TaCYP78A3, a gene encoding Cytochrome P450 CYP78A3 protein in wheat (*Triticum aestivum* L.), affects seed size

Meng Ma¹, Qian Wang¹, Zhanjie Li¹, Huihui Cheng¹, Zhaojie Li¹, Xiangli Liu¹, Weining Song²,³, Rudi Appels⁴, Huixian Zhao¹,²,⁵

¹ College of Life Sciences, Northwest A & F University, Yangling, Shaanxi 712100, China.
² State Key Laboratory of Crop Stress Biology in Arid Areas, Northwest A & F University, Yangling, Shaanxi 712100, China.
³ College of Agronomy, Northwest A & F University, Yangling, Shaanxi 712100, China.
⁴ School of Veterinary and Life Sciences, Murdoch University, South Street, Murdoch, WA 6150, Australia
⁵ Corresponding author: Huixian Zhao; Email: hxzhao212@nwsuaf.edu.cn; Telephone number: +86 029 87082262

Running title: *TaCYP78A3* influences wheat seed size.

Keywords: Cytochrome P450, *CYP78A*, *Triticum aestivum*, *TaCYP78A3*, seed size, BSMV-VIGS

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process which may lead to differences between this version and the Version of Record. Please cite this article as an 'Accepted Article', doi: 10.1111/tpj.12896

This article is protected by copyright. All rights reserved.
Summary

There are several studies describing quantitative trait loci (QTL) for seed size in wheat, but the relevant genes and molecular mechanisms remain largely unknown. Here, we report the functional characterization of the wheat \textit{TaCYP78A3} gene and its effect on seed size. \textit{TaCYP78A3} encoded wheat cytochrome P450 CYP78A3 and was specifically expressed in wheat reproductive organs. \textit{TaCYP78A3} activity was positively correlated with the final seed size. Its silencing caused cell number reduction in the seed coat, resulting in a 11\% decrease in wheat seed size, whereas \textit{TaCYP78A3} overexpression induced more cells in the seed coat, leading to a 11\%~48\% increase in \textit{Arabidopsis} seed size. In addition, the cell number of the final seed coat was determined by the \textit{TaCYP78A3} expression level, which affected the extent of integument cell proliferation in the developing ovule and seed. Unfortunately, \textit{TaCYP78A3} overexpression caused a reduced seed set due to an ovule developmental defect in \textit{Arabidopsis}. Moreover, \textit{TaCYP78A3} overexpression affected embryo development by promoting embryo integument cell proliferation during seed development, which also ultimately affected the final seed size in \textit{Arabidopsis}. In summary, our results indicated that \textit{TaCYP78A3} plays critical roles in influencing seed size by affecting the extent of integument cell proliferation. The present study provides direct evidence that \textit{TaCYP78A3} functions in affecting seed size in wheat and contributes to an understanding of the cellular basis of the gene influencing seed development.

Introduction

Seed size is an important agronomic trait that occupies a pivotal position in crop yield. There is a striking diversity of seed size among the plant species of the world, even within the same community (Harper \textit{et al.} 1970, Venable 1992). Seed size in crop is usually represented by the one thousand seed weight and acts as an important trait that has been selected during domestication and modern crop breeding (Shomura \textit{et al.} 2008).
Seed development involves complex processes, including three different growth programs: those of the diploid embryo, the triploid endosperm and the maternal integument, which coordinate to affect the final seed size (Berger et al. 2006, Ohto et al. 2005, Sundaresan 2005). Functional loss of HAiku2 (IKU2) and MINISEED3 (MINI3), encoding a receptor-like kinase and a WRKY transcription factor, respectively, produced small seeds with precocious endosperm cellularization and reduced embryo growth (Garcia et al. 2003, Luo et al. 2005).

IKU1, IKU2 and MINI3 were found to be in a genetic pathway regulated by SHORT HYPOCOTYL UNDER BLUE1 (SHB1), whose gain-of-function mutants produced larger seeds due to enhanced endosperm proliferation and delayed endosperm cellularization (Garcia, et al. 2003, Luo, et al. 2005, Wang et al. 2010, Zhou et al. 2009).

TRANSPARENT TESTA GLABRA 2 (TTG2) and APETALA 2 (AP2) have been reported to promote seed growth by increasing cell expansion in the integuments; the mutants of this gene have defective seed integuments and reduced seed size (Garcia et al. 2005, Jofuku et al. 2005, Johnson 2002). By controlling the seed integument cell proliferation, AUXINRESPONSE FACTOR2 (ARF2) and DA1 (DA means “large” in Chinese) affected the size of the seed coat, which physically limited seed size (Li et al. 2008, Schruff et al. 2006). ABSCISIC ACID DEFICIENT2 (ABA2) regulated embryo and endosperm size by promoting cell proliferation and early cellularization of the endosperm during early seed development and further controlled seed size (Cheng et al. 2014).

Cytochrome P450 (CYP) is one of the largest families of plant proteins (http://drnelson.uthsc.edu/cytochromeP450.html) (Nelson 2006, Nelson et al. 2004). Members of this family have been regarded as important genes in crop improvement (Feldmann 2001). Among the CYP family, the genes of the CYP78A class appear to be involved in plant-specific reactions (Nelson 1999), such as seed development (Adamski et al. 2009, Ito and Meyerowitz 2000). There are six and eight genes encoding CYP78A family proteins in Arabidopsis and rice (Nagasawa et al. 2013), respectively, and several genes have been
identified to affect the process of seed development (Adamski, et al. 2009, Fang et al. 2012, Ito and Meyerowitz 2000, Nagasawa, et al. 2013, Sotelo-Silveira et al. 2013). For example, Arabidopsis CYP78A5 acts independently of other tested maternal factors that influence integument cell proliferation, thereby determining seed size, and its activity level was a limit for seed growth in wild type (WT) (Adamski, et al. 2009). A tomato (Solanum lycopersicum) orthologous gene of CYP78A5/KLUH, SIKLUH, has also been shown to regulate fruit mass by shortening the cell division period (Chakrabarti et al. 2013, Monforte et al. 2014, van der Knaap et al. 2014). CYP78A5 and CYP78A7 played redundant roles in influencing embryo growth in Arabidopsis (Wang et al. 2008). The genetic relationship of CYP78A9 with its closest paralogs, CYP78A6 and CYP78A8, shows a redundant function in controlling floral organ growth and seed development by promoting cell proliferation (Fang, et al. 2012, Sotelo-Silveira, et al. 2013). In rice, it was found that the GE gene encoding CYP78A13 is required for proper size balance between embryo and endosperm (Nagasawa, et al. 2013, Xu et al. 2014, Yang et al. 2013). However, the function of other CYP78A family members, such as CYP78A3, remains unknown.

Wheat (Triticum aestivum L.) is one of the major staple crops for humans. With the increase in the global population, the shortage of foods has become increasingly serious. Therefore, wheat yield improvement has been a critical target in wheat breeding programs. There are several studies describing quantitative trait loci (QTL) for seed size in wheat (Nakamura et al. 2007, Somyong et al. 2011, Williams and Sorrells 2014). Only two seed size-associated genes TaGW2 (Su et al. 2011, Yang et al. 2012) and TaCwi (Ma et al. 2012a) were isolated and chromosome-located in wheat. To date, there is no report of wheat CYP78A genes influencing seed development. Here, we isolated and characterized a CYP78A member, TaCYP78A3 in wheat, and demonstrated its functions in affecting seed size. TaCYP78A3 silencing during wheat ovule and seed development caused a reduced cell number in the seed coat, resulting in a decrease in seed size. TaCYP78A3 overexpression in the developing ovule and/or seed induced more cells in the seed coat, leading to an enlarged seed in Arabidopsis.

This article is protected by copyright. All rights reserved.
Moreover, *TaCYP78A3* expression also influenced embryo development by promoting embryo integument cell proliferation during seed development in *Arabidopsis*. Our study provides direct evidence that *TaCYP78A3* affects seed size and contributes to the understanding of the cellular basis of the genetic influence of wheat seed development.

**Results**

Cloning and characterization of *TaCYP78A3* in wheat

To clone genes highly homologous to *Glycine max* *GmCYP78A3* (GenBank accession AF022463.1) in wheat, we used the protein sequence (GenBank accession AAB94592.1) of *GmCYP78A3* as a probe to BLAST search the Non-redundant protein sequences (nr) database, and we obtained a predicted full-length mRNA (GenBank accession KD229739.1) for the predicted protein *TuCYP78A3* (GenBank accession EMS50564.1) in *Triticum urartu*. The predicted mRNA sequence was used to design PCR primers to amplify *TaCYP78A3* in wheat cv. Shaan 512 with a large seed size. Then, we subcloned the PCR products and randomly chose 30 gDNA clones to sequence. Finally, three different sequences were obtained and temporarily named *TaCYP78A3*-1, *TaCYP78A3*-2 and *TaCYP78A3*-3.

The sequence alignment of the above three *TaCYP78A3* sequences with the chromosome-based draft sequence of hexaploid bread wheat (*Triticum aestivum*) genome (The International Wheat Genome Sequencing Consortium, http://www.wheatgenome.org/) (Mayer et al. 2014) showed that *TaCYP78A3*-1, *TaCYP78A3*-2 and *TaCYP78A3*-3 were located on wheat chromosome 7AS, 7BS and 7DS, respectively. Furthermore, the chromosomal localization results were verified by PCR using various wheat species containing the AA, AABB, DD and AABBD genome and Chinese Spring Nulli-tetrasomic lines (NT) and Double-ditelosomic lines (DT) (Figure S1). Therefore, the above three *TaCYP78A3* sequences were designed as *TaCYP78A3*-A, *TaCYP78A3*-B, *TaCYP78A3*-C.
and TaCYP78A3-D, respectively, and their GenBank accessions are KP768392, KP768393 and KP768394. TaCYP78A3-A, -B and -D all contained two exons and one intron (Figure 1a). All of the deduced proteins contained a hydrophobic region in the N terminus and putative oxygen and heme-binding domains that are characteristic of CYP78A family members (Figure 1a-b) (Nebert and Gonzalez 1987), and these characteristics were identical to those of other CYP78A family members. The gDNA size of TaCYP78A3-A, -B and -D were 2084, 2110 and 2072 bp, respectively, and the differences in the sequence length were caused by variations in introns and 3'-UTRs (Figure 1a). Thus, the entire deduced proteins of TaCYP78A3-A, -B and -D were exactly the same length (551 aa) and were highly similar in amino acid sequence (more than 97% identity) (Figure 1b, S2).

In the CYP78A family, several genes have been isolated (Figure 1c) and identified to influence the process of seed development (Adamski, et al. 2009, Sotelo-Silveira, et al. 2013). Phylogenetic analysis of TaCYP78A3 (putative amino acid sequences of TaCYP78A3-B) and other reported CYP78A family proteins from various plant species showed that TaCYP78A3 formed a cluster with Glycine max GmCYP78A3 and Triticum urartu TuCYP78A3 (GenBank accession EMS50564.1) (Ling et al. 2013), with 94% identity to TuCYP78A3 and 61% identity to GmCYP78A3 (Figure 1c). Moreover, an analysis plotting the BAC scaffold in the 7A assembly of T. aestivum (Chinese Spring) carrying the TaCYP78A3-A gene against the T. urartu assembly showed a high similarity (Figure S3), suggesting that the CYP78A3 gene was highly conserved in the two Triticum species. According to the above results, we concluded that TaCYP78A3 encoded a cytochrome P450 protein, which was the ortholog of GmCYP78A3 and TuCYP78A3.

Furthermore, expression sequence tags (ESTs) (for example, GenBank accession CK208119.1, 99% identity with TaCYP78A3) were found in the developmental stages of wheat spike and seed in the NCBI database. Similarly, the high expression of TaCYP78A3 was detected in young panicles (at 2 to 4 cm) and immature seeds

This article is protected by copyright. All rights reserved.
(0–5 days after flowering) of wheat cv. Shaan 512 by quantitative real-time PCR in the present study (Figure S4). These results suggested that TaCYP78A3 may play important roles in wheat seed development.

Silencing of TaCYP78A3 leads to a decrease in wheat seed size

To determine whether TaCYP78A3 affects seed development, we silenced the expression of all three TaCYP78A3 genes in wheat cv. Shaan 512 and Shaanmai 159, with large and small seed sizes, respectively, by using the Barley stripe mosaic virus-induced gene silencing (BSMV-VIGS) technique as described previously (Ma et al. 2012b). The BSMV construct (BSMV:TaCYP78A3) carrying a 344 bp fragment (TaCYP78A3-VIGS) from the TaCYP78A3–B cDNA sequence was used to silence all three of the genes in wheat seed (Figure S5, S6), and the infectivity and efficacy of BSMV-VIGS in both wheat cultivars was investigated as described previously (Ma, et al. 2012b). TaCYP78A3-specific silencing could be easily achieved due to the low similarity of nucleotide sequences among the CYP78A family (approximately 50%~70%). Ten plants of each wheat cv. were inoculated with BSMV:TaCYP78A3, BSMV:PDS (BSMV vector carrying 185 bp fragment of a barley phytoene desaturase gene PDS) (Ma, et al. 2012b) and BSMV:00 (empty vector) construct respectively, onto the spikes at the heading stage. The experiment was conducted five times. Seed photobleaching was observed in all the plants infected only with the BSMV:PDS construct at 20 days post inoculation (dpi) (the representative phenotype of wheat cultivar Shaan 512 is shown in Figure 2), suggesting that the endogenous PDS gene was silenced. RT-PCR analysis showed that the transcript levels of PDS and TaCYP78A3 were reduced in BSMV:PDS and BSMV:TaCYP78A3-infected plants, respectively, compared to those of BSMV:00-infected plants (Figure 2a). These results indicated that VIGS in wheat seeds was effective in our experiments. The mature seeds collected from the middle spikes of BSMV:TaCYP78A3-infected plants were shorter and narrower.
than those from BSMV:00 and BSMV:PDS-infected plants (Figure 2b). Statistical analysis showed that the size (represented by the projective area of a seed), length and width of the mature seed from BSMV:TaCYP78A3-infected plants were significantly reduced by ~11%, ~9% and ~7%, respectively, compared to those of BSMV:00-infected plants (Figure 2c, S7a, S7c). Moreover, the embryos from BSMV:TaCYP78A3-infected plants were also reduced in size (Figure 2d). TaCYP78A3 silencing caused the seed size to be reduced but did not affect the seed set in wheat. Similar results were also observed in wheat cv. Shaanmai 159 (Figure S7). Thus, it is plausible that TaCYP78A3 might influence seed size in wheat.

**Silencing of TaCYP78A3 reduces cell proliferation in wheat seed coats**

By controlling the extent of the integument cell proliferation in developing ovules and/or seeds, some CYP78A family members, such as CYP78A5, CYP78A6 and CYP78A9, determine the size of the seed coat, which ultimately appears to limit seed growth (Adamski, et al. 2009, Fang, et al. 2012, Sotelo-Silveira, et al. 2013). To verify whether the decreases in the seed size in TaCYP78A3-silenced plants were also caused by decreased cell numbers in the seed coat, we investigated the cell number in the outer layer of the seed coat (outer seed integument) in BSMV:00 and BSMV:TaCYP78A3-infected plants. Indeed, the cell number of the outer seed coat was reduced by ~11% in the seeds from BSMV:TaCYP78A3-infected plants at 20 dpi compared with those of BSMV:00-infected plants (Figure 3a-c). However, the length of the seed coat cell was slightly increased in BSMV:TaCYP78A3-infected plants (Figure 3c, 3f). Because CYP78As were involved in growth-promoting reactions by promoting cell proliferation, a plausible hypothesis is that the elongated cells in the seed coat are generated due to more space caused by reduced numbers of seed integument cells in seed development processes. Therefore, a 11% reduction of the final seed size in TaCYP78A3-silenced plants is most likely
attributable to an 11% decrease in the cell number, resulting in a small seed coat (seed coat size was represented by seed coat perimeter, which is calculated by the seed coat cell number times the cell length, n=10) (Figure 3d-f).

The cell number of the final seed coat is determined by the integument cell proliferation during ovule and seed development (Haughn and Chaudhury 2005). Therefore, we further investigated whether TaCYP78A3 silencing affected the integument cell proliferation in the developing ovule and seed by inoculating with BSMV:TaCYP78A3 construct onto wheat spikes at approximately 7 days prior to anthesis as described previously (Ma, et al. 2012b). The silencing effect covered almost all stages of the ovule (0-7 dpi) and seed development (8-25 dpi). The results showed that TaCYP78A3 silencing caused a slower integument cell proliferation during ovule and seed development than in control (BSMV:00 infected) plants (Figure 3h) and then resulted in a reduced seed size (Figure 3g). These results suggested that TaCYP78A3 likely played important roles in seed development and growth by affecting the extent of integument cell proliferation in the developing ovule and seed of wheat, which ultimately appeared to affect the final seed size.

**Overexpression of TaCYP78A3 causes seed and reproductive organ overgrowth in *Arabidopsis thaliana***

To further investigate the functions of TaCYP78A3 in influencing seed size, we explored the effects of TaCYP78A3 overexpression under three promoters with different expression patterns in developing Arabidopsis ovules or/and seeds. The CYP78A9 promoter (pCYP78A9) was specifically active in developing Arabidopsis ovules and seeds (Ito and Meyerowitz 2000, Sotelo-Silveira, et al. 2013), similar to the TaCYP78A3 promoter, which was specifically active in young panicles and immature wheat seeds (Figure S4). However, the 35S promoter showed relatively low activity in the ovules but drove high expression in the developing seeds (Hraška
et al. 2008, Jenik and Irish 2000). To further increase TaCYP78A3 expression only in ovules, we also expressed TaCYP78A3 by using the INNER NO OUTER promoter (pINO) (Villanueva et al. 1999). Consequently, at least three independent Arabidopsis transgenic lines bearing each construct were obtained. The mature seeds, developmental siliques and inflorescences of transgenic lines pCYP78A9::TaCYP78A3-3, 35S::TaCYP78A3-4 and pINO::TaCYP78A3-2 are shown as representatives for Arabidopsis transgenic lines bearing each construct (Figure 4a-c). The seed size of all TaCYP78A3 transgenic lines was significantly larger than that of WT plants (Figure 4a, 4i), and the seed length and width also increased accordingly (Figure S8a). These results suggested that the overexpression of TaCYP78A3 in ovules (as pNIO::TaCYP78A3 transgenic plants), in seeds (as 35S::TaCYP78A3 transgenic plants), or in both tissues (as pCYP78A9::TaCYP78A3 transgenic plants) was able to promote seed growth.

We also observed that the overexpression of TaCYP78A3 in all the three genotypes of transgenic plants caused different degrees of overgrowth in the reproductive organs, which manifested as wider, shorter and more plentiful siliques (Figure 4b, 4k, 4l); larger and more flowers (Figure 4c, 4d); and a longer inflorescence stem (Figure 5a) compared to WT. Similar cases were also observed in CYP78A5 and CYP78A9-overexpressed plants (Adamski, et al. 2009, Anastasiou et al. 2007, Sotelo-Silveira, et al. 2013). To conduct the histological localization of TaCYP78A3, TaCYP78A3–containing expression vectors were developed to express TaCYP78A3-GUS fusion proteins. GUS signals in the transgenic plants of the three genotypes were observed mainly in the developing ovules and/or seeds as mentioned above (Figure 4e-g). RT-PCR analysis further confirmed that TaCYP78A3 was significantly expressed in the inflorescence of transgenic T3 lines bearing different constructs (Figure 4h).

Moreover, we found that the extent of overgrowth was different among the TaCYP78A3 transgenic lines. Therefore, three pINO::TaCYP78A3 transgenic lines with different extents of overgrowth phenotypes (Figure
5a-b) were used for further investigation. RT-PCR analysis combined with phenotype investigation indicated that the effects of overgrowth were associated with TaCYP78A3 expression levels in these transgenic plants; the higher expression levels of TaCYP78A3 resulted in the longer main inflorescence stem, the shorter siliques per main inflorescence and the fewer but larger seeds per silique (Figure 5a-h, S8b). Moreover, it was also found that TaCYP78A3 transgenic plants had a longer growth period (2 weeks more than WT) (Figure S9). Therefore, TaCYP78A3 overexpression might increase longevity in Arabidopsis.

The overexpression of TaCYP78A3 caused an obviously increased seed size but significantly reduced seed set (Figure 5c, 5e, 5f). For example, pINO::TaCYP78A3-3 exhibited no fertility (Figure 5e-f). We demonstrated that the reduced seed set per silique in TaCYP78A3-overexpressing plants was caused by defects on the maternal side, as manually pollinating the transgenic plants with WT pollen could not rescue the seed number, and pollinating WT plants with pollen from the TaCYP78A3-overexpressing plants produced normal seed numbers per silique. To further explore the reason for the reduced seed set in TaCYP78A3-overexpressing plants, we observed the morphological and cytological characteristics of the developmental ovule. The outer integuments of all abnormal ovules from pINO::TaCYP78A3-2 transgenic plants consisted of more but smaller cells compared with those of WT (Figure S10a-b). Some defective ovules (38/50) showed different levels of wrinkle, asymmetry and reduction in size and then failed to further develop, resulting in female sterility (Figure 5c, S10a-b). Other ovules (12/50) showed a full shape with larger ovule integuments and further developed into mature seeds that were larger than those of the WT (Figure 5c, S10a-d). Therefore, the reduction of seed set was most likely caused by a developmental defect of the ovule, which was likely due to TaCYP78A3 overexpression promoting the overproliferation of ovule integument cells. Similar conclusions related to female sterility were reported in cyp78a5 mutants and transgenic plants of CYP78A5 (Zondlo and Irish 1999) or CYP78A9 overexpression (Ito and Meyerowitz 2000, Sotelo-Silveira, et al. 2013).

This article is protected by copyright. All rights reserved.
TaCYP78A3 affects seed size by promoting integument cell proliferation in Arabidopsis thaliana

In above experiments of TaCYP78A3 silencing in wheat, we speculated that the effects of TaCYP78A3 on seed size were mediated by promoting integument cell proliferation in the developing seed and ovule, which ultimately affected seed coat size. Seed coat size is determined by the cell number and area per cell in the seed coat. To confirm whether the increased seed size in TaCYP78A3-overexpressed Arabidopsis plants was caused by TaCYP78A3 promoting integument cell proliferation of the seed coat, we analyzed the cell number and cell area of the outer coat of mature seeds from different TaCYP78A3 transgenic lines. Indeed, the results showed that TaCYP78A3 overexpression led to increased cell numbers in the seed coat (Figure 6a). The increased cell number was positively associated with the TaCYP78A3 activity level and final seed size (Figure 4j, 6a, 6b, S8c) but was not correlated with the reduced overall number of seeds per plant (Figure S8d). Interestingly, the cell area of the seed coat in pINO::TaCYP78A3-2, -5 transgenic lines was reduced by approximately 9% compared with that of WT, but there were no significant changes between other transgenic lines (pCYP78A9::TaCYP78A3 or 35S::TaCYP78A3) and wild type plants (Figure 6c). Therefore, we postulated that the increased seed size in TaCYP78A3 transgenic plants was largely attributed to the increased cell number.

To further understand how TaCYP78A3 affects the cell number of the seed coat, we analyzed features of cell proliferation in the developing ovule and seed of pCYP78A9::TaCYP78A3-3 transgenic plants. TaCYP78A3 overexpression using the pCYP78A9 promoter caused enlarged outer ovule integument during ovule development (Figure S10a, S10d), which was attributed to a 8% increase in the cell numbers of the outer ovule integument (Figure S10b-c). A similar phenotype was observed in CYP78A5 or CYP78A9-overexpressing plants (Adamski, et al. 2009, Sotelo-Silveira, et al. 2013). Furthermore, we investigated the cell numbers of the outer seed coat from pCYP78A9::TaCYP78A3-3 transgenic lines and the WT plants during seed development at 1, 2, 3 days after flowering (DAF), and found that TaCYP78A3 overexpression caused a significant increase in cell number (Figure 6d, S8c).

This article is protected by copyright. All rights reserved.
4, 6 and 8 days after pollination (dap), which were in accordance with the stages described by Cheng et al. (Cheng, et al. 2014). The cell proliferation rate in the outer seed coat of pCYP78A9::TaCYP78A3-3 transgenic plants was faster than that of WT plants at the initial seed development stage (from 1 dap to 4 dap) but reduced to the same level as that of WT 6 dap (Figure 6d, 6e), leading to a 4% increase in the cell number of the seed coat in pCYP78A9::TaCYP78A3-3 transgenic plants (Figure 6f). Therefore, combined with the silencing of TaCYP78A3 leading to the reduced cell number of the ovule and seed integument in wheat (Figure 3h) and the overexpression of TaCYP78A3 promoting seed growth in Arabidopsis (Figure 4a, 4i), we concluded that the functions of TaCYP78A3 in influencing seed size were fulfilled by promoting integument cell proliferation during ovule and seed development. Similar results were also described in ARF2 (Schruff, et al. 2006) and DA1 (Li, et al. 2008), which act by restricting cell proliferation in the seed integuments to affect the final seed size.

In the above experiments, we found that TaCYP78A3 silencing led to reduced embryo size in wheat (Figure 2d), suggesting that the growth-promoting effects of TaCYP78A3 may also exist in the embryo. To further determine whether TaCYP78A3 influenced embryo development, we analyzed the number and size of the cotyledon epidermis cells of pCYP78A9::TaCYP78A3 and pINO::TaCYP78A3 transgenic lines. The cell number of the cotyledon epidermis in transgenic lines pCYP78A9::TaCYP78A3-3 and -7 were increased by approximately 13%, respectively, compared to WT, whereas no significant difference was observed in the cell size of the cotyledon epidermis between WT and transgenic lines (Figure 7a-7i), suggesting that embryo integument cell proliferation rather than cell elongation was the major cause of the enlarged embryo. However, further observation revealed that pINO::TaCYP78A3-2 and -5 transgenic lines with TaCYP78A3 expressed only in the ovule had a significantly enlarged cell size but a similar cell number in the embryo integument compared with WT (Figure 7d, 7f, 7h, 7i). This result suggests that the dramatically enlarged cotyledon from pINO::TaCYP78A3 transgenic plants was not the effect of TaCYP78A3 on embryo growth but rather the
enlarged seed size, which was caused by the effects of *TaCYP78A3* on promoting integument cell proliferation in the developing ovule.

Taken together, we concluded that *TaCYP78A3* influenced embryo development by promoting integument cell proliferation. Moreover, the increased seed size in *TaCYP78A3*-overexpressing plants was attributed to the combined effects of the enlarged embryo and seed coat, which were mediated by the *TaCYP78A3* activity level affecting the extent of integument cell proliferation in *TaCYP78A3*-overexpressing plants.

*TaCYP78A3* may function independently of other maternal factors that influence integument cell proliferation

By influencing maternal integument cell proliferation, *Dal*, *TTG2*, *ARF2* and *AP2* affect the size of the seed coat, which physically limits the final seed size (Garcia, et al. 2005, Li, et al. 2008, Ohto et al. 2009, Schruff, et al. 2006). To determine whether *TaCYP78A3* interacted genetically with these previously identified maternal factors known to affect seed growth, we compared the expression of these genes in *pINO::TaCYP78A3-2 and -5* transgenic lines with that in WT plants during seed development. The expression levels of these tested genes in the transgenic lines were similar to those in the WT plants (Figure S11), suggesting that *TaCYP78A3* may act independently of *Dal*, *TTG2*, *ARF2* and *AP2*.

It was reported that *CYP78A5* and *CYP78A9* acted maternally to promote cell proliferation in integuments (Adamski, et al. 2009, Sotelo-Silveira, et al. 2013). To test whether *TaCYP78A3* was redundant with *CYP78A5* and *CYP78A9* in influencing seed size, we obtained *TaCYP78A3* transgenic lines in homozygous *cyp78a5* (Salk_024697C) or *cyp78a9* (Salk_066588C) mutant (Figure S12) backgrounds, respectively. The overexpression of *TaCYP78A3* in homozygous *cyp78a5* and *cyp78a9* mutant backgrounds caused taller plants.
and larger seeds compared with the corresponding mutants (Figure S13a-d). However, the increased seed size of

\textit{pCYP78A5/9::TaCYP78A3;\textit{cyp78a5/9}} transgenic plants showed a simple additive effect but not a genetic
interaction between \textit{TaCYP78A3} and \textit{CYP78A5/9} (Figure S13c-d). In addition, the more branched phenotype of
the \textit{cyp78a5} mutant could not be rescued with \textit{TaCYP78A3} overexpression (Figure S13a). Thus, \textit{TaCYP78A3}
may function in influencing seed and reproductive organ development but is non-redundant with \textit{CYP78A9} and
\textit{CYP78A5}.

\textit{TaCYP78A3} expression level influences seed yield

To investigate the effect of \textit{TaCYP78A3} expression levels on seed yield, we further tested the seed weight of
\textit{TaCYP78A3}-silenced wheat plants and \textit{TaCYP78A3}-overexpressing transgenic lines of \textit{Arabidopsis}.
\textit{TaCYP78A3} silencing led to the 100-seed weight and total seed weight per panicle being reduced by
approximately 10\% due to the 11\% reduced seed size in wheat (Figure S14a, S14c). In contrast, \textit{TaCYP78A3}
overexpressing in a WT background also caused a reduction in the total \textit{Arabidopsis} seed weight per plant
(Figure S14d), although it caused greatly increased seed size (Figure 4i) and 100-seed weight (Figure S14b).
This result may occurred because the increased effects of \textit{TaCYP78A3} on seed size at the whole-plant level were
offset by the reduced seed number per silique and per plant (Figure 4j), and the reduced effects of \textit{TaCYP78A3}
on seed set were even stronger than the increased effects on seed size (Figure 5e, 5f). We further investigated the
harvest index (the proportion of total seed yield to total aerial biomass; \textit{n}=20) of \textit{TaCYP78A3}-silenced or
-overexpressed plants. \textit{TaCYP78A3} expression alteration led to the harvest index decreasing in both wheat and
\textit{Arabidopsis} (Figure S14e, S14f). Thus, either \textit{TaCYP78A3} silencing or overexpression would cause a reduction
in the seed yield due to the reduced seed size or seed set.

This article is protected by copyright. All rights reserved.
Discussion

CYP78A members act on reproductive organ development in plants


However, it was found that CYP78A family members were involved in reduced fertility; either the functional deletion or overexpression of CYP78As led to a reduced seed set (Adamski, et al. 2009, Fang, et al. 2012, Ito and Meyerowitz 2000, Sotelo-Silveira, et al. 2013). All these reductions in fertility involved the female reproductive organ ovule. For example, Adamski and coworkers (Adamski, et al. 2009) found that the reason for the reduced seed set per silique in plNO::KLU plants appeared to be a defect on the maternal side. A recent study indicated that CYP78A5 was involved in chromosome paring during female meiosis, most likely by affecting the normal expression pattern of DMC1, which encodes a RecA (Recombinase A) homolog and is expressed specifically in ovules during female meiosis (Zhao et al. 2014). Similarly, CYP78A9 overexpression or functional deletion resulted in female sterility in Arabidopsis due to ovule development defects (Ito and Meyerowitz 2000, Sotelo-Silveira, et al. 2013). In our experiment, the overexpression of TaCYP78A3 also led to decreased fertility due to ovule development defects that might be attributed to over-proliferation of integument cells during ovule development (Figure S10), resulting in seed set reduction. But the effects of TaCYP78A3 on...
seed size were not attributed to the reduced seed number in TaCYP78A3-overexpressing plants, because the limitation of seed number per plant did not increase the seed size in either WT or pINO::TaCYP78A3-1 plants (Figure S8e). However, TaCYP78A3 silencing did not obviously affect the seed set in wheat, which is likely due to TaCYP78A3 silencing occurring instantaneously and not completely in BSMV:TaCYP78A3-infected wheat plants. It was thought that the overexpression of TaCYP78A3 enhanced seed size in transgenic Arabidopsis by promoting integument cell proliferation, which resulted in the enlarged embryo and seed coat. The TaCYP78A3 expression levels affected the extent of integument cell proliferation, which influenced ovule development and resulted in a reduced seed set in TaCYP78A3-overexpressing plants. Thus, the expression levels of CYP78A family members significantly affected plant reproductive organ development and female fertility, which requires normal expression patterns of each CYP78A member during ovule development.

**Genes associated with seed size and yield in wheat**

Seed size and weight are important traits under selection during domestication and modern crop breeding (Shomura, et al. 2008). In wheat, a large number of QTLs controlling seed size and yield components were identified (Nakamura, et al. 2007, Somyong, et al. 2011, Williams and Sorrells 2014). For example, thirty QTLs for seed yield were detected on 10 chromosomes. Most of them were mapped onto the chromosomes of homologous groups 2, 5 and 7 (Su et al. 2006). Interestingly, the three TaCYP78A3 genes were also located on the short arms of the homologous group 7 chromosome. However, the genes TaGW2 (Su, et al. 2011, Yang, et al. 2012) and TaCwi (Ma, et al. 2012a), associated with seed size, were isolated and characterized by QTL analysis. TaGW2, like rice grain width-associated gene 2 (OsGW2) (Song et al. 2007), mainly affected the one-thousand seed weight for its involvement in seed development. TaCwi-A1, like the rice cell wall invertase gene (OsCwi) (Hirose et al. 2002), was characterized by allelic variations and QTL analysis. In the present study, this article is protected by copyright. All rights reserved.
three TaCYP78A3 genes, orthologues of GmCYP78A3 and TuCYP78A3, were isolated and characterized in wheat. Our results indicated that the seed size of the TaCYP78A3-silenced plant was 11% smaller than the seed size in the control plants due to a reduction cell number of seed coat in wheat (Figure 3), whereas the seed size of the TaCYP78A3-overexpressing plant was 11%~48% larger than the seed size of the wild type plant due to TaCYP78A3 promoting integument cell proliferation in Arabidopsis (Figure 6a-c). It was interesting that similar effects of TaCYP78A3 silencing on seed size were observed in both wheat cv. Shann 512 (large seed) and Shaanmai 159 (small seed) (Figure S7), suggesting that TaCYP78A3 play crucial roles in wheat seed development. Thus, TaCYP78A3 activity was positively associated with final seed size, and the effects of TaCYP78A3 on final seed size are mediated by promoting integument cell proliferation. The present study provided direct evidence that TaCYP78A3 affects the seed size of plants and contributes to the understanding of the cellular basis of genes influencing seed development in wheat.

**Experimental procedures**

**Plant material and growth**

The wheat cultivars used in this study were ‘Shaan 512’ and ‘Shaanmai 159’, with a large-seed and a small-seed phenotype, respectively. The growing conditions for the wheat seedlings and VIGS experiments were as described previously (Ma, et al. 2012b). Different wheat species, including Triticum urartu (AA), Aegilops tauschii (DD), Triticum turgidum ssp durum (AABB) and Triticum aestivum L. (AABBDD), were used to investigate the genomic origin of the cloned sequences by gene-specific PCR (Figure S1). Twenty-four nulli-tetrasomic (NT) lines and 16 double ditelosomic (DT) lines of wheat cv. “Chinese Spring” were used for chromosome location testing (Figure S1).

The Arabidopsis thaliana ecotype Columbia-0 (Col) was used as the wild type (WT). All mutants used in this study, including cyp78a5
(Salk_024697C) and cyp78a9 (Salk_066588C), were in the Columbia-0 background and were obtained from the ABRC (Ohio State University, Columbus). The protocol for Arabidopsis thaliana transformation was conducted according to Zhang et al. (Zhang et al. 2006).

Arabidopsis thaliana was grown under suitable conditions with a 16-h-light/8-h-dark cycle in 70% relative humidity (light intensity 170 to 200 mmol m$^{-2}$ s$^{-1}$) at 20 to 22°C for 70 d until the last siliques on the inflorescences were dry.

TaCYP78A3 cloning and sequence analysis

TaCYP78A3 was cloned by the homology cloning strategy. In brief, TaCYP78A3 was amplified with the genomic DNA of wheat cv. Shaan 512 and PCR primers (TaCYP78A3-F and TaCYP78A3-R), which were designed based on TaCYP78A3-predicted mRNA (GenBank accession KD229739.1), and the PCR products were subcloned. Thirty clones were randomly chosen to sequence.

To explore the chromosome location of the cloned sequences, genomic DNA samples from nulli-tetrasomic (NT) and double ditelosomic (DT) lines from Chinese Spring (CS) and wheat species containing the AA, DD, AABB and AABBDD genomes were used to perform gene-specific PCR (Figure S1).

Detailed protocols for sequence analysis are described in Methods S1.

Total RNA extraction and cDNA synthesis

To clone wheat TaCYP78A3 cDNA, total RNA was extracted from seeds one week post anthesis using RNAiso-mate (TaKaRa, Dalian, China). cDNA was recovered from total RNA using Oligotex-dT30 (TaKaRa, Dalian, China). TaCYP78A3 cDNA was cloned using the specific primers TaCYP78A3-R and TaCYP78A3-F with PrimeScript™ II 1st Strand cDNA Synthesis Kit (TaKaRa, Dalian, China). Three TaCYP78A3 cDNA sequences were obtained from 20 monoclones, with a complete open reading frame (ORF) encoding a polypeptide.

For analysis of gene silencing in BSMV-VIGS wheat spikes or seeds, the spikelets or seeds on the middle part of the BSMV-infected spikes were collected at 2, 4, 8, 15, 22 and 30 dpi. The total RNA of the spikelets or seeds was prepared as above.

BSMV-mediated TaCYP78A3 gene silencing

This article is protected by copyright. All rights reserved.
In vitro transcription of viral RNA and plant inoculations were performed as previously described (Ma, et al. 2012b). Ten plants of each wheat cultivar were infected with each BSMV+ target construct as the treatment groups, with inoculation with BSMV:00 as controls in each experiment. This experiment was repeated five times, and the data were averaged.

**RT-PCR and quantitative real-time RT-PCR**

Target gene expression was quantified by reverse transcription-PCR (RT-PCR). The expression patterns of \( \text{TaCYP78A3} \) in wheat were analyzed by quantitative real-time PCR (qRT-PCR) as previously described (Pfaffl et al. 2002). The details of the RT-PCR and qRT-PCR are described in Method S1.

**Morphological and cytological characterization and GUS staining of transgenic plants**

Morphological analysis of the plant growth was performed as previously described (Disch et al. 2006, Sotelo-Silveira, et al. 2013). Organ, embryo and seed sizes and seed set, as well as the plant growth dynamics, were measured as previously described (Disch, et al. 2006, Jiang et al. 2013). To measure the ovule size, the dissected pistil was stained using a 0.01% solution of Fluorescent Brightener 28 (SIGMA, China) as previously described (Adamski, et al. 2009). The cell number and cell size in the outer integument of the seed coat and cotyledons were measured as previously described (Alonso-Blanco et al. 1999). For the biomass and seed yield measurement, transgenic plants and the five WT plants were grown in identical pots and cultured under well-watered and appropriate lighting conditions. Values are represented as the means ± SE (bar charts). Each value represents measurements of at least 20 seeds from at least five plants. Detailed protocols for the measurement of seed, embryo, cell numbers and histochemical staining for \( \text{TaCYP78A3} \) activity and microscopy are described in Methods S1.

**PCR-based genotyping**

Identification of the \( \text{cyp78a5} \) or \( \text{cyp78a9} \) mutant allele was performed by PCR analysis using the LBa1 primer (5'-TGGTTACGTAGTGCCATCG -3') on the T-DNA left border. The \( \text{CYP78A5} \) wild-type allele was identified using the This article is protected by copyright. All rights reserved.
SALK_024697-LP and SALK_024697-RP primers; the CYP78A9 wild-type allele was identified using the SALK_066588-LP and SALK_066588-RP primers (Figure S12).

Acknowledgements

This work was financially supported by grants from the National Nature Science Foundation of China (30871578 and 31471482). We thank Shengwu Hu of the College of Agronomy in Northwest A & F University for helpful comments on the manuscript.

Short legends for Supporting Information

Method S1

Figure S1. Chromosome location of TaCYP78A3.

Figure S2. Multiple protein sequence alignment for three TaCYP78A3 genes, TaCYP78A3 and GmCYP78A3.

Figure S3. Dot-matrix plot of the CS (Chinese Spring) BAC scaffold in the 7A assembly (X axis) carrying the TaCYP78A3-A gene, with the corresponding region on the Triticum urartu assembly (Y axis).

Figure S4. TaCYP78A3 expression patterns in different tissues or organs (quantitative real-time PCR data).

Figure S5. Schematic organization of the Barley stripe mosaic virus (BSMV) genomes and the inserts used for BSMV-virus-induced gene silencing (VIGS).

Figure S6. Multiple DNA sequence alignment for TaCYP78A3-VIGS and the three TaCYP78A3 genes.

Figure S7. Comparison of the length and width of seeds from BSMV:00, BSMV:PDS and BSMV:TaCYP78A3-infected wheat plants.

Figure S8. Additional Arabidopsis thaliana seed measurements in response to increased TaCYP78A3 activity.

Figure S9. Overexpression of TaCYP78A3 may increase longevity in Arabidopsis.

Figure S10. TaCYP78A3 promotes cell proliferation in the ovule integument of Arabidopsis.

Figure S11. Quantitative real-time PCR analysis of the expression levels of Da1, TTG2, ARF2 and AP2 in wild-type Col and...
pINO::TaCYP78A3-2, -5 transgenic lines, respectively.

Figure S12. Identification of homozygous mutants in Arabidopsis.

Figure S13. Phenotype of Arabidopsis deletion mutant cyp78a5 or cyp78a9 could not be rescued with overexpression of TaCYP78A3.

Figure S14. TaCYP78A3 expression level leads to seed yield variation.

Table S1. Primers used for TaCYP78A3 analyses.

References


This article is protected by copyright. All rights reserved.


**FIGURE LEGENDS**

**Figure 1.** *TaCYP78A3* gene structure and phylogenetic analysis of the *CYP78A* family. (a) Schematic diagrams of *TaCYP78A3-A*, *TaCYP78A3-B* and *TaCYP78A3-D* gene structures. Black boxes: exons; black lines in the middle: introns; split: absent regions in *TaCYP78A3-A* and *TaCYP78A3-D* compared with *TaCYP78A3-B*; oval, rectangular and circular boxes: regions encoding hydrophobic domains, oxygen binding motifs and heme binding motifs, respectively. (b) Clustal W alignment of *TaCYP78A3-B* deduced protein with those of *TaCYP78A3-A* and *TaCYP78A3-D*. (c) The phylogenic tree of *TaCYP78A3* and other reported *CYP78A* family members. This article is protected by copyright. All rights reserved.
proteins from *Triticum urartu*, *Arabidopsis thaliana*, *Oryza sativa*, *Zea mays*, *Pinus radiata*, *Phalaenopsis* and *Glycine max* constructed by the neighbor-joining method.

**Figure 2.** The silence of *TaCYP78A3* reduced seed size in wheat cv. Shaan 512. (a) Reverse transcription-PCR analysis with total RNA isolated from the seeds of BSMV:00, BSMV:PDS (*phytoene desaturase*) and BSMV:TaCYP78A3-infected plants at 20 days post inoculation (dpi). (b) Images of mature seeds. (c) Comparison of the seed length and seed width of mature seed from the middle spikes; three plants from each of five replicates and two grains from each plant were used for measurement (n=30). (d) Mature embryos.

**Figure 3.** Effects of *TaCYP78A3* silencing on seed coat cell proliferation and seed size in wheat. (a) Seeds at 20 days post inoculation (dpi). (b) Cross-sections (dash lines in a) of seeds stained by Fluorescent Brightener. (c) Magnified view of the cross-section boxed in b. (d-f) Comparison of seed size (d) represented by the projective area of a seed, cell number (e) and cell length (f) of outer seed coat between BSMV:00 and BSMV:TaCYP78A3-infected plants at 20 dpi. (g–h) Comparison of ovule or seed size (g) represented by the projective area of a ovule or a seed and cell number of a cross section of the ovule or outer seed integument (h) between BSMV:00 and BSMV:TaCYP78A3-infected plants from 5 to 25 dpi. Data represent the mean ± SE from 10 independently collected seeds that were effectively detected by the target gene silencing; two plants from each of five replications and one seed from each plant were used for measurement. Asterisks (**) indicate significant differences at the P < 0.01 level (t-test). Bars = 2 mm in a; 200 μm in b; 100 μm in c.

**Figure 4.** Effects of *TaCYP78A3* overexpression on developing seeds and reproductive organs in *Arabidopsis*. From left to right in b, c and d) are Col wild-type (WT), *pCYP78A9::TaCYP78A3-3*, *35S::TaCYP78A3-4* and

This article is protected by copyright. All rights reserved.
*pINO::TaCYP78A3-2* plants as a representative for *Arabidopsis* transgenic lines bearing each construct. (a-c) Images of mature seeds (a), siliques (b) and inflorescences (c) from WT and *TaCYP78A3* transgenic lines. The uncoordinated development of androecium and gynoecium in the *pINO::TaCYP78A3-2* is visible (c, white arrow). (d) Col and *pINO::TaCYP78A3-2* flowers at stage 14. (e) A developing flower with sepals removed, and *TaCYP78A3* expression in the pistil and stamen using *pCYP78A9::TaCYP78A3::GUS* analysis at stage 12. (f) Developing seeds showing *TaCYP78A3* expression in the seed coat and embryo at 2 days after anthesis using *35S::TaCYP78A3::GUS* analysis (arrow heads). (g) A developing flower with sepals and stamens removed at stage 13, *TaCYP78A3* expression in the developing ovules inside the gynoecium using *pINO::TaCYP78A3::GUS* analysis. (h) Reverse transcription-PCR analysis with total RNA isolated from a mix of top inflorescence from 10 lines per genotype. (i-l) Quantification of seed and silique characteristics in response to changed *TaCYP78A3* activity. Thirty plant genotypes were grown without any assisted pollination and harvested for measurements. (i) Seed size represented by the projective area of a seed. (j) Seed number per silique. (k) Silique number per main inflorescence stem. (l) Silique length. Values represent means ± SE (n = 20). Asterisks (*) and (**) indicate significant differences from the wild type (Col) at P < 0.05 and P < 0.01 (t-test), respectively. Bars = 2 mm in a, b and d; 500 μm in c and g; 200 μm in e and h; 1 mm in f. Stages are in accordance with those described by Smyth et al. (Smyth et al. 1990).

**Figure 5.** Overexpression levels of *TaCYP78A3* influenced seed size and fertility in *Arabidopsis*. (a-b) The main inflorescences (a) and siliques (b) of wild-type (WT) Col and *pINO::TaCYP78A3-1, -2* and -3 transgenic lines. (c) Ethanol-decolorized siliques from WT and *pINO::TaCYP78A3-2* transgenic lines. (d) Reverse transcription-PCR analysis indicated different expression levels of *TaCYP78A3* in *pINO::TaCYP78A3-1, -2* and -3 transgenic lines. (e-f) Silique and seed characteristics in response to different *TaCYP78A3* overexpression

This article is protected by copyright. All rights reserved.
levels. (e) Seed size (the projective area of a seed). (f) Seed number per silique. (g) Silique number per main inflorescence stem. (h) Silique length. Values represent means ± SE (n = 20). Asterisks (*) and (**) indicate significant differences from the wild type (Col) at P < 0.05 and P < 0.01 (t-test), respectively. Bars = 2 cm in a; 2 mm in b.

**Figure 6.** Increased cell numbers in seed coats due to *TaCYP78A3* overexpression. (a) Fluorescence micrographs of mature seed outer coat stained by Fluorescent Brightener. (b-c) Quantification of cell number (b) and cell area (c) in outer integument of mature seeds from each genotype. (d-e) Light micrographs of alcohol-discolored seed from Col (WT, d) and *pCYP78A9::TaCYP78A3-3* transgenic lines (e) at 1 to 8 days after pollination (dap). (f) Cell numbers in outer integument of developing seed from Col and *pCYP78A9::TaCYP78A3-3* transgenic lines. Values represent means ± SE (n = 20). Asterisks (*) and (**) indicate significant differences from wild-type (WT) Col at P < 0.05 and P < 0.01 (t-test), respectively. Bar =100 μm.

**Figure 7.** *TaCYP78A3* promoted cell proliferation in the embryo integument. (a-c) Light micrographs of alcohol-discolored mature embryos from the wild-type (WT) Col (a), *pCYP78A9::TaCYP78A3-3* (b) and *pINO::TaCYP78A3-2* (c) plants. (d-f) Magnified view of the boxed areas in a, b and c, respectively. (g-i) Quantification characteristics of mature cotyledon epidermis cells in response to changed *TaCYP78A3* activity. Values represent means ± SE (n = 20). Asterisks (*) and (**) indicate significant differences from the wild type (Col) at P < 0.05 and P < 0.01 (t-test), respectively. Bars =100 μm in a, b and c; 10 μm in d, e and f.