Fig. 6A). As in the case of Arabidopsis, at least one AW-box per sequence was located close to the start codon, with one possible exception (MZ00044044/MZ00026553). Our analysis also showed that an AW-box can overlap with a second AW-box, both on the sense (Fig. 6B) and the antisense strand (Fig. 6C). The existence of a palindromic AW-box may hint at the binding of
ZmWri1a as a homodimer or as a heterodimer with another transcription factor of the AP2/EREBP family.

**Metabolic adjustments in ZmWri1a-OE kernels**

Transgenic ZmWri1a-OE kernels did not only show a significant increase in saturated and unsaturated fatty acids with 16 to 18 carbon atoms, but also a significant increase for several free amino acids (lys, glu, phe, ala, val), intermediates or co-factors of amino acid biosynthesis (pyroglutamate, aminoadipic acid, ornithine, norleucine) and intermediates of the tricarboxylic acid cycle (citric acid, succinic acid). Since the transcriptome analysis suggests that ZmWri1a essentially activates genes coding for enzymes in late glycolysis, fatty acid, CoA and TAG biosynthesis, and considering that no mis-regulated candidates participate in any additional pathways, the increase in amino acids and TCA intermediates probably reflects secondary adjustments of the carbon and nitrogen metabolism to the increased oil biosynthesis triggered by ZmWri1a. The three amino acids phenylalanine, alanine and valine are derived from PEP or pyruvate and their increase may simply be a by-product of a strongly increased carbon flux through glycolysis. Likewise, the increased amounts of citric and succinic acid may reflect an increased activity of the TCA cycle. Glutamate is the storage form of ammonia in the plant and the very steep increase in glutamate content of transgenic kernels may be indicative of an excess of nitrogen or rather an imbalance between nitrogen and available carbon skeletons. The increase in aminoadipic acid, ornithine and pyroglutamate is possibly linked to subsequent adjustments of the ammonia assimilation cycle between glutamine and glutamate and downstream pathways leading to proline and arginine synthesis.

**Use of ZmWri1a-OE plants in plant breeding**

Our results on greenhouse grown transgenic plants over-expressing ZmWri1a under the control of the constitutive CsVMV promoter confirm recent results on field grown transgenic plants using the embryo-preferred OLE promoter in showing that the increase in seed oil content is not linked to an increase in embryo size but in seed oil (Shen et al., 2010). This increase in oil does not compromise the other uses of the maize kernel, and in particular the use of starch for nutritional or industrial purposes, since neither study detected significant
changes in the overall starch content of the kernel, despite a decrease in the embryo, where very minor amounts of starch are stored (Shen et al., 2010). While it does not appear to be possible to produce oil rather than starch in maize endosperm by the simple over-expression of ZmWri1a in the starchy endosperm (Shen et al., 2010), increased yield of nutritionally improved oil from transgenic embryos is still of sufficient interest to the milling industry to warrant use of ZmWri1a-OE lines in plant breeding. In addition, the differences in oil quality seen in Arabidopsis between wild-type seed and mutant seed complemented with maize ZmWri1 genes may provide valuable information for the engineering of oil seed crops. Once it has been determined whether a heterochronic effect of the AT2S2 promoter or the altered coding sequence of the maize genes are responsible for the effect, the identified component could be transferred to rapeseed or other crops to produce oil with higher levels of polyunsaturated fatty acids.

Beyond nutritional aspects TAGs also present considerable interest for energy production. In vegetative tissues of Arabidopsis the ectopic expression of AtWRI1 leads to a 3-fold increase in TAG levels, which can be further increased by the concomitant repression of ADP-glucose pyrophosphorylase (Sanjaya et al., 2011). This approach can likely be transferred to maize since mRNA levels of all 18 target genes of ZmWri1a are significantly increased in leaves of ZmWri1a-OE plants (Table 2). And crosses to ADP-glucose pyrophosphorylase mutants may provide a means to shift carbon partitioning at least partially from starch to TAG biosynthesis not only in vegetative tissues but also in the endosperm.

MATERIALS AND METHODS

Plant material and growth conditions

The maize inbred line A188 and transgenic A188 plants over-expressing ZmWri1a were grown in a S2 greenhouse with a 16 h illumination period (100 Wm⁻²) at 24/19 °C (day/night) and without control of the relative humidity. Kernels were germinated in 0.2 l Favorit® MP Godets substrate (Eriterre, St. André de Corcy, France) and were transferred at 21 DAS to 10 l Favorit® Argile TM substrate (Eriterre, St. André de Corcy, France) supplemented with 4 g/l Osmocote® Exact® hi-end 15+9+12 fertilizer (Scotts, Marne la Vallée, France). All plants were propagated by hand pollination.
Seeds of the Arabidopsis *wri1-4* (Col-0 background) mutant line (Baud et al., 2007) were surface sterilized and germinated on Murashige and Skoog (MS) medium (Murashige and Skoog, 1962). After a cold treatment of 48 h at 4°C in the dark, plates were kept in a growth chamber (16/8 h light photoperiod at 150 µm m⁻² s⁻¹; 15°C night/20°C day temperature). After 10 d, the plantlets were transferred to compost, grown in a greenhouse under similar conditions, and irrigated twice a week with mineral nutrient solution.

**T-DNA constructs and plant transformation**

The coding sequences of *ZmWri1a* and *ZmWri1b* were amplified on cDNA from 7 DAP kernels (genotype A188) with primers FT126-F7 and FT126-R7 and on cDNA from 16 DAP kernels (genotype B73) with primers attB1-Wri2 and attB2-Wri2, respectively. After BP recombination into pDONZeo (Invitrogen) the resulting entry vectors L698 and L1129 were sequenced prior to LR recombination. Only L698 was recombined into the maize destination vector pBIOS886, which was based the backbone of vector pSB11 (Ishida et al., 1996) and contained between T-DNA borders a Basta resistance cassette for selection, a GFP cassette for the tracking of transgenics in segregating material as well as attR1 and attR2 sites downstream of the constitutive *CsVMV* promoter and upstream of the *Sac66* terminator. Primer sequences are given in Table S1.

*Agrobacterium*-mediated transformation of maize inbred line A188 with the resulting plasmid was based on a published protocol (Ishida et al., 2007). Among 11 independent transformation events, three events with single copy insertions and high expression levels were chosen for the present study.

For the construction of the *Pro*_{AT2S2}-*R1R2-HYGRO* destination vector the *AT2S2* (At4g27150) promoter) was amplified from with the proofreading Pfu Ultra DNA polymerase (Stratagene) on Arabidopsis genomic DNA using primers *Hind*III*pAT2S2up* and *AscI*pAT2S2low* (Table S1). The PCR product representing position -984 to -1 relative to the *AT2S2* translational start codon was digested with *Hind*III and *AscI* and cloned into *pMDC32* (Curtis and Grossniklaus, 2003) digested with *Hind*III and *AscI* to replace the 2X35S cassette. *ZmWri1a* and *ZmWri1b* were recombined from their respective entry clones into this vector by LR recombination.
The binary vectors were electroporated into *Agrobacterium tumefaciens* C58C1 strain and used for agroinfiltration of Arabidopsis *wri1-4* flower buds (Bechtold et al., 1993). Primary transformants were selected on MS medium containing hygromycin (50 mg l\(^{-1}\)) and then transferred to soil. Homozygous T3 lines producing 100% resistant plantlets were selected and used for complementation analyses. For each construct, seven independent transformants were considered.

**Sequence analysis**

The cDNA sequences corresponding to the 70 mers present on the micro-array were established by BlastN individual EST sequences or full length cDNA sequences at NCBI (www.ncbi.nlm.nih.gov/BLAST/). Consensus sequences were obtained using VectorNTI ContigExpress software (Invitrogen) and regularly updated. Genomic sequences were obtained by BlastN of the cDNA sequence against the maize genome (www.maizesequence.org). Deduced amino acid sequences were annotated by BlastP against the *Arabidopsis* genome at NCBI and screened for known, conserved domains using the CDS database.

After amino acid sequences alignment with ClustalW and conserved block selection with SeaView (http://pbil.univ-lyon1.fr/software/seaview.html) maximum likelyhood phylogenetic trees were generated with Treefinder software (www.treefinder.de) using the substitution model WAG_optimumG4 and 1000 bootstrap replicates.

**Quantitative reverse transcription-polymerase chain reaction**

Total RNA was extracted with the TRIzol\(^\text{®}\) reagent, treated with DNAsse and reverse transcribed as described previously (Javelle et al., 2010). Real time PCR was carried out with the Platinum\(^\text{®}\) SYBR\(^\text{®}\) Green qPCR SuperMix UDG (Invitrogen) on a StepOne™ Real-Time PCR System (Applied Biosystems). Dilutions series (2\(^n\) with n=0 to 7) of a mixture of all cDNA within a comparison were used to fix the C\(_T\) (threshold cycle). Gene expression levels relative to the *18S* rRNA reference gene were calculated by the \(\Delta\Delta C_T\) method (Schmittgen and Livak, 2008). The primers used are listed in supplementary Table S1.
Micro-array analysis

Hybridization of the maize 46 K (version 1) 70-mer oligo array printed on a single glass slide (http://www.maizearray.org) was performed in biological triplicate and with Cy5/Cy3 dye swap as technical replicates. Probes were synthesized from total RNA isolated from the aerial parts of 18 DAS maize seedlings from wild-type and ZmWri1a-OE plants was using an Eberwine-based amplification method and indirect labeling with dyes. Briefly, 2µg of total RNA were reverse-transcribed with an oligo dT primer bearing a T7 promoter. After second strand synthesis, double strand cDNA served as template for amplification by in vitro transcription with the Amino Allyl MessageAmp™ aRNA Kit (Ambion). The incorporated amino-allyl dUTP was then labelled by covalent binding with activated dye esters using Cy5 Mon-Reactive Dye Pack and Cy3 Mon-Reactive Dye Pack (GE Healthcare). Each microarray was hybridized with a mix of 3 µg of Cy3-labeled target and 3 µg of Cy5-labeled target resuspended in 60 µL of 1 x hybridization buffer (50% formamide, 5 x SSC, and 0.1% SDS, 0.4 µg/µl tDNA, 0.2 µg/µl salmon sperm DNA). Hybridizations were carried out overnight at 42°C in SlideBooster™ hybridization station (Advalytix, Munich, Germany). After slide washing (most stringent wash for 5 min in 0.05 x SSC at RT) slides were immediately scanned on a ScanArray 4000XL scanner (Perkin Elmer).

Raw data were normalized using a quantile normalization approach consisting in adjusting each individual signal distribution to a common one. Normalization was performed with the open source R software with packages dedicated to microarray analysis. Gene lists were constituted of differentially expressed genes presenting a minimum 2 fold change between the 2 conditions with a \( p \)-value statistical threshold at 0.01.

Metabolomic measurements

The extraction of metabolites from plant tissues and their analysis by gas chromatography coupled to time-of-flight mass spectrometry on a LECO Pegasus III with an Agilent 6890 N gas chromatography system and an Agilent 7683 automatic liquid sampler was performed as described (Tcherkez et al., 2009). The integration of peaks was performed using the LECOPegasus software. Because automated peak integration was occasionally erroneous, integration was verified manually for each compound in all analyses.
Fatty acid analyses were performed on pools of 20 seeds (Arabidopsis) and on meals of lyophilized grain tissues (maize) as previously described (Li et al., 2006). Starch content was determined with the Starch kit according to the instructions of the manufacturer (Roche). Amylose and amylopectin were detected simultaneously by multi-wavelength analysis (Sené et al., 1997).

REFERENCES


candidates, a study of the distribution of expressed sequence tags in organs, and a Web-based database. Plant Physiol 132: 681-697


Carrie C, Murcha MW, Millar AH, Smith SM, Whelan J (2007) Nine 3-ketoacyl-CoA thiolases (KATs) and acetoacetyl-CoA thiolases (ACATs) encoded by five genes in Arabidopsis thaliana are targeted either to peroxisomes or cytosol but not to mitochondria. Plant Mol Biol 63: 97-108


leucine zipper IV transcription factor Outer Cell Layer1 in maize identifies target genes involved in lipid metabolism and cuticle biosynthesis. Plant Physiol 154: 273-286


**ACKNOWLEDGMENTS**

We thank Isabelle Desbouchages, Alexis Lacroix and Priscilla Angelot for maize culture, Hervé Leyral and Claudia Bardoux for the preparation of buffers and media, the Biogemma transformation team for the production of transgenic maize plants, the Biogemma transcriptomics team for the hybridization and analysis of micro-arrays and the analytical facility "Plateforme Métabolisme-Métabolôme" of the "Institut Fédératif de Recherche 87" for metabolomics analysis. The work was financed in part by the successive Génoplante projects "MaizeTF" (GABI-GP-2003-6), "MaizeYield" (ANR-05-GPLA-031) and "HyperMaize" (ANR-07-GPLA-019) and the ANR project "Plant-TFcode" (ANR-07-BLAN-0211-02).

**FIGURE LEGENDS**

**Fig. 1: Phylogenetic tree of WRI proteins**

A maximum likelihood phylogenetic tree was generated using all full length grapevine (GSVIV, black), maize (GRMZM, red), rice (Os, blue) and sorghum (Sb, purple) amino acid
sequences related to AtWRI1, AtWRI2, AtWRI3 and AtWRI4 from Arabidopsis (At, green) using Treefinder software. Percentage values on each branch represent the corresponding bootstrap probability.

Fig. 2: Expression profiles of ZmWri1a and ZmWri1b in the maize plant

Expression profiles of ZmWri1a and ZmWri1b were established by qRT-PCR in major maize organs (A), during kernel development (B) and in dissected endosperm and embryo (C). Leaf blade (b) or sheath (s) of juvenile (j) or adult (a) leaves, roots of seedlings (s) or plantlets (p), immature (i) or mature (m) ears and tassels as well as kernels at 12 or 35 days after pollination were used in (A), kernels of the indicated days after pollination in (B) and dissected endosperms (End) or embryos (Emb) of the indicated days after pollination in (C). Error bars correspond to the standard deviation calculated from technical quadruplicates on organ or tissue pools of various size harvested from two plants.

Fig. 3: Complementation of the Arabidopsis wri1-4 mutant with maize ZmWri1a and ZmWri1b

(A) Images of mature seeds of wild-type (Col-0), mutant (wri1-4) and complemented mutant (wri1-4 + ZmWri1a, wri1-4 + ZmWri1b) (B, C) Seed fatty acid content of Col-0, wri1-4 and 6 independent wri1-4 lines complemented with ZmWri1a (B) or ZmWri1b (C) from maize. For each transformation event 5 distinct plants were cultured. Twenty seeds were harvested from each plant and pooled. Error bars correspond to the standard deviation calculated from 5 technical replicates per pool of 100 seeds.

Fig. 4: Expression profiles of ZmWri1a target genes in ZmWri1a-OE and wild-type kernels

Expression profiles of direct or indirect target genes of Wri1a were established by qRT-PCR in ZmWri1a-OE and wild-type kernels at four developmental stages indicated in DAP (days after pollination). Genes are identified by their MZ000xxxxx number used in Table 2. Error bars correspond to the standard deviation calculated from technical duplicates of the gene of interest and the reference gene on pools of 5 kernels. For each data point the
difference between *ZmWri1a*-OE and wild-type kernels is statistically significant with a p-value <0.05 in Student's t-test.

**Fig. 5: Putative metabolic functions of up-regulated ZmWri1a target genes**

Enzymatic functions encoded by up-regulated target genes of *ZmWri1a* and AtWR1 are in orange, those encoded only by *ZmWri1a* target genes in yellow and those encoded only by AtWR1 target genes in white.

**Fig. 6: AW-boxes in upstream sequences of ZmWri1a target genes**

(A) Schematic drawing of the sequence 3 kb upstream of the ATG start codon (arrow) of *ZmWri1a* target genes. AW-boxes in the sense strand are indicated by black vertical bars shifted to the top, AW-boxes in the antisense strand by blue vertical bars shifted to the bottom (B) Palindromic AW-boxes in the upstream region of MZ00014741 (C) Overlapping AW-boxes in the upstream region of MZ00042142. The fixed bases of the consensus sequence CnTnG(n)7GC are in upper case letters.
TABLES

Table 1: Metabolites with significantly increased or decreased content in ZmWri1a-OE kernels

<table>
<thead>
<tr>
<th>Metabolite</th>
<th>Trend in Wri1a-OE</th>
<th>Mean² Wri1a-OE</th>
<th>Mean² WT</th>
<th>SD Wri1a-OE</th>
<th>SD WT</th>
<th>Ratio Wri1a-OE / WT³</th>
</tr>
</thead>
<tbody>
<tr>
<td>Palmitic acid</td>
<td>UP</td>
<td>0.02362</td>
<td>0.01254</td>
<td>0.00217</td>
<td>0.00039</td>
<td>1.88**</td>
</tr>
<tr>
<td>Succinic acid</td>
<td>UP</td>
<td>0.00152</td>
<td>0.00105</td>
<td>0.00010</td>
<td>0.00006</td>
<td>1.45**</td>
</tr>
<tr>
<td>Linolenic acid</td>
<td>UP</td>
<td>0.00133</td>
<td>0.00069</td>
<td>0.00015</td>
<td>0.00006</td>
<td>1.94**</td>
</tr>
<tr>
<td>Lysine</td>
<td>UP</td>
<td>0.00188</td>
<td>0.00128</td>
<td>0.00011</td>
<td>0.00011</td>
<td>1.47**</td>
</tr>
<tr>
<td>Oleic acid</td>
<td>UP</td>
<td>0.00085</td>
<td>0.00039</td>
<td>0.00010</td>
<td>0.00007</td>
<td>2.18**</td>
</tr>
<tr>
<td>Glyceric acid</td>
<td>UP</td>
<td>0.00050</td>
<td>0.00032</td>
<td>0.00002</td>
<td>0.00004</td>
<td>1.53**</td>
</tr>
<tr>
<td>Stearic acid</td>
<td>UP</td>
<td>0.00230</td>
<td>0.00111</td>
<td>0.00017</td>
<td>0.00036</td>
<td>2.07**</td>
</tr>
<tr>
<td>Citric acid</td>
<td>UP</td>
<td>0.00411</td>
<td>0.00236</td>
<td>0.00052</td>
<td>0.00031</td>
<td>1.74**</td>
</tr>
<tr>
<td>Glutamic acid</td>
<td>UP</td>
<td>0.00049</td>
<td>0.00000</td>
<td>0.00001</td>
<td>0.00000</td>
<td>98.76**</td>
</tr>
<tr>
<td>Phosphoric acid</td>
<td>UP</td>
<td>0.00806</td>
<td>0.00446</td>
<td>0.00107</td>
<td>0.00073</td>
<td>1.81**</td>
</tr>
<tr>
<td>Phenylalanine</td>
<td>UP</td>
<td>0.00076</td>
<td>0.00049</td>
<td>0.00010</td>
<td>0.00003</td>
<td>1.54*</td>
</tr>
<tr>
<td>Arabinose</td>
<td>UP</td>
<td>0.00333</td>
<td>0.00244</td>
<td>0.00030</td>
<td>0.00020</td>
<td>1.36*</td>
</tr>
<tr>
<td>Linoleic acid</td>
<td>UP</td>
<td>0.00147</td>
<td>0.00081</td>
<td>0.00022</td>
<td>0.00016</td>
<td>1.82*</td>
</tr>
<tr>
<td>Decanoic acid</td>
<td>DOWN</td>
<td>0.00754</td>
<td>0.01061</td>
<td>0.00025</td>
<td>0.00132</td>
<td>0.71*</td>
</tr>
<tr>
<td>Pyroglutamic acid</td>
<td>UP</td>
<td>0.00723</td>
<td>0.00470</td>
<td>0.00117</td>
<td>0.00037</td>
<td>1.54*</td>
</tr>
<tr>
<td>Norleucine</td>
<td>UP</td>
<td>0.00024</td>
<td>0.00016</td>
<td>0.00003</td>
<td>0.00003</td>
<td>1.51*</td>
</tr>
<tr>
<td>Nicotinic acid</td>
<td>UP</td>
<td>0.00021</td>
<td>0.00009</td>
<td>0.00004</td>
<td>0.00005</td>
<td>2.39*</td>
</tr>
<tr>
<td>Alanine (peak A)</td>
<td>UP</td>
<td>0.09383</td>
<td>0.04840</td>
<td>0.00922</td>
<td>0.02302</td>
<td>1.94*</td>
</tr>
<tr>
<td>Valine</td>
<td>UP</td>
<td>0.02484</td>
<td>0.01708</td>
<td>0.00403</td>
<td>0.00160</td>
<td>1.45*</td>
</tr>
<tr>
<td>Aminoacidic acid</td>
<td>UP</td>
<td>0.00023</td>
<td>0.00011</td>
<td>0.00006</td>
<td>0.00004</td>
<td>2.14*</td>
</tr>
<tr>
<td>Nonanoic acid</td>
<td>DOWN</td>
<td>0.00001</td>
<td>0.00003</td>
<td>0.00000</td>
<td>0.00001</td>
<td>0.25*</td>
</tr>
<tr>
<td>Ornithine</td>
<td>UP</td>
<td>0.00124</td>
<td>0.00068</td>
<td>0.00034</td>
<td>0.00009</td>
<td>1.82*</td>
</tr>
</tbody>
</table>

1 Metabolites are in order of statistical significance of the difference between transgenic and wild-type kernels
2 Arbitrary units representing peak areas of biological triplicate
3 **, Student's t-test p-value < 0.01; *, Student's t-test p-value < 0.05
### Table 2: Expression level, annotation and predicted subcellular localization of ZmWri1a target genes

<table>
<thead>
<tr>
<th>#</th>
<th>ID oligonucleotide on array</th>
<th>Gene model (release 4a.53)</th>
<th>Annotation</th>
<th>Ratio OE/WT leaf (array)</th>
<th>Ratio OE/WT leaf (qRT-PCR)</th>
<th>Ratio OE/WT kernel (qRT-PCR)</th>
<th>Predicted subcellular localization (SignalP)</th>
<th>Predicted subcellular localization (Predotar)</th>
<th>Arabidopsis ortholog(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>MZ00042142</td>
<td>GRMZM2G019923_P01</td>
<td>phosphoenolpyruvate-protein phosphotransferase</td>
<td>2.29**</td>
<td>2.61*</td>
<td>4.00*</td>
<td>P</td>
<td>M</td>
<td>no hit</td>
</tr>
<tr>
<td>2</td>
<td>MZ00024552</td>
<td>GRMZM2G033526_P02</td>
<td>pyruvate kinase</td>
<td>1.73**</td>
<td>3.11**</td>
<td>4.20**</td>
<td>P</td>
<td>none</td>
<td>At1g32440^i</td>
</tr>
<tr>
<td>3</td>
<td>MZ00043500</td>
<td>GRMZM2G144730_P01</td>
<td>pyruvate kinase</td>
<td>1.79**</td>
<td>4.89**</td>
<td>3.63**</td>
<td>P</td>
<td>none</td>
<td>At1g32440^i</td>
</tr>
<tr>
<td>4</td>
<td>MZ00024718</td>
<td>GRMZM2G033894_P01</td>
<td>pyruvate dehydrogenase E1 alpha subunit</td>
<td>1.93**</td>
<td>27.33**</td>
<td>4.20**</td>
<td>P</td>
<td>P</td>
<td>At1g01090^i</td>
</tr>
<tr>
<td>5</td>
<td>MZ00016632</td>
<td>GRMZM2G127546_P01</td>
<td>pyruvate dehydrogenase E1 beta subunit</td>
<td>2.99**</td>
<td>1.97*</td>
<td>3.40**</td>
<td>P</td>
<td>P</td>
<td>At2g34590, At1g30120^i</td>
</tr>
<tr>
<td>6</td>
<td>MZ00014741</td>
<td>GRMZM2G088565_P01</td>
<td>pyruvate dehydrogenase E1 beta subunit</td>
<td>1.64**</td>
<td>3.40**</td>
<td>3.84**</td>
<td>P</td>
<td>P</td>
<td>At2g34590, At1g30120^i</td>
</tr>
<tr>
<td>7</td>
<td>MZ00017663</td>
<td>GRMZM2G165176_P01</td>
<td>pyruvate dehydrogenase E2 subunit</td>
<td>1.63**</td>
<td>1.92*</td>
<td>2.57**</td>
<td>none</td>
<td>none</td>
<td>At1g34430</td>
</tr>
<tr>
<td>8</td>
<td>MZ00043050</td>
<td>GRMZM2G058702_P01</td>
<td>pyruvate dehydrogenase E2 subunit</td>
<td>1.71**</td>
<td>2.09**</td>
<td>2.91**</td>
<td>P</td>
<td>P</td>
<td>At3g25860</td>
</tr>
<tr>
<td>9</td>
<td>MZ00017651</td>
<td>GRMZM2G038922_P01</td>
<td>lipoyltransferase</td>
<td>1.60**</td>
<td>3.27**</td>
<td>2.81**</td>
<td>P</td>
<td>P</td>
<td>At4g31050, At1g47580</td>
</tr>
<tr>
<td>10</td>
<td>MZ00056535</td>
<td>GRMZM2G127623_P01</td>
<td>3-ketoacyl-acyl carrier protein synthase III</td>
<td>1.62**</td>
<td>2.69**</td>
<td>5.06**</td>
<td>none</td>
<td>none</td>
<td>At1g62640^i</td>
</tr>
<tr>
<td>11</td>
<td>MZ00016866</td>
<td>GRMZM2G032878_P01</td>
<td>acyl carrier protein 3</td>
<td>1.82**</td>
<td>2.50**</td>
<td>3.14**</td>
<td>P</td>
<td>P</td>
<td>At1g54580, At1g54630</td>
</tr>
<tr>
<td>12</td>
<td>MZ00041636</td>
<td>GRMZM2G102878_P01</td>
<td>oleoyl-acyl carrier protein thioesterase</td>
<td>1.89**</td>
<td>6.48**</td>
<td>1.45**</td>
<td>none</td>
<td>none</td>
<td>At4g13050, At3g25110^i</td>
</tr>
<tr>
<td>13</td>
<td>MZ00017355</td>
<td>GRMZM2G117064_P01</td>
<td>long chain acyl-CoA synthetase 9</td>
<td>2.00**</td>
<td>6.03**</td>
<td>4.77**</td>
<td>P</td>
<td>ER</td>
<td>At1g77590</td>
</tr>
<tr>
<td>14</td>
<td>MZ00040095</td>
<td>GRMZM2G085474_P01</td>
<td>cytosolic acetyl-CoA acetyltransferase</td>
<td>1.60**</td>
<td>2.92**</td>
<td>2.96**</td>
<td>none</td>
<td>none</td>
<td>At5g47720, At5g48230</td>
</tr>
<tr>
<td>15</td>
<td>MZ00042163, MZ00016943</td>
<td>GRMZM2G010596_P01</td>
<td>ketopantoate hydroxymethyltransferase</td>
<td>1.96**</td>
<td>2.37**</td>
<td>3.42**</td>
<td>M</td>
<td>M</td>
<td>At2g46110, At3g61530</td>
</tr>
<tr>
<td>16</td>
<td>MZ00044044, MZ00026553</td>
<td>GRMZM2G145101_P02</td>
<td>beta-type carbonic anhydrase</td>
<td>1.94**</td>
<td>2.16**</td>
<td>5.17**</td>
<td>M</td>
<td>M</td>
<td>At1g58180, At4g33580</td>
</tr>
<tr>
<td>17</td>
<td>MZ00015977</td>
<td>not in release 4a.53</td>
<td>glycerol-3-phosphate dehydrogenase [NAD+]</td>
<td>2.06**</td>
<td>3.11**</td>
<td>9.22**</td>
<td>none</td>
<td>none</td>
<td>At5g40610</td>
</tr>
<tr>
<td>18</td>
<td>MZ00031529</td>
<td>GRMZM2G113875_P02</td>
<td>outer membrane lipoprotein</td>
<td>1.98**</td>
<td>5.00*</td>
<td>1.61**</td>
<td>P</td>
<td>too short</td>
<td>no hit</td>
</tr>
<tr>
<td>19</td>
<td>MZ00039375</td>
<td>GRMZM2G016004_P01</td>
<td>no blast hit</td>
<td>2.02**</td>
<td>1.79*</td>
<td>2.09*</td>
<td>none</td>
<td>no ATG</td>
<td>no hit</td>
</tr>
</tbody>
</table>

---

1 Identification number of the corresponding oligonucleotide deposited on the microarray
2 **, Student's t-test p-value<0.01; *, Student's t-test p-value<0.05
3 Value at 24 DAP
4 Published as target gene of AtWRI1 (Baud et al., 2009; Maeo et al., 2009)
Fig. 1: Phylogenetic tree of WRI proteins
A maximum likelihood phylogenetic tree was generated using all full length grapevine (GSVIV, black), maize (GRMZM, red), rice (Os, blue) and sorghum (Sb, purple) amino acid sequences related to AtWRI1, AtWRI2, AtWRI3 and AtWRI4 from Arabidopsis (At, green) using Treefinder software. Percentage values on each branch represent the corresponding bootstrap probability.
Fig. 2: Expression profiles of ZmWri1a and ZmWri1b in the maize plant

Expression profiles of ZmWri1a and ZmWri1b were established by qRT-PCR in major maize organs (A), during kernel development (B) and in dissected endosperm and embryo (C). Leaf blade (b) or sheath (s) of juvenile (j) or adult (a) leaves, roots of seedlings (s) or plantlets (p), immature (i) or mature (m) ears and tassels as well as kernels at 12 or 35 days after pollination were used in (A), immature (i) kernels of the indicated days after pollination in (B) and dissected endosperms (End) or embryos (Emb) of the indicated days after pollination in (C). Error bars correspond to the standard deviation calculated from technical quadruplicates on organ or tissue pools of various size harvested from two plants.
Fig. 3: Complementation of the Arabidopsis *wri1-4* mutant with maize *ZmWri1a* and *ZmWri1b*

(A) Images of mature seeds of wild-type (Col-0), mutant (*wri1-4*) and complemented mutant (*wri1-4 + ZmWri1a, wri1-4 + ZmWri1b*) (B, C) Seed fatty acid content of Col-0, *wri1-4* and 6 independent *wri1-4* lines complemented with *ZmWri1a* (B) or *ZmWri1b* (C) from maize. For each transformation event 5 distinct plants were cultured. Twenty seeds were harvested from each plant and pooled. Error bars correspond to the standard deviation calculated from 5 technical replicates per pool of 100 seeds.
Fig. 4: Expression profiles of ZmWri1a target genes in ZmWri1a-OE and wild-type kernels

Expression profiles of direct or indirect target genes of Wri1a were established by qRT-PCR in ZmWri1a-OE and wild-type kernels at four developmental stages indicated in DAP (days after pollination). Genes are identified by their MZ000xxxxx number used in Table 2. Error bars correspond to the standard deviation calculated from technical duplicates of the gene of interest and the reference gene on pools of 5 kernels. For each data point the difference between ZmWri1a-OE and wild-type kernels is statistically significant with a p-value <0.05 in Student's t-test.
Fig. 5: Putative metabolic functions of up-regulated ZmWri1a target genes

Enzymatic functions encoded by up-regulated target genes of ZmWri1a and AtWRI1 are in orange, those encoded only by ZmWri1a target genes in yellow and those encoded only by AtWR1 target genes in white.
Fig. 6: AW-boxes in upstream sequences of ZmWri1a target genes

(A) Schematic drawing of the sequence 3 kb upstream of the ATG start codon (arrow) of ZmWri1a target genes. AW-boxes in the sense strand are indicated by black vertical bars shifted to the top, AW-boxes in the antisense strand by blue vertical bars shifted to the bottom (B) Palindromic AW-boxes in the upstream region of MZ00014741 (C) Overlapping AW-boxes in the upstream region of MZ00042142. The fixed bases of the consensus sequence CnTnG(n),GC are in upper case letters.