Metabolic engineering of rice endosperm towards higher vitamin B1 accumulation

Simon Strobbe1, Jana Verstraete2, Christophe Stove3 and Dominique Van Der Straeten1,*

1Laboratory of Functional Plant Biology, Department of Biology, Ghent University, Gent, Belgium
2Laboratory of Toxicology, Department of Bioanalysis, Ghent University, Gent, Belgium

Received 4 August 2020; accepted 1 December 2020.
*Correspondence (Tel: +32 9 264 5185; email: Dominique.VanDerStraeten@UGent.be)

Summary

Rice is a major food crop to approximately half of the human population. Unfortunately, the starchy endosperm, which is the remaining portion of the seed after polishing, contains limited amounts of micronutrients. Here, it is shown that this is particularly the case for thiamin (vitamin B1). Therefore, a tissue-specific metabolic engineering approach was conducted, aimed at enhancing the level of thiamin specifically in the endosperm. To achieve this, three major thiamin biosynthesis genes, TH1C, TH1D and TH1E, controlled by strong endosperm-specific promoters, were employed to obtain engineered rice lines. The metabolic engineering approaches included ectopic expression of THIC alone, in combination with TH1D (bigenic) or combined with both TH1D and TH1E (trigenic). Determination of thiamin and thiamin biosynthesis intermediates reveals the impact of the engineering approaches on endosperm thiamin biosynthesis. The results show an increase of thiamin in polished rice up to threefold compared to WT, and stable upon cooking. These findings confirm the potential of metabolic engineering to enhance de novo thiamin biosynthesis in rice endosperm tissue and aid in steering future biofortification endeavours.

Keywords: metabolic engineering, vitamin B1, biofortification, rice, hidden hunger, micronutrient deficiency, nutrition, thiamine, endosperm, polishing.

Introduction

Rice is the most important staple for human consumption, serving as central part of the diet for over half of the global population (Zhao et al., 2010). Moreover, two billion people depend on rice as their primary source of calories (Ziska et al., 2015). As rice can be cheap and easily implemented in different climatic zones, it will retain its role as leading calorie supplier to a growing global population. On top of that, the extra pressure exerted by climate change and the increasing world population could make us more reliant on easily applicable staples such as rice (Li et al., 2015; Shrestha et al., 2016; Ye et al., 2015).

On a global scale, rice is mostly consumed as white rice (Paiva et al., 2016; Zhao et al., 2020). Polished white rice is retrieved from rice plants by several industrial processing steps. First, after collection of the rice seeds, the husk or hull is removed, yielding brown rice. At this point, the brown rice seeds still contain the rice bran, aleurone layer, embryo and endosperm. The bran layers are considered rich sources of fibre, fat, proteins and vitamins, while the embryo mainly contains more fat. This high-fat content is undesired since it negatively affects preservation of the rice product (Paiva et al., 2016). Therefore, in a second processing step, brown rice is polished by (industrial) milling, removing the outer layers and embryo, retaining the starchy endosperm, which improves storage potential (Babu et al., 2009). Unfortunately, this latter procedure coincides with the removal of a significant portion of micronutrients from the rice kernel, including B-vitamins and crucial minerals such as iron. For this reason, polished white rice is a very poor source of micronutrients and vitamins (Shobana et al., 2011). This low-nutritional status of white rice becomes particularly problematic in low-income countries considering the increasing dependence on rice (Steiger et al., 2014). Consequently, the population becomes more susceptible to micronutrient deficiencies. To cope with these problems, addition of (synthetic forms of) vitamins to white rice products, known as fortification, is advocated (Steiger et al., 2014). This industrial process, however, often fails to reach the poor rural population, where micronutrient delivery is required most (Blancquaert et al., 2017). These populations often still largely rely on rice as a main supplier of calories, causing a high incidence of micronutrient malnutrition. Interestingly, a study on the poor nutritional properties of polished white rice resulted in the discovery of the first vitamin (vital-amine), being thiamin or vitamin B1 (Semba, 2012).

Thiamin is a water soluble B-vitamin playing a pivotal role in central energy metabolism of all living organisms. In humans, thiamin deficiency (in severe cases referred to as ‘beriberi’) can cause degeneration of both the cardiovascular system (wet beriberi) and the nervous system (dry beriberi). Thiamin deficiency has been linked to cognitive impairment, dementia and depression (Pourhassan et al., 2019). Though severe thiamin deficiency outbreaks are considered rare, chronic thiamin malnutrition is a persisting problem with a severe impact on global human health, particularly affecting the elderly (Hoffman, 2016) and populations of low-income countries (Johnson et al., 2019). Thiamin deficiency, in the form of infantile beriberi, induces many casualties in children in Southeast Asia, being held responsible for up to 45% of all deaths of Cambodian children under the age of five (Johnson et al., 2019). Moreover, thiamin deficiency, likely coinciding with other clinical (and nutritional) pathologies, is...
often overlooked and therefore a largely underestimated concern, which requires adequate resolutions (Nazir et al., 2019). To avoid thiamin malnutrition, the World Health Organization recommends a daily thiamin intake (RDI) as high as 1.5 mg for lactating women (WHO, 2004). Educational interventions, focusing on dietary diversification, vitamin fortification and supplementation are proposed actions to combat this problem. On top of that, biofortification of staples which are highly consumed by the affected population, such as rice, can help decrease the burden of this micronutrient malnutrition (Bouis and Saltzman, 2017; Strobbe and Van Der Straeten, 2018; Van DerStraeten et al., 2020).

Biofortification, the enhancement of natural micronutrient levels, could be employed to deliver micronutrients, and therefore thiamin, in a cost-effective way, requiring only a one-time investment (Blancquaert et al., 2017; Strobbe and Van Der Straeten, 2018). The low levels of B1 present in polished rice might prove to be challenging, as limited variation in thiamin levels, could be employed to deliver micronutrients, and therefore thiamin, in a cost-effective way, requiring only a one-time investment (Blancquaert et al., 2017; Strobbe and Van Der Straeten, 2018). The low levels of B1 present in polished rice, combined with the (largely underestimated) health effects provoked by its deficiency propose thiamin as an ideal target in these biofortification endeavours (Dong et al., 2016). Breeding rice cultivars towards enhanced B1 content of polished rice might prove to be challenging, as limited variation in thiamin content has been observed in polished white rice (Sood et al., 2006), in contrast to higher thiamin variation (up to 10-fold) witnessed in brown rice (Kennedy and Burlingame, 2003).

Metabolic engineering approaches using genetic engineering, on the other hand, do not rely on variation in the available germplasm. Metabolic engineering does, however, require adequate fundamental knowledge on thiamin biosynthesis and metabolism.

The thiamin biosynthesis pathway in plants is shown in Figure 1a and b. In short, a pyrimidine moiety, 4-amino-2-methyl-5-hydroxymethylpyrimidine phosphate (HMP-P) and a thiazole moiety, 4-methyl-5-β-hydroxyethylthiazole phosphate (HET-P) are biosynthesized in the plastid by the action of HMP-P synthase (THIC) (Kong et al., 2008; Raschke et al., 2007) and HET-P synthase (TH1) (Godoi et al., 2006), respectively. Next, the bifunctional HMP-P kinase/TMP pyrophosphorylase (TH1) phosphorylates HMP-P to HMP-P and subsequently condenses HMP-P and HET-P to yield thiamin monophosphate (TMP) in the plastid (Ajawii et al., 2007b). Thiamin is considered to be created in the mitochondria as well as the cytosol by the phosphatase action of TMP phosphatase (TH2) (Hsieh et al., 2017; Mimura et al., 2016). Last, the active cofactor, thiamin pyrophosphate (TPP), is generated by the action of thiamin pyrophosphokinase (TPK) in the cytosol (Ajawii et al., 2007a). Fortunately, all these thiamin biosynthesis genes have (recently) been identified in Arabidopsis (Hsieh et al., 2017; Pourcel et al., 2013; Figure 1a), thereby facilitating the development of potential metabolic engineering approaches (Goyer, 2017; Strobbe and Van Der Straeten, 2018).

This study reports on the endosperm-specific ectopic expression of up to three thiamin biosynthesis genes in rice and its effect on thiamin metabolism in polished seeds.

Results

Generation of transgenic rice lines aimed at achieving endosperm-specific thiamin biofortification

To avoid detrimental side-effects on plant physiology (Bocobza et al., 2013; Wang et al., 2016), biofortification approaches are preferentially limited to the tissue of interest (Strobbe et al., 2018; Strobbe and Van Der Straeten, 2018), in the case of rice being the consumed starchy endosperm. To achieve enhanced accumulation of thiamin in polished white rice, we aimed at endosperm-specific enhancement of thiamin biosynthesis. Amongst the different chemical entities in vitamin B1 (including phosphorylated forms TMP and TPP), thiamin is considered an adequate and chemically stable source of the vitamin. Therefore, the goal of the metabolic engineering approach is the over-accumulation of this specific B1 vitamer. The committed enzymatic reactions as well as the related intermediate metabolites in the biosynthesis towards thiamin are depicted in Figure 1b. Thiamin is found to be the predominant form in which B1 vitamers are stored in rice seeds, explained by the high expression of thiamin de-phosphorylating enzymes in these tissues (Dong et al., 2016; Mangel, 2016). In this regard, stimulation of enzymes downstream of TMP (Figure 1a) is not expected to have an additional beneficial effect on total B1 levels in the endosperm tissue. This metabolic engineering strategy is aided by the existence of well-characterized endosperm-specific rice promoters, proven to yield satisfactory results in previous biofortification approaches (Blancquaert et al., 2015; Storzohenko et al., 2007).

The different strategies utilized in this study all encompass the thiamin biosynthesis enzyme THIC. The rationale behind this is twofold. First, THIC is known to catalyse a rate-limiting step in Arabidopsis thiamin biosynthesis (Bocobza et al., 2013; Pourcel et al., 2013), a feature which is likely conserved in rice endosperm as well. Second, the unique ability of the THIC pre-mRNA to sense the abundance of the thiamin biosynthesis end product, TPP, through the riboswitch located at its 3' UTR, resulting in alternative splicing of the pre-mRNA (Bocobza et al., 2013; Wachter et al., 2007), has the potential to limit THIC activity and thereby B1 biosynthesis flux. Introduction of transgenic THIC deprived of its riboswitch therefore allows increased expression, while abolishing feedback inhibition.

There are, however, some indications hinting towards additional limiting factors (Minhas et al., 2018; Pourcel et al., 2013), downstream of HMP generation and therefore likely persisting after introduction of THIC (Figure 1a and b). This includes the insufficient supply of HET-P, resulting from the chemical reaction governed by TH1. Being a suicidal enzyme, TH1 donates a sulphur residue to the thiazole moiety in thiamin (Chatterjee et al., 2011). A feeding study in Arabidopsis, supplementing plants with both thiamin biosynthesis intermediates, shows that HET shortage prevents thiamin accumulation when provided with sufficient HMP (Pourcel et al., 2013). These results indicate that the combined stimulation of both THIC and TH1 could lead to an increased flux towards thiamin. This notion should be dealt with caution, given that this strategy depends on sufficient activity of endogenous TH1, the step condensing the two aforementioned intermediates.

With these considerations in mind, constructs were created in which the T-DNA harbours: THIC (At1g29630) (single-gene approach, yielding ‘THIC’ lines), THIC and TH1 (At1g54770) (two-gene approach, yielding ‘TT’ lines) and last but not least THIC, TH1 and TH1 (At1g22940) (three-gene approach, yielding ‘TTT’-lines). All constructs involve the THIC sequences of the corresponding genes (see primers in Table S1), thereby facilitating the development of potential metabolic engineering approaches (Goyer, 2017; Strobbe and Van Der Straeten, 2018).

This study reports on the endosperm-specific ectopic expression of up to three thiamin biosynthesis genes in rice and its effect on thiamin metabolism in polished seeds.
Polishing reveals an unequal distribution of thiamin and its related metabolites in rice seeds

To allow a deeper insight in thiamin metabolism, specifically confined to the endosperm, removal of undesired tissues (embryo, aleurone) is essential. Polished white rice kernels were obtained using a Pearlest Rice Polisher followed by vigorous shaking on sandpaper (see Methods S1, Figures S1 and S2). By increasing the time in the polishing machine (M) and/or the shaking on sandpaper (SP), seeds with an increasing degree of...
ectopic expression of

... ectopic expression of ure 1c). Expression analysis of the transgenic lines confirmed the significantly lower expression compared to the other transgenic

... should be noted that lines THIC18.11 and THIC20.5 depicted a

... of thiamin. Indeed, polished white rice (5.6 µg/100 g) (Figure S2B) depic
ta more than 50-fold decrease in thiamin compared to brown rice (306 µg/100 g) (Figure S2C).

... This large drop in thiamin content was confirmed by measuring thiamin in the polishes (corresponding to the 10% of material removed during machine polishing, containing embryo and bran), which reached 2796 µg/100 g (Figure S2C).

... The metabolites HMP and HET are measured as representatives of the pyrimidine and thiazole biosynthesis intermediates, respectively. Unless specified otherwise, the levels of these metabolites that are referred to represent non-phosphorylated HMP and HET. The presence of phosphorylated forms of these metabolites can be derived by comparing the signals obtained for phosphatase- and non-phosphatase-treated samples. Similar to what was observed for thiamin, levels of HMP are relatively high in polishes and brown rice (Figure S2F), but lower in polished seeds (Figure S2E), though this difference is much lower compared to what was seen for thiamin. HET, on the other hand, displays limited variation upon polishing, indicating that it’s more evenly distributed over these different tissues (Figure S2H, I, J).

... The higher abundance of TPP as compared to TMP is more pronounced in polished seeds, indicating low-TMP levels in rice endosperm (Figure S3). However, the polishes do contain a substantial amount of TMP (21 µg/100 g) as well as TPP (86 µg/100 g) (Figure S3B).

... Overall, these results indicate that polishing has a severe effect on the thiamin-related metabolite composition of the seeds. Given these results and the fact that the starch-rich endosperm is the predominantly utilized tissue for human consumption, the effect of the metabolic engineering strategies on the metabolite levels will be primarily addressed by analysis of most thoroughly polished rice (60s M + 16h SP, 14% DOP), hereafter referred to as ‘polished white rice’.

**Ectopic THIC expression alters the biosynthetic flux towards thiamin in rice endosperm**

Twelve transgenic lines, resulting from independent transformation events, were created, harbouring the ‘THIC’ T-DNA (Figure 1c). Expression analysis of the transgenic lines confirmed the ectopic expression of Arabidopsis THIC in rice endosperm, although it should be noted that lines THIC18.11 and THIC20.5 depicted a significantly lower expression compared to the other transgenic lines analysed (Figure 2a). LC-MS/MS analysis did not reveal a difference between HMP and total HMP (the latter also including HMP-P or HMP-PP) in polished kernels, indicating that phosphorylated HMP metabolites are either absent or very low abundant in endosperm tissue. Therefore, measurement of HMP levels as total HMP was performed as a means to estimate the accumulation of intermediates of the pyrimidine branch of thiamin biosynthesis (controlled by THIC), as HMP-P is the actual intermediate presumed to be converted into HMP upon its accumulation. The ectopic expression of THIC did not result in a generally detectable accumulation of HMP, although some lines depicted a slight increase in HMP (Figure 2b), which was found to be significant in line THIC18.11. Assessment of the level of the thiazole intermediate in thiamin biosynthesis (HET-P) is also relevant, as it could potentially indicate changes in flux through the biosynthesis pathway and/or reveal potential regulatory effects of ectopic THIC expression. To assess the status of the thiazole intermediate, both HET and HET-total (sum of HET and its phosphorylated form HET-P) were measured. In WT and transgenic THIC lines, HET and HET-total levels were found to be approximately the same, indicating the absence of HET-P. Several lines (including THIC18.11 and THIC27.7) had increased levels of HET, while others (THIC24.7, THIC26.5 and THIC29.5) depicted a significant decrease in the level of this thiazole entity as compared to WT (Figure 2c).

... Changes in thiamin metabolism, caused by introduction of THIC, appear to affect the downstream biosynthesis pathway, as multiple lines displayed increased thiamin content (over twofold thiamin enhancement in eight out of twelve lines; Figure 2d).

**Ectopic expression of both THIC and THI1 in rice endosperm augments the level of the thiazole intermediate as well as thiamin**

Sixteen transgenic lines, resulting from independent transformation events, were created, harbouring the THIC-THI1 T-DNA (Figure 1c). These TT lines demonstrated ectopic expression of THI1 (highest in line TT50.6) as well as THIC (though low in TT54.7) (Figure 3a). Furthermore, the level of THIC expression was comparable with previously discussed THIC lines, solely expressing transgenic THIC. Similar to what was observed in THIC lines, no clear overall effect on HMP levels was seen, though lines TT45.4.1 and TT54.7 displayed significantly decreased HMP levels (Figure 3b). Particularly for line TT54.7, low-HMP levels are likely due to inadequate THIC activity. In the TT lines, the endosperm-specific stimulation of THI1 enabled augmentation of HET, the thiazole intermediate of the pathway. Interestingly, in contrast to WT and THIC lines, all TT-engineered lines had a higher level of HET-total as compared to HET (Figure 3c), suggesting the presence of elevated levels of HET-P in the TT lines (Figure S7B).

... Most notably, the content of HET-total (sum of HET-P and HET) in TT50.6 (28 µg/100 g) reached almost fivefold the WT level (5.6 µg/100 g). Moreover, the stimulation of both the pyrimidine and thiazole biosynthesis, via ectopic expression of THIC and THI1, respectively, induced an increase in thiamin levels in polished seeds (Figure 3d).

**Additional introduction of THI1 in rice endosperm causes a drop in thiazole levels, while ensuring an adequate flux towards thiamin**

Fourteen independent transgenic lines were created to achieve ectopic expression of THIC, THI1 and THI1 (TTT lines). The TTT lines exhibited expression of all three transgenes (with the exception of TT58.4.1.1) (Figure 4a), in which THIC and THI1 expression was...
comparable to previously described THIC- and TT-engineered lines. Line TTT69.7 exhibited a strong expression of all three transgenes, while TTT58.4.1 showed no transgene expression (though confirmed genotypically, and thus included in the analysis). The ectopic expression of these three genes did not appear to have a significant effect on the HMP content of polished seeds (Figure 4b). In contrast to the higher levels of HET seen in TT lines as compared to WT (5.6 µg/100 g) (Figure 3c), TTT lines depicted lowered HET levels, more than threefold reduced in TTT10.5.1 (1.6 µg/100 g) (Figure 4c). However, our data indicated that the phosphorylated biosynthetic thiazole entity, HET-P, was still obviously present in the polished seeds of these engineered lines (see also Figure S7), reflected by the higher levels of HET-total as compared to HET (Figure 4c). Boosting of thiamin biosynthesis via this trigenic approach was, in many lines, apparent from the observed thiamin levels of the polished seeds (Figure 4d), as thiamin content of the polished seeds reached a more than twofold increase compared to WT in lines TTT8.14 (2.6 fold), TTT10.5.1 (2.4 fold), TTT73.8.1 (2.5 fold) and TTT19.1 (2.4-fold).

Altered metabolite profiles of the engineered lines are confirmed in brown rice

The changes in composition of thiamin and thiamin biosynthesis related metabolites in polished white grains raise the question whether or not this effect is also reflected in the composition of the full grain. Therefore, metabolite levels were checked in brown rice (unpolished) of selected engineered rice lines. Lines depicting the most interesting metabolite profiles (e.g. higher content of thiamin, HMP or HET) were chosen to examine the unpolished
brown rice. Evaluation of these lines did indicate a change in full grain HMP content, as compared to WT, of several THIC (e.g. THIC18.11) and TTT lines (e.g. TTT56.4) (Figure 5a). These lines, THIC18.11 and TTT56.4, also displayed the highest level of HMP in polished seeds of THIC (Figure 2b) and TTT (Figure 4b) engineered lines, respectively. The highest HMP level in brown rice was observed in line TTT56.4 (8 µg/100 g), displaying a more than twofold increase as compared to WT (3.9 µg/100 g). Similar to polished seeds, no difference in total and non-phosphorylated HMP was observed in the unpolished seeds, suggesting the absence of significant levels of HMP-P or -PP. Determination of both HET (non-phosphorylated) and HET-total confirmed the observations made in polished seeds, suggesting the absence of significant levels of HMP-P or -PP. Determination of both HET (non-phosphorylated) and HET-total confirmed the observations made in polished seeds, although the drop in HET content in TT lines was only observed in one line (TTT10.5.1) (Figure 5b). Indeed, in accordance with what was observed in polished seeds (Figures 3c, 4c), analysis of brown rice confirmed the accumulation of HET-P in TT and TTT lines (Figure 5b). Furthermore, the observation that some THIC lines display increased HET levels (THIC18.11, THIC27.7), while a decrease was observed in others (THIC26.5) (Figure 2b), was confirmed in the unpolished seeds (Figure 5b). Interestingly, the higher thiamin levels witnessed in polished seeds (Figure 2d, 3d, 4d) were also reflected in the thiamin levels of unpolished brown seeds for multiple lines (Figure 5c), though that effect was less pronounced in brown rice. In lines THIC17.5 and THIC34.6.5, the higher thiamin content of the endosperm was also observed in brown rice, although this increase was not significant due to the much higher thiamin content in brown rice. Similarly, lines that did not display a significant thiamin increase in polished seeds (e.g. THIC18.11, THIC20.5, THIC26.5), also appeared to have thiamin levels resembling WT in brown rice. Finally, many engineered lines depicted a modest increase in TPP levels in brown rice (Figure S7C).

The enhanced thiamin levels in polished seeds remain stable upon cooking as well as storage

As the eventual goal of the metabolic engineering strategy is to obtain rice lines harbouring a nutritional benefit concerning thiamin content, the material should be evaluated resembling the state in which it is consumed, that is cooked white rice (polished), likely after storage periods of at least up to 4 months at higher temperatures (Blancquaert et al., 2015; Ding et al., 2016). Assessing the content of thiamin and thiamin-related metabolites in polished cooked rice allows to evaluate thiamin stability after cooking. Moreover, assessment of these metabolites in samples exposed to prolonged storage at adequate conditions.
temperature (28°C), allows insights in the effect of storage on metabolite content. Therefore, transgenic lines were chosen for each of the three engineering strategies, exhibiting high thiamin (and in some cases also HMP and/or HET) to assess the effect of cooking as well as storage. Metabolite levels are displayed in µg /100 g of untreated polished rice seeds (weight before cooking or storage procedure), allowing adequate comparison between treated (cooked or stored) and unprocessed polished seeds. Overall, adequate stability of HMP, HET and thiamin was found upon cooking as well as storage (Figure 6). Phosphorylated metabolites such as HET-P (Figure S7F, 7c) and TPP (Figure S7D) displayed a more obvious drop upon cooking. Most notably, thiamin levels were not significantly affected by the (30 min) cooking procedure nor by 4-month storage (Figure 6d). Though the levels of thiamin in cooked transgenic seeds appeared higher, they were not found to be significantly different from those of raw seeds of the corresponding line. Nevertheless, this could provide some insight in the chemical state of the vitamin in the engineered lines (e.g. bound by thiamin binding proteins). Overall, the augmented metabolite levels show to be in a stable state in the engineered lines, withstanding higher temperatures, demonstrating no substantial losses upon cooking or storage.

**Discussion**

The practice of polishing causes a tremendous drop in thiamin content of rice seeds

The vitamin B1 profile in brown rice, consisting of thiamin (306 µg/100 g or 11.63 nmol/g), TMP (3.8 µg/100 g or 0.11 nmol/g) and TPP (7.1 µg/100 g or 0.17 nmol/g), shows that thiamin makes up almost 98% of the B1 pool (see Figure S8; Figures S2, S3). The observation that B1 predominantly exist in the form of (non-phosphorylated) thiamin in brown rice was already reported 30 years ago (Shimizu et al., 1990). Moreover, the level of thiamin measured in brown rice is similar to what has been observed in different rice varieties (Kennedy and Burlingame, 2003), reaching very similar levels as described in both Japonica (Dong et al., 2016) and Indica (308 µg/100 g) (Swamy et al., 2019) varieties.

Polishing of rice seeds has an immense effect on the thiamin content, dropping to 5.6 µg/100 g (Figure S2b), corresponding

---

**Figure 4** Transgene expression and metabolite levels in TTT-engineered polished rice. In the different panels, the pyrimidine branch (THIC and HMP) is indicated in blue, the thiazole branch (THI1 and HET(-P)) in orange/red and downstream the pathway towards B1 vitamers (TH1 and thiamin) in green (a) Relative transgene expression of AtTHIC, AtTH1 and AtTH1 in rice seeds of homozygous T2 plants originating from independent plant transformation events revealed by real-time quantitative PCR. Values are means of two independent biological repeats (±SE). Data analysis and normalization were performed using the qBASE software, based on the 2−ΔΔCt method (Hellemans et al., 2007; Livak and Schmittgen, 2001) using OsActin (LOC_Os03g50885.1), LOC_Os11g43900 and LOC_Os07g02340 as reference genes. Specific information regarding primers and conditions can be retrieved from Table S1. Level (mean ± SE) of 4-amino-2-methyl-5-hydroxymethylpyrimidine (HMP) (b), 4-methyl-5-[β]-hydroxyethylthiazole (HET) and HET-total (sum of non-phosphorylated and phosphorylated HET) (c) and thiamin (d) as measured by quantitative LC-MS/MS analysis. Metabolites were measured in thoroughly polished white rice (60s M + 16h SP, see Figure S2). Analysis of metabolites was performed on at least three biological repeats (transgenic lines) and 15 biological repeats (WT). Statistical difference (see Methods S2) in mean metabolite level compared to WT is indicated to be significant (single asterisk, P < 0.05) or very significant (double asterisks, P < 0.01).
to a 98% loss. The degree of polishing is similar to what has been described to acquire clean white rice (10%–15% DOP; Liang et al., 2008; Monks et al., 2013). Assessment of thiamin levels in polished rice reveals the existence of a thiamin barrier, with exceptionally high levels in the embryo and outer layers of the grain contrasting the low levels of thiamin in the endosperm. Indeed, the germ and bran (polishings) exhibit a 500-fold higher thiamin content as compared to the endosperm (Figure S2B, C, D). This thiamin barrier lays between the starchy, white endosperm and the lipid-rich aleurone and embryo, with a thickness of a few hundred micrometer (see Figure S1B). This discrepancy in thiamin content likely reflects the different physiological/metabolic roles of these tissues, in which the embryo requires high-thiamin levels to acquire TPP for germination, as TPP is needed in the catabolism of the energy supply provided by the endosperm (Golda et al., 2004; Mitsunaga et al., 1987). Interestingly, this coincides with the occurrence of the proposed thiamin binding proteins in the rice germ, either causing or aiding the accumulation of thiamin in this tissue (Mitsunaga et al., 1987; Shimizu et al., 1996).

Concerning the levels of the thiamin biosynthesis intermediates, both HET and HMP content vary less dramatically upon polishing as compared to thiamin (Figure S2D, G, J). However, HMP levels in rice bran are up to 16-fold higher as compared to...
the endosperm (Figure S2E, F). These results also demonstrate the molar concentration of HMP in polished seeds to be slightly lower than thiamin, whereas the concentration of HET is approximately twice as high as thiamin (Figure S2, Table S3, Figure S8). The molar concentration of these metabolites seems to indicate that the biosynthetic flux towards the pyrimidine and thiazole moiety is not limiting; however, insufficient salvage towards the phosphorylated entities could still be a limiting factor. The presence of biosynthetic intermediates suggests that elevating the biosynthetic flux towards thiamin (through TH1 function) holds the potential to increase endosperm thiamin content, as the intermediates are either synthesized or transported towards the endosperm.

The observed high levels of thiamin (2796 µg/100 g, Figure S2C) in the rice bran (polishings) indicate its strong nutritional potential. Indeed, the required amount of thiamin to reach the RDI (lactating women, 1.5 mg thiamin/day) is present in only 54 g of rice bran material, further strengthening the call to utilize the maximal nutritional potential of this food product (Bodie et al., 2019; Sharif et al., 2014). To acquire higher thiamin intake as well as harness other nutritional benefits attributed to rice bran, this ‘by-product’ could be consumed as an additive in certain baked food products (Esa et al., 2013). Nevertheless, it should be noted that metabolic engineering of thiamin content in endosperm is required, as polished white rice is preferably consumed by populations affected by thiamin deficiency, given its storage potential.

**Thiamin biofortification of rice endosperm**

The results generated in this study indicate that thiamin metabolism in rice endosperm can be guided towards higher thiamin formation. However, it should be noted that the metabolic flexibility allowing to cope with this increase in flux towards thiamin is limited. Determination of thiamin levels as well as thiamin biosynthesis intermediates allows a proper assessment of the effects of the different metabolic engineering approaches. Furthermore, the effect of the engineering approaches on the different thiamin-related metabolites is reflected in unpolished brown rice (Figure 5) and persists upon cooking as well as storage.
Hence, the assessment of these engineered lines exposes novel aspects in thiamin metabolism of rice endosperm. To understand the rationale and impact of metabolic engineering of rice endosperm for thiamin content, insights in endosperm thiamin metabolism is required. Thiamin present in rice endosperm is presumed to be made by de novo biosynthesis during early endosperm development, likely enforced by pyrimidine salvage (Goyer, 2017). Indeed, the endosperm does exhibit expression of genes responsible for de novo thiamin biosynthesis (Dong et al., 2016; Mangel, 2016). However, it is proposed that additional thiamin can be acquired via import of either thiamin itself or its biosynthesis intermediate HMP from the surrounding tissues, similar to what is known in maize endosperm (Minhas et al., 2018). Here, real-time quantitative PCR analysis revealed expression of the native thiamin biosynthesis genes (OsTHIC, OsTHI1, OsTHI2 and the recently characterized OsTH2 (Hsieh et al., 2020)) in WT and transgenic rice seeds. The engineering approaches appeared to have no clear effect on the expression of the endogenous thiamin biosynthesis genes (Figure S10). This analysis also indicated that the levels of transgenic expression in engineered lines exceeded the levels of the corresponding native transcripts (Figure S11).

Ectopic expression of the thiamin biosynthesis gene THIC shows the potential to raise thiamin levels in both polished (Figure 2d) and brown rice (Figure 5c). Transgenic TH1 expression on the other hand, can augment the levels of HET in TT-engineered lines (Figure 3c). Moreover, these lines exhibit accumulation of phosphorylated HET (Figures 3c, 5b), which holds true for lines engineered with the downstream biosynthesis gene TH1 as well (Figures 4c, 5b). The link between TH1 expression and HET-P accumulation is further confirmed by the observation that the highest TH1 expressing TT lines (TT50.6 and TT33.14) depict the highest level of total HET as well as HET-P (Figures 3a, c, S7B). Similarly, line TTT78.10, expressing the 3 key genes in the thiamin pathway, exhibits low expression of THI1 and no apparent differences in HET and total HET levels (Figures 4a, c, S7B). Transgenic introduction of the downstream biosynthesis gene TH1 also demonstrates a clear effect, lowering HET content of polished seeds (Figure 4c), indicative of an activation of the downstream thiamin biosynthesis pathway mediated by TH1. An exception to this is line TTT69.7, in which high THI1 expression could explain the remaining elevated HET pool (Figure 4a, c), hinting that reaching even higher expression levels of TH1 could aid in avoiding a depleted pool of HET-P.

Despite the apparent success of the THIC-engineering approach, there are multiple indications of persisting bottlenecks in thiamin biosynthesis.

Figure 7 Schematic representation of the flux through the B1 biosynthesis pathway in endosperm of WT and engineered rice. A simplified vitamin B1 pathway, with the three major enzymes responsible for its biosynthesis, utilized in the metabolic engineering approaches, is depicted. Engineered enzymes are depicted in a larger font and are accompanied by an arrow pointing up. Similarly, relevant intermediates are displayed in the colour corresponding to the branch they are part of (pyrimidine, blue; thiazole, orange; B1 vitamers, green). Arrows next to the metabolites indicate an increase (full arrow up), decrease (full arrow down) or other potential outcomes (hollow arrows) as a result of the engineering strategy. The distribution of thiamin in seeds is shown in green, intensified in engineered lines. Higher HET-P levels in TT lines is depicted by a more intense orange colouring of the endosperm. Flux through the thiamin biosynthesis pathway in WT (a), THIC-engineered (b), TT-engineered (c), and TTT-engineered (d) lines.
First, inadequate flux in thiamin biosynthesis, downstream of pyrimidine and thiazole formation, could be hampering B1 accumulation in rice endosperm. These steps include endogenous TH1 and TH2 activity as well as salvage of the pyrimidine and thiazole intermediates. The observation that both pyrimidine and thiazole precursors are present in polished seeds (Figure S8), which do not depict a severe decline in content upon polishing as seen for thiamin (Figure S2B, E, H), indicates that inadequate TH1 activity could be limiting B1 accumulation in WT seeds. This could raise the question whether transgenic expression of TH1 only could aid in achieving higher B1 levels in seeds. The low-molar content of pyrimidine and thiazole precursors in WT seeds (Figure S8), however, is indicative for a limited success of such engineering endeavour. Substantial increase in rice seed B1 content therefore requires an increased supply of both biosynthetic intermediates (Minhas et al., 2018). Evaluation of the THIC-engineered lines provides some indication of bottlenecks downstream of pyrimidine and thiazole biosynthesis. Lines THIC18.11 and THIC36.1 depict increased levels of HMP and HET as compared to WT, while thiamin levels remain as low as in WT (Figure 2b, c and d). This could be caused by instability of the phosphorylated HMP intermediates (HMP-P, HMP-PP) in the (maturing) endosperm or inadequate capability to salvage the biosynthetic intermediates as well as a potential inadequate endogenous TH1 activity. Indeed, insufficient endogenous TH1 activity (catalysing HMP-P condensation with HET-P to form TMP) is also suggested by the high-HET (and HET-P) levels in TT-engineered lines (Figure 3c), which drop upon ectopic TH1 expression (Figure 4c). However, the drop in HET levels witnessed in multiple THIC-engineered lines, likely caused by the increased supply of the pyrimidine moiety (Figure 2c), combined with the higher thiamin levels in various THIC and TT-engineered lines, reflects limited endogenous TH1 activity. This effect is likely explained by the timing of transgene activation, as native TH1 expression is proposed to drop during endosperm maturation (Dong et al., 2016). These limitations could potentially also be attributed to the inadequate downstream conversion of TMP (the product of TH1 action) to thiamin, carried out by the cytosolic TMP phosphatase, TH2 (Hsieh et al., 2020). Fortunately, the recent characterization of this essential step in rice thiamin biosynthesis allowed to acquire a deeper insight into rice thiamin metabolism. Here, the transcript levels of OstH2 were found to be roughly of the same order of magnitude as the other endogenous thiamin biosynthesis genes (Figure S11). This bottleneck could result in a temporary buildup of TMP, which might not be detected due to its higher instability compared to thiamin. Therefore, poor salvage of these breakdown products such as pyrimidine derivatives, removed by pyrimidine salvage enzymes (Zallot et al., 2014), could be hampering B1 biosynthesis.

Second, insufficient TH1 activity can hinder metabolic engineering approaches, shown by the drop in HET content in TT-engineered lines as well as multiple THIC-engineered lines. The lowered HET levels in TT lines (Figure 4c) indicate the ability of TH1 to influence the metabolic equilibrium, in which the thiazole intermediate might become the depleted factor, posing a strong contrast to TT-engineered lines. However, it should be noted that an insufficient supply of the thiazole intermediate is not confined to inadequate TH1 activity, as HET salvage by thiazole kinase is presumably insufficient to allow satisfactory HET-P recovery (Dong et al., 2016; Yazdani et al., 2013). The notion that the enzymatic reaction, executed by TH1, is the rate-limiting step in many engineered lines, is not surprising, given the suicidal nature of this enzyme. Unfortunately, increasing TH1 abundance might impose a metabolic burden, considering that its turnover has been estimated to consume 2% to 12% of the plant maintenance energy (energy cost to maintain basic metabolism) (Hanson et al., 2018; Sun et al., 2019). The implications thereof, with respect to future biofortification approaches, are detrimental, which further stresses the necessity to employ precise metabolic engineering confined to the target tissue of interest. This is exacerbated by the near-suicidal nature of the THIC enzyme (Hanson et al., 2018). Applying non-suicidal variants of these enzymes has been proposed as a solution to this problem (Hanson et al., 2018). This could be feasible by utilization of prokaryotic or even plant specific non-suicidal TH1 analogs (Joshi et al., 2020; Sun et al., 2019). Given their higher energetic efficiency, these more efficient non-suicidal TH1 variants could potentially even provide a yield benefit (Joshi et al., 2020).

Third, the substrates required for the biosynthesis of the pyrimidine moiety (executed by THIC), which include S-adenosylmethionine (SAM) and 5-aminimidazole ribonucleotide (AIR) (Figure 1a), might be limited in the endosperm tissue. This has indeed been proposed to be the case in rice endosperm (Goyer, 2017), similar to what was suggested in maize endosperm (Guan et al., 2014). In this respect, the more pronounced increase of HMP in brown rice as compared to polished white rice could potentially be explained by a higher supply of substrates in this tissue, enabling synthesis of the pyrimidine moiety upon leaky transgene expression in the developing bran (potentially containing stabilizing proteins). The higher HMP levels in rice bran could result from a yet uncharacterized transport towards these tissues, partially counteracting the engineering approaches. These hypotheses of inadequate supply of substrate to allow formation of the pyrimidine moiety as well as depletion via transport can explain the limited increase in HMP observed in polished THIC-engineered lines (Figure 2b). Moreover, pinpointing precursors for pyrimidine biosynthesis as a limiting factor in thiamin accumulation upon THIC stimulation would explain the observation that TTT (and TT) engineered lines do not clearly outperform TT lines in regard to thiamin accumulation. Another explanation for these observations could be the inadequate localization of the THIC enzyme to the plastids of the developing rice endosperm. This could result in insufficient supply of biosynthetic precursors as well as pyrimidine formation in a non-desired subcellular location. However, the measured impact on thiamin metabolism of the introduced THIC, as well as TH1 and TH1 transgenes, confirms that the corresponding enzymes are able to execute their enzymatic reactions. Nevertheless, it cannot be excluded that the engineering approaches can be strengthened by more efficient subcellular trafficking of the engineered enzymes.

The low-thiamin levels in the endosperm, combined with the high-thiamin levels in seed aleurone and embryo spark the idea that biofortification approaches could greatly benefit from altered thiamin transport, breaking the aforementioned thiamin barrier. Indeed, the first attempts to achieve this, by usage of constitutive expression of the known long-distance thiamin transporter PUT3 (AtSg05630) (Martinis et al., 2016), resulted in a 2.3-fold increase of polished grain thiamin (Mangel, 2016). This shows the high potential of metabolic engineering approaches using thiamin reallocation to fill the maturing seeds with ex situ biosynthesized thiamin. These endeavours should, however, be approached with caution, as aberrant thiamin transport in rice (constitutive approach) might severely alter plant physiology (Bocobza et al.,...
2013; Chai et al., 2020). Therefore, a combined approach, stimulating de novo (in situ) thiamin biosynthesis in both endosperm as well as adjacent tissues, combined with enhancing thiamin transport on the endosperm-bran barrier, is a promising way to achieve thiamin biofortification in rice endosperm. Similarly, elevating thiamin import into the endosperm (sink) by engineering its transport from photosynthetic tissues (source), which are assumed to have sufficient thiamin biosynthesis (Guan et al., 2014), might enable high-thiamin accumulation in polished seeds (sink). In such metabolic engineering approaches, transporters of intermediates such as HMP (Beaudoin et al., 2018) could also be adapted to be used as intercellular transporters. Thus, by importing de novo synthesized thiamin (and/or HMP) from neighbouring tissues, the proposed limitation that biosynthesis precursors might exert on endosperm-specific thiamin biosynthesis could be circumvented.

On top of acquiring satisfactory metabolite levels, it is important to monitor the influence of metabolic engineering approaches on the general metabolism of the engineered tissue, as well as on the overall plant physiology, which is particularly important in the light of yield. It should be noted that some transgenic lines clearly depict lowered seed weight (Figure S5C). Moreover, significant deviations from WT were also observed for plant length, shoot mass and grain yield (Figure S5A, S5B, SSD). It could indeed be speculated that the metabolic engineering approaches have a negative impact on grain weights, though this could indeed be speculated that the metabolic engineering approach. Although the thiamin supplied by these engineered transgenic lines is still fairly low, they provide valuable insight in thiamin THIC, TH1, THI1 and TH1, all exert a specific effect on thiamin-related metabolites in the rice endosperm, allowing augmentation of B1 in this tissue, which is demonstrated to be stable upon both cooking and storage procedures. Taken together, these results illustrate the possibility to enhance de novo biosynthesis of thiamin in endosperm. As trigenic engineered lines (TTT) show a decline in HET levels, confirming increased flux towards B1, achieving higher TH1 activity or enhanced salvage of the thiazole moiety are potential targets in future biofortification strategies. As the results highlight a tremendous loss of thiamin in the process of polishing, future biofortification approaches could also be targeting adequate transport of thiamin and/or its intermediates, overruling the existing barrier between contrasting thiamin levels in endosperm versus bran. Although an increase in thiamin content in cooked polished rice seeds of over 5-fold compared to WT has been observed (Figure 6d), consumption of around 6 kg would be required to reach the RDI, considering a bioavailability of 50%. Future research should utilize the acquired insights and tackle the remaining bottlenecks to pursue thiamin-rich rice. These insights in metabolic engineering of thiamin content in rice endosperm represent an important step forward towards creation of high-thiamin rice to combat the persisting global burden of thiamin malnutrition.

Experimental procedures

For information regarding rice seed polishing, statistical analysis, cooking and storage of rice samples as well as harvesting and assessment of yield parameters, we refer to Methods S1, S2, S3 and S4, respectively.

Molecular cloning and construct design

Arabidopsis thaliana (Columbia-0) cDNA was used to amplify the coding sequences of THIC, TH1 and TH1 genes. These were cloned into pDONR221 (via BP recombination reaction) using the Gateway cloning system (primers depicted in Table S1). Subsequently, the fragments were transferred (via LR recombination reaction) into pre-existing destination vectors harbouring endosperm-specific promoters (Glutelin B1, Globulin) and Tnos terminator (Storozhenko et al., 2007). These transcriptional units (TU) were delivered (classical restriction–ligation) to the pMOD35S (binary vector containing hygromycin resistance marker gene, (Storozhenko et al., 2007)) via restriction–ligation reactions (PI-Psp1 for the THIC-TU, I-ceu for the TH1-TU and I-Ppo for the TH1-TU). This resulted in the plant transformation binary vectors harbouring the T-DNA segments as represented in Figure 1c.

Plant transformation and growth

Oryza sativa subsp. japonica var. Nipponbare was used in this project. Scutellum-derived rice calli (Rueb et al., 1994) were transformed with the binary vectors using Agrobacterium tumefaciens strain LBA4404. Plant transformation was performed according to the protocol of Hiei and coworkers (Hiei et al., 1994; Scarpella et al., 2000). Rice seeds were germinated on half strength MS medium for two weeks, after which plantlets were transferred to soil. Rice plants were grown on soil at 80% relative humidity and 28°C. Plants were grown in Structural Type I universal soil (Snebbout, Kaprijke, Belgium), supplemented with iron sulphate (1 g/kg dry soil) and ammonium sulphate (0.5 g/kg dry soil). Plants were grown in pots of 0.8 L (13 cm diameter, 12 cm height), with 2 plants per pot. Light regime (420 μmol/ m²/s) was set at 12 h/12 h. In these conditions, mature seeds were harvested at 4 months (120 days) of growth on soil.

Genotyping, RNA extraction and expression analysis

Genomic DNA was isolated from rice leaves with the Invitiorb Spin Plant Mini Kit (Invitrek) according to the manufacturer’s protocol. The genomic DNA enabled confirmation of the presence of the T-DNA construct in T0-generation, as well as selecting homozygous lines in T1 generation. This was achieved via quantitative real-time PCR (qPCR) using specific primer and probe combinations to both HPT1 and rice sucrose-phosphate synthase 1 (OsSPS) (specific conditions and primers/probes, see Table S2; Storozhenko et al., 2007). Accurate visualization of HPT1 copies (residing on the T-DNA) was done via qPCR on genomic DNA, using the
endogenous genomic OsSSPS for normalization. This methodology allowed selection of homozygous plants from T1 plants, the seeds of which were used to grow the T2 generation to be analysed in this study (shown in Figure S4 using line TT663 as an example).

Endosperm mRNA was extracted from five seeds at the milky stage of seed development (14-day post-anthesis). RNA was extracted using Trizol reagent (Invitrogen), followed DNAse treatment (ThermoFisher Scientific) and additional cleanup procedure using the GeneJET plant RNA purification kit (ThermoFisher Scientific). A detailed protocol of these combined procedures is shown in Figure S9. A Bio-Rad Script cDNA synthesis kit (containing oligo dT as well as random primers) was used to convert 1 µg of RNA to cDNA. Subsequently, qPCR was conducted on these cDNA samples using qPCR primers shown in Table S1. qPCRBIo SYGreen Mix with Fluorescein (Sopachem) was utilized in the reaction mix. A real-time quantitative PCR run comprised an initial denaturation at 95 °C for 3 min and 40 cycles of denaturation at 95 °C for 5s, annealing for 20s (variable temperature, see Table S1), and extension at 60 °C for 6s. Data analysis and normalization were performed using the qBASE software, based on the 2−ΔΔCt method (Hellemons et al., 2007; Livak and Schmittgen, 2001). Normalization was obtained by amplification of rice reference genes OsActin (LOC_Os03g50885.1) (Ji et al., 2014), LOC_Os11g43900 and LOC_Os07g02340 (Blancquaert et al., 2015).

LC-MS/MS determination

Quantification of metabolites (HMP, total HMP, HET, total HET, thiamin, TMP and TPP) was performed using a validated LC-MS/MS method, as described by (Verstraete et al., 2020a; Verstraete et al., 2020b). The analysis requires 200 mg of freshly harvested rice material (polished, unpolished or cooked seeds (see below). The quantification involves splitting the sample in two, measuring one half of the sample following treatment with phosphatase, allowing the assessment of total HET, total HMP and total B1. The other half, not treated with phosphatase, allows assessing HET, HMP, thiamin, TMP and TPP. The ‘total’ therefore reflects the combined level of the non-phosphorylated metabolite together with its phosphorylated entity/entities. The total level is represented as µg/100 g of the non-phosphorylated metabolite. Insight into the levels of phosphorylated metabolites can be obtained by looking at the difference between phosphatase-treated (total) and non-treated samples. It should be noted that the methodology displayed a greater uncertainty on measurements of TMP content. Unless stated otherwise, the levels are represented as µg /100 g of freshly harvested, mature rice seeds.

Acknowledgements

SS was indebted to the Agency for Innovation by Science and Technology in Flanders (IWT) for a predoctoral fellowship. S.S. gratefully acknowledges the Bijzonder Onderzoeksfonds (BOF) for a Post-Doctoral fellowship (BOF-P.DO.2019.0008.01). J.V. would like to thank the Research Foundation Flanders(FWO) for granting her a PhD fellowship (1561167N). D.V.D.S. and C.S. are grateful to Ghent University for financial support (GOA 01G00409; BijzonderOnderzoeksfonds UGent). The authors thank Dieter Blanchquaert for helpful discussions in the first phase of the project.

Conflict of interest

Authors declare no conflict of interest.

Author contributions

D.V.D.S., S.S., J.V. and C.S. designed the experiments. S.S. conducted molecular cloning, plant transformation, genotypic analysis, expression analysis, phenotyping and sampling of rice material. Molecular data analysis was done by S.S. and D.V.D.S. J.V. operated LC–MS/MS and analysed the LC–MS/MS data together with C.S. S.S. and D.V.D.S wrote the manuscript, J.V. and C.S. commented and approved the manuscript. D.V.D.S. conceived and coordinated the project.

References


Supporting information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Figure S1 Rice seeds upon different polishing methodologies.
Figure S2 Rice polishing and its effect on the levels of thiamin and its related metabolites.
Figure S3 Effect of polishing on the levels of phosphorylated B1 entities.
Figure S4 Zygosity analysis of TTT63 T1 generation.
Figure S5 Assessing yield parameters in transgenic engineered rice lines.
Figure S6 Picture of rice plants, panicles and polished seeds.
Figure S7 Levels of metabolites TPP and phosphorylated HET (HET-P) in polished and brown rice seeds.
Figure S8 Schematic representation of the relative molar concentrations of the metabolites giving rise to thiamin in rice seeds.
Figure S9 Protocol for RNA extraction from rice seeds.
Figure S10 Relative expression of endogenous thiamine biosynthesis genes OsTHIC, OsTH1, OsTH1, and OsTH2.
Figure S11 Semi-quantitative comparison of endogenous and transgene expression in rice seeds.
Table S1 Primer Information.
Table S2 Information regarding real-time quantitative PCR on rice genomic DNA to determine zygosity.
Table S3 Conversion of metabolite levels from mass based (µg/100 g) to molar (fmol/g) based units.
Methods S1 Rice seed polishing.
Methods S2 Statistical analysis.
Methods S3 Cooking and storage of rice samples.
Methods S4 Harvest and assessment of yield parameters.